Representing uncertainty by possibility distributions encoding confidence bands, tolerance and prediction intervals

Mohammad Ghasemi Hamed, Mathieu Serrurier, Nicolas Durand

HAL Id: hal-00938794

https://hal-enac.archives-ouvertes.fr/hal-00938794

Submitted on 24 Apr 2014

Representing uncertainty by possibility distributions encoding confidence bands, tolerance and prediction intervals

Mohammad Ghasemi Hamed [1,2], Mathieu Serrurier[1], and Nicolas Durand[1,2]

[1] IRIT - 118 route de Narbonne 31062, Toulouse Cedex 9, France
[2] ENAC/MAIAA - 7 avenue Edouard Belin 31055 Toulouse, France

Abstract. For a given sample set, there are already different methods for building possibility distributions encoding the family of probability distributions that may have generated the sample set. Almost all the existing methods are based on parametric and distribution free confidence bands. In this work, we introduce some new possibility distributions which encode different kinds of uncertainties not treated before. Our possibility distributions encode statistical tolerance and prediction intervals (regions). We also propose a possibility distribution encoding the confidence band of the normal distribution which improves the existing one for all sample sizes. In this work we keep the idea of building possibility distributions based on intervals which are among the smallest intervals for small sample sizes. We also discuss the properties of the mentioned possibility distributions.

Keywords: imprecise probabilities, possibility distribution, confidence band, confidence region, tolerance interval, prediction interval, normal distribution, distribution free.

1 Introduction

In 1978, Zadeh introduced the possibility theory [31] as an extension of his theory on fuzzy sets. Possibility theory offers an alternative to the probability theory when dealing with some kinds of uncertainty. Possibility distribution can be viewed as a family of probability distributions. Then, the possibility distribution contains all the probability distributions that are respectively upper and lower bounded by the possibility and the necessity measure [10]. For a given sample set, there are already different methods for building possibility distribution which encodes the family of probability distributions that may have generated the sample set[3, 25, 2]. The mentioned methods are almost all based on parametric and distribution free confidence bands.

In this paper we review some methods for constructing confidence bands for the normal distribution and for constructing distribution free confidence bands (γ-C distribution). Then we propose a possibility distribution for a sample set drawn from an unknown normal distribution based on Frey [15] confidence band

which improves the existing possibility distribution proposed by Aregui et al. [2] for all sample sizes. We also introduce possibility distribution which encodes tolerance intervals, named as γ-Confidence Tolerance Possibility distribution (γ-CTP distribution). The proposed possibility distribution uses tolerance intervals to build the maximal specific possibility distribution that bounds each population quantile of the true distribution (with a fixed confidence level) that might have generated our sample set. The distribution obtained will bound each confidence interval of inter-quantiles independently. This latter is different from a possibility distribution encoding a confidence band, because a possibility distribution encoding a confidence band will simultaneously bounds all population quantiles of the true distribution (with a fixed confidence level) that might have generated our sample set. Finally, we consider possibility distributions encoding prediction intervals (prediction possibility distribution). In this case, each α-cut will contain the next observation with a confidence level equal to $1 - \alpha$. Each of the proposed possibility distributions encodes a different kind of uncertainty that is not expressed by the other ones. We show that γ-C distribution is always less specific than γ-CTP distribution which is itself less specific than the prediction possibility distribution. This is due to the fact that the distributions properties are less and less strong. Note that the confidence level is usually chosen by the domain expert.

This paper is structured as follows: we begin with a background on possibility theory. Then we review possibility distribution encoding confidence bands and their relationship with confidence regions. In this section we introduce a method which improves existing possibility distributions. Next we see how to encode tolerance intervals and prediction intervals by possibility distributions. Finally, we end with a discussion on the mentioned possibility distributions and some illustrations.

2 Background

2.1 Possibility theory

Possibility theory, introduced by Zadeh [31, 12], was initially created in order to deal with imprecision and uncertainty due to incomplete information. This kind of uncertainty may not be handled by probability theory, especially when a priori knowledge about the nature of the probability distribution is lacking. In possibility theory, we use a membership function π to associate a distribution on the universe of discourse Ω. In this paper, we only consider the case $\Omega = \mathbb{R}$.

Definition 1 *A possibility distribution π is a function from Ω to $(\mathbf{R} \to [0, 1])$.*

The definition of the possibility measure Π is based on the possibility distribution π such that:
$$\Pi(A) = sup(\pi(x), \forall x \in A). \tag{1}$$
The necessity measure is defined by the possibility measure

$$\forall A \subseteq \Omega, N(A) = 1 - \Pi(A^C) \tag{2}$$

where A^C is the complement of the set A. A distribution is normalized if : $\exists x \in \Omega$ such that $\pi(x) = 1$. When the distribution π is normalized, we have : $\Pi(\emptyset) = 0, \Pi(\Omega) = 1$.

