Various representations and algebraic structure of linear imprecision indices

Alexander E. Lepskiy
Taganrog State University
RUSSIA
lepskiy@mail.ru

Andrey G. Bronevich
Taganrog State University
RUSSIA
brone@mail.ru

Abstract

The paper is devoted to the investigation of imprecision indices, introduced in [7]. They are used for evaluation of uncertainty (or more exactly imprecision), which is contained in information given by fuzzy (non-additive) measures, in particular, by lower or upper probabilities. We argue that there exist various types of uncertainty, for example, randomness, investigated in probability theory, imprecision, described by interval calculi, inconsistency, incompleteness, fuzziness and so on. In general these types of uncertainty have very complex behavior, caused by their interaction. Therefore, the choice of uncertainty measures is not unique, and is defined by the problems addressed. The classical uncertainty measures are Shannon's entropy and Hartley's measure. In the paper imprecision indices and their linear representatives are introduced axiomatically. The system of axioms enables to define various imprecision indices. So we investigate the algebraic structure of all imprecision indices and describe their families with best properties.

Keywords: Imprecision indices lower and upper probabilities, uncertainty based information.

1 Introduction

Measuring of uncertainty plays a major role in uncertainty theories, in particular, probability theory, information theory, fuzzy sets theory and so on. There are some ways of defining such measures in the theory of evidence, in the theory of fuzzy (non-additive) measures and in the theory of imprecise probabilities. However, one can see that in such general theories an uncertainty measure with the best properties has no yet been found. This situation is explained by very complex interaction among various types of uncertainty, including randomness, inconsistence, imprecision, incompleteness of analyzed information. We recall classical uncertainty measures, used in information theory and probability theory. Let X be a finite set of alternatives. Assigning to each alternative $x \in X$ some probability $P(\{x\})$, we have information, which is described by probability measure P, and in this case Shannon's entropy $S(P) = -\sum_{x\in X} P(\{x\})\log_2 P(\{x\})$ can be used. Let we know only that the "true" alternative is in a non-empty set $B \subseteq X$. This situation can be described by the set function $\eta_{\langle B\rangle}(A) = \begin{cases} 1, B \subseteq A \\ 0, B \not\subseteq A \end{cases}$, $A \subseteq X$, which gives lower probability of an event A, and Hartley's measure $H(\eta_{\langle B\rangle}) = \log_2 |B|$ can be justified. It is easily seen that in the first case uncertainty has a type that one call randomness, and the second case is more connected with imprecision of the information. The generalization of these two cases consists in the following. Consider a pair $(\underline{g}, \overline{g})$ of set functions $\underline{g} : 2^X \to [0,1]$, $\overline{g} : 2^X \to [0,1]$ defined on the powerset 2^X. We suggest that $\underline{g}(A) \le \overline{g}(A)$ for all $A \in 2^X$, $\underline{g}(\varnothing) = \overline{g}(\varnothing) = 0$, and there is a "true" probability measure P on 2^X with $\underline{g}(A) \le P(A) \le \overline{g}(A)$ for all $A \in 2^X$. In other words, set functions $\underline{g}, \overline{g}$ give us upper and lower bounds of probabilities, and for any event $A \in 2^X$ we have only the interval $\left[\underline{g}(A), \overline{g}(A)\right]$ of possible values of a "true" probability $P(A)$. In practical issues it is sufficient to define the lower probability $\underline{g}$, the upper probability can be calculated by $\overline{g}(A)=1-\underline{g}\left(\overline{A}\right)$, where $A\in 2^X$ and $\overline{A}$ is the complement of A. Due to works of Klir, Higashi, Harmanec and others (see [4,5,6]), there are two important uncertainty measures, which show the best properties in a sense of obeying axioms, which are similar to the axioms of Shannon's entropy. They are generalized Hartley's measure, and aggregate measure of uncertainty. Let $\underline{g}$ be a belief function, i.e. it can be represented by

$$\underline{g}=\sum_{B\in 2^X} m(B)\eta_{\langle B\rangle},$$

where $m(\varnothing)=0$, $m(B)\geq 0$ for all $B\in 2^X$, and $\sum_{B\in 2^X} m(B)=1$. Then generalized Hartley's measure is defined by

$$GH\left(\underline{g}\right)=\sum_{B\in 2^X\backslash\{\varnothing\}} m(B)\log_2|B|.$$

The aggregate measure of uncertainty is calculated by

$$Au\left(\underline{g}\right)=\sup_{P\geq \underline{g}} S(P),$$

where sup is taken over all probability measures on 2^X, which are consistent with $\underline{g}$, i.e. $P(A)\geq \underline{g}(A)$ for all $A\in 2^X$. It is worth to mention that generalized Hartley's measure can be used for measuring imprecision and aggregate measure of uncertainty for total uncertainty. It is easy to check that aggregate measure of uncertainty coincides with Shannon's entropy for probability measures and with Hartley's measure for $\underline{g}=\eta_{\langle B\rangle}$, $B\neq\varnothing$. These uncertainty measures can be also used in the case, where $\underline{g}$ is a 2-monotone set function [2]. There is a possibility to extend our consideration to the case, where the set of probability measures, which are consistent with $\underline{g}$, is empty. Then we say that the information in our disposal is inconsistent and we should analyze three types of uncertainty: randomness, imprecision, and inconsistence.

The paper has the following structure. First we remind some definitions and results from the theory of non-additive measures and axiomatics of imprecision indices, formulated in [7]. Then we analyze so called linear imprecision indices on the set of upper and lower probabilities, giving their detailed description, and introducing their important families with symmetrical properties.

