Reliable Computing **10:** 273–297, 2004.

Probability-Possibility Transformations, Triangular Fuzzy Sets, and Probabilistic Inequalities

DIDIER DUBOIS
Institut de Recherche en Informatique de Toulouse IRIT, Universite Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France, e-mail: dubois@irit.fr

LAURENT FOULLOY, GILLES MAURIS
Laboratoire d'Informatique, Systemes, Traitement de l'Information et de la Connaissance, LISTIC, Universite de Savoie—BP 806, 74016 Annecy, France, e-mail: {foulloy, mauris}@univ-savoie.fr

and

HENRI PRADE
Institut de Recherche en Informatique de Toulouse IRIT, Universite Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France, e-mail: prade@irit.fr

(Received: 14 October 2002; accepted: 22 February 2003)

Abstract. A possibility measure can encode a family of probability measures. This fact is the basis for a transformation of a probability distribution into a possibility distribution that generalises the notion of best interval substitute to a probability distribution with prescribed confidence. This paper describes new properties of this transformation, by relating it with the well-known probability inequalities of Bienaymé-Chebychev and Camp-Meidel. The paper also provides a justification of symmetric triangular fuzzy numbers in the spirit of such inequalities. It shows that the cuts of such a triangular fuzzy number contains the "confidence intervals" of any symmetric probability distribution with the same mode and support. This result is also the basis of a fuzzy approach to the representation of uncertainty in measurement. It consists in representing measurements by a family of nested intervals with various confidence levels. From the operational point of view, the proposed representation is compatible with the recommendations of the ISO Guide for the expression of uncertainty in physical measurement.

1. Introduction

A crucial step in the design of a problem solving tool where fuzzy sets [39] are involved is the determination of the membership functions of these fuzzy sets. This aspect is closely related to the interpretation of membership degrees: according to the problem, it is mainly devised in terms of similarity, preference, or uncertainty [13]. In this paper, only the uncertainty semantics will be considered. In this scope, possibility theory, introduced by Zadeh [40] appears as a mathematical counterpart of probability theory, that deals with uncertainty by means of fuzzy sets.

However, because of the lack of canonical methods for constructing membership functions, fuzzy set methods have often been dismissed prematurely by practitioners who are unfamiliar with this topic, especially in areas such as measurement. Therefore, traditional probabilistic methods remain dominant in the field of measurement and instrumentation and for the representation of uncertainty at large.

When the available information is frequentist, resorting to probabilistic modelling is natural. However, a common practice is to extract, from a probability distribution, intervals containing the value under concern, with a prescribed confidence level. This practice corresponds to a major departure from the regular probabilistic representation, since such a "confidence interval" represents a reliable set of *possible* values for a parameter. It can be viewed as a probability-possibility transformation, quite the converse move with respect to the Laplacean indifference principle, which presupposes uniform probability distributions when there is equal possibility among cases. However the weak point of the interval representation is the necessity of choosing a confidence level. It is usually taken as 95% (which means a .05 probability for the value to be out of the interval). However this choice is rather arbitrary.

Possibility measures encode families of probability distributions [2], [14] and can be viewed as a particular case of random sets [1], [9], [38]. Hence it is tempting to try to generalise our notion of confidence interval using a probability-possibility transformation. The idea of viewing possibility distributions, especially membership functions of fuzzy numbers, as encoding confidence intervals, is actually not new. Well before the advent of fuzzy sets, in the late forties, Shackle [34] introduced the connection between confidence intervals and the measurement of possibility in his theory of potential surprise, which is a first draft of possibility theory. McCain [30] also independently pointed out that a fuzzy interval models a nested set of confidence intervals with a continuum of confidence levels. The idea of relating fuzzy sets to nested confidence sets via a probability-possibility transformation was first proposed by Dubois and Prade [9]. Doing so, it is clear that some information is lost (since a probability family is obtained). However it may supply a nested family of confidence intervals instead of a single one. The guiding principle for this transformation is to minimise informational loss. The corresponding transformation has already been proposed in the past [9], [15], [25], [31]. More recent results have been obtained by Lasserre [26], Mauris et al. [27], [29] and applied to the problem of representing physical measurements.

This paper further explores the connection between this probability-possibility transformation, confidence intervals, and other well-known concepts and results in probability theory such as probabilistic inequalities. It also demonstrates the peculiar role played by symmetric triangular fuzzy numbers, which appear as the natural fuzzy counterpart to uniform probability distributions on bounded intervals. Lastly, it is suggested that possibility distributions, and especially the so-called truncated triangle possibility distribution, are a natural tool for the non-parametric representation of uncertainty measurement. A new procedure is proposed for building a membership function representing uncertainty in measurement. The procedure is based on piling up all the confidence intervals of the statistical distribution of the measures considered. This approach is well founded at the theoretical level in the possibility and probability frameworks [1], [19]. At the operational level, it is compatible with the recommendations of the ISO Guide, for the expression of uncertainty in measurement [20].

The paper is organised as follows. The second section deals with our proposal to build a possibility representation of measurement uncertainty, thus making a bridge between probability theory and possibility theory via the notion of confidence intervals. Bridges with the well-known Bienaymé-Chebychev and Camp-Meidel probabilistic inequalities are established. In the third section, it is shown that the triangular possibility distribution is a legitimate transformation of the uniform probability distribution with the same support, and that it is an upper bound of all the possibility transforms associated with all the bounded symmetric unimodal probability distributions with the same support. Finally, operational considerations concerning the representation of measurement uncertainty are presented and related with the ISO guide recommendations in metrology.

2. Probability-Possibility Transformations

The problem of converting possibility measures into probability measures and conversely has received attention in the past, but not by so many scholars (See [32], for a recent comparative review). This question is philosophically interesting as part of the debate between probability and fuzzy sets. Indeed, as pointed out by Zadeh [40], the membership function of a fuzzy set can be used for encoding a possibility distribution, and a possibility degree can be viewed as an upper bound on a probability degree (see [14], for instance). Possibility-probability transformations can be useful in any problem where heterogeneous uncertain and imprecise data must be dealt with (e.g. subjective, linguistic-like evaluations and statistical data). However, as pointed out in Dubois et al. [15], the probabilistic representations and the possibilistic ones are not just two equivalent representations of uncertainty. Hence there should be no symmetry between the two mutual conversion procedures, contrary to some proposals equating both kinds of uncertainty [24]. The possibilistic representation is weaker because it explicitly handles imprecision (e.g. incomplete knowledge) and because possibility measures are based on an ordinal structure rather than an additive one. Turning a probability measure into a possibility measure may be useful in the presence of other weak sources of information, or when computing with possibilities is simpler than computing with probabilities [22]. Opposite transformations turning a possibility measure into a probability measure were also proposed in Dubois and Prade [8], [9] and are of interest in the scope of decision-making [35], [36].

Dubois et al. [15] suggest that each kind of transformation should be guided by a particular information principle: the principle of insufficient reason when going from possibility to probability, and the principle of maximum specificity when going from probability to possibility. The first principle aims at finding a probability measure which preserves the uncertainty of choice between outcomes, and symmetries observed in a given problem, while the second principle aims at finding the most informative possibility distribution, under the constraints dictated by the possibility/probability consistency principle. This paper focuses on transforming a frequentist probability distribution into a maximally specific possibility distribution whose associated set of probability measures contains the former. A valid frequentist probability distribution contains all the information that can be gathered by observing a random phenomenon, and the criterion of maximal specificity intends to preserve as much original information as possible. This criterion is not necessarily adapted to the transformation of a subjective probability distribution reflecting an expert opinion. One may question the Bayesian credo, that any state of an agent's knowledge is necessarily representable by a single probability, since the form of a subjective probability is dictated by the exchangeable betting framework. In that case, a different criterion was proposed [16], which selects a less precise possibility measure.

2.1. Basics of Possibility Theory

Possibility measures [5], [11], [40] are set functions similar to probability measures, but they rely on an axiom which only involves the operation "supremum." A possibility measure Π on a set X (e.g. the set of reals) is characterised by a possibility distribution $\pi : X \to [0, 1]$ and is defined by:

$$\forall A \subseteq X, \quad \Pi(A) = \sup\{\pi(x),\ x \in A\}. \tag{2.1}$$

On finite sets this definition reduces to:

$$\forall A \subseteq X, \quad \Pi(A) = \max\{\pi(x),\ x \in A\}. \tag{2.2}$$

To ensure $\Pi(X) = 1$, a normalization condition demands that $\pi(x) = 1$, for some $x \in X$. The basic feature of a possibility distribution is the preference ordering it induces on X. Namely π describes what is known on the value of a variable V and $\pi(x) > \pi(x')$ means that $V = x$ is more plausible than $V = x'$. When $\pi(x) = 0$ it means that x is an impossible value of the variable V to which π is attached. When $\pi(x) = 1$ it just means that x is one of the most plausible values of this variable. Due to the definition of the possibility of an event, the possibility representation can be purely qualitative [12]. It may only use the fact that the unit interval is a total ordering. The obtained theory of uncertainty is to a large extent less expressive than probability, but also less demanding in information. Especially, it perfectly captures ignorance, letting the possibility of any event be equal to 1, except for the ever-impossible one. A possibility distribution π is more informative than another one π' whenever $\pi' > \pi$. Indeed, the set of possible values of V according to π is more restricted than the set of possible values of V according to π'. We then say that π is more *specific* than π'. In terms of fuzzy sets, this is fuzzy set inclusion of π in π'.

