Michael Koller

Skript AK LV 2012

February 1, 2012

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Mathematics Department

Contents

1. A general life insurance model

1.1 Introduction

The life insurance market offers a wide range of different policies. It is, without expert knowledge, hardly possible to differentiate between all these policies. This is in particular due to the fact that the content of a life insurance is an abstract good.

A life insurance can always be understood as a bet: either one gets a benefit or one pays the premium without getting anything in return. From this point of view life insurance mathematics is a part of probability theory.

Since a life insurance deals with monetary benefits and premiums it is also part of the financial market and the economy. In this context one should note that insurances whose benefits are unit-linked, e.g. the payout depends on the performance of a fond, actually rely on modern theory of financial markets.

Form a legal point of view a life insurance is a contract between the policy holder and the insurer.

As we have note above, life insurances is characterised by its abstract matter and its diversity. Since its content is abstract its value is not intuitively obvious. This is particular due to the fact that a life insurance is usually only bought once or twice during life time. In contrast, for example one buys a loaf of bread on a regular basis, and thus one has acquired a feeling for its correct value.

A life insurance - in particular an individual policy - is a long term contract. Take for example a thirty year old man who buys a permanent life insurance. Now suppose he dies when he reaches ninety, then the contract period was sixty years.

Due to this long duration of the contract and the risks taken - think for example of changing fundamentals - it is necessary to calculate the price of an insurance with care and foresight.

In this chapter we are going to explain classical types of insurance policies. Furthermore we introduce a general model for life insurance, which can be used to price many of the available policies.

1.2 Examples

We start with a description of the most common types of life insurance and the different methods of financing a policy.

1.2.1 Types of life insurance

It is characteristic for every life insurance that the insured event is strongly related to the health of the insured. Thus one can classify life insurance as follows:

– insurance on life or death,

– insurance on permanent disability,

– health insurance.

For an insurance on life or death the essential event is the survival of the insured person up to a certain date or the death before a certain date, respectively. Furthermore these insurance types can be classified by the causes of death which yield a payout (e.g. a life insurance which pays only in the event of an accidental death). Especially, various kinds of survivor's pensions and pure endowments are insurances on life or death.

For a permanent disability insurance the essential criterion is the (dis)ability of the insured at a given date. These insurances have the special feature, that already a certain degree of disability might be sufficient for a claim.

For insurances on health the payout depends on the health of the insured. This class of insurances contains also modern types of policies, like a long term care insurance. The latter only provides benefits if the insured is unable to meet his basic needs (e.g. he is unable to dress himself).

Besides a classification based on the insured event one can also classify the insurance based on the benefit. This can either be paid in annuities or in a lump sum.

In the following we give some typical examples of life insurances.

Pension: A pension policy constitutes that the insurer has to pay annuities to the insured when he reaches a certain age (age of maturity of the policy). Then the pension is paid until the death of the insured. The payment of the annuities is usually done at regular intervals: monthly, quarterly or yearly. Moreover the payment can be done in advance (at the beginning of each interval) or arrears (at the and of each interval). Since the pension is only paid until death, one can additionally agree upon a minimum payment period. In this case the pension is paid at least for the minimum period. (This type of pension contract supplies the desire of the insured to get at least something back for the premiums paid in.)

Pure endowment: A pure endowment insurance provides a payment from the insurer to the insured, if he reaches the age of maturity of the policy. Otherwise there is no payment.

Term/permanent life insurance: A (term/permanent) life insurance is the counterpart to a pure endowment insurance. In contrast to the latter a life insurance does not yield a payout to the insured if the age of maturity of the policy is reached. In the popular case of a term life insurance there is no payout at all if the insured reaches the age of maturity of the policy. But if the insured dies before that age his heirs get a payment. A special case is the permanent life insurance, which yields a payout to the heirs no matter how old the insured is at the time of his death. This insurance is in some countries very popular, since it in a sense an investment into ones off springs.

Endowment: The endowment insurance is the classic example of a life insurance. It is the sum of a pure endowment insurance and a term or permanent life insurance. This means that it yields a payout in the case of an early death and also in the case of reaching the fixed age of maturity.

Widow's pension: A widow's pension is connected to the life of two persons. This is in contrast to the previous examples, where only the life of one person was considered. For a widow's pension there is the insured (the person whose life is insured, e.g. husband) and the beneficiary person (e.g. spouse). As long as both persons are alive no payment is due. If the insured dies and the beneficiary is still alive, then the beneficiary gets a pension until death. Also for this kind of insurance it is possible to fix a minimum period of payments in the policy.

Orphan's pension: After the death of the father or the mother their child gets a pension until it is of age or until death.

Insurance on two lives: For an insurance on two lives, as in the case of a widow's or orphan's pension, one has to consider the lives of two persons. Here the policy fixes a payment depending on the state of the two persons (insured, beneficiary)$\in \{(**), (*\dagger), (\dagger *), (\dagger\dagger)\}$. Obviously, a widow's or an orphan's pension is just a special case of the insurance on two lives. As before, also in this case one could agree upon a minimum period of payments.

Refund guarantee: The refund guarantee is an additional insurance which is often sold together with a pension or a pure endowment. It is a life insurance whose payment equals the paid-in premiums, possibly reduced by the already received payments. The refund guarantee supplies the same want as the minimum periods of payment for the pensions.

Now we have discussed the main insurance types on life and death. Next we want to give a short description of insurances on permanent disability. For these the ability to work is the main criterion. In this context one should note that the probability of becoming disabled strongly depends on the economic environment. This is due to the fact that in a good economic environment everyone finds a job. But a person with restricted health has a hard time finding a job during an economic downturn. In connection with disability the following types of insurances are most common:

Disability pension: In the case of disability, after an initial waiting period, the insured gets a pension until he reaches a fixed age (or until his death) or until he is able to work again. A disability pension without fixed age for a final payment is called permanent disability pension. Often an initial waiting period is introduced, since in most cases disability occurs after an accident or illness and the person actually recovers quickly thereafter. Thus the initial waiting period reduces the price of these policies. Typical waiting periods are three or six month and one or two years.

Disability capital: The disability capital insurance pays a lump sum to the insured in case of a permanent disability.

Premium waiver: The premium waiver is an additional insurance. It waives the obligation to pay further premiums for the insured in the case of disability. Waiting periods are also common for this type of insurance.

Disability children's pension: This pension is similar to the orphan's pension. The only difference is that the cause for a payment is the disability of the mother or father instead of their death.

1.2.2 Methods of financing

Previously we looked at the different types of insurances. Now we are going to discuss the different financing methods and the ideas they are based on. The main principle of life insurances states that the value of the benefits provided by the insurer is equivalent to the value of the policy for the insured. Obviously one has to discuss this equivalence relation in more detail. We will do this in the next chapters and provide a precise definition of the equivalence principle. Then it will be used to calculate the premiums, but for now we get back to the methods of financing. The two common types are:

– Financing by premiums,

– Financing by a single payment (single premium).

Financing by premiums requires the insured to pay premiums to the insurer at regular intervals. This obligation usually ends either when the age of maturity of the policy is reached or if the insured dies.

The other option to finance a life insurance is a single payment. Often a policy incorporates a mixture of both financing methods.

1.3 The insurance model

In this section we want to introduce the insurance model which we will use thereafter. We attempt to describe the real world by a model. Thus it is important to use a model class which is flexible enough to accommodate this. Figure 1.1 shows the general setup of an insurance model. Here we think of an insured person who, at every time t, is in a state $1, 2, \dots, n$. State 1 could for example indicate that the person is alive. The state of the person is then given by the stochastic process X with $X_t(\omega) \in S = \{1, 2, \dots, n\}$. When the insured remains in one state or switches its state a payment, as defined in the insurance policy, is due. For this there are functions $a_i(t)$ and $a_{ij}(t)$ given which correspond to the lines in the figure. They define the amount which the insured gets if he remains in state i (payment $a_i(t)$) or if he switches from state i to j at time t (payment $a_{ij}(t)$). In the following we are going to introduce the necessary concepts for this setup. One distinguishes between the continuous time model, where $(X_t)_{t\in T}$ is defined on an interval in $\mathbb{R}$, and the discrete time model, where $(X_t)_{t\in T}$ is defined on a subset of $\mathbb{N}$. The continuous time model yields the more interesting statements whereas the discrete model is very important in applications, therefore we will discuss both models.

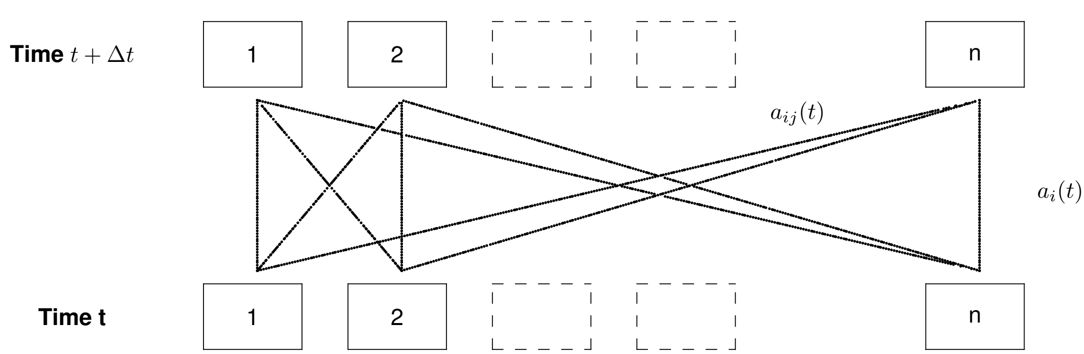

Figure 1.1. Policy setup from t to $t + \Delta t$

Definition 1.3.1 (State space). *We denote by S the state space which is used for the insurance policy. S is finite set.*

Example 1.3.2. For a life insurance or an endowment one often uses the state space $S = \{*, \dagger\}$.

Example 1.3.3. For a disability insurance one has to consider at least the states: alive (active), dead and disabled. Often one uses more states to get a better model. For example in Switzerland a model is used which uses the states $\{*, \dagger\}$ and the family of states $\{$person became disabled at the age of $x : x \in \mathbb{N}\}$.

With the states defined, it is now possible to derive a mathematical model for the payments/benefits. To define the model of the benefits, the so called policy functions, it is necessary to define the time set more precisely. We will define two different time sets since a discrete time set is often used in applications, but the continuous time set yields the neater results.

Definition 1.3.4. – $a_i(t)$ *denotes the sum of the payments to the insured up to time t, given that we know that he has always been in state i. The $a_i(t)$ are called the generalised pension payments. If this pension function is of bounded variation (see Def. 2.1.5) we can also write $a_i(t) = \int_0^t da_i(s)$.*

– $a_{ij}(t)$ *denotes the payments which are due when the state switches from i to j at time t. These benefits are called generalised capital benefits.*

– *In the case of a discrete time set $a_i^{Pre}(t)$ denotes the pension payment which is due at time t, given that the insured is at time t in i.*

– *In the case of a discrete time set $a_{ij}^{Post}(t)$ denotes the capital benefits which are due when switching from i at time t to j at time $t+1$. We are going to assume that the payment is transferred at the end of the time interval.*

The functions $a_i(t)$ are different in the continuous time model and the discrete time model. In the former $a_i(t)$ denotes the sum of the pension payments which are payed up to time t, similar to a mileage meter in a car. In the latter $a_i^{\mathrm{Pre}}(t)$ denotes the single pension payment at time t.

The following example illustrates the interplay between the state space and the functions which define the policy.

Example 1.3.5. Consider an endowment policy with 200,000 USD death benefit and 100,000 USD survival benefit. This insurance shall be financed by a yearly premium of 2,000 USD.

For a age at maturity of 65 the non trivial policy functions are:

$$
\begin{aligned}
a_*(x) &= \begin{cases} 0, & \text{if } x < x_0, \\ -\int_{x_0}^x 2000\,dt, & \text{if } x \in [x_0, 65], \\ -(65 - x_0) \times 2000 + 100000, & \text{if } x > 65, \end{cases} \\
a_{*\dagger}(x) &= \begin{cases} 0, & \text{if } x < x_0 \text{ or } x > 65, \\ 200000, & \text{if } x \in [x_0, 65], \end{cases}
\end{aligned}
$$

where x_0 is the age of entry into the contract, $*$ and $\dagger$ denote the states alive and dead, respectively.

2. Stochastic processes

2.1 Definitions

In this section we will recall basic definitions from probability theory. These will be used throughout the book.

To understand this chapter a basic knowledge in probability theory, measure theory and analysis is a prerequisite.

Definition 2.1.1 (Sets). *We are going to use the notations:*

$$\begin{aligned}
\mathbb{N} &= \textit{the set of the natural numbers including } 0, \\
\mathbb{N}_+ &= \{x \in \mathbb{N} : x > 0\}, \\
\mathbb{R} &= \textit{the set of the real numbers}, \\
\mathbb{R}_+ &= \{x \in \mathbb{R} : x \geq 0\}.
\end{aligned}$$

Furthermore we use the following notations for intervals. For $a, b \in \mathbb{R}, a < b$ we write

$$\begin{aligned}
[a,b] &:= \{t \in \mathbb{R} : a \leq t \leq b\}, \\
]a,b] &:= \{t \in \mathbb{R} : a < t \leq b\}, \\
]a,b[&:= \{t \in \mathbb{R} : a < t < b\}, \\
[a,b[&:= \{t \in \mathbb{R} : a \leq t < b\}.
\end{aligned}$$

Definition 2.1.2 (Indicator function). *For $A \subset \Omega$ we define the indicator function $\chi_A : \Omega \to \mathbb{R}, \omega \mapsto \chi_A(\omega)$ by*

$$\chi_A(\omega) := \begin{cases} 1, & \textit{if} \quad \omega \in A, \\ 0, & \textit{if} \quad \omega \notin A. \end{cases}$$

Furthermore δ_{ij} is Kronecker's delta, i.e., it is equal to 1 for $i = j$ and 0 otherwise.

Definition 2.1.3. *Let*

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x).$$

We define, if they exist, the left limit and the right limit of f at x by:

$$\begin{aligned}
f(x^-) &:= \lim_{\xi \uparrow x} f(\xi), \\
f(x^+) &:= \lim_{\xi \downarrow x} f(\xi).
\end{aligned}$$

Definition 2.1.4. *A real valued function $f : \mathbb{R} \to \mathbb{R}$ is said to be of order $o(t)$, if*

$$\lim_{t \to 0} \frac{f(t)}{t} = 0.$$

This is denoted by $f(t) = o(t)$.

Definition 2.1.5 (Function of bounded variation). *Let $I \subset \mathbb{R}$ be a bounded interval. For a function*

$$f : I \to \mathbb{R}, t \mapsto f(t)$$

the total variation of the function f on the interval I is defined by

$$V(f, I) = \sup \sum_{i=1}^{n} |f(b_i) - f(a_i)|,$$

where the supremum is taken with respect to all partitions of the interval I satisfying

$$a_1 \leq b_1 \leq a_2 \leq b_2 \leq \ldots \leq a_n \leq b_n.$$

The function f is of bounded variation *on I, if $V(f, I)$ is finite. Functions corresponding to a life insurance are usually defined on the interval $[0, \omega]$, where $\omega < \infty$ denotes the last age at which some individuals are alive.*

Properties of functions of bounded variation can be found for example in [DS57].

It is important to note, that functions of bounded variation form an algebra and a lattice. Thus, if f, g are functions of bounded variation and $\alpha \in \mathbb{R}$, then the following functions are also of bounded variation: $\alpha f + g$, $f \times g$, $\min(0, f)$ and $\max(0, f)$.

Definition 2.1.6 (Probability space, stochastic process). *We denote by $(\Omega, \mathcal{A}, P)$ a probability space which satisfies Kolmogorov's axioms.*

Let $(S, \mathcal{S})$ be a measurable space (i.e. S is a set and $\mathcal{S}$ is a σ-algebra on S) and T be a set.
The Borel σ-algebra on the real numbers will be denoted by $\mathcal{R} = \sigma(\mathbb{R})$.

A family $\{X_t : t \in T\}$ of random variables

$$X_t : (\Omega, \mathcal{A}, P) \to (S, \mathcal{S}), \omega \mapsto X_t(\omega)$$

is called stochastic process *on $(\Omega, \mathcal{A}, P)$ with* state space *S.*

For each $\omega \in \Omega$ a sample path of the process is given by the function

$$X_\cdot(\omega) : T \to S, t \mapsto X_t(\omega).$$

We assume that each sample path is right continuous and has left limits.

Definition 2.1.7 (Expectations). *Let X be a random variable on $(\Omega, \mathcal{A}, P)$ and $\mathcal{B} \subset \mathcal{A}$ be a σ-algebra. Then we denote by*

- $E[X]$ *the* expectation *of the random variable X,*
- $V[X]$ *the* variance *of the random variable X,*

– $E[X|\mathcal{B}]$ *the* conditional expectation *of* X *with respect to* $\mathcal{B}$.

Definition 2.1.8. *Let* $(X_t)_{t\in T}$ *be a stochastic process on* $(\Omega, \mathcal{A}, P)$ *taking values in a countable set* S. *We define for* $j \in S$ *the indicator function with respect to the process* $(X_t)_{t\in T}$ *at time* t *by*

$$I_j(t)(\omega) = \begin{cases} 1, & \textit{if} \quad X_t(\omega) = j, \\ 0, & \textit{if} \quad X_t(\omega) \neq j. \end{cases}$$

Analogous, we define for $j, k \in S$ *the number of jumps from* j *to* k *in the time interval* $]0, t[$ *by*

$$N_{jk}(t)(\omega) = \#\left\{\tau \in]0,t[: \, X_{\tau^-} = j \textit{ and } X_\tau = k\right\}.$$

Remark 2.1.9. In the following the function $I_j(t)$ is used to check if the the insured person is at time t in state j. Thus one can check if the pension $a_j(t)$ has to be paid. Similarly, a switch from i to j is indicated by an increase of $N_{ij}(t)$ by 1.

Definition 2.1.10 (Normal distribution). *A random variable* X *on* $(\mathbb{R}, \sigma(\mathbb{R}))$ *with density*

$$f_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R}$$

is called normal distributed *with expectation* μ *and variance* σ^2. *Such a random variable is denoted by* $X \sim \mathcal{N}(\mu, \sigma^2)$.

Examples of stochastic processes are:

Example 2.1.11 (Brownian motion). An example of a non trivial stochastic process is Brownian motion. Brownian motion $X = (X_t)_{t\geq 0}$ with continuous time set $(T = \mathbb{R}_+)$ and state space $S = \mathbb{R}$ is used to model many real world phenomena.

The process is characterised by the following properties:

1. $X_0 = 0$ almost surely.
2. X has independent increments: the random variables $B_{t_1} - B_{t_0}$, $B_{t_2} - B_{t_1}$, $\ldots$, $B_{t_n} - B_{t_{n-1}}$ are independent for all $0 \leq t_1 < t_2 < \ldots < t_n$ and all $n \in \mathbb{N}$.
3. X has stationary increments.
4. $X_t \sim \mathcal{N}(0, t)$.

One can show that almost all sample paths of X are continuous and nowhere differentiable.

Example 2.1.12 (Poisson process). The Poisson process $N = (N_t)_{t\geq 0}$ is a *counting process* with state space $\mathbb{N}$. For example it is used in insurance mathematics to model the number of incurred claims. This process also uses a continuous time set. The homogeneous Poisson process is characterised by the following properties:

1. $N_0 = 0$ almost surely.
2. N has independent and
3. N stationary increments.
4. For all $t > 0$ and all $k \in \mathbb{N}$ gilt: $P[N_t = k] = \exp(-\lambda\, t)\, \frac{(\lambda\, t)^k}{k!}$.

2.2 Markov chains on a countable state space

In the following S is a countable set.

Definition 2.2.1. *Let $(X_t)_{t\in T}$ be a stochastic process on $(\Omega, \mathcal{A}, P)$ with state space S and $T \subset \mathbb{R}$. The process X is called Markov chain, if for all*

$$n \geq 1,\ t_1 < t_2 < \ldots < t_{n+1} \in T,\ i_1, i_2, \ldots, i_{n+1} \in S$$

with

$$P[X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n] > 0$$

the following statement holds:

$$P[X_{t_{n+1}} = i_{n+1} | X_{t_k} = i_k \forall k \leq n] \quad = \quad P[X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n]. \tag{2.1}$$

Remark 2.2.2. 1. Equation (2.1) states that the conditional probabilities only depend on the *last* state. They do not depend on the path which led the chain into that state.

2. Markov chains are very versatile in their applications. This is due to the fact, that on the one hand they are very easy to handle and on the other hand they can model a wide range of phenomena. In the following we are going to model life insurances by Markov chains.

Example 2.2.3. 1. Let $(X_t)_{t\in T}$ be a stochastic process with $S \subset \mathbb{R}$ and $T = \mathbb{N}_+$, for which the random variables $\{X_t : t \in T\}$ are independent. This process is a Markov chain since

$$P[X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n] = \prod_{k=1}^{n} P[X_{t_k} = i_k]$$

for $n \geq 1,\ t_1 < t_2 < \ldots < t_{n+1} \in T,\ i_1, i_2, \ldots, i_{n+1} \in S$.

2. Based on the previous example we define $S_m = \sum_{k=1}^{m} X_k$, where $m \in \mathbb{N}$. This is also an example of a Markov chain.

Proof.

$$\begin{aligned} & P[S_{t_{n+1}} = i_{n+1} \,|\, S_{t_1} = i_1, S_{t_2} = i_2, \ldots, S_{t_n} = i_n] \\ &\quad = P[S_{t_{n+1}} - S_{t_n} = i_{n+1} - i_n] \\ &\quad = P[S_{t_{n+1}} = i_{n+1} \,|\, S_{t_n} = i_n]. \end{aligned}$$

Definition 2.2.4. *Let $(X_t)_{t\in T}$ be a stochastic process on $(\Omega, \mathcal{A}, P)$. Then*

$$p_{ij}(s,t) \quad := \quad P[X_t = j \,|\, X_s = i], \qquad \text{where } s \leq t \text{ and } i, j \in S,$$

is called the conditional probability to switch from state i at time s to state j at time t, or also transition probability for short.

The following theorem of Chapman and Kolmogorov is fundamental for the theory which we will present in the next chapters. The theorem states the relation of $P(s,t)$, $P(t,u)$ and $P(s,u)$ for $s \leq t \leq u$.

Theorem 2.2.5 (Chapman-Kolmogorov equation). *Let $(X_t)_{t\in T}$ be a Markov chain. For $s \le t \le u \in T$ and $i, k \in S$ such that $P[X_s = i] > 0$ the following equations hold:*

$$p_{ik}(s,u) = \sum_{j\in S} p_{ij}(s,t)\, p_{jk}(t,u), \tag{2.2}$$

$$P(s,u) = P(s,t) \times P(t,u). \tag{2.3}$$

This shows, that one can get $P(s,u)$ by matrix multiplication of $P(s,t)$ and $P(t,u)$ for $s \le t \le u \in T$.

Proof. Obviously, the equation holds for $t = s$ or $t = u$. Thus we can assume $s < t < u$ without loss of generality. We will use the following notation:

$$\begin{aligned} S^* &= \{j \in S : P[X_t = j \,|\, X_s = i] \neq 0\} \\ &= \{j \in S : P[X_t = j\,,\; X_s = i] \neq 0\}\,. \end{aligned}$$

(The last equality holds since $P[X_s = i] > 0$.) Now the Chapman-Kolmogorov equation can be deduced from the following equation:

$$\begin{aligned} p_{ik}(s,u) &= P[X_u = k \,|\, X_s = i] \\ &= \sum_{j\in S^*} P[X_u = k,\; X_t = j \,|\, X_s = i] \\ &= \sum_{j\in S^*} P[X_t = j \,|\, X_s = i] \times P[X_u = k \,|\, X_s = i,\; X_t = j] \\ &= \sum_{j\in S^*} p_{ij}(s,t) \times p_{jk}(t,u) \\ &= \sum_{j\in S} p_{ij}(s,t) \times p_{jk}(t,u), \end{aligned}$$

where we applied the Markov property to get equality in the forth line.

After proving the Chapman-Kolmogorov equation we are now able to introduce the abstract concept of transition matrices.

Definition 2.2.6 (Transition matrix). *A family $(p_{ij}(s,t))_{(i,j)\in S\times S}$ is called transition matrix, if the following four properties hold:*

1. $p_{ij}(s,t) \ge 0$.
2. $\sum\limits_{j\in S} p_{ij}(s,t) = 1$.
3. $p_{ij}(s,s) = \begin{cases} 1, & \text{if} \quad i = j, \\ 0, & \text{if} \quad i \neq j, \end{cases} \qquad$ *if* $P[X_s = i] > 0$.
4. $p_{ik}(s,u) = \sum_{j\in S} p_{ij}(s,t)\, p_{jk}(t,u)$ *for* $s \le t \le u$ *and* $P[X_s = i] > 0$.

Theorem 2.2.7. *Let $(X_t)_{t\in T}$ be a Markov chain. Then $(p_{ij}(s,t))_{(i,j)\in S\times S}$ is a transition matrix.*

Proof. This theorem is a direct consequence of the theorem by Chapman and Kolmogorov (Thm. 2.2.5).

Theorem 2.2.8. *A stochastic process* $(X_t)_{t\in T}$ *is a Markov chain, if and only if*

$$P[X_{t_1} = i_1, \ldots, X_{t_n} = i_n] \quad = \quad P[X_{t_1} = i_1] \prod_{k=1}^{n-1} p_{i_k, i_{k+1}}(t_k, t_{k+1}), \tag{2.4}$$

for all

$$n \geq 1,\ t_1 < t_2 < \ldots < t_{n+1} \in T,\ i_1, i_2, \ldots, i_{n+1} \in S.$$

Proof. Let $(X_t)_{t\in T}$ be a Markov chain satisfying

$$P[X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n] > 0.$$

Then the Markov property implies

$$P[X_{t_1} = i_1, \ldots, X_{t_n} = i_n] = P[X_{t_1} = i_1, \ldots, X_{t_{n-1}} = i_{n-1}] \cdot p_{i_{n-1}, i_n}(t_{n-1}, t_n).$$

This yields (2.4) by induction. The converse statement is trivial.

Theorem 2.2.9 (Markov property). *Let* $(X_t)_{t\in T}$ *be a Markov chain and* n, m *be elements of* $\mathbb{N}$. *Fix* $t_1 < t_2 < \ldots < t_n < t_{n+1} < \ldots < t_{n+m}$, $i \in S$ *and sets* $A \subset S^{n-1}$ *(where* S^{n-1} *denotes the* $n-1$ *times Cartesian product of the set* S*) and* $B \subset S^m$ *such that*

$$P\left[(X_{t_1}, X_{t_2}, \ldots, X_{t_{n-1}}) \in A, X_{t_n} = i\right] > 0.$$

Then the following equation (Markov property) *holds:*

$$\begin{aligned} P&\left[(X_{t_{n+1}}, X_{t_{n+2}}, \ldots, X_{t_{n+m}}) \in B \,|\, (X_{t_1}, X_{t_2}, \ldots, X_{t_{n-1}}) \in A, X_{t_n} = i\right] \\ &= P\left[(X_{t_{n+1}}, X_{t_{n+2}}, \ldots, X_{t_{n+m}}) \in B \,|\, X_{t_n} = i\right]. \end{aligned}$$

Proof. We use the notation $\boldsymbol{i}^n = (i_1, i_2, \ldots, i_n)$. An application of equation (2.4) yields:

$$\begin{aligned} P&\left[(X_{t_1}, \ldots, X_{t_{n-1}}) \in A, X_{t_n} = i\right] \\ &= \sum_{\boldsymbol{i}^{n-1} \in A, i_n = i} P\left[X_{t_1} = i_1\right] \times \prod_{k=1}^{n-1} p_{i_k, i_{k+1}}(t_k, t_{k+1}), \\ P&\left[(X_{t_1}, \ldots, X_{t_{n+m}}) \in A \times \{i\} \times B\right] \\ &= \sum_{\boldsymbol{i}^{n+m} \in A \times \{i\} \times B} P\left[X_{t_1} = i_1\right] \times \prod_{k=1}^{n+m-1} p_{i_k, i_{k+1}}(t_k, t_{k+1}). \end{aligned}$$

Finally these two equations imply

$$\begin{aligned} P&\left[(X_{t_{n+1}}, \ldots, X_{t_{n+m}}) \in B \,|\, (X_{t_1}, \ldots, X_{t_{n-1}}) \in A, X_{t_n} = i\right] \\ &= \sum_{(i_n, i_{n+1}, \ldots, i_{n+m}) \in \{i\} \times B} \prod_{k=n}^{n+m-1} p_{i_k, i_{k+1}}(t_k, t_{k+1}) \\ &\quad \times \frac{\sum_{\boldsymbol{i}^{n-1} \in A} P\left[X_{t_1} = i_1\right] \times \prod_{l=1}^{n-1} p_{i_l, i_{l+1}}(t_l, t_{l+1})}{\sum_{\boldsymbol{i}^{n-1} \in A} P\left[X_{t_1} = i_1\right] \times \prod_{l=1}^{n-1} p_{i_l, i_{l+1}}(t_l, t_{l+1})} \end{aligned}$$

$$= \sum_{(i_{n+1},\ldots,i_{n+m})\in B} \prod_{k=n}^{n+m-1} p_{i_k,i_{k+1}}(t_k,t_{k+1}) \frac{P\left[X_{t_n}=i\right]}{P\left[X_{t_n}=i\right]}$$

$$= P\left[(X_{t_{n+1}}, X_{t_{n+2}}, \ldots, X_{t_{n+m}}) \in B | X_{t_n} = i\right].$$

Definition 2.2.10. *A Markov chain $(X_t)_{t\in T}$ is called* homogeneous, *if it is time homogeneous, i.e., the following equation holds for all $s, t \in \mathbb{R}, h > 0$ and $i, j \in S$ such that $P[X_s = i] > 0$ and $P[X_t = i] > 0$:*

$$P[X_{s+h} = j \,|\, X_s = i] \quad = \quad P[X_{t+h} = j \,|\, X_t = i].$$

For a homogeneous Markov chain we use the notation:

$$\begin{aligned} p_{ij}(h) &:= p_{ij}(s, s+h), \\ P(h) &:= P(s, s+h). \end{aligned}$$

Remark 2.2.11. 1. A homogeneous Markov chain is characterised by the fact, that the transition probabilities, and therefore also the transition matrices, only depend on the size of the time increment.

2. For a homogeneous Markov chain one can simplify the Chapman-Kolmogorov equations to the semi-group property:

$$P(s+t) = P(s) \times P(t).$$

The semi-group property is popular in many different areas e.g. in quantum mechanics.

3. The mapping

$$P : T \to M_n(\mathbb{R}), t \mapsto P(t)$$

defines a one parameter *semi-group*.

2.3 Markov chains in continuous time and Kolmogorov's differential equations

In the following we will only consider Markov chains on a finite state space. Thus point wise convergence and uniform convergence will coincide on S. This enables us to give some of the proofs in a simpler form.

Definition 2.3.1. *Let $(X_t)_{t\in T}$ be a Markov chain with finite state space S and $T \subset \mathbb{R}$. For $N \subset S$ we define*

$$p_{jN}(s,t) := \sum_{k\in N} p_{jk}(s,t).$$

Definition 2.3.2 (Transition rates). *Let $(X_t)_{t\in T}$ be a Markov chain in continuous time with finite state space S. $(X_t)_{t\in T}$ is called* regular, *if*

$$\mu_i(t) \quad = \quad \lim_{\Delta t \searrow 0} \frac{1 - p_{ii}(t, t+\Delta t)}{\Delta t} \text{ for all } i \in S, \tag{2.5}$$

$$\mu_{ij}(t) \quad = \quad \lim_{\Delta t \searrow 0} \frac{p_{ij}(t, t+\Delta t)}{\Delta t} \text{ for all } i \neq j \in S \tag{2.6}$$

are well defined and continuous with respect to t.

The functions $\mu_i(t)$ *and* $\mu_{ij}(t)$ *are called* transition rates *of the Markov chain. Furthermore we define* μ_{ii} *by*

$$\mu_{ii}(t) \quad = \quad -\mu_i(t) \text{ for all } i \in S. \tag{2.7}$$

Remark 2.3.3. 1. In the insurance model the regularity of the Markov chain is used to derive the differential equations which are satisfied by the mathematical reserve corresponding to the policy (Thiele's differential equation, e.g. Theorem 5.2.1).

2. One can understand the transition rates as derivatives of the transition probabilities. For example we get for $i \neq j$:

$$\begin{aligned} \mu_{ij}(t) &= \lim_{\Delta t \searrow 0} \frac{p_{ij}(t, t+\Delta t)}{\Delta t} \\ &= \lim_{\Delta t \searrow 0} \frac{p_{ij}(t, t+\Delta t) - p_{ij}(t,t)}{\Delta t} \\ &= \frac{d}{ds} p_{ij}(t,s)\Big|_{s=t}. \end{aligned}$$

3. $\mu_{ij}(t)\, dt$ can be understood as the infinitesimal transition rate from i to j ($i \rightsquigarrow j$) in the time interval $[t, t+dt]$. Similarly, $\mu_i(t)\, dt$ can be understood as the infinitesimal probability of leaving state i in the corresponding time interval. Let us define

$$\Lambda(t) = \begin{pmatrix} \mu_{11}(t) & \mu_{12}(t) & \mu_{13}(t) & \cdots & \mu_{1n}(t) \\ \mu_{21}(t) & \mu_{22}(t) & \mu_{23}(t) & \cdots & \mu_{2n}(t) \\ \mu_{31}(t) & \mu_{32}(t) & \mu_{33}(t) & \cdots & \mu_{3n}(t) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{n1}(t) & \mu_{n2}(t) & \mu_{n3}(t) & \cdots & \mu_{nn}(t) \end{pmatrix}.$$

In a sense, Λ generates the behaviour of the Markov chain. That is, for a homogeneous Markov chain the following equation holds:

$$\Lambda(0) \quad = \quad \lim_{\Delta t \to 0} \frac{P(\Delta t) - 1}{\Delta t}.$$

$\Lambda := \Lambda(0)$ is called the *generator of the one parameter semi group*. We can reconstruct $P(t)$ by

$$P(t) = \exp(t\,\Lambda) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \Lambda^n.$$

4. In the remainder of the book we will only consider finite state spaces. This enables us to avoid certain technical difficulties with respect to convergence.

Based on the transition rates we can prove Kolmogorov's differential equations. These connect the partial derivatives of p_{ij} with μ:

Theorem 2.3.4 (Kolmogorov). *Let* $(X_t)_{t \in T}$ *be a regular Markov chain on a finite state space* S. *Then the following statements hold:*

1. *(Backward differential equations)*

$$\frac{d}{ds}p_{ij}(s,t) = \mu_i(s)p_{ij}(s,t) - \sum_{k\neq i}\mu_{ik}(s)p_{kj}(s,t), \tag{2.8}$$

$$\frac{d}{ds}P(s,t) = -\Lambda(s)P(s,t). \tag{2.9}$$

2. *(Forward differential equations)*

$$\frac{d}{dt}p_{ij}(s,t) = -p_{ij}(s,t)\mu_j(t) + \sum_{k\neq j}p_{ik}(s,t)\mu_{kj}(t), \tag{2.10}$$

$$\frac{d}{dt}P(s,t) = P(s,t)\Lambda(t). \tag{2.11}$$

Proof. The major part of the proof is based on the equations of Chapman and Kolmogorov.

1. We will prove the matrix version of the statement. This will help to highlight the key properties. Let $\Delta s > 0$ and set $\xi := s + \Delta s$.

$$\begin{aligned}\frac{P(\xi,t) - P(s,t)}{\Delta s} &= \frac{1}{\Delta s}\Big(P(\xi,t) - P(s,\xi)\,P(\xi,t)\Big)\\ &= \left(\frac{1}{\Delta s}(1 - P(s,\xi))\right) \times P(\xi,t)\\ &\longrightarrow -\Lambda(s)P(s,t) \text{ for } \Delta s \searrow 0,\end{aligned}$$

   where we used the Chapman-Kolmogorov equation and the continuity of the matrix multiplication.
2. Analogous one can prove the forward differential equation. Let $\Delta t > 0$.

$$\begin{aligned}\frac{P(s,t+\Delta t) - P(s,t)}{\Delta t} &= \frac{1}{\Delta t}\Big(P(s,t)P(t,t+\Delta t) - P(s,t)\Big)\\ &= P(s,t) \times \frac{1}{\Delta t}\Big(P(t,t+\Delta t) - 1\Big)\\ &\longrightarrow P(s,t)\Lambda(t) \text{ for } \Delta t \searrow 0.\end{aligned}$$

Remark 2.3.5. The primary application of Kolmogorov's differential equations is to calculate the transition probabilities p_{ij} based on the rates μ.

Definition 2.3.6. *Let $(X_t)_{t\in T}$ be a regular Markov chain on a finite state space S. Then we denote the conditional probability to stay during the time interval $[s,t]$ in j by*

$$\bar{p}_{jj}(s,t) := P\left[\bigcap_{\xi\in[s,t]}\{X_\xi = j\} \,|\, X_s = j\right]$$

where $s,t \in \mathbb{R}, s \leq t$ and $j \in S$.

In the setting of a life insurance this probability can for example be used to calculate the probability that the insured survives 5 years. The following theorem illustrates how this probability can be calculated based on the transition rates.

Theorem 2.3.7. *Let $(X_t)_{t\in T}$ be a regular Markov chain. Then*

$$\bar{p}_{jj}(s,t) = \exp\left(-\sum_{k\neq j}\int_s^t \mu_{jk}(\tau)d\tau\right) \tag{2.12}$$

holds for $s \le t$, if $P[X_s = j] > 0$.

Proof. We define $K_j(s,t)$ by $K_j(s,t) := \bigcap_{\xi\in[s,t]}\{X_\xi = j\}$. Let $\Delta t > 0$. We have $P[A\cap B\,|\,C] = P[B\,|\,C]\,P[A\,|\,B\cap C]$ and thus

$$\begin{aligned}
\bar{p}_{jj}(s,t+\Delta t) &= P[K_j(s,t)\cap K_j(t,t+\Delta t)\,|\,X_s=j] \\
&= P[K_j(s,t)\,|\,X_s=j]\,P[K_j(t,t+\Delta t)\,|\,X_s=j\cap K_j(s,t)] \\
&= P[K_j(s,t)\,|\,X_s=j]\,P[K_j(t,t+\Delta t)\,|\,X_t=j] \\
&= \bar{p}_{jj}(s,t)\,P[K_j(t,t+\Delta t)\,|\,X_t=j],
\end{aligned}$$

where we used the Markov property and the relation $\{X_s = j\}\cap K_j(s,t) = \{X_t = j\}\cap K_j(s,t)$. The previous equation yields

$$\begin{aligned}
\bar{p}_{jj}(s,t+\Delta t) - \bar{p}_{jj}(s,t) &= -\bar{p}_{jj}(s,t)\times\Big(1 - P[K_j(t,t+\Delta t)\,|\,X_t=j]\Big) \\
&= -\bar{p}_{jj}(s,t)\times\left(\sum_{k\neq j} p_{jk}(t,t+\Delta t) + o(\Delta t)\right),
\end{aligned}$$

where we used that the rates $\mu_{\cdot\cdot}$ are well defined.

Now taking the limit we get the differential equation

$$\frac{d}{dt}\bar{p}_{jj}(s,t) = -\bar{p}_{jj}(s,t)\times\sum_{k\neq j}\mu_{jk}(t).$$

Solving this equation with the boundary condition $\bar{p}_{jj}(s,s) = 1$ yields the statement of the theorem, (2.12).

2.4 Examples

In this section we want to illustrate the theory of the previous sections by some examples.

Example 2.4.1 (Life insurance). We start with a life insurance, which provides a sum of money to the heirs in case of the death of the insured. Usually one uses for this a model with either two states (alive $*$, dead $\dagger$) or three states (alive, dead (accident), dead (disease)). We will use the model with two states, and the death rate will be exemplary modelled by the function

$$\mu_{*\dagger}(x) = \exp(-9.13275 + 8.09438\cdot 10^{-2}x - 1.10180\cdot 10^{-5}x^2). \tag{2.13}$$

The death rate is the transition rate of the state transition $* \rightsquigarrow \dagger$. See section 4.3 for a derivation of the death rate. Based on the death rate and formula (2.12) we are now able to calculate the survival probability of a 35 year old man:

$$\bar{p}_{**}(35,x) = \exp\left(-\int_{35}^x \mu_{*\dagger}(\tau)d\tau\right), \qquad \text{for } x > 35.$$

Figure 2.1 shows the transition rate (dotted line) and the survival probability (continuous line) based on $x = 35$.

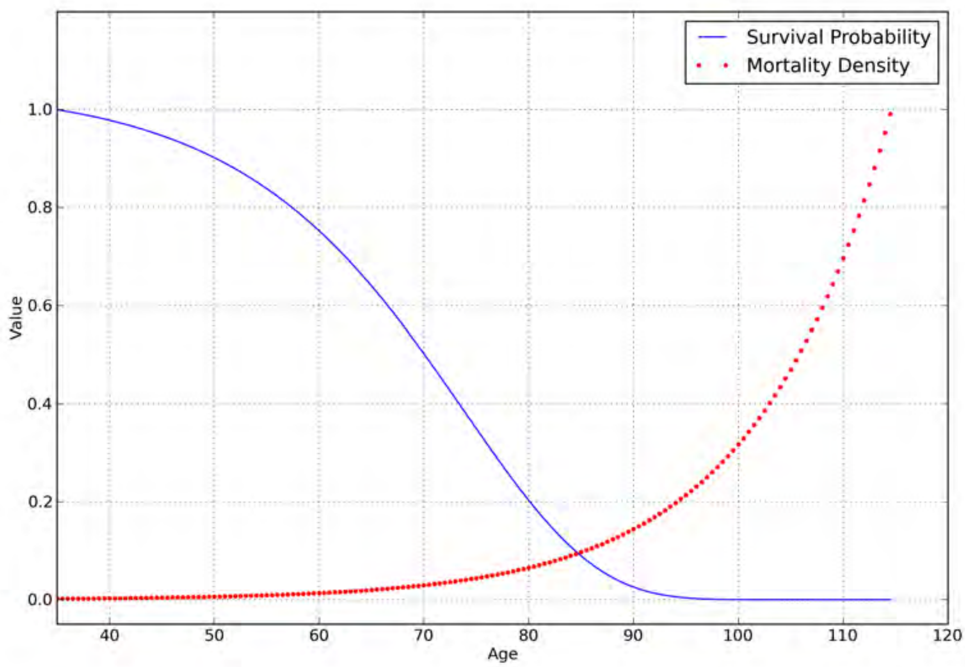

Figure 2.1. Mortality density $\mu_{*\dagger}(x)$ and survival probability $\bar{p}_{**}(35, x)$

Example 2.4.2 (Disability pension). We consider a model of a disability pension with the following three states:

state	**symbol**
active	$*$
disabled	$\diamond$
dead	$\dagger$

The transition rates are defined by

$$
\begin{aligned}
\sigma(x) &:= 0.0004 + 10^{(0.060\,x - 5.46)}, \\
\mu(x) &:= 0.0005 + 10^{(0.038\,x - 4.12)}, \\
\mu_{*\diamond}(x) &:= \sigma(x), \\
\mu_{*\dagger}(x) &:= \mu(x), \\
\mu_{\diamond\dagger}(x) &:= \mu(x).
\end{aligned}
$$

The transition rate σ is the infinitesimal probability of becoming disabled and μ is the corresponding probability of dying. We set the other transition rates equal to 0. Thus in particular this model does not incorporate the possibility of becoming active again ($\mu_{\diamond *} = 0$). Moreover one should note that in this model the mortality of disabled persons is equal to the mortality of active persons. This is a simplification, since in reality disabled persons have a higher mortality (they die earlier with a higher probability) than active persons. Therefore, this model yields an overpriced premium for the disability pension.

