An existence theorem for bounds (restrictions) on the expectation of a random variable. Its opportunities for utility and prospect theories

Harin, Alexander

Modern University for the Humanities

16 September 2015

An existence theorem for bounds (restrictions) on the expectation of a random variable. Its opportunities for utility and prospect theories

Alexander Harin
aaharin@gmail.com
Modern University for the Humanities

An existence theorem is proved for the case of a discrete random variable with finite support. If the random variable X takes on values in a finite interval $[a, b]$ and there is a lower non-zero bound on its n^{th} central moment $|E(X-E(X))^n| \geq \sigma_{n.Min}^n| > 0$, then non-zero bounds (restrictions) on its expectation $a < (a + r_{Expect}) \leq E(X) \leq (b - r_{Expect}) < b$ exist near the borders of the interval. In other words, under the above conditions, the non-zero "forbidden zones" exist near the borders a and b and of the interval $[a, b]$. Here $r_{Expect} = \frac{b-a}{2} - \sqrt{\left(\frac{b-a}{2}\right)^2 - \sigma^2_{Min.n}\left(\frac{|\sigma_{Min.n}|}{b-a}\right)^{n-2}}$. For the case of $n=2$ and for the minimal dispersion $\sigma_{Min} \to 0$, the bounds are $a < \left(a + \frac{\sigma^2_{Min}}{b-a}\right) \leq E(X) \leq \left(b - \frac{\sigma^2_{Min}}{b-a}\right) < b$. The theorem can be used in utility and prospect theories.

Contents

1. Introduction

Bounds on functions of random variables are considered in a number of works. At that, information of moments of random variables is used quite often. Bounds for probabilities and expectations of convex functions of discrete random variables with finite support are considered in Prékopa (1990). Inequalities on expectations of functions are considered in Prékopa (1992). The inequalities are based on the knowledge of moments of discrete random variables. A class of lower bounds on the expectation of a convex function using the first two moments of the random variable with a bounded support is considered in Dokov and Morton (2005).

Bounds on the exponential moments of $min(y, X)$ and $X1\{X < y\}$ using the first two moments of a random variable X are considered in Pinelis (2011).

Information of moments of a random variable can be used also for bounds on the expectation of this variable. These bounds can be of importance as well. In particular, they can be used in decision, utility and prospect theories (see, Section 5), including the analysis of Prelec's probability weighting function.

In the present article, the bounds on the expectation of a random variable are expressed in terms of its second or higher moments.

Due to the convenience of abbreviations and to the history of creation and development of the topic of this article, the term "bound" is often referred to here as the term "restriction," especially in mathematical expressions.

2. Preliminary notes

In the present article, the first and simplest case of a discrete random variable with finite support is considered. Other cases may be considered later.

Let us consider a discrete random variable X such that there is a probability space *(Ω, Æ, P)* and $X : \Omega \to R$. Let us suppose that

$$X = \{x_k\} : k = 1, 2, ..., K : 2 \le K < \infty$$

and

$$a \le x_k \le b : 0 < (b - a) < \infty$$

and the probability mass function is

$$f_X(x) = P(X = x) \equiv P(\{\omega \in \Omega : X(\omega) = x\}) \ .$$

Let us consider further the expectation of X

$$E(X) \equiv \sum_{k=1}^{K} x_k f_X(x_k) \equiv \mu \ ,$$

its central moments

$$E(X - \mu)^n = \sum_{k=1}^{K} (x_k - \mu)^n f_X(x_k)$$

and possible interrelationship between the expectation and the moments.

3. Maximality

Let us search for the probability mass function $f_X(x)$ such that a central moment of X attains the maximal possible modulus.

It is intuitively evident that the maximal possible absolute value of a central moment is obtained for the probability mass function, which is concentrated at the borders of the interval. Nevertheless, for the sake of mathematical rigor, this statement must be proved.

For the sake of simplicity, in this section, the probability mass function $f_X(x)$ will be used in simplified forms as $f \equiv f(x) \equiv f_X(x)$.

3.1. Pairs

In the scope of this section, let us analyze all the realizations (or observed values) x_k of the random variable X relative to μ.

Let us consider two possible realizations (points) x_a and x_b of the random variable X and the corresponding probabilities

$$f(x_a) \equiv f_X(x_a) \quad \text{and} \quad f(x_b) \equiv f_X(x_b).$$

For the purposes of this article, let us introduce a term "pair."

Sometimes, one may need to mark objects those are associated with pairs. Let us mark them by an additional subscript. To not confuse with the abbreviation of the term "probability," let us choose a subscript "C" ("couple").

Definition 3.1. Pair. Two realizations (points) x_a and x_b of the discrete random variable X, satisfying

$$a \le x_a \le \mu \le x_b \le b \ ,$$

will be called a "**pair**" (or a "couple")

$$X_{Pair} \equiv X_{Couple} \equiv X_C \equiv (x_a, x_b) \equiv (x_{C.a}, x_{C.b})$$

relative to μ if the balance

$$(\mu - x_a) f(x_a) = (x_b - \mu) f(x_b) \qquad (1)$$

is true, in other words, if $\mu \equiv E(X)$ is the expectation of x_a and x_b as well. At that, if X may be considered as a set, then a pair may be considered as a subset X_C of the set X, having the same expectation μ as X.

Note, if $x_a = x_b$ then the balance (1) can be also considered as true, though formally.

The sum of the probabilities $f(x_a)$ and $f(x_b)$ is assumed to be non-zero and (for the convenience of abbreviations, to not numerously use the long punctilious definition of the probability) can be named as the weight of the pair (couple) $w_{Pair} \equiv w_{Couple} \equiv w_C$ or simply w

$$w_{Pair} \equiv w_{Couple} \equiv w_C \equiv w > 0 \ .$$

$$f(x_a) + f(x_b) \equiv P(X = x_a) + P(X = x_b) \equiv w_C \equiv w > 0 \ .$$

The central moment $E_{Couple}(X_{Couple}-\mu)^n \equiv E_C(X_C-\mu)^n$ of this pair (couple) is

$$E_C(X_C - \mu)^n \equiv (x_a - \mu)^n f(x_a) + (x_b - \mu)^n f(x_b).$$

Its absolute value is limited by the sum of the absolute values of its components

$$\begin{aligned} | E_C(X_C - \mu)^n | &\le | (x_a - \mu)^n f(x_a) | + | (x_b - \mu)^n f(x_b) | = \\ &= (\mu - x_a)^n f(x_a) + (x_b - \mu)^n f(x_b) \end{aligned}.$$

3.2. Limiting function

Let us define a bounding function for a central moment of the pair. To not confuse the abbreviation of this function with the point b, this function will be named the limiting function L.

From the expressions of the balance and weight of the pair (couple)

$$f(x_b)=\frac{\mu-x_a}{x_b-\mu}f(x_A)=w_C-f(x_a)$$

and

$$\begin{aligned}f(x_a)+f(x_b)&=\frac{\mu-x_a+x_b-\mu}{x_b-\mu}f(x_a)=\\&=\frac{x_b-x_a}{x_b-\mu}f(x_a)=w_C\end{aligned},$$

one may replace $f(x_a)$ and $f(x_b)$ by functions of only x_a, μ, x_b and w_C

$$f(x_a)=\frac{x_b-\mu}{x_b-x_a}w_C \quad \text{and} \quad f(x_b)=\frac{\mu-x_a}{x_b-x_a}w_C$$

and obtain

$$\begin{aligned}|E_C(X-\mu)^n|&\le(\mu-x_a)^n f(x_a)+(x_b-\mu)^n f(x_b)=\\&=(\mu-x_a)^n\frac{x_b-\mu}{x_b-x_a}w_C+(x_b-\mu)^n\frac{\mu-x_a}{x_b-x_a}w_C\equiv\\&\equiv(\mu-x_a)^n\frac{x_b-\mu}{x_b-x_a}w+(x_b-\mu)^n\frac{\mu-x_a}{x_b-x_a}w\end{aligned}.$$

Definition 3.2. Limiting function. One may define a **limiting function** $L_C(x_a, \mu, x_b, n, w_C)$ or, abbreviated, $L(x_a, \mu, x_b, n, w)$ or simply L_C or L for a central moment of a pair (couple). This function depends only on x_a, μ, x_b, n, w_C

$$\begin{aligned}&L_{Couple}(x_a,\mu,x_b,n,w_{Couple})\equiv L_C(x_a,\mu,x_b,n,w_C)\equiv L(x_a,\mu,x_b,n,w)\equiv\\&\equiv(\mu-x_a)^n\frac{x_b-\mu}{x_b-x_a}w_C+(x_b-\mu)^n\frac{\mu-x_a}{x_b-x_a}w_C\end{aligned}.$$

Note, here x_a and x_b are the variables, but μ, n, but w_C are the parameters.

