A non-zero dispersion leads to the non-zero bias of mean

Harin, Alexander

Modern University for the Humanities

11 June 2013

A non-zero dispersion leads to the non-zero bias of mean

Alexander Harin
aaharin@yandex.ru
Modern University for the Humanities

A theorem of existence of the non-zero restrictions for the mean of a function on a finite numerical segment at a non-zero dispersion of the function is proved. The theorem has an applied character. It is aimed to be used in the probability theory and statistics and further in economics. Its ultimate aim is to help to answer the Aczél-Luce question whether *W(1)=1* and to explain, at least partially, the well-known problems and paradoxes of the utility theory, such as the underweighting of high and the overweighting of low probabilities, the Allais paradox, the four-fold pattern paradox, etc., by purely mathematical methods.

Contents

Introduction

The research, which first part is presented in this article, was initiated by the paradoxes of the decision and utility theories. The analysis of such paradoxes was started in 1738 by Bernoulli in [1]. The examples of these paradoxes are the Allais paradox [2], the Ellsberg paradox [3], the "four-fold pattern" paradox (see, e.g. [4]), etc. In 2002 Kahneman got the Prize in Memory of Nobel for the research in this field. In 2006 in [5], Kahneman and Thaler pointed out the preferences inconsistencies in the paradoxes are still not overcome adequately.

One of possible ways of solution of these paradoxes was proposed in [6] and in other works (see, e.g. [7] and [8]). The essence of this way consists in a proper attention to noises, imprecision and other reasons those may cause the dispersion, scattering, variation, spread of data.

Aczél and Luce [9] stated a fundamental question (problem) whether $W(1)=1$ (whether the Prelec's weighting function is equal to 1 at $p=1$). This question opens one more way which consists in a proper attention to boundaries.

The research partially presented in this article combines these two ways. That is to say, it considers a dispersion of data near boundaries.

The research has an applied character. Its aim is to provide a mathematical support for works those are based on the dispersion of data and for works those concern the Aczél-Luce question.

This article, as the first part of the research, deals with the general case of the restrictions for the mean of a function on finite numerical segments in the presence of a non-zero dispersion of the function.

The second part of the research will deal with the estimations of restrictions values.

The third part of the research will deal with the restrictions for the probability estimation and for the probability.

The fourth part of the research will deal with possible explanations of the abovementioned paradoxes of the utility theory and with the Aczél-Luce question.

An illustrative example of restrictions

Two points

Let us suppose a numerical segment *[A; B]* (see figure 1). Let us suppose two points are determined on this segment: a left point x_{Left} and a right point x_{Right} : $x_{Left} < x_{Right}$. The coordinates of the middle, mean point may be calculated as $M=(x_{Left}+x_{Right})/2$.

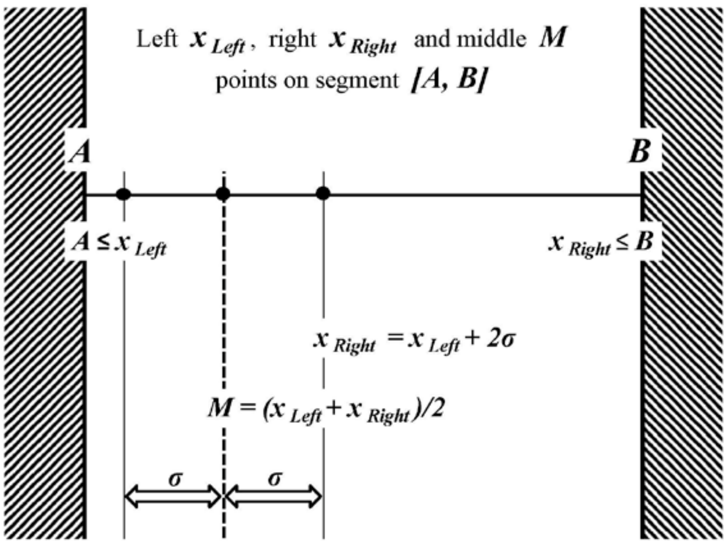

Figure 1. A segment *[A, B]*. Left x_{Left}, right x_{Right} and middle, mean M points on it

Suppose the points can not escape outside the borders of this segment. This means $A \le x_{Left}$ and $x_{Right} \le B$.

Suppose the points can not approach each other closer than a non-zero distance of two sigma $2\sigma > 0$. This means $x_{Right} \ge x_{Left}+2\sigma$ or $x_{Left} \le x_{Right}-2\sigma$. At that, $M-x_{Left}=x_{Right}-M \ge \sigma > 0$.

For the sake of simplisity and obviousness, figures 1-3 represent a case: $x_{Right}=x_{Left}+2\sigma$ and $x_{Left}=x_{Right}-2\sigma$ and $M-x_{Left}=x_{Right}-M=\sigma$.

One can easily see two types of zones can exist on the segment:

The mean point M can be located only in the zone which may be named "allowed" (see figure 2).

The mean point M can not be located in the zones which may be named "forbidden" (see figure 3).

Allowed zone

Due to the conditions of the example, the left point x_{Left} may not be located more left than the left border of the segment $A \le x_{Left}$ and the right point x_{Right} may not be located more right than the right border of the segment $x_{Right} \le B$. For M we have $M-x_{Left}=x_{Right}-M=\sigma$.

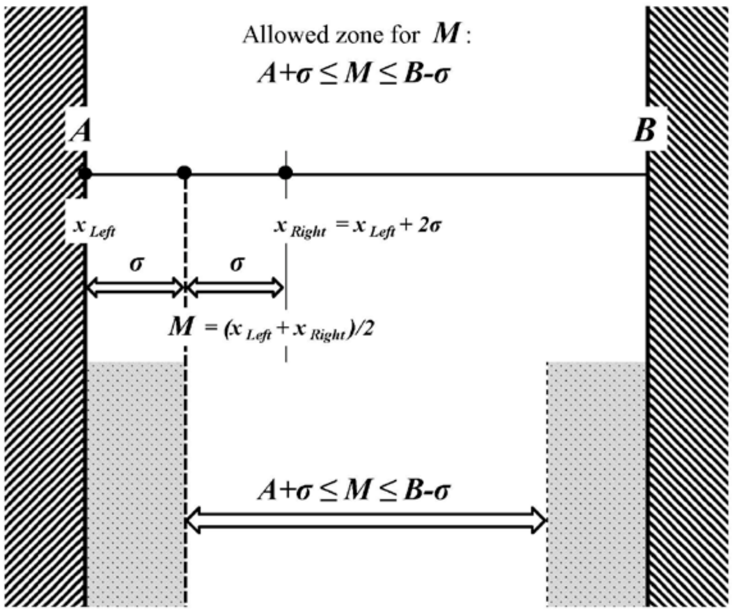

Figure 2. Allowed zone for M

The allowed zone for M is equal to $(B-A)-2\sigma$. It is less than the segment on 2σ. If the distance 2σ between the left x_{Left} and right x_{Right} points is non-zero then the difference between the allowed zone and the segment is non-zero also.