While probability measures are self-dual in the sense that $P(A) = 1 - P(A^c)$ where A^c is the complement of A, possibility measures are not so and $N(A) = 1 - \Pi(A^c)$ is the degree of necessity* of A [11]. Necessity measures satisfy similar properties as possibility measures with respect to set-intersection, for instance $N(A) = \inf\{N(X\backslash\{x\}),\ x \notin A\}$, noticing that $1 - \pi(x) = N(X\backslash\{x\})$.

Possibility and probability do not capture the same facets of uncertainty. The basic feature of probabilistic representations of uncertainty is additivity. Probability measures use the full strength of the algebraic structure of the unit interval. Uniform probability distributions on finite sets often model randomness. However, if considered as modelling belief, uniform probability distributions are also used in the case of total ignorance. In this situation, the probability of an outcome only depends on the number of such outcomes. Yet, only a set function that would assign the same degree to each non-impossible event (elementary or not) can model ignorance (the lack of knowledge) in a faithful way. The uniform possibility distribution is such a set function. But there does not exist a probability measure of this kind [17]. Uniform probability distributions only capture the idea of indecisiveness in front of a choice between outcomes. Hence while probability theory offers a quantitative model for randomness and indecisiveness, possibility theory offers a qualitative model of incomplete knowledge.

As it turns out, a numerical possibility measure, restricted to measurable subsets, can also be viewed as an upper probability function [2], [14], [22]. Formally, such a real-valued possibility measure Π is equivalent to the family $\boldsymbol{P}(\Pi)$ of probability measures such that $\boldsymbol{P}(\Pi) = \{P,\ \forall A \text{ measurable},\ P(A) \leq \Pi(A)\}$. Equivalently, that $\boldsymbol{P}(\Pi) = \{P,\ \forall A \text{ measurable},\ P(A) \geq N(A)\}$. The embedding of fuzzy sets into random set theory as done by Goodman and Nguyen [19], Wang Peizhuang [38], is also worth noticing. A possibility measure is actually a nested random set **S** whose realisations are its α-cuts $A_\alpha = \{x,\ \pi(x) \geq \alpha\}$, that is, $\pi(x)$ is the probability that the known value x belongs to the unknown set **S** (see also Dubois et al. [8], De Cooman and Aeyels [1]).

A numerical possibility distribution $\pi : R \to [0, 1]$ is called a fuzzy interval as soon as its α-cuts are (closed) intervals. When the modal value of π (x^m such that $\pi(x^m) = 1$) reduces to a singleton, it is also called a fuzzy number. Then, if π is continuous,

$$N(A_\alpha) = 1 - \alpha,\ \forall\alpha \in (0, 1],\ \text{and}\ \pi(x) = \sup\{\Pi\big((A_\alpha)^c\big),\ x \in A_\alpha,\ \alpha \in (0, 1]\}.$$

2.2. Basic Transformation Principles

The conversion problem between possibility and probability has roots in the possibility/probability consistency principle of Zadeh [39], that he proposed in the

* This definition of necessity measures via duality makes sense if the range of Π is upperbounded. Otherwise, necessity measures must be separately defined (as when possibility and necessity measures are used as approximations of infinite additive measures [22]).

paper founding possibility theory in 1978. This principle claims that an event must be possible prior to being probable, hence suggesting that degrees of possibility, whatever they are, cannot be less than degrees of probability (Dubois and Prade [7], Delgado et al. [3]). This is coherent with the fact that possibility measures can encode upper probabilities. In the following, a probability measure P and a possibility measure Π are said to be *consistent* if and only if $P \in \boldsymbol{P}(\Pi)$. This is the natural encoding of the principle of probability-possibility consistency. It looks natural to pick the result of transforming a possibility measure Π into a probability measure P in the set $\boldsymbol{P}(\Pi)$, and conversely to choose the possibility measure Π obtained from a probability measure P in such a way that $P \in \boldsymbol{P}(\Pi)$.

The starting point for devising transformation principles is to acknowledge the informational differences between possibility and probability measures. It is clear from the above discussion that by going from a probabilistic representation to a possibilistic representation, some information is lost because we go from point-valued probabilities to interval-valued ones. The converse transformation from possibility to probability adds information to some possibilistic incomplete knowledge.

More precisely, the probability-possibility transformation leads to find a bracketing of $P(A)$ for any measurable $A \subseteq X$ in terms of an interval $[N(A), \Pi(A)]$. When $[N(A), \Pi(A)]$ serves as a bracketing of $P(A)$, Π is consistent with P. Because $N(A) > 0 \Rightarrow \Pi(A) = 1$, this bracketing is never tight since it is always of the form $[\alpha, 1]$ or $[0, \beta]$. In order to keep as much information as possible, one should get the tightest intervals. It is easy to see that the fuzzy set with membership function π should be minimal in the sense of inclusion so that π is maximally specific (while respecting the constraint $\Pi(A) \geq P(A)$). A refinement in this specificity ordering consists in requesting that this fuzzy set be of minimal cardinality, i.e. that the value $\sum_{x \in X} \pi(x)$ be minimal (in the finite case).

Moreover, the possibility distribution π obtained from p should satisfy the constraint:

$$\pi(x) > \pi(x') \quad \text{if and only if} \quad p(x) > p(x') \quad \textit{(order preservation)}$$

since the ordering of elementary values is the basic information retained in possibilistic representations. However this condition may be weakened into:

$$p(x) > p(x') \quad \text{implies} \quad \pi(x) > \pi(x') \quad \textit{(weak order-preservation)},$$

in the sense that equally probable events need not be equally possible. The above principles for possibility/probability transformations sound reasonable but alternative ones have been proposed. These alternative views are discussed in [15], [32]. The most prominent one, due to Klir [18], [24], [25] is based on a principle of information invariance. In Klir's view, the transformation should be based on three assumptions:

- ∞ A scaling assumption that forces each value π_i to be a function of p_i / p_1 (where $p_1 \geq p_2 \geq \cdots \geq p_n$) that can be ratio-scale, interval scale, Log-interval scale transformations, etc.
- ∞ An uncertainty invariance assumption according to which the entropy $H(p)$ should be numerically equal to the measure of information $E(\pi)$ contained in the transform π of p. In order to be coherent with the probabilistic entropy, $E(\pi)$ can be the logarithmic imprecision index of Higashi and Klir [23], for instance.
- ∞ Transformations should satisfy the consistency condition $\pi(u) \geq p(u)$, $\forall u$, stating that what is probable must be possible.

Klir's assumptions are debatable. The uncertainty invariance equation $E(\pi) = H(p)$, along with a scaling transformation assumption (e.g., $\pi(x) = \alpha p(x) + \beta$, $\forall x$), reduces the problem of computing π from p to that of solving an algebraic equation with one or two unknowns. Then, the scaling assumption leads to assume that $\pi(x)$ is a function of $p(x)$ only. This pointwiseness assumption may conflict with the probability/possibility consistency principle that requires $\Pi \geq P$ for all events. See Dubois and Prade [7, pp. 258–259] for an example of such a violation. Then, the nice link between possibility and probability, casting possibility measures in the setting of upper and lower probabilities cannot be maintained.

The second and the most questionable prerequisite assumes that possibilistic and probabilistic information measures are commensurate. The basic idea is that the choice between possibility and probability is a mere matter of translation between languages "neither of which is weaker or stronger than the other" (quoting Klir and Parviz [25]). It means that entropy and imprecision capture the same facet of uncertainty, albeit in different guises. Our approach does not make this assumption.

2.3. Probability-Possibility Transformations in the Finite Case

The problem of turning a probability distribution p defined by probability values $p_1 \geq p_2 \geq \cdots \geq p_n$ into a possibility distribution π on a finite set $X = x_1, x_2, \quad , x_n$ is thus stated as follows:

Find a possibility distribution $\pi = (\pi_1, \pi_2, \quad , \pi_n)$ such that

$P(A) \leq \Pi(A) \quad \forall A \subseteq X$

p and π are order-equivalent

and π is maximally specific (any other solution π' is such that $\pi \leq \pi'$).

The solution to this problem exists and is unique. It already appears in [3], [9], and is given by:

$$\begin{aligned} \pi_1 &= 1, \\ \pi_i &= \sum_{j=i,n} p_j \quad \text{if } p_{i-1} > p_i \\ &= \pi_{i-1} \qquad \text{otherwise.} \end{aligned}$$

The proof is easy, noticing that if $p_{i-1} > p_i$, $\sum_{j=i,n} p_j$ is the minimal possible value for $\Pi(\{x_i, \quad , x_n\}) = \pi_i$.