The absolute value of a central moment, say $|E_C(X_C-\mu)^n|$, of the pair (couple) is, by definition, limited (bounded) by this limiting function $L_C(x_a, \mu, x_b, n, w_C)$

$$|E_C(X_C-\mu)^n|\le L_C(x_a,\mu,x_b,n,w_C)\ .$$

3.3. Search for the maximum. Derivatives

Let us find the maximum of the limiting function $L_C(x_a, \mu, x_b, n, w_C)$ for x_a and x_b.

3.3.1. Differentiation with respect to x_a

Let us differentiate $L(x_a, \mu, x_b, n, w)$ with respect to x_a

$$\frac{\partial L(x_a,\mu,x_b,n,w)}{\partial x_a}=$$

$$=\frac{\partial\left((\mu-x_a)^n\dfrac{x_b-\mu}{x_b-x_a}w+(x_b-\mu)^n\dfrac{\mu-x_a}{x_b-x_a}w\right)}{\partial x_a}=$$

$$=\{[-n(x_b-x_a)+(\mu-x_a)](\mu-x_a)^{n-1}+$$

$$+[-(x_b-x_a)+(\mu-x_a)](x_b-\mu)^{n-1}\}\frac{x_b-\mu}{(x_b-x_a)^2}w=$$

$$=\{[(\mu-x_a)-n(x_b-x_a)](\mu-x_a)^{n-1}-(x_b-\mu)^n\}\frac{x_b-\mu}{(x_b-x_a)^2}w$$

At $n\geq 1$, if $(\mu-x_a)<(x_b-x_a)$, that is, if $x_b>\mu$ and $x_b-x_a>0$, then

$$(\mu-x_a)-n(x_b-x_A)<0$$

and

$$\frac{\partial L(x_a,\mu,x_b,n,w)}{\partial x_a}<0\,.$$

So, at $n\geq 1$, for $\mu<x_b\leq b$ (and, as can easily be seen, for $a\leq x_a<\mu$) the first derivative with respect to x_a is strictly less than zero. That is, for $a\leq x_a<\mu<x_b\leq b$ or for $[a, b]$ except for the specific point μ, we have

$$L(a,\mu,x_b,n,w)>L(x_a,\mu,x_b,n,w)\ .$$

If $(\mu-x_a)=(x_b-x_a)$, that is, if $x_b=\mu$, then from

$$(\mu-x_a)f(x_a)=(x_b-\mu)f(x_b)\ ,$$

we obtain

$$(\mu-x_a)=(\mu-\mu)\frac{f(x_b)}{f(x_a)}=0$$

or $x_a=\mu$.

To include the specific point μ into the ranges of variation of the arguments x_a and x_b of the inequality

$$L(a,\mu,x_b,n,w) > L(x_a,\mu,x_b,n,w) \ ,$$

let us estimate the derivative $\partial L(x_a, \mu, x_b, n, w)/\partial x_a$ for both $x_a \rightarrow \mu$ and $x_b \rightarrow \mu$. One may impose some natural conditions of non-zero values of probabilities: $f(x_a)>0$ and $f(x_b)>0$.

Let, say, μ-x_a be the basic term. Then

$$(x_b-\mu)=\frac{f(x_a)}{f(x_b)}(\mu-x_a)$$

and

$$x_b-x_a=(x_b-\mu)+(\mu-x_a)= \\ =\left(\frac{f(x_a)}{f(x_b)}+1\right)(\mu-x_a)=\frac{w}{f(x_b)}(\mu-x_a) \ .$$

If $x_a \rightarrow \mu$ then the derivative

$$\{[(\mu-x_a)-n(x_b-x_a)](\mu-x_a)^{n-1}-(x_b-\mu)^n\}\frac{x_b-\mu}{(x_b-x_a)^2}w=$$

$$=\left\{\left[1-n\frac{w}{f(x_b)}\right]-\left(\frac{f(x_a)}{f(x_b)}\right)^n\right\}(\mu-x_a)^n\frac{f(x_a)}{f(x_b)}\left(\frac{f(x_b)}{w}\right)^2 w(\mu-x_a)^{-1}= .$$

$$=\left\{\left[1-n\frac{w}{f(x_b)}\right]-\left(\frac{f(x_a)}{f(x_b)}\right)^n\right\}\frac{f(x_a)f(x_b)}{w}(\mu-x_a)^{n-1}\xrightarrow[n>1;\ x_a\to\mu]{}0$$

So (at $n>1$, if μ-x_a tends to 0, then the derivative)

$$\frac{\partial L(x_a,\mu,x_b,n,w)}{\partial x_a}\xrightarrow[n>1;\ x_a\to\mu]{}0 \ .$$

Therefore, for $a \le x_a \le \mu \le x_b \le b$, the derivative $\partial L(x_a, \mu, x_b, n, w)/\partial x_a \le 0$.

Let us include the point μ into the ranges of variation of the arguments x_a and x_b of the inequality $L(a, \mu, x_b, n, w)>L(x_a, \mu, x_b, n, w)$. Let us consider an intermediate point, say $x_a=(a+\mu)/2$.

If, for $a\le x_a\le\mu\le x_b\le b$, the derivative $\partial L(x_a, \mu, x_b, n, w)/\partial x_a\le 0$, then, for $a\le x_a\le\mu\le x_b\le b$, the function $L(x_a, \mu, x_b, n, w)\ge L(\mu, \mu, x_b, n, w)=L(\mu, \mu, \mu, n, w)$ (and $L((a+\mu)/2, \mu, x_b, n, w)\ge L(\mu, \mu, \mu, n, w)$).

If, for $a\le x_a<\mu<x_b\le b$, the derivative $\partial L(x_a, \mu, x_b, n, w)/\partial x_a<0$ then, for $a<x_a<\mu<x_b\le b$, the function $L(a, \mu, x_b, n, w)>L(x_a, \mu, x_b, n, w)$ and $L(a, \mu, x_b, n, w)>L((a+\mu)/2, \mu, x_b, n, w)$.

Therefore,

$$L(a,\mu,x_b,n,w) > L\left(\frac{a+\mu}{2},\mu,x_b,n,w\right) \ge L(\mu,\mu,\mu,n,w)$$

or

$$L_C(a,\mu,x_b,n,w) > L_C(\mu,\mu,\mu,n,w) \ .$$

We have included the specific point μ into the ranges of variation of arguments of the inequality $L(a, \mu, x_b, n, w)>L(x_a, \mu, x_b, n, w)$ and the inequality is true for $a\le x_a\le\mu\le x_b\le b$.

So, at $n>1$, the limiting function $L_C(x_a, \mu, x_b, n, w_C)$ has a maximum

$$Max(L_C(x_a,\mu,x_b,n,w_C)) = L_C(a,\mu,x_b,n,w_C) \ .$$

for x_a for the total interval $[a, b]$.

3.3.2. Differentiation with respect to x_b

Let us differentiate $L(x_a, \mu, x_b, n, w)$ with respect to x_b

$$\frac{\partial L(x_a,\mu,x_b,n,w)}{\partial x_b}=$$

$$=\frac{\partial\left((\mu-x_a)^n\frac{x_b-\mu}{x_b-x_a}w+(x_b-\mu)^n\frac{\mu-x_a}{x_b-x_a}w\right)}{\partial x_b}=$$

$$=\{[(x_b-x_a)-(x_b-\mu)](\mu-x_a)^{n-1}+$$

$$+[n(x_b-x_a)-(x_b-\mu)](x_b-\mu)^{n-1}\}\frac{\mu-x_a}{(x_b-x_a)^2}w=$$

$$=\{(\mu-x_a)^n+[n(x_b-x_a)-(x_b-\mu)](x_b-\mu)^{n-1}\}\frac{\mu-x_a}{(x_b-x_a)^2}w .$$

At $n\geq 1$, if $(x_b-x_a)>(x_b-\mu)$, that is, if $x_a<\mu$, then

$$n(x_b-x_a)-(x_b-\mu)>0$$

and (if $x_b-x_a>0$)

$$\frac{\partial L(x_a,\mu,x_b,n,w)}{\partial x_b}>0 .$$

If $(x_b-x_a)=(x_b-\mu)$, that is, if $x_a=\mu$, then $x_b=\mu$ (see above).

So, at $n\geq 1$, for $a\leq x_a<\mu<x_b<b$ the first derivative with respect to x_b is strictly greater than zero. That is, for $a\leq x_a<\mu<x_b<b$ or for $[a, b]$ except for the specific point μ, we have

$$L(x_a,\mu,x_b,n,w)<L(x_a,\mu,b,n,w) .$$

To include the specific point μ into the ranges of variation of the arguments x_a and x_b, let us estimate the derivative $\partial L(x_a, \mu, x_b, n, w)/\partial x_b$ for both $x_b\rightarrow\mu$ and $x_a\rightarrow\mu$ under the same natural conditions of non-zero values of probabilities: $f(x_a)>0$ and $f(x_b)>0$.

