Safety margins for unsystematic biometric risk in life and health insurance

Marcus C. Christiansen

Preprint Series: 2011-04

Fakultät für Mathematik und Wirtschaftswissenschaften

UNIVERSITÄT ULM

Safety margins for unsystematic biometric risk in life and health insurance

Marcus C. Christiansen, University of Ulm, Germany

Abstract: In multistate life and health insurances, the pattern of states of the policyholder is random, thus exposing the insurer to an unsystematic biometric risk. For this reason safety margins are added on premiums and reserves. But in contrast to non-life insurance, traditionally the safety margins are not chosen explicitly but implicitly in form of a valuation basis of first order. If we define the implicit margins bottom-up, i.e. on the basis of the probability distribution of the random pattern of states of the policyholders, we are not able to control the level of safety that we finally reach for premiums and reserves. If we use a top-down approach, that means that we directly calculate explicit margins for premiums and reserves and then choose implicit safety margins that correspond to the explicit margins, we are able to control the total portfolio risk, but we have the problem that it is unclear how to allocate the total margin to partial margins for different transitions at different ages. Although the allocation of the total margin to the partial (implicit) margins is not relevant for the total portfolio risk, we have to pay attention since it has a great effect on the calculation of surplus.

In this paper we calculate asymptotic probability distributions for premiums and reserves of second order by using the functional delta method. As a result, we can not only determine the actual level of safety that is induced by given implicit safety margins, but we can also linearly decompose the total randomness of a portfolio to contributions that the different transition rates at different ages make to the total uncertainty. As a result we do not only get new insight into the sources of unsystematic biometric risk, but we also obtain a useful tool that allows to construct reasonable principles for the allocation of the total safety margin to implicit margins with respect to transitions and ages.

Keywords: implicit safety margins; valuation basis of first order; functional delta method; life insurance; health insurance

1. Introduction

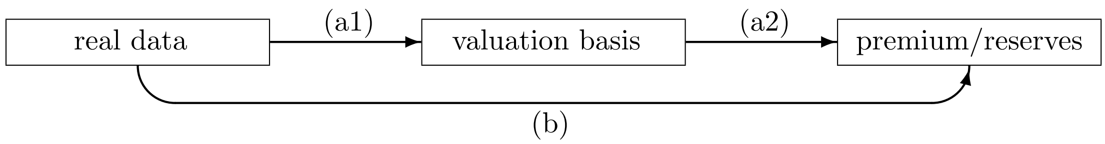
The insurance risk of a life insurer consists mainly of the three following risk components:

- **unsystematic biometric risk**: the probability distribution of the random pattern of states of the policyholder is perfectly known, but payments between insurer and policyholder may deviate from expected values,
- **biometric estimation risk**: the probability distribution of the random pattern of states of the policyholder is estimated from data, but because of the finite sample size the estimated probabilities may deviate from the true probabilities,

- **systematic biometric risk**: the probability distribution of the random pattern of states of the policyholder is perfectly known for the past and the present, but there may be systematic demographic changes that are not anticipated today.

This three risk components are similarly relevant for a health insurer whose calculation techniques are pursued on a similar technical basis to that of life insurance. For each of the three risks, actuaries add safety margins to premiums and reserves in order to make sure that insurers are able to meet their future obligations. In the present paper we solely focus on safety margins for the unsystematic biometric risk.

In non-life insurance, safety margins are traditionally calculated explicitly by means of the probability distribution of the future portfolio loss. In life insurance, the usual way is to add safety margins implicitly by adding margins on the valuation basis, that is the set of assumptions underlying the calculation of premiums and reserves. The idea is to overestimate and underestimate future interest rates, transition probabilities, and costs in such a way that they represent a worst case development for the insurer. The result is a so called valuation basis of first order, which is then used for the calculation of premiums and reserves on the safe side. The disadvantages of this implicit approach are, first, that it is generally not clear which level of safety is really reached by employing the first order valuation basis and, second, that it hardly takes respect of the size of impact that changes of different transition frequencies at different points of time really have on a portfolio. The following figure illustrates the main difference between (a) the implicit and (b) the explicit approach.



The implicit approach has two steps (a1) and (a2). First, the valuation basis of first order is estimated from real data. This can be a point estimation or, preferably, a confidence estimation. However, for the calculation of premiums and reserves in the next step (a2) we need a single scenario and can not deal with a confidence set of scenarios. Therefore, we identify a single scenario within the confidence set that represents some worst case. On the basis of this worst case, we then calculate premiums and reserves on the safe side. Note that in general it can be rather difficult to identify a worst case scenario within the confidence set, see Christiansen (2010). Once we have and apply such a worst-case scenario, we know that after the second step (a2) the confidence level is not less than the confidence level of the first step (a1). Even if we can cope with the difficulties to find a worst-case valuation basis, the two-step concept still has the disadvantage that the confidence estimation in the first step (a1) is not necessarily performed optimally in the sense that premiums and reserves are minimal with respect to a prespecified confidence level. We can be way above the wanted confidence level without knowing.

A reaction to that problem is the top-down method (see, for example, Bühlmann (1985)). The idea is to start with step (b) and then to do step (a2) in reverse in such a ways that we end up with a valuation basis of first order that meets the prespecified confidence level. In order to perform the top-down method we have to solve two problems. First, for doing step (b), we must know the probability distribution of the future portfolio loss (at least approximately). Second, for doing a reversed step (a2), we need a concept that tells us how to allocate the total safety margin among the margins on different transitions at different ages. For simple life insurances where the only possible transition is from active to dead, we find answers for both problems for example in

Pannenberg (1997) or Pannenberg and Schütz (1998). In case of insurance contracts with multiple states, the insurance literature still offers methods for the first problem, but, unfortunately, no satisfying answers to the second problem. In this paper we present solutions to both problems in a general multistate case by applying the functional delta method.

The idea to use the delta-method for calculating the probability distribution of reserves already appeared in Kalashnikov and Norberg (2003). The two authors applied the vector-valued delta method on parametric models for the valuation basis. By applying the functional delta method on a nonparametric model for the transition intensities, we do not only avoid the risk of choosing the wrong parametric model, but moreover obtain a decomposition of the total randomness to the randomness contributions of the different transitions at different ages. On the basis of that decomposition we are able to construct concepts for a top-down approach (in particular a reversed step (a2)) for policies with multiple states.

2. Motivating examples

In order to make the problems of common calculation methods more clear, we start with three examples for which we discuss safety margins for unsystematic biometric risk.

At first we look at a homogeneous portfolio of $m \in \mathbb{N}$ **pure endowment insurances** with a contract period of n years and a survival benefit of 1. All that we need to know about the insured is their time of death. Let x be the starting age for all policies and $T_x^1, ..., T_x^m$ the corresponding remaining lifetimes, which shall be stochastically independent and identical distributed with survival function $S_x(t) = {}_tp_x = P(T_x^1 > t)$. For a given constant interest rate r, the present portfolio value of benefits is at time zero equal to

$$ {}_nB_x^{(m)} = \sum_{p=1}^{m} \mathbf{1}_{T_x^p > n} \, (1+r)^{-n} \, . $$

The corresponding portfolio lump sum premium according to the classical principle of equivalence is

$$ m \, {}_nE_x^* = \mathbb{E}({}_nB_x^{(m)}) = \sum_{p=1}^{m} S_x^*(n) \, (1+r)^{-n} \, , $$

where $S_x^*(t)$ is some survival function of first-order. In order to be on the safe side with respect to unsystematic biometric risk, the implicit method seeks to define $S_x^*(n)$ in such a way that

$$ P\bigg(\sum_{p=1}^{m} \mathbf{1}_{T_x^p > n} \leq m \, S_x^*(n) \bigg) = \alpha \tag{2.1} $$

for some given confidence level $\alpha \in (0,1)$. The implicit safety margin s^{impl} is then defined by $S_x^*(n) = (1 + s^{impl}) \, S_x(n)$. Since we know that $\sum_{p=1}^{m} \mathbf{1}_{T_x^p > n}$ has a binomial distribution with parameters m and $S_x(n)$, the parameter s^{impl} can be easily calculated. As

$$ P\bigg(\sum_{p=1}^{m} \mathbf{1}_{T_x^p > n} \leq m \, S_x^*(n) \bigg) = P\big({}_nB_x^{(m)} \leq m \, {}_nE_x^* \big) \, , \tag{2.2} $$

the confidence level α of the valuation basis $S_x^*(n)$ is also the confidence level of the lump sum net premium ${}_nE_x^*$, and the implicit safety margin s^{impl} equals the explicit safety margin s^{expl}, where the latter is defined by ${}_nE_x^* = (1 + s^{expl})\, {}_nE_x$. Thus, ${}_nE_x^*$ is the minimal net portfolio premium which at least reaches confidence level α. All in all, the equivalence of implicit and explicit safety margin justifies the use of the implicit method instead of the explicit method.

The situation gets more complex if we look at a portfolio of **annuities** that pay constant yearly benefits of 1,

$$\ddot{a}_x = \sum_{k=0}^{\infty} {}_kE_x \,.$$

According to the implicit method, we choose a first-order survival function $S_x^*(k) = (1+s^{impl}(k))\, S_x(k)$ that satisfies

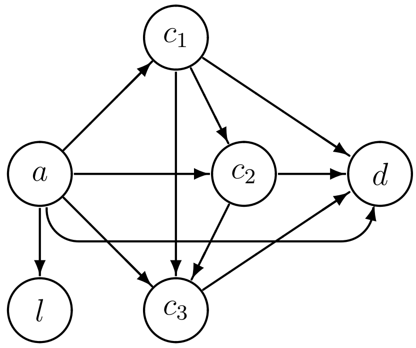
$$P\bigg(\sum_{p=1}^{m} \mathbf{1}_{T_x^p > k} \leq m\, S_x^*(k) \,, k = 1, 2, ... \bigg) = \alpha \,. \tag{2.3}$$

In contrast to (2.1), we usually have not only one but infinitely many solutions. For these solutions the corresponding first-order net portfolio premiums $m\, \ddot{a}_x^* = m\, (1+s^{expl})\, \ddot{a}_x$ lead to different levels of safety

$$P\bigg(\sum_{p=1}^{m} \sum_{k=1}^{\infty} \mathbf{1}_{T_x^p > k}\, (1 + r)^{-n} \leq m\, \ddot{a}_x^* \bigg) = \widetilde{\alpha} \,. \tag{2.4}$$

We can prove that $\widetilde{\alpha} \geq \alpha$, but we are not necessarily close to α. The lower bound property $\widetilde{\alpha} \geq \alpha$ can still justify the use of the implicit method, but our net premium might be significantly greater then needed, and the bad thing is that we do not realize that. A possible way out is the top-down method as described in Pannenberg and Schütz (1998). The idea is to start with (2.4) and to find just that $\ddot{a}_x^*$ for which $\widetilde{\alpha} = \alpha$. In order to do that, the distribution of the double sum (2.4) could by approximated with the help of the Central Limit Theorem (see, for example, Pannenberg (1997)). In a second step we choose the survival function $S_x^*(k)$ for all $k \in \mathbb{N}$ in such a way that the corresponding net premium equals our calculated $\ddot{a}_x^*$. Again we have the problem that there are infinitely many possible solutions for the survival function. Thus, we need a principle that tells us how to allocate the total margin among the partial margins $s^{impl}(k)$, $k \in \mathbb{N}$. For example we could say that the partial margins $s^{impl}(k)$ should be proportional to the contributions that the partial sums $\sum_{p=1}^{m} \mathbf{1}_{T_x^p > k}\, (1+r)^{-n}$ make to the variance of $\sum_{p=1}^{m} \sum_{k=1}^{\infty} \mathbf{1}_{T_x^p > k}\, (1+r)^{-n}$. However, the partial sums are not stochastically independent and it is not absolutely clear how to measure the contributions of the partial sums to the total variance. Furthermore, beside the variance there are many other risk measures that we could use. Pannenberg and Schütz (1998) give a long list of reasonable principles for the allocation of the total safety margin to the implicit margins $s^{impl}(k)$, $k \in \mathbb{N}$. In the present paper we will give a linear decomposition of the total randomness of the future portfolio loss to stochastically independent contributions with respect to age. On the basis of that result, we will be able to present a clear and intuitive principle for the allocation of the total margin to implicit margins.

Our third example is a **long term care insurance**. The need for care due to the frailty of an insured is classified according to the individuals ability to take care of himself. In Germany three different levels of frailty are commonly used. Lets say we have a state space of $\mathcal{S} = \{a =$ active/healthy$, c_1 =$ need for basic care$, c_2 =$ need for medium care$, c_3 =$ need for comprehensive care$, l =$ lapsed/canceled$, d =$ dead$\}$.



The pattern of states of the policyholders are modeled as Markovian jump processes $(X_t^p)_{t\geq 0}$ that for any time t give the actual state of insured p. Suppose that constant yearly annuities of $R_1 < R_2 < R_3$ are paid as long as the policyholder is in state c_1, c_2, c_3. Then the present portfolio value of benefits is at time zero equal to

$$B_x = \sum_{p=1}^{m} \sum_{i\in\{1,2,3\}} \sum_{k=1}^{\infty} R_i \, \mathbf{1}_{X_k^p = c_i} \, (1+r)^{-k} \, . \tag{2.5}$$

Given that the policyholders are all in state a at beginning of the contracts, its mean is

$$\mathbb{E}(B_x) = \sum_{p=1}^{m} \sum_{i\in\{1,2,3\}} \sum_{k=1}^{\infty} R_i \, P(X_k^p = c_i | X_0^p = a) \, (1+r)^{-k} \, .$$

According to the implicit method, we start with calculating first-order probabilities $P^*(X_t^p = j | X_s^p = i)$ for all times s, t and all states i, j (or, equivalently, the corresponding transition intensities) by using a generalized multistate version of (2.3). Note that we have to pay attention here to the question whether to overestimate or to underestimate the second-order transition probabilities (or second-order transition intensities). On the basis of the first-order probabilities we can calculate a first-order lump sum portfolio premium $m\,\mathbb{E}^*(B_x) = m\,(1 + s^{expl})\,\mathbb{E}(B_x)$. However, we do not know the level of safety that is implied by $m\,\mathbb{E}^*(B_x)$. If we want to meet a prespecified confidence level, we might again try to perform a top-down method. With the help of the Central Limit Theorem, we can (approximately) choose $m\,\mathbb{E}^*(B_x)$ as a quantile of (2.5). The task is then to allocate the total safety margin to implicit margins on the transition probabilities $P^*(X_t^p = j | X_s^p = i)$ (or the corresponding transition intensities). It seems a natural idea that the implicit safety margin in $P^*(X_t^p = j | X_s^p = i)$ should be somehow related to the risk contribution that the uncertainty about the number of transitions from i to j during $[s, t]$ makes to the total randomness of (2.5). But in our multistate example it is not clear how to decompose (2.5) with respect to transitions (i, j) and time intervals $[s, t]$, and we do not know of satisfying answers in the literature. In the present paper we will derive a useful decomposition with respect to transitions and time from the functional delta method. On the basis of that decompositions we will be able to construct reasonable top-down principles.

