Table of Contents

List of Figures

List of Tables

Acknowledgment

First, I thank the members of the jury who honored me by judging my dissertation. I am grateful for the time they spent to do this.

There are so many people I would also like to thank, those who have really provided me with their support during the achievement of this work.

I would like to express my gratitude to all my supervisors : Dr. Boutheina Ben Yaghlene not only for believing in my abilities and supervising my first steps into to the research area of uncertainty, but also for her continuous guidance and her invaluable help and support. Pr. Arnaud Martin for hosting me in his laboratory, and also for answering my different questions when guiding and supporting me during the elaboration of this work. I would also like to thank Pierre-Emmanuel Doré for his wise remarks and suggestions,for his carefulness, and his patience when introducing me to the world of the continuous belief functions, I want to say how much I have appreciated working, talking with them, all along these months, hoping to continue my researches with these gentle and helpful people. I am really thankful for their for your continuous supervision during the preparation of this work.

I wish to thank also all members of my laboratory $LARODEC$ at the $IHECCarthage$ and all members of the laboratories $CORDIAL$ and $PILGRIM$ in the $universityofRennes1$ for their hospitality during my training in France.

Last but by not the least, I want to express my gratitude to my parents for their devotion and unreserved love, my sisters my brothers, for listening to my complaints and for their encouragements, my nieces for disturbing me and my friends and all the persons who love me. Without your support, this work would never have been achieved.

Résumé

Le travail de ce master de recherche traite un problème qui est très important qui est la caractérisation des relations existant entre les fonctions de croyances continues. La première approche élaborée au cours de travail est d'évaluer la similarité entre deux fonctions de croyance continues. La deuxième est de mesurer l'inclusion d'une fonction dans une autre. La première contribution a été de proposer une mesure de similarité basée sur le calcul d'une distance entre deux ou plusieurs de fonctions de croyance. Plus précisément, nous construisons ainsi des densités de probabilité introduites par des distributions normales ou exponentielles représentant des fonctions de croyances continues. Notre méthode mesurant ainsi ne distance entre ces distributions. Chaque distribution étant caractérisée par ses propres paramètres, nous avons étudié le comportement de cette mesure de similarité pour enfin se focaliser sur les éléments pouvant avoir un impact sur le comportement de notre distance. Notre deuxième proposition est une mesure d'inclusion entre croyances continues. Pour cela, nous proposons deux typologies : une inclusion stricte et une inclusion partielle. Nous avons implémenté nos deux approches en utilisant MATLAB pour effectuer les simulations. Ces dernières ont été une base essentielle pour déterminer le rôle de chaque paramètre d'une distribution agissant sur tout aussi bien la mesure de similarité que sur les deux types d'inclusion.

Abstract

This mater thesis concerns an important issue which is the characterization of relations between continuous belief functions. The first method elaborated during this work is how to evaluate the similarity between two continuous belief functions. The second one is about measuring the inclusion of a function in another. The first contribution is to propose a similarity measure based on the computing of a distance between two or several belief functions. More precisely, we build probability density functions induced by normal or exponential distributions representing continuous belief functions. Our method, measures a distance between these distributions. Our second approach is an inclusion measure between belief functions. To do so, we propose two types : the strict and the partial one.

Each distribution is characterized by its own parameters ; we have analyzed the behavior of this continuous distance, and the inclusion so that we were able to determine the elements that could have an impact on the behavior of our similarity and inclusion measures.

We have implemented our two approaches using MATLAB for the experimentations. These experimentations are a fundamental base to determine the role of each parameter of a distribution having an impact on both the similarity measure and the types of inclusion.

Notations

Ω	frame of discernment
ω	elements of the frame of discernment
n	number of elements of Ω
m	mass function Ω as frame of discernment
$F(m)$	set of focal elements X,Y
focal elements	
bel	credibility function defined in the space 2^{Θ}
pl	plausibility function defined in the space 2^{Θ}
q	commonality function defined in the space 2^{Θ}
$betP$	pignistic probability defined in the space 2^{Θ}
$\mid X \mid$	cardinality of X
s	specialization operator
g	generalization operator
$\sqsubseteq_s$	specialization ordering
$\sqsubseteq_{pl}$	plausibility ordering
$\sqsubseteq_q$	commonality ordering

$\emptyset$	empty set
$\overline{X}$	complement of X in the defined space of X
$\overline{\mathbb{R}}$	extended set of reals
$Betf$	pignistic probability density
$Biso(Bet)$	set of isopignistic probability densities
$\vee$	maximal value of two reals
$\wedge$	minimum value of two reals
μ	mean of a normal distribution
σ	standard variation of a normal distribution
θ	mean of an exponential distribution
IncStr	Strict inclusion
IncPar	Partial inclusion
$\delta_{IncStr}(f_1, f_2)$	Degree of strict inclusion
$\delta_{IncPar}(f_1, f_2)$	Degree of partial inclusion

Introduction

The field of Knowledge Management have known a successful expansion these last years. This phenomena is due to the various new ways of data collecting. Within the collected data we have to deal with imperfections that represent the real-world issues. These imperfections present a number of challenges in terms of collecting, modeling, representing and managing.

To overcome these limitations, many researches have been done to adapt standard techniques to this kind of environment. The idea was to introduce theories managing uncertainty and/or imprecision, such as probability theory, belief functions theory, fuzzy set theory and possibility theory.

Then, managing imperfect information became a well explored field in Artificial Intelligence. This kind of information have to be handled and included in information systems instead of rejecting or replacing them (which can be an easy solution).
Among the theories over listed, we are interested in the belief functions theory, which offers a natural and easy way to handle imperfections. It is considered as an appropriate way to let experts express their beliefs in a flexible and natural manner. However, imperfections might result from using unreliable information sources or aggregating opinions of different experts We treat a special type of imperfection, namely, uncertainty which is well modeled within the belief functions theory framework.

And more precisely, we are interested on the field of continuous belief function has been marginalized comparing to the discrete case. This situation can be explained by the complex nature of these functions. In order to be able to treat them, we need to use distributions to model the continuous beliefs which can be a tedious situation when managing continuous belief functions.

Once the modeling issue solves, how do we describe relations existing between continuous

belief function ? How to model them ?
Our aim, through this work, is to investigate and define some relations that can exist between continuous belief functions and their proprieties.

It is important to mention that some relations under uncertainty and/or imprecision have been proposed in the literature but all these works are different from what we propose in this dissertation because they treat only the discrete belief functions.

This dissertation is organized in three chapters :

In chapters 1 : Belief Functions Theory : An overview we will present the necessary background regarding the different concepts of the belief functions theory. Then, a more attention is given to the continuous belief functions, their representation, and the different approaches to model them in the literature. In chapters 2 and 3, we present our research propositions and an overview of the contributions.
Chapter 2 : A distance between continuous belief functions : This chapter examines an important concept in the belief functions theory which is the similarity. Within this chapter, we provides different ways of similarity measure in the evidential theory, and then we propose our first contribution which is an adaptation of the distance of Jousselme using the formalism of Smets for modeling continuous belief functions. The major results that we have developed in this part are published in (Attiaoui 2012).
Chapter 3 : The inclusion of continuous belief functions : Another natural way to define a relation between belief functions which is the measurement of inclusion. Two approaches are presented : A strict and a partial inclusion. This measure will be defined and its proprieties presented. Later we will show how similarity can be involved and play a role in defining the inclusion between continuous belief functions. The major results that we have developed in this part are published in
Two appendices complete this master thesis. The first one briefly recalls other related uncertainty theories such as probability theory, and possibility theory. The second one will present the behavior of the functions presented in Chapter 2 and Chapter 3

Finally, a general conclusion summarizes the major achievements of this thesis and presents possible future developments.

Chapter 1

Theory of belief functions : An Overview

Information is the resolution of uncertainty
Claude Shannon

In this chapter we will introduce the different forms of imperfect information. Then, like we are interested by the belief functions theory, we will start by a general historical introduction. Later, we will explain the most important notions within this theory by setting out the discrete belief functions and giving a representation of the continuous belief function and how are we able to model them.

1 Introduction

Real-world problems always involve imperfect informations, that must be considered in information systems in order to model reality. Many research area took in consideration these imperfections : Information fusion, multi-criteria decision making, medical diagnosis, military applications, nuclear energy...
As defined by Dubois and Prade (Dubois Prade 2006), an information is any collection of symbols or signs provided either by the observation of natural (or artificial) phenomena or by the human cognition. It is intended for the understanding of the real-world and for supporting decision making, etc
A piece of information could be objective, i.e., stemming from measurements or sensors (it is not dependent of the information source). It could also be subjective, i.e., expressing individuals opinions without turning to direct observation (it is dependent of the information source).

In both cases, the information could be imperfect : this can be a result of faulty reading instruments, unreliable sensors, lack of background knowledge of individuals leading to poor opinions, etc.
We will consider the Belief Function Theory, also known as the Dempster-Shafer theory or Evidential Theory. It is a general framework for reasoning under uncertainty.

By reasoning, we mean the way we will manipulate the available information in order to provide interesting and useful conclusions.
By under uncertainty, we mean that we are in presence of information that does not represent perfectly the real world, and that we are uncertain about the exact values or state.

The belief function theory is a well known tool for representing imperfect information that are uncertain and imprecise such as expressed by humans

In this chapter, we will exhibit the foundations of the evidential theory by defining among other basics, the frame of discernment, focal elements and the different belief functions. Moreover, we will introduce the principle of the least commitment and how to measure the specificity of the functions. After that, we will lay out how to estimate the reliability of a belief and proceed to the discounting. Also some combination rules will be set out like the conjunctive, disjunctive rules, the RCR... Finally we will focus on continuous belief functions by describing the different methods to model them by presenting the formalism of Smets. In the other sections, we present other approaches to define continuous belief functions.

2 Typology of imperfect information

When we are in presence of information, characterizing the real world, we are sure that they are imperfect. In this section we

2.1 Imprecise information

It is characterized by the information content. In other words, it is related to the information itself.
Let's consider, for example, the age of a person. We say "John's age is between 20 and 25".
Formally $age(Johon) \in \{20, 21, 22, 23, 24, 25\}$, represents an imprecise information. In this case, we are unable to determine the exact age of John.
Thus, when we deal with this kind of information, we are not able to fully grasp the real world situation.

2.2 Uncertain information

It is the result of a lack of information about the real world. It is related to the source providing the information.
Uncertainty corresponds to partial or total ignorance of knowledge. Otherwise, it describes incompletely the reality.

Any uncertain information has an uncertain scorer which can on the one hand be numeric (The probability that China will win more medals than the other countries during the Olympic Games in London 2012 is 0.7), and on the other hand symbolic or linguistic (I belief that John is 24 years old).

2.3 Inconsistent information

Where no value is compatible with the information. For example "John is single", and "John's wife is Marry". John can not be married and single at the same time. The conflict between the two informations can lead us to an inconsistent deduction

3 Basics on the theory belief functions

The belief function theory started with the work of A. Dempster (Dempster 1967,1968). The aim of his researches was to model mathematically information that can not be described by a precise probabilistic distribution. To do so, he developed the notions of the lower and upper probabilities framing the exact distribution. Using that, he was able to represent more precisely the observed data.
Later, in his book "A mathematical Theory of evidence" (Shafer 1976), the author presented the information defined by an expert, where basic belief assignments have two functions : a credibility and a plausibility function corresponding respectively to the lower and upper probabilities of Dempster.
The theory was further developed by Smets (Smets1990a, Smets 1994), who proposed the Transferable Belief Model (TBM). This model presents a pignistic probability induced by a belief function which is built by defining a uniform probability from each positive mass. Moreover, in terms of upper and lower probabilities, it can be considered as the center of gravity of the set of probabilities dominating the belief functions. He also introduced new tools for information fusion and decision making.
The objetcive of the evidential theory is to represent information transmitted by a source concerning an event. A belief function must take in consideration all the possible events on which a source can describe a belief. Based on that, we can define the frame of discernment

Definition 1.1 Frame of discernment
The frame of discernment is a finite set of disjoint elements noted Ω *where* $\Omega = \{\omega_1, ..., \omega_n\}$*. This theory allows us to affect a mass on a set of hypotheses not only a singleton like in the probabilistic theory. Thus, we are able to represent ignorance, imprecision...*

Definition 1.2 Basic belief assignment (*bba*)
A bba is defined on the set of all subsets of Ω*, named power set and noted* 2^Ω*. It affects a real value from* $[0, 1]$ *to every subset of* 2^Ω *reflecting sources amount of belief on this subset. A bba* m *verifies :*

$$\sum_{X \subseteq \Omega} m(X) = 1. \tag{1.1}$$

A very common assumption advocates the existence of a closed world. In other words, all the possibilities are represented in Ω, and defined as :

$$m(\emptyset) = 0. \tag{1.2}$$

On the contrary, if we accept the case that, there existence any other possibility that is unrepresented in Ω, we have :

$$m(\emptyset) > 0. \tag{1.3}$$

Here, Smets supposed that we are in an open world, which means that decisions are not exhaustive. This was introduced by Smets.

Definition 1.3 Focal elements *Considering a set A in 2^{Ω} is a focal element with a mass m if an elementary mass is positive $m(A) > 0$. The set of focal elements of m is noted $\mathbb{F}(m)$.*

We can consider some special cases of mass functions :

- **Categorical mass function** : the function has only one focal element A. If this element is Ω, the frame of discernment, we can state that the function is empty and we are representing the total ignorance.
- **Non dogmatic mass function** : the function has a non null mass on the ignorance Ω.
- **Baysian mass function** : where focal elements are disjoints by pairs, it can be considered as a probability distribution.
- **Consonant mass function** : a function that has nested focal elements.