Definition 2 *The α-cut A_α of a possibility distribution $\pi(\cdot)$ is the interval for which all the point located inside have a possibility membership $\pi(x)$ greater or equal to α.*

$$A_\alpha = \{x | \pi(x) \geq \alpha, x \in \Omega\}, \tag{3}$$

2.2 Possibility distribution encoding a family of probability distribution

In fact, one interpretation of possibility theory, based on Zadeh's [31] consistency principle of possibility "what is probable should be possible", is to consider a possibility distribution as a family of probability distributions (see [10] for an overview). In the following, we denote f as density function, F as its Cumulative Distribution Function (CDF) and P as its probability measure. Thus, a possibility distribution π will represent the family of the probability distributions Θ for which the measure of each subset of Ω's will be bounded by its possibility measures :

Definition 3 *A possibility measure Π is equivalent to the family Θ of probability distribution F such that*

$$\Theta = \{F | \forall A \in \Omega, P(A) \leq \Pi(A)\}, A \subseteq \Omega. \tag{4}$$

Definition 4 *Given a probability density function $f(\cdot)$ with finite number of modes, we define the interval I^*_β defined below as "smallest β-content interval" of f.*

$$I^*_\beta = \{x | x \in f^{-1}[d], \forall d \in [c, +\inf)\} \tag{5}$$

where $\int_{\{x|f(x) \geq c\}} f(x)dx = \beta$ and $c > 0$.

We know that $Pr(I^*_\beta) = \beta$ and this interval is unique only if f has a finite number of modes. Now let θ be a set of Cumulative Distribution Function (CDF) F defined by a possibility distribution function $\pi(\cdot)$. Thus, an alternative to equations (4) is:

$$\forall \alpha \in [0, 1], \forall F \in \Theta, I^*_{F,\beta} \subseteq A_{\pi,\alpha} \tag{6}$$

where $\beta = 1 - \alpha$ and $A_{\pi,\alpha}$ is the α-cut of possibility distribution $\pi(\cdot)$. Thus, a possibility distribution encodes a family of probability distributions for which each quantile is bounded by a possibility α-cut. By considering the definition of necessity, we obtain the following inequalities:

$$N(A) \leq P(A) \leq \Pi(A), A \subset \Omega. \tag{7}$$

Thus by using the possibility and necessity measures, like in the Dempster-Shafer theory, we can define upper and lower values to describe how an event is likely to occur.

2.3 Probability-possibility transformation

In many cases it is desirable to move from the probability framework to the possibility framework. This is why several transformations based on various principles such as consistency (what is probable is possible) or information invariance have already been proposed [8, 9, 22, 13, 11]. Dubois et al.[14] suggest that when moving from probability to possibility framework we should use the "maximum specificity" principle which aims at finding the most informative possibility distribution. Formally the maximum specificity principle is defined as follow:

Definition 5 *Given the maximum specific possibility distribution (m.s.p.d) π^* that encodes the probability distribution function F (i.e. $\forall A \subseteq \Omega, N^*(A) \leq P(A) \leq \Pi^*(A)$) we have, for all π such as $\forall A \subseteq \Omega, N(A) \leq P(A) \leq \Pi(A)$, $\pi^*(x) \leq \pi(x), \forall x \in \Omega$.*

Because the possibility distribution explicitly handles the imprecision and is also based on an ordinal structure rather than an additive one, it has a weaker representation than the probability one. This kind of transformation (probability to possibility) may be desirable when we are in presence of weak source of knowledge or when it is computationally harder to work with the probability measure than with the possibility measure. The "most specific" possibility distribution is defined for a probability distribution having a finite number of modes and the equation is as below [11] :

$$\pi_t(x) = sup(1 - P(I_\beta^*), x \in I_\beta^*) \tag{8}$$

where π_t is the "most specific" possibility distribution, I_β^* is the smallest β-content interval [11]. Then, in the spirit of equation 6, given f and its transformation π^* we have :

$$A_\alpha^* = I_\beta^* \quad where \ \alpha = 1 - \beta.$$

Figure (1) presents the maximum specific transformation (in blue) of a normal probability distribution (in green) with mean and variance respectively equal to $0, 1$ ($\mathcal{N}(0, 1)$).
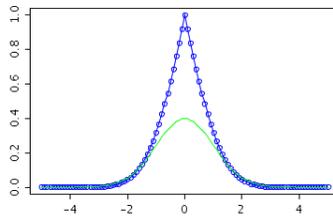

Fig. 1. The m.s.p.d for $\mathcal{N}(0, 1)$.

Definition 6 *The interval between the lower and upper quantiles of the same level are called inter-quantiles. The inter-quantile at level p is defined by*

$$[F^{-1}(p), F^{-1}(1-p)], 0 < p < 0.5 \tag{9}$$

where $F^{-1}(\cdot)$ is the inverse function of the continuous strictly-monotone CDF $F(\cdot)$.

