COMPARISON MEANINGFUL DISPERSION MEASURES AND RELATED MEASURES OF LOCATION OF SIMPLE RANDOM VARIABLES

Mikhail V. Sokolov
The St. Petersburg State University
St. Petersburg
Russia
April 2009

Correspondence: Mikhail V. Sokolov. Russia, 191123, St. Petersburg, Chaikovskogo st., 62. Tel: +7-812-272-75-34. Fax: +7-812-273-40-50. E-mail address: sos-homepage@yandex.ru.

Abstract

Adopting the approach of Calvo et al. (2004, section III), in this paper we consider a particular class of dispersion measures (penalty function based dispersion measures) and related measures of location (penalty function based means) of a simple random variable. Characterizations of comparison meaningful penalty function based dispersion measures and related means in nominal, ordinal, and some numerical scales are obtained by elementary methods. Subject to scale types under consideration, they correspond to such well-known means and dispersion measures as the mode, quantiles, quasiarithmetic means, the average absolute deviation, the variance, and others. An application of the topic to decision making under risk is also given.

Keywords: measures of dispersion; measures of location; comparison meaningfulness; invariance.

1. Introduction

Different classes of dispersion measures (measures of spread) and measures of location (central tendency) of random variables are widely used in different fields of applied science: statistics, decision making under risk, expert choice, multi-criterion evaluation, the economic theory of index numbers, utility theory, sociology, etc. But a variety of functional forms of dispersion measures and measures of location sometimes leads to arbitrariness in their use.

The representational theory of measurement (Luce et al., 1990; Narens, 2002; Roberts, 1979) is capable to partly organize the situation by focusing on those algorithms of data analysis, that lead to conclusions that are stable to a change of a measurement scale of variables under consideration. In the literature this postulate is known as the requirement of meaningfulness. The concept of meaningfulness is often formalized in terms of invariance with respect to admissible transformations that define a particular scale type. Informally, one shall say that a statement involving scales is meaningful if its truth or falsity is unchanged when admissible transformations are applied to all of the scales in the statement (Roberts, 1979, p. 59). The requirement of meaningfulness is intuitive and topical in connection with that the choice of a certain measurement scale is subjective. And a researcher has an opportunity to manipulate results of analysis if its conclusions appear unstable with respect to an admissible transformation of a scale.

The requirement of meaningfulness is often applied to characterize possible functional forms of means (e.g., see Aczél and Roberts, 1989; Marichal, 1998; Orlov, 1979, chapters 3.3, 3.4; Ovchinnikov, 1996). In this paper we employ this technique to characterize meaningful dispersion measures among a particular class of dispersion measures.

Adopting the approach of Calvo et al. (2004, section III), we consider a particular class of dispersion measures (penalty function based dispersion measure) and related sets of measures of location (penalty function based means) of a random variable. To avoid technical difficulties associated with measurability of functions of random variables and the existence of considered mathematical expectations we restrict ourselves to simple random variables (discrete random variables with finite sample spaces). Characterizations of meaningful penalty function based dispersion measures in nominal, ordinal, and some numerical scales are given. The structure of the considered class is such that meaningfulness of dispersion measure implies invariance of related sets of measures of location. Hence the considered topic is related to the problems, which are extensively studied in the literature (Aczél and Roberts, 1989; Fodor and Roubens, 1995; Marichal, 1998; Marichal et al., 2005; Orlov, 1979, chapters 3.3, 3.4; Ovchinnikov, 1996; Ovchinnikov and Dukhovny, 2002).

The paper is organized as follows. In section 2 we list scale types used in the paper and their admissible transformations. In section 3 we adopt the approach of Calvo et al. (2004, section III) to introduce penalty function based dispersion measures and related measures of location. In section 4 we specialize meaningfulness property for such measures, deduce the functional equation for a penalty function of a meaningful dispersion measure, and establish invariance of the set of related measures of location. In section 5 we characterize meaningful dispersion measures in particular scales (nominal, ordinal, and some numerical scales). Finally, in section 6 an application of the subject to decision making under risk is given.

2. Scale types under consideration and their admissible transformations

In the following we denote the sets of real numbers, negative real numbers, nonnegative real numbers, and positive real numbers by R, R_-, $\overline{\mathrm{R}}_+$, and R_+, respectively.

Let $\mathrm{X} \subseteq \mathrm{R}$ be an open interval. A family $\mathrm{T}(\mathrm{X})$ of bijections of X onto itself is called a *group of admissible transformations* if $\mathrm{T}(\mathrm{X})$ forms a group with respect to the functional composition operation $\circ$. A group of admissible transformations $\mathrm{T}(\mathrm{X})$ is said to be n*-point homogeneous* (Narens, 2002, p. 54), where n is a natural number, if for any $x_1 < ... < x_n$ and $x_1' < ... < x_n'$ in X there exists $T \in \mathrm{T}(\mathrm{X})$ such that $T(x_i) = x_i'$ for $i = 1,...,n$. A group of admissible transformations $\mathrm{T}(\mathrm{X})$ is n*-point unique* if for any $x_1 < ... < x_n$ and $x_1' < ... < x_n'$ in X there exists at most one $T \in \mathrm{T}(\mathrm{X})$ such that $T(x_i) = x_i'$ for $i = 1,...,n$. A group of admissible transformations is said to be *ordered* if it consists of order preserving functions. Obviously, an element of an ordered group of admissible transformations is continuous.

In the case of *nominal* and *ordinal* scales, the group of admissible transformations forms, respectively, the group of bijections

$$\mathrm{T}_N(\mathrm{X}) = \{T : T \text{ is a bijection of } \mathrm{X} \text{ onto itself}\}$$

and the automorphism group

$$\mathrm{T}_O(\mathrm{X}) = \{T : T \text{ is an increasing bijection of } \mathrm{X} \text{ onto itself}\}$$

of the interval X. Another important group of admissible transformations is (see Luce et al., 1990, chapter 20; Narens, 2002, section 2.3) the group

$$\mathrm{T}_{\mathrm{A}}^{(f)}(\mathrm{X}) = \{T : T(x) = f^{-1}(af(x) + b),\ a \in \mathrm{A},\ b \in \mathrm{R},\ x \in \mathrm{X}\},$$

where A is a subgroup of R_+ under multiplication and f is an increasing bijection of X onto R. If A is the trivial subgroup of R_+, then $\mathrm{T}_{\mathrm{A}}^{(f)}(\mathrm{X})$ corresponds to a *difference* scale when f is the identity function and scales conjugate to a difference scale for a general f. If A is a proper nontrivial subgroup of R_+, then $\mathrm{T}_{\mathrm{A}}^{(f)}(\mathrm{X})$ corresponds to a *discrete interval* scale when f is the identity function and scales conjugate to a discrete interval scale for a general f. Finally, the case $\mathrm{A}=\mathrm{R}_+$ corresponds to an *interval* scales when f is the identity function and scales conjugate to an interval scale for a general f.

3. Penalty function based dispersion measures and related means

Let $\mathrm{X} \subseteq \mathrm{R}$ be an open interval. By $\Im$ denote the set of all random variables whose distribution is concentrated on a set of finite number of elements of X. Random variables will usually be denoted by capitals; however, for $x \in \mathrm{X}$, $x_{\text{a.s.}}$ will denote the deterministic random variable concentrated on x. Given a random variable $X \in \Im$, the probability distribution associated with X will be denoted $\{x_1, p_1; \ldots; x_n, p_n\}$, where $x_i \in \mathrm{X}$, $p_i = \Pr\{X = x_i\} > 0$, $i = 1, \ldots, n$. Without loss of generality, in what follows it is assumed that $x_1 < \ldots < x_n$.

To define the terms "dispersion measure a random variable" and "measure of location of a random variable", we adopt a minor alteration of the approach of Calvo et al. (2004, section III) to so called penalty function based aggregation (see also Dershem, 1975; Mesiar, 2007; Orlov, 1979, chapter 4.4; Ostasiewicz and Ostasiewicz, 2000, section 4.2 for the related approaches).

