A new transformation of continuous unimodal asymmetric probability distributions into possibility distributions

Laurent Foulloy, Gilles Mauris

Version of Record: https://www.sciencedirect.com/science/article/pii/S0165011420305157
Manuscript_4932d87419b9b0e31e96d5b7e3dde8f3

A new transformation of continuous unimodal asymmetric probability distributions into possibility distributions

Laurent Foulloy, Gilles Mauris

Laurent.Foulloy@univ-smb.fr, Gilles.Mauris@univ-smb.fr

University Savoie Mont Blanc, LISTIC, Annecy, France

Contact information for the corresponding author:

Professor Laurent Foulloy

University Savoie Mont Blanc

Polytech Annecy-Chambéry, LISTIC

5 Chemin de bellevue

74940 Annecy

France

Abstract

This paper deals with a new transformation, so-called two-sided normalized (TSN), of continuous unimodal asymmetric probability distributions into possibility distributions. Many properties are derived and interpretations are discussed. A comparison with the optimal transformation is provided. In particular, the respective positions of right or left branches relative to the resulting optimal and TSN possibility distributions are given. It is also shown that the TSN transformation is the optimal transformation for the particular family of two-piece skewed distributions. The preservation of the asymmetry property is then analyzed and illustrated for several conventional asymmetric distributions and counter-examples of asymmetry preservation are provided. A multilinear approximation of the TSN transformation is finally proposed.

Keywords: uncertainty representation, asymmetric possibility distributions, probability possibility transformations, two-sided normalized transformation, coverage intervals

1. Introduction

The connection between probability theory and possibility theory has been considered since the beginning of possibility theory [40] [8], and then a long standing controversial debate took place in the literature [11] [22]. The issue of the transformation of a probability distribution into a possibility distribution is a key point of probability possibility relationships [26]. It has its roots in the consistency principle stated by Zadeh [40] and it was then developed according to two main different approaches founded on different opposing principles [30] [36].

On the one hand, Klir and coworkers [20] [15] [21] [18] [19] proposed a probability possibility transformation based on the so-called information invariance principle. This principle requires that the numbers expressing uncertainty in one theory should be transformed into corresponding ones in the other theory by an appropriate scale, and that the amount of uncertainty (the entropy $H(p)$ of the probability density p and the non-specificity $NS(\pi)$ for the possibility distribution π) should also be preserved in the transformation. One problem with this approach is the assumed equality $H(p) = NS(\pi)$ between uncertainty quantities with debatable commensurability. Furthermore, the scaling assumption leads to presuppose that $\pi(x)$ is a function of $p(x)$. This point-wise relationship may conflict with the consistency principle that requires $\Pi(A) \geq P(A)$, where $\Pi(A)$ and $P(A)$ are respectively the possibility and probability measures of the event A, which has to be true for any event and not only for singletons. Therefore, casting possibility measures into upper and lower probabilities [12] is no longer feasible.

On the other hand, Dubois-Prade and coworkers [9] [28] [10] [2] founded their transformation on a probability possibility consistency principle that can be summarized as: "the possibility of an event A is always greater than or equal to the probability of the event A", i.e. $\Pi(A) \geq P(A)$ for each A. Satisfying this inequality by a number of possibility distributions, they proposed to select the most informative one (in the sense of the possibility specificity) that, in addition, preserves the "shape", i.e. it preserves the order derived on the support, thus leading to the so-called optimal transformation. Note that when transforming probabilistic representation into a possibilistic one, some information is lost because point-value probabilities are transformed into interval valued ones. For the continuous case, which is the concern of this paper, the optimal possibility distribution can be achieved by deriving the level sets of the density f by computing the probability of this level sets: $\pi^{opt}(x) = \int_{y/\, p(y) \leq p(x)} p(y)dy$. The optimal possibility is thus the staking up of the shortest coverage intervals of the probability distribution coinciding with the alpha-cuts of the possibility distribution. For symmetric unimodal distributions, the alpha-cuts of the optimal possibility distribution bracket the narrowest coverage interval and the transformation takes the following simple form: $\pi^{opt}(x) = 2\min(F(x), 1 - F(x))$ with F the cumulative distribution function associated to f.

The optimal possibility distribution for the unimodal symmetric case has been widely used in applications [26] [14] unlike the asymmetric case that, until now, has only been considered by a few authors [2] [25]. While the optimal transformation is applicable to the asymmetric case,

it raises two main issues: the expression is not so simple and the resulting possibility distribution is not guaranteed to be convex on each side in all the cases, hence the alpha cuts of the possibility distribution do not necessarily bracket the coverage intervals. A few contributions have been made in the literature in deriving a simple approximation [25] and in considering convex/concave probability density [2], but the whole issue of the asymmetric case has not been solved, and this is the main object of this paper.

In this paper, by adopting the Dubois-Prade and coworkers line of thought and basing developments on a new probability possibility transformation proposed in the discrete context [27], we propose a new transformation of continuous asymmetric probability distributions into possibility distribution satisfying consistency, maximum specificity and dominance of coverage intervals. The paper is organized as follows: section 2 gives fundamental elements about coverage intervals and the optimal probability possibility transformation. Section 3 exposes the new so-called Two-Sided Normalized Transformation (TSN), its interpretation and properties. Section 4 focuses on the concept of asymmetry, its expression and applications to conventional asymmetric distributions used in applications as well as counter-examples of asymmetry preservation. Finally, section 5 presents a multi-linear approximation of the TSN transformation.

2. Building possibility distributions from coverage intervals

2.1. Coverage intervals

If $F:\mathbb{R} \to [0,1]$ is the cumulative distribution function of a random variable, the quantile function $Q(p)$ for all $p \in [0,1]$ for this random variable is

$$Q(p) = \inf\{x \in \mathbb{R} : p \le F(x)\}. \quad (1)$$

A coverage interval of level of confidence $1-\alpha$, $\alpha \in [0,1]$ is any interval $[Q(\beta), Q(\beta+1-\alpha)]$ with $0 \le \beta \le \alpha$ (see sections 3.12 to 3.16 in [17] for definitions of the terms and [4] [23] for estimations of coverage intervals).

When the cumulative distribution function is continuous and strictly increasing, its quantile function is simply $Q = F^{-1}$. Coverage intervals of level of confidence $1-\alpha$ are thus the intervals $U_{1-\alpha} = [F^{-1}(\beta), F^{-1}(\beta+1-\alpha)]$. In other words, a coverage interval of level of confidence $1-\alpha$ is any interval of $\mathbb{R}$ where the random variable takes its value with a probability $1-\alpha$.

A coverage interval of level of confidence $1-\alpha$ is said to be symmetric (or central or equal tailed) if $\beta = \frac{\alpha}{2}$.

A coverage interval of level of confidence $1-\alpha$ is said to be optimal if its length is the shortest among all coverage intervals of level of confidence $1-\alpha$. In the case of a unimodal symmetric distribution, the optimal coverage interval is the symmetric coverage interval [17]. This property will be recovered, as a particular case, in section 2.3.

2.2. From coverage intervals to possibility distributions

Let m be the mode of a unimodal continuous probability density function $f: X \subseteq \mathbb{R} \to \mathbb{R}^+$ with a cumulative distribution function F. Let X_L and X_R be two subsets of X such that $X_L = \{x \in X \,|\, x \leq m\}$ and $X_R = \{y \in X \,|\, y > m\}$. Let ρ be a function from $[0,1]$ to $[0,1]$. We consider coverage intervals $U_{1-\alpha}$ such that $\beta = \rho(\alpha)$, i.e.,

$$U_{1-\alpha} = [F^{-1}(\beta), F^{-1}(\beta + 1 - \alpha)] = [F^{-1}(\rho(\alpha)), F^{-1}(\rho(\alpha) + 1 - \alpha)] = [u(\alpha), v(\alpha)]. \quad (2)$$

It was shown in [10] that coverage intervals provide a means to build the transformation t of a probability density function f into a possibility distribution $\pi = t \circ f$. The transformation t is obtained by considering possibility distributions π such that their cuts of level α, for all $\alpha \in [0,1]$, are coverage intervals of level of confidence $1-\alpha$, that is

$$\pi(u(\alpha)) = \pi(v(\alpha)) = \alpha. \quad (3)$$

For unimodal continuous probability density functions considered in this paper, we have $u(\alpha) \in X_L$ and $v(\alpha) \in X_R$. For $\alpha = 1$, the coverage interval is reduced to the point $m = u(1) = v(1)$ and we have $\pi(m) = 1$.

Let $\psi: [0,1] \to \mathbb{R}$ be a strictly increasing function such that $\lim_{\alpha \to 0} \psi(\alpha) = -\infty$[1] and $\psi(1) = m$. Assuming that $\psi(\alpha) = F^{-1}(\rho(\alpha))$, if $F(\psi(\alpha)) + 1 - \alpha$ is strictly decreasing from 1 to $F(m)$ then coverage intervals defined by (2) provides a means to build possibility distributions. Indeed, using $u(\alpha) = \psi(\alpha)$ and $v(\alpha) = F^{-1}(F(\psi(\alpha)) + 1 - \alpha)$ we have for all $\alpha \in [0,1]$

$$\pi(u(\alpha)) = \pi(v(\alpha)) = \alpha. \quad (4)$$

For all $x \in X_L$, solving $x = u(\alpha) = \psi(\alpha)$ with respect to α gives the left hand side (LHS) part of the possibility distribution, i.e., $\pi_L(x) = \psi^{-1}(x)$. The right hand side (RHS) part of the possibility distribution has an analytic expression $\pi_R(y)$ for all $y \in X_R$ if the equation $F(y) = F^{-1}(F(\psi(\alpha)) + 1 - \alpha)$ can be solved analytically with respect to α.

For example, taking $\psi(\alpha) = F^{-1}(\frac{\alpha}{2})$ for all $\alpha \in [0,1]$ leads to symmetric coverage intervals $U_{1-\alpha} = [u(\alpha), v(\alpha)]$ such that

[1] If the cumulative distribution function is such that $F: [k, +\infty[\to \mathbb{R}^+$ the condition on ψ in 0 becomes $\psi(0) = k$.

$$\begin{cases} u(\alpha) = F^{-1}(\frac{\alpha}{2}), \\ v(\alpha) = F^{-1}(1-\frac{\alpha}{2}). \end{cases} \tag{5}$$

Since $\psi(1)=m=F^{-1}(1/2)$ the mode m is also the median of the distribution f. These coverage intervals are the shortest ones if the probability distribution is unimodal and symmetric. Solving each part of (5) with respect to α leads to the well-known optimal transformation of unimodal symmetric probability density functions proposed by Dubois, Prade and co-workers [9] [3] [10]

$$\begin{cases} \pi_L^{DPco}(x) = 2F(x), \forall x \in X_L, \\ \pi_R^{DPco}(y) = 2(1-F(y)), \forall y \in X_R. \end{cases} \tag{6}$$

Because $F(x)$ is defined and increasing for all $x \in X$, the two parts of (6) can be merged into a unique equation which defines the transformation t_{DPco}

$$\pi^{DPco}(x) = (t_{Dco} \circ f)(x) = t_{Dco}(f(x)) = 2\min(F(x), 1-F(x)). \tag{7}$$

For a given level of confidence, coverage intervals are not unique, thus an infinite number of non optimal possibility distributions can possibly be generated. For example, Fig. 1 shows the possibility distributions generated from the standard normal distribution $N(0,1)$ for $\psi_1(\alpha) = F^{-1}(\frac{\alpha}{2})$ and $\psi_2(\alpha) = \ln(\alpha)$.

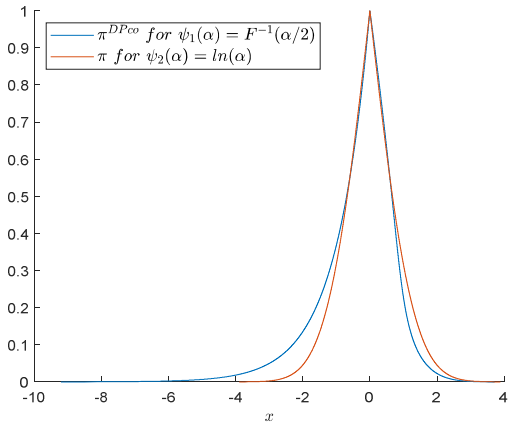

Fig. 1. Generation of possibility distributions from the standard normal distribution $N(0,1)$

2.3. Optimal transformation of unimodal continuous probability distributions

Let $[x \in X_L, y \in X_R]$ be a coverage interval with a level of confidence $1-\alpha$, i.e., $F(y) - F(x) = 1-\alpha$. The optimal transformation t_{Opt} of the density function f gives the possibility distribution $\pi^{Opt} = t_{Opt} \circ f$. It is obtained by minimizing the length of coverage interval $[x, y]$ for the given level of confidence $1-\alpha$. Using (2) we have

$$y-x=F^{-1}(\beta+1-\alpha)-F^{-1}(\beta). \tag{8}$$

The derivative of the length of the coverage interval with respect to β is

$$\frac{d}{d\beta}(y-x)=\frac{1}{f(F^{-1}(\beta+1-\alpha))}-\frac{1}{f(F^{-1}(\beta))}=\frac{1}{f(y)}-\frac{1}{f(x)}. \tag{9}$$

The length is minimal, i.e., the possibility distribution is optimal, when $f(x)=f(y)$ which corresponds to an horizontal cut of the probability density function f. Since the probability density function is unimodal and continuous, the condition $f(x)=f(y)$ defines $y\in X_R$ as a function of $x\in X_L$. Let φ be this function. The function φ is decreasing and such that $\varphi(m)=m$.