2 Basic definitions and problem statement

Let X be a finite set. In the sequel we will use the following notations:

1. $M(X)$ is the set of all real-valued set functions on the powerset 2^X;

2. $M_0(X)=\{g\in M(X)\,|\,g(\varnothing)=0\}$;

3. We write $g_1\leq g_2$ for $g_1,g_2\in M(X)$ if $g_1(A)\leq g_2(A)$ for all $A\in 2^X$.

4. $M_{mon}(X)\subset M_0(X)$ is the set of all normalized monotone set functions on 2^X. It means that $g\in M_0(X)$ implies $g(\varnothing)=0$, $g(X)=1$, and $g(A)\leq g(B)$ if $A\subseteq B$.

5. $M_{\Pr}(X)$ is the set of all probability measures on 2^X;

6. $M_{low}(X)=\{g\in M_0\,|\,\exists P\in M_{\Pr}: g\leq P\}$ is the set of all lower probabilities on 2^X.

6. $M_{up}(X)=\{g\in M_0\,|\,\exists P\in M_{\Pr}: g\geq P\}$ is the set of all upper probabilities on 2^X.

7. Let $g\in M(X)$ then the dual of g is denoted by $\overline{g}$ and by definition $\overline{g}(A)=g(X)-g\left(\overline{A}\right)$, $A\in 2^X$.

8. $M_{bel}(X)$ is the set of all belief functions on 2^X. Any $g\in M_{bel}(X)$ has the following unique representation: $g=\sum_{B\in 2^X} m(B)\eta_{\langle B\rangle}$, where $m(B)\geq 0$ for all $B\in 2^X$, $m(\varnothing)=0$, and $\sum_{B\in 2^X} m(B)=1$.

9. $M_{pl}(X)$ is the set of all plausibility functions on 2^X. Any $g\in M_{pl}(X)$ is uniquely represented by $g=\sum_{B\in 2^X} m(B)\overline{\eta}_{\langle B\rangle}$, where $m(B)\geq 0$ for all $B\in 2^X$, $m(\varnothing)=0$, and $\sum_{B\in 2^X} m(B)=1$.

We can consider the set $M(X)$ (or $M_0(X)$) as a linear space w.r.t. to usual sum of set functions and usual product of set functions and real numbers. In the non-additive measure theory, the basis consisting of functions $\eta_{\langle B\rangle}$, $B\in 2^X$, is of interest. Let $g\in M(X)$ and $g=\sum_{B\in 2^X} m_g(B)\eta_{\langle B\rangle}$ then the set function m_g is called the Möbius transform of g.

The function m_g is expressed by $m_g(B)=\sum_{A:A\subseteq B}(-1)^{|B\setminus A|}g(A)$. We will also use so-called dual Möbius transform of g. This transform is connected with the basis consisting of set functions $\eta^{\langle B\rangle}$, $B\in 2^X$, defined by $\eta^{\langle B\rangle}(A)=\eta_{\langle\overline{B}\rangle}(\overline{A})$. Let $g=\sum_{B\in 2^X}m^g(B)\eta^{\langle B\rangle}$ then the set function m^g is called the dual Möbius transform of g. It is calculated by $m^g(B)=\sum_{A:B\subseteq A}(-1)^{|A\setminus B|}g(A)$.

We remind now some definitions, introduced in [7].

Definition 1. *A functional* $f:M_{low}(X)\to[0,1]$ *is called imprecision index if the following conditions are fulfilled: 1)* $g\in M_{\Pr}(X)$ *implies* $f(g)=0$*; 2)* $f(g_1)\geq f(g_2)$ *for all* $g_1,g_2\in M_{low}(X)$ *such that* $g_1\leq g_2$*; 3)* $f\left(\eta_{\langle X\rangle}\right)=1$. Some important examples of imprecision indices are:

1) $v_B(g)=\overline{g}(B)-g(B)$ for a fixed $B\in 2^X/\{\varnothing,X\}$ (We call it a primitive imprecision index in the sequel.);

2) $v_p(g)=\left(2^{|X|}-2\right)^{-1/p}\left(\sum_{B\in 2^X}\left|\overline{g}(B)-g(B)\right|^p\right)^{1/p}$, $p\geq 1$;

3) $v_\infty(g)=\sup\left\{\left|\overline{g}(B)-g(B)\right| \mid B\in 2^X\right\}$.

It is clear that there are many ways for defining imprecision indices. One class of them consisting of linear imprecision indices is described in the following definition.

Definition 2. An imprecision index f on $M_{low}(X)$ is called *linear* if for any linear combination $\sum_{j=1}^k\alpha_j g_j\in M_{low}(X)$, $\alpha_j\in\mathbb{R}$, $g_j\in M_{low}(X)$, $j=1,...,k$, we have $f\left(\sum_{j=1}^k\alpha_j g_j\right)=\sum_{j=1}^k\alpha_j f\left(g_j\right)$.

3. An investigation of linear imprecision indices

We notice first that any linear functional f on $M(X)$ is defined uniquely by its values on a chosen basis of $M(X)$. It enables to define f by a set function $\mu_f:2^X\to\mathbb{R}$ with the following property $\mu_f(B)=f\left(\eta_{\langle B\rangle}\right)$, $B\in 2^X$. Since any $g\in M_{low}$ is represented as a linear combination of $\left\{\eta_{\langle B\rangle}\right\}_{B\in 2^X\setminus\{\varnothing\}}$, we take by definition that $\mu_f(\varnothing)=0$ (or $f\left(\eta_{\langle\varnothing\rangle}\right)=0$) for any linear imprecision index f.

Proposition 1. *Let* f *be a linear imprecision index on* M_{low} *then* $\mu_f\in M_{mon}(X)$ *with* $\mu_f\left(\{x\}\right)=0$ *for any* $x\in X$.

Proof. By definition $\mu_f(\varnothing)=0$ and $\mu_f(X)=1$; $\mu_f\left(\{x\}\right)=0$ for any $x\in X$ because $\eta_{\langle\{x\}\rangle}\in M_{\Pr}(X)$. Further we see that $\eta_{\langle B\rangle}\in M_{low}$ for $B\neq\varnothing$ and $\eta_{\langle B\rangle}\geq\eta_{\langle C\rangle}$ if $B\subseteq C$. It implies that $\mu_f(B)\leq\mu_f(C)$ for $B\subseteq C$, i.e. μ_f is monotone. This finishes the proof of the proposition.