A map $L : \mathrm{X}^2 \to \overline{\mathrm{R}}_+$ is said to be a *penalty function* if:

(i) $L(x, y) = 0$ if and only if $x = y$;

(ii) $L(x, y) \le L(x, z)$ whenever $x \le y \le z$ or $x \ge y \ge z$.

The map $D_L : \Im \times \mathrm{X} \to \overline{\mathrm{R}}_+$ defined by

$$D_L[X; y] = \sum_{i=1}^{n} p_i L(x_i, y), \quad X \in \Im, \; y \in \mathrm{X} \tag{1}$$

is called the *penalty function based dispersion measure* (for short, L*-dispersion*). The expected penalization (1) is the L*-dispersion of the random variable* X *about the point* y.

The number

$$V_L[X] = \inf_{y \in \mathrm{X}} D_L[X; y] \tag{2}$$

is said to be the L*-dispersion of the random variable* X.

Elements of the set

$$M_L[X] = \{y \in \mathrm{X} : D_L[X; y] = V_L[X]\}$$

are called the *penalty function based means* (for short, L*-means*) *of the random variable* X.

For example, the penalty function $L(x, y) = (x - y)^2$ yields the variance and expectation of a random variable X as $V_L[X]$ and $M_L[X]$, respectively.

In general, infimum in (2) may not be attained on X; the set $M_L[X]$ may be empty.

The term “mean” is quite natural in the above definition. Indeed, by (i) and (ii), it follows that each element of $M_L[X]$ is a Cauchy mean of a random variable X. In other words, $M_L[X] \subseteq [x_1, x_n]$.

The introduced class of dispersion measures and the related measures of location is quite rich. In particular, it includes the following means:

- the mode, induced by the penalty function

$$L(x,y) = \begin{cases} 0 & \text{if } x = y \\ d & \text{otherwise} \end{cases}, \tag{3}$$

where d is a positive constant;

- quantiles, induced by the penalty function

$$L(x,y) = \begin{cases} x - y & \text{if } x \geq y \\ d(y-x) & \text{otherwise} \end{cases}, \tag{4}$$

where d is a positive constant;

- quasiarithmetic means weighted by the weight function (Aczél and Daróczy, 1963), induced by the penalty function

$$L(x,y) = g(x)\left(f(x) - f(y)\right)^2,$$

where f is a strictly monotonic continuous function on X, g is a positive function on X;

- some types of so called *M*-estimators (Huber, 1964);
- some other implicit (or quasideviation) (Fishburn, 1986; Páles, 1988) and distance (or metric) based (Ostasiewicz and Ostasiewicz, 2000, section 4.2) means.

4. Meaningfulness of penalty function based dispersion measures and invariance of related means

Suppose that elements of the set $\Im$ are measured on a scale with a nontrivial group of admissible transformations $\mathrm{T(X)}$.

An L-dispersion on $\Im$ is called *comparison meaningful with respect to a group of admissible transformations* $\mathrm{T(X)}$ (for short, $\mathrm{T(X)}$-*meaningful*) if for each $T \in \mathrm{T(X)}$

$$D_L[X;y] \geq D_L[X;z],\ X \in \Im,\ y,z \in \mathrm{X} \Rightarrow D_L[T(X);T(y)] \geq D_L[T(X);T(z)]. \tag{5}$$

The given definition asserts preservation of the ordinal structure of L-dispersions under an admissible transformation and is equivalent to 1-meaningfulness (Narens, 2002, section 2.6) of the statement “$D_L[X;y] \geq D_L[X;z]$”. Since $\mathrm{T(X)}$ is a group, the assertion “$\Rightarrow$” in (5) can be replaced by “$\Leftrightarrow$”.

Lemma 1.

Let an L-dispersion on $\Im$ be meaningful with respect to a group of admissible transformations $\mathrm{T}(\mathrm{X})$. Then the set $M_L[X]$ is $\mathrm{T}(\mathrm{X})$-*invariant*. That is,

$$M_L[T(X)] = T(M_L[X]),\ T \in \mathrm{T}(\mathrm{X}),\ X \in \Im. \tag{6}$$

Proof.

Given $X \in \Im$. Let $z \in M_L[X]$, then

$D_L[X;y] \geq D_L[X;z]$ for any $y \in \mathrm{X}$.

By (5),

$D_L[T(X);T(y)] \geq D_L[T(X);T(z)]$ for any $T \in \mathrm{T}(\mathrm{X})$.

Since T is a bijection, we obtain that $T(z)$ is an element of $M_L[T(X)]$. ■

The reverse is false: the median is a counterexample in ordinal scales. Relation (6) states that L-means are measured on the same scale as the random variable X. We notice that in contrast of conventional definition of invariance (e.g., see Ovchinnikov and Dukhovny, 2002) in (6) we deal with the equality of sets rather than values.

The described property of meaningfulness substantially restricts the possible functional form of a penalty function. The following proposition formalizes this observation.

Proposition 1.

Let $\mathrm{T}(\mathrm{X})$ be a group of continuous admissible transformations. An L-dispersion on $\Im$ is $\mathrm{T}(\mathrm{X})$-meaningful if and only if for each $T \in \mathrm{T}(\mathrm{X})$ there exists a constant $q_T > 0$ such that

$$L(T(x), T(y)) = q_T L(x, y) \quad \forall x, y \in \mathrm{X}. \tag{7}$$

The map q that takes each $T \in \mathrm{T}(\mathrm{X})$ to q_T is a group homomorphism from $\mathrm{T}(\mathrm{X})$ to the subgroup $\{q_T : T \in \mathrm{T}(\mathrm{X})\}$ of positive real numbers under multiplication:

$$q_{T \circ T'} = q_T q_{T'},\ T, T' \in \mathrm{T}(\mathrm{X}). \tag{8}$$

To prove Proposition 1, we need the following lemma.

Lemma 2.

Let L and L' be two penalty functions on X.

$$D_L[X;y] \geq D_L[X;z],\ X \in \Im,\ y, z \in \mathrm{X} \ \Rightarrow\ D_{L'}[X;y] \geq D_{L'}[X;z] \tag{9}$$

if and only if there exists a positive constant q such that

$$L'(x,y) = qL(x,y),\ \forall x,y \in \mathrm{X}. \quad (10)$$

Proof.

If L and L' satisfy (10) then we obviously have (9).

Suppose L and L' satisfy (9); then

$$D_L[X;y] = D_L[X;z] \Leftrightarrow D_{L'}[X;y] = D_{L'}[X;z],\ X \in \mathfrak{I},\ y,z \in \mathrm{X}. \quad (11)$$

Given $x,y,z \in \mathrm{X}$, $x < y < z$. If $L(x,y) = L(x,z)$ then the application of (11) with $X = x_{\text{a.s.}}$ yields $L'(x,y) = L'(x,z)$. If $L(x,y) < L(x,z)$ then consider a binary random variable X with the probability distribution $\{x,p;z,1-p\}$, where the probability $p = p(x,y,z) \in (0,1)$ is chosen such that

$$D_L[X;y] = pL(x,y) + (1-p)L(z,y) = pL(x,z) = D_L[X;z]. \quad (12)$$

By (i) and (ii), such defined p exists and is unique. Using (11), we get

$$D_{L'}[X;y] = pL'(x,y) + (1-p)L'(z,y) = pL'(x,z) = D_{L'}[X;z]. \quad (13)$$

Combining (12) and (13), we have

$$\frac{L(x,z) - L(x,y)}{L(z,y)} = \frac{1-p}{p} = \frac{L'(x,z) - L'(x,y)}{L'(z,y)}. \quad (14)$$

(14) remains valid in the case $L(x,y) = L(x,z)$.

Now consider a ternary random variable Y with probability distribution $\{x,p;y,q;z,1-p-q\}$. Let S be the set of solutions (p,q) of the equation

$$D_L[Y;x] = qL(y,x) + (1-p-q)L(z,x) = pL(x,z) + qL(y,z) = D_L[Y;z] \quad (15)$$

that satisfy $p > 0$, $q > 0$, and $p + q < 1$.

