FROM REGULATORY LIFE TABLES TO STOCHASTIC MORTALITY PROJECTIONS: THE EXPONENTIAL DECLINE MODEL

Michel Denuit
Institute of Statistics, Biostatistics and Actuarial Science
Université Catholique de Louvain (UCL)
Louvain-la-Neuve, Belgium

Julien Trufin
Department of Mathematics
Université Libre de Bruxelles (ULB)
Bruxelles, Belgium

October 12, 2016

Abstract

Often in actuarial practice, mortality projections are obtained by letting age-specific death rates decline exponentially at their own rate. Many life tables used for annuity pricing are built in this way. The present paper adopts this point of view and proposes a simple and powerful mortality projection model in line with this elementary approach, based on the recently studied mortality improvement rates. Two main applications are considered. First, as most reference life tables produced by regulators are deterministic by nature, they can be made stochastic by superposing random departures from the assumed age-specific trend, with a volatility calibrated on market or portfolio data. This allows the actuary to account for the systematic longevity risk in solvency calculations. Second, the model can be fitted to historical data and used to produce longevity forecasts. A number of conservative and tractable approximations are derived to provide the actuary with reasonably accurate approximations for various relevant quantities, available at limited computational cost. Besides applications to stochastic mortality projection models, we also derive useful properties involving supermodular, directionally convex and stop-loss orders.

Keywords and phrases: life tables, risk measures, longevity risk, comonotonicity, life annuity, supermodular order, directionally convex order, increasing convex order.

1 Introduction and motivation

Many reference life tables used by actuaries for annuity pricing or reserving are based on the assumption that each age-specific one-year death probability or death rate declines at its own rate. Formally, denoting as $q_{x,t}$ the one-year death probability at integer age x during calendar year t, this means that

$$\frac{q_{x,t}}{q_{x,t-1}} = \rho_x \text{ for some } 0 < \rho_x < 1. \tag{1.1}$$

Mortality projections are then obtained starting from a reference life table (or base table) corresponding to calendar year t_0, say, and applying yearly age-specific mortality improvement rates ρ_x. Precisely, the one-year death probability at age x during calendar year t_0+k is obtained from

$$q_{x,t_0+k} = q_{x,t_0}\rho_x^k, \quad k = 1, 2, \ldots. \tag{1.2}$$

The same approach can be adopted for death rates $m_{x,t}$, turning (1.1)-(1.2) into:

$$\frac{m_{x,t}}{m_{x,t-1}} = \rho_x \text{ and } m_{x,t_0+k} = m_{x,t_0}\rho_x^k, \quad k = 1, 2, \ldots. \tag{1.3}$$

The German DAV "R" life tables used for annuity pricing are obtained in this way. See, e.g. Kruger and Pasdika (2006). Also in Austria and Switzerland, mortality projections for the annuity business are obtained from a base life annuity table with age-specific functions relating to future mortality improvements. In the UK, the CMI released a series of projected life tables, all built from age-specific improvement rates applied to a reference life table. In Belgium, the insurance law refers to the mortality projections produced by the Federal Planning Bureau, which are obtained in the very same way. Whereas most reference projected life tables are displayed in terms of $q_{x,t}$, the Danish FSA benchmark consists in age-specific yearly improvement rates applied to a set of death rates m_{x,t_0} corresponding to some base year t_0. We refer to Jarner and Moller (2015) for more details. Therefore, the specifications (1.1)-(1.2)-(1.3) cover a wide variety of situations often encountered in actuarial practice.

However, all these life tables are deterministic by nature, and so disregard the systematic longevity risk induced by the random departures from the reference forecast. It is well documented that ignoring the uncertainty in future mortality leads to underestimations of the amount of risk capital. We refer the reader to Denuit and Frostig (2007) for a detailed analysis in the Lee-Carter framework. This analysis extends to other one-factor mortality projection models, provided conditional survival probabilities are monotonic functions of the single time index. For proper risk evaluation, the exponential decline model proposed in this paper retains the reference forecast but incorporates uncertainty around the decreasing trends.

The remainder of this paper is organized as follows. Section 2 presents an exponential decline model allowing departures from the central scenario in order to account for the systematic longevity risk. In Section 3, we derive in our new model comonotonic approximations for the conditional survival probabilities and for the conditional expected present value of annuity payments. Also, we investigate conservative approximations for the consecutive numbers of survivors in a portfolio of life annuities. Next, Section 4 illustrates how to turn a regulatory life table into a stochastic mortality projection model. Specifically, we consider Belgian mortality statistics for male together with the reference projected life table published by the Federal Planning Bureau. In this setting, we also perform numerical illustrations assessing the accurateness of the approximations derived in the previous section. In Section 5, we show how to calibrate the exponential decline model on historical mortality data, without reference to a specific set of mortality improvement factors. To this end, we use a mixed Poisson likelihood. Finally, Section 6 concludes the present study.

2 Exponential decline model

Starting from (1.1) or (1.3), we now allow for stochastic departures from the central scenario. Formally, we assume that the exponential decline only holds on average and we replace (1.1) and (1.3) respectively with

$$\frac{q_{x,t}}{q_{x,t-1}} = \rho_x \Lambda_t \text{ and } \frac{m_{x,t}}{m_{x,t-1}} = \rho_x \Lambda_t \tag{2.1}$$

where the positive random variables Λ_t may be serially correlated, with unit mean. Hence, each death rate decreases at its own rate ρ_x subject to random shocks Λ_t. As the same Λ_t impacts the whole portfolio, these random factors account for the systematic longevity risk. The insured lifetimes are no more independent but only conditionally independent, given the future Λ_t.

Notice that the specification (2.1) differs from the model proposed in Schinzinger et al. (2014) who allowed for age-specific departures from the common trend. As we aim to recognize the systematic risk superposed to a given, deterministic mortality projection, such departures are replaced here with a single, age-independent sequence of Λ_t. Applying the same shock Λ_t to all ages results in comonotonic mortality improvement rates. For some specific applications, it may appear useful to relax this strong assumption. Consider for instance the case where the actuary has to assess the potential benefits obtained from natural hedging. It may then be preferable to allow for correlated (but not necessarily perfectly correlated, or comonotonic) shocks applying to different age classes, typically the pre-retirement ages where most of the insurance business including death benefits is concentrated versus the post-retirement ages where the longevity risk mainly remains. This is because comonotonic mortality improvement rates across all ages tend to overstate the benefits obtained from natural hedging mechanisms.

Suppose that we are now at time t_0. Let $T_1, \ldots, T_n$ be the remaining lifetimes for n policyholders aged x at time t, $t > t_0$. For simplicity, x and t are assumed to be integer. The collection of yearly shocks impacting the central scenario, i.e. the base table for calendar year t_0 and the mortality improvement rates ρ_x, is denoted as

$$\boldsymbol{\Lambda} = \Big(\Lambda_{t_0+1}, \Lambda_{t_0+2}, \ldots\Big).$$

Given $\boldsymbol{\Lambda}$, $T_1, \ldots, T_n$ are assumed to be independent and identically distributed. Henceforth, we denote the common k-year conditional survival probability as

$$\mathrm{P}[T_i > k|\boldsymbol{\Lambda}] = {}_kP_x(t|\boldsymbol{\Lambda}), \quad k \in \{0, 1, 2, \ldots\}.$$

Then, for any $k_1, \ldots, k_n \in \{0, 1, 2, \ldots\}$,

$$\begin{aligned} \mathrm{P}[T_1 > k_1, \ldots, T_n > k_n] &= \mathrm{E}\left[\prod_{i=1}^{n} \mathrm{P}[T_i > k_i | \boldsymbol{\Lambda}]\right] \\ &= \mathrm{E}\left[\prod_{i=1}^{n} {}_{k_i}P_x(t|\boldsymbol{\Lambda})\right]. \end{aligned}$$

In our setting, we have from (2.1) that

$$ {}_kP_x(t|\boldsymbol{\Lambda}) = \prod_{j=0}^{k-1} \left(1 - q_{x+j,t_0} \rho_{x+j}^{t-t_0+j} \prod_{l=1}^{t-t_0+j} \Lambda_{t_0+l}\right), \quad k \in \{0, 1, 2, \ldots\} \tag{2.2}$$

when mortality improvements are applied to one-year death probabilities, and

$$ {}_kP_x(t|\boldsymbol{\Lambda}) = \prod_{j=0}^{k-1} \exp\left(-\mu_{x+j,t_0} \rho_{x+j}^{t-t_0+j} \prod_{l=1}^{t-t_0+j} \Lambda_{t_0+l}\right), \quad k \in \{0, 1, 2, \ldots\} \tag{2.3}$$

when mortality improvements are applied to death rates. We see that in both cases, $\boldsymbol{\lambda} \mapsto {}_kP_x(t|\boldsymbol{\lambda})$ is decreasing, where we denote by $\boldsymbol{\lambda}$ the realizations of $\boldsymbol{\Lambda}$. Notice that whatever the realizations of the Λ_{t_0+l}, the specification (2.3) always produces admissible probability values (i.e. values comprised between 0 and 1) whereas it may be necessary to bound (2.2) to avoid realizations exceeding 1. As placing such additional constraints does not modify the results obtained in the next sections, we do not impose them explicitly in the present paper.