The following proposition gives the expression of any linear functional through the values of the transformed set function.

Proposition 2. *Let* f *be a linear functional on* $M(X)$ *then* $f(g)=\sum_{B\in 2^X}m^{\mu_f}(B)g(B)$ *for any* $g\in M(X)$.

Proof. By definition $\mu_f=\sum_{B\in 2^X}m^{\mu_f}(B)\eta^{\langle B\rangle}$ and $g=\sum_{C\in 2^X}m_g(C)\eta_{\langle C\rangle}$, therefore,

$$f(g)=\sum_{C\in 2^X}m_g(C)\mu_f(C)$$

$$=\sum_{C\in 2^X}m_g(C)\sum_{B\in 2^X}m^{\mu_f}(B)\eta^{\langle B\rangle}(C)$$

$$=\sum_{C\in 2^X}m_g(C)\sum_{B\in 2^X}m^{\mu_f}(B)\eta_{\langle C\rangle}(B)$$

$$=\sum_{B\in 2^X}m^{\mu_f}(B)\sum_{C\in 2^X}m_g(C)\eta_{\langle C\rangle}(B)$$

$$=\sum_{B\in 2^X}m^{\mu_f}(B)g(B)$$

The following theorem gives necessary and sufficient conditions on a linear functional to be an imprecision index through the dual Möbius transform of μ_f.

Theorem 1. *Let* f *be a linear functional on* $M(X)$ *then it is an imprecision index on* $M_{low}(X)$ *iff*

a) $m^{\mu_f}(X)=1$*;* $\sum_{D\in 2^X}m^{\mu_f}(D)=0$*;*

b) $\sum_{D:\,x\in D}m^{\mu_f}(D)=0$ *for all* $x\in X$*;*

c) $m^{\mu_f}(D)\leq 0$ *for all* $D\in 2^X\setminus\{\varnothing,X\}$.

Proof. It is clear that the condition a) guarantees that $f\left(\eta_{\langle X\rangle}\right)=1$ and $f\left(\eta_{\langle\varnothing\rangle}\right)=0$. It is easy to show that b) is the necessary and sufficient condition that $f(g)=0$ for any $g\in M_{\Pr}(X)$. Indeed, since

$\eta_{\langle\{x\}\rangle} \in M_{\Pr}(X)$ then $\mu_f(\{x\}) = f\left(\eta_{\langle\{x\}\rangle}\right) = 0$ and $\mu_f(\{x\}) = \sum_{D:\, x\in D} m^{\mu_f}(D) = 0$. On the other hand, any $g \in M_{\Pr}(X)$ can be represented as a convex combination of $\eta_{\langle\{x\}\rangle}$, i.e. $g = \sum_{x\in X} m_g(\{x\})\eta_{\langle\{x\}\rangle}$, therefore,

$$f(g) = \sum_{x\in X} m_g(\{x\}) f\left(\eta_{\langle\{x\}\rangle}\right)$$

$$= \sum_{x\in X} m_g(\{x\})\mu_f(\{x\}) = 0.$$

So b) is proved. c) is the sufficient and necessary condition of antimonotonicity of f on $M_{low}(X)$. Let c) be fulfilled and $g_1 \le g_2$ for $g_1, g_2 \in M_{low}(X)$ then by Proposition 2

$$f(g_1) - f(g_2) = \sum_{B\in 2^X} m^{\mu_f}(B)\left(g_1(B) - g_2(B)\right)$$

$$= \sum_{B\in 2^X\setminus\{\varnothing,X\}} m^{\mu_f}(B)\left(g_1(B) - g_2(B)\right).$$

Since $g_1(B) - g_2(B) \le 0$ for any $B \in 2^X$ and $m^{\mu_f}(B) \le 0$ for any $B \in 2^X \setminus \{\varnothing, X\}$, we get $f(g_1) \ge f(g_2)$, i.e. c) implies antimonotonicity of f. Vice versa, let f be antimonotone on $M_{low}(X)$ then for any $D \in 2^X \setminus \{\varnothing, X\}$ we can always find such $g_1, g_2 \in M_{low}(X)$ with $g_1(B) = g_2(B)$ for all $B \ne D$, and $g_1(D) < g_2(D)$. According to Proposition 2 $0 \le f(g_1) - f(g_2) = m^{\mu_f}(D)(g_1(D) - g_2(D))$, i.e. $m^{\mu_f}(D) \le 0$. The theorem is proved.

Conditions of Theorem 1 can be transformed to the expression, which is very close to the condition "avoiding sure loss" from the theory of imprecise probabilities [10]. It enables to get the implicit expression for an arbitrary linear imprecision index. We will further use the functions 1_B, $B \subseteq X$, on X defined by $1_B(x) = 1$ if $x \in B$, and $1_B(x) = 0$ otherwise.

Theorem 2. *Any linear imprecision index f on M_{low} can be uniquely represented by*

$$f(g) = 1 - \sum_{B\in 2^X} q(B)g(B),$$

where the set function q obeys the following conditions:

1) $q(\varnothing) = 0$, $q(X) = 0$, $q(B) \ge 0$ for all $B \in 2^X$;

2) $\sum_{B\in 2^X} q(B)1_B = 1_X$.

