Possibility transformation of the sum of two symmetric unimodal independent/comonotone random variables

Gilles Mauris
Laboratoire d’Informatique, Systèmes, Traitement de l’Information et de la Connaissance, LISTIC
Polytech Annecy-Chambéry, Domaine Universitaire, BP 80439, 74944 Annecy Le Vieux Cedex, France,
gilles mauris@univ-savoie.fr

Abstract

The paper extends author’s previous works on a probability/possibility transformation based on a maximum specificity/possibility principle to the case of the sum of two identical unimodal symmetric random variables. This transformation requires the knowledge of the dependency relationship between the two added variables. In fact, the comonotone case is closely related to the Zadeh’s extension principle. It often leads to the worst case in terms of specificity of the corresponding possibility distribution, but it may arise that the independent case is worse than the comonotone case, e.g. for symmetric Pareto probability distributions. When no knowledge about the dependence is available, a least specific possibility distribution can be obtained from Fréchet bounds.

Keywords: Possibility theory, uncertainty propagation, maximum specificity principle, independence, comonotonicity.

1. Introduction

In a lot of applications, a probability distribution or an interval are frequently used in order to give an account of uncertainty of observations or of a parameter. The first representation provides a quite rich information but it is not always easy to determine, unlike the second representation that is more simple to obtain but also generally more rough. A third uncertainty representation based on the possibility theory [1] has been proposed in the late 70’s by Dubois and Prade and further developed within the fuzzy community, among ones the author who has proposed a probability-possibility transformation identifying the dispersion intervals of level \(1-\alpha\) of a probability distribution to the \(\alpha\)-cuts of a fuzzy subset seen as a possibility distribution [2]. This identification is closely related to the transformation principles proposed originally by Dubois, Prade and Sandri [3] (consistency, preference preservation, least commitment). Following this way, the maximum specific possibility distribution is obtained, i.e. the least one in the sense of fuzzy subset inclusion without adding arbitrary information.

In other respects, as many practical engineering situations involve many variables, it is important to consider the problem of the possibility representation of a probability distribution associated to a function of many random variables not being necessarily independent. Since this issue is not so simple as it seems at a first sight, we will limit ourselves in this paper to the case of the probability-possibility transformation of the sum of two identical continuous unimodal symmetric random variables. After having recalled the basic principles of the previously proposed probability possibility transformation and associated properties, we will study the case where the two random variables are independent and the case where they are comonotone. It will be shown that this later case is closely related to the Zadeh’s extension principle conventionally used for the propagation of possibility distributions [4]. Finally, we will give some elements about the case where no knowledge about the dependency is available. The fundamental developments will be illustrated with widely used probability distributions: uniform and Gaussian ones, but also with less used ones (nevertheless important in specific areas): the Cauchy and symmetric Pareto distributions. The paper sheds some new lights at the conditions in which the Zadeh’s extension principle leads to least specific possibility distributions. Let us note that in this paper we consider the dependency concept under the conventional umbrella of random variables, and we do not consider other dependency notions introduced beyond the conventional probability theory [5][6].

2. Probability-possibility transformation recalls

2.1. Basics of possibility theory

A fundamental notion of the possibility theory [4] is the possibility distribution, denoted \(\pi\). Here, we consider possibility distributions defined on the real line, i.e. \(\pi\) is an upper semi-continuous mapping from \(R\) to the unit interval to the unit interval such that \(\pi(x) = 1\) for some \(x\) belonging to \(R\). Thus \(\pi\) is a fuzzy subset but with specific semantics for the membership function. Indeed, a possibility distribution describes the more or less plausible values of an uncertain variable \(X\). The possibility theory provides two evaluations of the likelihood of an event, for instance whether the value of a real variable \(X\) does lie within a given interval: the possibility \(\Pi\) and the necessity \(N\). The normalized measures of possibility \(\Pi\) and necessity \(N\) are defined from the possibility distribution \(\pi: R \rightarrow [0,1]\) such that \(\sup_{x \in R} \pi(x) = 1\) as follows:
∀ A ⊂ R, Π( A ) = sup_{x, y} π(x) and