However, if probabilities $p_i = p_{i+1}$ for some i, requesting weak order preservation only no longer preserves the unicity of the most specific possibility distribution consistent with p. Namely, we may choose between

$$\pi_i^1 = \sum_{j=i,n} p_j, \qquad \pi_{i+1}^1 = \sum_{j=i+1,n} p_j$$

and

$$\pi_i^2 = \sum_{j=i+1,n} p_j, \qquad \pi_{i+1}^2 = \sum_{j=i,n} p_j.$$

In particular there are $n!$ most specific, weakly order-equivalent possibilistic transforms of the uniform probability distribution, each obtained by means of an arbitrary permutation σ of elements of X. Namely $\forall i$, $\pi_i^\sigma = i / n$. The obtained possibility distribution always has minimal cardinality $\sum_{j=1,n} \pi_j$.

3. Continuous Probability-Possibility Transformations

Physical measurements generally are values in the set $\boldsymbol{R}$ of real numbers. Probability and possibility distributions considered here will be defined on $\boldsymbol{R}$. In the continuous case, our fundamental proposition is to derive a possibility distribution from a continuous density by means of a nominal (representative) value and the whole set of confidence intervals (with level ranging from 0 to 1) built around this nominal value. Only unimodal probability densities are considered. In the case of measurements, the nominal value will be the modal value x^m of the acquired data. For symmetric densities, the mode is equal to the mean and to the median and therefore this choice is natural. However, for asymmetric densities this choice is debatable. This paper nevertheless suggests the mode of the distribution as the most natural nominal value.

3.1. Main Results

Let us first recall our notion of a confidence interval: let p be a unimodal probability density and x^* be a "one-point" estimation of the "real" value (for example the mode or the mean value of the probability density). An interval is defined around the "one-point" estimation, and its confidence level corresponds to the probability that this interval contains the "real" value. For a confidence level α, such an interval, denoted I_α^* is called a confidence interval, and its confidence level is $P(I_\alpha^*) = \alpha$ (95%, 99% are values often used in the measurement area); $1 - P(I_\alpha^*)$ is the risk level, that is, the probability for the real value to be outside the interval. In the following, a nested family $\{I_\alpha^*\}$ of such confidence intervals all containing x^*, is assumed to be given.

DEFINITION 3.1. The fuzzy confidence interval induced by a continuous probability density p around x^* is the possibility distribution (denoted π^*) whose α-cuts are the closed confidence interval I^*_α of confidence level $P(I^*_\alpha) = \alpha$ around the nominal value x^* computed from p.

According to the definition, a possibility distribution π^* can be defined as follows:

$$\pi^*(x) = \sup\{1 - P(I^*_\alpha),\ x \in I^*_\alpha\}.$$

The possibility distribution π^* is continuous and encodes the whole set of confidence intervals in its membership function. Moreover, $\pi^*(x^*) = 1$. It can be proved that $p \in \boldsymbol{P}(\Pi^*)$ where Π^* is the possibility measure associated with π^*.

THEOREM 3.1. *For any probability density p, the possibility distribution π^* in Definition 3.1 is consistent with p, that is: $\forall A$ measurable, $\Pi^*(A) \geq P(A)$, Π^* and P being the possibility and probability measures associated respectively to π^* and p.*

Proof. For any measurable set $A \subseteq \boldsymbol{R}$, define the set $C = \{x \in \boldsymbol{R}, \pi^*(x) \leq \Pi^*(A)\}$. Obviously, $A \subseteq C$, because $\forall A$ measurable, $\Pi^*(A) = \sup_{x\in A} \pi^*(x) = \Pi^*(C)$. Now, $P(C) = \Pi^*(A)$. Indeed C^c is the cut of level $\Pi^*(A)$ of π^*, therefore, $P(C^c) = 1 - \Pi^*(A)$, due to Definition 3.1. Finally, $\Pi^*(A) \geq P(A)$ since $A \subseteq C$. □

A similar result was pointed out in the finite setting by Dubois and Prade [6]. A more general result in the infinite setting is proved by Jamison and Lodwick [22]. In the sequel, we show that ensuring the preservation of the maximal amount of information in π^* can motivate the choice of the nominal value as the mode x^m of the probability density. This is justified by the following lemma. In this lemma, the length of a measurable subset of the reals is its Lebesgue measure.

LEMMA 3.1. *For any continuous probability density p having a finite number of modes, any minimal length measurable subset I of the real line such that $P(I) = \alpha \in (0, 1]$, is of the form $\{x, p(x) \geq \beta\}$ for some $\beta \in [0, \text{pmax}]$ where $\text{pmax} = \sup_x p(x)$. It thus contains the modal value(s) of p.*

Proof. Let $I = \{x,\ p(x) \geq \beta\}$. I is a closed interval or a finite union thereof. Assume that there exists another measurable subset J of R such that $P(J) = P(I)$ with length(J) < length(I). Considering the three following disjoint domains of R: $I \cap J$, $I\backslash J$ and $J\backslash I$, we find that since $P(J) = P(I)$ by assumption: $P(J) - P(I) = \int_{J\setminus I} p(x)\,\mathrm{d}x - \int_{I\setminus J} p(x)\,\mathrm{d}x = 0$. Now, for $x \in I\backslash J$, $p(x) \geq \beta$, and for $x \in J\backslash I$, $p(x) < \beta$, therefore: $\int_{I\setminus J} \mathrm{d}x = \text{length}(I\backslash J) \leq \int_{J\setminus I} \mathrm{d}x = \text{length}(J\backslash I)$. Hence, $\text{length}(I\backslash J) + \text{length}(I \cap J) = \text{length}(I) \leq \text{length}(J\backslash I) + \text{length}(I \cap J) = \text{length}(J)$ which contradicts the assumption. □

Remark 3.1. Lemma 3.1 can be easily extended to continuous probability distributions on $\boldsymbol{R}^d$ by replacing the length by the Lebesgue measure on $\boldsymbol{R}^d$, i.e the hyper-volume, in the above proof.

This lemma does not require that the support of P be bounded. It has been proved in [15] for unimodal probability densities. Here, the proof is valid for any continuous probability density with a finite number of modes. However the unicity of the minimal length set I_α such that $P(I_\alpha) = \alpha \in (0, 1]$ is not always ensured. It holds for unimodal continuous probability densities with no range of constant value. It is also obvious from Lemma 3.1 that for any confidence level α, the smallest sets I_α such that $P(I_\alpha) = \alpha \in (0, 1]$ are nested. The lemma proves that these most informative (that is, with minimal length) confidence sets are cuts of the probability density. We call such intervals the *confidence intervals around the mode*. The corresponding possibility distribution is denoted π^{xm} and

$$\pi^{xm}(x) = 1 - P\big(\{y,\ p(y) \geq p(x)\}\big).$$

Since the minimal length sets I_α contain the modal values of p, i.e. x^m such that $p(x^m) = \text{pmax}$, whatever the probability density, it gives a justification for choosing $x^* = x^m$ and building the confidence intervals around modal values even for asymmetrical or multi-modal densities. Choosing confidence sets of minimal length ensures that this possibility distribution will be maximally specific. The degree of imprecision of π is defined by $\int_R \pi(y)\,\mathrm{d}y = \int_{[a,b]} \pi(y)\,\mathrm{d}y$ (if $[a, b]$ is the support of π). It is also equal to $\int_{[0,1]} \text{length}(A_\alpha)\,\mathrm{d}\alpha$, due to Fubini's theorem. Thus, minimising the size of the cuts of π consistent with p comes down to minimising the imprecision of π.

The notion of confidence intervals has been introduced in probability theory for a long time [21]. In the paper, we use the terminology "confidence interval" for reliable interval substitutes to probability distributions. It does not correspond to the traditional terminology. In statistics, a confidence interval has a different, although related, meaning [21]. Given a parameterized family $\{p_\theta\}$ of probability measures, and an observation x_0, the 95% confidence interval is the plausible range of the parameter θ, defined as $\{\theta,\ x_0 \in I_\theta\}$ where I_θ is a suitably defined interval $[a_\theta, b_\theta]$ such that $P_\theta(I_\theta) \geq 0.95$. Here we call I_θ a confidence interval associated to p_θ. Our notion of confidence interval is much closer to Fisher's fiducial interval (see again [21]).

A closed form expression of the possibility distribution induced by confidence intervals around the mode $x^* = x^m$ is obtained for unimodal continuous probability densities strictly increasing on the left and decreasing on the right of x^m:

$$\forall x \in [-\infty, x^m],\ \pi^{xm}(x) = \pi^{xm}\big(\ (x)\big) = \int_{(-\infty,x]} p(y)\,\mathrm{d}y + \int_{(\ (x),+\infty]} p(y)\,\mathrm{d}y, \quad (3.1)$$

where is the mapping defined by: $\forall x \in [-\infty, x^m]$, $(x) = y \geq x^m$ such that $p(x) = p(y)$. The function is continuous and strictly decreasing, therefore a one-to-one mapping, and from (3.1) is clear that is π^{xm} is continuous and differentiable, since p is continuous.

