FIRST-ORDER MORTALITY RATES AND SAFE-SIDE ACTUARIAL CALCULATIONS IN LIFE INSURANCE

BY

MARCUS C. CHRISTIANSEN AND MICHEL M. DENUIT

ABSTRACT

In this paper, we discuss how to define conservative biometric bases in life insurance. The first approach is based on cumulative hazard (or survival probabilities), the second one on the hazard itself, and the third one on the rate of increase of the hazard. The second case has been studied in the literature and the sum-at-risk plays a central role in defining safe-side requirements. The two other cases appear to be new and concepts related to sum-at-risk are defined.

KEYWORDS

Variations in the technical basis; calculating on the safe side; Solvency II; first-order basis; second-order basis; sum at risk.

1. INTRODUCTION AND MOTIVATION

The calculation of premiums and reserves on the safe side has always attracted a lot of interest in life insurance. Life insurance calculations are performed either with first-order technical bases or with second-order technical bases. First-order bases include a safety margin whereas second-order ones do not contain any margin and are assumed to be close to reality. Appropriate first-order bases are essential to the life insurance business. The reason is that these bases justify the use of expected present values without explicit safety loading (the loading being implicitly contained in the prudential life table and interest rate).

Practical experience shows that mortality rates can change significantly within one decade. Typically, we are in a situation as exemplified by Figure 1.1. The real mortality rate differs from the estimated one (black solid line) because of, for example, an unforeseen catastrophe (upper dashed line) or a longevity effect (lower dashed line). By applying statistical methods on data of the past, we can usually narrow future uncertainties down to a confidence band (grey area with grey solid curves as bounds). Premiums and reserves should now be chosen in such a way that they are on the safe side with respect to all kinds of mortality scenarios that are within that confidence band.

Astin Bulletin 40(2), 587-614. doi: 10.2143/AST.40.2.2061129

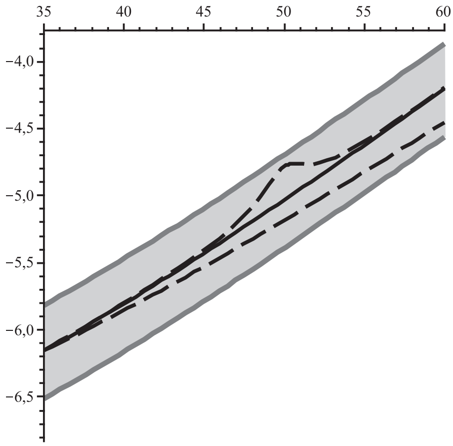

FIGURE 1.1: Log mortality rates: best estimate (black solid curve), alternative scenarios (black dashed curves), and confidence band (grey area with grey solid curves as bounds.

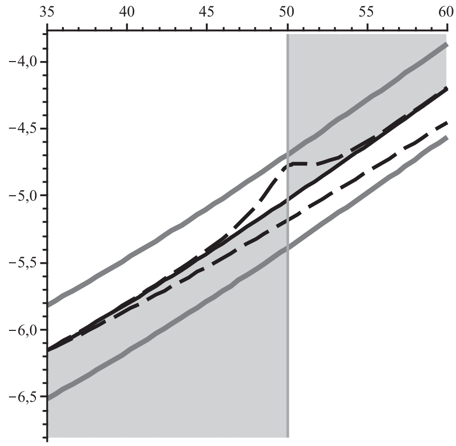

FIGURE 1.2: Log mortality rates: best estimate (black solid curve) with confidence bounds (grey solid curves), alternative scenarios (black dashed curves), and safe side area with respect to the best estimate (grey area).

So far, the literature offers three concepts for the construction of first-order mortality scenarios. First, there is the method based on the sum-at-risk, which was developed by Lidstone (1905), Norberg (1985), Hoem (1988), Ramlau-Hansen (1988), and Linnemann (1993). The sum-at-risk quantifies the financial consequence of a death occurring at time t, in which case the insurer has to pay the death benefit and the reserve is released. Lidstone (1905) and Norberg (1985) analyze the effects of variations in the valuation basis while maintaining the equivalence principle, whereas Hoem (1988), Ramlau-Hansen (1988), and Linnemann (1993) study the emergence of surplus. In the present paper we follow the surplus concept of the last three authors. For a given first-order mortality rate with corresponding sum-at-risk, all of these authors showed that reserves are on the safe side if the second-order mortality rate is smaller at ages for which the sum-at-risk is positive and if the second-order mortality rate is greater at ages for which the sum-at-risk is negative. (From Lidstone (1905) to Linnemann (1993) the complexity grew from simple single life insurance policies to portfolios of policies with multiple states.) This is exemplified in Figure 1.2. Assume that the sum-at-risk — here calculated on the basis of the best estimate (black solid line) — is positive until the policyholder reaches age 50 and negative afterwards. Think, for example, of a combination of a pure endowment insurance and a temporary life insurance. The sum-at-risk method yields now that premiums and reserves are on the safe side with respect to any second-order mortality rate within the grey area. Unfortunately, we can not say anything about our two alternative scenarios (dashed lines), because they do not completely lie within the grey area.

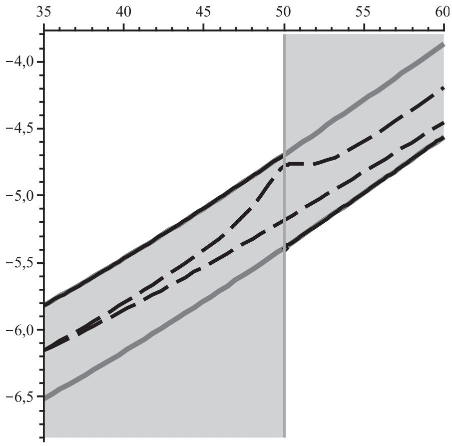

FIGURE 1.3: Log mortality rates: first-order basis (black solid curve), confidence bounds (grey solid curves), alternative scenarios (black dashed curves), and **desired** safe side area with respect to the first-order basis (grey area).

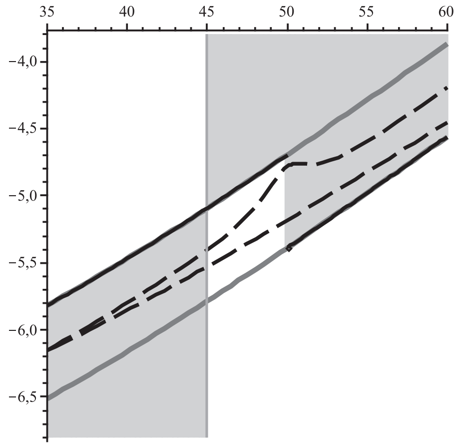

FIGURE 1.4: Log mortality rates: first-order basis (black solid curve), confidence bounds (grey solid curves), alternative scenarios (black dashed curves), and **true** safe side area with respect to the first-order basis (grey area).

The worst mortality scenario that can be found within the confidence band by the sum-at-risk method is shown in Figure 1.3. It is on the upper and lower bound where ever the sum-at-risk is positive and negative, respectively. Now we take that worst-scenario as our new first-order basis. Seemingly, we expanded the safe-side area in a way we wished for. The safe side area presumably contains now the whole confidence band, illustrated in Figure 1.3. Unfortunately, Figure 1.3 is in general wrong, because changing the first-order basis changes at the same time the sum-at-risk. The effect is shown in Figure 1.4. The switching point between positive and negative sums-at-risk moved in our example to age 45, and the grey area illustrates the actual safe-side area according to the sum-at-risk method with respect to our new first-order basis. We see that between age 45 and age 50 our confidence band is even completely outside the safe side area, which means that for all scenarios within the confidence band the sum-at-risk method can not decide whether they are on the safe side or not. To summarize, the sum-at-risk method does not yield a first-order basis that is definitely on the safe-side with respect to all scenarios within a confidence band.

The second method that can be found in the literature is based on derivatives. References using such an approach include Dienst (1995), Bowers et al. (1997), Kalashnikov and Norberg (2003), Christiansen and Helwich (2008), or Christiansen (2008a, 2008b). The problem is here that differentiation in general is a local concept. Strictly speaking, we can only study infinitesimal changes of the mortality rate. We get good approximations for realistic changes of the mortality rate if the confidence band for the second-order basis is not too wide, but still the approximation error is generally difficult to control.

Thus, the method based on derivatives works only for narrow confidence bands and yields not exact but only approximative results.

A third method for the construction of first-order mortality scenarios is given in Christiansen (2010). Based on Thiele's integral equation, another integral equation is developed whose solution yields the maximal prospective reserve with respect to all cumulative mortality intensities whose rate of increase is within some confidence band. In contrast to the first and the second method, the third method yields a first-order basis that is definitely on the safe-side with respect to a confidence band, and the results are always exact regardless of the width of the confidence band. However, by bounding the rate of increase of the cumulative mortality intensity and not the cumulative mortality intensity itself, it may happen that we exclude mortality scenarios that can occur in reality. On the other hand, the method of Christiansen (2010) includes scenarios that might be seen as rather unrealistic.

In the present paper we describe three approaches for the calculation of a first-order basis. All of them yield scenarios that are definitely on the safe side with respect to a confidence band, and they all offer exact results regardless of the width of the confidence band. They mainly differ in the sets of mortality scenarios that are included and excluded. Specifically,

(a) in approach 1, we allow for any cumulative hazard rate within a lower and an upper bound
(b) in approach 2, we allow for cumulative hazard rates whose rate of increase is within a lower and an upper bound. In case of differentiability, that is equivalent to have a lower and an upper bound for the hazard rate.
(c) in approach 3, we allow for cumulative hazard rates whose rate of acceleration is within a lower and an upper bound. In case of twice differentiability, that is equivalent to having a lower and an upper bound for the derivative of the hazard rate.

