Iterative Adjustment of Survival Functions by Composed Probability Distortions

Alexis Bienvenüe, Didier Rullière

Iterative adjustment of survival functions by composed probability distortions*

Alexis Bienvenüe and Didier Rullière†

Abstract

We introduce a parametric class of composite probability distortions which can be combined to converge to a target survival function. These distortions respect analytic invertibility and stability, which are shown to be relevant in many actuarial fields. We study the asymptotic impact of such distortions on hazard rates. The paper provides an estimation methodology, including hints for initialization. Some applications to survival data bring results for catastrophic event impact modeling. We also obtain accurate parametric representations of the mortality trend over years. At last, we suggest a prospective mortality simulation model which comes naturally from the above analysis.

Key-words: Probability distortions, mortality, iterated compositions, hyperbolic transform, risk measure, survival function transformation, conversion function.

1 Introduction

In an insurance company, many problems may occur when analyzing data mortality. First, it may be necessary to use a reference mortality table, especially when there is a lack of data at some ages, or when the construction of a whole mortality table is excluded. In this case, the reference mortality table lies on a population with a specific risk, distinct from the one of the insurance company. These differences of risk-exposed population require an adaptation of one table given the other, which can be expressed as a parametric deformation. Second, a precise representation of mortality over ages shows some local phenomena, leading to a non monotone hazard rate, which may require a relatively complex parametric shape. Third, the analysis of the evolution of mortality rates over time requires a model that can stay reliable after years.

A large literature deals with these problems. To adapt a mortality insurance table given a reference one, one may use *Proportional Hazard* transform or Wang transform (see Wang, 1996). Heligman, Pollard (1980) studied the precise structure of mortality as a function of the age. Lee, Carter (1992) described the evolution of the mortality over time, and many other authors suggest different parametric representations of mortality and its evolution (see Pitacco, 2004).

Nevertheless, these classical parametric solutions have several drawbacks:

- These solutions do not improve data adequation, and adding parameters is relatively tricky. This way, considering Wang transforms (Wang, 1996), the use of several successive transforms does not extend the class of transformed survival functions; the adaptation of one table given another with a single parameter may remain insufficiently accurate, and parameters adjunction could denature such a transform. Among other models, such the ones of Heligman, Pollard (1980) or Lee, Carter (1992), potential extensions may lead to very different expressions depending on the number of parameters that we wish to add, and the convergence properties of such transformation when increasing the number of parameters are unknown.
- The use of several parameters in order to fit data may cause important estimation problems, this estimation being numerically feasible only in the presence of initial values sufficiently close to the solution. Adding parameters or introducing a prospective framework requires the knowledge of initial values that may be hard to obtain.
- Practical simulations of random death dates are sometimes generated from easily invertible survival functions in order to speed up simulations. This choice leads away from previously presented classical models to favor simple, easily invertible laws. The good representation of mortality tables is then reduced with the use of laws having few parameters, like the

* This work has been funded by ANR Research Project ANR-08-BLAN-0314-01.

†Laboratoire SAF, Institut de Science Financière et d'Assurances, Université de Lyon, Université Lyon 1, EA 2429, 50 Avenue Tony Garnier, Lyon F-69007, France. `Alexis.Bienvenue@univ-lyon1.fr`; `Didier.Rulliere@univ-lyon1.fr`

Gompertz one. Thus, parametric inverse distribution functions are sometimes used to obtain stochastic simulations, but the adequacy of a set of mortality tables will not be able to exceed a given precision.

Many parametric expressions have been suggested to deal with each of those problems but they assume different forms, and it is interesting to look for a common parametric form, which may be used for probability distortions, for static and prospective mortality tables, and for inverse distribution function intended for stochastic simulations. Moreover, depending on desired accuracy, the choice of the number of parameters, without modifying the nature of the adjustment, is a question of great importance which is difficult to solve with classical tools.

Trying to give a helpful tool for all the issues we have introduced, it is natural to suggest the use of probability distortions, and to consider the composition of these distortions. Composed distortions allow us to get accurate and easily invertible adjustments of survival functions, with the possibility of increasing the number of parameters in order to converge to a target law. This choice can be useful to many issues, such as pricing or risk measuring.

In this paper, we show how our distortions modify random variables (Prop. 1, linked with Accelerated Failure Time models), hazard rates (Prop. 2) and stop-loss premiums in the regular variation case (Prop. 3). The main result of this paper is to establish that some particular distortions reduce the number of parameters (Th. 4), that these distortions allow an initial survival function to converge to any target survival function (Th. 5), and that accurate initialization values can be given for parameter estimation (Prop. 6).

The paper is structured as follows: in section 2, we introduce some general uses of probability distortions in the actuarial field, and the more specific constraints that we have chosen for our distortions. In section 3, we deal with the general form of these distortions. Some initial results on distorted random variables are given here. In particular, section 3.3 gives specific examples of distortions, mainly smoothed and composed versions of a basic class of *angle* functions. The estimation problem and the convergence demonstration of chosen distortions to any survival function target is explained in section 4. Lastly, some applications are given in the specific field of multiple mortality tables adjustment in section 5.

2 Probability distortions and constraints

There are many different aims when using probability distortions, including:

- Obtaining a parametric form for quantity of interest, improving the fit of a reference with real data (adjusting an official mortality table to business data, adjusting claims distribution on a segment given a global distribution).
- Explaining a phenomenon by the considered distortion, the parametric distortion being the main center of interest (e.g. explaining the evolution of a phenomenon over time).
- Applying a prudential rule, adding weight to the distribution's tail, or more generally to take into account phenomena that are not observed into data (carrying out a loading which preserve bracket pricing, giving a solvability margin).

The first use of probability distortions can be attributed to d'Alembert, J. Le Rond (1768). Amount distortions by way of utility functions appeared in Bernoulli's treaties (see Bernoulli, D., 1731). A few years later, d'Alembert suggested distorting probabilities themselves (see Pradier, 1998). Ironically, his intention was not to take into account a prudential constraint, but on the contrary to lessen rare events, in order to answer to the well-known Saint-Petersbourg's paradox.

More recently, probability distortions gained interest. In the economics, as utility functions modify amount perception while keeping probabilities unchanged, the dual theory from Yaari (1987) keeps amounts unchanged while distorting probabilities (see Bleichrodt, Eeckoudt, 2006, for applications in the actuarial situations). These different points of view can be seen as heirs of antagonistic views from d'Alembert and Bernoulli. In the actuarial field, probability distortions have been popularized by Wang's work. He used different distortions for pricing, and for risk measurement (see Wang, 1996; Wirch, Hardy, 1999). Risk measure evaluation for financial assets are also concerned by probability distortions, as illustrated in Wang (2000) or Hamada, Sherris (2003). Constraints can nevertheless appear in such an evaluation, as detailed in Pelsser (2007).

Generally speaking, risk measurement is framed by numerous axioms or principles on probability distortions (one can refer to Bühlman, 1980; Artzner, Delbaen, Eber, Heath, 1997; Landsman, Sherris, 2001; Goovaerts et al., 2004, and to article quoted therein). Thus, distortions are usually suggested as a viewpoint on prudential and risk analysis, following an axiomatic set of constraints characteristic of this field.

When one needs distortions likely to fit to data as closely as desired, and able to maintain some key properties like analytic invertibility of survival functions, one faces some deeply different constraints. Some authors use distortions to model the temporal evolution of risk, like mortality. As an example, the article of De Jong, Marshall (2007) is based on the evolution of Wang's transform parameters, and give projections of mortality tables. Nevertheless, some properties that seem helpful to us are not satisfied with the transforms they use, such as the ability of a transformation to be iterated in order to get as close as wanted to business data.