3 Comonotonic approximations

3.1 Risk measures

Let us recall the definition of the risk measures considered in this paper. For more details, we refer the reader to Denuit et al. (2005). Given a risk X, i.e. a random amount of benefit to be paid in execution to one or several insurance contracts, with distribution function F_X, and a probability level $p \in (0, 1)$, the corresponding Value-at-Risk, abbreviated VaR, is defined as

$$\mathrm{VaR}[X; p] = F_X^{-1}(p) = \inf\{x \in \mathbb{R} | F_X(x) \geq p\}.$$

The corresponding Tail-VaR is then defined as

$$\mathrm{TVaR}[X; p] = \frac{1}{1-p} \int_p^1 \mathrm{VaR}[X; \xi] \, d\xi.$$

The approximations derived in this section are conservative in that they overestimate the true Tail-VaR. We know from Section 3.4 in Denuit et al. (2005) that given two risks X and

Y with finite means, the following equivalences hold true:

$$
\begin{aligned}
& \mathrm{TVaR}[X;p] \leq \mathrm{TVaR}[Y;p] \text{ for all } 0 \leq p \leq 1 \\
\Leftrightarrow\ & \mathrm{E}[(X-t)_+] \leq \mathrm{E}[(Y-t)_+] \text{ for all } t \\
\Leftrightarrow\ & \mathrm{E}[g(X)] \leq \mathrm{E}[g(Y)] \text{ for all non-decreasing and convex functions } g.
\end{aligned}
$$

The risk X is then said to precede Y in the stop-loss or increasing convex order, denoted as $X \preceq_{\mathrm{SL}} Y$.

3.2 Perfectly dependent departures from the common trend

Actuaries face computational problems with most stochastic mortality projection models. Analytical calculations are generally impossible under (2.2)-(2.3) so that numerical procedures are needed. In case Monte Carlo simulations are used, the actuary must first simulate the future path $\boldsymbol{\lambda}$ of the random departures $\boldsymbol{\Lambda}$ and then, given the resulting life table ${}_kP_x(t|\boldsymbol{\lambda})$, realizations of the insured lifetimes $T_1, \ldots, T_n$. This nested simulation procedure may require considerable computational power so that tractable, conservative approximations appear to be useful for practice (not necessarily to replace more accurate computations but to get at least reasonably accurate orders of magnitude in a very fast way).

Effective approximations have been designed under the standard Lee-Carter model, after Denuit and Dhaene (2007) who considered conditional probabilities and conditional expectations (life expectancies or expected present value) given the unknown life table applying in the future. See also Denuit (2007), Denuit et al. (2010, 2013), and Gbari and Denuit (2014). Considering these works, we can see that the stated results are not specific to the Lee-Carter model but generally hold true as soon as the conditional survival probabilities are monotonic functions of the time index. This applies for instance to the new approach proposed by Cadena and Denuit (2015). In this section, we derive similar approximations for our specification (2.2)-(2.3).

The approximations derived in this paper are based on the following simple and intuitive idea. Even if the Λ_t are mutually independent, their products entering formulas (2.2)-(2.3) are obviously positively correlated. It turns out that making these products perfectly dependent provides the actuary with conservative approximations that are reasonably accurate.

Recall that the d-dimensional random vector $\boldsymbol{X}^+$ is comonotonic if there exist a random variable Z and non-decreasing functions h_i such that $\boldsymbol{X}^+$ is distributed as $(h_1(Z), h_2(Z), \ldots, h_d(Z))$. Now, define $\boldsymbol{\Pi}^+$ as the comonotonic version of the random vector $\boldsymbol{\Pi}$ gathering the products of Λ_l involved in (2.2)-(2.3). Formally, we define

$$
\Pi_j = \prod_{l=1}^{t-t_0+j} \Lambda_{t_0+l} \text{ for } j = 0, 1, 2, \ldots
$$

and, given a random variable Z uniformly distributed over the unit interval $[0,1]$,

$$
\Pi_j^+ = F_{\Pi_j}^{-1}(Z) \text{ for } j = 0, 1, 2, \ldots.
$$

Clearly, the random variables Π_j^+ are comonotonic as they are all obtained as increasing transformations of the same underlying random variable Z.

The two following properties of comonotonic random vectors are central to the present work. For a proof, we refer the reader to Denuit et al. (2005). Recall that the function $g : \mathbb{R}^2 \to \mathbb{R}$ is supermodular if

$$g(b_1, b_2) - g(a_1, b_2) - g(b_1, a_2) + g(a_1, a_2) \geq 0$$

for all $a_1 \leq b_1, a_2 \leq b_2$. Thus, such functions put more weight on (b_1, b_2) and (a_1, a_2) expressing positive dependence because both components are simultaneously large or small, than on (a_1, b_2) and (b_1, a_2) expressing negative dependence because they both mix one large component with one small component. This intuitively explains why supermodularity plays an important role in many problems involving positively dependent random variables. If g is differentiable then

$$g \text{ supermodular } \Leftrightarrow \frac{\partial^2}{\partial x_1 \partial x_2} g \geq 0.$$

The function $g : \mathbb{R}^d \to \mathbb{R}$ is supermodular if it is supermodular viewed as a function of (x_i, x_j) with fixed x_k, $k \neq i, j$, for any $i \neq j \in \{1, \ldots, d\}$. A twice differentiable function g is supermodular if

$$\frac{\partial^2}{\partial x_i \partial x_j} g \geq 0 \text{ for all } i \neq j \in \{1, \ldots, d\}.$$

The next result recalls the two fundamental properties of comonotonic random variables that will be used throughout this paper. For more details, we refer the interested reader to Property 2.3.3 and Proposition 6.3.7 in Denuit et al. (2005).

Property 3.1. *Let $F_1, F_2, \ldots, F_n$ be a collection of n distribution functions for random variables $X_1, X_2, \ldots, X_n$. Let $\boldsymbol{X}^+$ be a comonotonic random vector such that $X_i^+ = F_i^{-1}(Z)$, with Z uniformly distributed on the unit interval. Then,*

(i) for every probability level p,

$$VaR\left[\sum_{i=1}^n X_i^+; p\right] = \sum_{i=1}^n VaR\left[X_i^+; p\right] = \sum_{i=1}^n F_i^{-1}(p).$$

(ii) for every supermodular function g,

$$E[g(X_1, X_2, \ldots, X_n)] \leq E[g(X_1^+, X_2^+, \ldots, X_n^+)]$$

provided the expectations exist.

3.3 Conditional survival probabilities

Henceforth, considering (2.2), we denote ${}_kP_x(t|\boldsymbol{\Lambda})$ as ${}_kP_x(t|\boldsymbol{\Pi})$ where

$${}_kP_x(t|\boldsymbol{\Pi}) = \prod_{j=0}^{k-1} P_{x+j}(t+j|\Pi_j) \text{ with } P_{x+j}(t+j|\Pi_j) = 1 - q_{x+j,t_0} \rho_{x+j}^{t-t_0+j} \Pi_j.$$

Similarly, we define

$$ {}_kP_x(t|\mathbf{\Pi}^+) = \prod_{j=0}^{k-1} P_{x+j}(t+j|\Pi_j^+) \text{ with } P_{x+j}(t+j|\Pi_j^+) = 1 - q_{x+j,t_0}\rho_{x+j}^{t-t_0+j}\Pi_j^+. $$

Turning to (2.3), we then have

$$ {}_kP_x(t|\mathbf{\Pi}) = \prod_{j=0}^{k-1} P_{x+j}(t+j|\Pi_j) \text{ with } P_{x+j}(t+j|\Pi_j) = \exp\big(-\mu_{x+j,t_0}\rho_{x+j}^{t-t_0+j}\Pi_j\big) $$

and

$$ {}_kP_x(t|\mathbf{\Pi}^+) = \prod_{j=0}^{k-1} P_{x+j}(t+j|\Pi_j^+) \text{ with } P_{x+j}(t+j|\Pi_j^+) = \exp\big(-\mu_{x+j,t_0}\rho_{x+j}^{t-t_0+j}\Pi_j^+\big). $$

Let us now establish that the approximations proposed to the conditional survival probabilities are conservative. The next result shows that the supermodular order can be characterized by means of the stop-loss order for any increasing supermodular transform of the components of the random vectors. It complements the existing characterization of the supermodular order and appears to be of independent interest.