Remark 3.2. When p is unimodal, it is increasing before x^m and decreasing after x^m. When these monotonicity properties are in the wide sense only, (3.1) no longer makes sense because is no longer a one-to-one mapping. When p is still strictly monotonic on both sides of x^m, but has discontinuities, function may still makes sense if defined as $(x) = \max\{y \mid p(y) \geq p(x)\}$, but it may not be strictly decreasing any longer.

Remark 3.3. Using the closed form expression (3.1), the confidence interval takes the following form: $I_\alpha^{xm} = [(\pi_-^{xm})^{-1}(1-\alpha), \ \ \big((\pi_-^{xm})^{-1}(1-\alpha)\big)]$ where $(\pi_-^{xm})^{-1}$ is the inverse function of the increasing part of π^{xm}.

LEMMA 3.2 [15]. *If the unimodal density p has a bounded support* $\mathrm{supp}(p) =$ $[a, b]$*, then* $\forall c \in [a, b]$*,* $\forall \varphi : [a, c] \to [c, b]$ *such that* $\varphi(c) = c$*,* φ *is decreasing, let* $\pi_{\varphi, c}$ *be the possibility distribution defined by:*

$$\pi_{\varphi, c}(x) = \pi_{\varphi, c}\big(\varphi(x)\big) = \int_{(-\infty, x]} p(y)\,\mathrm{d}y + \int_{(\varphi(x), +\infty]} p(y)\,\mathrm{d}y.$$

Then $\pi_{\varphi, c}$ *is consistent with* p*.*

Proof. $\forall A$ such that $c \in A$, $\Pi(A) = 1 \geq P(A)$ when A is measurable; if $\sup A = x < c$, and since π is continuous, $\Pi(A) = \Pi((-\infty, x]) = \pi_{\varphi, c}(x) \geq P((-\infty, x]) \geq P(A)$. The same holds if $x = \inf A > c$ using $[x, +\infty)$. Other cases are proved similarly. So $\pi_{\varphi, c}$ is consistent with p. □

This result also follows from Theorem 3.1 and can be easily adapted to densities with infinite support. Joining the results of Lemmas 3.1 and 3.2 yields:

LEMMA 3.3. *The least specific possibility distribution consistent with a probability distribution P with unimodal continuous density p (in the sense that $P \in \boldsymbol{P}(\Pi)$), and that satisfies the order preservation condition is π_{φ, x^*} where* $\varphi(x) = \max\{y \mid p(y) \geq p(x)\}$*, and* $x^* = x^m$*.*

Proof. The order preservation condition forces the cuts of π to be of the form $\{x \mid p(x) \geq p(y)\}$, $\forall y$, and the core of π to be $\{x^m\}$, the modal value of p. From Lemma 3.2, π_{φ, x_m} with $\varphi(x) = \max\{y \mid p(y) \geq p(x)\}$ is consistent with p. Now if π' is such that $\pi'(x) < \pi_{\varphi, x_m}(x)$, for $x < x^m$ and π' satisfies order preservation, we clearly see that $\pi'(x) = \pi'(\varphi(x))$ and $\Pi'((-\infty, x] \cup [\varphi(x), +\infty)) <$ $\Pi_{\varphi, x_m}((-\infty, x] \cup [\varphi(x), +\infty)) = P((-\infty, x] \cup [\varphi(x), +\infty))$, i.e. π' is not consistent p. Note that φ obtained here is the function in (3.1). □

The situation can be summarised as follows: In order for π to be consistent with p, we need that $\forall I = [x, y]$, $\Pi(I^c) \geq P(I^c) = \int_{(-\infty, x]} p(t)\,\mathrm{d}t + \int_{(y, +\infty]} p(t)\,\mathrm{d}t$ for the complement of I. Lemma 3.2 says that if this condition is fulfilled for a nested family of intervals, that are the cuts of π, then π is consistent with p. To minimise the area under π, it is enough to minimise the size of these intervals, and Lemma 3.1 tells us that they should be taken as the cuts of the probability density itself. Lemma 3.3 also points out that this is equivalent to requesting that the ordering induced by p be preserved by π. If p is symmetric with mode x^m, then the possibility distribution $\pi^{xm}(x)$ is then easily defined as

$$\forall x \in [-\infty, x^m], \quad \pi^{xm}(x) = \pi^{xm}(2x^m - x) = 1 - P([x, 2x^m - x]).$$

So we have proved the following theorem.

THEOREM 3.2. *For unimodal continuous probability densities p with no range of constant value, if x^* is taken as the mode x^m, the possibility distribution induced by confidence intervals around the mode (cuts of p) is identical to the one obtained by the maximal specificity probability-possibility transformation, which verifies the consistency principle and the order preservation condition.*

Remark 3.4. For unimodal continuous densities which have ranges of constant value, Definition 3.1 may not define the possibility distribution everywhere, especially if the confidence intervals are given from cuts of p. Indeed there may be values β in $]0, 1]$ such that $P(I) \neq \beta$ for any cut I of p. Respecting the order preservation property leads to assign a constant value to $\pi^{xm}(x)$ in the intervals where π^{xm} is undefined. Then, π^{xm} is not continuous. Then Lemma 3.3 gives another approach which is applicable to when p has ranges of constant unequal values. It preserves the continuity of π^{xm} if these ranges are on the left-hand side of x^m but only the weak order preservation property holds. This technique, based on (3.1), can be extended to when there are ranges of constant unequal values on each side of π^{xm} (the domain function in (3.1) has been arbitrarily chosen on the left side of x^m). When there are ranges of constant *equal* values of p one on each side of x^m (e.g. symmetric p's), the continuity of π^{xm} is no longer ensured by using cuts of p, nor by (3.1).

Dubois et al. [15] already pointed out the relationship between the probability-possibility transformation based on cuts of p and confidence intervals. However this connection is but a consequence of their proposal. Here, the converse approach is used: we start from the notion of confidence intervals as a basis for defining possibility distributions.

EXAMPLE 3.1. For the triangular probability density with $x^{\text{mean}} = 0$ and $\sigma = 1$ defined by $\forall x \in [-\sqrt{6}, +\sqrt{6}]$, $p(x) = 1 / \sqrt{6} - |x / 6|$, a double-parabolic-shaped possibility distribution $\pi(x) = 1 + x^2 / 6 + 2|x / \sqrt{6}|$, $\forall x \in [-\sqrt{6}, +\sqrt{6}]$, is obtained (see Figure 3). The possibility distribution transforms of the reduced Gaussian and double exponential distributions ($x^{\text{mean}} = 0$ and $\sigma = 1$) are also plotted on Figure 3.

In summary, a statistical interpretation of possibility distributions in terms of confidence intervals is thus available. Definition 3.1 is more general because no assumptions are required, but the closed form (3.1) is more operational though it mainly concerns unimodal distributions with no range of constant density value.

3.2. From Possibility to Probability: The Converse Transformation

The closed form (3.1) enables the converse transformation (possibility-probability) to be considered. Differentiating $\pi^{xm}(x)$ stated by (3.1), we obtain for the derivative π':

$$\forall x \in [-\infty, x^m], \quad \pi'(x) = p(x) - p\big(\;(x)\big)\;\;'(x),$$
$$\forall y \in [x^m, +\infty], \quad \pi'(y) = -p(y) + p\big(\;^{-1}(y)\big) \,/\; \;'\big(\;^{-1}(y)\big).$$

The latter also reads $\pi'(y) = -p(\;(x)) + p(x) \,/\;\;'(x)$, since $y = \;(x)$, and $\;'(x) \neq 0$ due to the strict monotonicity of $p(x)$ before and after x^m. Eliminating $\;'(x)$ between the two expressions, we obtain:

$$\forall x \in [-\infty, x^m], \quad p(x) = \Big(\pi'(x) \cdot \pi'\big(\;(x)\big)\Big) \Big/ \Big(\pi'\big(\;(x)\big) - \pi'(x)\Big).$$

This expression allows to define the possibility-probability transformation of a unimodal possibility distribution π by defining a function g as follows:

$$\forall x \in [-\infty, x^m], \quad g(x) = y \geq x^m \quad \text{such that} \quad \pi(x) = \pi(y).$$

Therefore $\pi'(x) = \pi'(g(x))g'(x)$ because $\pi(x) = \pi(g(x))$, and thus

$$\begin{aligned} &\forall x \in [-\infty, x^m], \quad p(x) = \pi'(x) \,/\, \big(1 - g'(x)\big), \qquad \text{and} \\ &\forall y \in [x^m, +\infty], \quad p(y) = p(x). \end{aligned} \tag{3.2}$$

Therefore, $\int_{(x^m, +\infty]} p(y)\,\mathrm{d}y = -\int_{(-\infty, x^m]} p(x) g'(x)\,\mathrm{d}x$ and the function p so defined satisfies the normalization condition:

$$\begin{aligned} &\int_{(-\infty, x^m]} p(x)\,\mathrm{d}x + \int_{(x^m, +\infty]} p(y)\,\mathrm{d}y \\ &\quad = \int_{(-\infty, x^m]} \pi'(x) \,/\, \big(1 - g'(x)\big)\,\mathrm{d}x - \int_{(-\infty, x^m]} \pi'(x) \cdot g'(x) \,/\, \big(1 - g'(x)\big)\,\mathrm{d}x \\ &\quad = \int_{(-\infty, x^m]} \pi'(x)\big(1 - g'(x)\big) \,/\, \big(1 - g'(x)\big)\,\mathrm{d}x \\ &\quad = \int_{(-\infty, x^m]} \pi'(x)\,\mathrm{d}x \\ &\quad = \pi(x^m) - \pi(-\infty) = 1 - 0 = 1. \end{aligned}$$

When p is a continuous unimodal probability distribution, applying successively (3.1) and (3.2) retrieves p. Note that this possibility-probability transformation is different from the pignistic transformation proposed by Smets [35]. This is not surprising because the latter is based on another information principle, i.e. the insufficient reason principle, and consists in replacing each cut of π by uniformly distributed densities on this cut.

