

LIFE INSURANCE MATHEMATICS WITH RANDOM LIFE TABLES

Michel Denuit* and Esther Frostig[†]

ABSTRACT

When the insurer sells life annuities, projected life tables incorporating a forecast of future longevity must be used for pricing and reserving. To fix the ideas, the framework of Lee and Carter is adopted in this paper. The Lee-Carter model for mortality forecasting assumes that the death rate at age x in calendar year t is of the form $\exp(\alpha_x + \beta_x \kappa_t)$, where the time-varying parameter κ_t reflects the general level of mortality and follows an ARIMA model. The future lifetimes are all influenced by the same time index κ_t in this framework. Because the future path of this index is unknown and modeled as a stochastic process, the policyholders' lifetimes become dependent on each other. Consequently the risk does not disappear as the size of the portfolio increases: there always remains some systematic risk that cannot be diversified, whatever the number of policies. This paper aims to investigate some aspects of actuarial mathematics in the context of random life tables. First, the type of dependence existing between the insured life lengths is carefully examined. The way positive dependence influences the need for economic capital is assessed compared to mutual independence, as well as the effect of the timing of deaths through Bayesian credibility mechanisms. Then the distribution of the present value of payments under a closed group of life annuity policies is studied. Failing to account for the positive dependence between insured lifetimes is a dangerous strategy, even if the randomness in the future survival probabilities is incorporated in the actuarial computations. Numerical illustrations are performed on the basis of Belgian mortality statistics. The impact on the distribution of the present value of the additional variability that results from the Lee-Carter model is compared with the traditional method of mortality projection. Also, the impact of ignoring the dependence that arises from the model is quantified.

1. INTRODUCTION AND MOTIVATION

During the twentieth century, human mortality globally declined. These mortality improvements pose a challenge for the planning of public retirement systems as well as for the private life annuities business. When long-term living benefits are concerned, the calculation of expected present values (for pricing or reserving) requires an appropriate mortality projection to avoid underestimation of future costs. Actuaries therefore have to resort to life tables including a forecast of the future trends in mortality (the so-called projected life tables).

Different approaches for building projected life tables have been developed; see, for example, Pitacco et al. (2009) for an extensive review. To fix the ideas, we use in this paper the model introduced by Lee and Carter (1992). These authors proposed a simple model for describing the secular change in mortality as a function of a single time index. Precisely, the force of mortality at age x in calendar year

* Michel Denuit, Ph.D., professor of probability, statistics and actuarial mathematics, Institut de Sciences Actuarielles & Institut de Statistique, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium, michel.denuit@uclouvain.be.

[†] Esther Frostig, Ph.D., professor, Department of Statistics, University of Haifa, Haifa, Israel, frostig@stat.haifa.ac.il.

t is of the form $\exp(\alpha_x + \beta_x \kappa_t)$, where the stochastic process $\{\kappa_t, t \geq 0\}$ reflects the general level of mortality. The method describes the log of a time series of age-specific death rates as the sum of an age-specific component α_x that is independent of time and another component $\beta_x \kappa_t$ that is the product of a time-varying parameter reflecting the general level of mortality, and an age-specific component that represents how rapidly or slowly mortality at each age varies when the general level of mortality changes. This model is fitted to historical data. The resulting estimate of the time-varying parameter κ_t is then modeled and projected as a stochastic time series using standard Box-Jenkins methods. From this forecast of the general level of mortality, the actual age-specific rates are derived using the estimated age effects. Often a simple random walk with drift is used to project the time index to the future. Section 2 briefly summarizes the Lee-Carter approach. For more details we refer the interested readers to Lee (2000) and Pitacco et al. (2009).

This paper purposes to study the consequences of the dependence induced by projected life tables between the insured lifetimes. The textbook assumption for pricing life insurance contracts is that mortality risk is completely diversifiable and therefore not priced by markets in economic equilibrium. The law of large numbers is invoked to argue that a large enough portfolio effectively eliminates any idiosyncratic mortality risk. This paradigm is challenged when dependence that cannot be diversified away by selling more contracts is induced by the unknown evolution of future death rates. This is the case when the death rates are of the form $\exp(\alpha_x + \beta_x \kappa_t)$, where κ_t follows some stochastic process. Because all the future life lengths are influenced by the same κ_t 's, tending to shorten when the κ_t 's increase provided the β_x 's are positive, they are positively dependent. The type of dependence existing between the lifetimes will be carefully assessed in Section 3, expanding on previous results obtained by Denuit and Frostig (2007).

A standard approach to actuarial risk management consists of determining the amount of premium and the capital serving as a buffer against unexpected losses in accordance with a Value-at-Risk (VaR) principle. Specifically, the distribution of the present value of the insurer's payments is determined, and a high enough quantile is then calculated. Occasionally mixtures of VaRs (including Tail-VaR) may be preferable. This paper complies with this standard approach, and the results are phrased in terms of risk measures. The proofs are based on the theory of stochastic orderings, which are closely related to risk measures as explained in Denuit et al. (2005).

Section 4 considers the present value of future life annuity payments in a closed homogeneous portfolio. Often in practice, actuaries work with projected life tables as if they were known with certainty. This means that they enter the Lee-Carter point forecasts in their computations and neglect the correlation existing between the annuitants' lifetimes. Section 4 examines the consequences of this approach on the basis of numerical illustrations carried out on Belgian mortality statistics. In particular, we obtain the distribution function of the present value of life annuity payments in the Lee-Carter framework and compare it to the erroneous approaches consisting of either neglecting the positive correlation between the lifetimes or neglecting the randomness in the future death rates. Section 5 concludes.

2. STOCHASTIC MODELING FOR DYNAMIC MORTALITY

2.1 Notation and Assumption

We analyze the changes in mortality as a function of both age x and time t . Denote as $T_x(t)$ the remaining lifetime of an individual aged x on January 1 of year t ; this individual will die at age $x + T_x(t)$ in year $t + T_x(t)$. The mortality force at age x and time t , denoted $\mu_x(t)$, is defined as

$$\mu_x(t) = \lim_{\Delta \searrow 0} \frac{\Pr[x < T_0(t-x) \leq x + \Delta | T_0(t-x) > x]}{\Delta}.$$

Actuaries have traditionally calculated premiums and reserves using a deterministic mortality intensity. Here the forces of mortality will be described by a stochastic process as explained next.

2.2 Lee-Carter Model

Let us recall the basic features of the model proposed by Lee and Carter (1992). In this paper we assume that the age-specific mortality rates are constant within bands of age and time, but allowed to vary from one band to the next. Specifically, given any integer age x and calendar year t , it is assumed that given κ_t , the force of mortality at age $x + \xi$ and time $t + \tau$ is given by

$$\mu_{x+\xi}(t + \tau | \mathbf{\kappa}) = \mu_x(t | \mathbf{\kappa}) = \exp(\alpha_x + \beta_x \kappa_t) \quad \text{for } 0 \leq \xi, \tau < 1. \quad (2.1)$$

Interpretation of the parameters involved in the specification (2.1) is quite simple. When fitted by least-squares imposing that the κ_t 's be centered, the estimated value of α_x is an average of the empirical log death rates over time t so that $\exp \hat{\alpha}_x$ is the general shape of the mortality schedule. The actual forces of mortality change according to an overall mortality index κ_t modulated by an age response β_x . The shape of the β_x profile tells which rates decline rapidly and which slowly over time in response to change in κ_t . The time factor κ_t is intrinsically viewed as a stochastic process, and Box-Jenkins techniques are then used to model and forecast κ_t .

Throughout the paper we assume that all the β_x 's are positive. This is typically the case when the parameters are estimated from empirical mortality data. Positive β_x 's imply that the death rates are increasing in the κ_t 's. If the κ_t 's increase, then the life lengths tend to shorten.

To forecast, Lee and Carter (1992) assume that the α_x 's and β_x 's remain constant over time and forecast future values of κ_t using a standard univariate time series model. We will assume in this paper that the values $\kappa_1, \dots, \kappa_{t_0}$ are known but that the κ_{t_0+k} 's, $k = 1, 2, \dots$, are unknown and have to be projected from an ARIMA time series model. Bootstrap procedures can be used to account for estimation errors in the parameters; see Brouhns, Denuit, and Van Keilegom (2005).

3. IMPACT OF THE DEPENDENCE OF THE INSURED LIFETIMES

3.1 Exchangeable Lifetimes

In this paper we consider a group of n individuals, all aged x_0 at time t_0 . Denoting as ω the ultimate age of the life table, that is, the maximum life span, we need to consider the random vector $(\kappa_{t_0}, \dots, \kappa_{t_0+\omega-x_0})^T$ to describe the survival of all the members of the group. Henceforth, $\mathbf{\kappa}$ will denote the vector $(\kappa_{t_0}, \dots, \kappa_{t_0+\omega-x_0})^T$.