Property 3.2. *Given two random vectors* $(X_1, X_2, \ldots, X_n)$ *and* $(Y_1, Y_2, \ldots, Y_n)$, $(X_1, X_2, \ldots, X_n)$ *precedes* $(Y_1, Y_2, \ldots, Y_n)$ *in the supermodular order, i.e. the inequality*

$$ E[g(X_1, X_2, \ldots, X_n)] \leq E[g(Y_1, Y_2, \ldots, Y_n)] $$

holds for every supermodular function g *for which the expectations exist, if, and only if,* $X_i =_d Y_i$ *for* $i = 1, 2, \ldots, n$ *and*

$$ TVaR[\Psi(X_1, X_2, \ldots, X_n); p] \leq TVaR[\Psi(Y_1, Y_2, \ldots, Y_n); p] \text{ for all } p $$

for every non-decreasing and supermodular function Ψ.

Proof. Considering the "$\Leftarrow$" part, $\mathrm{E}[\Psi(X_1, X_2, \ldots, X_n)] \leq \mathrm{E}[\Psi(Y_1, Y_2, \ldots, Y_n)]$ obviously holds for every non-decreasing and supermodular function Ψ since $\mathrm{TVaR}[\Psi(X_1, X_2, \ldots, X_n); 0] = \mathrm{E}[\Psi(X_1, X_2, \ldots, X_n)]$ and $\mathrm{TVaR}[\Psi(Y_1, Y_2, \ldots, Y_n); 0] = \mathrm{E}[\Psi(Y_1, Y_2, \ldots, Y_n)]$. Theorem 3.4 in Muller and Scarsini (2000) then allows us to conclude as the increasing supermodular order between random vectors sharing the same univariate marginals is equivalent to the supermodular order.

Turning to the "$\Rightarrow$" part, let us show that the composition $h \circ \Psi$ is supermodular for every non-decreasing and convex function h and non-decreasing and supermodular function Ψ. Considering Theorem 3.2 in Denuit and Muller (2002), we know that we can restrict our analysis to differentiable functions. It is then easily seen that for any $i \neq j$

$$ \begin{aligned} \frac{\partial^2}{\partial x_i \partial x_j} h\big(\Psi(x_1, x_2, \ldots, x_n)\big) &= \frac{\partial}{\partial x_j}\left(h'\big(\Psi(x_1, x_2, \ldots, x_n)\big)\frac{\partial}{\partial x_i}\Psi(x_1, x_2, \ldots, x_n)\right) \\ &= h''\big(\Psi(x_1, x_2, \ldots, x_n)\big)\frac{\partial}{\partial x_i}\Psi(x_1, x_2, \ldots, x_n)\frac{\partial}{\partial x_j}\Psi(x_1, x_2, \ldots, x_n) \\ &\quad + h'\big(\Psi(x_1, x_2, \ldots, x_n)\big)\frac{\partial^2}{\partial x_i \partial x_j}\Psi(x_1, x_2, \ldots, x_n) \\ &\geq 0. \end{aligned} $$

Hence, $\Psi(X_1, X_2, \ldots, X_n)$ precedes $\Psi(Y_1, Y_2, \ldots, Y_n)$ in the increasing convex (or stop-loss) order, which is equivalent to the announced ranking between the respective TVaRs as recalled in Section 3.1. $\square$

Remark 3.3. The result holds under the weaker condition $\mathrm{E}[X_i] = \mathrm{E}[Y_i]$ for $i = 1, 2, \ldots, n$, replacing $X_i =_d Y_i$ for $i = 1, 2, \ldots, n$. This is because $\Psi(X_1, X_2, \ldots, X_n) \preceq_{\mathrm{SL}} \Psi(Y_1, Y_2, \ldots, Y_n)$ ensures that $\mathrm{E}[\Psi(X_1, X_2, \ldots, X_n)] \leq \mathrm{E}[\Psi(Y_1, Y_2, \ldots, Y_n)]$ holds for every non-decreasing and supermodular function Ψ. In particular, we have that $\mathrm{E}[h(X_i)] \leq \mathrm{E}[h(Y_i)]$ holds for every non-decreasing function h so that $\mathrm{P}[X_i > t] \leq \mathrm{P}[Y_i > t]$ holds for all t and $i = 1, 2, \ldots, n$. Now, together with $\mathrm{E}[X_i] = \mathrm{E}[Y_i]$, this implies that $X_i =_d Y_i$ must hold.

We are now in a position to derive the main result of this section.

Proposition 3.4. *For every positive integer d,*

$$TVaR\left[g\Big(P_x(t|\Pi_0), P_{x+1}(t+1|\Pi_1), \ldots, P_{x+d}(t+d|\Pi_d)\Big); p\right]$$

$$\leq TVaR\left[g\Big(P_x(t|\Pi_0^+), P_{x+1}(t+1|\Pi_1^+), \ldots, P_{x+d}(t+d|\Pi_d^+)\Big); p\right] \text{ for all } p,$$

for every non-decreasing (or non-increasing) supermodular function g.

Proof. Clearly, the one-year conditional survival probabilities $P_{x+j}(t+j|\Pi_j)$ and $P_{x+j}(t+j|\Pi_j^+)$ are identically distributed. As the random variables $P_{x+j}(t+j|\Pi_j^+)$ are comonotonic (they are all decreasing functions of Z), we know from Property 3.1(ii) that the inequality

$$\mathrm{E}\left[g\Big(P_x(t|\Pi_0), P_{x+1}(t+1|\Pi_1), \ldots, P_{x+d}(t+d|\Pi_d)\Big)\right]$$

$$\leq \mathrm{E}\left[g\Big(P_x(t|\Pi_0^+), P_{x+1}(t+1|\Pi_1^+), \ldots, P_{x+d}(t+d|\Pi_d^+)\Big)\right]$$

holds for every supermodular function g. The stated result now follows from Property 3.2. $\square$

As a direct consequence of the preceding result, we are in a position to compare Tail-VaR risk measures for exact and approximate conditional survival probabilities.

Corollary 3.5. *For every positive integer d,*

$$TVaR\left[{}_dP_x(t|\mathbf{\Pi}); p\right] \leq TVaR\left[{}_dP_x(t|\mathbf{\Pi}^+); p\right] \text{ for all } p.$$

Proof. Applying Proposition 3.4 to the supermodular function

$$g(s_1, s_2, \ldots, s_d) = \prod_{l=1}^{d} s_l, \quad s_i \geq 0 \text{ for all } i \in \{1, \ldots, d\},$$

proves the announced inequality. $\square$

Compared to Denuit et al. (2013), who made the Λ_t comonotonic for different times t, we are here able to derive conservative approximations to the conditional survival probabilities, as shown in Corollary 3.5. As it can be seen from Property 1 in Denuit et al. (2013), making each factor Λ_t comonotonic allows to derive conservative approximations for death rates but not necessarily for the conditional survival probabilities (see Section 3 in that paper for more details).

Let us now demonstrate why the computations carried with the approximations built from $\mathbf{\Pi}^+$ become so easy. Consider the VaR of this conditional probability. Recall that

$$F^{-1}_{g(X)}(p) = \begin{cases} g\big(F^{-1}_X(p)\big) \text{ if } g \text{ is continuous and increasing,} \\ \\ g\big(F^{-1}_X(1-p)\big) \text{ if } g \text{ is continuous and decreasing.} \end{cases}$$

It is then easy to write under (2.2) that

$$\begin{aligned} \text{VaR}\left[{}_kP_x(t|\mathbf{\Pi}^+);p\right] &= \text{VaR}\left[\prod_{j=0}^{k-1}\Big(1-q_{x+j,t_0}\rho_{x+j}^{t-t_0+j}F^{-1}_{\Pi_j}(Z)\Big);p\right] \\ &= \prod_{j=0}^{k-1}\Big(1-q_{x+j,t_0}\rho_{x+j}^{t-t_0+j}F^{-1}_{\Pi_j}(1-p)\Big). \end{aligned}$$

A similar result holds under (2.3). Then,

$$\text{TVaR}\left[{}_kP_x(t|\mathbf{\Pi}^+);p\right] = \frac{1}{1-p}\int_p^1\prod_{j=0}^{k-1}\Big(1-q_{x+j,t_0}\rho_{x+j}^{t-t_0+j}F^{-1}_{\Pi_j}(1-\xi)\Big)\,d\xi$$

which easily follows by numerical integration.