3.1 Belief Functions

Once, the masses defined, we can build different belief functions. the most famous one is the $credibility function.$

Credibility : measures the strength of the evidence in favor of a set of propositions for all $X \in 2^{\Omega}$:

$$bel(X) = \sum_{Y \subseteq X, Y \neq \emptyset} m(Y). \tag{1.4}$$

.

The credibility is interpreted as a degree of justified support given to proposition X by the available evidence.

Plausibility : quantifies the maximum amount of belief that could be given to a X of the frame of discernment for all $X \in 2^{\Omega}$:

$$pl(X) = \sum_{Y \in 2^{\Omega}, Y \cap X \neq \emptyset} m(Y). \tag{1.5}$$

The plausibility function is a measure of the maximum potential support that could be given to a subset X of the universe of discourse, of course if further evidence became available. In other words, it contains those parts of belief that do not contradict X.

The plausibility is considered as the dual function of the credibility :

$$pl(X) = 1 - bel(\overline{X}) \tag{1.6}$$

Commonality : measures the set of *bbas* affected to the focal elements included in the studied set, for all $X \in 2^{\Omega}$:

$$q(X) = \sum_{Y \supseteq X} m(Y). \tag{1.7}$$

Shafer showed for consonant belief function these different proposals : If m is a consonant belief function then :

1. $bel(A \cap B) = min(bel(A), bel(B), \forall A, B \subseteq \Omega$
2. $pl(A \cup B) = max(pl(A), p(B), \forall A, B \subseteq \Omega$
3. $pl(A) = max_{\omega \in A} pl(\omega), \forall$ non empty $A \subseteq \Omega$
4. $q(A) = min_{\omega \in A} pl(\omega) = min_{\omega \in A} q(\omega), \forall$ non empty $A \subseteq \Omega$

Decision making in belief functions theory

In (Smets and Kennes 1994), the authors introduced the Transferable Belief Model (TBM). They afforded new method to to represent, manipulate and combine information from different sources. The TBM is based on two levels :

– **The credal level** (*in Latin credo : I belief*), where beliefs are studied and combined using belief functions in order to preserve as much information as possible during the combination aiming at decision making.

- **The pignistic level** (*in latin pignus : a bet*), where beliefs are used to make decisions and represented by probability functions called the **pignistic probabilities**

When choosing the maximum of the credibility might be pessimistic, on the other hand decision made with the maximum of plausibility can be considered too optimistic. Based on that, the best solution is to use the **pignistic probability** choose the credal level, where for all $X \in 2^{\Omega}$, with $X \neq 0$:

$$betP(X) = \sum_{Y \in 2^{\Omega}, Y \neq 0} \frac{\mid X \cap Y \mid m(Y)}{\mid Y \mid 1 - m(\emptyset)}. \tag{1.8}$$

3.2 Combination rules

Belief functions are used to represent information provided by different sources, it is natural than to combine them for an issue of decision making.
Many combination rules have been proposed taking in consideration the nature of the sources.

Dempster's combination rule

The first one was proposed by Dempster in 1967 which is a conjunctive normalized combination rule also called $theorthogonalsum$
Given two mass functions m_1 and m_2, for all $X \in 2^{\Omega}$, $X \neq \emptyset$ the Dempster's rule is defined by :

$$m_D(X) = \frac{1}{1-k} \sum_{Y_1 \cup Y_2 = X} m_1(Y_1)m_2(Y_2) \tag{1.9}$$

where $k = \sum_{Y_1 \cap Y_2 = \emptyset} m_1(Y_1)m_2(Y_2)$ is the inconstancy of the fusion (or of the combination) can also be called the conflict or $globalconflict$.
$1 - k$ is the normalization factor of the combination in a closed world.

The conjunctive combination rule

In order to consider the issues of the open world, like introduced in (Smets 1990), where the author proposed the conjunctive combination rule. Considering two mass functions m_1 and m_2, for all $X \in 2^{\Theta}$ defined by :

$$m_{conj}(X) = \sum_{Y_1 \cap Y_2 = X} m_1(Y_1) m_2(Y_2) \tag{1.10}$$

We can note $m_{conj} = m_1 \oplus m_2$, and consider $k = m_{conj}(\emptyset)$ as an unexpected solution.
The operator $\oplus$ is associative and commutative but not idempotent.
This conjunctive combination rule can also be expressed by the relation of commonality for an easier implementation.

$$q_{1\oplus 2}(X) = q_1(Y_1) q_2(Y_2) \tag{1.11}$$

This combination rule can be extended to N mass functions m_i. We obtain $\oplus_{i\in[1,N]} m_i$, for all $X \subseteq \Omega$:

$$\oplus_{i\in[1,N]}\, m(X) = \sum_{Y_1 \cap \ldots \cap Y_N = X} \prod_{i\in[1,N]} m_i(Y_i) \tag{1.12}$$

This combination rule, not normalized, allows us to consolidate agreeing sources and discount the others, it also affects a mass on the empty set. This mass can be explained by the following factors :

- Disagreement between different sources of information, leading to the conflict
- Combining a source to itself (the conjunctive combination rule is non idempotent) : the auto conflict introduced by *Martin and al.* (Martin 2008)
- Inconstancy of the belief function in other terms combining non independent information sources

The disjunctive combination rule

First introduced by (Dubois and Prade 1986), then by (Smets 1993), we can considering two basic belief assignments m_1 and m_2, after proceeding to a disjunctive combination expressed like follows for all $X \subseteq \Omega$:

$$m_{disj}(X) = \sum_{Y_1 \cap Y_2 = X} m_1(Y_1) m_2(Y_2) \tag{1.13}$$

The disjunctive combination rule can be used when one of the sources is reliable or when we have no knowledge about their reliability.

Robust Combination Rules

Florea and al., (Florea 2009), presented the "Robust Combination Rules"(RCR). They are called robust, because of their abilities, to adapt to unreliable sources of information even when corresponding degree of reliability is unknown.
The RCR, are particular weighted sum of the conjunctive and disjunctive rules with weighting coefficients which are conflict functions.
They are defined by an adaptative weighting between conjunctive and disjunctive rules. The RCR act like the conjunctive rule when the conflict is low (k close to 0) or when the sources are both reliable. Meanwhile, in the presence of high conflict (k close to 1) or when at least one of the sources is unreliable, such as in this case, the RCR behaves like the disjunctive rule.
Let m_1 and m_2 two *bbds* defined in Ω, the RCR, between those two masses is defined such as for all $X \subseteq \Omega$, $X \neq \emptyset$:

$$m_{RCR}(X) = \alpha(k)m_{conj}(X) + \beta(k)m_{disj}(X) \tag{1.14}$$

Where m_{conj} is the conjunctive combination rule, and m_{disj} is the disjunctive one and $m_{RCR}(\emptyset) = 0$. α and β are functions of conflict $k = m_{conj}(\emptyset)$, they at the same time must satisfy :

$$\alpha(k) \sum_{X \subset \Omega} m_{conj}(X) + \beta(k) \sum_{X \subset \Omega} m_{disj}(X) = 1 \tag{1.15}$$

The equation (1.24) can also be expressed by :

$$\alpha(k) + (1 - k)\beta(k) = 1 \tag{1.16}$$

when $k = 0$, the sources are in total agreement. Otherwise when $k = 1$, they are in total conflict.

3.3 Belief functions' ordering

In order to define a principle that will allow us to choose one belief function over an other in a set of many functions, when we are in presence of associated belief function. We need to classify them. Thus, we define a relation that will specify a standard criterion that can ease this decision. Hence, in the evidential framework, some operators for specialization and generalization operators have been defined. Then a principle of least commitment based on that,shows according to which pattern one can choose the belief function needed.

Specialization and generalization of a belief function

Specialization and generalization operators transform belief functions making them more (respectively less) specified by using a mass transfer. Once one of these operations made, it becomes easier to choose the most (or less) committed belief function.

Specialization

The specification operator s is an application of $2^{\Omega} * 2^{\Omega}$ in $[0, 1]$ that verifies for all A and X in Ω :

$$(A \nsubseteq X) \Rightarrow s(A, X) = 0 \tag{1.17}$$

$$\sum_{A \subseteq \Omega} \sum_{X \subseteq \Omega} s(A, X) = 1 \tag{1.18}$$

We say that m_2 is a specialization of m_1 if and only if it exits a specialization operator that for all A in Ω

$$m_2(A) = \sum_{X \subseteq \Omega} s(A, X) m_1(X) \tag{1.19}$$

A specialization operator distributes the mass affected by a mass function on a set of its subsets. Let's consider $\mathcal{S}_s(m)$ the set of belief functions resulting from the specialization s of m.

Generalization

The generalization operator s is an application of $2^{\Omega} * 2^{\Omega}$ in $[0, 1]$ that verifies for all A and X in Ω :

$$(A \not\subseteq X) \Rightarrow g(A, X) = 0 \tag{1.20}$$

$$\sum_{A \subseteq \Omega} \sum_{X \subseteq \Omega} g(A, X) = 1 \tag{1.21}$$

We say that m_2 is a generalization of m_1 if and only if it exits a generalization operator that for all A in Ω

$$m_2(A) = \sum_{X \subseteq \Omega} g(A, X) m_1(X) \tag{1.22}$$

A generalization operator distributes the affected mass of a function to their oversets.

Specialization operator

Let's consider the three solutions proposed in (Dubois and Prade 1986, with the aim of ordering belief functions. So as to characterize their commitment, we suppose the existence of an ordering relation between them.

The first one, is associated to the specialization, if m_2 is a specialization of m_1, noted :

$$m_1 \sqsubseteq_s m_2 \tag{1.23}$$

It is also possible to represent this ordering operation using the plausibility (respectively the commonality) : we say if for all $X \subset \Omega$, the plausibility (respectively the commonality) of a belief function is lower than an other, that it is less committed in plausibility (respectively the commonality) noted :

$$(\vee X \subseteq \Omega pl_1(X) \leq pl_2(X)) \Leftrightarrow (m_1 \sqsubseteq_{pl} m_2) \tag{1.24}$$

$$(\vee X \subseteq \Omega q_1(X) \leq q_2(X)) \Leftrightarrow (m_1 \sqsubseteq_q m_2) \tag{1.25}$$

Denoeux, shows in (Denoeux 2008), that among these three ordering relations, $\sqsubseteq_s$ is the strongest one, because it integrates both of the plausibility and the commonality ordering.

Otherwise, the most used one is the third one $\sqsubseteq_q$. In (Dubois, Prade and Smets 2001), the more amount of belief remains unassigned, i.e. the bigger the focal elements having a high mass assignment, the higher the commonality degrees and the less informative is the belief function.

Principle of least commitment : PLC

In (Dubois and Prade 1986), this principle was introduced for the evidential reasoning. This principle is to belief function, what the maximum entropy principle is to the probabilistic theory.

It allows us to construct a partial ordering relation on the set of belief functions.
In a set of equally justified bodies of evidences, we can use the PLC presented in (Hsia 91), Hsia, shows how this principle supports the idea of choosing the belief function that involves the least an expert.
It can be considered as a natural approach to select the less committed *bba* from a subset. In that way no one gives more support than justified to any subset of Ω.

3.4 Discounting

In the belief function framework, knowledge about the reliability of a source of information (or sensor) is achieved by the discounting operation, which transforms each belief function provided by a source into a weaker, less informative one. The discounting operation is controlled by a discount rate in taking values in $[0, 1]$: if $\alpha = 0$, the belief function is unchanged ;
if $\alpha = 1$, the belief function is transformed into the vacuous belief function. This transformation means that the information provided by the sensor is completely discarded.
Smets in (Smets 1993), shows that the discounting operation is not *adhoc*, it can be derived from a simple model of sensor reliability. In this model, the sensor can be in two states : reliable or not. In the first case, when we know that the sensor is reliable, the belief function which is provided is accepted without any modification, otherwise when we know that it is not reliable, the information is considered as irrelevant.

$$\begin{cases} \mathrm{m}^{\alpha}(X) = (1-\alpha)m(X), & \text{if X} \subset \Omega \\ \mathrm{m}^{\alpha}(X) = (1-\alpha)m(X) + \alpha, & \text{if } X = \Omega. \end{cases} \tag{1.26}$$

4 Continuous belief functions

In the first part of this chapter, we have presented different specification of discrete belief function. Unfortunately, these functions do not allow us to manipulate continuous data that can be provided by sensors in different areas like : search and rescue problems (Dor'2010) , classification issues, information fusion...
Some researches were interested in representing belief functions in continuous frame of discernment like Strat and Smets.
In these sections, we will present how to extend these functions on the real numbers. To do so, we will focus on Smets' approach to represent continuous belief functions by using probability densities...

The principle purpose in section 4.1 is to provide some methods that will allow us to have the needed characteristics in order to build and manipulate belief functions on real numbers. Section 4.4 will present the basic belief densities associated to pignistic probability. Later, we will introduce a graphical representation of these function proposed by Vannobel following the formalism of Smets. Finally, other approaches that have been provided in the literature to characterize continuous belief functions.