The remaining lifetimes of these n individuals are denoted as T_1, \dots, T_n all distributed as $T_{x_0}(t_0)$. Let us first consider $0 < \xi < 1$. Given $\mathbf{\kappa}$, the survival probability over $(0, \xi)$ can be written as

$$\begin{aligned} {}_{\xi}p_{x_0}(t_0 | \mathbf{\kappa}) &= \Pr[T_{x_0}(t_0) > \xi | \mathbf{\kappa}] \\ &= \exp(-\xi \mu_{x_0}(t_0 | \mathbf{\kappa})) \\ &= \exp(-\xi \exp(\alpha_{x_0} + \beta_{x_0} \kappa_{t_0})) \end{aligned}$$

under the assumption (2.1). Let us now consider $\xi > 1$, and let us denote as $\lfloor \xi \rfloor$ the integer part of ξ . The ξ -year conditional survival probability is given by

$$\begin{aligned}
{}_x p_{x_0}(\boldsymbol{\kappa}) &= \Pr[T_{x_0}(t_0) > \xi | \boldsymbol{\kappa}] \\
&= \left(\prod_{j=0}^{\lfloor \xi \rfloor - 1} p_{x_0+j}(t_0 + j | \boldsymbol{\kappa}) \right) {}_{\xi - \lfloor \xi \rfloor} p_{x_0 + \lfloor \xi \rfloor}(t_0 + \lfloor \xi \rfloor | \boldsymbol{\kappa}) \\
&= \left(\prod_{k=0}^{\lfloor \xi \rfloor - 1} \exp(-\mu_{x_0+k}(t_0 + k | \boldsymbol{\kappa})) \right) \\
&\quad \times \exp(-(\xi - \lfloor \xi \rfloor) \mu_{x_0 + \lfloor \xi \rfloor}(t_0 + \lfloor \xi \rfloor | \boldsymbol{\kappa})) \\
&= \exp\left(-\sum_{k=0}^{\lfloor \xi \rfloor - 1} \exp(\alpha_{x_0+k} + \beta_{x_0+k} \boldsymbol{\kappa}_{t_0+k})\right) \\
&\quad \times \exp(-(\xi - \lfloor \xi \rfloor) \exp(\alpha_{x_0 + \lfloor \xi \rfloor} + \beta_{x_0 + \lfloor \xi \rfloor} \boldsymbol{\kappa}_{t_0 + \lfloor \xi \rfloor})). \tag{3.1}
\end{aligned}$$

Unconditionally, we then have $\Pr[T_{x_0}(t_0) > \xi] = \mathbb{E}[{}_x p_{x_0}(\boldsymbol{\kappa})]$.

Note that ${}_x p_{x_0}(\boldsymbol{\kappa})$ is a random variable because it involves the random $\boldsymbol{\kappa}_{t_0+j}$'s. The dependence is induced by the sharing of the same $\boldsymbol{\kappa}$ by all the T_i 's. It is assumed to disappear if the path of the time index becomes known to the actuary. So the T_i 's are independent and identically distributed given $\boldsymbol{\kappa}$. Unconditionally, the lifetimes remain identically distributed but become dependent; precisely, the T_i 's are exchangeable random variables. This means that their joint distribution function does not factor as in the independent case but is invariant under permutation. Specifically, for any permutation $\{i_1, \dots, i_n\}$ of $\{1, \dots, n\}$, we have

$$\begin{aligned}
\Pr[T_1 \leq t_1, \dots, T_n \leq t_n] &= \Pr[T_1 \leq t_{i_1}, \dots, T_n \leq t_{i_n}] \\
&= \mathbb{E} \left[\prod_{i=1}^n \Pr[T_i \leq t_i | \boldsymbol{\kappa}] \right] \\
&= \mathbb{E} \left[\prod_{i=1}^n (1 - {}_{t_i} p_{x_0}(\boldsymbol{\kappa})) \right] \tag{3.2}
\end{aligned}$$

for any $t_1, \dots, t_n \geq 0$. The individuals with lifetimes T_1, \dots, T_n are thus "identical" but subject to the same time index $\boldsymbol{\kappa}$.

We expect that the T_i 's are positively related when they have joint distribution function (3.2). This is because the same time index influences the size of all the T_i 's. Denuit and Frostig (2007) examined some consequences of the type of dependence existing between the T_i 's. Let us continue this analysis here.

The Lee-Carter setting falls in the scope of Shaked and Spizzichino (1998). The lifetimes T_1, \dots, T_n with joint distribution function (3.2) are weakened by failure in the terminology used in reliability theory. This positive dependence notion is stronger than the concept of association used in Denuit and Frostig (2007) and has proven its usefulness in reliability theory. Translated in an actuarial context, it basically states that the death of a policyholder decreases the survival probability for the others. This illustrates the type of positive dependence existing between the lifetimes when the forces of mortality are described by (2.1). As the positive dependence of the T_i 's is inherited from the positive dependence between the components of the random vector $\boldsymbol{\kappa}$, it is natural to start with a careful analysis of the properties of $\boldsymbol{\kappa}$.

The vector $\boldsymbol{\kappa}$ is multivariate Normal when classical ARIMA time series models are used to describe the dynamics of the time index. The analysis of Denuit and Frostig (2007) is based on the fact that $\boldsymbol{\kappa}$ is associated, which means in the Normal case that the variance-covariance matrix of $\boldsymbol{\kappa}$ has all non-negative elements. In many cases $\boldsymbol{\kappa}$ is not only associated but even multivariate totally positive of order 2 (MTP2, in short), which means in the Normal case that all the off-diagonal components of the inverse of the variance-covariance matrix are nonpositive (see, e.g., Tong 1990, Theorem 4.3.2). The following example shows that MTP2 is fulfilled if $\boldsymbol{\kappa}$ obeys the standard random walk with drift model.

EXAMPLE 3.1

Assume that the κ_t 's obey

$$\kappa_t = \kappa_{t-1} + d + \xi_t, \tag{3.3}$$

where the ξ_t 's are independent and Normally distributed with mean 0 and variance σ^2 . In (3.3), d is known as the drift parameter. Then $\boldsymbol{\kappa}$ is multivariate Normal, with mean vector $\mathbf{m} = (\kappa_{t_0} + d, \kappa_{t_0} + 2d, \dots, \kappa_{t_0} + (\omega - x_0)d)^T$ and variance-covariance matrix $\boldsymbol{\Sigma} = \sigma^2 \mathbf{A}$, where \mathbf{A} is a square matrix of dimension $\omega - x_0$ with element ij given by $\min\{i,j\}$. It can be checked that the inverse of the matrix \mathbf{A} is given by

$$\mathbf{A}^{-1} = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix},$$

so that $\boldsymbol{\kappa}$ obeying to (3.3) is indeed MTP_2 .

Let $\mathbf{e} = (1, \dots, 1)^T$, where the length depends on the context. For $J = \{i_1, \dots, i_k\} \subset \{1, \dots, n\}$, let $\boldsymbol{\tau}_J$ denote $(\tau_{i_1}, \dots, \tau_{i_k})^T$ and let the complement of J be denoted as \bar{J} . Let h_t denote a realization of the T_i 's up to time t , that is, an event of the form

$$h_t = \{T_J = \boldsymbol{\tau}_J, T_{\bar{J}} > t\mathbf{e}\}, \tag{3.4}$$

where $\mathbf{0} \leq \boldsymbol{\tau}_J \leq t\mathbf{e}$. In words, h_t means that the policyholders in J died before t , at times $\tau_{i_1}, \dots, \tau_{i_k}$, respectively, and that the policyholders in \bar{J} are still alive at time t . Now, let us fix $J, i \in \bar{J}, t \geq 0, \boldsymbol{\tau}_J \leq t\mathbf{e}$ such that $(\boldsymbol{\tau}_J, \tau_i)$ is in the support of (T_J, T_i) and consider

$$h'_t = \{T_J = \boldsymbol{\tau}_J, T_i = t \text{ and } T_{\bar{J} \setminus \{i\}} > t\mathbf{e}\}. \tag{3.5}$$

In words, the policyholders in J died at times $\tau_{i_1}, \dots, \tau_{i_k}$ under both h_t and h'_t . Those in \bar{J} are still alive at time t under h_t , whereas policyholder i dies at time t and those in $\bar{J} \setminus \{i\}$ outlive him under h'_t . The only difference between h_t and h'_t is thus that one of the survivors in h_t dies at time t under h'_t .