Example 3.6. It is convenient to consider LogNormally distributed Λ_t, which ensures that products of Λ_t are still LogNormal. Assume that each Λ_t is LogNormally distributed with unit mean, i.e. $\ln \Lambda_t$ is Normally distributed with mean $-\sigma^2/2$ and variance σ^2. To fix the ideas, we assume that the Λ_t are mutually independent (but the analysis is easily extended to any covariance structure). Making the products Π_j comonotonic for different values of j means that we replace Π_j with

$$\Pi_j^+ = \exp\left(-\frac{t-t_0+j}{2}\sigma^2+\sqrt{t-t_0+j}\,\sigma Z\right),\ \ j=0,1,2,\dots,$$

where Z is Normally distributed with zero mean and unit variance. Notice that the same Z defines all the products Π_j^+ so that the latter are indeed comonotonic.

3.4 Conditional present values

Corollary 3.5 is useful for pure endowments, for which the conditional expected present value of insurance benefits is proportional to the d-year conditional survival probabilities. Life annuities can be obtained as sums of pure endowments and thus involve conditional survival probabilities up to different time horizons. This is why we now extend the preceding result to sequences of survival probabilities. To this end, recall that the function $g:\mathbb{R}^d\to\mathbb{R}$ is directionally convex if it is supermodular and componentwise convex, i.e. convex viewed as a function of x_i with fixed x_k, $k\neq i$, for any $i\in\{1,\ldots,d\}$. A twice differentiable function g is directionally convex if, and only if,

$$\frac{\partial^2}{\partial x_i \partial x_j} g \geq 0 \text{ for all } i,j\in\{1,\ldots,d\}.$$

Let us first establish a useful property that connects the increasing directionally convex and the stop loss orders for any increasing directionally convex transform of the components of the random vectors. As it was the case for Property 3.2, the following result is of independent interest.

Property 3.7. *Given two random vectors* $(X_1,X_2,\ldots,X_n)$ *and* $(Y_1,Y_2,\ldots,Y_n)$, $(X_1,X_2,\ldots,X_n)$ *precedes* $(Y_1,Y_2,\ldots,Y_n)$ *in the increasing directionally convex order, i.e. the inequality*

$$E[g(X_1,X_2,\ldots,X_n)]\leq E[g(Y_1,Y_2,\ldots,Y_n)]$$

holds for every non-decreasing and directionally convex function g *for which the expectations exist, if, and only if,*

$$TVaR[\Psi(X_1,X_2,\ldots,X_n);p]\leq TVaR[\Psi(Y_1,Y_2,\ldots,Y_n);p] \text{ for all } p$$

for every non-decreasing and directionally convex function Ψ.

Proof. The "$\Leftarrow$" part is obviously true. Considering the "$\Rightarrow$" part, let us show that $h\circ\Psi$ is non-decreasing and directionally convex for every non-decreasing and convex function h and non-decreasing and directionally convex function Ψ. As in the proof of Property 3.2, we can restrict ourselves to differentiable functions. It is then easily seen that $\frac{\partial}{\partial x_i}h\big(\Psi(x_1,x_2,\ldots,x_n)\big)\geq 0$ and $\frac{\partial^2}{\partial x_i\partial x_j}h\big(\Psi(x_1,x_2,\ldots,x_n)\big)\geq 0$ for any $i=1,\ldots,n$ and $j=1,\ldots,n$, which ends the proof. □

Let us now apply this result to sequence of conditional survival probabilities.

Proposition 3.8. *For every integer* $d\geq 2$, *we have*

$$TVaR\left[g\Big(P_x(t|\mathbf{\Pi}),{}_2P_x(t|\mathbf{\Pi}),\ldots,{}_dP_x(t|\mathbf{\Pi})\Big);p\right]$$

$$\leq TVaR\left[g\Big(P_x(t|\mathbf{\Pi}^+),{}_2P_x(t|\mathbf{\Pi}^+),\ldots,{}_dP_x(t|\mathbf{\Pi}^+)\Big);p\right] \text{ for all } p,$$

for every non-decreasing directionally convex function g.

Proof. The reasoning is similar to the one leading to Proposition 3.1 in Gbari and Denuit (2016). Define the function $\Psi:[0;1]^d\longrightarrow[0;1]^d$ as

$$\Psi(p_1,\ldots,p_d)=\left(p_1,p_1p_2,\ldots,\prod_{l=1}^d p_l\right).$$

Then, the vector of conditional survival probabilities can be written as

$$\Big(P_x(t|\mathbf{\Pi}), {}_2P_x(t|\mathbf{\Pi}), \ldots, {}_dP_x(t|\mathbf{\Pi})\Big) = \Psi\Big(P_x(t|\mathbf{\Pi}), P_{x+1}(t+1|\mathbf{\Pi}), \ldots, P_{x+d}(t+d|\mathbf{\Pi})\Big).$$

Notice that the coordinate functions of Ψ, i.e. the functions mapping $(p_1, p_2, \ldots, p_d)$ to the product $p_1 p_2 \ldots p_l$ are all directionally convex, so that the function Ψ is also directionally convex. Hence, the random vector of the k-year conditional survival probabilities is a directionally convex transform of the random vector of the one-year conditional survival probabilities. As Proposition 3.2 ensures that the random vectors are also ordered in the increasing directionally convex order, i.e. that the inequality

$$\begin{aligned} &\mathrm{E}\left[g\Big(P_x(t|\Pi_0), P_{x+1}(t+1|\Pi_1), \ldots, P_{x+d}(t+d|\Pi_d)\Big)\right] \\ &\leq \mathrm{E}\left[g\Big(P_x(t|\Pi_0^+), P_{x+1}(t+1|\Pi_1^+), \ldots, P_{x+d}(t+d|\Pi_d^+)\Big)\right] \end{aligned}$$

holds for every non-decreasing directionally convex function g, Theorem 7.A.30 from Shaked and Shanthikumar (2007) allows us to conclude that the inequality

$$\begin{aligned} &\mathrm{E}\left[g\Big(P_x(t|\mathbf{\Pi}), {}_2P_x(t|\mathbf{\Pi}), \ldots, {}_dP_x(t|\mathbf{\Pi})\Big)\right] \\ &\leq \mathrm{E}\left[g\Big(P_x(t|\mathbf{\Pi}^+), {}_2P_x(t|\mathbf{\Pi}^+), \ldots, {}_dP_x(t|\mathbf{\Pi}^+)\Big)\right] \end{aligned}$$

holds for every non-decreasing directionally convex function g. The announced result then follows from Property 3.7. □

As an application, consider a basic life annuity contract paying one monetary unit at the end of each year, as long as the annuitant survives. Also, denote ξ rounded from below as $\lfloor \xi \rfloor$, i.e. $\lfloor \xi \rfloor$ is the largest integer smaller than, or equal to ξ. Then, $\lfloor T_i \rfloor$ is the curtate remaining lifetime for policyholder i. Let $v(s,t)$ be the present value at time s of a unit payment made at time t, $s \leq t$. The present value of all the payments made to annuitant i is

$$a_{\overline{T_i}|} = \sum_{k=1}^{\lfloor T_i \rfloor} v(t, t+k),$$

with the convention that the empty sum is zero.

The random variable

$$a_x(t|\mathbf{\Pi}) = \mathrm{E}[a_{\overline{T_1}|}|\mathbf{\Pi}] = \sum_{k\geq 1} {}_kP_x(t|\mathbf{\Pi}) v(t, t+k)$$

corresponds to the conditional expected present value of the payments made to an annuitant aged x in calendar year t whose survival obeys the exponential decline mortality projection model. Here, the discount factors can be deduced from an appropriate yield curve and are thus treated as known deterministic values. This random variable can be seen as the residual risk per annuity contract in an infinitely large portfolio where only systematic longevity risk remains. Defining

$$a_x(t|\mathbf{\Pi}^+) = \sum_{k\geq 1} {}_kP_k(t|\mathbf{\Pi}^+) v(t, t+k),$$

Proposition 3.7 allows us to write

$$\text{TVaR}\Big[a_x(t|\mathbf{\Pi});p\Big] \leq \text{TVaR}\Big[a_x(t|\mathbf{\Pi}^+);p\Big] \tag{3.1}$$

for all probability levels p, so that the comonotonic approximation provides an upper bound on the Tail-VaR of the systematic longevity risk in a large life annuity portfolio. Thus, we have derived a conservative approximation for the risk borne by the annuity provider when the portfolio is sufficiently large. Moreover, computations are very easy as Property 3.1(i) allows us to write

$$\text{VaR}\Big[a_x(t|\mathbf{\Pi}^+);p\Big] = \sum_{k\geq 1} \text{VaR}\Big[{}_kP_x(t|\mathbf{\Pi}^+);p\Big]v(t,t+k) \tag{3.2}$$

$$\text{TVaR}\Big[a_x(t|\mathbf{\Pi}^+);p\Big] = \sum_{k\geq 1} \text{TVaR}\Big[{}_kP_x(t|\mathbf{\Pi}^+);p\Big]v(t,t+k). \tag{3.3}$$

