Mean Values: A Multicriterial Analysis

Vladislav V. Podinovski and Andrey P. Nelyubin

1 Introduction

Mean values are widely used in management, economics, sociology, engineering and other areas of theory and practice. In statistics (see, for example, [8, 22]), mean values are aggregate representations of the varying characteristics of a group of homogeneous objects. Mean values cancel out random variations of a particular characteristic and tend to represent the effect caused by the main factors affecting it. Mean values allow us to compare the levels of the same characteristic in different groups of objects and to investigate the causes of such differences.

It is known that it is impossible to define a universally applicable notion of the mean value which satisfies all desirable properties [1, 8]. Instead, different notions of the mean value are required for different problems and situations. However, in some applications, it may be unclear which of the known mean values should be used, and different means may point to different conclusions. Policy recommendations in such situations may become problematic [6, 8, 10, 12, 16].

Grabisch et al. [7] regarded mean values as idempotent aggregation functions and concluded that the class of such functions "is huge, making the problem of choosing the right function (or family) for a given application a difficult one".

In this paper, we consider new approaches to the definition of the mean value based on the ideas and methods of multicriteria optimization. Such means turn out to be multi-valued, i.e., represented by sets of points. These allow two interpretations,

V. V. Podinovski
National Research University Higher School of Economics, Moscow, Russian Federation
e-mail: podinovski@mail.ru

A. P. Nelyubin (✉)
Mechanical Engineering Research Institute RAS, Moscow, Russian Federation
e-mail: nelubin@gmail.com

B. Goldengorin, S. Kuznetsov (eds.), *Data Analysis and Optimization*, Springer Optimization and Its Applications 202,
https://doi.org/10.1007/978-3-031-31654-8_17

either as the range of possible mean values in some specific situations (characterized by scale properties, such as equal importance or ordinality, and/or transfer principles), or as whole sets for the given sample.

2 Definition of Mean Values as Nondominated Points

Let X be the set of real numbers consisting of at least $n \geq 2$ elements referred to as data or points. These elements are typically obtained as a result of measurement of some characteristic:

$$X = \{x_1, x_2, \ldots, x_n\}\,. \tag{1}$$

These data are assumed homogeneous in the sense that they are obtained by utilizing the same scale of measurement [21, 23]. We assume that the data (1) are quantitative, i.e., the measurement is performed either on the interval scale or on the ratio scale [15].

The elements of the set (1) can be ranked in the non-decreasing and non-increasing order.

$$X_{\uparrow} =< x_{(1)}, x_{(2)}, \ldots, x_{(n)} >;\ X_{\downarrow} =< x_{[1]}, x_{[2]}, \ldots, x_{[n]} >, \tag{2}$$

where $x_{(1)} \leq x_{(2)} \leq \ \ldots \ \leq x_{(n)}$ and $x_{[1]} \geq x_{[2]} \geq \ \ldots \ \geq x_{[n]}$. In statistics, the set (1) is typically referred to as a sample and its non-decreasing sequence $X_{\uparrow}$ as a variational series.

Let x be an arbitrary fixed number (a point in Re). Its distance from any point x_i from X is given by $y_i = |x - x_i|$. Then the distance from x to the dataset X can be characterized by the vector $y = (y_1, y_2, \ \ldots, y_n)$. We can view this vector as the value of the vector criterion $f(x) = (f_1(x), f_2(x), \ \ldots, f_n(x))$, where $f_i(x) = |x - x_i|$, which is an element of the nonnegative quadrant $\mathrm{Re}^n_+ = [0, +\infty)^n$.

Let P^{Γ} be a preference relation (strict partial order) on Re^n_+, where Γ is information about the preferences with respect to distance: if $yP^{\Gamma}y'$, where $y = f(x)$ and $y' = f(x')$, then the point x is closer to the dataset X than x'.

The relation P^{Γ} generates the corresponding relation P_{Γ} on the numeric axis Re: $xP_{\Gamma}x' \iff f(x)P^{\Gamma}f(x')$.

Therefore, any candidate that we may choose as the closest to X and representing the set X must be nondominated under P_{Γ}. If the set $G^{\Gamma}(X)$ of nondominated under P_{Γ} points is externally stable, we refer to all such points as pn-means (principal new means) and, more specifically (reflecting the information Γ), as the means with respect to P_{Γ}.

If there is no further information about the preferences of the DM on Re^n_+, we obtain the Pareto relation $P^{\varnothing}$ defined as follows:

$$yP^{\varnothing}z \iff y_i \leq z_i, i = 1, 2, \ldots, n;\ y \neq z.$$

Relation $P^{\varnothing}$ generates the Pareto relation $P_{\varnothing}$ on Re: $xP_{\varnothing}x' \iff f(x)P^{\varnothing}f(x')$.

Theorem 1 The set of all means of the dataset (1) with respect to $P_{\varnothing}$ is the segment $G^{\varnothing}(X) = \overline{X} = [x_{(1)}, x_{(n)}]$, where $x_{(1)} = \min_{i \in N} x_i$ and $x_{(n)} = \max_{i \in N} x_i$, $N = \{1, 2, \ldots, n\}$. This set is externally stable.