We are now in a position to formalize the idea that $\boldsymbol{\kappa}$ gets larger under h'_t relative to h_t . Given a random vector S and an event A , let us denote as $\mathcal{L}(S|A)$ the conditional distribution of S given A . Recall the definition of the likelihood ratio order, which is used here to express the fact that $\boldsymbol{\kappa}$ gets "larger." Given two n -dimensional random vectors X and Y , with probability density functions f_X and f_Y , respectively, X is said to be smaller than Y in the likelihood ratio order, written as $X \preceq_{LR} Y$, if the inequality

$$f_X(\mathbf{x})f_Y(\mathbf{y}) \leq f_X(\mathbf{x} \wedge \mathbf{y})f_Y(\mathbf{x} \vee \mathbf{y}) \tag{3.6}$$

holds for all \mathbf{x} and \mathbf{y} in \mathbb{R}^n , where the lattice operators \vee and \wedge are defined as

$$\mathbf{x} \vee \mathbf{y} = (\max\{x_1, y_1\}, \dots, \max\{x_n, y_n\})$$

and

$$\mathbf{x} \wedge \mathbf{y} = (\min\{x_1, y_1\}, \dots, \min\{x_n, y_n\}).$$

Property 3.2

Consider the events h_t and h'_t defined in (3.4) and (3.5), and assume that $\boldsymbol{\kappa}$ is MTP_2 . Then we have $\mathcal{L}(\boldsymbol{\kappa}|h_t) \preceq_{LR} \mathcal{L}(\boldsymbol{\kappa}|h'_t)$.

PROOF

Let $G(\cdot)$ and $g(\cdot)$ denote the distribution function and the probability density function of $\mathbf{\kappa}$, respectively. Similarly, let us denote as $G(\cdot|h_t)$ and $g(\cdot|h_t)$ the conditional distribution function and the conditional probability density function of $\mathbf{\kappa}$ given h_t , respectively. We then have

$$g(\mathbf{\kappa}|h_t) = \frac{g(\mathbf{\kappa})}{c(h_t)} \left(\prod_{j \in \bar{J}} \Pr[T_j \geq \tau_j | \mathbf{\kappa}] \mu_{x_0 + \lfloor \tau_j \rfloor}(t_0 + \lfloor \tau_j \rfloor | \mathbf{\kappa}) \right) \left(\prod_{j \in \bar{J}} \Pr[T_j \geq t | \mathbf{\kappa}] \right),$$

where

$$c(h_t) = \int \cdots \int \left(\prod_{j \in \bar{J}} \Pr[T_j \geq \tau_j | \xi] \mu_{x_0 + \lfloor \tau_j \rfloor}(t_0 + \lfloor \tau_j \rfloor | \xi) \right) \left(\prod_{j \in \bar{J}} \Pr[T_j \geq t | \xi] \right) g(\xi) d\xi.$$

Similarly,

$$g(\mathbf{\kappa}|h'_t) = \frac{g(\mathbf{\kappa})}{c(h'_t)} \left(\prod_{j \in \bar{J}} \Pr[T_j \geq \tau_j | \mathbf{\kappa}] \mu_{x_0 + \lfloor \tau_j \rfloor}(t_0 + \lfloor \tau_j \rfloor | \mathbf{\kappa}) \right) \times \Pr[T_i \geq t | \mathbf{\kappa}] \mu_{x_0 + \lfloor t \rfloor}(t_0 + \lfloor t \rfloor | \mathbf{\kappa}) \left(\prod_{j \in \bar{J} \setminus \{i\}} \Pr[T_j \geq t | \mathbf{\kappa}] \right)$$

with

$$c(h'_t) = \int \cdots \int \left(\prod_{j \in \bar{J}} \Pr[T_j \geq \tau_j | \xi] \mu_{x_0 + \lfloor \tau_j \rfloor}(t_0 + \lfloor \tau_j \rfloor | \xi) \right) \Pr[T_i \geq t | \xi] \mu_{x_0 + \lfloor t \rfloor}(t_0 + \lfloor t \rfloor | \xi) \left(\prod_{j \in \bar{J} \setminus \{i\}} \Pr[T_j \geq t | \xi] \right) g(\xi) d\xi.$$

We thus have

$$g(\mathbf{\kappa}|h'_t) = \frac{c(h_t)}{c(h'_t)} g(\mathbf{\kappa}|h_t) \mu_{x_0 + \lfloor t \rfloor}(t_0 + \lfloor t \rfloor | \mathbf{\kappa}).$$

Provided $\mathbf{\kappa}$ is MTP_2 , $g(\cdot|h_t)$ and $g(\cdot|h'_t)$ are both MTP_2 functions because the conditional survival functions and probability density functions of the T_i 's are products of univariate functions involving a single time index. Considering two vectors $\mathbf{\kappa}'$ and $\mathbf{\kappa}''$, we can then write

$$\begin{aligned} &g(\mathbf{\kappa}' \vee \mathbf{\kappa}''|h'_t) g(\mathbf{\kappa}' \wedge \mathbf{\kappa}''|h_t) \\ &= \frac{c(h_t)}{c(h'_t)} g(\mathbf{\kappa}' \vee \mathbf{\kappa}''|h_t) g(\mathbf{\kappa}' \wedge \mathbf{\kappa}''|h_t) \mu_{x_0 + \lfloor t \rfloor}(t_0 + \lfloor t \rfloor | \mathbf{\kappa}' \vee \mathbf{\kappa}'') \\ &\geq \frac{c(h_t)}{c(h'_t)} g(\mathbf{\kappa}'|h_t) g(\mathbf{\kappa}''|h_t) \mu_{x_0 + \lfloor t \rfloor}(t_0 + \lfloor t \rfloor | \mathbf{\kappa}' \vee \mathbf{\kappa}'') \\ &\geq \frac{c(h_t)}{c(h'_t)} g(\mathbf{\kappa}'|h_t) g(\mathbf{\kappa}''|h_t) \mu_{x_0 + \lfloor t \rfloor}(t_0 + \lfloor t \rfloor | \mathbf{\kappa}') \\ &= g(\mathbf{\kappa}'|h'_t) g(\mathbf{\kappa}''|h_t), \end{aligned}$$

where the last inequality comes from the fact that the forces of mortality increase with $\mathbf{\kappa}$. □

Property 3.2 illustrates how the time index $\mathbf{\kappa}$ reacts to the history of the death occurrences. The updated $\mathbf{\kappa}$ is then integrated into the posterior distribution of the survival times in a Bayesian way.

Now that we have shown that the time index was monotone in the timing of deaths (in the \preceq_{LR} -sense), it seems reasonable to believe that the remaining lifetimes become shorter as deaths occur

prematurely. Indeed, early deaths suggest large κ_t 's, which in turn tend to decrease lifetimes. The next result makes this intuition precise, by means of stochastic dominance. Recall the following definition for stochastic dominance \preceq_{ST} : given two random vectors X and Y , X is said to be smaller than Y in the stochastic dominance, written as $X \preceq_{ST} Y$, if $\mathbb{E}[\Psi(X)] \leq \mathbb{E}[\Psi(Y)]$ for every nondecreasing function $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$ provided the expectations exist. It can be shown that $X \preceq_{LR} Y \Rightarrow X \preceq_{ST} Y$.

We are now in a position to state the following result.

Property 3.3

Assume that κ is MTP_2 and consider the events h_t and h'_t defined in (3.4) and (3.5). Then we have

$$\mathcal{L}(T_{\bar{j}\setminus\{i\}} - te|h'_t) \preceq_{ST} \mathcal{L}(T_{\bar{j}\setminus\{i\}} - te|h_t).$$

PROOF

Consider a nondecreasing function Ψ and write

$$\mathbb{E}[\Psi(T_{\bar{j}\setminus\{i\}} - te)|h_t] = \int \cdots \int \mathbb{E}[\Psi(T_{\bar{j}\setminus\{i\}} - te)|\kappa]g(\kappa|h_t) d\kappa.$$

Because $T_{\bar{j}\setminus\{i\}}$ has independent components given κ , and is decreasing in the \preceq_{ST} -sense in κ , the function $\kappa \mapsto \Psi^*(\kappa) = \mathbb{E}[\Psi(T_{\bar{j}\setminus\{i\}} - te)|\kappa]$ is nonincreasing. Hence, because $\mathcal{L}(\kappa|h_t) \preceq_{LR} \mathcal{L}(\kappa|h'_t) \Rightarrow \mathcal{L}(\kappa|h_t) \preceq_{ST} \mathcal{L}(\kappa|h'_t)$, we have that

$$\mathbb{E}[\Psi(T_{\bar{j}\setminus\{i\}} - te)|h_t] = \mathbb{E}[\Psi^*(\kappa)|h'_t] \leq \mathbb{E}[\Psi^*(\kappa)|h_t] = \mathbb{E}[\Psi(T_{\bar{j}\setminus\{i\}} - te)|h'_t],$$

whence the announced result follows. □

Property 3.3 shows that the conditional residual lives of the survivors at any time $t \geq 0$ are \preceq_{ST} -ordered, given two up-to-time- t realizations that are identical except that in one realization there is a death at time t , whereas there is no death at that time in the other realization. Lifetimes possessing this feature are said to be weakened by failure in the literature. Because the financial obligations of the insurance company increase in the remaining lifetimes in the context of life annuities, this result gives the conditions under which the payments to annuitants increase in the \preceq_{ST} -sense.