An inter-quantile at level p contains β proportion of the distribution where $\beta = 1 - 2p$. We will call a β-content inter-quantile I_β, the interval that contains β proportion of the underlying distribution, we have $Pr(I_\beta) = \beta$.

Proposition 1 *The maximum specific possibility distribution (m.s.p.d) $\pi^*(\cdot)$ of unimodal symmetric probability density function $f(\cdot)$ can be built just by calculating the β-content inter-quantile I_β of $f(\cdot)$ for all the values of β, where $\beta \in [0, 1]$.*

It is evident that for any unimodal symmetric p.d.f $f(\cdot)$, *the smallest β-content interval I_β^* of f is also its inter-quantile at level $\frac{1-\beta}{2}$.* Therefore the proposition 1 can be used.

3 Possibility distribution encoding confidence bands

3.1 Confidence Band

Definition 7 *The confidence band for a CDF F is a function which associates to each x an interval $[L(x), U(x)]$ such as :*

$$P(\forall x, L(x) \le F(x) \le U(x)) \ge \gamma \text{ where } \forall x, 0 \le L(x) \le U(x) \le 1 \tag{10}$$

In frequentist statistics, a confidence band is an interval defined for each value x of the random variable X such that for a repeated sampling, the frequency of $F(x)$ located inside the interval $[L(x), U(x)]$ for all the values of X tends to the confidence coefficient γ.

Note that given any γ level confidence band, we can use it to infer confidence intervals of the quantile function $Q(\beta) = F^{-1}(\beta) = \inf\{x \in R : \beta \le F(x)\}$, and that for all $\beta \in (0, 1)$. In other word the confidence band gives simultaneously confidence intervals for all $F^{-1}(\beta), \beta \in (0, 1)$. Therefore such confidence intervals derived from confidence bands are *Simultaneous confidence Intervals (SMI) for all population quantiles*. We can take advantage of this property to derive simultaneous γ-confidence intervals for β-content inter-quantiles of the unknown CDF $F(\cdot)$ and we will denote them by I_β^C.

By using proposition (1) and tables of confidence band stated in the statistic literature [21, 15, 7, 5, 20, 1], we can encode simultaneous γ-confidence intervals for β-content inter-quantiles I_β^C, of an unknown CDF $F(\cdot)$ by a possibility distribution represented by π_γ^C:

$$\pi_\gamma^C(x) = 1 - \max_{x \in I_{1-\alpha}^C} (\alpha) \text{ where } A_\alpha = I_\beta^C, \beta = 1 - \alpha \tag{11}$$

By construction, the obtained distribution has the following property:

Proposition 2 *Let π_γ^C be a possibility distribution obtained by equation (11) we have:*

$$P(\forall x, \forall \alpha \in (0,1), P(x \in A_\alpha) \geq 1 - \alpha) \geq \gamma$$

Suppose that K is a set of estimated CDF $\hat{F}$ for F, and that in a repeated sampling the frequency of the function $\hat{F}$ being equal to the true CDF F ($\forall x, \hat{F}(x) = F(x)$), tends to $1 - \alpha$. The function set K will be a confidence band for $F(x)$. This is expressed formally below:

$$P(\exists \hat{F} \in K, F = \hat{F}) = 1 - \alpha \tag{12}$$

Equations (10) and (12) are two different views representing confidence band. In equations (10), the goal is to find a set K composed of estimated cumulative distribution functions being one "choice" for estimating the true CDF F. When F belongs to a parametric family, we can use the confidence region of its parameter vector to construct its confidence band [21, 7, 15]. Therefore, confidence bands build by considering confidence regions are described by equation (12).

3.2   Possibility distribution encoding normal confidence bands

As we saw above, we can use the confidence region of parameter of a probability distribution to infer its confidence band. Cheng and Iles [7] and kanofsky [21] used this approach to infer the confidence band of the normal distribution. Aregui et al. [3], proposed to construct possibility distributions for sample set drawn from a known parametric family with an unknown parameter vector. Their possibility distribution encoded the Cheng et al. [7] confidence band. In another paper, Aregui et al. [2] used confidence region for the parameters of the normal distribution to infer a possibility distribution. This possibility distribution encodes all normal distributions having their parameters inside the desired confidence region of the "true parameters" belonging to the "real distribution" that has generated this sample set. As we saw previously, encoding the confidence region of parameters results in a possibility distribution which encodes the whole confidence band of the normal distribution. The band encoded by their method was built by the "Smallest Mood exact" confidence region [4]. The "Smallest Mood exact" region contains exactly the desired confidence level and it was the the second smallest confidence region (after the "likelihood-ratio test") in [4]. This region is easy to obtain and is particularly useful for small sample sizes. (Note that if we want to construct the m.s.p.d encoding confidence band of a normal distribution, we have to find the smallest possible SMIs which leads us to use the tightest confidence band.) In [15], Frey proposed the minimum-area confidence region and the minimum area based confidence band for the normal distribution. She showed that her minimum area confidence band improves other bands for all sample sizes. In the same way we propose a possibility distribution which encodes the Frey confidence band. In figures (2,3) we compared our possibility distribution named "0.95 Frey C.P.D." (0.95 Frey Confidence Possibility Distribution) which is encoded in blue with the Mood based and Smallest Mood based confidence possibility distribution.

