Continuous m-dimensional distorted probabilities

Vicenç Torra[a], Montserrat Guillen[b], Miguel Santolino[b]

[a] *School of Informatics, University of Skövde, Sweden*
[b] *Department of Econometrics, Riskcenter-IREA, University of Barcelona, Spain*

Abstract

Fuzzy measures, also known as non-additive measures, monotonic games, and capacities, have been used in many contexts. For example, they have been applied in economics, risk analysis, computer science in computer vision and machine learning and, in general, in mathematics.

However, when looking at applications, one of the problems that still needs to be solved is how the measure should be defined in an easy and intuitive way. When the reference set is finite, a few families of measures have been established, e.g. distorted probabilities, k-additive and decomposable measures. But, when the reference set is infinite, the only family is distorted probabilities.

In this paper we give a definition for m-dimensional distorted probabilities in the case that the reference set is not finite, and we study some properties of this family. We also give a definition for the hierarchically decomposable m-dimensional distorted probability that relates to another family of measures defined for the finite case.

Keywords: Fuzzy measure, Non-additive measures, m-dimensional distorted probabilities

1. Introduction

Distorted probabilities modify the shape of the distribution of a random variable. So, they provide a suitable framework to input information on the parts of the domain that, for some reason, seem to be more relevant than others. This is especially interesting in areas like risk management, where the main focus is on the extremes of the distribution rather than on the central part. Similarly, in business analytics, one may be willing to analyze profits and losses, but would rather be more worried about losses, or negative outcomes, than profits or positive gains. In fact, the distortion function must somehow reflect the concept of ”importance” or ”relevance”, whereas the original distribution is nothing more than reflecting the random behaviour.

The distorted probability is a convolution of a probability distribution and a distortion function that produces a measure which combines the stochastic nature in each

Email addresses: vtorra@his.se (Vicenç Torra), mguillen@ub.edu (Montserrat Guillen), msantolino@ub.edu (Miguel Santolino)

part of the domain of the distribution with the weight (or relative importance) of each region in the domain.

Many have contributed to this field [15, 17]. In the insurance field, Denneberg [2] (1989 and 1990) was a pioneer in addressing the fact that large economic losses are more important than small ones. He suggested that this simple idea can easily be implemented by looking at the distribution function in a different perspective by means of a distortion.

We aim at developing a theory for m-dimensional distorted probabilities using the results from the discrete case and fuzzy measures. In our examples we address two different ideas. Firstly, we simply show that distortions can operate on multidimensional regions, as opposed to intervals in the one dimensional case. Secondly, we show that the m-dimensional case can be used to cope with situations where a measure cannot be generated by a single one-dimensional distorted probability. It is known that in the discrete case fuzzy measures representable as one-dimensional distorted probabilities are a small fraction of the total. See, e.g., [8, 6] for details on how this is computed.

Our construction presumes that defining probability distributions on regions on continuous domains is easy, but that defining fuzzy measures on such continuous domains is not. So, an implicit goal is to use these distributions to define fuzzy measures by means of the distortion function, as we do in the discrete case. Observe that one can find a few ways of defining probability distributions on continuous domains. There are, among others, the multivariate normal distribution, multivariate t-distribution, spherical and elliptical distributions. Some distributions based on the Choquet integral have been even introduced [11].

In this paper, we also propose a definition for hierarchically decomposable m-dimensional distorted probabilities. This type of measures can be seen as a natural generalization of m-dimensional distorted probabilities and of hierarchically decomposable fuzzy measures. This latter type of measures were previously defined in a discrete setting in [10].

The paper is organized as follows. Section 2 presents notation and definitions. Section 3 provides the definition of distorted probabilities on random vectors, and introduces the new families of measures. Some properties of these new measures are analyzed. Section 4 shows some examples of splitting probability reference sets and Section 5 concludes.

2. Preliminaries

In this section we review some basic definitions on fuzzy measures and integrals. For more detailed description of the theory behind them the reader is referred to [3, 5, 12, 14].

Definition 1. *Let X be a finite reference set. A set function $\mu : X^n \to [0,1]$ is a fuzzy measure if it satisfies the following axioms:*

(i) $\mu(\emptyset) = 0$, $\mu(X) = 1$ *(boundary conditions)*
(ii) $A \subseteq B$ *implies* $\mu(A) \leq \mu(B)$ *for* $A, B \subseteq X$ *(monotonicity)*

These functions are also known with the names capacities, non-additive measures, and monotonic games.

Similar definitions exist when X is not finite. Then, we consider a measurable space $(X, \mathscr{X})$ and $A, B \in \mathscr{X}$. For example, it can be considered the measurable space $(\mathbb{R}, \mathscr{B})$ where $\mathbb{R}$ is the real line and $\mathscr{B}$ is a Borel σ-algebra. The axiom of continuity is often also required to the fuzzy measure, particularly when X is not finite [9, 16].