4.1 Continuous belief function on real numbers

Discrete belief functions, presented in the last chapter can not handle information on the set of real numbers.
Strat in (Strat 1984), and Smets (Smets 2005) proposed a definition of continuous belief functions. Strat, proposed to represent a mass function of a discretized variable in a triangular matrix. This representation reduces the number of the propositions and simplifies the computation of the intersections by focusing only on the contiguous intervals.
But unfortunately, Strat unlike Smets, did not have the theoretical background proposed on the Transferable Belief Model (TBM).
Smets based on the TBM's background, used the same representation than Strat, and proposed the belief functions in the extended set of reals noted $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.
However, using the belief function framework to model information in a continuous frame is not an easy task mainly to the complex nature of the focal elements. Comparing to the discreet domain, on real numbers, in (Smets 2005) *bba* becomes *basic belief densities* (*bbd*) defined on an interval $[a, b]$ of $\mathbb{R}$.

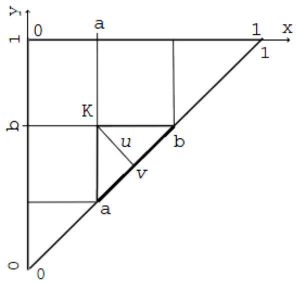

Figure 1.1 – Smets : The interval $[a, b]$ representation.

In this section, we will present an extension of these belief functions to intervals on $\mathbb{R}$.

Definition 1.4 Frame of discernment
Formally, in the case of continuous belief functions, the domain is limited to a spacial subset of power set defined as a Borel sigma-algebra generated by the collection of closed intervals on the real numbers, which is the domain of the functions.

Definition 1.5 The intervals
To model the set of intervals $\mathbb{I}$ *on* $\overline{\mathbb{R}}$ *contains the classical intervals of* $\mathbb{R}$ *among which* $\emptyset$ *; and the intervals* $[-\infty, y], [x, \infty]$ *and* $[-\infty, \infty]$ *are considered. Note that* $[x, y] = \emptyset$ *; whenever* $x > y$*.*

Definition 1.6 Probability density function
Comparing a mass density on $\overline{\mathbb{R}}$ *to a probability density function pdf on all the pairs* (x, y)*. Thus, the point* K *which is defined by with the pair of points* (a, b) *represents the interval* $[a, b]$ *like shown in Figure 1.1. This pair is defined by* u *and* v *where*

- *u is the distance from (a, b) to the perpendicular projection of (a, b) on the diagonal that contains the intervals*
- *v is the coordinate of this projection along this diagonal*

Definition 1.7 Basic belief densities
A generalization of the classical bba into a basic belief density (bbd) noted m^{I} *on the interval* I.*He defined the bbd where all focal elements are closed intervals or* $\emptyset$*.*
Given a normalized bbd m^{I}*, Smets defined an other function* f *on* $\mathbb{R}^2$*, where :*

$$\begin{cases} f(a,b) = m^{\mathrm{I}}([a,b]), & a \leq b\,; \\ f(a,b) = 0, & a > b. \end{cases}$$

f is called a probability density function (*pdf*) on $\mathbb{R}^2$.

4.2 Belief functions

To model the set of intervals $\mathbb{I}$ on $\overline{\mathbb{R}}$ contains the classical intervals of $\mathbb{R}$ among which $\emptyset$; and the intervals $[-\infty, y], [x, \infty]$ and $[-\infty, \infty]$ are considered. Note that $[x, y] = \emptyset$; whenever $x > y$.

By analogy to discrete belief functions, we have :

– Credibility :

$$bel^{\overline{\mathbb{R}}}([a,b]) = \int_{x=a}^{x=b} \int_{y=x}^{y=b} m^{\overline{\mathbb{R}}}([x,y])dydx. \tag{1.27}$$

– Plausibility :

$$pl^{\overline{\mathbb{R}}}([a,b]) = \int_{x=-\infty}^{x=b} \int_{y=max(a,x)}^{y=+\infty} m^{\overline{\mathbb{R}}}([x,y])dydx. \tag{1.28}$$

– Commonality :

$$q^{\overline{\mathbb{R}}}([a,b]) = \int_{x=-\infty}^{x=a} \int_{y=b}^{y=+\infty} m^{\overline{\mathbb{R}}}([x,y])dydx. \tag{1.29}$$

4.3 Combination of continuous belief functions

To combine two continuous belief functions, there exists many combination rules, the most famous and used one is the conjunctive combination rule presented in the last section but transposed to the continuous domain.
Given two bbd $m_1^{\overline{\mathbb{R}}}, m_2^{\overline{\mathbb{R}}}$ induced by two distinct pieces of evidence. For the conjunctive rule of combination, the product $m_1^{\Re}([a_1, b_1)], m_2^{\overline{\mathbb{R}}}([a_2, b_2)]$ is allocated to the interval $[a_1, b_1] \cap [a_2, b_2] = [a_1 \vee a_2, b_1 \wedge b_2]$ which may be empty.

$$\begin{aligned} m_{1\oplus 2}^{\overline{\mathbb{R}}} = & \int_{x=-\infty}^{a} \int_{y=b}^{y=+\infty} m_1^{\overline{\mathbb{R}}}([x,b]) m_2^{\overline{\mathbb{R}}}([a,y])dxdy \\ & + \int_{x=-\infty}^{a} \int_{y=b}^{y=\infty} m_1^{\overline{\mathbb{R}}}([a,y]) m_2^{\overline{\mathbb{R}}}([x,b])dxdy \\ & + m_1^{\overline{\mathbb{R}}}([a,b]) \int_{x=-\infty}^{a} \int_{y=b}^{y=\infty} m_2^{\overline{\mathbb{R}}}([x,y])dxdy \\ & + m_2^{\overline{\mathbb{R}}}([a,b]) \int_{x=-\infty}^{a} \int_{y=b}^{y=\infty} m_2^{\overline{\mathbb{R}}}([x,y])dxdy \end{aligned} \tag{1.30}$$

or equivalently, this combination can be expressed with the communality function :

$$q_{1\oplus 2}([a,b]) = q_1([a,b])q_2([a,b]) \tag{1.31}$$

4.4 Belief function associated to a pignistic probability

As we stated in the previous sections, when a decision must be made, according to Smets, we must build a probability function using the pignistic transformation. For a basic belief density *bbd* related to its density function f, the relation for the pignistic probability function becomes for $a < b$:

$$BetP([a,b]) = \int_{x=-\infty}^{x=\infty} \int_{y=x}^{y=\infty} \frac{y \wedge b - x \vee a}{y - x} f([x,y]) dy dx \tag{1.32}$$

The notations $\vee$ represents a function that gives the maximal of two reals, and $\wedge$ reflects the minimum of these reals.
The pignistic density transformation leads to a pignistic density $Betf$ which is defined by for each interval [a,b] in $\mathbb{R}$:

$$BetP([a,b]) = \int_a^b Betf(x)dx \tag{1.33}$$

Belief functions having a prescribed pignistic probability are called isopignistic where belief densities are having the same $Betf$. This set of isopignistic densities is denoted : $BIso(BetP)$ Within the belief functions theory, we need a method to build a density from the pignistic density. In this case, we apply the least commitment principle. As we have already said, it suggests to favour among all the isopignistic belief densities, the one that maximizes the commonality function. The q-least committed belief function is a consonant belief density, Which means that each focal interval is contained by the following one.

Smets shows that this q-least committed density generated from the pignistic transformation is a graphe having a bell shape form like shown in Figure 1.2 :

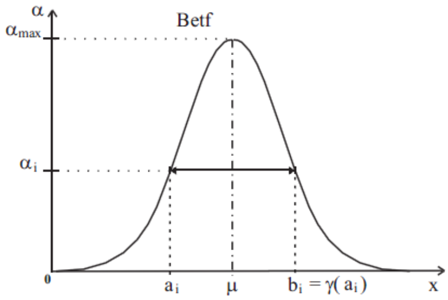

Figure 1.2 – Smets : A Bell Shape distribution.

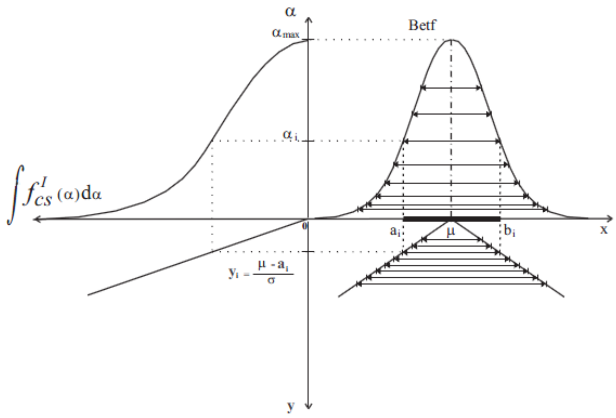

Figure 1.3 – Vannobel : Index function for continuous belief functions.

4.5 Graphical presentation of continuous belief functions

In (Vannobel 2012), proposed a graphical representation of the continuous belief functions using an index function in Smets formalism.

Let's consider $betf$ a probability density associated to a bell shape distribution like presented in Figure **??** The focal intervals $[a_i, b_i]$ are defined by *alphacuts* of the *pdf* such as :

$$betf(a_i) = betf(b_i) = \alpha_i \tag{1.34}$$

Vannobel,(Vannobel 2010, 2012) considers the horizontal cuts of the $betf_i$ as consonant $pdfs$ $betf_i$ in case of nested intervals. Thus, he can define a single *bbd*'s expression for a hole family of *pdf*. In the case of symmetrical $pdfs$ like Gaussian, focal intervals are defined such as $m(X) \neq 0$, and can be labeled by an index z such that $A^z = [A^{z-}, A^{z+}]$ with :

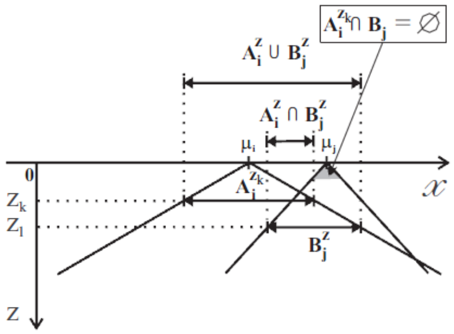

Figure 1.4 – Vannobel :Intersection and union of focal intervals resulting from two Gaussian *pdf*.

$$z = \frac{\mid x - \mu \mid}{\sigma}, z \in \mathbb{R}^+ \tag{1.35}$$

$$A^{z-} = \mu - \sigma z, A^{z-} \in [\Omega^-, \mu] \tag{1.36}$$

$$A^{z+} = \mu - \sigma z, A^{z+} \in [\mu, \Omega^+] \tag{1.37}$$

The Figure 1.4, illustrates the domain representing the focal intervals corresponding to a *pdf* $Betf$, ordered according to the label z

4.6 Continuous belief function using a multivalued mapping

Nguyen introduced in (Nguyen 1978) the notions of a source constituted by a probability space and a multivalued mapping which is able to define the lower probability. Let's consider (X,A), (S,B), ($P(S)$,B) as a measurable space, $P(S)$ the collection of all subsets of the set S and a multivaluated mapping where

$$\Gamma : X \rightarrow P(S) \tag{1.38}$$

This function Γ can hold at the same time two notions : On one hand, it defines both of the lower and upper probability, on the other hand, considers random sets. We can say that Γ as a multivalued mapping is strongly measurable with respect to the spaces (A,B).

Nguyen, considers a source (X, A, P),$\Gamma : X \rightarrow P(S)$, and while assuming the propriety of strong measurability presented above, he can propose that the lower probability measure P_* is deduced from the probability distribution of Γ considered as a random set.
Moreover, he supposes that Γ is a measurable mapping, then it is a random set by specifying its probability distribution . Thus the probability distribution of a random set Γ is precisely the basic probability assignment.
We say that there is a correspondence established between belief functions on a source S and the probability distribution of random sets. This relation can be expressed by if we want to construct we only need to construct its density on $P(S)$ which is already defined .

Doré *et al.* in (Doré 2011) proposed a similar approach founded on an index function that can be assumed as Γ. This function can describe the set of focal elements of a continuous belief function. In this case, every index has its own probability measure where there is an allocated weight to a set of focal elements using a credal measure.

Smets formalism takes into consideration only to closed intervals, in (Dor'2010) extended his work ti continuous belief functions where focal elements are not represented by intervals. He uses $\alpha cuts$ to measure to area of the portions of multimodal distributions.
In order to build these continuous belief functions he defined two cases involving the source of information :

- **Objective source** : usually intended to give information in a probabilistic frame, applying then the principle of maximal necessity. (Dubois 2001, 2004) suggests to use the most committed possibility distribution checking two conditions : The first one needs a possibility value higher than the probability ($\Pi(X) \geq P(X)$). The second condition, claims that the same ordering between two events : $P(A) \geq P(A') \Rightarrow \Pi(A) \geq \Pi(A')$. See Appendix A for the relation between the belief functions theory and the possibility theory.
- **Subjective source** : An expert models a phenomena using a probability, Doré uses this transformation to associate to a probability a belief function. Supposing that an expert's opinion can be subjective

5 Conclusion

We started this chapter by clarifying the difference between imprecision and uncertainty by giving some definitions and examples. Then, we have introduced many important concepts relative to the belief functions theory. We have presented the necessary theoretical background, by showing its strength in modeling imprecise, uncertain or incomplete information. Dealing with all these imperfections no matter how many sources we combine is an efficient manner to provide the best information in the process of decision making. However, discrete belief functions are unable to express continuous phenomena, so for that reason, continuous belief functions were introduced. We have presented the formalism of Smets for representing continuous belief functions on the extended set of reals. We also showed how model the basic belief densities within this chapter. Based on this description, we are now able to define continuous belief functions. In the next chapter, we will consider belief functions on the set of extended reals of Smets to build our continuous belief functions and develop our approaches.