So, the mean point M can not be located in any position of the segment.

Forbidden zones, restrictions

If $A \le x_{Left}$, $x_{Right} \le B$ and $x_{Right} - x_{Left} \ge 2\sigma$, then there are the restrictions of one sigma σ between the mean point and the borders of the segment. So, the mean point M can not be located in two zones located near the borders of the segment. These zones may be named forbidden zones or restrictions.

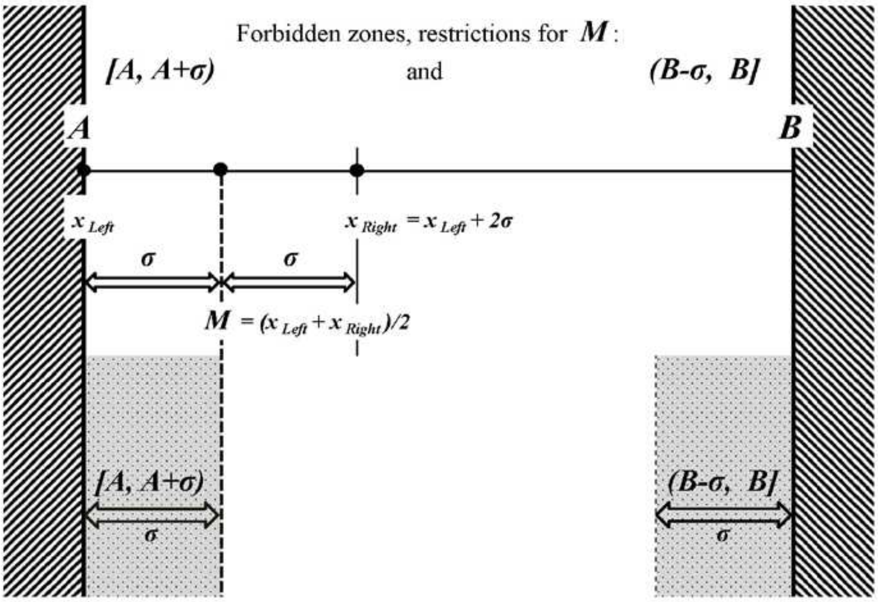

Figure 3. Forbidden zones, restrictions for M

As we can easily see, the restrictions exist between the allowed zone of the mean M and the borders of the segment A and B. The width of every restriction is equal to σ. If the distance 2σ between the left x_{Left} and right x_{Right} points is non-zero then the forbidden zones, restrictions for M are non-zero also.

1. Preliminary notes

1.1. Segment and function

Let us suppose a both finitely big and finitely small numerical segment $X=[A, B] : 0<Const_{AB}\leq(B-A)<\infty$, a set of points $x_k : k=1, 2, \dots K : 2\leq K\leq\infty$, and a finitely big non-negative function $f_K(x_k) :$ for $x_k<A$ and $x_k>B$ the statement $f_K(x_k)\equiv 0$ is true; for $A\leq x_k\leq B$ the statement $0\leq f_K(x_k)<\infty$ is true, and

$$\sum_{k=1}^{K} f_K(x_k) = W_K ,$$

where a constant W_K (the total weight of $f_K(x_k)$) is such that

$$0 < W_K < \infty .$$

Keeping generality, the function $f_K(x_k)$ may be normalized so that $W_K=1$.

1.2. Analog of moments

Definition 1.1. Let us define an analog of a moment of n-th order of the function $f_K(x_k)$ relative to a point x_0 as

$$E(X - X_0)^n = \frac{1}{W_K}\sum_{k=1}^{K}(x_k - x_0)^n f_K(x_k) = \sum_{k=1}^{K}(x_k - x_0)^n f_K(x_k) .$$

Let us further in this article, for brevity, name the analog of a moment of n-th order or the n-th order moment analog as simply the moment of n-th order.

Let us suppose the mean $M\equiv E(X)$ of the function $f_K(x_k)$ exists

$$E(X) \equiv \frac{1}{W_K}\sum_{k=1}^{K} x_k f_K(x_k) = \sum_{k=1}^{K} x_k f_K(x_k) \equiv M .$$

Let us suppose at least one central moment $E(X-M)^n : 2\leq n<\infty$, of the function $f_K(x_k)$ exists

$$E(X - M)^n = \frac{1}{W_K}\sum_{k=1}^{K}(x_k - M)^n f_K(x_k) = \sum_{k=1}^{K}(x_k - M)^n f_K(x_k) .$$

2. Maximality

Let us search a function which guarantees the maximal central moment and let us prove this choice. The intuitively evident maximal possible absolute value of a central moment is obtained for the function which is concentrated at the borders of the segment.

2.1. A couple of elements

Let us consider the mean M of the function $f_K(x_k)$, a couple of points x_A and x_B, such as

$$A \le x_A \le M \le x_B \le B \ ,$$

and a couple of elements $f_K(x_A)$ and $f_K(x_B)$ such as they are tied together by the conditions of the constant total weight f and the constant joint point M

$$f_K(x_A) + f_K(x_B) = f \ ,$$

$$(M - x_A) f_K(x_A) = (x_B - M) f_K(x_B) \ .$$

A central moment $E_{Couple}(X\text{-}M)^n$ of this couple of elements may be written as

$$E_{Couple}(X - M)^n \equiv (x_A - M)^n f_K(x_A) + (x_B - M)^n f_K(x_B) .$$

Its absolute value do not exceed the sum of absolute values of its parts

$$\begin{aligned} | E_{Couple}(X - M)^n | &\le | (x_A - M)^n | f_K(x_A) + | (x_B - M)^n | f_K(x_B) = \\ &= (M - x_A)^n f_K(x_A) + (x_B - M)^n f_K(x_B) \end{aligned} .$$