Let, say, $x_b-\mu$ be the basic term. Then

$$(\mu-x_a)=\frac{f(x_b)}{f(x_a)}(x_b-\mu)$$

and

$$x_b-x_a=\left(1+\frac{f(x_b)}{f(x_a)}\right)(x_b-\mu)=\frac{w}{f(x_a)}(x_b-\mu) .$$

If $x_b \rightarrow \mu$, then the derivative

$$\{(\mu - x_a)^n + [n(x_b - x_a) - (x_b - \mu)](x_b - \mu)^{n-1}\}\frac{\mu - x_a}{(x_b - x_a)^2} w =$$

$$= \left\{ \left(\frac{f(x_b)}{f(x_a)} \right)^n + \left[n \frac{w}{f(x_a)} - 1 \right] \right\} (x_b - \mu)^{n-1} \frac{f(x_b)}{f(x_a)} \left(\frac{f(x_a)}{w} \right)^2 w = \quad .$$

$$= \left\{ \left(\frac{f(x_b)}{f(x_a)} \right)^n + \left[n \frac{w}{f(x_a)} - 1 \right] \right\} \frac{f(x_b) f(x_a)}{w} (x_b - \mu)^{n-1} \xrightarrow[n>1;\ x_b \to \mu]{} 0$$

So (for $n>1$, if x_b (and x_a) tend to μ, then)

$$\frac{\partial L(x_a, x_b, x_b, n, w)}{\partial x_b} \xrightarrow[n>1;\ x_b \to x_b]{} 0 \ .$$

Therefore, for $a \le x_a \le \mu \le x_b \le b$, the derivative $\partial L(x_a, \mu, x_b, n, w)/\partial x_b \ge 0$.

Let us include the specific point μ into the ranges of variation of the arguments x_a and x_b of the inequality $L(x_a, \mu, b, n, w) > L(x_a, \mu, x_b, n, w)$. Let us consider an intermediate point, say $x_b = (\mu+b)/2$.

If, for $a \le x_a \le \mu \le x_b \le b$, the derivative $\partial L(x_a, \mu, x_b, n, w)/\partial x_b \ge 0$ then, for $a \le x_a \le \mu \le x_b \le b$, the function $L(x_a, \mu, \mu, n, w) = L(\mu, \mu, \mu, n, w) \le L(x_a, \mu, x_b, n, w)$ (and $L(\mu, \mu, \mu, n, w) \le L((x_a, \mu, (\mu+b)/2, n, w))$.

If, for $a \le x_a < \mu < x_b \le b$, the derivative $\partial L(x_a, \mu, x_b, n, w)/\partial x_b > 0$ then, for $a \le x_a < \mu < x_b < b$, the function $L(x_a, \mu, x_b, n, w) < L(x_a, \mu, b, n, w)$ and $L((a+\mu)/2, \mu, x_b, n, w) < L(x_a, \mu, b, n, w)$.

Therefore,

$$L(\mu, \mu, \mu, n, w) \le L\left(x_a, \mu, \frac{a+\mu}{2}, n, w \right) < L(x_a, \mu, b, n, w)$$

or

$$L_C(\mu, \mu, \mu, n, w) < L_C(x_a, \mu, b, n, w) \ .$$

We have included the specific point μ into the ranges of variation of arguments of the inequality $L(x_a, \mu, x_b, n, w) < L(x_a, \mu, b, n, w)$ and the inequality is true for $a \le x_a \le \mu \le x_b \le b$.

So, at $n>1$, the limiting function $L_C(x_a, \mu, x_b, n, w_C)$ has a maximum

$$Max(L_C(x_a, \mu, x_b, n, w_C)) = L_C(x_a, \mu, b, n, w_C) \ .$$

for x_b for the total interval $[a, b]$.

3.3.3. The maximum

So, at $n>1$, for $a \le x_a \le \mu \le x_b \le b$, the limiting function

$$L_C(x_a,\mu,x_b,n,w) = (\mu - x_a)^n \frac{x_b - \mu}{x_b - x_a} w + (x_b - \mu)^n \frac{\mu - x_a}{x_b - x_a} w$$

attains its maximum at the borders $x_a = a$ and $x_b = b$ of the interval $[a, b]$

$$Max(L_C(x_a,\mu,x_b,n,w_C)) = L_C(a,\mu,b,n,w_C) =$$
$$= (\mu - a)^n \frac{b-\mu}{b-a} w_C + (b-\mu)^n \frac{\mu - a}{b-a} w_C .$$

So, at $n>1$, the absolute value $|E_{Couple}(X-\mu)^n| \equiv |E_C(X-\mu)^n|$ of a central moment of the pair (x_a, x_b) is limited by the maximal limiting function L_C, that is concentrated at the borders $x_a=a$ and $x_b=b$ of the interval $[a, b]$

$$|E_C(X-\mu)^n| \le L_C(a,\mu,b,n,w_C) =$$
$$= (\mu - a)^n \frac{b-\mu}{b-a} w_C + (b-\mu)^n \frac{\mu - a}{b-a} w_C .$$

3.4. Representation by pairs. Succession of situations

3.4.1. Preliminary considerations

Let us analyze whether the total probability (weight)

$$\sum_{k=1}^{K} f(x_k) \equiv P(\Omega) \equiv W_K ,$$

and central moments

$$E(X-\mu)^n = \sum_{k=1}^{K} (x_k-\mu)^n f(x_k) .$$

of the variable X of Section 2 can be exactly represented by those of a set of pairs.

The final goal of this section is to exactly represent the modulus of any central moment of any variable X of Section 2 by a sum of moduli of central moments of a set of pairs of the same variable and to estimate this sum by the limiting functions.

The discrete random variable X can be treated as a set of points $\{x_k\}$. The probability mass function f of Section 2 can be also treated as a set of values $\{f(x_k)\}$ associated with $\{x_k\}$. A pair (x_a, x_b) defined in this section is a subset of the set $\{x_k\}$. If there are $K.C : K.C \geq 1$ pairs then, if there is a need, one can denote the $k.C^{th}$ pair (couple), such that $k.C=1, \ldots, K.C$, as $\{x_{k.C.a}, x_{k.C.b}\}$. The weight of this pair can be denoted as $w_{k.C}$. (The multiple notation, e.g. $x_{k.C.a}$, is used to avoid numerous three-storey and even four-storey indices in the text).

In this subsection we should often distinguish between points, values, objects, etc. associated with pairs (couples) and values, objects, etc. those are (still) not associated with pairs. To do this, let us denote points, values, objects, etc. associated with pairs (couples) as points, values of pairs (couples), pair's value, pairs' values, etc. and the objects, values, etc. those are (still) not associated with pairs as the **original** points, values, original objects, etc.

Linearity of sums

Let us mention the linearity of sums of weights and moments.

The total weight

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

and moments

$$E(X - x_0)^n = \sum_{k=1}^{K} (x_k - x_0)^n f(x_k).$$

of X depend linearly on the values $f(x_k)$. The sum is their linear function also. Therefore:

1) the total weight of a sum equals the sum of the weights and
2) the moment of a sum equals the sum of the moments.

The sum of the central moments of the pairs is limited by the sum of the maximal limiting functions (those are linear functions of $f(x_k)$ as well) of these pairs. One can see, indeed, that if for $k.C^{th}$ pair

$$|E_{k.C}(X_{k.C} - \mu)^n| \le L_{k.C}(a, \mu, b, n, w_{k.C}) ,$$

then for $K.C$ pairs

$$\sum_{k.C=1}^{K.C} |E_{k.C}(X_{k.C} - \mu)^n| \le \sum_{k.C=1}^{K.C} L_{k.C}(a, \mu, b, n, w_{k.C}) .$$

3.4.2. Situations

Let us divide the points x_k into three groups:

1) $x_{k.a} < \mu$,
2) $x_{k.\mu} = \mu$ (zero central moment(s)),
3) $x_{k.b} > \mu$.

Let us introduce the numbers $K.a$, $K.\mu$ and $K.b$, such that $k.a \le K.a$, $k.\mu \le K.\mu$, $k.b \le K.b$ and

$$K.a + K.\mu + K.b = K .$$

Owing to $x_{k.\mu} - \mu \equiv 0$, an arbitrary non-zero central moment depends only on $K.a$ and $K.b$. Let us consider in turn situations with various numbers

$$K.ab \equiv K.a + K.b .$$

from $K.ab = 0$ to the general situation.

Situation $K.ab=0$

Due to the condition $K \geq 2$ of Section 2 and $K.\mu \leq 1$, the case $K.ab<1$ cannot exist.

Nevertheless, let us consider optionally more general (or fictitious) cases of $K=1$ and of mutually coincident points $\{x_{k.\mu}=\mu\} : k.\mu=1, \ldots, K.\mu : K.\mu \geq 2$.

If $K.ab=0$, then only one point μ (or mutually coincident points $\{x_{k.\mu}=\mu\}$) and the corresponding value $f(\mu)$ (or the values $f(x_{k.\mu})$) can exist. Evidently, the value $f(\mu)$ (or the values $f(x_{k.\mu})$) do not contribute to the non-zero central moments.

