Tatra Mt. Math. Publ. **16** (1999), 295–310

TATRA MOUNTAINS
Mathematical Publications

APPLICATIONS OF NON-ADDITIVE SET FUNCTIONS IN DECISION MAKING AND INSURANCE

STANISŁAW HEILPERN

ABSTRACT. In this paper we study the relationships between different non-additive set functions and Choquet integrals with respect to such set functions. We investigate Bayesian inference problems and the statistical robust issues connected with them. We also apply the Choquet integral in simple decision making problems to calculate expected payoffs and use it for the construction of an insurance premium.

1. Introduction

This paper is devoted to monotone set functions and their application in decision making and insurance. First we recall some notions, types and properties of monotone set functions which we shall use in the next part of the paper. Next we study the relationships between distorted probabilities, decomposable measures and plausibility functions. We also give the definition and some properties of Choquet integrals with respect to monotone set functions.

In the second part of this paper we present applications of these theoretical investigations in decision making and insurance. We study Bayesian inference problems and the statistical robust issues connected with them. We also apply the Choquet integral in simple decision making problems to calculate expected payoffs and use it for the construction of an insurance premium.

2. Set functions

Let S be a complete, separable and metrizable space and let $\mathcal{J}$ be a σ-algebra of its subsets. In the case of a finite space S we assume that $\mathcal{J} = 2^S$. In

2000 Mathematics Subject Classification: 28E10, 90A05, 62P05.
Keywords: nonadditive set function, Choquet integral, decision making, insurance premium.

the infinite case we usually use the Borel σ-field. We shall study normalized *monotone set functions* $\mu\colon \mathcal{J} \to [0,1]$, i.e., set functions satisfying the following conditions:

$$\mu(\emptyset) = 0\,, \quad \mu(S) = 1\,,$$
$$A \subset B \Rightarrow \mu(A) \le \mu(B) \qquad \text{for all} \quad A, B \in \mathcal{J}\,.$$

We recall that the set function μ is

a) *additive* if for all $A, B \in \mathcal{J}$

$$A \cap B = \emptyset \Rightarrow \mu(A \cup B) = \mu(A) + \mu(B)\,,$$

b) k*-monotone* if for all $A_1, \dots, A_k \in \mathcal{J}$,

$$\mu\Big(\bigcup_{i=1}^{k} A_i\Big) \ge \sum_{\substack{I \subset \{1,\dots,k\} \\ I \ne \emptyset}} (-1)^{|I|+1} \mu\Big(\bigcap_{i \in I} A_i\Big)\,,$$

c) k*-alternating* if for all $A_1, \dots, A_k \in \mathcal{J}$

$$\mu\Big(\bigcap_{i=1}^{k} A_i\Big) \le \sum_{\substack{I \subset \{1,\dots,k\} \\ I \ne \emptyset}} (-1)^{|I|+1} \mu\Big(\bigcup_{i \in I} A_i\Big)\,,$$

d) *a capacity* if for an arbitrary nondecreasing sequence of subsets $A_n \in \mathcal{J}$, such that $A = \bigcup_n A_n$

$$\lim_{n\to\infty} \mu(A_n) = \mu(A)\,,$$

and for an arbitrary nonincreasing sequence of closed subsets F_n, such that $F = \bigcap_n F_n$,

$$\lim_{n\to\infty} \mu(F_n) = \mu(F)\,.$$

A 2-monotone set function is also called convex or submodular and a 2-alternating set function is called *concave* or supermodular.

In this paper we shall use the following types of monotone set functions:

i) *belief functions*, which are k-monotone for all $k \ge 2$ (or totally monotone),

ii) *plausibility functions*, which are k-alternating for all $k \ge 2$ (or totally alternating),

iii) *necessity measures*, which satisfy the condition

$$\mu(A \cap B) = \min\big(\mu(A), \mu(B)\big)\,,$$

iv) *possibility measure*, which satisfy the condition

$$\mu(A \cup B) = \max\big(\mu(A)\,,\,\mu(B)\big)\,,$$

v) *decomposable functions*, if there exists an operator $\perp\colon [0,1] \times [0,1] \to [0,1]$, such that:

$$A \cap B = \emptyset \ \Rightarrow\ \mu(A \cup B) = \mu(A) \perp \mu(B)\,.$$

Belief and plausibility functions are strictly connected with random sets and evidence theory [4, 16]; necessity and possibility measures are connected with fuzzy sets.

In many applications the operator $\perp$ used in the definition of a decomposable set function is a *t-conorm*, i.e., it is a non-decreasing, commutative and associative operator satisfying the condition $a \perp 0 = 0$ for all $a \in (0,1)$ [7]. A continuous *Archimedian* t-conorm, meaning that $a \perp a > a$ for all $a \in (0,1)$, can be characterized by an additive generator.

