TWMS J. Pure Appl. Math., V.4, N.2, 2013, pp.146-158

AN ANALOGUE OF THE CENTRAL LIMIT THEOREM FOR SOFT PROBABILITY

D.A. MOLODTSOV[1]

Abstract. Under the assumptions of the stability of sliding means and of the finiteness of spread, bounds for the soft probability of the deviation of the arithmetic mean from the interval mean are calculated.

Keywords: soft sets, soft probability, central limit theorem, soft hypothesis.

AMS Subject Classification: 03E72, 03E99, 60A05, 60F05, 62A99, 62A01, 60A99.

1. Introduction

Modern applied mathematics pays increasing attention to problems with undetermined multipliers. This is mainly caused by enhanced attention to control problems under uncertainty conditions. This class of problems includes the problems of controlling a financial portfolio, controlling a communication network, prediction, and many other applied problems.

The basic apparatus for dealing with uncertainties is the apparatus of classical probability theory based on Kolmogorov's axiomatics. However the application of probability theory in practice often involves difficulties, which have been noted by various authors since long ago [1]–[4].

On our opinion, the main problem is that verifying the hypothesis of stochastic stability, on which the practical application of probability theory is based, requires infinitely many trials. Any practical problem deals with a finite, sometimes small, number of trials. Thus, it is impossible to verify the convergence and stability of frequencies in the form of a limit. It remains to imagine the lacking experiments, which may lead to inadequate conclusions.

Various attempts to overcome this difficulty have been made, both in the framework of classical probability [6, 7] as and in other probability theories [8, 9], [13, 14]. One of such directions is soft probability theory [9]-[11], which is based on the notion of a soft set [12]. The counterpart of mathematical expectation in this theory is a family of sliding means. Clearly, averaging alone cannot lead to any instructive interesting results. Analyzing these means, we can accept various hypotheses on the future behavior variables. Of interest is to find out what constraints on the future behavior of variables are imposed by various hypotheses. This paper is devoted to one of such problems.

We consider the hypotheses of the stability of sliding means (which is the counterpart of the existence of mathematical expectation) and of the finiteness of mean spread (which is the counterpart of the existence of variance). Assuming that these hypotheses are true, we study the soft probability of the deviation of the arithmetic mean of trial results from the counterpart of mathematical expectation. In essence, the statement of the problem is very close to that of

[1]A.A. Dorodnitsyn Computational Center, Russian Academy of Sciences, Moscow, Russia,
e-mail: mold@ccas.ru, dmitri_molodtsov@mail.ru

Manuscript received November 2012.

the central limit theorem in classical probability theory; however, there are some differences. The main difference is that the trial results are not assumed to be independent.

Since the research apparatus differs substantially from that in the classical case, we give rather detailed proof, although the apparatus itself is quite elementary.

2. Notation and conventions

The initial data set to be dealt with is a sequence of real numbers $(x_1, \dots, x_n)$, where $x_i \in E$ and E denotes the set of real numbers. We can regard this sequence as an element of the nth power of the set E, that is, assume that $(x_1, \dots, x_n) \in E^n$; however, this is insufficient for our purposes.

The point is that we need to consider various subsequences of the initial sequence $(x_1, \dots, x_n)$ and construct other subsequences from them; thus, in addition to the space E^n, we use mappings of the form $\Phi(I) = \{f : I \to E\}$, where $I \subseteq \{1, \dots, n\}$. We denote values of mappings of this type by using two arguments as $f(I, i)$, where $I \subseteq \{1, \dots, n\}$ and $i \in I$. The notation $f(I, \cdot)$ stands for the vector

$$f(I, \cdot) = (f(I, i_1), f(I, i_2), \dots, f(I, i_k)) \in E^{|I|}, \tag{1}$$

where $(i_1, i_2, \dots, i_k) = I$, the indices $i_1, i_2, \dots, i_k$ are arranged in increasing order, and $|A|$ denotes the cardinality of a set A. The functions in the set $\Phi(I)$ are called f-vectors. Given a subset $(x_1, \dots, x_n) = I \subseteq \{1, \dots, n\}$, where $i_1 \le i_2 \le \dots \le i_k$ and $(x_1, \dots, x_{|I|}) \in E^{|I|}$, we define an f-vector $f[I, x] \in \Phi(I)$ by $f[I, x](I, i_j) = x_j$. Obviously, $f[I, x](I, \cdot) = x$.

For subsets of the form $\{k, k+1, \dots, m\} \subseteq \{1, \dots, n\}$, we introduce the special notation $\{k, \dots, m\} = I_m^k$.

Given a vector $x \in E^n$, we set $x_k^m = (x_k, x_{k+1}, \dots, x_m)$.

We introduce the function $S(x) = \sum\limits_{i=1}^{m} x_i$, where m is any positive integer and $x \in E^m$ is any vector. Given $J \subseteq \{1, \dots, n\}$, by $E(J)$ we denote the subspace of E spanned by the axes with indices belonging to J. If $I, J \subseteq \{1, \dots, n\}$, $I \cap J = \emptyset$, and $f \in \Phi(I)$ and $h \in \Phi(J)$ are two f-vectors, then the join $f \oplus h \in \Phi(I \cup J)$ of these f-vectors is defined by

$$f \oplus h(I \cup J, i) = \begin{cases} f(I, i), & i \in I; \\ h(J, i), & i \in J. \end{cases}$$

The section of a set $W \subseteq E^n$ by an f-vector $h \in \Phi(J)$, where $J \subset \{1, \dots, n\}$, is defined by

$$Sec[W, J, H] = \{f \in \Phi(\{1, \dots, n\} \backslash J) | f \oplus h(\{1, \dots, n\}, \cdot)\}.$$