We try to detail more precisely these constraints, aiming in particular at an invertible parametric form for a quantity of interest. The demands of analytic invertibility emanates from the pragmatic desire for ease of simulation of continuous random variables, conditional on their belonging to a given set. Here distortions are simple real functions, applied to survival functions, and the problem of the composition of such functions is also addressed. Ideally, the result is a representation of a survival function as a composition of several parametric functions. This aim is similar to the idea of a wavelet decomposition of a function: getting a class of functions large enough to generate (here by composition) target functions in several kind of problems, relevant enough to necessitate only a restricted number of parameters. These functions should also preserve some properties likely to be helpful to actuarial problems. Here, we present some distortions which properties seem to us interesting, and which are efficient in our numerical applications.

Specific constraints

We try in this paper to restrict the huge set of possible choices for probability distortions by suggesting a set of constraints which are relevant for many actuarial issues. We consider a class of distortions $\mathcal{T}$, which will be applied on survival functions from a class $\mathcal{S}$, so that each distorted function is also a survival function:

$$\forall T \in \mathcal{T}, \forall S \in \mathcal{S},\ T \circ S \in \mathcal{S}\,.$$

The class of distorsions $\mathcal{T}$ consists of the set of real functions T_θ, for some vector of parameters $\theta \in \Theta$, $\Theta \subset \mathbb{R}^p$, $p \in \mathbb{N}^*$:

$$\mathcal{T} = \{T_\theta : [0,1] \to [0,1]\}_{\theta \in \Theta}\ .$$

We will try to find a distortion with a reduced number of parameters and with an analytic expression likely to be easily computed with common computer languages. We set five constraints for these distortions, detailed below.

C1. Invertibility Simulation techniques being very commonly used in actuarial work, the preservation of the invertible character of a survival function arise from the knowledge of the analytic expression of the inverse distortion function T_θ: this knowledge allows easy simulation of random variables from the distorted law, given that this random variable belongs to a given set. Such a simulation is straightforward when applying the inverse survival function to an uniform random variable on a subset interval of $]0,1[$, but requires easy computation of the inverse function. The choice of working on survival functions may be explained by the presence, in life or non-life insurance, of conditioning on overshooting a given threshold by considered random variables.

$$\forall u \in]0,1[, \forall T_\theta \in \mathcal{T},\ \exists! v \in]0,1[,\ T_\theta(v) = u\,.$$

C2. Stability Ideally, we try to preserve the intuitive interest of being able to distort a function in a direction or its opposite, by demanding that inverse distortions belong to the same class as original distortions. This helps symmetry properties, as well as computer coding of distortions and their inverse functions. Under this constraint, exchanging the target function and the initial one will modify distortion parameters, but not the distortion expression itself. This seem logical without a priori information on the shape of target or approximated functions.

$$\forall T_\theta \in \mathcal{T}, \exists \theta' \in \Theta, T_\theta^{-1} = T_{\theta'} \in \mathcal{T} \,.$$

C3. Regularity Explaining the distortion is a pragmatic constraint, as is being able to estimate its parameters. We try, for example, to determine the influence of each parameter on some commonly used quantities (expectancies, stop-loss premiums...), to identify the consequences of fixing minimal or maximal possible values for each parameter. This leads us to establish some constraints on parameters, that is on the components of θ vectors, $\theta \in \Theta$. To get some quantitative arguments when a parameter is varying, and for the sake of clarity, we prefer that the set of parameters values be an open hyperrectangle of $\mathbb{R}^p$. As well, interpreting the impact of a parameter on the distortion should not lead to separate the analysis into several cases, and should be interpreted logically; this leads us to formulate continuity and differentiability conditions:

$$\begin{gathered} \Theta \text{ open hyperrectangle of } \mathbb{R}^p \,, \\ \forall x \in [0,1],\, \theta \mapsto T_\theta(x) \text{ continuously differentiable} \,, \\ \forall \theta \in \Theta,\, x \mapsto T_\theta(x) \text{ continuously differentiable} \,. \end{gathered}$$

C4. Convergence In order to better fit a reference survival function or business data, we set a convergence constraint. Applying distortions iteratively should lead us to reduce a specified distance (in the following, L^1 distance) between any target survival function and any initial survival function: iterated transformed functions must converge to the target survival function. We suppose that when the initial survival function is identical to the target function, the distortion does not change this function, so that the identity function belongs to the class of considered distortions.

$$\begin{gathered} \mathrm{Id} \in \mathcal{T} \,, \\ \forall S_0, S_1 \in \mathcal{S},\, \exists \text{ a series } (T_i)_{i\in\mathbb{N}} \text{ of elements of } \mathcal{T},\, T_n \circ \cdots \circ T_1(S_0) \xrightarrow[n\to+\infty]{L^1} S_1 \,. \end{gathered}$$

C5. Parameterization It is possible to change the parameterization of a distortion with a bijection $\mathcal{H}$ from the set Θ of all parameters to a new set $\tilde{\Theta}$. This way, one can replace a distortion T_θ by $\tilde{T}_\theta = T_{\mathcal{H}(\theta)}$. The set of all distortions is then obviously the same, but the parameters meaning, the constraints on parameters, and the ease of estimation could be modified. We prefer the parameters of an inverse distortion when their expression is a simple direct function of the parameters of the initial distortion. Among these preferred parameterizations, we present a particular class which can be expressed more formally: from axiom C2, there exists a bijection I_T which for all $\theta \in \Theta$ associate a $\theta' \in \Theta$ such that $T_\theta^{-1} = T_{\theta'}$, and we present parameterizations leading to

$$T_\theta^{-1} = T_{I_T(\theta)}, \text{ with } I_T(\theta) = D_T \cdot \theta \,,$$

where D_T is a diagonal matrix, with diagonal $\vec{d} = (d_1, \ldots, d_p)$, $d_1, \ldots, d_p \in \{-1, 1\}$. We call such a parameterization a *symmetrical parameterization*. When switching to an inverse distortion, the ith parameter is unchanged if $d_i = 1$. We then call it a *position* parameter. Its sign will change if $d_i = -1$. We then call it a *distance* parameter. The parameterization is said to be *entirely symmetrical* when $\Theta = \mathbb{R}^p$ and $T_\theta^{-1} = T_{-\theta}$ for all $\theta \in \Theta$. This implies in particular $T_{\vec{0}} = \mathrm{Id}$. This can facilitate the interpretation of the change of parameters when deriving the estimation. Entirely symmetrical parameterizations offer the possibility of suppressing a parameter while keeping inverse distortions in the same class, by simply choosing 0 for the value of suppressed parameter.