3.5 Numbers of survivors

Assume now that the insurer sells $L_0 = n$ life annuities to individuals aged x in calendar year t with identically distributed remaining lifetimes $T_1, T_2, \ldots, T_n$. More precisely, given $\mathbf{\Lambda}$, we assume that these lifetimes are independent and subject to the common survival probabilities (2.2) or (2.3). Define the number L_k of survivors to age $x+k$ for $k = 1, 2, \ldots$, starting from L_0. Formally, L_k is given by

$$L_k = \sum_{i=1}^{n} \text{I}[T_i > k] = L_{k+1} + \sum_{i=1}^{n} \text{I}[k < T_i \leq k+1],$$

where $\text{I}[\cdot]$ stands for the indicator function. Thus, even if the lifetimes were independent, the L_k are expected to be strongly positively dependent. Therefore, making them perfectly dependent, given the random variable Z controlling the conditional survival probabilities ${}_kP_x(t|\mathbf{\Pi}^+)$, is expected to produce a conservative and reasonably accurate approximation. This leads us to define

$$L_k^+ = \sum_{i=0}^{n-1} \text{I}\left[U > \sum_{j=0}^{i} \binom{n}{j} \big({}_kP_x(t|\mathbf{\Pi}^+)\big)^j \big(1 - {}_kP_x(t|\mathbf{\Pi}^+)\big)^{n-j}\right]$$

where U is a unit uniform random variable independent of Z defining $\mathbf{\Pi}^+$. Given Z, the random variables L_k^+ are comonotonic, being all obtained as increasing transformations of the same unit uniform U.

The next result shows that replacing L_k with L_k^+ provides the actuary with upper bounds on a variety of quantities.

Proposition 3.9. *For every integer $d \geq 2$, we have*

$$TVaR\left[g\Big(L_1, \ldots, L_d\Big);p\right] \leq TVaR\left[g\Big(L_1^+, \ldots, L_d^+\Big);p\right] \text{ for all } p,$$

for every non-decreasing directionally convex function g.

Proof. Given a function $g:\{0,1,\ldots,n\}^d\to\mathbb{R}$, define the auxiliary function $g^\star:[0,1]^d\to\mathbb{R}$ as

$$
\begin{aligned}
g^\star(p_x,\ldots,p_{x+d-1}) &= \mathrm{E}[g(L_1,\ldots,L_d)|P_{x+k-1}(t|\mathbf{\Pi})=p_{x+k-1},\ k=0,1\ldots,d] \\
&= \sum_{l_1=0}^{n}\sum_{l_2=0}^{l_1}\cdots\sum_{l_d=0}^{l_{d-1}} g(l_1,l_2,\ldots,l_d) \\
&\quad \binom{n}{l_1}p_x^{l_1}q_x^{n-l_1}\binom{l_1}{l_2}p_{x+1}^{l_2}q_{x+1}^{l_1-l_2}\ldots\binom{l_{d-1}}{l_d}p_{x+d-1}^{l_d}q_{x+d-1}^{l_{d-1}-l_d}.
\end{aligned}
$$

Now, we know from the proof of Proposition 3.2 in Gbari and Denuit (2014) that the auxiliary function $g^\star$ is supermodular provided g is increasing and directionally convex. Then

$$
\begin{aligned}
\mathrm{E}[g(L_1,\ldots,L_d)] &= \mathrm{E}[g^\star(P_x(t|\mathbf{\Pi}),\ldots,P_{x+d-1}(t|\mathbf{\Pi}))] \\
&\leq \mathrm{E}[g^\star(P_x(t|\mathbf{\Pi}^+),\ldots,P_{x+d-1}(t|\mathbf{\Pi}^+))] \text{ by Proposition 3.2} \\
&= \mathrm{E}[g(L_1^+,\ldots,L_d^+)].
\end{aligned}
$$

This ends the proof. □

In particular, define

$$V=\sum_{i=1}^{n} a_{\overline{T_i|}}=\sum_{k\geq 1}L_k v(0,k)$$

and

$$V^+=\sum_{k\geq 1}L_k^+ v(0,k).$$

Then, we have

$$\mathrm{TVaR}\big[V;p\big]\leq \mathrm{TVaR}\big[V^+;p\big] \text{ for all } p. \tag{3.4}$$

4 Turning a regulatory life table into a stochastic mortality projection model

4.1 Model calibration

Assume that the insurer is subject to regulatory requirements specifying the base table q_{x,t_0} and the age-specific improvement rates ρ_x. Specifically, the one-year death probabilities q_{x,t_0+k} are assumed to be given by (1.1)-(1.2). It is clear that performing the actuarial calculations using these q_{x,t_0+k} as if they were known with certainty, doing as if the remaining lifetimes $T_1,\ldots,T_n$ were independent, may greatly underestimate the risk borne by the insurer. The reason is clear: future one-year death probabilities are unknown and any departure from the best estimate q_{x,t_0+k} impacts the whole portfolio. In addition to random deviations from a known life table, the insurer also covers the systematic risk associated to the life table itself. To recognize the randomness in the future mortality, i.e. the systematic longevity risk, the actuary decides to impact the deterministic ρ_x with random shocks Λ_{t_0+k} before making the calculations.

The dynamics of these shocks can be determined from their empirical counterparts. Specifically, empirical $\widehat{\Lambda}_t$ can be obtained by averaging $\widehat{q}_{x,t}/(\widehat{q}_{x,t-1}\rho_x)$ over ages x, where $\widehat{q}_{x,t}$ is the observed one-year death probability. A time series analysis can be performed on the resulting $\widehat{\Lambda}_t$ to test for the presence of serial correlation.

As an application, we consider Belgian mortality statistics for males together with the reference projected life table published by the Federal Planning Bureau. The FPB model specifies $q_{x,t} = \exp(\alpha_x + \beta_x(t - t_0))$, $t > t_0$, where β_x is the rate of decrease of $q_{x,t}$ over time. Here, t_0 varies according to the successive forecasts made by the FPB. Thus, each age-specific death probability is assumed to decline at its own exponential rate, which is in line with the specification (1.1) defining

$$\rho_x = \exp(\beta_x).$$

The empirical $\widehat{\Lambda}_t$ are obtained for the period 2001-2012 by averaging $\widehat{q}_{x,t}/(\widehat{q}_{x,t-1}\rho_x^\star)$ over ages 25-100, where $\rho_x^\star$ is the mortality improvement rate computed by FPB in 2001 (as we aim to calibrate random shocks from the values that were predicted at the beginning of the period considered, i.e. 2001-2012). In Figure 4.1 we display the resulting $\widehat{\Lambda}_t$.

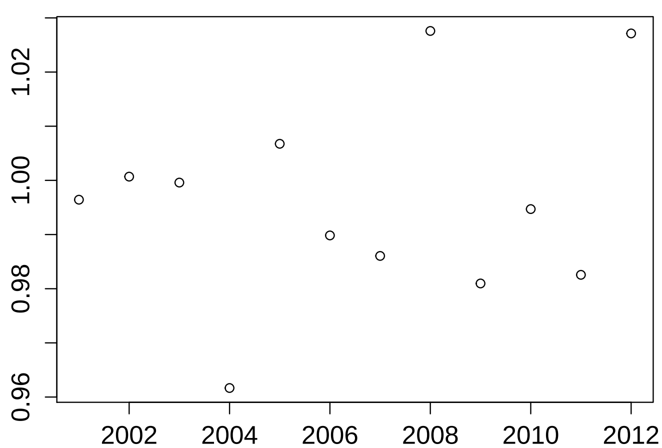

Figure 4.1: Empirical $\widehat{\Lambda}_t$ obtained for the period 2001-2012 by averaging the $\widehat{q}_{x,t}/(\widehat{q}_{x,t-1}\rho_x^\star)$ over ages $25 - 100$, where $\rho_x^\star$ is the mortality improvement rate computed by FPB in 2001.

Figure 4.2 displays the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the empirical $\widehat{\Lambda}_t$. The ACF at lag k gives the correlation coefficient between the observations made at times t and $t+k$. It measures the linear predictability of the series at time $t+k$ using only the value at time t. As it can be seen from Figure 4.2, there is no significant serial correlation between the time indices at various lags.