∀ A ⊂ R, N( A ) = 1 − Π( A ) = inf_{x, y} ( 1 − π(x) )

The possibility measure Π verifies :

∀ A, B ⊂ R, Π( A ∪ B ) = max( Π( A ), Π( B ) )

The necessity measure N verifies :

∀ A, B ⊂ R, N( A ∩ B ) = min( N( A ), N( B ) )

To qualify the informativeness of a possibility distribution, the concept of specificity can be used [7]. Indeed, a possibility distribution π that is said more specific than π₂ as soon as ∀ x, π₁(x) ≤ π₂(x) (it is the usual definition of inclusion of fuzzy sets), i.e. π₁ is more informative than π₂ (or dually less spread). If π(x) = 1 for some x and π(y) = 0 for all y ≠ x, then π is totally specific (fully precise and certain knowledge), if π(x) = 1 for all x then π is totally non specific (complete ignorance).

In fact, a numerical degree of possibility can be viewed as an upper bound to a probability degree [8][9]. Namely, with every possibility distribution π one can associate a non-empty family of probability measures dominated by the possibility measure :

\[ \mathcal{P}(\pi) = \{ P, \forall A \subset R, P(A) \leq \Pi(A) \}. \]

This provides a bridge between probability and possibility, and there is also a bridge with interval calculus. Indeed, a unimodal numerical possibility distribution may also be viewed as a nested set of intervals, which are the α-cuts of π : \[ [\pi_{\alpha}, \pi_{\beta}] = \{ x, \pi(x) \geq \alpha \}. \]

In fact, these α-cuts can be identified with the dispersion intervals of probability level \( \beta = 1 - \alpha \) of a probability distribution. For every possible probability level \( \beta \in [0, 1] \), the corresponding dispersion interval (called a coverage interval in the metrology area [10]) is defined as an interval that contains a portion of X with probability \( \geq \beta \). In others words, a dispersion interval of probability level \( \beta \) (denoted \( I_\beta \) ) is defined as an interval for which the probability \( P_{\alpha} \) to be outside this interval \( I_\beta \) does not exceed \( \alpha = 1 - \beta \). In fact, the dispersion intervals give information about the concentration of the probability distribution associated to a random variable.

For the same probability density function and for the same probability level, we can have different types of dispersion intervals. Indeed, we can impose the dispersion intervals to be defined around a same point \( x^* \), generally a typical central point of the probability density, e.g. the mode, the mean or the median. Obviously, the dispersion intervals built around the same point \( x^* \) are nested. It has been proven in [11] that stacking dispersion intervals on top of one another leads to a possibility distribution (denoted \( \pi^* \) having \( x^* \) as modal value). In fact, in this way, the α-cuts of \( \pi^* \), i.e., \( A_{\alpha} = \{ x, \pi^*(x) \geq \alpha \} \) are identified with the dispersion interval \( I_\alpha \) of probability level \( \beta = 1 - \alpha \) around the nominal value \( x^* \).

Thus, the possibility distribution \( \pi^* \) encodes the whole set of coverage intervals in its membership function. Moreover, this possibility distribution satisfies :

\[ \forall A \subset R, \Pi^*(A) \geq P(A), \] with \( \Pi^* \) and \( P \) the possibility and probability measures associated respectively to \( \pi^* \) and \( P \) (the underlying probability density function of the measurements). It is proved in [11] that the most informative (i.e. with minimal length) dispersion intervals are level cuts of the probability density, which are in the multi-modal case constituted of unions of intervals. The corresponding possibility distribution, denoted \( \tilde{\pi} \), has therefore the following expression:

\[ \pi^*(x) = 1 - \int_{p(y) < p(x)} p(y) dy \] (1)

This possibility distribution is the most specific possibility distribution among those that dominate \( P \), i.e. \( \forall A \subset R, \Pi^*(A) \geq P(A) \).