Definition 2. *Let μ be a fuzzy measure on a measurable space $(X, \mathscr{A})$. The fuzzy measure μ is continuous if it satisfies the following axiom. If $A_n \in \mathscr{A}$ and A_n is monotone, then $lim_{n\to\infty}\mu(A_n) = \mu(lim_{n\to\infty}A_n)$.*

The definition of fuzzy measures generalizes additive measures (and probabilities), both in the case of finite and not finite X.

Fuzzy measures are used in combination with the fuzzy integrals (e.g., the Choquet and Sugeno integrals [3, 9]). Their definition follows.

Definition 3. *Let X be a reference set, let $(X, \mathscr{A})$ be a measurable space, let μ be a fuzzy measure on $(X, \mathscr{A})$, and let f be a measurable function $f : X \to [0,1]$; then, the Choquet integral of f with respect to μ is defined by*

$$C_\mu(f) := \int_0^\infty \mu_f(r)dr,$$

where $\mu_f(r) := \mu(\{x|f(x) > r\})$.

Definition 4. *Let X be a reference set, let $(X, \mathscr{A})$ be a measurable space, let μ be a fuzzy measure on $(X, \mathscr{A})$, and let f be a measurable function $f : X \to [0,1]$; then, the Sugeno integral of f with respect to μ is defined by*

$$S_\mu(f) := \sup_{r\in[0,1]} [r \wedge \mu_f(r)],$$

where $\mu_f(r) := \mu(\{x|f(x) > r\})$ and $\wedge$ stands for the minimum.

The Choquet integral of an additive measure results into the Lebesgue integral. Accordingly, the Choquet integral of a function f with respect to a fuzzy measure can be seen as a kind of generalization of the expectation of f.

For finite X, the definition of a fuzzy measure requires that we assign to each $A \subseteq X$ a value satisfying the constraints above. This means that we need to consider $2^{|X|} - 2$ values (taking into account boundary conditions). A few families of measures have been proposed in the literature. Some examples are k-additive measures [4], $\perp$-decomposable fuzzy measures, Sugeno λ-measures [9] and hierarchically decomposable fuzzy measures [10]. These families have been introduced to simplify the burden of supplying the $2^{|X|} - 2$ values that are required.

2.1. Distorted probabilities

A distorted probability can be seen as a fuzzy measure on a set X. A distorted probability is defined in terms of a probability distribution and a distortion function of this probability distribution. Its definition follows.

Definition 5. *A fuzzy measure μ on a reference set X is a distorted probability if it can be expressed in terms of a non-decreasing function $g : [0,1] \to [0,1]$ and a probability P as $\mu = g \circ P$. That is, $\mu(A) = g(P(A))$ for all $A \subseteq X$.*

Decision makers are often interested in some parts of the domain. For instance, when looking at risk, we would rather weight the extremes much more than the other cases. This is the reason why distorted probabilities appear at the heart of disciplines such as risk management.

In the case of a discrete X, when $|X| = 2$ all measures are distorted probabilities, but for $|X| = 5$ the *number* of measures representable as distorted probabilities becomes rather small. In order to solve this inconvenience, [8] introduced m-dimensional distorted probabilities. The definition follows.

Definition 6. *[8] Let X be a discrete reference set, let $\{X_1, X_2, \cdots, X_m\}$ be a partition of X and let $(X_i, \mathscr{X}_i)$ be measurable spaces $(i = 1, \cdots, m)$, then we say that μ is an at most m dimensional distorted probability if there exists a function f on $[0,1]^m$ and probability P_i on $(X_i, \mathscr{X}_i)$ such that:*

$$\mu(A) = f(P_1(A \cap X_1), P_2(A \cap X_2), \cdots, P_m(A \cap X_m)) \tag{1}$$

[8] studies some properties. One of them is that for any fuzzy measure μ there exists a dimension d such that μ can be represented as a d-dimensional distorted probability. This implies that all fuzzy measures can be expressed as d-dimensional distorted probabilities.

2.2. Hierarchically decomposable fuzzy measures

The hierarchically decomposable fuzzy measure is a type of measure introduced for discrete domains in [10], in which the measure of the singletons is combined by means of t-conorms [7] in a hierarchical way. To do so, we first need to define a hierarchy of the elements in X. We do so below.

Definition 7. *H is a hierarchy of elements X if and only if the following conditions are fulfilled:*

(i) All the elements in X belong to the hierarchy, and the corresponding nodes are the leaves of the hierarchy:
For all x in X, $\{x\} \in H$.

(ii) There is only one root in the hierarchy, and it is denoted by root*. A node is the root if it is not included in any other node:*
if root $\in H$*, then there is no other node $m \in H$ such that* root *is m.*

(iii) All nodes belong to one and only one node, except for the root:
if $n \in H$ and $n \neq$ root*, then there exists a single $m \in H$ such that n is m.*

(iv) All nodes that contain only one element are singletons:
if $|h| = 1$, then there exists $x \in X$ such that $h = \{x\}$ for all $h \in H$.