The explicit knowledge of the transition probabilities p_{ij} is useful for many formulas in insurance mathematics. For the current model they can be calculated by Kolmogorov's differential equations. We get

$$\begin{aligned}
p_{**}(x,y) &= \exp\left(-\int_x^y [\mu(\tau)+\sigma(\tau)]\,d\tau\right), \\
p_{*\diamond}(x,y) &= \exp\left(-\int_x^y \mu(\tau)d\tau\right) \times \left(1-\exp\left(-\int_x^y \sigma(\tau)d\tau\right)\right), \\
p_{\diamond\diamond}(x,y) &= \exp\left(-\int_x^y \mu(\tau)d\tau\right),
\end{aligned}$$

which solve Kolmogorov's differential equations for this model:

$$\begin{aligned}
\frac{d}{dt}p_{**}(s,t) &= -p_{**}(s,t) \times (\mu(t)+\sigma(t)), \\
\frac{d}{dt}p_{*\diamond}(s,t) &= -p_{*\diamond}(s,t)\,\mu(t) + p_{**}(s,t)\,\sigma(t), \\
\frac{d}{dt}p_{*\dagger}(s,t) &= (p_{**}(s,t)+p_{*\diamond}(s,t)) \times \mu(t), \\
\frac{d}{dt}p_{\diamond *}(s,t) &= 0, \\
\frac{d}{dt}p_{\diamond\diamond}(s,t) &= -p_{\diamond\diamond}(s,t)\,\mu(t), \\
\frac{d}{dt}p_{\diamond\dagger}(s,t) &= p_{\diamond\diamond}(s,t)\,\mu(t), \\
\frac{d}{dt}p_{\dagger\dagger}(s,t) &= 0,
\end{aligned}$$

with the boundary conditions $p_{ij}(s,s) = \delta_{ij}$. Note that, if one uses a model with a positive probability of becoming active again, one has to modify the first, second, forth and fifth equation. Obviously one can solve these equations with numerical methods, the solutions for the given example are listed in Table 2.1.

Exercise 2.4.3. Consider the above system of differential equations.

1. Find an exact solution.
2. Find a numerical approximation to the solution.

Table 2.1. Transition probabilities for the disability insurance

Initial age	$x_0 = 30$
Algorithm	Runge-Kutta of order 4
Step width	0.001

age x	$p_{**}(x_0, x)$	$p_{*\diamond}(x_0, x)$	$p_{*\dagger}(x_0, x)$	$p_{\diamond\diamond}(x_0, x)$	$p_{\diamond\dagger}(x_0, x)$
30.00	1.00000	0.00000	0.00000	1.00000	0.00000
35.00	0.98743	0.00354	0.00903	0.99097	0.00903
40.00	0.96998	0.00850	0.02152	0.97849	0.02152
45.00	0.94457	0.01620	0.03923	0.96077	0.03923
50.00	0.90624	0.02903	0.06474	0.93526	0.06474
55.00	0.84725	0.05106	0.10169	0.89831	0.10169
60.00	0.75677	0.08832	0.15491	0.84509	0.15491
65.00	0.62287	0.14700	0.23013	0.76987	0.23013

3. Interest rate

3.1 Introduction

An important part of every insurance contract is the underlying interest rate. The so called technical interest rate describes the interest which the insurer guarantees to the insured. It is a significant factor for the size of the premiums. If the technical interest rate is too low it yields inflated premiums, if it is too high it might yield to insolvency of the insurance company.

For the technical interest rate one can use a deterministic or a stochastic model. In the latter case the interest rate will be coupled to the bond market. In the following we are going to define the main parameters connected to the interest rate and present their relations.

3.2 Definitions

Example 3.2.1. Suppose we put 10,000 USD on a bank account on the first of January. If at the end of the year there are 10,500 USD on the account, then the underlying interest rate was 5 %.

Definition 3.2.2 (Interest rate). *We denote by i the* yearly interest rate. *Furthermore we assume, that it depends on time and write $i_t, t \geq 0$. If we use a stochastic model for the interest rate, then i is a stochastic process $(i_t(\omega))_{t\geq 0}$.*

One should note that this definition is useful in particular for the discrete time model. Since in this case one would use the time intervals which are given by the discretisation. The calculation of the future capital based on an interest rate is done by

$$B_{t+1} = (1 + i_t) \times B_t.$$

Here B_t denotes the value of the account at time t. Often also the inverse of this relation is relevant. Thus one defines the discount rate.

Definition 3.2.3 (Discount rate). *Let i_t be the interest rate in year t. Then*

$$v_t = \frac{1}{1 + i_t}$$

is the discount rate *in year t.*

The discount rate can be used to calculate the net present value (the present value of the future benefits). If the interest rate is a stochastic process the previous considerations lead to the following problem: suppose we are going to receive 1 USD in one year, what is its present value? Generally there are two possible ways to find an answer.

Valuation principle A: If the interest rate i is known, the present value X is

$$X = \frac{1}{1+i}$$

and thus its mean is

$$X_A = E\left[\frac{1}{1+i}\right].$$

Valuation principle B: If the interest rate is known, the value of the account at the end of the year is $X(1+i)$. Thus

$$1 = E\left[X(1+i)\right] = X \times E\left[1+i\right]$$

holds and the mean is

$$X_B = \frac{1}{E\left[1+i\right]}.$$

But in general X_A is not equal to X_B. To overcome this problem one needs to fix by an assumption the valuation principle of the interest rate: we are always going to determine the net present value by valuation principle A (cf. [Büh92]).

After solving this paradox situation one understands the importance of the discount rate. Obviously, the same problem also arises in continuous time. For models in continuous time we assume that the interest is also payed continuously such that

$$B_{t+s} = \exp\left(\int_t^{t+s} \delta(\xi) d\xi\right) \times B_t.$$

Definition 3.2.4 (Interest intensity). *The* interest intensity *at time t is denoted by $\delta(t)$.*

A yearly interest rate i yields

$$e^{\delta} = 1 + i$$

and thus

$$\delta = \ln(1+i).$$

In continuous time the discount rate (from t to 0) is

$$v(t) = \exp\left(-\int_0^t \delta(\xi)\, d\xi\right).$$

Here the discount rate is modelled from t to 0 which is contrary to the discrete setting. The following relation holds:

$$v_t = \exp\left(-\int_t^{t+1} \delta(\xi)\, d\xi\right).$$

For a stochastic interest rate the interest intensity δ is also a stochastic process $(\delta_t(\omega))_{t\geq 0}$. Also in the continuous time setting we are going to use valuation principle A.

Finally it is worth to note the difference between $v(t)$ which denotes the discounting back to time 0 and v_t which denotes the discount from time $t+1$ to t. This later quantity is commonly used in actuarial sciences for recursion formulae.

3.3 Models for the interest rate process

Now we want to describe the stochastic behaviour of the interest rate. We start with an analysis of the average interest rate of Swiss government bonds in Swiss franc. Figure 3.1 shows the rate from 1948 up to 2009.

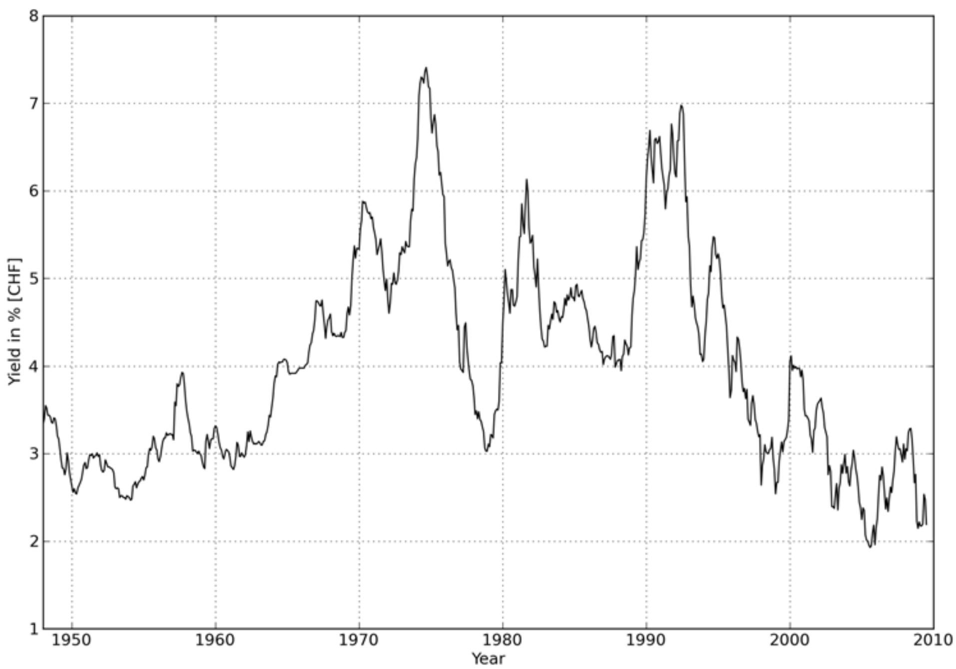

Figure 3.1. Average yield of government bonds in %

The figure shows that during the given period the interest rate of bonds was subject to huge fluctuations. The minimal rate of 1.93 % was reached in 2005. The maximum of 7.41 % was recorded during the "oil crisis" in the autumn of 1974. It is interesting to note, that in the first printing of this book (1999) the minimum was actually reached in 1954 at about 2.5 %. At this time nobody expected that the interest rates in Switzerland would drop that far again. Based on this observation the inherent risk of the technical interest rate in insurance policies becomes obvious.

After seeing these values one starts to wonder how the technical interest rate should be determined. First of all, this depends on the purpose of the model in use. One has to differentiate between short term and long term relations. Moreover, the interest rate might only be used for marketing or forecasting purposes. In any case it should be emphasised that it might be risky to fix a constant interest rate above the level of the observed minimum. Since in this case there might be periods during which the interest yield of the assets does not cover the liabilities. Thus one has to take care when fixing the technical interest rate.

Another method to determine the technical interest rate is based on the analysis of the yield-curve or the forward-curve. This curves allow to measure the interest rate structure. The yield-curve can be used to determine the interest rate which one would get for a bond with a fixed investment period. Figure 3.2 shows the yield-curve for various currencies, it indicates that the interest rate is smaller for a bond with a short investment period than for a bond with a long investment period. This is called a 'normal' interest rate structure. Conversely, one speaks of an 'inverse' interest rate structure if the bonds with a short investment period provide a better yield than those with a long investment period.

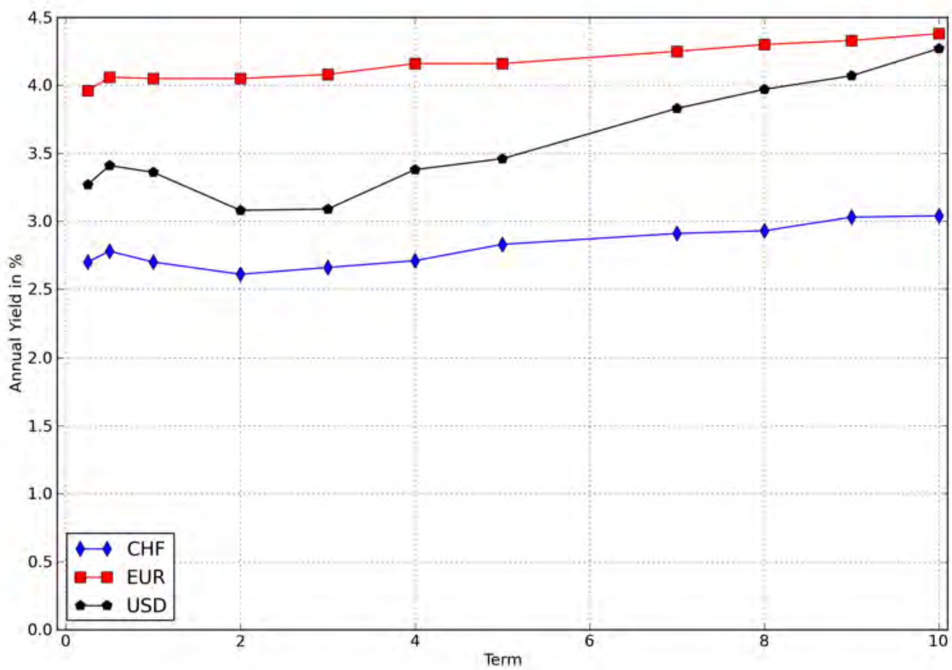

Figure 3.2. Yield curves as at 1.1.2008

This point of view provides a realistic evaluation of the interest rate. To utilise this we are going to define the so called zero coupon bond.

Definition 3.3.1 (Zero coupon bond). *Let $t \in \mathbb{R}$. Then the* zero coupon bond *with contract period t is defined by*

$$\mathcal{Z}_{(t)} = (\delta_{t,\tau})_{\tau \in \mathbb{R}^+}.$$

Thus the zero coupon bound is a security, which has the value 1 at time t.

Definition 3.3.2 (Price of a zero coupon bond). *Let $t \in \mathbb{R}$. Then the* price *of a zero coupon bond $\mathcal{Z}_{(s)}$ at time t with contract period s is denoted by*

$$\pi_t\left(\mathcal{Z}_{(s)}\right).$$

Based on these curves one can calculate the forward rates, i.e. the interest rate for year $n \rightsquigarrow n+1$, of the corresponding investment:

$$(1+i_k) = \frac{\pi_t\left(\mathcal{Z}_{(k)}\right)}{\pi_t\left(\mathcal{Z}_{(k+1)}\right)}.$$

Here i_k is called the forward rate at time t for the contract period $[k, k+1[$, it is the *expected* rate for this time interval. Therefore the discount rate is given by

$$v_k = \frac{\pi_t\left(\mathcal{Z}_{(k+1)}\right)}{\pi_t\left(\mathcal{Z}_{(k)}\right)}.$$

Thus one can use a time dependent technical interest rate and adapt the necessary elements in the expectation based on the liabilities. This is especially useful for short term contracts with cash flows which are accessible. Consider for example the acquisition of a pension portfolio. In this case one takes over the obligation to pay the pensions of a given pension fund. The described method reduces the risks taken by fixing the technical interest rate.

The third method to determine the technical interest rate uses a stochastic interest rate model. This is interesting for practical and theoretical purposes. On the one hand this method is useful for policies which are tied to the performance of funds and which provide guaranties. On the other hand it enables us to derive models which provide a tool to measure the risk of an insurance portfolio with respect to changes of the interest rate It turns out that for models with stochastic interest rate the corresponding risk does not vanish when the number of policies increases. This is a major difference to the deterministic model. Furthermore it means that the risk induced by the interest rate has a systemic, and thus dangerous, component. The construction of these models requires an analysis of the returns of several investment categories. Figure 3.3 shows the performance of two indices. These are measurements of the mean value of the return on investment in a given category. The indices in Figure 3.3 are the following

SPI	Swiss Performance Index: Swiss shares
SWISBGB	Swiss government bonds

Observing the performances of these indices (Figure 3.3) we find a significant difference between shares and bonds. The expected return is larger for shares, but they also have a larger volatility (greater variance).

For a model with stochastic interest rate one has to model processes as depicted in the Figures 3.1, 3.2 and 3.3. The main difficulty in this context is the fact that there is no common standard model. Thus the actuary is responsible for selecting an appropriate model for the given problem.

In the following section we will describe some popular models. The reader interested in further details on the financial market, in particular interest rate models, is referred for example to [Hul97].

3.4 Stochastic interest rate

In the previous section we have seen several possibilities to determine the value of a cash flow based on the interest rate. Furthermore, the basics of a stochastic interest rate model were introduced. In this section we will present explicitly several stochastic interest rate models.

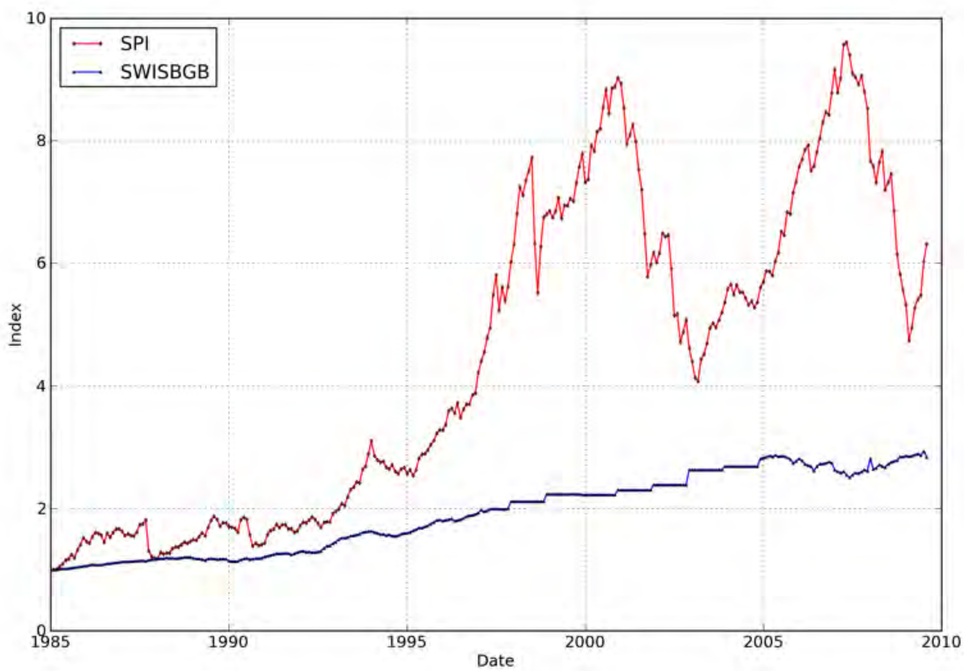

Figure 3.3. Performance of different indices

First of all one has to understand the difference between the stochastic behaviour of the interest rate and the stochastic component of the mortality. Both create a risk for an insurance company. On the one hand there is the risk induced by the fluctuation of the interest rate and on the other hand there is risk based on the individual mortality. Changes of the interest rate affect all policies to the same degree. But the variation of the risk based on the individual mortality decreases when the number of policies increases. This is due to the law of large numbers and the independence of individual lifetimes.

Now we give a brief survey of stochastic interest rate models. We will concentrate on a description of these models without rating them. Nevertheless one should note that some models (e.g. the random walk model) are unsuitable to realistically describe an interest rate process.

3.4.1 Discrete time interest rate models

Random walk: Let $\mu \in \mathbb{R}_+$ and $t \geq 0$. The interest rate i_t is defined by

$$\begin{aligned} i_t &= \mu + X_t, \mu \in \mathbb{R}, \\ X_t &= X_{t-1} + Y_t, \\ Y_t &\sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.} \end{aligned}$$

Note that this model is too simple to capture the real behaviour.

AR(1)-model: In this model the interest rate is an auto-regressive process of order 1:

$$\begin{aligned} i_t &= \mu + X_t, \\ X_t &= \phi\, X_{t-1} + Y_t, \text{ with } |\phi| < 1, \\ Y_t &\sim \mathcal{N}(0, \sigma^2). \end{aligned}$$

The model is mostly used by actuaries in England. The main idea is to start with an AR(1)-process as a model for the inflation. Then, in a second step, the models of the other economic values are based on this inflation. The constructed models have many parameters and are difficult to fit. References: [BP80] [Wil86] [Wil95].

3.4.2 Continuous time interest rate models

Brownian motion: $\delta_t = \delta + \sigma\, W_t$, where W_t is a standard Brownian motion.

Vasiček-model: The interest intensity is defined by the following stochastic differential equation

$$d\delta_t = -\alpha\,(\delta_t - \delta)\,dt + \sigma\,dW_t.$$

References: [Vas77].

Cox-Ingersoll-Ross: The interest intensity is defined by the following stochastic differential equation

$$d\delta_t = -\alpha\,(\delta_t - \delta)\,dt + \sigma\,\sqrt{\delta_t}\,dW_t.$$

References: [CIR85].

Markovian interest intensities: This model ([Nor95b]) uses a Markov chain $(X_t)_{t\geq 0}$ on a finite state space and a deterministic function $\delta_j(t)$ for each $j \in S$. Then the interest intensity is defined by

$$\delta_t = \sum_{j\in S} \chi_{\{X_t=j\}}\delta_j(t).$$

This means, that the interest intensity in a given state j at time t is determined by the corresponding deterministic function $\delta_j(.)$ evaluated at t. The model has the advantage, that it can be integrated into the Markov model. Furthermore it is very flexible due to its general state space. Therefore we will focus on this model in the following.

One should note that the Vasiček-model and the CIR-model feature mean reversion. Thus the interest intensity without stochastic noise (dW) converges in the long run toward the the mean intensity δ, since the differential equation without stochastic noise

$$d\delta_t = -\alpha\,(\delta_t - \delta)\,dt$$

has the solution

$$\delta_t = \gamma \times \exp(-\alpha\,t) + \delta.$$

The Vasiček-model and the Cox-Ingersoll-Ross-model are often used to model interest rate processes in applications.

Brownian motion and the Vasiček-model are problematic, since they allow negative interest rates with positive probability. In the Cox-Ingersoll-Ross model this can be prevented by an appropriate choice of the parameters.

In the following we assume that the presented stochastic differential equations have a solution.

We have seen above various models which are based on fundamentally different ideas. But in addition to the risk in the choice of the model there are further relevant systemic risks which affect the interest rate. These are:

The interest rate paid on an investment is not purely random. It also depends on political decisions. For example a monetary union causes the interest rates to converge, since in this case only one currency with one (random) interest rate exists (e.g. the European monetary union).

4. Cash flows and the mathematical reserve

4.1 Introduction

In the previous two chapters we introduced several types of insurances and their setup. Based on this we will now answer several fundamental questions.

First of all we will decide which general model we are going to use. Afterwards we will explain how to value and price an insurance policy.

The present value of an insurance policy, the so called mathematical reserve, has to be determined by an insurance company on a yearly basis for the annual statement. This is necessary since the company has to reserve this value. The mathematical reserve is also important for the insured when he wants to cancel his policy before maturity.

In the remainder of the chapter the insurance model from Chapter 1 will be combined with the stochastic models of Chapter 2. Obviously Markov chains with a countable state space are not the only possible stochastic model, but we will focus on these. On the one hand they are general enough to model many phenomena. On the other hand the corresponding formulas are simple enough to perform explicit calculations.

4.2 Examples

In this section we present some examples which motivate the use of the Markov chain model for insurance policies.

Example 4.2.1 (Life insurance). Usually the state space of a permanent life insurance consists of the states "dead" and "alive". Thus we use for the policy setup and for the stochastic process the state space $S = \{*, \dagger\}$, where $*$ denotes "alive" and $\dagger$ denotes "dead". Based on the benefits of such a policy, as described in Chapter 1, one has to model the corresponding stochastic process. We will use the exemplary life insurance from Chapter 2. A typical sample path of the stochastic process corresponding to this policy is shown in Figure 4.1. It indicates that at the time of death (here at $x = 45$) the corresponding payment (e.g. 200,000 USD) is due. The mortality at that time is:

$$\begin{aligned} \mu_{*\dagger}(x)\big|_{x=45} &= \exp(-9.13275 + 0.08094x - 0.000011x^2)\big|_{x=45} \\ &= 0.00404. \end{aligned}$$

This means that on average 4 out of 1000 forty-five year old men die per year.

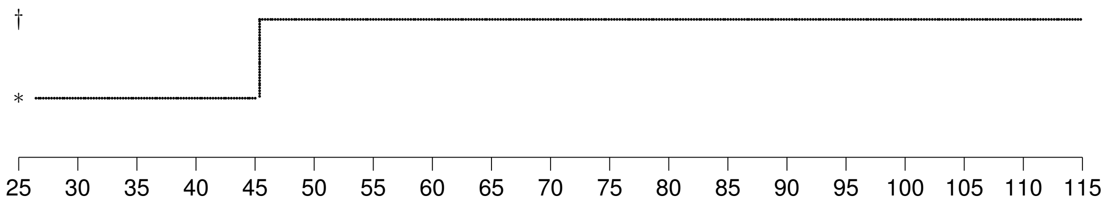

Figure 4.1. Trajectory of a mortality cover

Up to now we are not able to calculate the premiums for the policy in the example above. But we already notice the interplay between the payments and the stochastic processes.

Example 4.2.2 (Temporary disability pension). In this example we consider a policy of a disability pension which corresponds to the sample path in Figure 4.2. We want to record the various cash flows which it induces. For this we take the transition intensities from Example 2.4.2 with the additional assumption $\mu_{\diamond *}(x) = 0.05$. Then the sample path presented in Figure 4.2 causes the cash flows listed in Table 4.1.

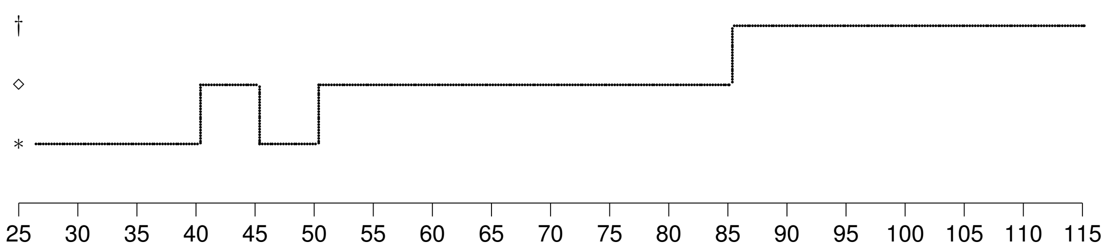

Figure 4.2. Trajectory of a disability cover

4.3 Fundamentals

In order to derive realistic models we have to know the fundamentals (underlying probabilities and biometric quantities). They are especially needed for the calculation of the premiums and mathematical reserves. As actuary one can look up the fundamentals in published tables. These list the probability of given events, for example the probability to die at a given age.

The tables used by insurance companies often incorporate a certain spread. For example one increases the probability of dying at a certain age if one calculates a life insurance. Conversely, one increases the

Table 4.1. Example of cash flows for a disability pension

time	**state**	**cash flow**	μ
$x \in\ [0, 40[$	active ($*$)	premiums	
$x = 40$	becomes disabled	disability capital	$\mu_{*\diamond} = 0.00214$
$x \in]40, 45[$	disabled ($\diamond$)	disability pension	
$x = 45$	becomes active	—	$\mu_{\diamond *} = 0.05000$
$x \in\ [45, 50[$	active ($*$)	premiums	
$x = 50$	becomes disabled	(maybe) disability capital	$\mu_{*\diamond} = 0.00387$
$x \in]50, 85[$	disabled ($\diamond$)	disability pension	
from 65		pension	
$x = 85$	dead	sum payable at death	$\mu_{\diamond\dagger} = 0.12932$

survival rate if one calculates a pension. This spread is used to decrease the risk of default and to cover possible demographic trends. That this is necessary illustrate for example the Tables 4.2 and 4.3. They list the average life expectancy for several generations, as given in the Swiss mortality tables. The average life expectancy is the number of years which a person of given age has on average still ahead. These tables clearly show that the average life expectancy increased during the last one hundred years. Therefore a spread is clearly necessary to cover this trend. Other western countries experience a similar increase of the average life expectancy.

Table 4.2. Average life expectancy based on Swiss mortality tables (male)

Alter	1881-88	1921-30	1939-44	1958-63	1978-83	1988-93	1998-03
1	51.8	61.3	64.8	69.4	72.1	73.8	76.6
20	39.6	45.2	47.9	51.5	53.8	55.3	58.0
40	25.1	28.3	30.4	32.8	35.1	36.8	39.0
60	12.4	13.8	14.8	16.2	17.9	19.3	21.1
75	5.6	6.2	6.6	7.5	8.5	9.2	10.3

Table 4.3. Average life expectancy based on Swiss mortality tables (female)

Alter	1881-88	1921-30	1939-44	1958-63	1978-83	1988-93	1998-03
1	52.8	63.8	68.5	74.5	78.6	80.5	82.2
20	41.0	47.6	51.3	56.2	60.1	61.8	63.4
40	26.7	30.9	33.4	37.0	40.7	42.5	43.8
60	12.7	15.1	16.7	19.2	22.4	24.0	25.2
75	5.7	6.7	7.4	8.6	10.7	11.9	12.8

We note that demographic quantities are constantly changing, like the average life expectancy. But where does this data actually come from and how where these tables calculated?

For the mortality tables one uses either the samples which are owned by the given insurance company or a collection of samples which is obtained jointly by several insurers. Then, to calculate the mortality rate, one counts the number of persons at risk and the number of died subjects for a given period of time (e.g. five years). The following example is based on data obtained by a large Swiss insurance company

[PT93]. Figure 4.3 shows the number of alive and dead people at a given age. Figure 4.4 shows the raw mortality and the smoothed mortality.

The smoothed mortality is obtained by a smoothing algorithm. We are not going to discuss these algorithms. But we want to note, that there are various algorithms which greatly differ in their complexity.

On the raw curve one notices for example the accident-bump (i.e. the increased mortality) between 15 and 25 years. This is not visible in the smooth curve. Thus one has to adjust in this region the smoothed mortality.

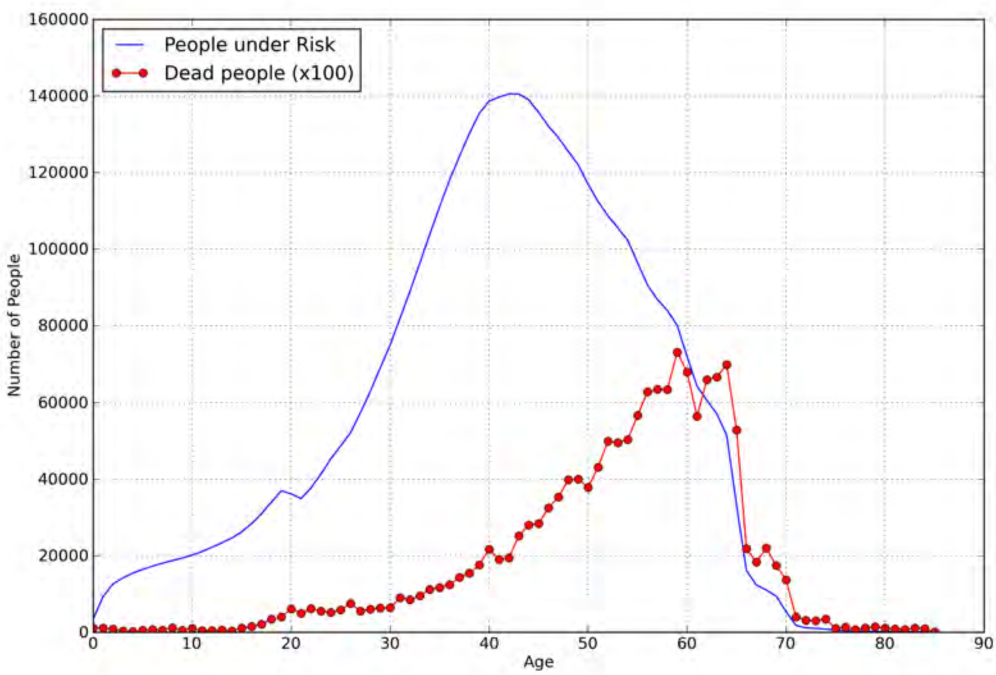

Figure 4.3. Inforce and number of dead people

In this example a polynomial of degree two was fitted to $\log(\mu_{*\dagger})$:

$$\mu_{*\dagger}(x) = \exp(-7.85785 + 0.01538 \cdot x + 5.77355 \cdot 10^{-4} \cdot x^2).$$

Analogous other demographic quantities which are relevant for the calculation of the premiums are obtained. Also for these a smoothed curve is obtained by an application a smoothing algorithm to the raw data.

The relevant probabilities and biometric quantities are collected in a catalogue. Then, with the help of such a catalogue, one can calculate the premiums and values of various products.

Table 4.4 lists the typically relevant quantities.

Table 4.4. Typical quantities for the calculation of premiums

variable	**meaning**
q_x	mortality, possibly separate for accidents and illness,
i_x	probability of becoming disabled, possibly separate for accidents and illness,
r_x	probability of becoming active again, possibly partitioned by the lengths of the disability period,
g_x	average degree of disability,
h_x	probability of being married at the time of death,
y_x	average age of the surviving marriage partner at the death of the insured.

Further details about the calculation of mortality tables and disability tables for the German and European market can be found in [DAV09].

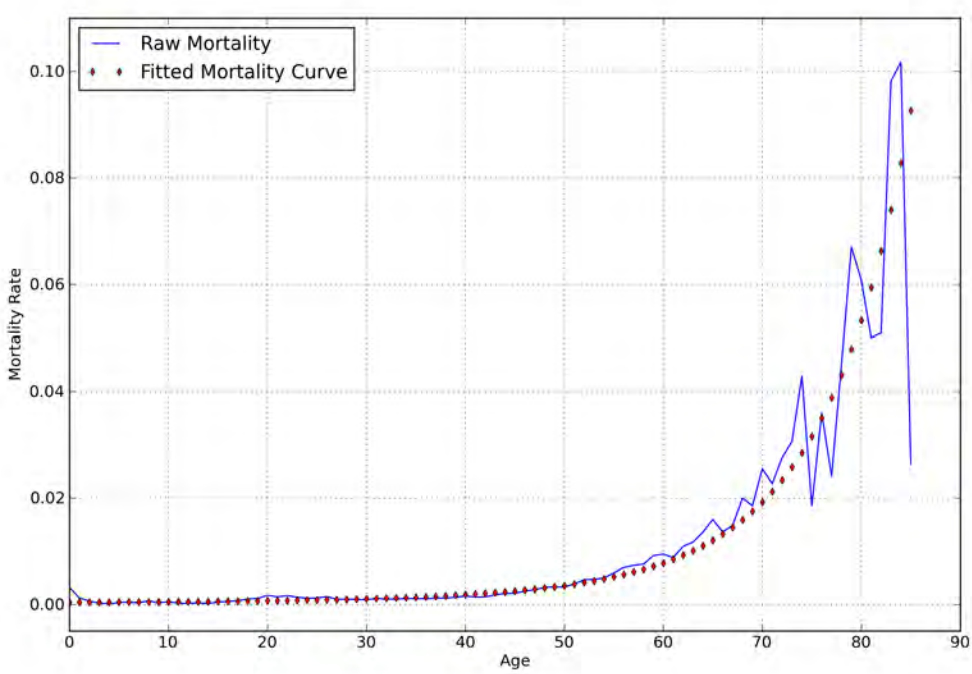

Figure 4.4. Mortality man

4.4 Deterministic cash flows

Definition 4.4.1 (Payout function). *A* deterministic payout function A *is a function*

$$A : T \to \mathbb{R}, t \mapsto A(t),$$

on $T \subset \mathbb{R}$ *with the following properties:*

1. *A is right continuous,*
2. *A is of bounded variation.*

The value $A(t)$ represents the total payments from the insurer to the insured up to time t. The payout functions are those functions in the policy setup which represent the benefits for the insured.

Example 4.4.2 (Disability insurance). We continue Example 4.2.2 by calculating the corresponding payout function. For this we assume that the policy does not contain a waiting period and that the disability pension is fixed to 20,000 USD per year (until the age of 65) with premiums of 2,500 USD per year until 65. Furthermore we suppose that the insurance was contracted at the age $x_0 = 25$. The payout function of this example is shown in Figure 4.5.

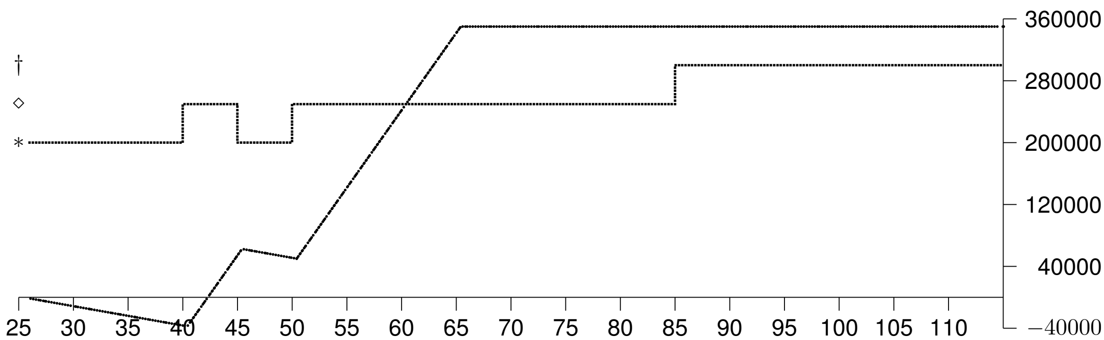

Figure 4.5. Cumulative payout of a disability annuity

Exercise 4.4.3. Derive the payout function for Example 4.4.2.

Remark 4.4.4 (Functions of bounded variation). Functions of bounded variation have the following properties [DS57]:

1. A function A of bounded variation corresponds to a measure on $\sigma(\mathbb{R})$. We will denote the measure again by A. This measure is called *Stieltjes measure*. In our setting it is also called payout measure.
2. Let A be a function of bounded variation on $\mathbb{R}$. Then there exist two positive, increasing and bounded functions B and C such that $A = B - C$. In the insurance model we can interpret B as inflow and C as outflow of cash. This representation is unique if one assumes that the measures corresponding to B and C have disjoint support. (Exercise: Calculate B and C for Example 4.4.2.)

3. Let A be the measure corresponding to a function of bounded variation on $\mathbb{R}$. Then A can be decomposed uniquely into a discrete measure μ and a continuous measure ψ. Furthermore one can decompose ψ into a part which is absolute continuous with respect to the Lebesgue measure and a remainder part. The support of μ is a countable set since A is finite on bounded sets.
4. Let A be a function of bounded variation and $T \in \sigma(\mathbb{R})$. Then $A \times \chi_T$ is also a function of bounded variation. (Here the function χ_T is the indicator function introduced in Definition 2.1.2.)

The above properties also hold for payout functions, since these are just functions of bounded variation. The decomposition of the Stieltjes measures will be useful later on. Thus we introduce the following notation:

Definition 4.4.5 (Decomposition of measures). *Let f be a function of bounded variation with corresponding Stieltjes measure A. Then we define*

$$\mu_f := A.$$

We know that we can decompose this measure uniquely into $A = B - C$, where B and C are positive measures with disjoint support. Therefore we define:

$$\begin{aligned} A^+ &:= B, \\ A^- &:= C. \end{aligned}$$

Furthermore the Stieltjes measure A can be decomposed uniquely into $A = D + E$, where D is discrete and E is continuous. Therefore we define:

$$\begin{aligned} A^{atom} &:= D, \\ A^{cont} &:= E. \end{aligned}$$

Furthermore, let μ be a measure which is absolute continuous with respect to the Lebesgue measure λ. Then we denote by $\frac{d\mu}{d\lambda}$ the Radon-Nikodym density *of μ with respect to λ.*

Above we have seen the most important properties of deterministic cash flows. This enables us to define their values with the help of the discount rate. Recall, the discount rate is given by

$$v(t) \quad = \quad \exp(-\int_0^t \delta(\tau)\, d\tau).$$

Then the present value of a cash flow is defined as follows.

Definition 4.4.6 (Value of a cash flow). *Let A be a deterministic cash flow and $t \in \mathbb{R}$. We define:*

1. *The value of a cash flow A at time t is*

$$V(t, A) := \frac{1}{v(t)} \int_0^\infty v(\tau)\, dA(\tau).$$

2. *The value of the future cash flow is*

$$V^+(t, A) := V(t, A \times \chi_{]t,\infty]}).$$

It is also called prospective value of the cash flow *or* prospective reserve.

Concerning these definitions one should note:

Remark 4.4.7. 1. The idea of the prospective reserve is to calculate the present value of the future cash flows. Thus a payment of ζ which is due in two years contributes $v(2) \times \zeta$ to the present value. Initially the reserves are defined for deterministic cash flows. To define them also for random cash flows one uses the corresponding conditional expectations.

2. The definition implicitly requires that $v(t)$ is integrable with respect to the measure A, i.e. $v \in L^1(A)$.

3. The equation $A = A^{\text{atom}} + A^{\text{cont}}$ also implies $V(t, A) = V(t, A^{\text{atom}}) + V(t, A^{\text{cont}})$. This decomposition allows us to use different methods of proof for the discrete and the continuous part of the measure.

Example 4.4.8. We want to calculate $V^+(t, A)$ for the cash flow defined in Example 4.4.2 with $\delta(\tau) = \log(1.04)$. The first step is to calculate A^+ and A^-:

$$\begin{aligned} dA^+ &= 20000\,(\chi_{[40,45[} + \chi_{[50,65[})d\tau, \\ dA^- &= 2500\,(\chi_{[25,40[} + \chi_{[45,50[})d\tau. \end{aligned}$$

Then we get for $t \in [25, 65[$

$$\begin{aligned} V^+(t, A) &= 20000 \int_t^{65} (1.04)^{-(\tau-t)}\,(\chi_{[40,45[} + \chi_{[50,65[})d\tau \\ &\quad -2500 \int_t^{65} (1.04)^{-(\tau-t)}\,(\chi_{[25,40[} + \chi_{[45,50[})d\tau. \end{aligned}$$

4.5 Stochastic cash flows

Definition 4.5.1 (Stochastic cash flow). *A* stochastic cash flow *or a* stochastic process of bounded variation *is a stochastic process $(X_t)_{t \in T}$ for which almost all sample paths are functions of bounded variation.*

Let A be a stochastic process of bounded variation such that $t \mapsto A_t(\omega)$ is right continuous and increasing for each $\omega \in \Omega$. Then it is possible to calculate the integral $\int f(\tau) d\mu_{A.(\omega)}(\tau)$ for a bounded Borel function f. Similarly, one can define P-almost everywhere the integral $\int f(\tau, \omega) d\mu_{A.(\omega)}(\tau)$ if $F_t = f(t, \omega)$ is a bounded function which is measurable with respect to the product sigma algebra. The construction of these integrals can be extended to general processes of bounded variation by decomposing the sample path, a function of bounded variation, into its positive (increasing) and negative (decreasing) part.

Definition 4.5.2. *Let $(A_t)_{t \in T}$ be a process of bounded variation on $(\Omega, \mathcal{A}, P)$ and $F : \mathbb{R} \times \Omega \to \mathbb{R}$ be a bounded and product measurable function. Then the description above yields the following definition*

$$(F \cdot A)_t(\omega) = \int_0^t F(\tau, \omega) dA\tau(\omega).$$

We also write this relation in the symbolic notation of stochastic differential equations:

$$d(F \cdot A) = F\,dA.$$

This definition allows us to give a precise definition of the stochastic cash flows in our insurance model.

Definition 4.5.3 (Policy cash flows). *We consider an insurance policy with state space S and payout functions $a_{ij}(t)$ and $a_i(t)$.*[1] *Based on Definition 2.1.8 we can define the stochastic cash flows corresponding to an insurance policy by*

$$\begin{aligned} dA_{ij}(t,\omega) &= a_{ij}(t)\, dN_{ij}(t,\omega), \\ dA_i(t,\omega) &= I_i(t,\omega)\, da_i(t), \\ dA &= \sum_{i\in S} dA_i + \sum_{(i,j)\in S\times S, i\neq j} dA_{ij}. \end{aligned}$$

The quantity $A_{ij}(t,\omega)$ is the sum of the random cash flows which are induced by transitions from state i to state j up to time t. Similarly, $A_i(t,\omega)$ represents the sum of the random cash flows up to time t which are pension payments for being in state i.

Remark 4.5.4. 1. The quantity $dA_{ij}(t,\omega)$ corresponds to the increase of the liabilities by a transition $i \rightsquigarrow j$. Therefore $A_{ij}(t,\omega)$ increases at time t by the capital benefit $a_{ij}(t)$ if at time t a transition $i \rightsquigarrow j$ takes place, i.e. if $N_{ij}(t)$ increases by 1. Similarly, $dA_i(t)$ corresponds to the increase of the liabilities caused by the insured being in state i.

2. The integrals appearing above are well defined since the corresponding processes are, by definition, of bounded variation. Moreover also the payout functions have the required regularity.

3. The quantities $(F \cdot A)_t$ are measurable for each t since F was assumed to be product measurable. Therefore also the expectation $E[(F \cdot A)_t]$ is well defined. Similarly, the conditional expectations $E[(F \cdot A)_t \mid \mathcal{F}_s]$ are well defined.

4. Thus one can apply Definition 4.4.6 (value of a cash flow) point wise (i.e. for each sample) to a stochastic cash flow. This yields the equation

$$\begin{aligned} dV(t,A) &= v(t)\, dA(t) \\ &= v(t) \left[\sum_{i\in S} I_i(t) da_i(t) + \sum_{(i,j)\in S\times S, i\neq j} a_{ij}(t) dN_{ij}(t). \right] \end{aligned}$$

5. In the discrete Markov model at most two cash flows occur during a time interval $[t, t+1[$. Firstly, if the policy is in state i, $a_i^{\text{Pre}}(t)$ is paid at the beginning of the interval. Secondly, if there is a transition $i \rightsquigarrow j$, $a_{ij}^{\text{Post}}(t)$ is due at the end of the interval. Hence the following equations can be used to calculate the total cash flows:

$$\Delta A_{ij}(t,\omega) = \Delta N_{ij}(t,\omega) a_{ij}^{\text{Post}}(t), \tag{4.1}$$

$$\Delta A_i(t,\omega) = I_i(t,\omega) a_i^{\text{Pre}}(t), \tag{4.2}$$

$$\Delta A(t,\omega) = \sum_{i\in S} \Delta A_i(t,\omega) + \sum_{i,j\in S} \Delta A_{ij}(t,\omega), \tag{4.3}$$

where $\Delta A(t)$ (and similarly ΔA_{ij}, and ΔA_i, respectively) stands for the change of $A(t)$ from t to $t+1$, eg $\Delta A(t) := A(t+1) - A(t)$.