The projection of a set $W \subseteq E^n$ on a subspace $\Phi(I)$ is

$$Pr[W, J] = \{h \in \Phi(J) | Sec[W, J, h] \neq \emptyset\}.$$

We refer to a sequence $f^n = \{f_1^n, f_2^n, \dots, f_n^n\}$, where $f_k^n : E^k \to E$, as a function with memory of dimension n, or simply as a function with n-memory. The soft mean of order m for a function with memory f^n of dimension n is, by definition, the interval function

$$\mu(f^n, x_1^n, m) = [\underline{\mu}(f^n, x_1^n, m) \bar{\mu}(f^n, x_1^n, m)],$$

where

$$\lambda_j^m(f^n, x_1^{j+m-1}) = \frac{1}{m} \sum_{i=j}^{j+m-1} f_i^n(x_1^i),$$

$$\underline{\mu}(f^n, x_1^n, m) = \min_{1 \le j \le n-m+1} \lambda_j^m(f^n, x_1^{j+m-1}), \bar{\mu}(f^n, x_1^n, m) = \max_{1 \le j \le n-m+1} \lambda_j^m(f^n, x_1^{j+m-1}).$$

3. Statement of the problem

Consider the function

$$g^n(x,a,\delta,\varepsilon)=\{g_1^n(x_1^n,a,\delta,\varepsilon),\dots,g_k^n(x_1^k,a,\delta,\varepsilon),\dots,g_n^n(x_1^n,a,\delta,\varepsilon)\},$$

with n-memory defined as follows: given $k=\{1,\dots,n\}$ and $\varepsilon>0$,

$$g_k^n(x_1^k,a,\delta,\varepsilon)=1 \text{ if } S(x_1^k)\geq k(a+\delta)+\varepsilon\varphi(k),$$
$$g_k^n(x_1^k,a,\delta,\varepsilon)=0 \text{ if } S(x_1^k)< k(a+\delta)+\varepsilon\varphi(k).$$

Here $a,\delta,\varepsilon\in E$, $\delta\geq 0$, $\varepsilon>0$, and φ is a strictly increasing function of a positive integer argument taking positive real values. In essence, the function g^n is the characteristic vector function of the event

$$S(x_1^k)\geq k(a+\delta)+\varepsilon\varphi(k), k=\{1,\dots,n\}.$$

Note that the function introduced above that the obvious property $g_k^n(x_1^k,a,\delta,\varepsilon)=g_k^k(x_1^k,a,\delta,\varepsilon)$ for $k\leq n$.

We assume that the given sequence $(x_1,\dots,x_n)$ satisfies the following two conditions (hypotheses):

- The mean hypothesis $HM(x_1^n,a,m,\delta)$ means that

$$|\frac{1}{m}S(x_j^{j+m-1})-a|\leq\delta \qquad \text{for} \quad j=1,\dots,n-m+1.$$

- The spread hypothesis (variance) $HD(x_1^n,a,m,\delta,\Delta)$ means that

$$\frac{1}{m}\sum_{i=j}^{j+m-1}\max\{|x_i-a|-\delta,0\}\leq\Delta \qquad \text{for} \quad j=1,\dots,n-m+1.$$

We denote the set of sequences $x_1^n=(x_1,\dots,x_n)$ satisfying the hypotheses $HM(x_1^n,a,m,\delta)$ and $HD(x_1^n,a,m,\delta,\Delta)$ by $X(n,m,a,\delta,\Delta)$; thus,

$$X(n,m,a,\delta,\Delta)=\{x_1^n||\tfrac{1}{m}S(x_j^{j+m-1})-a|\leq\delta,$$
$$\tfrac{1}{m}\sum_{i=j}^{j+m-1}\max\{|x_i-a|-\delta,0\}\leq\Delta, j=1,\dots,n-m+1\}.$$

It is easy to see that if $x_1^n\in X(n,m,a,\delta,\Delta)$, then $x_1^k\in X(k,m,a,\delta,\Delta)$ for $k=m,\dots,n-1$.

The soft mean of order m for a function g^n is the counterpart of probability.

Problem. Determine the range of variation of $\mu(g^n(\cdot,a,\delta,\varepsilon),x_1^n,m)$ as a function of the argument x_1^n on the set $X(n,m,a,\delta,\Delta)$ at fixed values of the parameters $a,\delta,\varepsilon\in E,\delta\geq 0,\varepsilon>0$ for a given function φ.

To solve this problem, we must find the two values

$$\max_{x_1^n\in X(n,m,a,\delta,\Delta)}\bar{\mu}(g^n(\cdot,a,\delta,\varepsilon),x_1^n,m)$$

and

$$\min_{x_1^n\in X(n,m,a,\delta,\Delta)}\underline{\mu}(g^n(\cdot,a,\delta,\varepsilon),x_1^n,m)$$

The latter value is determined in a trivial way. Obviously, $a_1^n=(a,\dots,a)\in X(n,m,a,\delta,\Delta)$, whence $g_k^n(a_1^k,a,\delta,\varepsilon)=0$ for $k=1,\dots,n$. Therefore, $\min\limits_{x_1^n\in X(n,m,a,\delta,\Delta)}\underline{\mu}(g^n(\cdot,a,\delta,\varepsilon),x_1^n,m)=0$.

Let us calculate the first value.

4. Calculation the upper soft probability

Note that

$$\max_{x_1^n \in X(n,m,a,\delta,\Delta)} \bar{\mu}(g^n(\bullet, a, \delta, \varepsilon), x_1^n, m) = \max_{1 \le j \le n-m+1} \bar{\theta}(g^n(\cdot, a, \delta, \varepsilon), j, m),$$

where $\bar{\theta}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \max_{x_1^n \in X(n,m,a,\delta,\Delta)} \lambda_j^m(g^n(\cdot, a, \delta, \varepsilon), x_m^{j+m-1})$.

Since the function g^n takes only two values, it follows that the function $\lambda_j^m(g^n(\cdot, a, \delta, \varepsilon), x_m^{j+m-1})$ takes finitely many values; therefore, it attains its maximum and minimum values on the set $X(n, m, a, \delta, \Delta)$. Let us describe some properties of the set $X(n, m, a, \delta, \Delta)$.