Proof. By Proposition 2 $f(g) = \sum_{B\in 2^X} m^{\mu_f}(B)g(B)$, since $m^{\mu_f}(X) = 1$, $g(X) = 1$, $g(\varnothing) = 0$, we get the required representation, choosing $q(B) = -m^{\mu_f}(B)$ for $B \in 2^X \setminus \{\varnothing, X\}$. The condition b) from Theorem 1 is reformulated for q as $\sum_{B:\, x\in B} q(B) = 1$ for all $x \in X$. We show that this condition is equivalent to 2). Actually, $\sum_{B\in 2^X} q(B)1_B = 1_X$ iff $\sum_{B\in 2^X} q(B)1_B(x) = 1$ for any $x \in X$, on the other hand, $\sum_{B\in 2^X} q(B)1_B(x) = \sum_{B:\, x\in B} q(B)$. The theorem is proved.

Remark 1. The condition of "avoiding sure loss" from the theory of imprecise probabilities can be formulated with the help of the set function q from the Theorem 2 as follows: let $g \in M_0(X)$ then $g \in M_{low}(X)$ iff for any set function q obeying 1), 2) from Theorem 2, we have $\sum_{B\in 2^X} q(B)g(B) \le 1$.

Theorem 3. *Let f be a linear functional on $M(X)$ then it is an imprecision index on $M_{low}(X)$ iff $\mu_f = a\mu - b\overline{\eta}_{\langle X\rangle}$, where $b > 0$, $a = 1 + b$, and $\mu \in M_{pl}(X)$ with $\mu(\{x\}) = b/a$ for all $x \in X$.*

Proof. *Necessity.* Let f be a linear imprecision index on $M_{low}(X)$ then

$$\mu_f(B) = \sum_{A\in 2^X} m^{\mu_f}(A)\eta_{\langle\overline{A}\rangle}(\overline{B})$$

$$= \sum_{A\in 2^X\setminus\{X,\varnothing\}} m^{\mu_f}(A)\eta_{\langle\overline{A}\rangle}(\overline{B})$$

$$+ m^{\mu_f}(X)\eta_{\langle\varnothing\rangle}(\overline{B}) + m^{\mu_f}(\varnothing)\eta_{\langle X\rangle}(\overline{B}),$$

where $m^{\mu_f}(A) \le 0$ for any $A \in 2^X \setminus \{X, \varnothing\}$ and $\eta_{\langle\varnothing\rangle} \equiv 1$, $m^{\mu_f}(X) = 1$. Let $a = -\sum_{A\in 2^X\setminus\{X,\varnothing\}} m^{\mu_f}(A)$ then taking $q(A) = -\frac{1}{a} m^{\mu_f}(\overline{A})$ for $A \in 2^X \setminus \{X, \varnothing\}$ and $q(A) = 0$ for $A \in \{X, \varnothing\}$, we get

$$\mu_f(B) = -a\sum_{A\in 2^X} q(\overline{A})\eta_{\langle\overline{A}\rangle}(\overline{B}) + 1 + m^{\mu_f}(\varnothing)\eta_{\langle X\rangle}(\overline{B})$$

$$= a\sum_{A\in 2^X} q(A)\left(1 - \eta_{\langle A\rangle}(\overline{B})\right)$$

$$- m^{\mu_f}(\varnothing)\left(1 - \eta_{\langle X\rangle}(\overline{B})\right) + m^{\mu_f}(\varnothing) + 1 - a.$$

It is clear $m^{\mu_f}(\varnothing) + 1 - a = \sum_{A\in 2^X} m^{\mu_f}(A) = \mu_f(\varnothing) = 0$, hence, we get the representation required

$$\mu_f(B) = a\sum_{A\in 2^X} q(A)\overline{\eta}_{\langle A\rangle}(B) - b\overline{\eta}_{\langle X\rangle}(B),$$

where $\mu = \sum_{A\in 2^X} q(A)\overline{\eta}_{\langle A\rangle}$, $b = m^{\mu_f}(\varnothing)$, $a = 1 + b$.

It is easy to show that $\mu(\{x\})=b/a$, $x\in X$, and $b>0$. Actually, by Proposition 1 $\mu_f(\{x\})=0$ for all $x\in X$, i.e. $\mu(\{x\})=b/a$ for all $x\in X$. On the other hand,

$$\mu_f(\{x\})=a\sum_{A:x\in A}q(A)-b=0,$$

i.e. $b\geq 0$ and if $b=0$ then $q\equiv 0$ and this contradicts to the definition of imprecision index.

Sufficiency. Assume that we have the representation of μ_f from the theorem. We prove sufficiency if we check all conditions from Theorem 1. We see that $\mu_f(\varnothing)=0$, $\mu_f(X)=1$, and $\mu_f(\{x\})=0$ for all $x\in X$, i.e. conditions a), b) are true. We will further prove that $m^{\mu_f}(A)\leq 0$ for all $A\in 2^X\setminus\{\varnothing,X\}$.

Since μ is a plausibility function, it is represented by $\mu=\sum_{A\in 2^X}m(A)\overline{\eta}_{\langle A\rangle}$, where $m(A)\geq 0$ for all $A\in 2^X$, $m(\varnothing)=0$, and $\sum_{A\in 2^X}m(A)=1$. We can write

$$\mu_f(B)=a\sum_{A\in 2^X}m(A)\overline{\eta}_{\langle A\rangle}(B)-b\overline{\eta}_{\langle X\rangle}(B)$$

$$=a\sum_{A\in 2^X}m(A)\left(1-\eta_{\langle A\rangle}(\overline{B})\right)-b\left(1-\eta_{\langle X\rangle}(\overline{B})\right)$$

$$=a\sum_{A\in 2^X}m(\overline{A})\left(1-\eta_{\langle \overline{A}\rangle}(\overline{B})\right)-b\left(1-\eta_{\langle X\rangle}(\overline{B})\right).$$

The last expression implies $m^{\mu_f}(A)=-am(\overline{A})\leq 0$ for all $A\in 2^X\setminus\{\varnothing,X\}$, i.e. c) is also true. The theorem is proved.