Definition 4.5.5. *Let A and (optionally) also v be stochastic processes on $(\Omega, \mathcal{A}, P)$ which are adapted to the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. In this case the* prospective reserve *is defined by:*

[1] Thus the functions are of bounded variation and, in particular, bounded.

$$V_{\mathbb{F}}^{+}(t, A) = E\left[V^{+}(t, A) \,|\, \mathcal{F}_t\right].$$

One should note that also these reserves, like the usual expectations, might not exist, i.e., they might be infinite. In the following we will always assume that $V_{\mathbb{F}}^{+}(t, A)$ and the other quantities exist. This assumption is always satisfied in applications.

For a Markov chain the conditional expectation with respect to $\mathcal{F}_t$ depends only on the state at time t. Thus we additionally define

$$V_j^{+}(t, A) = E\left[V^{+}(t, A) \,|\, X_t = j\right].$$

The following definition fixes our assumptions on the regularity of an insurance model.

Definition 4.5.6 (Regular insurance model). *A regular insurance model consists of:*

1. *a regular Markov chain $(X_t)_{t \in T}$ with a state space S,*
2. *payout functions $a_{ij}(t)$ and $a_i(t)$,*
3. *right continuous interest intensities $\delta_i(t)$ of bounded variation.*

4.6 Mathematical reserve

The mathematical reserve is the amount of money an insurance company has to reserve for the expected liabilities in order to remain solvent. We assume that the interest intensity δ has the following structure: $\delta_t = \sum_{j \in S} I_j(t)\, \delta_j(t)$. Then the required reserves for the cash flows are defined by:

Definition 4.6.1 (Mathematical reserve). *The mathematical reserve for being in state $g \in S$ within a time interval $T \in \sigma(\mathbb{R})$ under the condition $X_t = j$ is defined by*

$$V_j(t, A_{gT}) = E\left[\frac{1}{v(t)} \times \int_T v(\tau)\, dA_g(\tau) \,|\, X_t = j\right].$$

Similarly, for transitions from g to $h \in S$, we define

$$V_j(t, A_{ghT}) = E\left[\frac{1}{v(t)} \times \int_T v(\tau)\, dA_{gh}(\tau) \,|\, X_t = j\right].$$

We use the notation $V_j(t, A_g)$ and $V_j(t, A_{gh})$ for $V_j(t, A_{g\mathbb{R}})$ and $V_j(t, A_{gh\mathbb{R}})$, respectively.

Remark 4.6.2. The definitions of the mathematical reserve can be translated to the discrete model. One just has to replace the integrals by the corresponding sums:

$$V_j(t, A_{gT}) = E\left[\frac{1}{v(t)} \times \sum_{\tau \in T} v(\tau)\, \Delta A_g(\tau) \,|\, X_t = j\right].$$

Analogous, for transitions from g to $h \in S$ we set:

$$V_j(t, A_{ghT}) = E\left[\frac{1}{v(t)} \times \sum_{\tau \in T} v(\tau + 1)\, \Delta A_{gh}(\tau) \,|\, X_t = j\right],$$

where we assumed that the payments always take place at time $\tau + 1$.

Therefore the total reserve (or mathematical reserve) for a given state j is

$$V_j(t,A) = \sum_{g \in S} V_j(t, A_g) + \sum_{g,h \in S, g \neq h} V_j(t, A_{gh})$$

for the continuous time model and

$$V_j(t,A) = \sum_{g \in S} V_j(t, A_g^{\text{Pre}}) + \sum_{g,h \in S} V_j(t, A_{gh}^{\text{Post}})$$

for the discrete time model. Thus we have defined the mathematical reserves. The next step is to calculate their values. Let us consider the relevant cash flows. On the one hand there are flows of the form $dA_1(t) = a(t) dN_{jk}(t)$ and on the other hand are flows of the form $dA_2(t) = I_j(t) dA(t)$.

The first step is to calculate the integrals $\int dA$ for the partial cash flows. Afterwards we will derive explicit formulas for the mathematical reserves.

Theorem 4.6.3. *Let $(X_t)_{t \in T}$ be a regular Markov chain on $(\Omega, \mathcal{A}, P)$ (cf. Def. 2.3.2). Furthermore let $i, j, k \in S$, $s < t$ and $T \in \sigma(\mathbb{R})$ where $T \subset [s, \infty]$. Then the following statements hold:*

1.
$$E\left[\int_T a(\tau)\, dN_{jk}(\tau) \,|\, X_s = i\right] = \int_T a(\tau)\, p_{ij}(s,\tau)\, \mu_{jk}(\tau) d\tau$$
for $a \in L^1(\mathbb{R})$.

2. *Let A be a function of bounded variation, then*
$$E\left[\int_T I_j(\tau)\, dA(\tau) \,|\, X_s = i\right] = \int_T p_{ij}(s,\tau)\, dA(\tau).$$

Proof. 1. The step functions are dense in L^1. Therefore it is enough to show the equality for functions of the form $\chi_{[a,b]}$. Moreover, the Borel σ-algebra is generated by the intervals in $\mathbb{R}_+$. Thus we can take $T = [c, d]$. Further we can set $c = a$ and $d = b$ without loss of generality, since the indicator function is equal to zero outside the interval $[a, b]$.

Define the function
$$h(t) := E\left[N_{jk}(t) \,|\, X_s = i\right].$$

Based on this definition we get

$$\begin{aligned} h(t+\Delta t) - h(t) &= E\left[N_{jk}(t+\Delta t) - N_{jk}(t) \,|\, X_s = i\right] \\ &= \sum_{l \in S} E\left[\chi_{\{X_t = l\}}(N_{jk}(t+\Delta t) - N_{jk}(t)) \,|\, X_s = i\right] \\ &= \sum_{l \in S} E\left[N_{jk}(t+\Delta t) - N_{jk}(t) \,|\, X_t = l\right] \times p_{il}(s,t). \end{aligned}$$

Now we observe that all the terms where $j \neq l$ are of order $o(\Delta t)$. Thus we get

$$= p_{ij}(s,t) \times \mu_{jk}(t) \times \Delta t + o(\Delta t).$$

Therefore $h'(t) = p_{ij}(s,t)\, \mu_{jk}(t)$, and an integration of this equation with initial condition $h(0) = 0$ yields the first statement of the theorem.

2. For the second statement one has to interchange the order of integration, which is allowed by Fubini's theorem.

Remark 4.6.4. Also these statements can be translated easily to the discrete model. One gets the equations

$$E\left[\sum_{\tau\in T} a(\tau)\,\Delta N_{jk}(\tau)\,|\,X_s = i\right] = \sum_{\tau\in T} a(\tau)\,p_{ij}(s,\tau)p_{jk}(\tau,\tau+1)$$

and

$$E\left[\sum_{\tau\in T} I_j(\tau)\,\Delta A(\tau)\,|\,X_s = i\right] = \sum_{\tau\in T} p_{ij}(s,\tau)\,\Delta A(\tau).$$

Exercise 4.6.5. Complete the proof of Theorem 4.6.3.

An important consequence of Theorem 4.6.3 is the next theorem.

Theorem 4.6.6. *Let the assumptions of Theorem 4.6.3 be satisfied. Then*

$$dM_{ij}(t) := dN_{ij}(t) - I_i(t)\,\mu_{ij}(t)dt$$

is a martingale.

Proof. We have

$$N_{ij}(t) \in L^1(\Omega, \mathcal{A}, P)$$

and

$$\int_0^t I_j(\tau)\,\mu_{ij}(\tau)d\tau \in L^1(\Omega, \mathcal{A}, P),$$

which implies

$$M_{ij}(t) \in L^1(\Omega, \mathcal{A}, P).$$

Next, we have to prove the equality $E[M_{ij}(t)\,|\,\mathcal{F}_s] = M_{ij}(s)$ for $s < t$. But, since the processes M, N and I are all derived from $(X_t)_{t\in T}$, it is enough to prove $E[M_{ij}(t)\,|\,X_s = k] = M_{ij}(s)$. But this is true, since

$$\begin{aligned} E[M_{ij}(t)\,|\,X_s = k] - M_{ij}(s) &= E\left[\int_s^t dM_{ij}(\tau)\,|\,X_s = k\right] \\ &= E\left[\int_s^t dN_{ij}(t) - I_i(t)\,\mu_{ij}(t)dt\,|\,X_s = k\right] \\ &= 0, \end{aligned}$$

where we used Theorem 4.6.3 in the last step.

Another application of Theorem 4.6.3 yields the following equations for the mathematical reserves in our insurance model:

Theorem 4.6.7. *Let a_{ij} and a_i be payout functions and $(X_t)_{t\in T}$ be a regular Markov chain on $(\Omega, \mathcal{A}, P)$. Then the following equations hold for fixed interest intensities (i.e., $\delta_i = \delta$):*

$$\begin{aligned}
&E[V(t, A_{jT})|\, X_s = i] \\
&\quad = \frac{1}{v(t)} \int_T v(\tau)\, p_{ij}(s,\tau) da_j(\tau), \\
&E[V(t, A_{jkT})|\, X_s = i] \\
&\quad = \frac{1}{v(t)} \int_T v(\tau)\, a_{jk}(\tau)\, p_{ij}(s,\tau)\, \mu_{jk}(\tau) d\tau, \\
&E[V(t, A_{jS}) V(t, A_{lT})|\, X_s = i] \\
&\quad = \frac{1}{v(t)^2} \int_{T\times S} v(\theta) v(\tau) \Big\{ \chi_{\{\theta \leq \tau\}} p_{ij}(s,\theta) p_{jl}(\theta,\tau) \\
&\qquad\qquad + \chi_{\{\theta > \tau\}} p_{il}(s,\tau) p_{lj}(\tau,\theta) \Big\} da_j(\theta) da_l(\tau), \\
&E[V(t, A_{jkS}) V(t, A_{lmT})|\, X_s = i] \\
&\quad = \frac{1}{v(t)^2} \Bigg[\int_{T\times S} v(\theta) v(\tau) \Big\{ \chi_{\{\theta \leq \tau\}} p_{ij}(s,\theta) p_{kl}(\theta,\tau) \\
&\quad + \chi_{\{\theta > \tau\}} p_{il}(s,\theta) p_{mj}(\theta,\tau) \Big\} \mu_{jk}(\theta) \mu_{lm}(\tau) a_{jk}(\theta) a_{lm}(\tau) d\theta d\tau \\
&\quad + \delta_{jk,lm} \int_{T\cap S} v(\tau)^2 p_{ij}(s,\tau) \mu_{jk}(\tau) a_{jk}^2 d\tau \Bigg], \\
&E[V(t, A_{jS}) V(t, A_{lmT})|\, X_s = i] \\
&\quad = \frac{1}{v(t)^2} \int_{T\times S} v(\theta) v(\tau) \Big\{ \chi_{\{\theta \leq \tau\}} p_{ij}(s,\theta) p_{jl}(\theta,\tau) \\
&\quad + \chi_{\{\theta > \tau\}} p_{il}(s,\tau) p_{mj}(\tau,\theta) \Big\} da_j(\theta) \mu_{lm}(\tau) a_{lm}(\tau) d\tau.
\end{aligned}$$

Proof. The first two equations are a direct consequence of Theorem 4.6.3. For a proof of the remaining equations we refer to [Nor91].

Remark 4.6.8. Also this theorem can easily be translated to the discrete setting. The following equalities hold:

$$\begin{aligned}
E[V(t, A_{jT})|\, X_s = i] &= \frac{1}{v(t)} \sum_{\tau \in T} v(\tau)\, p_{ij}(s,\tau) a_j^{\text{Pre}}(\tau), \\
E[V(t, A_{jkT})|\, X_s = i] &= \frac{1}{v(t)} \sum_{\tau \in T} v(\tau+1)\, p_{ij}(s,\tau)\, p_{jk}(\tau,\tau+1)\, a_{jk}^{\text{Post}}(\tau),
\end{aligned}$$

where we used that, for a transition $j \rightsquigarrow k$, the payments $a_{jk}^{\text{Post}}(\tau)$ are made the end of the period.

Exercise 4.6.9. Complete the proof of Theorem 4.6.7.

Given the transition probabilities one can use Theorem 4.6.7 to calculate the expectations and variances of the prospective reserves for each cash flow. Then, based on these partial reserves, one can calculate the total prospective reserves by the following result.

Theorem 4.6.10. *Let a regular insurance model (Definition 4.5.6) with deterministic interest intensities be given. Then the prospective reserves are given by*

$$V_j^+(t) = \frac{1}{v(t)} \int_{]t,\infty[} v(\tau) \sum_{g\in S} p_{jg}(t,\tau)$$

$$\times \left\{ da_g(\tau) + \sum_{S \ni h \neq g} a_{gh}(\tau) \mu_{gh}(\tau) d\tau \right\}.$$

Remark 4.6.11. The formula of the previous theorem is not very useful, since one has to calculate integrals based on the transition probabilities p_{ij}. This becomes even more complicated by the fact that in applications often only the μ_{ij} are given. In the next section we will find a more elegant way to calculate this quantity.

4.7 Recursion formulas for the mathematical reserves

In this section we will derive a recursion formula for the reserves based on their integral representation. This recursion can be used in two ways. On the one hand it can be used to prove Thiele's differential equation. On the other hand it can be applied to the discrete model. Thereby it provides a way to calculate the values for various types of insurances. We will see in the remaining sections that these recursion equations, difference equations and differential equations are a extremely useful tools for explicit calculations. We adapt our definition of mathematical reserves, in order to simplify the proofs:

Definition 4.7.1. *We define for a regular insurance model (Definition 4.5.6):*

$$W_j^+(t) \quad := \quad v(t)\, V_j^+(t).$$

The difference between W and the usual mathematical reserve V is only the discount factor. V resembles the value of the cash flow at time t, whereas W is the value of the cash flow at time 0. Thus W is only a new notation which will help to keep the proofs simple. Based on this we are now able to derive a recursion formula for the prospective reserve.

Lemma 4.7.2. *Let $j \in S$, $s < t < u$ and $(X_t)_{t \in T}$ be a regular insurance model in continuous time with deterministic interest intensities. Then the following equation holds:*

$$\begin{aligned} W_j^+(t) &= \sum_{g \in S} p_{jg}(t,u)\, W_g^+(u) \\ &\quad + \int_{]t,u]} v(\tau) \sum_{g \in S} p_{jg}(t,\tau) \left\{ da_g(\tau) + \sum_{S \ni h \neq g} a_{gh}(\tau) \mu_{gh}(\tau) d\tau \right\}. \end{aligned}$$

Proof. The proof is based on the Chapman-Kolmogorov equation. We get

$$\begin{aligned} W_j^+(t) &= \int_{]t,\infty]} v(\tau) \sum_{g \in S} p_{jg}(t,\tau) \left\{ da_g(\tau) + \sum_{S \ni h \neq g} a_{gh}(\tau) \mu_{gh}(\tau) d\tau \right\} \\ &= \left(\int_{]t,u]} + \int_{]u,\infty]} \right) v(\tau) \sum_{g \in S} p_{jg}(t,\tau) \\ &\qquad \times \left\{ da_g(\tau) + \sum_{S \ni h \neq g} a_{gh}(\tau) \mu_{gh}(\tau) d\tau \right\} \end{aligned}$$

$$
\begin{aligned}
&= \int_{]t,u]} v(\tau) \sum_{g\in S} p_{jg}(t,\tau) \left\{ da_g(\tau) + \sum_{S\ni h\neq g} a_{gh}(\tau)\mu_{gh}(\tau)d\tau \right\} \\
&\quad + \int_{]u,\infty]} v(\tau) \sum_{g\in S} \left(\sum_{k\in S} p_{jk}(t,u) p_{kg}(u,\tau) \right) \\
&\qquad \times \left\{ da_g(\tau) + \sum_{S\ni h\neq g} a_{gh}(\tau)\mu_{gh}(\tau)d\tau \right\} \\
&= \int_{]t,u]} v(\tau) \sum_{g\in S} p_{jg}(t,\tau) \left\{ da_g(\tau) + \sum_{S\ni h\neq g} a_{gh}(\tau)\mu_{gh}(\tau)d\tau \right\} \\
&\quad + \sum_{k\in S} p_{jk}(t,u) \Bigg(\int_{]u,\infty]} v(\tau) \sum_{g\in S} p_{kg}(u,\tau) \\
&\qquad \times \left\{ da_g(\tau) + \sum_{S\ni h\neq g} a_{gh}(\tau)\mu_{gh}(\tau)d\tau \right\} \Bigg) \\
&= \sum_{g\in S} p_{jg}(t,u)\, W_g^+(u) \\
&\quad + \int_{]t,u]} v(\tau) \sum_{g\in S} p_{jg}(t,\tau) \left\{ da_g(\tau) + \sum_{S\ni h\neq g} a_{gh}(\tau)\mu_{gh}(\tau)d\tau \right\}.
\end{aligned}
$$

The recursion formula can also be translated to the discrete model. For this one assumes that the payments are done at discrete times rather than in continuous time. For example pensions are paid at the beginning of the interval and death capital is paid at the end of the interval. We denote the payments at the beginning of the year by $a_i^{Pre}(t)$ and those at the end of the year by $a_{ij}^{Post}(t)$. Thus, in particular we assume that a transition between states can only occur at the end of the year.

Setting $\Delta t = 1$ in the previous lemma yields the following recursion for the reserves in the discrete setting.

Theorem 4.7.3 (Thiele's difference equation). *For a discrete time Markov model the prospective reserve satisfies the following recursion:*

$$
V_i^+(t) \quad = \quad a_i^{Pre}(t) + \sum_{j\in S} v_t\, p_{ij}(t) \big\{ a_{ij}^{Post}(t) + V_j^+(t+1) \big\}.
$$

Proof. We know that $A(t) = \sum_{k\leq t} \Delta A(k)$ and also that

$$
\Delta V(t,A) = v(t) \left[\sum_{j\in S} I_j(t) \times a_i^{\mathrm{Pre}}(t) + \sum_{(i,j)\in S\times S} \Delta N_{ij}(t) \times a_{ij}^{\mathrm{Pre}}(t) \right].
$$

Hence we have

$$
\begin{aligned}
V_i^+(t) &= \frac{1}{v(t)} \mathbb{E}\left[\sum_{\tau=t}^{\infty} v(\tau) \times \Delta A(\tau) \,|\, X_t = i \right] \\
&= \frac{1}{v(t)} \mathbb{E}\left[\sum_{j\in S} I_j(t+1) \times \sum_{\tau=t}^{\infty} v(\tau) \times \Delta A(\tau) \,|\, X_t = i \right],
\end{aligned}
$$

remarking that $\sum_{j \in S} I_j(t+1) = 1$. If we now consider all the terms in $\Delta A(t)$ for a given $I_j(t+1)$ for $j \in S$, it becomes obvious that the Markov chain changes from $i \to j$ and in consequence only $N_{ik}(t)$ increases by one for $k = j$. If we furthermore use the projection property and the linearity of the conditional expected value and the fact that $\mathbb{E}\left[I_j(t+1) \,|\, X_t = i\right] = p_{ij}(t, t+1)$, together with the Markov property, we get the formula if we split $V_i^+(t)$ as follows:

$$
\begin{aligned}
V_i^+(t) &= \frac{1}{v(t)} \mathbb{E}\left[\sum_{\tau=t}^{\infty} v(\tau) \times \Delta A(\tau) \,|\, X_t = i\right] \\
&= \frac{1}{v(t)} \mathbb{E}\left[\left\{\sum_{\tau=t}^{t} + \sum_{\tau=t+1}^{\infty}\right\} v(\tau) \times \Delta A(\tau) \,|\, X_t = i\right].
\end{aligned}
$$

Doing this decomposition we get for the first part:

$$
\mathrm{Part}_1 = a_i^{Pre}(t) + \sum_{j \in S} v_i(t)\, p_{ij}(t) a_{ij}^{Post}(t),
$$

and for the second:

$$
\mathrm{Part}_2 = \sum_{j \in S} v_i(t)\, p_{ij}(t) V_j^+(t+1).
$$

Adding the two parts together we get the desired result:

$$
V_i^+(t) = a_i^{Pre}(t) + \sum_{j \in S} v_i(t)\, p_{ij}(t) \big\{ a_{ij}^{Post}(t) + V_j^+(t+1) \big\}.
$$

More concretely we have

$$
\begin{aligned}
V_i^+(t) &= \frac{1}{v(t)} \mathbb{E}\left[\sum_{j \in S} I_j(t+1) \times \sum_{\tau=t}^{\infty} v(\tau) \times \Delta A(\tau) \,|\, X_t = i\right] \\
&= a_i^{Pre}(t) + \sum_{j \in S} \mathbb{E}\left[I_j(t+1) \times \sum_{\tau=t}^{\infty} \frac{v(\tau)}{v(t)} \times \Delta A(\tau) \,|\, X_t = i\right] \\
&= a_i^{Pre}(t) + \sum_{j \in S} \mathbb{E}\Bigg[I_j(t+1) v_i(t) \Bigg\{ a_{ij}^{Post} + \\
&\quad + \mathbb{E}\Bigg[\sum_{\tau=t+1}^{\infty} \frac{v(\tau)}{v(t+1)} \times \Delta A(\tau) \,\Bigg|\, X_t = i, X_{t+1} = j\Bigg]\Bigg\} \Bigg|\, X_t = i\Bigg] \\
&= a_i^{Pre}(t) + \sum_{j \in S} v_i(t)\, p_{ij}(t) \big\{ a_{ij}^{Post}(t) + V_j^+(t+1) \big\}.
\end{aligned}
$$

Remark 4.7.4. – The formula shows that errors which are introduced by the discretisation of time are due to payments between the discretisation times.

– The recursion formula of the mathematical reserve is very important for applications, since it provides a way to calculate a single premium and yearly premiums. In fact it is the most important formula for explicit calculations.

– To solve a differential equation or a difference equation one needs a boundary condition. For example, if one calculates a pension, the boundary condition is given by the fact that the reserve has to be equal to zero at the final age ω.

4.8 Calculation of the premiums

In this section we are going to calculate single premiums and yearly premiums for several types of insurance policies. The calculations in the examples are based on the discrete recursion (Theorem 4.7.3). We start with an endowment policy.

Example 4.8.1 (Endowment policy in discrete time). We consider the insurance defined in Example 4.2.1. Thus there is a death benefit of 200,000 USD. Moreover we assume an endowment of 100,000 USD and a starting age of 30 with 65 as fixed age at maturity.

– How much is a single premium for this insurance, given a technical interest rate of 3.5%?

– How much are the corresponding yearly premium?

We use the mortality rates given by (2.13). First we calculate the single premium. The following payout functions are given:

$$
\begin{aligned}
a_{*\dagger}^{\text{Post}}(x) &= \begin{cases} 200000, & \text{if} \quad x < 65, \\ 0, & \text{otherwise,} \end{cases} \\
a_{**}^{\text{Post}}(x) &= \begin{cases} 0, & \text{if} \quad x < 64, \\ 100000, & \text{if} \quad x = 64, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
$$

An application of Theorem 4.7.3 yields the results presented in Table 4.5. Here one has to note that the mathematical reserves for the case of survival and the case of death have to be calculated separately. The reserve in the case of survival is $V_*(t, A_{**\mathbb{R}})$ and the reserve in the case of death is $V_*(t, A_{*\dagger\mathbb{R}})$ (cf. Definition 4.6.1).

The table of the mathematical reserves indicates that the recursion formula was used with the boundary condition $x = 65$. Figure 4.6 shows the necessary reserves for several values of the technical interest rate.

After the calculation of the single premium we will now consider the case of yearly premiums. The yearly payment of the premiums is modelled by the following payout function:

$$
a_{*}^{\text{Pre}}(x) = \begin{cases} -P, & \text{if} \quad x < 65, \\ 0, & \text{otherwise.} \end{cases}
$$

P has to be calculated such that the value of the insurance is equal to zero at the beginning of the policy. (Equivalence principle: The expected value of the benefits provided by the insurer and the value of the

Table 4.5. Reserves for an endowment

age	q_x	**res. for endowment**	**res. for death benefit**	**sums reserve**
65	0.01988	**100000**	**0**	**100000**
64	0.01836	94844	3548	98392
63	0.01696	90083	6647	96730
62	0.01566	85674	9348	95022
61	0.01446	81579	11696	93275
60	0.01336	77768	13730	91498
55	0.00897	62086	20275	82360
50	0.00602	50444	22766	73210
45	0.00404	41470	22956	64426
40	0.00271	34362	21874	56236
35	0.00181	28624	20135	48759
30	0.00121	23928	18116	42044

expected premium payments by the insured coincide.) The simplest method to determine the value of P is to consider V_x^{payout} (given by the first two payout functions of the policy) and V_x^{premiums} (given by the third function of the policy, i.e. by the premiums paid in). The total mathematical reserve is then given by $V_x = V_x^{\text{payout}} + V_x^{\text{premiums}}$. But we know that $V_x^{\text{premiums}} = P \times V_x^{\text{premiums},P=1}$ holds. Thus we can calculate P by the formula

$$P = -V_x^{\text{payout}} / V_x^{\text{premiums},P=1},$$

since V_x at inception of the policy is 0, as a consequence of the equivalence principle. For our example we get

$$P = 2.129,15 \text{ USD per year.}$$

Table 4.6. Reserves for an endowment with yearly premiums

age	**present value premiums**	**present value payout**	**reserve**
65	0	100000	100000
64	-2129	98392	96263
63	-4149	96730	92581
62	-6069	95022	88952
61	-7901	93275	85374
60	-9653	91498	81845
50	-23928	73210	49282
40	-34292	56236	21943
30	-42044	42044	0

Table 4.6 lists the reserves for this insurance with yearly premiums. Figure 4.7 illustrates the same data in a graph.

Exercise 4.8.2. Do the calculation for the previous example.

In the next example we will consider the simple disability insurances model which we looked at earlier. We will show how to model the disability pension with and without an exemption from payment of premiums.

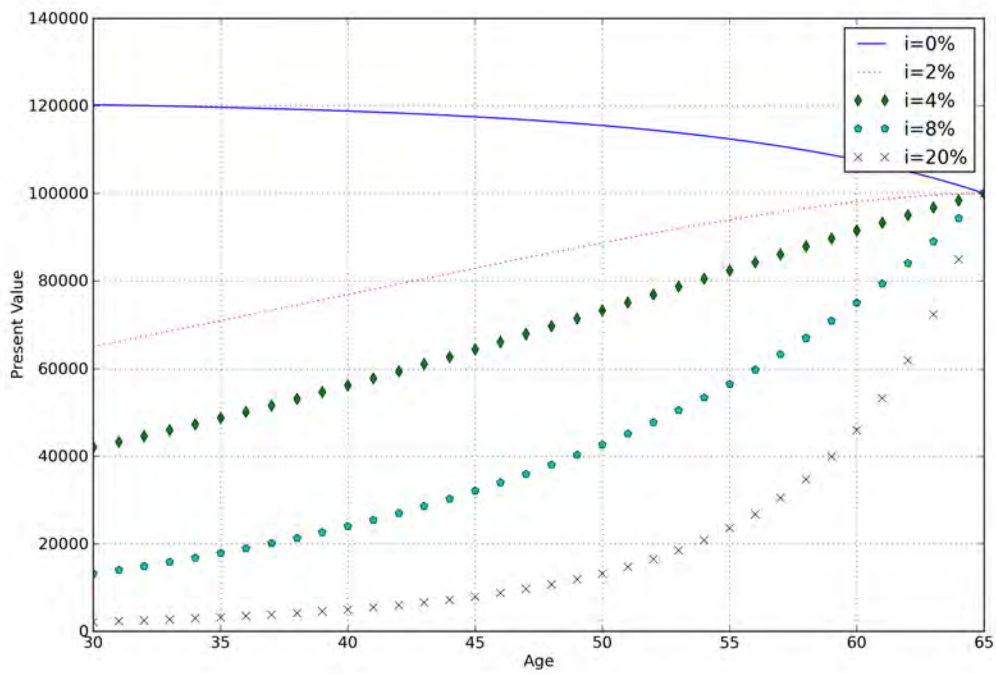

Figure 4.6. Mathematical reserves as a function of interest rates

Example 4.8.3 (Disability insurance). We use the model for the disability insurance introduced in Example 2.4.2. Thus in particular we do not incorporate the possibility that insured becomes active again. Moreover we also do not model a waiting period.

– Calculate for a 30-year old man the present value of a (new) disability pension based on 65 as age at maturity and a technical interest rate of 4%.

– Compare for the same person the present value of the premiums for a policy with exemption from payment of premiums and for a policy without this option.

First we calculate the present value of the payouts of the future disability pension. In this case the non-trivial functions which model the policy are: (where we assumed that the disability pension is payable in advance and has the value 1.)

$$a_{\diamond}^{\mathrm{Pre}} = \begin{cases} 1, & \text{if} \quad x < 65, \\ 0, & \text{otherwise.} \end{cases}$$

The boundary conditions are zero in this case, i.e., there is no payment when the age of maturity is reached. Table 4.7 lists the calculated values for this example. Thus one has to pay 4,396.8 USD as single premium for a future disability pension of 10,000 USD. Furthermore, the reserve is 170,790 USD for a disabled man of 35 years with the same disability pension as above.

Next we consider the present values of the premiums. We have to treat the following two cases separately. On the one hand there is the present value of a policy without exemption from payment of premiums

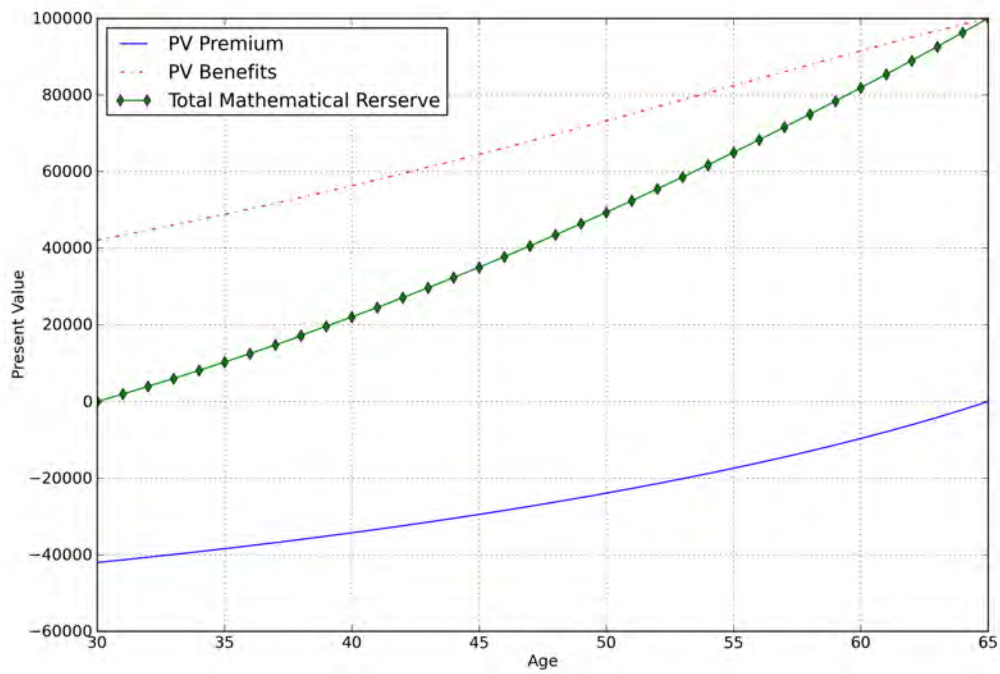

Figure 4.7. Endowment policy against regular premium payment

$$a_*^{\text{Pre}} = \begin{cases} 1, & \text{if} \quad x < 65, \\ 0, & \text{otherwise}, \end{cases}$$

$$a_\diamond^{\text{Pre}} = \begin{cases} 1, & \text{if} \quad x < 65, \\ 0, & \text{otherwise}. \end{cases}$$

On the other hand there is the present value of a policy with exemption from payment of premiums ("premium raider")

$$a_*^{\text{Pre}} = \begin{cases} 1, & \text{if} \quad x < 65, \\ 0, & \text{otherwise}, \end{cases}$$

$$a_\diamond^{\text{Pre}} = \begin{cases} 0, & \text{if} \quad x < 65, \\ 0, & \text{otherwise}. \end{cases}$$

The difference between these two policies is, that in the first case also with status "disabled" the insured has to pay the premiums. In both cases the boundary condition at the age of 65 is 0. An application of Thiele's difference equations yields the values listed in Table 4.8. We note that the value of the paid-in premiums is smaller in the case of an exemption of premiums. This can also be seen in the plot of these values in Figure 4.8.

Now we are able to calculate the yearly premium for a disability insurance of 10,000 USD with exemption from premiums

$$P = 4396.8/18.09044 = 243.05 \text{ USD per year}.$$

Table 4.7. Reserves of a disability pension

age	$p_{*\dagger}$	$p_{*\diamond}$	$V_*(x)$	$V_\diamond(x)$
65	0.02289	0.02794	**0.00000**	**0.00000**
64	0.02101	0.02439	0.00000	1.00000
63	0.01929	0.02129	0.02047	1.94299
62	0.01772	0.01860	0.05372	2.83515
61	0.01628	0.01625	0.09427	3.68174
60	0.01495	0.01420	0.13828	4.48719
55	0.00983	0.00732	0.34176	8.01299
50	0.00653	0.00387	0.46175	10.90260
45	0.00439	0.00214	0.50531	13.31967
40	0.00301	0.00127	0.50178	15.35782
35	0.00212	0.00084	0.47493	17.07904
30	0.00155	0.00062	0.43968	18.53012

Table 4.8. Present value of the premiums

age	$V(x)$ without premium raider	$V(x)$ with premium raider
65	**0.00000**	**0.00000**
64	1.00000	1.00000
63	1.94299	1.92251
62	2.83515	2.78144
61	3.68174	3.58747
60	4.48719	4.34891
55	8.01299	7.67123
50	10.90260	10.44085
45	13.31967	12.81436
40	15.35782	14.85604
35	17.07904	16.60411
30	18.53012	18.09044

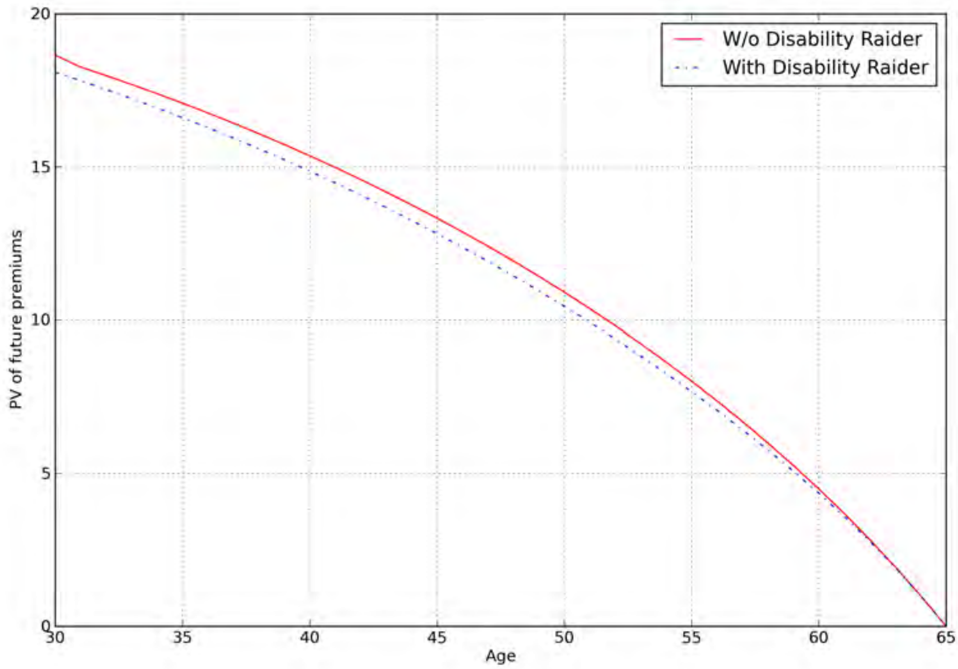

Figure 4.8. Present value of premiums

Exercise 4.8.4. 1. Do the calculations of the above example also for a model which includes the possibility of reactivation (cf. Example 4.2.2).

2. Extend the model by incorporating a waiting period of one year.

Next we consider an insurance on two lives. There are several possible states for which the policy could guarantee a pension.

Example 4.8.5 (Pension on two lives). We start with the calculation of a single premium for several types of an insurance on two lives. For this we assume that the two persons have the same mortality as given in Example 4.2.1 and that $x_1 = 30$ and $x_2 = 35$ are fixed. We set the technical interest rate to 3.5 %, and $\omega = 114$ be the maximal possible age of a living person.

There are three possible pensions: (we denote by x_1 the age of the first person)

state	type	formula
$**$	both persons are alive	$a_{**}^{\text{Pre}}(x) = \alpha_{**} \begin{cases} 0, & \text{if } x_1 < 65, \\ 1, & \text{otherwise,} \end{cases}$
$*\dagger$	second person is dead	$a_{*\dagger}^{\text{Pre}}(x) = \alpha_{*\dagger} \begin{cases} 0, & \text{if } x_1 < 65, \\ 1, & \text{otherwise,} \end{cases}$
$\dagger *$	first person is dead	$a_{\dagger *}^{\text{Pre}}(x) = \alpha_{\dagger *} \begin{cases} 0, & \text{if } x_1 < 65, \\ 1, & \text{otherwise.} \end{cases}$

The definitions of the pensions above are particular, since the pension for the second life (i.e. for $\dagger *$) is paid at the age of 65. Usually this pension would be paid immediately after the death of the first person.

We set $x = (x_1, x_2)$ and suppose that the two insured die independently. In this case the recursion takes the form

$$\begin{aligned}
V_{**}(x) &= a_{**}^{\text{Pre}}(x) + p_{x_1}\, p_{x_2}\, v\, V_{**}(x+1) + p_{x_1}\,(1-p_{x_2})\, v\, V_{*\dagger}(x+1) \\
&\quad + (1-p_{x_1})\, p_{x_2}\, v\, V_{\dagger *}(x+1), \\
V_{*\dagger}(x) &= a_{*\dagger}^{\text{Pre}}(x) + p_{x_1}\, v\, V_{*\dagger}(x+1), \\
V_{\dagger *}(x) &= a_{\dagger *}^{\text{Pre}}(x) + p_{x_2}\, v\, V_{\dagger *}(x+1).
\end{aligned}$$

This recursion yields the values listed in Table 4.9.

Table 4.9. Mathematical reserves (res.) for an insurance on two lives

Notations

$V_{*\cdot}$(R1)	res. for the pension of 1th life, independent of 2nd life
$V_{\cdot *}$(R2)	res. for the pension of 2nd life, independent of 1th life
V_{**}(R1)	res. for the pension of 1th life, if $X_t = (**)$,
V_{**}(R2)	res. for the pension of 2th life, if $X_t = (**)$,
V_{**}(R $(**)$)	res. for the pension of joint lives, if $X_t = (**)$,.

Alter 1	Alter 2	$V_{*\cdot}$(R1)	$V_{\cdot *}$(R2)	V_{**}(R1)	V_{**}(R2)	V_{**}(R $(**)$)
115	120	**0.00000**		**0.00000**	**0.00000**	**0.00000**
114	119	1.00000		0.00000	0.00000	1.00000
113	118	1.11505		0.11505	0.00000	1.00000
112	117	1.19991		0.19991	0.00000	1.00000
111	116	1.28640		0.28640	0.00000	1.00000
110	**115**	1.37771	**0.00000**	0.37771	0.00000	1.00000
109	114	1.47450	1.00000	0.45824	0.00000	1.01626
108	113	1.57710	1.11505	0.52973	0.06845	1.04736
90	95	4.64366	3.53796	2.04747	0.94178	2.59618
75	80	9.05696	7.43141	3.39065	1.76510	5.66631
65	70	12.54173	10.77780	3.88770	2.12377	8.65403
55	60	7.78663	6.26733	3.37939	1.86010	4.40724
40	45	4.30964	3.34208	2.13044	1.16288	2.17920
30	35	3.00101	2.30680	1.52353	0.82932	1.47748

Exercise 4.8.6. 1. Calculate the present values of the premiums for the insurance on two lives. Note that also in this calculation one has to consider three different cases.

2. Create a model for an orphan's pension. For this one has to consider three persons: farther, mother and child. Define the payout functions for a policy which pays 5,000 USD to the child if one of parent dies and 10,000 USD if both die. Assume that the policy matures if the child is 25 years old.

5. Difference equations and differential equations

5.1 Introduction

In this chapter we focus on the Markov model in continuous time. The differential equations are the continuous counter part to the difference equations of the discrete model.

These differential equations where first proved for simple insurance models by Thiele at the end of the 19th century. We are going to derive these equations for the Markov model. They are useful in two ways. On the one hand they help to deepen our understanding of the model. On the other hand they can be used to calculate the premiums for a policy.

5.2 Thiele's differential equations

In this section we are going to derive Thiele's differential equations for the mathematical reserve. For simplicity we consider in this chapter only reserves without jumps. Later on we will also allow jumps, but then the proofs become more involved.

Theorem 5.2.1 (Thiele's differential equation). *Let $(X_t)_{t \in T}$, a_{ij}, a_i and δ_t be a regular insurance model (Definition 4.5.6). Moreover, $da_g(t)$ be absolute continuous with respect to the Lebesgue measure λ, i.e. $da_g(t) = a_g(t)\, d\lambda$. (Thus the payout function $A_g(t)$ is continuous.) Then, assuming a deterministic interest intensity, the following statements hold:*

1. *$W_g^+(t)$ is continuous for all $g \in S$.*
2. $\frac{\partial}{\partial t} W_j^+(t) = -v(t) \left\{ a_j(t) + \sum_{j \neq g \in S} \mu_{jg}(t) a_{jg}(t) \right\}$
   $+\mu_j(t) W_j^+(t) - \sum_{j \neq g \in S} \mu_{jg}(t) W_g^+(t)$. *(Thiele's differential equation)*
3. $V_j^+(t) = \frac{1}{v(t)} \left[\begin{array}{c} \int_t^u v(\tau) \overline{p}_{jj}(t,\tau) \{ a_j(\tau) + \sum\limits_{j \neq g \in S} \mu_{jg}(\tau) (a_{jg}(\tau) \\ + V_g^+(\tau))\} d\tau + v(u) \overline{p}_{jj}(t,u) V_j^+(u^-) \end{array} \right]$.

Proof. The proof of the first statement is left as an exercise to the reader. To prove the second statement we fix $j \in S, t \in \mathbb{R}$ and $\Delta t > 0$. Then Lemma 4.7.2 implies

$$W_j^+(t) \quad = \quad v(t) \left\{ a_j(t) + \sum_{j \neq g \in S} \mu_{jg}(t) a_{jg}(t) \right\} \Delta t$$

$$
\begin{aligned}
&+(1-\mu_j(t)\Delta t)W_j^+(t+\Delta t)\\
&+\sum_{j\neq g\in S}\mu_{jg}(t)W_g^+(t+\Delta t)\,\Delta t+o(\Delta t),
\end{aligned}
$$

where we used the following facts

$$
\begin{aligned}
p_{jj}(t,t+\Delta t) &= 1-\Delta t\,\mu_j(t)+o(\Delta t),\\
p_{jk}(t,t+\Delta t) &= \Delta t\,\mu_{jk}(t)+o(\Delta t).
\end{aligned}
$$

The above equation yields

$$
\begin{aligned}
\frac{W_j^+(t+\Delta t)-W_j^+(t)}{\Delta t} &= -v(t)\,\{a_j(t)+\sum_{j\neq g\in S}\mu_{jg}(t)a_{jg}(t)\}\\
&+\ \mu_j(t)\,W_j^+(t+\Delta t)\\
&-\ \sum_{g\neq j}\mu_{jg}(t)W_g^+(t+\Delta t)+\frac{o(\Delta t)}{\Delta t}.
\end{aligned}
$$

Letting $\Delta t\longrightarrow 0$ we get

$$
\begin{aligned}
\frac{\partial}{\partial t}W_j^+(t) &= -v(t)\left\{a_j(t)+\sum_{j\neq g\in S}\mu_{jg}(t)a_{jg}(t)\right\}+\mu_j(t)W_j^+(t)\\
&- \sum_{j\neq g\in S}\mu_{jg}(t)W_g^+(t).
\end{aligned}
$$

For the proof of the third statement we use Thiele's differential equation:

$$
\begin{aligned}
&\exp(-\int_o^t\mu_j(\tau)d\tau)\left(-v(t)\{a_j(t)+\sum_{j\neq g\in S}\mu_{jg}(t)a_{jg}(t)\}\right.\\
&\qquad\left.-\sum_{j\neq g\in S}\mu_{jg}(t)W_g^+(t)\right)\\
&= \exp(-\int_o^t\mu_j(\tau)d\tau)\left(\frac{\partial}{\partial t}W_j^+(t)-\mu_j(t)W_j^+(t)\right)\\
&= \frac{\partial}{\partial t}\left(\exp(-\int_o^t\mu_j(\tau)d\tau)W_j^+(t)\right).
\end{aligned}
$$

An integration $\int_t^u$ of both sides yields

$$
\begin{aligned}
&\exp(-\int_o^t\mu_j(\tau)d\tau)\left(\exp(-\int_t^u\mu_j(\tau)d\tau)W_j^+(u)-W_j^+(t)\right)\\
&=\int_t^u\exp(-\int_o^t\mu_j(\xi)d\xi)\exp(-\int_t^\tau\mu_j(\xi)d\xi)
\end{aligned}
$$

$$\times \Big[- v(\tau)\Big\{ a_j(\tau) + \sum_{j\neq g \in S} \mu_{jg}(\tau) a_{jg}(\tau) \Big\} - \sum_{j\neq g\in S} \mu_{jg}(\tau) W_g^+(\tau) \Big] d\tau.$$

Hence, we get

$$\begin{aligned} V_j^+(t) &= \frac{1}{v(t)} \Big[\int_t^u v(\tau) \overline{p}_{jj}(t,\tau) \{ a_j(\tau) + \sum_{j\neq g\in S} \mu_{jg}(\tau) \\ & \quad \times (a_{jg}(\tau) + V_g^+(\tau)) \} d\tau + v(u) \overline{p}_{jj}(t,u) V_j^+(u^-) \Big], \end{aligned}$$

where we used that $\overline{p}_{jj}(t,\tau) = \exp(-\int_t^\tau \mu_j(\xi) d\xi)$.