Theorem 1. [15]. *The function $\perp\colon [0,1] \times [0,1] \to [0,1]$ is a continuous Archimedian t-conorm iff there exists an additive generator, i.e., a strictly increasing, nonnegative and continuous function $g\colon [0,1] \to R$, $g(0) = 0$ such that: $a \perp b = g^{(-1)}\big(g(a) + g(b)\big)$, where the pseudo-inverse*

$$g^{(-1)}(x) = \begin{cases} g^{-1}(x)\,, & 0 \le x < g(1)\,,\\ 1\,, & x \ge g(1)\,. \end{cases}$$

A t-conorm is called *strong* if it is strictly increasing in $(0,1) \times (0,1)$. A continuous Archimedian t-conorm, which is not strong is called *nilpotent*. For a strong t-conorm $\perp$ we have $g(1) = \infty$ and for a nilpotent one $g(1) = M < \infty$.

Let $S = \{s_1, \dots, s_n\}$ be finite a space, $A \subset S$, $J_A = \{i\colon s_i \in A\}$ and $\mu\big(\{s_i\}\big) = a_i$. The t-conorm is associative, then

$$\mu(S) = a_1 + \dots + a_n = \mathop{\perp}_{i=1}^{n} a_i = g^{(-1)}\left(\sum_{i=1}^{n} g(a_i)\right) = 1\,,$$

$$\mu(A) = \mathop{\perp}_{i \in J_A} a_i = g^{(-1)}\left(\sum_{i \in J_A} g(a_i)\right)$$

for an arbitrary decomposable measure μ induced by the t-conorm $\perp$. When this t-conorm is strong then $g(1) = \infty$ and

$$g(1) = g\big(\mu(S)\big) = \sum_{i=1}^{n} g(a_i)\,.$$

In this case there exists s_i such that $g(a_i) = \infty$, so $a_i = 1$.

The monotone set function μ defined on the finite space S induces another set function, called the *basic probability assignment* or *Möbius inverse* $m\colon A \to [0,1]$, such that

$$m(\emptyset) = 0\,, \qquad \sum_{A\in\mathcal{J}} \mu(A) = 1\,, \qquad \mu(A) = \sum_{B\subset A} m(B)\,.$$

The basic probability assignment is not a monotone set function. The set function μ is a belief function iff the induced basic probability assignment is non-negative, i.e., $\mu(A) \geq 0$ for all $A \in \mathcal{J}$.

Dually for any set function μ defined on the algebra $\mathcal{J}$ and for all $A \in \mathcal{J}$ we may define, the *conjugate* set function:

$$\mu'(A) = 1 - \mu(A^c)\,,$$

where A^c is the complement of the subset A. Such a pair of set functions: μ and μ' will play an important part in our investigations.

If the set function μ is convex (concave) then its conjugate set function μ' is concave (convex) and we obtain the following important relation:

$$\mu \leq \mu' \qquad (\mu' \leq \mu)\,.$$

Let $\mathscr{M}$ be the family of all probability measures defined on an algebra $\mathcal{J}$. The convex set function μ induces a subfamily of probability measures:

$$\operatorname{core}(\mu) = \{P \in \mathscr{M} : \mu \leq P \leq \mu'\}\,.$$

This family, called the *core*, will be also labelled by the pair of conjugate set functions (μ, μ'). It is a nonempty set and it is a compact convex subset of the set of all probability measures. [12].

When the set function μ is a probability measure then the core is a one-element subset because additivity implies the equality $\mu = \mu'$. The conjugate set function of a belief function is a plausibility function and for a necessity function is a possibility measure.

LEMMA 1. *Let μ be a necessity measure and μ' its conjugate possibility measure. Then*

(i) *if $\mu'(A) < 1$ then $\mu(A) = 0$,*

(ii) *if $\mu(A) > 0$ then $\mu'(A) = 1$.*

Proof.

(i) The set function μ is a necessity measure, then

$$\min\{\mu(A)\,,\ \mu(A^c)\} = \mu(A \cap A^c) = 0\,.$$

If $\mu'(A) < 1$ then $\mu(A^c) > 0$ by the definition. The above facts imply that $\mu(A) = 0$.

(ii) Is proved similarly. □

On the other hand, the subfamily of probability measures $\mathcal{A} \in \mathscr{M}$ induces lower $\underline{P}$ and upper $\overline{P}$ probabilities by the formulas

$$\underline{P}(A) = \inf_{P \in \mathcal{A}} P(A), \qquad \overline{P}(A) = \sup_{P \in \mathcal{A}} P(A).$$

In modern utility theory and actuarial science authors often use another kind of set function, namely *distorted probability* [2, 5, 14, 18]. This is a set function induced by a probability measure P and a nondecreasing function $f\colon [0,1] \to [0,1]$, such that $f(0) = 0$, $f(1) = 1$ by the formula

$$\mu(A) = f\big(P(A)\big).$$

When the function f is concave (convex) we obtain a concave (convex) monotone set function μ.