Proposition 3 *The Frey normal confidence band improves the confidence band resulted by the "Smallest Mood exact" region, in area sense, and this for all sample sizes.*

Sketch of proof : The "smallest Mood confidence region" is found by the optimal confidence levels ψ_1 and ψ_2 for the mean and variance confidence intervals and also for each sample size such that $(1-\psi_1)(1-\psi_1) = \gamma = 0.95$ [4]. We can observe that the Frey confidence band for $\gamma = 0.95$ gives confidence band with smaller area (integral over all confidence intervals of inter-quantiles) than the band issued from a Mood confidence region $(1-\psi)^2 = 0.95^2$ where $(1-\psi) = 0.95$ and this for all sample sizes. We know that we the confidence band built from Mood confidence region $(1-\psi)^2$ where $(1-\psi) = 0.95$ is everywhere smaller than the confidence band built from the "smallest Mood confidence region" for $\gamma = 0.95$. The statement above holds for every value of $\gamma \in (0.01, 1)$ ∎

3.3 Possibility distribution encoding distribution free confidence bands

For distribution free confidence bands, the most known method is the Kolmogorof [5] statistic for small sample sizes and the Kolmogorof-Smirnof test for large sample sizes. Some other methods have also been suggested based on the weighted version of the Kolmogorof-Smirnof test [1]. Owen also proposed a nonparametric likelihood confidence band for a distribution function. Remark that Owen's nonparametric likelihood bands are narrower in the tails and wider in the center than Kolmogorov-Smirnov bands and are asymmetric on the empirical cumulative distribution function. Frey [20] suggested another approach in which the upper and lower bounds of the confidence band are chosen to minimize a narrowness criterion and she compared her results to other methods. The optimal bands have a nice property : by choosing appropriate weights, you may obtain bands that are narrow in whatever region of the distribution is of interest. Masson et al. [25], suggested simultaneous confidence intervals of the the multinomial distribution to build possibility distributions and in another paper, Aregui et al. [3] proposed the Kolmogorof confidence band [5] to construct predictive belief functions[26] for sample set drawn from an unknown distribution. Thus, we propose to use the Frey band to construct the possibility distribution since it allows to have narrower α-cuts for the α's of interest.

4 Possibility distribution encoding tolerance interval

A tolerance interval, is an interval which guarantees with a specified confidence level γ, to contain a specified proportion β of the population. Confidence bounds or limits are endpoints within which we expect to find a stated proportion of the population. As the sample set grows, a parameter's confidence interval downsizes toward zero. In the same way, increasing the sample size leads the tolerance

interval bounds to converge toward a fixed value. We name a $100\%\beta$ tolerance interval(region) with confidence level $100\%\gamma$, a β-content γ-coverage tolerance interval (region) and we represent it by $I_{\gamma,\beta}^T$.

Having a sample set which come from a CDF $F(\cdot)$ with unknown parameters and for a given confidence level γ, we encode all the β-content γ-coverage tolerance intervals of $F(\cdot)$, $\forall \beta \in (0,1)$, by a possibility distribution and we name it "γ-confidence tolerance possibility distribution" (γ-CTP distribution represented by π_γ^{CTP}). When we do not know the distribution of the sample set, we can use β-content γ-coverage distribution free tolerance intervals ,$\forall \beta \in (0,1)$, of the unknown probability distribution in order to build Distribution Free γ-Confidence Tolerance Possibility (γ-DFCTP distribution represented by π_γ^{DFCTP}) distribution. The possibility distributions π_γ^{CTP} and π_γ^{DFCTP} will have by construction, the following property:

Proposition 4 *Let π_γ^{CTP} (or π_γ^{DFCTP}) be a possibility distribution that encodes tolerance intervals, we have:*

$$\forall \alpha \in (0,1),\ P(\forall x,\ P(x \in A_\alpha) \geq 1 - \alpha) \geq \gamma \text{ where } A_\alpha = I_{\gamma,\beta}^T, \beta = 1 - \alpha$$

Note that, it may also be interesting to fix the proportion β and make the confidence coefficient vary , $\gamma \in (0,1)$, to have a β-content tolerance possibility distribution.