Denote

$$A = 1 + \frac{L(x,z)}{L(z,x)},\ B = 1 + \frac{L(y,z) - L(y,x)}{L(z,x)},$$

then (15) can be written in the form

$$Ap + Bq = 1. \quad (16)$$

By (i), $A > 1$. Hence, the set S is infinite.

Using (11), we get

$$A'p + B'q = 1 \text{ whenever } (p,q) \in S, \quad (17)$$

where

$$A' = 1 + \frac{L'(x,z)}{L'(z,x)},\ B' = 1 + \frac{L'(y,z) - L'(y,x)}{L'(z,x)}.$$

The system of linear equations (16) and (17) has infinitely many solutions if and only if

$$\operatorname{rank}\begin{pmatrix} A & B \\ A' & B' \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A & B & 1 \\ A' & B' & 1 \end{pmatrix} = 1,$$

where $\operatorname{rank}(\cdot)$ is the rank of a matrix. Hence $A = A'$, $B = B'$, and

$$\frac{L(x,z)}{L(z,x)} = A - 1 = A' - 1 = \frac{L'(x,z)}{L'(z,x)}, \tag{18}$$

$$\frac{L(y,z) - L(y,x)}{L(z,x)} = B - 1 = B' - 1 = \frac{L'(y,z) - L'(y,x)}{L'(z,x)}. \tag{19}$$

Relation (18) is valid for any $x < z$. Thus, equalities (14) and (19) can be written in the form

$$\frac{L(x,z) - L(x,y)}{L(y,z)} = \frac{L'(x,z) - L'(x,y)}{L'(y,z)} \tag{20}$$

and

$$\frac{L(y,z) - L(y,x)}{L(x,z)} = \frac{L'(y,z) - L'(y,x)}{L'(x,z)}, \tag{21}$$

respectively.

Multiplying (20) and (21), an easy calculation gives

$$\frac{L(x,z)L(y,x) + L(x,y)L(y,z) - L(x,y)L(y,x)}{L(y,z)L(x,z)} =$$
$$= \frac{L'(x,z)L'(y,x) + L'(x,y)L'(y,z) - L'(x,y)L'(y,x)}{L'(y,z)L'(x,z)}.$$

Therefore,

$$\frac{\dfrac{L(y,x)}{L(x,y)} + \dfrac{L(y,z) - L(y,x)}{L(x,z)}}{\dfrac{L(y,z)}{L(x,y)}} = \frac{\dfrac{L'(y,x)}{L'(x,y)} + \dfrac{L'(y,z) - L'(y,x)}{L'(x,z)}}{\dfrac{L'(y,z)}{L'(x,y)}}. \tag{22}$$

Applying (18) and (21) to the numerator of the right part of (22), we obtain

$$\left[\frac{L(y,z)}{L(x,y)} - \frac{L'(y,z)}{L'(x,y)}\right]\left(\frac{L(y,x)}{L(x,y)} + \frac{L(y,z) - L(y,x)}{L(x,z)}\right) = 0. \tag{23}$$

Taking into account (i) and (ii), we get that the second brackets of (23) are positive and can be cancelled. Combining (23), (20), and (18), we have

$$\frac{L(z,y)}{L'(z,y)} = \frac{L(y,z)}{L'(y,z)} = \frac{L(x,y)}{L'(x,y)} = \frac{L(x,z)}{L'(x,z)} = \frac{L(z,x)}{L'(z,x)} = q,$$

where q is a positive constant. (10) now follows from the arbitrariness of the choice $x < y < z$. ■

Proof of Proposition 1.

Obviously, for any penalty function L such that (7) is satisfied, we have (5). We have to prove our statement only in the other direction.

Let an L-dispersion on $\mathfrak{I}$ be $\mathrm{T}(\mathrm{X})$-meaningful. Given $T \in \mathrm{T}(\mathrm{X})$, since T is continuous and therefore strictly monotone, the map $L'(\cdot,\cdot) = L(T(\cdot),T(\cdot))$ is a penalty function on X. Relation (7) now follows from Lemma 2.

Given $T, T' \in \mathrm{T}(\mathrm{X})$, iterating (7), we obtain

$$q_{T \circ T'} L(x,y) = L(T \circ T'(x), T \circ T'(y)) = q_T L(T'(x),T'(y)) = q_T q_{T'} L(x,y),\; x,y \in \mathrm{X}. \tag{24}$$

(24) with $x \neq y$ yields (8). ■

If L is a metric on X then Proposition 1 states that $\mathrm{T}(\mathrm{X})$ is a subgroup of the dilation group (the group of similarity transformations) of the metric space (X, L).

5. Meaningful penalty function based dispersion measures and related means in particular scales

The mode and the median are known as preferable Cauchy means for nominally and ordinally measured variables, respectively. Characterizations of the order statistics as the only invariant symmetric Cauchy means in an ordinal scale (under some regularity assumptions) are well known (Fodor and Roubens, 1995; Marichal et al., 2005; Orlov, 1979, chapter 3.3; Ovchinnikov, 1996). Invariant Cauchy (and some other) means in ratio and interval scales are characterized by Aczél and Roberts (1989) (see also Marichal, 1998) and form quite a rich class. In this section we characterize possible functional forms of meaningful L-dispersions measures in these scales.

Proposition 2.

Let an L-dispersion on $\mathfrak{I}$ be $\mathrm{T}_{\mathrm{A}}^{(f)}(\mathrm{X})$-meaningful. By $r(\mathrm{A})$ denote the generating rank of A.

1. If $r(\mathrm{A}) = 0$, then L has the form

$$L(x,y) = h(f(x))\, g(f(y) - f(x)),$$

where the function $g : \mathrm{R} \to \mathrm{R}_+$ is non-decreasing on R_+, non-increasing on R_-, $g(x) = 0$ if and only if $x = 0$; h is an arbitrary positive solution of the exponential Cauchy functional equation

$$h(u+v) = h(u)h(v),\; (u,v) \in \mathrm{R}^2. \tag{25}$$

2. If $r(\mathrm{A}) = 1$, then L has the form

$$L(x,y) = d_{\mathrm{sgn}(x-y)}\left(\ln\left| f(x) - f(y)\right|\right) \left| f(x) - f(y)\right|^c,\; x \neq y,$$

where c is a nonnegative constant, $d_i : \mathrm{R} \to \mathrm{R}_+$, $i = -1,1$ are $\ln a_0$-periodic functions, where a_0 is a unique element of a linearly independent generating set of A.

3. If $r(\mathrm{A}) \geq 2$ then L has the form

$$L(x,y) = d_{\mathrm{sgn}(x-y)} \left| f(x) - f(y) \right|^c, \ x \neq y, \tag{26}$$

where $d_{-1}, d_1 > 0$, and $c \geq 0$ are some constants.

Proof.

In the case $\mathrm{T}(\mathrm{X}) = \mathrm{T}_{\mathrm{A}}^{(f)}(\mathrm{X})$ the constant q_T in Proposition 1 depends only on $a \in \mathrm{A}$ and $b \in \mathrm{R}$, i.e. $q_T = q(a,b)$. Denote $G(x,y) = L\left(f^{-1}(x), f^{-1}(y)\right)$, $(x,y) \in \mathrm{R}^2$, then equations (8) and (7) go over into the forms

$$q(a_1 a_2, a_1 b_2 + b_1) = q(a_1, b_1) q(a_2, b_2), \ a_1, a_2 \in \mathrm{A}, \ b_1, b_2 \in \mathrm{R} \tag{27}$$

and

$$G(ax+b, ay+b) = q(a,b) G(x,y), \ a \in \mathrm{A}, \ x, y, b \in \mathrm{R}, \tag{28}$$

respectively. We notice, that G is a penalty function on R.

1. In the case $r(\mathrm{A}) = 0$ (28) is up to changing the variables the generalized homogeneity functional equation. Its the general solution is given by (e.g., see Aczél and Roberts, 1989, case 2)

$$G(x,y) = h(x) g(y-x),$$

where g is an arbitrary function and the function h defined by $h(b) = q(1,b)$ is an arbitrary positive solution of the exponential Cauchy functional equation (25). From (i) and (ii), it follows that g is nonnegative, non-decreasing on R_+, non-increasing on R_-, and $g(x) = 0$ if and only if $x = 0$.