Let us then consider independent Λ_t that are LogNormally distributed with unit mean, i.e. $\ln \Lambda_t$ is Normally distributed with mean $-\sigma^2/2$ and variance σ^2 (see Example 3.6). By maximizing the corresponding LogNormal likelihood computed from $\widehat{\Lambda}_{2001}, \ldots, \widehat{\Lambda}_{2012}$ we get the parameter estimate $\widehat{\sigma} = 0.0184$.

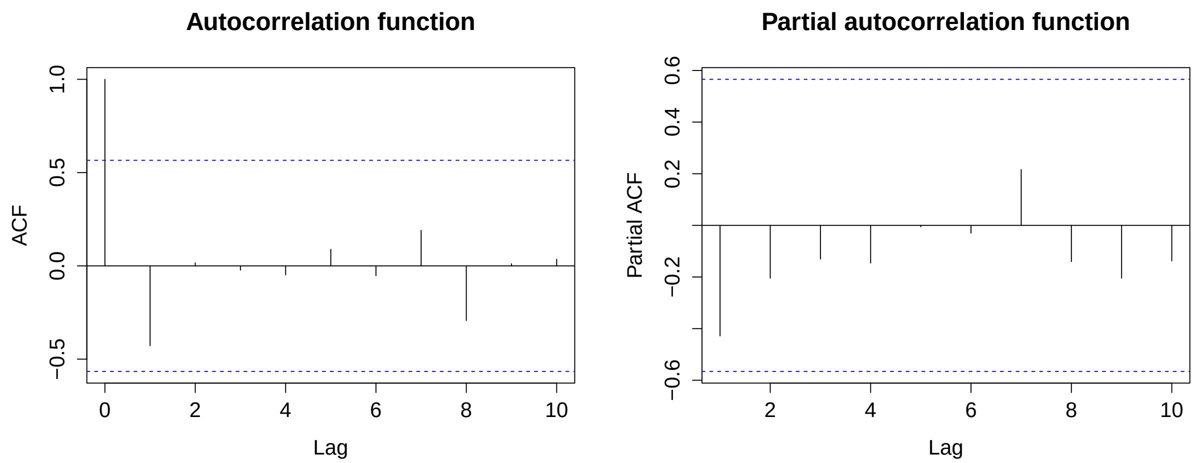

Figure 4.2: ACF and PACF of the empirical $\widehat{\Lambda}_t$ shown in Figure 4.1.

4.2 Life annuity contract

We consider an annuitant aged 65 in calendar year 2016. The constant technical interest rate is 2%. Figure 4.3 displays the distribution function of $a_x(t|\mathbf{\Pi})$ obtained from 100 000 simulations of the path of the random departures $\mathbf{\Lambda}$. Comonotonic approximations (3.2) and (3.3) together with $\mathrm{VaR}[a_x(t|\mathbf{\Pi}); p]$ and $\mathrm{TVaR}[a_x(t|\mathbf{\Pi}); p]$ computed from the simulated $a_x(t|\mathbf{\Pi})$ are depicted in Figure 4.4 for several probability levels p.

We clearly see there that $p \mapsto \mathrm{VaR}[a_x(t|\mathbf{\Pi}^+); p]$ crosses $p \mapsto \mathrm{VaR}[a_x(t|\mathbf{\Pi}); p]$ exactly once, around $p = 50\%$ and dominates it after the unique crossing point. Together with the ordering of the respective means $\mathrm{E}[a_x(t|\mathbf{\Pi}^+)]$ and $\mathrm{E}[a_x(t|\mathbf{\Pi})]$, this supports the ranking of the Tail-VaR that is known to hold from (3.1). The comonotonic approximation $\mathrm{TVaR}[a_x(t|\mathbf{\Pi}^+); p]$ always dominates $\mathrm{TVaR}[a_x(t|\mathbf{\Pi}); p]$, in accordance with (3.1). Furthermore, we see that $\mathrm{VaR}[a_x(t|\mathbf{\Pi}^+); p]$ (resp. $\mathrm{TVaR}[a_x(t|\mathbf{\Pi}^+); p]$) is quite close to the simulated value $\mathrm{VaR}[a_x(t|\mathbf{\Pi}); p]$ (resp. $\mathrm{TVaR}[a_x(t|\mathbf{\Pi}); p]$) since their relative differences vary between -0.498% (resp. 0.048%) for $p = 5\%$ and 0.834% (resp. 0.944%) for $p = 99.5\%$ (see Figure 4.5). This suggests that the proposed comonotonic approximations are reasonably accurate and provide the actuary with the correct order of magnitude at reduced computational cost.

4.3 Numbers of survivors

We consider a portfolio of $n = 100$ (small portfolio) or $n = 1\,000$ (large portfolio) annuitants aged $x = 65$ in 2016 with constant technical interest rate 2%. For each of the 100 000 simulated paths for $\mathbf{\Lambda}$, we generate a realization of the remaining lifetimes $T_1, T_2, \ldots, T_n$ from the corresponding survival probabilities ${}_kP_x(t|\mathbf{\Pi})$. Then, we get the number L_k of survivors to age $x + k$ for $k = 1, 2, \ldots$, starting from $L_0 = n$. We also make 100 000 simulations of the pair (U, Z) producing the numbers L_k^+ for $k = 1, 2, \ldots$ with $L_0^+ = n$.

Figures 4.6 ($n = 100$) and 4.7 ($n = 100\,00$) show the distribution functions of V and V^+ together with those of the large portfolio approximations $na_x(t|\mathbf{\Pi})$ and $na_x(t|\mathbf{\Pi}^+)$. We observe that the distribution function of V crosses the distribution function of V^+ only once

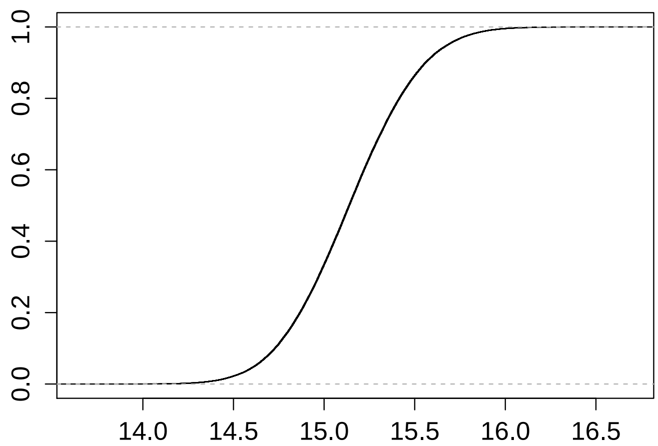

Figure 4.3: Empirical distribution function of $a_x(t|\mathbf{\Pi})$ obtained by simulating 100 000 paths of the random departures $\mathbf{\Lambda}$.

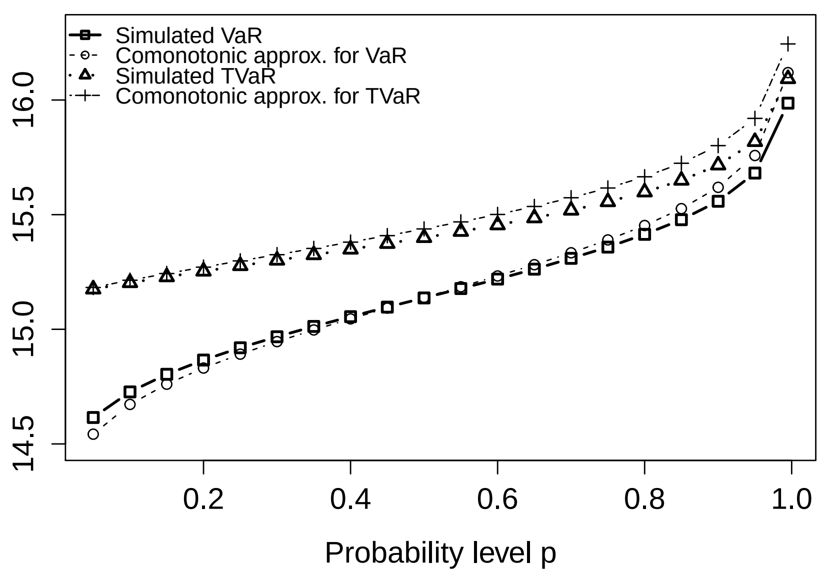

Figure 4.4: Comonotonic approximations $\text{VaR}[a_x(t|\mathbf{\Pi}^+);p]$ and $\text{TVaR}[a_x(t|\mathbf{\Pi}^+);p]$ together with $\text{VaR}[a_x(t|\mathbf{\Pi});p]$ and $\text{TVaR}[a_x(t|\mathbf{\Pi});p]$ deduced from the simulated $a_x(t|\mathbf{\Pi})$ for probability levels $p \in \{5\%, 10\%, 15\%, 20\%, \ldots, 90\%, 95\%, 99.5\%\}$.