3.3. Probabilistic Inequalities Viewed as Probability-Possibility Transformations

Several well-known notions in statistics exist [21] that define bracketing approximations of confidence intervals for unknown probability distributions. When the probability law of considered measurements is unknown, the confidence intervals can be supplied by the Bienaymé-Chebychev inequality that can be written as follows. Let X be the quantity to be estimated and x^{mean} the mean value of its probability distribution, σ its standard deviation:

$$P(X \in [x^{\text{mean}} - k\sigma, x^{\text{mean}} + k\sigma]) \geq 1 - 1 / k^2, \qquad \text{for } k \geq 1.$$

It allows to define confidence intervals of confidence value $1 / k^2$ centred on the mean-value x^{mean} of an unknown probability law whose associated probability measure is denoted P. Note that P only needs to be known by its mean-value x^{mean} and its standard deviation σ, and it does not necessarily need to be unimodal nor symmetric.

If the probability law is known to be unimodal and symmetric, a tighter bound, the Camp-Meidel inequality can be applied [21]:

$$P(X \in [x^{\text{mean}} - k\sigma, x^{\text{mean}} + k\sigma]) \geq 1 - 1 / 2.25k^2, \qquad \text{for } k \geq 1.$$

The preceding probability inequalities suggest a simple method to build distribution free possibility approximations for probability distributions, due to Theorem 3.1, letting $\pi(x^{\text{mean}} - k\sigma) = \pi(x^{\text{mean}} + k\sigma) = 1 / k^2$, for $k \geq 1$ for instance. The confidence intervals for the Camp-Meidel inequality have a smaller length than those obtained by the Bienaymé-Chebychev inequality. This difference is due to the richer knowledge available for the probability distribution in the former case. In some sense, our results in the previous subsections give a systematic method for building the tightest inequalities adapted to each probability distribution.

4. Symmetric Triangular Fuzzy Numbers

The most usual possibility distribution found in applications of fuzzy sets is the triangular fuzzy number completely defined by its support $[x_1, x_2]$ and its modal value x^m such that $\pi(x^m) = 1$. It is often used in granular representations of subsets of the real line using fuzzy partitions. It is also very often used in fuzzy interval analysis, because it is a good trade-off between expressiveness and simplicity. Some authors have tried to justify the use of the triangular shape using some more elaborate, but still ad hoc rationale [33]. In contrast, the above results lead to a very natural interpretation of the symmetric triangular fuzzy number as yielding the optimal distribution-free confidence intervals for symmetric probability distributions with bounded support.

4.1. Triangular Fuzzy Numbers Yield a Probabilistic Inequality for Bounded Support Symmetric Densities

First, consider the case of the uniform probability distribution on a bounded interval $[x_1, x_2]$. This is the most natural probabilistic representation of incomplete knowledge when only the support is known. It is non-committal in the sense of maximal entropy, for instance, and it applies Laplace indifference principle stating that what is equipossible is equiprobable. The following result shows that triangular fuzzy numbers are obtained from the probability-possibility transformation under the weak order preservation condition:

THEOREM 4.1. *The possibility distribution transform of the uniform probability distribution on a bounded interval* $[x_1, x_2]$ *around* $x^* = x^{\mathrm{mean}}$ *is the triangular possibility distribution (denoted t.p.d.) of mode* x^{mean} *and whose support is the support of the probability distribution.*

Proof. It is straightforward because integrating a constant (as per (3.1)) gives a linear function on each side of x^{mean}. □

So the set of confidence intervals of the uniform probability distribution around its mean-value yields a symmetric triangular fuzzy number. Note that choosing x^* as any other value in the support is possible and yields the triangular fuzzy number with support $[x_1, x_2]$ and modal value x^*. It is also consistent with p. This is because if we only request the weak order preservation condition, the most specific possibility distribution containing the uniform probability distribution is not unique any longer. Requiring the full order preservation condition yields the non-fuzzy interval $[x_1, x_2]$, the uniform possibility distribution. The choice of $x^* = x^{\mathrm{mean}}$ is very natural from a symmetry argument. Moreover, the uniform probability distribution suggests the idea that even if as likely as the other values, the central value of the interval is in some sense more plausible than the other ones. Indeed, it would be strange to select a representative value of the uniform probability distribution closer to one boundary than to the other one. On such a basis, we can prove the main result of this section:

THEOREM 4.2. *The triangular symmetric possibility distribution of support* $[x_1, x_2]$ *and of mode* $x^m = (x_1 + x_2)/2$ *is the least upper bound of all the possibility transforms of symmetric probability distributions of mode* x^m *and support* $[x_1, x_2]$.

The proof requires the following lemma (see Figure 1).

LEMMA 4.1. *The possibility transform of any unimodal symmetric probability density has no inflexion point (with vanishing second derivative), and its shape is convex on the left and the right sides of the modal value.*

Proof. From the symmetry assumption, the expression in (3.1) becomes:

$$\forall x \in [-\infty, x^m), \quad \pi^{xm}(x) = \pi^{xm}\big(\ (x)\big) = 2\int_{(-\infty, x]} p(y)\,\mathrm{d}y,$$
$$\forall x \in [x^m, +\infty), \quad \pi^{xm}(x) = \pi^{xm}\big(\ ^{-1}(x)\big).$$

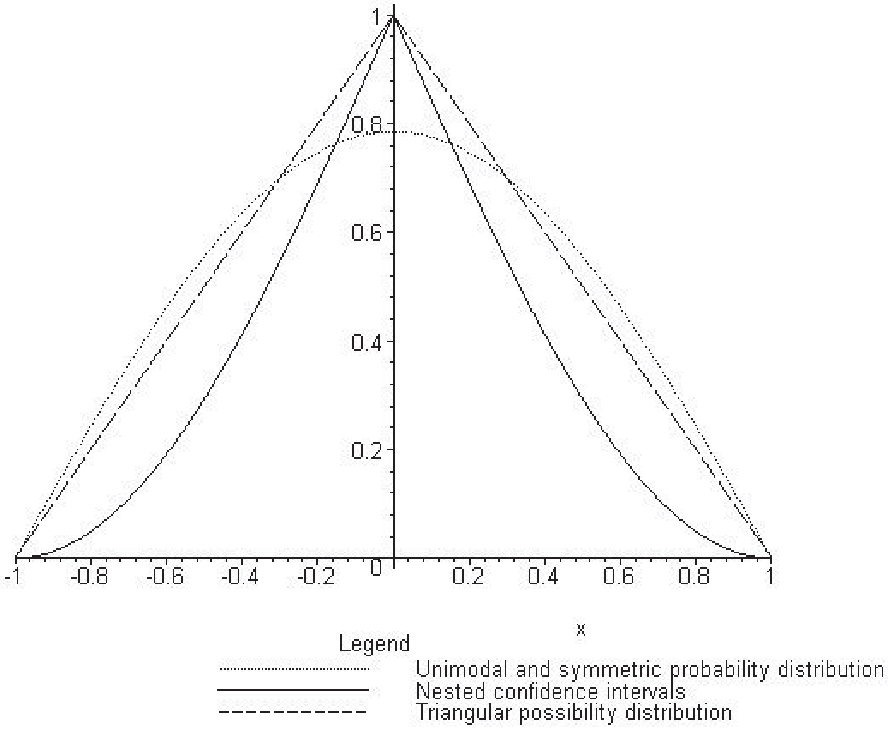

Figure 1. Possibility transformations of support-bounded probability distributions.