Let f_L and f_R be the restrictions of the probability density function to the sets X_L and X_R. Fig. 2 illustrates the restrictions for the probability density function for $Weibull(0.5,1.5)$.

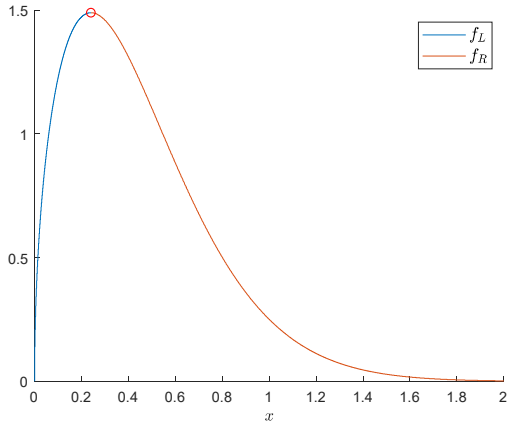

Fig. 2. Restrictions of the probability density function for $Weibull(0.5,1.5)$.

For all $x\in X_L$, the interval $[x,\varphi(x)]$ represented in Fig. 3, where $\varphi(x)=f_R^{-1}(f_L(x))$, is a coverage interval with the level of confidence

$$\beta_L(x)=F(\varphi(x))-F(x). \tag{10}$$

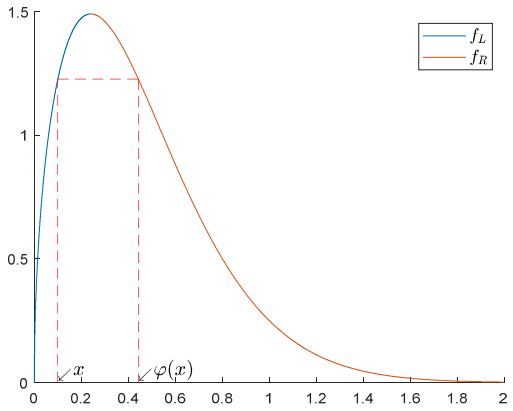

Fig. 3. Coverage interval for a given $x \in X_L$.

According to (3) the LHS of the optimal possibility distribution $\pi_L^{Opt}(x)$ for all $x \in X_L$ is

$$\pi_L^{Opt}(x) = 1 - \beta_L(x). \tag{11}$$

The same approach can be applied to build the RHS of the possibility distribution. In this case, for all $y \in X_R$ the interval $[\varphi^{-1}(y), y]$, where $\varphi^{-1}(y) = f_L^{-1}(f_R(y))$, is a coverage interval with the level of confidence

$$\beta_R(y) = F(y) - F(\varphi^{-1}(y)). \tag{12}$$

Thus, the RHS of the optimal possibility distribution is

$$\pi_R^{Opt}(y) = 1 - \beta_R(y), \forall y \in X_R. \tag{13}$$

In the general case, the inverse of the restrictions f_L and f_R have no analytic expression and must be numerically computed as shown in Fig. 4 for the Weibull distribution.

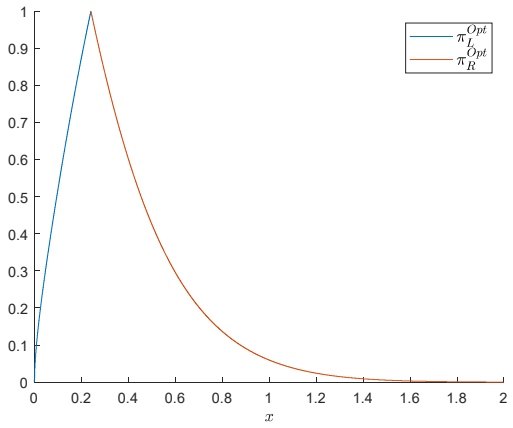

Fig. 4. Optimal possibility distribution for $Weibull(0.5,1.5)$.

Remark. When f is a symmetric probability density function, we have $\varphi(x) = 2m - x$. Since the cumulative distribution function is symmetric with respect to the point $(m, 1/2)$, it leads to $F(\varphi(x)) = 1 - F(x)$ and, using (10), $\beta_L(x) = F(\varphi(x)) - F(x) = 1 - 2F(x)$. Finally, using (11) the LHS of the possibility distribution is

$$\pi_L^{Opt}(x) = 2F(x). \tag{14}$$

Using (4) we have $2F(x) = \alpha$, thus the function ψ is such that $\forall \alpha \in [0,1], \psi(\alpha) = F^{-1}(\frac{\alpha}{2})$ leading to the transformation t_{DPco} given in (7). It shows that the shortest coverage intervals are the symmetric coverage intervals for symmetric distributions as stated in [17].

3. The two-sided normalized transformation

Let f be a density function and F the cumulative distribution function, the transformation t_{TSN} introduced in [27] and defined by

$$\pi^{TSN}(x) = t_{TSN}(f(x)) = \min(\frac{F(x)}{F(m)}, \frac{1-F(x)}{1-F(m)}), \forall x \in X, \tag{15}$$

is called Two-Sided Normalized (TSN) transformation.

This transformation was originally motivated by considerations on the preservation of the asymmetry in the transformation of the probability density function. Indeed, the transformation given by (7), when applied to an asymmetric distribution with a mode different from the median, modifies the position of the mode for the possibility distribution since $\pi^{DPco}(x) = 1$ for x such that $F(x) = \frac{1}{2}$. The weighting introduced in t_{TSN} for $F(x)$ and $1 - F(x)$, i.e., $F(m)$ and $1 - F(m)$, ensures to keep the same mode for f and π^{TSN}.

The next sub-sections give an interpretation and a justification of (15) in terms of coverage intervals while section 4 emphasizes the issue of asymmetry preservation.

3.1. Interpretation of the TSN transformation

The first family of asymmetric probability distributions was introduced by Fechner in 1897 and further developed under the name two-piece skewed distribution (see [38] for historical landmarks and analysis). Recent developments have been made by Fernandez et al. [13]. The two-piece name is justified by the fact that the two-piece skewed distribution mixes the LHS and the RHS of a unimodal symmetric probability density function denoted by *g* to obtain the skewed probability density function

$$f_s(x) = A\, g(x\gamma^{-sign(x)}) \text{ with } A = \frac{2}{\gamma + \frac{1}{\gamma}}. \tag{16}$$

As explained in [38], the initial idea relies on a probability density function f composed by two half of normal distributions around the mode m, i.e.,

$$f(x)=\begin{cases} f_L(x)=\dfrac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-m)^2}{2\sigma_1^2}} & ,if\ \ x\le m, \\ f_R(y)=\dfrac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-m)^2}{2\sigma_2^2}} & ,if\ \ y\ge m. \end{cases} \tag{17}$$

Then f_L and f_R are respectively scaled by $\dfrac{2\sigma_1}{\sigma_1+\sigma_2}$ and $\dfrac{2\sigma_2}{\sigma_1+\sigma_2}$ to generate the skewed continuous distribution f_s, such that $f_s(m)=A=\dfrac{2}{\sqrt{2\pi}(\sigma_1+\sigma_2)}$. The effect of this transformation is to move the median value of f away from the mode as it can be seen on the cumulative distribution functions in Fig. 5.

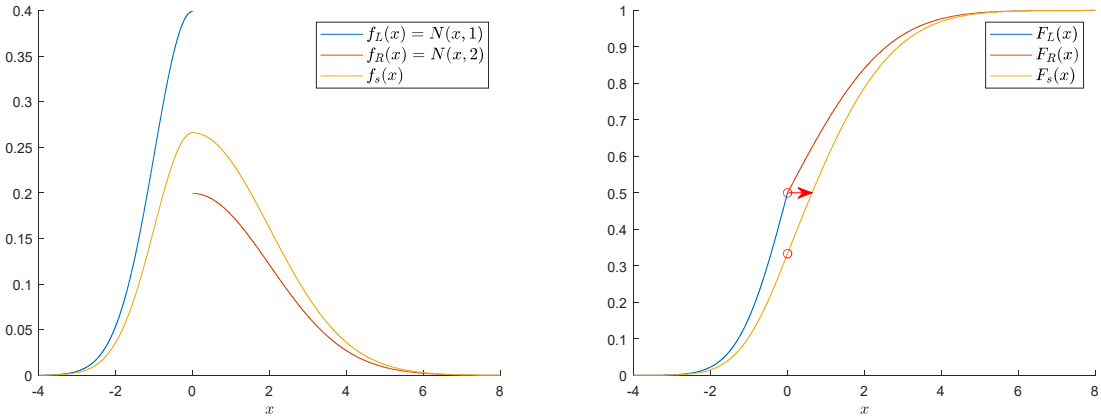

Fig. 5. Probability and cumulative distribution functions of a two-piece skewed normal distribution.

The TSN transformation can be interpreted as the reverse process. Starting from an asymmetric probability density function, its cumulative distribution function is modified such that the mode becomes the median. By doing so, the resulting probability density function becomes discontinuous at the mode, which is also the case for f in (17). This process is shown in Fig. 6 for the distribution $Weibull(0.5,1.5)$.

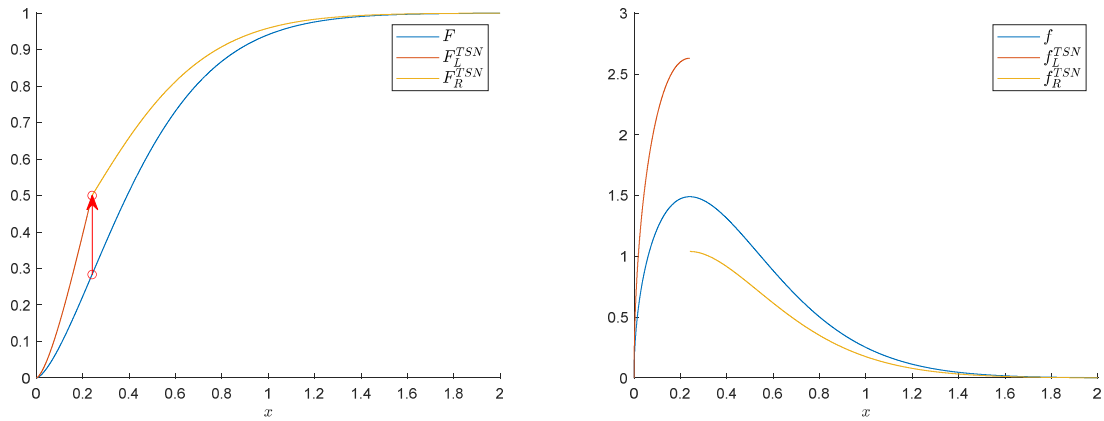

Fig. 6. Moving the mode towards the median for $Weibull(0.5,1.5)$.

The simplest solution to move the mode towards the median is to scale the LHS of the cumulative distribution function by $\frac{1}{2F(m)}$, i.e., $F_L^{TSN}(x)=\frac{1}{2F(m)}F(x),\forall x\le m$, but other transformations could be considered. From the inverse of the composition of two functions we have $\left(F_L^{TSN}\right)^{-1}(\alpha)=F^{-1}(2\alpha F(m)),\forall\,\alpha\in[0,1]$. Using (5) with $\psi(\alpha)=(F_L^{TSN})^{-1}(\frac{\alpha}{2})$ gives

$$\begin{cases} u(\alpha)=(F_L^{TSN})^{-1}(\frac{\alpha}{2})=F^{-1}(\alpha F(m)), \\ v(\alpha)=F^{-1}(\alpha(F(m)-1)+1). \end{cases} \tag{18}$$

Solving each part of (18) with respect to α leads to

$$\begin{cases} \pi_L^{TSN}(x)=\frac{F(x)}{F(m)},\forall x\in X_L, \\ \pi_R^{TSN}(y)=\frac{1-F(y)}{1-F(m)},\forall y\in X_R. \end{cases} \tag{19}$$

Finally, π_L^{TSN} is increasing and π_R^{TSN} is decreasing, they can be merged into the unique equation given in (15) which defines the TSN transformation (another proof is given in [26]).