Remark 5.2.2. We derived the following integral equation from Thiele's differential equation:

$$V_j^+(t) = \frac{1}{v(t)} \left[\begin{array}{c} \int_t^u v(\tau)\overline{p}_{jj}(t,\tau)\{ \underbrace{a_j(\tau)}_{I} + \sum_{j\neq g\in S} \overbrace{\underbrace{\mu_{jg}(\tau)(a_{jg}(\tau)}_{IIa} + V_g^+(\tau))}^{IIb} \} d\tau \\ + \underbrace{v(u)\overline{p}_{jj}(t,u)V_j^+(u^-)}_{III} \end{array} \right].$$

This formula shows the structure of the reserve. The components of the reserve are:

I) reserve for payments in state j (pensions and premiums),
II) reserves for state transitions composed of
 IIa) transition cost (e.g. death benefit) and
 IIb) necessary reserves in the new state,
III) reserve, for the case that the insured is still in j after $[t, u]$.

5.3 Examples - Thiele's differential equation

In this section we look at examples related to those in the discrete setting. In the first example the differential equations have an explicit solution.

Example 5.3.1 (Term life insurance). We consider a term life insurance with death benefit b, which is financed by a premium of size c. In this situation the differential equations take the following form:

$$\begin{aligned} \frac{\partial}{\partial t} W_*(t) &= v^t(c - \mu_{x+t}\, b) + \mu_{x+t} W_*(t) - \mu_{x+t} W_\dagger(t), \\ \frac{\partial}{\partial t} W_\dagger(t) &= 0 \end{aligned}$$

with the boundary condition $W_*(s-x) = W_\dagger(s-x) = 0$, where s denotes the age of maturity of the policy.

Next, we are going to calculate the mathematical reserve. The above equations obviously imply $W_{\dagger}(t) \equiv 0$. Thus we only have to calculate $W_*(t)$. The homogeneous part of the equation satisfies

$$\frac{dW_*(t)}{W_*(t)} = \mu_{x+t} dt$$

and therefore

$$L_h(t) = A \times \exp\Big(\int_0^t \mu_{x+\tau}\, d\tau\Big).$$

By variation of constants we get

$$\begin{aligned}
L_p(t) &= A(t) \times L_h(t),\\
\frac{d}{dt} L_p &= A' \times L + A \times L'\\
&= A' \times L + A \times L\\
&= A' \times L + L_p,\\
A' \times L &= v^t (c - \mu_{x+t}\, b),\\
A' &= v^t (c - \mu_{x+t}\, b) \exp\Big(-\int \mu_{x+\tau}\, d\tau\Big)\\
&= v^t (c - \mu_{x+t} b)\, {}_t p_x,\\
A(t) &= \int_0^t v^\tau (c - \mu_{x+\tau} b)\, {}_\tau p_x d\tau.
\end{aligned}$$

Finally, the boundary condition $W_*(s-x) = 0$ yields

$$\begin{aligned}
W_*(s-x) &= A(s-x) \times L(s-x)\\
&= \left[\int_0^{s-x} v^\tau (c - \mu_{x+\tau} b)\, {}_\tau p_x d\tau\right] \times \left[\exp\Big(\int_0^{s-x} \mu_{x+\tau} d\tau\Big)\right],\\
c &= b\, \frac{\int_0^{s-x} v^\tau\, {}_\tau p_x\, \mu_{x+\tau} d\tau}{\int_0^{s-x} v^\tau\, {}_\tau p_x d\tau}.
\end{aligned}$$

Example 5.3.2 (Endowment policy). We consider the endowment policy defined in Example 4.8.1. Thus it contains a death benefit of 200,000 USD and an endowment of 100,000 USD. We consider a 30 year old man and 65 as the age of maturity of the policy.

– How much is a single premium for this insurance if the technical interest rate is 3.5%?

– How do these results compare to the values in the corresponding example in discrete time?

We use the mortality rates given by (2.13). For the single premium the following payout function defines the policy:

$$a_{*\dagger}(x) = \begin{cases} 200000, & \text{if} \quad x < 65,\\ 0, & \text{otherwise.} \end{cases}$$

Now Thiele's differential equations are

$$\begin{aligned}
\frac{\partial}{\partial t} W_*(t) &= v^t (c - \mu_{x+t}\, a_{*\dagger}(x+t)) + \mu_{x+t}\, W_*(t) - \mu_{x+t}\, W_\dagger(t), \\
\frac{\partial}{\partial t} W_\dagger(t) &= 0,
\end{aligned}$$

with the boundary conditions $W_*(s-x) = 100000 \times v(s)$ and $W_\dagger(s-x) = 0$. Then Theorem 5.2.1 yields the results listed in Table 5.1.

Table 5.1. Discretisation error for an endowment policy

age	$\mu_{*\dagger}(x)$	**reserve discrete model**	**reserve cont. model**	**diff. in %**
65	0.01988	**100000**	**100000**	
64	0.01836	98392	98512	0.12
63	0.01696	96730	96955	0.23
62	0.01566	95022	95341	0.34
61	0.01446	93275	93678	0.43
60	0.01336	91498	91975	0.52
55	0.00897	82360	83092	0.89
50	0.00602	73210	74059	1.16
45	0.00404	64426	65308	1.37
40	0.00271	56236	57096	1.53
35	0.00181	48759	49566	1.65
30	0.00121	42044	42782	1.75

Note that the difference of the reserves in the discrete and continuous model have always the same sign. This is caused by the fact, that people only die at the end of the year in the discrete model. Therefore the required single premium is smaller than in the continuous model.

Exercise 5.3.3. – Calculate yearly premiums for the previous example.

– What happens to the discretisation error, if we suppose that people die in the discrete model only at the middle of the year?

– What happens, if we suppose that the interest rates drops linearly from 6% at the age of 30 to 3% at the age of 65?

Exercise 5.3.4. Calculate with the continuous model the premiums for the policy defined in Example 4.8.3.

Example 5.3.5 (Pensions on two lives). 1. Derive Thiele's differential equation for pensions on two lives.

2. Calculate the premiums for the pensions based on the assumption that the husband and his wife die independently.
3. What happens if the the mortality rate for the state $(**)$ (or $(*\dagger) \cup (\dagger *)$) decreases (or increases) by 15 % for each life? (Empirical studies show that the mortality rate of widows and widowers is increased in comparison to the rest of the population.)

For this example we only derive Thiele's differential equations and present the results in a figure. The calculations are done in the setting of Example 4.8.5. (Δt denotes the difference in age of the husband and his wife.)

$$\begin{aligned}
\frac{\partial W_{**}^{+}(t)}{\partial t} &= -v^t\, a_{**}(t) + (\mu_{x+t}^{husband} + \mu_{x+t+\Delta t}^{wife}) W_{**}^{+}(t) \\
&\qquad - \mu_{x+t}^{husband} W_{\dagger *}^{+}(t) - \mu_{x+t+\Delta t}^{wife} W_{*\dagger}^{+}(t), \\
\frac{\partial W_{*\dagger}^{+}(t)}{\partial t} &= -v^t\, a_{*\dagger}(t) + \mu_{x+t}^{wife} \left(W_{*\dagger}^{+}(t) - W_{\dagger\dagger}^{+}(t) \right), \\
\frac{\partial W_{\dagger *}^{+}(t)}{\partial t} &= -v^t\, a_{\dagger *}(t) + \mu_{x+t+\Delta t}^{wife} \left(W_{\dagger *}^{+}(t) - W_{\dagger\dagger}^{+}(t) \right), \\
\frac{\partial W_{\dagger\dagger}^{+}(t)}{\partial t} &= 0.
\end{aligned}$$

Figure 5.1 shows the relation of the present values of the benefits for a change in mortality by ±15%. The results are what we expected: a pension on the joint lives becomes more expensive, whereas a pension on the 2nd life becomes cheaper.

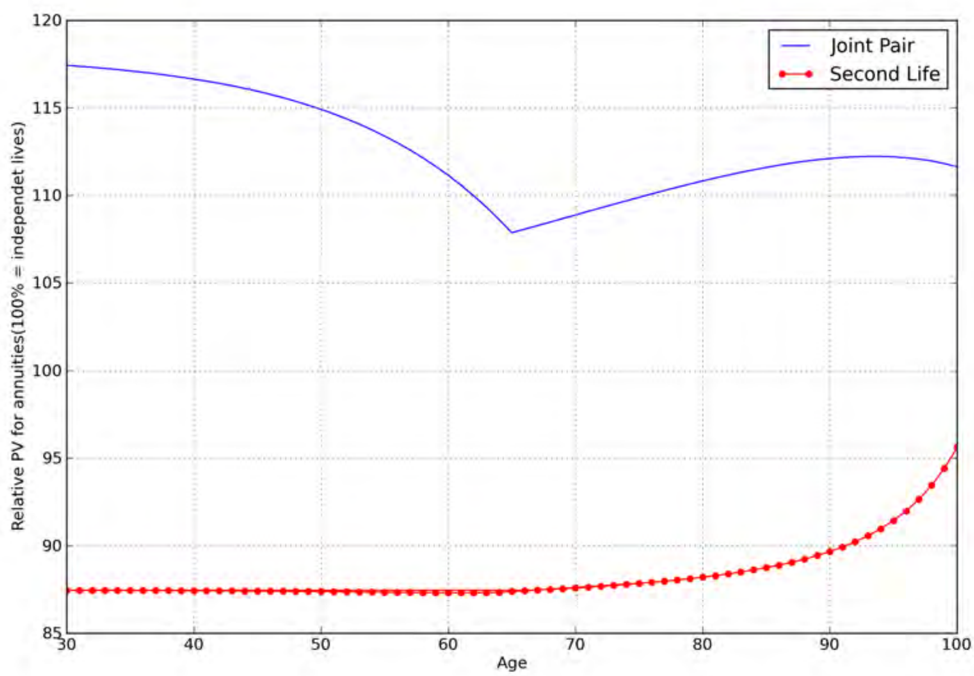

Figure 5.1. Ratio between the present values of benefits for annuities on two lives (100% = independent mortality probabilities).

Exercise 5.3.6. Complete the previous example.

Exercise 5.3.7. 1. Calculate the present value of the premiums for the insurance on two lives. (Also in the continuous setting one has to treat three cases separately.)

2. Create a model for an orphan's pension. In the model one has to consider three persons: farther, mother and child. Define the payout functions for a policy which pays 5,000 USD to the child if one of parent dies and 10,000 USD if both die. Assume that the policy matures if the child is 25 years old.

5.4 Differential equations for moments of higher order

Thiele's differential equations characterise the mathematical reserve of an insurance policy. In this section we look at the moments of the mathematical reserve. This will enable us for example to calculate the variance of the reserve, which is a measure for its variation. Furthermore, it can be used to analyse the risk structure of an insurance policy.

We start with the difference equations corresponding to the discrete model. The payout functions in the discrete Markov model have the following form:

$$\Delta B_t = \sum_{j \in J} I_j(t) a_j^{Pre}(t) + \sum_{j,k \in J} \Delta N_{jk}(t)\, v_t\, a_{jk}^{Post}(t),$$

where v_t is the yearly discount from $t+1$ to t, with $v_t = \sum_{j \in J} I_j(t) v_j(t)$.

The prospective reserves are

$$\begin{aligned} V_t^+ &= \sum_{\xi=t}^{\infty} \Big(\prod_{k=t}^{k<\xi} v_k\Big)\, \Delta B_\xi \\ &= \sum_{\xi=t}^{\infty} \Big(\prod_{k=t}^{k<\xi} v_k\Big) \left\{ \sum_{j \in J} I_j(\xi) a_j^{Pre}(\xi) + \sum_{j,l \in J} \Delta N_{jl}(\xi)\, v_\xi\, a_{jl}^{Post}(\xi) \right\}. \end{aligned}$$

Now our aim is to calculate the expectation of the p-th power of the mathematical reserve $((V_t^+)^p)$ conditioned on $\mathcal{F}_t$. The linearity of the integral yields the difference equation

$$V_t^+ = v_t \sum_{j \in J} I_j(t+1) V_{t+1}^+ + \sum_{j \in J} I_j(t) a_j^{Pre}(t) + \sum_{j,k \in J} \Delta N_{jk}(t)\, v_t\, a_{jk}^{Post}(t).$$

This formula indicates that the future reserve is composed of the payments in the period $]t, t+1]$, the payments at time $\{t\}$ and the payments in the period $]t+1, \infty[$. To keep the calculations simple we assume that there is no payment at time $\{t\}$. Furthermore we will use the notation

$$L_t = \sum_{j,k \in J} \Delta N_{jk}(t)\, v_t\, a_{jk}^{Post}(t).$$

Therefore we can simplify the recursion to

$$V_t^+ = v_t \sum_{j \in J} I_j(t+1) V_{t+1}^+ + L_t = \sum_{j \in J} I_j(t+1) \left(v_t\, V_{t+1}^+ + L_t \right).$$

Now the p-th moment is given by

$$\begin{aligned}(V_t^+)^p &= \left(\sum_{j\in J} I_j(t+1)(v_t\, V_{t+1}^+ + L_t)\right)^p \\ &= \sum_{k=0}^p \binom{p}{k}\left(\sum_{j\in J} I_j(t+1)v_t\, V_{t+1}^+\right)^k L_t^{p-k} \\ &= \sum_{k=0}^p \binom{p}{k}\sum_{j\in J} I_j(t+1)(v_t\, V_{t+1}^+)^k L_t^{p-k},\end{aligned}$$

where we used the fact that $I_\alpha(t+1)I_\beta(t+1) = \delta_{\alpha\beta}I_\alpha(t+1)$. Next, we can use $P[A\cap B|C] = P[A|B\cap C]\times P[B|C]$ to simplify the recursion for the expectation. We get

$$\begin{aligned}&E\left[(V_t^+)^p \mid X_t = i\right] \\ &\quad= E\left[\sum_{k=0}^p \binom{p}{k}(v_t^i)^k \sum_{j\in J} I_j(t+1)(V_{t+1}^+)^k L_t^{p-k} \mid X_t = i\right] \\ &\quad= \sum_{k=0}^p \binom{p}{k}(v_t^i)^k \sum_{j\in J} E\left[I_j(t+1)(V_{t+1}^+)^k L_t^{p-k} \mid X_t = i\right] \\ &\quad= \sum_{k=0}^p \binom{p}{k}(v_t^i)^k \sum_{j\in J} E\left[I_j(t+1)(V_{t+1}^+)^k \left(v_t\, a_{ij}^{Post}(t)\right)^{p-k} \mid X_t = i\right] \\ &\quad= (v_t^i)^p \sum_{j\in J} p_{ij}(t,t+1)\sum_{k=0}^p \binom{p}{k}\left(a_{ij}^{Post}(t)\right)^{p-k} E\left[(V_{t+1}^+)^k \mid X_{t+1} = j\right].\end{aligned}$$

This is the difference equation for the higher order moments of the reserve, if there are no payments in advance. Note that the integration of the payments in advance $a_j^{Pre}(t)$ is not complicated. Nevertheless, to keep the presentation clear we continue without them. We summarise our findings in the following theorem.

Theorem 5.4.1 (Differential equation for moments of higher order). *Under the above assumptions the higher order moments of the reserve satisfy the recursion*

$$\begin{aligned}&E\left[(V_t^+)^p \mid X_t = i\right] \\ &\quad= (v_t^i)^p \sum_{j\in J} p_{ij}(t,t+1)\sum_{k=0}^p \binom{p}{k}\left(a_{ij}^{Post}(t)\right)^{p-k} E\left[(V_{t+1}^+)^k \mid X_{t+1} = j\right].\end{aligned}$$

Exercise 5.4.2. Derive the above formula for a model which includes payments in advance.

After treating the discrete case we are now going to derive the analogue statements for the continuous setting. The proofs will become more involved, since the model is more general. Before stating the theorem we recall some definitions:

$$\begin{aligned}dB &= \sum_{j\in S} I_j(t)dB_j + \sum_{j\neq k} dB_{jk}, \\ dv_t &= -v_t \times \delta_t\, dt, \\ \delta_t &= \sum_{j\in S} I_j(t)\delta_j(t).\end{aligned}$$

The p-th moment of the prospective reserve is defined by

$$\begin{aligned} V_j^{(p)}(t) &:= E\left[(V_t^+)^p \mid X_t = j\right] \\ &= E\left[\left(\frac{1}{v_t}\int_t^\infty v\, dB\right)^p \mid X_t = j\right], \end{aligned}$$

where we implicitly assumed that $V_t^+ \in L^p(\Omega, \mathcal{A}, P)$ and that the functions δ, a_i and a_{jk}, μ_{jk} are piecewise continuous. Then the following theorem holds.

Theorem 5.4.3 (Differential equations for moments of higher order). *Under the above assumptions the functions $V_j^{(p)}(t)$ satisfy the differential equation*

$$\begin{aligned} \frac{\partial}{\partial t} V_j^{(p)}(t) &= \Big(p\,\delta_j(t) + \sum_{S \ni k \neq j} \mu_{jk}(t)\Big) V_j^{(p)}(t) - p\, a_j(t) V_j^{(p-1)}(t) \\ &- \sum_{j \neq k \in S} \mu_{jk}(t) \sum_{k=0}^{p} \binom{p}{k} (a_{jk}(t))^{p-k}\, V_j^{(k)}(t) \end{aligned}$$

for all $t \in]0, n[\setminus \mathcal{D}$ with the boundary condition

$$V_j^{(p)}(t^-) = \sum_{k=0}^{p} \binom{p}{k} (\Delta a_j(t))^{p-k}\, V_j^{(k)}(t)$$

for all $t \in \mathcal{D}$. Here $\mathcal{D}$ is the set of discontinuities of the payout function B.

Remark 5.4.4. – The differential equation given above also holds at points where the functions $V_j^{(p)}(t)$ are not differentiable. In these points one gets a valid interpretation by considering the differentials which are given by a formal multiplication with the factor dt.

– The idea of the proof is to represent a suitable martingale in two different ways. These representations will then be used to find a stochastic differential equation for the martingale. Then, since the drift term of the differential equation is zero for a martingale, one obtains an ordinary differential equation.

Proof. It turns out to be more convenient to show the differential equation for $W_j^{(p)}(t) = v_t^p\, V_j^{(p)}(t)$. We have

$$\begin{aligned} dW_j^{(p)}(t) &= d(v_t^p)\, V_j^{(p)}(t) + v_t^p\, dV_j^{(p)}(t) && (5.1) \\ &= -p v_t^p \sum_{j \in S} I_j(t) \delta_j(t) dt\, V_j^{(p)}(t) + v_t^p\, dV_j^{(p)}(t). && (5.2) \end{aligned}$$

Now we define the martingale

$$\begin{aligned} M^{(p)}(t) &:= E\left[\left(\int_0^\infty v\, dB\right)^p \mid \mathcal{F}_t\right] \\ &= E\left[\left\{\left(\int_0^t + \int_t^\infty\right) v\, dB\right\}^p \mid \mathcal{F}_t\right] \end{aligned}$$

and the function

$$U_t = \int_0^t v\, dB.$$

The Markov property implies

$$E\left[\left\{\int_t^\infty v\, dB\right\}^{p-k} \mid \mathcal{F}_t\right] = \sum_{j \in S} I_j(t) W_j^{(p-k)}(t),$$

and using the Binomial Theorem we get

$$M^{(p)}(t) = \sum_{k=0}^{p} \binom{p}{k} \sum_{j \in S} U_t^k I_j(t) W_j^{(p-k)}(t).$$

By choosing a right continuous modification of $M^{(p)}$ we can ensure that U and $I_j(t)$ are right continuous. Now we want to simplify the differential form

$$dM^{(p)}(t) = \sum_{k=0}^{p} \binom{p}{k} \sum_{j \in S} d\left(U_t^k I_j(t) W_j^{(p-k)}(t)\right).$$

Recall that for a function of bounded variation A we denote by A^{cont} the continuous part and by A^{atom} the discontinuous part. An application of Itô's formula yields

$$\begin{aligned} & d\left(U_t^{(k)} I_j(t) W_j^{(p-k)}(t)\right) \\ & \quad = k U_t^{(k-1)} dU_t^{cont} I_j(t) W_j^{(p-k)}(t) \\ & \quad + U_t^k dI_j^{cont}(t) W_j^{(p-k)}(t) + U_t^k I_j(t) dW_j^{(p-k),cont}(t) \\ & \quad + \left\{U_t^k I_j(t) W_j^{(p-k)}(t) - U_{t-}^k I_j(t^-) W_j^{(p-k)}(t^-)\right\}. \end{aligned} \tag{5.3}$$

To simplify this formula we use for the first line the identities

$$\begin{aligned} dU_t^{cont} &= v_t \sum_{l \in S} I_l(t) a_l(t), \\ I_\alpha(t) I_\beta(t) &= \delta_{\alpha\beta} I_\alpha(t). \end{aligned}$$

For the second line of (5.3) note that the continuous part of $I_\alpha(t)$ vanishes. Finally we have to deal with the jump part in the third line. The jumps have two possible origins. On the one hand they might be caused by a transition:

$$\sum_{j\neq l\in S}\left(\{U_{t^-}+v_t a_{jl}(t)\}^k\, W_l^{(p-k)}(t)-U_{t^-}^k W_l^{(p-k)}(t^-)\right)dN_{jl}(t).$$

On the other hand they might be due to a jump in a pension:

$$\sum_{l\in S} I_l(t)\left(\{U_{t^-}+v_t \Delta a_l(t)\}^k\, W_l^{(p-k)}(t)-U_{t^-}^k W_l^{(p-k)}(t^-)\right).$$

Moreover, we know that jumps can only occur on the set $\mathcal{D}$ and on this set one can replace $I_l(t)$ by $I_l(t^-)$, since these coincide with probability 1. We also know that a simultaneous jump of both components occurs with probability 0. Finally, $W_l^{(p-k)}(t)$ is continuous and does not induce any jumps, since $\int_t^n v\, dB$ is almost surely continuous on $t\notin\mathcal{D}$. Therefore we get

$$\begin{aligned}
d&\left(U_t^{(k)} I_j(t) W_j^{(p-k)}(t)\right)\\
&=\sum_{l\in S} I_l(t)\left(kU_t^{(k-1)} v_t a_j(t) W_j^{(p-k)}(t)+U_t^k dW_j^{(p-k),cont}(t)\right)\\
&+\sum_{j\neq l\in S}\left(\{U_{t^-}+v_t a_{jl}(t)\}^k\, W_l^{(p-k)}(t)-U_{t^-}^k W_l^{(p-k)}(t^-)\right)dN_{jl}(t)\\
&+\sum_{l\in S} I_l(t^-)\left(\{U_{t^-}+v_t \Delta a_l(t)\}^k\, W_l^{(p-k)}(t)-U_{t^-}^k W_l^{(p-k)}(t^-)\right).
\end{aligned}$$

Applying the fact $X_{t^-}dt=X_t\,dt$ and the previous formula to (5.3) we derive

$$\begin{aligned}
dM^{(p)}(t)&-\sum_{j\neq l\in S}\sum_{k=0}^{p}\binom{p}{k}\Big(\{U_{t^-}+v_t a_{jl}(t)\}^k\, W_l^{(p-k)}(t)\\
&\qquad\qquad -U_{t^-}^k W_l^{(p-k)}(t^-)\Big)dM_{jl}(t)\\
&=\sum_{j\in S} I_j(t)\sum_{k=0}^{p}\binom{p}{k}\Bigg[kU_t^{(k-1)} v_t a_j(t) W_j^{(p-k)}(t)dt+U_t^k dW_j^{(p-k),cont}(t)\\
&+\sum_{j\neq l\in S}\left(\{U_t+v_t a_{jl}(t)\}^k\, W_l^{(p-k)}(t)-U_t^k W_j^{(p-k)}(t)\right)\mu_{jl}(t)dt\Bigg]\\
&+\sum_{j\in S} I_l(t^-)\sum_{k=0}^{p}\binom{p}{k}\left(\{U_{t^-}+v_t \Delta a_j(t)\}^k\, W_j^{(p-k)}(t)-U_{t^-}^k W_j^{(p-k)}(t^-)\right),
\end{aligned}\tag{5.4}$$

where we used the identity

$$dM_{ij}(t) \;=\; dN_{ij}(t)-I_i(t)\,\mu_{ij}(t).$$

Note that

$$dW_j^{(p-k),cont}(t) \;=\; -(p-k)v_t^{(p-k)}\delta_j(t)\,dt\,V_j^{(p-k)}(t)+v_t^{(p-k)}dV_j^{(p-k),cont}(t). \tag{5.5}$$

The left hand side of (5.4) is the differential of a sum of martingales. Thus also the right hand side is the differential of a martingale. Now this has to be constant, since it is previsible and of bounded variation. Therefore the increments of the continuous and the discrete part have to be equal to zero. But this is only possible if

$$
\begin{aligned}
0 \quad = \quad & \sum_{k=1}^{p} \binom{p}{k} k U_t^{(k-1)} v_t a_j(t) W_j^{(p-k)}(t) dt \\
& + \sum_{k=0}^{p} \binom{p}{k} U_t^{k} W_l^{(p-k)}(t) \\
& + \sum_{k=0}^{p} \binom{p}{k} \sum_{j \neq l \in S} \sum_{r=0}^{k} \binom{k}{r} U_t^{(r)} \left\{ v_t a_{jl}(t) \right\}^{k-r} W_l^{(p-k)}(t) \mu_{jl}(t) dt \\
& - \sum_{k=0}^{p} \binom{p}{k} U_t^{k} W_j^{(p-k)}(t) \left(\sum_{j \neq l \in S} \mu_{jl}(t) \right) dt
\end{aligned} \tag{5.6}
$$

holds for all $j \in S$ and all $t \in]0, n[\setminus \mathcal{D}$. For $x \in \mathcal{D}$ we get

$$
0 = \sum_{k=0}^{p} \binom{p}{k} \left(\sum_{r=0}^{k} \binom{k}{r} U_{t^-}^{(r)} \left\{ v_t \Delta a_l(t) \right\}^{k-r} W_j^{(p-k)}(t) - U_{t^-}^{(k)} W_j^{(p-k)}(t^-) \right).
$$

Using the identity

$$
\binom{p}{k} k = \binom{p}{k-1} (p-(k-1))
$$

we can transform the first line of Equation (5.6) into

$$
\begin{aligned}
& \sum_{k=1}^{p} \binom{p}{k-1} (p-(k-1)) U_t^{(k-1)} v_t a_j(t) W_j^{(p-1-(k-1))}(t) dt \\
& \qquad = \sum_{k=0}^{p} \binom{p}{k} (p-k) U_t^{(k)} v_t a_j(t) W_j^{(p-1-k)}(t) dt,
\end{aligned}
$$

where we set $W_j^{(-1)} \equiv 0$. Hence the third line of (5.6) becomes

$$
\sum_{k=0}^{p} U_t^{(r)} \sum_{r=0}^{k} \binom{p}{k} \binom{k}{r} \sum_{j \neq l \in S} \left\{ v_t a_{jl}(t) \right\}^{k-r} W_l^{(p-k)}(t) \mu_{jl}(t) dt.
$$

This can be transformed by the identity

$$
\binom{p}{k} \binom{k}{r} = \binom{p}{r} \binom{p-r}{k-r}
$$

into the representation

$$
\sum_{r=0}^{p} \binom{p}{r} U_t^{(r)} \sum_{k=r}^{p} \binom{p-r}{k-r} \sum_{j \neq l \in S} \left\{ v_t a_{jl}(t) \right\}^{r} W_l^{(p-k-r)}(t) \mu_{jl}(t) dt,
$$

i.e.,

$$\sum_{k=0}^{p} \binom{p}{k} U_t^{(k)} \sum_{r=0}^{p-k} \binom{p-k}{r} \sum_{j \neq l \in S} \{v_t a_{jl}(t)\}^r \, W_l^{(p-k-r)}(t) \mu_{jl}(t) dt.$$

If we now gather the powers of U_t we get

$$0 = \sum_{k=0}^{p} \binom{p}{k} U_t^{(k)} dQ_j^{(p-k)}(t),$$

where

$$\begin{aligned} dQ_j^{(q)}(t) &= q v_t a_j(t) W_j^{(q-1)}(t) dt + dW_j^{(q),cont}(t) \\ &\quad + \sum_{k=0}^{q} \binom{q}{k} \sum_{j \neq l \in S} \{v_t a_{jl}(t)\}^k \, W_l^{(q-k)}(t) \mu_{jl}(t) dt \\ &\quad - W_j^{(q)}(t) \sum_{j \neq l \in S} \mu_{jl}(t) \, dt. \end{aligned}$$

This equation implies $dQ_j^{(0)}(t) \equiv 0$ and thus $dQ_j^{(q)}(t) \equiv 0$ by induction. Finally, the formula above and Equation (5.1) and (5.5) imply the result.

5.5 The distribution of the mathematical reserve

The distribution function can be used to answer questions which only depend on the tail of the distribution. Thus it is important for the estimation of extreme risks.

This section has the same structure as the previous section. In the beginning, we solve the problem for the discrete time set. Afterwards we treat the continuous time model.

Recall that the cash flows in the discrete Markov model are given by

$$\Delta B_t = \sum_{j \in J} I_j(t) a_j^{Pre}(t) + \sum_{j,k \in J} \Delta N_{jk}(t) \, v_t \, a_{jk}^{Post}(t).$$

We want to calculate the distribution function of the discounted future cash flows

$$P_i(t,u) = P\left[\sum_{j=t}^{\infty} \left(\prod_{k=t}^{k<j} v_k\right) \Delta B_j < u | X_t = i \right].$$

The following identities hold:

$$\begin{aligned} P_i(t,u) &= P\left[\sum_{j=t}^{\infty} D_{t,j} \, \Delta B_j < u | X_t = i \right] \\ &= \sum_{l \in J} p_{il}(t) \, P\left[\sum_{j=t}^{\infty} D_{t,j} \, \Delta B_j < u | X_t = i, X_{t+1} = l \right] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{l \in J} p_{il}(t) P\Bigg[v_{i,t} \sum_{j=t+1}^{\infty} D_{t,j}\, \Delta B_j < u - a_i^{Pre}(t) \\
&\qquad\qquad - v_{i,t}\, a_{il}^{Post}(t) | X_{t+1} = l \Bigg] \\
&= \sum_{l \in J} p_{il}(t)\, P_l(t+1, v_{i,t}^{-1}(u - a_i^{Pre}(t)) - a_{il}^{Post}(t)), \text{ where} \\
D_{t,j} &= \prod_{k=t}^{k<j} v_k.
\end{aligned}
$$

These relations are summarised in the following theorem.

Theorem 5.5.1 (Distribution of the reserves). *The distribution function of the reserves satisfy the recursion*

$$
P_i(t,u) = \sum_{j \in J} p_{ij}(t)\, P_j(t+1, v_{i,t}^{-1}(u - a_i^{Pre}(t)) - a_{ij}^{Post}(t)).
$$

Besides the recursion formula also boundary conditions are required. These are, in contrast to the previous problems, now given in form of distributions rather than fixed values. For an insurance whose mathematical reserve is equal to zero at maturity the boundary condition is for example given by: (Here ω denotes the maximal age at which insured persons are alive.)

$$
P_i(\omega+1, u) = \begin{cases} 0, & \text{if } u \leq 0, \\ 1, & \text{if } u > 0. \end{cases}
$$

The distribution function of a pension in the discrete model is shown in Figure 5.2. The jumps, which are caused by the discrete model, are clearly visible.

In the previous sections we have seen that one can use differential equations to find the moments of the reserves. For cash flows of sufficient regularity one could also prove that the moments are differentiable.

Analogous one could try to find differential equations for the distribution functions. But the following example shows, that for distribution functions this is not an easy task.

Example 5.5.2. This example will illustrate that the distribution function can be discontinuous even for relatively simple insurance policies. We consider an endowment policy, with a death benefit of 100,000 USD and an endowment of 200,000 USD. Now we want to calculate the reserve for an insured of 30 years of age. Here T_x will denote the future life span. The following equations hold:

$$
V_{30} = \begin{cases} 100000 \times v^{T_x}, & \text{if } \; T_x < 35, \\ 200000 \times v^{65-30}, & \text{if } \; T_x \geq 35. \end{cases}
$$

Now assume a technical interest rate of 1.5 %. Then the following equation holds for $0 < \alpha \leq 100000$:

$$
\begin{aligned}
P[V_{30} < \alpha] &= P[100000 v^{T_x} < \alpha,\, T_x < 35] \\
&= P[35 > T_X > \log(\alpha/100000)/\log(v)] \\
&= {}_{\gamma}p_x - {}_{35}p_x, \text{where} \\
\gamma &= \log(\alpha/100000)/\log(v).
\end{aligned}
$$

This calculation shows that the distribution function has a jump of size ${}_{35}p_x$ at $200000 v^{35}$, and thus it is discontinuous.

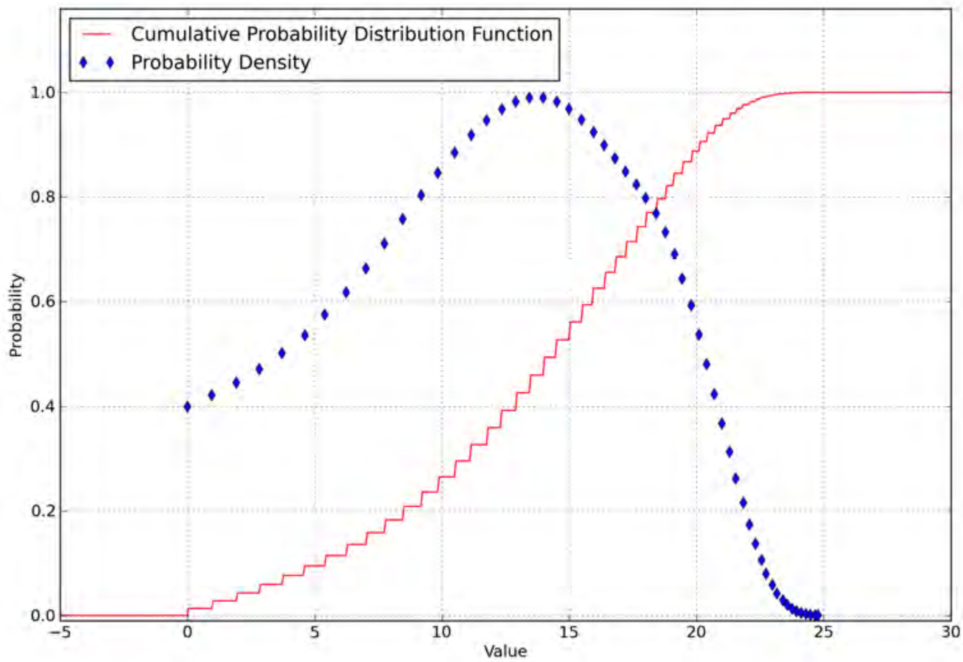

Figure 5.2. Probability distribution function for the present value of an immediate payout annuity (x=65)

The next theorem shows that the distribution functions satisfy an integral equation. Note that the recursion in discrete time provides an approximation to this integral equation.

Theorem 5.5.3. *The conditional distribution functions of the reserves*

$$P_j(t,u) = P\left[\int_t^\infty \exp\left(-\int_t^\xi \delta_\tau d\tau\right) dB(\xi) \le u | X_t = j\right]$$

satisfy the integral equation

$$\begin{aligned}P_j(t,u) = &\sum_{k\neq j}\int_t^\infty \exp\left(-\int_t^\xi \sum_{l\neq j}\mu_{jl}(\tau)d\tau\right)\mu_{jk}(\xi)\\ &\times P_k\left(\xi, \exp(\delta_j(\xi-t))u - \int_t^\xi \exp(\delta_j(\xi-\tau))dB_j(\tau) - a_{jk}(s)\right)d\xi\\ &+\exp\left(-\int_t^\infty \sum_{l\neq j}\mu_{jl}(\tau)d\tau\right)\chi_{\left[\int_t^n \exp(-\delta_j(\tau-t))dB_j(\tau)\le u\right]}.\end{aligned} \tag{5.7}$$

Proof. The proof is analogous to the discrete setting. One considers

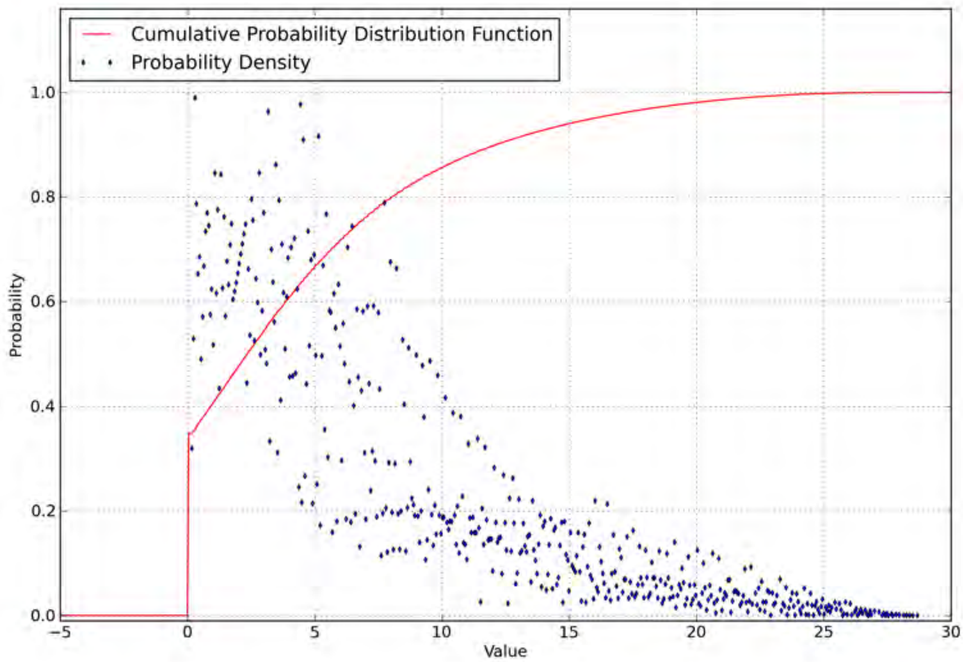

Figure 5.3. Probability distribution function of the present value of a deferred widows pension (x=65)

$$A = \left\{ \int_t^\infty \exp\left(- \int_t^\xi \delta \right) dB(\xi) \leq u \right\}$$

and treats, as in the discrete setting, the various cases separately. We leave the proof to the reader and refer to [HN96].

Exercise 5.5.4. Complete the proof of the previous theorem. [HN96]

After deriving the integral equations for the distribution function, we will modify these slightly. The equations are still hard to handle, since the right hand side depends on t. To overcome this problem we define

$$Q_j(t,u) := P_j\left(t, \exp(\delta_j t)\left(u - \int_0^t \exp(-\delta_j\, \tau) dB_j(\tau)\right)\right).$$

The mapping from P to Q can be inverted by

$$P_j(t,u) = Q_j\left(t, \exp(-\delta_j t)u + \int_0^t \exp(-\delta_j\, \tau) dB_j(\tau)\right).$$

Using Q_j one can easily derive the equation

$$
\begin{aligned}
&\exp\left(-\int_0^t \mu_j\right) Q_j(t,u) = \int_t^n \exp\left(-\int_0^s \mu_j\right) \sum_{k\neq j} \mu_{jk}(s) \\
&\quad \times Q_k\Bigg(s, \exp((\delta_j-\delta_k)s)u + \int_0^s \exp(-\delta_k\tau)dB_k(\tau) \\
&\quad - \exp((\delta_j-\delta_k)s)\int_0^s \exp(-\delta_j\tau)dB_j(\tau) - \exp(-\delta_k s)a_{jk}(s)\Bigg)ds \\
&\quad + \exp\left(-\int_0^n \mu_j\right)\chi_{\left[\int_0^n \exp(-\delta_j(\tau))dB_j(\tau)\leq u\right]}.
\end{aligned}
$$

Theorem 5.5.5. *The functions Q satisfy (in the sense of Stieltje's differentials) the following differential equations*

$$
\begin{aligned}
d_t Q_j(t,u) &= \mu_j\, dt\, Q_j(t,u) - \sum_{k\neq j} \mu_{jk}\, dt \\
&\times\ Q_k(t, \exp((\delta_j-\delta_k)t)\, u + \int_0^t \exp(-\delta_k\, \tau)dB_k(\tau) \\
&-\ \exp((\delta_j-\delta_k)t) \int_0^t \exp(-\delta_j\, \tau)dB_j(\tau) - \exp(-\delta_k\, t)\, a_{jk}(t)),
\end{aligned}
$$

with the boundary conditions

$$
Q_j(n,u) = \chi_{\left[\int_0^n \exp(-\delta_j\, \tau)dB_j(\tau)\leq u\right]}.
$$

Remark 5.5.6. Theorem 5.5.5 proves useful, since the equations given therein are easier to solve by numerical methods. The main idea is to derive Q in a first step and then calculate P based on Q.

6. Abstract Valuation

6.1 Framework: Valuation Portfolios

Definition 6.1.1 (Stochastic Cash Flows). *A* Stochastic Cash Flow *is a sequence* $x = (x_k)_{k\in\mathbb{N}} \in L^2(\Omega, \mathcal{A}, P)^{\mathbb{N}}$*, which is* $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ *adapted.*

Definition 6.1.2 (Regular Stochastic Cash Flows). *A* Regular Stochastic Cash Flow x with respect to $(\alpha_k)_{k\in\mathbb{N}}$*, with* $\alpha_k > 0 \forall k$ *is a stochastic cash flow such that*

$$Y := \sum_{k\in\mathbb{N}} \alpha_k X_k \in L^2(\Omega, \mathcal{A}, P).$$

We denote the vector space of all regular cash flows by $\mathcal{X}$*.*

Remark 6.1.3.
1. We note that for all $n \in \mathbb{N}_0$ the image of

$$\psi : L^2(\Omega, \mathcal{A}, P)^n \to \mathcal{X}, (x_k)_{k=0,\dots n} \mapsto (x_0, x_1, \dots, x_n, 0, 0 \dots)$$

   is a sub-space of $\mathcal{X}$.
2. $\mathcal{X}$ has been defined this way in order to capture cash flow streams where the sum of the cash flows is infinite with a finite present value. In this set up α_k can be interpreted as a majorant of the price of the payment 1 at time k.

Theorem 6.1.4.
1. *For* $x, y \in \mathcal{X}$*, we define the scalar product as follows:*

$$\begin{aligned} < x, y > &= \sum_{k\in\mathbb{N}} < \alpha_k\, x_k, \alpha_k\, y_k > \\ &= E[\sum_{k\in\mathbb{N}} \alpha_k^2\, x_k\, y_k], \end{aligned}$$

   and remark that the scalar product exists as a consequence of the Cauchy-Schwartz inequality .
2. $\mathcal{X}$ *equipped with the above defined scalar product is a Hilbert space with norm* $||x|| = \sqrt{< x, x >}$*.*

Proof. We leave the proof of this proposition to the reader.

In a next step we introduce the concept of a positive valuation functional and we closely follow [Büh95].

Definition 6.1.5 (Positivity).
1. $x = (x_k)_{k\in\mathbb{N}} \in \mathcal{X}$ *is called* positive *if* $x_k > 0$ *P-a.e. for all* $k \in \mathbb{N}$*. In this case we write* $x \geq 0$*.*

2. $x = (x_k)_{k \in \mathbb{N}} \in \mathcal{X}$ *is called* strictly positive *if* $x_k > 0$ *P-a.e. for all* $k \in \mathbb{N}$ *and there exists a* $k \in \mathbb{N}$*, such that* $x_k > 0$ *with a positive probability. In this case we write* $x > 0$.