3.2 Impact on Risk Capital

Now compare the amount of capital needed for a portfolio with lifetimes obeying the Lee-Carter model to that of another portfolio made of independent lifetimes with the same marginal distributions. The riskiness is measured here as the amount of capital needed in each case, as determined with the help of a risk measure.

Consider a risk X with distribution function F_X such that $F_X(0-) = 0$. We consider value-at-risk

$$\text{VaR}[X; \alpha] = F_X^{-1}(\alpha) = \inf\{x \in \mathbb{R} | F_X(x) \geq \alpha\}, \quad 0 < \alpha < 1,$$

and mixtures of VaR's (also called Yaari-Wang risk measures) of the form

$$\rho_g[X] = \int_0^{+\infty} g(\text{Pr}[X > t]) dt = \int_0^1 \text{VaR}[X; 1 - \varepsilon] dg(\varepsilon)$$

with the distortion function g nondecreasing with $g(0) = 0$ and $g(1) = 1$. Tail-VaR's defined as

$$\text{TVaR}[X; p] = \frac{1}{1-p} \int_p^1 \text{VaR}[X; \varepsilon] d\varepsilon$$

are typical examples of mixtures of VaR's.

Assume that each policyholder has a life annuity contract that pays one monetary unit at the end of each year provided he or she is still alive at that time. Let us denote as $\alpha_{\bar{q}}$ the present value of a sequence of $[\xi]$ unit payments made at times $1, \dots, [\xi]$, that is,

$$\alpha_{\overline{a}|} = \sum_{j=1}^{\lfloor \xi \rfloor} v(0, j),$$

where the $v(0, j)$'s are the discount factors: that is, $v(0, j)$ is the present value at time 0 of a unit payment made at time j . Then $\alpha_{\overline{a}|}$ is the present value of the payments made to annuitant i . Let T_1, \dots, T_n have joint distribution function (3.2) and let

$$V = \sum_{i=1}^n \alpha_{\overline{a}|}$$

be the present value of all the payments made by the insurer to the n annuitants.

Let us denote as $T_1^\perp, \dots, T_n^\perp$ the independent version of the Lee-Carter lifetimes T_1, \dots, T_n . Formally, $T_1^\perp, \dots, T_n^\perp$ have joint distribution function

$$\begin{aligned} \Pr[T_1^\perp \leq t_1, \dots, T_n^\perp \leq t_n] &= \prod_{i=1}^n \Pr[T_i \leq t_i] \\ &= \prod_{i=1}^n (1 - \mathbb{E}[{}_i p_{x_0}(\boldsymbol{\kappa})]) \end{aligned} \quad (3.7)$$

for any $t_1, \dots, t_n \geq 0$. Furthermore, define V^\perp as the present value of all the payments made to the n annuitants assuming that their respective lifetimes are mutually independent. We know from Property 4.5 in Denuit and Frostig (2007) that $\rho_g[V^\perp] \leq \rho_g[V]$ for any concave distortion g . This means that V^\perp precedes V in the stop-loss order, that is, that the stop-loss premiums $\mathbb{E}[(V^\perp - t)_+]$ are smaller than the corresponding $\mathbb{E}[(V - t)_+]$ whatever t . We see that the positive dependence induced by the projected life tables increases the need for economic capital when this capital is computed with the help of ρ_g with a concave g compared to a portfolio comprising independent lives. Note that the inequality remains valid for any risk measure agreeing with stop-loss order.

The first three moments of V are derived in the Appendix. They can be used to perform moment-based approximations (like the NP formula, for instance) to get a first idea of the behavior of V . In particular, we have $\mathbb{E}[V^\perp] = \mathbb{E}[V]$ and $\text{Var}[V^\perp] \leq \text{Var}[V]$.

3.3 Monitoring Experience

Assume now that we have observed the T_i 's for t years (from year t_0 to year $t_0 + t$, say). We have observed a number of deaths (and the corresponding ages at death), and we know how many individuals are still alive out of the n initial ones. Based on this information, we would like to revise the distribution of the κ_{t_0+k} 's.

Think, for instance, that there remain more survivors than expected from the a priori distribution of the time index. This suggests that the actual κ_{t_0+k} 's may be smaller than anticipated. This should be taken into account to revise the amounts of the reserves needed to cover future annuity payments (which should be increased). The information available at time $t_0 + t$ can be included in the actuarial applications by using the a posteriori distribution of the future κ_{t_0+t+k} 's given past information. The actual computation of this a posteriori distribution requires numerical methods, but the present section shows that the a posteriori distribution possesses some monotonicity property with respect to the timing of deaths.

Recall the definition of the events h_t and h'_t from (3.4) and (3.5). Intuitively speaking, h'_t is a worse survival history than h_t so that we expect that $\boldsymbol{\kappa}$ becomes "larger" under h'_t relative to h_t making the lifetimes shorter and reducing the liabilities of the insurance company. This is formalized in the next result.

Property 3.4

Assume that $\boldsymbol{\kappa}$ is MTP_2 and consider the events h_t and h'_t defined in (3.4) and (3.5). Let T_1, \dots, T_n have joint distribution function (3.2). Then

$$\rho_g \left[\sum_{k \in \bar{J} \setminus \{i\}} \alpha_{T_k-i} | h'_t \right] \leq \rho_g \left[\sum_{k \in \bar{J} \setminus \{i\}} \alpha_{T_k-i} | h_t \right]$$

for any distortion g , where $\rho_g[X|h_t]$ stands for the risk measure computed with the conditional distribution of X given h_t .

Property 3.4 directly follows from Property 3.3 because whatever the risks X and Y , $X \leq_{ST} Y \Rightarrow \rho_g[X] \leq \rho_g[Y]$ for any distortion g .

4. NUMERICAL ILLUSTRATION

Here we use the mortality projections built in Pitacco et al. (2009, chapter 5). We evaluate the common marginal survival function $\xi \rightarrow \mathbb{E}[\xi p_{65}(\mathbf{\kappa})]$ of the T_i 's by simulation. Specifically, 10,000 trajectories of the κ_t 's are randomly generated, and the resulting $\xi p_{65}(\mathbf{\kappa})$ are averaged to get the expectation $\mathbb{E}[\xi p_{65}(\mathbf{\kappa})]$. The histograms of the simulated values for the $\xi p_{65}(\mathbf{\kappa})$'s, $\xi = 5, 10, 15, \dots, 30$, are given in Figure 1. Figure 2 displays the survival function $\xi \mapsto \mathbb{E}[\xi p_{65}(\mathbf{\kappa})]$ with $\omega = 104$, together with the 5% and 95% quantiles of $\xi p_{65}(\mathbf{\kappa})$.

Let us now work with a constant annual interest rate equal to 4%. For $n = 100$, we get $\mathbb{E}[V] = 1,166.08$ and $\text{Var}[V] = 2,342.72$ using the formulas worked out in the Appendix. The upper bound for V is 1,979.28, and the skewness coefficient is $\gamma[V] = -0.06001539$.

Increasing n to 1,000 gives $\mathbb{E}[V] = 11,660.84$, $\text{Var}[V] = 49,455.71$. For the upper bound and the mean value, it is just 10 times the figures obtained for $n = 100$. Because of the positive dependence existing between the T_i 's, the variance for $n = 1,000$ is more than 10 times the variance obtained with $n = 100$. The skewness is $\gamma[V] = -0.02543851$ closer to 0 than for $n = 100$.

Because the exact computation of the distribution function of V may be time consuming, actuaries sometimes simplify the calculations by assuming that the lifetimes are (1) mutually independent and (2) subject to some deterministic death rates of the form

$$\mu_{x_0+k}^{\text{det}}(t_0 + k) = \exp(\alpha_{x_0+k} + \beta_{x_0+k} \mathbb{E}[\kappa_{t_0+k}])$$

obtained by replacing κ_{t_0+k} with its expected value. This means that $\mu_{x_0+k}^{\text{det}}(t_0 + k)$ is the median of the stochastic death rate $\mu_{x_0+k}(t_0 + k | \mathbf{\kappa})$. The use of the median death rates to generate deterministic mortality forecast was originally proposed by Lee and Carter (1992).