Remark. The TSN transformation can also be interpreted in terms of truncated distributions [35] [37]. Indeed, the probability distribution f can be split into its left and right truncated distributions $f_L^{Tr}(x)$ and $f_R^{Tr}(x)$ such that

$$\begin{cases} f_L^{Tr}(x)=\frac{f(x)}{F(m)}, \text{if } x\le m \text{ and 0 otherwise}, \\ f_R^{Tr}(x)=\frac{f(x)}{1-F(m)}, \text{if } x\ge m \text{ and 0 otherwise}. \end{cases} \tag{20}$$

The coverage intervals for f_L^{Tr} are $[u(\alpha),v(\alpha)]$, with $v(\alpha)=m$ for each α. Using (4) it gives

$$F_L^{Tr}(\psi(\alpha))+1-\alpha=1\Leftrightarrow x=\psi(\alpha)=\left(F_L^{Tr}\right)^{-1}(\alpha)\Leftrightarrow\pi_L^{Tr}(x)=F_L^{Tr}(x)=\frac{F(x)}{F(m)}. \tag{21}$$

The coverage intervals for f_R^{Tr} are $[u(\alpha),v(\alpha)]$, with $u(\alpha)=m$ for each α. Using (4) we have $\psi(\alpha)=m$ and $y=v(\alpha)=\left(F_R^{Tr}\right)^{-1}(1-\alpha)$ since $F_R^{Tr}(m)=0$. Thus, it leads to

$$\pi_R^{Tr}(y)=1-F_R^{Tr}(y)=1-\frac{F(y)-F(m)}{1-F(m)}=\frac{1-F(y)}{1-F(m)}. \tag{22}$$

Therefore, roughly speaking, the TSN transformation consists in splitting the probability distribution f into the left and right truncated distributions relative to its mode (note that the

TSN can be applied in the same way to other points than the mode, e.g. the mean). Then, the truncated distributions are respectively transformed into left and right possibility distributions. Finally, these possibility distributions are merged to give (15).

3.2.Properties of the TSN transformation

Four properties of the TSN transformation are given in this section:

- Proposition 1 is a straightforward property showing that the TSN transformation is the same as the transformation t_{DPco} when the probability density function is symmetric.
- The convexity of the LHS and RHS parts of π^{TSN} is shown in Proposition 2.
- The position of the LHS and RHS parts of π^{TSN} relative to π^{Opt} is given by Proposition 3.
- Proposition 4 shows that the link between the TSN transformation and the two-piece skewed distributions explained in Section 3.1 is stronger than the proposed interpretation. Indeed, it shows that the TSN transformation is the optimal transformation for this particular family of skewed distributions.

Proposition 1. *When the continuous unimodal probability density function* f *is symmetric,* $t_{TSN}(f(x)) = t_{DPco}(f(x)) = t_{Opt}(f(x))$ *for all* $x \in X$.

Proof. In the case of a symmetric probability density function we have $F(m) = \frac{1}{2}$ which gives the proposition.

■

Proposition 2. *The LHS and RHS parts of the result of the TSN transformation are always strictly convex.*

Proof. Since it is assumed that the probability density function f is continuous and unimodal, it is strictly increasing on X_L and decreasing on X_R. Therefore its cumulative distribution function F is convex on X_L and concave on X_R. The LHS and RHS parts of the possibility distribution π^{TSN} are obtained by the composition of increasing affine functions by, respectively, $F(x)$ and $-F(x)$ which are both convex functions. Affine functions are convex and the composition of increasing convex functions by convex functions is also convex.

■

In Fig. 4, the LHS part of the optimal transformation for $Weibull(0.5,1.5)$ is concave. As shown in Fig. 7, it is convex for the TSN transformation.

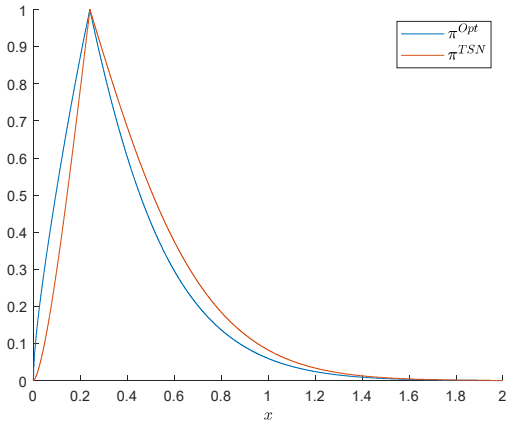

Fig. 7. Transformations t_{Opt} and t_{TSN} for $Weibull(0.5,1.5)$

Proposition 3. *If a continuous unimodal probability density function* $f(x)$ *with a mode m is differentiable for all* $x \in X$, *the LHS (resp. RHS) part of the two-sided normalized transformation is lower (resp. greater) than the LHS (resp. RHS) part of the optimal transformation if* $F(m) < \frac{1}{2}$.

Proof. See Appendix A.

■

Remark. Proposition 3 provides information about the relative positions of π^{TSN} and π^{Opt} but the optimal transformation π^{Opt} remains the most specific one, i.e., it has the shortest coverage intervals for all levels of confidence. Fig. 8 illustrates Proposition 3 with two sets of parameters for the Beta distribution.

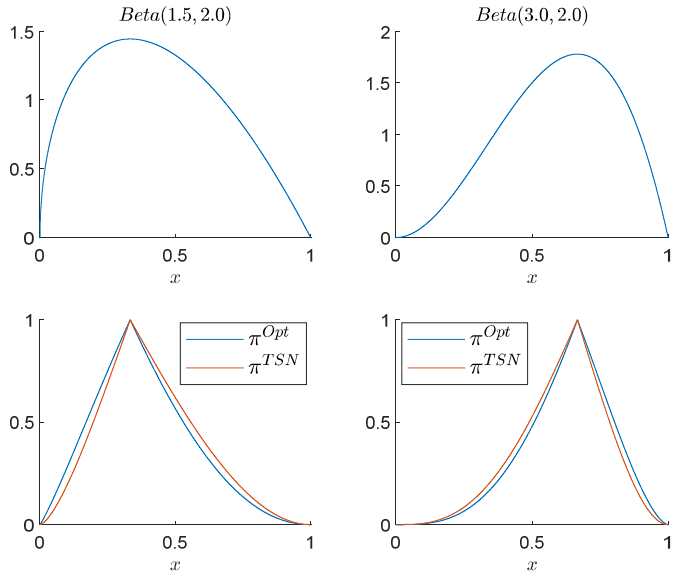

Fig. 8. Possibility distributions for the Beta distribution with two different sets of parameters

Proposition 4. *The two-sided normalized transformation is the optimal transformation for the family of two-piece skewed distributions defined by*

$$f_s(x) = A\, g(x\gamma^{-sign(x)}) \text{ with } A = \frac{2}{\gamma + \frac{1}{\gamma}} \tag{23}$$

where g is a unimodal symmetric probability density function.

Proof. See Appendix B.

■

Remark. Proof of Proposition 4 in Appendix B provides two ways of expressing π^{TSN}. The first one is given by (73) where π^{TSN} is defined from the cumulative distribution function F_s of the skewed distribution. The other one, based on the cumulative distribution function G of the symmetric distribution, is obtained, for all $x \in X$, by merging (64) and (69) which leads to

$$\pi^{TSN}(x) = \pi^{Opt}(x) = 2\min(G(\gamma x), 1 - G(\frac{x}{\gamma})). \tag{24}$$

4. Asymmetry and the two-sided normalized transformation

4.1. Definitions and properties

The measure of the asymmetry of a probability distribution goes back to Pearson and his moment coefficient of skewness [33]. Since then, besides the standardized third central moment, several coefficients have been proposed, see for example [7] [29] [16] [24] [39] [35] [5] [31]. In this section, function definitions based on the departure from symmetry (e.g. [6] [31]) are respectively used for the asymmetry of the probability density functions and for possibility distributions.

Definition 1. *A continuous probability density function f is said to be right asymmetric about m if* $asym_f(x) = f(x) - f(2m - x) < 0, \forall x < m.$

Definition 2. *A possibility distribution $\pi = t \circ f$, obtained from the transformation t of a probability density function f, is said to be right asymmetric about m if* $asym_\pi(x) = \pi(x) - \pi(2m - x) < 0, \forall x < m.$

Definition 3. *A probability-possibility transformation t preserves the asymmetry if* $asym_f(x) < 0 \Rightarrow asym_\pi(x) < 0, \forall x < m.$

If the TSN transformation is used to transform a continuous unimodal probability density function f with a mode m to a possibility distribution π^{TSN} then the four possibilities are:

- f is right asymmetric about m and π^{TSN} is right asymmetric about m,
- f is not right asymmetric about m and π^{TSN} is right asymmetric about m,
- f is not right asymmetric about m and π^{TSN} is not right asymmetric about m,
- f is right asymmetric about m and π^{TSN} is not right asymmetric about m.

From a quick glance at Fig. 7 or Fig. 8, it may seem like the TSN transformation always preserves the asymmetry about the mode. The next two propositions provide sufficient conditions for the preservation of the asymmetry, i.e., the first of the above possibilities holds true:

- Proposition 5 shows that the strict monotonicity of the function h, such that $h(x)=\frac{f(x)}{f(2m-x)}, \forall x<m$, is related to the preservation of the asymmetry by the TSN transformation.
- Proposition 6 provides a means to check the monotonicity of h from the convexity of the function ω, such that $\omega(x)=\frac{1}{f(x)}\frac{df(x)}{dx}, \forall x<m$. Since ω depends on the probability density function and its derivative, it gives analytic expressions often simpler than that of h.

Proposition 5. *Let f be a unimodal continuous probability density function with a mode m. Let h be the function such that $h(x)=\frac{f(x)}{f(2m-x)}, \forall x<m$. If h is strictly increasing from L to 1, where $L=0$ if f is defined on a left-bounded interval or $L=\lim_{x\to-\infty} h(x)<\frac{F(m)}{1-F(m)}$ if f is defined on $\mathbb{R}$, then the asymmetry about m is preserved by the TSN transformation.*

Proof. See Appendix C.

■

The functions h for the distribution $Weibull(0,5;1,5)$ and the skewed normal distribution $SKN(0,1,3) = 2\phi(x)\Phi(3x)$, where ϕ is the probability density function of the standard normal distribution $N(0,1)$, are given in Fig. 9 as illustrations of Proposition 5. According to Proposition 12 and Proposition 13, $L=0$ for these distributions. It is formally shown in section 4.2 that h is strictly increasing for these distributions.

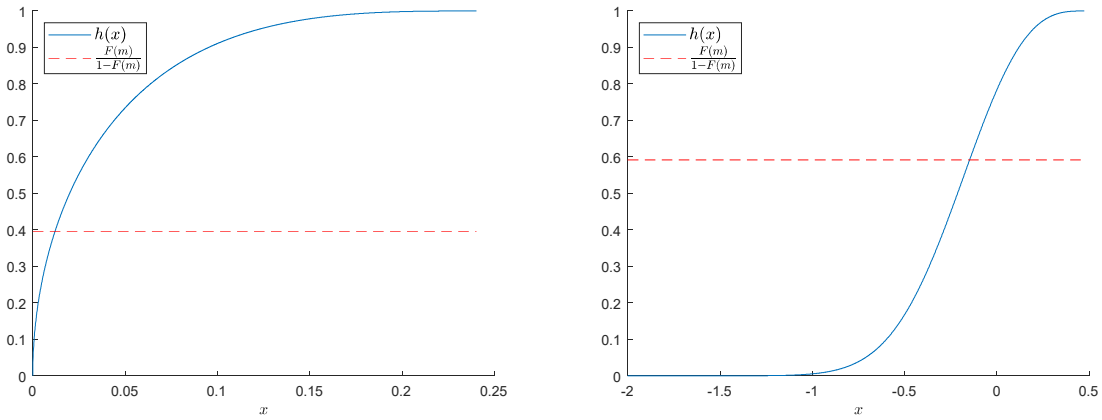

Fig. 9. Functions h for $Weibull(0.5,1.5)$ and $SKN(0,1,3)$

Proposition 6. *If a unimodal continuous probability density function* f *with a mode* m *is twice differentiable for all* $x \in X$ *and the function* $\omega(x) = \dfrac{1}{f(x)}\dfrac{df(x)}{dx}$ *is strictly convex for all* $x < m$ *then the asymmetry about m is preserved by the TSN transformation.*

Proof. For all $x < m$ we have

$$\begin{cases} \omega(x) = \dfrac{f'(x)}{f(x)}, \\ \omega(2m-x) = \dfrac{1}{f(2m-x)}\dfrac{df(2m-x)}{dx} = -\dfrac{f'(2m-x)}{f(2m-x)}. \end{cases} \tag{25}$$

Thus, if $\omega(x)$ is strictly convex then $\omega(2m-x)$ is strictly concave. The twice differentiability of ω gives $\lim_{x\to m}\omega(x) = \lim_{x\to m}\omega(2m-x) = 0$ and $\lim_{x\to m}\omega'(x) = \lim_{x\to m}\left[\omega(2m-x)\right]' = \dfrac{f''(m)}{f(m)} < 0$.