Possibility distribution encoding tolerance interval for the normal distribution When our sample set comes from a univariate normal distribution, the lower and upper tolerance bounds (x_l and x_u ,respectively) are calculated by formulas (13) and (14) where, $\bar{X}$ is the sample mean, S the sample standard deviation, $\chi_{1-\gamma,n-1}^2$ represents the p-value of the chi-square distribution with $n - 1$ degree of freedom and $Z_{1-\frac{1-\beta}{2}}^2$ is the squared of the critical value of the standard normal distribution with probability $(1 - \frac{1-\beta}{2})$ [19]. Hence, the boundaries of a β-content γ-coverage tolerance interval for a random sample of size n drawn from an unknown normal distribution are defined as follows:

$$x_l = \bar{X} - \mathbf{k}S,\ x_u = \bar{X} + \mathbf{k}S \tag{13}$$

$$\mathbf{k} = \sqrt{\frac{(n-1)(1+\frac{1}{n})Z_{1-\frac{1-\beta}{2}}^2}{\chi_{1-\gamma,n-1}^2}} \tag{14}$$

For more details on tolerance intervals see [16].

By using proposition (1), we can find the boundaries of the $(1-\alpha)$-cut $A_{1-\alpha} = [x_l, x_u]$ of the possibility distribution which are calculated by (13), then we obtain the possibility distribution π_γ^{CTP} as computed below, where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

$$\pi_\gamma^{CTP}(x) = 2(1 - \Phi\left(\sqrt{\frac{\chi_{(1-\gamma,n-1)}^2 (\frac{x-\bar{X}}{S})^2}{(n-1)(1+\frac{1}{n})}}\right)) \tag{15}$$

Possibility distribution encoding distribution free tolerance interval
Let $\{x_1, x_2, \cdots, x_n\}$ be n independent observations from the random variable X and let $f(x)$ be its continuous probability density function. A distribution free tolerance region is the region between two tolerance limits where the probability that this region contains β proportion of the unknown probability distribution function is equal to γ . The mentioned tolerance limits are functions $L_1(x_1, x_2, \cdots, x_n) = x_r$ and $L_2(x_1, x_2, \cdots, x_n) = x_s$ constructed based on the order statistics of the observations.

$$\int_{x_s}^{x_r} f(x)dx \geq \beta, \tag{16}$$

In order to find the distribution free β-content γ-coverage tolerance interval (region) of continuous random variable x, we have to find the smallest n and the order statistics x_r and x_s for which the probability that equation (16) holds is greater or equal to γ. Equation (16) has a sampling distribution which was first defined by Wilks [30] for a univariate random variable with symmetrical values of r and s. Wilks [30] definition puts the following constraints on r and s: $s = n - r + 1$ and $0 < r < s \leq n$. Wald [27] proposed a non-symmetrical multivariate generalization of the Wilks method. Note that because distribution free tolerance intervals are based on order statistics, the sample size required for a given distribution free tolerance interval may increase with the interval's confidence level (γ) or the interval's proportion β. For example, in order to have 95% 0.99-content tolerance interval between the first and last element of a sample set, using formula in [18], we need $n = 473$. For the calculation of the sample size requirement for tolerance intervals. The reader can refer to [16] and [18].

The construction of possibility distribution based on distribution free tolerance intervals (region) raises some problems, because for a given sample set there are many ways to choose the r and s order statistics. If we choose them symmetrically such that $r = n - s + 1$, then the possibility distribution which encodes these intervals does not guarantee that its α-cuts include the mode and the α-cuts are neither the smallest ones. In fact, for any symmetric unimodal distribution, if we choose r and s order statistics in a symmetrical way, we will have tolerance intervals which are also the smallest possible ones and also include the mode of the distribution (see proposition (1)). Thus the Distribution Free γ-Confidence Tolerance Possibility (π_γ^{DFCTP}) distribution is constructed by equation below where x_r and x_s are the limits for the distribution free $I_{\gamma,\beta}^T$ of our sample set.

$$\pi_\gamma^{DFCTP}(x) = 1 - \max_{x \in I_{\gamma,1-\alpha}^T} (\alpha) \; where \; A_\alpha = I_{\gamma,\beta}^T = [x_r, x_s], \beta = 1 - \alpha$$

5   Possibility distribution encoding prediction intervals

Let us now define a prediction interval and its associated possibility distribution. A prediction interval uses past observations to estimate an interval for what

the future values will be, however other confidence intervals and credible intervals of parameters give an estimate for the unknown value of a true population parameter.