3 Transformations

3.1 Definitions

Our transformations act on the logit scale, which has been shown to be relevant in various contexts. We focus here on distortions of real random variable survival functions. Let $\mathcal{S}$ be the set of real integrable random variable survival functions, so that functions $S \in \mathcal{S}$ are cadlag from $\mathbb{R}$ to $[0,1]$, $S(x) = 1$ for all $x \leqslant 0$ and $\int_0^{+\infty} S(t)\, dt < \infty$. For $S \in \mathcal{S}$ and f any bijective increasing function from $\mathbb{R}$ to $\mathbb{R}$, we denote T_f the function from $[0,1]$ to $[0,1]$ such that

$$T_f(u) = \begin{cases} 0 & \text{if } u = 0, \\ \text{logit}^{-1}(f(\text{logit}(u))) & \text{if } 0 < u < 1, \\ 1 & \text{if } u = 1. \end{cases}$$

We call f the *conversion function* of the distortion T_f. The logit function and its inverse, $\text{logit}(x) = \ln(x/(1-x))$ and $\text{logit}^{-1}(x) = 1/(1+e^{-x})$, are here used in a very classical way, so that for any f the distortion belongs to $[0,1]$. This choice is not crucial, since the survival function distortion mainly relies on f. The main advantage of the logit function is the simple analytic expression of its inverse. It can be rapidly evaluated, as exponential and logarithm functions are directly computable by the arithmetic coprocessor of modern computers. Note that any distribution function could have been chosen instead. One can easily switch from one setting to the other, modifying the conversion function: $\text{logit}^{-1}\left(f(\text{logit}(u))\right) = \Phi\left(\tilde{f}(\Phi^{-1}(u))\right)$, with $f(u) = \text{logit}(\Phi(\tilde{f}(\Phi^{-1}(\text{logit}^{-1}(u)))))$. In particular, the Wang transform (see Wang, 1996) could be accessed letting Φ be Gaussian distribution function, and $\tilde{f}(x) = x + \lambda$, $\lambda \in \mathbb{R}$.

Setting $\mathbb{T}_f(S)(x) = T_f(S(x))$ for all $x \in \mathbb{R}$, one gets $\mathbb{T}_g \circ \mathbb{T}_f = \mathbb{T}_{g \circ f}$ and $(\mathbb{T}_f)^{-1} = \mathbb{T}_{(f^{-1})}$.

3.2 Impact on random variables

Let X and $\widehat{X}$ be real random variables with respective survival functions $S \in \mathcal{S}$ and $\widehat{S} = \mathbb{T}_f(S)$. In this section, we observe how some characteristics of X are modified by the distortion.

Proposition 1 (From X to $\widehat{X}$) *Let $S \in \mathcal{S}$ be an invertible survival function. Then*

$$\widehat{X} \overset{\mathcal{L}}{=} S^{-1} \circ \text{logit}^{-1} \circ f^{-1} \circ \text{logit} \circ S(X)\,.$$

The proof is straightforward. This depiction of the distortion gives a direct link with Accelerated Failure Times models (AFT), see Bagdonavicius, Nikulin (2002).

Proposition 2 (Hazard rate) *Let $\mu(t)$ and $\widehat{\mu}(t)$ denote the respective hazard rates of one random variable and its transform: $\mu(t) = -S'(t)/S(t)$ and $\widehat{\mu}(t) = -\widehat{S}'(t)/\widehat{S}(t)$. Then, when $t \to \infty$,*

$$\frac{\widehat{\mu}(t)}{\mu(t)} \sim f'(\text{logit}(S(t)))\,. \tag{1}$$

When f has an asymptotic direction $f'(t) \to a$, the hazard rate is asymptotically multiplied by a.

Proof : When $S(t) < 1$, $\widehat{\mu}(t)/(1-\widehat{S}(t)) = f'(\text{logit}(S(t)))\mu(t)/(1-S(t))$, leading to the result. □

Proposition 3 (Regular variations) *Let $Z_0^*(x) = \mathrm{E}\,(X-x)_+$ be the average charge for Stop-Loss reinsurance treaty with priority x, and $\widehat{Z}_0^*(x)$ the same quantity for $\widehat{X}$. Suppose S is regularly varying with exponent $\rho \leqslant 0$, that is $S(tx)/S(t) \xrightarrow[t\to+\infty]{} x^\rho$, and f has an asymptote with slope a, that is $f(u) - (au+b) \xrightarrow[u\to-\infty]{} 0$. Then $\widehat{S}$ is regularly varying with exponent $a\rho$ and*

$$\widehat{Z}_0^*(x) \sim e^b \frac{(-(\rho+1))^a}{-(a\rho+1)} x^{1-a} Z_0^*(x)^a \quad \textit{when } x \to +\infty\,.$$

Proof : Note $Z_p^*(x) = \int_x^{+\infty} t^p S(t)\, dt = \mathrm{E}[(X^{p+1} - x^{p+1})_+]/(p+1)$ (for p such that the integral converges). When S is slowly varying, theorem 1 p. 281, from Feller's book (see Feller, 1968) provides us the following equivalency when $x \to +\infty$ and for $\rho + p + 1 < 0$ and $a\rho + p + 1 < 0$: $-(\rho+p+1)Z_p^*(x) \sim x^{p+1}S(x)$ and $-(a\rho+p+1)\widehat{Z}_p^*(x) \sim x^{p+1}e^b S(x)^a$. □

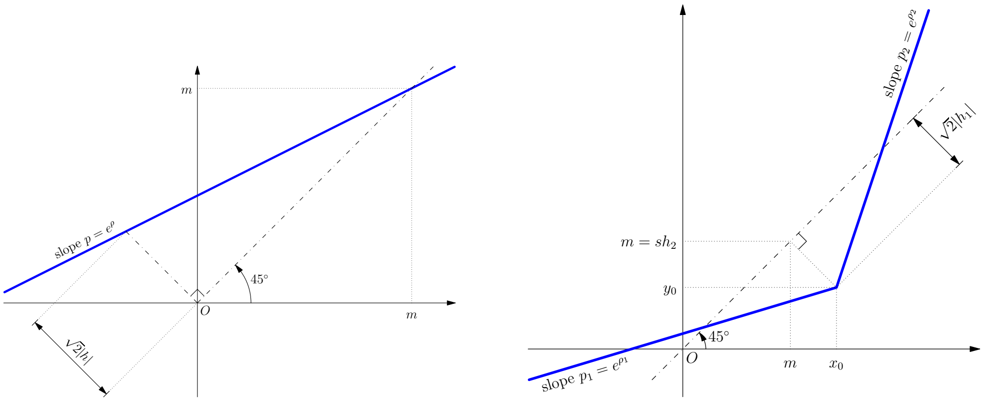

Figure 1: Affine and angle functions

3.3 Conversion functions

Affine functions These functions are defined by two parameters $p > 0$ and m:

$$D_{p,m}(x) = p(x - m) + m\,.$$

See figure 1 for function and parameter illustration. They are obviously invertible, with $(D_{p,m})^{-1} = D_{1/p,m}$. Parameter p is the slope, and m the threshold for which $D_{p,m}(m) = m$, separating the areas where the distorted survival function is increased or not. One can remark that for these functions, the induced distortion correspond to the Brass model (see Brass, 1969, 1974).

Choosing parameters $\rho = \ln p$ and m leads to one distance parameter and one position parameter (see axiom C5). Choosing $h = m(1-p)/(1+p)$ instead of m leads to the entirely symmetrical parameterization: for $h \in \mathbb{R}$ and $\rho \in \mathbb{R}$, $\bar{D}_{\rho,h}(x) = e^{\rho}(x+h) + h$ and $\bar{D}^{-1}_{\rho,h} = \bar{D}_{-\rho,-h}$. ρ is the logarithmic slope and h the height of the intersection with the diagonal $y = -x$.

Angle functions See figure 1 for function and parameter illustration. Angle functions have four parameters: the apex position (x_0, y_0), and two slopes $p_1 > 0$ and $p_2 > 0$. They can be written:

$$A_{x_0,y_0,p_1,p_2}(x) = \begin{cases} y_0 + p_1(x - x_0) & \text{if } x \leqslant x_0, \\ y_0 + p_2(x - x_0) & \text{if } x \geqslant x_0. \end{cases}$$

These functions are bijective functions, with $(A_{x_0,y_0,p_1,p_2})^{-1} = A_{y_0,x_0,1/p_1,1/p_2}$.