2.2. Modification of the basic expression

After substituting $f_K(x_B)$ by

$$f_K(x_B) = \frac{M - x_A}{x_B - M} f_K(x_A) = f - f_K(x_A)$$

and replacing $f_K(x_A)$ and $f_K(x_B)$ by functions of x_A and x_B

$$f_K(x_A) + f_K(x_B) = \frac{x_B - M + M - x_A}{x_B - M} f_K(x_A) = \frac{x_B - x_A}{x_B - M} f_K(x_A) = f \ ,$$

we obtain the function $f_K(x_k)$:

$$f_K(x_A) = \frac{x_B - M}{x_B - x_A} f \quad \text{and} \quad f_K(x_B) = \frac{M - x_A}{x_B - x_A} f \ .$$

At that, the expression for the central moment $E_{Couple}(X\text{-}M)^n$ of the couple may be reorganized to the expression which depends only on x_A and x_B

$$| E_{Couple}(X - M)^n | \le (M - x_A)^n \frac{x_B - M}{x_B - x_A} f + (x_B - M)^n \frac{M - x_A}{x_B - x_A} f \ .$$

2.3. Derivatives

Let us use the analysis of derivatives to find a maximum of the absolute value of central moments $|E_{Couple}(X\text{-}M)^n|$ of this couple of elements.

Let us differentiate the expression for the absolute value of a central moment by x_A

$$[(M-x_A)^n \frac{x_B-M}{x_B-x_A} f + (x_B-M)^n \frac{M-x_A}{x_B-x_A} f]'_{x_A} =$$

$$= \frac{-n(M-x_A)^{n-1}(x_B-x_A)+(M-x_A)^n}{(x_B-x_A)^2}(x_B-M)f +$$

$$+(x_B-M)^n \frac{-(x_B-x_A)+(M-x_A)}{(x_B-x_A)^2} f =$$

$$= \{[-n(x_B-x_A)+(M-x_A)](M-x_A)^{n-1} +$$

$$+[-(x_B-x_A)+(M-x_A)](x_B-M)^{n-1}\} \frac{(x_B-M)}{(x_B-x_A)^2} f$$

Since $n \geq 2$ hence:

If $(x_B-x_A)=(M-x_A)$ that is if $x_B=M$ then from

$$(M-x_A)f_K(x_A) = (x_B-M)f_K(x_B) \ .$$

we obtain

$$(M-x_A) = (M-M)\frac{f_K(x_B)}{f_K(x_A)} = 0$$

or $x_A=M$ also, hence all the central moments equal zero.

So, at $(x_B-x_A)=(M-x_A)$ that is at $x_B=M=x_A$

$$[(M-x_A)^n \frac{x_B-M}{x_B-x_A} f + (x_B-M)^n \frac{M-x_A}{x_B-x_A} f]'_{x_A} = 0 \, .$$

This is an indifferently stable equilibrium state. If $x_B=M=x_A$ then all the central moments are not depended upon x_A and are equal to zero.

At $(x_B-x_A)>(M-x_A)$ that is at $x_B>M$

$$[(M-x_A)^n \frac{x_B-M}{x_B-x_A} f + (x_B-M)^n \frac{M-x_A}{x_B-x_A} f]'_{x_A} < 0 \, .$$

So, for non-zero $E_{Couple}(X\text{-}M)^n$, the first derivative by x_A is strictly less than zero for any $A \leq x_A < M$ independently of x_B for any $M < x_B \leq B$. The closer is x_A to A the more is $|E_{Couple}(X\text{-}M)^n|$.

Hence, for any $x_B : M < x_B \leq B$, the maximums of the absolute values of a central moments $E_{Couple}(X\text{-}M)^n$ are attained at the minimal x_A, that is at $x_A=A$.

Let us differentiate the expression for the absolute value of a central moment $E_{Couple}(X\text{-}M)^n$ of the couple f by x_B

$$[(M-x_A)^n \frac{x_B-M}{x_B-x_A} f + (x_B-M)^n \frac{M-x_A}{x_B-x_A} f]'_{x_B} =$$

$$= (M-x_A)^n \frac{(x_B-x_A)-(x_B-M)}{(x_B-x_A)^2} f +$$

$$+ \frac{n(x_B-M)^{n-1}(x_B-x_A)-(x_B-M)^n}{(x_B-x_A)^2}(M-x_A) f = .$$

$$= \{[(x_B-x_A)-(x_B-M)](M-x_A)^{n-1} +$$

$$+[n(x_B-x_A)-(x_B-M)](x_B-M)^{n-1}\} \frac{(M-x_A)}{(x_B-x_A)^2} f$$

Since $n \geq 2$ hence:

If $(x_B\text{-}x_A)=(x_B\text{-}M)$ that is if $x_A=M$ then $x_A=M$ also, hence all the central moments equal zero.

So, at $(x_B\text{-}x_A)=(M\text{-}x_A)$ that is at $x_B=M=x_A$

$$[(M-x_A)^n \frac{x_B-M}{x_B-x_A} f + (x_B-M)^n \frac{M-x_A}{x_B-x_A} f]'_{x_A} = 0 .$$

This is an indifferently stable equilibrium state. If $x_B=M=x_A$ then all the central moments are not depended upon x_B and are equal to zero.

At $(x_B\text{-}x_A)>(x_B\text{-}M)$ that is at $x_A<M$

$$[(M-x_A)^n \frac{x_B-M}{x_B-x_A} f + (x_B-M)^n \frac{M-x_A}{x_B-x_A} f]'_{x_B} > 0 .$$

So, for non-zero $E_{Couple}(X\text{-}M)^n$, the first derivative by x_B is strictly more than zero for any $M<x_B\leq B$ independently of x_A for any $A\leq x_A<M$. The closer is x_B to B the more is $|E_{Couple}(X\text{-}M)^n|$.

Hence, for any $x_A : A\leq x_A<M$, the maximums of the absolute values of the central moments $E_{Couple}(X\text{-}M)^n$ are attained at the maximal x_B, that is at $x_B=B$.