Therefore, $\omega(x)$ and $\omega(2m-x)$ are decreasing functions with respect to x and are such that for all $x<m$, $\omega(x) > \omega(2m-x) \Leftrightarrow f'(x)f(2m-x) + f(x)f'(2m-x) > 0$.

Thus, $h'(x) = \dfrac{f'(x)f(2m-x) + f(x)f'(2m-x)}{f^2(2m-x)} > 0, \forall x < m$. Since h is strictly increasing, Proposition 5 holds true which completes the proof.

■

Remark. The function ω can also be written as $\omega(x) = \dfrac{d\log(f(x))}{dx}, \forall x < m$. Since ω is a decreasing function when it is convex, it means that f is a log-concave density function. This property is useful to identify distributions which are candidates for testing the convexity of ω but it only provides candidate distributions since f can be log-concave while ω being not convex.

4.2. Examples of right asymmetry preservation

In this sub-section, five conventional examples are given of unimodal continuous asymmetric distributions for which the TSN transformation preserves the asymmetry about the mode. They all verify Proposition 6, i.e., the function ω is strictly convex (four other examples are given in Appendix D). The sixth example concerns Pearson type IV distributions which provides a case where ω is not convex but h is a strictly increasing function and, therefore, the TSN transformation preserves the asymmetry. The last example of two-piece skewed distributions illustrates that the condition in Proposition 5, though sufficient, is not necessary for preserving asymmetry.

4.2.1. Weibull distribution

The Weibull distribution is a particular case of the Generalized Extreme Value distribution family and is defined by

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^{k}}, \forall x \geq 0, \tag{26}$$

with $k > 1$ in order to have a mode $m = \lambda\left(\frac{k-1}{k}\right)^{\frac{1}{k}}$.

It leads to $\omega(x) = -\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} + \frac{(k-1)}{x}$ and $\omega''(x) = (k-1)\frac{(2-k)k\left(\frac{x}{\lambda}\right)^{k} + 2}{x^3}$. For $1 < k \leq 2$, ω is strictly convex because $\omega''(x) > 0$. When $k > 2$, $\omega''(x) = 0$ for $x_1 = \lambda\left(\frac{2}{k(k-2)}\right)^{\frac{1}{k}}$. If $m < x_1$ then $\forall x < m$ the function ω is strictly convex which leads to

$$m < x_1 \Leftrightarrow \frac{k-1}{k} < \frac{2}{k(k-2)} \Leftrightarrow k < 3. \tag{27}$$

For $k \geq 3$ the function h is not strictly increasing which means that at least one x exists such that $f(2m - x) \geq f(x)$.

Whatever the parameter λ, the TSN transformation preserves the asymmetry of the Weibull distribution for $k \in]1,3[$.

4.2.2. Beta distribution

The beta distribution is defined, for all $x \in [0,1]$, by

$$f(x) = \frac{x^{1-\alpha}(1-x)^{1-\beta}}{B(\alpha,\beta)} \tag{28}$$

with $B(\alpha,\beta) = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du$.

We will consider $\alpha, \beta > 1$ in order to have a mode $m = \frac{\alpha - 1}{\alpha + \beta - 2}$.

We have $\omega(x) = \frac{(\alpha + \beta - 2)x + 1 - \alpha}{x(x-1)}$ and its second derivative is given by

$$\omega''(x) = \frac{2(\alpha+\beta-2)x^3 - 6(\alpha-1)x^2 + 6(\alpha-1)x - 2(\alpha-1)}{x^3(x-1)^3} = \frac{n(x)}{d(x)}. \quad (29)$$

For all $x \in [0,1], d(x) < 0$. The first derivate of n with respect to x is

$$n'(x) = 6(\alpha+\beta-2)x^2 - 12(\alpha-1)x + 6(\alpha-1). \quad (30)$$

The discriminant of n' is negative, therefore $n'(x) > 0$ since $\alpha, \beta > 1$. For all $x \in [0, m]$ the function n is strictly increasing from $n(0) = -(\alpha - 1)$ to $n(m) = \frac{(\alpha-\beta)(\alpha-1)(\beta-1)}{(\alpha+\beta-2)^2}$. Since for all $x \in [0,1], d(x) < 0$, we can conclude that if $n(x) < 0$, for all $x < m$, then $\omega''(x) > 0$ and ω is strictly convex.

The TSN transformation preserves the asymmetry of the beta distribution for all parameters such that $1 < \alpha < \beta$.

4.2.3. Gamma distribution

The gamma distribution is defined by

$$f(x) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\Gamma(k)\theta^k}, \forall x > 0. \quad (31)$$

The gamma function is defined by $\Gamma(z) = \int_0^\infty u^{z-1} e^{-u} du$ with z a complex number with a positive real part. With a shape parameter $k > 1$ and a scale parameter $\theta > 0$ the mode of the probability density function is $m = (k-1)\theta$.

The gamma distribution leads to $\omega(x) = -\frac{x - \theta(k-1)}{\theta x}$ and $\omega''(x) = \frac{2(k-1)}{x^3}$. The function ω is strictly convex if $k > 1$.

The TSN transformation preserves the asymmetry of the gamma distribution for all parameters such that $k > 1$ and $\theta > 0$.

4.2.4. Log-normal distribution

The log-normal distribution is defined by

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}, \forall x > 0, \text{ with } \sigma > 0. \quad (32)$$

It leads to $\omega(x) = -\frac{\sigma^2 - \mu + \ln(x)}{\sigma^2 x}$ and $\omega''(x) = \frac{2(\mu - \sigma^2) - 2\ln(x) + 3}{\sigma^2 x^3}$.

The mode of the log-normal distribution is $m = e^{\mu - \sigma^2}$ and we have $-2\ln(x) > -2\ln(m) = -2(\mu - \sigma^2)$. Thus, $2(\mu - \sigma^2) - 2\ln(x) + 3 > 3$ and since $x > 0$, ω is strictly convex.

Whatever the parameters μ and σ, the TSN transformation preserves the asymmetry of the log-normal distribution.

4.2.5. Skewed normal distribution

Let ϕ be the probability density function of the standard normal distribution and Φ its cumulative distribution function. The skewed normal distribution [1] is defined by

$$f(x) = 2\phi(x)\Phi(\alpha x), \forall x \in \mathbb{R}, \tag{33}$$

where α is a skewing parameter.

For this distribution, we have $\omega(x) = -x + \alpha\frac{\phi(\alpha x)}{\Phi(\alpha x)}$ and $\omega''(x) = \alpha\left(\frac{\phi(\alpha x)}{\Phi(\alpha x)}\right)''$.

Let r be the inverse Mill's ratio defined for all x by $r(x) = \frac{\phi(x)}{1 - \Phi(x)}$. The function r is strictly convex [34], i.e. $r''(\alpha x) > 0$, if $\alpha > 0$. Since $1 - \Phi(x) = \Phi(-x)$, we have $\omega''(-\alpha x) = \omega''(\alpha x) = r''(\alpha x)$ therefore ω is convex if $\alpha > 0$.

The TSN transformation preserves the asymmetry of the skew normal distribution when the skewing parameter α is positive.

4.2.6. Pearson type IV distribution as a non convex case

In 1895, Pearson analyzed asymmetrical frequency curves and introduced a classification into five types [33], then extended into twelve distribution types [32], as the solution of the differential equation

$$\omega(x) = \frac{1}{f(x)}\frac{df(x)}{dx} = \frac{a_0 + a_1 x + a_2 x^2}{c_0 + c_1 x + c_2 x^2 + c_3 x^3}. \tag{34}$$

The probability density function of the distributions known as Pearson type IV distributions is

$$f(x) = y_0 \frac{e^{-\nu \arctan(\frac{x}{a})}}{(1 + \frac{x^2}{a^2})^{\mu}}, \forall x \in \mathbb{R}, \text{ with } a > 0. \tag{35}$$

The normalizing factor y_0 exists only if $\mu > 1/2$. The function ω is

$$\omega(x)=\frac{-a\nu-2\mu x}{a^2+x^2},\forall x\in\square\ . \tag{36}$$

The mode m of f is such that $\omega(m)=0$, i.e. $m=-\frac{a\nu}{2\mu}$. Since ω has two extrema, respectively in $x=\frac{-a\nu+a\sqrt{4\mu^2+\nu^2}}{2\mu}$ and $x=\frac{-a\nu-a\sqrt{4\mu^2+\nu^2}}{2\mu}$, it cannot be convex for all $x<m$.

The function $h$is

$$h(x)=\frac{e^{-\nu\left(\arctan\left(\frac{x}{a}\right)+\arctan k(x)\right)}}{\left(\frac{x^2}{a^2}+1\right)^{\mu}}\left(k^2(x)+1\right)^{\mu},\forall x\in\square, \tag{37}$$

with $k(x)=\frac{x\mu+a\nu}{a\mu}$ and its derivative is

$$h'(x)=\frac{-a\nu\left(k^2(x)+1\right)^{\mu}\left(a\nu+2\mu x\right)^2 e^{-\nu\left(\arctan\left(\frac{x}{a}\right)+\arctan k(x)\right)}}{\left(\frac{x^2}{a^2}+1\right)^{\mu}\left(a^2+x^2\right)\left(a^2\left(\mu^2+\nu^2\right)+2a\mu\nu x+\mu^2x^2\right)},\forall x\in\square\ . \tag{38}$$

It can be easily shown that $a^2\left(\mu^2+\nu^2\right)+2a\mu\nu x+\mu^2x^2\geq a^2\mu^2$. Therefore, if $\nu<0$, $h'(x)$ is strictly positive and $h(x)$ is strictly increasing for all $x<m$.

The TSN transformation preserves the asymmetry of the distribution defined by (35) when $\nu<0$.

4.2.7. Two-piece skewed distributions

Given the two-piece skewed distribution f_s, without loss of generality, it is assumed that the mode of the symmetric probability density function g is $m=0$. Using the symmetry of g, from (23) we have $f_s(x)=Ag(\gamma x)$ and $f_s(2m-x)=Ag(\frac{-x}{\gamma})=Ag(\frac{x}{\gamma})$ for all $x<0$. Therefore, f_s is right asymmetric if $\gamma>1$. For the same reason, from (24) we have $\pi^{TSN}(x)=2G(\gamma x)$ and $\pi^{TSN}(2m-x)=2(1-G(\frac{-x}{\gamma}))=2G(\frac{x}{\gamma})$ for all $x<0$. Therefore, π^{TSN} is right asymmetric if $\gamma>1$. Thus, the TSN transformation preserves the asymmetry for the family of two-piece skewed distributions if $\gamma>1$ or, equivalently, $F(m)=F(0)<\frac{1}{2}$.

The asymmetry preservation is not related to the monotonicity of the function h, but it can be verified that many symmetric probability functions g lead to monotonically increasing functions h. Let us informally explain this observation by a counter-example. In order to have the function h non-monotonic, we can consider a probability density function $g$with a sinusoidal term such that its cumulative distribution function G is convex for all $x \le 0$, e.g.,

$$g(x) = \begin{cases} g_L(x) = -\dfrac{2(x-1) - \dfrac{\cos(\pi x)}{\pi}}{2\left((x-1)^2 - \dfrac{\sin(\pi x)}{\pi^2}\right)^2}, \text{if } x \le 0, \\ g_R(x) = g_L(-x), \text{if } x \ge 0. \end{cases} \tag{39}$$

Then, the function g is skewed according to (23). Even with such a specific distribution, the skewing parameter γ has to be tuned in order to obtain a non-monotonic function h as shown in Fig. 10. Therefore, it emphasizes, in an informal way, that it is much easier to obtain monotonic functions h than non-monotonic ones.