Therefore, the notion of the means with respect to $P_{\varnothing}$ is equivalent to the means in the sense of Cauchy.

Proofs of this and the following theorems can be found in [19, 20].

Let us note that, if the function φ is increasing on Re_+, then changing the original criteria $f_i(x) = |x - x_i|$ by $\varphi(f_i(x))$ does not change the set $G^{\varnothing}(X)$. For example, one can use "smooth" criteria $f_i(x) = (x - x_i)^2$. Therefore, the original use of formula $f_i(x) = |x - x_i|$ as a measure of distance is not essential and is not a limiting assumption for the suggested approach.

3 Mean Values for Equally Important Criteria

In this section, we assume that all criteria are equally important [17] and denote this information E. In this case, the distance from the point x to the dataset X is represented by the preference relation P_E on Re, which is defined by the two equivalent decision rules [17], where $f_i(x) = |x - x_i|$:

$$xP_Ex' \iff \left(f_{(1)}(x) \le f_{(1)}\left(x'\right), f_{(2)}(x) \le f_{(2)}\left(x'\right), \ldots, f_{(n)}(x) \le f_{(n)}\left(x'\right)\right),$$

and at least one of these inequalities is strict;

$$xP_Ex' \iff \left(f_{[1]}(x) \le f_{[1]}\left(x'\right), f_{[2]}(x) \le f_{[2]}\left(x'\right), \ldots, f_{[n]}(x) \le f_{[n]}\left(x'\right)\right),$$

and at least one of these inequalities is strict.

In this case, the pn-means (with respect to P_E) (elements of the set $G^E(X)$) are the points on the numerical axis which are nondominated under P_E.

Theorem 2 We have $G^E(X) \subseteq G^{\varnothing}(X) = \overline{X}$, and the set $G^E(X)$ is externally stable.

Note that, if the function φ is increasing on Re_+, then changing the original criteria $f_i(x)$ to criteria $\varphi(f_i(x))$ does not change the relation P_E and the set $G^E(X)$.

Let us consider examples of sets $G^E(X)$ constructed according to the methods described in Sect. 5.

Example 1 Let $n = 3$ and $X = \{1, 2, 5\}$. In this example, $G^E(X) = [1.5, 3]$.

In the above example, the set $G^E(X)$ is a single line segment. However, for large n, this set may be the union of several segments, excluding their endpoints.

Example 2 For $n = 6$ and different sets X, we have:

$$G^E(\{10, 11, 15, 61, 107, 110\}) = [10.5, 83) \cup (83.5, 85) \cup (106.5, 108);$$

$$G^E(\{10, 11, 40, 55, 70, 110\}) = [10.5, 18) \cup (18; 67.5) \cup (68, 75);$$

$$G^E(\{10, 57, 61, 64, 109, 110\}) = (56.5, 57.5) \cup (58.5, 88.5) \cup (108, 109.5].$$

Examples 1 and 2 also illustrate the following result.

Theorem 3 Let the distance between two adjacent elements $x_{(i)}$ and $x_{(i+1)}$ of the variational series (2) be the smallest among all other pairs of adjacent elements of this series, and let these two elements be uniquely defined. Then the midpoint $x^c = ½(x_{(i)} + x_{(i+1)})$ is an element of $G^E(X)$. Moreover, if $x_{(i)}$ is $x_{(1)}$ or if $x_{(i+1)}$ is $x_{(n)}$, then x^c is the left or, respectively, right, endpoint of the set $G^E(X)$.

If $x_{(1)} \neq x_{(n)}$ and $(x_{(1)}, x_{(n)}) \not\subset G^E(X)$, for some values of parameter s, the power mean.

$$g^s(X) = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i)^s\right)^{1\backslash s}, s \neq 0.$$

is not the mean with respect to P_E. This is because, as s increases on Re, the function $g^s(X)$, extended to preserve continuity, passes through all values from the interval $(x_{(1)}, x_{(n)})$ [8]. However, we have the following result:

Theorem 4 The arithmetic mean is a mean with respect to P_E, i.e., $g^1(X) \in G^E(X)$.

Example 3 According to Example 2, for $X = \{10, 57, 61, 64, 109, 110\}$ we have: $G^E(X) = (56.5, 57.5) \cup (58.5, 88.5) \cup (108, 109.5]$. In this example, the geometric mean $g^0(X) = 54.66 \notin G^E(X)$ and harmonic mean $g^{-1}(X) = 35.75 \notin G^E(X)$, and $g^1(X) = 68.5 \in G^E(X)$. In Example 1, for $X = \{1, 2, 5\}$, we have $G^E(X) = [1.5, 3]$. Here, the quadratic mean $g^2(X) = 3.162 \notin G^E(X)$, but $g^1(X) = 2.67 \in G^E(X)$.