All the mutually coincident points $\{x_{k.\mu}=\mu\}$ (or the single point) may be represented as only one aggregated point $x_{Aggr.\mu}=\mu$ and the corresponding value

$$f_{Aggr}(\mu) \equiv \sum_{k.\mu=1}^{K.\mu} f(x_{k.\mu}) \ .$$

We may formally divide the value $f_{Aggr}(x_{Aggr.\mu}) \equiv f(\mu)$ into two parts $f_{1.C}(\mu)$ and $f_{2.C}(\mu)$ satisfying $f_{1.C}(\mu)=f_{2.C}(\mu)=f(\mu)/2$. The balance formally remains

$$(\mu-\mu) f_{1.C}(\mu) = (\mu-\mu) f_{2.C}(\mu)$$

or

$$(\mu-\mu)\frac{f(\mu)}{2} = (\mu-\mu)\frac{f(\mu)}{2} \ .$$

Evidently, the total weight of this formal pair equals the total weight $f(\mu)$ (or the sum the weights $f(x_{k.\mu})$). The central moments equal zero for both the pair and the point μ (or the points $\{x_{k.\mu}\}$). So, the total weight and central moments of the point μ (or the points $\{x_{k.\mu}\}$) can be exactly represented by a pair of the previous subsections.

Further, as a rule, we shall not consider the point(s) $x_k=\mu$.

Situation $K.ab=1$

Here, only two possible cases can take place: the case $K.a=1$ and $K.b=0$ or the case $K.a=0$ and $K.b=1$.

Generally, the first central moment

$$\sum_{k=1}^{K}(x_k-\mu)f(x_k)\equiv 0$$

may be transformed to

$$\sum_{k=1}^{K}(x_k-\mu)f(x_k)=\sum_{k.a\le K.a}(x_{k.a}-\mu)f(x_{k.a})+$$
$$+\sum_{k.\mu\le K.\mu}(x_{k.\mu}-\mu)f(x_{k.\mu})+\sum_{k.b\le K.b}(x_{k.b}-\mu)f(x_{k.b})=0 ,$$

where the limits of the sums $k.a\le K.a$, $k.\mu\le K.\mu$ and $k.b\le K.b$ denote, that $K.\mu$ or $K.a$ or $K.b$ can equal zero. That is, generally, there can be cases with no members of the sum(s) of $k.\mu$ or $k.a$ or $k.b$.

Now, since

$$x_{k.\mu}-\mu\equiv 0 ,$$

this central moment may be transformed to the balance

$$\sum_{k.a\le K.a}(\mu-x_{k.a})f(x_{k.a})=\sum_{k.b\le K.b}(x_{k.b}-\mu)f(x_{k.b}) .$$

Suppose $K.a=1$ and $K.b=0$. Then

$$\sum_{k.a\le K.a}(\mu-x_{k.a})f(x_{k.a})=0 .$$

There are only two possible cases: $f(x_{k.a})>0$ and $f(x_{k.a})=0$. Evidently, for $K.ab=1$, the case $f(x_{k.a})>0$ cannot exist. If $f(x_{k.a})=0$ then the balance can formally hold, but this case does not contribute to the non-zero central moments $E(X-\mu)^n>0$.

The consideration of the case $K.a=0$ and $K.b\ge 1$ is fully analogous to the preceding one.

So, the case $K.a=0$ and $K.b\ge 1$ and the case $K.a\ge 1$ and $K.b=0$ either cannot occur or do not contribute to the non-zero central moments $E(X-\mu)^n>0$.

So, Situation $K.ab=1$ cannot occur or does not contribute to the non-zero central moments.

Further, as a rule, we shall not consider those cases that do not contribute to the non-zero central moments, namely $x_k : f(x_k)=0$ and $x_k=\mu$.

Situation *K.ab=2*

Here, the only possible case which contributes to the non-zero central moments, is the case *K.a=1* and *K.b=1*.

If *K.a=1* and *K.b=1*, then we have the balance

$$(\mu - x_{1.a}) f(x_{1.a}) = (x_{1.b} - v) f(x_{1.b}) \ .$$

Therefore, the original points $x_{1.a}$ and $x_{1.b}$ are the required pair (couple) of the previous subsections.

Evidently, the total weight and moments of the pair are equal to those of the original points.

So, the original total weight and central moments of Situation *K.ab=2* can be exactly represented by the total weight and central moments of a pair of the previous subsections.

Remark 3.3

Let us further, for definiteness, enumerate the points $x_{k.a}$ and $x_{k.b}$, for example, from those furthest from μ, to those closest to μ.

Divided sets

Let us define "divided" or "exactly divided" sets.

Definition 3.4. Let us suppose given an initial set of points $\{x_k\}$ and the initial set of values $\{f(x_k)\}$ associated with $\{x_k\}$ as in Section 2.

A divided or exactly **divided set of points** $\{x_k\}$ (with respect to the **initial set of points**) is defined as the same initial set of points $\{x_k\}$ such that at least one value $f(x_k)$ (associated with a point x_k) is divided into, at least, two parts $f_1(x_k)$ and $f_2(x_k)$ satisfying the equality

$$f(x_k) = f_1(x_k) + f_2(x_k) \ .$$

A divided or exactly **divided set of values** (with respect to the **initial set of values**) is the set of values associated with the divided set of points.

The notation of a divided value may be more complex, e.g.

$$f(x_k) \equiv f_{1(k)}(x_k) + f_{2(k)}(x_k)$$

or, more generally,

$$f(x_k) \equiv \sum_{d(k)=1}^{D(k)} f_{d(k)}(x_k) \quad : \ 2 \le D(k) < \infty \ .$$

More generally, every value $f(x_k)$ (that will be either divided or not divided) of the initial set of values $\{f(x_k)\}$ may be written via the values $f_{d(k)}(x_k)$ of the exactly divided set $\{f_{d(k)}(x_k)\}$, by definition, as

$$f(x_k) \equiv \sum_{d(k)=1}^{D(k)} f_{d(k)}(x_k) \quad : \ 1 \le D(k) < \infty \ .$$

Note, the divided set of points and the initial set of points are the same sets. The divided set of values and the set of initial values differ from each other. Because of these properties, there is a reason to distinguish between divided and initial sets of points by the associated sets of values.

Note, that a divided set of points can serve as the new initial set of points for a subsequent division, i.e., modification.

Evidently, the total weight and moments of the divided set of points are equal to those of the initial set of points.

Let us consider the total weight and moments of a divided set of points.

By the general definition (see above), the total weight of the divided values $f_{d(k)}(x_k)$ is equal to the initial value $f(x_k)$

$$f(x_k) = \sum_{d(k)=1}^{D(k)} f_{d(k)}(x_k)$$

for every initial value $f(x_k)$. Therefore, the total weight of the divided set is equal to that of the initial set.

Both the divided values $f_{d(k)}(x_k)$ and the initial value $f(x_k)$ are associated with the same point x_k. Therefore and by the general definition, the sum of moments of every divided point is equal to the moment of the initial point

$$\sum_{d(k)=1}^{D(k)} (x_k - x_0)^n f_{d(k)}(x_k) = (x_k - x_0)^n \sum_{d(k)=1}^{D(k)} f_{d(k)}(x_k) = (x_k - x_0)^n f(x_k) \ .$$

Therefore, the total moment of the whole divided set is equal to that of the whole initial set.

One can see, indeed, that, by definition, the total weight W_D of the exactly divided set of points is

$$W_D \equiv \sum_{k=1}^{K} \sum_{d(k)=1}^{D(k)} f_{d(k)}(x_k) \equiv \sum_{k=1}^{K} f(x_k) \equiv W_K$$

and the total moment $E_D(X-x_0)^n$ of the exactly divided set of points is

$$\begin{aligned} &E_D(X - x_0)^n \equiv \\ &\equiv \sum_{k=1}^{K} \sum_{d(k)=1}^{D(k)} (x_k - x_0)^n f_{d(k)}(x_k) \equiv \sum_{k=1}^{K} (x_k - x_0)^n f(x_k) \equiv \ . \\ &\equiv E(X - x_0)^n \end{aligned}$$

So, we have specified the properties of the divided sets: the total weight and moments of a divided set of points are equal to the total weight and moments of the initial set of points.

Situation $K.ab = 3$

Here, there are only two possible cases those can contribute to the non-zero central moments: the case of $K.a=2$ and $K.b=1$, or the case of $K.a=1$ and $K.b=2$.

Let us consider the case of $K.a=2$ and $K.b=1$.

Let us make the first step of the representation of the total weight and central moments of the original set of points by the total weight and central moments of the set of the pairs.