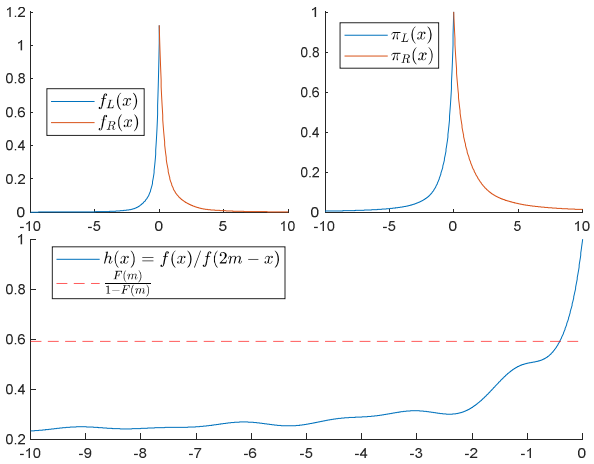

Fig. 10. Function g given by (39) skewed by $\gamma = 1.3$

4.3. Counter-examples of asymmetry preservation

This sub-section illustrates the three cases, presented in section 4.1, where the asymmetry is not preserved. Many conventional probability distributions have been tried but none of them could fit the three cases where asymmetry is not preserved. Therefore, special probability density functions have been built for this purpose using the following principle:

- Define two convex functions $\pi_L^{TSN}(x)$ for all $x \le m$ and $\pi_R^{TSN}(x)$ for all $x \ge m$. According to Proposition 10 to keep the possibility for f to be right asymmetric we must have $F(m) < \dfrac{1}{2}$. This condition is obtained if the derivatives of the convex

functions are such that $\pi_L'^{TSN}(m) > -\pi_R'^{TSN}(m)$. Indeed, since $\pi_L'^{TSN}(x) = \frac{f(x)}{F(m)}$ and $\pi_R'^{TSN}(x) = -\frac{f(x)}{1-F(m)}$ it leads to $F(m) < \frac{1}{2}$.

- Compute $F(m)$ from the derivatives $\pi_L'^{TSN}(m)$ and $\pi_R'^{TSN}(m)$.
- Generate $f_L(x) = F(m)\pi_L'^{TSN}(x)$ for all $x \le m$ and $f_R(x) = -(1-F(m))\pi_R'^{TSN}(x)$ for all $x \ge m$.

4.3.1. f not right asymmetric and π^{TSN} right asymmetric

Let π^{TSN} be the possibility distribution such that

$$\pi^{TSN}(x) = \begin{cases} \pi_L^{TSN}(x) = x^2, \text{ if } 0 \le x \le 1, \\ \pi_R^{TSN}(x) = \frac{1}{x}, \text{ if } x \ge 1. \end{cases} \tag{40}$$

It gives the probability density function

$$f(x) = \begin{cases} f_L(x) = 2xF(m), \text{ if } 0 \le x \le 1, \\ f_R(x) = -\frac{1-F(m)}{x^2}, \text{ if } x \ge 1, \end{cases} \tag{41}$$

with a modal value of $m = 1$ and $F(m) = 1/3$.

The convex functions $\pi^{TSN}(x) = \pi_L^{TSN}(x)$ and $\pi^{TSN}(2m-x) = \pi_R^{TSN}(2-x)$ have no intersecting point for $0 \le x < 1$, therefore π^{TSN} is right asymmetric and there is only one point where $h(x) = \frac{F(m)}{1-F(m)}$. From Proposition 12 we know that $h(0) = 0$, thus the asymmetry of f can be checked from the monotonicity of $h(x) = \frac{f(x)}{f(2m-x)} = x(2-x)^2$ for all $0 \le x < 1$. It can be easily verified that $h'(x) = 0$ for $x = 2/3$ and $h(2/3) = 32/27 > 1$ which means that f is not right asymmetric.

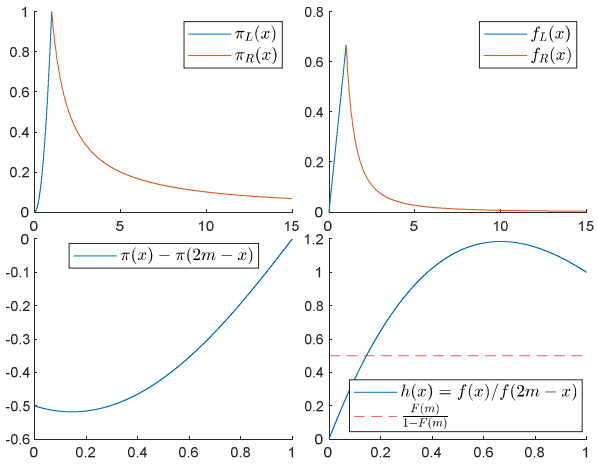

Fig. 11. f not right asymmetric and π^{TSN} right asymmetric

4.3.2. f not right asymmetric and π^{TSN} not right asymmetric

Let π^{TSN} be the possibility distribution defined by

$$\pi^{TSN}(x)=\begin{cases}\pi_L^{TSN}(x)=\dfrac{-1}{5x-1}, \text{ if } x\le 0,\\ \pi_R^{TSN}(x)=e^{-x}, \text{ if } x\ge 0.\end{cases} \tag{42}$$

It defines the probability density function

$$f(x)=\begin{cases}f_L(x)=F(m)\dfrac{5}{(5x-1)^2}, \text{if } x\le 0,\\ f_R(x)=(1-F(m))e^{-x}, \text{if } x\ge 0.\end{cases} \tag{43}$$

with a modal value of $m=0$ and $F(m)=0.1667$.

For $x<0$ the convex functions $\pi^{TSN}(x)=\pi_L^{TSN}(x)$ and $\pi^{TSN}(2m-x)=\pi_R^{TSN}(-x)$ have one intersecting point, therefore π^{TSN} is not right asymmetric.

One intersecting point means that $asym_{\pi^{TSN}}(x)$ has two extrema or, equivalently, two values exist such that $h(x)=\dfrac{F(m)}{1-F(m)}$. Since $\lim\limits_{x\to m}h(x)=1>\dfrac{F(m)}{1-F(m)}$, the function h cannot be strictly increasing, thus f is right asymmetric if $h(x)=\dfrac{f_L(x)}{f_R(2m-x)}<1$, for all $x<0$. Since e^x decreases faster than $\dfrac{1}{x}$ when $x\to-\infty$ we have $\lim\limits_{x\to-\infty}h(x)=+\infty$ and therefore f is not right asymmetric.

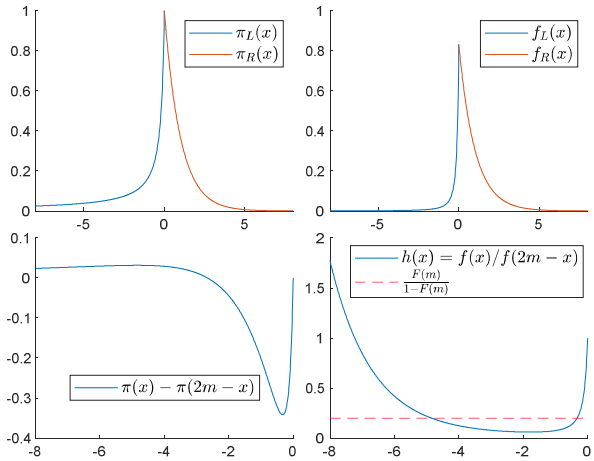

Fig. 12. f not right asymmetric and π^{TSN} not right asymmetric

4.3.3. f right asymmetric and π^{TSN} not right asymmetric

For this counter-example, the possibility distribution π^{TSN} is chosen such that for all $x<m$ the convex functions $\pi^{TSN}(x)=\pi_L^{TSN}(x)$ and $\pi^{TSN}(2m-x)=\pi_R^{TSN}(2m-x)$ have an infinite number of intersecting points, therefore π^{TSN} is not right asymmetric, e.g.,

$$\pi^{TSN}(x)=\begin{cases}\pi_L^{TSN}(x)=\dfrac{1}{(x-1)^2-\dfrac{\sin(\pi x)}{\pi^2}}, \text{if } x\le 0,\\ \pi_R^{TSN}(x)=\dfrac{1}{(x+1)^2}, \text{if } x\ge 0.\end{cases} \tag{44}$$

It defines the probability density function

$$f(x)=\begin{cases}f_L(x)=-F(m)\dfrac{2(x-1)-\dfrac{\cos(\pi x)}{\pi}}{\left((x-1)^2-\dfrac{\sin(\pi x)}{\pi^2}\right)^2}, \text{if } x\le 0,\\ f_R(x)=(1-F(m))\dfrac{2}{(x+1)^3}, \text{if } x\ge 0,\end{cases} \tag{45}$$

with a modal value of $m=0$ and $F(m)=0.4631$.

As in the previous counter-example, f is right asymmetric if $h(x)=\dfrac{f_L(x)}{f_R(2m-x)}<1, \forall x<0,$

but now, we have $\lim\limits_{x\to-\infty}h(x)=\dfrac{F(m)}{1-F(m)}$. Indeed, when $x\to-\infty$ the periodic term are bounded and the limit is given by the ratio of the terms of highest degree in x of $f_L(x)$ and $f_R(2m-x)$. Because $asym_{\pi^{TSN}}$ is periodic, so is the function h. Extrema of $h(x)$ are decreasing starting from $x=m$ where $\lim\limits_{x\to m}h(x)=\dfrac{f_L(m)}{f_R(m)}=1$. Thus, $h(x)<1, \forall x<0$, which leads f to be right asymmetric.

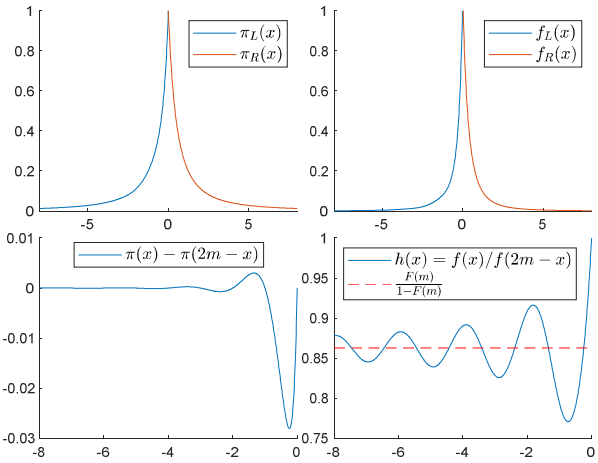

Fig. 13. f right asymmetric and π^{TSN} not right asymmetric

5. Multilinear approximation of the two-sided normalized transformation

The aim of this section is to provide approximations of the TSN transformation of any distribution for practical use in applications. For the sake of clarity, the RHS of the possibility distribution is considered and the same approach can be developed for the LHS part. Let $(u,\pi^{TSN}(u))$ and $(v,\pi^{TSN}(v))$ be two points on the RHS of the possibility distribution. The bilinear approximation of the possibility distribution for y between u and v is shown in Fig. 14.

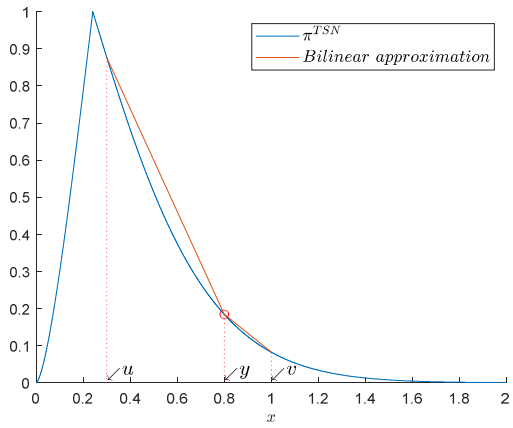

Fig. 14. Bilinear approximation of the RHS part of the two-sided normalized transformation for $Weibull(0.5,1.5)$

The surface S under the RHS approximation is a function of y, that is

$$S(y)=\frac{(\pi^{TSN}(u)+\pi^{TSN}(y))(y-u)}{2}+\frac{(\pi^{TSN}(y)+\pi^{TSN}(v))(v-y)}{2}. \tag{46}$$

Its second derivative with respect to y is

$$S''(y)=\frac{v-u}{2}\pi''^{TSN}(y). \tag{47}$$

According to Proposition 2, π^{TSN} is strictly convex on X_R, that is $\pi''^{TSN}(y)>0,\forall y\in X_R$ and, therefore, $S''(y)>0$. Thus S is a strictly convex function. Since $S(u)=S(v)$ the function S has a minimum. The value $y_{\min}$ for which the minimum is reached provides the best bilinear approximation in the sense of the maximum of specificity.

This property is used to build the algorithm provided in Fig. 15. Given the interval $[u,v]$ the difference e between surface below the linear approximation of π^{TSN} and the surface below the distribution itself can be computed. Due to the convexity of π^{TSN} this difference is always positive. If this difference is lower than a given threshold ε, the vector $\begin{bmatrix}a_1\\a_0\end{bmatrix}$ of the linear interpolation a_1y+a_0 is returned. Otherwise, $y_{\min}$ is computed and the process is recursively

called on the intervals $[u, y_{\min}]$ and $[y_{\min}, v]$. The algorithm returns the list of vectors $\begin{bmatrix} a_{i,1} \\ a_{i,0} \end{bmatrix}$ to be used in the linear interpolations $p_i(y) = a_{i,1} y + a_{i,0}$.