Theorem 5 The median is a mean with respect to P_E, i.e., if n is an odd integer and the median is unique, we have $\mu(X) = x_{\left(\frac{n+1}{2}\right)} \in G^E(X)$. If n is an even number, the median is not unique and we have $\mu(X) = \left[x_{\left(\frac{n}{2}\right)}, x_{\left(\frac{n}{2}+1\right)}\right] \subseteq G^E(X)$.

Examples 1 and 2 provide illustrations to the above theorem.

It should be noted the following peculiarity of the means with respect to P_E: if the points $x_i \in X$ and $x_j \in X$, $x_i < x_j$, are included in $G^E(X)$, then the point $x_k \in X$, such that $x_i < x_k < x_j$, may not belong to $G^E(X)$!

Example 4 For $n = 7$ and different sets X, we have:

$$X' = \{1, 2, 3, 6, 8, 9, 11\}, G^E(X') = [2; 3) \cup (3; 8.5];$$

$$X'' = \{1, 2, 3, 7, 8, 10, 11\}, G^E\left(X''\right) = [2; 3) \cup (3.5; 9].$$

Here $x_{(2)}$, $x_{(4)} \in G^E(X)$, whereas $x_{(3)} \notin G^E(X)$ for both sets $X = X'$ and $X = X''$. Moreover, the point $x_{(3)} = 3$ is a punctured point of the set $G^E(X')$ (it is dominated under P_E by the point $x_{(5)} = 8$, and in its arbitrarily small neighborhood there are points nondominated under P_E).

This feature clearly violates the very principle of constructing means as points closest to points from X, and is not consistent with the intuitive concept of a mean value. Therefore, the presence of this feature can be considered as a paradox of means with respect to P_E.

4 Mean Values for Equally Important Criteria Measured on the First Ordered Metric Scale

Let y be any vector estimate such that $y_i > y_j$. Consider any $\delta > 0$ such that $y_i - \delta \geq y_j + \delta$. Define the vector estimate z by replacing component y_i by $y_i - \delta$ and y_j by $y_j + \delta$, but $y_i - \delta \geq y_j + \delta$. Moving from y to z reduces the larger deviation y_i from one point in the sample and increases a smaller deviation y_j from a different point, by the same amounts δ. The resulting set of distances becomes closer to the ideal set of minimally possible equal deviations. Assume that, for any y and δ described above, the vector estimate z is preferred to the original vector estimate y, in the sense that z is "closer" to X than y and is therefore more suitable for the definition of the mean. Denote Δ the information about the described principle. Such approach is an analogue of Pigou-Dalton's principle of transfer for income distribution [2, 5]. This means that the equally important criteria have a common first ordered metric scale [4]. The preference relation $P_{E\Delta}$, generated on Re^n by the joint information E and Δ, is defined by the following decision rule [14, 18]:

$$x P_{E\Delta} x' \iff f_{[1]}(x) \leq f_{[1]}\left(x'\right), f_{[1]}(x) + f_{[2]}(x) \leq f_{[1]}\left(x'\right) + f_{[2]}\left(x'\right), \ldots$$
$$\ldots f_{[1]}(x) + f_{[2]}(x) + \cdots + f_{[n]}(x) \leq f_{[1]}\left(x'\right) + f_{[2]}\left(x'\right) + \ldots f_{[n]}\left(x'\right),$$

and at least one of these inequalities is strict. In this case, the pn-means are the points that are nondominated under $P_{E\Delta}$. Because $P_{E\Delta} \supset P_E$, we have $G^E(X) \supseteq G^{E\Delta}(X)$.

Theorem 6 The arithmetic mean is a mean with respect to $P_{E\Delta}$, i.e., $g^1(X) \in G^{E\Delta}(X)$.

Theorem 7 If n is odd, the median (which is uniquely defined), is a mean with respect to $P_{E\Delta}$, i.e., $\mu(X) \in G^{E\Delta}(X)$. If n is even and the median is not uniquely defined, we only have $\mu(X) \cap G^{E\Delta}(X) \neq \varnothing$.

Example 5 If $n = 5$ and $X = \{1, 2, 3, 5, 11\}$, we have $G^{E\Delta}(X) = [3, 6]$, $\mu(X) = 3$ and $g^1(X) = 4.4$. If $n = 4$ and $X = \{10, 11, 12, 110\}$, we have $G^{E\Delta}(X) = [11.5,$

60], $\mu(X)$ = [11, 12] and $g^1(X) = 35.75$. If $X = \{10, 11, 20, 110\}$, we have $G^{E\Delta}(X) = [15.5, 60]$, $\mu(X) = [11, 20]$ and $g^1(X) = 37.75$.

Let us define the set $H = \{1, 2, \ldots, h\}$, where $h = \lfloor(n+1)/2\rfloor$ is the integer part of $(n+1)/2$.