The value $f(x_{1.b})$ can be exactly divided into two parts $f_1(x_{1.b})$ and $f_2(x_{1.b})$ satisfying the balance

$$(\mu - x_{1.a}) f(x_{1.a}) = (x_{1.b} - \mu) f_1(x_{1.b})$$

and the equality of divided sets

$$f_2(x_{1.b}) = f(x_{1.b}) - f_1(x_{1.b}) \; .$$

Here, the points $x_{1.a}$ and $x_{1.b}$ are the initial set of points. The divided points are the same points. The values $f(x_{1.a})$ and $f(x_{1.b})$ are the initial set of values. The divided values are $f(x_{1.a})$, $f_1(x_{1.b})$ and $f_2(x_{1.b})$.

Due to the properties of the divided sets, the total weight and moments of the divided set of points are equal to those of the initial set of points.

The first portion of the original set of points is the set $x_{1.a}$ and $x_{1.b}$ of the divided set with the associated values $f(x_{1.a})$ and $f_1(x_{1.b})$. Since the balance

$$(\mu - x_{1.a}) f(x_{1.a}) = (x_{1.b} - \mu) f_1(x_{1.b})$$

is true, the two points $x_{1.a}$ and $x_{1.b}$ of the divided set with the associated values $f(x_{1.a})$ and $f_1(x_{1.b})$ are the required pair of the previous subsections. Evidently, the total weight and moments of the pair are equal to those of the first portion of the original set of points.

So, the first step of the representation has been done. The total weight and moments of the pair as of the first portion of the set of the pairs are equal to those of the first portion of the divided original set of points.

This can be seen in more detail for the central moments

$$\begin{aligned}
E(X-\mu)^n &= (x_{1.a}-\mu)^n f(x_{1.a}) + (x_{2.a}-\mu)^n f(x_{2.a}) + (x_{1.b}-\mu)^n f(x_{1.b}) = \\
&= (x_{1.a}-\mu)^n f(x_{1.a}) + (x_{1.b}-\mu)^n f_1(x_{1.b}) + \\
&+ (x_{2.a}-\mu)^n f(x_{2.a}) + (x_{1.b}-\mu)^n f_2(x_{1.b}) = \\
&= E_{1.C}(X-\mu)^n + \\
&+ (x_{2.a}-\mu)^n f(x_{2.a}) + (x_{1.b}-\mu)^n f_2(x_{1.b})
\end{aligned} \; .$$

Let us make the second step of the representation.

The balance remains

$$(\mu - x_{1.a}) f(x_{1.a}) + (\mu - x_{2.a}) f(x_{2.a}) = \\ = (x_{1.b} - \mu) f_1(x_{1.b}) + (x_{1.b} - \mu) f_2(x_{1.b}) \text{ ,}$$

and we come to Situation *K.ab=2* for $f(x_{2.a})$ and $f_2(x_{1.b})$

$$(\mu - x_{2.a}) f(x_{2.a}) = (x_{1.b} - \mu) f_2(x_{1.b}) \text{ .}$$

So, as a result of the first step, the number of unpaired values is diminished by one and we come to resulting Situation $K.ab_{Diminished}=K.ab-1=2$.

As has been proved above, the total weight and central moments of Situation *K.ab=2* can be exactly represented by the total weight and central moments of a pair of the previous subsections. So, this is the second and final step.

So, the final step of the representation has been done. The total weight and moments of the final portion of the set of the pairs of points are equal to those of the final portion of the divided original set of points.

This can be seen in more detail for the central moments

$$E(X - \mu)^n = E_{1.C}(X - \mu)^n + (x_{2.a} - \mu)^n f(x_{2.a}) + (x_{1.b} - \mu)^n f_2(x_{1.b}) = \text{ .} \\ = E_{1.C}(X - \mu)^n + E_{2.C}(X - \mu)^n \text{ .}$$

So, Situation *K.ab=3*, at *K.a=2* and *K.b=1*, can be represented by the sum of the first step and the final step.

So, the total weight and moments of the divided original set of points are equal to those of the initial original set of points. For every step, the total weight and moments of the portion of the set of the pairs are equal to those of the portion of the divided original set of points. Both the total weight and moments depend linearly on the values of the members of the sets. Therefore, the total weight and moments of the sum of the portions are equal to the sum of the constituent weights and moments correspondingly. Therefore, for whole Situation *K.ab=3*, the total weight and moments of the set of the pairs are equal to those of the original set of points.

If *K.a=1* and *K.b=2*, then the consideration is analogous to the preceding one.

So, the total weight and central moments of Situation *K.ab=3* can be exactly represented by the total weight and central moments of a set of pairs (couples) of the previous subsections.

General Situation *K.ab*. Suppose general Situation $K.ab \geq 4$, $K.a \geq 1$ and $K.b \geq 1$ (the case of $K.a=0$ and $K.b \geq 1$ and the case of $K.b=0$ and $K.a \geq 1$ cannot exist or do not contribute to the non-zero central moments.

Let us consider $f(x_{1.a})$ and $f(x_{1.b})$. There are only two possible variants: less possible but more easy Variant 1 (equality)

$$(\mu - x_{1.a}) f(x_{1.a}) = (x_{1.b} - \mu) f(x_{1.b}) \ .$$

and more possible but less easy Variant 2 (inequality)

$$(\mu - x_{1.a}) f(x_{1.a}) \neq (x_{1.b} - \mu) f(x_{1.b})$$

Let us make the first step of the representation of the total weight and moments of the original set of points by the total weight and central moments of the set of the pairs. Evidently, this first step may be implemented in one of the two forms depending on whether Variant 1 (equality) or Variant 2 (inequality) takes place.

Variant 1 (equality). If

$$(\mu - x_{1.a}) f(x_{1.a}) = (x_{1.b} - \mu) f(x_{1.b})$$

then the two points $x_{1.a}$ and $x_{1.b}$ are the required pair (couple) of the previous subsections. Therefore, the total weight and an arbitrary total moment of the pair are the same as those of the portion of the original set.

As a result of this first step within the scope of Variant 1 (equality), the number of unpaired (uncoupled) values is diminished by two and from Situation *K.ab* we come to Situation $K.ab_{Diminished}=K.ab-2$. Here, the number $K.ab_{Diminished}=K.ab-2$ is composed of $2, \ldots, K.a$ and $2, \ldots, K.b$.

Let us make the first step of the representation within the scope of Variant 2 (inequality).

Variant 2 (inequality). If

$$(\mu - x_{1.a}) f(x_{1.a}) \neq (x_{1.b} - \mu) f(x_{1.b}) ,$$

then there are only two possible cases as well:

$$(\mu - x_{1.a}) f(x_{1.a}) < (x_{1.b} - \mu) f(x_{1.b})$$

and

$$(\mu - x_{1.a}) f(x_{1.a}) > (x_{1.a} - \mu) f(x_{1.a}) .$$

Suppose, for example, that

$$(\mu - x_{1.A}) f(x_{1.A}) < (x_{1.B} - \mu) f(x_{1.B}) .$$

Then one should divide the value $f(x_{1.b})$ into two parts $f_1(x_{1.b})$ and $f_2(x_{1.b})$ satisfying the balance

$$(\mu - x_{1.A}) f(x_{1.A}) = (x_{1.B} - \mu) f_1(x_{1.B})$$

and the equality of divided sets

$$f_2(x_{1.B}) = f(x_{1.B}) - f_1(x_{1.B}) .$$

Here, the points $x_{1.a}$ and $x_{1.b}$ are the initial set of points. The divided points are the same points. The values $f(x_{1.a})$ and $f(x_{1.b})$ are the initial set of values. The divided values are $f(x_{1.a})$, $f_1(x_{1.b})$ and $f_2(x_{1.b})$.

Due to the properties of the divided sets, the total weight and moments of the divided set of points are equal to those of the initial set of points.

The first portion of the original set of points is the subset $(x_{1.a}, x_{1.b})$ of the divided set with the associated values $f(x_{1.a})$ and $f_1(x_{1.b})$. Since the balance

$$(\mu - x_{1.a}) f(x_{1.a}) = (x_{1.b} - \mu) f_1(x_{1.b})$$

is true, two points $x_{1.a}$ and $x_{1.b}$ of the divided set with the values $f(x_{1.a})$ and $f_1(x_{1.b})$ are the required pair of the previous subsections. Evidently, the total weight and moments of the pair are equal to those of the first portion of the original set of points.

So, within the scope of Variant 2 (inequality), the first step of the representation has been done. The total weight and moments of the pair as of the first portion of the set of the pairs are equal to those of the first portion of the divided original set of points.

As a result of this first step within the scope of Variant 2 (inequality), the number of unpaired (uncoupled) values is diminished by one (taking into account the part $f_2(x_{1.b})$ of the value $f(x_{1.b})$) and we come to Situation $K.ab_{Diminished}=K.ab-1$. Note, that the number $K.ab$ is composed of $1, \ldots, K.a$ and $1, \ldots, K.b$. And here, the number $K.ab_{Diminished}=K.ab-1$ is composed of $2, \ldots, K.a$ and $2, \ldots, K.b$ plus one.

So, we have considered the first step of diminishing the number $K.ab$ for general Situation $K.ab \geq 4$ within the scopes of both parallel variants. It diminishes $K.ab$ by one or two.