The second approach is based on the method of Christiansen (2010). The first and third approaches seem to be new in the literature. Suppose that confidence bands for (a), (b), and (c) are given. Then approach (a) includes the biggest set of mortality scenarios. In return we obtain premiums and reserves that have a strong safety loading, but the first-order basis is not necessarily a true cumulative hazard rate itself. Approach (b) makes stronger restrictions and includes less mortality scenarios than (a), thus the first-order basis is always a true cumulative hazard rate, and premiums and reserves now have a smaller safety loading. Approach (c) makes the strongest restrictions on the set of admissible mortality scenarios, now hazard rates are never decreasing, and in return we obtain the smallest safety loading for premiums and reserves. It is not obvious which of the restrictions of (a) to (c) on the set of admissible mortality scenarios are really satisfied in reality. Therefore, we present and compare in this paper all three approaches and leave it to the practitioner to decide which a priori assumptions he is willing to accept. The following table gives a condensed overview:

	Construction of a first-order basis on the safe side with respect to a confidence band	Cumulative hazard rates of second-order within the confidence band are ...
method based on sum-at-risk	**no**	
method based on derivatives	**approximately**, confidence band for (cumulative) hazard rate	arbitrary
method (a), section 3	**yes**, confidence band for cumulative hazard rate	arbitrary
method (b), section 4	**yes**, confidence band for hazard rate	increasing
method (c), section 5	**yes**, confidence band for rate of increase of hazard rate	increasing with increasing speed

2. Basic modeling

Consider a life insurance policy that is issued at time 0. We write x for the age of the policyholder at the beginning of the contract period, T for his or her total lifetime, and ω_x for the limiting age for individuals with age x at contract time zero.

The cash-flows of the contract are described by the following functions:

1. The lump sum $c(t)$ is payable upon death at time t. We assume that the function c has bounded variation on $[0, \omega_x]$ and is left-continuous (left-continuity ensures that when the death benefit corresponds to the reserve or to the part of a loan still to be reimbursed, the payment at the time of death is not taken into account).
2. The functions $B(t)$ and $\Pi(t)$ give the accumulated annuity benefits and premiums in case of survival up to t. We assume that B and Π have bounded variation on $[0, \omega_x]$ and are right-continuous.

We write $v(s,t)$ for the value at time s of a unit payable at time $t > s$ and assume that it has a representation of the form

$$v(s,t) = e^{-\int_{(s,t]} \phi(u)\,\mathrm{d}u}$$

with ϕ being the *interest intensity* (or *short interest rate*).

The *cumulative mortality intensity* (or cumulative hazard rate) is defined by

$$\Lambda_x(t) := -\ln P(T > x + t \mid T > x).$$

We assume that Λ_x is continuous. In order to distinguish between different cohorts, we do *not* further simplify this notation to $\Lambda_x(t) = \Lambda(x + t)$. If Λ_x is differentiable, we can also define a *mortality intensity* (or hazard rate),

$$\lambda_x(t) := \frac{\mathrm{d}}{\mathrm{d}t}\Lambda_x(t).$$

If Λ_x is even twice differentiable, we define

$$\alpha_x(t) := \frac{\mathrm{d}}{\mathrm{d}t}\lambda_x(t) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Lambda_x(t).$$

3. Worst-case if the cumulative hazard rate is bounded

The prospective reserve at time s is obtained as the expected present value of future benefits minus the expected present value of future premiums, that is,

$$\begin{aligned} V(s) &:= \mathbb{E}\Big[\int_{(s,T-x)} v(s,t)\,\mathrm{d}(B-\Pi)(t) + v(s,T-x)c(T-x)\,|\,T-x>s\Big] \\ &= \int_{(s,\omega_x]} e^{\Lambda_x(s)-\Lambda_x(t)} v(s,t)\,\mathrm{d}(B-\Pi)(t) - \int_{(s,\omega_x]} v(s,t)\,c(t)\,e^{\Lambda_x(s)}\mathrm{d}e^{-\Lambda_x(t)}. \end{aligned}$$

Now we regard $V(s)$ as a mapping of the conditional survival function, the latter being defined by

$$[s,\omega_x] \ni t \mapsto e^{\Lambda_x(s)-\Lambda_x(t)} = P(T>x+t\,|\,T>x+s).$$

What happens to the prospective reserve if the conditional survival function is shifted by an amount of $Q(\cdot)$ to $\exp\{\Lambda_x(s)-\Lambda_x(\cdot)\}+Q(\cdot)$? In the following we assume that $Q(\cdot)$ is right-continuous, has bounded variation on $[s,\omega_x]$, and is equal to zero at s and ω_x. Using the linearity of $V(s)$ with respect to the conditional survival function and applying Fubini's Theorem, we get in obvious notation

$$\begin{aligned} &V\Big(s, e^{\Lambda_x(s)-\Lambda_x(\cdot)} + Q(\cdot)\Big) - V\Big(s, e^{\Lambda_x(s)-\Lambda_x(\cdot)}\Big) \\ &\quad= \int_{(s,\omega_x]} Q(t)\,v(s,t)\,\mathrm{d}(B-\Pi)(t) - \int_{(s,\omega_x]} v(s,t)\,c(t)\,\mathrm{d}Q(t) \\ &\quad= \int \mathbf{1}_{(s,\omega_x]}(u)\Big(\int_{[u,\omega_x]} v(s,t)\,\mathrm{d}(B-\Pi)(t) - v(s,u)\,c(u)\Big)\,\mathrm{d}Q(u) \\ &\quad=: \int_{(s,\omega_x]} S_s(u)\,\mathrm{d}Q(u) \end{aligned} \tag{3.1}$$

where

$$S_s(u) := v(s,u)\Big(\int_{[u,\omega_x]} v(u,t)\,\mathrm{d}(B-\Pi)(t) - c(u)\Big) \tag{3.2}$$

can be seen as *cumulative survival cost at time s* for survival at and after u. To motivate that definition, look at the example where $Q = \varepsilon \mathbf{1}_{[t_0, \omega_x)}$ for some fixed $t_0 > s$ and an $\varepsilon > 0$. For a homogeneous portfolio that means that we have from time t_0 on throughout (100ε)% more policyholders that are still alive. According to (3.1), the effect of shift $Q = \varepsilon \mathbf{1}_{[t_0, \omega_x)}$ on the prospective reserve $V(s)$ is $\varepsilon S_s(t_0)$. Coming back to the homogeneous portfolio, $\varepsilon S_s(t_0)$ is the increase of the discounted cost per policy due to increasing the survival rate on $[t_0, \omega_x)$ by ε. We can get another interesting interpretation of function S_s after applying partial integration on the last term of (3.1), which gives

$$\begin{aligned} V\left(s, e^{\Lambda_x(s) - \Lambda_x(\cdot)} + Q(\cdot)\right) - V\left(s, e^{\Lambda_x(s) - \Lambda_x(\cdot)}\right) &= \int_{(s, \omega_x]} S_s(u)\, \mathrm{d}Q(u) \\ &= \int_{(s, \omega_x)} -Q(u)\, \mathrm{d}S_s(u). \end{aligned} \tag{3.3}$$

(Note that Q is right-continuous, S_s is left-continuous, and that we assumed that $Q(s) = Q(\omega_x) = 0$). Now we see that $-\mathrm{d}S_s(u)$ describes the effect that the increase $Q(u)$ of the survival function at time u has on the prospective reserve $V(s)$. Therefore, we denote $-\mathrm{d}S_s(u)$ as *survival cost at time s* for survival at time u, and by differentiating (3.2) we can show that

$$\begin{aligned} -\mathrm{d}S_s(u) &= v(s,u)(\mathrm{d}B(u) - \mathrm{d}\Pi(u) - \phi(u)c(u)\mathrm{d}u + \mathrm{d}c(u)) \\ &=: v(s,u)\mathrm{d}S(u) \end{aligned} \tag{3.4}$$

for all $u \geq s$, where $\mathrm{d}S(u)$ is denoted as *survival cost for survival at time u.* This representation allows for an intuitive interpretation: By infinitesimally delaying the death of the policyholder at time u, additional benefits of $\mathrm{d}B(u)$ fall due, additional premiums of $\mathrm{d}\Pi(u)$ are paid, the insurer gets a discounting advantage for the death benefit of $\phi(u)c(u)\mathrm{d}u$, and the contractual liabilities concerning death change by $\mathrm{d}c(u)$.

Now we assume that the conditional survival function $\exp\{\Lambda_x(s) - \Lambda_x(\cdot)\}$ has a lower and an upper bound,

$$e^{U_x(s) - U_x(t)} \leq e^{\Lambda_x(s) - \Lambda_x(t)} \leq e^{L_x(s) - L_x(t)}, \quad t \in [s, \omega_x], \tag{3.5}$$

where the bounds shall be continuous survival functions with respect to t. Instead of studying shifts of the survival function of the form $\exp\{\Lambda_x(s) - \Lambda_x(\cdot)\} + Q(\cdot)$ within the boundaries (3.5), we will study shifts of the form $\Lambda_x + H$ and use the equivalent bounds

$$L_x(t) - L_x(s) \leq \Lambda_x(t) - \Lambda_x(s) \leq U_x(t) - U_x(s), \quad t \in [s, \omega_x], \tag{3.6}$$

where the bounds have to be continuous cumulative hazard rates with limiting age ω_x. From now on we see the prospective reserve $V(s) = V(s, \Lambda_x)$ as a mapping of the cumulative hazard rate Λ_x.