Statement 4.1. *If $J = \{1, \dots, k\}$ and $m \le k < n$, then*

$$\bigcup_{f \in Pr[X(n,m,a,\delta,\Delta),J]} f(J, \cdot) = X(k, m, a, \delta, \Delta).$$

Proof. Take a vector $x_1^k \in X(k, m, a, \delta, \Delta)$ and the corresponding f-vector $f[J, x_1^k]$. Consider another f-vector $h \in \Phi(\{1, \dots, n\} \backslash J)$ defined by $h(\{1, \dots, n\} \backslash J, k + i) = x_{k-m+(i \mod m)}$.

Joining $f[J, x]$ with the f-vector $h \in E(\{1, \dots, n\} \backslash J)$, we obtain $f[J, x_1^k] \oplus h \in \Phi(\{1, \dots, n\})$. It is easy to see that $f[J, x_1^k] \oplus h(\{1, \dots, n\}, \cdot) \in X(n, m, a, \delta, \Delta)$. It follows that $Sec[X(n, m, a, \delta, \Delta), J, f[J, x_1^k]] \ne \emptyset$; therefore,

$$X(k, m, a, \delta, \Delta) \subseteq \bigcup_{f \in Pr[X(n,m,a,\delta,\Delta),J]} f(J, \cdot).$$

Let $f \in Pr[X(n, m, a, \delta, \Delta), J]$. This means that $f \in \Phi(J)$ and there exists an $h \in \Phi(\{1, \dots, n\} \backslash J)$ for which $f \oplus h(\{1, \dots, n\}, \cdot) \in X(n, m, a, \delta, \Delta)$. Thus, $f[J, \cdot] \in X(k, m, a, \delta, \Delta)$, that is,

$$X(k, m, a, \delta, \Delta) \supseteq \bigcup_{f \in Pr[X(n,m,a,\delta,\Delta),J]} f(J, \cdot).$$

This proves the required assertion. □

Statement 4.2. *If $m \le k < n$ and $1 \le j \le n - m + 1$, then*

$$\bar{\theta}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \bar{\theta}(g^{j+m-1}(\cdot, a, \delta, \varepsilon), j, m).$$

Proof. By definition,

$$\lambda_j^m(g^n(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}) = \frac{1}{m} \sum_{i=j}^{j+m-1} g_i^n(x_1^i, a, \delta, \varepsilon) =$$

$$= \frac{1}{m} \sum_{i=j}^{j+m-1} g_i^{j+m-1}(x_1^i, a, \delta, \varepsilon) = \lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}).$$

Therefore,

$$\bar{\delta}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \max_{x_1^n \in X(n,m,a,\delta,\Delta)} \lambda_j^m(g^n(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}) =$$

$$= \max_{x_1^n \in X(n,m,a,\delta,\Delta)} \lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}).$$

Since the function $\lambda_j^m(g^{j+m-1}(\cdot,a,\delta,\varepsilon),x_1^{j+m-1})$ does not depend on the components of the vector x_1^n with indices larger than $j+m-1$, it follows that

$$\max_{x_1^n\in X(n,m,a,\delta,\Delta)} \lambda_j^m(g^{j+m-1}(\cdot,a,\delta,\varepsilon),x_1^{j+m-1}) =$$

$$= \max_{f\in Pr[X(n,m,a,\delta,\Delta),\{1,\dots,j+m-1\}]} \lambda_j^m(g^{j+m-1}(\cdot,a,\delta,\varepsilon),f(\{1,\dots,j+m-1\},\cdot)) =$$

$$= \max_{x_1^{j+m-1}\in X(j+m-1,m,a,\delta,\Delta)} \lambda_j^m(g^{j+m-1}(\cdot,a,\delta,\varepsilon),x_1^{j+m-1}) =$$

$$= \bar\theta(g^{j+m-1}(\cdot,a,\delta,\varepsilon),j,m).$$

□

The required assertion follows.

Statement 4.3. *If* $1\le j\le n-m+1$ *and*

$$\vartheta(j)=\begin{cases} j \mod m, & j \mod m>0,\\ m, & j \mod m=0,\end{cases}$$

then

$$\bar\theta(g^{j+m-1}(\cdot,a,\delta,\varepsilon),j,m)\le \bar\theta(g^{\vartheta(j)+m-1}(\cdot,a,\delta,\varepsilon),\vartheta(j),m).$$

Proof. Take any vector $x_1^{j+m-1}\in X(j+m-1,m,a,\delta,\Delta)$.

Consider the f-vector

$$h=f[\{j-\vartheta(j)+1,\dots,j+m-1\},x_{j-\vartheta(j)+1}^{j+m-1}]\in\Phi(\{j-\vartheta(j)+1,\dots,j+m-1\}).$$

Obviously, $y=h(\{j-\vartheta(j)+1,\dots,j+m-1\},\cdot)\in X(m+\vartheta(j)-1,m,a,\delta,\Delta)$. It is sufficient to prove that

$$\lambda_j^m(g^{j+m-1}(\cdot,a,\delta,\varepsilon),x_1^{j+m-1})\le\lambda_{\vartheta(j)}^m(g^{\vartheta(j)+m-1}(\cdot,a,\delta,\varepsilon),y).$$

For this purpose, it suffices to show that, for any $k=0,1,\dots,m-1$, we have

$$g_{j+k}^{j+m-1}(x_1^{j+k},a,\delta,\varepsilon)\le g_{\vartheta(j)+k}^{\vartheta(j)+m-1}(y_1^{\vartheta(j)+k},a,\delta,\varepsilon).$$