Therefore, the expression of the derivative of the possibility transform is:

$$\forall x \in [-\infty, x^m), \quad \pi'(x) = \pi^{xm}\big(\ (x)\big) = 2p(x),$$
$$\forall x \in [x^m, +\infty), \quad \pi'(x) = -2p\big(\ ^{-1}(x)\big).$$

Since $p(x)$ is positive and strictly increasing on the interval $[x_1, x^m]$, the first and second order derivatives of π^{xm} are strictly positive. Therefore, the possibility transform is convex on this interval. By the same reasoning, the possibility transform is convex on the interval $[x^m, x_2]$. □

Proof of Theorem 4.2. Using Lemma 4.1, since $\pi^{xm}(x_1) = 0$ and $\pi^{xm}(x^m) = 1$, it is straightforward that π^{xm} lies under the triangular possibility distribution of modal value x^m, and of same support $[x_1, x_2]$. Moreover since π^{xm} coincides with it when the distribution is uniform and that the uniform distribution is the limit of families of unimodal distributions, the triangular possibility distribution of modal value x^m is the envelope of all π^{xm} for all symmetric distributions of bounded support $[x_1, x_2]$ see Figure 2. □

Theorems 4.1 and 4.2 provide a new justification of using symmetric triangular fuzzy sets for uncertainty expression, as expected. Under the assumption

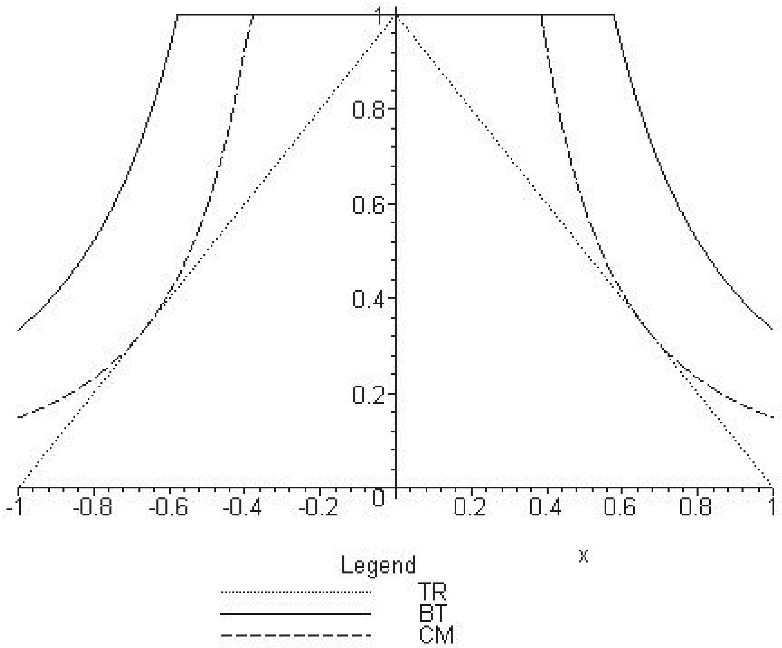

Figure 2. Nested family of confidence intervals obtained from probabilistic inequalities ($a = 1$).

of symmetry, their α-cuts can soundly be viewed as distribution-free confidence intervals of degree $1 - \alpha$ for quantities lying in a specified interval. Unfortunately this result does not carry over to distributions that would not be symmetric: A non-symmetric unimodal probability distribution is not necessarily consistent with the corresponding triangular possibility distribution having the same mode and the same support.

EXAMPLE 4.1. Consider the piecewise linear probability density of support $[-2, +2]$ defined by

$$\forall x \in [-2, -1.5], \quad p(x) = 0.6x + 1.2,$$
$$\forall x \in [-1.5, 0], \quad p(x) = (0.2 / 3)x + 0.4,$$
$$\forall x \in [0, 2], \quad p(x) = -0.2x + 0.4.$$

A piece-wise parabolic possibility distribution is obtained by applying (3.1). It gives $\pi^{xm}(-1.5) = 0.3 > \pi_{\text{triangle}}(-1.5) = 0.25$.

Remark 4.1. The technique used in this section can solve the difficulty found in Section 3 for symmetric densities with ranges of equal constant value on each side of the mode. Assume $x^m = 0$, without loss of generality. Let $[a, b]$ be a range of constant value for p on the right side of x^m. Then, $[-b, -a]$ is a range of constant (identical) value for p on the right side of x^m. Then for values x and $-x$, where $x \in [a, b]$, just let $\pi^{xm}(x) = \pi^{xm}(-x) = 1 - P([-x, x])$. This method can be adapted to such asymmetric densities with ranges of equal constant value on each side of the mode as well.

It is interesting to compare the possibility distributions obtained by the Bienaymé-Chebychev and Camp-Meidel inequalities with the triangular fuzzy numbers. The two classical inequalities are valid for larger classes of probability distributions hence they should yield less specific bounds than the triangular fuzzy number. Consider a uniform probability distribution with support $[-a, +a]$, for some positive value a. Its variance is $\sigma^2 = a^2 / 3$. Let π_{BT} and π_{CM} be the probability distributions stemming from the Bienaymé-Chebychev and Camp-Meidel inequalities respectively:

$$\begin{aligned} \pi_{\mathrm{BT}}(x) &= \min(1, a^2 / 3x^2), \\ \pi_{\mathrm{CM}}(x) &= \min(1, 4a^2 / 27x^2), \end{aligned}$$

while $\pi^{xm}(x) = \max(0, 1 - |x| / a)$. It is easy to verify that indeed $\pi_{\mathrm{BT}}(x) \geq \pi_{\mathrm{CM}}(x) \geq \pi^{xm}(x)$. More specifically, the differences $\pi_{\mathrm{BT}}(x) - \pi^{xm}(x)$ and $\pi_{\mathrm{CM}}(x) - \pi^{xm}(x)$ are both minimal for $x = \pm 2a / 3$. And $\pi_{\mathrm{CM}}(2a/3) = \pi^{xm}(2a/3)$: The Camp-Meidel distribution has a tangency point with the triangular distribution (see Figure 2). So, the triangular fuzzy number encodes strictly better distribution-free confidence intervals than the Camp-Meidel inequality for support-bounded symmetric distributions with prescribed support.

4.2. The Truncated Triangular Possibility Distribution and Unbounded Densities

Quite often, unbounded probability distributions are characterised by a shape depending on the mean value x^{mean} and the standard deviation σ (which are often estimated statistically from a set of measurements). The possibility transforms of the Gaussian and the Laplace (double-exponential) probability distributions (and also of the uniform and triangular ones) are plotted in Figure 3 for reduced variable, i.e. $x^{\mathrm{mean}} = 0$ and $\sigma = 1$.

An interesting observation is that these possibility transforms cross each other at the same point [28]: ($x^5 = x^m + 1.73\sigma$; $\varepsilon = 0.086$). This property can be used to define a simple upper bound of these possibility distributions by considering a truncated triangular possibility distribution (denoted t.t.p.d.) with a truncation point (x^ε; ε). Finally, in order to easily handle the possibility representation, especially for further computations, the latter has also been restricted to the interval $[x^m - 3.2\sigma,\ x^m + 3.2\sigma]$ which corresponds to the largest 99% confidence interval for the considered four probability distributions.

The t.t.p.d. thus defined provides a simple distribution-free possibility representation of measurements whose α-cuts bound the confidence intervals of the four

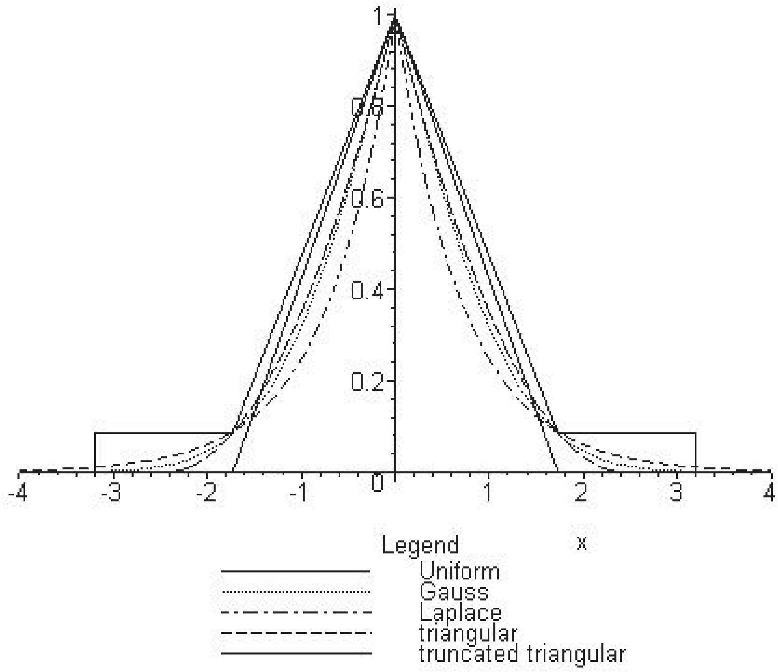

Figure 3. Possibility distributions associated to the four considered probability laws.

probability laws. An application of this possibility expression to measurements acquired by an ultrasonic range sensor has been presented in [27].

Note that the Bienaymé-Chebychev inequality supplies a possibility distribution that is far worse than the t.t.p.d. in terms of specificity. It is not unexpected because the t.t.p.d. is an approximation of the set of confidence intervals dedicated to the four considered probability distributions. However, the Bienaymé-Chebychev inequality gives a set of confidence intervals valid for any probability distribution one might choose.

The Camp-Meidel inequality supplies a possibility distribution that it is neither far from what an upper bound of the four optimal possibility distributions could be, nor far from the t.t.p.d. Anyway, the t.t.p.d. is more interesting in terms of specificity. Of course, the t.t.p.d. can only be used when the probability modelling of the measurement result can be described by one of these four probability distributions.