Theorem 8 The set $G^{E\Delta}(X)$ is externally stable and coincides with the segment $[\alpha, \beta]$, where

$$\alpha = \frac{1}{2}\min_{p\in H}\left(x_{(p)} + x_{(n+1-p)}\right), \beta = \frac{1}{2}\max_{p\in H}\left(x_{(p)} + x_{(n+1-p)}\right) \quad (3)$$

Example 6 For $n = 5$, we have $h = \lfloor(n+1)/2\rfloor = 3$ and $H = \{1, 2, 3\}$. For $X = \{1, 2, 7, 8, 11\}$, using Theorem 8, we have:

$$\alpha = \frac{1}{2}\min\left\{x_{(1)} + x_{(5)}, x_{(2)} + x_{(4)}, x_{(3)} + x_{(3)}\right\} = \frac{1}{2}\min\{1+11, 2+8, 7+7\}$$
$$= \frac{1}{2}\min\{12, 10, 14\} = 5;$$

$$\beta = \frac{1}{2}\max\left\{x_{(1)} + x_{(5)}, x_{(2)} + x_{(4)}, x_{(3)} + x_{(3)}\right\} = \frac{1}{2}\max\{12, 10, 14\} = 7;$$

Therefore, $G^{E\Delta}(X) = [\alpha, \beta] = [5, 7]$.

5 On the Construction of Sets of Mean Values

For the construction of the set $G^E(X)$, we can use known methods of multicriteria optimization developed for the construction of the sets of nondominated variants [17]. Such methods utilize families of functions that are increasing (decreasing), or at least non-decreasing (non-increasing) with respect to P_E. For example, we can solve a parametric program which minimizes the function of single variable $\psi(f(x)|c) = \min_{\pi\in\Pi} \max_{i\in N} \{f_{\pi(i)}(x) - c_i\}$ on the set X, by varying the vector parameter $c \in f\left(\overline{X}\right)$. However, even if n is not very large, the number $n!$ of terms of this function (with respect to which the maximization is performed) turns out unacceptably large.

Taking into account that the set X is one-dimensional, we can utilize a different approach. Namely, we can consider a dense grid with the small step h which covers the set X, and identify the nondominated (with respect to P_E) points of this grid by simple enumeration [9]. The step h depends on the required precision and can decrease in the process of calculations of the set $G^E(X)$. We used this approach for the construction of the set $G^E(X)$ in Examples 1 и 2.

Example 7 Let us demonstrate the construction of the set $G^E(X)$ for $X = \{1, 2, 5, 9, 11\}$. Using computer for the calculations, while reducing the step length h, we obtain the following results:

$h = 1$: [2, 7] ∪ [9, 9].
$h = 0.1$: [1.5, 7.4] ∪ [8.6, 9.4].
$h = 0.01$: [1.50, 7.49] ∪ [8.51, 9.49].
$h = 0.001$: [1.500, 7.499] ∪ [8.501, 9.499].
$h = 0.0001$: [1.5000, 7.4999] ∪ [8.5001, 9.4999]

Using the enumeration approach with $h = 0.01$, we found out that the point 4.5 dominates the points 7.5 and 8.5. Similarly, the point 2.5 dominates the point 9.5. Therefore, by Theorem 2, we have $G^E(X) = [1.5, 7.5) \cup (8.5, 9.5)$.

Let us highlight another result that may be useful in the construction of the set $G^E(X)$.

Theorem 9 Let vector estimates of all $x \in X$ be located at the points of some uniform grid covering X. Then, in order to test if any grid point is a mean with respect to P_E, it suffices to compare its vector estimate only with the vector estimates of all the other points of the grid.

Let us note that the uniform grid required by the conditions of Theorem 9 can always be constructed if all points in X are rational numbers. In practical applications, these would typically be integer numbers or decimal fractions.

It is worth noting that it is easier to construct the set of means $G^{E\Delta}(X)$ than the set $G^E(X)$. According to Theorem 8, the set $G^{E\Delta}(X)$ is easily found by calculating the endpoints α and β of the segment $[\alpha, \beta]$ using formulae (3) – see Example 6.

6 On Comparing Multi-valued Means

In practice, it is important that we can compare the mean values measured on the same scale. For the means that are uniquely defined, this is a simple task of comparing the two numerical values. In the case of multi-valued means, in statistics, it is common to substitute such means by a single number, e.g., in the case of a median when n is an even number.

The set $G^\Gamma(X)$ consists of l intervals with the endpoints $x^1, x^2; x^3, x^4; \ldots; x^{2l-1}, x^{2l}$, and these intervals do not intersect with each other. Define the length $D^\Gamma(X)$ of the set $G^\Gamma(X)$ as the sum of the lengths of all these intervals: $D^\Gamma(X) = \sum_{k=1}^{l} \left|x^{2k} - x^{2k-1}\right|$. Furthermore, define $D_x^\Gamma(X)$ the length of the part of the set $G^\Gamma(X)$ that is located to the right of the point x. It includes the (part) of one interval and all the other intervals located to the right of x. The relative length $d_x^{\Gamma}(X)$ is defined as the ratio $d_x^{\Gamma}(X) = D_x^{\Gamma}(X)$: $D^{\Gamma}(X)$.