Chapter 2

Similarity between continuous belief functions

Knowing ignorance is strength. Ignoring knowledge is sickness
$LaoTse(500B.C.)$

Based on the formalism of Smets, we will use this method to build our continuous belief functions. In this chapter we will introduce a distance to measure the similarity between two probability density functions. Different proprieties will be presented that this measure must satisfy. Moreover, we will make a comparison between our proposal and a distance computed using a classical scalar product. We will show that we new way to measure the similarity measure between continuous belief functions based induced by normal and exponential distributions.

1 Introduction

The concept of similarity is considered an attractive topic for many research fields and the comparison of objects represents a fundamental task in many real-world application areas such as, decision making, medicine, meteorology, psychology, molecular biology, data mining, case-based reasoning, etc.
Obviously, comparing these pieces of information could be of a great interest and should be accomplished effectively in order to solve many problems in the presence of imperfect data.
Terms such as closeness, affinity, resemblance, etc. can be considered as expressing the degree of similarity between objects or pieces of information.
Terms such as distance and divergence can be thought of as expressing dissimilarity. We can find many of these words in the literature and generally, they are expressing the same thing with a possibly slight semantic differences.
This chapter is organized as follows : Section 2 presents the different views of similarity, some are based on a divergence measure and some use a distance. Moreover it presents the properties that a distance should satisfy. After showing that existing similarity measure are not entirely able to measure a distance between continuous belief functions represented by probability density functions , we propose, in Section 3 , a distance that take into account the nature of these functions. We illustrate our proposed method and apply it to two types of distributions : A normal and an exponential distributions in section 5.

2 Similarity Vs Distance

2.1 The divergence of Kullback-lieber

The divergence of Kullback-lieber (or DKL) was introduced in (Kullback 1978) "Information Theory and Statistics" is a non-symmetric measure of the difference between two probability distributions P and Q ; Typically P represents the "true" distribution of data, observations, or a precisely calculated theoretical distribution. The measure Q typically represents a theory, model, description, or approximation of P.
The K-L divergence is not a true metric for example, it is not symmetric. Generally $D_{KL}(P\|Q) = D_{KL}(Q\|P)$.
For the discrete case, KullbackLeibler divergence is defined as follows :

$$D_{KL}(P \parallel Q) = \sum_i P(i) Ln \frac{p(i)}{q(i)} \tag{2.1}$$

It represents the average of the logarithmic difference between the probabilities P and Q, where the average is taken using the probabilities P. The equation (**??**) is accurate only if $P(i) > 0$, and $Q(i) > 0$.

For the continuous case, the KullbackLeibler divergence is represented as :

$$D_{KL}(P \parallel Q) = \int_{-\infty}^{+\infty} p(x) Ln \frac{p(x)}{q(x)} dx \tag{2.2}$$

Let's consider $X \mapsto N_1(\mu_x, \sigma_x)$ and $Y \mapsto N_2(\mu_y, \sigma_y)$ where N_1 and N_2 are normal distributions. In order to measure the divergence between $N_1(\mu_x, \sigma_x)$, we use the following equation [1] :

$$D_{KL}(X \parallel Y) = \frac{(\mu_1 - \mu_2)}{2\sigma_2} + \frac{1}{2}(\frac{\sigma_1^2}{\sigma_2^2} - 1 - Log_2(\frac{\sigma_1^2}{\sigma_2^2})) \tag{2.3}$$

This DKL satisfies the following properties :

P1. Non negativity $D_{KL}(P \parallel Q) \geq 0$

$D_{KL}(P \parallel Q) = 0$, if and only if $P = Q$, almost every time.

P2. Additivity $D_{KL}(P \parallel Q) = D_{KL}(P_1 \parallel Q_1) + D_{KL}(P_2 \parallel Q_2)$

2.2 Distances

Recently, distances between belief functions, have intrusted a large number of researches, in order to quantify interactions between them. Within the evidential theory, several approaches have been defined like Ristic and Smets who proposed a distance based on Dempster's conflict factor. Other researchers, proposed geometrical (Euclidian) distances, like Fixsen and Mahler, Jousselme and al. In the technical literature, two main areas can be distinguished regarding the use of distances according to (Jousselme 2012).

The first one is for algorithm evaluation and optimization like classification algorithms, belief functions approximation, and also for some properties estimation for combination rules. The second use of the distances is to refer to an agreement between information sources essentially

1. http ://www.allisons.org/ll/MML/KL/Normal/

for clustering, discounting...
In our case, a distance will be considered as a measure of similarity like introduced as the agreement.

In order to quantify how much two or more objects are different, we use a distance to measure the similarity between the belief functions.

The measure of Jaccard is a matric whose elements are considered, for all $A, B \in 2^{\Omega} \backslash \emptyset$

$$J(A,B) = \frac{\mid A \cap B \mid}{A \cup B} \tag{2.6}$$

In (Diaz 2006), the authors proposed to replace $J(A,B)$ by an other function of similarity S where the matrix becomes :

$$D(A,B) = \frac{2 \mid A \cap B \mid}{\mid A \mid + \mid B \mid} \tag{2.7}$$

where $\mid A \mid$ and $\mid B \mid$ denote respectively the cardinalities of A and B.
Otherwise, in (Diaz 2006), the author showed a distance between two *bbas*, in order to measure the similarity between them, has to consider the resemblance of the mass distributions of the focal elements and also the proximity of the set being compared to the frame of discernment itself including the ignorance.

Perry and Stephanou presented in (Perry and Stephanou 1991) a distance that measures the difference between the information present during the combination. Here, they see how closely an evidence matches a pre-defined prototype. The distance between two belief functions is measured by the difference in the amount available when considered separately and when they combined.
According to them, the similarity between two focal elements represents the degree to which the two focal elements have attributes in common which would tend to make them indistinguishable. They defined the similarity between A and B as :

$$Sim(A,B) = \frac{\{A \cap B\}}{\{A \cup B\}} \tag{2.8}$$

Once the similarity $Sim(A,B)$, they determined a distance that measure the divergence where it takes into consideration the information change when combining the belief functions.

Their work for classifying sensors ' information is based on three elements which are (1) the combination of the belief function using the Dempster's rule. (2) a degree of prototype that typifies a class, and finally (3) a belief measure of divergence.

2.3 Basic proprieties of a distance

Here we will namely consider a distance as a similarity measure.

In our case, comparing uncertain pieces of information comes down to comparing continuous belief functions represented by probability density functions. Hence, we need a measure to quantify the amount of similarity two or several belief functions represented by their distributions.
An important question must be considered : what are natural properties that such measures should satisfy ?

Let's consider f_1 and f_2 two distributions on $\overline{\mathbb{R}}$. A belief similarity measure or a distance denoted $d(f_1, f_2)$ should satisfy the following basic properties :

Property 1. Non-negativity

$d(f_1, f_2) \geq 0$

Namely, a distance (or similarity measure) between two distributions must never be negative.

Property 2. Symmetry

$d(f_1, f_2) = d(f_2, f_2)$

The distance between f_1,f_2 must be strictly the same.
Property 2 can be extended to the reflexivity $d(f_1, f_1) = 0$

Property 3. Upper bound

$\forall f_i, d(f_i, f_i) = 1$

This property implies a full dissymmetry, a maximal contradictory between the distributions

Property 4. Lower bound and Non-degeneracy

$\forall f_i, d(f_i, f_i) = 0$

This property implies a total similarity between the distributions.

Property 5. Triangle inequality
$d(f_1, f_2) \leq d(f_1, f_3) + d(f_2, f_3)$

2.4 Scalar product

On a defined set, a distance can be measured with a scalar product between two vectors f and g. A scalar product has a symmetric and a bilinear form, and is considered positive when having the following proprieties :

P1. Symmetry $\langle f, g\rangle = \langle g, f\rangle$

P2. Non negativity $\langle f, f\rangle > 0$

P3. the zero vector $\langle f, f\rangle = 0 \Rightarrow f = \overrightarrow{0}$

We have $\langle f, f\rangle = \|f\|^2$ which is the square norm of f.

To measure a distance using a scalar product, we use the following expression :

$$d_{SP}(f, g) = \sqrt{\frac{1}{2}(\|f\|^2 + \|g\|^2 - 2\langle f, g\rangle))} \tag{2.17}$$

2.5 Distance of Jousselme

Jousselme et al. in (Jousselme 2001), apply a classical similarity measure to achieve the comparison of the focal elements of two bbas, in order to define a distance in a vector space generated by the focal elements. A *bba* i can be seen as a vector $\overrightarrow{m}$ in the mentioned vector space and the normalized distance

The distance can be viewed as the most appropriate distance to measure the dissimilarity between two *bbas* m_1, m_2 according to (Jousselme 2001) after making a comparison of distances in belief functions theory. for two *bbas* : m_1, m_2 on 2^Ω :

$$d(m_1, m_2) = \sqrt{\frac{1}{2}((\overrightarrow{m_1} - \overrightarrow{m_2})^T \underline{\underline{D}}(\overrightarrow{m_1} - \overrightarrow{m_2})} \tag{2.18}$$

Where $\underline{\underline{D}}$ is a $2^N * 2^N$ is a matrix whose elements are $D(A, B) = \frac{|A \cap B|}{|A \cup B|}$. $|A \cap B|$ is a measure

of the conflict between A and B. If $|A \cap B| = 0$, means that A and B have no object in common which means that they are in great conflict.
Jousselme presents an other expression for her distance defined as :

$$d(m_1, m_2) = \sqrt{\frac{1}{2}(\|m_1\|^2 + \|m_2\|^2 - 2\langle m_1, m_2\rangle))} \tag{2.19}$$

and $\langle m_1, m_2\rangle$ is the scalar product defined by :

$$\langle m_1, m_2\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} m_1(A_i)m_2(A_j)\frac{|A_i \cap A_j|}{|A_i \cup A_{j|}} \tag{2.20}$$

where $n = |2^\Omega|$.

Therefore, $d(m_1, m_2)$ is considered as an illustration of the scalar product where the factor $\frac{1}{2}$ is needed to normalize d and guarantee that $0 \leq d(m_1, m_2) \leq 1$.

According to (Jousselme 2010, 2012) this metric distance respects all the properties expected by a distance and it can be considered as an appropriate measure of the difference or the lack of similarity between two *bbas*.
This distance is based on the dissimilarity of Jaccard defined as :

$$Jac(A, B) = \frac{|A \cap B|}{|A \cup B|}. \tag{2.21}$$

Moreover, $d(m_1, m_2)$ can be called a total conflict measure, which is an interesting property to measure the total conflict based on a measurable distance like presented in (Martin 2008).

3 A distance between continuous belief functions

Traditional distances that are used for the discrete case are totally useless due to the nature of the continuous belief functions. First of all, and as presented previously a scalar product is able to compute a distance between two *bbd*. To handle the problem in regard to the nature of these functions, a scalar product is defined on $\overline{\mathbb{R}}$ by :

$$\langle f, g\rangle = \int_{x=-\infty}^{+\infty}\int_{y=-\infty}^{+\infty} f([x, y])g([x, y])dxdy \tag{2.22}$$

3.1 Mathematical expression

In this section, we will present a new method to measure the similarity based on Jousselme's distance using Smets' formalism on continuous belief functions. Founded on the properties of belief functions on real numbers, we are now able to define a distance between two densities in a interval I.

$$\langle f_1, f_2\rangle = \int_{-\infty}^{+\infty}\int_{y_i=x_i}^{+\infty}\int_{-\infty}^{+\infty}\int_{y_j=x_j}^{y_j=+\infty} f_1(x_i, y_i)f_2(x_j, y_j)\delta(x_i, x_j, y_i, y_j)dy_j dx_j dy_i dx_i \tag{2.23}$$

The scalar product of the two continuous $pdfs$ is noted : $\langle f_1, f_2\rangle$ with a function δ defined as $\delta : \mathbb{R} \longrightarrow [0, 1]$

$$\delta(x_i, x_j, y_i, y_j) = \frac{\lambda([\![max(x_i, x_j), min(y_i, y_j)]\!])}{\lambda([\![max(y_i, y_j), min(x_i, x_j)]\!])} \tag{2.24}$$

where λ represents the Lebesgue measure used for the interval's length and

$\delta(x_i, x_j, y_i, y_j)$ is an extension of the measure of Jaccard applied for the intervals in the case of continuous belief functions.

$$[\![a, b]\!] = \begin{cases} \emptyset, & \text{if } a > b \\ \text{[a,b]}, & \text{otherwise.} \end{cases} \tag{2.25}$$

Therefore, the distance between two probability functions is defined by the following equation :

$$d(f_1, f_2) = \sqrt{\frac{1}{2}(\|f_1\|^2 + \|f_2\|^2 - 2\langle f_1, f_2\rangle)} \tag{2.26}$$

In fact, this distance can be used between more than two belief functions.

Let's consider σf a set of *bbds*. We measure the distance between one *bbd* and the $n-1$ other ones by :

$$d(f_i, \sigma f) = \frac{1}{n-1} \sum_{j=1, i \neq j}^{n} d(f_i, f_j) \tag{2.27}$$

3.2 Algorithm implementation

In this section, we present the implementation steps of continuous belief functions. To this end, we have developed programs in MATLAB.
The implemented functions are used to perform simulations of a distance that measure the similarity between them. Different results carried out from these simulations will be presented and analyzed in order to evaluate our proposed method.