Replacing (x_0, y_0) by $(m, h_1) = ((x_0 + y_0)/2, (y_0 - x_0)/2)$, m becomes a position parameter and h_1 a distance parameter. Next replace m by $h_2 = sm$, where s is a distance parameter, say $s = \text{sign}(p_1 - p_2)$, so that the angle symmetry is preserved. This leads to an entirely symmetrical parameterization: $\bar{A}_{\rho_1,\rho_2,h_1,h_2} = A_{x_0,y_0,p_1,p_2}$, where $p_1 = e^{\rho_1}$, $p_2 = e^{\rho_2}$, $x_0 = sh_2 - h_1$, $y_0 = sh_2 + h_1$, and $s = \text{sign}(\rho_1 - \rho_2)$, for which $\bar{A}^{-1}_{\rho_1,\rho_2,h_1,h_2} = \bar{A}_{-\rho_1,-\rho_2,-h_1,-h_2}$ and $\bar{A}_{0,0,0,0} = \text{Id}$. In the coordinate system $(O, \vec{i}, \vec{j})$, where $\vec{i} = (1,1)$ and $\vec{j} = (-1,1)$, h_1 is a measure of the vertical position of the apex, and h_2 of its horizontal position.

Hyperbolic functions Hyperbolic functions are smooth versions of the angle functions; see figure 2 for function and parameter illustration. They can be defined using five parameters: apex position (x_0, y_0), asymptotes rates p_1, p_2, and smoothing ϵ:

$$\begin{aligned} H_{x_0,y_0,p_1,p_2,\epsilon}(x) &= y_0 + \frac{p_1 + p_2}{2}(x - x_0) - \text{sign}(p_1 - p_2)\sqrt{((p_1 - p_2)(x - x_0)/2)^2 + \sqrt{p_1 p_2}\epsilon^2}\,, \\ H^{-1}_{x_0,y_0,p_1,p_2,\epsilon} &= H_{y_0,x_0,1/p_1,1/p_2,\epsilon}\,, \end{aligned}$$

with the convention $\text{sign}(0) = 0$. As expected, $H_{x_0,y_0,p_1,p_2,0} = A_{x_0,y_0,p_1,p_2}$. One can also use an entirely symmetrical parameterization: $\bar{H}_{\rho_1,\rho_2,h_1,h_2,e} = H_{x_0,y_0,p_1,p_2,\epsilon}$, with $p_1 = e^{\rho_1}$, $p_2 = e^{\rho_2}$, $x_0 = sh_2 - h_1$, $y_0 = sh_2 + h_1$, $\epsilon = se$, where $s = \text{sign}(\rho_1 - \rho_2)$. We get: $\bar{H}^{-1}_{\rho_1,\rho_2,h_1,h_2,e}(x) = \bar{H}_{-\rho_1,-\rho_2,-h_1,-h_2,-e}(x)$ and $\bar{H}_{0,0,0,0,0} = \text{Id}$.

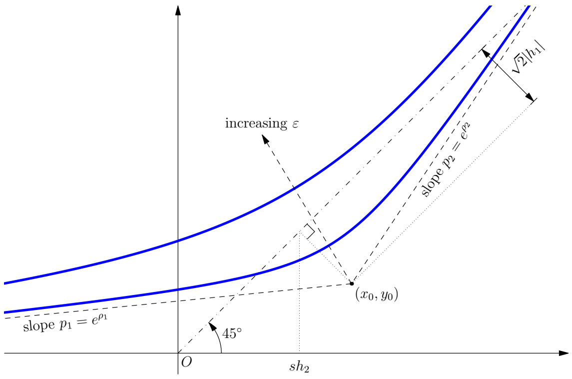

Figure 2: Hyperbolic function

Angle composition It may be useful to employ composite functions as one conversion function:

- The composition of several conversion functions may cause some parameters to be useless. As an example, the composition of n angle functions is entirely characterized by $2n+2$ parameters, which is less than n times the 4 parameters of an angle.
- A particular knowledge (e.g. known asymptotical direction $y = x$ if the transformation is to be local) may simplify the composite function expression and reduce the parameters number.
- Parameter meaning may be clearer with the composite function.

In order to better manage the successive composition of functions, it may be interesting to write a composition of n angles as a composition of one angle with 4 parameters and $n-1$ angles of two parameters, which gives the $2n+2$ degrees of freedom of the global composition.

Let us simply denote by A_4 an angle with 4 parameters, and A_2 an angle with two parameters (of kind $A_{x_0,x_0,1,p}$). We are interested in the form of a $A_4 \circ A_4' \cdots \circ A_4'''$ composition.

Theorem 4 *Any composition of n angles can be reduced to a composition of one angle with 4 parameters and $n-1$ angles with two parameters of kind $A_{x_0,x_0,1,p}$, whatever the position of the angle with 4 parameters. In particular, any composition of angles can be written in the form*

$$A_4^{\circ(n)} = A_2^{\circ(k)} \circ A_4 \circ A_2^{\circ(n-1-k)}, \quad 0 \leqslant k \leqslant n-1\,,$$

where $A_p^{\circ(k)} = A_p' \circ A_p'' \circ \cdots \circ A_p''''$ is the composition of k angles with p parameters, $p \in \{2,4\}$, and $A_p^{\circ(0)} = \mathrm{Id}$. All A_2 denote angles with two parameters, of the kind $A_{x_0,x_0,1,p}$, $x_0 \in \mathbb{R}, p > 0$, with their apex on the diagonal $y = x$, and their first slope equal to 1.

Proof : This derives from the fact that every composition of two angles can be written as a composition of two angles with 2 and 4 parameters, $A_4 \circ A_4' = A_4'' \circ A_2 = A_2' \circ A_4'''$, where all A_2 are angles with two parameters, of kind $A_{x_0,x_0,1,p}$, $x_0 \in \mathbb{R}$, $p > 0$. □

The $2n+2$ parameters that are necessary to characterize the composite function can be decomposed as $4 + 2(n-1)$, and no parameter is useless. The choice of the parameterization, which was a simple preference, is important: if the inverse function of a two parameters angle did not belong to the same class of function, like for example $A_{0,x_0,1,p_2}$, one could not establish the previous result without imposing that A_4 be in the last position.

Shift functions See figure 3 for function illustration. Shift functions have first increasing then decreasing derivative, in order to locally adjust hazard rates, through Proposition 2. They are defined as a smoothed version of two angles composite: $A_{x_0,x_0,1,p} \circ \bar{A}_{x_0',x_0',1,p'}$. Moreover, asymptotic directions are chosen to be one at $+\infty$ and $-\infty$, so that $p' = 1/p$. Finally, Shift functions are:

$$\begin{aligned} Z_{m,h,\rho,\epsilon} &= H_{m-h,m-h,1,e^{\rho},\epsilon} \circ H_{m+h,m+h,1,e^{-\rho},\epsilon}\,, \\ (Z_{m,h,\rho,\epsilon})^{-1} &= Z_{m,-h,\rho,\epsilon}\,. \end{aligned}$$