4.3. Additive Propagation of Symmetric Triangular Possibility Distributions

Like in the case of random variables, possibility distributions of functions of possibilistic variables can be computed by an appropriate tool called the extension principle [39]. According to this principle, fuzzy operations are identified with interval analysis for each cut. It thus provides a graded generalisation of interval calculus [4]. Moreover, the invariant forms are far more numerous for possibility propagation than in probability propagation, i.e. not only double exponential, and Gaussian distributions [10].

A general parameterized representation of fuzzy intervals has been first proposed in [7], [10]. Let L be any u.s.c. non-increasing mapping from $[0, +\infty)$ to $[0, 1]$ satisfying the following requirements: $\forall x > 0$, $L(x) < 1$; $\forall x < 1$, $L(x) > 0$; $L(0) = 1$. Two cases are considered: either $L(1) = 0$, or $L(x) > 0$, $\forall x$, and then we assume that $\lim_{x \to +\infty} L(x) = 0$.

Under these requirements, L is said to be a *shape* function. $L(x) = \max(1 - x^n, 0)$, $\max(1 - x, 0)^n$, for $n > 0$, e^{-x}, e^{-x^2}, $\frac{1}{x+1}$ are examples of shape-functions. We consider the class of u.s.c. fuzzy intervals whose membership function $A(\cdot)$ can be described by means of two shape functions L and R and four parameters: $x^a_{m^*}, x^{a\,*}_m \in \boldsymbol{R} \cup \{-\infty, +\infty\}$, $s, t \in [0, +\infty)$ having the form

$$\begin{aligned}
&A(x) = L\big((x^a_{m^*} - x) / s\big) && \forall x < x^a_{m^*},\\
&A(x) = 1 && \text{if } x \in [x^a_{m^*}, x^{a\,*}_m],\\
&A(x) = R\big((x - x^{a\,*}_m) / t\big) && \forall x > x^{a\,*}_m.
\end{aligned}$$

For $s = 0$ (resp. $t = 0$), $A(x) = 0$ when $x < x^a_{m^*}$ (resp. $x > x^{a\,*}_m$), by convention. A fuzzy interval A described as above is called an LR-fuzzy interval and the notation $A = (x^a_{m^*}, x^{a\,*}_m, s, t)_{LR}$ is adopted. L and R are called the left shape and the right shape functions, respectively. The parameters s and t are called the left spread and the right spread, respectively.

The addition of fuzzy numbers is a possibilistic counterpart to the convolution of random variables. The sum $A \oplus B$ of A and B is defined by the sup-min extension principle:

$$A \oplus B(z) = \sup_x \min\big(A(x), B(z - x)\big).$$

The following result was pointed out in [7], [10] for the addition of L-R fuzzy numbers of the form $A = (x^a_{m^*}, x^{a\,*}_m, s, t)_{LR}$ and $B = (x^b_{m^*}, x^{b\,*}_m, u, v)_{LR}$:

$$A \oplus B = (x^a_{m^*} + x^b_{m^*},\ x^{a\,*}_m + x^{b\,*}_m,\ s + u,\ t + v)_{LR}.$$

The cuts $(A \oplus B)_\alpha$ can be obtained by applying interval arithmetic to the α-cuts A_α and B_α. So, the addition of two triangular symmetric possibility distributions of support $[x^a_1, x^a_2]$ (respectively $[x^b_1, x^b_2]$) and of mode x^a_m (respectively x^b_m) is the triangular possibility distribution of support $[x^a_1 + x^b_1,\ x^a_2 + x^b_2]$ and of mode $x^a_m + x^b_m$.

THEOREM 4.3. *Let A and B be two symmetric triangular fuzzy numbers. The membership function of $A \oplus B$ is a possibility distribution consistent with the sum (by regular convolution) of any two symmetric probability distributions having the same supports as A and B respectively.*

Proof. The addition of A and B by the sup-min extension principle yields a symmetric triangular fuzzy number. Due to Theorem 4.2, the symmetric triangular fuzzy number is the maximally specific possibility distribution consistent with all the bounded symmetric probability distributions with the same support. Obviously, the probability density of the random addition of any two symmetric probability distributions p_a and p_b having supports $[x_1^a, x_2^a]$ and $[x_1^b, x_2^b]$ is symmetric, has support $[x_1^a + x_1^b,\ x_2^a + x_2^b]$, and therefore is consistent with the triangular distribution obtained by the addition of the two triangular possibility distributions. Hence $A \oplus B$ is less specific than the possibilistic transforms of the result of the convolution of p_a and p_b. □

Remark 4.2. The same theorem holds for the subtraction. However it generally will not hold for other type of operations, nor for any shape-function, nor for asymmetric fuzzy numbers.

5. Application to the Representation of Uncertainty in Measurement

Generally, the acquisition of information by measurement systems is not perfect, i.e. not totally corresponding to the observed phenomena. The reasons for imperfection are various: approximate definition of the measurand, limited knowledge of the environment context, variability of influence quantities, and so on. These effects lead to shifts and/or to fluctuations of the values, i.e. uncertainties. Representing these uncertainties is an old issue in science, but the norm to deal with these uncertainties is quite recent. Indeed, the ISO Guide for the expression of uncertainty in measurement [20] prepared under the aegis of the main international organisations in metrology:* BIPM, IEC, ISO, OIML, has been published in 1993.

According to this ISO Guide, the expression of measurement uncertainty must satisfy some requirements in order to be widely used at the practical level. The Guide recommends a parametric representation of the measurement uncertainty which:

a) characterises the dispersion of the observed values; for example the standard deviation or half the width of an interval at a given level of confidence,
b) can provide a confidence interval which contains an important proportion of the observed values,
c) can be easily propagated in further processing.

Thus, the ISO Guide proposed to characterise the measurement result by the best estimation of the measurand (i.e. in general the mean value) and the standard deviation. In fact, it simplifies the probability approach by considering only the first two moments (mean value, variance) of the probability distribution. In the

* BIPM: Bureau International des Poids et Mesures; IEC: International Electro-technical Committee; ISO: International Organization for standardization; OIML: International Organization of Legal Metrology.

fuzzy/possibility representation proposed before, the best estimation is simply the value which has membership degree 1, i.e. the modal value. Using the horizontal view of a fuzzy subset, each cut defines an interval of confidence $1 - \alpha$. Therefore, the fuzzy/possibility representation is compatible with the Guide, especially with the second aspect of point (*a*), and moreover provides the intervals for the whole set of levels of confidence, and not only for an arbitrary one, e.g. 99%. Note that there are many probability distributions corresponding to the prescribed mode and given confidence intervals, e.g. for 0.5 and 0.95. One practical difficulty with the choice of the mode for a representative value is that it is less easy to estimate than for instance the mean value. Recent proposals such as so called rough histograms (Strauss et al. [37]), where the partition underlying histograms is changed into a fuzzy partition, may provide robust estimators of the mode.

This is coherent with the fact that a possibility distribution is an upper bound of a family of probability distributions. The level of confidence $1 - \alpha$ is a lower bound of the probability that the value belongs to the interval I_α. This representation is thus not equivalent to the definition of one single probability distribution, and is particularly interesting for the expression of the so-called uncertainty of type B (i.e. those are evaluated by other means than statistical methods). As illustrated by many examples in the ISO Guide, an uncertainty of type B is often given by experts under the form of a few intervals (but not all of them) corresponding to particular levels of confidence, generally 0 and 1 [27]. Therefore, building a possibility distribution from these intervals sounds a more natural method than deriving the standard deviation since the latter requires assumptions on the shape of the probability law which is often a priori unknown in theses cases. Moreover, defining a possibility distribution by linear interpolation between the considered interval bounds is well-founded as demonstrated in Section 3.

6. Conclusion

This paper has proposed a systematic approach for the transformation of a continuous probability distribution into a maximally specific possibility distribution that enables upper bounds of probabilities of events to be computed. The obtained possibility distribution encodes a nested family of tightest confidence intervals around the mode of the statistical distribution considered. There are of course other transformations from probability to possibility [6], [15], [16], [31], [32]. The one proposed here aims at preserving as much of the information contained in the probability distribution one starts with. Each α-cut of the obtained possibility distribution is the smallest (hence the most informative) interval one may use in place of the probability distribution, with a guarantee that the probability that the unknown parameter of interest lies in this interval is at least $1 - \alpha$.

Moreover, a possibility-theoretic interpretation of probabilistic inequalities such as Bienaymé-Chebychev and Camp-Meidel inequalities has been suggested. Applied to support-bounded densities, this result provides a new justification of triangular fuzzy numbers for representing uncertainty within an interval. Indeed, the triangular possibility distribution is an optimal transform of the uniform probability distribution and it is the upper envelope of all the possibility distributions transformed from symmetric probability densities with the same support. It yields better confidence intervals for this class of distributions than the Camp-Meidel inequalities.