The *cumulative* and *decumulative* distribution functions of a measurable function $X\colon S \to \mathbb{R}$ are defined by the formulas

$$F(x) = \mu\{X \le x\}, \qquad G(x) = \mu\{X > x\}.$$

For a probability measure we obtain $G(x) = 1 - F(x)$, but for a general, non-additive set function this equality is false.

Let μ be a convex, monotone set function and let μ' be its conjugate (concave) set function. Then the cumulative and decumulative distribution functions F' and G' of the set function μ' are equal:

$$F'(x) = 1 - G(x), \qquad G'(x) = 1 - F(x)$$

and satisfy the following relations

$$F(x) \le F'(x), \qquad G(x) \le G'(x).$$

For the distorted probability $\mu(A) = f\big(P(A)\big)$ we obtain

$$G(x) = f\big(G_P(x)\big), \qquad G'(x) = 1 - f\big(1 - G_P(x)\big),$$

where G_P is the decumulative distribution function of the random variable X with respect to the probability measure P.

In the case of a finite space $S = s_1, \dots, s_n$, if we want to determine a distorted probability, decomposable or possibility measure, then it suffices to define only the n coefficients $\mu\big(\{s_1\}\big), \dots, \mu\big(\{s_n\}\big)$. In the general case, for example the belief function, we need to define the $2^n - 2$ coefficients corresponding to all nonempty subsets of the space S, keeping in mind the constraints determined by the properties of this monotone set function, for instance, the monotonicity, nonnegativity, concavity or alternativity.

3. Relationships between monotone set functions

Let the set function $\mu = f(P)$ be a distorted probability induced by the concave function f and the probability measure P, that is, a concave distorted probability. In this case, there exists $c \leq 1$ such that f is a strictly increasing, continuous function on the interval $[0,c]$ and the constant $f(x) = 1$ for $x \geq c$. We denote the restriction of the function f to $[0,c]$ by the symbol f_c.

Now we recall some theorems which give us a better characterization of such monotone set functions. The first theorem was proved by Chateauneuf in [2] for the finite space S, but it is also true for the general case.

Theorem 2. *A concave, distorted probability $\mu = f(P)$ is a decomposable set function generated by a continuous t-conorm with the additive generator: $g(x) = f_c^{-1}(x)$.*

Proof. The function g has the following properties:

$$g\big(f(x)\big) = \begin{cases} x\,, & 0 \leq x < c\,, \\ c\,, & c \leq x\,, \end{cases} \qquad \text{and} \quad g^{(-1)}(x) = f(x)\,.$$

Let $A, B \in \mathcal{J}$ and $A \cap B = \emptyset$. We shall show that

$$\mu(A \cup B) = \mu(A) \perp \mu(B)\,,$$

where $\perp$ is a t-conorm induced by the additive generator g. For this purpose we shall study two cases.

(i) $P(A) \geq c$ or $P(B) \geq c$.

Let $P(A) \geq c$, then $g\big(f\big(P(A)\big)\big) = c$ and

$$\begin{aligned}\mu(A) \perp \mu(B) &= g^{(-1)}\Big(g\big(f\big(P(A)\big)\big) + g\big(f\big(P(B)\big)\big)\Big) \geq g^{(-1)}\Big(g\big(f\big(P(A)\big)\big)\Big) \\ &= g^{(-1)}(c) = 1\,,\end{aligned}$$

because $g^{(-1)}$ is a nondecreasing function. On the other hand we have

$$\mu(A \cup B) \geq \mu(A) = f\big(P(A)\big) = 1\,.$$

(ii) $P(A) < c$ and $P(B) < c$.

In this case we have

$$\begin{aligned}\mu(A) \perp \mu(B) &= g^{(-1)}\Big(g\big(f\big(P(A)\big)\big) + g\big(f\big(P(B)\big)\big)\Big) = g^{(-1)}\big(P(A) + P(B)\big) \\ &= g^{(-1)}\big(P(A \cup B)\big) = f\big(P(A \cup B)\big) = \mu(A \cup B)\,.\end{aligned}$$

□

It is easy to see that the above theorem is true for more general distorted probabilities. It is sufficient that the function f be continuous and strictly increasing on the interval $[0,c]$. The above operator $\perp$ is a nilpotent t-conorm satisfying the condition $g(1) \leq 1$.