Definition 6.1.6 (Positive functionals). $Q : \mathcal{X} \to \mathbb{R}$ *is called a* positive, continuous and linear functional *if the following hold true:*

1. *If* $x > 0$*, we have* $Q[x] > 0$.
2. *If* $x = \lim_{n\to\infty} x_n$*, for* $x_n \in \mathcal{X}$ *we have* $Q[x] = \lim_{n\to\infty} Q[x_n]$.
3. *For* $x, y \in \mathcal{X}$ *and* $\alpha, \beta \in \mathbb{R}$ *we have* $Q[\alpha x + \beta y] = \alpha Q[x] + \beta Q[y]$.

Theorem 6.1.7 (Riesz representation theorem). *For* Q *a positive, linear functional as defined before, there exists* $\phi \in \mathcal{X}$*, such that*

$$Q[y] \quad = \quad <\phi, y> \forall y \in \mathcal{X}.$$

Proof. This is a direct consequence of Riesz representation theorem of continuous linear functionals of Hilbert spaces.

Definition 6.1.8 (Deflator). *The* $\phi \in \mathcal{X}$ *generating* $Q[\bullet]$ *is called* deflator.

Theorem 6.1.9. *For a positive functional* $Q : \mathcal{X} \to \mathbb{R}$*, with deflator* $\psi \in \mathcal{X}$ *we have the following:*

1. $\phi_k > 0$ *for all* $k \in \mathbb{N}$.
2. ϕ *is unique.*

Proof. 1. Assume $\phi_k = 0$. In this case we have $Q[(\delta_{kn})_{n\in\mathbb{N}}] = 0$ which is a contradiction.

2. Assume $Q[y] = <\phi, y> = <\phi^*, y>$ for all $y \in \mathcal{X}$. In this case we have $<\phi - \phi^*, y> = 0$, in particular for $y = \phi - \phi^*$. Hence we have $||\phi - \phi^*|| = 0$.

Definition 6.1.10 (Projections). *For* $k \in \mathbb{N}$ *we define the following projections:*

1. $p_k : \mathcal{X} \to L^2(\Omega, \mathcal{A}, P), x = (x_n)_{n\in\mathbb{N}} \mapsto (\delta_{kn}\, x_n)_{n\in\mathbb{N}}$*, the projection on the k-th coordinate.*
2. $p_k^+ : \mathcal{X} \to L^2(\Omega, \mathcal{A}, P), x = (x_n)_{n\in\mathbb{N}} \mapsto (\chi_{k\le n}\, x_n)_{n\in\mathbb{N}}$*, the projection starting on the k-th coordinate.*

Definition 6.1.11 (Valuation at time t, pricing functionals). *For* $t \in \mathbb{N}$ *we define the valuation of* $x \in \mathcal{X}$ *at time* t *by*

$$Q_t[x] \quad = \quad Q[x|\mathcal{F}_t] = \frac{1}{\phi_t} E[\sum_{k=0}^{\infty} \phi_k\, x_k | \mathcal{F}_t]$$

In the same sense as for mathematical reserves we define the value of the future cash flows at time t *by*

$$Q_t^+[x] \quad = \quad Q[p_t^+(x)].$$

The operators Q_t *and* Q_t^+ *are called pricing functionals.*

Definition 6.1.12 (Zero Coupon Bonds). *The Zero Coupon Bond* $\mathcal{Z}_{(k)} = (\delta_{kn})_{n\in\mathbb{N}}$ *is an element of* $\mathcal{X}$*. We remark that*

$$\pi_0\left(\mathcal{Z}_{(t)}\right) \quad = \quad Q[\mathcal{Z}_{(t)}] = E[\phi_t].$$

Theorem 6.1.13. *The cash flow $x = (x_k)_{k\in\mathbb{N}}$ in the discrete Markov model (cf. proposition 4.7.3) on a finite time interval $T \subset \mathbb{N}$ is given by:*

$$x_k = \sum_{(i,j)\in S^2} \Delta N_{ij}(k-1) a_{ij}^{Post}(k-1) + \sum_{i\in S} I_i(k) a_i^{Pre}(k),$$

where we assume that $\Delta N_{ij}(-1) = 0$.

Proof. The form of the cash flow follows from the calculations of the earlier chapters. It remains to show that $(x_k)_{k\in\mathbb{N}}$ is in L^2. This is however easy, since the benefit functions and the state space are finite. Given the fact that also the time considered for a life insurance is finite, the required property follows.

Theorem 6.1.14. *For $x \in \mathcal{X}$, as defined above we have the following:*

1. $E[\Delta N_{ij}(s)|X_t = k] = p_{ki}(t,s)p_{ij}(s,s+1)$,
2. $E[I_i(s)|X_t = k] = p_{ki}(t,s)$,
3. $E[x_s|X_t = k] =$

$$\sum_{(i,j)\in S^2} p_{ki}(t,s-1)p_{ij}(s-1,s)a_{ij}^{Post}(s-1) + \sum_{i\in S} p_{ki}(t,s)a_i^{Pre}(s),$$

and we assume that $p_{ki}(t,s-1) = 0$ if $t = s$.

Proof. We leave the proof of this proposition to the reader as an exercise.

Definition 6.1.15. *The abstract vector space of financial instruments we denote by $\mathcal{Y}$. Elements of this vector space are for example all zero coupon bonds, shares, options on shares etc.*

Remark 6.1.16. – Link to the arbitrage free pricing theory: If we assume that Q does not allow arbitrage we are in the set up of chapter 7. In proposition 7.2.15 we have seen that $\pi(X) = E^Q[\beta_T X]$, where β_T denotes the risk free discount rate. In the context of the above, we would have $\pi_0(x) = Q[x] = E^P[\phi_T x]$. Hence we can identify $\phi_T = \frac{dQ}{dP}\beta_T$. In consequence we can interpret a deflator as a discounted Radon-Nikodym density with respect to the two measures P and Q.

– In the same sense the concepts of definition 6.1.11 have a lot in common with the definition of the present values of a cash flow stream as defined in chapter 4.6.

$$\begin{aligned} B_t(s) &= \pi_t\left(\mathcal{Z}_{(s)}\right) \\ &= E\left[\exp(-\int_t^s r_u du) \times \exp(-\int_t^s \lambda^T dW - \frac{1}{2}\int_t^s \lambda^T \lambda\, du) \mid \mathcal{G}_t\right], \end{aligned}$$

and hence we have actually calculated the corresponding deflators as follows:

$$\phi_t(s) = \exp(-\int_t^s r_u du) \times \exp(-\int_t^s \lambda^T dW - \frac{1}{2}\int_t^s \lambda^T \lambda\, du).$$

Theorem 6.1.17. *Let Q be a positive, continuous functional $Q : \mathcal{X} \to \mathbb{R}$, and assume $Q[\bullet] =< \phi, \bullet >$, with $\phi = (\phi_t)_{t\in\mathbb{N}}$ $\mathbb{F}$ – adapted. In this case $(\phi_t\, Q_t[x])_{t\in\mathbb{N}}$ is an $\mathbb{F}$-martingale over P.*

Proof. Since $\mathcal{F}_t \subset \mathcal{F}_{t+1}$ and the projection property of the conditional expectation we have

$$\begin{aligned} E^P[\phi_{t+1}\, Q_{t+1}[x] | \mathcal{F}_t] &= E^P[E^P[\sum_{k\in\mathbb{N}} \phi_k x_k\, Q_{t+1}[x] | \mathcal{F}_{t+1}] | \mathcal{F}_t] \\ &= E^P[\sum_{k\in\mathbb{N}} \phi_k x_k\, Q_{t+1}[x] | \mathcal{F}_t] \\ &= \phi_t\, Q_t[x]. \end{aligned}$$

Example 6.1.18 (Replicating Portfolio Mortality). In this first example we consider a term insurance, for a 50 year old man with a term of 10 years, and we assume that this policy is financed with a regular premium payment. Hence there are actually two different payment streams, namely the premium payment stream and the benefits payment stream. For sake of simplicity we assume that the yearly mortality is $(1 + \frac{x-50}{10} \times 0.1)\%$. We assume that the death benefit amounts to 100,000 USD and we assume that the premium has been determined with an interest rate $i = 2\%$. In this case the premium amounts to $P = 1394.29$. The replicating portfolio in the sense of expected cash flows at inception is therefore given as follows (cf proposition 6.1.14). We remark that the units have been valued with two (flat) yield curves with interest rates of 2% and 4% respectively, and remark the the use of arbitrary yield curves does not imply additional complexity.

Age	Unit	Units for Mortality	Units for Premium	Total Units	Value $i = 2\%$	Value $i = 4\%$
50	$\mathcal{Z}_{(0)}$	–	-1394.28	-1394.28	-1394.28	-1394.28
51	$\mathcal{Z}_{(1)}$	1000.00	-1380.34	-380.34	-372.88	-365.71
52	$\mathcal{Z}_{(2)}$	1089.00	-1365.16	-276.16	-265.43	-255.32
53	$\mathcal{Z}_{(3)}$	1174.93	-1348.77	-173.84	-163.81	-154.54
54	$\mathcal{Z}_{(4)}$	1257.56	-1331.24	-73.67	-68.06	-62.97
55	$\mathcal{Z}_{(5)}$	1336.69	-1312.60	24.09	21.82	19.80
56	$\mathcal{Z}_{(6)}$	1412.12	-1292.91	119.20	105.85	94.21
57	$\mathcal{Z}_{(7)}$	1483.67	-1272.23	211.44	184.07	160.67
58	$\mathcal{Z}_{(8)}$	1551.18	-1250.60	300.57	256.54	219.62
59	$\mathcal{Z}_{(9)}$	1614.50	-1228.09	386.41	323.33	271.48
60	$\mathcal{Z}_{(10)}$	1673.52	–	1673.52	1372.87	1130.57
Total					**0.00**	**-336.47**

Exercise 6.1.19 (Replicating Portfolio Disability). Consider a disability cover and calculate the replicating portfolios for a deferred disability annuity and a disability in payment.

6.2 Cost of Capital

In section 6.1 we have seen how to abstractly valuate $x \in \mathcal{X}$ by means of a pricing functional Q . For some financial instruments $y \in \mathcal{Y}^\star$ we can directly observe $Q[y]$ such as for a lot of zero coupons bonds $\mathcal{Z}_{(\bullet)}$. On the other hand this is not always possible.

Definition 6.2.1. *We denote by $\mathcal{Y}^\star$ the set of all stochastic cash flows in $x \in \mathcal{Y}$ such that $Q[x]$ is observable. With $\tilde{\mathcal{Y}} = span < \mathcal{Y}^\star >$ we denote the vector space generated by $\mathcal{Y}^\star$ and we define:*

1. *$x \in \mathcal{Y}^\star$ is called of level 1.*

2. *$x \in \tilde{\mathcal{Y}}$ is called of level 2.*
3. *$x \in \mathcal{Y} \setminus \tilde{\mathcal{Y}}$ is called of level 3.*

Remark 6.2.2. It is clear that the model uncertainty and the difficulties to value assets or liabilities increases from level 1 to level 3. Since we are interested in market values only the valuation of level 1 assets and liabilities are really reliable. For level 2 assets and liabilities on has to find a sequence of $x_n = \sum_{k=1}^{n} \alpha_k\, e_k$ with $e_k \in \mathcal{Y}^\star$ such that $x = \lim_{n\to\infty} x_n$. Since we assume that Q is linear and continuous we can calculate

$$\begin{aligned} Q[x] &= \lim_{n\to\infty} Q[x_n] \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} Q[\alpha_k\, e_k] \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \alpha_k\, Q[e_k]. \end{aligned}$$

For level 3 assets and liabilities the situation is even more difficult, since there is no obvious way to do it. The best, which we can be done is to define $\tilde{Q}[x]$ such that $\tilde{Q}[x] = Q[x] \forall x \in \mathcal{Y}$ and hope that $\tilde{Q}[x] \approx Q[x]$ for the $x \in \mathcal{Y}$ we want to valuate. In most cases such $\tilde{Q}[\bullet]$ are based on first economic principles. In the following we want to see how the *Cost of Capital* concept works for insurance liabilities and how we can concretely implement it.

Definition 6.2.3 (Utility Assumption). *If we have $x, y \in L^2(\Omega, \mathcal{A}, P)^+$, with $x = E[y]$. A rational investor would normally prefer x, since there is less uncertainty. The way to understand this, is by using utility functions. For $x \in L^2(\Omega, \mathcal{A}, P)^+$ and u a concave function, the utility of x is defined as $E[u(x)]$. The idea behind utilities is that the first 10,000 USD are higher valued than the one 10,000 USD from 100,000 USD to 110,000 USD. Hence the increase of utility per fixed amount decreases if amounts increase. As a consequence of the Jensen's inequality, we see that the utility of a constant amount is higher than the utility of a random payout with the same expected value.*

Definition 6.2.4. *Let $x = (x_k)_{k\in\mathbb{N}} \in \mathcal{X}$ be an insurance cash flow, for example generated by a Markov model.*

1. *In this case we define the expected cash flows by*

$$CF(x) = (E[x_k])_{k\in\mathbb{N}}.$$

2. *The corresponding portfolio of financial instruments in the vector space $\mathcal{Y}$ we define by*

$$VaPo^{CF}(x) = \sum_{k\in\mathbb{N}} CF(x)_k \mathcal{Z}_{(k)} \in \mathcal{Y}$$

3. *By $R(x)$ we denote the residual risk portfolio given by*

$$\begin{aligned} R(x) &= x - VaPo^{CF}(x) \\ &= \sum_{k\in\mathbb{N}} (x_k - CF(x)_k)\, \mathcal{Z}_{(k)} \in \mathcal{Y} \end{aligned}$$

4. *For a given $x \in \mathcal{X}$ we denote by $VaPo^*(x)$ an approximation $y \in \tilde{\mathcal{Y}}$ of x, such that $||x - VaPo^*(x)|| \leq ||x - VaPo^{CF}(x)||$.*

Since we are sometimes interested in conditional expectations, we will also use the following notations for $A \in \mathcal{A}$:

$$\begin{aligned} CF(x \mid A) &= (E[x_k \mid A])_{k \in \mathbb{N}}, \\ VaPo^{CF}(x \mid A) &= \sum_{k \in \mathbb{N}} CF(x \mid A)_k \mathcal{Z}_{(k)} \in \mathcal{Y}, \end{aligned}$$

Theorem 6.2.5. *The value of $x \in \mathcal{X}$ can be decomposed in*

$$Q[x] = Q[VaPo^{CF}(x)] + Q[R(x)],$$

and we have

$$Q[VaPo^{CF}(x)] \geq Q[x]$$

if we use the utility assumption.

Remark 6.2.6. 1. We will denote $x \in \mathcal{X}$ with $x \leq 0$ as a liability. Proposition 6.2.5 hence tells us that we need to reserve more than $Q[VaPo^{CF}(x)]$ for this liability as a consequence of the corresponding uncertainty.

2. A risk measure is a functional (not necessarily linear) $\psi : \mathcal{X} \to \mathbb{R}$ which aims to measure the capital needs in an adverse scenario. There are two risk measures, which are commonly used the *Value at Risk* and the *Expected Shortfall* to a given quantile $\alpha \in \mathbb{R}$. The value at risk (VaR) is defined as the corresponding quantile minus the expected value. The expected shortfall is the conditional expectation of the random variable given a loss bigger than the corresponding loss, again minus the expected value. We can hence speak about a 99.5% VaR or a 99% expected shortfall. It is worthwhile to remark that these two concepts are normally applied to losses. Hence in the context introduced above one would strictly speaking calculating the VaR$(-x)$, when considering $x \in \mathcal{X}$. Furthermore in a lot of applications, such as Solvency II, we assume that there is a Dirac measure (aka stress scenario), which just represents the corresponding VaR-level for example. So concretely the stress scenarios, which are used under Solvency II should in principle represent the corresponding point (Dirac) measures at to the confidence level 99.5 %. In the concrete set up, one would for example assume that $q_x(\omega) \in L^2(\Omega, \mathcal{A}, P)$ is a stochastic mortality and one would define the $A, B \in \mathcal{A}$, as the corresponding probabilities in the average and in the tail. In consequence for a policy $x \in \mathcal{X}$, we would have two replicating portfolios, namely $VaPo^{CF}(x \mid A)$ for the average and $VaPo^{CF}(x \mid B)$ for the stressed event according to the risk measure chosen. The corresponding required risk capital is then given (in present value terms) by $Q[VaPo^{CF}(x \mid B) - VaPo^{CF}(x \mid A)]$.

Definition 6.2.7 (Required Risk Capital). *For a risk measure ψ_α such as VaR or expected shortfall to a security level α we define the required risk capital at time $t \in \mathbb{N}$ by*

$$RC_t(x) \quad = \quad \psi_\alpha(p_k(x - VaPo^{CF}(x))).$$

Remark 6.2.8. 1. If we use $VaR_{99.5\%}$ the required risk capital at time t corresponds to the capital needed to withstand a 1 in 200 year event.

2. The definition above could apply to individual insurance policies, but is normally applied to insurance portfolios $\tilde{x} = \sum_{k=1}^{n} x_k$, where x_k are the individual insurance policies.

3. What is more material than the diversifiable risk is the risk, which affects all of the individual insurance policies at the same time, such as a pandemic event, where the overall mortality could increase by 1 % in a certain year such as 1918.

Definition 6.2.9 (Cost of Capital). *For a unit cost of capital $\beta \in \mathbb{R}^+$ and an insurance portfolio $\tilde{x} \in \mathcal{X}$, we define:*

1. *The present value of the required risk capital by*

$$PVC(\tilde{x}) \quad = \quad Q[\sum_{k \in \mathbb{N}} RC_t(\tilde{x}) \mathcal{Z}_{(k)}].$$

2. *The cost of capital $CoC(\tilde{x})$ is given by:*

$$CoC(\tilde{x}) \quad = \quad \beta \times PVC(\tilde{x}),$$

and $\tilde{Q}$ is defined by $\tilde{Q}[\tilde{x}] = Q[VaPo^{CF}(\tilde{x})] + \beta PVC(\tilde{x})$.

Remark 6.2.10. 1. The concept as defined before is somewhat simplified, since one normally assumes that the required capital C from the shareholder is $\alpha \times C$ after tax and investment income on capital, Assume a tax-rate κ and a risk-free yield of i. In this case we have

$$\alpha \times C = i \times (1 - \kappa) \times C + \beta \times C,$$

and hence $\beta = \alpha - i \times (1 - \kappa)$. In reality the calculation can still become more complex since we discount future capital requirements risk-free and because of the fact that the interest rate i is not constant. In order to avoid these technicalities, we will assume for this book that i is constant.

2. We remark $\tilde{Q}[\tilde{x}]$ is not uniquely determined, but depends on a lot of assumptions such as $\psi\alpha$, α, β, ...

3. For the moment we did not yet see how to actually model $\tilde{x}$ and we remark that one is normally focusing on the non-diversifiable part of the risks within $\tilde{x}$.

Example 6.2.11. We continue with example 6.1.18 and we assume that the risk capital is given by a pandemic event where $\Delta q_x = 1\%$ for all ages. This roughly corresponds to the increase in mortality of 1918 as a consequence of the Spanish flu pandemic. The aim of this example is to calculate the required risk capital and the market value of this policy based on the cost of capital method using $\beta = 6\%$. The required risk capital in this context can be calculated as $\Delta q_x \times 100,000$ and we get the following results:

Age	Unit	Units for Risk Capital	Units for Benefits	Total Units	$-\tilde{Q}[x]$ $i = 2\%$	$-\tilde{Q}[x]$ $i = 4\%$
50	$\mathcal{Z}_{(0)}$	1000.00	-1394.28	-1334.28	-1334.28	-1334.28
51	$\mathcal{Z}_{(1)}$	990.00	-380.34	-320.94	-314.65	-308.60
52	$\mathcal{Z}_{(2)}$	979.11	-276.16	-217.41	-208.97	-201.01
53	$\mathcal{Z}_{(3)}$	967.36	-173.84	-115.80	-109.12	-102.95
54	$\mathcal{Z}_{(4)}$	954.78	-73.67	-16.38	-15.14	-14.00
55	$\mathcal{Z}_{(5)}$	941.41	24.09	80.57	72.98	66.22
56	$\mathcal{Z}_{(6)}$	927.29	119.20	174.84	155.25	138.18
57	$\mathcal{Z}_{(7)}$	912.45	211.44	266.19	231.73	202.28
58	$\mathcal{Z}_{(8)}$	896.94	300.57	354.39	302.47	258.95
59	$\mathcal{Z}_{(9)}$	880.80	386.41	439.26	367.55	308.61
60	$\mathcal{Z}_{(10)}$	–	1673.52	1673.52	1372.87	1130.57
Total					**520.69**	**143.98**

We remark that the value of the policy at inception becomes positive, which means nothing else, that the insurance company does need equity capital to cover the economic loss. It is obvious that this is the case for $i = 2\%$, since the premium principle did not allow for a compensation of the risk capital. More interestingly even at the higher interest rate the compensating effect is not big enough to turn this policy into profitability.

Exercise 6.2.12. In the same sense as for the mortality example calculate the respective risk capitals and the $\tilde{Q}$ for a disability cover.

6.3 Inclusion in the Markov Model

In this section we want to have a look how we could concretely use the recursion technique for the calculation of the cost of capital in a Markov chain similar environment. In order to do that we look at an insurance policy with a term of one year.

We assume that we have a mortality of q_x in case of a "normal" year with a probability of $(1-\alpha)$ and an excess mortality of Δq_x in an extreme year with probability α. We denote with $\Gamma = \frac{q_x + \Delta q_x}{q_x}$. Furthermore we assume a mortality benefit of 100,000. In this case we get the following by some simple calculations:

$$\begin{aligned}
VaPo^{CF}(x) &= (\delta_{1k}(q_x + \alpha(\Gamma - 1)q_x \times 100000))_{k\in\mathbb{N}} \\
RC_1(x) &= (\delta_{1k}(1-\alpha)(\Gamma - 1)q_x \times 100000)_{k\in\mathbb{N}} \\
\tilde{Q}[x] &= Q[(\delta_{1k}(q_x + \alpha(\Gamma - 1)q_x \times 100000 + \\
&+ \beta(1-\alpha)(\Gamma - 1) \times 100000))_{k\in\mathbb{N}}]
\end{aligned}$$

We see that the price of this insurance policy with only payments at time 1 can be decomposed into a part representing best estimate mortality:

$$\delta_{1k}\{q_x\,(1 + \alpha\,(\Gamma - 1))\},$$

where we can arguably say that this $\tilde{q}_x = q_x\,(1 + \alpha\,(\Gamma - 1))$ is our actual best-estimate mortality. On top of that we get a charge for the excess mortality Δq_x with an additional cost of β. Hence we get the following:

1. There is a contribution to the reserve from the people surviving the year with a probability p_x.
2. There is a contribution to the reserve from the people dying in normal years with probability q_x and the defined benefit $a_{*\dagger}^{\text{post}}$, and
3. There is finally a contribution of the people dying in extreme years with probability Δq_x and the additional cost of defined benefit of $\beta \times a_{*\dagger}^{\text{post}}$.

The interesting fact is that we can actually use the same recursion of the reserves for the Markov chain model as in proposition 4.7.3 with the exception that now the "transition probabilities" do not fulfil anymore the requirement that their sum equals 1. However this method provides a pragmatic way to implement the cost of capital in legacy admin systems.

The main problem for the determining of the corresponding Markov chain model is the underlying stochastic mortality model. For the QIS 5 longevity model a similar calculation can be used. In this model it is assumed that the mortality drops by 25 % in an extreme scenario. Hence the calculation goes along the following process:

1. Determine $x_1 = VaPo^{CF}(\tilde{x})$.
2. Determine $x_2 = VaPo^{CF}(\tilde{x})$ for stressed mortality.
3. $\tilde{Q}[x] = Q[x_1] + \beta\, Q[x_2 - x_1]$

Example 6.3.1. In this example we want to revisit the exercise 6.1.18 and we want again to calculate the market value of the insurance liability, but this time with the recursion. We get the following results:

Age	Benefit Normal	Benefit Premium	Excess Risk	Math Res. $i = 2\%$	Value $i = 2\%$	Value $i = 4\%$
50	100000	-1394.28	6000	**0.00**	**520.69**	**143.98**
51	100000	-1394.28	6000	426.43	901.09	542.82
52	100000	-1394.28	6000	765.56	1193.21	861.67
53	100000	-1394.28	6000	1015.22	1394.79	1096.96
54	100000	-1394.28	6000	1172.95	1503.20	1244.68
55	100000	-1394.28	6000	1235.88	1515.45	1300.33
56	100000	-1394.28	6000	1200.79	1428.16	1258.89
57	100000	-1394.28	6000	1064.00	1237.49	1114.74
58	100000	-1394.28	6000	821.42	939.18	861.64
59	100000	-1394.28	6000	468.45	528.45	492.63
60				0	0	0

We remark that this calculation was much faster to calculate since it is based on Thiele's difference equation for the mathematical reserves, and we get at the same time the corresponding results for the classical case and also for the case using the cost of capital approach.

As seen in the calculation above there is a small second order effect, which we can detect, when looking more closely. The results below correspond to the 2% valuation:

```
Direct Method  520.698380872792
Recursion      520.698380872793
```

Exercise 6.3.2. Perform the corresponding calculation for the disability example.

6.4 Asset Liability Management

Until now we have looked only at insurance liabilities as an $x \in \mathcal{X}$. An insurance company needs to cover its insurance liabilities $l = \sum x\iota \in \mathcal{X}$ with corresponding assets, which are also elements in $\mathcal{X}$.

Definition 6.4.1 (Assets and Liabilities). *An $x \in \mathcal{X}$ with a valuation functional Q is called*

1. *an asset if $Q[x] \geq 0$ and*
2. *a liability if $Q[x] \leq 0$.*

Definition 6.4.2 (Insurance balance sheet). *An insurance balance sheet consists of a set of assets $(a_i)_{i \in I}$ and a set of liabilities $(l_j)_{j \in J}$. The equity of an insurance balance sheet is defined as*

$$e = \sum_{i \in I} a_i + \sum_{j \in J} l_j.$$

The insurance entity is called bankrupt if $Q[e] < 0$.

Definition 6.4.3. *In an insurance market, each insurance company is required to hold an adequate amount of risk capital in order to absorb shocks. In order to do that, the regulator defines a risk measure $\psi\alpha$ to a security level α. In this context an insurance company is called solvent if:*

$$Q[e] \geq \psi\alpha(e).$$

Remark 6.4.4. Note that an insurance regulator may not want to use a market consistent approach. Never the less the above definition can be used, be suitably adjust ψ.

Definition 6.4.5 (Asset Liability Management). *Under Asset Liability Management we understand the process of analysing $(l_j)_{j\in J}$ and the (dynamic) management of $(a_i)_{i\in I}$ in order to achieve certain targets , such as remaining solvent.*

Definition 6.4.6. *For an insurance liability $l \in \mathcal{X}$ an asset portfolio $(a_i)_{i\in I}$ is called:*

1. matching *if $\sum_{i\in I} a_i + l = 0$, and*
2. cash flow matching *if $\sum_{i\in I} a_i + VaPo^{CF}(l) = 0$.*

Remark 6.4.7. We remark that is normally not feasible to do a perfect matching, and hence one normally uses a cash flow matching to a achieve a proxy for a perfect match. We also remark that in this case the shareholder equity needs still be able to absorb the basis risk $l - VaPo^{CF}(l)$.

Definition 6.4.8 (Duration). *The duration for an $x \in \mathcal{X}$ with $x = \sum_{i\in\mathbb{N}} \alpha_i \, \mathcal{Z}_{(i)}$ and $\alpha_i \geq 0$ is defined by*

$$d(x) = \frac{Q[\sum_{i\in\mathbb{N}} \alpha_i \times i \times \mathcal{Z}_{(i)}]}{Q[\sum_{i\in\mathbb{N}} \alpha_i \times \mathcal{Z}_{(i)}]}$$

We say that an asset portfolio $(a_i)_{i\in I}$ is duration matching a liability l if the following two conditions are fulfilled:

1. $Q[\sum_{i\in I} a_i + l] = 0$, *and*
2. $d(\sum_{i\in I} a_i) = d(-l)$.

Example 6.4.9. In this example we want to further elaborate on the example 6.1.18 and we want to see how the replicating scenario changes in case a pandemic occurs in year three, with an excess mortality of 1 %. We want also to have a look on what risk is implied in this, assuming that the pandemic at the same time leads to a reduction of interest rates down from 2% to 0.5 %. Finally we want to see an example how we could do a perfect cash flow matching portfolio and duration matched portfolio.

Definitions We assume that $A \in \mathcal{A}$ represents the information that we have going to have average mortality after year 3 and three and that the person survived until then (year 2). In the same sense we assume that $B \in \mathcal{A}$ represents the same as A but with the exception that we assume a pandemic event in the year 3 with an average excess mortality of 1%. For simplicity reasons (to avoid notation) we use $x, y \in \mathcal{X}$ as abbreviations for the corresponding conditional random variables.

Calculation of the Replicating Portfolios In a first step we will calculate the replicating portfolios (starting at time 2) with respect to both A and B. Doing this we get the following results for case A:

Age	Unit	Units for Mortality	Units for Premium	Total Units	Value $i = 2\%$	Value $i = 4\%$
52	$\mathcal{Z}_{(0)}$	–	-1394.28	-1394.28	-1394.28	-1394.28
53	$\mathcal{Z}_{(1)}$	1200.00	-1377.55	-177.55	-174.07	-170.72
54	$\mathcal{Z}_{(2)}$	1284.40	-1359.64	-75.24	-72.32	-69.57
55	$\mathcal{Z}_{(3)}$	1365.21	-1340.61	24.60	23.18	21.87
56	$\mathcal{Z}_{(4)}$	1442.25	-1320.50	121.75	112.48	104.07
57	$\mathcal{Z}_{(5)}$	1515.33	-1299.37	215.95	195.59	177.49
58	$\mathcal{Z}_{(6)}$	1584.27	-1277.28	306.99	272.59	242.61
59	$\mathcal{Z}_{(7)}$	1648.95	-1254.29	394.65	343.57	299.90
60	$\mathcal{Z}_{(8)}$	1709.23	–	1709.23	1458.81	1248.91
Total					**765.56**	**460.30**

For case B we get:

Age	Unit	Units for Mortality	Units for Premium	Total Units	Value $i = 2\%$	Value $i = 4\%$
52	$\mathcal{Z}_{(0)}$	–	-1394.28	-1394.28	-1394.28	-1394.28
53	$\mathcal{Z}_{(1)}$	1200.00	-1377.55	-177.55	-174.07	-170.72
54	$\mathcal{Z}_{(2)}$	**2272.40**	-1345.87	**926.52**	890.54	856.62
55	$\mathcal{Z}_{(3)}$	1351.38	-1327.03	24.35	22.95	21.65
56	$\mathcal{Z}_{(4)}$	1427.64	-1307.12	120.51	111.34	103.01
57	$\mathcal{Z}_{(5)}$	1499.97	-1286.21	213.76	193.61	175.70
58	$\mathcal{Z}_{(6)}$	1568.22	-1264.34	303.88	269.83	240.16
59	$\mathcal{Z}_{(7)}$	1632.24	-1241.58	390.65	340.09	296.86
60	$\mathcal{Z}_{(8)}$	1691.91	–	1691.91	1444.03	1236.26
Total					**1704.05**	**1365.28**

We note two things:

– The pandemic happens when the person is aged 53 and we see the impact in $\mathcal{Z}_{(2)}$ at age 54. This has to do with the convention that we assume that the deaths occur at the end on the year, hence just before the person gets 54.

– We see that the difference in reserves amounts to $1704.05 - 765.56 = 938.49$ which represents the economic loss as a consequence of the pandemic. The biggest contributor to this loss is the increased death benefit, e.g. $926.52 - 1284.40 = 962.87$.

Matching asset portfolios Based on the above it is now easy to calculate the cash flow matching portfolio, by just investing the different amounts of liabilities into the corresponding assets, such as buying $24.60\mathcal{Z}_{(3)}$. We remark that consequently we would have to sell $-177.55\mathcal{Z}_{(1)}$. In normal circumstances for mature businesses this will not occur, since it is a consequence that we consider a term insurance policy and not for example an endowment.

Mismatch in case of a pandemic The table below finally shows the cash flow mismatch as a consequence of the pandemic and we see that in this case the present values do not have a big impact since the main difference is at time 1.

Age	Unit	Units Normal	Units Stress	Difference Units	Value $i = 2\%$	Value $i = 0\%$
52	$\mathcal{Z}_{(0)}$	-1394.28	-1394.28	0.00	0.00	0.00
53	$\mathcal{Z}_{(1)}$	-177.55	-177.55	0.00	0.00	0.00
54	$\mathcal{Z}_{(2)}$	-75.24	926.52	1001.77	962.87	1001.77
55	$\mathcal{Z}_{(3)}$	24.60	24.35	-0.24	-0.23	-0.24
56	$\mathcal{Z}_{(4)}$	121.75	120.51	-1.23	-1.13	-1.23
57	$\mathcal{Z}_{(5)}$	215.95	213.76	-2.18	-1.98	-2.18
58	$\mathcal{Z}_{(6)}$	306.99	303.88	-3.11	-2.76	-3.11
59	$\mathcal{Z}_{(7)}$	394.65	390.65	-3.99	-3.48	-3.99
60	$\mathcal{Z}_{(8)}$	1709.23	1691.91	-17.31	-14.78	-17.31
Total					**938.49**	**973.67**

Example 6.4.10 (Lapses). In this example we want to see how lapses can influence the replicating portfolios. In order to do that we have to change the example 6.1.18 a little bit, as follows:

– We consider a term insurance, for a 50 year old man with a term of 10 years, and we assume that this policy is financed with a regular premium payment. Hence there are actually two different payment streams, namely the premium payment stream and the benefits payment stream. For sake of simplicity we assume that the yearly mortality is $(1 + \frac{x-50}{10} \times 0.1)\%$. We assume that the benefit amounts to 100.000 USD and we assume that the premium has been determined with an interest rate $i = 2\%$.

– In this case the premium amounts to $P = 9562.20$.

– In addition the policyholder can surrender the policy at any time and gets back 98 % of the expected future cash flows valued at the pricing interest rate of 2%. We remark here that this is a risk since the surrenders can happen in case the market value of the corresponding units is below the surrender value.

– We remark that the units have been valued with two (flat) yield curves with interest rates of 2% and 4% respectively.

In order to calculate this example we will perform the following steps:

1. Calculation of the cash flow matching portfolio in case of no surrenders.
2. Calculation of the cash flow including lapses with an average lapse rate of 7 %
3. Calculation of the cash flows at time 2, assuming average lapses, lapses at 25 % at time 2.

Calculation of the cash flow matching portfolio in case of no surrenders:

Age	Unit	Units for Mortality	Units for Premium	Total Units	Value $i = 2\%$	Value $i = 4\%$
50	$\mathcal{Z}_{(0)}$	–	-9562.20	-9562.20	-9562.20	-9562.20
51	$\mathcal{Z}_{(1)}$	1000.00	-9466.57	-8466.57	-8300.56	-8140.94
52	$\mathcal{Z}_{(2)}$	1089.00	-9362.44	-8273.44	-7952.17	-7649.26
53	$\mathcal{Z}_{(3)}$	1174.93	-9250.09	-8075.16	-7609.40	-7178.79
54	$\mathcal{Z}_{(4)}$	1257.56	-9129.84	-7872.27	-7272.76	-6729.25
55	$\mathcal{Z}_{(5)}$	1336.69	-9002.02	-7665.32	-6942.72	-6300.34
56	$\mathcal{Z}_{(6)}$	1412.12	-8866.99	-7454.87	-6619.71	-5891.69
57	$\mathcal{Z}_{(7)}$	1483.67	-8725.12	-7241.45	-6304.11	-5502.90
58	$\mathcal{Z}_{(8)}$	1551.18	-8576.79	-7025.61	-5996.29	-5133.54
59	$\mathcal{Z}_{(9)}$	1614.50	-8422.41	-6807.90	-5696.55	-4783.14
60	$\mathcal{Z}_{(10)}$	–	88080.30	88080.30	72256.53	59503.90
Total					**0**	**-7368.19**

We remark that the there is considerable value in the policy if we assume no lapses, in case we earn a higher interest rate, such as 4 %.

Calculation of the cash flow matching portfolio in case of 7% surrenders:

Age	Unit	Units for Mortality	Units for Premium	Total Units	Value $i = 2\%$	Value $i = 4\%$
50	$\mathcal{Z}_{(0)}$	–	-9562.20	-9562.20	-9562.20	-9562.20
51	$\mathcal{Z}_{(1)}$	1019.67	-8797.22	-7777.54	-7625.04	-7478.40
52	$\mathcal{Z}_{(2)}$	1594.01	-8084.65	-6490.63	-6238.59	-6000.95
53	$\mathcal{Z}_{(3)}$	2066.98	-7421.70	-5354.72	-5045.87	-4760.33
54	$\mathcal{Z}_{(4)}$	2449.23	-6805.70	-4356.47	-4024.71	-3723.93
55	$\mathcal{Z}_{(5)}$	2750.73	-6234.02	-3483.29	-3154.92	-2863.01
56	$\mathcal{Z}_{(6)}$	2980.77	-5704.13	-2723.35	-2418.26	-2152.31
57	$\mathcal{Z}_{(7)}$	3147.94	-5213.57	-2065.63	-1798.25	-1569.71
58	$\mathcal{Z}_{(8)}$	3260.18	-4759.99	-1499.81	-1280.07	-1095.89
59	$\mathcal{Z}_{(9)}$	3324.77	-4341.11	-1016.34	-850.42	-714.06
60	$\mathcal{Z}_{(10)}$	5486.26	42159.27	47645.53	39085.93	32187.61
Total					**-2912.44**	**-7733.21**

We remark that at that time, the company makes still some additional gains as a consequence of the 2% surrender penalty.

Calculation of the cash flow matching portfolio in case of high surrenders: We assume that there has been observed an exceptional lapse rate at time 2 of 25% of the portfolio.

Age	Unit	Units for Mortality	Units for Premium	Total Units	Value $i = 2\%$	Value $i = 4\%$
50	$\mathcal{Z}_{(0)}$	–	-9562.20	-9562.20	-9562.20	-9562.20
51	$\mathcal{Z}_{(1)}$	1019.67	-8797.22	-7777.54	-7625.04	-7478.40
52	$\mathcal{Z}_{(2)}$	1594.01	-8084.65	-6490.63	-6238.59	-6000.95
53	$\mathcal{Z}_{(3)}$	4773.16	-5966.47	-1193.30	-1124.48	-1060.84
54	$\mathcal{Z}_{(4)}$	1968.98	-5471.25	-3502.26	-3235.55	-2993.75
55	$\mathcal{Z}_{(5)}$	2211.37	-5011.66	-2800.29	-2536.31	-2301.63
56	$\mathcal{Z}_{(6)}$	2396.30	-4585.67	-2189.36	-1944.09	-1730.28
57	$\mathcal{Z}_{(7)}$	2530.70	-4191.30	-1660.60	-1445.65	-1261.92
58	$\mathcal{Z}_{(8)}$	2620.93	-3826.66	-1205.73	-1029.08	-881.01
59	$\mathcal{Z}_{(9)}$	2672.85	-3489.91	-817.05	-683.67	-574.05
60	$\mathcal{Z}_{(10)}$	4410.52	33892.74	38303.27	31422.02	25876.31
Total					**-4002.67**	**-7968.76**

ALM Risk of mass lapses Finally we want to look what happens when we have mass lapses as indicated before, but if we have invested in the cash flow matching portfolio according to average 7 % lapses. Hence we have to calculate the assets according to 7 % lapses and the liabilities according 25 % lapses.

Age	Unit	Units for Assets	Units for Liability	Total Units	Value $i = 2\%$	Value $i = 4\%$
52	$\mathcal{Z}_{(0)}$	-6490.63	6490.63	0	0	0
53	$\mathcal{Z}_{(1)}$	-5354.72	1193.30	-4161.42	-4079.82	-4001.36
54	$\mathcal{Z}_{(2)}$	-4356.47	3502.26	-854.21	-821.04	-789.76
55	$\mathcal{Z}_{(3)}$	-3483.29	2800.29	-682.99	-643.60	-607.18
56	$\mathcal{Z}_{(4)}$	-2723.35	2189.36	-533.99	-493.32	-456.45
57	$\mathcal{Z}_{(5)}$	-2065.63	1660.60	-405.02	-366.84	-332.90
58	$\mathcal{Z}_{(6)}$	-1499.81	1205.73	-294.08	-261.13	-232.41
59	$\mathcal{Z}_{(7)}$	-1016.34	817.05	-199.28	-173.48	-151.43
60	$\mathcal{Z}_{(8)}$	47645.53	-38303.27	9342.26	7973.53	6826.29
Total					**1134.26**	**254.76**

Now we see that the lapses induce quite a big risk for the company since they lose in case of mass lapses almost 1 % of the face value of the policy, more concretely $1134.26 - 254.76 = 879.50$.

The above example shows very clearly how the behaviour of the policyholders can change the cash flow matching portfolio and in consequence induces a risk. As a consequence the risk minimising portfolio in the sense of $VaPo^*(x)$ for an insurance portfolio $x \in \mathcal{X}$ does also consist of additional assets offsetting the corresponding risks. In the above example the corresponding asset would be a (complex) put option, which allows to sell the bond portfolio at the predefined (book-) values. So in reality insurance companies aim to model these risk in order to determine the corresponding assets to manage and reduce the undesired risk.

In the example above we have assumed that at a given year 25% of the policies in force lapse. In practise one models the dynamic lapse behaviours. Eg the lapse rate is a function of the interest differential between market and book yields. Normally the corresponding lapse rates stay below 1, which is interesting. Assuming a market efficient behaviour, one would expect that there is a binary decision of the policyholders to stick to the contract or to lapse as a function of the before mentioned interest differential. In consequence the underlying theory how to model such policyholder behaviour is not as crisp and transparent as with the arbitrage free pricing theory, since market efficient behaviours is normally not observed. As a corollary there is a lot of model risk intrinsic to these calculations and it is important to test the results from the models with different scenarios.

Remark 6.4.11. At the end of this section a remark on how to determine a $VaPo^*(x)$ for an $x \in \mathcal{X}$: One normally models an $l \in \mathcal{X}$ and simulates $l(\omega)$ together with some test assets $D \subset \mathcal{Y}$ observable prices and cash flows. We denote $D = \{d_1, \dots d_n\}$. Hence at the end of this process we have a vector

$$\mathcal{W} := (l(\omega_i), d_1(\omega_i), \dots d_n(\omega_i))_{i \in I}.$$

Now the process is quite canonical:

1. We define a distance between two $x, y \in \mathcal{X}$, for example by means of $||x||$ as defined.
2. We solve the numerical optimisation problem, for minimising the distance between l and the target $y \in \text{span} < \mathcal{D} >$.

We note two things:

– The numerical procedures to determine y can sometimes prove to be difficult since the corresponding design matrix can be near to a singular matrix, and hence additional care is needed.

– In case of the $|| \bullet ||$ defined before, we remark that it has been deducted from the Hilbert space $\mathcal{X}$. Hence what we actually doing is to use the projection $\tilde{p} : \mathcal{X} \to \text{span} < \mathcal{D} >$, which can be expressed by means of $< \bullet, \bullet >$. We remark that $y = \tilde{p}(x)$.

7. Unit-linked policies

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

7.1 Introduction

Up to now we have mostly considered models with deterministic interest rate or with an interest rate given by a Markov chain on a finite state space. This helped us to keep the calculations simple. In this chapter we have a look at some more general models. On the one hand we consider models for policies whose actual value depends on the performance of an underlying unit (usually a fond), on the other hand we will discuss further models with stochastic interest rate.

With respect to these two types of models one has to note, that the corresponding theory is still in development and no final presentation is available. Especially, up to now, there is no mutual consent on the models used for the interest rate processes. In fact there are various opinions on the ideal model for the stochastic interest rate, and each of these models yields a different numerical result.

What are the advantages and disadvantages of these models? As mentioned before, models try to resemble reality and they are more or less precise. Thus clearly the following relation holds: if the model has more stochastic elements, then

– it might resemble reality more precisely,

– it also becomes more complex with respect to assumptions, parameter fitting and numerical algorithms.

In our opinion, the main difficulty when using a stochastic interest rate is due to the diversity of ideas for the possible models. In the following we do not want to rate these models, but our aim is to describe a general approach which provides a fundamental knowledge of this area and of the required techniques.