3.1. Construction of a worst-case scenario

We are interested in the maximal value that the prospective reserve $V(s)$ can take if the conditional survival function may be chosen arbitrarily within the bounds (3.5) or, equivalently, if the cumulative mortality intensity may be chosen arbitrarily within the bounds (3.6). In other words, we are looking for the worst-case prospective reserve (or at least an upper bound for it) from the perspective of the insurer. Let Y and Z be random variables with survival functions $e^{\Lambda_x(s) - \Lambda_x(\cdot)}$ and $e^{\Lambda_x(s) + H(s) - \Lambda_x(\cdot) - H(\cdot)}$. If S_s is non-increasing or, equivalently, $\mathrm{d}S(t)$ is never negative, then the fact that

$$
\begin{aligned}
V(s, \Lambda_x) &= \int_{(s, \omega_x]} S_s(t)\, \mathrm{d}e^{\Lambda_x(s) - \Lambda_x(\cdot)}(t) + \int_{(s, \omega_x]} v(s,t)\, \mathrm{d}(B - \Pi)(t) \\
&= \mathbb{E}(-S_s(Y)) + \text{const}
\end{aligned}
$$

(for the first equality apply Fubini's Theorem similar to (3.1)) leads to

$$
\begin{aligned}
&P(Y > t) \geq P(Z > t) \text{ for all } t \quad \Longrightarrow \\
&V(s, \Lambda_x) = \mathbb{E}(-S_s(Y)) + \text{const} \geq \mathbb{E}(-S_s(Z)) + \text{const} = V(s, \Lambda_x + H)
\end{aligned}
$$

or equivalently

$$
\begin{aligned}
&H(s) - H(t) \leq 0 \text{ for all } t \in [s, \omega_x] \quad \Longrightarrow \\
&V(s, \Lambda_x) = \mathbb{E}(-S_s(Y)) + \text{const} \geq \mathbb{E}(-S_s(Z)) + \text{const} = V(s, \Lambda_x + H).
\end{aligned}
$$

The same relation holds for the prospective reserves if S_s is non-decreasing and $H(s) - H(t) \geq 0$ for all $t \in [s, \omega_x]$. Thus, we get that $\Lambda_x = L_x$ maximizes the prospective reserve if S_s is non-increasing and $\Lambda_x = U_x$ maximizes the prospective reserve if S_s is non-decreasing. In other words, the lower bound L_x and the upper bound U_x are worst-case scenarios if the survival cost $\mathrm{d}S$ is throughout non-negative and throughout non-positive, respectively. This result can be generalized to cases where $\mathrm{d}S$ may change its sign, as shown next.

Property 3.1. *Let* $\mathrm{d}S$ *be the survival cost according to* (3.4)*. Then, for all continuous functions* H *with bounded variation on* $[s, \omega_x]$*, we have*

$$
\begin{aligned}
\operatorname{sign}(H(s) - H(t)) &= \operatorname{sign}(\mathrm{d}S(t)) \textit{ for all } t > s \\
\Longrightarrow \quad V(s, \Lambda_x + H) &\geq V(s, \Lambda_x)
\end{aligned} \tag{3.7}
$$

and

$$\operatorname{sign}(H(s) - H(t)) = -\operatorname{sign}(\mathrm{d}S(t)) \text{ for all } t > s \implies V(s, \Lambda_x + H) \leq V(s, \Lambda_x). \tag{3.8}$$

PROOF. Because of (3.3), the difference

$$V(s, \Lambda_x + H) - V(s, \Lambda_x) = \int_{(s, \omega_x)} e^{\Lambda_x(s) - \Lambda_x(t)} \left(e^{H(s) - H(t)} - 1\right) v(s,t)\, \mathrm{d}S(t)$$

is always non-negative and non-positive under conditions (3.7) and (3.8), respectively. □

Property 3.1 allows us to calculate an upper bound for the prospective reserve:

Proposition 3.2. *Let* dS *be the survival cost according to* (3.4). *Then* $\bar{\Lambda}_x$ *defined by*

$$\bar{\Lambda}_x(t) - \bar{\Lambda}_x(s) := \begin{cases} L_x(t) - L_x(s) \ : \ S(t) > 0 \\ U_x(t) - U_x(s) \ : \ S(t) < 0 \end{cases} \tag{3.9}$$

and arbitrary values on $\{s\} \cup \{t \,|\, \mathrm{d}S(t) = 0\}$ *satisfies* $V(s, \bar{\Lambda}_x) \geq V(s, \Lambda_x)$ *for all cumulative mortality intensities* Λ_x *that are within the bounds* (3.6).

PROOF. Apply Property 3.1, and note that in the proof of Property 3.1 the function $e^{\Lambda_x(s) - \Lambda_x(t)} + Q$ is not necessarily a survival function but only has to be right-continuous and of bounded variation on $[s, \omega_x]$. □

We denote $\bar{\Lambda}_x$ as *worst-case scenario with respect to the bounds* (3.6). Note that $\bar{\Lambda}_x(\cdot)$ is not necessarily monotone and, hence, not always a true survival function.

REMARK 3.3. (Time invariance). What happens to the worst-case scenario of $V(s)$ when time s is moving forward? The worst-case scenario according to Proposition 3.2 depends only on the sign of dS which does not depend on s. That means that if we once calculated $\bar{\Lambda}_x$ at the beginning of the contract period $s = 0$, it remains to be a worst-case scenario during the whole contract time.

However, the approach presented in this section has a significant disadvantage. In many examples the worst-case scenario $\bar{\Lambda}_x$ is not monotone and, hence, not a true cumulative hazard rate anymore. This implies that the upper bound for the prospective reserve is in fact not sharp. This is why in Section 4 we bound the rate of increase of Λ_x instead of Λ_x itself.

4. Worst-case if the hazard rate is bounded

In contrast to (3.6), we assume now that the rate of increase of the cumulative hazard rate is bounded,

$$\mathrm{d}L_x(t) \le \mathrm{d}\Lambda_x(t) \le \mathrm{d}\,U_x(t), \quad t \in [s, \omega_x], \tag{4.1}$$

where L_x and U_x are continuous and increasing functions with bounded variation on $[s, \omega_x]$. In case of differentiability, that is equivalent to bounding the hazard rate,

$$l_x(t) \le \lambda_x(t) \le u_x(t), \quad t \in [s, \omega_x], \tag{4.2}$$

where l_x and u_x are the derivatives of L_x and U_x. The monotony of L_x implies that Λ_x is monotone and, hence, is always a true cumulative hazard rate.

The prospective reserve at time s can be written as

$$\begin{aligned} V(s) &= \int_{(s,\omega_x]} e^{\Lambda_x(s)-\Lambda_x(t)} v(s,t)\,\mathrm{d}(B-\Pi)(t) \\ &\quad + \int_{(s,\omega_x]} v(s,t)\,c(t)\,e^{\Lambda_x(s)-\Lambda_x(t)}\,\mathrm{d}\Lambda_x(t). \end{aligned} \tag{4.3}$$

Alternatively, we can see the prospective reserve as the unique solution of Thiele's integral equation

$$\begin{aligned} V(s) &= (B-\Pi)(\omega_x) - (B-\Pi)(s) - \int_{(s,\omega_x]} V(t-)\,\phi(t)\,\mathrm{d}t \\ &\quad + \int_{(s,\omega_x]} R(t)\,\mathrm{d}\Lambda_x(t) \end{aligned} \tag{4.4}$$

with initial value $V(\omega_x) = 0$, where $R(s) := c(s) - V(s) - \Delta(B-\Pi)(s)$ is the so-called *sum-at-risk for occurrence of dead at time s.*

What happens to the prospective reserve if Λ_x is shifted by an amount of H to $\Lambda_x + H$? By generalizing the ideas of Lidstone (1905), Norberg (1985), Hoem (1988), Ramlau-Hansen (1988), and Linnemann (1993) to a model with a *cumulative* mortality intensity, we obtain the following result, which is the basis for the 'sum-at-risk method'.

Property 4.1. *Let $R(s, \Lambda_x)$ be the sum-at-risk that corresponds to Λ_x. If the shifted cumulative mortality intensity $\Lambda_x + H$ is still a continuous cumulative hazard rate, then we have*

$$\begin{aligned} &\mathrm{sign}(\mathrm{d}H(t)) = \mathrm{sign}(R(t,\Lambda_x)) \text{ for all } t > s \\ &\quad\Longrightarrow\quad V(s,\Lambda_x + H) \ge V(s,\Lambda_x) \end{aligned} \tag{4.5}$$

and

$$\begin{aligned}&\operatorname{sign}(\mathrm{d}H(t)) = -\operatorname{sign}(R(t,\Lambda_x)) \textit{ for all } t > s \\ &\quad\Longrightarrow\quad V(s,\Lambda_x+H) \le V(s,\Lambda_x).\end{aligned} \tag{4.6}$$

PROOF. Let $W(s) := V(s, \Lambda_x + H) - V(s, \Lambda_x)$ be the difference between the prospective reserves at time s. By replacing $V(s, \Lambda_x + H)$ and $V(s, \Lambda_x)$ with the right hand side of (4.4), we get an integral equation for W,

$$\begin{aligned}W(s) &= -\int_{(s,\omega_x]} (V(t-,\Lambda_x+H) - V(t-,\Lambda_x))\,\phi(t)\,\mathrm{d}t \\ &\quad + \int_{(s,\omega_x]} \big(R(t,\Lambda_x+H)\,\mathrm{d}(\Lambda_x+H)(t) - R(t,\Lambda_x)\,\mathrm{d}\Lambda_x(t)\big)\end{aligned}$$

with initial value $W(\omega_x) = 0$. With defining C by

$$C(\omega_x) - C(s) = \int_{(s,\omega_x]} R(t,\Lambda_x)\,\mathrm{d}H(t),$$

we can write the above integral equation for W in the form

$$W(s) = C(\omega_x) - C(s) - \int_{(s,\omega_x]} W(t-)\,\phi(t)\,\mathrm{d}t + \int_{(s,\omega_x]} - W(t)\,\mathrm{d}(\Lambda_x+H)(t).$$

We can interpret this integral equation as a Thiele integral equation for a policy with no death benefits and accumulated annuity benefits and premiums of C. (In the actuarial literature, $C(s)$ is interpreted as the accumulated surplus at time s). Hence, the integral equation has the solution (see (4.3))

$$\begin{aligned}W(s) &= \int_{(s,\omega_x]} e^{\Lambda_x(s)+H(s)-\Lambda_x(t)-H(t)}\,v(s,t)\,\mathrm{d}C(t) \\ &= \int_{(s,\omega_x]} e^{\Lambda_x(s)+H(s)-\Lambda_x(t)-H(t)}\,v(s,t)\,R(t,\Lambda_x)\,\mathrm{d}H(t).\end{aligned} \tag{4.7}$$

Under the conditions in (4.5) and (4.6) we obtain $W(s) \ge 0$ and $W(s) \le 0$, respectively, and, hence, $V(s,\Lambda_x + H) - V(s,\Lambda_x) = W(s) \ge 0$ and $V(s, \Lambda_x + H) - V(s,\Lambda_x) = W(s) \le 0$.