From the proof of Theorem 3, we see that we can use the basis $\left\{\overline{\eta}_{\langle B\rangle}\right\}_{B\in 2^X\setminus\{\varnothing\}}$ of $M_0(X)$ for defining other sufficient and necessary conditions on linear imprecision index. We formulate them in

Corollary 1. *Let f be a linear functional on $M(X)$ and $\mu_f=\sum_{A\in 2^X\setminus\{\varnothing\}}m(A)\overline{\eta}_{\langle A\rangle}$ then f is an imprecision index iff 1) $\mu_f\in M_0(X)$; 2) $\mu_f(\{x\})=0$ for all $x\in X$; 3) $m(A)\geq 0$ for all $A\in 2^X\setminus\{\varnothing,X\}$.*

Theorem 4. *Let f be a linear functional on $M(X)$ then it is an imprecision index on $M_{low}(X)$ iff 1) $\mu_f\in M_0(X)$; 2) $\mu_f(\{x\})=0$ for all $x\in X$; 3) the set function $\mu_f^{\{x\}}$, defined by $\mu_f^{\{x\}}(B)=\mu_f(B\cup\{x\})$, $B\in 2^X$, is in $M_{pl}(X)$ for any $x\in X$.*

Proof. *Necessity.* Let f be a linear imprecision index on $M_0(X)$. It is sufficient to show that 3) is true. By Theorem 3 we can use the following representation

$$\mu_f=a\sum_{A\in 2^X}m(A)\overline{\eta}_{\langle A\rangle}(B)-b\overline{\eta}_{\langle X\rangle},$$

where $b>0$, $a=1+b$, $m(A)\geq 0$ for all $A\in 2^X$, $m(\varnothing)=0$, $m(X)=0$, and $\sum_{A\in 2^X}m(A)=1$. Then

$$\mu_f(B\cup\{x\})=a\sum_{A:x\in A}m(A)-b$$

$$+a\sum_{A:x\notin A}m(A)\overline{\eta}_{\langle A\rangle}(B\cup\{x\}).$$

Since $a\sum_{A:x\in A}m(A)-b=\mu_f(\{x\})=0$, $\overline{\eta}_{\langle A\cup\{x\}\rangle}(B)=\overline{\eta}_{\langle A\rangle}(B\cup\{x\})$, we get $\mu_f^{\{x\}}=a\sum_{A:x\notin A}m(A)\overline{\eta}_{\langle A\cup\{x\}\rangle}$, i.e. $\mu_f^{\{x\}}$ is a plausibility function for any $x\in X$.

Sufficiency. Assume that the all conditions of the theorem are fulfilled, however, f is not linear imprecision index on $M_{low}(X)$. In this case at least one of inequalities $m(A)\geq 0$, $A\in 2^X\setminus\{\varnothing,X\}$, from Corollary 1 is not true, and there is a $B\in 2^X\setminus\{\varnothing,X\}$ such that $m(B)<0$. Let $x\in X\setminus B$. In the similar way as in the proof of Theorem 3 we get

$$\mu_f^{\{x\}}=\sum_{A:x\notin A}m(A)\overline{\eta}_{\langle A\cup\{x\}\rangle},$$

Since in the last sum $m(A)<0$ for $A=B$, we conclude that $\mu_f^{\{x\}}$ is not a plausibility function. This implies that our assumption is wrong, and f is a linear imprecision index. The theorem is proved in the whole.

It seems to be logical in some problems that the quantity of imprecision in the situation, where we only know that the true alternative belongs to the set B, depends on $|B|$ and does not depend on other factors. In this case we assume that $f(\eta_{\langle B\rangle})=f(\eta_{\langle C\rangle})$ or $\mu_f(B)=\mu_f(C)$ if $|B|=|C|$, and we call such linear imprecision indices symmetrical. In the sequel we will use the fact that such symmetrical monotone set functions can be viewed as distorted probabilities [9]. Let P be a probability measure on $X=\{x_1,...,x_N\}$; let $\lambda:[0,1]\to[0,1]$ be non-decreasing function with $\lambda(0)=0$, $\lambda(1)=1$, then the set function $g=\lambda\circ P$ ($g(A)=\lambda(P(A))$, $A\in 2^X$) is called distorted probability. We are interested in the case, where $P(\{x_i\})=1/N$, $i=1,...,N$. Further we will use the following sufficient condition of total monotonicity [1]: let $g=\lambda\circ P$ then it is a belief function if λ is infinitely differentiable on $[0,1)$ and $d^n\lambda(t)/dt^n\geq 0$, $n=1,2,...$, for any $t\in[0,1)$.

Theorem 5. *Let f be a linear functional on $M(X)$ and $\mu_f=\lambda\circ P$, i.e. μ_f is a distorted probability,* *mentioned above, and* $P(\{x_i\}) = 1/N$, $i = 1,...,N$. *Then* f *is an imprecision index if*

1) $\lambda(1/N) = 0$;

2) λ *is infinitely differentiable on* $\left[\frac{1}{N},1\right)$ *and* $(-1)^{n-1} d^n\lambda(t)/dt^n \geq 0$, $n = 1,2,...,$ *for any* $t \in \left[\frac{1}{N},1\right)$.

Proof. We will check that the all conditions from Theorem 4 are true. It is clear that $\mu_f \in M_0(X)$ and $\mu_f(\{x\}) = 0$ for all $x \in X$. Now we prove that 3) is also true. In this case $\mu_f^{\{x\}}(B) = \lambda(P(B \cup \{x\}))$, $B \in 2^{X\setminus\{x\}}$, $\mu_f^{\{x\}}$ can be considered as a distorted probability on $2^{X\setminus\{x\}}$, and $\mu_f^{\{x\}} = \lambda_1 \circ P_1$, where $\lambda_1(t) = \lambda\left(\frac{1+t(N-1)}{N}\right)$, $t \in [0,1]$, and $P_1(\{y\}) = 1/(N-1)$, $y \in X \setminus \{x\}$. We find that $\overline{\mu_f^{\{x\}}}(A) = 1 - \lambda_1(P_1(\overline{A})) = 1 - \lambda_1(1 - P_1(A))$, i.e. $\overline{\mu_f^{\{x\}}} = \lambda_2 \circ P_1$ is a distorted probability and $\lambda_2(t) = 1 - \lambda_1(1-t) = 1 - \lambda\left(1 - \frac{t(N-1)}{N}\right)$. It is clear $\mu_f^{\{x\}} \in M_{pl}(X)$ iff $\overline{\mu_f^{\{x\}}} \in M_{bel}(X)$. Then we can argue that $\mu_f^{\{x\}}$ is a plausibility function if $d^n\lambda_2(t)/dt^n \geq 0$, $n = 1,2,...,$ for any $t \in [0,1)$, or $(-1)^{n-1} d^n\lambda(t)/dt^n \geq 0$, $n = 1,2,...,$ for any $t \in \left[\frac{1}{N},1\right)$.