In the case $r(\mathrm{A}) \geq 1$, by symmetry of the right part of (27),

$$q(a_1, a_1 b_2 + b_1) = q(a_1, b_1 + b_2). \tag{29}$$

For given $a_1 \in \mathrm{A} \setminus \{1\}$ and $b \in \mathrm{R}$, substituting $b_1 = b/(1-a_1)$ and $b_2 = -b/(1-a_1)$ in (29), we get

$$q(a_1, b) = q(a_1, 0), \tag{30}$$

i.e. $q(a,b)$ is independent of b whenever $a \neq 1$. Setting $a_1 = 1$ and $b_2 = 0$ in (27), by (30), it follows that $q(1, b_1) \equiv 1$ for all $b_1 \in \mathrm{R}$. Therefore, (30) holds for all $a_1 \in \mathrm{A}$ and $b \in \mathrm{R}$, and (28) reduces to

$$G(ax+b, ay+b) = q(a,0) G(x,y), \ a \in \mathrm{A}, \ x, y, b \in \mathrm{R}. \tag{31}$$

In what follows we consider (31) on the set $\{(x,y) \in \mathrm{R}^2 : x < y\}$. The case $\{(x,y) \in \mathrm{R}^2 : x > y\}$ can be dealt with in a similar way. Denote $g(y) = G(0,y)$, then substituting $x = 0$ and $b = 0$ in (31), we get

$$g(ay) = q(a,0) g(y), \ a \in \mathrm{A}, \ y \in \mathrm{R}_+. \tag{32}$$

2. If $r(\mathrm{A})=1$ then denote $c=\ln q(a_0,0)/\ln a_0$, where a_0 is a unique element of a linearly independent generating set of A. The general solution of (32) is given by

$$g(y)=d(\ln y)y^c,\ y\in \mathrm{R}_+,$$

where $d:\mathrm{R}\to\mathrm{R}_+$ is a $\ln a_0$-periodic function and c is a constant. Setting $a=1$ and $b=-x$ in (31), we get

$$G(x,y)=\frac{G(0,y-x)}{q(1,0)}=g(y-x)=d\big(\ln(y-x)\big)(y-x)^c,\ x<y.$$

By (i) and (ii), it follows that $c\geq 0$.

3. If $r(\mathrm{A})\geq 2$ then a generating set of A contains at least two elements a_0 and a_0' such that $\ln a_0/\ln a_0'$ is an irrational number. Without loss of generality, we can assume that $a_0>1$ and $a_0'>1$. As above, there exist a $\ln a_0$-periodic function $d:\mathrm{R}\to\mathrm{R}_+$, a $\ln a_0'$-periodic function $d':\mathrm{R}\to\mathrm{R}_+$, and nonnegative constants c, c' such that

$$g(y)=d(\ln y)y^c=d'(\ln y)y^{c'},\ y\in\mathrm{R}_+.$$

By (i) and (ii), g is positive and non-decreasing on R_+. Therefore, the functions defined by $\ln d(\ln y)=\ln g(y)-c\ln y$ and $\ln d'(\ln y)=\ln g(y)-c'\ln y$ are of bounded variation (as a difference of two non-decreasing functions) on the intervals $[1,a_0]$ and $[1,a_0']$, respectively. In particular, these functions are bounded. Hence, the left-hand side of the equality

$$\ln d'(\ln y)-\ln d'(\ln y)=(c-c')\ln y$$

is bounded, then so does the right-hand side. Thus, $c=c'$, $d=d'$, and

$$d(\ln a_0)=d(0)=d(\ln a_0').$$

Let $y_0\in\mathrm{R}_+$. Since g is non-decreasing and the function $y\to y^c$ is continuous, the values $d(\ln y_0-)=\lim_{y\uparrow y_0}d(\ln y)$, $d(\ln y_0+)=\lim_{y\downarrow y_0}d(\ln y)$ are well defined and

$$d(\ln y_0-)\leq d(\ln y_0)\leq d(\ln y_0+).$$

Since $\ln a_0$ and $\ln a_0'$ are incommensurable, the numbers $n\ln a_0+m\ln a_0'$, where n and m are integers, lie everywhere dense among R. Therefore,

$$d(\ln y_0-)=d(\ln y_0)=d(\ln y_0+)=d(0).$$

Thus, d is equal identically to a constant. ■

An L-dispersion $D_L[X;y]$ with the penalty function (26) is known as the c-th skew-symmetric absolute moment of a random variable $f(X)$ about a point $f(y)$. If f is linear, then $D_L[X;f^{-1}(0)]$ corresponds to the risk measure introduced by Luce (1980, Theorem 4). If f is linear and $d_{-1} = d_1$, then $D_L[X;y]$ is a subclass of the three-parameter family of risk measures introduced by Stone (1973); the related L-means were characterized by Bickel and Lehmann (1975, Theorem 2). If $c = 0$ and $d_{-1} = d_1$ then the related L-mean is equal to the mode, if $c = 1$ then it equals the quantile of order $d_1/(d_{-1} + d_1)$ regardless of f, if $c = 2$ then it equals the skew-symmetric quasiarithmetic mean with the generator f, and it tends to the quasimidrange $f^{-1}([f(x_1) + f(x_n)]/2)$ regardless of d_{-1}, d_1, and p_i as $c \to +\infty$. Obviously, for any $c \geq 0$ the set $M_L[X]$ of related L-means is nonempty, if $c > 1$ then it contains a single element, if $c \in [0,1)$ then $M_L[X] \subseteq \{x_1,..., x_n\}$.

The next proposition is based on the fundamental result of Alper and Narens (see Luce et al., 1990, chapter 20, theorem 5) and slightly generalizes the third part of Proposition 2.

Proposition 3.

An L-dispersion on $\Im$ is meaningful with respect to a 2-point homogeneous ordered group of admissible transformations $\mathrm{T}(\mathrm{X})$ if and only if there exist an increasing bijection f of X onto R and constants $d_{-1}, d_1 > 0$, $c \geq 0$ such that L has the form (26).

Proof.

Let the penalty function be given by (26). Then the L-dispersion on $\Im$ is $\mathrm{T}^{(f)}_{\mathrm{R}_+}(\mathrm{X})$-meaningful. We have to prove our statement only in the other direction.

First we recall that elements of $\mathrm{T}(\mathrm{X})$ are continuous and strictly increasing.

If $\mathrm{T}(\mathrm{X})$ is 2-point unique, then the application of the Alper–Narens theorem (Luce et al., 1990, chapter 20, theorem 5) yields that there exists an increasing bijection f of X onto R such that $\mathrm{T}(\mathrm{X}) = \mathrm{T}^{(f)}_{\mathrm{R}_+}(\mathrm{X})$. The result now follows from the third part of Proposition 2.

If $\mathrm{T}(\mathrm{X})$ is not 2-point unique then there exist $x_0 \neq y_0 \neq z_0$ in X and $T_1, T_2 \in \mathrm{T}(\mathrm{X})$ such that $T_1(x_0) = T_2(x_0)$, $T_1(y_0) = T_2(y_0)$, and either $T_1(x) \neq T_2(x)$ for all $x \in (y_0, z_0]$ (if $y_0 < z_0$), or $T_1(x) \neq T_2(x)$ for all $x \in [z_0, y_0)$ (if $y_0 > z_0$). Below we consider only the case $y_0 < z_0$, the other one can be dealt with in a similar way. Since $\mathrm{T}(\mathrm{X})$ is a group, the function

$$T_0 = T_1^{-1} \circ T_2$$

is an element of $\mathrm{T}(\mathrm{X})$. Obviously, $T_0(x_0) = x_0$, $T_0(y_0) = y_0$, and $T_0(x) \neq x$ for all $x \in (y_0, z_0]$. Without loss of generality, we can assume that $T_0(x) < x$ for all $x \in (y_0, z_0]$.[1]

[1] If the inverse inequality holds, then consider T_0^{-1} instead of T_0.

By Proposition 1, there exists a constant $q_{T_0} > 0$ such that

$L(x_0, y_0) = L(T_0(x_0), T_0(y_0)) = q_{T_0} L(x_0, y_0)$.