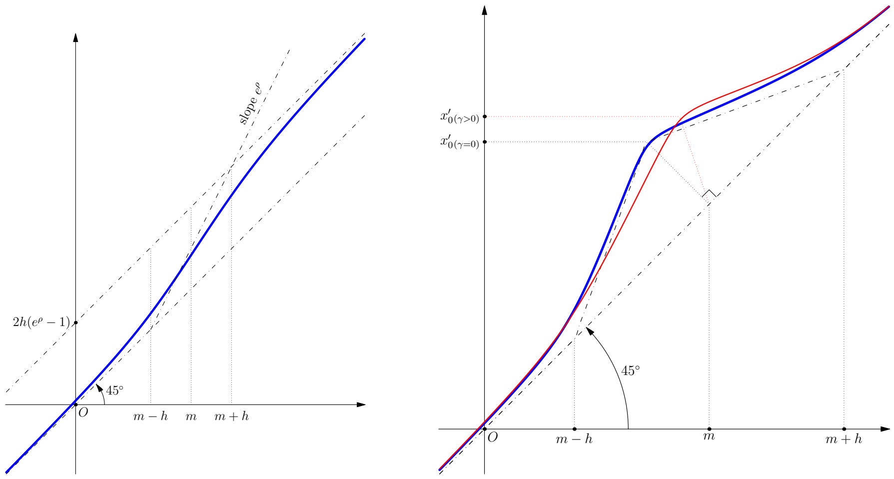

Figure 3: Shift function $Z_{m,h,\rho,\varepsilon}$, and bump function (here $0 \leqslant \gamma < \rho$ and $h > 0$)

For $h = 0$ we get the identity function, and h can be seen as a distance between the two asymptotes. Parameter m localizes the center, and ρ represents the shift speed from an asymptote to the other. Shift functions act in the same way as Wang's transform (see Wang, 1996), with a smooth transition between two levels. They may be useful in non-life insurance context, when only the distribution tail has to be changed.

Bump functions These functions are smooth versions of a three $A2$-angle composite, with fixed asymptotes with equation $y = x$ at $-\infty$ and $+\infty$, so that the adjustment is local; see figure 3 for an illustration. Without smoothing, these Bumps correspond to: $B^0_{x_0,p_2,x''_0,p''_2} = A_{x_0,x_0,1,p_2} \circ A_{x'_0,x'_0,1,p'_2} \circ A_{x''_0,x''_0,1,p_2}$, with $p'_2 = 1/(p_2 p''_2)$ and $x'_0 = (x''_0(p''_2 - 1) + x_0 p''_2(p_2 - 1))/(p_2 p''_2 - 1)$. Changing parameterization, we define a smooth version as (for $\rho \neq 0$):

$$\begin{aligned} B_{m,h,\rho,\gamma,\epsilon} &= H_{m+h,m+h,1,e^{\rho+\gamma},\epsilon} \circ H_{x'_0,x'_0,1,e^{-2\rho},\epsilon} \circ H_{m-h,m-h,1,e^{\rho-\gamma},\epsilon}\,, \\ \text{with } x'_0 &= m + h\left(\frac{e^{2\rho} - 2e^{\rho-\gamma} + 1}{e^{2\rho} - 1}\right). \end{aligned}$$

When $\rho \neq 0$, $B^{-1}_{m,h,\rho,\gamma,\epsilon} = B_{m,-h,-\rho,\gamma,\epsilon}$. The degenerate case $\rho = 0$ corresponds to the identity function. m represent the horizontal position of the Bump, h its height. Slopes p_2 and p''_2 acting on the left-hand side and right-hand side of the Bump, ρ can be seen as the return to asymptote speed, and γ as an symmetry coefficient.

Hyperbolic composite functions In some situations where increasing the number of parameters is needed for better accuracy, we compose several hyperolic functions. By Theorem 4, we compose smooth versions of a 4-parameters angle and $(n-1)$ 2-parameters angles. Choosing the same smoothing parameter for all these functions leads to:

$$G_{(x_0,y_0,p_1,p_2,e),(a_1,q_1),\ldots,(a_n,q_n)} = H_{a_n,a_n,1,q_n,e} \circ \cdots \circ H_{a_1,a_1,1,q_1,e} \circ H_{x_0,y_0,p_1,p_2,e}\,,$$

$$\left(G_{(x_0,y_0,p_1,p_2,e),(a_1,q_1),\ldots,(a_n,q_n)}\right)^{-1} = H_{y_0,x_0,\frac{1}{p_1},\frac{1}{p_2},e} \circ H_{a_1,a_1,1,\frac{1}{q_1},e} \circ \cdots \circ H_{a_n,a_n,1,\frac{1}{q_n},e}\,.$$

From Theorem 4, all increasing continuous stepwise linear functions with $n+1$ vertices can be written this way, with any possible position within the composition for the 5-parameter hyperbola. We will see in paragraph 4.3 that initialization parameters are easy to obtain. These functions are very well adapted for the determination of one monotone analytically invertible parametric function corresponding to a particular dataset. This kind of situation usually occurs when one needs to sample from a smoothed empirical distribution.

4 Estimation and convergence of iterative adjustment

4.1 Estimation methodology

Here, we aim at transforming a survival function $S_0 \in \mathcal{S}$ in order to get it close to another survival function $S \in \mathcal{S}$. We consider for this purpose the distance D on $\mathcal{S}$, defined by

$$D(S, S') = \int_0^{+\infty} |S(t) - S'(t)| \, dt \, .$$

This distance is finite for every couple of $\mathcal{S}^2$ due to the integrability hypothesis on elements of $\mathcal{S}$.

Remark 1 *Let X and X' be two nonnegative random variables with respective survival functions S and S'. Then $|\mathrm{E}\, X' - \mathrm{E}\, X| \leqslant D(S, S')$. Thus, distance D allows to control the difference between the expectations of these two random variables.*

Restricting on the family $(\mathbb{T}_{f_\theta})_\theta$ of transformations, with conversion functions $(f_\theta)_\theta$ parameterized with vector θ, one gets:

$$S_1 = \mathbb{T}_{f_{\theta_0^*}}(S_0) \, , \quad \text{where} \quad \theta_0^* = \arg\min_\theta D(\mathbb{T}_{f_\theta}(S_0), S) \, .$$

One can iterate this process, defining a survival functions sequence $(S_n)_n$:

$$S_2 = \mathbb{T}_{f_{\theta_{2,1}^*}}(\mathbb{T}_{f_{\theta_{2,0}^*}}(S_0)) \, , \quad \text{where} \quad (\theta_{2,0}^*, \theta_{2,1}^*) = \arg\min_{\theta_0, \theta_1} D(\mathbb{T}_{f_{\theta_1}}(\mathbb{T}_{f_{\theta_0}}(S_0)), S) \, , \quad \text{etc.}$$

4.2 Convergence

We try here to check that the constraint C4 holds for most of the conversion functions we have suggested. We write, for $S \in \mathcal{S}$, $M(S) = \{x \in \mathbb{R}, \ S(x) \in]0, 1[\}$. The set $M(S)$ can be seen as a support interval for the survival function S, or for its underlying random variable. The following theorem shows that any suitable initial survival function S_0 can be iteratively distorted so that the resulting survival function $\mathbb{T}_{f_1 \circ \cdots \circ f_n} S_0$ converges to S. Suitability conditions on S_0 are only depending on its natural support and on its strict monotony.

Theorem 5 (Convergence to any target) *Let $S \in \mathcal{S}$ be a given target survival function. Suppose that $S_0 \in \mathcal{S}$ is such that $M(S) \subset M(S_0)$ and S_0 is strictly decreasing on $M(S_0)$. Then, for the families $(A_\theta)_\theta$, $(H_\theta)_\theta$, and families built by composition from these ones, like $(G_\theta)_\theta$,*

$$\lim_{n \to \infty} D(S_n, S) = 0 \, .$$

In particular, any strictly decreasing $S_0 \in \mathcal{S}$ on $\mathbb{R}$ is suitable for any target $S \in \mathcal{S}$.