Function approximate(π^{TSN} **, u, v,** ε **)**

$$e = \frac{(\pi^{TSN}(u)+\pi^{TSN}(v))(v-u)}{2} - \int_u^v \pi^{TSN}(t)dt$$

$$y_{\min} = \min_{y\in[u,v]} \frac{(\pi^{TSN}(u)+\pi^{TSN}(y))(y-u)}{2} + \frac{(\pi^{TSN}(y)+\pi^{TSN}(v))(v-u)}{2}$$

if $e < \varepsilon$ **then**

return $\begin{bmatrix} \dfrac{\pi^{TSN}(v)-\pi^{TSN}(u)}{v-u} \\ \pi^{TSN}(u) - \dfrac{\pi^{TSN}(v)-\pi^{TSN}(u)}{v-u}u \end{bmatrix}$

else

return [approximate(π^{TSN} **, u,** $y_{\min}$**,** ε **), approximate(** π^{TSN} **,** $y_{\min}$**, v,** ε **)]**

Fig. 15. Recursive approximation of π^{TSN}

Because π_R^{TSN} (resp. π_L^{TSN}) is decreasing (resp. increasing) the slopes of the linear interpolations are ordered, i.e. $a_{i,1} < a_{i+1,1}$. Thus, the multilinear interpolation $\tilde{\pi}^{TSN}$ of π^{TSN} is

$$\begin{cases} \tilde{\pi}_L^{TSN}(x) = \max_i(0, p_i^l(x)) = \max_i(0, a_{i,1}^L x + a_{i,0}^L), \text{ if } x \in X_L, \\ \tilde{\pi}_R^{TSN}(y) = \max_i(0, p_i^R(y)) = \max_i(0, a_{i,1}^R y + a_{i,0}^R), \text{ if } y \in X_R. \end{cases} \tag{48}$$

Finally, since $a_{i,1}^L > 0$ and $a_{i,1}^R < 0$ both parts of (48) can be merged into the single equation

$$\tilde{\pi}^{TSN}(x) = \min(\tilde{\pi}_L^{TSN}(x), \tilde{\pi}_R^{TSN}(x)), \forall x \in X. \tag{49}$$

Fig. 16 shows multilinear approximations of the two-sided normalized transformation for $Weibull(0.5,1.5)$ and $\varepsilon = 0.01$ and $\varepsilon = 0.001$.

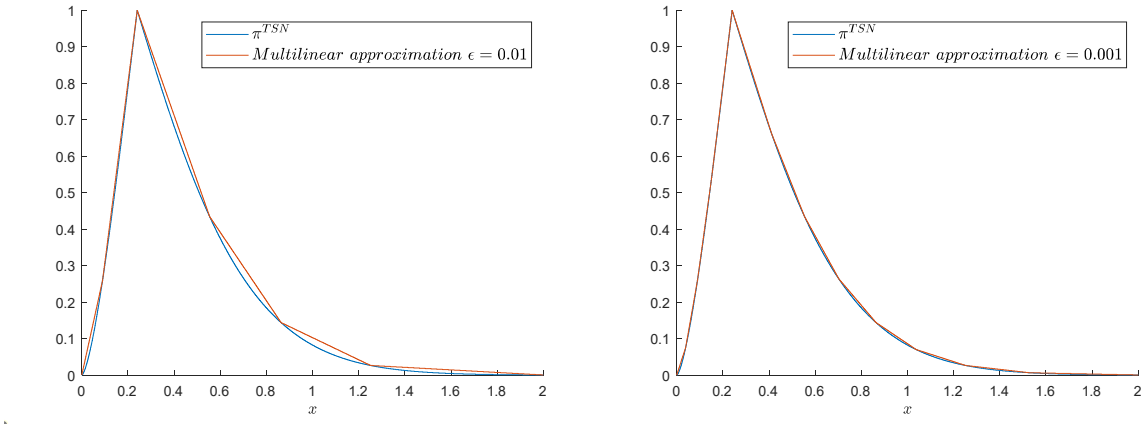

Fig. 16. Multilinear approximations of the two-sided normalized transformation for $Weibull(0.5,1.5)$

Conclusion

In this paper the links between coverage intervals and possibility distributions were highlighted. A new transformation of asymmetric probability density functions, called Two-Sided Normalized (TSN) transformation, was investigated. This transformation generalizes the transformation proposed by Dubois-Prade and co-workers in the sense that it keeps its simplicity and gives the same result when considering symmetric distributions.

Among the results presented in this paper, it was shown that the TSN transformation is the optimal transformation for the family of two-piece skewed distributions. Moreover, for this later family the asymmetry of the original probability distribution is preserved by the TSN. For other skewed distributions, an interesting result is that the monotonicity of the function $h(x) = \frac{f(x)}{f(2m-x)}$ with $x < m$ is a sufficient condition for preserving asymmetry. It was also shown that this condition is satisfied when the derivative of the logarithm of the density is convex.

Surprisingly many well-known distributions satisfy the convexity condition. Further analysis will be necessary to investigate if this sufficient condition can be turned into a necessary one. Another interesting question concerns the links between the transformation and the cumulative distribution function. Considering more complex functions than affine functions, as in the TSN transformation, will lead to new transformations. Finally, probability-possibility transformations of asymmetric distributions open new questions about the meaning of asymmetry and more generally on the meaning of the "shape" preservation in the transformation. More specifically, possible connections between the proposed transformation and preservation of asymmetry/symmetry might be provided by looking deeper in the relationships between the expression of the transformation and the skewness orderings considered in the literature. Connections with quantitative measures of skewness might also worth to be explored.

References

[1] A. Azzalini, "A Class of Distributions Which Includes the Normal Ones," *Scandinavian Journal of Statistics*, vol. 12, no. 2, pp. 171–178, 1985. [Online]. Available: http://-www.jstor.org/stable/4615982

[2] C. Baudrit and D. Dubois, "Practical representations of incomplete probabilistic knowledge," *Computational Statistics & Data Analysis*, vol. 51, no. 1, pp. 86 – 108, 2006, the Fuzzy Approach to Statistical Analysis. [Online]. Available: http://www.sciencedirect.com/-science/article/pii/S0167947306000454

[3] C. Baudrit, D. Dubois, and H. Fargier, "Représentation de la connaissance probabiliste incomplète," in *Actes des rencontres francophones sur la Logique Floue et ses Applications (LFA 2003)*, 2003, pp. 65–72.

[4] L.-A. Chen, J.-Y. Huang, and H.-C. Chen, "Parametric coverage interval," *Metrologia*, vol. 44, no. 2, pp. L7–L9, mar 2007. [Online]. Available: https://doi.org/10.1088/0026-1394/-44/2/n01

[5] F. Critchley and M. C. Jones, "Asymmetry and gradient asymmetry functions: Density-based skewness and kurtosis," *Scandinavian Journal of Statistics*, vol. 35, no. 3, pp. 415–437,

2008. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9469.2008.00599.x

[6] S. Das, P. K. Mandal, and D. Ghosh, "On homogeneous skewness of unimodal distributions," *Sankhyā: The Indian Journal of Statistics, Series B (2008-)*, vol. 71, no. 2, pp. 187–205, 2009. [Online]. Available: http://www.jstor.org/stable/41343028

[7] K. A. Doksum, "Measures of location and asymmetry," *Scandinavian Journal of Statistics*, vol. 2, no. 1, pp. 11–22, 1975. [Online]. Available: http://www.jstor.org/stable/4615571

[8] D. Dubois and H. Prade, *Fuzzy sets and systems : theory and applications*. Academic Press, N.Y., 1980.

[9] D. Dubois, S. Sandri, and H. Prade, "On possibility/probability transformation," in *Fuzzy Logic*, R. Lowen and M. Roubens, Eds. Kluwer Academic Press, 1993, pp. 103–112.

[10] D. Dubois, L. Foulloy, G. Mauris, and H. Prade, "Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities," *Reliable Computing*, vol. 10, pp. 273–297, 2004. [Online]. Available: https://doi.org/10.1023/B:REOM.0000032115.22510.b5

[11] D. Dubois, H. T. Nguyen, and H. Prade, "Possibility theory, probability and fuzzy sets: misunderstandings, bridges and gaps," in *Fundamentals of Fuzzy Sets* , ser. The Handbooks of Fuzzy Sets Series, D. Dubois and H. Prade, Eds. Boston, Mass.: Kluwer, 2000, pp. 343–438.

[12] D. Dubois and H. Prade, "When upper probabilities are possibility measures," *Fuzzy Sets and Systems*, vol. 49, no. 1, pp. 65 – 74, 1992. [Online]. Available: http://-www.sciencedirect.com/science/article/pii/016501149290110P

[13] C. Fernandez and M. F. J. Steel, "On Bayesian Modeling of Fat Tails and Skewness," *Journal of the American Statistical Association*, vol. 93, no. 441, pp. 359–371, 1998. [Online]. Available: http://www.jstor.org/stable/2669632

[14] A. Ferrero, J. M. Pou, M. Prioli, and S. Salicone, "Possibility and probability: application examples and comparison of two different approaches to uncertainty evaluation," in *17th International Congress of Metrology, Paris, France*, 2015.

[15] J. F. Geer and G. J. Kleer, "A mathematical analysis of information-preserving transformation between probabilistic and possibilistic formulation of uncertainty," *International Journal of General Systems*, vol. 20, no. 2, pp. 143–176, 1992. [Online]. Available: https://doi.org/10.1080/03081079208945024

[16] R. A. Groeneveld and G. Meeden, "Measuring skewness and kurtosis," *Journal of the Royal Statistical Society. Series D (The Statistician)*, vol. 33, no. 4, pp. 391–399, 1984. [Online]. Available: http://www.jstor.org/stable/2987742

[17] GUM. (2008) Evaluation of measurement data - Supplement 1 to the guide to the expression of uncertainty in measurement - Propagation of distributions using a Monte Carlo method. GUM. [Online]. Available: http://www.bipm.org/utils/common/documents/jcgm/-JCGM_101_2008_E.pdf

[18] G. Jumarie, "Possibility-Probability Transformation: a New Result via Information theory of Deterministic Functions," *Kybernetics*, vol. 23, no. 5, pp. 56–59, 1994. [Online]. Available: https://doi.org/10.1108/03684929410064509

[19] ——, "Further Results on Possibility-Probability Conversion via relative Information and Informational Invariance," *Cybernetics and Systems*, vol. 26, no. 1, pp. 111–128, 1995. [Online]. Available: https://doi.org/10.1080/01969729508927490

[20] G. J. Klir and B. Parviz, "Probability-possibility transformations: a comparison," *International Journal of General Systems*, vol. 21, no. 3, pp. 291–310, 1992. [Online]. Available: https://doi.org/10.1080/03081079208945083

[21] G. J. Klir and J. F. Geer, *Information-Preserving Probability-Possibility Transformations: Recent Developments*. Dordrecht: Springer Netherlands, 1993, pp. 417–428. [Online]. Available: https://doi.org/10.1007/978-94-011-2014-2_39

[22] M. Laviolette and J. W. Seaman, "The efficacy of fuzzy representations of uncertainty," *IEEE Transactions on Fuzzy Systems*, vol. 2, no. 1, pp. 4–15, Feb 1994. [Online]. Available: https://doi.org/10.1109/91.273116

[23] S.-H. Lin, W. Chan, and L.-A. Chen, "A non-parametric coverage interval," *Metrologia*, vol. 45, no. 1, pp. L1–L4, jan 2008. [Online]. Available: https://doi.org/10.1088/0026-1394/45/1/n01

[24] H. L. MacGillivray, "Skewness and asymmetry: Measures and orderings," *The Annals of Statistics*, vol. 14, no. 3, pp. 994–1011, 1986. [Online]. Available: http://www.jstor.org/stable/3035554

[25] G. Mauris, "Representing and approximating symmetric and asymmetric probability coverage intervals by possibility distributions," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 1, pp. 41–45, Jan 2009. [Online]. Available: https://doi.org/10.1109/TIM.2008.2004980

[26] ——, "A review of relationships between possibility and probability representations of uncertainty in measurement," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 3, pp. 622–632, March 2013. [Online]. Available: https://doi.org/10.1109/TIM.2012.2218057

[27] ——, "A possibilistic view of binomial estimation," in *Proceedings of the 15th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2014), Montpellier, France*, 2014, pp. 396–405.