(v) All non-singletons are defined in terms of nodes that are in the tree:
if $|h| \neq 1$, then, for all $h_i \in h$, $h_i \in H$.

Given a hierarchy of elements, it is assigned to each leaf in the hierarchy a real value in the unit interval, and, for each node that is not a leaf, a t-conorm. This is called *labeled hierarchy* and its definition is given below. Additionally, we need to know for each node h in the hierarchy H which are the elements that this node encompasses. This is given by the function extension $EXT(h)$ defined as follows.

Definition 8. *Let H be a hierarchy according to Definition 7 and let h be a node in H; then, the* extension *of h in H is defined as:*

$$EXT(h) := \begin{cases} h & \text{if } |h| = 1 \\ \cup_{h_i \in h} EXT(h_i) & \text{if } |h| \neq 1. \end{cases}$$

Definition 9. *Let H be a hierarchy according to Definition 7; then, a* labeled hierarchy *L for H is a tuple $L = < H, \perp, m >$, where $\perp$ is a function that maps each node $n \in H$ that is not a leaf into a t-conorm, and m is a function that maps each singleton into a value of the unit interval. For simplicity, we will express $\perp(h)$ by $\perp_h$.*

With all these elements, we can now define the hierarchically decomposable fuzzy measures.

Definition 10. *Let $L = < H, \perp, m >$ be a labeled hierarchy according to Definition 9, let EXT be the function in Definition 8; then, the corresponding* Hierarchically $\perp$-Decomposable Fuzzy Measure *(HDFM for short) of a set B is defined as $\mu(B) = \mu_{root}(B)$, where μ_A for a node $A = \{a_1, ..., a_n\}$ is defined recursively as*

$$\mu_A(B) = \begin{cases} 0 & \text{if } |B| = 0 \\ m(B) & \text{if } |B| = 1 \\ \perp_A(\mu_{a_1}(B_1), ..., \mu_{a_n}(B_n)) & \text{if } |B| > 1. \end{cases}$$

Here, $B_i = B \cap EXT(a_i)$ for all a_i in A.

When $\mu(X) = 1$, Definition 10 leads to a fuzzy measure.

3. Distorted probabilities on random vectors and new families of measures

In this section we consider vectors of random variables and extend the concept of distorted probabilities to dimensions higher than one. If we consider only one single random variable, the partition boils down to splitting the domain in disjoint intervals.

Before extending m-dimensional distorted probabilities to random vectors on continuous domains, let us look at some graphical presentation. Definitions are given on vectors of two variables but the definition can be modified straightforwardly to any n-dimensional random vector with $n > 2$. We have different regions in the domain (regions can have arbitrary shape), each one with its probability distribution, and a function to combine the probability measures when we have sets on different regions.

Figure 1 shows two general examples of a two-dimensional random vector domain in which it is displayed a circle representing a set A that is a part of a measurable space $\mathscr{A}$. Partitions are shown in the two graphs of the figure and they are delimited by solid

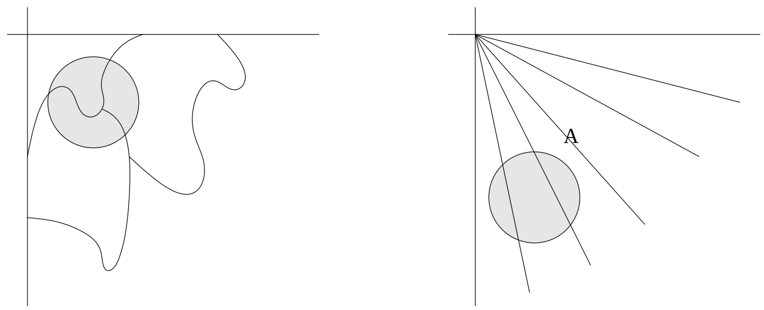

Figure 1: Representation of two different m-dimensional distorted probabilities in a two dimensional space.

borders. On the left hand side, partitions are exemplified as finite, whereas on the right hand side, we have plotted unbounded partitions. Probability distributions are defined on the plane for each region of the two random variables (which are not necessarily independent) and the distortion function is then implemented on the intersection of circle A with each of the partitions. In both graphs, A intersects with three different regions.

Definition 11. *Let Y_1, Y_2 be two random variables and $X = (Y_1, Y_2)$ be a vector of random variables (a random vector) with some joint probability distribution. Let $\{X_1, X_2, \cdots, X_m\}$ be a partition of X in the sense that X_i is composed of two random variables defined on a partition of the domain of X, and $(X_i, \mathscr{X}_i)$ be measurable spaces, then we say that μ is an at most m dimensional distorted probability if there exists a function f on $[0,1]^m$ and probabilities P_i on $(X_i, \mathscr{X}_i)$ such that:*

$$\mu(A) = f(P_1(A \cap X_1), P_2(A \cap X_2), \cdots, P_m(A \cap X_m)). \tag{2}$$

Note that m refers to the number of partition regions.