Definition 8 *Let $x_1, x_2, \cdots, x_n$ be a random sample drawn from an arbitrary distribution, then the interval $[x_l, x_u]$ is a $100\%(1-\alpha)$ prediction interval such that:*

$$P\left(x_l \leq x \leq x_u\right) = 1 - \alpha.$$

The prediction interval for the next future observation from a normal distribution is given by [17]:

$$\frac{x_{n+1} - \overline{X}_n}{S\sqrt{1+1/n}} \sim t_{n-1} \tag{17}$$

$$I_\beta^{Prev} = [\overline{X}_n - t_{(\frac{\alpha}{2}, n-1)} S\sqrt{1 + \frac{1}{n}}, \overline{X}_n + t_{(1-\frac{\alpha}{2}, n-1)} S\sqrt{1 + \frac{1}{n}}] \tag{18}$$

So the equation (18) gives a two tailed $1 - \alpha$ prediction interval for the next observation x_{n+1} , where $\overline{X}_n$ represents the estimated mean from the n past observations , $t_{(1-\frac{\alpha}{2}, n-1)}$ is the $100(\frac{1+p}{2})th$ quantile of Student's t-distribution with $n-1$ degrees of freedom and $\beta = 1 - \alpha$.

By using proposition (1) and equation (18), we can infer a prediction possibility (π^{Prev}) distributions for a sample set which comes from a normal distribution with an unknown mean and variance. π^{Prev} is computed as below where $T_{n-1}(\cdot)$ is the CDF of the Student distribution with $n-1$ degree of freedom. Equation (5) shows the properties of the α-cuts of π^{Prev}.

$$\pi^{Prev}(x) = 2(1 - T_{n-1}\left(\left|\frac{x_{n+1} - \overline{X}_n}{S\sqrt{1+1/n}}\right|\right)) \tag{19}$$

By construction, the obtained distribution has the following property:

Proposition 5 *Let π^{prev} be a possibility distribution that encodes prediction intervals using equation (19) build from a random sample set $X = \{x_1, \ldots, x_n\}$ we have:*

$$\forall \alpha \in (0,1), \ P(x_{n+1} \in A_\alpha) \geq 1 - \alpha \ \text{where} \ A_\alpha = I_\beta^{Prev}$$

For distribution free prediction intervals, the reader can find more information in [16],[23] and [6].

6   Discussion and Illustrations

We have seen three different types of intervals and their encoding possibility distributions. The most known approach is to choose the possibility distribution which is encoded by confidence bands. However, depending on the application, we might be interested to infer other possibility distributions than the one that

encodes conventional SMIs. We can deduce, from the propositions 2, 4 and 5, that:

$$\forall x, \ \pi_\gamma^C(x) \geq \pi_\gamma^{CTP}(x) \geq \pi^{Prev}(x).$$

The choice of the distributions depends on the application. In this work, because of the lack of space, we just focused on two-sided intervals. It might be useful to construct one-sided possibility distributions encoding one-sided intervals. One can also be interested to encode percentiles of a distribution with a possibility distribution, but note that instead of using possibility distribution encoding two-sided percentiles we can use π_γ^{CTP} (two-sided). For more information the reader can refer to [16]. The reviewed distributions can be used for different purpose in uncertainty management. Wallis [29] used the Wald et al.[28] normal tolerance limits to find tolerance intervals for linear regression. In the same way, we can use our γ-CTP distribution to build probabilistic regression which encodes tolerance bounds of the response variable. Note that we are not restricted to linear possibilistic linear regression with homoscedastic and normal errors. We can also apply our γ-CTP and γ-DFCTP distributions to do possibilistic non-parametric and parametric regression with homoscedastic or heteroscedastic errors.

Figure (2) shows the $\pi_{0.95}^C$ for a sample set of size 10 with sample mean and sample variance respectively equal to 0 and 1, figure (2) represents the same concept for $n = 25$. This figure illustrates the proposition 3. Indeed, we can see that our possibility distribution is more informative than the Aregui et al. possibility distribution. In figure (5) the blue color is used to represent π^{Prev}
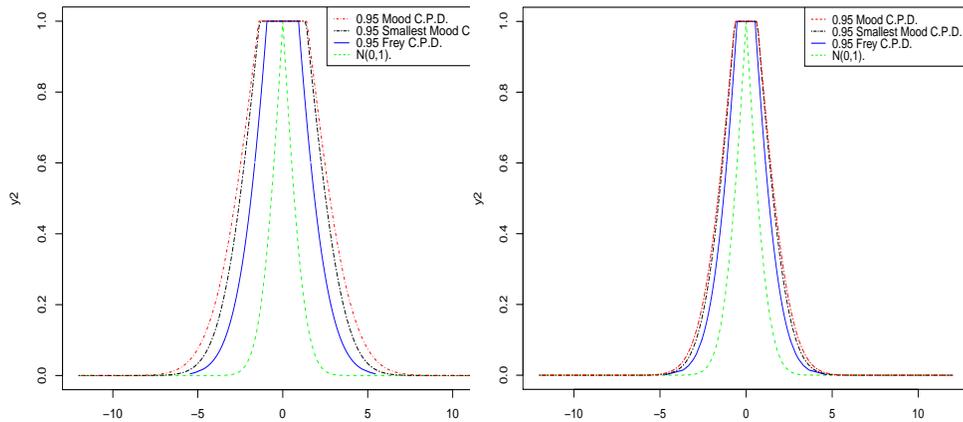

Fig. 2. Possibility distribution encoding normal confidence band for a sample set of size 10 having $(\bar{X}, S) = (0, 1)$.