So, for non-zero central moments $E_{Couple}(X\text{-}M)^n$ of a couple f of elements $f_K(x_A)$ and $f_K(x_B)$, the maximums of the absolute values of $E_{Couple}(X\text{-}M)^n$ are attained at $x_A=A$ and $x_B=B$. That is, they are attained for the functions that are concentrated at the borders of the segment $[A, B]$ and

$$Max(|E_{Couple}(X-M)^n|) \leq (M-A)^n \frac{B-M}{B-A} f + (B-M)^n \frac{M-A}{B-A} f .$$

2.4. Dividing by couples

Let us analyze whether any function of the chapter 1 and its central moments may be completely divided and represented by such couples of elements.

Let us divide the points x_k into three groups: $x_{k(A)}<M$, $x_{k(M)}=M$ (zero central moments) and $x_{k(B)}>M$. At that, $k(A)\leq K_A$, $k(M)\leq K_M$, $k(B)\leq K_B$ and

$$K_A + K_M + K_B = K .$$

Let us enumerate the points $x_{k(A)}$ and $x_{k(B)}$, for example, from utmost points and maximal weights to closest to M points and minimal weights. So, for $E(X\text{-}M)^2>0$ we have $k(A)=1, \ldots, K_A$ and $k(B)=1, \ldots, K_B$.

The definition of the mean

$$E(X) \equiv \sum_{k=1}^{K} x_k f_K(x_k) \equiv M$$

may be transformed to the first central moment

$$\sum_{k=1}^{K} (x_k - M) f_K(x_k) = 0 \ ,$$

or, if $K_A\geq 1$ and $K_B\geq 1$,

$$\sum_{k=1}^{K} (x_k - M) f_K(x_k) = \sum_{k(A)=1}^{K_A} (x_{k(A)} - M) f_K(x_{k(A)}) +$$

$$+ \sum_{k(M)=1}^{K_M} (x_{k(M)} - M) f_K(x_{k(M)}) + \sum_{k(B)=1}^{K_B} (x_{k(B)} - M) f_K(x_{k(B)}) = \ ,$$

$$= \sum_{k(A)=1}^{K_A} (x_{k(A)} - M) f_K(x_{k(A)}) + \sum_{k(B)=1}^{K_B} (x_{k(B)} - M) f_K(x_{k(B)}) = 0$$

and it may be transformed to a balance

$$\sum_{x_{k(A)}=1}^{K_A} (M - x_{k(A)}) f_K(x_{k(A)}) = \sum_{x_{k(B)}=1}^{K_B} (x_{k(B)} - M) f_K(x_{k(B)}) .$$

Let us consider cases with various numbers of elements $K_{AB}=K_A+K_B$.

The case 0. If $K_{AB}=0$, then $E(X\text{-}M)^n=0$.

The case 1. Evidently, due to the definition of the mean, the case $K_A=0$ and $K_B\geq 1$ and the case $K_A\geq 1$ and $K_B=0$ cannot exist.

The case 2. If $K_{AB}=2$, $K_A=1$ and $K_B=1$, then

$$(M - x_{1(A)}) f_K(x_{1(A)}) = (x_{1(B)} - M) f_K(x_{1(B)})$$

and the pair $f_K(x_{1(A)})$ and $f_K(x_{1(B)})$ is the couple of the previous subchapters.

The case 3. If $K_{AB}=3$, $K_A=2$ or $K_B=2$, for example, if $K_A=2$ and $K_B=1$, then we divide the element $f_K(x_{1(B)})$ into two parts $f_{K.1}(x_{1(B)})$ and $f_{K.2}(x_{1(B)})$ such as

$$f_K(x_{1(B)}) = f_{K.1}(x_{1(B)}) + f_{K.2}(x_{1(B)})$$

and

$$(M - x_{1(A)}) f_K(x_{1(A)}) = (x_{1(B)} - M) f_{K.1}(x_{1(B)}) \ .$$

The pair $f_K(x_{1(A)})$ and $f_{K.1}(x_{B1})$ is the couple. The balance remains

$$(M - x_{1(A)}) f_K(x_{1(A)}) + (M - x_{2(A)}) f_K(x_{2(A)}) =$$
$$= (x_{1(B)} - M) f_{K.1}(x_{1(B)}) + (x_{1(B)} - M) f_{K.2}(x_{1(B)}) \ ,$$

and we come to the case 2

$$(M - x_{2(A)}) f_K(x_{2(A)}) = (x_{1(B)} - M) f_{K.2}(x_{1(B)}) \ .$$

The pair $f_K(x_{2(A)})$ and $f_{K.2}(x_{B1})$ is the couple also.

The general case L. Suppose a case $K_{AB}=L\geq 4$, $K_A\geq 2$ or $K_B\geq 2$. If

$$(M - x_{1(A)}) f_K(x_{1(A)}) = (x_{1(B)} - M) f_K(x_{1(B)}) \ ,$$

then the pair $f_K(x_{1(A)})$ and $f_K(x_{1(B)})$ is the couple. The number of uncoupled elements is diminished by two and we come to the case L-2. If

$$(M - x_{1(A)}) f_K(x_{1(A)}) \neq (x_{1(B)} - M) f_K(x_{1(B)}) \ ,$$

then, as in the case 3, let us divide the appropriate element as in the case 3 and we diminish the number of uncoupled elements by one and come to the case L-1.

So, we may consecutively diminish the number of uncoupled elements from any L to 2 and, so, we may come to fully coupled elements. Hence, any function of the chapter 1 may be completely divided by couples of elements, except of $x_{k(M)}$.

So, any function of the chapter 1 and its central moments may be completely divided and represented by couples of elements except of points $x_{k(M)}$ which do not contribute to central moments. So, the function $f_{Max.K}(x_k)$, which possesses maximal central moments modules, should be concentrated at the borders $x_1=A$ and $x_2=B$ of the segment. At the condition of the unit norm of the chapter 1 and for the mean M, the function $f_{Max.K}(x_k)\equiv f_{Borders.K}(x_k)$ should have a form

$$f_{Borders.K}(A) = \frac{B-M}{B-A} \qquad \text{and} \qquad f_{Borders.K}(B) = \frac{M-A}{B-A} \ .$$

The central moments $E_{Borders}(X-M)^n$ of the function $f_{Borders.K}(x_k)$ are

$$E_{Borders}(X-M)^n = (A-M)^n \frac{B-M}{B-A} + (B-M)^n \frac{M-A}{B-A} .$$

The modules of the central moments $E_{Borders}(X-M)^n$ of the function $f_{Borders.K}(x_k)$ are not more, than

$$| E_{Borders}(X-M)^n | \leq (M-A)^n \frac{B-M}{B-A} + (B-M)^n \frac{M-A}{B-A} .$$

For the even orders $2n$ of the central moments $E_{Borders}(X-M)^n$ the inequality is reduced to the equality without the modules

$$E_{Borders}(X-M)^{2n} = (M-A)^{2n} \frac{B-M}{B-A} + (B-M)^{2n} \frac{M-A}{B-A} .$$

So, the modules of the central moments of any function $f_K(x_k)$ of the chapter 1 are not more, than

$$Max(| E(X-M)^n |) \leq (M-A)^n \frac{B-M}{B-A} + (B-M)^n \frac{M-A}{B-A} \qquad (2.1).$$

2.5. Two notes

Let us analyze the maximal central moments $E_{Borders}(X\text{-}M)^n$ for $M=(B\text{-}A)/2$ and for M which is near A or B.