We begin with a look at policies whose value is tied to a bond or a fond. The payout of theses so called "unit-linked policies" usually consists of a certain number of shares of a fond (the underlying unit) to the insured, in case of an occurrence of the insured event.

These policies have the characteristic feature that the benefits (endowments or death benefits) are not deterministic, but random. A unit-linked policy is usually financed by a single premium. This type of financing is preferred due to the management of these policies. Note that this is in contrast to traditional policies. Moreover one has to note, that the value at risk is constant for a traditional policy, but for a unit-linked policy it depends on the underlying fond.

To analyse unit-linked policies, we introduce the following notation:

$N(t)$	number of shares at time t,
$S(t)$	value of a share at time t.

We assume that $N(t)$ is deterministic. The relevant quantities for a life insurance in this setting and in the traditional setting are summarised in the following table:

	traditional	(pure) unit-linked
death benefit	$C(t) = 1$	$C(t) = S(t)$
value (time 0)	$\pi_0(t) = \exp(-\delta t)$	$\pi_0(t) = S(0)$ (assuming a "normal" economy)
single premium	$E\left[\int_0^T \pi_0(t) d(\chi_{T_x \leq t})\right]$ $= \int_0^T \exp(-\delta t)_t p_x \mu_{x+t} dt$	$E\left[\int_0^T \pi_0(t) d(\chi_{T_x \leq t})\right]$ $= S(0) \int_0^T {}_t p_x \mu_{x+t} dt$ $= (1 - {}_T p_x) S(0)$

The above terms indicate that the financial risk taken by the insurer is smaller for a unit-linked product than for a traditional product with fixed technical interest rate[1]. Furthermore, note that in the calculation of the single premium we implicitly assumed that the mean of the discounted (to time 0) value of the fond at time t coincides with the value of the fond at time 0. This means, that first of all we have to start with a discussion of the value or price of a fond.

The model also did not include any guarantees. But in general one would like to add a guarantee (e.g. a refund guarantee for the paid in premiums) to the policy. For example, the guarantee could be of the form

$$G(t) = \int_0^t \bar{p}(s) ds,$$

where $\bar{p}(s)$ denotes the density of the premiums at time s. More general one could be interested in a refund guarantee of the paid in premiums with an additional interest at a fixed rate:

$$G(t) = \int_0^t \exp(r(t-s)) \bar{p}(s) ds.$$

In these examples the payout function would be

$$C(t) = \max(S(t), G(t)).$$

Let us assume that the value of the fond is given by a stochastic process (with distribution P). What is, in this setting, the value of the discounted payment $C(t)$ at time t? A first guess might be

$$\pi_0\left(C(t)\right) = E^P\left[\max(S(t), G(t))\right].$$

But it is not that simple! If this would be the value of the payment, there would be the possibility to make a profit without risk (arbitrage). In order to prevent this possibility one has to change the measure P. In mathematical finance it is proved that an equivalent martingale measure exists, such that there is no arbitrage. Then in a "fair" market we have

$$\pi_0\left(C(t)\right) = E^Q\left[\max(S(t), G(t))\right],$$

where Q is a measure equivalent to P such that the discounted value of the underlying fond is a martingale.

Furthermore, in mathematical finance payments like $C(t)$ are called the payouts of an option. To determine the price of an option, one uses the arbitrage free pricing theory. A quick introduction to this theory will be given in the next section.

[1] Actually, this is not true in general. Here we implicitly assumed that the capital market risks of a unit-linked policy are minimised by an appropriate trading strategy. For a classical insurance such a trading strategy replicates the cash flows by zero coupon bonds with the corresponding maturities.

7.2 Pricing theory

In this section we have a look at modern financial mathematics. It is not our aim to give a comprehensive exposition with proofs of every detail, which would easily fill a whole book. We only want to give a brief survey which illustrates the theory. The reader interested in more details is referred to [Pli97], [HK79], [HP81] and [Duf92].

In this context we clearly also have to mention the paper of Black and Scholes [BS73] with their famous formula for pricing of options.

7.2.1 Definitions

First of all we start with an example which illustrates the use of the pricing theory. The price of a share, modelled by a geometric Brownian motion $(S_t(\omega))$, might develop as shown in Figure 7.1.

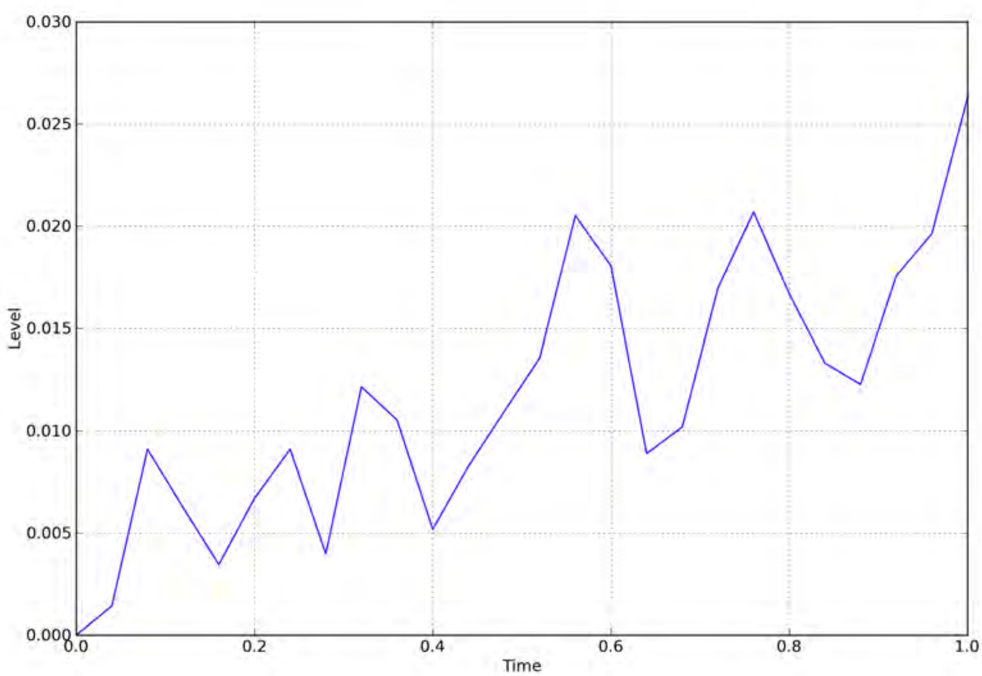

Figure 7.1. Movement of a share price

A European call option for a certain share is the right to buy these shares at a fixed price c (strike price) at a fixed time T. The value of this right at time T is

$$H = \max\left(S_T - c, 0\right).$$

Now a bank would like to know the value (i.e., the fair price) of this option at time 0. As noted in the previous section, taking the expectation would systematically yield the wrong values. In many cases this would provide the possibility to make profits without risk. Thus there would be arbitrage opportunities.

To simplify the exposition we will consider the most simple models of the economy, i.e. finite models. In particular also the time set will be discrete. The reader interested in the corresponding theorems for continuous time can for example find these in [HP81]. In the following the ideas and concepts of the pricing theory are presented.

Let $(\Omega, \mathcal{A}, P)$ be a probability space where Ω is a finite set. Moreover, we assume that $P(\omega) > 0$ holds for all $\omega \in \Omega$.

We also fix a finite time T, as the time at which all trading is finished. The σ-algebra of the observable events at time t is denoted by $\mathcal{F}_t$, and the shares are traded at the times $\{0, 1, 2, \dots, T\}$.

We suppose that there are $k < \infty$ stochastic processes, which represent the prices of the shares $1, \dots, k$, i.e.,

$$S = \{S_t, t = 0, 1, 2, \dots, T\} \text{ with components } S^0, S^1, \dots S^k.$$

As usual, we assume that each S^j is adapted to $(\mathcal{F}_t)_t$. Here S_t^j can be understood as the price of the jth share at time t. The fact that the price process has to be adapted reflects the necessity that one has to know at time t the previous price of S. The share S^0 plays a special role. We suppose that $S_t^0 = (1+r)^t$, i.e., we have the possibility to make risk free investments which provide interest rate r. The risk free discount factor is defined by

$$\beta_t = \frac{1}{S_t^0}.$$

Next, we are going to define what is meant by a trading strategy.

Definition 7.2.1. *A* trading strategy *is a previsible ($\phi_t \in \mathcal{F}_{t-1}$) process $\Phi = \{\phi_t, t = 1, 2, \dots, T\}$ with components ϕ_t^k.*

We understand ϕ_t^k as the number of shares of type k which we own during the time interval $[t-1, t[$. Therefore ϕ_t is called the portfolio at time $t-1$.

Notation 7.2.2. *Let X, Y be vector valued stochastic processes. Then we use the notations:*

$$\begin{aligned} < X_s, Y_t > &= X_s \cdot Y_t = \sum_{k=0}^{n} X_s^k \times Y_t^k, \\ \Delta X_t &= X_t - X_{t-1}. \end{aligned}$$

Next, we want to determine the value of the portfolio at time t:

time	value of the portfolio
$t-1$	$\phi_t \cdot S_{t-1}$
t^-	$\phi_t \cdot S_t$

Thus, the return in the interval $[t-1, t[$ is $\phi_t \cdot \Delta S_t$, and hence the total return is

$$G_t(\phi) = \sum_{\tau=1}^{t} \phi_\tau \cdot \Delta S_\tau.$$

We fix $G_0(\phi) = 0$, and $(G_t)_{t \ge 0}$ is called return process.

Theorem 7.2.3. *G is an adapted and real valued stochastic process.*

Proof. The proof is left as an exercise to the reader.

Definition 7.2.4. *A trading strategy is* self financing*, if*

$$\phi_t \cdot S_t = \phi_{t+1} \cdot S_t, \qquad \forall t = 1, 2, \ldots, T-1.$$

A self financing trading strategy is just a trading strategy where at no time further money is added to or deduced from the portfolio.

Definition 7.2.5. *A trading strategy is* admissible*, if it is self financing and*

$$V_t(\phi) := \begin{cases} \phi_t \cdot S_t, & \text{if} \quad t = 1, 2, \ldots, T, \\ \phi_1 \cdot S_0, & \text{if} \quad t = 0 \end{cases}$$

is non negative. (In other words, one is not allowed to become bankrupt.) The set of admissible trading strategies is denoted by Φ.

Remark 7.2.6. The idea of admissible trading strategies is to consider only portfolios which neither lead to bankruptcy nor allow an addition or deduction of money. This also indicates, that the value of the trading strategies remains constant when the portfolio is rearranged. Thus a trading strategy, which generates the same cash flow as an option, can be used to determine the value of the option.

Definition 7.2.7. *A contingent claim is a positive random variable X. The set of all contingent claims is denoted by $\mathcal{X}$.*

A random variable X is attainable, if there exists an admissible trading strategy $\phi \in \Phi$ which replicates it, i.e.

$$V_T(\phi) = X.$$

In this case one says "ϕ replicates X".

Definition 7.2.8. *The price of an attainable contingent claim, which is replicated by ϕ, is denoted by*

$$\pi = V_0(\phi)$$

(We will see later, that this price is not necessarily unique. It coincides with the initial value of the portfolio.)

7.2.2 Arbitrage

We say, the model offers arbitrage opportunities, if there exists

$$\phi \in \Phi \text{ with } V_0(\phi) = 0 \text{ and } V_T(\phi) \text{ positive and } P[V_T(\phi) > 0] > 0,$$

i.e., money is generated out of nothing. If such a strategy exists, one can make a profit without taking any risks. One of the axioms of modern economy says, that there are no arbitrage opportunities. This is fundamental for some important facts in the option pricing theory.

Now, we are going to define what is meant by a price system.

Definition 7.2.9. *A mapping*

$$\pi : \quad \mathcal{X} \to [0, \infty[, \quad X \mapsto \pi(X)$$

is called price system *if and only if the following conditions hold:*

- $\pi(X) = 0 \iff X = 0$,
- π *is linear.*

A price system is consistent, *if*

$$\pi(V_T(\phi)) = V_0(\phi) \qquad \text{for all } \phi \in \Phi.$$

The set of all consistent price systems is denoted by Π, and $\mathbb{P}$ denotes the set

$$\mathbb{P} = \{Q \text{ is a measure equivalent to } P, \text{ s.th. } \beta \times S \text{ is a martingale w.r.t. } Q\},$$

where β is the discount factor from time t to 0. The measures $\mu \in \mathbb{P}$ are called equivalent martingale measures.

Theorem 7.2.10. *There is a bijection between the consistent price systems $\pi \in \Pi$ and the measures $Q \in \mathbb{P}$. It is given by*

1. $\pi(X) = E^Q[\beta_T X]$.
2. $Q(A) = \pi(S_T^0 \chi_A)$ *for all* $A \in \mathcal{A}$.

Proof. Let $Q \in \mathbb{P}$. We define $\pi(X) = E^Q[\beta_T X]$. Then π is a price system, since P is strictly positive on Ω and Q is equivalent to P. Thus it remains to show, that π is consistent. For $\phi \in \Phi$ we get

$$\begin{aligned} \beta_T V_T(\phi) &= \beta_T \phi_T S_T + \sum_{i=1}^{T-1} (\phi_i - \phi_{i+1}) \beta_i S_i \\ &= \beta_1 \phi_1 S_1 + \sum_{i=2}^{T} \phi_i (\beta_i S_i - \beta_{i-1} S_{i-1}), \end{aligned}$$

where we used that ϕ is self financing. This yields

$$\begin{aligned} \pi(V_T(\phi)) &= E^Q[\beta_T V_T(\phi)] \\ &= E^Q[\beta_1 \phi_1 S_1] + E^Q\left[\sum_{i=2}^{T} \phi_i (\beta_i S_i - \beta_{i-1} S_{i-1})\right] \\ &= E^Q[\beta_1 \phi_1 S_1] + \sum_{i=2}^{T} E^Q\left[\phi_i E^Q[(\beta_i S_i - \beta_{i-1} S_{i-1}) | \mathcal{F}_{i-1}]\right] \\ &= \phi_1 E^Q[\beta_1 S_1] \\ &= \phi_1 \beta_0 S_0, \end{aligned}$$

since ϕ is previsible and βS is a martingale with respect to Q.

Thus, π is a consistent price system.

Now let $\pi \in \Pi$ be a consistent price system and Q be defined as above. Then $Q(\omega) = \pi(S_t^0 \chi_{\{\omega\}}) > 0$ holds for all $\omega \in \Omega$, since $S_t^0 \chi_{\{\omega\}} \neq 0$. Moreover, we have $\pi(X) = 0 \iff X = 0$ and therefore Q is absolutely continuous with respect to P.

In the next step, we are going to show that Q is a probability measure. We define

$$\phi^0 = 1 \qquad \text{and} \qquad \phi^k = 0 \quad \forall k \neq 0.$$

Hence, by the consistency of π, we get

$$\begin{aligned} 1 &= V_0(\phi) \\ &= \pi(V_T(\phi)) \\ &= \pi(S_T^0 \cdot 1) \\ &= Q(\Omega). \end{aligned}$$

The prices of positive contingent claims are positive and Q is additive. Therefore, Kolmogorov's axioms are satisfied, since Ω is finite. We have $Q(\omega) = \pi(S_T^0 \cdot \chi_{\{\omega\}})$ by definition. Hence, also

$$E[f] = \sum_{\omega} \pi(S_T^0 \cdot \chi_{\{\omega\}}) \cdot f(\omega) = \pi(S_T^0 \cdot \sum_{\omega} f(\omega)).$$

Thus, with $f = \beta_t\, X$, we have

$$E^Q[\beta_T\, X] = \pi(S_T^0 \cdot \beta_T \cdot X) = \pi(X).$$

Now we still have to show that $\beta_T\, S_T^k$ is a martingale for all k. Let k be a coordinate and τ be a stopping time, and set

$$\begin{aligned} \phi_t^k &= \chi_{\{t \leq \tau\}}, \\ \phi_t^0 &= \left(S_\tau^k / S_\tau^0\right) \chi_{\{t > \tau\}}. \end{aligned}$$

(We keep the share k up to time τ, then it is sold and the money is used for a risk free investment.) It is easy to show, that the strategy ϕ is previsible and self financing. Finally, for an arbitrary stopping time τ,

$$\begin{aligned} V_0(\phi) &= S_0^k, \\ V_T(\phi) &= \left(S_\tau^k / S_\tau^0\right) S_T^0 \\ &\text{and} \\ S_0^k &= \pi(S_T^0 \cdot \beta_\tau \cdot S_\tau^k) \\ &= E^Q\left[\beta_\tau \cdot S_\tau^k\right]. \end{aligned}$$

Thus $\beta_T\, S_T^k$ is a martingale with respect to Q.

Above we have proved one of the main theorems in the option pricing theory. Next, we will present further statements without proofs. They all can be found for example in [HP81].

Theorem 7.2.11. *The following statements are equivalent*

1. *There is no arbitrage opportunity,*
2. $\mathbb{P} \neq \emptyset$,
3. $\Pi \neq \emptyset$.

Lemma 7.2.12. *Suppose there exists a self financing strategy $\phi \in \Phi$ such that*

$$V_0(\phi) = 0,\ V_T(\phi) \geq 0,\ E[V_T(\phi)] > 0.$$

Then there exists an arbitrage opportunity.

Example 7.2.13. We are going to calculate the price of an option for a simple example. Consider a market with two shares $Z = (Z_1, Z_2)$ which are traded at the times $t = 0$, $t = 1$ and $t = 2$. Figure 7.2 shows the possible behaviour of these shares in form of a tree. To calculate the price of the option we suppose that all nine possibilities have the same probability.

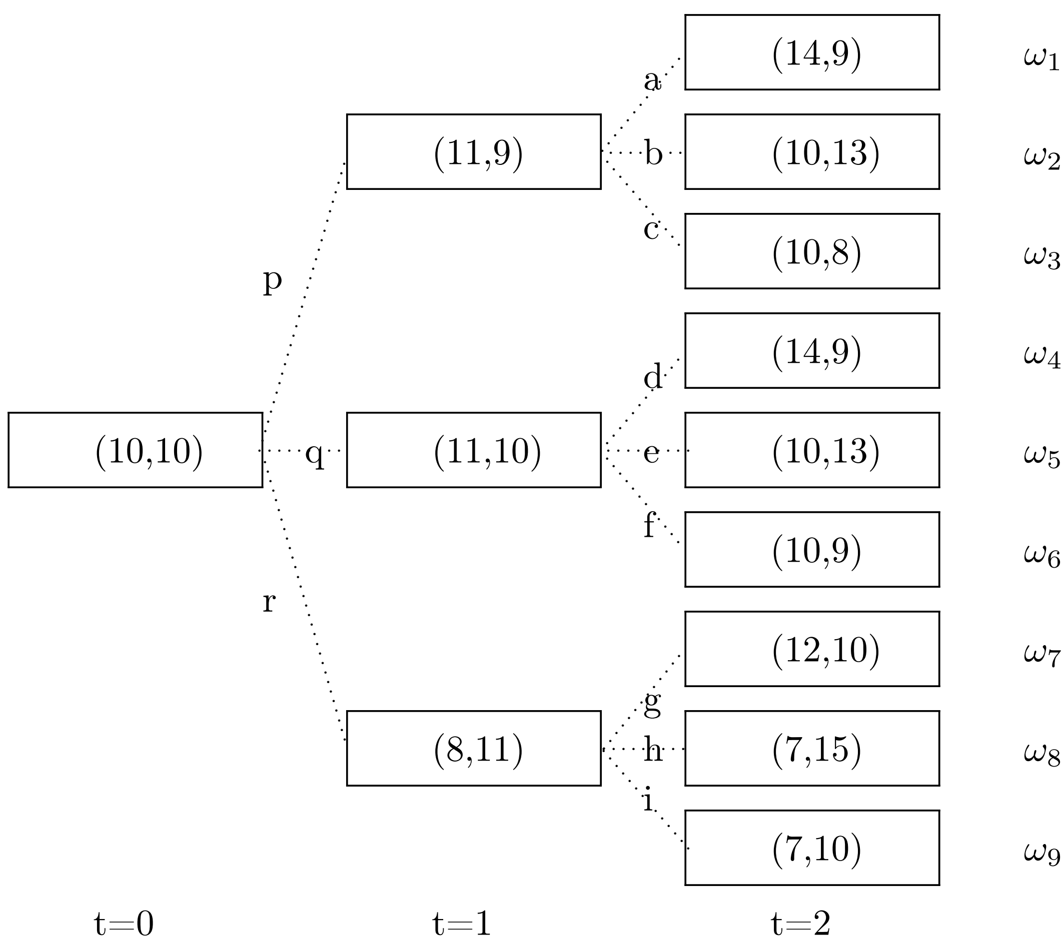

Figure 7.2. Example calculation of an option price

We want to calculate the price of a complex option given by

$$X = \{2Z_1(2) + Z_2(2) - [14 + 2\min(\min\{Z_1(t), Z_2(t)\}, 0 \le t \le 2)]\}^+ .$$

First of all, we have to find an equivalent martingale measure. Thus we have to solve for the times $t = 0$ and $t = 1$ the following equations:

$$\begin{aligned} 10 &= 11p + 11q + 8r, && \text{(martingale condition for } Z_1\text{)} \\ 10 &= 9p + 10q + 11r, && \text{(martingale condition for } Z_2\text{)} \\ 1 &= p + q + r. \end{aligned}$$

The solution to these equations is $p = q = r = \frac{1}{3}$.

Here we can see explicitly which circumstances imply the existence and uniqueness of a martingale measure. In this example the martingale measure is, from a geometric point of view, defined as the intersection of three hyper-planes. Depending on their orientation, there is either one or there are many or there is none equivalent martingale measure.

Next, we can derive the equations for the times $t = 1$ and $t = 2$. These are

$$\begin{aligned} 11 &= 14a + 10b + 10c, \\ 9 &= 9a + 13b + 8c, \\ 1 &= a + b + c, \\ \\ 11 &= 14d + 10e + 10f, \\ 10 &= 9d + 13e + 9f, \\ 1 &= d + e + f, \\ \\ 8 &= 12g + 7h + 7i, \\ 11 &= 10g + 15h + 10i, \\ 1 &= g + h + i, \end{aligned}$$

and they are solved by $(a, b, c) = (0.25,\ 0.15,\ 0.60)$, (d, e, f)
$= (0.25,\ 0.25,\ 0.50)$ and $(g, h, i) = (0.20,\ 0.20,\ 0.60)$.

Now we know the transition probabilities with respect to the martingale measure, which enables us to calculate the martingale measure Q itself. The results of these calculations are summarised in the following table:

state	$X(\omega_i)$	$Q(\omega_i)$
ω_1	5	1/12
ω_2	1	1/20
ω_3	0	1/5
ω_4	5	1/12
ω_5	0	1/12
ω_6	0	1/6
ω_7	4	1/15
ω_8	1	1/15
ω_9	0	1/5

Finally, we can calculate the price of the option as expectation with respect to Q. The result is $\frac{73}{60}$.

7.2.3 Continuous time models

For models in continuous time we restrict our exposition to the statements, the proofs can be found in the references mentioned before. A major difference between the discrete and the continuous setting is that we are going to *assume* that $\mathbb{P} \neq \emptyset$ holds for the continuous time model.

We start with some basic definitions.

Definition 7.2.14. – *A trading strategy ϕ is a locally bounded, previsible process.*

– *The value process corresponding to a trading strategy ϕ is defined by*

$$V : \Pi \to \mathbb{R}, \phi \mapsto V(\phi) = \phi_t \cdot S_t = \sum_{i=0}^{k} \phi_t^k \cdot S_t^k.$$

– *The return process G is defined by*

$$G : \Pi \to \mathbb{R}, \phi \mapsto G(\phi) = \int_0^\tau \phi \, dS = \int_0^\tau \sum_{i=0}^{k} \phi^k dS^k.$$

– *ϕ is self financing, if $V_t(\phi) = V_0(\phi) + G_t(\phi)$.*

– *To define admissible trading strategies we use the notation:*

$$\begin{aligned} Z_t^i &= \beta_t \cdot S_t^i, && \textit{discounted value of share } i \\ G^*(\phi) &= \int \sum_{i=1}^{k} \phi^i \, dZ^i, && \textit{discounted return} \\ V^*(\phi) &= \beta \, V(\phi) = \phi^0 + \sum_{i=1}^{k} \phi^i \, Z^i. \end{aligned}$$

A trading strategy is called admissible, if it has the following three properties:

1. *$V^*(\phi) \geq 0$,*
2. *$V^*(\phi) = V^*(\phi)_0 + G^*(\phi)$,*
3. *$V^*(\phi)$ is a martingale with respect to Q.*

Theorem 7.2.15. *1. The price of a contingent claim X is given by $\pi(X) = E^Q[\beta_T \, X]$.*

2. A contingent claim is attainable $\iff$ $V^* = V_0^* + \int H dZ$ for all H.

Definition 7.2.16. *The market is called complete, if every integrable contingent claim is attainable.*

Although this theory is very important, we only gave a brief sketch of the main ideas. Thus it is recommended that the reader extends his knowledge of financial mathematics by consulting the references.

7.3 The economic model

As we have seen in the previous sections we need an underlying economic model to calculate the price of an option. In principle one can use various different economic models. Exemplary we are going to consider the most common model: geometric Brownian motion.

The following reference are a good sources for various aspects of the economic model: [Dot90], [Duf88], [Duf92], [CHB89], [Per94], [Pli97].

Convention 7.3.1 (General conventions). *For the remainder of this chapter we will use the following notations and conventions:*

- T_x *denotes the future lifespan of an* x *year old person.*
- *The* σ*-algebras generated by* T_x *are denoted by* $\mathcal{H}_t = \sigma\left(\{T > s\}, 0 \leq s \leq t\right).$
- *We assume, that the values of the shares in the portfolio are given by standard Brownian motions* W*. (Compare with Figure 7.3.).*
- $\mathcal{G}_t$ *denotes the* σ*-algebra generated by* W *augmented by the* P*-null sets.*

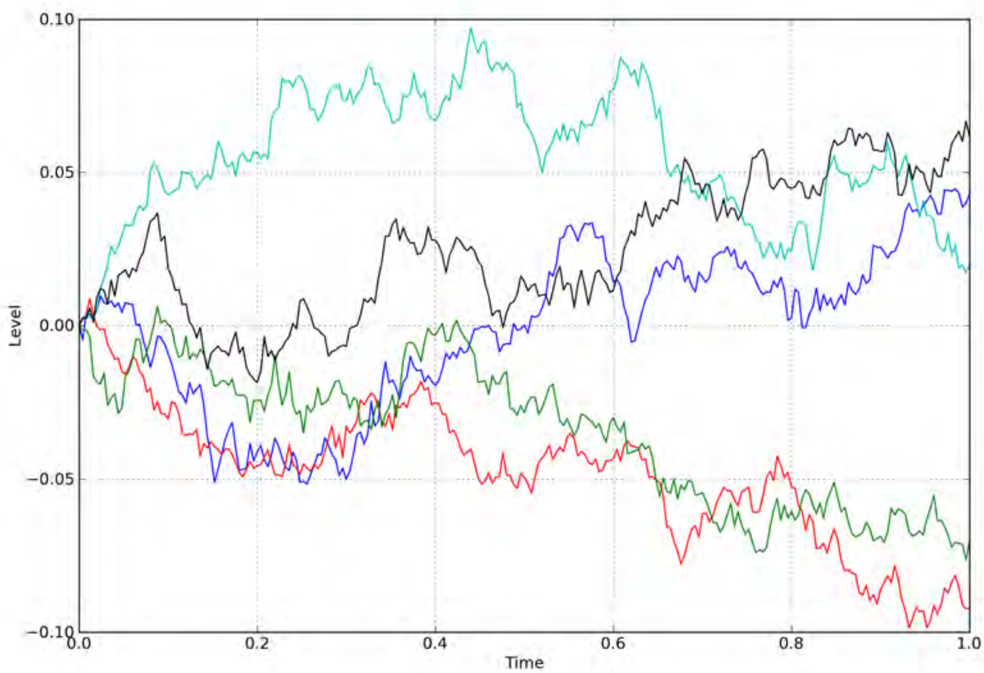

Figure 7.3. 5 Simulations of a Brownian motion

Convention 7.3.2 (Independence of the financial variables). – *We assume that* $\mathcal{G}_t$ *and* $\mathcal{H}_t$ *are stochastically independent. This means, that the financial variables are independent of the future lifespan.*

- $\mathcal{F}_t = \sigma\left(\mathcal{G}_t, \mathcal{H}_t\right)$ *denotes the* σ*-algebra generated by* $\mathcal{G}_t$ *and* $\mathcal{H}_t$*.*

Definition 7.3.3 (Black-Scholes model). *This economic model consists of two investment options:*

$B(t)$	$=$	$\exp(\delta\, t)$	*risk free investment.*
$S(t)$	$=$	$S(0)\,\exp\left[\left(\eta - \frac{1}{2}\sigma^2\right)\, t \,+\, \sigma\, W(t)\right]$	*shares, modeled by a geometric Brownian motion (cf. Fugue 7.4).*

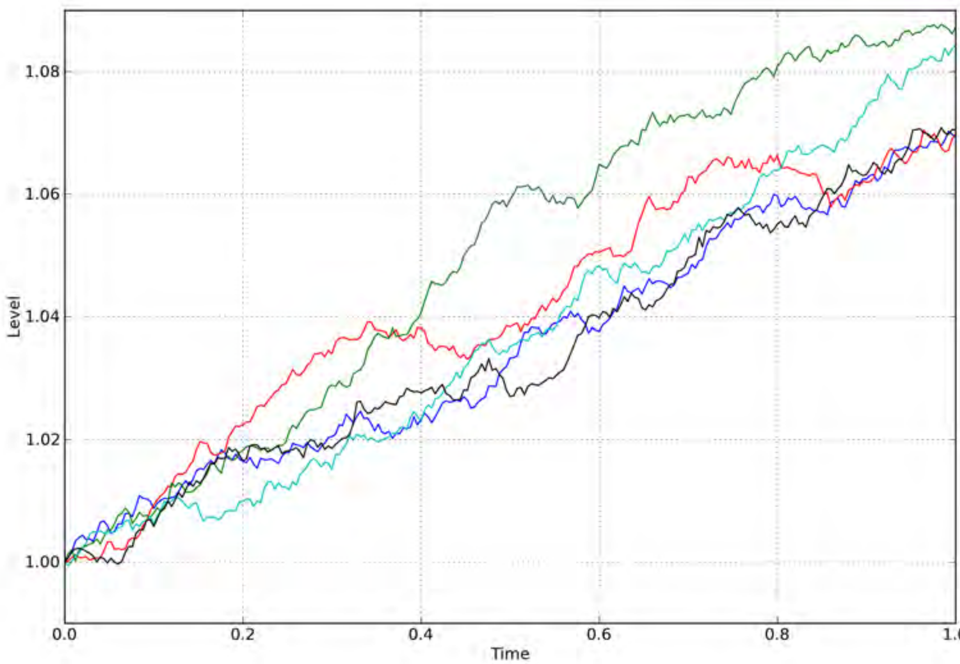

Figure 7.4. 5 Simulations of a geometric Brownian motion

S is the solution to the following stochastic differential equation:

$$dS = \eta\, S\, dt \,+\, \sigma\, S\, dW.$$

Exercise 7.3.4. Prove that S solves the stochastic differential equation given above.

Next we will calculate the discounted values of B and S:

$$B^*(t) = \frac{B(t)}{B(t)} = 1,$$
$$S^*(t) = \frac{S(t)}{B(t)} = S(0)\,\exp\left[\left(\eta - \delta - \tfrac{1}{2}\sigma^2\right)\, t \,+\, \sigma\, W(t)\right].$$

Thus we have defined the investment options. To calculate the option prices we need to find an equivalent martingale measure. That is, we have to find a measure Q such that S^* is a martingale with respect to Q. For this we define the following Radon-Nikodym density:

$$\xi_t = \exp\left(-\frac{1}{2}\left(\frac{\eta-\delta}{\sigma}\right)^2 t - \frac{\eta-\delta}{\sigma} W(t)\right) \qquad \text{for all } t \in [0,T].$$

Exercise 7.3.5. Prove the following statements:

1. $E[\xi_t] = 1$,
2. $Var[\xi_t] = \exp\left(\left(\frac{\eta-\delta}{\sigma}\right)^2 t\right) - 1$,
3. $\xi_t > 0$.

(Hint: $W(t) \sim \mathcal{N}(0,t)$.)

An application of Girsanov's theorem – a theorem in the theory of stochastic integration (e.g. [Pro90] Theorem 3.6.21) – shows that

$$\hat{W}_t = W(t) + \frac{\eta-\delta}{\sigma} t$$

is a Brownian motion with respect to $Q = \xi \cdot P$.

Naturally, after this transformation we want to prove that

$$S^*(t) = S(0) \exp\left(-\frac{1}{2}\sigma^2 t + \sigma \hat{W}(t)\right)$$

is a martingale with respect to Q. (Then the price of the option is given by its expectation with respect to Q.)

Proof. We have to show the equality

$$E^Q\left[S^*(u)|\mathcal{F}_t\right] = S^*(t)$$

for $t, u \in \mathbb{R}, u > t$. With the notation $u = t + \Delta t$, $W_u = W_t + \Delta W$ and $Z \sim \mathcal{N}(0,1)$ we have

$$\begin{aligned}
&E^Q\left[S^*(u)|\mathcal{F}_t\right] \\
&\quad = E^Q\left[S(0) \exp\left(-\frac{1}{2}\sigma^2 t + \sigma \hat{W}(t) + (-\frac{1}{2}\sigma^2 \Delta t + \sigma \Delta\hat{W})\right)|\mathcal{F}_t\right] \\
&\quad = S(0) \exp\left(-\frac{1}{2}\sigma^2 t + \sigma W(t)\right) E^Q\left[\exp\left(-\frac{1}{2}\sigma^2 \Delta t + \sigma\sqrt{\Delta t} Z\right)|\mathcal{F}_t\right] \\
&\quad = S^*(t).
\end{aligned}$$

Therefore the measure Q is equivalent to P, and S^* is a martingale with respect to Q. An economist would say, "it exists (at least) one consistent price system".

Theorem 7.3.6. *Let the economic model defined above be given, i.e. it is defined by $(\Omega, \mathcal{A}, P)$, S and B. Then at time t the price of a policy with death benefit $C(T)$ is*

$$\pi_t(T) = E^Q\left[\exp\left(-\delta(T-t)\right) C(T)|\mathcal{F}_t\right].$$

Remark 7.3.7. The main difference of this model in comparison to the classical model is the fact that one has to calculate the expectation with respect to Q and not with respect to P. Moreover one should note, that we have not proved the uniqueness of the price system.

The following two formulas are an important consequence of the previous considerations.

Theorem 7.3.8. *A single premium for a policy based on the economic model defined above is given by the following formulas.*

Endowment policy:

$$V(0) = E^Q\left[\exp(-\delta T)\, C(T)\right] \cdot {}_Tp_x.$$

Term life insurance:

$$V(0) = \int_0^T E^Q\left[\exp(-\delta t)\, C(t)\right] p_{**}(x, x+t)\, \mu_{*\dagger}(x+t) dt.$$

7.4 Calculation of single premiums

Up to now the calculations have been relatively simple, since we did not include any guarantees in our policy model. Next, we will consider a unit-linked policy with an additional guarantee. We recall some of the notations from the previous sections:

$N(\tau)$	Number of shares at time τ,
$S(\tau)$	value of a share at time τ,
$G(\tau)$	guaranteed benefits at time τ,
$C(\tau) = \max\{N(\tau)S(\tau), G(\tau)\}$	value of the insurance at time τ.

7.4.1 Pure endowment policy

Theorem 7.4.1. *Let the Black-Scholes model be given. Then the single premium for a pure endowment policy with payout*

$$C(T) = \max\{N(T)S(T), G(T)\}$$

is given by

$${}_TG_x = {}_Tp_x \left[G(T)\, \exp(-\delta T) \Phi(-d_2^0(T)) + S(0) N(T) \Phi(d_1^0(T))\right],$$

where

$$\begin{aligned}
\Phi(y) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \exp(-\frac{x^2}{2}) dx, \\
d_1^t(s) &= \frac{\ln\left[\frac{N(s)S(t)}{G(s)}\right] + \left(\delta + \frac{1}{2}\sigma^2\right)(s-t)}{\sigma\sqrt{s-t}}, (s > t), \\
d_2^t(s) &= \frac{\ln\left[\frac{N(s)S(t)}{G(s)}\right] + \left(\delta - \frac{1}{2}\sigma^2\right)(s-t)}{\sigma\sqrt{s-t}}, (s > t).
\end{aligned}$$

Proof. In the following we denote by J^* the discounted value of a random variable J. The value of the pure endowment policy at time zero is $E^Q[C^*(T)]$. We set $Z = S^*(T)$. Then the following equations hold

$${}_TG_x = {}_Tp_x\, E^Q\left[\max\{N(T)Z, G^*(T)\}\right]$$

and

$$Z = S(0)\, \exp\left(-\frac{1}{2}\sigma^2 T + \sigma \hat{W}(T)\right) \qquad \text{where} \qquad \hat{W}(T) \sim \mathcal{N}(0,T).$$

Thus we get

$$\begin{aligned} {}_TG_x &= {}_Tp_x \int_{-\infty}^{\infty} \max\left[N(T)S(0)\exp(-\frac{1}{2}\sigma^2 T + \sigma\,\xi), G^*(T)\right] f(\xi)d\xi, \\ f(\xi) &= \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{1}{2T}\xi^2\right). \end{aligned}$$

Next we define $\bar{\xi} = \frac{1}{\sigma}\left[\ln\left(\frac{G^*(T)}{N(T)S(0)}\right) + \frac{1}{2}\sigma^2 T\right]$ and note that $\xi > \bar{\xi}$ implies $N(T)Z > G^*(T)$. Therefore, the single premium is given by

$$\begin{aligned} {}_TG_x &= {}_Tp_x \Bigg(G^*(T) \int_{-\infty}^{\bar{\xi}} f(\xi)d\xi \\ &\qquad + N(T)S(0) \int_{\bar{\xi}}^{\infty} \exp(-\frac{1}{2}\sigma^2 T + \sigma\,\xi) f(\xi)d\xi \Bigg) \\ &= {}_Tp_x \Bigg(G^*(T) \int_{-\infty}^{\bar{\xi}} f(\xi)d\xi \\ &\qquad + N(T)S(0) \int_{\bar{\xi}}^{\infty} \frac{1}{\sqrt{2\pi T}} \exp(-\frac{1}{2T}(\xi - \sigma T)^2 d\xi \Bigg). \end{aligned}$$

This equation, with adapted notation, yields the statement of the theorem.

7.4.2 Term life insurance

Theorem 7.4.2. *Let the Black-Scholes model be given. Then the single premium for a term life insurance with death benefit*

$$C(t) = \max\{N(t)S(t), G(t)\}$$

is given by

$$G^1_{x:T} = \int_0^T \left(G(t)\exp(-\delta t)\Phi(-d_2^0(t)) + S(0)N(t)\Phi(d_1^0(t))\,{}_tp_x\mu_{x+t}\right) dt,$$

where

$$\begin{aligned} \Phi(y) &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx, \\ d_1^t(s) &= \frac{\ln\left[\frac{N(s)S(t)}{G(s)}\right] + \left(\delta + \frac{1}{2}\sigma^2\right)(s-t)}{\sigma\sqrt{s-t}}, \\ d_2^t(s) &= \frac{\ln\left[\frac{N(s)S(t)}{G(s)}\right] + \left(\delta - \frac{1}{2}\sigma^2\right)(s-t)}{\sigma\sqrt{s-t}}, \end{aligned}$$

for $s > t$.

Exercise 7.4.3. Prove the previous theorem by the same methods which we used for the pure endowment policy.

7.5 Thiele's differential equation

Now we want to derive Thiele's differential equation. For this we need to determine premiums for the policies. We introduce the notation $\bar{p}(t)$ for the density of the premiums at time t. Then the equivalence principle yields the following two equations:

$$ {}_T G_x = \int_0^T \bar{p}(t) \exp(-\delta\, t)\, {}_t p_x dt $$

and

$$ G^1_{x:T} = \int_0^T \bar{p}(t) \exp(-\delta\, t)\, {}_t p_x dt. $$

Also in this section the pure endowment policy and the term life insurance will be considered separately. The mathematical reserve for these policies is given by:

$$ \begin{aligned} \text{Pure endowment:} \quad V(t) = \;& {}_{T-t}p_{x+t} \pi_t(T) \\ & - \int_t^T \bar{p}(\xi) \exp(-\delta(\xi - t))\, {}_{\xi-t}p_{x+t}\, d\xi. \\ \text{Life insurance:} \quad V(t) = \;& \int_t^T \left(\pi_t(\xi)\mu_{x+\xi} - \bar{p}(\xi) \exp(-\delta(\xi - t))\right) \\ & \times {}_{\xi-t}p_{x+t}\, d\xi, \end{aligned} $$

where

$$ \begin{aligned} \pi_t(s) = \;& G(s) \exp(-\delta(s-t)) \Phi(-d_2^t(s)) \\ & + N(s) S(t) \Phi(d_1^t(s)), \\ d_1^t(s) = \;& \frac{\ln\left[\frac{N(s)S(t)}{G(s)}\right] + \left(\delta + \frac{1}{2}\sigma^2\right)(s-t)}{\sigma\sqrt{s-t}}, \\ d_2^t(s) = \;& \frac{\ln\left[\frac{N(s)S(t)}{G(s)}\right] + \left(\delta - \frac{1}{2}\sigma^2\right)(s-t)}{\sigma\sqrt{s-t}}, \end{aligned} $$

for $s > t$.

Remark 7.5.1. – In the classical setting the reserves were deterministic, but here they depend on the underlying share S.

– Note that we are beyond the deterministic theory of differential equations. In particular we have to use Itô's formula, which takes the following form for the purely continuous case of a standard Brownian motion W:

$$ df(W) = f'\, dW + \frac{1}{2} f''\, ds. $$

For the policies defined above we have the following theorem.

Theorem 7.5.2. *1. The differential equation for the price of a pure endowment policy is:*

$$ \frac{\partial V}{\partial t} = \bar{p}(t) + (\mu_{x+t} + \delta)\, V(t) - \frac{1}{2}\sigma^2 S(t)^2 \frac{\partial^2 V}{\partial S^2} - \delta\, S(t) \frac{\partial V}{\partial S}. $$

2. *The differential equation for the price of a term life insurance is:*

$$\frac{\partial V}{\partial t} = \bar{p}(t) + (\mu_{x+t} + \delta)\, V(t) - C(t)\mu_{x+t} - \frac{1}{2}\sigma^2 S(t)^2 \frac{\partial^2 V}{\partial S^2} - \delta\, S(t) \frac{\partial V}{\partial S}.$$

Before we prove this theorem, we want to make some comments on the formulas.

Remark 7.5.3. 1. One obtains Black-Scholes formula by setting $\mu_{x+t} = \bar{p}(t) = 0\, \forall t$.

2. The first terms in the differential equations in the theorem above coincide with the classical case, i.e. the dependence of the values on the premiums, on the mortality and on the interest rate. Due to the shares in the model a further term appears: $-\frac{1}{2}\sigma^2 S(t)^2 \frac{\partial^2 V}{\partial S^2} - \delta\, S(t)\frac{\partial V}{\partial S}$. It represents the fluctuations of the underlying shares.