Fig. 3. Possibility distribution encoding normal confidence band for a sample set of size 25 having $(\bar{X}, S) = (0, 1)$.

for different sample sets drawn from the normal distribution, all having the same sample parameters, $(\overline{X}, S) = (0,1)$ but different sample sizes. The green distribution represents the probability-possibility transformation of $\mathcal{N}(0,1)$.
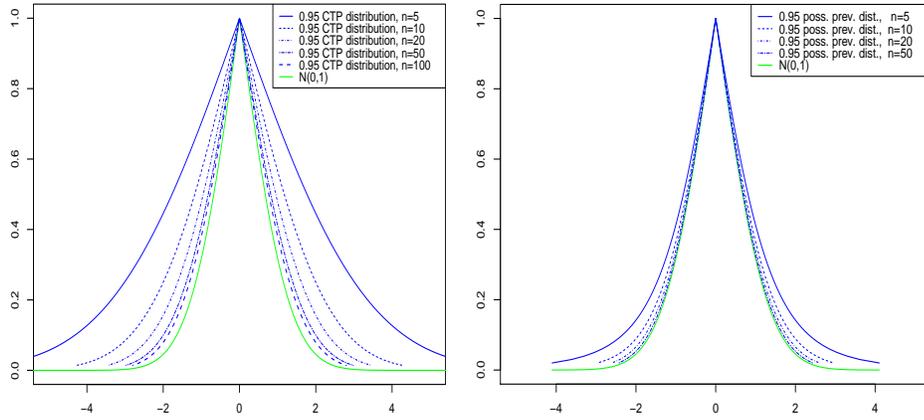

Fig. 4. 0.95-confidence tolerance possibility distribution for different sample sizes having $(\overline{X}, S) = (0,1)$.

Fig. 5. 0.95-confidence prevision possibility distribution for different sample sizes having $(\overline{X}, S) = (0,1)$.

In figure (4) we used the previous settings for the $\pi_{0.95}^{CTP}$. Note that, for $n \geq 100$, the tolerance interval is approximately the same as the maximum likelihood estimated distribution. In figure 6, the blue curves represents the $\pi_{0.95}^{DFCTP}$ for a sample set of size 450, drawn from $\mathcal{N}(0,1)$ and the green distribution represents the probability-possibility transformation for $\mathcal{N}(0,1)$. In figure (7), we used two different sample sets with $n = 194$ to build two different $\pi_{0.9}^{DFCTP}$. In this example, in order to reduce the required sample size, we restricted the biggest β to 0.98.

7 Conclusion

In this work, we proposed different possibility distributions encoding different kind of uncertainties. We also proposed a possibility distribution encoding confidence band of the normal distribution which improves the existing one for all sample sizes. Building possibility distributions which encode tolerance intervals and prediction intervals are also new concepts that we introduced in this work. For future works, we propose to build in the same way the possibility distributions encoding distribution free tolerance regions [27] and tolerance regions for
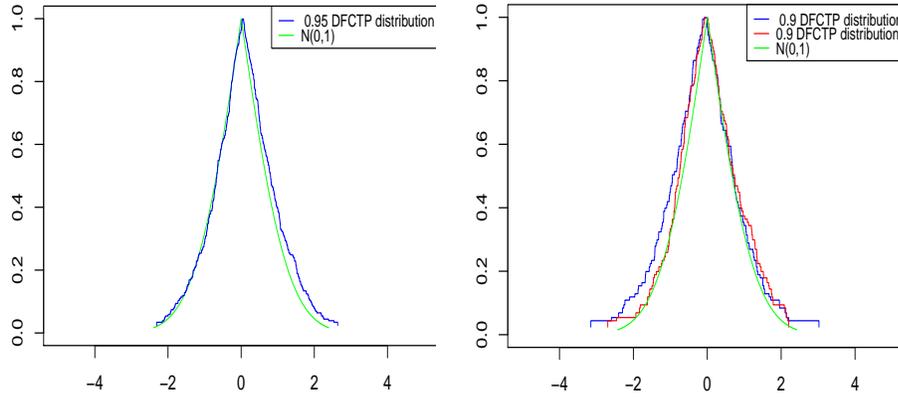
Fig. 6. Distribution free 0.95-confidence tolerance possibility distribution for a sample set with size 450 drawn from $\mathcal{N}(0,1)$.

Fig. 7. Distribution free 0.9-confidence tolerance possibility distributions for a sample set with size 194 drawn from $\mathcal{N}(0,1)$.

the multivariate normal distribution [24]. We also propose to use our distributions for possibilistic regression.

Acknowledgement We would like to thank Prof. Jesse Frey for her help and guideline during this work.