Regions of X_i can be closed or open, finite (see Figure 1 (left)) or infinite (see Figure 1 (right)), and each region have a probability distribution P_i. That is, we can compute for all $A \in \mathscr{X}_i$ the probability $P(A)$. The function f combines the probabilities of $P_i(A \cap X_i)$ in each region X_i. In Figure 1, the shaded circle corresponds to A, and it is clearly observed that this region intersects with three partition regions. On the left, the three regions are finite and on the right, they are infinite. Then the distorted probability of each part of the circle is implemented through f.

Taking into account the properties of fuzzy measures we can prove some properties on the function f involved in m-dimensional distorted probabilities. As it can be seen from the results below, we do not make any assumption on the properties of f but we derive them from the properties of non-additive measures. In other words, we prove in this paper a characterization of the function f. In this way, Definition 11 is the most general one as it has no constraints on f.

Proposition 12. *Let $\{X_1, X_2, \cdots, X_m\}$ be a partition of X and let $(X_i, \mathscr{X}_i)$ be measurable spaces. Let μ be a m-dimensional distorted probability generated by $P_1, \ldots, P_m$ and f. Then, we can prove the following properties for the function f:*

- *Unanimity in zero. I.e.,* $f(0,\ldots,0)=0$.
- *Unanimity in one. I.e.,* $f(1,\ldots,1)=1$.

Proof. Unanimity in zero follows from the fact that $\mu(\emptyset)=0$, and the fact that $P_i(\emptyset)=0$ for all i. Unanimity in one follows from the fact that $\mu(X)=1$ and that $P_i(X_i)=1$ for all X_i. □

Proposition 13. *Let* $\{X_1,X_2,\cdots,X_m\}$ *be a partition of* X *and let* $(X_i,\mathscr{X}_i)$ *be measurable spaces. Let* μ *be a m-dimensional distorted probability on* X *using these measurable spaces and generated by* $P_1,\ldots,P_m$ *and* f*. Then, if the density function* p_i *of* P_i *is continuous,* f *is monotonic.*

Proof. Recall that f is monotonic when $a\leq b$ implies $f(a)\leq f(b)$ for a,b vectors of $[0,1]^m$.

First note that for any a_i,b_i such that $0\leq a_i<b_i\leq 1$ we have that, there are sets $A_i\in\mathscr{X}_i$ with probability $a_i=P(A_i)$ and $C_i\in\mathscr{X}_i$ such that $A_i\cap C_i=\emptyset$ and $P(C_i)=c_i=b_i-a_i$. This is so because we have assumed that the density functions p_i are continuous.

Then, $A=\cup A_i$ and $C=\cup_i C_i$. Then, $C\cap A=\emptyset$. It is thus clear that $A\subseteq A\cup C$ implies $\mu(A)\leq\mu(A\cup C)$. Therefore, the following is also true.

$$\begin{aligned}
\mu(A) &= f(a_1,\ldots,a_m)\\
&= f(P_1(A\cap X_1),P_2(A\cap X_2),\cdots,P_m(A\cap X_m))\\
&\leq f(P_1(A\cup C\cap X_1),P_2(A\cup C\cap X_2),\cdots,P_m(A\cup C\cap X_m))\\
&= f(P_1(A\cap X_1)+P_1(C\cap X_1),P_2(A\cap X_2)+P_2(C\cap X_2),\cdots,P_m(A\cap X_m)+P_m(C\cap X_m))\\
&= f(a_1+c_1,\ldots,a_m+c_m)\\
&= f(b_1,\ldots,b_m)\\
&= \mu(A\cup C)
\end{aligned} \tag{3}$$

Therefore, the proposition is proven. □

We have required in this proposition that density functions are continuous. Note that, when density functions are not continuous, we can select a value a_i for which there is a set A_i such that $a_i=P(A\cap X_i)$ and a value b_i such that $a_i\leq b_i$, and then the function f assigns $f(a_i)>f(b_i)$. Observe that such non-monotonic function will lead to a fuzzy measure.

A function satisfying unanimity in zero, in one and monotonicity is an aggregation operator in the sense of [1, 5]. t-Norms, t-conorms, uninorms, copulas, means and aggregation operators in the sense of [12] satisfy these properties.

The reversal can be also proved. That is, if f is monotonic and satisfying unanimity in zero and one, then a partition $X=\{X_1,\ldots,X_m\}$ and probabilities $P_1,\ldots,P_m$ on the parts of X lead to a m-dimensional distorted probability.

Proposition 14. *Let $\{X_1, X_2, \cdots, X_m\}$ be a partition of X and let $(X_i, \mathscr{X}_i)$ be measurable spaces, let f be a monotonic function satisfying unanimity in zero and one. Then,*

$$\mu(A) = f(P_1(A \cap X_1), P_2(A \cap X_2), \cdots, P_m(A \cap X_m)) \tag{4}$$

is a m-dimensional distorted probability.