Defining

$$R_s(t) := e^{\Lambda_x(s)+H(s)-\Lambda_x(t)-H(t)}\,v(s,t)\,R(t,\Lambda_x) \tag{4.8}$$

as the *sum-at-risk at time s* for occurrence of death at time t, we get from (4.7) an expression similar to (3.3):

$$V(s,\Lambda_x+H) - V(s,\Lambda_x) = \int_{(s,\omega_x]} R_s(t)\,\mathrm{d}H(t). \tag{4.9}$$

While the survival cost $-\mathrm{d}S_s(t)$ describes the effect that a $Q(t)$ shift of the survival function has on $V(s)$, it is here $R_s(t)$ that quantifies the effect that a $\mathrm{d}H(t)$ shift of the mortality intensity has on $V(s)$. Property 4.1 is similar to Property 3.1. While $\mathrm{sign}(\mathrm{d}S(t))$ describes the direction of the effect that a $H(t)$ shift of the cumulative mortality intensity at time t has on $V(s)$, it is here sign $(R(t))$ that quantifies the direction of the effect that a $\mathrm{d}H(t)$ shift of the mortality intensity at time t has on $V(s)$. It is then tempting to believe that we can find a worst-case scenario analogously to Proposition 3.2 by letting $\mathrm{d}\Lambda_x$ be equal to $\mathrm{d}L_x$ and $\mathrm{d}U_x$ where ever $R(t, \Lambda_x)$ is negative and positive, respectively. As already indicated in the introduction, this idea does not work. The problem is here that quantity (4.8) depends on shift H, whereas $-\mathrm{d}S_s$ did not depend on shift Q. Therefore the worst-case problem is more complicated here.

4.1. Construction of a worst-case scenario

Property 4.1 guarantees that the valuation basis Λ_x is on the safe side with respect to all alternative mortality scenarios $\Lambda_x + H$ that meet condition (4.6). This safe side area usually does not contain the whole confidence band (4.1). (See also the explanations in the introduction.) But if we had a mortality scenario $\bar{\Lambda}_x$ that satisfies

$$\mathrm{d}\bar{\Lambda}_x(t) = \begin{cases} \mathrm{d}L_x(t) \; : \; R(t, \bar{\Lambda}_x) < 0 \\ \mathrm{d}U_x(t) \; : \; R(t, \bar{\Lambda}_x) > 0 \end{cases}, \tag{4.10}$$

then Property 4.1 would yield a safe side area for $\bar{\Lambda}_x$ that indeed contains the whole confidence band (4.1), because all possible shifts H meet (4.6). The natural questions are therefore:

- Does such a special scenario $\bar{\Lambda}_x$ always exist?
- If so, how do we find $\bar{\Lambda}_x$?

Answers to that questions can be found in Christiansen (2010). By replacing the cumulative mortality intensity in Thiele's integral equation (4.4) with the right hand side of (4.10), we get a new integral equation that does not directly depend on $\bar{\Lambda}_x$ anymore,

$$\begin{aligned} \bar{V}(s) = {} & (B-\Pi)(\omega_x) - (B-\Pi)(s) - \int_{(s,\omega_x]} \bar{V}(t-)\,\phi(t)\,\mathrm{d}t \\ & + \int_{(s,\omega_x]} \frac{\bar{R}(t) - |\bar{R}(t)|}{2}\,\mathrm{d}L_x(t) + \int_{(s,\omega_x]} \frac{\bar{R}(t) + |\bar{R}(t)|}{2}\,\mathrm{d}U_x(t) \end{aligned} \tag{4.11}$$

with initial value $\bar{V}(\omega_x) = 0$, where we use the short notation $\bar{V}(s) := V(s, \bar{\Lambda}_x)$ and $\bar{R}(s) := R(s, \bar{\Lambda}_x)$. Christiansen (2010) showed that the integral equation (4.11) has a unique solution in

$$\{V : [0, \omega_x] \to \mathbb{R} \mid V \text{ is right-continuous and has bounded variation, } V(\omega_x) = 0\}.$$

Once we have a solution $\bar{V}$ for (4.11), we can construct a worst-case mortality scenario as follows.

Property 4.2. *Let $\bar{V}$ be the unique solution of integral equation* (4.11) *with corresponding sum-at-risk $\bar{R}$. Then $\bar{\Lambda}_x$ defined by*

$$\mathrm{d}\bar{\Lambda}_x(t) = \begin{cases} \mathrm{d}L_x(t) \; : \; \bar{R}(t) < 0 \\ \mathrm{d}U_x(t) \; : \; \bar{R}(t) > 0 \end{cases} \tag{4.12}$$

and arbitrary but increasing values on $\{0\} \cup \{t : \bar{R}(t) = 0\}$ *is a cumulative mortality intensity with* $V(s, \bar{\Lambda}_x) \geq V(s, \Lambda_x)$ *for all* $s \in [0, \omega_x]$ *and all* Λ_x *that satisfy* (4.1).

Proof. Christiansen (2010) showed that $\bar{\Lambda}_x$ is indeed a cumulative mortality intensity. In the same way that we derived (4.11) from (4.10), we can verify that $V(\cdot, \bar{\Lambda}_x)$ is equal to the unique solution $\bar{V}$ of (4.11). Thus, we also have $\bar{R} = R(\cdot, \bar{\Lambda}_x)$, which means that (4.12) satisfies (4.10). By applying Property 4.1 now for each $s \in [0, \omega_x]$, we get the maximality of $V(s, \bar{\Lambda}_x)$ for all $s \in [0, \omega_x]$. □

We call $\bar{\Lambda}_x$ according to (4.12) the *worst-case scenario with respect to* (4.1).

Remark 4.3. (Time invariance & characterization of the worst-case). Note that the worst-case scenario $\bar{\Lambda}_x$ maximizes not only the prospective reserve at some fixed time s, but also at any other time $t \in [0, \omega_x]$. That means that if we once calculated $\bar{\Lambda}_x$ at the beginning of the contract period, it remains to be a worst-case scenario during the whole contract period. This implies that $R(t, \bar{\Lambda}_x) = c(t) - V(t, \bar{\Lambda}_x) - \Delta(B - \Pi)(t)$ is minimal for all t, and consequently

$$R(t, \bar{\Lambda}_x) = \inf_{\Lambda_x} \left(R(t, \Lambda_x) \right)$$

for all t. By interpreting a positive sum-at-risk as occurrence character and a negative sum-at-risk as survival character, we get that the worst-case scenario is always that scenario that has the biggest share of survival character during the contract period.

The worst-case method in this section fixes the problem of the previous section that the worst-case scenario $\bar{\Lambda}_x$ is in general no cumulative hazard rate anymore. However, we still get unrealistic scenarios where the mortality intensity jumps between extremes and where mortality rates can also fall with increasing age. Such scenarios make sense in risk management if one is interested not in usual but in extreme developments of mortality.

Still, we can ask the question if it is possible to calculate worst-case scenarios which additionally have the following two properties: (a) they never fall with increasing age and (b) they have no extreme jumps. An answer to that question is given in Section 5.

4.2. Alternative construction of a worst-case scenario

Earlier in this section we discussed that, in contrast to (3.3), formula (4.9) does not yield a construction method for a worst-case scenario because the integrand depends on shift H. The 'method based on derivatives' (see the introduction of this paper) gets rid of that dependence on H by just allowing for local shifts H. Christiansen (2008a) shows that

$$\left| V(s,\Lambda_x + H) - V(s,\Lambda_x) - \int_{(s,\omega_x]} e^{\Lambda_x(s) - \Lambda_x(t)} v(s,t) R(t,\Lambda_x) \, \mathrm{d}H(t) \right| = o(\|H\|), \tag{4.13}$$

where $\|H\|$ is the total variation of H on $[0,\omega_x]$. Thus, given that $o(\|H\|)$ is negligible, the prospective reserve $V(s,\Lambda_x + H)$ can be maximized by choosing $\mathrm{d}(\Lambda_x + H)$ equal to $\mathrm{d}L_x$ and $\mathrm{d}U_x$ where ever $R(t,\Lambda_x)$ is negative and positive, respectively. This is the same scenario as the one suggested by the sum-at-risk method. The difference to the true worst-case scenario rises with $o(\|H\|)$. Christiansen (2008a) interpreted the integrand of (4.13) as some form of generalized gradient

$$(\nabla_{\Lambda_x} V)(t) := e^{\Lambda_x(s) - \Lambda_x(t)} v(s,t) R(t,\Lambda_x)$$

which gives us a new idea for the construction of a worst-case scenario. In the same way that gradient ascent methods are used to find local maxima of differentiable functions on $\mathbb{R}^n$, we can do iterated small steps in direction of the generalized gradient $\nabla_{\Lambda_x} V$ in order to find a maximizing mortality scenario:

1. Choose a starting mortality scenario $\Lambda_x^{(0)}$.
2. Calculate a new scenario by using the iteration
$$\mathrm{d}\Lambda_x^{(n+1)}(t) := \mathrm{d}\Lambda_x^{(n)}(t) + K(\nabla_{\Lambda_x^{(n)}} V)(t)\, \mathrm{d}t.$$
   If the right hand side is below $\mathrm{d}L_x$ or above $\mathrm{d}U_x$, we cut $\mathrm{d}\Lambda_x^{(n+1)}$ off at $\mathrm{d}L_x$ or $\mathrm{d}U_x$, respectively. Here, $K > 0$ is some step size that has to be chosen.
3. Repeat step 2 until $|V(s,\Lambda_x^{(n+1)}) - V(s,\Lambda_x^{(n)})|$ is below some error tolerance.