For symmetric unimodal probability distributions with mode \( M \), the expression (1) becomes:

\[ \pi^*_M(x) = 2F(x), x \leq M \]
\[ \pi^*_M(x) = 2(1 - F(x)), x \geq M \]

and thus, the possibility distribution is simply related to the cumulative distribution function \( F \). The expression under the form of α-cut nested intervals is :

\[ \pi^*_\alpha = \left[ F^{-1}(\alpha/2), F^{-1}(1 - \alpha/2) \right] \] (2)

with \( F^{-1} \) the left inverse function (called also quantile function) of the corresponding cumulative distribution function \( F \). \( F^{-1}(y) = \inf \{ X(x) : y \} \).

The following figure illustrates the possibility distributions equivalent to the probability distributions of mode 0 and standard deviation for uniform Gaussian, triangular and double exponential densities.

---

Fig. 1: Examples of possibility distributions equivalent to probability distribution with \( M=0 \) and \( \sigma = 1 \).
In the following, for a sake of simplicity but without lack of generality we will consider distributions with a mode $M=0$.

### 2.2. Properties of transformation

In fact, the equivalent possibility distribution is a functional summary of the dispersion of the random variable. Thus, it has many relationships with the conventional dispersion parameters: mean absolute deviation, variance, interquantile intervals.

For instance, the specificity index of the possibility distribution equivalent to an unimodal symmetric probability distribution, is equal twice the mean absolute deviation [12]:

$$sp(\pi_x) = \int_{-\infty}^{\infty} \pi_x(x) dx = 2E|X|. $$

The relationship with the variance is not so direct because it involves the mean absolute deviation of the integral of the equivalent possibility distribution $\pi_i_x$:

$$\pi_i_x(x) = \int_{-\infty}^{\infty} \pi_x(y) dy$$

$$2V(X) = \int_{-\infty}^{\infty} |x|\pi_i_x(x) dx$$

Let us remark that when the mean absolute deviation or the variance is indefinite, the equivalent possibility distribution can still be computed (see hereafter the case of the Cauchy distribution).

One more general result allows to relate the specificity order of possibility distributions to order on the moments. Indeed, for every function $L$ having the form

$L(x, \theta) = h|\theta - \theta|$ ( $\theta$ being a fixed central value of the distribution, here the mode which for symmetric distribution is equal to the mean and to the median), with $h : [0, \infty) \rightarrow [0, \infty) $ a decreasing function, we have [12]:

$$\pi_x(x) \leq \pi_x(x) \Leftrightarrow E_x(L(x)) \leq E_x(L(x)), \forall L$$

The mean absolute deviation is obtained with

$L(x, \theta) = |x - \theta|$, and the variance with

$L(x, \theta) = |x - \theta|^2$.

Let us also note that Fréchet has introduced in 1940 [13] another dispersion measure, i.e. the half-width of the interquartile interval defined in the following way: $[F^{-1}(0.25), F^{-1}(0.75)]$. This interval is indeed the cut of level 0.5 of the equivalent possibility distribution $\pi_x$.

### 3. Transformation of the sum of two identical symmetric unimodal random variables

Let us consider two identical continuous symmetric unimodal random variables $X_1$ and $X_2$ represented by the distribution functions $F_1$ and $F_2$, and $Y$ the random variable associated to their sum $Y = X_1 + X_2$ represented by the distribution function $G$. To determine, the possibility distribution equivalent to $Y$, denoted $\pi_Y$, the knowledge of $G$ is required, and the associated expression depends on the dependence relationship between $X_1$ and $X_2$. Hereafter, we consider two extremal cases of dependency (independence and comonotonicity [14]), and also the case where no knowledge about the dependency is available. For these different cases, we determine the expression of $\pi_Y$ for $X_1$ and $X_2$ being identical uniform, Gaussian, Cauchy or symmetric Pareto distributions, and we compare them in terms of specificity. Without lack of generality, we consider distributions with a mode at 0, the results being invariant with translation.