[28] G. Mauris, V. Lasserre, and L. Foulloy, "A fuzzy approach for the expression of uncertainty in measurement," *Measurement*, vol. 29, no. 3, pp. 165–177, 2001. [Online]. Available: https://doi.org/10.1016/S0263-2241(00)00036-1

[29] H. Oja, "On location, scale, skewness and kurtosis of univariate distributions," *Scandinavian Journal of Statistics*, vol. 8, no. 3, pp. 154–168, 1981. [Online]. Available: http://www.jstor.org/stable/4615828

[30] M. Oussalah, "On the probability/possibility transformations: a comparative analysis," *International Journal of General Systems*, vol. 29, no. 5, pp. 671–718, 2000. [Online]. Available: https://doi.org/10.1080/03081070008960969

[31] P. N. Patil, P. P. Patil, and D. Bagkavos, "A measure of asymmetry," *Statistical Papers*, vol. 53, no. 4, pp. 971–985, Nov 2012. [Online]. Available: https://doi.org/10.1007/s00362-011-0401-6

[32] K. Pearson, "IX. Mathematical Contributions to the Theory of Evolution. - XIX. Second Supplement to a Memoir on Skew Variation," *Philosophical Transactions of the Royal Society of London*, vol. 216, 1916. [Online]. Available: https://doi.org/10.1098/rsta.1916.0009

[33] K. Pearson and O. M. F. E. Henrici, "X. Contributions to the mathematical theory of evolution. - II. Skew variation in homogeneous material," *Philosophical Transactions of the Royal Society of London*, vol. 186, 1895. [Online]. Available: https://doi.org/10.1098/-

rsta.1895.0010

[34] M. R. Sampford, "Some Inequalities on Mill's Ratio and Related Functions," *The Annals of Mathematical Statistics*, vol. 24, no. 1, pp. 130–132, 1953. [Online]. Available: https://-www.jstor.org/stable/2236360

[35] A. Stuart and K. Ord, *Kendall's Advanced Theory of Statitics - Volume I - Distribution Theory*. Wiley-Blackwell, 6th edition, 1994.

[36] T. Sudkamp, "On probability-possibility transformations," *Fuzzy Sets and Systems*, vol. 51, no. 1, pp. 73 – 81, 1992. [Online]. Available: http://www.sciencedirect.com/science/-article/pii/016501149290077H

[37] M. S. Tokmachev, "Modeling of truncated probability distributions," *IOP Conference Series: Materials Science and Engineering*, vol. 441, p. 012056, nov 2018. [Online]. Available: https://doi.org/10.1088/1757-899x/441/1/012056

[38] K. F. Wallis, "The Two-Piece Normal, Binormal, or Double Gaussian Distribution: Its Origin and Rediscoveries," *Statistical science*, vol. 29, no. 1, pp. 106–112, May 2014. [Online]. Available: http://www.jstor.org/stable/43288461

[39] L. Xiaojun and J. M. Morris, "On measuring asymmetry and the reliability of the skewness measure," *Statistics & Probability Letters*, vol. 12, no. 3, pp. 267 – 271, 1991. [Online]. Available: http://www.sciencedirect.com/science/article/pii/016771529190089A

[40] L. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, vol. 1, pp. 3–28, 1978. [Online]. Available: https://doi.org/10.1016/S0165-0114(99)80004-9

Appendix A. **Proof of Proposition 3**

Proposition 3. *If a continuous unimodal probability density function* $f(x)$ *with a mode m is differentiable for all* $x \in X$, *the LHS (resp. RHS) part of the two-sided normalized transformation is lower (resp. greater) than the LHS (resp. RHS) part of the optimal transformation if* $F(m) < \frac{1}{2}$.

Intermediate results are necessary and the appendix is organized as follows:

- Proposition 7 and Proposition 8 give a limit for the derivative of the lower bound of the α cut of the optimal possibility distribution respectively when the probability density function is differentiable in all $x \in X$ except in $x = m$ or in all $x \in X$.
- Proposition 9 establishes the respective position of the LHS and RHS of the TSN and optimal possibility distribution with respect to the previous limit.
- Finally, the proof of Proposition 3 is given as a corollary of Proposition 9.

Proposition 7. *If* $[u(\alpha), v(\alpha)]$ *is the* α *-cut of the optimal possibility distribution and* $f(x)$ *is differentiable for all* $x \in X$ *except in* $x = m$ *where it has left and right derivatives, respectively denoted by* $f'_L(m)$ *and* $f'_R(m)$, *then* $\left.\frac{du}{d\alpha}\right|_{\alpha=1} = \frac{f'_R(m)}{f(m)(f'_R(m) - f'_L(m))}$.

Proof. Since the probability density function is unimodal, there exists $\beta \geq 0$ such that $\beta = f(u(\alpha)) = f(v(\alpha))$. It defines $v(\alpha)$ as a function of $u(\alpha)$. Let φ be this function, i.e. $v(\alpha) = \varphi(u(\alpha))$. The function φ is decreasing and such that $\varphi(m) = m$. Thus, $f(u(\alpha)) = f(\varphi(u(\alpha)))$. Differentiating $F(v(\alpha)) - F(u(\alpha)) = 1 - \alpha$ with respect to α gives

$$f(v(\alpha))\frac{dv}{d\alpha} - f(u(\alpha))\frac{du}{d\alpha} = -1. \tag{50}$$

Since $f(v(\alpha)) = f(u(\alpha))$ and $\frac{dv}{d\alpha} = \varphi'(u(\alpha))\frac{du}{d\alpha}$, (50) can be written as

$$f(u(\alpha))(1 - \varphi'(u(\alpha)))\frac{du}{d\alpha} = 1. \tag{51}$$

Differentiating f with respect to $x = u(\alpha)$ gives

$$\forall x \neq m, f'(x) = f'(\varphi(x))\varphi'(x). \tag{52}$$

In m we have

$$f'_L(m) = f'_R(m)\varphi'(m). \tag{53}$$

Replacing (53) in (51) for $\alpha = 1$, i.e. $u(\alpha) = m$, leads to

$$\left.\frac{du}{d\alpha}\right|_{\alpha=1} = \frac{f_R'(m)}{f(m)(f_R'(m) - f_L'(m))}. \tag{54}$$

■

Proposition 8. *If* $[u(\alpha), v(\alpha)]$ *is the* α *-cut of the optimal possibility distribution and* $f(x)$ *is twice differentiable* $\forall x \in X$ *then* $\left.\frac{du}{d\alpha}\right|_{\alpha=1} = \frac{1}{2f(m)}$.

Proof. Let f'' be the second derivative of f with respect to $x = u(\alpha)$, i.e.,

$$f''(x) = f''(\varphi(x))(\varphi'(x))^2 + f'(\varphi(x))\varphi''(x). \tag{55}$$

Since $\varphi(m) = m$, f is unimodal and differentiable in all x, $f'(m) = 0$. Therefore, the second equation in (55) gives

$$f''(m) = f''(m)(\varphi'(m))^2. \tag{56}$$

Since φ is decreasing, $\varphi'(m) = -1$. From (53) it comes $f_L'(m) = -f_R'(m)$ and replacing in (54) completes the proof.

■

Proposition 9. *When* $\frac{F(m)}{f(m)} < \left.\frac{du}{d\alpha}\right|_{\alpha=1}$ *the LHS (resp. RHS) part of the two-sided normalized transformation is lower (resp. greater) than the LHS (resp. RHS) part of the optimal transformation. The property holds also when* $\frac{F(m)}{f(m)} > \left.\frac{du}{d\alpha}\right|_{\alpha=1}$ *by replacing LHS by RHS.*

Proof. Let $\left[u^{Opt}(\alpha), v^{Opt}(\alpha)\right]$ and $\left[u^{TSN}(\alpha), v^{TSN}(\alpha)\right]$ be respectively the α -cut of the optimal possibility distribution π^{Opt} and the α -cut of the two-sided normalized possibility distribution π^{TSN}. The α -cuts of π^{TSN} and π^{Opt} provide two coverage intervals with the same level of confidence. Since $F(v^{TSN}(\alpha)) - F(u^{TSN}(\alpha)) = F(v^{Opt}(\alpha)) - F(u^{Opt}(\alpha))$ and F is an increasing function, the respective position of the intervals are $u^{TSN}(\alpha) \le u^{Opt}(\alpha) \le v^{TSN}(\alpha) \le v^{Opt}(\alpha)$ or $u^{Opt}(\alpha) \le u^{TSN}(\alpha) \le v^{Opt}(\alpha) \le v^{TSN}(\alpha)$.

Let us consider the case where $u^{Opt}(\alpha) \le u^{TSN}(\alpha) \le v^{Opt}(\alpha) \le v^{TSN}(\alpha)$ and define $\Delta(\alpha) = F(u^{TSN}(\alpha)) - F(u^{Opt}(\alpha))$. For $\alpha = 0$ and $\alpha = 1$ we have respectively $F(u^{TSN}(0)) = F(u^{Opt}(0)) = 0$ and $F(u^{TSN}(1)) = F(u^{Opt}(1)) = F(m)$. Since F is a strictly increasing function, $\Delta(\alpha)$ has a maximum or a minimum. If the derivative of $\Delta(\alpha)$ with respect to α, evaluated for $\alpha = 1$, is negative, then $u^{Opt}(\alpha) \le u^{TSN}(\alpha) \le v^{Opt}(\alpha) \le v^{TSN}(\alpha)$

otherwise $u^{TSN}(\alpha) \le u^{Opt}(\alpha) \le v^{TSN}(\alpha) \le v^{Opt}(\alpha)$. According to the definition of π^{TSN}, $F(u^{TSN}(\alpha)) = \alpha F(m)$ and therefore the derivative of $\Delta(\alpha)$ with respect to α is

$$\frac{d\Delta(\alpha)}{d\alpha} = F(m) - F'(x(\alpha))\frac{dx}{d\alpha}. \tag{57}$$

Now, since $F'(m) = f(m)$, evaluating (57) for $\alpha = 1$ leads to

$$\frac{d\Delta(\alpha)}{d\alpha} < 0 \Leftrightarrow \frac{F(m)}{f(m)} < \left.\frac{dx}{d\alpha}\right|_{\alpha=1}. \tag{58}$$

■

Proof of Proposition 3: It is a corollary of Proposition 9 using Proposition 8 in (58).

Appendix B. **Proof of Proposition 4**

Proposition4. *The two-sided normalized transformation is the optimal transformation for the family of two-piece skewed distributions defined by*

$$f_s(x) = A\, g(x\gamma^{-sign(x)}) \text{ with } A = \frac{2}{\gamma + \frac{1}{\gamma}}$$

where g is a unimodal symmetric probability density function.

For the sake of simplicity and without loss of generally, the symmetric probability density function g is supposed to be zero-centered. Let f_{sL} and f_{sR} be the LHS and RHS parts of f_s, i.e.,

$$\begin{cases} f_{sL}(x) = Ag(\gamma x), \forall x \in X_L, \\ f_{sR}(y) = Ag(\frac{y}{\gamma}), \forall y \in X_r. \end{cases} \tag{59}$$

The LHS and RHS parts of the cumulative distribution function F_s of f_s are

$$\begin{cases} F_{sL}(x) = \int_{-\infty}^{x} Ag(\gamma t)\,dt = \frac{A}{\gamma}G(\gamma x), \forall x \in X_L, \\ F_{sR}(y) = F_{sL}(0) + \int_0^y Ag(\frac{t}{\gamma})\,dt = \frac{A}{\gamma}G(0) + A\gamma G(\frac{y}{\gamma}) - A\gamma G(0), \forall y \in X_r, \end{cases} \tag{60}$$

where G is the cumulative distribution function of the symmetric law.