3. Modeling the random pattern of states

For life insurances and those health insurances, that are pursued on a similar technical basis to that of life insurance, contractual guaranteed payments between insurer and policyholder are defined as deterministic functions of time and of the pattern of states of the policyholder. Therefore, at first we need a model for that pattern of states.

3.1. Single policy

Consider an insurance policy that is issued at time 0, terminates at a fixed finite time T, and is driven by a Markovian jump process $(X_t)_{t\geq 0}$ with finite state space $\mathcal{S}$ and deterministic initial state $X_0 \in \mathcal{S}$. Let $J := \{(j,k) \in \mathcal{S}^2 \,|\, j \neq k\}$ denote the transition space. Following Milbrodt and Stracke (1997), the probability distribution of the Markov process is uniquely described by cumulative transition intensities $q_{jk}(t)$, $(j,k) \in J$. According to Andersen et al. (1991, section II.6.), the transition probability matrix

$$p(s,t) := \Big(P(X_t = k \,|\, X_s = j)\Big)_{(j,k)\in\mathcal{S}^2}, \qquad 0 \leq s \leq t\,,$$

has the representation

$$p(s,t) = \prod_{(s,t]} \big(1 + \mathrm{d}q\big) := \lim_{\max|t_i - t_{i-1}| \to 0} \prod \big(\mathbb{I} + q(t_i) - q(t_{i-1})\big) \tag{3.1}$$

for partitions $s < t_0 < t_1 < \cdots < t_n = t$ if the q_{jk} are nondecreasing càdlàg (right-continuous with left-hand limits) functions, zero at time zero, $q_{jj} = -\sum_{k\neq j} q_{jk}$, and $\Delta q_{jj}(t) \geq -1$ for all t. If q is absolutely continuous, i.e.

$$q_{jk}(s) = \int_0^s \mu_{jk}(t)\,\mathrm{d}t\,, \quad s \geq 0\,,$$

for some integrable functions μ_{jk}, then the probability distribution of $(X_t)_{t\geq 0}$ can also be uniquely defined by the so-called transition intensities μ_{jk}, $(j,k) \in J$. In the present paper we throughout assume that the transitions intensities μ_{jk} always exist, but nevertheless we will still need the more general cumulative modeling concept when it comes to empirical probability distributions of $(X_t)_{t\geq 0}$.

On the basis of the Markovian jump process $(X_t)_{t\geq 0}$ we additionally define the indicator processes

$$I_i(t) := \mathbf{1}_{\{X_t = i\}}\,, \quad i \in \mathcal{S}\,,$$

and the counting processes

$$N_{jk}(t) := \#\big\{\tau \in (0,t] \,\big|\, X_\tau = k, X_{\tau-} = j\big\}\,, \quad (j,k) \in J\,,$$

which give the number of jumps from j to k till time t.

3.2. Homogeneous portfolio

Now let there be $m \in \mathbb{N}$ policies with stochastically independent and identically distributed jump processes $(X_t^1)_{t\geq 0}, ..., (X_t^m)_{t\geq 0}$. We use the same notation as in the previous section but add superscripts $p = 1, ..., m$ that signify the different policies. Moreover, let

$$I_i^{(m)}(t) := \sum_{p=1}^m I_i^p(t)\,, \quad N_{jk}^{(m)}(t) := \sum_{p=1}^m N_{jk}^p(t)\,.$$

We write $I^{(m)}(t)$ for the row vector with entries $I_i^{(m)}(t)$. The paths of the counting processes $N_{jk}^{(m)}(t)$ are almost surely of bounded variation on compacts since the jump processes $(X_t^1)_{t\geq 0}, ..., (X_t^m)_{t\geq 0}$

have almost surely just a finite number of jumps on compacts (see Milbrodt and Helbig (1999), p. 174). The so-called Nelson-Aalen estimator for the cumulative interest intensities q_{jk} is defined as

$$Q^{(m)}_{jk}(t) := \int_{(0,t]} \frac{1}{I^{(m)}_j(u-)}\,\mathrm{d}N^{(m)}_{jk}(u)\,, \tag{3.2}$$

see Andersen et al. (1991, section IV.1). Let $Q^{(m)}(t)$ be the quadratic matrix that has the entries $Q^{(m)}_{jk}(t)$ for $j \neq k$ and the diagonal entries $Q^{(m)}_{jj}(t) = -\sum_{l\in\mathcal{S}\setminus\{j\}} Q^{(m)}_{jl}(t)$.

Proposition 3.1. *We almost surely have*

$$I^{(m)}(t) = I^{(m)}(s) \prod_{(s,t]} (1 + \mathrm{d}Q^{(m)}) \tag{3.3}$$

for all $0 \leq s < t \leq T$.

Proof. Since the jump processes $(X^1_t)_{t\geq 0}, ..., (X^m_t)_{t\geq 0}$ have almost surely just a finite number of jumps on compacts, and because of the multiplicative structure of product integration, without loss of generality we may assume that $Q^{(m)}$ has only one jump in $(s,t]$. We suppose the jump is at $u \in (s,t]$. Then we have $I^{(m)}(t) = I^{(m)}(u)$ and $I^{(m)}(s) = I^{(m)}(u-)$, and (3.3) is equivalent to

$$I^{(m)}(u) = I^{(m)}(u-)\,(\mathbb{I} + \Delta Q^{(m)}(u))\,.$$

Indeed, for the i-th column on the right hand side we can show that

$$\begin{aligned}
&\big[I^{(m)}(u-)\,(\mathbb{I} + \Delta Q^{(m)}(u))\big]_i \\
&= \sum_{j:j\neq i} I^{(m)}_j(u-)\Delta Q^{(m)}_{ji}(u) + I^{(m)}_i(u-)\Big(1 - \sum_{j:j\neq i} \Delta Q^{(m)}_{ij}(u)\Big) \\
&= \sum_{j:j\neq i} \Delta N^{(m)}_{ji}(u) + I^{(m)}_i(u-) - \sum_{j:j\neq i} \Delta N^{(m)}_{ij}(u) \\
&= I^{(m)}_i(u)\,.
\end{aligned}$$

□

With the help of the functional delta method, it is possible to calculate an asymptotic distribution of $(Q^{(m)}_{jk})_{(j,k)\in J}$. According to Andersen et al. (1991, Example IV.1.9), we have the following proposition.

Proposition 3.2. *Let the cumulative transition intensity matrix q be absolutely continuous with density function μ, and let $P(X_t = j) > 0$ for all $j \in \mathcal{S}$ and $t \in (0,T)$. Then we have*

$$\sqrt{m}\big(Q^{(m)}_{jk} - q_{jk}\big)_{(j,k)\in J} \xrightarrow{d} (U_{jk})_{(j,k)\in J}\,, \quad m \to \infty\,, \tag{3.4}$$

on $[0,T]$, where $(U_{jk})_{(j,k)\in J}$ is a vector of stochastically independent Gaussian processes with $U_{jk}(0) = 0$ and covariance functions

$$\mathrm{Cov}(U_{jk}(s), U_{jk}(t)) = \int_{(0,s\wedge t]} \frac{1}{P(X_{u-} = j)}\,\mathrm{d}q_{jk}(u)\,. \tag{3.5}$$

4. Present values for individual policies and portfolios

Payments between insurer and policyholder are of two types:

(a) Lump sums are payable upon transitions $(j, k) \in J$ between two states and are specified by deterministic functions b_{jk} of bounded variation on $[0, T]$. To distinguish between the time of transition and the actual time of payment, Milbrodt and Stracke (1997) introduced an increasing function $DT : (0, \infty) \to (0, \infty)$, $DT(t) \geq t$, such that upon transition from j to k at time t the amount $b_{jk}(t)$ is payable at time $DT(t)$. To simplify notation we assume that $DT(T) = T$.

(b) Annuity payments fall due during sojourns in a state and are defined by deterministic and right-continuous functions B_j, $j \in \mathcal{S}$ of bounded variation on $[0, T]$, which for each time t specify the total amount $B_j(t)$ paid in $[0, t]$.

As the investment portfolio of the insurance company bears interest, our model needs a discounting function. We write $v(s, t)$ for the value at time s of a unit payable at time $t \geq s$, and we assume that the functions $v(s, \cdot)$ are right-continuous and of bounded variation and that $v(s, t)$ is positive and has a constant upper bound for all $0 \leq s \leq t \leq T$.

4.1. Single policy

Let $L(s)$ be the present value of future benefits minus future premiums with respect to time s. According to Milbrodt and Stracke (1997, Lemma 2.11), we have

$$L(s) = \sum_{j \in \mathcal{S}} \int_{(s,T]} v(s, t)\, I_j(t)\, \mathrm{d}B_j(t) + \sum_{(j,k) \in J} \int_{(s,T]} v(s, DT(t))\, b_{jk}(t)\, \mathrm{d}N_{jk}(t)\,.$$

Using formal matrix multiplication, we can also write

$$L(s) = \int_{(s,T]} v(s, t)\, I(t)\, \mathrm{d}B(t) + \int_{(s,T]} v(s, DT(t))\, \mathbf{e}^\top (b(t) * \mathrm{d}N(t))\, \mathbf{e}\,, \tag{4.1}$$

where $I(t)$ is a row vector with entries $I_j(t)$, $B(t)$ is a column vector with entries $B_j(t)$, $b(t)$ and $N(t)$ are quadratic matrices with entries $b_{jk}(t)$ and $N_{jk}(t)$ for $j \neq k$ and zeros for $j = k$, '$*$' means the componentwise product of two matrices, and $\mathbf{e}$ is a column vector with all entries equal to one.

Let $V_{\mathcal{F}}(s)$ be the conditional expected value of all discounted payments that fall due strictly past s given the information $\mathcal{F}_s = \sigma(X_u, u \leq s)$,

$$V_{\mathcal{F}}(s) = \mathbb{E}(L(s) \,|\, \mathcal{F}_s)\,.$$

Because of the Markov property of $(X_t)_{t \geq 0}$, we almost surely have $V_{\mathcal{F}}(s) = \mathbb{E}(L(s) \,|\, X_s)$, and, hence, we can write

$$V_{\mathcal{F}}(s) = I(s)\, V(s)\,, \tag{4.2}$$

where $V(s)$ shall be a column vector with entries $V_i(s) = \mathbb{E}(L(s) \,|\, X_s = i)$. The quantity $V_i(s)$ is usually denoted as *prospective reserve given state i at time s*. According to Milbrodt and Stracke (1997, Lemma 4.4(c)), we have

$$V_i(s) = \sum_{j \in \mathcal{S}} \int_{(s,T]} v(s, t)\, p_{ij}(s, t)\, \mathrm{d}B_j(t) + \sum_{(j,k) \in J} \int_{(s,T]} v(s, DT(t))\, b_{jk}(t)\, p_{ij}(s, t-)\, \mathrm{d}q_{jk}(t)\,, \tag{4.3}$$

for all $i \in \mathcal{S}$ and $s \in [0,T]$ with $P(X_s = i) > 0$. Under the assumptions made above, $V_i(s)$ is always finite. Using formal matrix multiplication, we can also write

$$V(s) = \int_{(s,T]} v(s,t)\, p(s,t)\, \mathrm{d}B(t) + \int_{(s,T]} v(s, DT(t))\, p(s,t-)\big(b(t) * \mathrm{d}q(t)\big)\, \mathbf{e}\,. \tag{4.4}$$

Another important quantity is the so-called sum at risk for a transition from j to k at time t,

$$R_{jk}(t) = v(t, DT(t))\, b_{jk}(t) + V_k(t) + \Delta B_k(t) - V_k(t) - \Delta B_j(t)\,, \quad j \neq k\,. \tag{4.5}$$

For matrix notation we define $R(t)$ as a quadratic matrix with entries $R_{jk}(t)$ for $j \neq k$ and zeros for $j = k$. We have $R_{jk}(t) = [v(t, DT(t))\, b(t) + (V(t) + \Delta B(t))\mathbf{e}^\top - \mathbf{e}(V(t) + \Delta B(t))^\top]_{jk}$ for all $j \neq k$.

Using (3.1), $V(s)$ can be seen as a vector-valued deterministic functional of the cumulative transition intensity matrix q,

$$V(s) = V(s; q)\,, \quad s \in [0,T]\,. \tag{4.6}$$

According to Christiansen (2008) it is well defined for all q that meet the assumptions that follow formula (3.1). In order to be able to apply the functional delta method later on, we now show the Hadamard differentiability of the functional $V(s;\cdot)$.

Proposition 4.1. *Let* $\overline{q}_n = q_n + \varepsilon_n h_n$ *with* $\|q_n - q\|_\infty \to 0$, $\|h_n - h\|_\infty \to 0$, $\varepsilon_n \to 0$, *and the functions* q, q_n, *and* $\overline{q}_n$ *of uniformly bounded variation. Then we have for all* $0 \leq s \leq T$

$$\left\| \frac{1}{\varepsilon_n}\big(V(s;\overline{q}_n) - V(s;q_n)\big) - \mathrm{D}_q V(s;h) \right\| \longrightarrow 0\,, \quad n \to \infty\,, \tag{4.7}$$

where $\mathrm{D}_q V(s;h)$ *is a supremum norm continuous linear mapping in* h *that equals*

$$\mathrm{D}_q V(s;h) = \int_{(s,T]} v(s,t)\, p(s,t-)(R(t) * \mathrm{d}h(t))\, \mathbf{e}\,, \tag{4.8}$$

interpreted for h *not of bounded variation by formal integration by parts according to* (4.11).