THEOREM 3. *A continuous, Archimedean t-conorm generates a distorted probability on a finite space S.*

P r o o f . First we assume that $\perp$ is a nilpotent t-conorm. The set function

$$P(\{s_i\}) = p_i = \frac{g(a_i)}{M},$$

where $M = g(1) = \sum_{i=1}^{n} g(a_i)$ is a probability measure and the function

$$f(x) = g^{(-1)}(Mx),$$

where $0 \le x \le 1$, satisfies the conditions: $f(0) = 0$, $f(1) = g^{(-1)}(M) = 1$ and

$$f\big(P(A)\big) = g^{(-1)}\Big(M \sum_{i \in J_A} p_i\Big) = g^{(-1)}\Big(\sum_{i \in J_A} g(a_i)\Big) = \mu(A).$$

Thus we obtain that the decomposable measure μ is a distorted probability $f(P)$.

Now, let $\perp$ be a strong t-conorm. In this case, $S = \{s_1, \dots, s_k, s_{k+1}, \dots, s_n\}$, where $a_i = 1$ for $i \le k$ and $a_i < 1$ for $i > k$. Let $S_1 = \{s_1, \dots, s_k\}$, m is a number satisfying the condition: $\frac{1}{k+1} < m < \frac{1}{k}$ and

$$p_i = \begin{cases} m, & i \le k, \\ \frac{1-km}{M} g(a_i), & i > k, \end{cases}$$

where $M = \sum_{i=k+1}^{n} g(a_i) < \infty$. The additive set function $P(\{s_i\}) = p_i$ is a probability measure. Let

$$f(x) = \begin{cases} g^{-1}\Big(\frac{M}{1-km} x\Big), & x < m, \\ 1, & x \ge m. \end{cases}$$

Then $f(0) = 0$ and $f(1) = 1$. Let A be a subset of S, such that $A \cap S_1 \ne \emptyset$. Then $\mu(A) = 1$, $P(A) = \sum_{i \in J_A} p_i \ge m$ and $f\big(P(A)\big) = 1$.

If $A \cap S_1 = \emptyset$, then

$$P(A) = \sum_{i \in J_A} p_i = \frac{1-km}{M} \sum_{i \in J_A} g(a_i) \le 1 - km < m.$$

In this case too, the decomposable measure μ is a distorted probability $f(P)$ because

$$f\big(P(A)\big) = g^{-1}\Big(\sum_{i \in J_A} g(a_i)\Big) = \mu(A).$$

□

The above theorem was proved by C h a t e a u n e u f in [2], but we show the form of the distorted probability for a strong t-conorm. The next theorem links the distorted probabilities with belief and plausibility functions.

Theorem 4. [22] *If the function f has derivatives of all orders and $(-1)^n f^{(n)} \le 0$ (resp. $f^{(n)} \ge 0$) then the distorted probability $\mu = f(P)$ is a plausibility (resp. belief) function.*

Then we shall study the case of concave distorted probabilities induced by simple power functions. We shall use them later in application to insurance.

Theorem 5. [11]

(i) *Let $f(x) = x^q$, where $0 < q \le 1$ and let P be a probability measure. The distorted probability $\mu = f(P)$ is a plausibility function.*

(ii) *Let $f(x) = 1-(1-x)^q$, where $q > 1$. The distorted probability $\mu = f(P)$ is a plausibility function if q is a natural number.*

Proof. This is implied directly by Theorem 4. □

Example 1.

(i) [11] Let $S = \{s_1, s_2, s_3\}$. The set function $P(\{s_1\}) = P(\{s_2\}) = 0.4$, $P(\{s_3\}) = 0.2$ is a probability measure. Let $f(x) = 1-(1-x)^{1.5}$. Then we obtain the following distorted probability μ, conjugate set function μ' and the basic probability assignment m

A	p	μ	μ'	m
$\{s_1\}$	0.4	0.535	0.253	0.253
$\{s_2\}$	0.4	0.535	0.253	0.253
$\{s_3\}$	0.2	0.284	0.089	0.089
$\{s_1, s_2\}$	0.8	0.911	0.716	0.21
$\{s_1, s_3\}$	0.6	0.747	0.465	0.122
$\{s_2, s_3\}$	0.6	0.747	0.465	0.122
S	1	1	1	-0.05

The value of $\mu(S) = -0.05$ is negative then the set function μ is not a plausibility.

(ii) Let $f(x) = x$, then $g(x) = x$. Thus the probability measure is the decomposable measure induced by the Łukasiewicz t-conorm: $x \perp y = \min(x+y, 1)$. See also Example 3.

(iii) Let $S = \{s_1, \dots, s_7\}$, $S_1 = \{s_1, s_2, s_3\}$ and let the decomposable measure μ be induced by the product t-conorm: $x \perp y = x+y-xy$ with additive generator $g(x) = -\ln(1-x)$ and it takes the following values on the singletons: $a_1 = a_2 = a_3 = 1$, $a_4 = 0.7$, $a_5 = 0.6$, $a_6 = 0.4$, $a_7 = 0.1$, where $a_i = \mu(\{s_i\})$. In this case, $k = 3$, $M = 2.736$ and choose $m = 0.3$.