Proof. By construction $\mu(\emptyset) = f(0, 0, \ldots, 0) = 0$ and $\mu(X) = f(1, 1, \ldots, 1)$. Then, $\mu(A) \leq \mu(B)$ when $A \subseteq B$ follows from monotonicity of f. □

Propositions 12, 13, and 14 imply that unanimity in zero and one, and monotonicity are necessary and sufficient conditions for a function f to generate a m-dimensional distorted probability.

We can also prove the proposition below which establishes what type of function f ensures that the measure is additive. The proposition concludes that the only type of function that is suitable is a weighted mean.

Proposition 15. *A m-dimensional distorted probability is additive if and only if*

$$f(a_1, a_2, \ldots, a_n) = \sum \alpha_i a_i$$

with α such that $\alpha_i \geq 0$ and $\sum \alpha_i = 1$. That is, f is a weighted mean.

Proof. Our approach is similar to the one of the proof of Proposition 13. Let us consider two pairs of disjoint sets $A_i \in \mathscr{X}_i$ with probability $a_i = P(A_i)$ and $C_i \in \mathscr{X}_i$ with probability $c_i = P(C_i)$ such that $A_i \cap C_i = \emptyset$. Note that $A = \cup A_i$ and $C = \cup C_i$. Then,

$$P(A \cup B) = P(A) + P(C)$$

and therefore,

$$f(a_1 + c_1, \ldots, a_n + c_n) = f(a_1, \ldots, a_n) + f(c_1, \ldots, c_n).$$

This functional equation has been studied in the literature. See, e.g., Proposition 3.4 in [12]. The solution of this functional equation is

$$f(a_1, \ldots, a_n) = \sum \alpha_i a_i$$

with α_i arbitrary. The condition that $f(1, 1, \ldots, 1) = 1$ implies that

$$f(1, 1, \ldots, 1) = \sum \alpha_i 1 = \sum \alpha_i = 1$$

and monotonicity implies that $\alpha_i \geq 0$.

Therefore, the proposition is proven. □

Given three functions f_1, f_2, and f_3 such that they satisfy unanimity in zero and one, and monotonicity, it is easy to see that the function f defined by

$$f(a_1, \ldots, a_n) = f_1(f_2(a_1, \ldots, a_r), f_3(a_{r+1}, \ldots, a_n))$$

satisfies also unanimity in zero and one, and monotonicity. This property is known for aggregation operators. See, e.g., [1, 5, 12].

Using this property, we can consider a hierarchical structure for f by means of functions f_s. Such measure will be thus defined by the partition $\{X_1, X_2, \cdots, X_m\}$ on X, the measurable spaces $(X_i, \mathscr{X}_i)$, and a hierarchical structure of f on the partition on X. This type of measure will be a m-dimensional distorted probability by definition. The following example presents a specific case of a five-dimensional distorted fuzzy measure.

Example 1. *Let $\{X_1, X_2, X_3, X_4, X_5\}$ be a partition on X, let $(X_i, \mathscr{X}_i)$ be measurable spaces, let f_1, f_2, f_3, and f_4 be monotonic functions satisfying unanimity in zero and one. Then, the measure μ defined by*

$$\mu(A) = f_4(f_2(f_1(P(A \cap X_1), P(A \cap X_2)), P(A \cap X_3)), f_3(P(A \cap X_4), P(A \cap X_5)))$$

is a 5-dimensional distorted fuzzy measure.

This example shows that a measure of this type is a natural extension of hierarchically decomposable fuzzy measures (see Definition 10). Note that t-conorms, which were used to define hierarchically decomposable fuzzy measures in Definition 10, are monotonic functions satisfying unanimity in zero and one.

We call a measure of this type a hierarchically decomposable m-dimensional distorted probability. A formal definition of this type of measure will follow the structure of Definition 10. We do not include it here for the sake of conciseness.

4. Application of m-dimensional distorted probabilities

In this section we discuss a few examples of m-dimensional distorted probabilities. We begin with an example to illustrate a case in which the measure can be represented by a 2-dimensional distorted probability but not by the traditional (one dimensional) distorted probability.

Example 2. *Let A, B, C be three disjoint subsets of X. Let us consider the case that $\mu(A) < \mu(B)$ but that $\mu(A \cup C) > \mu(B \cup C)$. For example, let $\mu(A) = 0.1$, $\mu(B) = 0.5$, $\mu(C) = 0.2$, $\mu(A \cup C) = 0.8$, and $\mu(B \cup C) = 0.6$.*

This situation can not be expressed with distorted probabilities. Note that as a distorted probability is such that $\mu = f \circ P$, $\mu(A) < \mu(B)$ implies $P(A) < P(B)$ and $\mu(A \cup C) > \mu(B \cup C)$ implies $P(A \cup C) = P(A) + P(C) > P(B \cup C) = P(B) + P(C)$ which implies $P(A) > P(B)$.