Hence, $q_{T_0} = 1$.

Denote $l(x) = \lim_{y \downarrow x} L(x, y)$. From (i) and (ii), it follows that the limit is well-defined and nonnegative. By Proposition 1, for each $T \in \mathrm{T}(\mathrm{X})$ there exists a constant $q_T > 0$ such that

$$l(T(x)) = q_T l(x) \quad \forall x \in \mathrm{X}. \tag{33}$$

Then

$0 < L(y_0, z_0) = q_{T_0} L(y_0, z_0) = L(T_0(y_0), T_0(z_0)) = L(y_0, T_0(z_0)) = L(y_0, T_0 \circ T_0(z_0)) = \ldots = l(y_0)$.

Combining this with (33) and taking into account 2-point homogeneity of $\mathrm{T}(\mathrm{X})$, we obtain that $l(x) > 0$ for all $x \in \mathrm{X}$. Let $T, T' \in \mathrm{T}(\mathrm{X})$ and $T(z') = T'(z')$ for some $z' \in \mathrm{X}$, then

$q_T l(z') = l(T(z')) = l(T'(z')) = q_{T'} l(z')$.

Hence, $q_T = q_{T'}$ whenever there is a point $z' \in \mathrm{X}$ such that $T(z') = T'(z')$.

Since $\mathrm{T}(\mathrm{X})$ is 2-point homogeneous, for a given $T \in \mathrm{T}(\mathrm{X})$ there exist $T' \in \mathrm{T}(\mathrm{X})$ and $x', y' \in \mathrm{X}$ such that $T'(x') = T(x')$ and $T'(y') = y'$. Therefore, $q_T = q_{T'} = q_I = 1$, where $I \in \mathrm{T}(\mathrm{X})$ is the identity transformation.

Thus,

$L(T(x), T(y)) = L(x, y)$ for all $x, y \in \mathrm{X}$, $T \in \mathrm{T}(\mathrm{X})$.

Taking into account 2-point homogeneity of $\mathrm{T}(\mathrm{X})$, we obtain (26) with $c = 0$. ■

Since the group $\mathrm{T}_O(\mathrm{X})$ is 2-point homogeneous, but not 2-point unique, we obtain

Corollary 1.

An L-dispersion on $\Im$ is $\mathrm{T}_O(\mathrm{X})$-meaningful if and only if the penalty function has the form

$$L(x, y) = d_{\mathrm{sgn}(x-y)}, \quad x \neq y, \tag{34}$$

where d_{-1}, d_1 are some positive constants.

An unexpected result of Corollary 1 is that the penalty function (4) doesn't induce a $\mathrm{T}_O(\mathrm{X})$-meaningful L-dispersion in spite of $\mathrm{T}_O(\mathrm{X})$-invariance of the corresponding L-means (the quantiles). However, it can easily be checked that such an L-dispersion satisfies $\mathrm{T}_O(\mathrm{X})$-meaningfulness if y and z in (5) are taken from an interval of the special form (see the next proposition).

Proposition 4.

Let D_L be an L-dispersion on $\Im$. For any nondegenerate random variable $X \in \Im$ there exists $i \in \{1,...,n-1\}$ such that for each $T \in \mathrm{T}_O(\mathrm{X})$

$$D_L[X;y] \ge D_L[X;z],\ y,z \in [x_i, x_{i+1}] \ \Rightarrow\ D_L[T(X);T(y)] \ge D_L[T(X);T(z)]; \tag{35}$$

if and only if either L has the form (34), or

$$L(x,y) = d_{\mathrm{sgn}(x-y)}\, g(\min\{x,y\}),\ x \ne y, \tag{36}$$

or

$$L(x,y) = d_{\mathrm{sgn}(x-y)} \left| f(x) - f(y) \right|,\ x \ne y, \tag{37}$$

where d_{-1}, d_1 are positive constants, $f : \mathrm{X} \to \mathrm{R}$, $g : \mathrm{X} \to \mathrm{R}_+$ are strictly decreasing functions.

Proof.

Obviously, if L is of the form (34), (36), or (37) then D_L satisfies (35). We have to prove our statement only in the other direction.

Let D_L be an L-dispersion satisfying aforementioned conditions. Fix arbitrary $x_0 < y_0 < z_0$ in X and denote

$$d = L(z_0, x_0) / L(x_0, z_0) > 0,\quad s = (L(x_0, z_0) - L(x_0, y_0)) / L(y_0, z_0) \ge 0.$$

Given $x < y < z$. Since the group $\mathrm{T}_O(\mathrm{X})$ is 3-point homogeneous, there exists $T_0 \in \mathrm{T}_O(\mathrm{X})$ such that $T_0(x) = x_0$, $T_0(y) = y_0$, and $T_0(z) = z_0$. Denote $L'(\cdot,\cdot) = L(T_0(\cdot), T_0(\cdot))$.

From relation (35) with a binary random variable it follows that relations (12)–(14) of Lemma 2 remain valid. Consider a binary random variable X with the probability distribution $\{x, p; z, 1-p\}$, where the probability $p = p(x,z) \in (0,1)$ is chosen such that

$$D_L[X;x] = (1-p)L(z,x) = pL(x,z) = D_L[X;z].$$

Arguing as above (see relations (12)–(14)), we get (18). Hence, (20) is also valid.

Using (18) and (20), we obtain

$$\frac{L(z,x)}{L(x,z)} = \frac{L'(z,x)}{L'(x,z)} = \frac{L(T_0(z), T_0(x))}{L(T_0(x), T_0(z))} = \frac{L(z_0, x_0)}{L(x_0, z_0)} = d, \tag{38}$$

$$\frac{L(x,z) - L(x,y)}{L(y,z)} = \frac{L(T_0(x), T_0(z)) - L(T_0(x), T_0(y))}{L(T_0(y), T_0(z))} = \frac{L(x_0, z_0) - L(x_0, y_0)}{L(y_0, z_0)} = s. \tag{39}$$

By (39), it follows that for any $w < x < y < z$ in X

$L(w,y) - L(w,x) = sL(x,y)$ and $L(w,z) - L(w,x) = sL(x,z)$.

Subtracting the first equation from the second one and using (39), we obtain

$$L(w,z) - L(w,y) = s(L(x,z) - L(x,y)) \ \Rightarrow\ sL(y,z) = s^2 L(y,z) \ \Rightarrow\ s = s^2.$$

Hence, $s \in \{0,1\}$.

If $s = 0$ then there exists a function $g : \mathrm{X} \to \mathrm{R}_+$ such that

$L(x,z) = g(x)$ for any $x < z$.

The application of (38) yields

$$L(z,x) = dg(x),\; x < z. \qquad (40)$$

From (ii) and (40), it follows that g is non-increasing. Fix arbitrary $w < x < y < z$ and consider a binary random variable X with the probability distribution $\{w, p; z, 1-p\}$, $p \in (0,1)$. By (35),

$$D_L[X;x] = pg(w) + (1-p)dg(x) \geq pg(w) + (1-p)dg(y) = D_L[X;y]$$

so that

$$D_L[T(X);T(x)] = pg(T(w)) + (1-p)dg(T(x)) \geq pg(T(w)) + (1-p)dg(T(y)) = D_L[T(X);T(y)],$$

for any $T \in \mathrm{T}_O(\mathrm{X})$. Since elements of $\mathrm{T}_O(\mathrm{X})$ are ordered,

$$g(x) \geq g(y),\; x < y \;\Rightarrow\; g(T(x)) \geq g(T(y)),\; T \in \mathrm{T}_O(\mathrm{X}). \qquad (41)$$

From (41) and 2-point homogeneity of $\mathrm{T}_O(\mathrm{X})$, it follows that either g is equal identically to a constant and, therefore, L has the form (34), or g is strictly decreasing and, therefore, L has the form (36).