Proof : Let $a, b \in \mathbb{R}_+$ such that $a < b$. Consider $\varepsilon > 0$. Let us first prove that there exists $n \in \mathbb{N}^*$ and a finite sequence $(t_i)_{0 \leqslant i \leqslant n}$ on $[a, b]$ such that for any survival function $S' \in \mathcal{S}$, if S' coincides with S on all t_i, then

$$\int_a^b |S'(t) - S(t)| \, dt < \varepsilon \, . \tag{2}$$

Let N be an integer such that $(2 + b - a)/N < \varepsilon$. We build the sequence $(t_i)_i$ by induction, and a subset J of $\mathbb{N}$: one first sets $t_0 = b$ and $J = \emptyset$, then:

- if $S(t_i^-) \geqslant S(t_i) + 1/N$, then we add item i to J, and set $t_{i+1} = \max(t_i - 1/N^2, a)$,
- else, we set $t_{i+1} = \max(\inf\{t, \ S(t) \leqslant S(t_i) + 1/N\}, a)$, so that (if $t_{i+1} > a$) $S(t_{i+1}) \leqslant S(t_i) + 1/N$ and $S(t_{i+1}^-) \geqslant S(t_i) + 1/N$.

We stop this induction as soon as a t_i reaches a, and we denote by n this final subscript. This way, the sequence $(t_i)_i$ is strictly decreasing, and for all $i < n - 2$, $S(t_{i+2}) \geqslant S(t_i) + 1/N$. This sequence is thus finite, with length at most $2N$.

Let $S' \in \mathcal{S}$ be such that, for all $0 \leqslant i \leqslant n$, $S'(t_i) = S(t_i)$. Since S' and S are decreasing,

$$\begin{aligned}\int_a^b |S'(t) - S(t)|\, dt &= \sum_{i=1}^{n} \int_{t_i}^{t_{i-1}} |S'(t) - S(t)|\, dt \leqslant \sum_{i=1}^{n} \int_{t_i}^{t_{i-1}} (S(t_i) - S(t_{i-1}))\, dt \\ &\leqslant \sum_{i \in J} \frac{1}{N^2} + \sum_{i \notin J} \frac{t_{i-1} - t_i}{N} \leqslant \frac{2}{N} + \frac{b-a}{N} < \varepsilon\,.\end{aligned}$$

To end the proof, the key point is that any piecewise affine function from $\mathbb{R}$ to $\mathbb{R}$ with finite number of apices can be seen as a composition of angle functions. To approach S, S_0 is then distorted so as to coincide with S at the points $x_1, \ldots, x_n$ from (2). □

4.3 Initialization values

We suggested using some particular conversion functions: angle compositions or smoothed versions of them like hyperbolas compositions. When estimating parameters of these compositions, it is necessary to start from a good initial value. It is possible to proceed in several ways:

- The initial parameter vector may correspond to an identity conversion function if the initial survival function is not too far from the target one. Nevertheless, this choice may lead many optimization algorithms to a solution far from the optimal solution.
- When composing multiple functions, it might be easier to estimate separately optimal parameters for each distortion, and this may lead to initial values for aggregate parameters of the composite function. This choice however has to cope with the case where two antagonistic effects compensate each other, as an example when a first conversion function creates a distance from the target, in order to ease the adjustment of a second conversion function.
- The simultaneous adjustment of all parameters of the composite function is the solution likely to give best results if initial parameters value is not too far from the optimal vector. We try here to suggest initial values which lead to a correct approximation of the target function.

In this paragraph, one consider conversion functions that are all hyperbolic conversion functions. The smoothing parameter does not seem to be the hardest to estimate, so we focus on the estimation of angle composite functions, with angles defined by 2 or 4 parameters.

We suppose that we start from a finite set of abscissas $\{x_i\}_i \in I$, with $I = \{1, \ldots, p\}$, for which are given the target survival function and its logit $l_i = \operatorname{logit} S(x_i)$, as well as the logit of the current survival function to be distorted $\alpha_i = \operatorname{logit} \widehat{S}(x_i)$. This scatter plot is a finite part of a curve that we write $l(\alpha)$. We are looking for f_θ so that points $\{(l_i, f_\theta(\alpha_i))\}_{i \in I}$ are as close as possible to the first diagonal Δ, defined by equation $y = x$. One possibility is to look for a function $\widehat{f_\theta}$ which could be able to associate l_i to some of the α_i, for $i \in I_k$, $I_k \subset I$. For those points, we could ensure $\{(l_i, f_\theta(\alpha_i))\}_{i \in I_k}$ are in Δ. Is it possible to find an angle composite function that reach all points $\{(l_i, \alpha_i)\}_{i \in I_k}$, $I_k \subset I$? It is relatively simple, thanks to the following Proposition.

Proposition 6 *Consider a set of successive points $\{(u_i, v_i)\}_{i \in \{1,..,3+k\}}$ of an increasing curve, with $u_1 \leqslant \ldots \leqslant u_{3+k}$ and $v_1 \leqslant \ldots \leqslant v_{3+k}$, $k \geqslant 0$. The angle composite functions*

$$G_\theta^{(0)} = A_{x_0,y_0,p_1,p_2} \qquad \textit{and} \qquad G_\theta^{(k)} = A_{a_k,a_k,1,q_k} \circ \cdots \circ A_{a_1,a_1,1,q_1} \circ A_{x_0,y_0,p_1,p_2} \quad \forall k \geqslant 1\,,$$

are such that: $G_\theta^{(k)}(u_i) = v_i$, for all $i \in \{1, \ldots, 3+k\}$, setting $x_0 = u_2$, $y_0 = v_2$, $p_1 = (v_2 - v_1)/(u_2 - u_1)$, $p_2 = (v_3 - v_2)/(u_3 - u_2)$,

$$a_k = v_{2+k} \qquad \textit{and} \qquad q_k = \left(\frac{v_{3+k} - v_{2+k}}{u_{3+k} - u_{2+k}}\right)\left(\frac{u_{2+k} - u_{1+k}}{v_{2+k} - v_{1+k}}\right), \quad k \geqslant 1\,.$$

Proof : One easily checks that $A_{x_0,y_0,p_1,p_2}(u_1) = v_1$, $A_{x_0,y_0,p_1,p_2}(u_2) = v_2$ and $A_{x_0,y_0,p_1,p_2}(u_3) = v_3$. One then checks by induction that for all $i \leqslant k$, $G^{(k)}(u_i) = v_i$. □

Consequently, we use the following process for the initialization of the parameters vector: a subset (u_i, v_i) is extracted from the set (l_i, α_i), taking care to choose points as far as possible from each

Sym. bump	Asym. bump	H_5	$H_2 \circ H_5$	$H_2 \circ H_2 \circ H_5$
$m = -0.13168$	$m = -1.56494$	$x_0 = 0.96975$	$x_0 = 1.10085$	$x_0 = 0.04447$
$h = 1.46117$	$h = -2.67358$	$y_0 = -0.47722$	$y_0 = -0.41017$	$y_0 = -1.01970$
$\rho = -1.60908$	$\rho = 0.86007$	$p_1 = 0.76112$	$p_1 = 0.74577$	$p_1 = 0.90951$
$\epsilon = 0.40543$	$\gamma = -1.18722$	$p_2 = 3.43476$	$p_2 = 5.44530$	$p_2 = 0.00408$
	$\epsilon = 0.26891$	$\epsilon = 0.02593$	$a_1 = 1.43665$	$a_1 = -1.02535$
			$q_1 = 0.17370$	$q_1 = 1882.67788$
			$\epsilon = 0.20335$	$a_2 = 1.41794$
				$q_2 = 0.06723$
				$\epsilon = 0.49316$
$I_Q = 1.68$	$I_Q = 2.25$	$I_Q = 1.97$	$I_Q = 2.22$	$I_Q = 2.66$