We can represent a measure of this type with a 2-dimensional distorted probability. Let X_1 and X_2 define a partition of X. Let $(X_1, \mathscr{X}_1)$ and $(X_2, \mathscr{X}_2)$ be measurable spaces such that $A \in \mathscr{X}_1$, $B \in \mathscr{X}_2$ and also $C \in \mathscr{X}_1$. Then, to build the 2-dimensional distorted probability we need to define P_1 on X_1, P_2 on X_2 and f.

Let us assume that we already have a probability distribution P on X. Then, let us define $P_1(A) = P(A|X_1)$ for all $A \in \mathscr{X}_1$ and $P_2(A) = P(A|X_2)$ for all $A \in \mathscr{X}_2$. Then, let $a = P_1(A)$, $b = P_2(B)$, and $c = P_1(C)$. Now, we can define f as follows: $f(0,0) = 0$, $f(0,b) = 0.5$, $f(a,0) = 0.1$, $f(a,b) = \mu(A \cup B)$, $f(c,0) = 0.2$, $f(c,b) = 0.6$, $f(a+c,0) = 0.8$ and $f(a+c,b) = \mu(A \cup B \cup C)$.

The definitions given for P_1, P_2 and f permit us to define the two-dimensional distorted probability μ in terms of f, P_1 and P_2.

In this example we have supposed that P was already known. If this is not the case, any arbitrary P_1 and P_2 could be used in this case as there are no other constraints than $\mu(A) = 0.1$, $\mu(B) = 0.5$, $\mu(C) = 0.2$, $\mu(A \cup C) = 0.8$, and $\mu(B \cup C) = 0.6$. Recall, however, that $f : [0,1]^2 \to [0,1]$ and that f needs to satisfy unanimity in zero, unanimity in one and monotonicity according to Propositions 12 and 13.

We can use this structure in situations as in the following example.

Example 3. *Let $X = \mathbb{R}$. Let x_0 be an element of X. Let $X_1 = \{x | x < x_0\}$ and $X_2 = \{x | x \geq x_0\}$. Given P_1 and P_2 probability distributions on X_1 and X_2 and given a function f, we can define a 2-dimensional distorted probability using Equation (2).*

Let us consider a measure that is highly superadditive in one region (say X_1) but with low values for small sets. E.g., as in Example 2 ($\mu(A) = 0.1$, $\mu(C) = 0.2$ but $\mu(A \cup C) = 0.8$), but at the same time there are other regions (say X_2) where the measure is large (e.g., $\mu(B) = 0.5$). In this case, if adding a region of X_2 to a region X_1 increases the measure only moderately, we may have a condition as in Example 2 $\mu(B \cup C) < \mu(A \cup C)$. In this case, a distorted probability cannot be used and we need a 2-dimensional distorted probability (as the one in Example 2).

We introduce now an example of the application of this type of measures to financial markets. We show that situations described in the previous two cases, may reflect the attitudes of investors in the financial markets.

Example 4. *Let us assume that X is a random variable associated with profits and losses. So, its domain is equal to $\mathbb{R}$, where profits take positive values and losses correspond to negative values. The decision-maker considers the partition $X = \{X_1, X_2\}$ where $X_1 \in (-\infty, k]$ and $X_2 \in (k, +\infty)$ with probability distributions P_1 and P_2 respectively, where k is a positive constant. The idea is that k reflects a maximum level of gains that an investor would consider the upper limit in a normal context.*

Let A, B, C be three disjoint subsets of X associated with very large losses (A), moderate profits (B) (i.e., positive values smaller than k) and high profits (C), so $A \subseteq X_1$, $B \subseteq X_1$ and $C \subseteq X_2$.

Let us consider the case of a conservative investor. Having some knowledge on the market, the investor would think that C is less plausible than B. So, he weights large profits lower than moderate profits, $\mu(C) < \mu(B)$, simply because he thinks the latter occurs more often than the former. However, his opinion varies when considering also the possibility of large losses. In that case, he thinks that $\mu(A \cup B) < \mu(A \cup C)$. Indeed, he thinks that when large losses occur there is more volatility in the market and therefore large profits are also more plausible than just moderate profits. This is the reason he weights the extreme scenario $A \cup C$ more than the asymmetric case $A \cup B$. Let us consider $a = P_1(A)$, $b = P_2(B)$, and $c = P_2(C)$, then:

$$\mu(B) = f(0,b) > f(0,c) = \mu(C)$$
$$\mu(A \cup B) = f(a,b) < f(a,c) = \mu(A \cup C).$$

So, the plausibility measure of the investor can be obtained as a result of a two-dimensional distortion measure rather than as a one-dimensional one.