The mean is in the center of the segment

Let us analyze the maximal central moments for $M=(B\text{-}A)/2$.

Let us differentiate the expression for the absolute value of a central moment $E_{Borders}(X\text{-}M)^n$ by M

$$\frac{1}{B-A}[(M-A)^n(B-M)+(B-M)^n(M-A)]'_M =$$
$$=\frac{1}{B-A}[n(M-A)^{n-1}(B-M)-(M-A)^n -$$
$$-n(B-M)^{n-1}(M-A)+(B-M)^n]$$

and, at $M=(B\text{-}A)/2$,

$$\frac{1}{B-A}[n(M-A)^{n-1}(B-M)-(M-A)^n -$$
$$-n(B-M)^{n-1}(M-A)+(B-M)^n] = \quad .$$
$$=\frac{1}{B-A}(\frac{B-A}{2})^n[n-1-n+1]=0$$

So, at $M=(B\text{-}A)/2$, for any $n\geq 2$ there is an extremum or a point of inflection.

Let us differentiate $E_{Borders}(X\text{-}M)^n$ once more

$$\frac{1}{B-A}[n(M-A)^{n-1}(B-M)-(M-A)^n-$$
$$-n(B-M)^{n-1}(M-A)+(B-M)^n]'_M=$$
$$=\frac{1}{B-A}[n(n-1)(M-A)^{n-2}(B-M)-n(M-A)^{n-1}-n(M-A)^{n-1}+$$
$$+n(n-1)(B-M)^{n-2}(M-A)-n(B-M)^{n-1}-n(B-M)^{n-1}]=$$
$$=\frac{1}{B-A}[n(n-1)(M-A)^{n-2}(B-M)-2n(M-A)^{n-1}+$$
$$+n(n-1)(B-M)^{n-2}(M-A)-2n(B-M)^{n-1}]$$

and, at $M=(B\text{-}A)/2$,

$$\frac{1}{B-A}[n(n-1)(M-A)^{n-2}(B-M)-2n(M-A)^{n-1}+$$
$$+n(n-1)(B-M)^{n-2}(M-A)-2n(B-M)^{n-1}]=$$
$$=\frac{1}{B-A}(\frac{B-A}{2})^{n-1}[n(n-1)-2n+n(n-1)-2n]=$$
$$=(\frac{B-A}{2})^{n-2}[n(n-3)]$$

That is, at $M=(B\text{-}A)/2$:

For $n=2$ there is a well-known maximum, the moment of inertia of two material points which weights are equal to each other

$$E_{Borders}(X-\frac{B-A}{2})^2=(M-A)^2\frac{B-M}{B-A}+(B-M)^2\frac{M-A}{B-A}=$$
$$=(\frac{B-A}{2})^2\frac{1}{2}+(\frac{B-A}{2})^2\frac{1}{2}=(\frac{B-A}{2})^2$$

For $n=3$ there is a point of inflection and for $n>3$ there are minimums.

The mean is near a border of the segment

Let us search maximums which are close to the borders of the segment.

Let us differentiate the absolute value of a central moment $E_{Borders}(X\text{-}M)^n$ by M for $M \approx A$ and $n>>1$

$$\frac{1}{B-A}[(M-A)^n(B-M)+(B-M)^n(M-A)]'_M =$$
$$=\frac{1}{B-A}[n(M-A)^{n-1}(B-M)-(M-A)^n -$$
$$-n(B-M)^{n-1}(M-A)+(B-M)^n] \approx$$
$$\approx \frac{1}{B-A}[(B-M)^n - n(B-M)^{n-1}(M-A)]$$

and

$$\frac{1}{B-A}[(B-M)^n - n(B-M)^{n-1}(M-A)] =$$
$$=\frac{(B-M)^{n-1}}{B-A}[(B-M)-n(M-A)] =$$
$$=\frac{(B-M)^{n-1}}{B-A}[(B-A-(M-A))-n(M-A)] =$$
$$=\frac{(B-M)^{n-1}}{B-A}[(B-A)-(n+1)(M-A)] = 0$$

and

$$\boldsymbol{M-A \approx \frac{B-A}{n+1}}.$$

The second derivative gives

$$\frac{1}{B-A}[n(M-A)^{n-1}(B-M)-(M-A)^n -$$
$$-n(B-M)^{n-1}(M-A)+(B-M)^n]'_M =$$
$$=\frac{1}{B-A}[n(n-1)(M-A)^{n-2}(B-M)-n(M-A)^{n-1}-n(M-A)^{n-1}+$$
$$+n(n-1)(B-M)^{n-2}(M-A)-n(B-M)^{n-1}-n(B-M)^{n-1}] = .$$
$$=\frac{1}{B-A}[n(n-1)(M-A)^{n-2}(B-M)-2n(M-A)^{n-1}+$$
$$+n(n-1)(B-M)^{n-2}(M-A)-2n(B-M)^{n-1}]$$

For $M \approx A$ and $n >> 1$

$$\frac{1}{B-A}[n(n-1)(M-A)^{n-2}(B-M)-2n(M-A)^{n-1}+$$
$$+n(n-1)(B-M)^{n-2}(M-A)-2n(B-M)^{n-1}]\approx$$
$$\approx n\frac{(B-M)^{n-2}}{B-A}[(n-1)(M-A)-2(B-M)]=$$
$$=n\frac{(B-M)^{n-2}}{B-A}[(n-1)(M-A)-2((B-A)-(M-A))]=$$
$$=n\frac{(B-M)^{n-2}}{B-A}[(n+1)(M-A)-2(B-A)]$$

and, for $M \approx A+(B-A)/(n+1)$ and $n >> 1$,

$$n\frac{(B-M)^{n-2}}{B-A}[(n+1)(M-A)-2(B-A)]\approx$$
$$\approx n\frac{(B-M)^{n-2}}{B-A}[(n+1)\frac{B-A}{n+1}-2(B-A)]=.$$
$$=n(B-M)^{n-2}[1-2]<0$$

So, the second derivative is negative and there are the maximums at the points $M \approx A+(B-A)/(n+1)$.