Basic belief densities induced by normal distributions

In this part, we will focus on a normal probability density function, as presented in (Ristic and Smets 2004) . The focal sets of the belief functions are the intervals $[\mu - x, \mu + x]$ of $\overline{\mathbb{R}}$, with μ : the mean of the normal distribution and $x \in \mathbb{R}^+$ and σ : the standard deviation. We consider a normal distribution $\mathcal{N}(x;\mu;\sigma)$, with $x \geq \mu$:

$$\varphi(x) = 2(x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi^3}} e^{\frac{(x-\mu)^2}{2\sigma}}, \tag{2.28}$$

where $\varphi(x)$ is the basic belief density associated to the Gaussian, when we apply the principle of least commitment where x is the representation of the intervals previously mentioned.

This function is null at $x = \mu$, increases with x and reaches a maximum of $4/(\sigma e\sqrt{2\pi})$ at $x = \mu + \sqrt{2\sigma}$, then decreases to 0 at x goes to infinity like presented in (Smets 2005) .

3.3 Implementation of continuous belief functions

In order to represent our belief densities according to a normal and an exponential distribution, we have to create our own distributions.

Implementation of a continuous belief function induced by a normal distribution

First of all, we have to define a function that will be called : **"Mass norm"**, where we are able to compute our generated mass functions.
Within this function, we generate the normal distribution of the *bbds* defined by a mean, a standard variation and a discretization step.
To model our distribution, we use the equation (2.28), that guarantees a symmetric normal distribution. After that, we allocate the mass function $bbd.M$ representing our probability density function and then create different vectors :

- $bbd.X$: a vector having the different x of the *bbds*
- $bbd.Y$: a vector having the different y of the *bbds*

Afterwards, we focus on the distance counting. Every distribution, bbd_1 and bbd_2 is characterized by respectively (X_1, Y_1), (X_2, Y_2). We put $[X_1, X_2]$, $[Y_1, Y_2]$ into two different vectors, in order to count the x and y representing our two *bbds*. Having those vectors, we measure our function δ by calculating term by term using the equation (3) the diffracts rations of the intersection of the intervals on their union.
These measures are assigned in a matrix which contains all the ratios of the lengths. To make this distance operate in the continuous case of belief functions and using MATLAB, we replace the integrals present in equation (2.23) by a double sum to be able to have a formula that will measure the similarity of continuous belief functions. Finally, we implement our distance based on Jousselme's using the formalism proposed by Smets to generate our distance. In that specific case, we are automatically dealing with consonant belief functions

Basic belief densities induced by exponential distributions

In this section, we will suppose that the probability distribution follows an exponential density. The expression used for a the probability density is the following :

$$f(y) = \frac{y}{\theta^2} e^{\frac{-y}{\theta}} \tag{2.29}$$

It is obtained when we use the Least Commitment Principle presented in Section 4.2, on a set of basic belief densities associated to the exponential distribution, where θ is the mean and the focal elements are in the intervals $[0, x]$.
To implement a continuous belief function induced by an exponential distribution, we apply the same method except that we use equation (2.29) to obtain the needed mass function.

pdf	1	2	3	4
μ	0	0	4	4
σ	1	0.5	1	0.5

Table 2.1 – Probability density distributions

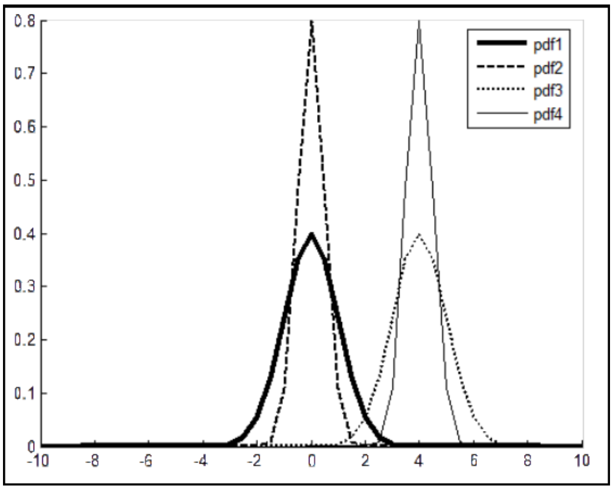

Figure 2.1 – Four $pdfs$ following normal distribution.

As mentioned in the previous section, the result of our work is a distance between two or several *bbds*.

In the next section we will consider the cases of two different kinds of distributions : the first one is a normal representation and the second is an exponential one. We will use these distributions to deduce the different *bbds* and then we are able to measure the distance between two or several continuous belief functions.

The aim is to have a probability distribution that includes uncertainty, that will be modeled by the basic belief densities.

4 Similarity evaluation

In figure 2.1 we consider four $pdfs$ having normal distributions like presented in table 2.1

In this section, we will measure and compare the similarity of the four distributions presented in Figure 2.1 using two methods. The first one will be based on the divergence of Kullback-Lieber presented in equation (2.3). In the second one, we will use our distance in equation (2.26)

D_{KL}	f_1	f_2	f_3	f_3
f_1	0	0.324	16.5	16
f_2	0.625	0	16.625	16
f_3	16	16.5	0	0.5
f_4	16.25	16	0.625	0

Table 2.2 – Distances measured using the divergence of Kullback-Lieber.

developed for continuous belief functions induced by normal distributions, then we will measure the average of the obtained distances like shown in equation (2.27).

4.1 The divergence of Kullback-Lieber for similarity of normal distributions

We applied the divergence of Kullback-Lieber for the fours $pdfs$ of the table 2.1

According to the properties of this divergence, we do not need to have a normalized value. When the distributions like f_1 and f_2 have closed values of means and standard deviation, so the divergence between them is not high. Contrary to f_3 and f_4 where they are characterized by biggest means than μ_1, where $\mu_3 = \mu_4 = 4 > \mu_1 = 0$, we are in the presence of a huge dissimilarity between the distributions.
More D_{KL} is high, more important is the divergence.
We also notice that the means of our evidential distributions play a role in measuring the similarity using the divergence of Kullback-Lieber. Moreover, this divergence measure as mentioned before, is not a distance, and confirmed by the fact that the similarity is not taken into account. When we compare $D_{KL}(f_1, f_2) = 0.324$ is different from $D_{KL}(f_2, f_1) = 0.625$.
This asymmetry is confirmed for the divergence between all the distributions where $D_{KL}(f_i, f_j) \neq D_{KL}(f_j, f_i)$

4.2 A continuous distance for similarity of normal distributions

First of all, we measure the distance between f_1 and the rest of the $pdfs$ using equation (2.26). We also completed all the distances between a *pdf* and the other ones. Next, the average of the distance between all the *bbds* was computed according to equation (2.27) and presented in table 2.3.

According to the Figure 2.1, the smaller is the distance, the more similar are the distributions. The distance between f_1 and f_4 is the biggest one. In other words, those two distributions

d	f_1	f_2	f_3	f_4
f_1	0	0.3873	0.7897	0.8247
f_2	0.3873	0	0.8247	0.8434
f_3	0.7897	0.8247	0	0.3873
f_4	0.8247	0.8434	0.3873	0

Table 2.3 – Distances measured using the distance for continuous belief functions.

Average distance	Value
$d(f_1, \sigma f)$	0.634
$d(f_2, \sigma f)$	0.680
$d(f_3, \sigma f)$	0.667
$d(f_4, \sigma f)$	0.685

Table 2.4 – Averages of the continuous distance.

are the farest from each other comparing to the distance between (f_1, f_2), and (f_1, f_3).

Otherwise, according to the obtained measures, we can say that f_2 is the most similar distribution to f_1 as it is the case between f_3 who f_4.

Once all the distances between the $4pdfs$ computed we are able to determine their averages using equation (2.27). The resulats are shown in the above table :

We notice that all the averages have close values. When we use the divergence of Kullback-Lieber, we can suggest that the difference between the means of the distributions plays an important role in similarity measurement according to the resulting values in table 2.2. This can be shown by the divergence values obtained between pdf_1 and pdf_3, pdf_4.

5 A distance between two continuous belief functions

In this part, we will make a comparison between two distances. One is measured using a classical scalar product and the other is our adaptation of Jousselme's distance for two and several continuous belief functions. On one hand, we start by exploring belief densities induced by normal distributions, then on the other hand we will focus on the case of exponential distributions

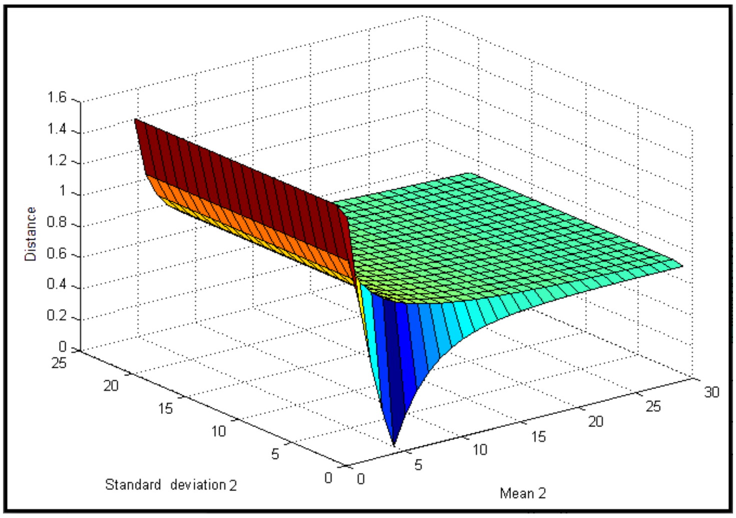

Figure 2.2 – Distance using a classical scalar product.

5.1 Belief densities induced by normal distribution :

In Figure 2.2, 2.3, we fix $\mu_1 = 0$, $\sigma_1 = 0.5$, and for the second *pdf*, $0 \leq \mu_2 \leq 10$ with a step $0, 5$ and $0.1 \leq \sigma_1 \leq 3$ with a step =0.1.
For the normal distribution, we only use focal elements where $y = 2\mu - x$ if $\mu_1 \neq \mu_2$, then the distance based on the classical scalar product is null. Else, $y = 2\mu = x$, so the scalar product presented in the continuous domain is :

$$\int_{y=\mu}^{+\infty} f(y)g(y)dxdy. \tag{2.30}$$

Classical Scalar Product for belief densities induced by normal distribution

For the second part of this analysis, we consider the distance presented in equation (2.26) for continuous belief functions.
The figure 2.2 shows the behavior of the resulting distance based on a classical scalar product. At the beginning we notice a remarkable drop in the value of the distance until reaching a null value followed by a discontinuity in the 3D representation. This discontinuity occurs after dealing with similar distributions ($\mu_1 = \mu_2, \sigma_1 = \sigma_2$) and having a distance based on the classical scalar product where $d = 0$.
We also notice that when the two probability densities have only same means ($\mu_1 = \mu_2 = 0$) the distance $d > 0$, we can say that the standard deviation must certainly paly an important

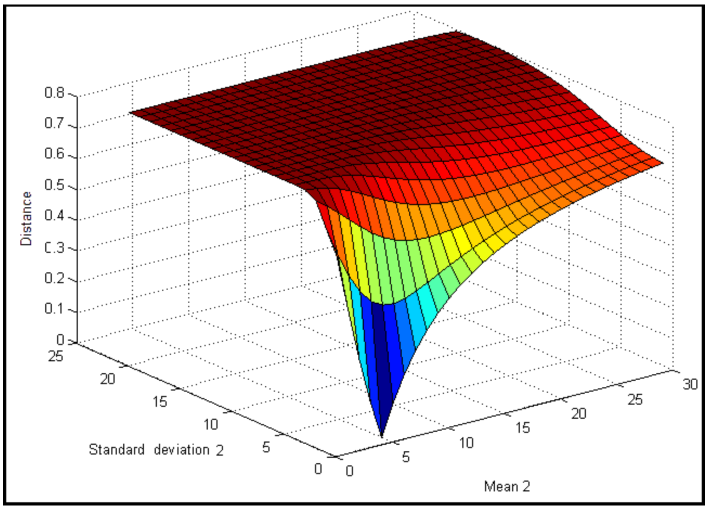

Figure 2.3 – Distance using Jaccard based on scalar product.

role when computing this similarity measure.
When $\sigma_1 > \sigma_2$, the value of the distance can be considered as high even $d > 1$. Like the second *pdf* has rising values until $\sigma_1 = \sigma_2$,the distance drops until becomes null. Otherwise, when the difference between the standard deviation grows when $\sigma_2 > \sigma_2$, the resulting distance using a scalar product also grows but still remains small comparing the case where $\sigma_1 > \sigma_2$.

Based on that, we can say that using a distance based on a classical scalar product is almost useless in the case of measuring the similarity because it gives us a bad representation of a distance.

Continuous distance for belief densities induced by normal distribution

As stated in the previous section, a distance based on a classical scalar product is not the best way to measure the similarity between two probability density function. That's way use will use the distance presented in equation (2.26)

By varying the values of μ_2 between $[0, 10]$ and σ_2 in $[0.1, 3]$ with a step of 0.5 we obtain figure 2.3.