Proof. We have

$$\pi_t^*(T) = \exp(-\delta t)\pi_t(T).$$

Hence, by the definition of V, we get

$$V(t) = {}_{T-t}p_{x+t}\pi_t^*(T)\exp(\delta t) - \int_t^T \bar{p}(\xi)\exp(-\delta(\xi - t))_{\xi-t}p_{x+t} d\xi$$

and

$$\pi_t^*(T) = \Psi(t)\left[V(t) + \int_t^T \bar{p}(\xi)\exp(-\delta(\xi - t))_{\xi-t}p_{x+t} d\xi\right],$$

where

$$\Psi(t) = \frac{\exp(-\delta t)}{{}_{T-t}p_{x+t}}.$$

Now we can apply Itô's formula to the function $\pi_t^*(t, S)$, since π_t^* is a function of S and t. We get

$$\begin{aligned} dY_t &= U(t + dt, X_t + dX_t) - U(t, X_t) \\ &= \left(U_t dt + \frac{1}{2}U_{xx}b^2 dt\right) + U_x dX_t \\ &= \left(U_t + \frac{1}{2}U_{xx}b^2\right) dt + U_x\, b\, dB_t \end{aligned}$$

and

$$d\pi^* = \left(\frac{\partial \pi^*}{\partial t} + \frac{\partial \pi^*}{\partial S}\, a + \frac{1}{2}\,\frac{\partial^2 \pi^*}{\partial S^2}\, b^2\right) dt + \frac{\partial \pi^*}{\partial S} b\, d\hat{W}.$$

Furthermore we know that

$$dS = \delta S(t) dt + \sigma S(t) d\hat{W},$$

and thus we have $a = \delta S(t)$ and $b = \sigma S(t)$. In the next step we want to determine the two terms:

$$\begin{aligned} \frac{\partial \pi_t^*}{\partial S} &= \Psi(t)\, \frac{\partial V}{\partial S}, \\ \frac{\partial^2 \pi_t^*}{\partial S^2} &= \Psi(t)\, \frac{\partial^2 V}{\partial S^2}. \end{aligned}$$

To get $\frac{\partial \pi^*}{\partial t}$, we start with

$$\begin{aligned}
\frac{\partial}{\partial t} {}_{\xi-t}p_{x+t} &= \mu_{x+t}\, {}_{\xi-t}p_{x+t}, \\
\frac{\partial}{\partial t}\Psi(t) &= \left(\frac{A}{B}\right)' = \frac{A'}{B} - \frac{A}{B^2}\, B' \\
&= -\left(\mu_{x+t} + \delta\right) \Psi(t).
\end{aligned}$$

Now, with the formula from above we get

$$\begin{aligned}
\frac{\partial \pi^*}{\partial t} &= \frac{\partial \Psi}{\partial t}\left(V(t) + \int_t^T \bar{p}(\xi)\exp(-\delta(\xi - t))_{\xi-t}p_{x+t}dt\right) \\
&+ \Psi(t)\left(\frac{\partial V}{\partial t} + \frac{\partial}{\partial t}\int_t^T \bar{p}(\xi)\exp(-\delta(\xi - t))_{\xi-t}p_{x+t}dt\right) \\
&= \Psi(t)\left(\frac{\partial V}{\partial t} - \left(\mu_{x+t} + \delta\right) V(t) - \bar{p}(t)\right),
\end{aligned}$$

where we applied the chain rule to

$$\frac{\partial}{\partial t}\int_t^T \bar{p}(\xi)\exp(-\delta(\xi - t))_{\xi-t}p_{x+t}dt.$$

Thus we get

$$\begin{aligned}
\pi_s^*(T) &= \pi_t^*(T) + \int_t^s \Psi(\xi)\frac{\partial V}{\partial S}\sigma S d\hat{W}(\xi) \\
&+ \int_t^s \Psi(\xi)\left[\frac{\partial V}{\partial S}\delta S + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - (\mu_{x+\xi} + \delta)V(\xi)\right. \\
&\qquad\qquad \left. + \frac{\partial V}{\partial t}(\xi)\bar{p}(\xi)\right] d\xi.
\end{aligned}$$

Now the drift term is equal to zero, since $\pi_\cdot^*(T)$ is a martingale. Therefore we finally get

$$\frac{\partial V}{\partial t} = \bar{p}(t) + \left(\mu_{x+t} + \delta\right) V(t) - \frac{1}{2}\sigma^2 S(t)^2 \frac{\partial^2 V}{\partial S^2} - \delta\, S(t)\frac{\partial V}{\partial S}.$$

Exercise 7.5.4. Prove the second part of the theorem above.

7.6 Variable Annuities (VA)

This section provides a short introduction into variable annuities, their risks and the way financial risks are hedged. A variable annuity is a special type of unit linked contract, with additional financial guarantees, such as a minimal performance guarantee, a guaranteed minimal death benefit etc. Hence the policyholder does not only buy the underlying funds, but also option like guarantees. These latter ones are the main characteristics of variable annuities and are the reason for a part of their complexity regarding the pricing and hedging of these long term guarantees. The underlying product features will be described in some greater detail in section 7.6.1.

In order to price and hedge these guarantees the concepts of arbitrage free pricing theory ("Black-Scholes") together with the so called Itô-Calculus are used to immunise and hedge the liabilities as much as possible with a dynamic hedging strategy. The value of the guarantees depends on financial market variables such as equity market, volatility and interest rates. It also depends on mortality and the various parameters regarding the policyholder behaviour. Since one aims to hedge the market variables, the values of the liabilities are represented in terms of a (Taylor) approximation with respect to the relevant variables such as equity returns, interest rate movement and volatility of the funds. This means that changes in value to each of the underlying quantities is approximated. These approximations of the liability values (and also assets) are called "Greeks".

To understand how these options are synthetically "constructed" one needs to understand the concept of a replicating portfolio. Hence one holds at every point in time a portfolio P_t with the aim that this portfolio matches at time T just the payout of the option mentioned above. In order to construct such portfolios one usually uses the "greeks". These greek letters represent the sensitivity of an option in case of a change to the underlying economic parameters such as equity price, interest rate levels, etc. We have the following relationships:

$$\begin{aligned}
\Delta_P &= \frac{\partial P}{\partial S} \\
&= \Phi(d_1), \\
\Gamma &= \frac{\partial^2 P}{\partial S^2} \\
&= \frac{\Phi'(d_1)}{S \times \sigma \times \sqrt{T}}, \\
\nu &= \frac{\partial P}{\partial \sigma} \\
&= S \times \Phi'(d_1) \times \sqrt{T-t}, \\
\rho &= \frac{\partial P}{\partial r} \\
&= -(T-t) \times K \times e^{-r \times (T-t)} \times \Phi(-d_2).
\end{aligned}$$

All these partial derivatives serve to approximate the change in value of the underlying derivatives when the fundamental parameters change. To do this one uses Itô-Formula for a given SDE such as

$$dX_t = \mu_t \, dt + \sigma_t \, dB_t,$$

follows:

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} dB_t,$$

where we clearly see the various greeks. In the concrete context the respective approximations looks as follows:

$$\begin{aligned}
\Delta f(S,t,r,\sigma) &= \underbrace{\frac{\partial f}{\partial S}}_{Delta} \Delta S + \frac{1}{2} \underbrace{\frac{\partial^2 f}{\partial S^2}}_{Gamma} (\Delta S)^2 + \underbrace{\frac{\partial f}{\partial t}}_{Theta} \Delta t \\
&\quad + \underbrace{\frac{\partial f}{\partial r}}_{Rho} \Delta r + \underbrace{\frac{\partial f}{\partial \sigma}}_{Vega} \Delta \sigma + \ldots
\end{aligned}$$

Note that this is the based for the hedging program on one side and allows to also determine the hedging error as a consequence of the higher order terms on the other.

When using assets with the same underlying value and Greeks, it means that the change in economic equity moves in parallel, thereby reducing the risk up to the higher order errors of both assets and liabilities. The concept of dynamic hedging foresees the updating of the underlying liability values and Greeks and the corresponding re-balancing of the assets in continuous time in order to result in an optimal hedge. In reality the recalculation of the liability values and the re-balancing is done on a less frequent basis, resulting in a tracking or hedging error. There is also a basis risk, when individual assets behave differently from the index chosen for hedging. The following example shows how this concept works in practice: Assume that our portfolio has the following characteristics:

Quantity	M $	Meaning
Value of Guarantee π	Net: 4526	The economic value of the liability.
δ (1%)	-85	The "Delta" is the sensitivity of the value of the liability with respect to the change of the underlying equity index. If the equity market would fall by say 2%, the value of the guarantee would increase by approximately $-2 \times -85 = 170$ M $.
ρ (1bp)	-23.6	The "Rho" is the sensitivity of the insurance liability with respect to interest rate movements. Assume for example that the interest rates would increase by 0.2%. In this case the value of the guarantee would decrease by approximately $20 \times -23.6 = -472$ M $.
ν (1%)	240	The "Vega" is the sensitivity with respect to the volatility of the equity index. If volatility in the market would increase from say 18% to 20%, this would mean that the value of the guarantee would increase by $2 \times 240 = 480$ M $.

As indicated above a dynamic hedging program aims to reduce/minimize the changes of the (economic) value of the guarantee by buying assets with offsetting characteristics. In order to do so the company, in a first instance decides which of the "Greeks" it wants to hedge. Some companies do not hedge anything, others use a 3-Greeks hedging programme. We need to see what a hedge could look like and what the corresponding hedge error would be. We assume for example (as above) a fall in equities of 2%, an increase in interest rates of 0.2% and increase in volatility of 2%. Then we get the following:

Quantity	Liabilities Sensitivity	Assets Sensitivity	Liabilities Value	Assets Value	Net Change in Equity
Value of Guarantee/Assets BoP π			4526	0	–
δ (Change -2%)	-85	-53	+170	+106	-64
ρ (Change +20bp)	-23.6	-0.3	-472	-6	+464
ν (Change +2%)	240	19	+480	+ 38	-442
Value of Guarantee/Assets EoP			4704	138	**-40**

This means that the hedge loss would equal to $(4526 - 4704) + (138 - 0) = -40$. In absence of a hedging the corresponding loss would equal to $(4526 - 4704) = -181$ and hence hedging reduced the loss in this example by 78% $(= 100\% \times (1 - \frac{40}{181}))$. Figure 9.3 shows how the change of the value of the liabilities is approximated linearly by its tangent.

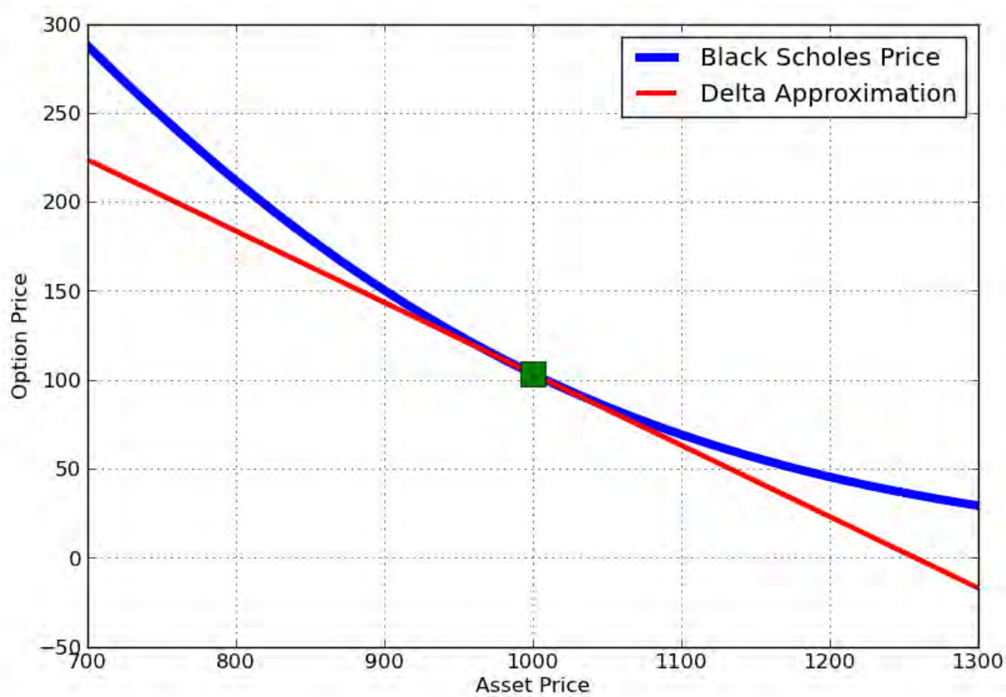

Figure 7.5. δ–Hedging

By now we see that we need to somehow calculate the value of these liabilities together with their corresponding Greeks in order to set up a dynamic hedging strategy. For simple forms of liabilities the underlying calculations can be performed by closed form formulas. In the concrete set up, the liabilities are more complex and hence the valuation and the calculation of the Greeks are done via simulation ("Monte

Carlo") and depends on the Economic Scenario Generator used. The expected values are estimated in the simulation context by taking suitable averages and the partial derivatives are estimated with a difference quotient with a suitable small shock.

Next, we need to understand the various types of guarantees offered under the umbrella of variable annuity policies. There are several types of performance guarantees for unit linked policies and one may often choose them a la carte, with higher risk charges for guarantees that are riskier for the insurance company. The first type is comprised of guaranteed minimum death benefits (GMDB), which can be received only if the owner of the contract dies. GMDBs come in various forms:

– Return of premium (a guarantee that you will not have a negative return),
– Roll-up of premium at a particular rate (a guarantee that you will achieve a minimum rate of return, greater than 0),
– Maximum anniversary value (looks back at account value on the anniversaries, and guarantees you will get at least as much as the highest values upon death),
– Greater of maximum anniversary value or particular roll-up.

Besides death benefits, which the contract holder generally can't time, there are guaranteed living benefits (GLB), which pose significant risks for insurance companies as contract holders will likely exercise these benefits when they are worth the most. Annuities with guaranteed living benefits (GLBs) tend to have higher fees commensurate with the additional risks underwritten by the issuing insurer. Some GLB examples, in no particular order:

– Guaranteed Minimum Income Benefit (GMIB, a guarantee that one will get a minimum income stream upon annuitisation at a particular point in the future)
– Guaranteed Minimum Accumulation Benefit (GMAB, a guarantee that the account value will be of a certain amount at a certain point in the future)
– Guaranteed Minimum Withdrawal Benefit (GMWB, a guarantee similar to the income benefit, but one that doesn't require annuitising, and where one can withdraw flexibly a maximal amount per calender year until the "Guarnteed Withdrawal Balance" (GWB) is nil.)
– Guaranteed-for-life Income Benefit (This is an enriched version of a GMWB, where the minimal annual withdrawal can also continue when the GWB is depleted until the death of the insured person.)

7.6.1 Product Design

Overall the products need to be understood in the context that there are two phases (figure 7.6):

1. Accumulation phase without withdrawal, and
2. The phase of withdrawal until the person dies during which the policyholder still has the ability to invest in risky assets.

Before explaining the various variants in detail, it makes sense to have a closer look at figure 7.6. In the concrete set up we are looking at a "virtual" product where the policyholder buys a VA at age 60 and invests 100'000 $ in a 100 % equity funds. For the first 5 years (until age 65) there are no additions or withdrawals from this fund and the value changes as a consequence of the market movement. At age 65 the person starts to withdraw (according to the contract) up to 5 % of the maximum of the fund value

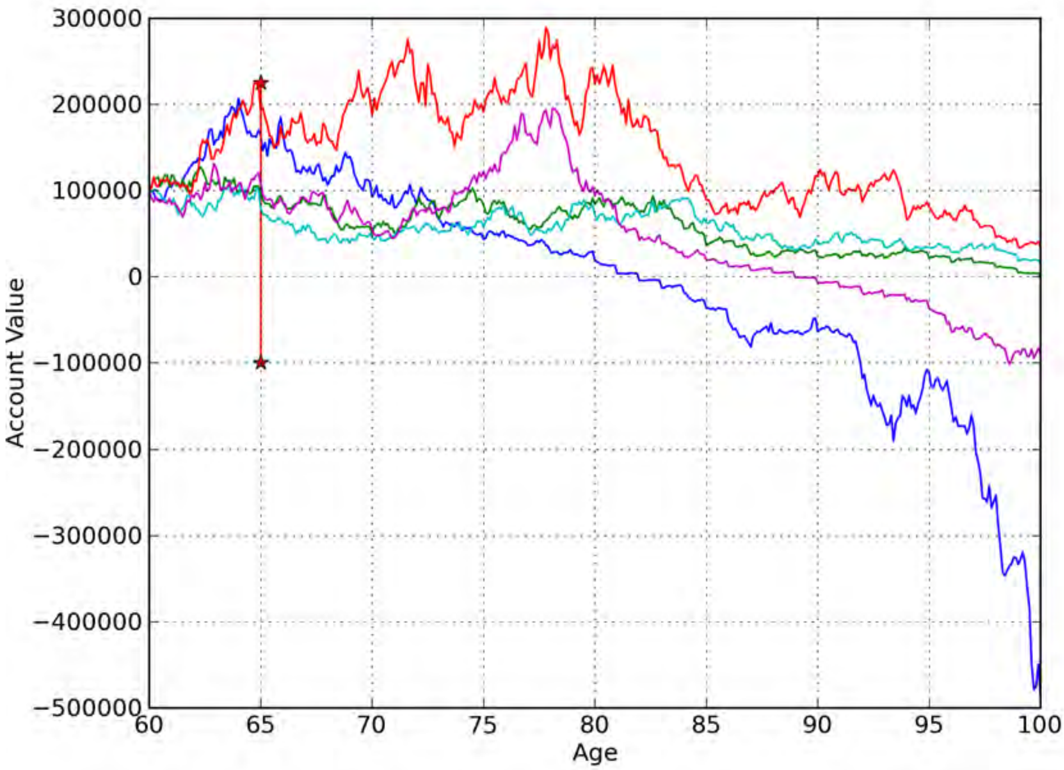

Figure 7.6. Phases of a VA product for a GLB

at that time and the original amount. Hence, depending on the fund's past performance he has the right to withdraw no more than the minimally guaranteed 5'000 $ in case of an adverse fund performance. Depending on the fund performance there is now a risk for the insurance company since the fund, still investing into equities might run out of money and the company has to carry on paying the respective amounts.

For illustratory purposes we have assumed in the example that the insured person lives until the age of 100, which is obviously high. In reality the random life span of the policyholder makes this product still more complex since mutualisation between policyholders enters. It is, however clear that the risk for the insurance company is an increased life span and hence a longer consumption period.

There are two types of complexities intrinsic to these variable annuity products, namely the many possible options which can be granted and sold to the policyholder and hence the understanding of the different features. It becomes obvious from figure 7.6 that market developments could lead to losses for the insurance company. The technique to mitigate these losses and to replicate these guarantees is to set up a corresponding hedging programme.

GMDB. The Guaranteed Minimal Death Benefit is to some degree the easiest option to each VA. The aim of this benefit is to ensure a minimal death benefit in case the policyholder dies. Assume, for example, that the policyholder has invested 100'000 $ at age 60 and dies at age 80. We assume also that the fund value has dropped to 50'000 $. The easiest GMDB is to guarantee a minimal fixed amount in case of

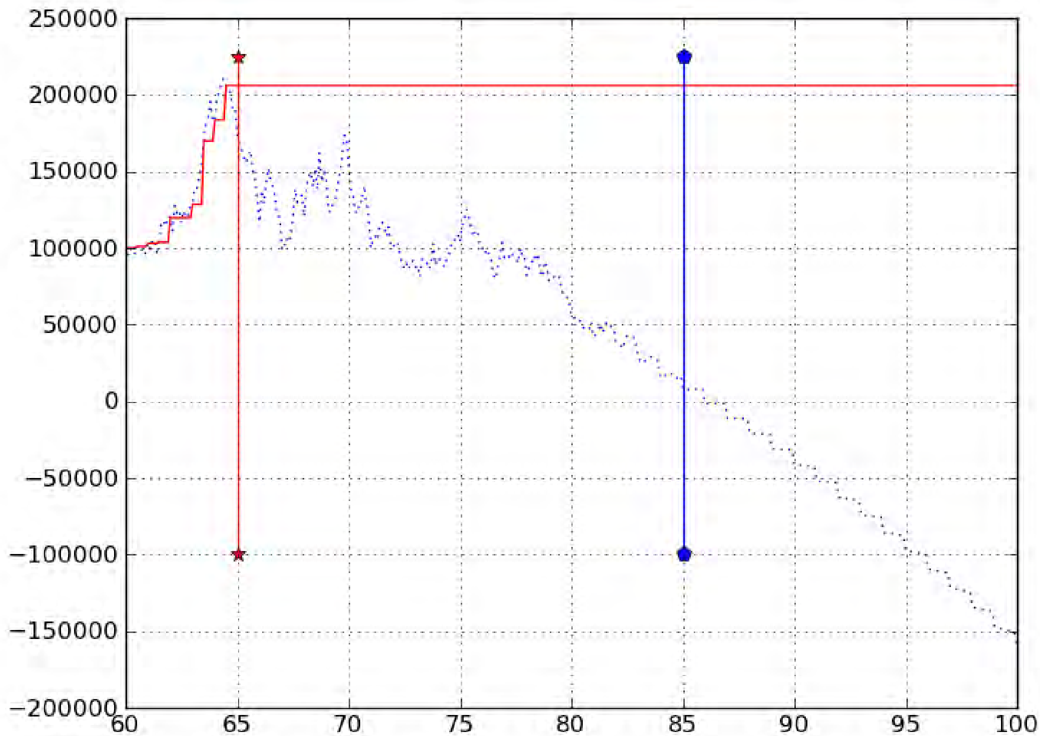

Figure 7.7. Sum at Risk for a GMDB trajectory

death, say 100'000 \$ for arguments sake. In figure 7.7 we assume that there is a ratchet until age 65, and that the fund value has ratched up to 220'000 \$. In this case the loss to the insurance company is 170'000 \$ (= 220000 − 50000). From a financial point of view the insurance company has written a put option at age 80 with a strike price of 220'000 \$. In consequence one can consider a GMDB portfolio as put option portfolio, where the strike prices correspond to the respective death benefits at a certain time and where the amount of the put is weighted according to the expected number of people dying at this time.

GMWB. Considering a GMWB written for a single premium, this process involves the random movement of the funds. Figure 7.6 shows this, assuming that the insured person dies at age 100. In a next step the trajectories with losses need to be weighted with the corresponding people alive who have not surrendered at a certain point in time. Figure 7.8 shows this with one sample trajectory. In the concrete case the company has to honour the withdrawal payments because the funds is depleted at age 86.

After having looked at this sample trajectory one needs to look at the expected cash flow profile stemming from the guarantee. As before we start with a 60 year old person with a deferral period of 5 years (hence starting to withdraw at age 65 at 5% p.a.) with a income phase starting at age 85. We furthermore assume a single premium of USD 100'000. Figure 7.9 shows this. The red curve illustrates the expected losses and compares them with a guaranteed (fixed) annuity where the money is invested risk free. Two things become obvious:

– The guarantee starts to bite in most cases after some 25 years when the people are 85+ years old.

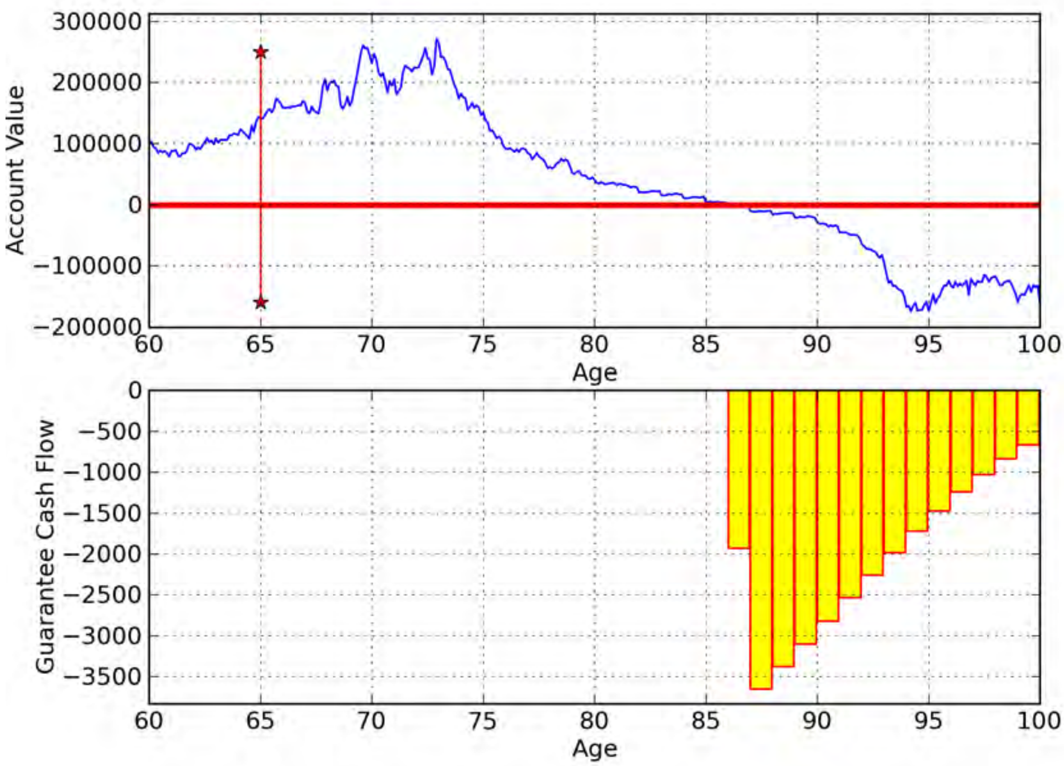

Figure 7.8. Guarantee Cashflows for one sample trajectory

– The percentage of the guarantee (on the lower part of the figure) increases with increasing age and exceeds 20% at age 80.
– In average this guarantee equals about 14 % of the funds and has a rather long duration (of 26).

From a formal point of view the following quantities are important:

Funds Value: $F(t)$ denotes the funds value,

GAWA: $GAWA(t)$ is the maximal amount which can be consumed under the GMWB policy,

GWA: $GWA(t)$ is the maximal amount which can be withdrawn per year without needing a partial surrender.

It is worth noting that both quantities $GAWA$ and GWA depend on the time and (potentially) also on the funds performance, in the sense that an additional bonus could be allocated to the funds. Each time a withdrawal is made, the $GAWA$ is reduced by this amount. Depending on the contract specificities, the guarantee consists of being able to continue withdrawing (until a certain point in time or until death) even when the maximal amount is exceeded.

The following aspects are worth noting:

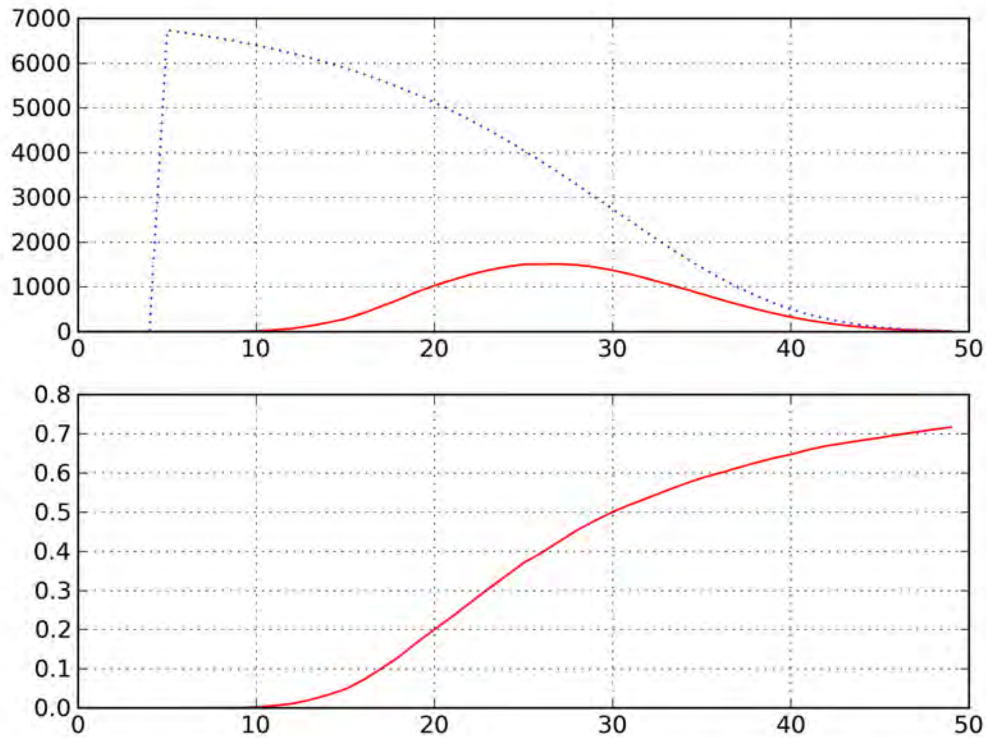

Figure 7.9. Expected P/L

- The initial $GAWA$ and GWA are a function of the funds value at the first time the policyholder decides to withdraw money. This particularly means that policyholder behaviour plays an important role.
- In some instances the $GAWA$ has also a ratchet feature, which could mean that the $GAWA$ is locked in at each quarter end and can not fall thereafter (before withdrawing). Normally the GWA is then expressed as a fixed percentage of the $GAWA$.

Formulae:

In the following section I summarise the economics of this sort of contract. Assume the following:

- Person aged x_0 purchases such a GMWB and pays a single premium EE;
- Assume that the person starts to withdraw at age sw and that the income phase starts at age s;
- We use the following notation:
  - $F(t)$: Funds value at time t. To be more precise we denote with $F(t)^-$ and $F(t)^+$ the value of the funds before and after withdrawal of the annuity, respectively. In consequence we have $F(t)^+ = F(t)^- - R(t)$.

– $GWB(t)$: GWB ("Guaranteed Withdrawal Balance ") value at time t. To be more precise we denote with $GWB(t)^-$ and $GWB(t)^+$ the value of the funds before and after withdrawal of the annuity, respectively. In consequence we have $GWB(t)^+ = GWB(t)^- - R(t)$.

– $f(x)$: GAWA percentage if person starts to withdraw at age x;

– $GAWA(t)$: maximal allowable withdrawal benefit (usually $= GWB \times f(x)$).

– $R(t)$: actual amount withdrawn. Note that we have the following: $R(\xi) = 0$ for $\xi < sw$, $0 \leq R(\xi) \leq GAWA(x)$, for all $\xi \in [sw, s[$, and $R(\xi) = GAWA(x)$, for all $\xi \geq s$, assuming that the "for life option" is in place and identifying x and t in the obvious way, eg $x(t) = t - t_0 + x_0$

– With $\eta(t, \tau) \in \mathbb{R}$, we denote the fund performance during the time interval $[t, \tau]$, with $\tau > t$.

– We do not allow for changes in funds and lapses at this time and also do not consider the death of the person insured. This would add some complexity, where actually the annuities need to be weighted with the respective probabilities ${}_tp_x$ and in the same sense the respective death cover weighted with ${}_tp_x\, q_x$. For the moment assume that $\sum \tau \geq 0$ stands for "until death".

– By $X(t)$ we denote the loss at time t occurring from GMWB guarantees. It is obvious that under these premises the value of the total guarantee $Y = \sum_{\tau \geq 0} (1 + r(\tau))^{-\tau} X(\tau)$, where $r(\tau)$ represents the risk free interest between $[0, \tau]$.

As described above we assume for example the following:

$$f(x) = \begin{cases} 5\% & \text{if } x \in [55, 74], \\ 6\% & \text{if } x \in [75, 84], \\ 7\% & \text{if } x \geq 85. \end{cases}$$

For the recursion we have at time $t_0 = 0$

$$\begin{aligned} F(0) &= EE, \\ GWB(0) &= EE, \\ GAWA(0) &= f(sw) \times GWB(0). \end{aligned}$$

Afterwards from time $t - 1 \rightsquigarrow t$ we have the following, assuming that the withdrawals takes place before the end of the first quarter, once a year.

$$\begin{aligned} F(t)^- &= (1 + \eta(t-1, t)) \times F(t-1)^+, \\ GWB(t)^- &= \max\{GWB(t-1)^+, \max_{k=0,1,\ldots,4} \{1 + \eta(t-1, t-1+\frac{k}{4})\} \times F(t-1)^+\}, \\ GAWA(t) &= \max(GAWA(t-1), f(sw) \times GWB(t)^-), \\ F(t)^+ &= \max(0, F(t)^- - R(t)), \\ GWB(t)^+ &= \max(0, GWB(t)^- - R(t)), \\ X(t) &= \max(0, R(t) - F(t)^-), \\ \pi(Y) &= E^Q \left[\sum_{\tau \geq 0} (1 + r(\tau))^{-\tau} X(\tau) \right]. \end{aligned}$$

If we now pick a given, mortality cover – the simplest one – namely the payment of $GWB(t)$ in case of death, we can calculate the value of the insurance option as follows:

$$\pi(Y) \quad = \quad E^Q\left[\sum_{\tau\geq 0}^{\infty}(1+r(\tau))^{-\tau}X(\tau)\times {}_\tau p_{x_0}\right],$$

where we assume that no guarantee fee is charged. The inclusion of a guarantee fee can be viewed as an additional annuity which consumes the funds, but only until its depletion.

Next we look at a concrete example, once with and once without ratchet guarantee (aka "step-up"). Figures 7.10 and 7.11 show the way the guarantee manifests in the case with and without the ratchet option. The figure 7.10 plots the quantiles (with respect to a risk neutral simulation) of the guarantees X. Moreover figure 7.11 shows the additional guarantee fee (green) and the GMDB payment (blue). Note that all guarantee payments occur in later years as a consequence that the ratchet does not affect the funds value, but rather the (virtual) GWB account. Only when the funds run out of money the guarantee starts to bite.

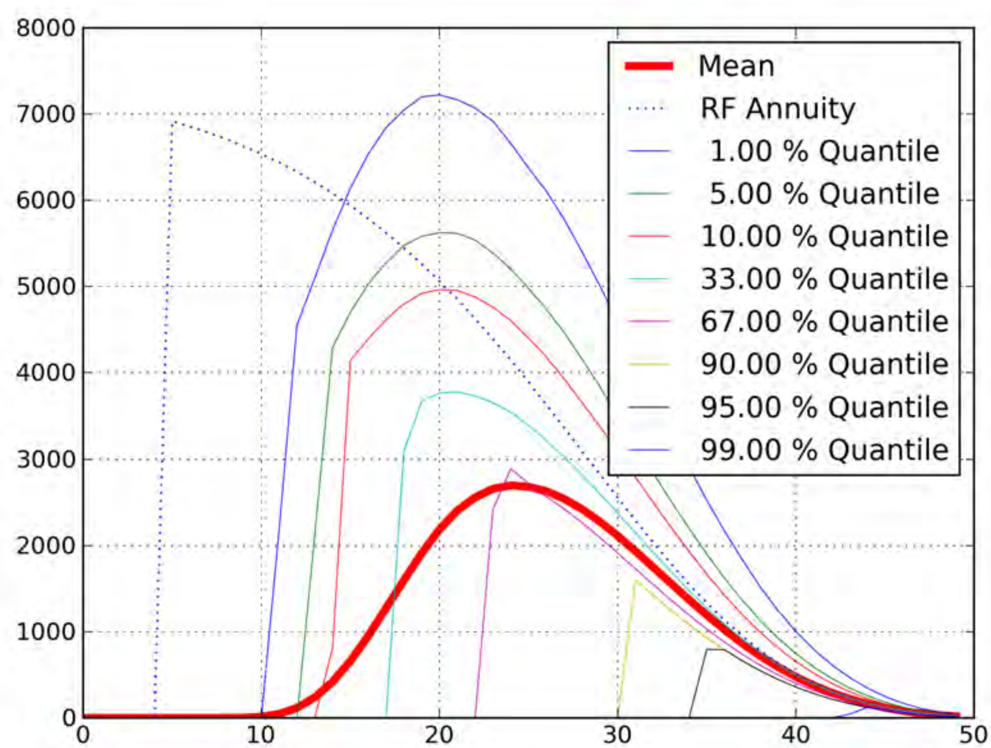

Figure 7.10. Guarantees for a GMWB with ratchet; x: time; y: payout of respective option

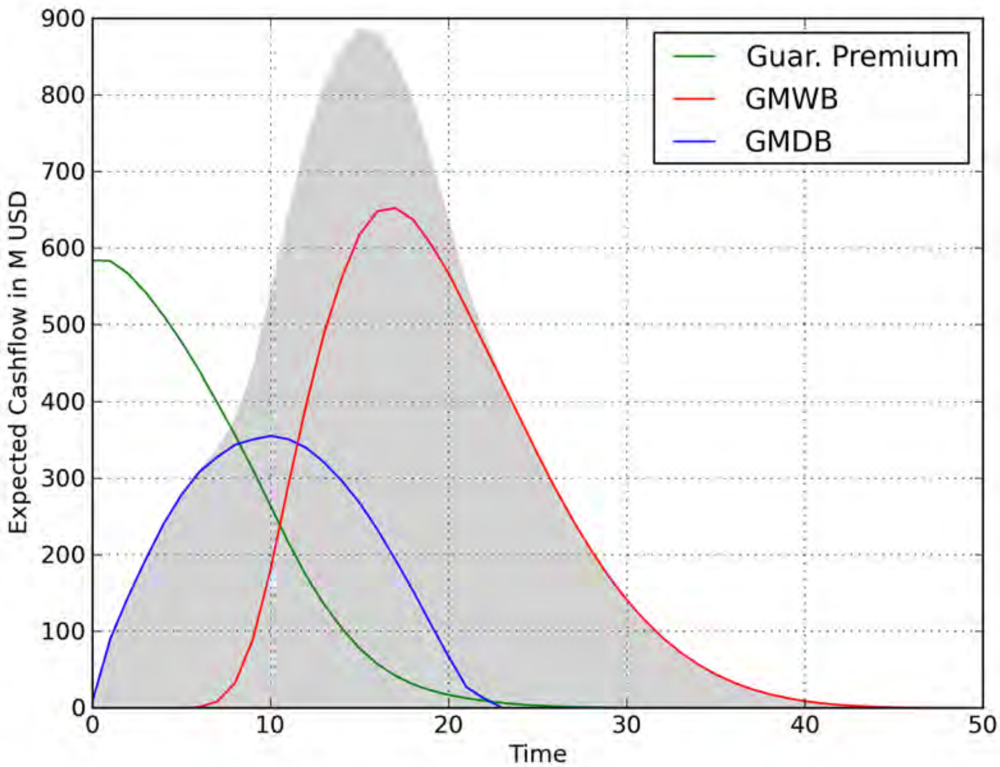

Figure 7.11. Guarantees for a GMWB without ratchet; x: time; y: payout of respective option

7.6.2 Hedge Attribution Analysis

The following table provides a structure of a possible hedge attribution analysis. Such an analysis could for example be performed monthly.

	Position M $	Assets	Liabilities	Net
1	**MV BoP**	**xxx.x**	**xxx.x**	**xxx.x**
2	Market (δ)	xxx.x	xxx.x	xxx.x
3	Market (γ)	xxx.x	xxx.x	xxx.x
4	Rho (ρ)	xxx.x	xxx.x	xxx.x
5	Volatility (ν)	xxx.x	xxx.x	xxx.x
6	Theta (θ)	xxx.x	xxx.x	xxx.x
7	Dividends	xxx.x	–	xxx.x
8	Basis Risk	–	xxx.x	xxx.x
9	Knock Out	xxx.x	xxx.x	xxx.x
10	Unexplained	xxx.x	xxx.x	xxx.x
11	**MV EoP**	**xxx.x**	**xxx.x**	**xxx.x**
12	**Total PnL**	**xxx.x**	**xxx.x**	**xxx.x**

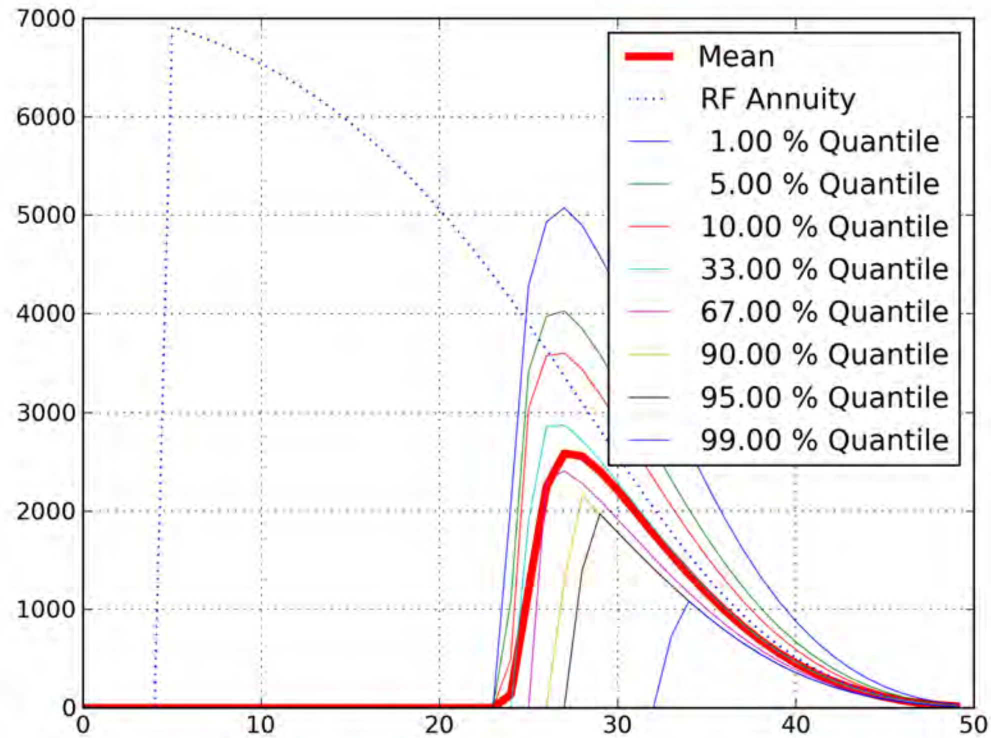

Figure 7.12. Guarantees for a GMWB with ratchet including the funds transfer feature; x: time; y: payout of respective option

Note the following:

- In contrast to the current hedge attribution analysis one would also calculate the volatility component for the liabilities. This does not necessarily mean that this component is also hedged.
- In the position "Knock Out" one would calculate for Liabilities the deviation between expected and effective policyholder behaviour. If more people would lapse or void their guarantees compared with assumptions this position would show the experience variance. One could think of showing the whole PH behaviour aspect in one position or eventually split it further.
- The following relations hold $\sum(1)\ldots(10) = (11)$ and $(12) = (11) - (1)$.

8. Policyholder Bonus Mechanism

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

The aim of this chapter is to introduce the concept of policyholder bonus and the corresponding effects.

8.1 Concept of Surplus: Traditional Approach and Legal Quote

Lets reconsider example 6.4.10. We have used in this context a interest rate for pricing of 2%. According to the EU laws the technical interest rate must normally not exceed 60% of the average government bond yields. This would mean that at that time the average government bond yield was above 3.3 % – assume 4 %. If we redo the corresponding calculations, we get the following:

Age	Unit	Units for Mortality	Units for Premium	Total Units	Value $i = 2\%$	Value $i = 4\%$
50	$\mathcal{Z}_{(0)}$	–	-9562.20	-9562.20	-9562.20	-9562.20
51	$\mathcal{Z}_{(1)}$	1000.00	-9466.57	-8466.57	-8300.56	-8140.94
52	$\mathcal{Z}_{(2)}$	1089.00	-9362.44	-8273.44	-7952.17	-7649.26
53	$\mathcal{Z}_{(3)}$	1174.93	-9250.09	-8075.16	-7609.40	-7178.79
54	$\mathcal{Z}_{(4)}$	1257.56	-9129.84	-7872.27	-7272.76	-6729.25
55	$\mathcal{Z}_{(5)}$	1336.69	-9002.02	-7665.32	-6942.72	-6300.34
56	$\mathcal{Z}_{(6)}$	1412.12	-8866.99	-7454.87	-6619.71	-5891.69
57	$\mathcal{Z}_{(7)}$	1483.67	-8725.12	-7241.45	-6304.11	-5502.90
58	$\mathcal{Z}_{(8)}$	1551.18	-8576.79	-7025.61	-5996.29	-5133.54
59	$\mathcal{Z}_{(9)}$	1614.50	-8422.41	-6807.90	-5696.55	-4783.14
60	$\mathcal{Z}_{(10)}$	–	88080.30	88080.30	72256.53	59503.90
Total					**0**	**-7368.19**

Looking at these results we see that the difference in value using the technical pricing rate of 2 % and the market yield of 4 % amounts to 7368.19, which represents 7.3 % of the maturity benefit. Since the corresponding differences can also become bigger than that consumer protection regulation was introduced in the form of legal quotes. The idea there was that for example 90 % of the profit has to allocated to the policyholder and the shareholder receives maximally the remaining 10 %. Assume that in a concrete year, we have:

– Mathematical reserve: 45'000 USD,

– Investment return: 4.2 %,

– Required technical interest rate: 2.0 %.

In this case we get the following income statement:

Item	Amount
Investment Income	1890 USD
Technical Interest	- 900 USD
Gross Profit	990 USD
of which PH	891 USD
of which SH	99 USD

and we remark the following:

1. The biggest part of the gross profit is allocated to the policyholder and the shareholder receives only the smaller part.
2. Since the return on the assets is random, this is also true for the bonus payment, which however can not become negative.
3. As a consequence the shareholder assumes an additional downside risk if he operates an investment strategy which also allows investment returns below the technical interest rate with a positive probability.
4. It is worth mentioning that the insurance company can also give a higher policyholder bonus than the legal minimum defined by the legal quote. Obviously the shareholder return is then reduced accordingly.

After the split of the gross investment return into policyholder bonus and shareholder return, one needs to look how the bonus is used. Here there are different possibilities. The bonus can be used to

– Reduce the premium. Hence assuming a regular premium of 4900 USD, the policyholder would only need to pay 4001 USD.
– Accumulate it on a bank account. The bonus is invested in a bank account type of investment where a yearly interest rate is credited. At the end of the policy the insured receives the value of this account.
– Increase the benefits. In this case the policyholder bonus is used as a single premium to increase the insurance benefits.