If $s = 1$ then (39) reduces to the Sincov functional equation (e.g., see Aczél, 1966, section 5.1.2)

$$L(x,z) = L(x,y) + L(y,z) \qquad (42)$$

on the domain $\{(x,y,z) \in \mathrm{X}^3 : x \leq y \leq z\}$. Fix arbitrary $x_0 \in \mathrm{X}$ and denote

$$f(y) = \begin{cases} L(y,x_0) & \text{if } y \leq x_0 \\ -L(x_0,y) & \text{if } y > x_0 \end{cases}.$$

Given $x' \leq y'$ in X. If $x_0 \leq x' \leq y'$ then substituting $x = x_0$, $y = x'$, $z = y'$ in (42), we get

$$L(x',y') = L(x_0,y') - L(x_0,x') = f(x') - f(y'). \qquad (43)$$

If $x' \leq x_0 \leq y'$, then substituting $x = x'$, $y = x_0$, $z = y'$ in (42), we have

$$L(x',y') = L(x',x_0) + L(x_0,y') = -f(y') + f(x'). \qquad (44)$$

Finally, if $x' \leq y' \leq x_0$, then substituting $x = x'$, $y = y'$, $z = x_0$ in (42), we get

$$L(x',y') = L(x',x_0) - L(y',x_0) = f(x') - f(y'). \qquad (45)$$

Combining (43), (44), and (45), we obtain

$L(x',y') = f(x') - f(y')$ whenever $x' \leq y'$.

The application of (38) yields

$L(x',y') = dL(y',x') = d(f(y') - f(x'))$ whenever $x' \leq y'$.

By (i) and (ii), it follows that f is strictly decreasing. Finally, we get (37). ■

Corollary 2.

An L-dispersion on $\Im$ is $\mathrm{T}_N(\mathrm{X})$-meaningful if and only if its penalty function has the form (3).

6. An application

Economic decisions on a set of random outcomes are sometimes taken on the basis of only two parameters, which are often referred to as "return" and "risk". The first parameter $M[X]$ is some measure of location and the second one $V[X]$ is some measure of dispersion of an underlying random outcome X. This approach to economic decisions can be considered (under certain conditions) as a slight generalization of ordinal utility theory. Indeed, consider a (complete and transitive) preference relation $\succeq'$ on $\Im$. Let each random outcome $X \in \Im$ have a unique certainty equivalent, and let M be the functional that takes each random variable X to its certainty equivalent. If the preference relation $\succeq'$ conforms to the relation of first-order stochastic dominance, then M is a utility function for $\succeq'$:

$X \succeq' Y$ if and only if $M[X] \geq M[Y]$

and $M[X]$ is a Cauchy mean of X. Hence $M[X]$ can be referred to as "return" of X.

This approach reduces the problem of choosing a rational economic decision from the set $\Im$ to the analysis of a preference relation on the set of ordered pairs $\{(M[X], V[X]),\ X \in \Im\}$. A well-known example of such an approach is mean-variance portfolio analysis of Markowitz (1952) and Tobin (1958), where the mathematical expectation and the variance of a random outcome are used as "return" and "risk", respectively.

In this section we will consider the particular case of this approach with $M = M_L$ and $V = V_L$ and will try to characterize meaningful preference relations on the set $\{(M_L[X], V_L[X]),\ X \in \Im\}$.

Let L be a penalty function on X such that for each $X \in \Im$ $M_L[X]$ is a one-element set. In the following we identify $M_L[X]$ with its unique element. Denote

$\mathrm{S}_L = \{(M_L[X], V_L[X]),\ X \in \Im\}$.

Consider a (complete and transitive) preference relation $\succeq$ on S_L. By $\sim$ and $\succ$ we denote the symmetric (the *indifference* relation) and asymmetric (the *strict preference* relation) parts of $\succeq$, respectively.

A preference relation $\succeq$ is *nondegenerate* if there exist $(m,v),(m',v')\in S_L$ such that $(m,v)\succ(m',v')$.

A preference relation $\succeq$ is said to be *monotone* if a decision maker prefers more profitable alternatives to less:

$(m,0)\succeq(m',0)$ whenever $m\geq m'$.

A preference relation $\succeq$ is called *representable* if there exists a utility function $U:S_L\to R$ (defined up to an order-preserving transformation) such that for any $(m,v),(m',v')\in S_L$

$(m,v)\succeq(m',v')$ if and only if $U(m,v)\geq U(m',v')$.

A utility function U of a representable preference relation is *idempotent* if $U(m,0)=m$ for all $m\in X$.

A preference relation $\succeq$ is said to be *meaningful with respect to a group of admissible transformations* $T(X)$ (for short, $T(X)$-*meaningful*) if for any $X,Y\in\Im$ and each $T\in T(X)$

$(M[X],V[X])\succeq(M[Y],V[Y])\;\Rightarrow\;(M[T(X)],V[T(X)])\succeq(M[T(Y)],V[T(Y)])$.

The given definition asserts preservation of the preference structure under transformations $T\in T(X)$ and is equivalent to 1-meaningfulness of the statement "$(M[X],V[X])\succeq(M[Y],V[Y])$".

Using the results of sections 4 and 5, we will characterize meaningful preference relations on the set S_L under the following additional assumption: for all $X,Y\in\Im$ and $T\in T(X)$

$$M_L[X]\geq M_L[Y]\;\Rightarrow\;M_L[T(X)]\geq M_L[T(Y)];\qquad V_L[X]\geq V_L[Y]\;\Rightarrow\;V_L[T(X)]\geq V_L[T(Y)].\tag{46}$$

This assumption can be interpreted as comparison meaningfulness of M_L and V_L separately. For ordered scales (46) can be considered as a generalization of (5) (see (6) and Proposition 1). However, under some regularity assumptions (46) is reduced to the previously considered target setting.

Lemma 3.

Let the L-dispersion on $\Im$ and the related L-mean satisfy (46) for a group of continuously differentiable admissible transformations $T(X)$. If the penalty function L is symmetric and continuously differentiable, for each $X\in\Im$ there exists a unique solution y_0 of the equation

$$\sum_{i=1}^{n}p_iL_2'(x_i,y_0)=0,\tag{47}$$

and there exist continuous non-zero mixed derivatives $L_{12}''(x,y)$, $L_{21}''(x,y)$, $x\neq y$; then (6) holds and for each $T\in T(X)$ there exists a constant $q_T>0$ such that (7) holds.

Proof.

By assumptions, the partial derivative L_2' is continuous and for each binary random variable X with the probability distribution $\{x_1, p; x_2, 1-p\}$ the equation

$$pL_2'(x_1, y_0) + (1-p)L_2'(x_2, y_0) = 0$$

has a unique solution y_0. Thus, $\operatorname{sgn} L_2'(x, y) = \operatorname{sgn}(y - x)$ and for each fixed $x_1 < x_2$ the function $y \mapsto L_2'(x_2, y)/L_2'(x_1, y)$ is strictly increasing with respect to $y \in (x_1, x_2)$. Hence, L_2' is a *quasideviation* on X and $M_L[X] = y_0$, where y_0 is a unique solution of (47), is a *quasideviation mean* of a random variable $X \in \Im$ generated by L_2' (see Páles, 1988, definition 1, definition 2).

By (46), it follows that

$$M_L[X] = M_L[Y],\ X, Y \in \Im \ \Rightarrow\ M_L[T(X)] = M_L[T(Y)],\ T \in \mathrm{T}(\mathrm{X}). \tag{48}$$

As it is well known (e.g., see Ovchinnikov, 1996, proposition 4.1), a Cauchy mean M_L satisfies (48) if and only if (6) holds. Combining (47) and (6), we obtain that the equality

$$\sum_{i=1}^{n} p_i L_2'(T(x_i), T(y_0))T'(y_0) = 0,\ T \in \mathrm{T}(\mathrm{X}) \tag{49}$$

holds at the same y_0.