Let us assume that the annual financial results of the business line X follow a normal distribution with mean $m = 0.15$ and variance, $\sigma^2 = 0.30$. In this example, k is equal to 0.30. Let A, B, C be three disjoint subsets of X such that $A \subseteq X_1$, $B \subseteq X_1$ and $C \subseteq X_3$, representing large losses, moderate profits and high profits. In particular $A \in (-\infty, -0.30]$ and $B \in [0.10, 0.15]$ and $C \in [0.60, \infty)$. In this case, $P_1(A) = 0.07$, $P_1(B) = 0.07$ and $P_2(C) = 0.07$. A conservative manager thinks that moderate profits occur 25% more often than than moderate profits, $\mu(C) < \mu(B)$. However, his opinion varies when considering also the possibility of large losses. In that case, he thinks that $\mu(A \cup C)$ is 25% more often than $\mu(A \cup B)$:

$$\mu(B) = f(0, b) = 0.0875 > 0.07 = f(0, c) = \mu(C)$$
$$\mu(A \cup B) = f(a, b) = 0.14 < 0.175 = f(a, c) = \mu(A \cup C).$$

Our last example is simply a case where two-dimensional distortions arise naturally.

Example 5. *Let the random vector $X = (Y_1, Y_2)$ consists of two random variables Y_1 and Y_2 which are associated with the annual financial results of two business lines, whose domain is $\mathbb{R}^2$. The business analyst that only uses one-dimensional concepts would think that losses may have a bad impact on the firm, whereas profits always have a good impact. However, in the two-dimensional scenario, one can imagine that the worst case is a conjunction of the two business lines having losses, or even when the profits on one line cannot compensate the losses of the other. So, one should weight the region of the sum of the two random variables being negative as having much more importance than when they are both positive, or even when one line is negative but can be compensated by the other. One would establish the following three regions in the domain:*

- *All lines are profitable*

$$X_1 = \{Y_{11}, Y_{21}\} \in [0, +\infty)^2$$

- *The sum is profitable, but one line is negative*

$$X_2 = \{Y_{12}, Y_{22}\} \in \{(y_1, y_2) | -y_2 < y_1 < 0\} \cup \{(y_1, y_2) | -y_1 < y_2 < 0\}$$

- *The sum is not profitable*

$$X_3 = \{Y_{13}, Y_{23}\} \in \{(y_1, y_2) | y_1 + y_2 < 0\}.$$

In this case the 3-dimensional distortion probability is quite natural and could reflect different attitudes of the management in each region of the partition. So, more weight could be given to the third case, when trying to reflect the importance of having an aggregate business that is not profitable.

Let us assume that we know the probability distribution P on X. Then, let us define $P_1(D) = P(D|X_1)$ for all $D \in \mathscr{X}_1$, $P_2(D) = P(D|X_2)$ for all $D \in \mathscr{X}_2$ and $P_3(D) =$

$P(D|X_3)$ for all $D \in \mathscr{X}_3$. Let A,B,C be three disjoint subsets of X such that $A \subseteq X_1$, $B \subseteq X_2$ and $C \subseteq X_3$. Let consider the case that $P_1(A) = P_2(B) = P_3(C)$. A conservative manager would have a risk averse attitude giving more weight to the third case than the other two scenarios, as follows:

$$\mu(A) = f(a,0,0) < f(0,0,c) = \mu(C)$$
$$\mu(B) = f(0,b,0) < f(0,0,c) = \mu(C)$$

where $a = P_1(A)$, $b = P_2(B)$, $c = P_3(C)$. However, a risk neutral agent would weight the three subsets with the same importance,

$$\mu(A) = f(a,0,0) = \mu(B) = f(0,b,0) = \mu(C) = f(0,0,c).$$

Finally, a decision-maker with risk appetite would consider more plausible the first subset and would give more weight to this scenario than the others,

$$\mu(A) = f(a,0,0) > f(0,b,0) = \mu(B)$$
$$\mu(A) = f(a,0,0) > f(0,0,c) = \mu(C).$$

Now, we will consider the concrete numerical example that was designed in Example 4. Let us assume that an additional business line is considered with mean and variance equal to 0.1. *In this case, the annual financial results of the two business lines X follow a bivariate normal distribution with mean vector $m = (0.15, 0.10)$ and covariance matrix,*

$$\Sigma = \begin{bmatrix} 0.3 & -0.05 \\ -0.05 & 0.1 \end{bmatrix}.$$

Let A,B,C be three disjoint subsets of X such that $A \subset X_1$, $B \subset X_2$ and $C \subset X_3$. We consider the case that

- *All lines are profitable*

$$A = \{Y_{11}, Y_{21}\} \in [0.21, +\infty)^2$$

- *The sum is profitable, but one line is negative*

$$B = \{Y_{12}, Y_{22}\} \in \{(y_1, y_2) | y_1 < -0.21, y_2 > 0.24\}$$

- *The sum is not profitable*

$$C = \{Y_{13}, Y_{23}\} \in \{(y_1, y_2) | y_1 < -0.21, y_2 < 0.2\}.$$

In this case, $P_1(A) = 0.12$, $P_2(B) = 0.12$ and $P_3(C) = 0.12$.