As an application, a new procedure for a fuzzy/possibility expression of measurement uncertainty has been outlined, based on these results. It uses the truncated triangular possibility distribution obtained as an approximation of four standard densities. At the operational level, the proposed expression is compatible with the ISO Guide of expression of uncertainty in measurement. When the uncertainty is expressed by a standard deviation only, then a better specificity is obtained by using the triangular possibility distribution than the ttpd. But, when the uncertainty is only expressed by the range of the set of measures, then a better specificity is obtained by using the truncated triangular possibility distribution. It is then useless to compute a standard deviation from the supplied range of measures in the latter case. And finally, when the range and the standard deviation are both supplied, then the greater the standard-deviation/range ratio will be, the more useful the triangular possibility distribution will be.

Further research should consider the propagation of the proposed fuzzy/possibility expression of measurement uncertainty, and the extent to which it can be used as a substitute to a plain random variable propagation. If this question is simple (as shown above) for operations like the addition or the subtraction which preserve the symmetry of probability distribution, it is more difficult for operations like division which do not preserve symmetry. It presupposes a deeper study of the possibility transforms of unimodal but asymmetric probability distributions. Can asymmetric triangular fuzzy numbers be used? As mentioned before, an important point is then to choose the nominal value, i.e. the value x^* used in definition 1. Must it be the modal, the mean or the median value ? It seems that this choice must be made before building a distribution-free possibility representation in order to be able to find an adequate family of nested intervals to be used in an operational definition of a consistent possibility distribution.

Acknowledgement

Authors are indebted to V. Lasserre for her early contribution to some parts of this work.

References

1. De Cooman, G. and Aeyels, D.: A Random Set Description of a Possibility Measure and Its Natural Extension, *IEEE Trans on Systems, Man and Cybernetics* **30** (2000), pp. 124–131.
2. De Cooman, G. and Aeyels, D.: Supremum-Preserving Upper Probabilities, *Information Sciences* **118** (1999), pp. 173–212.

3. Delgado, M. and Moral, S.: On the Concept of Possibility-Probability Consistency, *Fuzzy Sets and Systems* **21** (1987), pp. 311–318.
4. Dubois, D., Kerre, E., Mesiar, R., and Prade, H.: Fuzzy Interval Analysis, in: Dubois, D. and Prade, H. (eds), *Fundamentals of Fuzzy Sets*, The Handbooks of Fuzzy Sets Series, Kluwer Academic Publishers, Boston, USA, 2000, pp. 483–581.
5. Dubois, D., Nguyen, H. T., and Prade, H.: Possibility Theory, Probability and Fuzzy Sets: Misunderstandings, Bridges and Gaps, in: Dubois, D. and Prade, H. (eds), *Fundamentals of Fuzzy Sets*, The Handbooks of Fuzzy Sets Series, Kluwer Academic Publishers, Boston, USA, 2000, pp. 343–438.
6. Dubois, D. and Prade, H.: Consonant Approximations of Belief Functions, *Int. J. Approximate Reasoning* **4** (1990), pp. 419–449.
7. Dubois, D. and Prade, H.: *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.
8. Dubois, D. and Prade, H.: Fuzzy Sets, Probability and Measurement, *European Journal of Operational Research* **40** (1989) pp. 135–154.
9. Dubois, D. and Prade, H.: On Several Representations of an Uncertain Body of Evidence, in: Gupta, M. M. and Sanchez, E. (eds), *Fuzzy Information and Decision Processes*, North-Holland, 1982, pp. 167–181.
10. Dubois, D. and Prade, H.: Operations on Fuzzy Numbers, *Int. J. Systems Science* **9** (1978) pp. 613–626.
11. Dubois, D. and Prade, H.: *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York, 1988.
12. Dubois, D. and Prade, H.: Qualitative Possibility Theory and Its Applications to Constraint Satisfaction and Decision under Uncertainty, *Int. Journal of Intelligent Systems* **14** (1999), pp. 45–61.
13. Dubois, D. and Prade, H.: The Three Semantics of Fuzzy Sets, *Fuzzy Sets and Systems* **90** (1997), pp. 141–150.
14. Dubois, D. and Prade, H.: When Upper Probabilities Are Possibility Measures, *Fuzzy Sets and Systems* **49** (1992), pp. 65–74.
15. Dubois, D., Prade, H., and Sandri, S.: On Possibility/Probability Transformations, in: Lowen, R. and Roubens, M. (eds), *Fuzzy Logic*, Kluwer Academic Publishers, Dordrecht, 1993, pp. 103–112.
16. Dubois, D., Prade, H., and Smets, P.: New Semantics for Quantitative Possibility Theory, in: *Proc.of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ESQARU 2001), Toulouse, France, LNAI 2143*, Springer-Verlag, Berlin, 2001, pp. 410–421.
17. Dubois, D., Prade, H., and Smets, P.: Representing Partial Ignorance, *IEEE Trans. on Systems, Man and Cybernetics* **26** (1996), pp. 361–377.
18. Geer, J. F. and Klir, G.: A Mathematical Analysis of Information-Preserving Transformations between Probabilistic and Possibilistic Formulations of Uncertainty, *Int. J. of General Systems* **20** (1992), pp. 143–176.
19. Goodman, I. R. and Nguyen, H. T.: *Uncertainty Models for Knowledge-Based Systems*, North-Holland, Amsterdam, 1985.
20. *Guide for the Expression of Uncertainty in Measurement*, ISO 1993, 1993.
21. Kendall, M. and Stuart, A.: *The Advanced Theory of Statistics*, Ed. Griffin and Co., 1977.
22. Jamison, K. D. and Lodwick, W. A.: The Construction of Consistent Possibility and Necessity Measures, *Fuzzy Sets and Systems* **132** (2002), pp. 65–74.
23. Higashi, M. and Klir, G.: Measures of Uncertainty and Information Based on Possibility Distributions, *Int. J. General Systems* **8** (1982), pp. 43–58.
24. Klir, G. J.: A Principle of Uncertainty and Information Invariance, *Int. Journal of General Systems* **17** (1990), pp. 249–275.
25. Klir, G. J. and Parviz, B.: Probability-Possibility Transformations: A Comparison, *Int. J. of General Systems* **21** (1992), pp. 291–310.
26. Lasserre, V.: *Modélisation floue des incertitudes de mesures de capteurs*, Ph. D. Thesis, University of Savoie, Annecy, France, 1999.

27. Mauris, G., Berrah, L., Foulloy, L., and Haurat, A.: Fuzzy Handling of Measurement Errors in Instrumentation, *IEEE Trans. on Measurement and Instrumentation* **49** (2000), pp. 89–93.
28. Mauris, G., Lasserre, V., and Foulloy, L.: A Fuzzy Approach for the Expression of Uncertainty in Measurement, *Int. Journal of Measurement* **29** (2001), pp. 165–177.
29. Mauris, G., Lasserre, V., and Foulloy, L.: Fuzzy Modeling of Measurement Data Acquired from Physical Sensors, *IEEE Trans. on Measurement and Instrumentation* **49** (2000), pp. 1201–1205.
30. McCain, R. A.: Fuzzy Confidence Intervals, *Fuzzy Sets and Systems* **10** (1983), pp. 281–290.
31. Moral, S.: Constructing a Possibility Distribution from a Probability Distribution, in: Jones, A., Kauffmann, A., and Zimmermann, H. J. (eds), *Fuzzy Sets Theory and Applications*, D. Reidel, Dordrecht, 1986, pp. 51–60.
32. Oussalah, M.: On the Probability/Possibility Transformations: A Comparative Analysis, *Int. Journal of General Systems* **29** (2000), pp. 671–718.
33. Pedrycz, W.: Why Triangular Membership Functions?, *Fuzzy Sets and Systems* **64** (1994), pp. 21–30.
34. Shackle, G. L. S.: *Decision Order and Time in Human Affairs*, Cambridge University Press, 1961.
35. Smets, P.: Constructing the Pignistic Probability Function in a Context of Uncertainty, in: Henrion, M., Schachter, R. D., Kanal, L. N., and Lemmer, J. F. (eds), *Uncertainty in Artificial Intelligence*, North-Holland, Amsterdam, 1990, pp. 29–40.
36. Smets, P.: Decision Making in a Context Where Uncertainty Is Represented by Belief Functions, in: Srivastava, R. P. (ed.), *Belief Functions in Business Decisions*, Physica-Verlag, Heidelberg, 2001.
37. Strauss, O., Comby, F., and Aldon, M.-J.: Rough Histograms for Robust Statistics, in: *IEEE Int. Conf. On Pattern Recognition (ICPR'00)*, Barcelona, 2000, pp. 2684–2687.
38. Wang, P. Z.: From the Fuzzy Statistics to the Falling Random Subsets, in: Wang, P. P. (ed.), *Advances in Fuzzy Sets, Possibility Theory and Applications*, Plenum Press, New York, 1983, pp. 81–96.
39. Zadeh, L. A.: Fuzzy Sets, *Information and Control* **8** (1965), pp. 338–353.
40. Zadeh, L. A.: Fuzzy Sets as a Basis for a Theory of Possibility, *Fuzzy Sets and Systems* **1** (1978), pp. 3–28.