REMARK 4.1

Deterministic projected life tables could also be obtained by replacing $\mu_{x_0+k}(t_0 + k | \mathbf{\kappa})$ with its expected value

$$\mathbb{E}[\mu_{x_0+k}(t_0 + k | \mathbf{\kappa})] = \exp \left(\alpha_{x_0+k} + \beta_{x_0+k} \mathbb{E}[\kappa_{x_0+k}] + \frac{\beta_{x_0+k}^2}{2} \text{Var}[\kappa_{t_0+k}] \right) > \mu_{x_0+k}^{\text{det}}(t_0 + k). \quad (4.1)$$

However, this natural approach consisting of using mean death rates as predictions biases the actuarial computations, as shown next. Let $T_{x_0}(t_0)$ be as defined in Section 2.1, and let $\tilde{T}_{x_0}(t_0)$ have piecewise constant death rates $\mathbb{E}[\mu_{x_0+k}(t_0 + k | \mathbf{\kappa})]$, $k = 1, 2, \dots$. Then $\tilde{T}_{x_0}(t_0) \leq_{ST} T_{x_0}(t_0)$ when the κ_{t_0+k} 's are associated. In this case the expectation of the product appearing in the expression of $\xi p_{x_0}(\mathbf{\kappa})$ is larger than the product of the corresponding expectations, that is,

$$\begin{aligned} \Pr[T_{x_0}(t_0) > \xi] &\geq \prod_{k=0}^{\lfloor \xi \rfloor - 1} \mathbb{E}[\exp(-\exp(\alpha_{x_0+k} + \beta_{x_0+k} \kappa_{t_0+k}))] \\ &\times \mathbb{E}[\exp(-(\xi - \lfloor \xi \rfloor) \exp(\alpha_{x_0+\lfloor \xi \rfloor} + \beta_{x_0+\lfloor \xi \rfloor} \kappa_{t_0+\lfloor \xi \rfloor}))]. \end{aligned}$$

Figure 1
Histograms of 10,000 simulated values for ${}_{\xi}p_{65}(\kappa)$'s, $\xi = 5, 10, 15, \dots, 30$

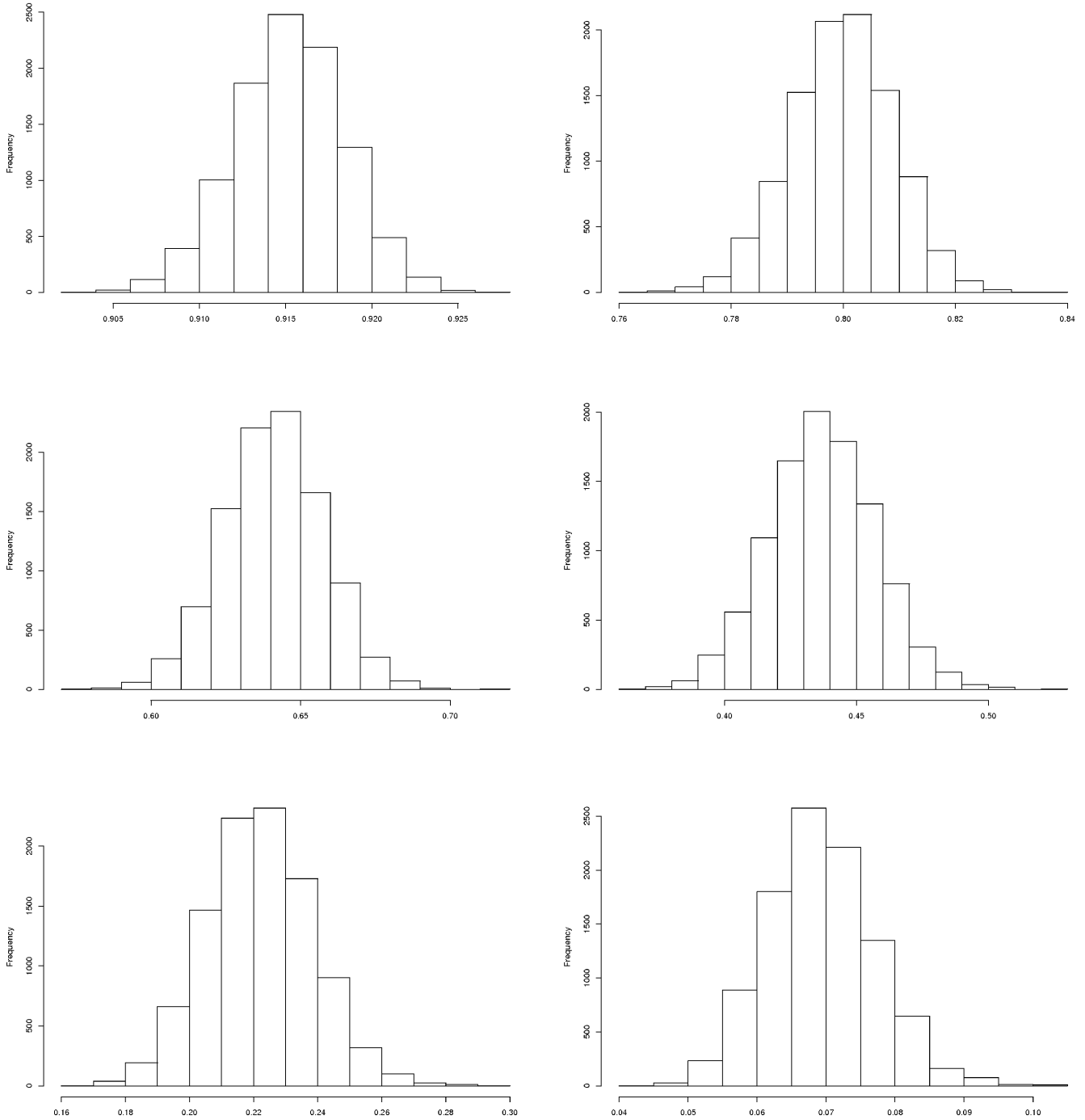
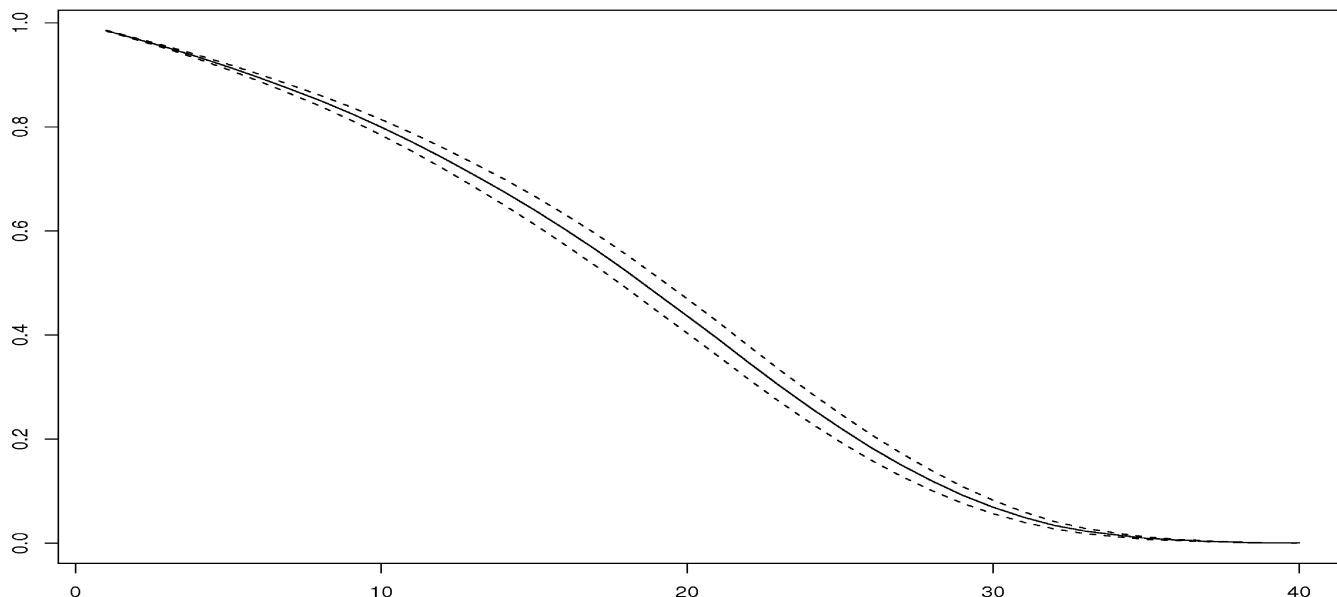