The analog of the central moments $E_{Borders}(X-M)^n$ of the function $f_{Borders.K}(x_k)$ for $M-A \approx (B-A)/(n+1)$ and $n >> 1$ may be counted as

$$|E_{Borders}(X-M)^n| \le (M-A)^n\frac{B-M}{B-A}+(B-M)^n\frac{M-A}{B-A}\approx$$
$$\approx (B-M)^n\frac{M-A}{B-A}\approx (B-A-\frac{B-A}{n+1})^n\frac{1}{n+1}= \qquad .$$
$$=(B-A)^n(1-\frac{1}{n+1})^n\frac{1}{n+1}$$

For $n >> 1$

$$(1-\frac{1}{n+1})^n=(1-\frac{1}{n+1})^{n+1}(1-\frac{1}{n+1})^{-1}\approx\frac{1}{e}.$$

So, for $M \approx A+(B-A)/(n+1)$ and $n >> 1$, the maximums (those are attained for even n) of $E_{Borders}(X-M)^n$ are curiously

$$Max(|E_{Borders}(X-(A+\frac{B-A}{n+1}))^n|)\approx\frac{1}{e}\frac{(B-A)^n}{n+1}.$$

Evidently, for $M \approx B-(B-A)/(n+1)$, at $n >> 1$, the maximums (those are attained for even n) of $E_{Borders}(X-M)^n$ are analogously

$$Max(|E_{Borders}(X-(B-\frac{B-A}{n+1}))^n|)\approx\frac{1}{e}\frac{(B-A)^n}{n+1}.$$

3. Theorem

3.1. General lemma about tendency to zero for central moments

Lemma 3.1. If, for the function $f_K(x_k)$, defined in the section 1, $M \equiv E(X)$ tends to A or to B,

then, for $n : 2 \le n < \infty$, $E(X-M)^n$ tends to zero.

Proof. For $M \to A$, the estimation (2.1) gives

$$
\begin{aligned}
&|E(X-M)^n| \le (M-A)^n \frac{B-M}{B-A} + (B-M)^n \frac{M-A}{B-A} = \\
&= [(M-A)^{n-1} + (B-M)^{n-1}] \frac{(M-A)(B-M)}{B-A} < \\
&< [(B-A)^{n-1} + (B-A)^{n-1}] \frac{(M-A)(B-M)}{B-A} \le \\
&\le 2(B-A)^{n-1}(M-A) \xrightarrow[M \to A]{} 0
\end{aligned}
.
$$

This rough estimation is already sufficient for the purpose of this article. But a more precise estimation may be obtained:

Let us transform

$$
\begin{aligned}
&[(M-A)^{n-1} + (B-M)^{n-1}] \frac{(M-A)(B-M)}{B-A} = \\
&= [(\frac{M-A}{B-A})^{n-1} + (\frac{B-M}{B-A})^{n-1}](B-A)^{n-1} \frac{(M-A)(B-M)}{B-A}
\end{aligned}
.
$$

Let us consider the terms $(M-A)/(B-A)$ and $(B-M)/(B-A)$. Keeping in mind $A \le M \le B$ we obtain $0 \le (M-A)/(B-A) \le 1$ and $0 \le (B-M)/(B-A) \le 1$. For $n \ge 2$ we have

$$
\begin{aligned}
&(\frac{M-A}{B-A})^{n-1} + (\frac{B-M}{B-A})^{n-1} \le \\
&\le \frac{M-A}{B-A} + \frac{B-M}{B-A} = \frac{B-A}{B-A} \equiv 1
\end{aligned}
.
$$

So,

$$
\begin{aligned}
&[(\frac{M-A}{B-A})^{n-1} + (\frac{B-M}{B-A})^{n-1}](B-A)^{n-1} \frac{(M-A)(B-M)}{B-A} \le \\
&\le (B-A)^{n-1} \frac{(M-A)(B-M)}{B-A} \le (B-A)^{n-1}(M-A)
\end{aligned}
.
$$

So,

$$
|E(X-M)^n| \le (B-A)^{n-1}(M-A) \xrightarrow[M \to A]{} 0 \qquad (3.1).
$$

For $M \to B$, the proof is similar and gives

$$
|E(X-M)^n| \le (B-A)^{n-1}(B-M) \xrightarrow[M \to B]{} 0 \qquad (3.2).
$$

So, if $(B-A)$ and n are finite and $M \to A$ or $M \to B$, then $E(X-M)^n \to 0$.

The lemma has been proved.

3.2. General theorem of existence of restrictions for mean

Definition 3.1. Let us define the term "restriction for dispersion of n-th order" $r_{Dispersion.n} \equiv r_{Disp.n} > 0$ (where dispersion is implied in the broad sense, as scattering, spread, variation, etc) as a minimal absolute value of the analog of the n-th order central moment $E(X-M)^n$ such as $|E(X-M)^n| \geq r^n{}_{Disp.n} > 0$.

For $n=2$ the restriction for dispersion of second order is equal to the minimal possible standard deviation $r_{Disp.2} = \sigma_{Min}$.

Note, $r_{Disp.n} < (B-A)$. This follows from

$$E(X-M)^n = \sum_{k=1}^{K}(x_k - M)^n f_K(x_k) < (B-A)^n \sum_{k=1}^{K} f_K(x_k) = (B-A)^n .$$

Theorem 3.2. If, for the finite nonnegative discrete function $f_K(x_k)$ defined in the section 1, with the mean $M \equiv E(X)$ of the function and the analog of an n-th $: 2 \leq n < \infty$, order central moment $E(X-M)^n$ of the function, a non-zero restriction for dispersion of the n-th order $r_{Disp.n} : |E(X-M)^n| \geq r^n{}_{Disp.n} > 0$, exists

then the non-zero restriction $r_{Mean} > 0$ for the mean $E(X)$ exists such as $A < (A + r_{Mean}) \leq M \equiv E(X) \leq (B - r_{Mean}) < B$.