Because none of the points of the set $G^{\Gamma}(X)$ has any advantages (in the sense of representing the sample) compared to its other points, any of them may be regarded as an equally valid candidate for the choice of the mean. This is analogous to the principle of insufficient reason for decision making under ignorance [13]. Using first-order stochastic dominance [11], we say that the mean $G^{\Gamma}(X')$ is not less than the mean $G^{\Gamma}(X'')$ and state this as $G^{\Gamma}(X')) \succsim G^{\Gamma}(X'')$, if $d_x^{\Gamma}\left(X'\right) \geq d_x^{\Gamma}\left(X''\right)$ for each $x \in$Re. If the latter inequality is strict for at least one $x \in$Re, the former mean is greater than the latter. This relationship between the means ("is not less than") is a partial quasi-order. The corresponding relation "is greater than" is denoted $\succ$ and is a partial strict order (it is irreflexive and transitive). This strict relation is essentially a probabilistic dominance relation, or a strict first-order stochastic dominance relation [11]. Note that we have $d_x^{\Gamma}(X) = 1-F(x)$, where $F(x)$ is the cumulative distribution function corresponding to the uniform distribution with the density equal to 1 / $D^{\Gamma}(X)$ on $G^{\Gamma}(X)$ and equal to zero outside $G^{\Gamma}(X)$.

It is clear that the relation $\succcurlyeq$ is weak in the sense that it would typically not result in a definitive comparison of the means. Relation $\succcurlyeq$ can be extended using the ideas of second-order stochastic dominance, but this approach does not appear to be sufficiently effective in practice either.

Another approach would be to "compress" the means that are not uniquely defined to single-valued means. However, this would lead to a loss of information, and the results of comparison would be approximate. For example, let the mean $G^{\Gamma}(X)$ consist of several not intersecting intervals defined by the endpoints x^1, x^2; $x^3, x^4;\ \ldots; x^{2l\text{-}1}, x^{2l}$. We can represent this mean by its the centre of mass $x^{\Gamma}(X)$ and refer to it as the centroid mean.

Example 8 Let $G^E(X') = [1, 2) \cup (5, 8)$ and $G^E(X'') = [1.5, 4.5] \cup (8, 9]$. We have:

$$x^E\left(X'\right) = (1.5 \cdot 1 + 6.5 \cdot 3)/4 = 5.25;\ x^E\left(X''\right) = (3 \cdot 3 + 8.5 \cdot 1)/4 = 4.375.$$

Because 5.25 > 4.375, we can accept that the mean $G^E(X')$ is greater than $G^E(X'')$.

It is useful to note that, if $G^{\Gamma}(X') \succ G^{\Gamma}(X'')$, then $x^{\Gamma}(X') > x^{\Gamma}(X'')$ [11].

It is worth noting that it easier to compare the means $G^{E\Delta}(X')$ and $G^{E\Delta}(X'')$ than the means $G^E(X')$ and $G^E(X'')$, because the former are the segments $[\alpha', \beta']$ and $[\alpha'', \beta'']$ respectively. Because the graph of the function $d_x^{E\Delta}(X)$ is a broken line consisting of the single segment $[\alpha, \beta]$ on which it decreases from 1 to 0, $G^{E\Delta}(X')$ $G^{E\Delta}(X'')$ is true if and only if $\alpha' \geq \alpha''$ and $\beta' \geq \beta''$.

For the simplified application of the mean $G^{E\Delta}$, we can represent the segment $[\alpha, \beta]$ by its midpoint $\gamma = ½\,(\alpha + \beta)$, which can be referred to as the centroid mean (with respect to $P_{E\Delta}$).

Example 9 The means of the real GDP per capita in Europe calculated based on the data from Eurostat [3] are shown in Table 1 and Fig. 1.

Table 1 Mean real GDP per capita in Europe (in Euro)

Year	2012	2013	2014	2015	2016	2017	2018	2019
m	25,741	25,825	26,257	27,063	27,564	28,274	28,909	29,249
μ	21,780	20,400	20,250	21,020	22,270	23,200	24,120	24,570
α	19,720	19,745	20,250	21,020	21,995	22,840	23,245	23,485
β	40,840	41,310	42,015	42,970	43,850	43,750	44,335	44,545
γ	30,280	30,528	31,133	31,995	32,923	33,295	33,790	34,015