### 3.1. Independent random variables

In this widely used case, $Y$ is defined by the conventional convolution product of $X_1$ with $X_2$:

$$G(y) = \int_{-\infty}^{\infty} F_1(y-x) dF_2(x)$$

For unimodal and symmetric random variables $X_1$ and $X_2$, $Y$ is also unimodal and symmetric, thus $\pi_Y$ can be written in the form:

$$\pi_Y(y) = 2G(y), y \leq 0$$

$$\pi_Y(y) = 2(1-G(y)), y \geq 0$$

(3)

For arbitrary $X_1$ and $X_2$, it is not easy to obtain an analytical expression for $G$, but the later is available for frequently encountered cases. Corresponding equivalent possibility distributions are plotted in the figures of the following subsection 3.2 in a way such that the comparison with the comonotone case will be easy.

#### Gaussian distributions

One of the most important property which explains the abundant use in a lot of areas of the Gaussian (or normal) distribution is that it is stable under addition, i.e. the sum of two independent Gaussian random variables $N_1(m_1, \sigma_1)$ an $N_2(m_2, \sigma_2)$ is also a Gaussian random variable $N_y(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

#### Cauchy distribution

This distribution will be denoted $C(M,a)$ ($M$ being the location parameter, and $a$ the scale parameter,) and the associated distribution function has the following expression:

$$F(x) = \frac{1}{\pi} \arctg(\frac{x-M}{a})$$

This distribution has no mean nor variance (but the median and the mode are definite). The shape invariance by the convolution product is also satisfied (it also belongs to the family of stable distribution) and thus we have:

$$G(y) = \frac{1}{\pi} \arctg(\frac{y-M}{a_1 + a_2})$$
Uniform distribution

This distribution will be denoted \( U(a,b) \), it has a constant probability density on the interval \([a,b]\). Therefore, the distribution function is a straight line. A remarkable result is that the equivalent possibility distribution around the middle of the interval is a triangular symmetric possibility distribution [2]. The convolution product of two identical uniform distributions gives a triangular probability density, and consequently a parabolic possibility distribution (see figure 2).

Let us remark that thanks to a theorem from Birnbaum [15], the equivalent possibility distribution of the sum \( \mathcal{P}_Y \) of two identical independent unimodal symmetric random variables \( X_1 \) and \( X_2 \) is more specific than the sum \( \mathcal{P}_{Y'} \) (\( \mathcal{P}_{Y'}(y) \leq \mathcal{P}_Y(y) \)) of two identical independent symmetric random variables \( X'_1 \) and \( X'_2 \) those equivalent possibility distributions are respectively more specific than those of \( X_1 \) and \( X_2 \). Thus, the possibility distribution equivalent to the sum of two independent unimodal symmetric distributions having a bounded support is included in the equivalent possibility distribution of the sum of two uniform distributions having the same support. Indeed, the uniform distribution is the least specific one among the symmetric unimodal distributions with bounded support [11].

Symmetric Pareto distribution

The Pareto distribution is a particular case of power laws which is applied in physical and social sciences. More precisely, we consider the following symmetric Pareto probability density:

\[
p_{SP}(x) = \frac{1}{2(1+|x|)^2}
\]

Due to the exponent 2 in the denominator, this distribution has no variance (and also no mean). An analytical expression for the convolution product is computable even if it is not so easy (see some calculus elements in the Annex). The expression for the distribution function of the sum of two such symmetric Pareto distributions is thus:

\[
G(y) = \frac{1}{2} + \frac{1}{2} \ln(1+y) - \frac{1}{4} \left( \frac{2+y}{y.(1+y)} + \frac{1+y}{2+1+y} \right)
\]

The corresponding equivalent possibility distribution is then deduced according to equation (3).