Proof. From the conditions of the theorem and from (3.1) for $M \to A$,

$$0 < r^n{}_{Disp.n} \leq |E(X-M)^n| \leq (B-A)^{n-1}(M-A)$$

and

$$0 < \frac{r^n{}_{Disp.n}}{(B-A)^{n-1}} \leq (M-A)$$

So,

$$(M-A) \geq r_{Mean} \equiv \frac{r^n{}_{Disp.n}}{(B-A)^{n-1}} > 0 \qquad (3.3).$$

For $M \to B$, the proof is similar and gives

$$(B-M) \geq r_{Mean} \equiv \frac{r^n{}_{Disp.n}}{(B-A)^{n-1}} > 0 \qquad (3.4).$$

So, as long as $r_{Disp.n}$ is finitely small, n is finitely big, $(B-A)$ is both finitely big and finitely small,

then r_{Mean} is finitely small and both $(M-A) \geq r_{Mean} > 0$ and $(B-M) \geq r_{Mean} > 0$.

The theorem has been proved.

Note. This estimation is an ultra-reliable one. It is, in a sense, as ultra-reliable as the Chebyshev inequality. Preliminary calculations [10] which were performed for real cases, such as normal, uniform and exponential distributions with the minimal values $\sigma^2{}_{Min}$ of the analog of the dispersion (in the particular sense), gave the restrictions r_{Mean} for the mean of the function, those are not worse than

$$r_{Mean} \geq \frac{\sigma_{Min}}{3} .$$

Comments to the theorem

We may reformulate the essence of the theorem in some variants:

If the analog of a finite ($n<\infty$) central moment $E(X-M)^n$ of a finite non-negative function, which is defined for a finite segment, cannot approach zero closer, than by a non-zero value $|E(X-M)^n|\geq r^n{}_{Disp.n}>0$, then the mean of the function also cannot approach any border of this segment closer, than by the non-zero value $r_{Mean}>0$.

More particular: If the analog of the dispersion (in the particular sense) $E(X-M)^2$ of a finite non-negative function, which is defined for a finite segment, cannot approach zero closer, than by a non-zero value $E(X-M)^2\geq\sigma^2{}_{Min}>0$, then the mean of the function also cannot approach any border of this segment closer, than by the non-zero value $r_{Mean}>0$.

In other words: If for a finite non-negative function, which is defined for a finite segment, a non-zero restriction $r_{Disp.n}>0$ exists between the zone of possible values of the analog of a finite ($n<\infty$) central moment $E(X-M)^n$ of the function and zero $|E(X-M)^n|\geq r^n{}_{Disp.n}>0$, then the non-zero restrictions $r_{Mean}>0$ also exist between the zone of possible values of the mean of this function and any border of the segment.

More particular: If for a finite non-negative function, which is defined for a finite segment, a non-zero restriction $\sigma_{Min}>0$ exists between the zone of possible values of the analog of the dispersion (in the particular sense) $E(X-M)^2$ of the function and zero $E(X-M)^2\geq\sigma^2{}_{Min}>0$, then the non-zero restrictions $r_{Mean}>0$ also exist between the zone of possible values of the mean of this function and any border of the segment.

In other words: If there is zero restriction $r_{Disp.n}=0$ for dispersion (in the broad sense) $E(X-M)^n$ of a function then there are zero restrictions $r_{Mean}=0$ for the mean of the function. The more restriction $r_{Disp.n}>0$ for the dispersion the more restrictions $r_{Mean}>0$ for the mean.

So, a restriction $r_{Disp.n}>0$ for the dispersion biases the boundaries of the zone of possible values of the mean from the borders of the segment to the middle of the segment.

So, a restriction $r_{Disp.n}>0$ for the dispersion biases the mean from the borders to the middle of the segment.

Simplified: A non-zero dispersion of a finite non-negative function leads to the non-zero restrictions for the mean of this function.

More simplified: A non-zero dispersion leads to the non-zero bias of the mean.

Most simplified: Dispersion biases mean.

4. Applications of the theorem in economics

The theorem has been preliminary proved (see, e.g., [11]) in the probability theory and statistics for probability estimation and for the probability as the limit of the probability estimation. In the presence of data dispersion, scattering, spread, variation, the restrictions can exist for probability estimation and for the probability near the borders of the probability scale.

Further, the theorem has been preliminary used in economics (see, e.g., [12]) and has explained the well-known problems and paradoxes of decision theory and utility theory, such as the underweighting of high and the overweighting of low probabilities, the four-fold pattern paradox, etc.

In the presence of a data dispersion, the restrictions, those can exist for the probability near the borders of the probability scale, can bias the results of experiments in comparison with no data dispersion. The preliminary researches, including considerations of the restrictions as a hypothesis, showed this bias can explain (at least partially) the well-known problems and paradoxes of decision and utility theories. It should be noted, this explanation is true not only for a particular combination of parameters but both for high and low probabilities and both for gains and losses (see, e.g., [12]).

The new field of applications of the theorem may be concerned with the Aczél-Luce question [9] whether $W(1)=1$ (whether the Prelec's weighting function is equal to 1 at $p=1$).

Conclusions

Possibility of existence of non-zero restrictions in the presence of a non-zero dispersion (both in the particular sense, as the analog of the second central moment, and in the broad sense, as the scattering, spread) has been analyzed in this article.