Having pdf_2 with changing values , we notice different behaviors of our continuous distance based on Jaccard measure. First of all, when $\sigma_1 > \sigma_2$, we have a high value related to our distance that decreases as long as σ_2 increases. The distance still decreasing until reaching a null

value when $\sigma_1 = \sigma_2$. In this special case when the standard deviations are equal, same situation for the means, we are dealing with total similar distributions that generates a distance with a null value.
Once $\sigma_1 < \sigma_2$, the difference between these two parameters rises, so does the distance. This growth involves both of μ_2 and σ_2. But, it is really the standard deviation that really plays an important role. We notice even whenμ_2 reaches its maximal values, creating thus a big difference between the means of the two distributions, this does not have a remarkable impact on the similarity measure.
Contrary, the gradual increasing values of σ_2, do really influence the obtained result. When we are making the distance having its maximal value, which value is normalized, responding then to the properties presented for a distance. Every time, when the value of the distance grows in figure 2.3, the similarity between the two distributions drops.

Comparing to the results obtained with a distance based on a classical scalar product, we observe that our distance takes into consideration the standard deviation, that does have a real impact when computing this measure of similarity.

The difference between them, is the function δ presented in (17). δ, that allows us to have a more specific distance between two belief functions and use wisely the mean to have a better measure of the distance.
It is more useful, then and accurate to use our distance proposed in equation (2.26). This distance responds to the properties from Proprety 1 to preoprety 6, including : normalization, upper and lower bound and also the similarity.

5.2 Belief densities induced by exponential distributions

We measure the distance between two exponential distributions according to a classical scalar product definition, and our adaptation of Jousselme's distance for continuous belief functions.

Figures 2.4 and figure 2.5 show respectively the results obtained after computing the two distances. We are dealing with two exponential distributions f_1, f_2, where $\theta_1 = 1$ and $\theta_2 \in [0.1, 10]$ with a discretization step $= 0.1$.

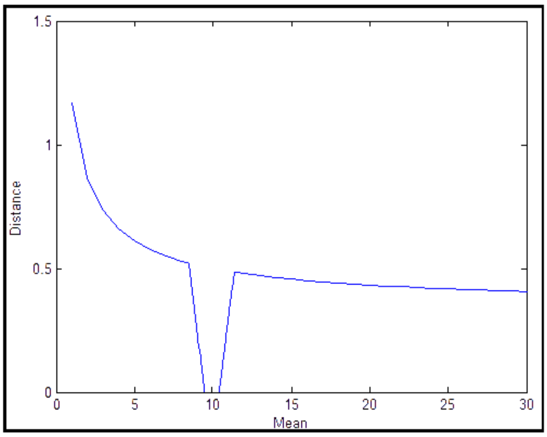

Figure 2.4 – Distance using Jaccard based on scalar product.

Classical Scalar Product for belief densities induced by exponential distributions :

At the beginning of figure 2.4, when $\theta_1 = 1$ and $\theta_2 = 0.1$, the distance based on classical scalar product has the highest value. When the value of θ_2 increases, the distance between the distributions decreases continually until it makes a discontinuity when $\theta_1 = \theta_2$ where it has a null value, this means that $f_1 = f_2$.

After that, the value of θ_2 gets far from θ_1, the probability distribution f_2 becomes different from f_1 and that generates a non null distance that increases suddenly just when θ_2 gets bigger, then the distance based on classical scalar product decreases and gets stable and reaches a value near to 0.4. Based on these observations, the distance using the classical scalar product is not really adapted for continuous belief functions because it creates a discontinuity, and except the point where $\theta_1 = \theta_2$, the variation of this distance does not reflect the difference between θ_1 and θ_2.

Continuous distance for belief densities induced by exponential distributions :

The behavior of the distance in figure 2.5 is starting at the highest value $d(f_1, f_2) = 0.6$ with the similar exponential distribution presented previously.

When the value of θ_2 gets closer to θ_1, $d(f_1, f_2)$ decreases and reaches a null value where $f_1 = f_2$ (the distributions are similar to each other). Unlike the distance based on classical scalar product, the increase of $d(f_1, f_2)$ is gradual as θ_2 gets bigger values, the distance grows continually.

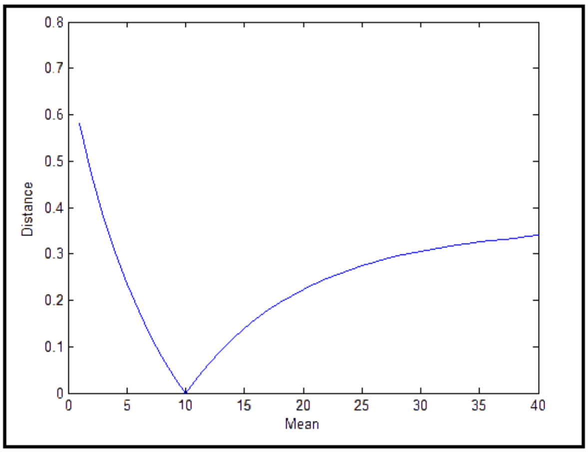

Figure 2.5 – Distance using Jaccard based on scalar product.

6 Conclusion

In this chapter, we have provided an analysis of natural properties that a distance between belief functions should satisfy. We have also shown different methods used to measure the similarity between these functions. We proposed a distance for continuous belief functions based on the proposal of Jousselme using the formalism of Smets. To do so, we simulated both normal and exponential distributions to show that our continuous similarity measure satisfies all properties for a distance. Moreover, this analysis led us to the conclusion that the standard deviation of a normal distribution representing a continuous belief function has an impact on the similarity measure.

Chapter 3

The inclusion of continuous belief functions

After defining a relation of similarity based on a distance, we will focus on an other situation characterizing continuous belief functions. In this chapter we will propose two forms of inclusion : The strict and the partial inclusion between continuous belief function induced by normal distributions also based on the formalism of Smets.

1 Introduction

Characterizing the nature of the relation that exists between belief functions can be considered as an important step in multiple tasks as information fusion and conflict managing. As already proved in the discrete case, the two more or several belief functions are included in each other, the less they are in conflict. It is an important step that will be considered when combining different sources of information and will have an impact during the process of depiction making.

In this dissertation, we emphasize on two types of inclusions between continuous belief functions : the strict and the partial inclusion. In a given or simulated distributions describing our $pfds$,

The rest of the chapter is organized as follows : Section 2 presents the notion of inclusion for the conflict measurement between discrete belief functions. Next we will present our second approach for defining a relation of inclusion between continuous belief function by presenting both of the partial and the strict inclusion. Here,we will introduce the proprieties that this relation should satisfy, the mathematical expression. Later, in section 4 we will illustrate by using normal distributions. Finally we will make an overview between the measure of inclusion and the distance presented in the previous chapter.

2 The inclusion within discrete belief functions

In this section we will present the notion of inclusion between discrete belief functions used for the aim of conflict measure.

Recently Martin (Martin 2012) defined a degree of inclusion as involved in the measurement made in order to determine the conflict during the combination of two discrete belief functions. He presented an index of inclusion has binary values where $Inc(X_1, Y_2) = 1$ si $X_1 \subseteq Y_2$ and 0 otherwise with X_1, Y_2 being respectively the focal elements of m_1 and m_2. This index is then used to measure the degree of inclusion of the two mass functions and defined as :

$$d_{inc} = \frac{1}{|F_1||F_2|} \sum_{X_1 \in F_1} \sum_{Y_2 \in F_1} Inc(X_1, Y_2) \tag{3.1}$$

$$\sigma_{inc}(m_1, m_2) = max(d_{inc}(m_1, m_2), d_{inc}(m_2, m_1)) \tag{3.2}$$

Where d_{inc} is the degree of inclusion of m_1 in $_2$ and inversely.
This inclusion is used as a conflict measure for two mass functions, using it like presented :

$$Conf(m_1, m_2) = (1 - \sigma_{inc}(m_1, m_2)d(m_1, m_2)) \tag{3.3}$$

where $d(m_1, m_2)$, is the distance of Jousselme presented in section

3 The inclusion within continuous belief functions

Let's consider the following proprieties that the measure of inclusion between continuous belief functions should satisfy.

3.1 Proprieties

The inclusion defined between two intervals A and B in a set I satisfies the following requirement :

Property 1. Non-negativity

$Inc(f_1, f_2) \geq 0$

Namely, the inclusion of a distribution in the other must never be negative.

Property 2. asymmetry

$\delta_{Inc}(f_1, f_2) \neq \delta_{Inc}(f_1, f_2)$, $\forall \mathrm{f}_1 \neq f_2$

no need for the inclusion to be symmetric ....

Property 3. Upper bound

$Inc(f_i, f_i) = 1$

This property implies a total inclusion of a distribution on the other.

Property 3. lower bound

$Inc(f_i, f_i) = 0$

This property implies the absence of any intersection or inclusion of a distribution on the other.

3.2 The strict inclusion

Definition

In this section, we will define the strict inclusion between two continuous belief functions represented by two basic belief densities *bbd*.
We use these distributions to deduce first a degree of inclusion between the *bbds* and then we are able to measure the inclusion of our continuous belief functions.
Let's consider two continuous $pdfs$: f_1 and f_2. If one of the distributions is included in an other distribution, then we define to measure the strict inclusion the following equation :

$$IncStr(f_1, f_2) = \int_{-\infty}^{+\infty} \int_{y_1=x_1}^{+\infty} \int_{x_2=-\infty}^{x_1=+\infty} \int_{y_2=x_2}^{y_2=+\infty} \delta_{IncStr}([x_1, y_1], [x_2, y_2]) f_1([x_1, y_1]) f_2([x_2, y_2]) dy_2 dx_2 dy_1 dx_1 \tag{3.8}$$

Where $X = [x_1, y_1]$, $Y = [x_2, y_2]$ are the considered intervals and $\delta_{IncStr}(X, Y)$ is the degree of inclusion that will allow us to measure the inclusion of the first interval in the second one.
We will consider that $\delta_{IncStr}(X, Y)$ having binary values :

$$\delta_{IncStr}(X, Y) = \begin{cases} 1, & \text{if X} \subseteq Y \\ 0, & \text{otherwise.} \end{cases} \tag{3.9}$$

If we are in presence of two distributions that do touch each other, there is an intersection between them. The $\delta_{IncStr}(X, Y)$ will have the value 1, and the strict inclusion will be weighted by the masses of our continuous belief functions. Otherwise $\delta_{IncStr}(X, Y)$ will be null.

3.3 The partial inclusion

Considering two *bbds* represented by two intervals X and Y : We say that X is partially included in Y or inversely, if and only if their intersection is different of $\emptyset$.
To represent the partial inclusion we define :

$$IncPar(f_1, f_2) = \int_{-\infty}^{+\infty} \int_{y_1=x_1}^{+\infty} \int_{x_i=-\infty}^{x=+\infty} \int_{y_2=x_2}^{y_i=+\infty} \delta_{IncPar}([x_1, y_1], [x_2, y_2]) f_1([x_1, y_1]) f_2([x_2, y_2]) dy_2 dx_2 dy_1 dx_1 \tag{3.10}$$

with $\delta_{IncP}(X, Y)$ is the degree of partial inclusion :

$$\delta_{IncPar}([x_1, y_1],[x_2, y_2]) = \frac{max(0, min(y_2, y_2) - max(x_1, x_2))}{y_1 - x_1} \tag{3.11}$$

Where $\delta_{IncP}(X, Y)$ representing the length of the intersection of the two probability density functions f_1 and f_2 on the length of f_1 if we are measuring $IncPar(f_1, f_2)$ and the length of f_2 if we have to calculate the partial inclusion of f_2 in f_1 : $IncPar(f_2, f_1)$

4 Inclusion between two continuous belief functions

To illustrate both of the strict and partial inclusion, which are two forms of relation between two continuous belief functions that are induced using a normal distribution. We decided to fix the value of the first distribution pdf_1, where, it is characterized by its mean $\mu_1 = 0$ and its standard deviation $\sigma_1 = 1$. For the second distribution pdf_2, we will vary the values of μ_2 in $[0, 5]$ and σ_2 in $[0.1, 3]$.
Here, our purpose is to see the behavior of the inclusion when we modify one of the $pdfs$, and the parameters that infer in the obtained results.

4.1 Strict inclusion between belief densities induced by normal distributions

This strict inclusion is a natural approach, that allows us to perceive if there exists any intersection between two distributions, how they behave and the parameters that interfere during this process. In this case, when a degree of inclusion has binary values of 0 and 1, we will study this property between two continuous belief functions.

At the beginning of the experimentation, when $\mu_1 = \mu_2$ and $\sigma_1 > \sigma_2$, we remark no intersection between the two distributions. In other words, there is then no inclusion. This situation is expressed by $IncStr(f_1, f_2) = 0$ like shown in Figure 3.1. This value is due to the fact that the degree of our strict intersection $\delta_{IncStr}(X, Y) = 0$ havs presented in the equation (3.9).