Evidently, such a step may be a general intermediate one.

Let us suppose a general intermediate situation such that there are already *g.C.ab : 1 < g.C.ab < K.C.ab*, pairs (couples) which represent the total weight and moments of some original points and there are still *K.a - g.a + 1* of $x_{k.a}$ points and *K.b - g.b + 1* of $x_{k.b}$ points, the total weight and moments of which are still not represented by those of pairs. Let us represent the total weight and moments for this general intermediate situation. For illustrativeness, examples of this general intermediate step may be written via formulae.

The total weight for the general intermediate situation before the general intermediate step can be represented as

$$W_{K.ab} = \sum_{k.C.ab=1}^{g.C.ab} w_{k.C.ab} + \sum_{k.a=g.a}^{K.a} f(x_{k.a}) + \sum_{k.b=g.b}^{K.b} f(x_{k.b}) \ .$$

The central moments for the general intermediate situation before the general intermediate step can be represented as

$$E(X-\mu)^n =$$
$$= \sum_{k.C.ab=1}^{g.C.ab} E_{k.C.ab}(X-\mu)^n + \sum_{k.a=g.a}^{K.a} (x_{k.a}-\mu)^n f(x_{k.a}) + \sum_{k.b=g.b}^{K.b} (x_{k.b}-\mu)^n f(x_{k.b}) \ .$$

Variant 1 (equality). The general intermediate step can be seen in more detail for the total weights

$$W_{K.ab} = \sum_{k.C.ab=1}^{g.C.ab} w_{k.C.ab} + f(x_{g.a}) + f(x_{g.b}) + \sum_{k.a=g.a+1}^{K.a} f(x_{k.a}) + \sum_{k.b=g.b+1}^{K.b} f(x_{k.b}) =$$
$$= \sum_{k.C.ab=1}^{g.C.ab+1} w_{k.C.ab} + \sum_{k.a=g.a+1}^{K.a} f(x_{k.a}) + \sum_{k.b=g.b+1}^{K.b} f(x_{k.b}) \ .$$

The general intermediate step can be seen in more detail for the central moments

$$E(X-\mu)^n =$$
$$= \sum_{k.C.ab=1}^{g.C.ab} E_{k.C.ab}(X-\mu)^n + (x_{gk.a}-\mu)^n f(x_{g.a}) + (x_{g.b}-\mu)^n f(x_{g.b}) +$$
$$+ \sum_{k.a=g.a+1}^{K.a} (x_{k.a}-\mu)^n f(x_{k.a}) + \sum_{k.b=g.b+1}^{K.b} (x_{k.b}-\mu)^n f(x_{k.b}) =$$
$$= \sum_{k.C.ab=1}^{g.C.ab} E_{k.C.ab}(X-\mu)^n + \sum_{k.a=g.a+1}^{K.a} (x_{k.a}-\mu)^n f(x_{k.a}) + \sum_{k.b=g.b+1}^{K.b} (x_{k.b}-\mu)^n f(x_{k.b}) \ .$$

Variant 2 (inequality). The general intermediate step can be seen in more detail for the total weights

$$W_{K.ab} = \sum_{k.C.ab=1}^{g.C.ab} w_{k.C.ab} + f(x_{g.a}) + f_1(x_{g.b}) + f_2(x_{g.b}) +$$
$$+ \sum_{k.a=g.a+1}^{K.a} f(x_{k.a}) + \sum_{k.b=g.b+1}^{K.b} f(x_{k.b}) =$$
$$= \sum_{k.C.ab=1}^{g.C.ab+1} w_{k.C.ab} + \sum_{k.a=g.a+1}^{K.a} f(x_{k.a}) + f_2(x_{g.b}) + \sum_{k.b=g.b+1}^{K.b} f(x_{k.b})$$

.

The general intermediate step can be seen in more detail for the central moments

$$E(X-\mu)^n == \sum_{k.C.ab=1}^{g.C.ab} E_{k.C.ab}(X-\mu)^n +$$
$$+ (x_{gk.a}-\mu)^n f(x_{g.a}) + (x_{g.b}-\mu)^n f_1(x_{g.b}) + (x_{g.b}-\mu)^n f_2(x_{g.b}) +$$
$$+ \sum_{k.a=g.a+1}^{K.a} (x_{k.a}-\mu)^n f(x_{k.a}) + \sum_{k.b=g.b+1}^{K.b} (x_{k.b}-\mu)^n f(x_{k.b}) =$$
$$= \sum_{k.C.ab=1}^{g.C.ab} E_{k.C.ab}(X-\mu)^n + \sum_{k.a=g.a+1}^{K.a} (x_{k.a}-\mu)^n f(x_{k.a}) +$$
$$+ (x_{g.b}-\mu)^n f_2(x_{g.b}) + \sum_{k.b=g.b+1}^{K.b} (x_{k.b}-\mu)^n f(x_{k.b})$$

.

So, we have considered the general step of diminishing the number $K.ab$ for general Situation $K.ab \geq 4$ within the scopes of both parallel variants. It diminishes $K.ab$ by one or two.

Evidently, this general step may be repeated as many times as needed to reach final Situations $K.ab_{Diminished}=3$ or $K.ab_{Diminished}=2$.

If the original set of points is divided, then the total weight and moments of the original set of points are equal to those of the initial original set of points. For every step, the total weight and moments of the portion of the set of the pairs are equal to those of the portion of the divided original set of points. Both the total weight and moments depend linearly on the values of the members of the sets. Therefore, the total weight and moments of the sum of the portions are equal to the sum of the constituent weights and moments correspondingly. Therefore, for whole general Situation $K.ab$, the total weight and moments of the set of the pairs are equal to those of the original set of points.

So, in general Situation $K.ab : K.ab \geq 4$, at $K.a \geq 1$ and $K.b \geq 1$, the total weight and central moments of a discrete random variable X of Section 2 may be exactly represented by the total weight and central moments of the pairs of this section.

3.5. General limitations

3.5.1. Weights

Let us consider the weights (probabilities) of groups of realizations (points) x_k of X, of groups of pairs (couples) and general limitations on them.

Remembering

$$K.a + K.\mu + K.b = K$$

of the preceding subsection, the total weights of these groups may be denoted as W_a, W_μ and W_b such that

$$W_a \equiv \sum_{k.a \le K.a} f(x_{k.a}) \ , \ W_\mu \equiv \sum_{k.\mu \le K.\mu} f(x_{k.\mu}) \ , \ W_b \equiv \sum_{k.b \le K.b} f(x_{k.b})$$

and the sum of the weights (probabilities) is

$$W_a + W_\mu + W_b = W_K \equiv P(\Omega) = 1 \ .$$

Let us denote the total weight of the total set of all the pairs (couples) as $W_{Couple} \equiv W_C$, the weight of the set of the formal pairs $(\mu_{k.C.\mu}, \mu_{k+1.C.\mu})$ as $W_{C.\mu}$ and the total weight of the set of the pairs $(x_{k.C.a}, x_{k.C.b})$ as $W_{C.ab}$. By this definition, the weight of, e.g., a $k.C.ab^{\text{th}}$ pair (couple) $(x_{k.C.a}, x_{k.C.b})$, is denoted as $w_{k.C.ab}$ and

$$\sum_{k.C=1}^{K.C} w_{k.C} \equiv W_C \ , \quad \sum_{k.C.\mu \le K.C.\mu} w_{k.C.\mu} \equiv W_{C.\mu} \ , \quad \sum_{k.C.ab=1}^{K.C.ab} w_{k.C.ab} \equiv W_{C.ab} \ ,$$

and we have

$$W_C = W_{C.\mu} + W_{C.ab} \ .$$

Evidently,

$$W_{C.\mu} = W_\mu \ ,$$

and, due to the preceding subsection,

$$W_{C.ab} = W_a + W_b$$

and

$$W_C = W_K \equiv P(\Omega) = 1 \ .$$

3.5.2. The general limiting function

Let us consider the central moments

$$E(X-\mu)^n = \sum_{k.Couple=1}^{K.Couple} E_{k.Couple}(X-\mu)^n = \sum_{k.C=1}^{K.C} E_{k.C}(X-\mu)^n =$$
$$= \sum_{k.C.\mu \le K.C.\mu} E_{k.C.\mu}(X-\mu)^n + \sum_{k.C.ab=1}^{K.C.ab} E_{k.C.ab}(X-\mu)^n = \qquad .$$
$$= \sum_{k.C.ab=1}^{K.C.ab} E_{k.C.ab}(X-\mu)^n$$

The maximal limiting functions $L_{k.C.ab}(a, \mu, b, n, w_{k.C.ab})$ satisfying

$$| E_{k.C.ab}(X-\mu)^n | \le L_{k.C.ab}(a,\mu,b,n,w_{k.C.ab}) \ ,$$

allow estimating the central moments $E(X\text{-}\mu)^n$ of the random variable X

$$| E(X-\mu)^n | \le | \sum_{k.C.ab=1}^{K.C.ab} E_{k.C.ab}(X-\mu)^n | \le$$
$$\le \sum_{k.C.ab=1}^{K.C.ab} L_{k.C.ab}(a,\mu,b,n,w_{k.C.ab}) \qquad .$$