A similar result can be already found in Christiansen (2008), but on the basis of the total variation norm. As we will need Proposition 4.1 in a situation where h is some Gaussian process that has almost surely an infinite total variation, the results in Christiansen (2008) are of no help here.

Proof. We have to show that

$$\frac{1}{\varepsilon_n}\big(V(s;\Phi,\overline{q}_n) - V(s;\Phi,q_n)\big) - \mathrm{D}_q V(s;h)$$

converges to zero. By doing formal matrix operations, we obtain

$$(R(t) * \mathrm{d}h(t))\, \mathbf{e} = v(t, DT(t))\, (b(t) * \mathrm{d}h(t))\, \mathbf{e} + \mathrm{d}h(t)\, (V(t) + \Delta B(t))\,. \tag{4.9}$$

Inserting this property into (4.8), for the second addend on the right hand side of (4.9) we get

$$
\begin{aligned}
&\int_{(s,T]} v(s,t)\, p(s,t-)\, \mathrm{d}h(t)\, (V(t) + \Delta B(t)) \\
&= \int_{(s,T]} v(s,t)\, p(s,t-)\, \mathrm{d}h(t) \Big(\int_{[t,T]} v(t,u) p(t,u)\, \mathrm{d}B(u) \\
&\qquad\qquad + \int_{(t,T]} v(t, DT(u))\, p(t,u-) \big(b(u) * \mathrm{d}q(u)\big)\, \mathbf{e} \big) \Big) \\
&= \int_{(s,T]} v(s,u) \Big(\int_{(s,u]} p(s,t-)\, \mathrm{d}h(t)\, p(t,u) \Big)\, \mathrm{d}B(u) \\
&\quad + \int_{(s,T]} v(s, DT(u)) \Big(\int_{(s,u)} p(s,t-)\, \mathrm{d}h(t)\, p(t,u-) \Big) \big(b(u) * \mathrm{d}q(u)\big)\, \mathbf{e} \big)\,,
\end{aligned}
$$

and for the first addend on the right hand side of (4.9) we have

$$
\int_{(s,T]} v(s, DT(t))\, p(s,t-) (b(t) * \mathrm{d}h(t))\, \mathbf{e}\,.
$$

Hence, we obtain

$$
\begin{aligned}
&\frac{1}{\varepsilon_n}\big(V(s;\Phi,\overline{q}_n) - V(s;\Phi,q_n)\big) - \mathrm{D}_q V(s;h) \\
&= \int_{(s,T]} v(s,t) \Big(\frac{1}{\varepsilon_n}\big(\overline{p}_n(s,t) - p_n(s,t)\big) - \int_{(s,t]} p(s,u-)\, \mathrm{d}h(u)\, p(u,t) \Big)\, \mathrm{d}B(t) \\
&\quad + \int_{(s,T]} v(s,DT(t)) \Big(\frac{1}{\varepsilon_n}\big(\overline{p}_n(s,t-) - p_n(s,t-)\big) - \int_{(s,t)} p(s,u-)\, \mathrm{d}h(u)\, p(u,t-) \Big) \big(b(t) * \mathrm{d}q(u)\big)\, \mathbf{e} \big) \\
&\quad + \int_{(s,T]} v(s,DT(t)) \frac{1}{\varepsilon_n}\big(\overline{p}_n(s,t-) - p(s,t-)\big) \big(b(t) * \mathrm{d}(\overline{q}_n - q)(u)\big) \mathbf{e} \\
&\quad - \int_{(s,T]} v(s,DT(t)) \frac{1}{\varepsilon_n}\big(p_n(s,t-) - p(s,t-)\big) \big(b(t) * \mathrm{d}(q_n - q)(u)\big) \mathbf{e} \\
&\quad + \int_{(s,T]} v(s,DT(t))\, p(s,t-) \Big(b(t) * \mathrm{d}\Big(\frac{1}{\varepsilon_n}\big(\overline{q}_n - q_n\big) - h\Big)(u) \Big) \mathbf{e}\,.
\end{aligned}
\tag{4.10}
$$

The second and third line converge to zero because of (A.2). By applying (A.3) with $q_n = q$, we see that the integrand in the fourth line is uniformly bounded. The integrator $\overline{q}_n - q = (q_n - q) + \varepsilon_n h_n$ in the fourth line converges to zero in supremum norm, which is not a sufficient condition for the convergence of the corresponding integrals. But by formal integration by parts similar to (4.11), we can transform the integral in the fourth line of (4.10) to a term where $\overline{q}_n - q$ is not an integrator anymore and the sumpremum norm convergence is indeed sufficient. Thus, the integral in the fourth line converges to zero. Similarly we can show that also the integral in the fifth line of (4.10) converges to zero. For the integral in the sixth line of (4.10), we can show convergence to zero by again using formal integration by parts similar to (4.11) and the supremum norm convergence to zero of $\frac{1}{\varepsilon_n}(\overline{q}_n - q_n) - h = h_n - h$. □

We now give an alternative representation for (4.8) that can be used as a definition if h is not of bounded variation. The i-th component $\mathrm{D}_q V_i(s;h)$ of $\mathrm{D}_q V(s;h)$ is of the form

$$
\sum_{(j,k)\in J} \int_{(s,T]} v(s,t)\, p_{ij}(s,t-)\, R_{jk}(t)\, \mathrm{d}h_{jk}(t)\,.
$$

The functions $v(s,\cdot)$, $p_{ij}(s,\cdot-)$, and R_{jk} are of bounded variation, and, thus, their product $f_{ijk} := v(s,\cdot)\,p_{ij}(s,\cdot-)\,R_{jk}$ can be decomposed to a sum of a left-continuous function $f_{ijk}^{\rightarrow} = [v(s,\cdot)\,p_{ij}(s,\cdot-)\,R_{jk}]^{\rightarrow}$ and a right-continuous function $f_{ijk}^{\leftarrow} = [v(s,\cdot)\,p_{ij}(s,\cdot-)\,R_{jk}]^{\leftarrow}$ of bounded variation. With the help of Fubini's theorem, we get

$$\begin{aligned}
&\int_{(s,T]} f_{ijk}(t)\,\mathrm{d}h_{jk}(t)\\
&= \int_{(s,T]} \Big(f_{ijk}(s+) + \int_{(s,t]} \mathrm{d}f_{ijk}^{\rightarrow}(u) + \int_{(s,t]} \mathrm{d}f_{ijk}^{\leftarrow}(u) \Big) \mathrm{d}h_{jk}(t)\\
&= f_{ijk}(s+) \int_{(s,T]} \mathrm{d}h_{jk}(t) + \Big(\int_{(s,T]} \int_{[u,T]} \mathrm{d}h_{jk}(t)\,\mathrm{d}f_{ijk}^{\rightarrow}(u) + \int_{(s,T]} \int_{[u,T]} \mathrm{d}h_{jk}(t)\,\mathrm{d}f_{ijk}^{\leftarrow}(u) \Big)\\
&= f_{ijk}(s+)\,\big(h_{jk}(T) - h_{jk}(s)\big) + \int_{(s,T]} \big(h_{jk}(T) - h_{jk}(u-)\big)\,\Big[\mathrm{d}f_{ijk}^{\rightarrow}(u) + \mathrm{d}f_{ijk}^{\leftarrow}(u)\Big] .
\end{aligned}$$

Wit the help of definition (3.1), we see that $f_{ijk}(s+)$ is equal to $\mathbf{1}_{\{i=j\}}\,R_{jk}(s+)$. All in all, the i-th component $\mathrm{D}_q V_i(s;h)$ of $\mathrm{D}_q V(s;h)$ can be written in the form

$$\begin{aligned}
\mathrm{D}_q V_i(s;h) &= \sum_{(j,k)\in J} \Big(\mathbf{1}_{\{i=j\}}\,R_{jk}(s+)\,\big(h_{jk}(T) - h_{jk}(s)\big) + \int_{(s,T]} \big(h_{jk}(T) - h_{jk}(u-)\big)\\
&\qquad \times \Big[\mathrm{d}[v(s,\cdot)\,p_{ij}(s,\cdot-)\,R_{jk}]^{\rightarrow}(u) + \mathrm{d}[v(s,\cdot)\,p_{ij}(s,\cdot-)\,R_{jk}]^{\leftarrow}(u) \Big] \Big)\\
&= \sum_{(j,k)\in J} \Big(\mathbf{1}_{\{i=j\}}\,R_{jk}(s+)\,\big(h_{jk}(T) - h_{jk}(s)\big) + \int_{(s,T]} \big(h_{jk}(T) - h_{jk}(u-)\big)\\
&\qquad \times\ \mathrm{d}[v(s,\cdot)\,p_{ij}(s,\cdot-)\,R_{jk}](u) \Big)
\end{aligned} \tag{4.11}$$

and, hence, we do not need bounded variation for h.

The following Corollary shows that (4.7) implies also some form of Taylor expansion of first order for the functional $V(s;\cdot)$.

Corollary 4.2. *Under the assumptions of Theorem 4.1, we also have*

$$V(s;\overline{q}_n) - V(s;q_n) = \mathrm{D}_q V(s;\overline{q}_n - q_n) + o(\|\overline{q}_n - q_n\|_\infty)\,. \tag{4.12}$$

Proof. As $\mathrm{D}_q V(s;\cdot)$ is a supremum norm continuous mapping, we can substitute $\mathrm{D}_q V(s;h)$ in formula (4.7) by $\mathrm{D}_q V(s;h_n)$. With the linearity of $\mathrm{D}_q V(s;\cdot)$, we then transform (4.7) to

$$\frac{1}{\varepsilon_n} \Big\| V(s;\overline{q}_n) - V(s;q_n) - \mathrm{D}_q V(s;\overline{q}_n - q_n) \Big\| \longrightarrow 0\,.$$

By defining $\varepsilon_n := \|\overline{q}_n - q_n\|_\infty$, we obtain (4.12). □

4.2. Homogeneous portfolio

Now let there be $m \in \mathbb{N}$ policies with stochastically independent and identically distributed jump processes $(X_t^1)_{t\geq 0}, ..., (X_t^m)_{t\geq 0}$. We assume here that the payment functions $B = B^p$ and $b = b^p$

of the policies $p = 1, ..., m$ are all identical and that the policyholders are all from the same generation. The present portfolio value of future benefits minus future premiums is

$$\begin{aligned} L^{(m)}(s) &= \sum_{p=1}^{m} L^p(s) \\ &= \int_{(s,T]} v(s,t) \Big(\sum_{p=1}^{m} I^p(t) \Big) \, \mathrm{d}B(t) + \int_{(s,T]} v(s, DT(t)) \, \mathbf{e}^\top \Big(b(t) * \mathrm{d}\Big(\sum_{p=1}^{m} N^p(t) \Big) \Big) \, \mathbf{e} \, . \end{aligned}$$

By applying (3.2), (3.3), and (3.1), we can transform $L^{(m)}(s)$ to

$$\begin{aligned} L^{(m)}(s) &= \int_{(s,T]} v(s,t) I^{(m)}(t) \, \mathrm{d}B(t) \\ &\qquad + \int_{(s,T]} v(s, DT(t)) \, I^{(m)}(s-) \, (b(t) * \mathrm{d}Q^{(m)}(t)) \, \mathbf{e} \\ &= I^{(m)}(s) \Big(\int_{(s,T]} v(s,t) \prod_{(s,t]} (1 + \mathrm{d}Q^{(m)}) \, \mathrm{d}B(t) \\ &\qquad + \int_{(s,T]} v(s, DT(t)) \prod_{(s,t)} (1 + \mathrm{d}Q^{(m)}) \, (b(t) * \mathrm{d}Q^{(m)}(t)) \, \mathbf{e} \Big) \\ &= I^{(m)}(s) \, V(s; Q^{(m)}) \, . \end{aligned}$$

That means that the difference between the present portfolio value of future payments at time s and its expected value given the information till time s has the representation

$$\begin{aligned} L^{(m)}(s) - \mathbb{E}(L^{(m)}(s) \,|\, \mathcal{F}_s) &= I^{(m)}(s) \big(V(s; Q^{(m)}) - V(s; q) \big) \\ &= I^{(m)}(s) \, \mathrm{D}_q V(s; Q^{(m)} - q) + o(\|Q^{(m)} - q\|_\infty) \, . \end{aligned} \tag{4.13}$$

The second equality follows from (4.12).

5. Net premium

5.1. Single policy

According to the classical principle of equivalence, the mean present value of all net premium payments shall be equal to the mean present value of all benefits, i.e.

$$0 \overset{!}{=} \mathbb{E}\Big(L(0) + \Delta B_{X_0}(0) \Big) = V_{X_0}(0) + \Delta B_{X_0}(0) \, .$$

In general, this principle does not lead to a unique net premium unless we fix a premium scheme and allow only an overall premium level $\pi > 0$ to vary. By decomposing $L(0) + \Delta B_{X_0}(0)$ to a benefits and a premiums part

$$L(0) + \Delta B_{X_0}(0) = L^{benef}(0) + \Delta B^{benef}_{X_0}(0) + \pi \left(L^{prem}(0) + \Delta B^{prem}_{X_0}(0) \right) ,$$

the equivalence equation has the form

$$\begin{aligned} 0 &= \mathbb{E}\Big(L^{benef}(0) + \Delta B^{benef}_{X_0}(0) + \pi \left(L^{prem}(0) + \Delta B^{prem}_{X_0}(0) \right) \Big) \\ &= V^{benef}_{X_0}(0) + \Delta B^{benef}_{X_0}(0) + \pi \left(V^{prem}_{X_0}(0) + \Delta B^{prem}_{X_0}(0) \right) . \end{aligned} \tag{5.1}$$

Here, $V_{X_0}^{benef}(0)$ and $\pi\, V_{X_0}^{prem}(0)$ denote the addends of a decomposition of $V_{X_0}(0)$ to a benefits and a premiums part. Recall that the initial state X_0 was assumed to be deterministic. From equation (5.1), we get an explicit formula for the premium level,

$$\pi = -\frac{V_{X_0}^{benef}(0) + \Delta B_{X_0}^{benef}(0)}{V_{X_0}^{prem}(0) + \Delta B_{X_0}^{prem}(0)}\,. \tag{5.2}$$

In the following we see π as a functional of the cumulative interest intensity q,

$$\pi(q) = -\frac{V_{X_0}^{benef}(0;q) + \Delta B_{X_0}^{benef}(0)}{V_{X_0}^{prem}(0;q) + \Delta B_{X_0}^{prem}(0)}\,. \tag{5.3}$$

This functional is well defined for all q that meet the assumptions that follow formula (3.1) and for which the denominator is not zero. The following proposition gives its Hadamard differential.