The more the value of the second standard deviation grows, the more difference between σ_1 and σ_2 develops. Here, the strict inclusion between the two distributions, which generates a growth in the behavior of the obtained curve. We are then, in the presence of a growth of the

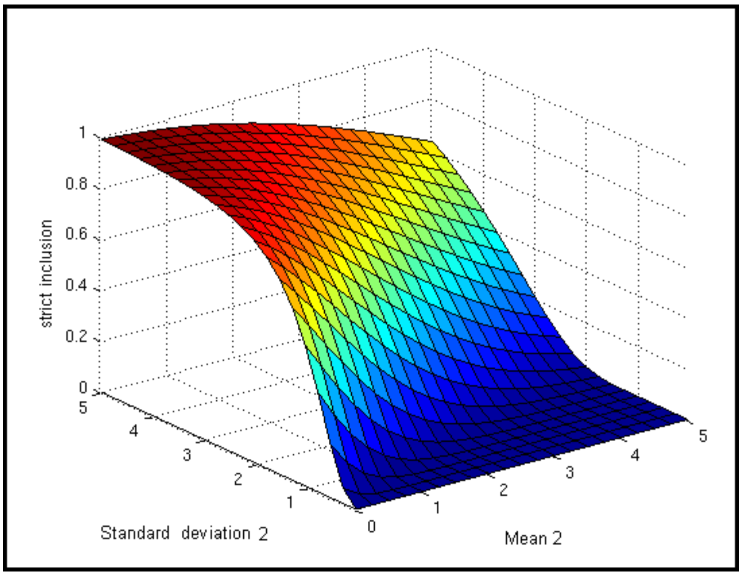

Figure 3.1 – Strict inclusion of two belief normal distributions .

strict inclusion of pdf_1 in pdf_2 due to the variation of the second distribution.
Otherwise, we notice in Figure 3.1, relative to the strict inclusion, that it is not the gap between the means μ_1 and μ_2 that generates this growth. It is the difference between both the standard deviations, that is in the origin of this growing phenomenon of inclusion. As shown when $\sigma_1 > \sigma_2$ and μ_2 having its maximal value with $\mu_2 = 5$, the strict inclusion is null. This value is due to the fact that the degree of strict inclusion $\delta_{IncStr}(X, Y) = 0$, which can be explained by the non-existence of any intersection between pdf_1 and pdf_2.
The part of the curve where the difference between the two standard deviations is low and at the same time the difference between the means is high. We have a small growing strict inclusion for example when $\mu_2 = 3$ and $\sigma_2 = 1,5$, the strict inclusion $IncStr(f_1, f_2) = 0,1$. At the meantime, even when $\mu_1 = \mu_2 = 0$ $\sigma_1 = \sigma_2 = 1$, there is an intersection between the two distributions and $IncStr(f_1, f_2) > 0$. Moreover, when the gap between σ_1 and σ_2 gets higher, the curve increases, generating a bigger strict inclusion between pdf_1 and pdf_2, until rising its maximal value where $IncStr(f_1, f_2) = 1$, with $\delta_{IncStr}(X, Y) = 1$. In this specific case, we are in presence of a total inclusion of the first distribution pdf_1 in the second one pdf_2.

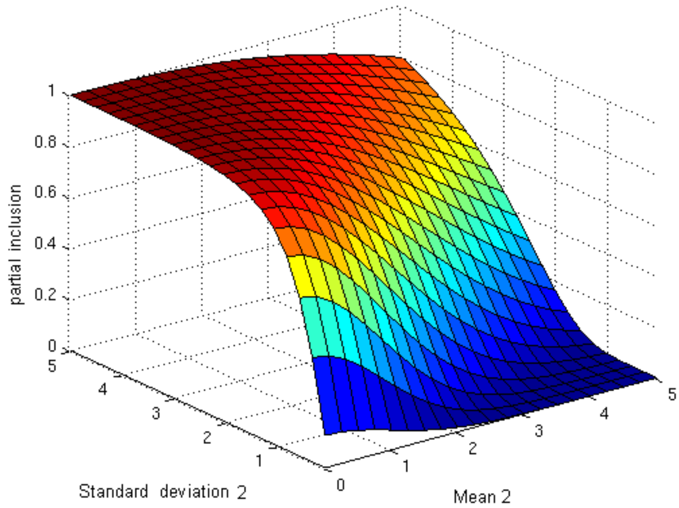

Figure 3.2 – Partial inclusion of two belief normal distributions

4.2 Partial inclusion between belief densities induced by normal distributions

The partial inclusion is defined in order to give us the proportion of the intersection between two $pdfs$.

During this experimentation, we keep the same values used for the strict inclusion, in which, the first pdf $\mu_1 = 0$ and its standard deviation $\sigma_1 = 1$. For the second distribution pdf_2, μ_2 in $[0, 5]$ and σ_2 in $[0, 3]$. We then obtain the result shown in Figure 3.2.
As long as $\sigma_1 > \sigma_2$, the partial inclusion is null, then there is no inclusion between the distributions. We notice that, when we are dealing with similar distribution where $\mu_1 = \mu_2 = 0$ and $sigma_1 = sigma_2 = 1$, the value of the partial inclusion is inconsiderable. It is possible to state that when we are in presence of two distributions having the same values, there is not necessarily any total inclusion between them.
We also take note, as the difference between σ_1 and σ_2 rises, due to the of pdf_2, the figure obtained grows faster, and reaches its maximal value $IncPar(f_1, f_2) = 1$, generating a curve more arched than the strict inclusion
When the difference between μ_1 and μ_2 increases because of the variation of pdf_2, the partial inclusion reaches a value of $IncPar(f_1, f_2) = 0.85$, which becomes lower when the gap between two standard deviations is the highest ($\sigma_2 = 3$), we obtain the maximal value for the partial

inclusion :$IncPar(f_1, f_2) = 1$ like presented in Figure 3.2.
In this specific case, we witness of a full and total inclusion of pdf_1 in pdf_2. This is similar to what we have presented regarding the strict inclusion.

Here we have non-negative inclusions, that respect the lower and upper bounds where the values are $[0, 1]$. Besides, the property of asymmetry is also respected because, the inclusion (what ever is strict or partial) of a distribution in the other does not necessarily involve the inverse case with the same value.

In order to focus on the role played by the growing standard deviation of pdf_2, we will deal a particular case for the partial inclusion. We keep the values of the first distribution, and we decide to have $\mu_1 = \mu_2 = 0$, and make σ_2 vary between $[0, 3]$, we then obtain Figure **??**.

As presented, as long as $\sigma_1 > \sigma_2$, the value of the inclusion is null, and gradually when σ_2 rises, the inclusion has a similar behavior and increases too until reaching $IncPar = 1$. Finally, we can assure that it is not the mean of the normal distribution that has an impact on the characterization of the relation of inclusion between pdf_1 and pdf_2, but it is the standard deviation or the difference between σ_1 and σ_2 that affects on both the strict and partial inclusion. When $IncPar = 1$, we have a total inclusion of the first distribution pdf_1 in the second one pdf_2, this situation is considered as a strict inclusion where, pdf_1 is fully included in pdf_2.

Comparing the results obtained in Figure 3.1 and Figure 3.2, where we have the same values for the two distributions, we notice that the partial inclusion reaches the maximal value faster that the strict one. Besides, its area is bigger and larger. The partial inclusion is dominating. Based on the curve representing the partial inclusion, when we decide to modify and make only σ_2 change, we can assure that it is the standard deviation that infers on the intersection between two distributions. The difference between the means μ_1 and μ_2, even when the latter is growing does not really have a considerable impact on the relation of inclusion between pdf_1 and pdf_2.

5 Inclusion and asymmetry

In order to develop the propriety of asymmetry previously exposed, we will use the distributions presented in the previous chapter in table 2.1. In this regard, we will consider the case of the partial inclusion. The partial inclusion is dominating and can be a special case of the strict inclusion that's why we will focus on it.

$IncP$	pdf_1	pdf_2	pdf_3	pdf_4
pdf_1	0	0.5498	0.0253	0.0013
pdf_2	0.9595	0	0.0041	1.9 e^{-007}
pdf_3	0.0253	0.8247	0	0.9595
pdf_4	0.0041	1.9 e^{-007}	0.5498	0

Table 3.1 – Partial inclusion and asymmetry.

Using the equation (3.10), we measure the partial inclusion of the four $pdfs$ we obtain the following results :

The propriety of asymmetry between two continuous belief functions can be confirmed when we observe the measures of inclusion in the table 3.1.

Obviously, the partial inclusion of a distribution with itself gives us a null value, which is a natural result. Otherwise, $IncP(pdf_1, pdf_2) = IncP(pdf_3, pdf_4)$, which are the highest values. This situation of a large inclusion of pdf_1 in pdf_2 and pdf_3 in pdf_4 consolidates what is shown in Figure 2.1.

Meanwhile, for the case of inclusion of pdf_1 in pdf_4 and pdf_4 in pdf_1, where the difference between the means is very important we are dealing with very small values, this confirms what we have already witnessed in the previous section. In other words the distributions do not have a remarkable impact on the measurement of the partial inclusion. This situation is corroborated, when $IncP(pdf_2, pdf_1) = IncP(pdf_3, pdf_4) = 0.9595$. Knowing that $\sigma_1 = \sigma_3$ and $\sigma_2 = \sigma_4$, when computing these partial inclusions, the difference between the standard deviations is maximal, generates a considerable value.

6 Inclusion and similarity

Within this section, we will study the impact of the similarity of two continuous belief functions when we include the notion of inclusion between them. Based on the results obtained in table 2.3 and table 3.1, we will explain the way similarity influences on the relation of inclusion.

We have already shown that when we measure the similarity between continuous belief functions induced by normal distribution that the mean does not have a real impact on the resulting distance. It is the standard deviation with a high value characterizing a distribution that plays a remarkable role for identifying this relation. So is the case, when we are dealing with the

inclusion.
For the same values of pdf_1 and the variation of pdf_2, we obtained Figure 3.2 for the partial inclusion and Figure 2.3 for the distance between continuous belief function. We already proved that the similarity measure is symmetric, contrary to the inclusion. We notice that as long as the distance between two $pdfs$ grows and gets next to its maximal value ($d = 1$), the value of the partial inclusion drops. We also can say that when the distributions are similar the value of the inclusion increases. Then, like it is the standard variation that influences both of these phenomenon, we can say that these relations are very tied.

7 Conclusion

In this chapter, we have emphasized the evaluation of the relation of inclusion between continuous belief functions induced by a normal distribution. We have defined two forms of inclusions : the strict and the partial one. Before that, we have provided all preliminary things that will allow us to experiment this kind of relation. We have also provided the approach on which the evaluation of the inclusions will be based.
Besides, we have also shown the way we have built our distributions and how we were able to measure the intersection, then the degree of inclusion ($\delta_{IncStr}, \delta_{IncPar}$), and finally how to calculate the inclusion.
The obtained results have shown that the standard deviation does really have an important impact on how a normal distributions can recover a second one. When the value of σ_2 increases, the covering rises too and reaches the biggest value of 1. Moreover, comparing both relations, we can say by analyzing the figures that the partial inclusion dominates the strict one on its area, the speed of growth, the appearance.

Conclusion and perspectives

In this report, we have defined two methods to characterize relations in an uncertain environment, namely the belief functions theory. In this work, we proposed two relations between continuous belief function. These approaches have been respectively named a distance between continuous belief functions and, strict and partial measurement for uncertain continuous distributions.

For each one of these approaches, we have provided the different component of the building procedures : the construction of the distributions, values of the means, standard deviations. Moreover, we have provided the mathematical expressions and the different properties that each measure should satisfied.

Using the formalism of Smets, continuous belief functions are defined by closed intervals. To achieve that, we have build our functions by inducing them with normal(gaussian), or exponential distributions.

For every approach proposed, we have proposed a function so called δ. This important function has an essential role when defining a relation between probability density functions. within our continuous distance, this δ function measures the intersection of intervals divided by their union. It is the extension Jaccard applied for evidential distributions.
After analyzing the behavior of our distance between two normal distributions, we can say that the value of standard deviation has a remarkable impact one similarity measure. The inclusion measurement, also has its own δ function. Nevertheless, it has a different interpretation in this case because it takes in consideration the length of the intersection of tow intervals on the length of one of them. It depends on the fact that if we want to have the inclusion of f_1 in f_2 or f_2 in f_1.

Anyway, once this measure applied for the case of normal distributions, we noticed that the means are not as important as the standard deviations. This latter, plays a fundamental role when computing both of the strict and partial inclusion.

Defining and analyzing these relations between continuous belief functions induced by normal distributions helped as to conclude that the standard deviation do have a remarkable influence on these measures.

Information fusion is a field of research well explored. During the combination of belief function, we have to deal with an important phenomena which is the conflict. A conflict between two experts : the contradiction contained in the response expressed as mass of two functions (Yager 1983). Then the global conflict is considered as the sum of partial conflicts from the empty intersections of focal elements of different experts combined.
Martin (Martin 2008), instead of having to conflict as the global conflict generated by the mass of $\emptyset$, after the combination of evidences, he defined a measure able to quantify the conflict between an expert and n other experts expressing their beliefs on the same observation. This conflict is then expressed by a distance between the experts. If their opinions are far from each other, he considers that they are in conflict. This conflict measure is based on Jousselme's distance. In order to model the conflict, Martin and Osswald (2006, 2007) proposed a combination rule that deal with partial conflict named $PCR6$ Proportional conflict Redistribution. This rule can manage the conflict generated between s experts. This rule redistributes the masses of focal elements proportionally to their initial masses.

Recently, Destercke and Burger (2012), revisited the notion of conflict. Considering that Dempster's rule of combination plays a central role. After defining some proprieties of conflict between sets, they extend it to belief functions where conflict should not be measured according to some unverified (in)dependence assumption between sources. To Do so, they propose two solutions if dependence structure

- Consider sets of possible dependencies.
- Consider the least-commitment principle.