In some cases it is suitable to define symmetrical μ_f by a non-decreasing function $\varphi : [1,+\infty) \to [0,+\infty)$ with $\varphi(1) = 0$ assuming that $\mu_f(A) = \varphi(|A|)/\varphi(|X|)$ for $A \neq \varnothing$. Then $\lambda(t) = \varphi(tN)/\varphi(N)$ for $t \in \left[\frac{1}{N},1\right]$, where $N = |X|$. It is easy to see that according to Theorem 5, μ_f determines a linear imprecision index if φ is is infinitely differentiable on $[1,N)$ and $(-1)^{n-1} d^n\varphi(t)/dt^n \geq 0$, $n = 1,2,...,$ for any $t \in [1,N)$.

Example 1. Let $\varphi(t) = \ln(t)$ then $\mu_f(A) = \ln(|A|)/\ln(|X|)$. In this case the corresponding linear imprecision index can be considered as the analog of generalized Hartley's measure. We see that $(-1)^{n-1} d^n \ln(t)/dt^n = t^{-n} \geq 0$ for $t \geq 1$, i.e. μ_f determines a linear imprecision index on $M_{low}(X)$.

4. The algebraic structure of the set of all linear imprecision indices

Let f_1, f_2 be linear functionals on $M(X)$ then their linear combination $f = af_1 + bf_2, a,b \in \mathbb{R}$ is also a linear functional. If we take into consideration set functions $\mu_{f_1}, \mu_{f_2}, \mu_f$, we see that $\mu_f = a\mu_{f_1} + b\mu_{f_2}$, i.e. the set of all linear functionals on $M(X)$ is a linear space and this space is isomorphic to the linear space $M(X)$ of all set functions on 2^X. It is easy to show that if f_1, f_2 are linear imprecision indices then their convex sum $f = af_1 + bf_2$, where $a,b \geq 0$, $a+b=1$, is also linear imprecision index, i.e. the set of all linear imprecision indices is a convex set. We denote by $M_I(X)$ the set of all set functions μ_f, which correspond to linear imprecision indices on $M_{low}(X)$. One can say that we understand the algebraic structure of a convex set if we have description of its extreme points. The following theorem gives the necessary and sufficient condition on an arbitrary $\mu \in M_I$ to be an extreme point.

Theorem 6. *Let* $\mu \in M_I(X)$, $\mu = \sum_{A\in\mathfrak{B}} m(A)\overline{\eta}_{\langle A\rangle} - b\overline{\eta}_{\langle X\rangle}$, *where* $\mathfrak{B} \subseteq 2^X \setminus \{\varnothing, X\}$, $m(A) > 0$ *for all* $A \in \mathfrak{B}$, $b > 0$, *then* μ *is an extreme point of* $M_I(X)$ *iff functions* $\{1_A\}_{A\in\mathfrak{B}}$ *are linearly independent.*

Proof. Notice first that any $\mu \in M_I(X)$ has the representation $\mu = \sum_{A\in\mathfrak{B}} m(A)\overline{\eta}_{\langle A\rangle} - b\overline{\eta}_{\langle X\rangle}$ by Corollary 1, $b > 0$, and $\mathfrak{B}$ is not empty. Secondly, $\mu(\{x\}) = 0$ for all $x \in X$, i.e.

$$\sum_{A\in\mathfrak{B}} m(A)1_A = b1_X .$$

We will show that μ is not an extreme point of $M_I(X)$ iff functions $\{1_A\}_{A\in\mathfrak{B}}$ are linearly dependent. This implies evidently the theorem statement. Assume that functions $\{1_A\}_{A\in\mathfrak{B}}$ are linearly dependent. Then there exist two different solutions of

$\sum_{A\in\mathfrak{B}} \alpha_A 1_A = 1_X$ w.r.t. α_A, $A \in \mathfrak{B}$. We choose one of them as $\alpha_A^{(1)} = m(A)/b$, $A \in \mathfrak{B}$. Since $\alpha_A^{(1)} > 0$ for all $A \in \mathfrak{B}$, we can choose another solution $\alpha_A^{(2)}$ with $\alpha_A^{(2)} \geq 0$, $A \in \mathfrak{B}$. Let $b_2 = 1/\left(\left(\sum_{A\in\mathfrak{B}} \alpha_A^{(2)}\right) - 1\right)$, then it is easy to see that $b_2 > 0$ and the set function μ_2, defined by

$$\mu_2 = \sum_{A\in\mathfrak{B}} b_2\alpha_A^{(2)}\overline{\eta}_{\langle A\rangle} - b_2\overline{\eta}_{\langle X\rangle},$$

is in M_I. Defining

$$c = \sup\left\{ r \in \mathbb{R} \mid rb_2\alpha_A^{(2)} \le m(A), A \in \mathfrak{B}, rb_2 \le b \right\},$$

we confirm that $c \in (0,1)$, $\mu \ge c\mu_2$. Then

$$\mu_2 = \frac{1}{1-c}\left(\mu - c\mu_1\right) = \sum\nolimits_{A\in\mathfrak{B}} m_1(A)\overline{\eta}_{\langle A\rangle} - b_1\overline{\eta}_{\langle X\rangle}.$$

where $m_1(A) = \frac{1}{1-c}\left(m(A) - cb_2\alpha_A^{(2)}\right)$, $b_1 = \frac{1}{1-c}\left(b - cb_2\right)$, is in $M_I(X)$. We see that $\mu \ge (1-c)\mu_1 + c\mu_2$, i.e. we have proved that μ is not an extreme point of $M_I(X)$.