and dominates it after the unique crossing point. This supports the stochastic inequality (3.4). Of course, when the portfolio size n increases, we also see that the distribution

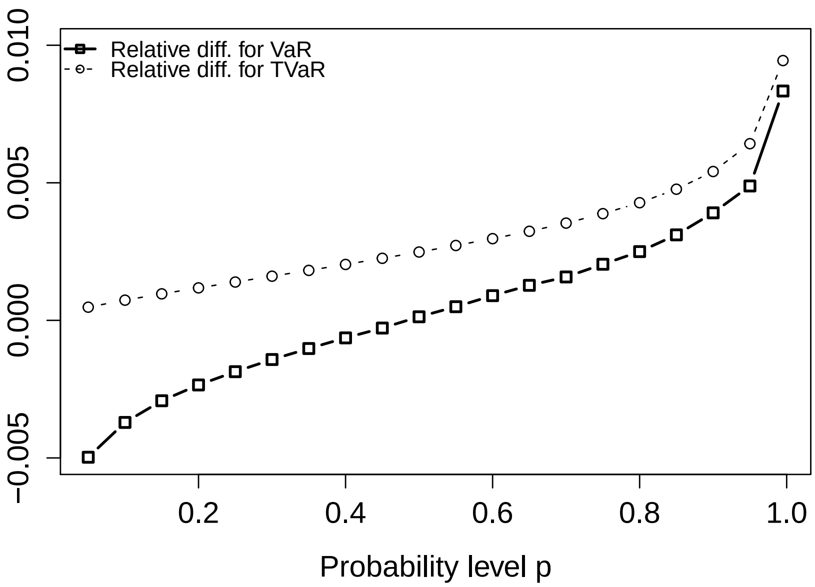

Figure 4.5: Relative difference between the comonotonic approximation $\mathrm{VaR}[a_x(t|\mathbf{\Pi}^+); p]$ (resp. $\mathrm{TVaR}[a_x(t|\mathbf{\Pi}^+); p]$) and the simulated value $\mathrm{VaR}[a_x(t|\mathbf{\Pi}); p]$ (resp. $\mathrm{TVaR}[a_x(t|\mathbf{\Pi}); p]$) for probability levels $p \in \{5\%, 10\%, 15\%, 20\%, \ldots, 90\%, 95\%, 99.5\%\}$.

function of the large portfolio approximation $na_x(t|\mathbf{\Pi})$ (resp. $na_x(t|\mathbf{\Pi}^+)$) gets closer to the distribution function of V (resp. V^+).

5 Calibration on historical data

5.1 Mortality improvement rates

Assume that we have at our disposal age-specific mortality statistics for calendar years $t_1, \ldots, t_n$ and that we aim to predict mortality for future years $t_n + k$, $k \geq 1$. Let $D_{x,t}$ be the number of deaths recorded at age x in calendar year t, from an exposure-to-risk $\mathrm{ETR}_{x,t}$. Henceforth, we work with death rates, i.e. we consider specification (2.3). Let $\mu_{x,t}$ be the force of mortality at age x in calendar year t, assumed to be constant on each square of the Lexis diagram (but allowed to vary between squares). We then have $\mu_{x,t} = m_{x,t}$.

Recently, several authors have introduced and investigated parametric mortality projection methods based on mortality improvement rates (as opposed to mortality rates). See, e.g., Haberman and Renshaw (2012, 2013), Mitchell et al. (2013) and Schinzinger et al. (2014). In this section, we explain how to fit model (2.3) based on mortality improvement rates to historical data.

Let us decompose ρ_x into $\beta_x \delta$, imposing $\beta_x \geq 0$ and $\sum_x \beta_x = 1$. Precisely, we assume that mortality at age x improves at yearly rate $\beta_x \delta \Lambda_t$, where

- the coefficients $\beta_x \geq 0$ measure the sensitivity of mortality at the different ages x,

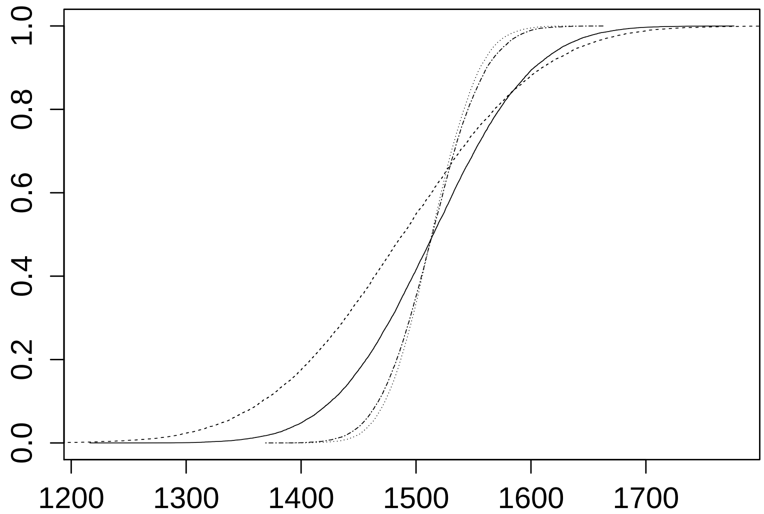

Figure 4.6: Distribution functions of V (—) and V^+ (- - -) together with those of the large portfolio approximations $na_x(t|\mathbf{\Pi})$ (······) and $na_x(t|\mathbf{\Pi}^+)$ (· - · -) for $n = 100$.

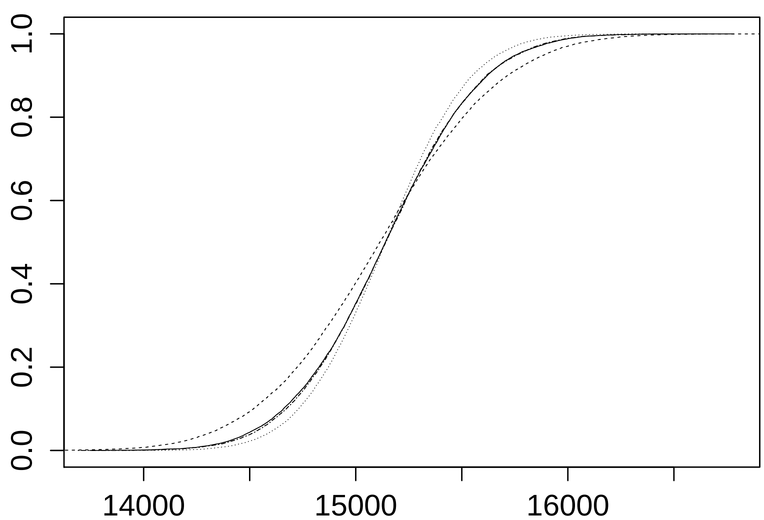

Figure 4.7: Distribution functions of V (—) and V^+ (- - -) together with those of the large portfolio approximations $na_x(t|\mathbf{\Pi})$ (······) and $na_x(t|\mathbf{\Pi}^+)$ (· - · -) for $n = 1000$.

subject to the identifiability constraint $\sum_x \beta_x = 1$;

- the positive random variables Λ_t, with unit mean, account for possible departures from the expected trend caused by circumstances specific to calendar year t;
- the parameter δ can be interpreted as the yearly expected global improvement factor because the aggregate all-age improvement is equal on average to

$$\mathrm{E}\left[\sum_x \beta_x \delta \Lambda_t\right] = \delta.$$

To fit the model to observations relating to calendar years $t_1, \ldots, t_n$, we consider a base year $t_0 < t_1$. Then,

$$\mathrm{E}[D_{x,t}|\mathbf{\Lambda}] = \mathrm{ETR}_{x,t}\big(\beta_x\delta\big)^{t-t_0}\left(\prod_{j=t_0}^{t-1}\Lambda_j\right)\mu_{x,t_0}, \quad t > t_0,$$

applies to all $t \in \{t_1, \ldots, t_n\}$. Notice that our starting point corresponds to year t_0, and working backward is not equivalent as $\mathrm{E}[\Lambda_t^{-1}] \neq 1$ when $\mathrm{E}[\Lambda_t] = 1$.

5.2 Mixed Poisson likelihood

The Poisson specification for death counts proposed by Brouhns et al. (2002) has now been largely adopted in actuarial mortality studies. In order to estimate the parameters, we assume that $D_{x,t}$ are independent Poisson random variables, given $\mathbf{\Lambda}$. Dealing with Poisson mixtures, convenient choices for the distribution of Λ_t include the Gamma, LogNormal and Inverse Gaussian distributions. Here, we consider LogNormally distributed Λ_t, which ensures that products of Λ_t are still LogNormal. Henceforth, we assume that the Λ_t are independent and identically distributed and that each Λ_t is LogNormally distributed with unit mean, i.e. $\ln\Lambda_t$ is Normally distributed with mean $-\sigma^2/2$ and variance σ^2.