By assumptions, $T' > 0$. Therefore, the function $(x, y) \mapsto L_2'(T(x), T(y))T'(y)$ is a quasideviation, too; (47) and (49) are the same quasideviation means. The equality problem of quasideviation means is solved by Páles (1988, theorem 8): for a given $T \in \mathrm{T}(\mathrm{X})$ there exists a continuous function $q_T : \mathrm{X} \to \mathrm{R}_+$, depending on T, such that

$$L_2'(T(x), T(y))T'(y) = q_T(y)L_2'(x, y) \quad \forall x, y \in \mathrm{X}. \tag{50}$$

Differentiating (50) with respect to x, we get

$$q_T(y) = \frac{L_{21}''(T(x), T(y))T'(y)T'(x)}{L_{21}''(x, y)},\ x \neq y. \tag{51}$$

L_{21}'' is a symmetric function:

$$L_{21}''(x, y) = L_{12}''(y, x) = L_{21}''(y, x),\ x \neq y,$$

where the first equality follows from symmetry of L and the second one holds due to continuity of L_{12}'' and L_{21}''.

Hence, the right-hand side of (51) is symmetric, so does the left-hand side. This is possible only if q_T is equal identically to a constant. Integrating (50) with respect to y and taking into account (i), we obtain (7). ■

The next technical result states that a meaningful preference relation has an idempotent utility function.

Lemma 4.

Suppose that the preference relation $\succeq$ satisfies the following conditions:

(a) $\succeq$ is nondegenerate;

(b) $\succeq$ is representable;

(c) $\succeq$ is monotone;

(d) $\succeq$ is meaningful with respect to a group of admissible transformations $\mathrm{T}(\mathrm{X})$ such that $\mathrm{T}_{\{0\}}^{(f)}(\mathrm{X}) \subseteq \mathrm{T}(\mathrm{X})$ for some increasing bijection f of X onto R;

(e) for any $(m,v) \in \mathrm{S}_L$ there exist $x, y \in \mathrm{X}$ such that $(x,0) \succeq (m,v) \succeq (y,0)$.

Then there exists an idempotent utility function representing $\succeq$.

Proof.

By (b), there exists a utility function U representing $\succeq$. By assumption (c), the function u defined by

$u(x) = U(x,0)$, $x \in \mathrm{X}$

is non-decreasing.

We are going to prove that for each $X \in \mathfrak{I}$ there exists a unique $x \in \mathrm{X}$ such that

$$U(M_L[X], V_L[X]) = u(x). \tag{52}$$

Indeed, if (52) holds, then the utility function $u^{-1} \circ U$ is well defined and idempotent.

First, we prove that u is strictly increasing, i.e. for each $X \in \mathfrak{I}$ there exists at the most one x such that (52) holds. Since $\succeq$ is nondegenerate (a), there exist $y < z$ such that $u(y) < u(z)$. Let $x_0, x_1 \in \mathrm{X}$ with $x_0 < x_1$. Since the group of admissible transformations $\mathrm{T}_{\{0\}}^{(f)}(\mathrm{X})$ is 1-point homogeneous, there exists $T \in \mathrm{T}_{\{0\}}^{(f)}(\mathrm{X})$ such that $T(x_0) = x_1$. Obviously $T(x) > x$ for all $x \in \mathrm{X}$. Define a sequence $(..., x_{-1}, x_0, x_1, ...)$ inductively as follows

$x_{k+1} = T(x_k)$, $x_{-k-1} = T^{-1}(x_{-k})$, $k = 0,1,...$

By Luce et al. (1990, lemma 5, p. 130), it follows that there exist integers $k_1 < k_2$ such that $x_{k_1} \le y < z \le x_{k_2}$. Thus, $u(x_{k_1}) < u(x_{k_2})$. Combining this with (d), we get $u(x_0) < u(x_1)$. Hence, u is strictly increasing.

To prove that for each $X \in \mathfrak{I}$ there exists $x \in \mathrm{X}$ such that (52) holds assume the converse. Suppose that there exists $X \in \mathfrak{I}$ such that

$U(M_L[X], V_L[X]) \neq u(x)$ for all $x \in \mathrm{X}$.

By assumptions (c) and (e), there exists $x^* \in \mathrm{X}$ such that

$$u(x^* - \varepsilon) < U(M_L[X], V_L[X]) < u(x^* + \varepsilon) \text{ for any } \varepsilon > 0. \quad (53)$$

Obviously, u has a jump discontinuity at the point x^* (if it is not, then tending ε to zero, we obtain a contradiction: $U(M_L[X], V_L[X]) = u(x^*)$).

Let $x \in \mathrm{X}$. Since $\mathrm{T}_{\{0\}}^{(f)}(\mathrm{X})$ is 1-point homogeneous, there exists $T \in \mathrm{T}(\mathrm{X})$ such that $T(x^*) = x$. Using (d) and (53), we have

$$u(T(x^* - \varepsilon)) < U(M_L[T(X)], V_L[T(X)]) < u(T(x^* + \varepsilon)) \text{ for any } \varepsilon > 0.$$

Arguing as above, we see that $u \circ T$ is discontinuous at the point x^*. Since T is continuous, then u is discontinuous at the point $T(x^*) = x$. By the arbitrariness of the choice of x, u is discontinuous at each point of the interval X. But a monotone function has at most a countable set of discontinuity points. This contradiction proves (52). ■

The next proposition is a variant of a well-known result in the theory of invariant means (e.g., see Ovchinnikov, 1996); it deduces the functional equation for a utility function of a meaningful preference relation.

Proposition 5.

Under the conditions of Lemma 3 and Lemma 4, there exist an idempotent utility function U and a subgroup $\{q_T : T \in \mathrm{T}(\mathrm{X})\}$ of positive reals under multiplication such that (7) holds and

$$U(T(m), q_T v) = T \circ U(m, v), \; (m, v) \in \mathrm{S}_L, \; T \in \mathrm{T}(\mathrm{X}). \quad (54)$$

Proof.

From Lemma 3, it follows that for each $T \in \mathrm{T}(\mathrm{X})$ there exists a constant $q_T > 0$ such that (7) holds.

Let U be an idempotent utility function representing $\succeq$ (Lemma 4). Let $X \in \Im$ and denote $x = U(M_L[X], V_L[X])$, then

$$U(T(M_L[X]), q_T V_L[X]) = U(M_L[T(X)], V_L[T(X)]) = U(T(x), 0) = T(x) = T \circ U(M_L[X], V_L[X]),$$

where the first equality follows from Lemma 3, the second one holds by meaningfulness of $\succeq$, and the third one follows form idempotence of U. ■

Proposition 5 and the results of section 5 make possible to characterize penalty functions and utility functions of representable preference relations, that are meaningful in particular scales. As an example, consider the case $\mathrm{T}(\mathrm{X}) = \mathrm{T}^{(f)}_{\mathrm{R}_+}(\mathrm{X})$.

Proposition 6.

Under the conditions of Lemma 3 and Lemma 4 with $\mathrm{T}(\mathrm{X}) = \mathrm{T}^{(f)}_{\mathrm{R}_+}(\mathrm{X})$, where the function f is continuously differentiable and $f' \neq 0$, there exist constants $c > 1$, $d > 0$, and r such that

$$L(x, y) = d\left| f(x) - f(y)\right|^c \tag{55}$$

and one of utility functions has the form

$$U(m,v) = v^{1/c} r + f(m),\ (m,v) \in \mathrm{S}_L = \mathrm{X} \times \overline{\mathrm{R}}_+ . \tag{56}$$

Proof.

(55) follows from Lemma 3 and the third part of Proposition 2. Since $f(\mathrm{X}) = \mathrm{R}$, we have $\mathrm{S}_L = \mathrm{X} \times \overline{\mathrm{R}}_+$.

(54) goes over into the form

$$U\left(f^{-1}(af(m) + b), a^c v\right) = f^{-1}\left(af \circ U(m,v) + b\right),\ m \in \mathrm{X},\ v \in \overline{\mathrm{R}}_+,\ a \in \mathrm{R}_+,\ b \in \mathrm{R} . \tag{57}$$

Denote $G(m,v) = f \circ U(m,v)$. If $v > 0$, then substituting a for $v^{-1/c}$ and b for $-f(m)v^{-1/c}$ in (57), we have

$$G(m,v) = \left(G\left(f^{-1}(af(m) + b), a^c v\right) - b\right)/a = v^{1/c} G\left(f^{-1}(0),1\right) + f(m) . \tag{58}$$

If $v = 0$ then $G(m,0) = f(m)$. Hence, (58) holds for all $(m,v) \in \mathrm{X} \times \overline{\mathrm{R}}_+$. Denote $r = G\left(f^{-1}(0),1\right)$, then (58) can be written in the form (56). ■

The special case of (56) with affine f (i.e. $\mathrm{T}(\mathrm{X})$ defines an interval scale) and $c = 2$ is a widely known utility function – a linear combination of the mathematical expectation and the standard deviation of a corresponding random outcome X.