In order to increase the speed of convergence, we could try to increase K to infinity. As the sign of $(\nabla_{\Lambda_x^{(n)}} V)(t)$ is equal to the sign of $R(t, \Lambda_x^{(n)})$, and since we cut $\mathrm{d}\Lambda_x^{(n+1)}$ off at $\mathrm{d}L_x$ and $\mathrm{d}U_x$, we obtain the following algorithm:

1. Choose a starting mortality scenario $\Lambda_x^{(0)}$.
2. Calculate a new scenario by using the iteration

$$\mathrm{d}\Lambda_x^{(n+1)} := \begin{cases} \mathrm{d}L_x(t) & : R(t,\Lambda_x^{(n)}) < 0 \\ \mathrm{d}\,U_x(t) & : R(t,\Lambda_x^{(n)}) > 0 \\ \mathrm{d}\Lambda_x^{(n)}(t) & : R(t,\Lambda_x^{(n)}) = 0 \end{cases}. \tag{4.14}$$

3. Repeat step 2 until $V(s,\Lambda_x^{(n+1)}) - V(s,\Lambda_x^{(n)})|$ is below some error tolerance.

For $n = 0$, step 2 yields the same scenario as the one that is suggested by the sum-at-risk method or the method based on derivatives in order to maximize the prospective reserve $V(s)$. Hence, the second algorithm is just an iteration of the sum-at-risk method or the method based on derivatives. The question is whether that algorithm converges to the true worst-case. The following result provides the answer.

Proposition 4.4. *Let $\Lambda_x^{(0)}, \Lambda_x^{(1)}, \Lambda_x^{(2)}, \ldots$ be a series of cumulative mortality intensities calculated by iterating* (4.14). *Then*

$$\lim_{n\to\infty} V(s,\Lambda_x^{(n)}) = V(s,\bar{\Lambda}_x),$$

where $\bar{\Lambda}_x$ is the worst-case scenario according to (4.12).

Note that Proposition 4.4 states the convergence of the prospective reserves $V(s,\Lambda_x^{(n)})$ and not the convergence of the scenarios $\Lambda_x^{(n)}$. The latter do not necessarily converge to $\bar{\Lambda}_x$ on $\{t : R(t,\bar{\Lambda}_x) = 0\}$.

PROOF. Since definition (4.14) implies that

$$\mathrm{sign}\Big(\mathrm{d}\big(\Lambda_x^{(n+1)} - \Lambda_x^{(n)}\big)(t)\Big) = \mathrm{sign}\Big(R\big(t,\Lambda_x^{(n)}\big)\Big),$$

we can apply Property 4.1 and obtain that

$$V\big(t,\Lambda_x^{(n)}\big) \le V\big(t,\Lambda_x^{(n+1)}\big)$$

for all $t \ge s$ and all integers $n \ge 1$. Because the prospective reserves must be finite, the limit $\lim_{n\to\infty} V(t,\Lambda_x^{(n)})$ exists pointwise for all t. With the series $V(t,\Lambda_x^{(n)})$ being monotonic increasing, the series $R(t,\Lambda_x^{(n)}) = c(t) - V(t,\Lambda_x^{(n)}) - \Delta(B-\Pi)(t)$ must be monotonic decreasing. In view of definition (4.14), we can then conclude that $\mathrm{d}\Lambda_x^{(n)}(t) \ge \mathrm{d}\Lambda_x^{(n+1)}(t)$ for all t and that the limit $\lim_{n\to\infty} \Lambda_x^{(n)}$ exists in the Banach space of continuous functions with support in $[0,\omega_x]$ and the total variation as its norm. The differentiability of the prospective reserves with respect to the cumulative mortality intensity (see (4.13)) implies also continuity. Therefore we have

$$\lim_{n\to\infty} V\Big(s, \Lambda_x^{(n)}\Big) = V\Big(s, \lim_{n\to\infty} \Lambda_x^{(n)}\Big).$$

One can easily show now that the limit $\lim_{n\to\infty}\Lambda_x^{(n)}$ meets property (4.10) and therefore corresponds to the unique solution of (4.11). Hence, we have $\lim_{n\to\infty}\Lambda_x^{(n)}(t) = \bar{\Lambda}_x(t)$ on $\{t : \bar{R}(t) \neq 0\}$ and, consequently, $V(s, \lim_{n\to\infty}\Lambda_x^{(n)}) = V(s, \bar{\Lambda}_x)$. □

We see that the sum-at-risk method yields an approximation of the true worst-case, and we can improve this approximation by just iterating the sum-at-risk method.

5. Worst-case if the rate of increase of the hazard rate is bounded

Here we generally assume that λ_x exists. In contrast to (4.2) we do not bound the mortality intensity but the rate of increase of the mortality intensity. Specifically, assume that the inequalities

$$\mathrm{d}l_x(t) \leq \mathrm{d}\lambda_x(t) \leq \mathrm{d}u_x(t), \quad t \in [s, \omega_x], \tag{5.1}$$

are valid where l_x and u_x are hazard rates, and let $\lambda_x(s)$ be an arbitrary but fixed starting value at present (time s). The specification of λ_x at time s is needed since the derivative $\mathrm{d}\lambda_x$ on its own does not uniquely determine λ_x. If Λ_x is twice differentiable, (5.1) is equivalent to bounding α_x on $[s, \omega_x]$. If we choose a lower bound $\mathrm{d}l_x$ that is positive, then the mortality intensity is always monotonic increasing on $[s, \omega_x]$,

$$\lambda_x(t) - \lambda_x(s) \geq \int_{(s,t]} \mathrm{d}l_x(u).$$

Remark 5.1. For some practical applications it might be desirable to relax the assumption that $\lambda_x(s)$ is a fixed constant. In fact, we can easily expand our model to the more general case where $\lambda_x(s)$ is just restricted to an interval with finite bounds $l_x(s)$ and $u_x(s)$,

$$l_x(s) \leq \lambda_x(s) \leq u_x(s).$$

For that purpose, we transform the more general problem into an equivalent problem with a constant starting value: Expand the time interval $[s, \omega_x]$ to $[s - \varepsilon, \omega_x]$, let $\lambda_x(s - \varepsilon) = 0$, and define $\mathrm{d}l_x$ and $\mathrm{d}u_x$ on $(s - \varepsilon, s]$ in such a way that $\int_{(s-\varepsilon,s]} \mathrm{d}l_x = l_x(s)$ and $\int_{(s-\varepsilon,s]} \mathrm{d}u_x = u_x(s)$. Further, set $v(s-\varepsilon,s)$ to 1 and redefine the payment functions B, Π, c in such a way that there are no payments on $[s-\varepsilon,s]$. Now we can apply all results of this section on the new equivalent worst-case problem with fixed starting value $\lambda_x(s-\varepsilon) = 0$.

We now regard the prospective reserve at time s as a mapping of the mortality intensity λ_x and are looking for a scenario within the bounds (5.1) and with fixed starting value $\lambda_x(s)$ that maximizes $V(s,\lambda_x)$. Let λ_x be some starting point that is shifted by a function h that is right-continuous and has bounded variation on $[s,\omega_x]$. By applying Fubini's Theorem, (4.9) can be transformed to

$$\begin{aligned} V(s,\lambda_x+h) - V(s,\lambda_x) &= \int_{(s,\omega_x]} R_s(t)\Big(\int_{(s,t]} \mathrm{d}h(u)\Big)\mathrm{d}t \\ &= \int_{(s,\omega_x]}\Big(\int_{[u,\omega_x]} R_s(t)\mathrm{d}t\Big)\mathrm{d}h(u). \end{aligned}$$

With writing H for the cumulative version of h, we denote

$$CR_s(u) := \int_{[u,\omega_x]} R_s(u)\mathrm{d}t = \int_{[u,\omega_x]} e^{\Lambda_x(s)-\Lambda_x(t)+H(s)-H(t)}\, v(s,t)\, R(t,\Lambda_x)\mathrm{d}t$$

as *cumulative sum-at-risk at time s for occurrence of death at and after u*. This gives an expression similar to (3.3) and (4.9), that is,

$$V(s,\lambda_x+h) - V(s,\lambda_x) = \int_{(s,\omega_x]} CR_s(u)\mathrm{d}h(u). \tag{5.2}$$

The cumulative sum-at-risk $CR_s(u)$ describes the effect that a change $\mathrm{d}h(u)$ of the rate of increase of the mortality intensity (the rate of acceleration of the cumulative mortality intensity) has on $V(s)$. Analogously to Property 3.1 and Property 4.1 we get the following result:

Property 5.2. *If the shifted mortality intensity $\lambda_x + h$ is a regular hazard rate, then*

$$\operatorname{sign}(\mathrm{d}h(t)) = \operatorname{sign}(CR_s(t)) \text{ for all } t > s \implies V(s,\lambda_x+h) \geq V(s,\lambda_x) \tag{5.3}$$

and

$$\operatorname{sign}(\mathrm{d}h(t)) = -\operatorname{sign}(CR_s(t)) \text{ for all } t > s \implies V(s,\lambda_x+h) \leq V(s,\lambda_x). \tag{5.4}$$

For the proof just apply (5.2). In contrast to $\mathrm{d}S(t)$ in Property 3.1 and $R(t)$ in Property 4.1, the sign of $CR_s(t)$ depends on shift h. Thus we can not just transform the ideas of the previous sections in order to find a maximizing scenario. Again, we can get rid of the dependence on h by allowing just for local shifts. Applying Fubini's Theorem, the integral in (4.13) can be transformed to

$$\begin{aligned} &\int_{(s,\,\omega_x]} e^{\Lambda_x(s)-\Lambda_x(t)} v(s,t)\, R(t,\Lambda_x) \Big(\int_{(s,t]} \mathrm{d}h(u) \Big) \mathrm{d}t \\ &= \int_{(s,\,\omega_x]} (\nabla_{\lambda_x} V)(u) \mathrm{d}h(u), \end{aligned} \tag{5.5}$$

where $\nabla_{\lambda_x} V$ is interpreted as some form of generalized gradient defined by

$$(\nabla_{\lambda_x} V)(u) := \int_{[u,\,\omega_x]} e^{\Lambda_x(s)-\Lambda_x(t)} v(s,t)\, R(t,\Lambda_x) \mathrm{d}t.$$

If $o(\|H\|)$ in (4.13) is negligible, then the prospective reserve $V(s,\lambda_x + h)$ can be maximized by choosing $\mathrm{d}(\lambda_x + h)$ equal to $\mathrm{d}l_x$ and $\mathrm{d}u_x$ where ever $(\nabla_{\lambda_x} V)$ is negative and positive, respectively. If the confidence band (5.1) is not very narrow, $o(\|H\|)$ might not be negligible. But the first-order Taylor expansion (4.13) in the version of (5.5) allows at least to formulate a characteristic of global maxima.