Table 1: Adjustment of survival function for death year 1915 with distortions of the survival function for death year 1913.

other. We take for example the minimum abscissa point, the maximum abscissa point, and $k+1$ intermediate points regularly spaced (that is $k+2$ intermediate intervals):

$$(u_j, v_j) = (\alpha_{s(j)}, l_{s(j)}) \quad \text{with} \quad s(j) = \inf_{i \leqslant p} \left\{ L(i) \geqslant \frac{L(n)}{k+2}(j-1) \right\}, \quad j = 1, \dots, k+3 .$$

The function L can be chosen as $L(i) = |l_i - l_1|$, or any other increasing function. By Proposition 6, we deduce initialization values for all previously presented compositions $A_2 \circ \dots A_2 \circ A_4$. By choosing for the smooth parameter ϵ a small positive value, for example starting from $\epsilon = 1$ in our applications, we obtain initialization values for conversion functions of kind $H_2 \circ \dots H_2 \circ H_4$. The choice of an initialization value $\epsilon \neq 0$ can be explained by continuous differentiability conditions, which ease the convergence of main optimization algorithms.

5 Numerical applications

We have chosen here to present applications to survival data analysis. We use in this section some mortality tables from Internet website Human Mortality Database (2008). These tables are given by death year, for the United States and France (tables for an age bracket of one year, and a death year bracket of one year, denoted 1×1, for the whole population, men and women). We call these two tables respectively USA and France.

In the discrete case, it is necessary to provide a distance measure between the target function and its adjustment, and this distance should be adapted to the discrete character of the problem. We have here a set of points which correspond to the values of the survival function at step n, $s_i^n = S_n(x_i)$ and the set of values of the target survival function $s_i = S(x_i)$, for different points x_i, $i \in \{1, \dots, p\}$. We measure the quality of the adjustment at step n with the following quality index $I_Q^n = -\log_{10}(p^{-1} \sum_{i=1}^{p} |s_i^n - s_i|)$.

5.1 Catastrophic event modelling

For this application, we distort the French table for the death year 1913, in order to get closer to the French table for the death year 1915, in the very first year of the First World War. The distortion we get aims at showing the ability of conversion functions to model a catastrophic change on a given table, even when this one affects differently the survival probabilities at different ages, and concerns in particular young adults. Survival functions correspond to a product of annual survival probabilities, as if mortality was always in accordance with that of the considered year, 1913 or 1915. The applied model is:

$$S^{1915} = \mathbb{T}_f S^{1913} .$$

With 4 parameters, we found a quality index close to 2 with a symmetric bump, or a little bit better with more parameters. We suggest here adjustments which minimize the distance between the target and the adjusted distribution function. This choice could lead to differences, for example, in annual death probabilities. Depending on the further use of the distortion, it may be preferable to choose other optimization criteria. We do not here develop such other criteria.

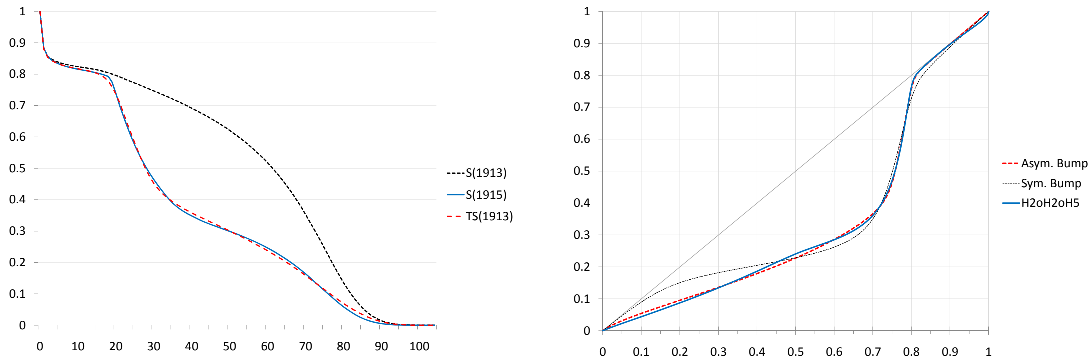

Figure 4: Adjustment of the 1915 survival function by distortion of 1913 survival function with an asymmetrical bump conversion function ($\mathbb{T}_f S^{1913}$ dashed, $I_Q \simeq 2.25$), left, and corresponding distortions T_f using a symmetric bump (thin dotted line), an asymmetric bump (bold dashed line) and a composed hyperbolic distortion, right.

We have used an age bracket from 0 to 104 years old (taking a wider bracket would artificially increase the quality index), and results are given in the table 1. For other adjustments with less than five parameters, it seems here better to use bump functions, which benefit from the fact that the table is not deeply modified for higher ages. This last bump function is illustrated in figure 4. This quality index with this last 5 parameters function is better than the one obtained with 7 parameters hyperbolic composed distortions. Thanks to the knowledge of particular properties of the conversion function, we were able to benefit from two parameters less. To improve precision even more, we need to use composed hyperbolic distortions.

The Figure 4 shows starting and ending curves, and distortion functions are shown in Figure 4 (right). As it appears in Table 1, the composed hyperbolic distortion $H_2 \circ H_2 \circ H_5$ is the better distortion to apply. One can see in Figure 4 (right), the symmetry constrainst which gives its special shape to the symmetric bump, and the need for an asymmetry coefficient which makes the asymmetric bump close to the $H_2 \circ H_2 \circ H_5$ distortion.

5.2 Prospective model, dynamic distortion

A second model consists in representing each table with a hyperbolic or composed hyperbolic distortion, where all parameters are evolving with time. The model is the following one:

$$S(x,t) = T_{f_{\theta(t)}}(S^0(x)) \,.$$

Where $\theta(t)$ is depending on t, and S^0 is the survival function of an exponential law with parameter 1. In order to get a quality index close to 3, we use f_{θ_t} of kind $H_2 \circ H_5$ for conversion function. Indeed, this function gave good results over one unique year.

$$f_{\theta(t)} = H_{a_1^t, a_1^t, 1, q_1^t, \epsilon^t} \circ H_{x_0^t, y_0^t, p_1^t, p_2^t, \epsilon^t} \,.$$

For the evolution of the parameter $\theta(t)$, we make the simple choice of a linear evolution. One may notice that, due to this choice, results could depend on the chosen parameterization. In particular, a linear evolution on a slope p would not have the same effect than a linear evolution on the logarithm ρ of this slope.

Parameters are given with respect to a reference year. For tables available on death year 1975-2000 , we take the middle of the bracket as reference year, that is $t_0 = 1990$. The chosen age bracket includes all available ages, from 0 to 110 years old. The f_{θ_t} function's parameters are supposed to evolve linearly with considered death year:

$$\theta_t = (x_0^t, y_0^t, p_1^t, p_2^t, \epsilon^t, a_1^t, q_1^t) = (x_0^0 + t\delta_{x_0}, y_0^0 + t\delta_{y_0}, p_1^0 + t\delta_{p_1}, p_2^0 + t\delta_{p_2}, \epsilon^0 + t\delta_\epsilon, a_1^0 + t\delta_{a_1}, q_1^0 + t\delta_{q_1}) \,,$$

t represents the difference between the considered death year and the reference year t_0.