We also note that for a large class of parametric families of random variables the value $U\left(M_L[X], V_L[X]\right)$ with the utility function of the form (56) and the penalty function (55) is equal to the quantile of X (in financial literature the quantile is often referred to as Value at Risk (VaR)). Indeed, since the utility function is defined up to an order-preserving transformation, (56) can be also written in the form

$$U(m,v) = f^{-1}\left(v^{1/c} r + f(m)\right),\ (m,v) \in \mathrm{X} \times \overline{\mathrm{R}}_+ .$$

Given a group of admissible transformations $\mathrm{T}_{\mathrm{R}_+}^{(f)}(\mathrm{X})$, let $X_{1,0}$ be a continuous random variable, and let $F_{1,0}$ be the cumulative distribution function (c.d.f.) of $X_{1,0}$. Suppose that the support of $F_{1,0}$ is the convex hull of X, $V_L[X_{1,0}]=1$, and $f \circ M_L[X_{1,0}]=0$, where L is defined by (55). By definition, put

$$X_{a,b} = T(X_{1,0}) = f^{-1}(af(X_{1,0})+b),\ T \in \mathrm{T}_{\mathrm{R}_+}^{(f)}(\mathrm{X}).$$

The c.d.f. of $X_{a,b}$ is given by

$$F_{a,b}(x) = F_{1,0} \circ f^{-1}((f(x)-b)/a),$$

$$V_L^{1/c}[X_{a,b}] = a \text{, and } f \circ M_L[X_{a,b}] = b.$$

By $\Im_f$ denote the set $\Im_f = \{X_{a,b},\ a \in \mathrm{R}_+,\ b \in \mathrm{R}\}$. Then

$$U(M_L[X_{a,b}], V_L[X_{a,b}]) = f^{-1}(V_L^{1/c}[X_{a,b}]r + f \circ M_L[X_{a,b}]) = f^{-1}(ar+b) =$$
$$= f^{-1}(af \circ F_{1,0}^{-1}(\alpha)+b) = F_{a,b}^{-1}(\alpha),\ X_{a,b} \in \Im_f,$$

where $\alpha = F_{1,0} \circ f^{-1}(r) \in (0,1)$. Hence, $U(M_L[X], V_L[X])$ reduces to VaR of order α whenever $X \in \Im_f$.

In particular, in the case of the family of normally distributed random variables, $f(x)=x$, and $c=2$, (56) is reduced to VaR of order $\Phi(r)$, where Φ is the c.d.f. of the standard normal distribution.

7. Conclusion and open problems

The technical results established in this paper may serve as an additional argument to use the considered dispersion measures and the related measures of location for random variables, measured on nominal, ordinal, interval, or difference scales.

We conclude by noting the two possible generalizations of the cited construction. For some applications the "only if" part of assumption (i) in the definition of a penalty function may be found too restrictive. For example, in difference scales one may tolerate the existence of an interval of length l regarding to that x and y are indistinguishable: $L(x,y)=0$ if and only if $|x-y| \le l$. Another example is a pseudometric on X. The mentioned variation of the term "penalty function" does not allow to apply arguments used in Lemma 2 and requires a little bit more elaborated technique.

Intuitive definitions of the terms "dispersion measure" and "measure of location" are given by Dershem (1975); the idea of the approach is related to the theory of best approximation (e.g., see

Singer, 1974). Let $(\Im, \rho)$ be a metric space. The distance $\rho[X; y_{\text{a.s.}}]$ is called the *dispersion of the random variable* $X \in \Im$ *about the point* $y \in \mathrm{X}$. The distance

$$V[X] = \inf_{y \in \mathrm{X}} \rho[X; y_{\text{a.s.}}]$$

is the *dispersion measure of the random variable* X. Elements of the set

$$M[X] = \{y \in \mathrm{X} : \rho[X; y_{\text{a.s.}}] = V[X]\}$$

are called *means of the random variable* X (to ensure the set $M[X]$ contains exactly one element, we may restrict ourselves with strictly convex reflexive Banach spaces (e.g., see Singer, 1974, p. 36)). It would be attractive to obtain analogous (to those, obtained in sections 4 and 5) characterization results under these definitions of "dispersion measure" and "mean".

Acknowledgement

The author wishes to thank Lebedev Anton and an anonymous referee for their valuable suggestions, comments, and corrections.

References

Aczél, J. (1966). *Lectures on functional equations and their applications*. New York: Academic Press.

Aczél, J., & Daróczy, Z. (1963). Über verallgemeinerte quasilineare Mittelwerte, die mit Gewichtsfunktionen gebildet sind. *Publicationes Mathematicae Debrecen*, 10, 171–190.

Aczél, J., & Roberts, F.S. (1989). On the possible merging functions. *Mathematical Social Sciences*, 17, 205–243.

Bickel, P.J., & Lehmann, E.L. (1975). Descriptive statistics for nonparametric models. II. Location. *The Annals of Statistics*, 3(5), 1045–1069.

Calvo, T., Mesiar, R., & Yager, R.R. (2004). Quantitative weights and aggregation. *IEEE Transactions on Fuzzy Systems*, 12(1), 62–69.

Dershem, H.L. (1975). A relationship between approximation theory and statistical measurements. *Pi Mu Epsilon Journal*, 6(2), 69–74.

Fishburn, P.C. (1986). Implicit mean value and certainty equivalence. *Econometrica*, 54(5), 1197–1206.

Fodor, J.C., & Roubens, M. (1995). On meaningfulness of means. *Journal of Computational and Applied Mathematics*, 64, 103–115.

Huber, P.J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35, 73–101.

Luce, R.D. (1980). Several possible measures of risk. *Theory and Decision*, 12(3), 217–228. Correction: Luce, R.D. (1981). Correction to 'Several possible measures of risk'. *Theory and Decision*, 13(4), 381.

Luce, R.D., Krantz, D.H., Suppes, P., & Tversky, A. (1990). *Foundations of measurement* (Vol. III). San Diego: Academic Press.

Marichal, J.-L. (1998). *Aggregation operations for multicriteria decision aid*. University of Liège: Ph.D. Thesis.

Marichal, J.-L., Mesiar, R., & Rückschlossová, T. (2005). A complete description of comparison meaningful functions. *Aequationes Mathematicae*, 69(3), 309–320.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.

Mesiar, R. (2007). Fuzzy set approach to the utility, preference relations, and aggregation operators. *European Journal of Operational Research*, 176(1), 414–422.

Narens, L. (2002). *Theories of meaningfulness*. London: Lawrence Erlbaum Associates.

Orlov, A.I. (1979). *Stability of Socio-Economical Models*. Moscow: Nauka (in Russian).

Ostasiewicz, S., & Ostasiewicz, W. (2000). Means and their applications. *Annals of Operations Research*, 97, 337–355.

Ovchinnikov, S. (1996). Means on ordered sets. *Mathematical Social Sciences*, 32(1), 39–56.

Ovchinnikov, S., & Dukhovny, A. (2002). On order invariant aggregation functionals. *Journal of Mathematical Psychology*, 46(1), 12–18.

Páles, Zs. (1988). General inequalities for quasideviation means. *Aequationes Mathematicae*, 36(1), 32–56.

Roberts, F.S. (1979). *Measurement theory, with applications to decisionmaking, utility, and the social sciences*. London: Addison-Wesley.

Singer, I. (1974). *The theory of best approximation and functional analysis*. Philadelphia: SIAM.

Stone, B.K. (1973). A general class of three-parameter risk measures. *The Journal of Finance*, 28(3), 675–685.

Tobin, T. (1958). Liquidity preferences as a behavior toward risk. *Review of Economic Studies*, 25, 65−85.