The functions f_{sL} and f_{sR} have inverse, therefore $\varphi(x)$ reduces to the simple expression

$$\varphi(x) = f_{sR}^{-1}(f_{sL}(-x)) = -\gamma^2 x. \tag{61}$$

Thus, the LHS part of the optimal transformation is given by

$$\pi_L^{Opt}(x) = 1 - F_s(-\gamma^2 x) + F_s(x), \forall x \in X_L. \tag{62}$$

Using (60) in (62) leads to

$$\pi_L^{Opt}(x) = 1 + A(\gamma - \frac{1}{\gamma})G(0) - A\gamma G(-\gamma x) + \frac{A}{\gamma} G(\gamma x), \forall x \in X_L. \tag{63}$$

Since G is symmetric, $G(-\gamma x) = 1 - G(\gamma x), \forall x \in X_L$, and $G(0) = \frac{1}{2}$, (63) can be written as

$$\pi_L^{Opt}(x) = 2G(\gamma x), \forall x \in X_L. \tag{64}$$

Extracting $G(\gamma x)$ from the first equation of (60) gives

$$2G(\gamma x) = \frac{2\gamma}{A} F_{sL}(x) = \frac{\gamma^2}{1+\gamma^2} F_{sL}(x), \forall x \in X_L. \tag{65}$$

Since $F_{sL}(0) = A\gamma G(0) = \frac{A\gamma}{2} = \frac{\gamma^2}{1+\gamma^2}$, (64) leads to

$$\pi_L^{Opt}(x) = 2G(\gamma x) = \frac{F_{sL}(x)}{F_{sL}(0)}, \forall x \in X_L. \tag{66}$$

The same approach can be developed for the RHS, i.e. $y \in X_R$, of the possibility distribution by considering now (13) with $\varphi^{-1}(y) = f_{sL}^{-1}(f_{sR}(-y)) = -\frac{y}{\gamma^2}$. It leads to

$$\pi_R(y) = 1 - F_s(y) + F_s(-\frac{y}{\gamma^2}), \forall y \in X_R. \tag{67}$$

Using (60) in (67) gives

$$\pi_R(y) = 1 + A(\gamma - \frac{1}{\gamma})G(0) - A\gamma G(\frac{y}{\gamma}) + \frac{A}{\gamma} G(-\frac{y}{\gamma}), \forall y \in X_R. \tag{68}$$

Since G is symmetric, $G(-\frac{y}{\gamma}) = 1 - G(\frac{y}{\gamma}), \forall y \in X_R$, and (68) becomes

$$\pi_R^{Opt}(y) = 2(1 - G(\frac{y}{\gamma})), \forall y \in X_R. \tag{69}$$

Extracting $G(\frac{y}{\gamma})$ from the second equation of (60) gives

$$G(\frac{y}{\gamma}) = \frac{1}{A\gamma}(F_{sR}(y) + A(\gamma - \frac{1}{\gamma})G(0)), \forall y \in X_R. \quad (70)$$

Replacing $G(0) = \frac{1}{2}$ in (70) leads to

$$2(1 - G(\frac{y}{\gamma})) = \frac{2}{A\gamma}(1 - F_{sR}(y)) = \frac{\gamma^2 + 1}{\gamma^2}(1 - F_{sR}(y)), \forall y \in X_R. \quad (71)$$

Since $1 - F_{sR}(0) = 1 - \frac{A}{\gamma}G(0) = 1 - \frac{A}{2\gamma} = \frac{\gamma^2}{1+\gamma^2}$, (69) can be written as

$$\pi_R^{Opt}(y) = 2(1 - G(\frac{y}{\gamma})) = \frac{1 - F_{sR}(y)}{1 - F_{sR}(0)}, \forall y \in X_R \ . \quad (72)$$

Finally, merging (66) and (72) completes the proof, that is

$$\pi^{Opt}(x) = \min(\frac{F_s(x)}{F_s(0)}, \frac{1 - F_s(x)}{1 - F_s(0)}) = \pi^{TSN}(x), \forall x \in X. \quad (73)$$

■

Appendix C. **Proof of Proposition 5**

Proposition 5. *Let* f *be a unimodal continuous probability density function with a mode* m. *Let* h *be the function such that* $\forall x < m, h(x) = \frac{f(x)}{f(2m - x)}$. *If* h *is strictly increasing from* L *to* 1, *where* $L = 0$ *if* f *is defined on a left-bounded interval or* $L = \lim_{x \to -\infty} h(x) < \frac{F(m)}{1 - F(m)}$ *if* f *is defined on* $\mathbb{R}$, *then the asymmetry about* m *is preserved by the TSN transformation.*

Intermediate results are necessary and this appendix is organized as follows:

- Proposition 10 provides a relation between f and $F(m)$ when f is right asymmetric.
- Proposition 11 gives a sufficient condition for f to be right asymmetric.
- Proposition 12 and Proposition 13 provide the lower bounds for h in the case of probability density functions defined on a left-bounded interval or in the case of skewed right-asymmetric probability density functions.
- Proposition 14 gives results about the extrema of $asym_{\pi^{TSN}}$.
- Finally, the proof of Proposition 5 is given.

Proposition 10. *If* f *is right asymmetric about* m *then* $F(m) < \frac{1}{2}$.

Proof. If f is right asymmetric about m, we have $F(x) < 1 - F(2m - x)$ with F increasing. Thus, in m we have $F(m) < 1 - F(m)$.

■

Proposition 11. *If for all $x < m$, h is strictly increasing, then f is right asymmetric about m.*

Proof: Since f is continuous $\lim_{x \to m} f(x) = \lim_{x \to m} f(2m - x) = f(m)$, thus $\lim_{x \to m} h(x) = 1$. If h is strictly increasing $\forall x < m$ then $h(x) < 1$ which leads to $f(x) - f(2m - x) < 0$.

■

Proposition 12. *If f is defined on a bounded interval whose left bound is b and h is strictly increasing for all $x < m$ then $L = \lim_{x \to b} h(x) = 0$.*

Proof. We already know that $f(x) < f(2m - x)$ from Proposition 11. The left bound gives $\lim_{x \to b} f(x) = 0 < \lim_{x \to b} f(2m - x) = f(2m - b)$, thus $L = \lim_{x \to b} h(x) = \lim_{x \to b} \frac{f(x)}{f(2m - x)} = 0$.

■

Proposition 13. *If f is a skewed probability density function defined for all $x \in \mathbb{R}$ by $f(x) = 2\phi(x)\Phi(\alpha x)$, with ϕ a symmetric density function defined on $\mathbb{R}$, and h is strictly increasing for all $x < m$ then $L = \lim_{x \to -\infty} h(x) = 0$.*

Proof. We have $\lim_{x \to -\infty} h(x) = \lim_{x \to -\infty} \frac{\phi(x)\Phi(\alpha x)}{\phi(-x)\Phi(-\alpha x)}$. Since ϕ is symmetric, $\phi(x) = \phi(-x)$ and $L = \lim_{x \to -\infty} h(x) = \lim_{x \to -\infty} \frac{\Phi(\alpha x)}{1 - \Phi(\alpha x)} = 0$.

■

Remark. For other distributions defined on $\mathbb{R}$, no general answer can be provided since $L = \lim_{x \to -\infty} h(x)$ is under the indeterminate form $0/0$.

Proposition 14. *If f is right asymmetric about m and $asym_{\pi^{TSN}}$ has only one extremum, then the asymmetry about m is preserved.*

Proof. Without loss of generality, the probability density function is assumed to be defined on $\mathbb{R}$ and we have the two properties

$$\begin{cases} \lim_{x \to -\infty} asym_{\pi^{TSN}}(x) = 0, \\ \lim_{x \to m} asym_{\pi^{TSN}}(x) = 0. \end{cases} \tag{74}$$

The derivative of $asym_{\pi^{TSN}}$ for all $x < m$ is

$$asym'_{\pi^{TSN}}(x) = \frac{f(x)}{F(m)} - \frac{f(2m-x)}{1-F(m)}. \tag{75}$$

Its limits are

$$\begin{cases} \lim_{x\to-\infty} asym'_{\pi^{TSN}}(x) = 0, \\ \lim_{x\to m} asym'_{\pi^{TSN}}(x) = \dfrac{f(m)(1-2F(m))}{F(m)(1-F(m))}. \end{cases} \tag{76}$$

If f is defined on a bounded interval whose left bound is denoted by b, since we consider probability density functions such that $\lim_{x\to b} f(x) = 0$, we have $\lim_{x\to-\infty} asym'_{\pi^{TSN}}(x) < 0$.

If f is right asymmetric then $\dfrac{f(m)(1-2F(m))}{F(m)(1-F(m))} > 0$ because $F(m) < \dfrac{1}{2}$ according to Proposition 10.

According to the limits of $asym_{\pi^{TSN}}$ and the limits of $asym'_{\pi^{TSN}}$, the function $asym_{\pi^{TSN}}$ has at least one extremum. If this extremum is unique, since the limit of $asym'_{\pi^{TSN}}$ in m is positive, this extremum is the minimum and $asym_{\pi^{TSN}}$ is negative.

■

Proof of Proposition 5. Extrema of the function $asym_{\pi^{TSN}}$ are reached at the points where $asym'_{\pi^{TSN}}(x) = 0$ or equivalently at the points where $h(x) = \dfrac{F(m)}{1-F(m)}$.

If h is strictly increasing, f is right asymmetric according to Proposition 11. Then, according to Proposition 10, $F(m) < \dfrac{1}{2}$ and therefore $\dfrac{F(m)}{1-F(m)} < 1$.

We have $\lim_{x\to m} h(x) = 1$ and $h > 0$ since f is a density function. Thus if h is strictly increasing, only one point exists where $h(x) = \dfrac{F(m)}{1-F(m)}$. The function $asym_{\pi^{TSN}}$ is minimum in this point. Then, using Proposition 14 completes the proof.

■

Appendix D. **Other examples of asymmetry preservation**

This section provides four other examples of distributions for which the TSN transformation preserves the asymmetry because the function ω is strictly convex.

1. Gumbel distribution

The Gumbel distribution is also a particular case of the Generalized Extreme Value distribution family and is defined by

$$f(x)=\frac{ze^{-z}}{\beta},\forall x\in\mathbb{R}\text{, with } z=e^{-\frac{x-\mu}{\beta}}. \tag{77}$$

It leads to $\omega(x)=\frac{e^{-\frac{x-\mu}{\beta}}-1}{\beta}$ and $\omega''(x)=\frac{e^{-\frac{x-\mu}{\beta}}}{\beta^3}>0$.

Whatever the parameters μ and β, the TSN transformation preserves the asymmetry of the Gumbel distribution.

2. Fréchet distribution

The Fréchet distribution is another case of the Generalized Extreme Value distribution defined by

$$f(x)=\frac{\alpha}{s}\left(\frac{x-m^-}{s}\right)^{-1-\alpha} e^{-\left(\frac{x-m^-}{s}\right)^{-\alpha}}, \forall x > m^-, \tag{78}$$

with $\alpha>0$ and m^- the lower bound of X.

The function ω and its second derivative are respectively $\omega(x)=-\dfrac{(\alpha+1)\left(\dfrac{x-m^-}{s}\right)^{\alpha}-\alpha}{(x-m^-)\left(\dfrac{x-m^-}{s}\right)^{\alpha}}$

and $\omega''(x)=(\alpha+1)\dfrac{\alpha(\alpha+2)-2\left(\dfrac{x-m^-}{s}\right)^{\alpha}}{\left(x-m^-\right)^3\left(\dfrac{x-m^-}{s}\right)^{\alpha}}$.

The denominator of ω'' if always positive since $x > m^-$. Let x_1 be the point where the numerator is equal to 0, i.e.,

$$\alpha(\alpha+2)-2\left(\frac{x_1-m^-}{s}\right)^{\alpha}=0 \Leftrightarrow x_1=m^-+s\left(\frac{\alpha(\alpha+2)}{2}\right)^{\frac{1}{\alpha}}. \tag{79}$$

The mode of the Fréchet distribution is $m=m^-+s\left(\dfrac{\alpha}{1+\alpha}\right)^{\frac{1}{\alpha}}$. Thus, the function ω is strictly convex if $m<x_1$, that is

$$m<x_1 \Leftrightarrow \frac{\alpha}{1+\alpha}<\frac{\alpha(\alpha+2)}{2} \Leftrightarrow 1<\frac{(\alpha+2)(1+\alpha)}{2}. \tag{80}$$

Since $\alpha>0$, (80) is always true.

Whatever the parameters α and m^-, the TSN transformation preserves the asymmetry of the Fréchet distribution.

3. Rayleigh distribution

The probability density function of the Rayleigh distribution is defined by

$$f(x)=\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}, \forall x \geq 0, \text{ with } \sigma>0. \tag{81}$$

For this distribution, we have $\omega(x)=\dfrac{\sigma^2-x^2}{\sigma^2 x}$ and $\omega''(x)=\dfrac{2}{x^3}$. Thus, ω is strictly convex.

The TSN transformation preserves the asymmetry of the Rayleigh distribution whatever the value of the parameter σ .

4. Inverse Gaussian distribution

The inverse Gaussian distribution is defined by

$$f(x)=\left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} e^{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}}, \forall x>0, \text{ with } \mu>0 \text{ and } \lambda>0. \tag{82}$$

It leads to $\omega(x)=-\dfrac{\lambda x^2+3\mu^2 x-\lambda\mu^2}{2\mu^2 x^2}$ and $\omega''(x)=\dfrac{3(\lambda-x)}{x^4}$. The mode of the probability density function is $m=\mu\left(\sqrt{1+\dfrac{9\mu^2}{4\lambda^2}}-\dfrac{3\mu}{2\lambda}\right)$. For all $x<m$ the function ω is strictly convex if $m<\lambda$.

Let $a=\dfrac{\mu}{\lambda}$, then $m<\lambda \Leftrightarrow a\sqrt{1+\dfrac{9}{4}a^2}<1+\dfrac{3}{2}a^2 \Leftrightarrow 0<1+2a^2$ which is always true.

The TSN transformation preserves the asymmetry of the inverse Gaussian distribution whatever the parameters $\lambda>0$ and $\mu>0$.

Acknowledgement

The authors would like to thank Prof. Pierre Baras, LAMA, University Savoie Mont Blanc, for his help in Propositions 7 and 8.