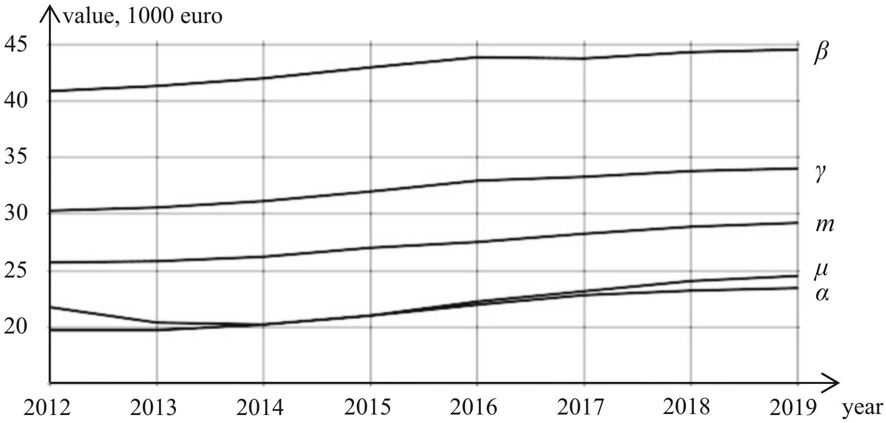

Fig. 1 Means of the GDP per capita in Europe

Table 1 shows that the GDP per capita m is increasing in the period from 2012 to 2019, but the median GDP per capita μ is decreasing until 2014 and is increasing afterwards. Therefore, it is impossible to make a definite conclusion about the GDP growth in the given period. However, the mean $G^{E\Delta}$ (defined by its boundaries α and β) is increasing (there is only an insignificant decrease of β in 2017), and the condensed mean γ is increasing over the whole period. This observation supports the conclusion that the GDP per capita in Europe has been increasing in the given period.

7 On Stability of Pn-Means

The question of stability of the means with respect to small perturbations of the data (1) is important from both theoretical and practical points of view.

Because $G^{\varnothing}(X) = \overline{X} = [x_{(1)}; x_{(n)}]$, a small change of the values x_i may lead only to small changes of $x_{(1)}$ and $x_{(n)}$. Therefore, the set of the means $G^{\varnothing}(X)$ is stable.

However, the mean with respect to P_E may not be stable in the sense that a very small perturbation of a single point in X may lead to a noticeable change of the set $G^E(X)$. The following examples illustrate this possibility.

Example 10 For $X = \{1, 2, 3\}$, we have $G^E(X) = [1.5, 2.5]$. However, for $X^\varepsilon = \{1, 2 - \varepsilon, 3\}$, where $\varepsilon > 0$ is very small, we have $G^E(X^\varepsilon) = [1.5 – 0.5\varepsilon, 2]$. The right endpoint of the set of the means with respect to P_E has changed by 0.5.

Example 11 For $X = \{10, 25, 40, 110\}$, we have $G^E(X) = [25, 60]$. However, for $X^\varepsilon = \{10, 25, 40 + \varepsilon, 110\}$, where $\varepsilon > 0$ is very small, we have $G^E(X^\varepsilon) = [17.5, 60]$. It is interesting that, although only one point in X has increased by a very small ε, the left endpoint of the set of the means (with respect to P_E) has decreased by 7.5.

Let us now consider the issue of stability of the mean with respect to $P_{E\Delta}$.

Example 12 In the setting of Example 10, we have $G^{E\Delta}(X) = \{2\}$ and $G^{E\Delta}(X^\varepsilon) = [2 - \varepsilon, 2]$. Here, a change of one of the data points in X by ε leads to the change of one of the endpoints of the set of the means with respect to $P_{E\Delta}$ by the same ε.

Example 13 Under the conditions of Example 11, we have $G^{E\Delta}(X) = [32.5, 60]$ и $G^{E\Delta}(X^\varepsilon) = [32.5 + 0.5\varepsilon, 60]$. In this example, a change of one of the data points in X by ε results in the change of one of the endpoints for the set of the means by 0.5ε.

Consider the general case. Suppose that the dataset X stated by (1) has changed to the set $X^\varepsilon = \{x_1 + \varepsilon_1, x_2 + \varepsilon_2, \ldots, x_n + \varepsilon_n\}$, where $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are arbitrary numbers.

Theorem 10 The mean with respect to $P_{E\Delta}$ is stable in the following sense: If X is changed to X^ε, the endpoints of the set of the means $G^{E\Delta}(X) = [\alpha, \beta]$ do not change by more than the following value:

$\max\{|\varepsilon_1|, |\varepsilon_2|, \ldots, |\varepsilon_n|\}$.

Therefore, the means with respect to $P_\varnothing$ and $P_{E\Delta}$ are stable with respect to small perturbations of the dataset (1), while the means with respect to P_E may be noticeably unstable.

8 The Case of Data with Repetitions

Assume that the dataset allows repetitions, i.e., the point x_1 occurs β_1 times, x_2 occurs β_2 times, ..., x_n occurs β_n times. In this case, the dataset (1) is replaced by Table 2.

Table 2 Data with repetitions

Value x_i	x_1	x_2	...	x_n
Weight β_i	β_1	β_2	...	β_n

In statistics, the numbers β_i are referred to as weights or (absolute) frequencies, and they are used for the calculation of weighted means.

All our results obtained above, starting with the definition of pn-means, are extended to the described more general case. For this, we consider the dataset consisting of the repeating values x_1 (β_1 times), x_2 (β_2 times) and so on, i.e., we restate the data in Table 2 as follows:

$$\left(\underbrace{x_1, \ldots, x_1}_{\beta_1}, \underbrace{x_2, \ldots, x_2}_{\beta_2}, \ldots, \underbrace{x_n, \ldots, x_n}_{\beta_n}\right).$$

The use of the described methods of construction of pn-means in this case may be computationally demanding as the dimension of the problem becomes very large for large values β_i. To overcome this problem, we may use decision rules developed in theory of qualitative criteria importance measured on continuous scale [18]. In this approach, we treat the integer numbers β_i as quantitative coefficients reflecting the importance of criteria and use notation P^{β} and $P^{\beta\Delta}$ to denote the corresponding relations instead of P^E and $P^{E\Delta}$.