We obtain the following probability measure P: $p_1 = p_2 = p_3 = 0.3$, $p_4 = 0.044$, $p_5 = 0.033$, $p_6 = 0.019$, $p_7 = 0.004$. Choose the function

$$f(x) = \begin{cases} 1 - e^{-27.36x}, & x < 0.3\,, \\ 1, & x \geq 0.3\,. \end{cases}$$

For example $\mu(\{s_4, s_5\}) = 0.7 + 0.6 - 0.42 = 0.88$ and $f(P(\{s_4, s_5\})) = 1 - e^{-2.12} = 0.88$.

Now we shall study the relation between distorted probabilities and possibility measures.

Theorem 6. *Let f be a strictly increasing function and let P be a probability measure such that there exists $B \neq S$ and $0 < P(B) < 1$. Then the distorted probability $\mu(A) = f(P(A))$ is not a possibility measure or a necessity measure.*

Proof. If the function f is a strictly increasing function then $\mu(B) = f(P(B)) > 0$. From Lemma 1 we have that if μ is a necessity measure then $\mu'(B) = 1$. On the other hand we obtain

$$\mu'(B) = 1 - \mu(B^c) = 1 - f(1 - P(B))\,.$$

If $\mu'(B) = 1$ then $1 - P(B) = 0$ because the function f is strictly increasing. Then we obtain $P(B) = 1$, so we have a contradiction. The proof for the possibility measure is similar. $\square$

4. Choquet integral

Now we shall introduce and investigate the integral of a measurable, bounded function (random variable) $X\colon S \to \mathbb{R}$ with respect to a monotone set function μ. This integral is based on the notion of the decumulative distribution function.

Definition 1. [3] The *Choquet integral* of a non-negative function X with respect to the monotone set function μ is given by

$$\int X\,d\mu = \int_0^\infty G(x)\,dx\,,$$

where the integral on the right hand side is in the sense of Riemann.

For an arbitrary function $X = X^+ - X^-$ with positive and negative parts X^+ and X^- we define the integral in the following manner

$$\int X\,d\mu = \int X^+ d\mu - \int X^- d\mu'\,.$$

Definition 2. The functions X and Y are comonotonic if there exists function Z and weakly increasing functions g and h such that $X = g(Z)$ and $Y = h(Z)$.

Now we list the properties of the Choquet integral connected with additivity:

i) If the functions X and Y are comonotonic then

$$\int (X+Y)\,d\mu = \int X\,d\mu + \int Y\,d\mu\,.$$

ii) If the set function μ is concave then

$$\int (X+Y)\,d\mu \le \int X\,d\mu + \int Y\,d\mu \qquad \text{(subadditivity)}\,.$$

iii) If the set function μ is convex then

$$\int (X+Y)\,d\mu \ge \int X\,d\mu + \int Y\,dm \qquad \text{(superadditivity)}\,.$$

For a distorted probability $\mu = f(P)$ the integral is

$$\int X\,d\mu = \int_0^\infty f\big(G_p(x)\big)\,dx\,,$$

where G_P is the decumulative distribution function of the random variable X with respect to the probability measure P.

5. Decision making

Many authors used the non-additive set function in decision making problems [6, 8, 9, 10, 13, 14, 23]. In many cases they were generalizations of the Neumann-Morgerstern or Savage approaches and some authors used distorted probability for this purpose [13, 23]. We now study a simple decision problem. Let S be a space of states of nature, let A be a set of actions and let l be a payoff function

$$l\colon S \times A \to \mathbb{R}\,.$$

We introduce the following notation

$$l_a(s) = l(s,a)$$

for a fixed action $a \in A$, where $s \in S$ and we assume that the function l is bounded.

When we have incomplete, inaccurate or uncertain knowledge about of the space states of nature S we may say that we have a *distribution of uncertainty* on the space S [9, 10]. There is a mathematical model of such knowledge. It may be a probability measure P, a fuzzy set which induces the two set functions: necessity Nec and possibility Pos measures, a random set (belief structure) with belief Bel and plausibility Pl functions or a family of probability measures generated by *lower* $\underline{P}$ and *upper* $\overline{P}$ *probabilities*. We may treat the above distributions of uncertainty as the family of probability measures Π, which is the core induced by a pair of conjugate set functions: convex μ and concave μ', i.e.,

$$\Pi = (\mu, \mu').$$

We may determine the *interval of expected payoff* for an action $a \in A$ under the distribution of uncertainty Π in two ways. First we define it as

$$M_a(\Pi) = \left[\int l_a \, d\mu \, , \ \int l_a \, d\mu' \right].$$

The second way is based on lower $\underline{E}(l_a) = \inf\limits_{P\in\Pi} \int l_a \, dP$ and upper $\overline{E}(l_a) = \sup\limits_{P\in\Pi} \int l_a \, dP$ expected payoffs and the interval of expected payoffs is

$$E_a(\Pi) = \left[\underline{E}(l_a), \ \overline{E}(l_a)\right].$$

The first way is more conservative. We obtain that

$$E_a(\Pi) \subset M_a(\Pi).$$

We may give the conditions of equality of such intervals of expected values.