This estimate can be easily simplified. From

$$L_{k.C.ab}(a,\mu,b,n,w_{k.C.ab}) =$$
$$= (\mu-a)^n \frac{b-\mu}{b-a} w_{k.C.ab} + (b-\mu)^n \frac{\mu-a}{b-a} w_{k.C.ab} =$$
$$= \left[(\mu-a)^n \frac{b-\mu}{b-a} + (b-\mu)^n \frac{\mu-a}{b-a} \right] w_{k.C.ab}$$

there follows

$$\sum_{k.C.ab=1}^{K.C.ab} L_{k.C.AB}(a,\mu,b,n,w_{k.C.ab}) =$$
$$= \sum_{k.C.ab=1}^{K.C.ab} \left[(\mu-a)^n \frac{b-\mu}{b-a} + (b-\mu)^n \frac{\mu-a}{B-a} \right] w_{k.C.ab} = \quad .$$
$$= \left[(\mu-a)^n \frac{b-\mu}{b-a} + (b-\mu)^n \frac{\mu-a}{b-a} \right] \sum_{k.C.ab=1}^{K.C.ab} w_{k.C.ab}$$

Since

$$\sum_{k.C.ab=1}^{K.C.ab} w_{k.C.ab} \equiv W_{C.ab}$$

and because here the weight is a convenient denotation of the probability, then

$$W_{C.ab} = W_a + W_b \le P(\Omega) = 1$$

and then we have

$$\sum_{k.Couple.ab=1}^{K.Couple.ab} L_{k.Couple.ab}(a, \mu, b, n, w_{k.Couple.ab}) \le$$

$$\le (\mu - a)^n \frac{b-\mu}{b-a} + (b-\mu)^n \frac{\mu - a}{b-a}$$

and

$$| E(X-\mu)^n | \le \sum_{k.C.ab=1}^{K.C.ab} L_{k.C.ab}(a, \mu, b, n, w_{k.C.ab}) \le$$

$$\le (\mu - a)^n \frac{b-\mu}{b-a} + (b-\mu)^n \frac{\mu - a}{b-a} .$$

So, we have considered a discrete random variable X with finite support. X takes on values in a finite interval *[a, b]*. We have proved that the maximal possible modulus of a central moment of this variable is attained for the probability mass function which is concentrated at the borders of the interval. We have also obtain the estimate of this maximal possible modulus of a central moment of X

$$| E(X-\mu)^n | \le (\mu - a)^n \frac{b-\mu}{b-a} + (b-\mu)^n \frac{\mu - a}{b-a} \qquad (1)$$

4. Theorem

4.1. Preliminary considerations

Remark 4.1. Simplification. Let us simplify the inequality (1) for a central moment of X

$$|E(X-\mu)^n| \le (\mu-a)^n \frac{b-\mu}{b-a} + (b-\mu)^n \frac{\mu-a}{b-a} =$$
$$= [(\mu-a)^{n-1} + (b-\mu)^{n-1}] \frac{(\mu-a)(b-\mu)}{b-a}$$

and

$$[(\mu-a)^{n-1} + (b-\mu)^{n-1}] \frac{(\mu-a)(b-\mu)}{b-a} =$$
$$= \left[\left(\frac{\mu-a}{b-a}\right)^{n-1} + \left(\frac{b-\mu}{b-a}\right)^{n-1}\right](b-a)^{n-1} \frac{(\mu-a)(b-\mu)}{b-a}$$

Keeping in mind $a \le \mu \le b$ we have $0 \le (\mu-a)/(b-a) \le 1$ and $0 \le (b-\mu)/(b-a) \le 1$. For $n \ge 2$ we have

$$\left(\frac{\mu-a}{b-a}\right)^{n-1} + \left(\frac{b-\mu}{b-a}\right)^{n-1} \le$$
$$\le \frac{\mu-a}{b-a} + \frac{b-\mu}{b-a} = \frac{b-a}{b-a} \equiv 1 .$$

So,

$$\left[\left(\frac{\mu-a}{b-a}\right)^{n-1} + \left(\frac{b-\mu}{b-a}\right)^{n-1}\right](b-a)^{n-1} \frac{(\mu-a)(b-\mu)}{b-a} \le$$
$$\le (b-a)^{n-1} \frac{(\mu-a)(b-\mu)}{b-a} =$$
$$= (b-a)^{n-2}(\mu-a)(b-\mu) .$$

So,

$$|E(X-\mu)^n| \le (b-a)^{n-2}(\mu-a)(b-\mu)$$

Let us define two terms for the purposes of this article:

Definition 4.2. Bound (restriction) on the expectation.

A "**non-zero bound (restriction) on the expectation** $\boldsymbol{restriction_{Expectation}} \equiv \boldsymbol{r_{Expect}} \equiv \boldsymbol{r}$" signifies the impossibility for the expectation to be located closer to a border of the interval than some non-zero distance.

In other words, a non-zero bound designates the existence of a non-zero distance from a border of the interval. Within this distance, it is impossible for the expectation to be located.

This bound may be denoted also as a "forbidden zone" for the expectation near a border of the interval.

The "bound" for one border and the "bound" for another border constitute the "bounds" for the borders.

The value of a non-zero bound (or the width of a non-zero "forbidden zone") signifies the minimal possible distance between the expectation and a border of the interval. For brevity, the term "the value of a bound" may be shortened to "the bound."

Definition 4.3. A non-zero bound on a central moment.

At the beginning, let us define a "non-zero bound on the dispersion $\sigma^2_{Min.2} \equiv \sigma^2_{Min}$" to be the minimal value of the dispersion $E(X-\mu)^2$ satisfying $E(X-\mu)^2 \geq \sigma^2_{Min.2} > 0$.

Let us define analogously a general "**non-zero bound on the** $\boldsymbol{n}^{\mathbf{th}}$ **order central moment** $|\boldsymbol{\sigma^n_{Min.n}}|$" to be the minimal absolute value of the n^{th} order central moment $E(X-\mu)^n$ satisfying $|E(X-\mu)^n| \geq |\sigma^n_{Min.n}| > 0$.

4.2. Theorem and notes

4.2.1. Theorem

Theorem. Suppose, a discrete random variable X with finite support takes on values in an interval $[a, b] : 0<(b-a)<\infty$. If there is a non-zero lower bound $|\sigma^n_{Min.n}| > 0$ on the modulus of a central moment $|E(X-\mu)^n| \geq |\sigma^n_{Min.n}| : 2\leq n<\infty$, then the non-zero bounds (restrictions) $restriction_{Expectation} \equiv r_{Expect} > 0$ on the expectation exist near the borders of the interval and

$$a < (a + r_{Expect}) \leq \mu \equiv E(X) \leq (b - r_{Expect}) < b \ .$$

Proof. From the conditions of the theorem and from Remark (4.1) we have

$$0 < |\sigma^n_{Min.n}| \leq |E(X-\mu)^n| \leq (b-a)^{n-2}(\mu - a)(b-\mu) \ .$$

This can be rewritten by $r \equiv r_{Expect} \equiv \mu - a$ as

$$0 < |\sigma^n_{Min.n}| \leq (b-a)^{n-2} r(b - a - (\mu - a)) =$$
$$= (b-a)^{n-2} r((b-a) - r)$$

and

$$0 < \frac{|\sigma^n_{Min.n}|}{(b-a)^{n-2}} \leq r(b-a) - r^2 \ .$$

So, we have the inequality

$$r^2 - (b-a)r + \frac{|\sigma^n_{Min.n}|}{(b-a)^{n-2}} < 0 \qquad (2).$$

For the equation

$$r^2 - (b-a)r + \frac{|\sigma^n_{Min.n}|}{(b-a)^{n-2}} = 0 \qquad (3).$$

its roots are

$$r_{1,2} = \frac{b-a}{2} \pm \sqrt{\left(\frac{b-a}{2}\right)^2 - \frac{|\sigma^n_{Min.n}|}{(b-a)^{n-2}}}$$

or

$$r_{1,2} = \frac{b-a}{2} \pm \sqrt{\left(\frac{b-a}{2}\right)^2 - \sigma^2_{Min.n}\left(\frac{|\sigma_{Min.n}|}{b-a}\right)^{n-2}} \qquad (4).$$

or

$$r_{1,2} = \frac{b-a}{2}\left(1 \pm \sqrt{1 - 4\left(\frac{|\sigma_{Min.n}|}{b-a}\right)^n}\right) \ .$$

Let us consider a function

$$\Phi \equiv r^2 - (b-a)r + \frac{|\sigma^n{}_{Min.n}|}{(b-a)^{n-2}} \ .$$

Its derivatives are

$$\frac{\partial \Phi}{\partial r} = 2r - (b-a) \quad \text{and} \quad \frac{\partial^2 \Phi}{\partial r^2} = 2 > 0 \ .$$

The first derivative is equal to zero and the function has its minimum at

$$r_0 = \frac{b-a}{2} \ .$$

The point r_0 is located between the points of the roots of the equation (3). The function is equal to zero at the roots. Therefore, the values of the function are less than zero when r is located between the points of the roots.