Property 5.3. *Let $\bar{\lambda}_x$ be a scenario within the bounds* (5.1) *that maximizes $V(s, \bar{\lambda}_x)$. Then $\bar{\lambda}_x$ satisfies*

$$\mathrm{d}\bar{\lambda}_x(t) = \begin{cases} \mathrm{d}l_x(t) & : (\nabla_{\bar{\lambda}_x} V)(t) < 0 \\ \mathrm{d}u_x(t) & : (\nabla_{\bar{\lambda}_x} V)(t) > 0 \end{cases} \tag{5.6}$$

on $(s, \omega_x]$.

Proof. Assume that $\bar{\lambda}_x$ does not satisfy (5.6). Then

$$\int_{(s,\omega_x] \cap \{(\nabla_{\bar{\lambda}_x} V)(t) < 0\}} (\nabla_{\bar{\lambda}_x} V)(t) \mathrm{d}(l_x - \bar{\lambda}_x)(t) + \int_{(s,\omega_x] \cap \{(\nabla_{\bar{\lambda}_x} V)(t) > 0\}} (\nabla_{\bar{\lambda}_x} V)(t) \mathrm{d}(u_x - \bar{\lambda}_x)(t)$$

is strictly positive. Let $\tilde{\lambda}_x$ be defined by the right hand side of (5.6). Applying (4.13) in the version of (5.5), we obtain for $\varepsilon > 0$

$$V\big(s, \bar{\lambda}_x + \varepsilon(\tilde{\lambda}_x - \bar{\lambda}_x)\big) = V(s, \bar{\lambda}_x) + \varepsilon \underbrace{\int_{(s,\,\omega_x]} (\nabla_{\bar{\lambda}_x} V)(u)\, \mathrm{d}(\tilde{\lambda}_x - \bar{\lambda}_x)(u)}_{>0} + o(\varepsilon).$$

Now choose ε small enough such that the integral (linear Taylor term) is greater than the absolute value of the remainder $o(\varepsilon)$. Then we have $V(s, \bar{\lambda}_x + \varepsilon(\tilde{\lambda}_x - \bar{\lambda}_x)) > V(s, \bar{\lambda}_x)$, which means that $\bar{\lambda}_x$ is not maximal. □

If $\bar{\lambda}_x$ is a maximizing scenario, then the characteristic (5.6) that $\bar{\lambda}_x$ is mainly on the bounds is analogous to (3.9) and (4.12) (or (4.10)). However, a worst-case integral equation similar to (4.11) seems to be out of reach here. The crux of (4.11) is that the discontinuities at $R(t) = 0$ of the integrator (4.10) in the last integral in (4.4) are annihilated by the integrand $R(t)$. We do not have such a property for (5.6) since the signs of $CR_s(t)$ and $(\nabla_{\bar{\lambda}_x} V)(t)$ can differ. However, in order to find a worst-case scenario, we can at least design gradient ascent methods similar to the algorithms in Section 4.2:

1. Choose a starting mortality scenario $\lambda_x^{(0)}$.
2. Calculate a new scenario by using the iteration

$$d\lambda_x^{(n+1)}(t) := d\lambda_x^{(n)}(t) + K(\nabla_{\lambda_x^{(n)}} V)(t)\, dt.$$

   If the right hand side is below dl_x or above du_x, we cut $d\lambda_x^{(n+1)}$ off at dl_x or du_x, respectively. Here, $K > 0$ is some step size that has to be chosen.
3. Repeat step 2 until $|V(s, \lambda_x^{(n+1)}) - V(s, \lambda_x^{(n)})|$ is below some error tolerance.

If we increase K to infinity, we obtain the following algorithm:

1. Choose a starting mortality scenario $\lambda_x^{(0)}$.
2. Calculate a new scenario by using the iteration

$$d\lambda_x^{(n+1)} := \begin{cases} dl_x(t) & : (\nabla_{\lambda_x^{(n)}} V)(t) < 0 \\ du_x(t) & : (\nabla_{\lambda_x^{(n)}} V)(t) > 0 \\ d\lambda_x^{(n)}(t) & : (\nabla_{\lambda_x^{(n)}} V)(t) = 0 \end{cases}. \tag{5.7}$$

3. Repeat step 2 until $V(s, \lambda_x^{(n+1)}) - V(s, \lambda_x^{(n)})|$ is below some error tolerance.

If these algorithms converge, the limit satisfies (5.6). The second algorithm makes better use of the fact that maximizing scenarios are always of the form (5.6).

Remark 5.4 (Bounds for higher order derivatives). In sections 3 and 4 and in this section we looked for worst-case mortality rates with respect to confidence bounds for Λ_x, $d\Lambda_x$, and $d(\Lambda'_x) = d\lambda_x$. The ideas in this section can be generalized to confidence bounds of higher order $d(\Lambda''_x)$, $d(\Lambda'''_x)$, ... by applying Fubini's Theorem in (5.5) not only once but consecutively in order to obtain generalized gradients $(\nabla_{\Lambda''_x} V)$, $(\nabla_{\Lambda'''_x} V)$, etc. With these gradients, one can find characterizations of the maximizing scenario similar to (5.6) and design iteration methods similar to (5.7).

6. Numerical illustrations

6.1. Confidence bands for the underlying hazard

For the numerical illustrations, we assume that the hazard rate is piecewise constant and obeys the Lee-Carter model. Specifically, this means that for any integer age x and calendar year t

$$\lambda_{x+\xi}(t+\chi) = \lambda_x(t) = \exp(\alpha_x + \beta_x \kappa_t) \tag{6.1}$$

for any $0 < \xi < 1$, where the parameters β_x and κ_t are subject to constraints ensuring model identification. Here, the parameters are estimated from the mortality surface available from the Belgian Federal Planning Bureau. The ages considered here range from 35 to $\omega = 115$, and the observation period is 1970-2006. The estimated parameters are displayed in Table 1.

Lee & Carter (1992) reported that for life expectancy forecasts, it is reasonable to restrict attention to the errors in forecasting the mortality index κ_t and to ignore those in fitting the mortality matrix, even for short run forecasts. The same comment applies to actuarial present values. Therefore, we disregard the sampling errors in the α_x and β_x and concentrate on the variability relating to the mortality index κ_t.

We use Box-Jenkins techniques to forecast κ_t within an ARIMA times series model. For the data set under consideration, an ARIMA(0,1,0) process is appropriate to describe the dynamics of the κ_t. This means that the κ_t obey

$$\kappa_t = \kappa_{t-1} + \theta + \xi_t, \tag{6.2}$$

where θ is the drift parameter and the ξ_t are independent and normally distributed with common mean 0 and variance σ^2. The estimated parameters are $\hat{\theta} = -1.061997$ and $\hat{\sigma}^2 = 1.097764$. The projected κ_t are then obtained from $\hat{\kappa}_{2006}$ by adding a linear trend with slope $\hat{\theta}$.

Consider the cohort reaching age x_0 in year t_0. We take the central forecast produced by the Lee-Carter approach as reference,

$$\lambda^{\text{ref}}_{x_0+k} = \exp\Big(\alpha_{x_0+k} + \beta_{x_0+k}\big(\kappa_{t_0} + k\theta\big)\Big).$$

For our specific cohort cohort, we determine the band $(\pi_{\text{low}}\lambda^{\text{ref}}_{x_0+k}, \pi_{\text{up}}\lambda^{\text{ref}}_{x_0+k})$ such that

$$P\Big(\exp\big(\alpha_{x_0+k} + \beta_{x_0+k}\kappa_{t_0+k}\big) \notin \Big(\pi_{\text{low}}\lambda^{\text{ref}}_{x_0+k}, \pi_{\text{up}}\lambda^{\text{ref}}_{x_0+k}\Big) \text{ for some } k = 1,2,\ldots,\omega - x_0\Big) \le \varepsilon$$

for some probability level ε small enough. In order to fix the values of π_{low} and π_{up}, we require that

TABLE 1

ESTIMATED LEE-CARTER PARAMETERS FOR BELGIAN MALES, GENERAL POPULATION

Age x	α_x	β_x	Age x	α_x	β_x	Age x	α_x	β_x
35	–6.49800	0.01025	62	–4.02154	0.02170	89	–1.49859	0.00812
36	–6.44594	0.01102	63	–3.92972	0.02224	90	–1.41131	0.00780
37	–6.39713	0.01135	64	–3.83591	0.02247	91	–1.32524	0.00750
38	–6.31816	0.01165	65	–3.73619	0.02250	92	–1.24031	0.00721
39	–6.22791	0.01187	66	–3.63617	0.02283	93	–1.15656	0.00693
40	–6.13930	0.01170	67	–3.54186	0.02308	94	–1.07388	0.00665
41	–6.04465	0.01161	628	–3.44848	0.02295	95	–0.99240	0.00637
42	–5.95458	0.01188	69	–3.34902	0.02265	96	–0.91183	0.00611
43	–5.868239	0.01221	70	–3.24567	0.02232	97	–0.83212	0.00587
44	–5.77236	0.01195	71	–3.14967	0.02160	98	–0.75337	0.00563
45	–5.67342	0.01161	72	–3.05698	0.02080	99	–0.67555	0.00540
46	–5.57691	0.01182	73	–2.96229	0.02039	100	–0.59857	0.00517
47	–5.47885	0.01211	74	–2.86884	0.01991	101	–0.52240	0.00495
48	–5.37766	0.01255	75	–2.77536	0.01905	102	–0.44697	0.00473
49	–5.27580	0.01333	76	–2.67986	0.01806	103	–0.37224	0.00452
50	–5.18100	0.01436	77	–2.58440	0.01712	104	–0.29813	0.00432
51	–5.08784	0.01548	78	–2.48886	0.01632	105	–0.22463	0.00412
52	–4.98878	0.01644	79	–2.39115	0.01578	106	–0.15161	0.00393
53	–4.89342	0.01684	80	–2.29443	0.01516	107	–0.07907	0.00375
54	–4.80459	0.01691	81	–2.20478	0.01410	108	–0.00686	0.00357
55	–4.70873	0.01757	82	–2.11702	0.01288	109	0.06503	0.00340
56	–4.60393	0.01896	83	–2.02564	0.01204	110	0.13677	0.00323
57	–4.50356	0.01953	84	–1.93736	0.01112	111	0.20835	0.00307
58	–4.40951	0.01904	85	–1.85548	0.00984	112	0.28002	0.00291
59	–4.31335	0.01947	86	–1.76928	0.00901	113	0.35179	0.00276
60	–4.21347	0.02087	87	–1.67766	0.00870	114	0.42390	0.00261
61	–4.11588	0.02145	88	–1.58707	0.00844	115	0.49639	0.00246