Moreover, it suffices to prove this inequality only for those $k=0,1,\dots,m-1$ which satisfy the condition $g_{j+k}^{j+m-1}(x_1^{j+k},a,\delta,\varepsilon)=1$; for other k, the required inequality is obvious. This condition means that $S(x_1^{j+k})\ge(j+k)(a+\delta)+\varepsilon\varphi(j+k)$. By assumption, we have $1\le j\le n-m+1$, whence $j\ge\vartheta(j)$. Obviously, in the case $j=\vartheta(j)$, the required assertion is valid. Consider the case where $j>\vartheta(j)$. Let us decompose the sum $S(x_1^{j+k})$ into two parts as

$$S(x_1^{j+k})=S(x_1^{j-\vartheta(j)})+S(x_{j-\vartheta(j)+1}^{j+k})$$

By the construction of the vector y, we have $S(x_{j-\vartheta(j)+1}^{j+k})=S(y_1^{k+\vartheta(j)})$, and since $j-\vartheta(j)$ is a multiple of m, we have

$$S(x_1^{j-\vartheta(j)})\le(j-\vartheta(j))(a+\delta).$$

This implies

$$(j-\vartheta(j))(a+\delta)+S(y_1^{k+\vartheta(j)})\ge S(x_1^{j-\vartheta(j)})+S(y_1^{k+\vartheta(j)})=S(x_1^{j+k})\ge(j+k)(a+\delta)+\varepsilon\varphi(j+k).$$

Elementary transformations yield

$$S(y_1^{k+\vartheta(j)})\ge(k+\vartheta(j))(a+\delta)+\varepsilon\varphi(j+k).$$

Since the function φ increases, we obtain

$$S(y_1^{k+\vartheta(j)}) \geq (k+\vartheta(j))(a+\delta) + \varepsilon\varphi(k+\vartheta(j)),$$

which means that $g_{\vartheta(j)+k}^{\vartheta(j)+m-1}(y_1^{\vartheta(j)+k}, a, \delta, \varepsilon) = g_{\vartheta(j)+k}^{\vartheta(j)+k}(y_1^{\vartheta(j)+k}, a, \delta, \varepsilon) = 1.$ □

Statement 4.4. *The following relation holds:*

$$\max_{x_1^n \in X(n,m,a,\delta,\Delta)} \bar{\mu}(g^n(\cdot, a, \delta, \varepsilon), x_1^n, m) = \max_{1\leq j\leq m} \bar{\theta}(g^n(\cdot, a, \delta, \varepsilon), j, m),$$

This assertion follows from Statement 3.

Together with the set $X(n, m, a, \delta, \Delta)$, we consider its modification, which is determined by fewer constraints:

$$Y(n, m, a, \delta, \Delta) = \{x_1^n \big| |\tfrac{1}{m} S(x_j^{j+m-1}) - a| \leq \delta,$$
$$\tfrac{1}{m}\sum_{i=j}^{j+m-1} \max\{|x_i - a| - \delta, 0\} \leq \Delta, j = \overline{1, n-m+1}\}.$$

Obviously, $X(n, m, a, \delta, \Delta) \subseteq Y(n, m, a, \delta, \Delta)$, and $X(m, m, a, \delta, \Delta) = Y(m, m, a, \delta, \Delta)$.

Consider also the function

$$\bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \max_{x_1^{j+m-1}\in Y(j+m-1,m,a,\delta,\Delta)} \lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}),$$

which is an analogue of $\bar{\theta}(g^n(\cdot, a, \delta, \varepsilon), j, m)$.

Obviously, we have

$$\bar{\theta}(g^n(\cdot, a, \delta, \varepsilon), j, m) \leq \bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m)$$

for $j = 2, \dots, n-m+1$ and

$$\bar{\theta}(g^n(\cdot, a, \delta, \varepsilon), 1, m) = \bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), 1, m).$$

We also introduce the solution set of the problem $\bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m)$:

$$\Psi(g^n(\cdot, a, \delta, \varepsilon), j, m) = \{x_1^{j+m-1} \in Y(j+m-1, m, a, \delta, \Delta)|,$$
$$\lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}) = \bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m)\}.$$

Statement 4.5. *If $m \geq j > 1$, then there exists an $x_1^{j+m-1} \in \Psi(g^n(\cdot, a, \delta, \varepsilon), j, m)$ for which $S(x_1^{j-1}) \leq S(x_{m+1}^{j+m-1})$.*

Proof. First, let us prove that the set $\Psi(g^n(\cdot, a, \delta, \varepsilon), j, m)$ is compact. Note that the set $Y(j+m-1, m, a, \delta, \Delta)$ is compact. The condition

$$\lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}) = \bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m)$$

is equivalent to

$$\lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}) \geq \bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m).$$

Since the function $g^n(\cdot, a, \delta, \varepsilon)$ takes only the values 0 and 1, there exists a nonnegative integer k such that $\bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \frac{k}{m}$. We set

$$Y^*(j+m-1, m, a, \delta, \Delta, J) = \{x_1^{j+m-1} \in Y(j+m-1, m, a, \delta, \Delta)|g_j^j(x_1^j, a, \delta, \varepsilon) = 1, j \in J\}.$$

By the definition of the function g^n, we have

$$Y^*(j+m-1, m, a, \delta, \Delta, J) = \{x_1^{j+m-1} \in Y(j+m-1, m, a, \delta, \Delta)|S(x_1^j) \geq j(a+\delta)+\varepsilon\varphi(j), j \in J\},$$

which implies the compactness of $Y^*(j+m-1, m, a, \delta, \Delta, J)$.

Since

$$\Psi(g^n(\cdot,a,\delta,\varepsilon),j,m) = \bigcup_{\substack{J\subset\{1,\dots,j+m-1\}\\ |J|=k}} Y^*(j+m-1,m,a,\delta,\Delta,J),$$

it follows that the set $\Psi(g^n(\cdot,a,\delta,\varepsilon),j,m)$ is compact as well.

Choose $x_1^{j+m-1} \in \Psi(g^n(\cdot,a,\delta,\varepsilon),j,m)$ so that $S(x_{m+1}^{j+m-1}) \geq S(y_{m+1}^{j+m-1})$ for any $y_1^{j+m-1} \in \Psi(g^n(\cdot,a,\delta,\varepsilon),j,m)$. We can do this because of the continuity of the function S and the compactness of the set Ψ.

If both conditions

$$|\frac{1}{m}S(x_j^{j+m-1}) - a| \leq \delta \quad \text{and} \quad \frac{1}{m}\sum_{i=j}^{j+m-1} \max\{|x_i - a| - \delta, 0\} \leq \Delta$$

in the definition of the set $Y(j+m-1,m,a,\delta,\Delta)$ hold as strict inequalities, then we can slightly increase the component x_{j+m-1} without violating these conditions; this means that the sum $S(x_{m+1}^{j+m-1})$ is not maximum. Therefore, at least one of the conditions on the vector x_1^{j+m-1} holds as an equality.