Proposition 5.1. *Let $\overline{q}_n = q_n + \varepsilon_n h_n$ with $\|q_n - q\|_\infty \to 0$, $\|h_n - h\|_\infty \to 0$, $\varepsilon_n \to 0$, and the functions q, q_n, and $\overline{q}_n$ of uniformly bounded variation. Then we have for all $0 \le s \le T$*

$$\left\| \frac{1}{\varepsilon_n}\big(\pi(\overline{q}_n) - \pi(q_n)\big) - \mathrm{D}_q\pi(h) \right\| \longrightarrow 0\,, \quad n \to \infty\,, \tag{5.4}$$

where $\mathrm{D}_q\pi(h)$ is a supremum norm continuous linear mapping in h that equals

$$\mathrm{D}_q\pi(h) = -\,\frac{1}{V_{X_0}^{prem}(0;q) + \Delta B_{X_0}^{prem}(0)}\Big(\mathrm{D}_q V_{X_0}^{benef}(0;h) - \pi(q)\,\mathrm{D}_q V_{X_0}^{prem}(0;h)\Big)\,, \tag{5.5}$$

interpreted for h not of bounded variation by formal integration by parts according to (4.11)*.*

A similar result can be already found in Christiansen (2008), but on the basis of the total variation norm. As we will need Proposition 4.1 in a situation where h is some Gaussian process that has almost surely an infinite total variation, the results in Christiansen (2008) are of no help here.

Proof. From Proposition 4.1 we know that $V_{X_0}^{prem}(0;q)$ and $V_{X_0}^{benef}(0;q)$ are Hadamard differentiable with differentials $\mathrm{D}_q V_{X_0}^{prem}(0;\cdot)$ and $\mathrm{D}_q V_{X_0}^{benef}(0;\cdot)$. Furthermore, we can show that also the mapping

$$g : \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R}\,, \quad g(x,y) = -\frac{x}{y}$$

is Hadamard differentiable with differential

$$\mathrm{D}_{(x,y)} g(h_x, h_y) = -\frac{1}{y}\big(h_x - g(x,y)\, h_y\big)\,.$$

Now we just need to apply the chain rule of Hadamard differentiation in order to obtain (5.5). □

Remark 5.2. If we write

$$V_{X_0}(0;q) = V_{X_0}^{benef}(0;q) - \pi(q)\, V_{X_0}^{prem}(0;q)$$

for the prospective reserve that corresponds to premium level $\pi(q)$, then formula (5.5) has the representation

$$\mathrm{D}_q\pi(h) = -\,\frac{1}{V^{prem}_{X_0}(0;q) + \Delta B^{prem}_{X_0}(0)}\,\mathrm{D}_q V_{X_0}(0;h)\,.$$

That means that the Hadamard differentials at q of prospective reserve $V_{X_0}(0;\cdot)$ and premium level $\pi(\cdot)$ just differ in a constant factor. Since premium payments have a negative sign, the constant factor $-1/(V^{prem}_{X_0}(0;q) + \Delta B^{prem}_{X_0}(0))$ is positive. Without loss of generality, it is convenient to standardize the premium schemes in such a way that

$$-\big(V^{prem}_{X_0}(0;q) + \Delta B^{prem}_{X_0}(0)\big) \overset{!}{=} 1\,,$$

since this leads to the equality

$$\mathrm{D}_q\pi(h) = \mathrm{D}_q V_{X_0}(0;h)\,.$$

5.2. Homogeneous portfolio

The premium level $\pi = \pi^p$, $p = 1, ..., m$, (identical for all contracts in our homogeneous portfolio) is perfectly balancing benefits and premiums if

$$0 \overset{!}{=} L^{(m),benef}(0) + m\,\Delta B^{benef}_{X_0}(0) + \pi\,\big(L^{(m),prem}(0) + m\,\Delta B^{prem}_{X_0}(0)\big)\,.$$

From this equation and by using $L^{(m)}(0) = m\,V(0;Q^{(m)})$, we obtain that the perfectly balancing premium (or second order premium) is equal to $\pi(Q^{(m)})$ with $\pi(\cdot)$ defined by (5.3). However, as $\pi(Q^{(m)})$ is random and, thus, not perfectly known for the future, the actuary chooses the premium level π on the basis of expected values instead, see (5.1). The difference between the perfectly balancing premium $\pi(Q^{(m)})$ and the equivalence premium (5.2) is of the form

$$\pi(Q^{(m)}) - \pi(q) = \mathrm{D}_q\pi(Q^{(m)} - q) + o(\|Q^{(m)} - q\|_\infty)\,. \tag{5.6}$$

To see that, apply the ideas of the proof of (4.12) on Proposition 5.1.

6. Asymptotic probability distribution of present portfolio values

6.1. Homogeneous portfolio

On the basis of the results of the previous sections, we now apply the functional delta method with the aim to obtain an asymptotic distribution for the present portfolio value $L^{(m)}(s)$.

Theorem 6.1. *Let the cumulative transition intensity matrix q be absolutely continuous with density function μ, and let $P(X_t = j) > 0$ for all $j \in \mathcal{S}$ and $t \in (0,T)$. Then we have*

$$\sqrt{m}\big(V_i(s;Q^{(m)}) - V_i(s;q)\big)_{i\in\mathcal{S}} \overset{d}{\longrightarrow} \big(\mathrm{D}_q V_i(s;U)\big)_{i\in\mathcal{S}}\,, \quad m \to \infty\,, \tag{6.1}$$

on $[0,T]$, where $U = (U_{jk})_{(j,k)\in J}$ is a vector of independent Gaussian processes according to Proposition 3.2. The addends of

$$D_q V_i(s;U) = \sum_{(j,k)\in J} \int_{(s,T]} v(s,t)\,p_{ij}(s,t)\,R_{jk}(t)\,\mathrm{d}U_{jk}(t)$$

(recall that the integrals are defined by integration by parts according to (4.11)*) are stochastically independent and normally distributed random variables with zero mean and variances*

$$\sigma^2_{ijk}(s) = \int_{(s,T]} \frac{1}{P(X_{t-}=j)} \Big(v(s,t)\, p_{ij}(s,t-)\, R_{jk}(t) \Big)^2 \, \mathrm{d}q_{jk}(t)\,. \tag{6.2}$$

Proof. With the Hadamard differentiability of $V(s;\cdot)$ according to Proposition 4.1 and with property (3.4), equation (6.1) follows from the functional delta method according to Andersen et al. (1991, Theorem II.8.2). As $\mathrm{D}_q V_i(s;\cdot)$ is a linear and continuous mapping and U is Gaussian, the functional value $\mathrm{D}_q V_i(s;U)$ is normally distributed (cf. van der Vaart and Wellner (1996), section 3.9.2). From the linearity of $\mathrm{D}_q V_i(s;\cdot)$ and the zero mean of U we also get that the mean of $\mathrm{D}_q V_i(s;U)$ is zero. The stochastic independence of the addends of $D_q V_i(s;U)$ is a consequence of the independence of the Gaussian processes U_{jk}. For the calculation of the variances (6.2) we use representation (4.11),

$$\begin{aligned}
&\mathbb{E}\Big(\int_{(s,T]} v(s,t)\, p_{ij}(s,t)\, R_{jk}(t)\, \mathrm{d}U_{jk}(t) \Big)^2 \\
&= \mathbb{E}\Big(\mathbf{1}_{\{i=j\}}\, R_{jk}(s+) \big(U_{jk}(T) - U_{jk}(s) \big) \Big)^2 \\
&\quad + \mathbb{E}\Big(\mathbf{1}_{\{i=j\}}\, R_{jk}(s+) \big(U_{jk}(T) - U_{jk}(s) \big) \\
&\qquad\qquad \times \int_{(s,T]} \big(U_{jk}(T) - U_{jk}(u-) \big)\, \mathrm{d}[v(s,\cdot)\, p_{ij}(s,\cdot-)\, R_{jk}](u) \Big) \\
&\quad + \mathbb{E}\Big(\int_{(s,T]} \big(U_{jk}(T) - U_{jk}(u-) \big)\, \mathrm{d}[v(s,\cdot)\, p_{ij}(s,\cdot-)\, R_{jk}](u) \Big)^2 .
\end{aligned}$$

Formula (3.5), Fubini's Theorem, and $\mathbf{1}_{\{i=j\}} = v(s,s+)\, p_{ij}(s,s+)$ lead to

$$\begin{aligned}
&\big(\mathbf{1}_{\{i=j\}}\, R_{jk}(s+) \big)^2 \int_{(s,T]} \frac{1}{P(X_{r-}=j)}\, \mathrm{d}q_{jk}(r) \\
&+ \mathbf{1}_{\{i=j\}}\, R_{jk}(s+) \int_{(s,T]} \int_{[u,T]} \frac{1}{P(X_{r-}=j)}\, \mathrm{d}q_{jk}(r)\, \mathrm{d}[v(s,\cdot)\, p_{ij}(s,\cdot-)\, R_{jk}](u) \\
&+ \int_{(s,T]} \int_{(s,T]} \int_{(u_1 \vee u_2, T]} \frac{1}{P(X_{r-}=j)}\, \mathrm{d}q_{jk}(r) \\
&\qquad \times \mathrm{d}[v(s,\cdot)\, p_{ij}(s,\cdot-)\, R_{jk}](u_1)\, \mathrm{d}[v(s,\cdot)\, p_{ij}(s,\cdot-)\, R_{jk}](u_2) \\
&= \int_{(s,T]} \frac{1}{P(X_{r-}=j)} \Big(\mathbf{1}_{\{i=j\}}\, R_{jk}(s+) + \int_{(s,r]} \mathrm{d}[v(s,\cdot)\, p_{ij}(s,\cdot-)\, R_{jk}](u) \Big)^2 \mathrm{d}q_{jk}(r) \\
&= \int_{(s,T]} \frac{1}{P(X_{r-}=j)} \Big(v(s,r)\, p_{ij}(s,r-)\, R_{jk}(r) \Big)^2 \mathrm{d}q_{jk}(r)\,.
\end{aligned}$$

□

According to Theorem 4.1, the random variable $V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)$, which can be interpreted as the total unsystematic biometric risk of a portfolio, is asymptotically equal to the (uncountable) weighted sum of the increments $\mathrm{d}(Q^{(m)}_{jk} - q_{jk})(t)$ with weights $v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)$,

$$V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q) \approx \sum_{(j,k)\in J} \int_{(0,T]} v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)\, \mathrm{d}(Q^{(m)}_{jk} - q_{jk})(t)\,. \tag{6.3}$$

As the randomness of $\mathrm{d}(Q^{(m)}_{jk} - q_{jk})(t)$ uniquely corresponds to the uncertainty about transition (j,k) at time t, formula (6.3) offers an intuitive linear decomposition of the total unsystematic biometric risk to its risk sources. Theorem 6.1 says now that the approximation error goes to zero if the portfolio size m increases to infinity and that the random contributions of the different risk sources are (asymptotically) independent and normally distributed. Thus, (6.3) is a decomposition to a weighted sum of independent random contributions with well-known distributions, and it is just the decomposition that we were missing in the second and third example in section 2.

Recalling that the stochastic process $(X_t)_{t\geq 0}$ has a deterministic starting value X_0, for $s=0$ Theorem 6.1 yields that

$$\frac{\sqrt{m}}{\sigma^2_{X_0}}\big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\big) \xrightarrow{d} N(0,1)\,, \quad m\to\infty\,, \tag{6.4}$$

where $\sigma^2_{X_0}$ is defined by

$$\sigma^2_{X_0} := \sum_{(j,k)\in J} \sigma^2_{X_0 jk}(0) = \sum_{(j,k)\in J} \int_{(0,T]} \Big(v(0,t)\,R_{jk}(t)\Big)^2\, p_{X_0 j}(0,t-)\,\mathrm{d}q_{jk}(t)\,.$$

We now show that (6.4) is also true in a stronger sense.

Theorem 6.2. *Under the assumptions of Theorem 6.1 and in case of $\sigma^2_{X_0} > 0$, we have*

$$\mathbb{E}\left|\frac{\sqrt{m}}{\sigma^2_{X_0}}\big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\big)\right|^r \longrightarrow \mathbb{E}|N(0,1)|^r\,, \quad m\to\infty\,,\ r\in\mathbb{N}\,,$$

for a standard normally distributed random variable $N(0,1)$.