	France 1975-2005		USA 1975-2005	
	value at t_0	annual variation $\delta_{\dots}$	value at t_0	annual variation $\delta_{\dots}$
x_0	-41.538638	-0.842434	-93.314998	-0.286203
y_0	28.769735	0.221127	0.278864	0.057883
p_1	0.520306	0.007222	0.448850	0.009311
p_2	0.356317	0.017845	0.111053	0.000217
ϵ	5.035559	0.067616	3.714785	0.118081
a_1	0.528556	0.004680	4.684051	0.077014
q_1	0.119153	-0.001669	0.145247	-0.000537
I_Q average	2.72		2.70	
I_Q min. (year)	2.42 (2004)		2.47 (1975)	
I_Q max. (year)	3.07 (1995)		3.04 (2001)	

Table 2: Simultaneous adjustments of survival functions, for death years 1975-2005, by dynamic distortions of an exponential law of parameter 1, with linear evolution of distortions parameters.

The results of these dynamic distortions are given in the table 2. One might be afraid of a greater instability of the quality index, but here no death year leads to an index less than 2.4: even for the worst adjustment, the curves of distorted and target survival functions are almost identical. Most adjustments have a quality index greater than 2.6, which represent an error of order $2 \cdot 10^{-3}$ on survival functions.

Lastly, considering narrower age brackets (here from 0 to 110 years old), narrower death year bracket (here from year 1975 to 2005), would lead, as previously, to an appreciable improvement in the quality index. The parameters number is here equal to 14, and some parameters have a reduced utility. This number is not so big when looking to the quantity of data. Some parameters remain very stable with death year. The study of which parameters we shall keep is not developed at this time, since the aim of this paragraph is just to show the faculty of some conversion functions to adapt themselves in a prospective framework. We could have suggested another model, by representing all tables in the 1975-2005 period by a distortion of the table of the reference year, e.g. 1990:

$$S(x,t) = T_{f_{\theta(t)}}(S(x,1990))\,.$$

It seems far easier to adapt one mortality table from year 1990 rather than an exponential law of parameter 1, deeply unsuited to human mortality. To give an illustration, the quality index we get by adjusting the 2005 death year table with the 1990 one, for a hyperbolic conversion function like $H_2 \circ H_5$, is 3.13 for French tables, and 3.32 for American tables. The improvement of the quality index is close to 0.1 or 0.2 only, compared to the distortion of an exponential law. We decided here to keep a continuous parametric expression, with an easy analytic inverse function.

5.3 Stochastic simulations

Let us suppose first that one can easily compute the inverse function of the initial survival function S of X. The invertibility constraint on distortions allows easy simulation (by inversion method) of the law of the distorted variable $\widehat{X}$: if U is a random variable with a uniform law on $]0,1[$,

$$\widehat{X} \stackrel{\mathcal{L}}{=} S^{-1}(T_{f^{-1}}(U))\,.$$

When the conversion function depends on an explanatory parameter vector or does evolve with the passage of time (as in the section 5.2), this method allows simulation of residual lifetimes in accordance with a mortality table depending on one or several parameters, like a prospective mortality table. As an example, let us denote by f_t a conversion function to be applied to a survival reference function S to model the law of the residual lifetime of someone born at a time t. For someone from this birth date aged u, a random sample of a survival residual lifetime X_u^t can be obtained from a random variable V generated from a uniform distribution between 0 and 1, by:

$$\widehat{X}_u^t \stackrel{\mathcal{L}}{=} S^{-1}\left[T_{f_t^{-1}}\Big(V \cdot T_{f_t}(S(u))\Big)\right] - u\,.$$

6 Conclusion

Starting from a given initial survival function, iterative distortions allow to converge to any target survival function. This is of great importance when looking for a parametric representation of the distribution of a random variable, especially when distortions parameters change over time. We proposed readily invertible distortions that help simulation of the distorted distributions. Finding the best number of parameters, or investigating risk-measure properties of the distorted random variable are natural perspectives of this work.

Aknowledgements

We wish to thank the anonymous referee and the editor whose comments helped us in improving the content and readibility of our paper.

References

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1997), *Thinking Coherently.* Risk 10, 68-71.

Bagdonavicius, V., Nikulin, M. (2002), *Accelerated Life Models.* Monographs on Statistics and Applied Probability 94.

Bernoulli, D. (1731), *Specimen theoriae novae de mensura sortis.* English trad. Exposition of a new theory on the measurement of risk, Econometrica, XXI, pp. 223 sqq., 1954.

Bleichrodt H., Eeckoudt, L. (2006), *Survival risks, intertemporal consumption, and insurance: The case of distorded probabilities.* Insurance: Mathematics and Economics, 38, 335-346.

Brass, W. (1969), *A generation method for projecting death rates.* Population growth and the brain dead. Bechhofer F. editor, Birmingham, Edinburgh University Press: 75-91.

Brass, W. (1974), *Mortality models and their uses in demography.* Trans. Faculty Actuaries, 33. 122-133.

Bühlman, H., (1980) *An economic premium principle.* ASTIN Bulletin 11, 52-60.

d'Alembert (1768), *Vingt-troisième mémoire. V. Sur le calcul des probabilités,* Opuscules Mathématiques, vol.IV, pp.74-79, Paris, David.

De Jong, P., Marshall, C., (2007), *Mortality projection based on the Wang transform,* ASTIN Bulletin, 37 (1), 149-161.

Feller, W. (1968) *An Introduction to Probability Theory and Its Applications,* John-Wiley & Sons.

Goovaerts, M.J., Kaas, R., Dhaene, J., Tang, Q., (2004) *Some new classes of consistent risk measures,* Insurance: Mathematics and Economics, 34, 505-516.

Hamada, M., Sherris, M. (2003) *Contingent claim pricing using probability distorsion operators: methods from insurance risk pricing and their relationship to financial theory,* Applied Mathematical Finance, 10, 19-47

Heligman, L, Pollard, J.H. (1980) *The age pattern of mortality.* J. Inst. of Actuaries, 107, 49-80.

Human Mortality Database, University of California, Berkeley (USA), and Max Planck Institute for Demographic Research, Rostock (Germany). http://www.mortality.org.

Landsman, Z., Sherris, M., (2001) *Risk measures and insurance premium principles,* Insurance: Mathematics and Economics, 29 , 103-115.

Lee, R.D., Carter, L.W. (1992) *Modelling and forecasting U.S. mortality (with discussion).* Journal of the American Statistical Association, 87(419), 659-675.

Pelsser, A. (2007), *On the applicability of the Wang Transform for Pricing Financial Risks,* University of Amsterdam and NETSPAR, working paper.

Pitacco, E. (2004), *Survival models in a dynamic context: a survey.* Insurance: Mathematics and Economics, 35, 2, 279-298.

Pradier, P.-C. (1998), *Concepts et mesures du risque en théorie économique - essai historique et critique,* Thèse ENS-Cachan.

Wang, S.S. (1996), *Premium calculation by transforming the layer premium density,* ASTIN Bulletin 26 (1), 71-92.

Wang, S.S. (2000), *A class of distorsion operators for pricing financial and insurance risks,* The Journal of Risk and Insurance 67(2), 15-36.

Wirch, J.L., Hardy, M.R. (1999), *A synthesis of risk measures for capital adequacy,* Insurance: Mathematics and Economics, 25, 337-347.

Yaari, M.E. (1987), The dual theory of choice under risk. Econometrica 55 (1), 95-115.