Let $|\frac{1}{m}S(x_j^{j+m-1}) - a| = \delta$. It is easy to show that the cases $\frac{1}{m}S(x_j^{j+m-1}) - a = -\delta$ and $\delta > 0$ are impossible too.

If $\frac{1}{m}S(x_j^{j+m-1}) - a = \delta$, then we obtain

$$\frac{1}{m}S(x_1^m) - a \leq |\frac{1}{m}S(x_1^m) - a| \leq \delta = \frac{1}{m}S(x_j^{j+m-1}) - a,$$

or $S(x_1^m) \leq S(x_j^{j+m-1})$.

It follows that

$$S(x_1^m) = S(x_1^{j-1}) + S(x_j^m) \leq S(x_j^{j+m-1}) = S(x_j^m) + S(x_{m+1}^{j+m-1}),$$

that is, $S(x_1^{j-1}) \leq S(x_{m+1}^{j+m-1})$.

Now, consider the case where the vector satisfies the following constraint (as an equality):

$$\frac{1}{m}\sum_{i=j}^{j+m-1} \max\{|x_i - a| - \delta, 0\} = \Delta.$$

For $l = m+1,\dots,j+m-1$, we have holds $|x_i - a| - \delta = x_i - a - \delta \geq 0$. Indeed, otherwise, we could slightly increase the corresponding component x_i, and the sum $S(x_{m+1}^{j+m-1})$ would not be maximum. Therefore,

$$\frac{1}{m}\sum_{i=j}^{m} \max\{|x_i - a| - \delta, 0\} + \frac{1}{m}\sum_{i=m+1}^{j+m-1}(x_i - a - \delta) = \Delta m.$$

Hence we have

$$\sum_{i=j}^{m} \max\{|x_i - a| - \delta, 0\} + \sum_{i=m+1}^{j+m-1}(x_i - a - \delta) =$$

$$= \Delta m \geq \sum_{i=1}^{m} \max\{|x_i - a| - \delta, 0\} = \sum_{i=1}^{j-1} \max\{|x_i - a| - \delta, 0\} + \sum_{i=j}^{m} \max\{|x_i - a| - \delta, 0\} \geq$$

$$\geq \sum_{i=1}^{j-1}(|x_i - a| - \delta) + \sum_{i=j}^{m} \max\{|x_i - a| - \delta, 0\} \geq \sum_{i=1}^{j-1}(x_i - a - \delta) + \sum_{i=j}^{m} \max\{x_i - a - \delta, 0\}.$$

It follows that

$$\sum_{i=m+1}^{j+m-1}(x_i - a - \delta) \geq \sum_{i=1}^{j-1}(x_i - a - \delta),$$

or

$$S(x_{m+1}^{j+m-1}) = \sum_{i=m+1}^{j+m-1} x_i \geq \sum_{i=1}^{j-1} x_i = S(x_1^{j-1}).$$

□

This completes the proof of the statement.

Statement 4.6. *The following relation holds:*

$$\max_{1\leq j\leq m} \bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), 1, m).$$

Proof. Let $m \geq j > 1$. Consider the problem

$$\bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \max_{x_1^{j+m-1}\in Y(j+m-1,m,a,\delta,\Delta)} \lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}).$$

According to Statement 5, we can choose a vector $x_1^{j+m-1} \in \Psi(g^n(\cdot, a, \delta, \varepsilon), j, m)$ for which $S(x_1^{j-1}) \leq S(x_{m+1}^{j+m-1})$. Recall that

$$\bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}).$$

Let us partition the index set $\{1, \ldots, j + m - 1\}$ into the three disjoint sets $J_1 = \{1, \ldots, j - 1\}$, $J_2 = \{j, \ldots, m\}$, and $J_3 = \{m + 1, \ldots, j + m - 1\}$. Note that the sets J_1 and J_3 have the same cardinality.

Let $y = x_{m+1}^{j+m-1}$; then $y_1 = x_{m+1}, \ldots, y_{j-1} = x_{j+m-1}$. Let us prove that

$$g_{m+k}^{j+m-1}(x_1^{m+k}, a, \delta, \varepsilon) \leq g_k^k(y_1^k, a, \delta, \varepsilon)$$

for any $k \in J_1$. It suffices to consider only those indices $k \in J_1$ for which

$$g_{m+k}^{j+m-1}(x_1^{m+k}, a, \delta, \varepsilon) = 1,$$

that is, $S(x_1^{m+k}) \geq (m + k)(a + \delta) + \varepsilon\varphi(m + k)$. We decompose the sum $S(x_1^{m+k})$ into two parts as $S(x_1^{m+k}) = S(x_1^m) + S(x_{m+1}^{m+k}))$. Since $y = x_{m+1}^{j+m-1}$, we have $S(x_{m+1}^{m+k}) = S(y_1^k))$. It follows from the condition $x_1^{j+m-1} \in Y(j + m - 1, m, a, \delta, \Delta)$ that $|\frac{1}{m}S(x_1^m) - a| \leq \delta$, whence $S(x_1^m) \leq m(a + \delta)$. Therefore, we have the chain of inequalities

$$m(a + \delta) + S(y_1^m) \geq S(x_1^m) + S(x_{m+1}^{m+k}) \geq (m + k)(a + \delta) + \varepsilon\varphi(m + k).$$

Performing elementary transformations and taking into account the function φ being increasing, we obtain

$$S(y_1^m) \geq k(a + \delta) + \varepsilon\varphi(m + k) \geq k(a + \delta) + \varepsilon\varphi(k).$$

This inequality means that $g_k^k(y_1^k, a, \delta, \varepsilon) = 1$. Thus, we have proved the inequality

$$g_{m+k}^{j+m-1}(x_1^{m+k}, a, \delta, \varepsilon) \leq g_k^k(y_1^k, a, \delta, \varepsilon)$$

for $k = 1, \ldots, j - 1$.