To state the relevant decision rules for the vector estimates y and z, define the following set and values:

$$\begin{aligned} W(y, z) &= \{y_1\} \cup \{y_2\} \cup \cdots \cup \{y_m\} \cup \{z_1\} \cup \{z_2\} \cup \cdots \cup \{z_m\} \\ &= \{w_1, w_2, \ldots, w_q\}, w_1 > w_2 > \cdots > w_q; \end{aligned}$$

$$\begin{aligned} &= \sum\nolimits_{i: y_i \geq w_k} \beta_i, \quad b_k(z) = \sum\nolimits_{i: z_i \geq w_k} \beta_i \; b_k(y) = \sum\nolimits_{i: y_i \geq w_k} \beta_i \\ &= \sum\nolimits_{i: z_i \geq w_k} \beta_i, k = 1, 2, \ldots, q-1; \end{aligned}$$

$$d_k(y) = \sum_{j=1}^{k} b_j(y) \left(w_j - w_{j+1}\right), k = 1, 2, \ldots, q-1.$$

Decision rule for P^{β}:

$$y P^{\beta} z \iff b_k(y) \leq b_k(z), k = 1, 2, \ldots, q-1, \tag{4}$$

and at least one of these inequalities is strict.

Decision rule for $P^{\beta\Delta}$:

$$y P^{\beta\Delta} z \iff d_k(y) \leq d_k(z), k = 1, 2, \ldots, q-1, \tag{5}$$

Table 3 Data in Example 14

Value x_i	1	2	4	5	7	9	11
Weight β_i	2	1	4	1	2	3	1

Table 4 Data for Example 15

Number i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Value $x_{(i)}$	1	1	2	4	4	4	4	5	7	7	9	9	9	11

and at least one of these inequalities is strict.

Example 14 Consider the dataset in Table 3.

Using decision rules (4) and (5), let us compare points 5 and 3 which have the following vector estimates: $y = f(5) = (4, 3, 1, 0, 2, 4, 6)$ and $z = f(3) = (2, 1, 1, 2, 4, 6, 8)$. In this case, $W = (8, 6, 4, 3, 2, 1, 0)$. Therefore, $q = 7$. We have:

$b(y) = (b_1(y), b_2(y), \ldots, b_6(y)) = (0, 1, 6, 7, 9, 13);$
$b(z) = (b_1(z), b_2(z), \ldots, b_6(z)) = (1, 4, 6, 6, 9, 14);$
$d(y) = (d_1(y), d_2(y), \ldots, d_6(y)) = (0, 2, 8, 15, 24, 37);$
$d(z) = (d_1(z), d_2(z), \ldots, d_6(z)) = (2, 10, 16, 22, 31, 45).$

Note that $b_1(y) = 0 < b_1(z) = 1$ but $b_4(y) = 7 > b_1(z) = 6$. According to (4), neither $yP^{\beta}z$ nor $zP^{\beta}y$ is true. However, because all 6 inequalities (5) are true and at least one of them is strict, we have $yP^{\beta\Delta}z$.

Note that formula (3) is easier to use if we first rearrange data with repetitions in the form (1).

Example 15 Consider the data from Table 3 of Example 14. We can rearrange these data as in Table 4 in which we specify the ordinal number i for each point and the corresponding value $x_{(i)}$.

Using formulae (3), we consecutively calculate:

$$\alpha = \frac{1}{2}\min\{x_{(1)} + x_{(14)}, x_{(2)} + x_{(13)}, x_{(3)} + x_{(12)}, x_{(4)} + x_{(11)}, x_{(5)} + x_{(10)}, x_{(6)} + x_{(9)}, x_{(7)} + x_{(8)}\} =$$
$$= \frac{1}{2}\min\{1 + 11, 1 + 9, 2 + 9, 4 + 9, 4 + 7, 4 + 7, 4 + 5\}$$
$$= \frac{1}{2}\min\{12, 10, 11, 13, 11, 9\} = 4.5;$$

$$\beta = \frac{1}{2}\max\{x_{(1)} + x_{(14)}, x_{(2)} + x_{(13)}, x_{(3)} + x_{(12)}, x_{(4)} + x_{(11)}, x_{(5)} + x_{(10)}, x_{(6)} + x_{(9)}, x_{(7)} + x_{(8)}\} =$$
$$= \frac{1}{2}\max\{12, 10, 11, 13, 11, 11, 9\} = 6.5.$$

Therefore, $G^{\beta\Delta}(X) = [4.5, 6.5]$.

9 Conclusion

In this paper, we introduced new notions of the means based on unifying ideas of multicriteria optimization. These notions do not require certain properties of the means, which are typically assumed by the conventional approaches in statistics and which can sometimes complicate the choice of a suitable mean in some problems [8]. Instead, our approach utilizes the distance from a current point to each point of the dataset. The proximity from a current point to all points in the dataset is characterized by the vector components of which are the distances between the current point and each point of the dataset. The means are defined as the points which are nondominated with respect to the preference relation among the vectors of distances characterized by scale properties, such as equal importance or ordinality, and/or transfer principles.