3.2. Comonotone random variables

The comonotonicity of two random variables corresponds to an extremal case of positive dependence which is defined by the existence of two increasing functions \( h_1 \) and \( h_2 \) such that \( X_1 = h_1(U) \) and \( X_2 = h_2(U) \), \( U(0,1) \) being an uniform distribution function on \([0,1]\). The comonotonicity implies maximal correlation coefficient if the later exists. One fundamental property of the comonotonicity is that the quantile function of the sum is the sum of the quantile functions [14]:

\[
G^{-1}(\alpha) = F^{-1}_1(\alpha) + F^{-1}_2(\alpha), \forall \alpha \in [0,1].
\]

This property allows to directly determine the possibility distribution equivalent to \( Y \) from its alpha-cuts by using equation (2). In fact, the alpha-cuts of the sum is the sum of the alpha-cuts that is equivalent to interval calculus and similar to Zadeh’s extension principle. Let us also note that the resulting shape invariance for distributions is larger for comonotonicity than for independence. In particular shape invariance is satisfied for the four types of distribution considered in the following illustrations.

Uniform distribution

In this conventional case, we obtain the same resulting possibility distributions as the ones obtained by applying the Zadeh’s extension principles to the sum of the two equivalent triangular possibility distributions.

\[
\text{Fig. 2. Possibility distributions equivalent to the sum of two uniform distributions with a standard deviation of 1.}
\]

Gaussian distributions

\[
\text{Fig. 3. Possibility distributions equivalent to the sum of two Gaussian distributions with a standard deviation of 1.}
\]
In the comonotone case, the standard deviation of the sum is the sum of the standard deviation that corresponds to a maximal correlation coefficient of 1.

**Cauchy distribution**

In this very particular case, we obtain the same distribution function, and thus the same equivalent possibility distribution for the sum either for independence or comonotonicity.

![Graph of Cauchy distribution](image)

**Symmetric Pareto distributions**

In this case, independence leads to an equivalent possibility distribution less specific that the one obtained with comonotonicity.

![Graph of Symmetric Pareto distributions](image)

**Remark**

Let us note that for the first three distributions, the equivalent possibility distributions of the sum in the independence case are more specific that the ones obtained for the comonotonic case, in fact equal for the Cauchy distribution. On the other hand, it is the opposite for the symmetric Pareto distribution: independence leads to a less specific possibility distribution (i.e., larger dispersion intervals) than comonotonicity. Let us remark that the considered symmetric Pareto distribution has as well as the Cauchy distribution no mean nor variance, but it has a heavier tail.

### 3.3. Unknown dependency

In the case where no information about dependency is available, it is of value to find the possibility distribution associated to $Y$ which contains all the equivalent possibility distributions issued from all the dependence relationships. For this purpose, Frank, Nelsen et Schweizer results [16] can be used. Indeed, they provide bounds on the distribution function from Fréchet bounds on the joint distribution [17]. These bounds constitute in fact a $p$-box [18], from which it is possible to build a possibility distribution [9] (it is not mandatory optimal). For uniform, Gaussian and Cauchy distributions analytic expressions for the bounds of the sum of two identical random variables are available:

$$
\begin{align*}
G_{\text{min}}(y) & \leq G(y) \leq G_{\text{max}}(y), \\
G_{\text{min}}(y) &= 0, y \leq 0 \\
G_{\text{max}}(y) &= 2F(y/2), y \leq 0 \\
&= 1, y \geq 0
\end{align*}
$$

As we have $G_{\text{min}}(0) = 0$ and $G_{\text{max}}(0) = 1$, we obtain the following expression for the equivalent possibility distribution [9]:

$$
\begin{align*}
\pi_{\text{e}}(y) &= \min(G_{\text{min}}(y), 1 - G_{\text{max}}(y)), \text{ and} \\
\pi_{\text{e}}(y) &\leq 2F(y/2), y \leq 0 \\
&\leq 2(1 - F(y/2)), y \geq 0.
\end{align*}
$$

It is important to notice that the worst dependency relationship corresponds to the comonotonicity for the uniform, Gaussian and Cauchy distributions with equality with independence for the later distribution.