Time t	κ_t	Time t	κ_t	Time t	κ_t	Time t	κ_t
1970	16.75929	1980	9.62899	1990	–3.48992	2000	–12.04364
1971	16.57651	1981	8.24881	1991	–4.16356	2001	–13.49205
1972	16.08087	1982	6.94314	1992	–4.64450	2002	–14.04277
1973	15.55938	1983	7.22406	1993	–4.56888	2003	–15.24265
1974	13.54655	1984	5.61148	1994	–6.14456	2004	–18.81096
1975	14.04627	1985	5.07589	1995	–5.79617	2005	–19.34756
1976	13.68657	1986	3.34483	1996	–7.75518	2006	–21.47259
1977	9.97551	1987	0.16340	1997	–9.24917		
1978	11.02624	1988	–0.59693	1998	–10.28659		
1979	9.85718	1989	–1.62033	1999	–10.58694		

$$P\Big(\ln \pi_{\text{up}} \geq \beta_{x_0+k}\Big(\kappa_{t_0+k} - \big(\kappa_{t_0} + k\theta\big)\Big) \geq \ln \pi_{\text{low}} \text{ for all } k = 1, 2, \ldots, \omega - x_0\Big) \geq 1 - \varepsilon.$$

The bounds can now be determined as quantiles of the random vector

$$\Big(\beta_{x_0+1}\Big(\kappa_{t_0+1} - \big(\kappa_{t_0} + \theta\big)\Big), \ldots, \beta_\omega\Big(\kappa_{t_0+\omega-x_0} - \Big(\kappa_{t_0} + \big(\omega - x_0\big)\theta\Big)\Big)\Big)^T,$$

which is multivariate normal with zero mean and variance-covariance matrix

$$\tilde{\Sigma} = \begin{pmatrix} \sigma^2\beta^2_{x_0+1} & \sigma^2\beta_{x_0+1}\beta_{x_0+2} & \cdots & \sigma^2\beta_{x_0+1}\beta_\omega \\ \sigma^2\beta_{x_0+1}\beta_{x_0+2} & 2\sigma^2\beta^2_{x_0+2} & \cdots & 2\sigma^2\beta_{x_0+2}\beta_\omega \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2\beta_{x_0+1}\beta_\omega & 2\sigma^2\beta_{x_0+2}\beta_\omega & \cdots & (\omega - x_0)\sigma^2\beta^2_\omega \end{pmatrix}.$$

Consider the generation aged $x_0 = 35$ in year $t_0 = 2006$. Imposing that the future life table should be in the band $(\pi_{\text{low}}\lambda^{\text{ref}}_{x_0+k}, \pi_{\text{up}}\lambda^{\text{ref}}_{x_0+k})$ with a probability of at least 99%, we get $\pi_{\text{low}} = 0.9396105$ and $\pi_{\text{up}} = 1.064271$. These values have been found using the qmvnorm function of the R package mvtnorm.

6.2. Annuity with death benefits

Consider an annuity insurance with additional death benefits. A constant premium is paid yearly in advance from age 35 on till retirement at age 65. From then on a constant annuity benefit of 1 is paid yearly in advance till death. The functions Π and B are thus given by

$$\Pi(t) = \text{const} \sum_{k=0}^{34} \mathbf{1}_{[k,\infty)}(t), \;\; B(t) = \sum_{k=35}^{\omega_x} \mathbf{1}_{[k,\infty)}(t).$$

If the policyholder dies before age 65, a death benefit is paid that has the size of the prospective reserve just before retirement. If the policyholder dies after retirement but before age 85, a death benefit is paid that equals the prospective reserve at that time. The function c is thus given by

$$c(t) = V(30-) \cdot \mathbf{1}_{[0,30)}(t) + V(t-) \cdot \mathbf{1}_{[30,50)}(t).$$

We assume that the yearly interest rate is at 2.25% and that interest is paid continuously with an intensity of $\phi(t) = \ln(1.0225)$. For our exemplary calculations we use the mortality intensity $\lambda^{\text{ref}}_{x_0+k}$ derived from (6.1) with parameters estimated from Belgian mortality statistics as best estimate and confidence bands where

(A) the lower and upper bound are -6.03895% and $+6.4271\%$ below and above the best estimate as obtained from the coefficients π_{low} and π_{up}.
(B) the lower and upper bound are -25% and $+15\%$ below and above the best estimate as suggested in Consultation Paper no. 49 of the Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS) for the Solvency II project.

The equivalence principle and the best estimate mortality rate yield a constant yearly premium of 0.51059236. Figure 6.1 shows the death benefit function

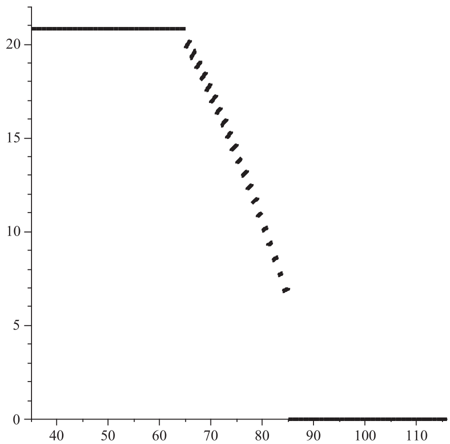

FIGURE 6.1: Death benefit function $c(t)$.

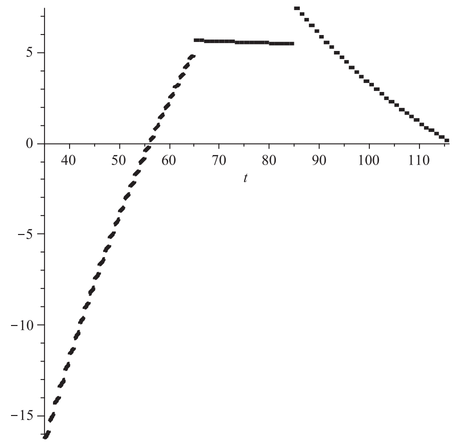

FIGURE 6.2: Approach I: Cumulative survival cost $S_0(t)$.

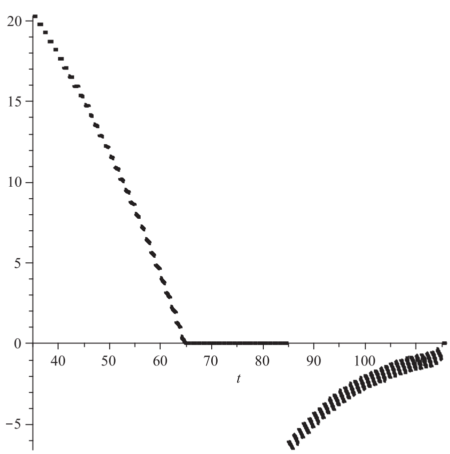

FIGURE 6.3: Approach II: Sum-at-risk $R(t)$ with respect to the best estimate mortality rate

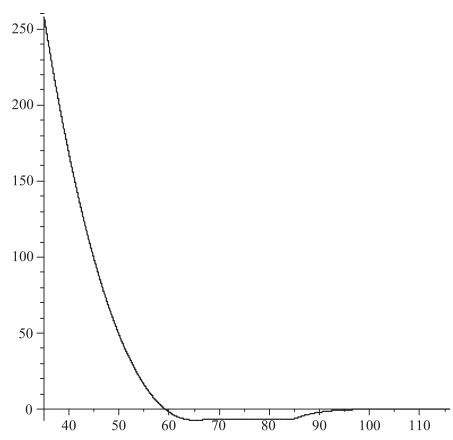

FIGURE 6.4: Approach III: Cumulative sum-at-risk $CR_0(t)$ with respect to the best estimate mortality rate and zero shift

$c(t)$. Figures 6.2, 6.3, and 6.4 illustrate the cumulative survival cost $S_0(t)$ at time zero for survival at and after time t, the sum-at-risk $R(t)$ for occurrence of death at time t, and the cumulative sum-at-risk $CR_0(t)$ at time zero for occurrence of death at and after time t. All illustrations are based on the best estimate mortality scenario. For the calculation of $CR_0(t)$, we assumed that the shift H is zero, which implies that the cumulative sum-at-risk is equal to the generalized gradient $(\nabla_{\lambda_x} V)(t)$ here. The following table shows the prospective reserve $V(0-)$ before beginning of the contract with respect to different mortality scenarios:

valuation basis	prospective reserve $V(0-)$ w.r.t. confidence band (A)	prospective reserve $V(0-)$ w.r.t. confidence band (B)
best estimate	0.000000	0.000000
lower bound	0.007703	0.084987
upper bound	–0.001460	0.005674
separated contract	0.324404	1.112690
worst-case method I	0.163750	0.594622
worst-case method II	0.084025	0.323160
worst-case method III	0.079010	0.296063
sum-at-risk method	0.082006	0.285200
2 × sum-at-risk method	0.082772	0.297097
3 × sum-at-risk method	0.084012	0.322808