References

1. T. W. Anderson and D. A. Darling. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. of Math. Stat.*, 23(2):193–212, 1952.
2. A. Aregui and T. Denoeux. Consonant belief function induced by a confidence set of pignistic probabilities. In *ECSQARU*, pages 344–355. Springer Berlin / Heidelberg, 2007.
3. A. Aregui and T. Denœux. Constructing predictive belief functions from continuous sample data using confidence bands. In *ISIPTA*, pages 11–20, July 2007.
4. B.C. Arnold and R.M. Shavelle. Joint confidence sets for the mean and variance of a normal distribution. *Am. Stat.*, 52(2):133–140, 1998.
5. Z. W. Birnbaum. Numerical tabulation of the distribution of kolmogorov's statistic for finite sample size. *J. Amer. Statistical Assoc.*, 47(259):425–441, 1952.
6. S. Chakraborti and P. van der Laan. Precedence probability and prediction intervals. *J. Roy. Stat.Society. Series D (The Statistician)*, 49(2):219–228, 2000.
7. R. C. H. Cheng and T. C. Iles. Confidence bands for cumulative distribution functions of continuous random variables. *Technometrics*, 25(1):77–86, 1983.

8. M. R Civanlar and H. J. Trussell. Constructing membership functions using statistical data. *Fuzzy Sets Syst.*, 18:1–13, January 1986.

9. M. Delgado and S. Moral. On the concept of possibility-probability consistency. *Fuzzy Set. Syst.*, 21(3):311–318, 1987.

10. D. Didier. Possibility theory and statistical reasoning. *Compu. Stat. Data An.*, 51:47–69, 2006.

11. D. Dubois, L. Foulloy, G. Mauris, and H. Prade. Probability-possibility transformations, triangular fuzzy sets and probabilistic inequalities. *Rel. Comp.*, 2004.

12. D. Dubois and H. Prade. *Fuzzy sets and systems - Theory and applications.* Academic press, New York, 1980.

13. D. Dubois and H. Prade. Fuzzy sets and probability : Misunderstandings, bridges and gaps. In *IEEE Fuzzy Sys.*, pages 1059–1068, 1993.

14. D. Dubois, H. Prade, and Sandra Sandri. On possibility/probability transformations. In *IFSA*, pages 103–112, 1993.

15. J. Frey, O. Marrero, and D. Norton. Minimum-area confidence sets for a normal distribution. *J. Stat. Plan. Inf.*, 139:1023–1032, 2009.

16. G. J. Hahn and W. Q. Meeker. *Statistical Intervals: A Guide for Practitioners.* John Wiley and Sons, 1991.

17. G.J. Hahn. Factors for calculating two-sided prediction intervals for samples from a normal distribution. *JASA*, 64(327):878–888, 1969.

18. D. L. Hanson and D. B. Owen. Distribution-free tolerance limits elimination of the requirement that cumulative distribution functions be continuous. *Technometrics*, 5(4):518–522, 1963.

19. W. G. Howe. Two-sided tolerance limits for normal populations, some improvements. *J. Amer. Statistical Assoc.*, 64(326):610–620, 1969.

20. Jesse and Frey. Optimal distribution-free confidence bands for a distribution function. *J. Stat. Plan. Inf*, 138(10):3086 – 3098, 2008.

21. Paul Kanofsky and R. Srinivasan. An approach to the construction of parametric confidence bands on cumulative distribution functions. *Biometrika*, 59(3):623–631, 1972.

22. G.J. Klir. A principle of uncertainty and information invariance. *Internat. J. General Systems*, 17(23):249–275, 1990.

23. H. S. Konijn. Distribution-free and other prediction intervals. *Am. Stat.*, 41(1):11–15, 1987.

24. K. Krishnamoorthy and Thomas Mathew. Comparison of approximation methods for computing tolerance factors for a multivariate normal population. *Technometrics*, 41(3):234–249, 1999.

25. M. Masson and T. Denoeux. Inferring a possibility distribution from empirical data. *Fuzzy Sets Syst.*, 157:319–340, February 2006.

26. P. Smets. The transferable belief model and other interpretations of dempster-shafer's model. pages 326–333, 1990.

27. A. Wald. An extension of wilks' method for setting tolerance limits. *Ann. of Math. Stat.*, 14(1):45–55, 1943.

28. A. Wald and J. Wolfowitz. Tolerance limits for a normal distribution. *The Annals of Mathematical Statistics*, 17(2):208–215, 1946.

29. W. A. Wallis. Tolerance intervals for linear regression. In *Proc. Second Berkeley Symp. on Math. Statist. and Prob.*, pages 43–51, 1951.

30. S. S. Wilks. Determination of sample sizes for setting tolerance limits. *Ann. of Math. Stat.*, 12(1):91–96, 1941.

31. L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Set. Syst.*, 1(1):3–28, 1978.