Proof. By the Almost Sure Representation Theorem there exist versions of $V_{X_0}(0;Q^{(m)})$ and $N(0,1)$ such that

$$\frac{\sqrt{m}}{\sigma^2_{X_0}}\big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\big) \xrightarrow{a.s.} N(0,1)\,, \quad m\to\infty\,.$$

We obtain the convergence of the moments if we have uniform integrability of the sequence $|W_m|^k = |\sqrt{m}\,(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q))|^k$, i.e.

$$\sup_m \int_{\{|W_m|^k\geq\alpha\}} |W_m|^k\,\mathrm{d}P \longrightarrow 0\,, \quad \alpha\to\infty\,.$$

By applying Hölders inequality and Markovs's inequality, we get

$$\begin{aligned}
0 &\leq \int_{\{|W_m|^k\geq\alpha\}} |W_m|^k\,\mathrm{d}P\\
&= \mathbb{E}\Big(|W_m|^k\,\mathbf{1}_{\{|W_m|^k\geq\alpha\}}\Big)\\
&\leq \Big(\mathbb{E}(|W_m|^{2k})\Big)^{1/2}\Big(\mathbb{E}(\mathbf{1}_{\{|W_m|^k\geq\alpha\}})\Big)^{1/2}\\
&\leq \Big(\mathbb{E}(|W_m|^{2k})\Big)^{1/2}\Big(\frac{\mathbb{E}(|W_m|^{2k})}{\alpha^2}\Big)^{1/2}\\
&= \frac{1}{\alpha}\,\mathbb{E}(|W_m|^{2k})\,.
\end{aligned}$$

That means that we are done if we can show that the sequence $\mathbb{E}(|W_1|^{2k}), \mathbb{E}(|W_2|^{2k}), \ldots$ is uniformly bounded. According to (4.13) we have

$$W_m = \sqrt{m}\,(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)) = \frac{1}{\sqrt{m}}\sum_{p=1}^{m}\big(L^p(0) - \mathbb{E}(L^p(0))\big)\,.$$

By applying Proposition A.5 with $Y_p := L^p(0) - \mathbb{E}(L^p(0))$ and using Proposition A.4, we obtain a finite upper bound for $\mathbb{E}(|W_m|^{2k})$ that does not depend on m,

$$\begin{aligned}\mathbb{E}(|W_m|^{2k}) &= \frac{1}{m^k}\mathbb{E}(|Y_1 + \cdots + Y_m|^{2k}) \\ &\leq 2^k\, k^{k+1}\,(2k)!\,\Big(\sup_{l=2,\ldots,2k}\mathbb{E}(|Y_1|^l)\Big)^k\,.\end{aligned}$$

□

Remark 6.3 (Progress of time)**.** So far, without saying we implicitly supposed that $s = 0$ is the present time. If we instead suppose that we are currently at some arbitrary but fixed time $s \in (0,T]$, then the states $X_s^1, \ldots, X_s^m$ of the policyholders at present time s are well-known and can be regarded as deterministic. Theorem 6.2 can then be rewritten to

$$\mathbb{E}\bigg|\frac{\sqrt{m}}{\sigma_{X_s}^2}\big(V_{X_s}(s;Q^{(m)}) - V_{X_s}(s;q)\big)\bigg|^r \longrightarrow \mathbb{E}|N(0,1)|^r\,, \quad m\to\infty\,,\, r\in\mathbb{N}\,,$$

with

$$\sigma_{X_s}^2 = \sum_{(j,k)\in J}\int_{(s,T]}\Big(v(s,t)\,R_{jk}(t)\Big)^2\, p_{X_s j}(s,t-)\,\mathrm{d}q_{jk}(t)\,.$$

The proof is completely analogous.

In a homogeneous portfolio with stochastically independent policyholders, the present values $L^1(0), \ldots, L^m(0)$ are also stochastically independent. Furthermore we showed that $L^{(m)}(0) = m\,V_{X_0}(0;Q^{(m)})$ and $\mathbb{E}(L^{(m)}(0)) = m\,V_{X_0}(0;q)$. Thus, we get that the present value $L^1(0)$ of the single policy $p = 1$ has a total variance of

$$\begin{aligned}\mathrm{Var}(L^1(0)) &= \frac{1}{m}\mathrm{Var}\Big(\sum_{p=1}^{m}L^p(0)\Big) \\ &= \frac{1}{m}\mathrm{Var}\Big(m\big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\big)\Big) \\ &= \mathrm{Var}\Big(\sqrt{m}\big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\big)\Big)\,.\end{aligned}$$

Taking limits on both sides, with Theorem 6.2 and equation (6.2) we get

$$\mathrm{Var}(L^1(0)) = \sum_{(j,k)\in J}\int_0^T\big(v(0,t)\,R_{jk}(t)\big)^2 p_{X_0 j}(0,t)\,\mathrm{d}q_{jk}(t)\,. \tag{6.5}$$

This variance formula is just the variance formula that we know from Hattendorf's Theorem (cf. Milbrodt and Helbig (1999), Folgerung 10.39). It was not clear from Hattendorf's Theorem but it is now clear from Theorem 6.1 and interpretation (6.3) that the (uncountably many) addends

$$\big(v(0,t)\,R_{jk}(t)\big)^2 p_{X_0 j}(0,t)\,\mathrm{d}q_{jk}(t)$$

uniquely correspond to the contributions that the (asymptotically) independent random addends

$$\sqrt{m}\,v(0,t)\,p_{X_0 j}(0,t-)\,R_{jk}(t)\,\mathrm{d}(Q_{jk}^{(m)}-q_{jk})(t)\approx v(0,t)\,p_{X_0 j}(0,t-)\,R_{jk}(t)\,\mathrm{d}U_{jk}(t)$$

make to the total variance. This additive structure of the total variance is a very useful feature that can be used to measure the impact that the different risk sources (transitions $(j,k)\in J$ at times $t\in(0,T]$) have on the total unsystematic risk.

6.2. General portfolio

Now we say that the payment functions B^p and b^p of the policies $p=1,...,m$ may differ and that the policyholders are not necessarily from the same generation. But we still asssume that the cumulative transition intensity matrix q is the same for all policyholders and that the portfolio $\{1,...,m\}$ consists of n homogeneous subportfolios with $m_1,...,m_n$ policies. Inhomogeneity does not only come from the different payment functions B^p and b^p but also from time lags between the calendar times at which the contracts start. In order to keep the notation simple, we say that possible time lags are accounted for by redefining the payment functions B^p and b^p in such a way that $s=0$ is the present time for all contracts. We write $L^{(m_1)}(0),...,L^{(m_n)}(0)$ for the present values of the future liabilities of the subportfolios and $V_i^1(0),...,V_i^n(0)$ for the prospective reserves of single policies in subportfolios $l=1,...,n$.

Corollary 6.4. *Under the assumptions of Theorem* 6.1*, for the total present portfolio value* $L^{(m_1,...,m_n)}(0)=L^{(m_1)}(0)+...+L^{(m_n)}(0)$ *we have*

$$\frac{1}{\sqrt{m}}\Big(L^{(m_1,...,m_n)}(0)-\mathbb{E}(L^{(m_1,...,m_n)}(0))\Big)\xrightarrow{d}\sum_{l=1}^{n}\sqrt{c_l}\,D_q V_{X_0}^l(0;U^l)$$

for $m\to\infty$ *and sequences* $m_l=m_l(m)$ *with*

$$\frac{m_l}{m}\longrightarrow c_l\,,\quad l\in\{1,...,n\}\,,$$

where the U^l *are stochastically independent vectors of Gaussian processes according to Proposition* 3.2*.*

Proof. Using the homogeneity of Hadamard differentiation, we can extend Proposition 4.1 to

$$\frac{1}{\varepsilon_m}\sqrt{\frac{m_l}{m}}\big(V^l(s;\overline{q}_m)-V^l(s;q_m)\big)\longrightarrow\sqrt{c_l}\,\mathrm{D}_q V^l(s;h)$$

and generalize Theorem 6.1 to

$$\sqrt{\frac{m_l}{m}}\sqrt{m_l}\big(V_i^l(s;Q^{(m_l)})-V_i^l(s;q)\big)_{i\in\mathcal{S}}\xrightarrow{d}\big(\sqrt{c_l}\,\mathrm{D}_q V_i^l(s;U^l)\big)_{i\in\mathcal{S}}$$

for all $l = 1, ..., n$. By applying this result on the addends of

$$\frac{1}{\sqrt{m}}\Big(L^{(m_1,\dots,m_n)}(0) - \mathbb{E}(L^{(m_1,\dots,m_n)}(0))\Big) = \sum_{l=1}^{n} \frac{\sqrt{m_l}}{\sqrt{m}} \frac{1}{\sqrt{m_l}}\Big(L^{(m_l)}(0) - \mathbb{E}(L^{(m_l)}(0))\Big)$$
$$= \sum_{l=1}^{n} \frac{\sqrt{m_l}}{\sqrt{m}} \sqrt{m_l}\Big(V^l_{X_0}(0; Q^{(m)}) - V^l_{X_0}(0; q)\Big)$$

(see also (4.13)), where the right hand sides are stochastically independent sums as the random pattern of states of the policyholders are stochastically independent, we obtain the result of the corollary. □

7. Asymptotic distribution of the net premium

7.1. Homogeneous portfolio

We now apply the functional delta method with the aim to obtain an asymptotic distribution for the perfectly balancing (or second order) premium $\pi(Q^{(m)})$.

Theorem 7.1. *Let the cumulative transition intensity matrix q be absolutely continuous with density function μ, and let $P(X_t = j) > 0$ for all $j \in \mathcal{S}$ and $t \in (0, T)$. Then we have*

$$\sqrt{m}\big(\pi(Q^{(m)}) - \pi(q)\big) \xrightarrow{d} \mathrm{D}_q\pi(U)\,, \tag{7.1}$$

where $U = (U_{jk})_{(j,k)\in J}$ is a vector of independent Gaussian processes according to Proposition 3.2. The random variable $D_q\pi(U)$ can be decomposed into a sum of stochastically independent and normally distributed random variables with zero mean and variances

$$\int_{(0,T]} \Big(v(0,t)\left(R^{benef}_{jk}(t) - \pi(q)\, R^{prem}_{jk}(t)\right)\Big)^2 p_{X_0 j}(0,t)\,\mathrm{d}q_{jk}(t)\,. \tag{7.2}$$

If we write $R_{jk}(t)$ for $R^{benef}_{jk}(t) - \pi(q)\, R^{prem}_{jk}(t)$ in the sense of Remark 5.2, then (7.2) is just equal to $\sigma^2_{X_0}$.

Proof. The proof is analogous to the proof of Theorem 6.1. □

Theorem 7.2. *Under the assumptions of Theorem 7.1 and given that the variance σ^2_π of $D_q\pi(U)$ is strictly positive, we have*

$$\mathbb{E}\left|\frac{\sqrt{m}}{\sigma^2_\pi}\big(\pi(Q^{(m)}) - \pi(q)\big)\right|^r \longrightarrow \mathbb{E}|N(0,1)|^r\,, \quad m \to \infty\,,\ r \in \mathbb{N}\,,$$

for a standard normally distributed random variable $N(0, 1)$.

Proof. The proof is analogous to the proof of Theorem 6.2. □

7.2. General portfolio

Analogously to section 6.2, we now say that the payment functions B^p and b^p of the policies $p = 1, ..., m$ may differ and that the policyholders are not necessarily from the same generation. But we still suppose that the cumulative transition intensity matrix q is the same for all policyholders and that the portfolio $\{1, ..., m\}$ consists of n homogeneous subportfolios with $m_1, ..., m_n$ policies. Inhomogeneity does not only come from the different payment functions B^p and b^p but also from time lags between the calendar times at which the contracts started. In order to keep the notation simple, we say that possible time lags are accounted for by redefining the payment functions B^p and b^p in such a way that $s = 0$ is the present time for all contracts. Let $\pi(Q^{(m_1,...,m_n)}) = (\pi^1(Q^{(m_1)}), ..., \pi^n(Q^{(m_n)}))$ be the perfectly balancing (or second order) premiums for single policies in the homogeneous subportfolios $l = 1, ..., n$.

Corollary 7.3. *Under the assumptions of Theorem* 6.1*, for the premium vector* π *we have*

$$\sqrt{m}\Big(\pi(Q^{(m_1,...,m_n)}) - \mathbb{E}(\pi(Q^{(m_1,...,m_n)}))\Big) \longrightarrow \big(\frac{1}{\sqrt{c_l}}\, D_q\pi^l(U^l)\big)_{l\in\{1,...,n\}}$$

for $m \to \infty$ *and sequences* $m_l = m_l(m)$ *with*

$$\frac{m_l}{m} \longrightarrow c_l\,, \quad l \in \{1, ..., n\}\,,$$

where the U^l *are stochastically independent vectors of Gaussian processes according to Proposition* 3.2*.*

Proof. Analogously to the proof of Corollary 6.4, using the homogeneity of Hadamard differentiation and applying Proposition 5.1 on the vector

$$\sqrt{m}\Big(\pi(Q^{(m_1,...,m_n)}) - \mathbb{E}(\pi(Q^{(m_1,...,m_n)}))\Big)\,,$$

whose entries are stochastically independent since the random pattern of states of the policyholders are independent, we obtain the result of the corollary. □

8. Top-down safety margins

We now discuss the construction of explicit and implicit safety margins for unsystematic biometric risk. The unsystematic biometric risk is the risk that payments between insurer and policyholder may deviate from expected values even if the probability distribution of the random pattern of states of the policyholder is perfectly known. Therefore, in the following we assume that the correct q is known.

8.1. Explicit safety margins for a homogeneous portfolio

We are interested in the random fluctuations of the present portfolio value of future payments $L^{(m)}(0)$ and the perfectly balancing premium $\pi(Q^{(m)})$ around their expected values. The idea is to approximate $L^{(m)}(0)$ by

$$L^{(m)}(0) \overset{d}{\approx} \mathbb{E}(L^{(m)}(0)) + \sqrt{m}\, D_q V_{X_0}(0; U)$$

according to Theorem 6.1 and to approximate $\pi(Q^{(m)})$ by

$$\pi(Q^{(m)}) \overset{d}{\approx} \mathbb{E}(\pi(Q^{(m)})) + \frac{1}{\sqrt{m}} D_q \pi(U)$$

according to Theorem 7.1. The right hand sides are normally distributed and the corresponding expectations and covariances are completely known. Alternatively we can write

$$P\left(\sqrt{\frac{1}{m\,\sigma_{X_0}^2}}\left(L^{(m)}(0)) - \mathbb{E}(L^{(m)}(0))\right) \geq x\right) \approx 1 - \Phi(x)\,,$$
$$P\left(\sqrt{\frac{m}{\sigma_\pi^2}}\left(\pi(Q^{(m)}) - \pi(q)\right) \geq x\right) \approx 1 - \Phi(x)\,, \quad x \in \mathbb{R}\,.$$

The approximation error decreases to zero if we increase the sample size m to infinity. Let α be a given confidence level and u_α be the α-quantile of the standard normal distribution. Then we obtain the upper confidence bound

$$L^*(0) = \mathbb{E}(L^{(m)}(0)) + u_\alpha \sqrt{m\,\sigma_{X_0}^2(q)} \tag{8.1}$$

for the present portfolio value $L^{(m)}(0)$, the upper confidence bound

$$V_{X_0}^*(0) = V_{X_0}(0;q) + u_\alpha \sqrt{\frac{\sigma_{X_0}^2(q)}{m}} \tag{8.2}$$

for the prospective reserve $V_{X_0}(0;Q^{(m)})$, and the upper confidence bound

$$\pi^* = \pi(q) + u_\alpha \sqrt{\frac{\sigma_\pi^2(q)}{m}} \tag{8.3}$$

for the perfectly balancing premium $\pi(Q^{(m)})$. The relative differences between $\mathbb{E}(L^{(m)}(0))$ and $L^*(0)$, between $V_{X_0}(0;q)$ and $V_{X_0}^*(0)$, and between $\pi(q)$ and π^* are the explicit safety margins for the unsystematic biometric risk.