'Separated contracts' means that the policy is unbundled into an annuity policy and a temporary life insurance policy, and the two parts are valuated on the basis of the lower bound and the upper bound. The results '2× sum-at-risk method' and '3× sum-at-risk method' are obtained by applying the sum-at-risk method iteratively. The rows 'worst-case method I, II, and III' refer to the methods of sections 3, 4, and 5. The bounds are given by

$$L_x(t) - L_x(s) = \pi_{\text{low}}(\Lambda_x(t) - \Lambda_x(s)),\ U_x(t) - U_x(s) = \pi_{\text{up}}(\Lambda_x(t) - \Lambda_x(s))$$

for approach I, by

$$l_x(t) = \pi_{\text{low}}\,\lambda_x(t),\ u_x(t) = \pi_{\text{up}}\,\lambda_x(t)$$

for approach II and the sum-at-risk method, and by

$$\mathrm{d}l_x(t) = \Delta l_x(t) = \min\{\pi_{\text{up}}\Delta\lambda_x(t), \pi_{\text{low}}\Delta\lambda_x(t)\},$$
$$\mathrm{d}u_x(t) = \Delta u_x(t) = \max\{\pi_{\text{up}}\Delta\lambda_x(t), \pi_{\text{low}}\Delta\lambda_x(t)\}$$

for integers $t = 1, \ldots, \omega_x$, and $\mathrm{d}\,l_x(t) = \mathrm{d}u_x(t) = 0$ otherwise for approach III. The last approach additionally needs a starting value for λ_x at outset. Using Remark 5.1, we let $\lambda_x(0) \in [\pi_{\text{low}}\lambda_x(0), \pi_{\text{up}}\lambda_x(0)]$. Although section 6.1 generates π_{up} and π_{low} only for the hazard concept (approach II), it still makes sense to employ them also for the other two concepts: In case of approach I, we justify that with the linearity of integration. In case of approach III, the restriction $\lambda_x(0) \in [\pi_{\text{low}}\lambda_x(0), \pi_{\text{up}}\lambda_x(0)]$ and the fact that $\Delta\lambda_x(t) \geq 0$ (at least in our example) imply that the set of scenarios that we are maximizing over for approach III is a true subset of the set of scenarios that we are maximizing over for approach II. Hence, π_{up} and π_{low} represent a conservative choice for approach III.

The numbers in the table above can be seen as (maximal) losses due to mortality rate fluctuations. (For their interpretation note that the yearly annuity benefit is 1). We see that the throughout lower and throughout upper mortality scenario have a rather small impact compared to the mixed scenarios that are produced by the worst-case methods or the sum-at-risk method. As a consequence of this, actuaries often separate policies with both, survival and death benefits, into two parts with either survival or death benefits. But we see in our example that the separation concept significantly overestimates the risk. The maximal losses of the worst-case scenarios I to III show a downward order, which is due to the fact that (at least in our example) the set of scenarios that we are maximizing over is monotonic decreasing. As predicted at the end of section 4, the iteration of the sum-at-risk method shows a convergence to the worst-case according to approach II.

The following table shows the (numerically calculated) worst-case mortality scenarios, which all are at any time t either equal to a bound or equal to the best estimate:

ages where the valuation basis is equal to ...	... the upper bound of confidence band (A)	... the best estimate	... the lower bound of confidence band (A)
best estimate		(35,116)	
worst-case method I	(35,65] ∪ (84,85]		(65,84] ∪ (85,116]
worst-case method II	(35,64.792)		(64.792, 116)
worst-case method III	{35, …, 59}		{60, …, 116}
sum-at-risk method	(35,65)	(65,85)	(85,116)
ages where the valuation basis is equal to ...	**... the upper bound of confidence band (B)**	**... the best estimate**	**... the lower bound of confidence band (B)**
best estimate		(35,116)	
worst-case method I	(35,65] ∪ (84,85]		(65,84] ∪ (85,116]
worst-case method II	(35,63.925)		(63.925, 116)
worst-case method III	{35, …, 57}		{58, …, 116}
sum-at-risk method	(35,65)	(65,85)	(85,116)

Worst-case method I sets the cumulative mortality intensity equal to the upper and lower bound where ever the cumulative survival cost S_0 is increasing and decreasing. The shifting times are independent of the confidence band, because S_0 does not depend on the bounds. Note that we have two negative jumps for the cumulative mortality intensity. The sum-at-risk method sets the mortality intensity equal to the upper and lower bound where ever the sum-at-risk R (with respect to the best estimate) is positive and negative. We see that the result differs from the true worst-case calculated by worst-case method II. For the sum-at-risk method, the shifting times do not depend on the confidence band, but for worst-case method II they do. For worst-case method III, our choice of the bounds $\mathrm{d}l_x$ and $\mathrm{d}u_x$ allows only for changes of λ_x at integer times. The shift between high and low rate of increase of the mortality intensity is for both confidence bands earlier than the shift between high and low mortality intensity in method II, which is a result from the facts that the cumulative sum-at-risk aggregates the future sums-at-risk and that the sum-at-risk is first throughout positive and then throughout negative. While the mortality intensity of method II shows an extreme jump from u_x to l_x near the age of retirement, the mortality intensity of method III evolves gradually with small steps Δl_x or Δu_x.

7. Conclusion

In this paper, we have shown how to construct a first-order life table that represents the worst mortality scenario from the insurer's point of view within all scenarios contained in a given confidence region. By worst mortality scenario we mean in this paper the one that maximizes the prospective reserve for given benefits and premiums. Three approaches are offered for the calculation of first-order mortality bases, which differ in their restrictions on the set of admissible mortality scenarios. The confidence region can be given by confidence bands for either the cumulative hazard, its rate of increase (i.e., the hazard rate itself), or its acceleration (i.e. the rate of increase of the hazard rate). We leave it to the practitioner to decide which a priori assumptions he is willing to accept.

Contrarily to other methods that can be found in the literature, the first-order mortality basis determined in this paper are on the safe side with respect to a confidence band. This allows for an explicit link with statistical inference to quantify the risk that the insurer suffers adverse mortality experience (by means of the confidence level).

The maximization problems that arise in the first and second approach are fully solved, whereas the maximization problem of the third approach is only tackled with an approximative procedure. The derivation of an exact method also for the third approach could also be of theoretical interest.

The worst-case calculations performed in this paper are based on single insurance contracts. The results remain still valid for homogeneous portfolios, but we do not consider heterogeneous portfolios. Extending the results to heterogeneous portfolios is an interesting field for future research, especially in view of applications such as Solvency II.

Section 6.1 demonstrates a well-known technique for the calculation of hazard rate bounds of the form (4.2). Under some reasonable additional conditions, the results for the hazard rate can be used to create also bounds for cumulative hazard rates (cf. (3.6)) or for the rate of increase of hazard rates (cf. (5.1)). However, it is still an open question how to obtain *sharp* bounds in the latter case.

Acknowledgements

The Authors would like to express their gratitude to an associate editor and two anonymous referees whose comments have been extremely useful to revise a previous version of the present work.

This work has been performed while the first author visited the UCLouvain supported by a DFG research grant. The warm hospitality of the Institute of Statistics is gratefully acknowledged.

Michel Denuit acknowledges the financial support of the *Banque Nationale de Belgique* under grant "Risk measures and Economic capital", and of the *Onderzoeksfonds K.U.Leuven* (GOA/07: Risk Modeling and Valuation of Insurance and Financial Cash Flows, with Applications to Pricing, Provisioning and Solvency).

References

Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A. and Nesbitt, C.J. (1997) *Actuarial Mathematics, 2nd edition*. The Society of Actuaries, Schaumburg, Illinois.

Christiansen, M.C. (2008a) A sensitivity analysis concept for life insurance with respect to a valuation basis of infinite dimension. *Insurance: Mathematics and Economics* **42**, 680-690.

Christiansen, M.C. (2008b) A sensitivity analysis of typical life insurance contracts with respect to the technical basis. *Insurance: Mathematics and Economics* **42**, 787-796.

Christiansen, M.C. (2010) Biometric worst-case scenarios for multi-state life insurance policies. *Insurance: Mathematics and Economics* **47**, 190-197.

Christiansen, M.C. and Helwich, M. (2008) Some further ideas concerning the interaction between insurance and investment risks. *Blätter der DGVFM* **29(2)**, 253-266.

Dienst, H.-R. (1995) *Zur aktuariellen Problematik der Invaliditätsversicherung*. Verlag Versicherungswirtschaft, Karlsruhe.

Hoem, J.M. (1988) The versatility of the Markov chain as a tool in the mathematics of life insurance. *Transactions of the 23rd International Congress of Actuaries R*, 171-202.

Kalashnikov, V. and Norberg, R. (2003) On the Sensitivity of Premiums and Reserves to Changes in Valuation Elements. *Scandinavian Actuarial Journal* **3**, 238-256.

Lee, R.D. and Carter, L.R. (1992) Modelling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* **87**, 659-671.

Lidstone, G.J. (1905) Changes in pure premium values consequent upon variations in the rate of interest or rate of mortality. *Journal of the Institute of Actuaries* **39**, 209-252.

Linnemann, P. (1993) On the application of Thiele's differential equation in life insurance. *Insurance: Mathematics and Economics* **13**, 63-74.

Norberg, R. (1985) Lidstone in the continuous case. *Scandinavian Actuarial Journal*, 27-32.

RAMLAU-HANSEN, H. (1988) The emergence of profit in life insurance. *Insurance: Mathematics and Economics* **7**, 225-236.

MARCUS C. CHRISTIANSEN
Institut für Versicherungswissenschaften
Universität Ulm
D-89069 Ulm
Germany

MICHEL M. DENUIT
Institut de Statistique, Biostatistique et Sciences Actuarielles
Université Catholique de Louvain
B-1348 Louvain-la-Neuve
Belgium