Remark 5.1. As an alternative to independent and identically distributed Λ_t we could also consider that $\Lambda_t = \exp(Z_t)$ with Z_t obeying an AR(1) process. In the exponential decline model, mean reverting behavior is meaningful for $Z_t = \ln\Lambda_t$, in the sense that deviations of Z_t from its expected value $-\frac{\sigma^2}{2}$ are corrected at the next step so that Z_t has a long term trend level:

$$Z_t = -\frac{\sigma^2}{2} + \alpha\left(Z_{t-1} + \frac{\sigma^2}{2}\right) + \Delta_t.$$

The direct maximization of the exact Poisson-LogNormal likelihood is out of reach as it involves multiple integrals over each Λ_t comprised in the observation period. Therefore, we opt for an approximate likelihood that retains the mixed Poisson marginal distributions for the death counts $D_{x,t}$ as well as part of their dependence structure.

The sample analog to

$$\frac{\mathrm{E}[D_{x,t-1}|\mathbf{\Lambda}]}{\mathrm{ETR}_{x,t-1}} = \big(\beta_x\delta\big)^{t-t_0-1}\left(\prod_{j=t_0}^{t-2}\Lambda_j\right)m_{x,t_0}$$

is the crude death rate

$$\frac{D_{x,t-1}}{\mathrm{ETR}_{x,t-1}} = \widehat{m}_{x,t-1}.$$

In order to estimate the parameters $(\beta_x, \delta, \sigma^2)$, we use the approximation

$$\mathrm{E}[D_{x,t}|\boldsymbol{\Lambda}] \approx \mathrm{ETR}_{x,t}\widehat{m}_{x,t-1}\beta_x\delta\Lambda_{t-1}$$

and we maximize the corresponding Poisson-LogNormal likelihood. Notice that $\mathrm{ETR}_{x,t}\widehat{m}_{x,t-1}$ is the expected number of deaths at age x in calendar year t if the mortality conditions prevailing in year $t-1$ still apply in year t, thus in the absence of longevity improvements.

5.3 Application

In practice, the mean specification in the likelihood under log-link is

$$\exp\left(\ln\left(\frac{\mathrm{ETR}_{x,t}}{\mathrm{ETR}_{x,t-1}}D_{x,t-1}\right) + \ln\delta + \ln\beta_x\right)$$

so that $\ln\left(\frac{\mathrm{ETR}_{x,t}}{\mathrm{ETR}_{x,t-1}}D_{x,t-1}\right)$ is treated as an offset, $\ln\delta$ is the intercept, and age is treated as a factor suitably constrained.

The model is fitted to Belgian male mortality experience for the period $1970-2012$ and age range 25-100. The data comprise the annual numbers of recorded deaths and matching exposures to risk.

Figure 5.1 plots the coefficients β_x with the constraint $\sum_x \beta_x = 1$. Also, $\widehat{\delta} = 75.039$ and $\widehat{\sigma} = 0.055$. Based on this set of estimated $\rho_x = \beta_x\delta$ and σ, we can perform all the actuarial calculations described in Section 4, using simulations of comonotonic approximations.

6 Discussion

In this paper, we have studied a simple mortality projection model where random shocks impact on a set of yearly age-specific mortality improvement rates. This model encompasses many existing reference life tables produced by regulators or professional bodies. It recognizes the systematic risk induced by the uncertainty surrounding future mortality. The model can also be calibrated on historical data using a Poisson-LogNormal likelihood. Comonotonic approximations have been derived for various quantities of interest, providing the actuary with reasonably accurate values at reduced computational cost.

Acknowledgements

Michel Denuit acknowledges the financial support from the contract "Projet d'Actions de Recherche Concertées" No 12/17-045 of the "Communauté française de Belgique", granted by the "Académie universitaire Louvain".

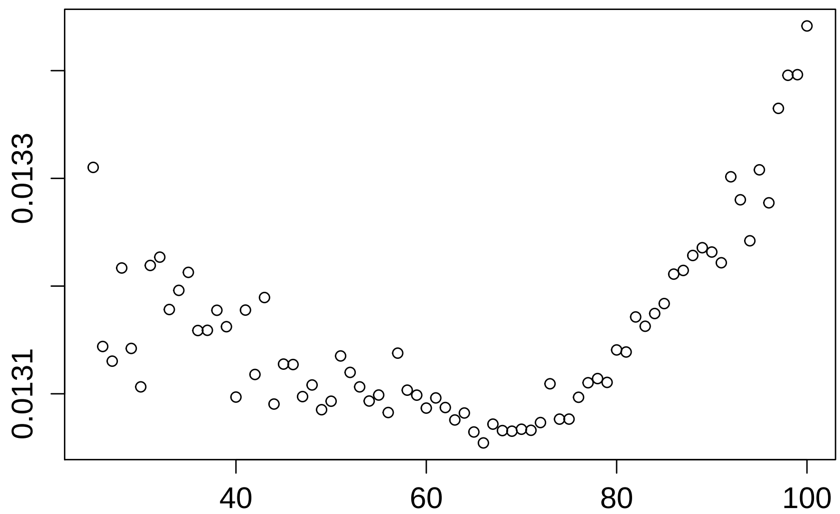

Figure 5.1: Estimated coefficients β_x for the age range 25-100.

References

- Brouhns, N., Denuit, M., Vermunt, J.K. (2002). A Poisson log-bilinear approach to the construction of projected life tables. Insurance: Mathematics and Economics 31, 373-393.

- Cadena, M., Denuit, M. (2015). Semi-parametric accelerated hazard relational models with applications to mortality projections. ISBA Discussion Paper - 2015/24.

- Denuit, M. (2007). Distribution of the random future life expectancies in log-bilinear mortality projection models. Lifetime Data Analysis 13, 381-397.

- Denuit, M., Dhaene, J. (2007). Comonotonic bounds on the survival probabilities in the Lee-Carter model for mortality projection. Computational and Applied Mathematics 203, 169-176.

- Denuit, M., Dhaene, J., Goovaerts, M.J., Kaas, R. (2005). Actuarial Theory for Dependent Risks: Measures, Orders and Models. Wiley, New York.

- Denuit, M., Frostig, E. (2007). Association and heterogeneity of insured lifetimes in the Lee-Carter framework. Scandinavian Actuarial Journal 2007, 1-19.

- Denuit, M., Haberman, S., Renshaw A.E. (2010). Comonotonic approximations to quantiles of life annuity conditional expected present values: extensions to general ARIMA models and comparison with the bootstrap. ASTIN Bulletin 40, 331-349.

- Denuit, M., Haberman, S., Renshaw, A. (2013). Approximations for quantiles of life expectancy and annuity values using the parametric improvement rate approach to modelling and projecting mortality. European Actuarial Journal 3, 191-201.

- Denuit, M., Muller, A. (2002). Smooth generators of integral stochastic orders. Annals of Applied Probability 12, 1174-1184.

- Gbari, S., Denuit, M. (2014). Efficient approximations for numbers of survivors in the Lee-Carter model. Insurance: Mathematics and Economics 59, 71-77.

- Gbari, S., Denuit, M. (2016). Stochastic approximations in CBD mortality projection models. Journal of Computational and Applied Mathematics 296, 102-115.

- Haberman, S., Renshaw A.E. (2012). Parametric mortality improvement rate modelling and projecting. Insurance: Mathematics and Economics 50, 309-333.

- Haberman, S., Renshaw, A. (2013). Modelling and projecting mortality improvement rates using a cohort perspective. Insurance: Mathematics and Economics 53, 150-168.

- Jarner, S. F., Moller, T. (2015). A partial internal model for longevity risk. Scandinavian Actuarial Journal 2015, 352-382.

- Kruger, R., Pasdika, U. (2006). Coping with longevity - the new German annuity valuation table DAV 2004 R. In Proceedings of the 28th International Congress of Actuaries, Paris, Vol. 28, pp. 12-14.

- Mitchell, D., Brockett, P., Mendoza-Arriaga, R., Muthuraman, K. (2013). Modeling and forecasting mortality rates. Insurance: Mathematics and Economics 52, 275-285.

- Muller, A., Scarsini, M. (2000). Some remarks on the supermodular order. Journal of Multivariate Analysis 73, 107-119.

- Schinzinger, E., Denuit, M., Christiansen, M. (2014). An evolutionary credibility model of Lee-Carter type for mortality improvement rates. ISBA Discussion Paper - 2014/21.

- Shaked, M., Shanthikumar, J.G. (2007). Stochastic Orders. Springer, New York.