8.2. Implicit safety margins for a homogeneous portfolio

In order to be on the safe side with respect to unsystematic biometric risk, the idea of the implicit approach is to calculate premiums and reserves on the basis of a first-order basis q^* instead of the correct cumulative transition intensity matrix q. Let μ^* and μ be the corresponding transition intensity matrices. We denote

$$\mu_{jk}^*(t) - \mu_{jk}(t)$$

as the *implicit safety margin for transition* $(j,k) \in J$ *at time* $t \in [0,T]$. The differences can be positive and negative, depending on the risk structure of the policy. In order to achieve a prespecified level of safety, the top-down approach says that the first-order basis μ^* shall be chosen in such a way that the corresponding premium and reserves are increased by approximately the same amount as by the explicit safety margin of above.

Principle 8.1. The implicit safety margins $\mu_{jk}^*(t) - \mu_{jk}(t)$ for transitions $(j,k) \in J$ and times $t \in (0,T]$ should increase premiums and reserves by (approximately) the same amount as the explicit safety margins in (8.1) to (8.3). In mathematical terms that means that

$$V_{X_0}^*(0) = V_{X_0}(0;q^*) \quad \text{and} \quad \pi^* = \pi(q^*)\,. \tag{8.4}$$

From Corollary 4.2 we know that

$$V_{X_0}(0;q^*) - V_{X_0}(0;q) = \mathrm{D}_q V_{X_0}(0;q^* - q) + o(\|q^* - q\|_\infty)\,.$$

That means that the total (or explicit) safety margin $V_{X_0}(0;q^*) - V_{X_0}(0;q)$ for the prospective reserve at time zero is asymptotically equal to the (uncountable) weighted sum of the implicit safety margins $\mu^*_{jk}(t) - \mu_{jk}(t)$ with weights $v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)$,

$$V_{X_0}(0;q^*) - V_{X_0}(0;q) \approx \sum_{(j,k)\in J} \int_{(0,T]} v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t) \left(\mu^*_{jk}(t) - \mu_{jk}(t)\right) \mathrm{d}t\,. \tag{8.5}$$

Under the standardization notion for premium schemes according to Remark 5.2, we get a similar approximation for the total (or explicit) safety margin of the premium,

$$\pi(q^*) - \pi(q) \approx \sum_{(j,k)\in J} \int_{(0,T]} v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t) \left(\mu^*_{jk}(t) - \mu_{jk}(t)\right) \mathrm{d}t\,. \tag{8.6}$$

Because of the equality of the right hand sides of (8.5) and (8.6), the two equations in (8.4) are asymptotically equivalent.

Corollary 8.2. *Under the assumptions of Remark 5.2, we have*

$$V_{X_0}(0;q^*) - V_{X_0}(0;q) = \pi(q^*) - \pi(q) + o(\|q^* - q\|_\infty)\,,$$

which means that the equations $V^*_{X_0}(0) = V_{X_0}(0;q^*)$ and $\pi^* = \pi(q^*)$ of (8.4) *are asymptotically equivalent.*

Because of this corollary, in the following we only study safety margins for the prospective reserve.

The problem with Principle 8.1 is that it does not lead to a unique μ^* but allows for many different solutions. Therefore we have to think about further principles for defining μ^* in a top-down manner. The random variable $V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)$, which can be interpreted as the total unsystematic biometric risk, is according to Theorem 6.1 asymptotically equal to the (uncountable) weighted sum of increments $\mathrm{d}U_{jk}(t)$ with weights $v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)$,

$$V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q) \overset{d}{\approx} \sum_{(j,k)\in J} \int_{(0,T]} v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)\, \mathrm{d}U_{jk}(t)\,. \tag{8.7}$$

The randomness of $\mathrm{d}U_{jk}(t) \overset{d}{\approx} \sqrt{m}\, \mathrm{d}(Q^{(m)}_{jk} - q_{jk})(t)$ can be interpreted as the uncertainty about transition (j,k) at time t. Interestingly, the increments $\mathrm{d}U_{jk}(t)$ are stochastically independent for all $(j,k) \in J$ and $t \in [0,T]$. Thus, formula (8.7) says that the total unsystematic biometric risk can be asymptotically decomposed to a weighted sum of independent random contributions that uniquely correspond to the randomness of transition $(j,k) \in J$ at time $t \in (0,T]$. This additivity property is very useful for defining implicit safety margins. If we value the unsystematic biometric risk by a risk measure γ that is additive for independent sums, we can asymptotically create an additive decomposition of the total safety margin to partial safety margins that uniquely correspond to the uncertainty in transition $(j,k) \in J$ at time $t \in (0,T]$. The implicit safety margins $\mu^*_{jk}(t) - \mu_{jk}(t)$ should then be defined in such a way that they just create those γ-valued partial safety margins.

Corollary 8.3. *Let γ be a risk measure that is additive for independent sums and for which*

$$\gamma\Big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\Big) = \gamma\Big(\mathrm{D}_q V_{X_0}(0;U)\Big) + o(r_m)\,. \tag{8.8}$$

Then

$$v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)\, (\mu^*_{jk}(t) - \mu_{jk}(t))\, \mathrm{d}t = \gamma\Big(\frac{1}{\sqrt{m}}\, v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)\, \mathrm{d}U_{jk}(t)\Big) \tag{8.9}$$

is a sufficient condition for the asymptotic equality of implicit safety margins and γ-valued safety margins,

$$V_{X_0}(0;q^*) - V_{X_0}(0;q) + o(\|q^* - q\|_\infty) = \gamma\Big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\Big) + o(r_m)\,.$$

Proof. Because of the additivity of γ in case of independent sums, we have

$$\begin{aligned}&\gamma\Big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\Big) + o(r_m)\\ &= \sum_{(j,k)\in J} \int_{(0,T]} \gamma\Big(\frac{1}{\sqrt{m}}\, v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)\, \mathrm{d}U_{jk}(t)\Big)\,.\end{aligned}$$

The right hand side is equal to the right hand side of (8.5) if we define $(\mu^*_{jk}(t) - \mu_{jk}(t))$ by (8.9). □

Example 8.4 (Risk measures)**.** Here we give some examples of risk measures γ that satisfy the assumptions of Corollary 8.3. According to the so-called expectation principle, we define

$$\gamma_1(\cdot) := C\,\mathbb{E}(\cdot)$$

for some constant $C > 0$. The desired additivity follows from the linearity of the expectation operator, and by applying Theorem 6.2 we get a convergence rate of $o(r_m) = o(m^{-1/2})$. If we use a variance principle,

$$\gamma_2(\cdot) := C\,\mathrm{Var}(\cdot)$$

for some constant $C > 0$, we also have the additivity property for independent sums and get from Theorem 6.2 a convergence rate of $o(r_m) = o(m^{-1})$. If the unsystematic biometric risk is traded on a market, we can also think of

$$\gamma_3(\cdot) := \mathbb{E}^Q(\cdot)$$

for some risk-neutral measure Q.

Principle 8.5. Let γ be a risk measure for the valuation of unsystematic biometric risk that is additive for stochastically independent addends and that satisfies (8.8). Then the contribution of the implicit safety margin $\mu^*_{jk}(t) - \mu_{jk}(t)$ for transition (j,k) at time t should asymptotically be equal to the γ-contribution that the randomness of transition (j,k) at time t makes to the total γ-safety margin. In mathematical terms, let

$$v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)\, (\mu^*_{jk}(t) - \mu_{jk}(t))\, \mathrm{d}t = \gamma\Big(\frac{1}{\sqrt{m}}\, v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)\, \mathrm{d}U_{jk}(t)\Big) \tag{8.10}$$

for all $(j,k) \in J$ and $t \in (0,T]$.

In contrast to Principle 8.1, Principle 8.5 uniquely defines the safety margins $\mu^*_{jk}(t) - \mu_{jk}(t)$ for all transitions $(j,k) \in J$ at all times $t \in (0,T]$. The calculations are made at time zero, the beginning of the contract period. What happens when time is progressing and we are currently at time $s \in (0,T]$? At time s a proportion of $p_{X_0 i}(0,s)$ policies is in state i, and for each of them we need a reserve of $V_i(s;Q^{(m)}) - V_i(s;q)$ in order to meet all liabilities after s. Thus, the discounted total reserve that we need for future unsystematic biometric risk (all risk after time s) is

$$W(s) := v(0,s)\, p_{X_0 \bullet}(0,s)\, (V(s;Q^{(m)}) - V(s;q)) ,$$

where $p_{X_0 \bullet}(0,s)$ is a row vector with entries $p_{X_0 i}(0,s)$, $i \in \mathcal{S}$. If we calculate according to Principle 8.5, we have that the total implicit safety margin $V_{X_0}(0;q^*) - V_{X_0}(0;q)$ is asymptotically equal to the γ-valued safety margin $\gamma(W(0))$. At time s, we have a proportion of $p_{X_0 i}(0,s)$ policyholders in state $i \in \mathcal{S}$ for which we reserve a discounted safety margin of $v(0,s)\,(V_i(s;q^*) - V_i(s;q))$. That means that the mean discounted total safety margin at time s is

$$v(0,s)\, p_{X_0 \bullet}(0,s)\, (V_i(s;q^*) - V_i(s;q)) .$$

On the other hand, according to our risk measure γ, we should have a discounted safety margin of $\gamma(W(s))$ at time s.

Corollary 8.6. *Under the assumptions of Principle 8.5 but with* (8.8) *extended to any reference time $s \in [0,T]$, we have*

$$v(0,s)\, p_{X_0 \bullet}(0,s)\, (V(s;q^*) - V(s;q)) + o(\|q^* - q\|_\infty) = \gamma(W(s)) + o(r_m) .$$

That means that the implicit safety margins according to Principle 8.5 are asymptotically commensurate with risk on the whole time interval $[0,T]$.

Proof. By using $V(s;q^*) - V(s;q) + o(\|q^* - q\|_\infty) = \mathrm{D}_q V(s; q^* - q)$, the left hand side can be transformed to

$$\sum_{(j,k) \in J} \int_{(s,T]} v(0,t)\, p_{X_0 j}(0,t-)\, R_{jk}(t)\, (\mu^*_{jk}(t) - \mu_{jk}(t))\, \mathrm{d}t .$$

On the other hand we have

$$\begin{aligned} \gamma(W(s)) + o(r_m) &= \gamma\Big(v(0,s)\, p_{X_0 \cdot}(0,s)\, (V(s;Q^{(m)}) - V(s;q))\Big) + o(r_m) \\ &= \gamma\Big(\sum_{(j,k) \in J} \int_{(s,T]} v(0,t)\, p_{X_0 j}(0,t-)\, \mathrm{d}U_{jk}(t) \Big) \\ &= \sum_{(j,k) \in J} \int_{(s,T]} \gamma\Big(v(0,t)\, p_{X_0 j}(0,t-)\, \mathrm{d}U_{jk}(t) \Big) . \end{aligned}$$

Then from (8.10) we get the desired equality. □

Example 8.7 (Expectation as risk measure)**.** If we choose the expectation principle $\gamma = \gamma_1 = C\,\mathbb{E}(\cdot)$ (cf. Examples 8.4), then (8.10) yields implicit risk margins $\mu^*_{jk}(t) - \mu_{jk}(t) = 0$ that are constantly zero. The reason is that the unsystematic biometric risk is diversifiable and the expectation principle always yields zero for diversifiable risks.

Example 8.8 (Variance as risk measure)**.** In case of a variance principle $\gamma = \gamma_2 = C\,\mathrm{Var}(\cdot)$ (cf. Examples 8.4), from (8.10) and Theorem 6.1 we obtain

$$\mu^*_{jk}(t) = \mu_{jk}(t)\Big(1 + C\,\frac{1}{m}\,v(0,t)\,R_{jk}(t)\Big)\,, \qquad (j,k) \in J\,,\ t \in (0,T]\,.$$

The implicit safety margins depend on the size of the portfolio and converge to zero if we increase the number of policies to infinity, which is the diversification effect. Further on, the safety margins are proportional to the corresponding discounted sums at risk $R_{jk}(t)$, which means that we obtained some form of generalization of the classical safety margin condition

$$\mu^*_{jk}(t) \gtrless \mu_{jk}(t) \Longleftrightarrow R_{jk}(t) \gtrless 0\,, \qquad (j,k) \in J\,,\ t \in (0,T]\,,$$

see, for example, Ramlau-Hansen (1988). While the classical result only gives an answer to the question where to overestimate and where to underestimate the transition rates, our result additionally suggests a size for the implicit margins. The missing constant can be derived, for example, from principle (8.4). With

$$\begin{aligned} &C\,\mathrm{Var}\Big(V_{X_0}(0;Q^{(m)}) - V_{X_0}(0;q)\Big)\\ &= C\,\mathrm{Var}\Big(\frac{1}{\sqrt{m}}\mathrm{D}_q V_{X_0}(0;U)\Big) + o(m^{-1})\\ &= C\,\frac{1}{m}\,\sigma^2_{X_0} + o(m^{-1}) \end{aligned}$$

according to Theorem 6.2 and the asymptotic version

$$\mathrm{D}_q V_{X_0}(0;q^* - q) + o(\|q^* - q\|_\infty) = u_\alpha \sqrt{\frac{\sigma^2_{X_0}(q)}{m}}$$

of (8.4), we asymptotically get for C the equation

$$C\,\frac{1}{m}\,\sigma^2_{X_0} = u_\alpha \sqrt{\frac{\sigma^2_{X_0}(q)}{m}}\,,$$

and, finally, we obtain

$$\mu^*_{jk}(t) = \mu_{jk}(t)\Big(1 + \frac{u_\alpha}{\sqrt{\sigma^2_{X_0}}}\,\frac{1}{\sqrt{m}}\,v(0,t)\,R_{jk}(t)\Big)\,, \qquad (j,k) \in J\,,\ t \in (0,T]\,. \tag{8.11}$$

Example 8.9 (Market risk measure)**.** If we choose a market risk measure $\gamma = \gamma_3 = \mathbb{E}^Q(\cdot)$ (cf. Examples 8.4), equation (8.10) leads to

$$\mu^*_{jk}(t)\,\mathrm{d}t = \mu_{jk}(t)\,\mathrm{d}t + \frac{1}{\sqrt{m}}\,\mathbb{E}^Q(\mathrm{d}U_{jk}(t))\,, \qquad (j,k) \in J\,,\ t \in (0,T]\,.$$

The implicit safety margins still depend on the size of the portfolio and converge to zero if we increase the number of policies to infinity. But, interestingly, the risk structure of the insurance contract plays absolutely no role. The implicit safety margins can be chosen without taking account of the type of policy they are defined for. Depending on the market price of unsystematic biometric risk, the total safety margin can here also get negative.