We consider the case of a conservative manager. He would have a risk averse attitude giving lower weight to the first case than the second scenario. He decided to increase the probability of the second scenario by 20%*.*

$$\mu(A) = f(a,0,0) = 0.12 < 0.14 = f(0,b,0) = \mu(B)$$

The manager knows that $P(A \cup C) = 0.24$ *and* $P(B \cup C) = 0.24$*. However, he expects a highly volatile market in next future. Consequently, his decision is to increase* $P(A \cup C)$ *by* 20%*.*

$$\mu(A \cup C) = f(a,0,c) = 0.29 > 0.24 = f(0,b,c) = \mu(B \cup C)$$

A risk neutral agent would weight the three subsets with the same importance, approximately

$$\mu(A) = \mu(B) = \mu(C) = 0.12.$$

Finally, a decision-maker with risk appetite would consider more plausible the first subset and would increase the likelihood of scenario by 20%*, so*

$$\mu(A) = 0.14 > 0.12 = \mu(B)$$
$$\mu(A) = 0.14 > 0.12 = \mu(C).$$

5. Conclusions

In this paper we have introduced m-dimensional distorted probabilities on random vectors for continuous domains. We have given some results to characterize their construction and provided some examples. We have also related them to hierarchically decomposable fuzzy measures, and defined hierarchically decomposable fuzzy measures on continuous domains, as a subtype of m-dimensional distorted probabilities. Our definition is based on our previous definition for finite domains. Nevertheless, note that our characterization of the function f is only applicable to the continuous case. Otherwise, the domain and the range of f are finite.

The application of fuzzy measures to real problems needs families of measures that are easy to define. To this end, a large number of families exist on discrete domains but this is not the case for continuous ones. Distorted probabilities are probably the only such family.

In this work we have contributed with two families that naturally extend distorted probabilities and make the use of fuzzy measures easier. These results apply to fuzzy measures defined to pairs of (or in general, to the product of t) random variables.

As stated in the introduction the usefulness of our approach is partially based on the understanding that there are already parametric distributions on continuous domains and, therefore, they can be used to define fuzzy measures.

Future work includes further results on these types of m-dimensional distorted probability measures, the definition of other families of transformed-probability measures, and the extension to functions related to probability distributions such as the cumulative distribution function, the survival function or the characteristic function. We also aim at development of software to compute efficiently the Choquet integral of functions with respect to these families of measures, in the line of [13].

References

[1] Calvo, T., Kolesárová, A., Komorníková, M., Mesiar, R. (2002) Aggregation operators: properties, classes and construction methods, in T. Calvo, G. Mayor, R. Mesiar (eds.) Aggregation Operators, Physica-Verlag, 3-104.

[2] Denneberg, D. (1989) Distorted probabilities and insurance premiums. In: Proceedings of the 14th SOR, Ulm. Athenaüm, Frankfurt.

[3] Denneberg, D. (1994) Non Additive Measure and Integral, Kluwer Academic Publishers.

[4] Grabisch, M., (1996), *k*-order additive fuzzy measures, Proc. 6th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU), 1345-1350, Granada, Spain.

[5] Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2009) Aggregation Functions, Cambridge University Press, Encyclopedia of Mathematics and its Applications, No 127.

[6] Honda, A., Nakano, T., Okazaki, Y. (2002) Distortion of fuzzy measures, Proc. of the SCIS/ISIS conference.

[7] Klir, G. J., Yuan, B. (1995) Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall, UK

[8] Narukawa, Y., Torra, V. (2005) Fuzzy measure and probability distributions: distorted probabilities, IEEE Trans. on Fuzzy Systems 13:5 617-629.

[9] Sugeno, M., (1974), Theory of fuzzy integrals and its applications, Ph. D. Dissertation, Tokyo Institute of Technology, Tokyo, Japan.

[10] Torra, V., (1999), On hierarchically S-decomposable fuzzy measures, Int. J. of Intel. Systems, 14:9, 923-934.

[11] Torra, V. (2014) Distributions based on the Choquet integral and non-additive measures, RIMS Kokyuroku 1906 136-143.

[12] Torra, V., Narukawa, Y. (2007) Modeling decisions: information fusion and aggregation operators, Springer.

[13] Torra, V., Narukawa, Y. (2016) Numerical integration for the Choquet integral, Information Fusion 31 137-145.

[14] Torra, V., Narukawa, Y., Sugeno, M. (eds.) (2013) Non-additive measures: theory and applications, Springer.

[15] Wang, S. S., Young, V. R. (1998) Risk-adjusted credibility premiums using distorted probabilities, Scandinavian Actuarial Journal1998:2 143-165.

[16] Wang, Z., Klir, G.J. (1992) Fuzzy Measure Theory, Springer.

[17] Young, V. R., Wang, S. S. (1998) Updating non-additive measures with fuzzy information, Fuzzy sets and systems 94:3 355-366.