Vice versa assume that μ is not an extreme point of $M_I(X)$. Then there exist set functions $\mu_1, \mu_2 \in M_I(X)$ such that $\mu = a\mu_1 + b\mu_2$, where $a, b > 0$ and $a + b = 1$. Since $\mu_1, \mu_2 \in M_I(X)$ we have $\sum\nolimits_{A\in\mathfrak{B}} m_i(A)1_A = b_i 1_X$, where $b_i > 0$, $i = 1,2$. Therefore, the equation $\sum\nolimits_{A\in\mathfrak{B}} \alpha_A 1_A = 1_X$ has more than one solution w.r.t. $\alpha_A \in \mathbb{R}$, $A \in \mathfrak{B}$, hence, functions $\{1_A\}_{A\in\mathfrak{B}}$ are linearly dependent if μ is not an extreme point of M_I. The theorem is proved.

Theorem 6 implies that the set $M_I(X)$ has the finite number of extreme points. According to the Theorem by Krein-Milman [8], any $\mu \in M_I(X)$ can be represented as a convex sum of extreme points. However, it is a very hard problem to describe such extreme points explicitly. Further we consider one convex subset of $M_I(X)$, for which this problem can be solved.

Definition 3. Let f be a linear imprecision index on $M_{low}(X)$, then we call it *complementarily symmetrical* if $m^{\mu_f}(A) = m^{\mu_f}(\overline{A})$ for all $A \in 2^X \setminus \{\varnothing, X\}$.

Important examples of complementarily symmetrical linear imprecision indices are primitive imprecision indices. We see that

$$\nu_B(g) = g(X) - g(B) - g(\overline{B}) + g(\varnothing),$$

$$\begin{aligned}\mu_{\nu_B}(A) &= \eta_{\langle A\rangle}(X) - \eta_{\langle A\rangle}(B) - \eta_{\langle A\rangle}(\overline{B}) + \eta_{\langle A\rangle}(\varnothing)\\ &= \eta^{\langle\varnothing\rangle}(A) - \eta^{\langle\overline{B}\rangle}(A) - \eta^{\langle B\rangle}(A) + \eta^{\langle X\rangle}(A).\end{aligned}$$

Therefore, $m^{\mu_{\nu_B}}(A) = 1$ if $A \in \{\varnothing, X\}$, $m^{\mu_{\nu_B}}(A) = -1$ if $A \in \{B, \overline{B}\}$, and $m^{\mu_{\nu_B}}(A) = 0$ otherwise. We can also express μ_{ν_B} through plausibility functions. In this case

$$\begin{aligned}\mu_{\nu_B}(A) &= 1 - \eta_{\langle\overline{B}\rangle}(\overline{A}) - \eta_{\langle B\rangle}(\overline{A}) + \eta_{\langle X\rangle}(\overline{A})\\ &= \overline{\eta}_{\langle B\rangle}(A) + \overline{\eta}_{\langle\overline{B}\rangle}(A) - \overline{\eta}_{\langle X\rangle}(A).\end{aligned}$$

By Theorem 6 it is easy to show that primitive indices ν_B, $B \in 2^X \setminus \{\varnothing, X\}$, are extreme points of $M_I(X)$. Actually, it follows from the fact that functions $\{1_B, 1_{\overline{B}}\}$ are linearly independent.

The role of primitive indices for describing the set of all complementarily symmetrical linear indices shows the following theorem.

Theorem 7. *The set of all complementarily symmetrical linear indices is convex. Any complementarily symmetrical linear index can be uniquely represented by a convex sum of primitive indices.*

Proof. Let f_1, f_2 be complementarily symmetrical imprecision indices, then

$$f_i(g) = \sum\nolimits_{B\in 2^X} m^{f_i}(B)g(B),\ i = 1,2,\ g \in M_{low}(X),$$

and by Definition 3 $m^{\mu_{f_i}}(B) = m^{\mu_{f_i}}(\overline{B})$, $i = 1,2$, for all $B \in 2^X \setminus \{\varnothing, X\}$. Let $f = af_1 + cf_2$, where $a, c \ge 0$, and $a + c = 1$. Then it is easy to see that $m^{\mu_f} = am^{\mu_{f_1}} + cm^{\mu_{f_2}}$. This implies $m^{\mu_f}(B) = m^{\mu_f}(\overline{B})$ for all $B \in 2^X \setminus \{\varnothing, X\}$, i.e. the set of all complementarily symmetrical linear indices is convex.

Now we will prove that any complementarily symmetrical linear index can be represented by a convex sum of primitive indices. Let f be a complementarily symmetrical linear index and $g \in M_{low}(X)$ then

$$f(g) = \sum\nolimits_{B\in 2^X} m^{\mu_f}(B)g(B),$$

where $m^{\mu_f}(B) = m^{\mu_f}(\overline{B})$ for all $B \in 2^X \setminus \{\varnothing, X\}$. Let $\mathfrak{D} = \left\{B \in 2^X \setminus \{X\} \mid x \in B\right\}$, $\overline{\mathfrak{D}} = \left\{B \in 2^X \mid \overline{B} \in \mathfrak{D}\right\}$ for some $x \in X$ then $\mathfrak{D} \cup \overline{\mathfrak{D}} = 2^X \setminus \{\varnothing, X\}$, $\mathfrak{D} \cap \overline{\mathfrak{D}} = \varnothing$.