The theorem of existence of the non-zero restrictions for the mean of a discrete finite non-negative function on a segment $X=[A, B]$ at a non-zero analog of a central moment of the function has been proved. The theorem states if there is a non-zero restriction $r_{Dispersion.n} \equiv r_{Disp.n} > 0$ for the analog of the n-th central moment $|E(X-M)^n|$: $\infty > n \geq 2$, of a discrete finite nonnegative function such as $|E(X-M)^n| \geq r^n_{Disp.n} > 0$, then the non-zero restriction $r_{Mean} > 0$ exists for the mean of this function. The value of the restriction r_{Mean} at A is (see (3.3))

$$(M - A) \geq r_{Mean} \equiv \frac{r_{Disp.n}^{\;n}}{(B - A)^{n-1}} > 0 \ .$$

The value of the restriction r_{Mean} at B is also (see (3.4))

$$(B - M) \geq r_{Mean} \equiv \frac{r_{Disp.n}^{\;n}}{(B - A)^{n-1}} > 0 \ .$$

For $n=2$ the analog of the central moment is the analog of the dispersion (in the particular sense) and r_{Mean} at A may be rewritten for the minimum σ_{Min} of the analog of the standard deviation σ such as $\sigma \geq \sigma_{Min} \equiv r_{Disp.2} > 0$ as

$$(M - A) \geq r_{Mean} = \frac{r_{Disp.2}^{\;2}}{(B - A)} \equiv \frac{\sigma_{Min}^{\;2}}{(B - A)} > 0 \ .$$

The value of the restriction r_{Mean} at B may be also rewritten for the minimum σ_{Min} of the analog of the standard deviation σ such as $\sigma \geq \sigma_{Min} \equiv r_{Disp.2} > 0$ as

$$(B - M) \geq r_{Mean} = \frac{r_{Disp.2}^{\;2}}{(B - A)} \equiv \frac{\sigma_{Min}^{\;2}}{(B - A)} > 0 \ .$$

The function, which ensures the maximal absolute values of the analog of the central moments relative to the mean M, is the function $f_{Borders.K}(x_k)$, which is concentrated at the opposite borders of the numerical segment $X=[A, B]$

$$f_{Borders.K}(A) = \frac{B - M}{B - A} \qquad \text{and} \qquad f_{Borders.K}(B) = \frac{M - A}{B - A} \ .$$

For the module of the analog of the central moments $E_{Borders}(X-M)^n$

$$| E_{Borders}(X - M)^n | \leq (M - A)^n \frac{B - M}{B - A} + (B - M)^n \frac{M - A}{B - A} .$$

of the function $f_{Borders.K}(x_k)$ we have:

For $M=(B-A)/2$:

for $n=2$ there is a well-known maximum, for $n=3$ there is a point of inflection and for $n>3$ there are minimums.

For $M\approx A$ and for $M\approx B$, at $n>>1$:

there are the maximums of the analog of the modules of the central moments $E_{Borders}(X-M)^n$ at

$$x \approx A + \frac{B-A}{n+1} \quad \text{and} \quad x \approx B - \frac{B-A}{n+1} .$$

For $M\approx A+(B-A)/(n+1)$, the maximums (those are attained for even n) of $E_{Borders}(X-M)^n$ are (curiously with $\approx 1/e$ coefficient)

$$Max(|E_{Borders}(X-(A+\frac{B-A}{n+1}))^n|) \approx \frac{1}{e}\frac{(B-A)^n}{n+1}.$$

For $M\approx B-(B-A)/(n+1)$, the maximums (those are attained for even n) of $E_{Borders}(X-M)^n$ are also

$$Max(|E_{Borders}(X-(B-\frac{B-A}{n+1}))^n|) \approx \frac{1}{e}\frac{(B-A)^n}{n+1}.$$

The above estimations for the restrictions r_{Mean} for the mean are, in a sense, as ultra-reliable as the Chebyshov inequality. For real cases such as normal distribution, for the minimal values σ^2_{Min} of the analog of the dispersion (in the particular sense), the preliminary calculations [10] give the restrictions r_{Mean} for the mean, those are no worse than

$$r_{Mean} \geq \frac{\sigma_{Min}}{3} .$$

The theorem may have a significant practical value (It is considered and proved here mainly due to this value):

The theorem has been preliminary proved [11] in the probability theory and statistics for probability estimation and for the probability as the limit of the probability estimation.

The hypothesis of the restrictions and the preliminary proof of the theorem have been used in economics and have qualitatively explained the well-known problems and paradoxes of decision theory and utility theory, such as the underweighting of high and the overweighting of low probabilities, the four-fold pattern paradox, etc. (see, e.g., [12]).

New applications of the theorem may be concerned with researches of the Aczél-Luce question [9] whether the Prelec's weighting function is equal to 1 at $p=1$.

References

[1] Bernoulli D. (1738), “Specimen Theoriae Novae de Mensura Sortis”, Commentarii Academiae Scientiarum Imperialis Petropolitanae, 5 (1738): 175-192. English translation: “Exposition of a New Theory on the Measurement of Risk”, Econometrica, 22(1): 23-36.

[2] Allais, M. (1953) Le comportement de l'homme rationnel devant le risque: Critique le postulats et axioms de L'École Américaine. *Econometrica*, 21: 503-546.

[3] Ellsberg L. (1961), “Risk, Ambiguity, and the Savage Axioms”, The Quarterly Journal of Economics, 75 (4): 643-669.

[4] Tversky, A. and Wakker, P. (1995) Risk attitudes and decision weights. *Econometrica*, 63: 1255-1280.

[5] Kahneman, D. and Thaler, R. (2006) "Anomalies: Utility Maximization and Experienced Utility" Journal of Economic Perspectives, 20, #1: 221-234.

[6] Hey, J., and Orme C. (1994), “Investigating Generalizations of Expected Utility Theory Using Experimental Data”, Econometrica, 62 (6): 1291-1326.

[7] Chay, K., McEwan, P., and Urquiola M. (2005), “The Central Role of Noise in Evaluating Interventions that Use Test Scores to Rank Schools”, American Economic Review, 95 (4): 1237-1258.

[8] Butler, D. J., and Loomes G. C. (2007), “Imprecision as an Account of the Preference Reversal Phenomenon”, American Economic Review, 97 (1): 277-297.

[9] Aczél, János and R. Duncan Luce (2006), “A behavioral condition for Prelec’s weighting function on the positive line without assuming W(1)=1” Journal of Mathematical Psychology. Volume 51, Issue 2, April 2007, Pages 126–129.

[10] Harin A. (2009), “About existence of ruptures in the probability scale: Calculation of ruptures’ values”, Proceedings of the Ninth International Scientific School "Modelling and Analysis of Safety and Risk in complex systems", 9, 458-464, (2009) (in Russian).

[11] Harin A. (2010), “Theorem of existence of ruptures in the probability scale”, Proceedings of the 9th International conference "Financial and Actuarial Mathematics and Eventoconverging Technologies", 9, 312-315, (2010) (in Russian).

[12] Harin A. (2012), “Economics. Paradoxes. Ruptures. Eventology”, Proceedings of the 11th International conference "Financial and Actuarial Mathematics and Eventology of Safety", 11, 368-371, (2012) (in Russian).