Therefore, the expectation can be located only between $(a + r_{Expect})$ and $(b - r_{Expect})$ as

$$a < (a + r_{Expect}) \le E(X) \le (b - r_{Expect}) < b \ ,$$

which proves the theorem.

4.2.2. Symmetry

The expression for the roots of the equation (3) is symmetric with respect to the mean point $(b-a)/2$ of the interval. So, evidently, it may be used both as

$$r_{1,2} = \frac{b-a}{2} \pm \sqrt{\left(\frac{b-a}{2}\right)^2 - \sigma^2{}_{Min.n} \frac{|\sigma^{n-2}{}_{Min.n}|}{(b-a)^{n-2}}}$$

or as the minor root

$$r_2 = \frac{b-a}{2} - \sqrt{\left(\frac{b-a}{2}\right)^2 - \sigma^2{}_{Min.n} \frac{|\sigma^{n-2}{}_{Min.n}|}{(b-a)^{n-2}}} \ ,$$

e.g., in the expression

$$a < (a + r_2) \le E(X) \le (b - r_2) < b \ .$$

4.2.3. Dispersion

For the most important case of $n = 2$ and the dispersion $|\sigma^n{}_{Min.n}| = \sigma^2{}_{Min}$, denoting the half of the length of the interval $[a, b]$ as

$$h \equiv h_{Half} \equiv \frac{b-a}{2} \ ,$$

one can laconically rewrite the inequality (2)

$$r^2 - 2hr + \sigma^2{}_{Min} < 0 \qquad (5)$$

and the roots of the equation $r^2 - 2hr + \sigma^2{}_{Min} = 0$ as

$$r_{1,2} = h \pm \sqrt{h^2 - \sigma^2{}_{Min}} \ ,$$

or, denoting $r \equiv r_2$ as the minor root,

$$r = h - \sqrt{h^2 - \sigma^2{}_{Min}} \qquad (7).$$

The maximal possible dispersion is $\sigma^2 \le ((b-a)/2)^2$. So, denoting the maximal possible standard deviation as

$$\sigma_{Max} = \frac{b-a}{2} \ ,$$

we have

$$r = \sigma_{Max} - \sqrt{\sigma^2{}_{Max} - \sigma^2{}_{Min}} \qquad (8).$$

or in the form of, e.g.,

$$r = \sigma_{Max}\left(1 - \sqrt{1 - \frac{\sigma^2{}_{Min}}{\sigma^2{}_{Max}}}\right) \qquad (9).$$

4.2.4. Infinitesimal case

For the important case of $\sigma_{Min.n} \rightarrow 0$ one can easily obtain

$$a < \left(a + \frac{|\sigma^n{}_{Min.n}|}{(b-a)^{n-1}}\right) \le E(X) \le \left(b - \frac{|\sigma^n{}_{Min.n}|}{(b-a)^{n-1}}\right) < b$$

and for $n = 2$ and $\sigma_{Min.n} = \sigma_{Min.2} = \sigma_{Min}$

$$a < \left(a + \frac{\sigma^2{}_{Min}}{b-a}\right) \le E(X) \le \left(b - \frac{\sigma^2{}_{Min}}{b-a}\right) < b \ .$$

5. Opportunities of the theorem for utility and prospect theories

The dispersion is a common measure of a scattering. The scattering can be caused by noise and/or uncertainty, measurement errors, etc.

So, the theorem can be used in researches of the influence of the scatter of experimental data on their expectations near the borders of intervals.

There is a way of researches in utility and prospect theories.

Noise and uncertainty are widespread phenomena in economics, in particular in decision, utility and prospect theories. Their analysis is one of ways of researches (see, e.g., Schoemaker and Hershey, 1992, Butler and Loomes, 2007).

There is another way of researches.

It consists in the analysis of Prelec's probability weighting function at the probabilities $p \sim 1$ (see Steingrimsson and Luce, 2007, Aczél and Luce, 2007and Harin 2014).

The theorem synthesizes these two ways.

Sketches of versions of the above existence theorem have at least partially explained the problems, including underweighting of high and the overweighting of low probabilities, risk aversion, the "four-fold pattern" paradox, etc. (see, e.g., Harin 2012). So, the theorem can be used also in decision, utility and prospect theories, especially in researches of Prelec's weighting function.

6. Conclusions

Suppose a discrete random variable $X=\{x_k\} : k=1, 2, \dots K : 2 \le K < \infty$, takes on values in a finite interval $[a, b]$ and there is a non-zero lower bound on the modulus of its central moment $|E(X-E(X))^n|$ (this bound is denoted as $|\sigma^n{}_{Min.n}|$, so, $|E(X-E(X))^n| \ge |\sigma^n{}_{Min.n}| > 0$). Under these conditions, the existence theorem is proved for non-zero bounds (restrictions) $restriction_{Expectation} \equiv r_{Expect} \equiv r > 0$ on its expectation $E(X)$ near the borders of the interval.

The main bounding inequality of the present article is

$$a < (a + r_{Expect}) \le E(X) \le (b - r_{Expect}) < b \ ,$$

In other words, under the above conditions, the non-zero "forbidden zones" (those widths are equal to r_{Expect}) are proved to exist near the borders a and b of the interval $[a, b]$.

In this inequality the bounds on the expectation are

$$r_{Expect} = \frac{b-a}{2} - \sqrt{\left(\frac{b-a}{2}\right)^2 - \sigma^2{}_{Min.n}\left(\frac{|\sigma_{Min.n}|}{b-a}\right)^{n-2}} \ .$$

For the most important case of $n=2$ (for the minimum $\sigma^2{}_{Min} > 0$ of the dispersion σ^2), the bounds $r \equiv r_{Expect}$ on the expectation can be written laconically, denoting the half of the interval as $h_{Half} \equiv h \equiv (b-a)/2$,

$$r = h - \sqrt{h^2 - \sigma^2{}_{Min}} \ ,$$

or, denoting the maximal possible standard deviation as $\sigma_{Max} = (b-a)/2$,

$$r = \sigma_{Max} - \sqrt{\sigma^2{}_{Max} - \sigma^2{}_{Min}} \ ,$$

or, e.g., in the form of

$$r_{Expect} = \sigma_{Max}\left(1 - \sqrt{1 - \frac{\sigma^2{}_{Min}}{\sigma^2{}_{Max}}}\right) \ .$$

The main bounding inequality can be rewritten for $\sigma_{Min} \to 0$ as

$$a < \left(a + \frac{\sigma^2{}_{Min}}{b-a}\right) \le E(X) \le \left(b - \frac{\sigma^2{}_{Min}}{b-a}\right) < b \ .$$

The theorem for the dispersion can be used in researches of the influence of the scatter of experimental data on their expectations near the borders of finite intervals; utility and prospect theories, especially in researches of Prelec's weighting function.

References

Aczél, J., and R. D. Luce, "A behavioral condition for Prelec's weighting function on the positive line without assuming $W(1)=1$," *Journal of Mathematical Psychology* **51** (2007), pp. 126–129.

Dokov, S. P., D.P. Morton. 2005. Second-Order Lower Bounds on the Expectation of a Convex Function. *Mathematics of Operations Research* **30**(3) 662–677

Harin, A., "The random–lottery incentive system. Can $p\sim 1$ experiments deductions be correct?" *16th conference on the Foundations of Utility and Risk*, 2014.

Harin, A., (2012), "Data dispersion in economics (II) – Inevitability and Consequences of Restrictions", *Review of Economics & Finance* **2** (2012), no. 4: 24–36.

Kahneman, D., and R. Thaler, "Anomalies: Utility Maximization and Experienced Utility," *Journal of Economic Perspectives* **20** (2006), no. 1, 221–234.

Pinelis, I. 2011. Exact lower bounds on the exponential moments of truncated random variables. *Journal of Applied Probability* **48**(2) 547-560.

Prékopa, A. 1990. The discrete moment problem and linear programming. *Discrete Applied Mathematics* **27**(3) 235–254.

Prékopa, A. 1992. Inequalities on Expectations Based on the Knowledge of Multivariate Moments. *Lecture Notes-Monograph Series* **22** *Stochastic Inequalities* 309-331.

Schoemaker, P., and J. Hershey, "Utility measurement: Signal, noise, and bias," *Organizational Behavior and Human Decision Processes* **52** (1992) no. 3, 397–424.

Steingrimsson, R., and R. D. Luce, "Empirical evaluation of a model of global psychophysical judgments: IV. Forms for the weighting function," *Journal of Mathematical Psychology* **51** (2007), pp. 29–44.