Figure 2
Graph of Survival Function $\xi \mapsto \mathbb{E}[\xi p_{65}(\kappa)]$ (Solid Line), Together with 5% and 95% Quantiles of $\xi p_{65}(\kappa)$ (Dotted Lines), Based on 10,000 Simulations



By Jensen’s inequality we get

$$\begin{aligned} \Pr[T_{x_0}(t_0) > \xi] &\geq \prod_{k=0}^{\lfloor \xi \rfloor - 1} \exp(-\mathbb{E}[\exp(\alpha_{x_0+k} + \beta_{x_0+k} \kappa_{t_0+k})]) \\ &\quad \times \exp(-(\xi - \lfloor \xi \rfloor) \mathbb{E}[\exp(\alpha_{x_0+\lfloor \xi \rfloor} + \beta_{x_0+\lfloor \xi \rfloor} \kappa_{t_0+\lfloor \xi \rfloor})]) \\ &= \Pr[\tilde{T}_{x_0}(t_0) > \xi]. \end{aligned}$$

In words, working with $\mathbb{E}[\mu_{x_0+k}(t_0 + k|\kappa)]$ in lieu of $\mu_{x_0+k}(t_0 + k|\kappa)$ makes the lifetimes artificially shorter. Using expected death rates thus shortens the annuitants’ lifetimes. This procedure has an obvious effect on the risk capital.

We know from Section 3.2 that approximation (1) is not conservative for computing risk measures ρ_g with g concave (or, more generally, when the quantity of interest is monotonic with respect to the stop-loss order). Let us now demonstrate that approximation (1) may have a huge impact on the VaR, especially in large portfolios, whereas approximation (2) has little consequences. To this end, let us quantify the impact of ignoring the dependence that arises from the Lee-Carter model. Specifically, we compute the distribution function of V under the following scenarios:

SCENARIO 1

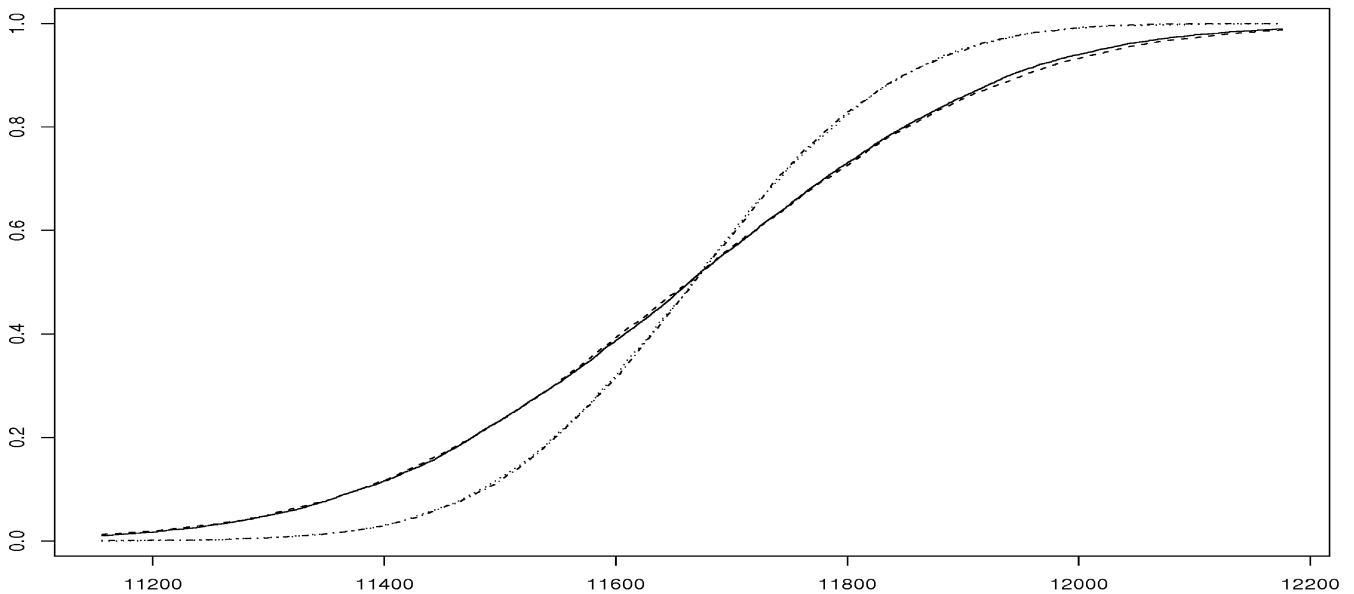
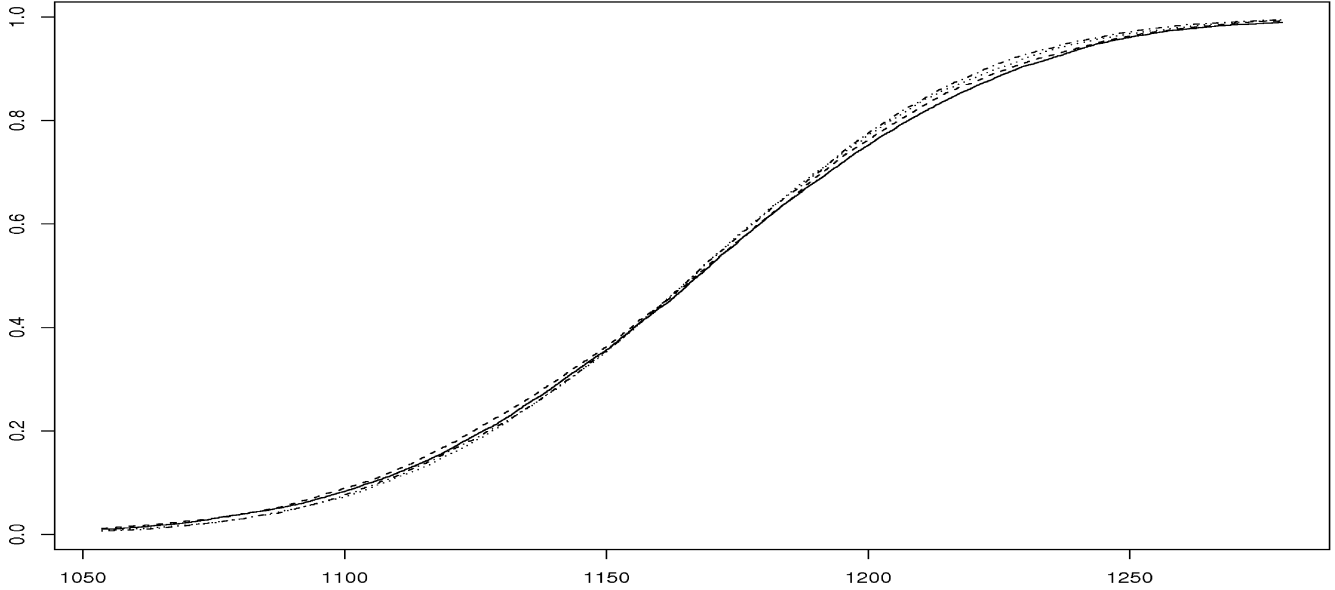
The lifetimes T_1, \dots, T_n have joint distribution function (3.2): that is, the lifetimes are conditionally independent given κ with common ξ -year conditional survival probability $g_{p_{x_0}}(\kappa)$. This complies with the assumptions of the Lee-Carter model.

SCENARIO 2

The lifetimes have the copula corresponding to the joint distribution function (3.2) but common ξ -year survival probability corresponding to the deterministic death rates $\mu_{x_0+k}^{\det}(t_0 + k)$. More precisely, the copula C corresponding to the joint distribution function (3.2) is given by

Figure 3

Distribution Function of V with $n = 100$ (Top Panel) and $n = 1,000$ (Bottom Panel) Displayed in a Solid Line for Scenario 1, Together with the Distribution Functions Computed under Scenario 2 (Broken Line), Scenario 3 (Dotted Line), and Scenario 4 (Dashed-Dotted Line)



$$C(u_1, \dots, u_n) = \mathbb{E} \left[\prod_{i=1}^n (1 - F^{-1}(u_i) p_{x_0}(\boldsymbol{\kappa})) \right] \quad \text{with } u_1, \dots, u_n \in [0, 1].$$

where F is the common marginal distribution function of the lifetimes. Under Scenario 2 the lifetimes have joint distribution function

$$t \mapsto C(G(t_1), \dots, G(t_n)) \quad \text{where } G(t) = 1 - \exp \left(- \int_0^t \mu_{x_0+\tau}^{\text{det}}(t_0 + \tau) d\tau \right).$$

Compared to Scenario 1, the lifetimes have the same dependence structure (i.e., the same copula), but the marginal behavior is modified. Instead of the true marginal distribution function $\xi \mapsto 1 - \mathbb{E}[{}_t p_{x_0}(\boldsymbol{\kappa})]$, we use the approximate distribution G inherited from the pointwise mortality projection obtained with the Lee-Carter model (i.e., using the deterministic death rates $\mu_{x_0+k}^{\text{det}}(t_0 + k)$).