Theorem 7. [19] *If μ' is the concave capacity and the distribution of uncertainty Π is the core (μ, μ'), then*

$$E_a(\Pi) = M_a(\Pi).$$

Proof. Let $E_P(l_a) = \int l_a \, dP$ and $\overline{M}(l_a) = \int l_a \, d\mu'$. Generally we have $\overline{M}(l_a) \geq \overline{E}(l_a) = \sup\limits_{P\in\Pi} E_P(l_a)$. First we assume that l_a is semicontinuous. We know that for every upper semicontinuous function h there is a $P \in \Pi$ such that $G_{P,h} = G_{\mu',h}$, where $G_{\mu,h}$ is the decumulative function of the function h under the set function μ [12]. Then

$$\overline{M}(l_a) = \int\limits_0^\infty G_{\mu',l_a}(x) \, dx = \int\limits_0^\infty G_{P,l_a}(x) \, dx = E_P(l_a) \leq \overline{E}(l_a),$$

thus $\overline{M}(l_a) = \overline{E}(l_a)$.

For a measurable bounded function l_a and probability measure $P \in \Pi$ there exists a nondecreasing sequence of semicontinuous functions l_n which approximates the expectation $E_P(l_a)$ from below: $\lim\limits_{n\to\infty} E_P(l_n) = E_P(l_a)$. [1]. This fact implies that for a general bounded payoff function l_a we have equality $\overline{M}(l_a) = \overline{E}(l_a)$. □

6. Bayesian inference

In classical Bayesian inference we have the *prior distribution* P defined on the space of states of nature S and the *posterior distribution* P_x on S obtained after observing the sample x. The value of the posterior probability is

$$P_x(A) = \frac{\int L_{A,x}(s)\,dP(s)}{\int L_{A,x}(s)\,dP(s) + \int L_{A^c,x}(s)\,dP(s)},$$

where $A \in \mathcal{J}$, $L_{A,x}(s) = f(x|s)\mathbf{1}_A(s)$ and $f(x|s)$ is the *likelihood function.*

Now we assume that prior knowledge about a space of states of nature S is modelled not by a single probability measure P, but by a family of probability measures Π.

This *prior distribution of uncertainty*, observation x and likelihood function $f(x|s)$ induce the *posterior distribution of uncertainty*

$$\Pi_x = \{P_x \colon P \in \Pi\}$$

with lower and upper probabilities, $\underline{P}_x$ and $\overline{P}_x$.

The computing of the boundaries of the posterior distribution of uncertainty is the main aim of our investigation. In the general case we may give the following estimate.

Theorem 8. [19] *Let $\underline{P}$ be the lower and $\overline{P}$ be the upper probability. Then*

$$\underline{P}_x(A) \ge \frac{\underline{E}(L_{A,x})}{\underline{E}(L_{A,x}) + \overline{E}(L_{A^c,x})} \ge \frac{\int L_{A,x}\,d\underline{P}}{\int L_{A,x}\,d\underline{P} + \int L_{A^c,x}\,d\overline{P}},$$

$$\overline{P}_x(A) \le \frac{\overline{E}(L_{A,x})}{\overline{E}(L_{A,x}) + \underline{E}(L_{A^c,x})} \le \frac{\int L_{A,x}\,d\overline{P}}{\int L_{A,x}\,d\overline{P} + \int L_{A^c,x}\,d\underline{P}}.$$

But in the case when the prior distribution of uncertainty P satisfies certain regular conditions we obtain the equalities in the above formula.

Theorem 9. [19] *If the upper prior probability $\overline{P}$ is continuous and the concave set functions and the prior distribution of uncertainty $\Pi = (\underline{P}, \overline{P})$ then*

$$\underline{P}_x(A) = \frac{\underline{E}(L_{A,x})}{\underline{E}(L_{A,x}) + \overline{E}(L_{A^c,x})} = \frac{\int L_{A,x}\,d\underline{P}}{\int L_{A,x}\,d\underline{P} + \int L_{A^c,x}\,d\overline{P}},$$

$$\overline{P}_x(A) = \frac{\overline{E}(L_{A,x})}{\overline{E}(L_{A,x}) + \underline{E}(L_{A^c,x})} = \frac{\int L_{A,x}\,d\overline{P}}{\int L_{A,x}\,d\overline{P} + \int L_{A^c,x}\,d\underline{P}}.$$

We obtain a simpler form of lower and upper posterior probabilities when the prior distribution of uncertainty is generated by a random set, i.e., the multifunction $W \colon (\Omega, \mathcal{A}, \mathrm{Pr}) \to \mathcal{J}$.