Let $z = x_j^m$ and consider the vector $w = (f[J_1, y] \oplus f[J_2, z])(\{1, \dots, m\}, \cdot)$. Simply speaking, the vector w is the concatenation of the vectors y and z. The construction of w is schematically shown in Figure 1.

$$
\begin{array}{l}
x = \boxed{1, \dots, j-1,} \quad \overset{z}{\boxed{j, \dots, m}} \quad \overset{y}{\boxed{m+1, \dots, j+m-1}} \\
w = (y, z) \\
w = \boxed{m+1, \dots, j+m-1} \quad \boxed{j, \dots, m}
\end{array}
$$

Fig. 1.

It is easy to see that $w \in Y(m, m, a, \delta, \Delta)$.

Let us prove that

$$g_{j+k}^{j+m-1}(x_1^{j+k}, a, \delta, \varepsilon) = g_{j+k}^{j+k}(x_1^{j+k}, a, \delta, \varepsilon) \le g_{j+k}^{j+k}(w_1^{j+k}, a, \delta, \varepsilon)$$

for $k = 0. \dots, m-j$. These inequalities follow from the estimates

$$S(x_1^{j+k}) = S(x_1^{j-1}) + S(x_j^{j+k}) \le S(x_{m+1}^{j+m-1}) + S(x_j^{j+k}) = S(y_1^{j-1}) + S(z_1^{1+k}) = S(w_1^{j+k}),$$

which use the inequality $S(x_1^{j-1}) \le S(x_{m+1}^{j+m-1})$. Thus, we have shown that

$$\sum_{k=0}^{m-1} g_{j+k}^{j+m-1}(x_1^{j+k}, a, \delta, \varepsilon) \le \sum_{k=0}^{m-1} g_{1+k}^{m}(w_1^{1+k}, a, \delta, \varepsilon).$$

This means that

$$\bar{\psi}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \lambda_j^m(g^{j+m-1}(\cdot, a, \delta, \varepsilon), x_1^{j+m-1}) \le \lambda_j^m(g^m(\cdot, a, \delta, \varepsilon), w_1^m),$$

which proves the required assertion. $\square$

The following statement follows directly from Statement 6.

Statement 4.7. *The following relation holds:* $\max\limits_{1 \le j \le m} \bar{\theta}(g^n(\cdot, a, \delta, \varepsilon), j, m) = \bar{\theta}(g^m(\cdot, a, \delta, \varepsilon), 1, m)$.

Now, let us calculate $\bar{\theta}(g^m(\cdot, a, \delta, \varepsilon), 1, m)$. Recall that

$$\bar{\theta}(g^m(\cdot, a, \delta, \varepsilon), 1, m) = \max_{x_1^m \in X(m,m,a,\delta,\Delta)} \lambda_1^m(g^m(\cdot, a, \delta, \varepsilon), x_1^m)$$

$$X(m, m, a, \delta, \Delta) = \{x_1^m \in E^m \big| |\frac{1}{m} S(x_1^m) - a| \le \delta, \frac{1}{m} \sum_{i=1}^{m} \max\{|x_i - a| - \delta, 0\} \le \Delta\}.$$

Statement 4.8. *If $x_1^m \in X(m, m, a, \delta, \Delta)$ and there exists an index $k \in 1, \dots, m-1$ for which $g_k^m(x_1^k, a, \delta, \varepsilon) = 0$ and $g_{k+1}^m(x_1^{k+1}, a, \delta, \varepsilon) = 1$, then there exists a vector $y_1^m \in X(m, m, a, \delta, \Delta)$ for which $\lambda_1^m(g^m(\cdot, a, \delta, \varepsilon), x_1^m) < \lambda_1^m(g^m(\cdot, a, \delta, \varepsilon), y_1^m)$.*

Proof. The conditions $g_k^m(x_1^k, a, \delta, \varepsilon) = 0$ and $g_{k+1}^m(x_1^{k+1}, a, \delta, \varepsilon) = 1$ imply the relations

$$S(x_1^k) < k(a + \delta) + \varepsilon\varphi(k),$$

$$S(x_1^k) + x_{k+1} \ge (k+1)(a + \delta) + \varepsilon\varphi(k+1).$$

Rewriting these relations in the form

$$S(x_1^k) = k(a + \delta) + \varepsilon\varphi(k) - \beta, \beta > 0,$$

$$S(x_1^k) + x_{k+1} = (k+1)(a + \delta) + \varepsilon\varphi(k+1) + \gamma, \gamma \ge 0,$$

we see that

$$x_{k+1} = a + \delta + \varepsilon[\varphi(k+1) - \varphi(k)] + \gamma + \beta.$$

Let $y_1^m \in E^m$ be the vector in which $y_k = x_k + \beta$, $y_{k+1} = x_{k+1} - \beta$, and all of the remaining components y_1^m are the same as in the vector x_1^m. It follows from the definition of y_1^m that $S(x_1^j) = S(y_1^j)$, $j \in \{1, \dots, m\} \backslash \{k\}$, and $S(y_1^k) = S(x_1^k) + \beta = k(a+\delta) + \varepsilon\varphi(k)$. This means that $g_j^m(x_1^j, a, \delta, \varepsilon) = g_j^m(y_1^j, a, \delta, \varepsilon)$ for $j \in \{1, \dots, m\} \backslash \{k\}$ and $g_k^m(y_1^k, a, \delta, \varepsilon) = 1$. It remains to verify that $y_1^m \in X(m, m, a, \delta, \Delta)$. The condition $|\frac{1}{m}S(x_1^m) - a| \le \delta$ does hold, because $S(y_1^m) = S(x_1^m)$. To verify the second condition $\sum\limits_{i=1}^{m} \max\{|x_i - a| - \delta, 0\} \le \Delta m$, it suffices to consider only the terms corresponding to the indices k and $k+1$. We have the obvious estimates

$$\max\{|x_{k+1} - a| - \delta, 0\} = \varepsilon[\varphi(k+1) - \varphi(k)] + \gamma + \beta,$$
$$\max\{|y_{k+1} - a| - \delta, 0\} = \varepsilon[\varphi(k+1) - \varphi(k)] + \gamma = \max\{|x_{k+1} - a| - \delta, 0\} - \beta,$$
$$\max\{|y_k - a| - \delta, 0\} = \max\{|x_k + \beta - a| - \delta, 0\} \le \max\{|x_k - a| + \beta - \delta, 0\} \le$$
$$\le \max\{|x_k - a| - \delta, -\beta\} + \beta \le \max\{|x_k - a| - \delta, 0\} + \beta.$$