8.3. Explicit safety margins for general portfolios

The idea is here to approximate $L^{(m_1,\dots,m_n)}(0)$ by

$$L^{(m_1,\dots,m_n)}(0) \overset{d}{\approx} \mathbb{E}(L^{(m_1,\dots,m_n)}(0)) + \sum_{l=1}^{n} \sqrt{m_l}\, D_q V^l_{X_0}(0; U^l)$$

according to Corollary 6.4 and to approximate $\pi(Q^{(m_1,\dots,m_n)})$ by

$$\pi(Q^{(m_1,\dots,m_n)}) \overset{d}{\approx} \mathbb{E}(\pi(Q^{(m_1,\dots,m_n)})) + \big(\frac{1}{\sqrt{m_l}}\, D_q \pi^l(U^l)\big)_{l\in\{1,\dots,n\}}$$

according to Corollary 7.3. The right hand sides are normally distributed and the corresponding expectations and covariances are completely known. The approximation error decreases (to zero) if we increase the portfolio sizes $m_1, \dots, m_n$. In order to achieve a level of safety of α, we calculate α-quantiles of the normally distributed random variables on the right hand side. Alternatively, we can also use calculation principles that base on moments such as expectation principle, variance principle, or standard deviation principle since Theorems 6.2 and 7.2 state the convergence of all corresponding moments.

8.4. Numerical examples

In this section we numerically illustrate the implicit safety margins that we get from (8.11). At first we look at a homogeneous portfolio of $m = 1000$ **pure endowment insurances** with a contract period of 30 years, a starting age of $x = 35$, and a survival benefit of 1. Let the discounting factor $v(s,t)$ and the mortality intensity $\mu_{ad}(t)$ be given by $v(s,t) = (1.0225)^{t-s}$ and

$$\mu_{ad}(x+t) = 0.0005 + 0.000075858 \cdot 10^{0.038\,(x+t)}\,,$$

respectively. A constant premium of 0.01915 is paid yearly in advance. Figures 8.1 and 8.2 illustrate the implicit safety margins according to (8.11) for a safety level of 0.95. The first-order mortality rate is smaller than the second-order mortality rate since pure endowment insurances have a survival character. The difference between first-order and second-order mortality rate is close to zero at the beginning of the contract time and strongly widens with increasing age because the absolute amount of the discounted sum-at-risk is strongly rising with increasing age. Now assume that we have **temporary life insurances** instead of pure endowment insurances. Suppose the death benefit is 1 and the constant yearly premium is 0.0069695. Figures 8.3 and 8.4 illustrate the corresponding implicit safety margins according to (8.11) with safety level 0.95. The first-order mortality rate is greater than the second-order mortality rate since temporary life insurances have an occurrence character. But in contrast to the pure endowment insurance portfolio we have here that the factor between first-order and second-order mortality rates is at its peak at the beginning of the contract period and stays clearly away from 1 throughout the whole contract time. However, the absolute difference between first-order and second-order mortality rate is still rising with increasing age. If we combine a death benefit of 1 with a survival benefit of 2 and set the constant yearly premium to 0.045273, the resulting **endowment insurances** neither have a pure survival character nor a pure occurrence character. Figures 8.5 and 8.6 illustrate the corresponding implicit safety margins. The combined contracts show an occurrence character till age 53 and a survival character after age 53. At the beginning of the contract, the first-order mortality intensity is clearly above the second-order mortality intensity, but it is slower in growth such that it ends clearly below the second-order mortality intensity.

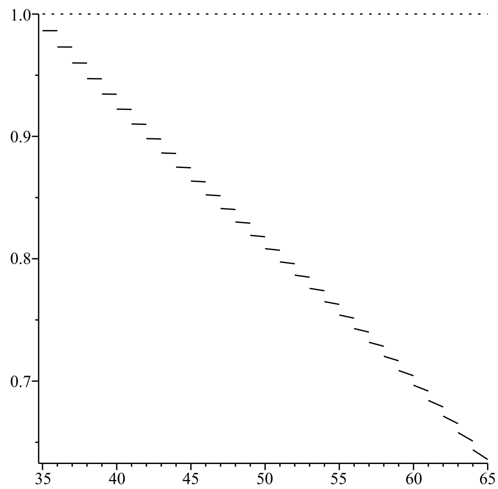


Figure 8.1: Factor $\mu^*_{ad}(t)/\mu_{ad}(t)$ between first-order and second-order mortality for the pure endowment insurances portfolio

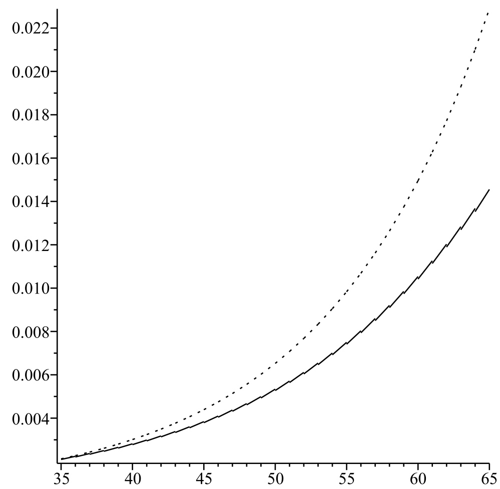


Figure 8.2: First-order mortality $\mu^*_{ad}(t)$ (solid line) and second-order mortality $\mu_{ad}(t)$ (dotted line) for the pure endowment insurances portfolio

At next we look at a **permanent health insurance** with a state space of $\mathcal{S} = \{a = \text{active}, i = \text{invalid}, d = \text{dead}\}$. A constant yearly disability benefit of 1 is paid in case the policyholder is in state invalid. A yearly premium of 0.085063 has to be paid during sojourns in state active till age 63. For the mortality in state active, the mortality in state invalid, the disability rate, and the reactivation rate we used the tables 2008 of the Federal Statistical Office Germany and the tables DAV 1997 TI, DAV 1997 I, and DAV 1997 RI (ultimate tables) of the German Actuarial Society. The three latter tables are in fact tables of first-order, but for our exemplary calculations we ignore that and use them as tables of second-order. All four tables just give yearly transition probabilities. In order to obtain a time-continuous model with transition intensity matrix μ we suppose that all transition intensities are constant in between integer times. Figures 8.7 to 8.10 show the implicit safety margin factors according to (8.11) for a safety level of 0.95. For the transitions (a, i), (i, a), and (i, d) the safety margin factors are all at their peaks/bottoms at the beginning of the contract and go monotonously to 1 towards the end of the contract period. The factors for (i, a) and (i, d) are nearly similar and are approximately equal to the inverses of the factors of (a, i). The implicit safety margins for the transition (a, d) are comparatively small, have the opposite direction compared to (i, d), and are farthest away from 1 in the middle of the contract period.

9. Concluding remarks

The definition of implicit safety margins is a critical point in life and health insurance, not only because of the risk that premiums and reserves are not chosen adequately, but also because the implicit safety margins are essential for the calculation of the surplus. The top-down approach is a convincing concept for the definition of implicit safety margins if we want to control the confidence level of premiums and reserves of first order. Although we find several papers in the insurance

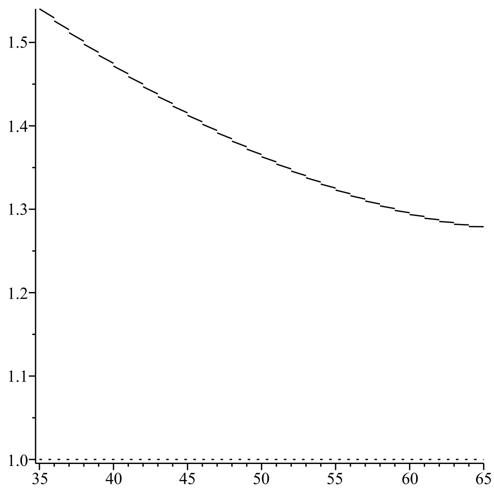


Figure 8.3: Factor $\mu^*_{ad}(t)/\mu_{ad}(t)$ between first-order and second-order mortality for the temporary life insurances portfolio

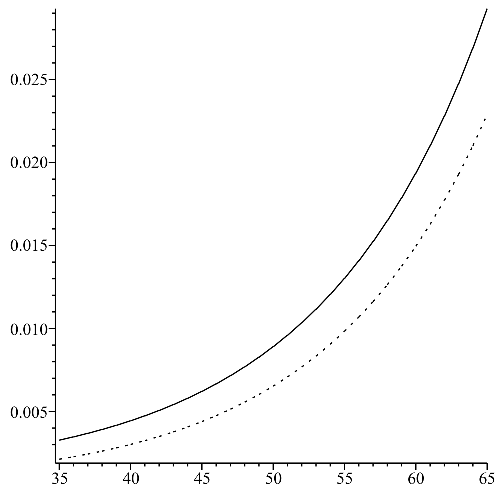


Figure 8.4: First-order mortality $\mu^*_{ad}(t)$ (solid line) and second-order mortality $\mu_{ad}(t)$ (dotted line) for the temporary life insurances portfolio

literature that address the problems that come with the top-down approach, the question of the allocation of the total safety margin to partial (implicit) margins is far from being comprehensively answered. In particular for multistate policies it is difficult to identify the contributions that the different transitions at different ages make to the total unsystematic biometric risk. We presented a new decomposition of the total unsystematic risk that can be a useful tool for modeling just those contributions. Our tool comes with several nice features.

- The addends of the decomposition uniquely correspond to the different transitions at different ages.
- The decomposition is linear and has stochastically independent addends.
- The probability distribution of the addends of the decomposition is completely known.
- The decomposition of the randomness of the prospective reserve and of the premium level are consistent among each other.
- The total variance of the decomposition formula exactly equals the total portfolio variance according to Hattendorf's Theorem, although the decomposition formula is just asymptotic.

Again we want to emphasize that our decomposition formula is just asymptotically true, but for realistic portfolio sizes the approximation error is rather small. On the basis of the decomposition formula we defined allocation principles for a top-down concept. In case of the variance as risk measure, our formula for the implicit safety margins turned out to be a natural generalization of the classical sum-at-risk method. However, in this paper we only dealt with the unsystematic biometric risk. It seems to be an interesting field for further research to develop similar methods for other risk components such as the estimation risk and other valuation basis parameters such as costs.

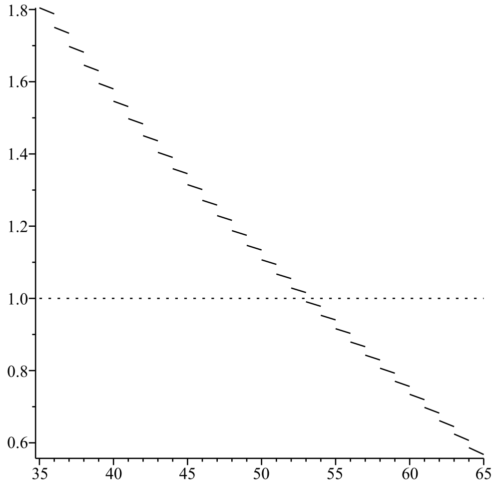


Figure 8.5: Factor $\mu^*_{ad}(t)/\mu_{ad}(t)$ between first-order and second-order mortality for the endowment insurances portfolio

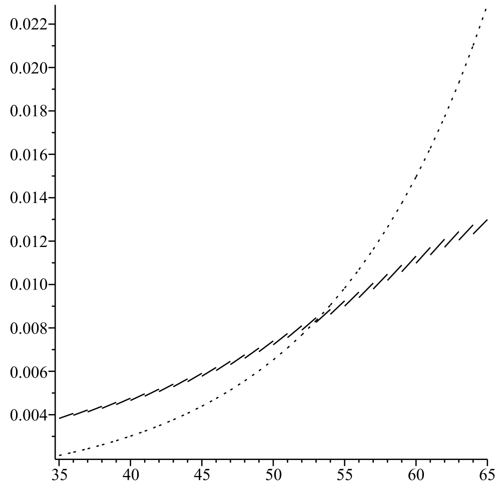


Figure 8.6: First-order mortality $\mu^*_{ad}(t)$ (solid line) and second-order mortality $\mu_{ad}(t)$ (dotted line) for the endowment insurances portfolio

A. Appendix

A.1. Product integral

Proposition A.1 (Formal integration by parts)**.** *Let $\overline{q}$, q, and h be càdlàg functions with bounded variation. Then we have for all $0 \leq s < t$ that*

$$\begin{aligned}
&\int_{(s,t]} \prod_{(s,u)} (1 + \mathrm{d}\overline{q}) \, \mathrm{d}h(u) \prod_{(u,t]} (1 + \mathrm{d}q) \\
&= (h(t) - h(s)) + \int_{(s,t]} \prod_{(s,u)} (1 + \mathrm{d}\overline{q}) \, \mathrm{d}\overline{q}(u) \, (h(t) - h(u)) \\
&\quad + \int_{(s,t]} (h(u-) - h(s)) \, \mathrm{d}q(u) \prod_{(u,t]} (1 + \mathrm{d}q) \\
&\quad + \int_{(s,t]} \int_{(s,r)} \prod_{(s,u)} (1 + \mathrm{d}q) \, \mathrm{d}q(u) \, (h(r-) - h(u)) \, \mathrm{d}\overline{q}(r) \prod_{(r,t]} (1 + \mathrm{d}\overline{q}) \, .
\end{aligned} \tag{A.1}$$

Proof. See the proof of formula (3) on page 19 in Gill (1994). □

Since the right hand side of (A.1) does not have h as integrator anymore, we can use (A.1) as a definition for the left hand side if h is bounded but not of bounded variation.