Theorem 10. [20, 21] *If the prior distribution of uncertainty* Π *is the family of probability measures* $(\mathrm{Bel}, \mathrm{Pl})$, *where* Bel *and* Pol *are the belief and plausibility functions generated by some random set* W *then*

$$\underline{P}_x(A) = \frac{E(\underline{L}_{A,x})}{E(\underline{L}_{A,x}) + E(\overline{L}_{A^c,x})}\,,$$

$$\overline{P}_x(A) = \frac{E(\underline{L}_{A,x})}{E(\overline{L}_{A,x}) + E(\underline{L}_{A^c,x})}\,,$$

where $E(\underline{L}_{A,x}) = \int \underline{L}_{A,x}(\omega)\, d\,\mathrm{Pr}(\omega)$, $\underline{L}_{A,x}(\omega) = \inf\limits_{s\in W(\omega)} L_{A,x}(s)$ *and* $\overline{L}_{A,x}(\omega) = \sup\limits_{s\in W(\omega)} L_{A,x}(s)$.

In this formula we have an ordinary integral and we may compute the inf and sup over a subset of the space S, not over the subfamily of probability measures as in the general case.

We may use the above investigations in certain statistical robust problems and in decision making problems, where we obtain the outer information x.

Example 2. (ε-contamination class) [12, 19, 21] Suppose that we have an initial prior probability P_0 on the space of states of nature S but we are not completely confident about the prior so we form the following family of probability measures

$$\Pi = \{P\colon P = (1-\varepsilon)P_0 + \varepsilon\mathcal{Q},\ \mathcal{Q} \in \mathscr{M}\}\,,$$

where $0 \le \varepsilon \le 1$. This prior distribution of uncertainty is generated by some random set and the lower and upper posterior probabilities have the following form:

$$\underline{P}_x(A) = \frac{(1-\varepsilon)\int L_{A,x}(s)\, dP_0(s)}{(1-\varepsilon)\int L_{S,x}(s)\, dP_0(s) + \varepsilon \sup\limits_{s\in A^c} L(s)}\,,$$

$$\overline{P}_x(A) = \frac{(1-\varepsilon)\int L_{A,x}(s)\, dP_0(s) + \varepsilon \sup\limits_{s\in A} L(s)}{(1-\varepsilon) L_{S,x}(s)\, dP_0(s) + \varepsilon \sup\limits_{s\in A} L(s)}\,.$$

7. The insurance premium

The insurance premium calculation is the basic problem in actuarial science. The *insurance risk* X is defined as a non-negative random variable and *insurance premium* is a functional

$$H\colon \mathcal{X} \to [0, \infty)\,,$$

where $\mathcal{X}$ is the class of all nonnegative random variables on S.

If $a, h > 0$, the layer at $(a, a+h]$ of the risk X is defined in the following manner:

$$I_{(a,a+h]}(X) = \begin{cases} 0\,, & 0 \le X < a\,, \\ X - a\,, & a \le X < a + h\,, \\ h\,, & a + h \le X\,. \end{cases}$$

The notion of layer of risk is used in many reinsurance problems [17].

The functional correctly representing the insurance premium ought to satisfy the following basic properties: [5, 17, 18]

i) nonnegative safety loading: $H(X) \ge E(X)$,

ii) no rip-off: $H(X) \le \sup X$,

iii) consistency: $H(X + c) = H(X) + c,\ c \in R$,

iv) proportionality: $H(cX) = cH(X),\ c \ge 0$,

v) preserving the first stochastic dominance: if $G_X(t) \le G_Y(t)$ for all $t > 0$ then $H(X) \le H(Y)$,

vi) subadditivity: $H(X + Y) \le H(X) + H(Y)$,

vii) no unjustified risk loading: $P(X = b) = 1 \Rightarrow H(X) = b$,

viii) comonotonic additivity: if the risks X, Y are comonotonic then $H(X + Y) = H(X) + H(Y)$,

ix) layer additive: $H(X) = \sum\limits_{i=1}^{\infty} H(I_{(x_{i-1},x_i]})$,

for any division $0 = x_0 < x_1 < \dots < x_n < \dots$

The two last properties are strictly connected with reinsurance problems.

The insurance premium functional defined using the Choquet integral

$$H(X) = \int X\, d\mu\,,$$

where μ is a concave set function [5, 18], satisfies the above conditions. For instance, the concave distorted probability $\mu = f(P)$ is an example of such set functions.