For the continuous belief function, a distance is considered as a similarity measure. Then we are unable in that case to consider it as conflict measure. As future work, we consider different tracks of refinement that we classify into two major research directions :

- Defining a conflict between continuous belief functions : finding a new way to define and weight the conflict between our beliefs based on these relations
- Modeling the conflict, see is the standard deviation does also have an impact as it is the case for similarity and the inclusion measures

Bibliography

Attiaoui, D., Doré, P.-E., Martin, A., and Ben Yaghlane, B. (2012). A distance between continuous belief functions. 6th Scalable Uncertainty Management (SUM).

Destercke, S. and Thomas, B. (2012). Revisiting the notion of conflicting belief functions. In *Belief Functions*, volume 164, pages 153–160. Springer.

Diaz, J., Rifqi, M., and Bouchon-Meunier, B. (2006). A similarity measure between basic belief assignments. In *9th International Conference Information Fusion*.

Doneoeux, T. (2008). Conjunctive and disjunctive combination for belief functions induced by nondistinct bodies of evidence. *Artificial Intelligence*, 172 :234–264.

Doré, P.-E., Osswald, C., Martin, A., Jousselme, A.-L., and Maupin, P. (2011). Continuous belief functions to qualify sensors performances. In *ECSQARU*, volume 6717 of *Lecture Notes in Computer Science*, pages 350–361. Springer.

Dubois, D., Prade, H., and Smets, P. (2001). New semantics for quantitative possibility theory. In *ISIPTA*, pages 152–161.

Florea, M.-C., Jousselme, A.-L., Bossé, E., and Dominicc, G. (2009). Robust combination rules for evidence theory. *Information Fusion*, 10(2) :183–197.

Hsia, Y.-T. (1991). Characterizing belief with minimum commitment. pages 1184–1189.

Jousselme, A.-L., D, G., and Éloi, B. (2001). A new distance between two bodies of evidence. *Information Fusion*, 2(2) :91–101.

Jousselme, A.-L. and Maupin, P. (2012). Distances in evidence theory : Comprehensive survey and generalizations. *Int. J. Approx. Reasoning*, 53(2) :118–145.

Martin, A. (2012). About conflict in the theory of belief functions. *Belief 2012*.

Martin, A., Jousselme, A.-L., and Osswald, C. (2008). Conflict measure for the discounting operation on belief functions. *Information Fusion*, 9.

Martin, A. and Osswald, C. (2007). Generalized proportional conflict redistribution rule applied to sonar imagery and radar targets classification. *CoRR*.

Nguyen, H. T. (1978). On random sets and belief functions. In *Classic Works of the Dempster-Shafer Theory of Belief Functions*, volume 65, pages 531–542. Springer.

P-E, D., Martin, A., Abi-Zeid, I., Jousselme, A.-L., and Maupin, P. (2010). Belief functions induced by multimodal probability density functions, an application to the search and rescue problem. *RAIRO - Operations Research*, 44(4) :323–343.

Ristic, B. and Smets, P. (2004). Belief theory on the continuous space with an application to model based classification. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 1119–1126, Perugia, Italy.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press.

Smets, P. (1989). Constructing the pignistic probability function in a context of uncertainty. pages 29–40. North-Holland.

Smets, P. (1990). The Combination of Evidence in the Transferable Belief Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5) :447–458.

Smets, P. (1993). Belief Functions : the Disjunctive Rule of Combination and the Generalized Bayesian Theorem. *International Journal of Approximate Reasoning*, 9 :1–35.

Smets, P. (2005). Belief functions on real numbers. *Int. J. Approx. Reasoning*, 40(3) :181–223.

Smets, P. and Kennes, R. (1994). The Transferable Belief Model. *Artificial Intelligent*, 66 :191–234.

Strat, T. (1984). Continuous belief function for evidential reasoning. 4th National Conference on Artificial Intelligence.

Vannobel, J.-M. (2010). Consonant continuous belief functions conflicts calculation. In *IPMU*, volume 6178, pages 706–715. Springer.

Vannobel, J.-M. (2012). Continuous belief functions : Focal intervals properties. In *Belief Functions*, volume 164, pages 93–100. Springer.

Yager, R. R. (2008). Entropy and specificity in a mathematical theory of evidence. In *Classic Works of the Dempster-Shafer Theory of Belief Functions*, volume 219, pages 291–310. Springer.

Appendix A

In this appendix, we recall some other theories handling imperfect information. We will start by presenting the fundamental concepts of the probability theory then the possibility theory.

1 Probability theory

Probability theory is surely the oldest theory allowing to model uncertainty and deal with imperfect information. Probability theory could have two main interpretations :

- The frequentist probability : The probability of an event is the ratio of the number of favorable cases to the number of all possible cases.
- The Bayesian probability : The probability of a event is the degree of confidence that someone grants to the occurrence of a realization of an event A.

Here are presented some important concepts in probability theory :

Probability measure : A probability measure P is defined as :

$P : 2^{\Omega} \mapsto [0, 1]$

$P(\emptyset) = 0$

$P(\Omega) = 1$

$\forall A, B \in 2^{\Omega}, P(A \cup B) = P(A) + P(B) - P(A \cap B).$

Probability distribution : is a function defined

$P : \Omega \mapsto [0, 1]$

$\sum_{X \subseteq \Omega} P(X) = 1.$

Equiprobability : guarantees the same probability to all possible realizations of an event where $\forall x_i \in \Omega$:

$$P : \Omega \mapsto [0, 1]$$
$$p(x_1) = p(x_2) = ... = p(x_N) = \frac{1}{n}.$$

Probabilistic conditioning : The probability of the realization of an event B can change according to an other event A using :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Bayes Theorem : is a mathematical representation of how the conditional probability of event A given B is related to the converse conditional probability of B given A :

$$P(A|B) = \frac{P(B \cap A) * P(A)}{P(B)}.$$

2 Possibility theory

Possibility theory represents a non-classical uncertainty theory, first introduced by Zadeh in (Zadeh, 1978) and then developed by several authors,e.g., Dubois and Prade.

Possibility theory offers a flexible tool for representing imperfect information that is uncertain and imprecise such as expressed by humans. Before presenting the basics of possibility theory, let us first give some meanings of the term **possibility** :

- feasibility or realizability is the first interpretation of possibility, as we can say "it is possible to finish this work".
- plausibility refers to the degree to which an event can occur as in the sentence it is possible that it will rain this night.
- Logically consistent with the available information(i.e. not contradicting) :

Possibility distribution : denoted π a function which associates to each element ω of the universe of discourse Ω a value from a given possibilistic scale L such as :

$$\begin{aligned} &\pi \mapsto L \\ &\omega \mapsto \pi_x(\omega) \end{aligned}$$

Necessity measure (N or Nec) : expresses the necessity of the event A : the certainty degree that the value of x belongs to A.

$$N(A) = min_{\omega \in A}(1 - \pi(\omega))$$

Possibility measuredenoted Π : the possibility degree of A, it evaluates to what extent it is possible that the actual value of x belongs to A.

It is important to mention that there are many other important concepts in possibility theory which have not been presented above (e.g. discounting, Non specificity, information fusion etc.).

3 Belief functions theory and links with possibility theory

Dubois and Prade showed that when focal elements F_i of a given basic belief assignment m are nested (representing consonant belief functions), a possibility distribution π can be recovered from m. They proved that a belief function is a necessity measure ($bel = N$) and a plausibility function is a possibility measure ($pl = \Pi$)

Let m be a basic belief assignment with nested focal elements : $F_1 \subset F_2 \subset, ...F_n$ The possibility distribution π is derived using the following equation (Dubois Prade, 1986) :

$$\pi(x_i) = \sum_{j=i}^{n} m(F_i).$$

Appendix B

In this appendix, we make an analysis of two important functions. $\delta(x_i, x_j, y_i, y_j)$ which has been used for the distance. The second one is $\delta_{IncPar}([x_1, y_1], [x_2, y_2])$ for the partial inclusion. We will define more precisely these two δ functions because of the role played when measuring both of the similarity and the inclusion.

4 The distance measurement

We have defined in chapter 2 a distance that is able to measure the similarity between two or several belief function. The equation used for that aim is equation (2.26). Within this equation, we used the factor $\frac{1}{2}$ for normalization issues. Besides, a scalar product between the probability density functions belongs to this distance expressed by : $\langle f_1, f_2 \rangle$. Within this measure, we introduced a δ function defined by :

$$\delta(x_i, x_j, y_i, y_j) = \frac{\lambda([\![max(x_i, x_j), min(y_i, y_j)]\!])}{\lambda([\![max(y_i, y_j), min(x_i, x_j)]\!])}$$

where λ is considered as the measure of Lebesgue wich is length of a bounded interval I (open, closed, half-open) with endpoints a and b where $(a < b)$ is defined by : $\lambda = b - a$.

- $[\![max(x_i, x_j), min(y_i, y_j)]\!])$ is the expression of the intersection of two intervals defining our $pdfs$.
- $[\![max(y_i, y_j), min(x_i, x_j)]\!]$ is the intervals union.

Our function $\delta(x_i, x_j, y_i, y_j)$ takes into account, the intervals lengths, their intersection and also their union.

We will study the behavior of our δ function for the simulation presented in chapter 2, dealing with two distributions. For the first one we fix the mean $\mu_1 = 0$ and the standard

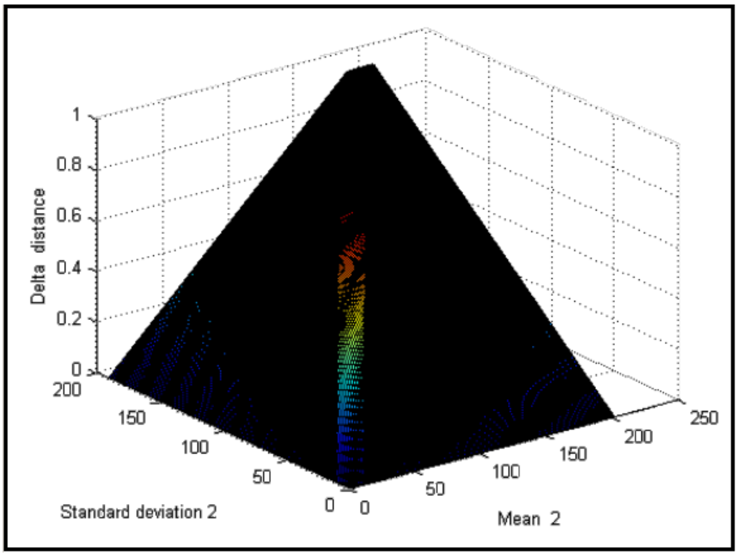

Figure A.1 – δ of the distance.

deviation $\sigma_1 = 0.5$. For the second distribution, we make its parameters increase for the mean $\mu_2 \in [0, 10]$ and $\sigma_2 \in [0, 3]$.

We obtain the result shown in figure A.1.

As we have witnessed in chapter 2, as long as the value of the standard deviation σ_2 increases, we obtain the same behavior for the δ. When difference between μ_1 and μ_2 is maximal and meanwhile $\sigma_1 > \sigma_2$, δ function for the distance does not have a considerable value. Contrary, when the difference between the standard deviation is the highest, here in this specific case, we notice that the value of δ is very close to 1.

5 The partial inclusion measurement

As presented previously, we have also defined a δ function for both of the partial and strict inclusion. The strict inclusion is defined by :

$$\delta_{IncStr}(X, Y) = \begin{cases} 1, & \text{if } \mathrm{X} \subseteq Y \\ 0, & \text{otherwise.} \end{cases}$$

Hence, if an intersection exists between the intervals, $\delta_{IncStr} = 1$, otherwise its value is null. It is not interesting to study this case. On the other hand, we will focus on the δ_{IncPar} behavior.

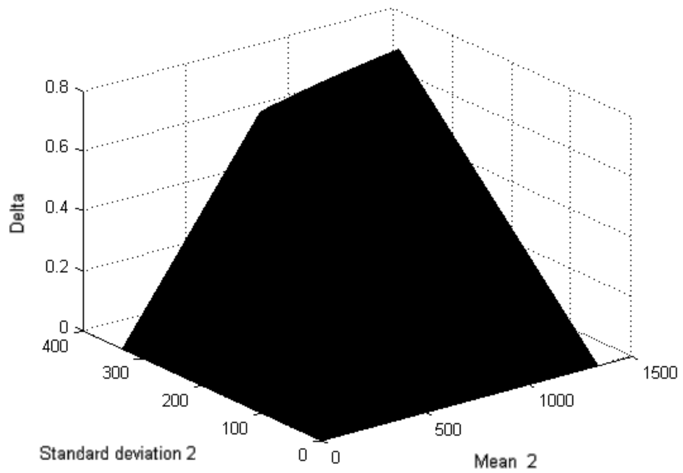

Figure A.2 – δ of the partial inclusion.

To do so, we obtain the figure A.2 using the following equation :

$$\delta_{IncPar}([x_1, y_1],[x_2, y_2]) = \frac{max(0, min(y_2, y_2) - max(x_1, x_2))}{y_1 - x_1} \quad (3)$$

$max(0, min(y_2, y_2) - max(x_1, x_2)$ represents the intervals intersection. The other part of this equation depends on the fact if we want to measure the inclusion of f_1 in f_2 or the contrary. Then we have the length of the chosen interval.

As we have stated, it is this δ_{IncPar} function that plays an important role when measuring the partial inclusion. Besides, wa can confirm the fact that its is the difference between the standard deviations that interferes in the behavior of our obtained function, which affects directly the value of the partial inclusion.