SCENARIO 3

The lifetimes T_1^+, \dots, T_n^+ have the joint distribution function (3.7): that is, the lifetimes are independent with common ξ -year survival probability $\mathbb{E}[{}_t p_{x_0}(\boldsymbol{\kappa})]$. Compared with the assumptions underlying the Lee-Carter model, the actuary now neglects the positive dependence between the lifetimes, but accounts for the uncertainty in the future death rates (as the survival probability is obtained by averaging over all the future paths of the time index $\boldsymbol{\kappa}$).

Scenario 3 keeps the exact marginal distribution from Scenario 1 but replaces the copula C used in Scenarios 1–2 with mutual independence. The positive dependence between the T_i 's is thus disregarded.

SCENARIO 4

The lifetimes are independent with common ξ -year survival probability corresponding to the deterministic death rates $\mu_{x_0+k}^{\text{det}}(t_0 + k)$. Compared with the assumptions underlying the Lee-Carter model, the actuary now neglects both the positive dependence between the lifetimes and the uncertainty in the future death rates.

Actuaries often work under Scenario 4, using deterministic projected life tables furnished by professional bodies or regulatory authorities as if they were known with certainty.

The resulting distribution functions for V are displayed in Figure 3. We see that the positive dependence does impact the tails of the distribution function, with comparatively higher quantiles under Scenarios 1–2. The impact of positive dependence is much higher in a large portfolio: for $n = 1,000$, the distribution functions obtained with Scenarios 1–2 greatly differ from those corresponding to Scenarios 3–4, whereas they are much closer for $n = 100$. This shows that the systematic risk is comparatively more important in a larger portfolio. The results obtained under Scenarios 1–2 are very close. The same comment applies to the results obtained under Scenarios 3–4. Hence, the difference essentially comes from the copula and is less affected by the marginal behavior. To sum up, the computations based on a set of deterministic projected life tables (or reduction factors) corresponding to Scenarios 3–4 may greatly underestimate the insurer's liabilities in large portfolios.

5. CONCLUSION

This paper discusses some of the consequences of the use of projected life tables. In the Lee-Carter framework, a single time index drives mortality forecasts. This index generally obeys an ARIMA time series model. The randomness in the future trajectory of this process induces dependence between policyholders' remaining lifetimes, which are no more independent but conditionally independent, given the time index.

Our contribution in this paper is twofold:

1. We make precise the type of dependence existing between the lifetimes subject to the Lee-Carter model, and we derive several consequences, namely, credibility updating formulas.
2. Because neglecting this positive dependence is not conservative, we perform numerical illustrations to quantify the riskiness of the approach commonly adopted by actuaries in practice (Scenario 4 in Section 4).

It has been shown that the actuary should account for the positive dependence induced by the common sharing of the time index κ , because failing to do so leads to underestimation of risk measures ρ_g with g concave. Also, numerical illustrations based on Belgian mortality statistics show that ignoring the positive dependence may result in underevaluation of high quantiles.

APPENDIX

DERIVATION OF THE MOMENTS OF V

In this appendix, we provide explicit expressions for the first three moments of V (allowing for NP approximations, for instance). As the ruin probability is obtained from the present value of all the payments made to the annuitants up to time j , we even compute the moments of V_j , defined as

$$V_j = \sum_{i=1}^n \alpha_{\min\{T_i, j\}},$$

where $\alpha_{\min\{T_i, j\}}$ is the random present value of a sequence of unit payments made at times $1, 2, \dots, j$ provided policyholder i is still alive at the payment dates.

A.1 Mean of V_j

Let us decompose V_j as $V_j = \sum_{i=1}^n V_{ij}$, where the contribution V_{ij} of policyholder i to V_j is the present value of all the payments made to the i th annuitant up to time j . Denoting as $\mathbb{1}[T_i > h]$ the indicator of the event $T_i > h$, that is,

$$\mathbb{1}[T_i > h] = \begin{cases} 1, & \text{if policyholder } i \text{ is still alive at time } h, \\ 0, & \text{otherwise,} \end{cases}$$

then,

$$V_{ij} = \sum_{\ell=1}^j \mathbb{1}[T_i > \ell] v(0, \ell).$$

The mean of V_j is then given by

$$\mathbb{E}[V_j] = \sum_{i=1}^n \mathbb{E}[V_{ij}] = \sum_{i=1}^n \sum_{\ell=1}^j v(0, \ell) \mathbb{E}[\ell p_{x_0}(\kappa)].$$

A.2 Variance of V_j

The variance of V_j writes

$$\text{Var}[V_j] = \sum_{i=1}^n \text{Var}[V_{ij}] + 2 \sum_{i=1}^{n-1} \sum_{k=i+1}^n \text{Cov}[V_{ij}, V_{kj}].$$

The variance of V_{ij} is obtained from

$$\begin{aligned}\mathbb{E}[V_{ij}^2] &= \mathbb{E} \left[\sum_{\ell=1}^j \mathbb{I}[T_i > \ell] \mathfrak{v}(0, \ell) \sum_{m=1}^j \mathbb{I}[T_i > m] \mathfrak{v}(0, m) \right] \\ &= \sum_{\ell=1}^j \sum_{m=1}^j \mathfrak{v}(0, \ell) \mathfrak{v}(0, m) \mathbb{E}[\mathbb{I}[T_i > \max\{\ell, m\}]] \\ &= \sum_{\ell=1}^j (\mathfrak{v}(0, \ell))^2 \mathbb{E}[_{\ell}p_{x_0}(\boldsymbol{\kappa})] + 2 \sum_{\ell=1}^{j-1} \sum_{m=\ell+1}^j \mathfrak{v}(0, \ell) \mathfrak{v}(0, m) \mathbb{E}[_m p_{x_0}(\boldsymbol{\kappa})]\end{aligned}$$

and

$$\begin{aligned}(\mathbb{E}[V_{ij}])^2 &= \sum_{\ell=1}^j \mathfrak{v}(0, \ell) \mathbb{E}[_{\ell}p_{x_0}(\boldsymbol{\kappa})] \sum_{m=1}^j \mathfrak{v}(0, m) \mathbb{E}[_m p_{x_0}(\boldsymbol{\kappa})] \\ &= \sum_{\ell=1}^j (\mathfrak{v}(0, \ell))^2 (\mathbb{E}[_{\ell}p_{x_0}(\boldsymbol{\kappa})])^2 + 2 \sum_{\ell=1}^{j-1} \sum_{m=\ell+1}^j \mathfrak{v}(0, \ell) \mathfrak{v}(0, m) \mathbb{E}[_{\ell}p_{x_0}(\boldsymbol{\kappa})] \mathbb{E}[_m p_{x_0}(\boldsymbol{\kappa})]\end{aligned}$$

so that

$$\begin{aligned}\text{Var}[V_{ij}] &= \sum_{\ell=1}^j (\mathfrak{v}(0, \ell))^2 \mathbb{E}[_{\ell}p_{x_0}(\boldsymbol{\kappa})] (1 - \mathbb{E}[_{\ell}p_{x_0}(\boldsymbol{\kappa})]) \\ &\quad + 2 \sum_{\ell=1}^{j-1} \sum_{m=\ell+1}^j \mathfrak{v}(0, \ell) \mathfrak{v}(0, m) \mathbb{E}[_m p_{x_0}(\boldsymbol{\kappa})] (1 - \mathbb{E}[_{\ell}p_{x_0}(\boldsymbol{\kappa})]).\end{aligned}$$

Now, for $k \neq j$, the covariance between V_{ij} and V_{kj} is obtained from

$$\begin{aligned}\text{Cov}[V_{ij}, V_{kj}] &= \text{Cov} \left[\sum_{\ell=1}^j \mathbb{I}[T_i > \ell] \mathfrak{v}(0, \ell), \sum_{m=1}^j \mathbb{I}[T_k > m] \mathfrak{v}(0, m) \right] \\ &= \sum_{\ell=1}^j \sum_{m=1}^j \mathfrak{v}(0, \ell) \mathfrak{v}(0, m) \text{Cov}[\mathbb{I}[T_i > \ell], \mathbb{I}[T_k > m]] \\ &= \sum_{\ell=1}^j \sum_{m=1}^j \mathfrak{v}(0, \ell) \mathfrak{v}(0, m) \text{Cov}[_{\ell}p_{x_0}(\boldsymbol{\kappa}), _m p_{x_0}(\boldsymbol{\kappa})]\end{aligned}$$

because

$$\begin{aligned}\text{Cov}[\mathbb{I}[T_i > \ell], \mathbb{I}[T_k > m]] &= \mathbb{E}[\text{Cov}[\mathbb{I}[T_i > \ell], \mathbb{I}[T_k > m] | \boldsymbol{\kappa}]] \\ &\quad + \text{Cov}[\mathbb{E}[\mathbb{I}[T_i > \ell] | \boldsymbol{\kappa}], \mathbb{E}[\mathbb{I}[T_k > m] | \boldsymbol{\kappa}]] \\ &= \text{Cov}[_{\ell}p_{x_0}(\boldsymbol{\kappa}), _m p_{x_0}(\boldsymbol{\kappa})],\end{aligned}$$

where the first term disappeared invoking the conditional independence of T_i and T_k given $\boldsymbol{\kappa}$.