It turns out that such means are typically not unique and that their sets may have a complex structure. This potentially complicates the calculation of such means for large samples. However, the advances in computer and software technologies make this computational issue less problematic.

The suggested means allow two different interpretations, either as the range of possible mean values in some specific situations characterized by scale properties, or as whole sets that characterize the chosen sample.

Among the new means introduced in this paper, the means defined with respect to relation $P_{E\Delta}$ should be of the most practical interest. The set $G^{E\Delta}(X)$ of such means has a simple structure (it is a segment $[\alpha, \beta]$), and it is stable with respect to small perturbations of the dataset. Furthermore, there exists a simple exact method for the calculation of the set $G^{E\Delta}(X)$. Namely, we have suggested analytical formulae for the calculation of the endpoints α и β.

In applications, the comparison of different multi-valued means developed in our paper may be uninteresting because they usually turn out to be incomparable under the corresponding partial preference relation. However, in some problems, the described multi-valued approach has advantages over the use of known means (see, e.g., Example 9). If, instead of the set of pn-means, we consider their corresponding centres of mass, then such centroid means are uniquely defined. The latter are equally operational as the conventional means and but are less informative than the original pn-means. For example, instead of the mean $G^{E\Delta}(X) = [\alpha, \beta]$, we may use the corresponding centroid mean (with respect to $P_{E\Delta}$) $\gamma = ½\,(\alpha + \beta)$.

The suggested new means are a useful complement to the range of conventional means used in statistics. Among further research avenues arising from our paper, let us note development of new pn-means under different assumptions about the properties of the scales of measurement and corresponding computational methods.

Acknowledgments This work is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University).

References

1. Beliakov, G., Pradera, A., Calvo, T.: Aggregation functions: a Guide for Practitioners. Springer, Berlin (2007)
2. Dalton, H.: The measurement of the inequality of incomes. Economic J. **30**, 348–361 (1920)
3. Eurostat: Real GDP per capita. https://ec.europa.eu/eurostat/databrowser/view/sdg_08_10/default/table?lang=en. (2020).
4. Fishburn, P.C.: Decision and Value Theory. Wiley, New York (1964)
5. Fishburn, P.C., Willig, R.D.: Transfer principles in income redistribution. J. Public Econom. **25**, 323–328 (1984)
6. Foster, C.: Being mean about the mean. Math. Sch. **43**, 32–33 (2014)
7. Grabisch, M., Marichal, J.-L., Mesiar, R., Pap, E.: Aggregation functions: means. Inf. Sci. **181**, 1–22 (2011)
8. Gini, C.: Le Medie. Ulet, Torino (1957)
9. Knuth, D.E.: The Art of Computer Programming: Vol. 3: Sorting and Searching, 2nd edn. Addison-Wesley, New York (1998)
10. Kricheff, R.S.: Means move – analyze the averages. In: That Doesn't Work Anymore – Retooling Investment Economics in the Age of Discruption, pp. 37–42. De Gruyter, Berlin (2018)
11. Levy, H.: Stochastic dominance and expected utility: survey and analysis. Manag. Sci. **38**, 555–593 (1992)
12. Lewontin, R., Levins, R.: The politics of averages. Capital. Nat. Social. **11**, 111–114 (2000)
13. Luce, R.D., Raiffa, H.: Games and Decisions. Wiley, New York (1957)
14. Marshall, A.W., Olkin, I.: Inequalities: Theory of Majorization and its Applications. Academic, New York (1979)
15. Mirkin, B.G.: Group choice. Wiley, New York (1979)
16. Nelson, L.S.: Some notes on averages. J. Quality Technology. **30**, 100–101 (1998)
17. Podinovskii, V.V.: Multicriterial problems with uniform equivalent criteria. USSR Comp. Math. Math. Physics. **15**, 47–60 (1975)
18. Podinovski, V.V.: On the use of importance information in MCDA problems with criteria measured on the first ordered metric scale. J. of Multi-Crit. Decis. Anal. **15**, 163–174 (2009)
19. Podinovski, V.V., Nelyubin, A.P.: Mean quantities: a multicriteria approach. Control Sciences. #. **5**, 3–16 (2020) (In Russian)
20. Podinovski, V.V., Nelyubin, A.P.: Mean quantities: a multicriteria approach. II. Control Sciences. **2**, 33–41 (2021) (In Russian)
21. Roberts, F.S.: Measurement Theory: with Applications to Decisionmaking, Utility, and Social Sciences (Encyclopedia of Mathematics and its Applications). Cambridge University Press, Cambridge (1984)
22. Smith, M.J.: Statistical Analysis. Handbook. A Comprehensive Handbook of Statistical: Concepts, Techniques and Software Tools. The Winchelsea Press, Edinburgh (2018)
23. Stevens, S.S.: On the theory of scales of measurement. Sci. New Series. **103**, 677–680 (1946)