It follows that

$$\max\{|y_k - a| - \delta, 0\} + \max\{|y_{k+1} - a| - \delta, 0\} \le \max\{|x_k - a| - \delta, 0\} + \beta + \max\{|x_{k+1} - a| - \delta, 0\} - \beta.$$

□

According to Statement 8, an optimal solution of the problem $\bar{\theta}(g^m(\cdot, a, \delta, \varepsilon), 1, m)$ is to be sought among those vectors $x_1^m \in X(m, m, a, \delta, \Delta)$ for which there exists a $k \in \{1, \dots, m-1\}$ such that $g_j^m(x_1^j, a, \delta, \varepsilon) = 1$ for any $j \in \{1, \dots, k\}$. Solving the maximization problem for this k, we obtain a solution of the initial problem. Seeking a solution requires knowing solvability conditions for some elementary systems of inequalities.

Statement 4.9. *The system of inequalities*

$$A \le \sum_{i=1}^{m-k} y_i \le B, \sum_{i=1}^{m-k} \max\{|y_i - a| - \delta, 0\} \le C$$

with respect to a vector $y \in E^{m-k}$ is solvable if and only if

$$(m-k)(a-\delta) - C \le B \quad \textit{and} \quad A \le (m-k)(a+\delta) + C.$$

Proof. It is easy to see that the rang of $\sum\limits_{i=1}^{m-k} y_i$, when the vector y satisfies the inequality $\sum\limits_{i=1}^{m-k} \max\{|y_i - a| - \delta, 0\} \le C$ is the interval

$$[(m-k)(a-\delta) - C, (m-k)(a+\delta) + C].$$

Therefore, the system is solvable if and only if the required inequalities hold, q.e.d.

First, we solve first the auxiliary problem of determining conditions on the first k components of a vector $x \in E^m$ under which this vector can be extended to a vector belonging to the set $X(m, m, a, \delta, \Delta)$, that is, finding the projection $Pr[X(m, m, a, \delta, \Delta), \{1, \dots, k\}]$. For convenience, we set $y = x_{k+1}^m \in E^{m-k}$. Calculating the projection reduces to determining necessary and sufficient conditions for the solvability of the following system of three inequalities with respect to y:

$$m(a-\delta) \le S(x_1^k) + S(y_1^{m-k}) \le m(a+\delta)$$
$$\sum_{i=1}^{k} \max\{|x_i - a| - \delta, 0\} + \sum_{i=1}^{m-k} \max\{|y_i - a| - \delta, 0\} \le \Delta m.$$

Applying Statement 9, we see that the required projection is described by the inequalities

$$\sum_{i=1}^{k}\max\{|x_i-a|-\delta,0\}+S(x_1^k)\le m(a+\delta)+\Delta m-(m-k)(a-\delta),$$

$$\sum_{i=1}^{k}\max\{|x_i-a|-\delta,0\}-S(x_1^k)\le -m(a-\delta)+\Delta m+(m-k)(a+\delta).$$

These two inequalities can be rewritten in the equivalent form

$$\sum_{i=1}^{k}\max\{|x_i-a|-\delta,0\}+|S(x_1^k)-ka|\le m(2\delta+\Delta)-k\delta.$$

Now, it remains to require that $g_j^m(x_1^j,a,\delta,\varepsilon)=1$ for any $j\in\{1,\dots,k\}$. These conditions mean that

$$S(x_1^j)=j(a+\delta)+\varepsilon\varphi(j)+u_i,\quad u_i\ge 0,\quad j\in\{1,\dots,k\}.$$

The last relations can be treated as the change of variables

$$x_1=a+\delta+\varepsilon\varphi(1)+u_1,\quad x_j=S(x_1^j)-S(x_1^{j-1})=a+\delta+\varepsilon[\varphi(j)-\varphi(j-1)]+u_j-u_{j-1},j\in\{2,\dots,k\},$$

which reduces the problem to the solvability of the system of inequalities

$$u_1+\sum_{j=2}^{k}\max\{|\delta+\varepsilon[\varphi(j)-\varphi(j-1)]+u_j-u_{j-1}|-\delta,0\}+u_k\le$$
$$\le m(2\delta+\Delta)-2k\delta-\varepsilon[\varphi(k)+\varphi(1)],u_i\ge 0,i\in\{1,\dots,k\}.$$

A solvability condition for this system is the inequality

$$\min_{\substack{u_j\ge 0\\ j=1,\dots,k}}\left\{u_1+\sum_{j=2}^{k}\max\{|\delta+\varepsilon[\varphi(j)-\varphi(j-1)]+u_j-u_{j-1}|-\delta,0\}+u_k\right\}\le$$
$$\le m(2\delta+\Delta)-2k\delta-\varepsilon[\varphi(k)+\varphi(1)].$$

We write min rather than inf because, as we shall prove in what follows, the minimum is attained. Since the variables change independently, we minimize the function with respect to its variables successively, starting with the variable u_1. This variable is contained only in the first two terms, and the function to be optimized has the form

$$G_1(u_1)=u_1+\max\{|\delta+\varepsilon[\varphi(2)-\varphi(1)]+u_2-u_1|-\delta,0\}.$$