A.3 Skewness of V_j

The skewness coefficient of V_j is defined as $\mathbb{E}[(V_j - \mathbb{E}[V_j])^3] / (\text{Var}[V_j])^{3/2}$. Clearly,

$$\mathbb{E}[(V_j - \mathbb{E}[V_j])^3] = \mathbb{E}[V_j^3] - 3\mathbb{E}[V_j]\mathbb{E}[V_j^2] + 2(\mathbb{E}[V_j])^3.$$

Now,

$$\begin{aligned}\mathbb{E}[V_j^3] &= \mathbb{E} \left[\left(\sum_{i=1}^n V_{ij} \right)^3 \right] \\ &= \sum_{i=1}^n \mathbb{E}[V_{ij}^3] + 3 \sum_{i < h} \mathbb{E}[V_{ij}^2 V_{hj}] + 3 \sum_{i < h} \mathbb{E}[V_{ij} V_{hj}^2] + 6 \sum_{i < h < k} \mathbb{E}[V_{ij} V_{hj} V_{kj}].\end{aligned}$$

Raising V_{ij} to the third power and taking the expectation gives

$$\begin{aligned}\mathbb{E}[V_{ij}^3] &= \mathbb{E} \left[\left(\sum_{\ell=1}^j \mathbb{I}[T_i > \ell] v(0, \ell) \right)^3 \right] \\ &= \mathbb{E} \left[\sum_{\ell=1}^j \mathbb{I}[T_i > \ell] (v(0, \ell))^3 + 3 \sum_{\ell < m} \mathbb{I}[T_i > m] (v(0, \ell))^2 v(0, m) \right. \\ &\quad \left. + 3 \sum_{\ell < m} \mathbb{I}[T_i > m] v(0, \ell) (v(0, m))^2 + 6 \sum_{\ell < m < r} \mathbb{I}[T_i > r] v(0, \ell) v(0, m) v(0, r) \right] \\ &= \sum_{\ell=1}^j \mathbb{E}[\ell p_{x_0}(\mathbf{k})] (v(0, \ell))^3 + 3 \sum_{\ell < m} \mathbb{E}[m p_{x_0}(\mathbf{k})] (v(0, \ell))^2 v(0, m) \\ &\quad + 3 \sum_{\ell < m} \mathbb{E}[m p_{x_0}(\mathbf{k})] v(0, \ell) (v(0, m))^2 + 6 \sum_{\ell < m < r} \mathbb{E}[r p_{x_0}(\mathbf{k})] v(0, \ell) v(0, m) v(0, r).\end{aligned}$$

The other expectations involved in $\mathbb{E}[V_j^3]$ can be obtained from

$$\begin{aligned}\mathbb{E}[V_{ij}^2 V_{hj}] &= \mathbb{E} \left[\mathbb{E} \left[\left(\sum_{\ell=1}^j \mathbb{I}[T_i > \ell] v(0, \ell) \right)^2 \sum_{o=1}^j \mathbb{I}[T_h > o] v(0, o) \mid \mathbf{k} \right] \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\left(\sum_{\ell=1}^j \mathbb{I}[T_i > \ell] v(0, \ell) \right)^2 \mid \mathbf{k} \right] \mathbb{E} \left[\sum_{o=1}^j \mathbb{I}[T_h > o] v(0, o) \mid \mathbf{k} \right] \right] \\ &= \mathbb{E} \left[\left(\sum_{\ell=1}^j \ell p_{x_0}(\mathbf{k}) (v(0, \ell))^2 + 2 \sum_{\ell < m} m p_{x_0}(\mathbf{k}) v(0, \ell) v(0, m) \right) \sum_{o=1}^j o p_{x_0}(\mathbf{k}) v(0, o) \right] \\ &= \sum_{\ell=1}^j \sum_{o=1}^j (v(0, \ell))^2 v(0, o) \mathbb{E}[\ell p_{x_0}(\mathbf{k}) o p_{x_0}(\mathbf{k})] \\ &\quad + 2 \sum_{\ell < m} \sum_{o=1}^j v(0, \ell) v(0, m) v(0, o) \mathbb{E}[m p_{x_0}(\mathbf{k}) o p_{x_0}(\mathbf{k})]\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[V_{ij} V_{kj} V_{hj}] &= \mathbb{E} \left[\mathbb{E} \left[\sum_{\ell=1}^j \mathbb{I}[T_i > \ell] v(0, \ell) \sum_{m=1}^j \mathbb{I}[T_k > m] v(0, m) \sum_{o=1}^j \mathbb{I}[T_h > o] v(0, o) \mid \mathbf{k} \right] \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\sum_{\ell=1}^j \mathbb{I}[T_i > \ell] v(0, \ell) \mid \mathbf{k} \right] \mathbb{E} \left[\sum_{m=1}^j \mathbb{I}[T_k > m] v(0, m) \mid \mathbf{k} \right] \right. \\ &\quad \left. \times \mathbb{E} \left[\sum_{o=1}^j \mathbb{I}[T_h > o] v(0, o) \mid \mathbf{k} \right] \right] \\ &= \sum_{\ell=1}^j \sum_{m=1}^j \sum_{o=1}^j v(0, \ell) v(0, m) v(0, o) \mathbb{E}[\ell p_{x_0}(\mathbf{k}) m p_{x_0}(\mathbf{k}) o p_{x_0}(\mathbf{k})].\end{aligned}$$

ACKNOWLEDGMENTS

Both authors would like to thank Rob Kaas as well as an associate editor and two anonymous referees for interesting comments made on earlier versions of this manuscript. Michel Denuit would like to thank Pascal Schoenmaekers from Munich Re for stimulating exchanges about the Lee-Carter approach to mortality forecasting. Michel Denuit acknowledges the financial support of the *Communauté française de Belgique* under contract “Projet d’Actions de Recherche Concertées” ARC 04/09-320, as well as the financial support of the *Banque Nationale de Belgique* under grant “Risk Measures and Economic Capital”. Esther Frostig thanks the Zimerman Foundation for the study of Banking and Finance for their financial support.

REFERENCES

- BROUHNS, N., M. DENUIT, AND I. VAN KEILEGOM. 2005. Bootstrapping the Poisson Log-Bilinear Model for Mortality Projection. *Scandinavian Actuarial Journal*, 212–24.
- DENUIT, M., J. DHAENE, M. J. GOOVAERTS, AND R. KAAS. 2005. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. New York: Wiley.
- DENUIT, M., AND E. FROSTIG. 2007. Association and Heterogeneity of Insured Lifetimes in the Lee-Carter Framework. *Scandinavian Actuarial Journal* 107: 1–19.
- DENUIT, M., AND E. FROSTIG. 2008. First-Order Mortality Basis for Life Annuities. *Geneva Risk and Insurance Review*, 33: 75–89.
- LEE, R. D. 2000. The Lee-Carter Method of Forecasting Mortality, with Various Extensions and Applications. *North American Actuarial Journal* 4: 80–93.
- LEE, R. D., AND L. CARTER. 1992. Modelling and Forecasting the Time Series of U.S. Mortality. *Journal of the American Statistical Association* 87: 659–71.
- PITACCO, E., M. DENUIT, S. HABERMAN, AND A. OLIVIERI. 2009. *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford: Oxford University Press.
- SHAKED, M., AND F. SPIZZICHINO. 1998. Positive Dependence Properties of Conditionally Independent Random Lifetimes. *Mathematics of Operations Research* 23: 944–59.
- TONG, Y. L. 1990. *The Multivariate Normal Distribution*. New York: Springer.

Discussions on this paper can be submitted until January 1, 2010. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.