$$\begin{aligned}f(g) &= m^{\mu_f}(X)g(X) + m^{\mu_f}(\varnothing)g(\varnothing)\\ &\quad + \sum\nolimits_{B\in\mathfrak{D}} m^{\mu_f}(B)\left(g(B) + g(\overline{B})\right)\\ &= -\sum\nolimits_{B\in\mathfrak{D}} m^{\mu_f}(B)\left(g(X) - g(B) - g(\overline{B}) + g(\varnothing)\right)\\ &\quad + \sum\nolimits_{B\in\mathfrak{D}} m^{\mu_f}(B)\left(g(X) + g(\varnothing)\right)\\ &\quad + m^{\mu_f}(X)g(X) + m^{\mu_f}(\varnothing)g(\varnothing).\end{aligned}$$

We see $\sum\nolimits_{B\in\mathfrak{D}} m^{\mu_f}(B) = \sum\nolimits_{B:x\in B} m^{\mu_f}(B) - m^{\mu_f}(X) = -m^{\mu_f}(X) = -1$. The equality $\sum\nolimits_{B\in 2^X} m^{\mu_f}(B) = 0$ implies that $m^{\mu_f}(\varnothing) = -\sum\nolimits_{B\in\mathfrak{D}}\left(m^{\mu_f}(B) + m^{\mu_f}(\overline{B})\right) -$

$m^{\mu_f}(X)=1$. Hence,

$$f(g)=\sum_{B\in\mathfrak{D}}(-1)m^{\mu_f}(B)\nu_B ,$$

where $(-1)m^{\mu_f}(B)\geq 0$ for all $B\in\mathfrak{D}$, and $\sum_{B\in\mathfrak{D}}(-1)m^{\mu_f}(B)=1$, i.e. f can be represented by a convex sum of primitive indices.

We prove that the found representation is unique if we show that system $\{\nu_B\}_{B\in\mathfrak{D}}$ of all primitive indices is linearly independent in $M(X)$, or we show the same property for set functions $\left\{\mu_{\nu_B}\right\}_{B\in\mathfrak{D}}$. It is easy to see that set functions $\mu_{\nu_B}=\overline{\eta}_{\langle B\rangle}+\overline{\eta}_{\langle\overline{B}\rangle}-\overline{\eta}_{\langle X\rangle}$, $B\in\mathfrak{D}$, are linearly independent, this follows immediately from the fact that set functions $\left\{\overline{\eta}_{\langle B\rangle}\right\}_{B\in 2^X\backslash\{\varnothing\}}$ are also linear independent in $M(X)$. The theorem is proved in the whole.

Example 2. Let $\xi: X\to\mathbb{R}$, $\max_{x\in X}\xi(x)-\min_{x\in X}\xi(x)=1$. Then we can define the linear imprecision index with the help of Choquet integral [3] $f(g)=\int_X \xi d\overline{g}-\int_X \xi dg$, where $g\in M_{low}$. Then $\mu_f(B)=\max_{x\in B}\xi(x)-\min_{x\in B}\xi(x)$ for $B\neq\varnothing$. It is easy to show that such defined index f is complementarily symmetrical. It is worth to mention that in the theory of imprecise probabilities $\int_X \xi d\overline{g}$ can be viewed as an upper estimation of the expectation $E[\xi]$, and $\int_X \xi dg$ as a lower estimation of the expectation $E[\xi]$.

Conclusion

Although, measuring uncertainty plays a central role in various uncertainty theories, there is no possibility to find one true uncertainty measure. This can be explained by the fact that there are many various types of uncertainty, they have different interpretations; it is very difficult to understand their mutual interaction. One way for overcoming this problem is to find families of suitable uncertainty measures, satisfying some justified properties. The choice of the best uncertainty measure depends considerably on the problem solved. In this paper we have proposed how imprecision can be measured if uncertain information is described by lower probabilities. We have treated the case, where uncertainty consists of some randomness and imprecision. The introduced axiomatics enables us to give detailed description of linear imprecision indices, and investigate some of them with symmetrical properties.

Acknowledgments

The authors express their sincere thanks to the anonymous reviewers, whose excellent work allows to increase the paper's quality. Dr. Bronevich on his behalf expresses his gratitude to the Fulbright Program, Binghamton University, and Prof. George Klir for research opportunity provided.

References

[1] A. G. Bronevich, On the closure of families of fuzzy measures under eventwise aggregations, Fuzzy sets and systems 153 (2005) 45-70.

[2] A. Chateauneuf, J. Y. Jaffray, Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion, Mathematical Social Sciences 17 (1989) 263-283.

[3] D. Denneberg, Non-additive measure and integral, basic concepts and their role for applications, in M. Grabisch, T. Murofushi, M. Sugeno (Eds.), *Fuzzy Measures and Integrals - Theory and Applications.* Physical-Verlag, Heidelberg (Studies in Fuzzyness and Soft Computing, Vol. 40), 2000.

[4] D. Harmanec, G. J. Klir, Measuring total uncertainty in Dempster-Shafer theory: A novel approach, International Journal of General Systems 22(4) (1994) 405-419.

[5] M. Higashi, G. J. Klir, Measures of uncertainty and information based on possibility distributions, International Journal of General Systems 9(1) (1983) 43-58.

[6] G. J. Klir, Uncertainty and information: foundations of generalized information theory, John Wiley & Sons, Inc., 2006.

[7] A. E. Lepskiy, The linear imprecision indices on the lower probabilities, Proc. of the 11th IPMU Conference (Paris, July 2-7, 2006) 1724-1731.

[8] R. R. Phelps, Lectures on Choquet's theorem, Springer-Verlag, Heidelberg, 2000.

[9] D. Schmeidler, Subjective probability and expected utility without additivity. Econometrica 57 (1989).

[10] P. Walley, Statistical reasoning with imprecise probabilities, Chapman & Hall, London, 1991.