Proposition A.2 (Hadamard differential of a product integral)**.** *Let $\overline{q}_n$, q_n, q, h_n, and h be càdlàg functions, and let $\overline{q}_n = q_n + \varepsilon_n h_n$ with $\|q_n - q\|_\infty \to 0$, $\|h_n - h\|_\infty \to 0$, $\varepsilon_n \to 0$, and the*

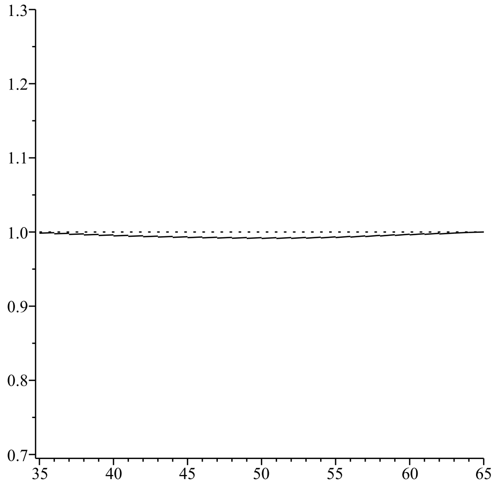


Figure 8.7: Factor $\mu^*_{ad}(t)/\mu_{ad}(t)$ for the permanent health insurances portfolio

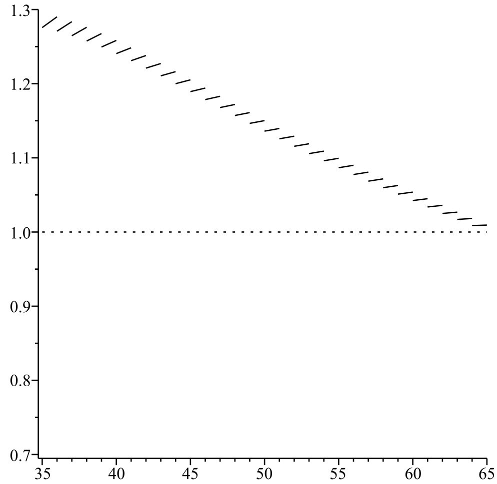


Figure 8.8: Factor $\mu^*_{ai}(t)/\mu_{ai}(t)$ for the permanent health insurances portfolio

functions q, q_n, and $\overline{q}_n$ of uniformly bounded total variation. Then we have

$$\sup_{0\le s<t\le T}\left\|\frac{1}{\varepsilon_n}\Big(\prod_{(s,t]}(1+\mathrm{d}\overline{q}_n)-\prod_{(s,t]}(1+\mathrm{d}q_n)\Big)-\int_{(s,t]}\prod_{(s,u)}(1+\mathrm{d}q)\,\mathrm{d}h(u)\prod_{(u,t]}(1+\mathrm{d}q)\right\|\to 0\,, \quad \text{(A.2)}$$

where the third addend is a supremum norm continuous linear mapping in h, interpreted for h not of bounded variation by formal integration by parts according to (A.1)*.*

Proof. See the proof of formula (5) on page 21 in Gill (1994). □

Corollary A.3. *Under the assumptions of Proposition A.2 we also have*

$$\sup_{0\le s<t\le T}\left\|\prod_{(s,t]}(1+\mathrm{d}\overline{q}_n)-\prod_{(s,t]}(1+\mathrm{d}q_n)\right\|\le C\,\varepsilon_n \qquad \text{(A.3)}$$

for some positive constant $C<\infty$.

Proof. With the help of the linear decomposition

$$\begin{aligned}&\frac{1}{\varepsilon_n}\Big(\prod_{(s,t]}(1+\mathrm{d}\overline{q}_n)-\prod_{(s,t]}(1+\mathrm{d}q_n)\Big)\\&=\bigg(\frac{1}{\varepsilon_n}\Big(\prod_{(s,t]}(1+\mathrm{d}\overline{q}_n)-\prod_{(s,t]}(1+\mathrm{d}q_n)\Big)-\int_{(s,t]}\prod_{(s,u)}(1+\mathrm{d}q)\,\mathrm{d}h(u)\prod_{(u,t]}(1+\mathrm{d}q)\bigg)\\&\quad+\int_{(s,t]}\prod_{(s,u)}(1+\mathrm{d}q)\,\mathrm{d}h(u)\prod_{(u,t]}(1+\mathrm{d}q)\,,\end{aligned}$$

we get the result of the corollary by applying (A.2) on the term in the second line and by using the boundedness of the term in the third line. □

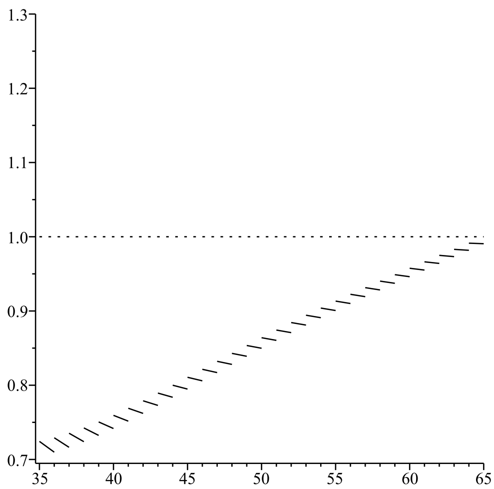


Figure 8.9: Factor $\mu^*_{ia}(t)/\mu_{ia}(t)$ for the permanent health insurances portfolio

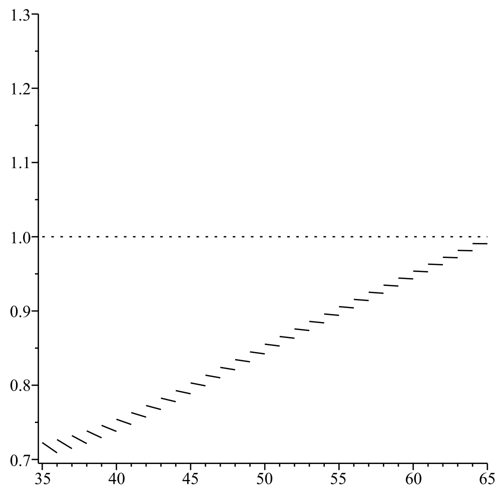


Figure 8.10: Factor $\mu^*_{id}(t)/\mu_{id}(t)$ for the permanent health insurances portfolio

A.2. Moments

Proposition A.4. *For all $k \in \mathbb{N}$ we have $\mathbb{E}(|L(0)|^k) < \infty$.*

Proof. Using representation (4.1) for $L(0)$, we can show that

$$|L(0)| \leq \sup_{0<t\leq T}\{v(0,t)\} \int_{(0,T]} \mathbf{e}^\top \mathrm{d}|B|(t) + \sup_{0<t\leq T}\{v(0,t)\} \sum_{(i,j)\in J} \sup_{0<t\leq T}\{b_{ij}(t)\}\, N_{ij}(T)\,.$$

According to section 4, we have that $\sup_{0<t\leq T}\{v(0,t)\} \leq C_\Phi < \infty$. The finite variation of the B_i and the b_{ij} implies that $\int_{(0,T]} \mathbf{e}^\top \mathrm{d}|B|(t)$ and $\sup_{0<t\leq T}\{b_{ij}(t)\}$ are finite. Thus, we obtain

$$|L(0)| \leq \mathrm{const} + \mathrm{const} \sum_{(i,j)\in J} N_{ij}(T)\,.$$

From Hölder's inequality we get

$$\Big| \sum_{(i,j)\in J} N_{ij}(T)\Big|^k \leq |J|^{k-1} \sum_{(i,j)\in J} |N_{ij}(T)|^k\,,$$

which means that we are done if we can show that $\mathbb{E}(|N_{ij}(T)|^k) < \infty$ for all $(i,j) \in J$ and $k \in \mathbb{N}$. According to Milbrodt and Helbig (1999, sections 4.B and 12.B), the stochastic process N_{ij} is a semimartingale with decomposition of the form

$$N_{ij}(t) = M_{ij}(t) + \int_{(0,t]} \mathbf{1}_{X_{s-}=i}\, \mathrm{d}q_{ij}(s)\,,$$

where the first addend is a square-integrable martingale and the second addend is a predictable cadlag process. By using Ito's formula we can show that

$$\begin{aligned}(N_{ij}(t))^k &= \int_{(0,t]} k\,(N_{ij}(s-))^{k-1}\,\mathrm{d}N_{ij}(s)\\ &\qquad + \sum_{0<s\leq t} (N_{ij}(s))^k - (N_{ij}(s-))^k - k\,(N_{ij}(s-))^{k-1}\,\Delta N_{ij}(s)\\ &= \sum_{l=0}^{k-1} \int_{(0,t]} \binom{k}{l}\,(N_{ij}(s-))^l\,\mathrm{d}N_{ij}(s)\\ &= \sum_{l=0}^{k-1} \binom{k}{l} \int_{(0,t]} \Big[(N_{ij}(s-))^l\,\mathrm{d}M_{ij}(s) + (N_{ij}(s-))^l\,\mathbf{1}_{X_{s-}=i}\,\mathrm{d}q_{ij}(s)\Big]\,.\end{aligned}$$

Taking expectation on both sides and using the monotony of N_{ij}, we obtain

$$\begin{aligned}\mathbb{E}(|N_{ij}(t)|^k) &= \sum_{l=0}^{k-1} \binom{k}{l} \int_{(0,t]} \mathbb{E}\Big((N_{ij}(s-))^l\,\mathbf{1}_{X_{s-}=i}\Big)\mathrm{d}q_{ij}(s)\\ &\leq \sum_{l=0}^{k-1} \binom{k}{l}\,\mathbb{E}(|N_{ij}(t)|^l)\,q_{ij}(t)\,.\end{aligned}$$

That means that $\mathbb{E}(|N_{ij}(t)|^k)$ is finite if $\mathbb{E}(|N_{ij}(t)|^1), \ldots, \mathbb{E}(|N_{ij}(t)|^{k-1})$ are finite. Hence, with $\mathbb{E}(|N_{ij}(t)|^1) < \infty$ according to Milbrodt and Helbig (1999, Folgerung 4.38) and complete induction we arrive at the result of the proposition. □

Proposition A.5. *Let $Y_1, Y_2, \ldots$ be independent and identically distributed random variables with zero mean. For all $n, k \in \mathbb{N}$ we have*

$$\mathbb{E}(|Y_1 + \ldots + Y_n|^{2k}) \leq n^k\, 2^k\, k^{k+1}\, (2k)! \Big(\sup_{l=2,\ldots,2k} \mathbb{E}(|Y_1|^l)\Big)^k\,. \tag{A.4}$$

Proof. By using the multinomial formula and the independence and identical distribution property, we get

$$\begin{aligned}\mathbb{E}(|Y_1 + \ldots + Y_n|^{2k}) &= \sum_{l_1+\cdots+l_n=2k} \frac{(2k)!}{l_1!\cdots l_n!}\mathbb{E}(Y_1^{l_1}\cdots Y_n^{l_n})\\ &= \sum_{l_1+\cdots+l_n=2k} \frac{(2k)!}{l_1!\cdots l_n!}\mathbb{E}(Y_1^{l_1})\cdots\mathbb{E}(Y_1^{l_n})\\ &= \sum_{\substack{l_1+\cdots+l_n=2k\\ l_1,\ldots,l_m\neq 1}} \frac{(2k)!}{l_1!\cdots l_n!}\mathbb{E}(Y_1^{l_1})\cdots\mathbb{E}(Y_1^{l_n})\,,\end{aligned}$$

where the last equality follows from the fact that $\mathbb{E}(Y_1^1) = 0$. Since the number s of variables in $\{l_1, \ldots, l_n\}$ that are greater or equal to 2 at the same time can not be greater than k, the last line of above equals

$$\sum_{s=1}^{k} \binom{n}{s} \sum_{\substack{l_1+\cdots+l_s=2k \\ l_1,\ldots,l_s \geq 2}} \frac{(2k)!}{l_1! \cdots l_s!} \mathbb{E}(Y_1^{l_1}) \cdots \mathbb{E}(Y_1^{l_s})$$

$$\leq \sum_{s=1}^{k} n^s \sum_{\substack{l_1+\cdots+l_s=2k \\ l_1,\ldots,l_s \geq 2}} (2k)! \Big(\sup_{l=2,\ldots,2k} \mathbb{E}(|Y_1|^l) \Big)^s$$

$$\leq k\, n^k \,(2k)^k \,(2k)! \Big(\sup_{l=2,\ldots,2k} \mathbb{E}(|Y_1|^l) \Big)^k .$$

□

References

Andersen, P.K., Borgan, Ø.,Gill, R.D., Keiding, N., 1991. Statistical Models Based on Counting Processes. Springer Series in Statistics, New York.

Bühlmann, H., 1985. Premium calculation from top down. ASTIN Bulletin 15(2), 89-101.

Christiansen, M.C., 2008. 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., 2010. Biometric worst-case scenarios for multi-state life insurance policies. Insurance: Mathematics and Economics 47, 190-197.

Gill, R., 1994. Lectures on Survival Analysis. In: Bakry, D., R. Gill, Molchanov, S., Lectures on Probability Theory. Lecture Notes in Mathematics 1581. Springer, Berlin.

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

Milbrodt, H., Helbig, M., 1999. Mathematische Methoden der Personenversicherung. Walter de Gruyter, Berlin.

Milbrodt, H., Stracke, A., 1997. Markov models and Thiele's integral equations for the prospective reserve. Insurance: Mathematics and Economics 19, 187-235.

Pannenberg, M., 1997. Statistische Schwankungszuschläge für biometrische Rechnungsgrundlagen in der Lebensversicherung. Blätter der DGVFM 23(1), 35-64.

Pannenberg, M., Schütz, E.U., 1998. Sicherheitszuschläge für biometrische Rechnungsgrundlagen in der Lebensversicherung. Transactions of the 26th International Congress of Actuaries 6, 481-511.

Ramlau-Hansen, H., 1988. The emergence of profit in life insurance. Insurance: Mathematics and Economics 7, 225-236.

van der Vaart, A,W., Wellner, J.A., 1996. Weak Convergence and Empirical Processes. Springer, New York.