### 3.4. Discussion

The results of the preceding sub-section are only partly in accordance with the idea that the extremal positive dependence, i.e., comonotonicity, between two random variables leads to the least specific possibility distribution (i.e., the most spread in terms of dispersion intervals). They point out that the notion of non interactive variable introduced in the Zadeh’s extension principle fuzzy theory, and widely used in the fuzzy community, is indeed equivalent to the comonotonicity of random variables (this connection is mentioned in the Ph. D. thesis of Williamson [19]). Accordingly, the Zadeh’s extension principle is conservative for unimodal symmetric random variables with bounded support, but not necessarily if the support is indefinite. Indeed, the example of symmetric Pareto distributions shows that co-
monotonocity is not always the worst case. The fact that for the Cauchy distribution comonotonicity leads to the same result as independence suggest that for distributions with indefinite mean, independence is worse than comonotonicity. However, if we conjecture by analogy with the bounds of Frank et al., that the symmetric Pareto distribution satisfies equation (4), then independence would not be the worst case (see fig. 5), but this point has to be further studied. Finally let us also mention that for unimodal asymmetric random variables with bounded support, the Zadeh’s extension principle is not necessarily conservative as shown by an example described in [20].

4. Conclusion

The paper has addressed the issue of probability/possibility transformation of the sum of two identical continuous symmetric unimodal random variables considering either independence or comonotonicity of the variables. We emphasize the fact that the Zadeh’s extension principle applied to possibility distribution propagation leads to the same result as the one obtain by considering comonotone random variables. It has been shown that for the uniform, Gaussian and Cauchy distributions the worst case in terms of specificity is obtained with the comonotonicity. For the Cauchy distribution, independence and comonotonicity leads to the same equivalent possibility distribution, and for the symmetric Pareto distribution, independence leads to a worse equivalent possibility distribution than comonotonicity. When no information about dependency is available, Frank et al.’s bounds can be used to build a possibility distribution that is less specific than the ones obtained with assuming independence or comonotonicity.

The study has considered a lot of conditions for the sum of distributions functions (identical, symmetric, unimodal). Therefore further works have to be done to obtain more general results. In particular, to determine conditions in order Zadeh’s extension principle leads to the worst case, and also to determine the most specific possibility distribution dominating the optimal distribution for all dependency cases.

References


Annex

Hereafter, we present some calculus elements for the determination of the distribution function of the sum of two identical independent symmetric Pareto distributions:

\[ p_{SP}(x) = \frac{1}{2(1 + |x|)} \]

gives by integration:

\[ F_{SP}(x) = \frac{1}{2(1 - x)}, x \leq 0; \]

\[ F_{SP}(x) = 1 - \frac{1}{2(1 + x)}, x \geq 0 \]

The expression of the distribution function of the sum \( Y \) for \( x \geq 0 \) is:

\[ F_{SP}(x) = \frac{1}{2(1 - x)}, x \leq 0; \]

\[ F_{SP}(x) = 1 - \frac{1}{2(1 + x)}, x \geq 0 \]

The expression of the distribution function of the sum \( Y \) for \( x \geq 0 \) is:

\[ G_Y(y) = \int_0^y F_{SP}(y - x)p_{SP}(x)dx \]

\[ + \int_0^y F_{SP}(y - x)p_{SP}(x)dx + \int_0^\infty F_{SP}(y - x)p_{SP}(x)dx \]

Further, each of the three integrals is computed by making a decomposition in simple elements that leads to the following result:

\[ G(y) = \frac{1}{2} + \frac{1}{2\ln(1 + y)} + \frac{(2 + y)}{4y(1 + y)} + \frac{1}{2(1 + y)} - \frac{1}{2\ln(1 + y)} \]

\[ - \frac{1}{2\ln(1 + y)} + \frac{1}{4(1 + y)(3 + 2y + y^2)} \]