This is a piecewise linear function of the form

$$G_1(u_1)=\varepsilon(\varphi(2)-\varphi(1))+u_2,\quad \text{under}\quad u_1\in[0,\varepsilon(\varphi(2)-\varphi(1))+u_2],$$
$$G_1(u_1)=u_1,\quad \text{under}\quad u_1\in[\varepsilon(\varphi(2)-\varphi(1))+u_2,\varepsilon(\varphi(2)-\varphi(1))+u_2+2\delta],$$
$$G_1(u_1)=2u_1-u_2-\varepsilon(\varphi(2)-\varphi(1))-2\delta,\quad \text{under}\quad u_1\in[\varepsilon(\varphi(2)-\varphi(1))+u_2+2\delta,+\infty).$$

The minimum of this function is attained at zero, that is, at $u_1=0$. Let us substitute the found value of u_1 and consider the function of the argument u_2. It has the form

$$G_2(u_2)=\max\{|\delta+\varepsilon[\varphi(2)-\varphi(1)]+u_2|-\delta,0\}+\max\{|\delta+\varepsilon[\varphi(3)-\varphi(2)]+u_3-u_2|-\delta,0\}=$$
$$=\varepsilon[\varphi(2)-\varphi(1)]+u_2+\max\{|\delta+\varepsilon[\varphi(3)-\varphi(2)]+u_3-u_2|-\delta,0\}.$$

This function is of the same type as in the first case; therefore, its minimum is attained again at zero. At all other steps, except at the last one, the results are similar. Consider the last step. The function to be minimized has the form

$$G_k(u_k) = \max\{|\delta + \varepsilon[\varphi(k) - \varphi(k-1)] + u_k| - \delta, 0\} + u_k = \varepsilon[\varphi(k) - \varphi(k-1)] + 2u_k.$$

It is easy to see that the minimum is again attained at zero. Now, the solvability condition takes the form

$$\sum_{i=2}^{k} \varepsilon[\varphi(j) - \varphi(j-1)] \leq m(2\delta + \Delta) - 2k\delta - \varepsilon[\varphi(k) + \varphi(1)],$$

and elementary transformations yield

$$\varepsilon\varphi(k) + k\delta \leq m(\delta + \frac{\Delta}{2}).$$

□

We have obtained the following result.

Theorem 4.1. *Suppose that a sequence $(x_1, \ldots, x_n)$ satisfies the mean and the spread hypothesis, that is, $(x_1, \ldots, x_n) \in X(n, m, a, \delta, \Delta)$. Let φ be a strictly increasing function of a positive integer argument taking positive real values. Suppose that $a, \delta, \varepsilon \in E$, $\delta \geq 0$, and $\varepsilon > 0$. Then the range of variation of the function $\mu(g^n(\cdot, a, \delta, \varepsilon), x_1^n, m)$, which determines the soft probability of the event*

$$S(x_1^k) \geq k(a + \delta) + \varepsilon\varphi(k), \quad k = 1, \ldots, n,$$

is given by

$$\min_{x_1^n \in X(n,m,a,\delta,\Delta)} \underline{\mu}(g^n(\cdot, a, \delta, \varepsilon), x_1^n, m) = 0,$$

$$\max_{x_1^n \in X(n,m,a,\delta,\Delta)} \bar{\mu}(g^n(\cdot, a, \delta, \varepsilon), x_1^n, m) = \frac{k^*}{m},$$

where k^* is the maximum solution of the inequality $\varepsilon\varphi(k) + k\delta \leq m(\delta + \frac{\Delta}{2})$ on the positive integer interval $k \in [1, \ldots, m-1]$. If there are no solutions in this interval, then the corresponding quantity vanishes.

Note that possible bounds for the soft probability of deviations do not depend of the size of the database, that is, on n.

References

[1] Alimov, Yu.I., Kravtsov, Yu.A., (1992), Is probability a "normal" physical quantity?, Sov. Phys. Usp., 35(7), pp.606-622.

[2] Alimov, Yu.I., (1980), An Alternative to the Method of Mathematical Statistics, Znanie, Moscow (in Russian).

[3] Alimov, Yu.I., (1979), Realism and fantasy in applications of probability theory revisited, Avtomatika, 4, pp.83-90.

[4] Chaikovskii, Yu.V., (2004), On the Nature of Randomness, Tsentr Sist. Issled., Moscow (in Russian).

[5] El'yasberg, P.E., (1983), Measurement Data: How Much is Needed? How to Process it?, Nauka, Moscow (in Russian).

[6] Kolmogorov, A.N., (1983), On Logical Foundations of Probability Theory, Lecture Notes in Math., 1021, Probability Theory and Mathematical Statistics, ed. by K. Itô and J. V. Prokhorov, Springer-Verlag, Berlin, Heidelberg, pp.1-5.

[7] Kolmogorov, A.N., (1963), On tables of random numbers, Sankhyā: Indian J. of Statistics, Ser. A, 25(4), pp.369-376.

[8] Kuznetsov, V.P., (1991), Interval Statistical Models, Radio i Svyaz', Moscow (in Russian).

[9] Molodtsov, D.A., (2004), Theory of Soft Sets, URSS, Moscow (in Russian).

[10] Molodtsov, D.A.,(2008), Soft uncertainty and probability, Nechetkie Sist. Myagkie Vychisl., 3(1), pp.7-30.
[11] Molodtsov, D.A., (2010), Frequency and Probability Finite Stability, Nechetkie Sist. Myagkie Vychisl., 5(1).
[12] Molodtsov, D.A., (1999), Soft set theory: First results, Comput. Math. Appl., 37(4-5), pp.19-31.
[13] Von Mises R., (1964), The Mathematical Theory of Probability and Statistics, Academic, London.
[14] Wally, P., (1991), Statistical Reasoning with Imprecise Probabilities, Chapman & Hall, London.

Dmitry Molodtsov was born in 1949 in Moscow, Russia. He is a senior researcher in the Computing Centre of Russian Academy of Sciences. D. Molodtsov graduated from Lomonosov Moscow State University in 1971, and received his Ph.D. (Mathematician) in 1974 from this University. He received Doctor of Science (Mathematician) degree from Moscow State University in 1990. He is a founder of the soft set theory. His current research interests are in the areas of soft set theory and its applications (soft probability, soft portfolio management, soft predictions, soft analysis, soft game theory, soft optimization).