EXAMPLE 3.

i) Let the function $f(x) = x^q$, where $0 < q \le 1$. Then we obtain the *proportional hazards transform*, i.e., the simple and natural risk-adjusted premium defined by Wang in [18]. In this case, the functional of the insurance premium is

$$H(X) = \int\limits_0^{\infty} (G_p(t))^q dt\,.$$

The set function $\mu = f(P)$, where P is a probability measure, is a plausibility function (Theorem 5i) and decomposable measure with additive generator $g(x) = x^p$, where $p = 1/q$, (Theorem 2) induced by family of Yager's t-conorms [24]

$$Y_p(x, y) = \min\left\{1, (x^p + y^p)^{1/p}\right\}.$$

ii) Let $f(x) = 1 - (1 - x)^q$, where $q > 1$. This transformation, called the *dual-power function*, is another function used in actuarial theory in premium calculation [5, 17]. The set function $\mu = f(P)$ is a plausibility function for integer q (Theorem 5ii) and the decomposable measure with additive generator $g(x) = 1 - (1 - x)^p$, where $p = 1/q$ and $0 < p < 1$.

REFERENCES

[1] BUJA, A.: *Simultaneously least favourable experiments. I: Upper standard functionals and sufficiency.*, Z. Wahrsch. Verw. Gebiete **65** (1984), 367–384.

[2] CHATEAUNEUF, A.: *Decomposable capacities, distorted probabilities and concave capacities*, Math. Social. Sci. **31** (1996), 19–37.

[3] CHOQUET, G.: *Theory of capacities*, Ann. Inst. Fourier **5** (1953), 131–295.

[4] DEMPSTER, A. P.: *Upper and lower probabilities induced by a multivalued mapping*, Ann. Math. Stat. **38** (1967), 325–339.

[5] DENNEBERGD.: *Premium calculation: why standard deviation should be replaced by absolute deviation*, Astin Bulletin **20** (1990), 181–190.

[6] GILBOA, I.: *Expected utility with purely subjective nonadditive probability*, J. Math. Econom. **56** (1987), 65–88.

[7] DUBOIS, D.—PRADE, H.: *A class of fuzzy measures based on triangular norms*, Internat. J. Gen. Syst. **8** (1982), 43–61.

[8] GRABISCH, M.—NGUYEN, H. T.—WALKER, E. A.: *Fundamentals of Uncertainty Calculi, with Application to Fuzzy Inference*, Kluwer Academic, Dordrecht, 1995.

[9] HEILPERN, S.: *Distribution of uncertainty and decision making*, in: Cybernetics and Systems Research (R. Trappl, ed.), Vol. 1., World Scientific, Singapure, 1992.

[10] HEILPERN, S.: *Decision making under uncertainty*, Wydawnictwo AE, Wrocław, 1992. (In Polish.)

[11] HEILPERN, S.: *Distorted probabilities, decomposable measures and insurance premium*, in: Proc. Seventh IFSA World Congress, Prague 1977, Vol. IV. pp. 3–8.

[12] HUBER, P. J., STRASSEN, V.: *Minimax tests and Neyman-Pearson lemma for capacities*, Ann. Statist. **1** (1965), 251–263.

[13] PUPPE, C.: *Distorted Probabilities and Choice under Risk*, Springer-Verlag, Berlin, 1991.

[14] SCHMEIDLERD.: *Subjective probability and expected utility without additivity*, Econometrica **57** (1988), 571–587.

[15] SCHWEIZER, B.—SKLAR, A.: *Associative functions and statistical triangle inequalities*, Publ. Math. Debrecen **8** (1961), 169–186.

[16] SHAFER, G.: *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, 1976.

[17] WANG, S.: *Premium calculation by transforming the layer premium density*, Astin Bulletin **26** (1996), 71–92.

[18] WANG, S.: *Risk-adjusted credibility premiums using distorted probabilities*, University of Waterloo, Res. Report 4, 1996.

[19] WASSERMAN, L. A.—KADANE, J. B.: *Bayes' theorem for Choquet capacities*, Ann. Statist. **18** (1990).

[20] WASSERMAN, L. A.: *Belief functions and statistical inference*, Canad. J. Statist. **18** (1990).

[21] WASSERMAN, L. A.: *Prior envelopes based on belief functions*, Ann. Statist. **18** (1990).

[22] WIERZCHOŇ, S. T.: *The use of Dempster's upper and lower probabilities in decision-making*, in: Proc. Polish Symp. Interval & Fuzzy Math., Poznań 1983, Poznań, 1985, pp. 228–242.

[23] YAARI, M. E.: *The dual theory of choice under risk*, Econometrica **55** (1987), 95–115.

[24] YAGER, R. R.: *Multiple objectives decision-making using fuzzy sets*, Internat. J. Man-Machine Stud. **9** (1977), 375–382.

Received March 25, 1998

Institute of Economic Cybernetics
Academy of Economics
ul. Komandorska 118–120
PL–53-345 Wrocław
POLAND

E-mail: heilpern@manager.ae.wroc.pl.