We finally want to revisit the above example with different investment returns:

Return	-5%	0%	2%	5%
Investment Income	-2250	0	900	2250
Technical Interest	-900	-900	-900	-900
Gross Profit	-3150	-900	0	1350
of which PH	0	0	0	1215
of which SH	-3150	-900	0	135

Finally we want to remark that there are two slightly different approaches to legal quotes. In the UK context the insurance company balance sheet is split into policyholder and shareholder funds. The legal quote applies to the policyholder funds. In case of an under-coverage of the policyholder funds, the shareholder is required to compensate for this just as in the example above for an investment return of -5%. In continental Europe, such as in Germany and France the legal quote is applied to the whole balance sheet. The difference between the two concepts is that in the UK context the shareholder receives the full return on his assets. On the other hand the shareholder has to share the return on "his" assets with the policyholder according to the legal quote requirements. In a next step we will look at insurance policies from a different point of view, namely from the policyholders' point of view. Also here we want to apply a market consistent approach and hence the value of the premium before and after paying them to the insurance company is equal, but it is allocated to different stakeholders. We want to illustrate this concept based on an example. We assume that

1. the policy is defined as above as an endowment policy with the parameters set of example 6.4.10, and that
2. the pricing interest rate amounts to $i = 2\%$ with a market interest rate of $i = 4\%$.
3. Furthermore we assume that there is a legal quote of 90 % of the difference in present values as shown above.
4. Finally we assume that the present value of the administration charges amounts to 0.2 % of the present value of the benefits and that the tax rate of the insurance company amounts to 35 %.

In this case we know that:

1. The premium amounts to $P = 1.02 \times 9562.20 = 9753.44$.
2. The present value of the premiums amounts (at 4%) to -1.02×58033.16.
3. The present value of death benefits amounts (at 4%) to 7285.59.
4. The present value of maturity benefits amounts (at 4%) to 28481.29.
5. The present value of surrender payments amounts (at 4%) to 14533.06.
6. The present value of gross profit amounts (at 4%) to 7368.19.

Based on this information we can now decompose the present value of the payments of the policyholder (-59k USD) into its parts:

Type	Stakeholder	Amount	%-age
Mortality	PH	7285.59	12.3 %
Maturity	PH	28481.29	48.1 %
Surrender	PH	14897.09	25.1 %
PH Bonus	PH	6631.37	11.2 %
Subtotal PH	PH	57295.34	**96.7 %**
Admin	Employees of Insurer	1160.66	1.9 %
Tax	Tax authorities	257.88	0.4 %
Profit	SH	478.93	**0.8 %**
Total	(equals PV Prem.)	59193.82	100.0 %

After having seen why policyholder bonus is generated, we need to better understand what is done with the bonus allocated to a policy, since this materially impacts the underlying investment strategies and the corresponding risks. There are from a high level point of view two different ways how gross bonus can be used, namely for the benefit of an individual policyholder or for the entire collective of policyholders together. In the first case the individual policyholder benefits from the excess funds allocated to his policy, be in the form of a reduction of premiums or be it that his benefits are increased. Besides the direct allocation of the gross bonus to individual policyholders, there is also the possibility that a part of the money is used for the entire insurance portfolio together, in the sense of mutualisation of insurance risk. This is known as reserve strengthening and is a consequence that the mathematical reserves for an insurance portfolio are estimated based on statistical methods, which involve some uncertainty. In case a new estimation of the expected reserves needed turns out to be higher than the reserves within the balance sheet, it is necessary to strengthen them in order to be able to honour the corresponding commitments. Now there are two possibilities how such a reserve strengthening is financed, by the shareholder, or as mentioned before by the collective of policyholders by using a part of the gross profits stemming from the in-force portfolio.

8.2 Portfolio Calculations

When doing ALM it is normally important to use efficient calculation processes since a lot of simulations are needed. Let's look at the moment at an insurance portfolio with 1 million polices. During the year end calculation, one needs normally to calculate the mathematical reserves for the past, the current and the next year, in order to save these values in the data base to be able to interpolate them for a possible policy surrender. Assume for the moment that your tariff engine is performing such a calculation in 0.01 seconds. Hence the whole year end calculation takes you 8h 20min. When doing ALM this is obviously too long when requiring 10000 simulations, since this would result in about 3500 years run-time on the same infrastructure. Some acceleration can be gained by using a grid, but even then it seems to make sense to look for faster methods to do the above task. In this section we will look at such approaches, which can concretely be implemented. If we consider a simple set up we have a set of policies $\mathcal{P}$, and each policy can be characterised by its Markov representation, eg the state space S_i, the discount factors (which is normally the same for all polices), the transition probabilities (which are normally structurally similar per different state space) and finally the benefits vectors $a_{ij}(t)$ and $a_i(t)$. In a lot of cases one can also restrict the state space to a common one, which we call S. In order to be concrete we want to have a look at the set of all insurance policies on one life, be it a lump sum or an annuity. In this case we choose $S = \{*, \dagger, \ddagger\}$ as corresponding state space for a person with age ξ, where $\ddagger$ represents the state of a surrendered policy. In this set up we introduce the linear vector space of all insurance policies for this person ξ by

$$\mathcal{F}_\xi = \{x_\xi = (a_{ij}(t), a_i(t)) : i, j \in S \text{ and } t \in \mathbb{N}\}.$$

We now remark that both the mathematical reserve and the expected cash flow operators

$$\begin{aligned} \Phi_{t,j}(x_\xi) &: \mathcal{F}_\xi \to \mathbb{R}, x_\xi \mapsto V_j(t)[x_\xi] \\ \Psi_{t,j}(x_\xi) &: \mathcal{F}_\xi \to \mathbb{R}, x_\xi \mapsto \mathbb{E}[CF(s)[x_\xi] | X_{t_0} = j] \end{aligned}$$

are linear (continuous) functionals from $\mathcal{F}_\xi \to \mathbb{R}$, where ξ denotes the policy considered, t and s the respective times and $j \in S$ a state. When recognising this fact we can now construct the space of all insurance policies for a given portfolio $\mathcal{P} = \{\xi_1, \xi_2, \ldots \xi_n\}$ by defining the respective Cartesian product such as:

$$\begin{aligned} S &= \prod_{i \in \mathcal{P}} S_i, \text{ and} \\ \mathcal{F} &= \prod_{i \in \mathcal{P}} \mathcal{F}_i. \end{aligned}$$

In the same sense we can now define the mathematical reserve and expected cash flow operator of the whole insurance portfolio as the corresponding sum:

$$\Phi_t(x) : \mathcal{F} \to \mathbb{R}, x((\xi_i)_{i \in \mathcal{P}}) \mapsto \sum_{i \in \mathcal{P}} V_{j(\xi_i)}(t)[x_{\xi_i}] \text{ and}$$

$$\Psi_t(x) : \mathcal{F} \to \mathbb{R}, x((\xi_i)_{i \in \mathcal{P}}) \mapsto \sum_{i \in \mathcal{P}} \mathbb{E}[CF(s)[x_{\xi_i}] | X_{t_0} = j(\xi_i)].$$

Until now we have gained nothing except for a more complex representation of what we already know. The way we can now make the whole thing much more efficient is to use the given structure and the fact that the two operators defined above are linear. In the concrete set up where each policy is characterised by the three states above we can define a new “pseudo” state space

$$\tilde{S} = \{x0, x1, \dots, x120, y0, y1, \dots, \dagger, \ddagger\},$$

where the states $x0, \dots, x120$ stand for males which are alive and have the respective age at time t_0. Similarly $y0, \dots, y120$ stand for the respective females. There is only a need to map the respective $x_\xi \in \mathcal{F}_\xi$ into $\tilde{\mathcal{F}}$. Assume that $\mathcal{R}50 \subset \mathcal{P}$ is a homogeneous set of policies representing the males ages 50. In this case we have the following benefit functions:

$$\begin{aligned}
a_{x50}^{Pre} &= \sum_{\xi \in \mathcal{R}50} a_{*}^{Pre}(\xi) \\
a_{x50,x50}^{Post} &= \sum_{\xi \in \mathcal{R}50} a_{*,*}^{Post}(\xi) \\
a_{x50,\dagger}^{Post} &= \sum_{\xi \in \mathcal{R}50} a_{*,\dagger}^{Post}(\xi) \\
a_{x50,\ddagger}^{Post} &= \sum_{\xi \in \mathcal{R}50} a_{*,\ddagger}^{Post}(\xi)
\end{aligned}$$

Please note that in order to implement the approach described above, one needs to be careful with respect to the definition of time. In the usual Markov model the time t is calculated with respect to the age of each policyholder in a certain year. For the above purpose, it is useful to enumerate by the number of years into the future, starting at $t = 0$. Hence one needs to correspondingly adjust the time of the benefit functions. We remark that for all the other states in $\tilde{S} \setminus \{\dagger, \ddagger\}$ the same approach is used. It is also worth noting that this “pseudo” Markov chains can be interpreted as “normal” Markov chains, where the initial state of the portfolio is given by a probability distribution, over which one integrates. After having done so, we can now perform a lot of the calculations much faster. After having applied the above mapping into $\tilde{\mathcal{F}}$, the calculations do not depend anymore on the actual size of the portfolio. By this we gain considerable amounts of times when doing the actual calculations. Looking at the sample portfolio above we would possibly use 1h calculation time to perform the mapping on $\tilde{\mathcal{F}}$, including the data base queries. Once this is done the calculation of the mathematical value and expected cash flow operator take some 1 to 2 seconds on a common laptop. Hence doing ALM in a lot of cases result in run-times of several minutes. In the same sense stress scenarios can be calculated much faster. Finally I would like to mention that this approach has been applied concretely for the examples in section 8.3 using 402 states for $\tilde{\mathcal{F}}$ by also splitting annuities from capital insurance. Using this approach one can also map deferred widows pension using a collective approach and hence one can cover the vast majority of traditional life insurance covers sold by a life insurance company. Since structurally the method is a slight variation of the Markov recursion, this method can actually be implemented using the same core Markov calculation objects.

8.3 Portfolio Dynamics and ALM

The aim of this section is to look at the portfolio dynamics induced by the relationship between assets and liabilities and the corresponding asset liability management (ALM). To this end we fix $(\Omega, \mathcal{A}, P)$, together with a filtration $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{N}}$. We assume that we have represented the benefits and premiums in a suitable ("pseudo") Markov model with state space $\tilde{S}$ and benefits vector space $\tilde{F}$. In this case we can represent the benefits for the entire portfolio by $x = ((a_{i,j}^{Post}(t))_{i,j\in\tilde{S}}, (a_i^{Pre}(t))_{i\in\tilde{S}})_{t\in\mathbb{N}} \in \tilde{F}$, and we know that we can decompose this into

$$x = x^{Benefits} + x^{Premium},$$

where $x^{Benefits}$ and $x^{Premium}$ represent the corresponding benefit vectors in $\tilde{\mathcal{F}}$ for benefits and premiums, respectively. At this point the bonus concept enters. If for example the benefits are increased, the corresponding $x^{Benefits}$ is increased accordingly. Formally one introduces hence a random vector α_t which represents the relative benefit level. This means that $\alpha_0 = 1$ and also that α is previsible. For traditional bonus promises, the bonus allocated to the individual policy become a guarantee, means that α is increasing in t for each trajectory and hence we define the new benefit representation of the policy at time t as follows:

$$\hat{x}^{Benefits} = ((\alpha_t \times a_{i,j}^{Post}(t))_{i,j\in\tilde{S}}, (\alpha_t \times a_i^{Pre}(t))_{i\in\tilde{S}})_{t\in\mathbb{N}},$$

remarking that this is now a random quantity, since α is a random vector. In consequence the entire portfolio after bonus allocation is given by

$$\hat{x} = \hat{x}^{Benefits} + x^{Premium}.$$

The mechanism to increase α is performed by using the allocated bonus as a single premium. In a next step we want have a look at a concrete example. The first step is to define a stochastic model, which generates the corresponding states of the world. In the concrete set up we are assuming a world with a constant interest rate with one risky asset (say a share) with has a non constant, stochastic volatility. We are using the Heston model, which is given by

$$\begin{aligned} dV_t &= \kappa(\theta - V_t)dt + \eta\sqrt{V_t}dW_t^1 \\ dS_t &= \mu S_t\,dt + \sqrt{V_t}\,S_t\,dW_t^2. \end{aligned}$$

This model is given by two stochastic differential equations, where the first equation describes the volatility of the shares. This volatility process has a structural similarity with the interest rate models we have seen before, in the sense that the volatility process is also mean reverting. In order to solve this stochastic differential equation system we can use a numerical method such as the Milstein scheme (see for example [KP92]). After having done this, it is important to understand how the simulation works. In principle one does first a loop over the different simulation and calculates the quantities, which need to be analysed. The following code performs the corresponding task:

```
1  # 4. Stepper
2  # ---------------------------
3  for i in 1... n:
4      sim.vNewTrajectory()
5     (pl, cf, ... ) = CalcPV(sim, lvp, lvl)
    ... keep and analyse results for run i
```

Hence first one generates a new trajectory of the world (line 4, in this case based on the Heston model), and secondly on calculates the development of the insurance portfolio, as outlined above in the subroutine `CalcPV` (line 5). Hence it makes sense to look at this part of the code in order to understand what happens:

```
... code left away
 1  for i in 1... MaxTime:
 2      TargetSurplus  = ... target surplus level required
 3      perf = ((1+iRF) * (1-EquityProportion) +
            EquityProportion* sim.dGetValue(1,float(i)) / sim.dGetValue(1,float(i-1.)))
 4      cf[i] = psi[i-1] * lvl.dGetCF(i) + lvp.dGetCF(i)
 5      l[i] = psi[i-1] * lvl.dGetDKTilde(i) + lvp.dGetDKTilde(i) - cf[i]
 6      a[i] = a[i-1] * perf - cf[i]
 7      ExAssets = a[i] - l[i]
```

```
8     # Calculate Excess Assets and Pl, afterwards adjust assets by pl for SH
9      if ExAssets > 0.:
10         ExAssets = max(0., ExAssets - TargetSurplus)
11         Bonus = ExAssets * RelBonusAllocation * LegalQuote
12         pl[i-1] = ExAssets * RelBonusAllocation * (1-LegalQuote)
13     else:
14         Bonus = 0.
15         pl[i-1] = ExAssets
16     a[i] -= pl[i-1]
17     e[i] += pl[i-1]
18     psi[i] = psi[i-1] + Bonus / (psi[i-1] * (lvl.dGetDKTilde(i) - lvl.dGetCF(i)))
```

What this code does, is the following:

1. In a first step, the minimal required surplus of the policyholder funds (line 1) and the performance of the assets (line 2) is calculated. We see that in the concrete set up we have a mixture of assets. A part of them yielding risk free and the reminder, the equity portion having an equity yield.
2. In lines 4,5 and 6 the cash-flows and the assets and liabilities are calculated at the end of the period. We see that

$$\begin{aligned} \Psi_i(\hat{x}^{Benefits}) &= \texttt{psi[i-1] * lvl.dGetCF(i)}, \text{ and} \\ \Psi_i(x^{Premium}) &= \texttt{lvp.dGetCF(i)}. \end{aligned}$$

3. In a next (lines 7 to 15) step the excess assets are calculated in order to determine, whether there is a bonus in the corresponding year. If there are excess assets (line 9), the bonus is calculated.
4. Finally the benefit level for the subsequent year is calculated (line 18)

We remark that the initialisation of the code plus the analysis have been left in order to focus on the essential parts of the calculation. After having done this, we want to have a look at some sample output of the program. Figure 8.1 and 8.2 show the mathematical reserves and the expected cash flows for the benefits and the premiums respectively. Moreover figure 8.3 show the quantiles of the profit and loss over time for the 5%, 10%, 33%, 50 %, 67%, 90% and 95% quantiles for both the profit and loss account and the corresponding dividends. We remark that this figure shows that the underlying portfolio is insufficiently financed and suffers considerable losses between the time 5 and 15, and also that the losses start earlier in cases where we observe an adverse equity performance. The output of the program looks as follows:

```
Input Parameter / Main Results
------------------------------
Simulations                                 5000
>>   PV Benefits ..... .....DK 1   =  22,786,998,031
>>   PV Premium. ..... .....DK 0   =  -4,091,249,171
>>   Mathematical Reserves..DK     =  18,695,748,860
>>   Underlying Assets .....A0     =  22,635,000,000
>>   Shareholder Equity.....E0     =   1,584,450,000
Distribution of Economic Profit
-------------------------------
Min Return                                0.0962
Max Return                                0.2473
Average                                   0.1792
 0.5%  Quantile                           0.1288
 1.0%  Quantile                           0.1324
 5.0%  Quantile                           0.1452
```

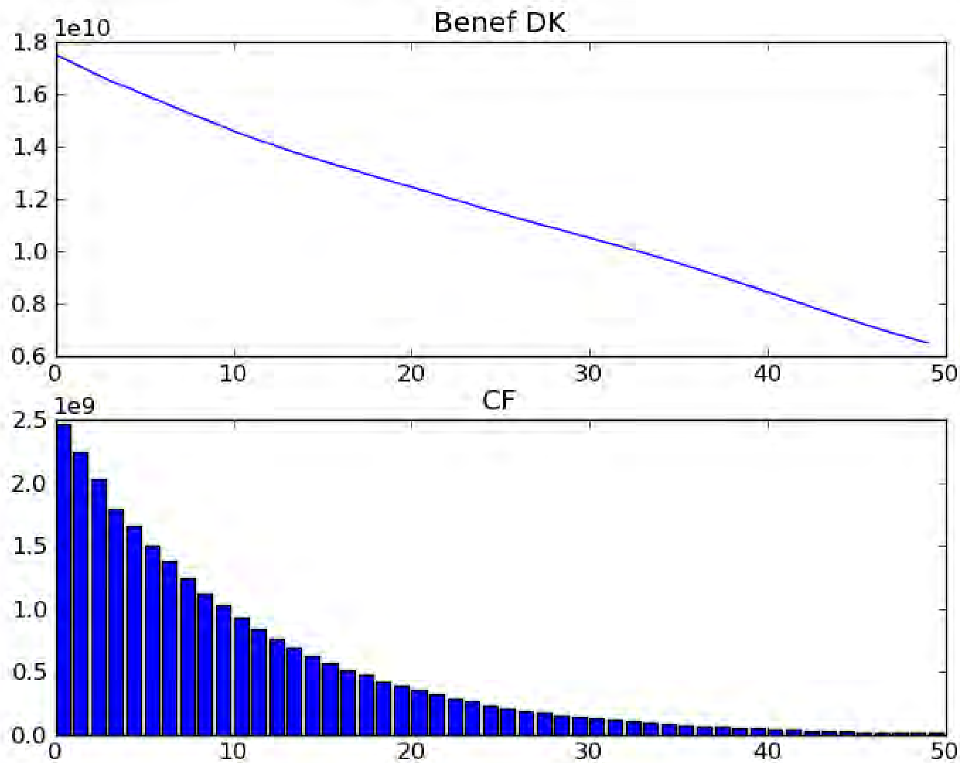

Figure 8.1. Mathematical reserve and expected cash flows of portfolio (benefits only)

```
10.0%  Quantile                                0.1529
25.0%  Quantile                                0.1654
99.5%  Quantile                                0.2289
Time used
---------
Step              Calc EP Time used     2.3515 s
Step              Calc PF Time used     3.6497 s
Step              Calc PF Time used     3.2491 s
Step            A0 and E0 Time used     0.0190 s
Step   Simulation Stepper Time used    81.8312 s
Time used for preparation               9.2 s
Time used for simulations              81.8 s
Time per simulation                     0.0163 s
```

We see in particular that the implementation via the "pseudo" Markov model results in an extremely fast calculation of the portfolio in about 9 seconds. Also the simulation is performed very fast, despite the fact that a relatively slow laptop was used. The output of the above run is summarised also in figure 8.4.

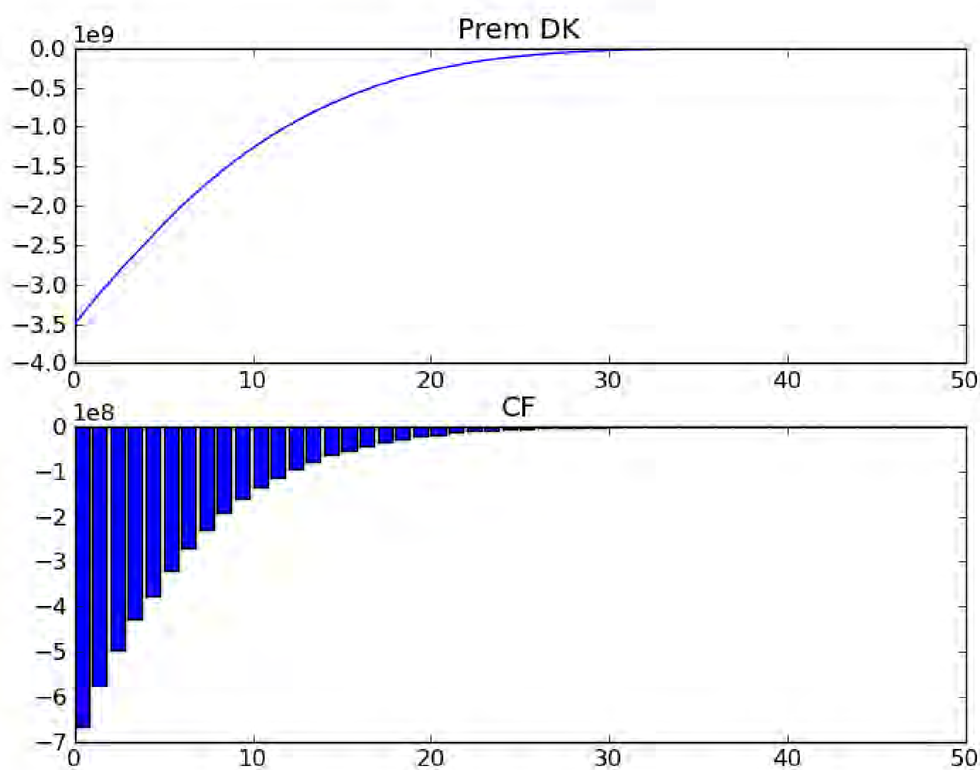

Figure 8.2. Mathematical reserve and expected cash flows of portfolio (premiums only)

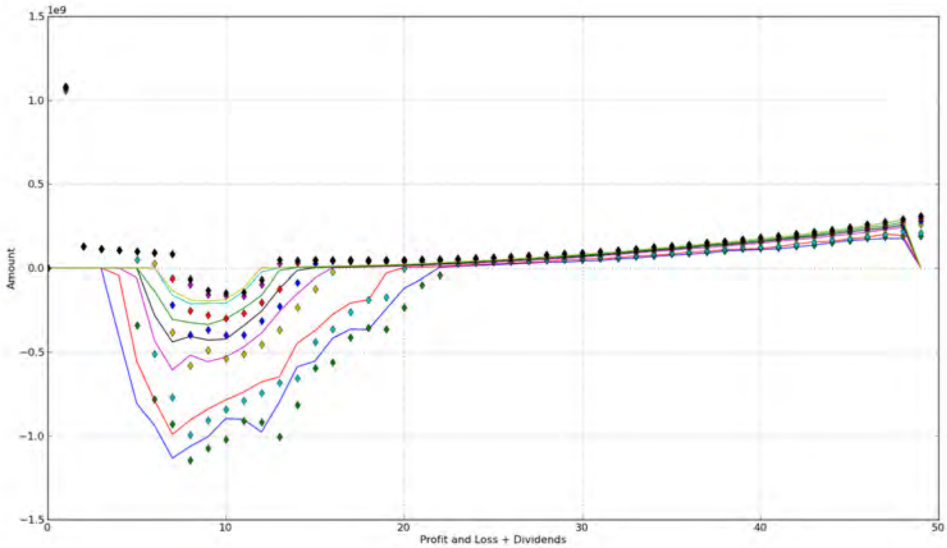

Figure 8.3. Quantiles for profits and losses

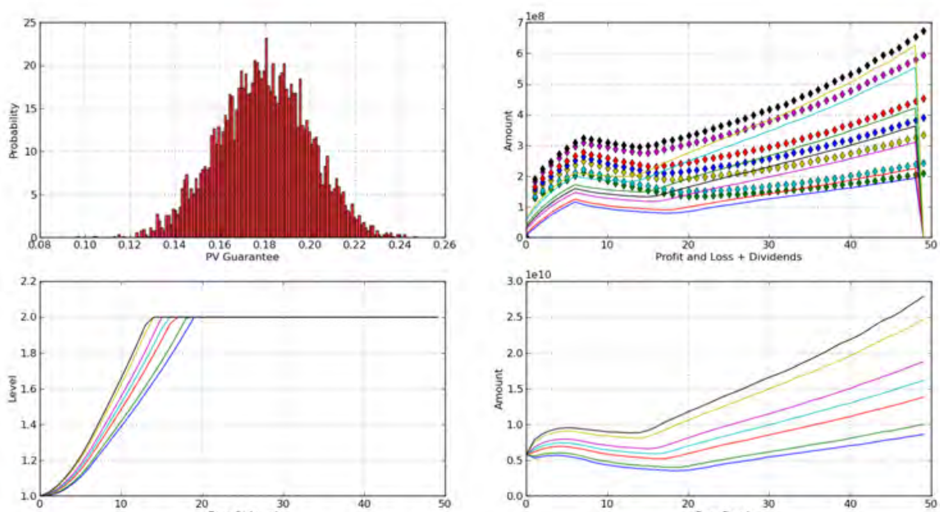

Figure 8.4. Summary of ALM analysis

9. Financial Risks and their Modelling

The aim of this chapter is to educate the readers, in order that they understand the basics of financial risk management and so that they can interpret the numbers within this report. For the underlying abstract valuation concept we refer to chapter 6.

9.1 The Model underlying Financial Risks

In order to develop a model for managing and measuring financial risks we have a look at the balance sheet, which have seen earlier in this book:

Balance sheet	Book **A**	Book **L**	Market **A**	Market **A**	
Cash	6200	47100	6200	48513	MR
Bonds	35700	**2200**	37842	**3569**	SHE
Shares	4400		4800		
Properties	1100		1300		
Loans	1400		1400		
Alternatives	500		540		
Total	49300	49300	52082	52082	

It is clear that we need to decouple the valuation π_t from the underlying asset. So formally the balance sheet consists of assets $(\mathcal{A}_i)_{i \in S_A}$ and Liabilities $(\mathcal{L}_i)_{i \in S_L}$ and we assume that both index sets S_A and S_L are finite. Now assume we have 1000 shares from HSBC. We could say that these 1000 shares are "one" asset. On the other hand we could model the same holding as holding 1000 pieces of the asset "1 HSBC share". Therefore we denote by $(\alpha_i)_{i \in S_A}$ and $(\lambda_i)_{i \in S_L}$ the number of units which we own at the certain point of time. Furthermore we want to separate the shareholder equity from the liabilities and we denote it $\mathcal{E}$.

If we write $\alpha_1 \, \mathcal{A}_1$ we assume that we are holding α_1 units of the asset $\mathcal{A}_1$. Hence our portfolio is an abstract finite dimensional linear vector space $\mathcal{Y} = \operatorname{span}\{(\mathcal{A}_i)_{i \in S_A}, (\mathcal{L}_i)_{i \in S_L}, \mathcal{E}\}$. In this context our balance sheet is a point $x = \sum_{i \in S_A} \alpha_i \mathcal{A}_i + \sum_{i \in S_L} \lambda_i \mathcal{L}_i \in \mathcal{Y}$.

As seen before some assets and liabilities can be further decomposed in simpler assets and liabilities and hence we can find a suitable basis for the vector space $\mathcal{Y} = \operatorname{span}\{e_1, \ldots, e_n\}$, where $(e_k)_{k \in \mathbb{N}_n}$ is its basis, and we remark that we can also write our balance sheet as $x = \sum_{k \in \mathbb{N}_n} \gamma_k \, e_k$.

The idea to introduce $\mathcal{Y}$ is to have a normalised vector space. Assume for example that we hold some ordinary bonds. In this case we would use as $e_k = \mathcal{Z}_{(k)}$, the corresponding zero coupon bonds, etc.

We finally remark that the balance sheet $x \in \mathcal{Y}$ actually represents a random cash flow vector, and hence we strictly have x_t or $X_t(\omega) \in \mathcal{X}$ if we assume that the changes of the portfolio follow a stochastic process (cf. appendix 7). For measuring the risk of the actual balance sheet it is normally sufficient to assume that $y \in \mathcal{Y}$ does not change.

Next we need to look at the second part, namely the valuation π_t, and we remark that:

- The valuation is dependent on time.
- We assume that the valuation is a linear functional $\pi_t : \mathcal{Y} \to \mathbb{R}$ which allocates to each asset its value (see also appendix 6).
- A liability $\mathcal{L}$ is characterised by $\pi(\mathcal{L}) \leq 0$. In the same sense and asset has a positive value. As a consequence an $x \in \mathcal{Y}$ can in principle be both an asset or a liability, depending on the economic environment and also depending on the valuation functional.

After having defined the different parts we need to have a closer look at what equity or capital ($\mathcal{E}$) means. In the context of the balance sheet we observe that the sum of the value of all assets equals the sum of the value of all liabilities (neglecting the sign). Hence we have the following:

$$
\begin{aligned}
x &= \sum_{i \in S_A} \alpha_i \mathcal{A}_i + \sum_{i \in S_L} \lambda_i \mathcal{L}_i + \mathcal{E} \in \mathcal{X}, \text{ and} \\
\pi(x) &= \pi \left(\sum_{i \in S_A} \alpha_i \mathcal{A}_i + \sum_{i \in S_L} \lambda_i \mathcal{L}_i + \mathcal{E} \right) \\
&= 0, \text{ and hence} \\
SHE = \pi(\mathcal{E}) &= -\pi \left(\sum_{i \in S_A} \alpha_i \mathcal{A}_i + \sum_{i \in S_L} \lambda_i \mathcal{L}_i \right).
\end{aligned}
$$

This means that we can always calculate the value of the shareholders' equity if we know the value of all other assets and liabilities.

Finally we want to show how to tackle the stochastic valuation functional π_t. Since we live in a linear vector space $\mathcal{Y}$ with a basis $(e_k)_{k \in \mathbb{N}_n}$, it is sufficient to define the price $\pi_t(e_k)$. The idea is to decouple the operator from the economy and the corresponding set up is to define the state of the economy by a stochastic process $(R_t)_{t \in \mathbb{R}} \in \mathbb{R}^m$. You could think that one of the components could be inflation, another could be the level of the 10 year interest rate, etc. In this setup we can define:

$$
\pi_t(e_k) = f_k(R_t),
$$

where $f_k : \mathbb{R}^m \to \mathbb{R}$ is a sufficiently regular function. If we assume for example that $R_t[10]$ is the interest rate for the 10 year bond, then we have (depending on our definition of π)

$$
\pi_t(\mathcal{Z}_{(10)}) = (1 + R_t[10])^{-10}.
$$

The idea of financial risk management is to assess and control the change of the value of the shareholder equity, e.g. the profit and loss induced by this change. If we assume for the moment that the time t is denoted in years, one is normally interested in the following quantity:

$$PL_T = (\pi_T(\mathcal{E}) - \pi_0(\mathcal{E}))\,.$$

The loss which we encounter within the time interval $[0,T]$. Banks normally look at one week, eg $T = 1/52$, Solvency II looks at $T = 1$. One measures the risk, as indicated before based on the random variable PL_T.

Here again is a more formal environment: In order to assess the financial risk of an insurance company the following steps are needed.

1. Define the valuation methodology π_t,
2. Define (note this is a big model assumption) which stochastic process R_t models the economy,
3. Define the universe of all assets and liabilities $\mathcal{Y}$,
4. Define and calculate the functions $(f_k)_{k\in\mathbb{N}_n}$,
5. Analyse the possible balance sheets $x \in \mathcal{Y}$ and decompose each $\mathcal{A}_i$ and $\mathcal{L}_i$ into the basis $(e_k)_{k\in\mathbb{N}_n}$,
6. Define the risk measure to be used such as VaR, etc.,
7. Implement the model.

The implementation of the above steps in its purest form is very complex and therefore one normally has to make approximations.

9.2 Approximations

A common approximation starts with the simplification of the function f_k, by using a *Taylor approximation.* Since we are interested in

$$\begin{aligned} PL_T &= \pi_T(\mathcal{E}) - \pi_0(\mathcal{E}) \\ &= [\pi_T - \pi_0] \circ \left(\sum_{i\in S_A} \alpha_i \mathcal{A}_i + \sum_{i \in S_L} \lambda_i \mathcal{L}_i \right), \end{aligned}$$

we use the following first order Taylor approximation

$$\begin{aligned} \pi_T(e_k) - \pi_0(e_k) &= f_k(R_T) - f_k(R_0) \\ &\approx \nabla f_k(x)\|_{x=R_0} \times \Delta(R). \end{aligned}$$

If we apply this formula to all assets and liabilities we get a model where the gains and losses are linear in the risk factors R. If there is a balance sheet $x = \sum_{k\in\mathbb{N}_n} \gamma_k\, e_k$ we can obviously sum over the different e_k and we get the following approximation:

$$\begin{aligned} \pi_T(x) - \pi_0(x) &= \sum_{k\in\mathbb{N}_n} \gamma_k \times (f_k(R_T) - f_k(R_0)) \\ &\approx \delta^T \times \Delta(R), \end{aligned}$$

where

$$\delta = \sum_{k \in \mathbb{N}_n} \gamma_k \times \nabla f_k(x)\|_{x=R_0},$$

and where we denote with x^T the transposed of a matrix or vector.

Another simplification is to use a stochastic process, which is analytically easy to tackle. Both the risk metrics method and also the Swiss solvency test use a multi-dimensional normal distribution for $Z = \Delta R$.

Hence we have

$$Z \sim \mathcal{N}(0, \Sigma),$$

where we Σ denotes the *covariance matrix*. One can express this matrix by the standard deviation vector s for each of the risk factors and the correlation matrix ρ. In a first step we define the matrix $S = (v_i \times \delta_{ij})_{i,j}$. Furthermore we need to know that if $X_1 \sim \mathcal{N}(\mu_1, \Sigma_1)$ is a multidimensional normal distribution and A and b are a matrix and a vector, respectively, we then know that $X_2 := A \times X_1 + b \sim \mathcal{N}(\mu_1 + b, A \times \Sigma \times A^T)$. Using this formula we finally get the following relationship:

$$\Sigma = S \times \rho \times S,$$

keeping in mind that $S = S^T$.

If we use the two approximations, the calculation of the VaR at a level α (eg 99.5%) can be calculated as follows. In a first step we denote by

$$\zeta = F^{-1}_{\mathcal{N}(0,1)}(\alpha),$$

and we get in consequence:

$$\begin{aligned} VaR_{PL}(\alpha) &= F^{-1}_{\mathcal{N}(0,1)}(\alpha), \\ &= \zeta \times \sqrt{(s \times \delta)\rho(s \times \delta)^T}. \end{aligned}$$

Hence the value at risk can easily be calculated using some simple matrix multiplications. The example which follows is based on these approximations.

At this point it is important to remark that every model has flaws and hence it is of utmost importance to understand the *limitations* of a model. The risk to choose a "wrong" or "inaccurate" model is called *model risk*. Here it important how a model is constructed. Figure 9.1 aims to show this. In principle there are the reality (left hand side of the figure) which one tries to model in order to answer "difficult" questions which can not be answered directly. In order to do that one creates a model (right hand side of the figure) and one should be able to answer the corresponding questions within the model. Next one translates the results back to reality and "hopes" that the diagram is commutative. From this point of view the model risk is the missing "approximate" commutativity of the model. As a corollary one needs to acknowledge that each model is suited and best adapted for a certain purpose and that it is dangerous to use the model outside that.

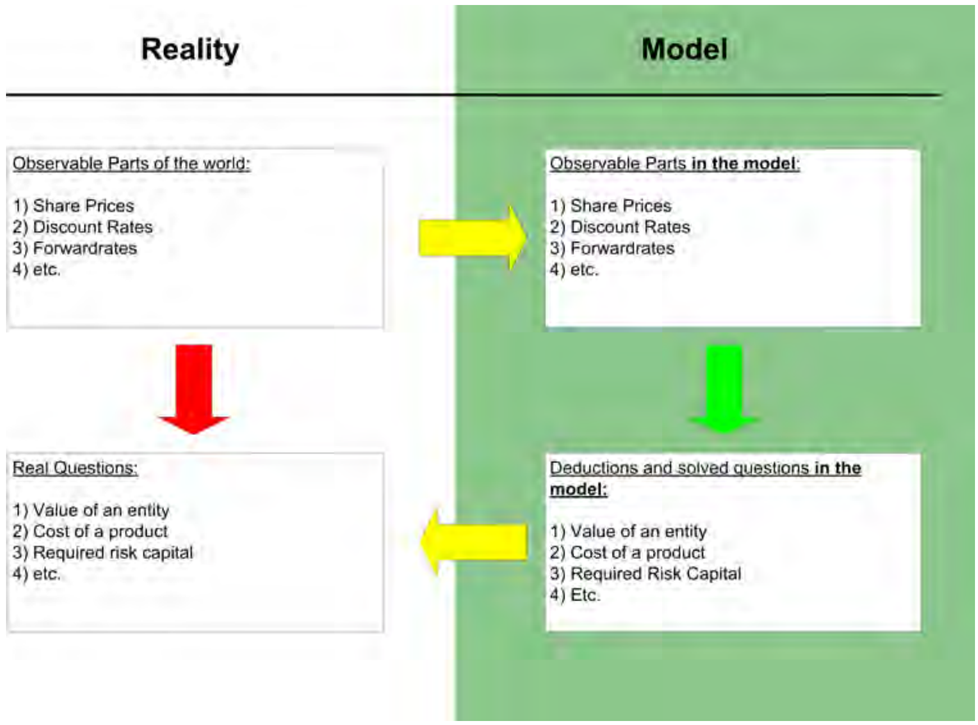

Figure 9.1. Models and Model Risk

Another interesting aspect with respect to model risk is the fact that one can, from time to time, observe difficult and lengthy discussions between experts on which model is better. Such discussions can stem from the fact that these people do not distinguish between reality and the model and hence these discussions can end up in religion like beliefs.

In the same sense the results of every model depend on the parameters chosen. The risk of inaccurate model parameters is called *parameter risk*. An easy example is the equity volatility, which is for example used for the Black-Scholes model. The value of the corresponding options is heavily dependent of the volatility chosen. As remarked before the volatility for equity market indexes is normally in the region of say 17 %. In case of market disruptions this parameter can spike up to 30 % and above. Hence it is crucial to exactly know how the model behaves with respect to different parameters.

Finally it is worth noting that the distinguishing between model and parameter risk is not always clear.

9.3 Concrete Implementation

For the concrete implementation of a risk model for financial risk there are, in principle, the following three different approaches:

1. *Analytical approach*, such as the one used in the Swiss solvency test: Here the required risk capital is calculated based on a closed formula. The advantage here is the fast calculation times because this approach is only feasible for a limited class of model.
2. *Model based simulation* (aka Monte Carlo approach): One can, in principle, use whichever model is deemed to be adequate and one simulates the corresponding random variables. Here one can also

use sophisticated methods to link variables together such as the copula method. This method is very flexible - for the price of having normally longer running times, since one requires normally a sizable amount of simulations in order to determine the tail probabilities with an adequate accuracy. Assume for example that we are interested in the 99.5% VaR. In this case we have only 500 simulations which are beyond this level for a sample of 100000.

3. *Historical Simulation*: In this case one uses past observed financial data to predict the future. The big advantage is the fact that we do not need to assume which is the correct distribution. In this class of methods we can either run through the past time series or one can use the boot-strapping method. The problem with this method is the fact that there are only quite short time series (say 50 years) for the underlying financial data. Since one is normally interested in rare events such as the one in a 200 year event one needs to amend this method correspondingly. Furthermore, one needs also to remark that the behaviour of some financial variables has changed considerably over the past 50 years, such as foreign exchange rates, which were fixed until the 1970s and are now floating.

As seen above there are different methods on which we can implement financial risk management. In this section we will have a closer look at the multi-normal model, as used in the analytic part of the Swiss solvency test. First we need to look at the risk factors used and then to calculate the risk capital for the balance sheet introduced above, based on a simplified model.

The Swiss solvency test uses the following risk factors:

- Zero coupon prices for CHF, EUR, GBP and USD, for 13 time buckets,
- Interest rate volatility,
- Credit Spreads for four different rating categories,
- Four different currencies vs CHF: EUR, GBP, USD and YEN,
- Seven equity indexes,
- Equity volatility,
- Real estate, hedge funds and private equity indexes,

each of which is modelled as a normal distribution. Before making a concrete example we want to have a look on how big the different quantities are. Since there are 81 risk factors, this would result in a 81x81 covariance matrix. In consequence we will have a look at a part of it. Firstly we want to look at the corresponding standard deviations (as of 31/12/08).

Risk Factor RF_i		Quantity	$\sigma(RF_i)$
EUR	1	bps	61.82
EUR	2	bps	72.08
EUR	3	bps	73.00
EUR	4	bps	73.12
EUR	5	bps	83.53
EUR	6	bps	70.43
EUR	7	bps	68.09
EUR	8	bps	65.93
EUR	9	bps	64.88
EUR	10	bps	63.54
EUR	15	bps	58.91
EUR	20	bps	60.94
EUR	30	bps	59.95
EURO STOXX		in%	18.78
Credit	AAA	bps	11.08
Credit	AA	bps	12.00
Credit	A	bps	23.80
Credit	BBB	bps	52.60

From the above table we see that the volatility for equities was about 19% and the standard deviation for spread risk increases if the credit quality deteriorates. In a next step we want to have a look at the (simplified) correlation matrix:

$\rho_{i,j}$	EUR 5	EUR 10	EUR 20	STOXX	AA	BBB
EUR 5	1.00	0.89	0.66	0.36	-0.14	-0.23
EUR 10		1.00	0.73	0.32	-0.17	-0.21
EUR 20			1.00	0.16	-0.09	-0.09
STOXX				1.00	-0.45	-0.50
AA					1.00	0.61
BBB						1.00

Looking at figure 9.2 we see clearly how the different risk factors are situated in the matrix. One sees the four times 13 risk factors relating to interest rates as first block, which is highly correlated between themselves and slightly less between different currencies. Afterwards one sees the correlation with the credit block, followed by the equity–like investments etc.

From the above table it becomes obvious that the credit spreads have a quite high negative correlation with stock market prices. This means that credit spreads increase normally at the same time when equity markets fall. One can observe that increasing stock market prices normally imply increasing interest rates. These two remarks are for example valid for the credit market crisis in 2008. Here we observed decreasing stock market indexes, reduced interest rate levels and increased credit spreads.

In order to make a concrete example based on the above data we need to base the calculation on a balance sheet and we assume:

Item	**EUR**	**Term**	**Rating**
MR	-10000	10	Government
Bond 1	5000	5	BBB
Bond 2	4000	20	AA
STOXX	2000		
Capital	1000		

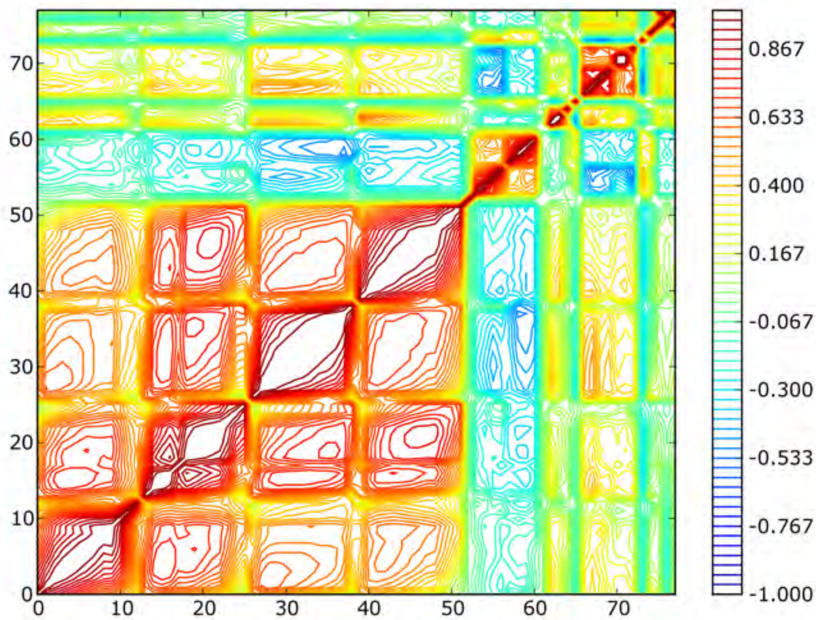

Figure 9.2. Correlation Matrix

So as a first step we need to calculate the sensitivities regarding our risk factor vector EUR 5,EUR 10, EUR 20, STOXX, AA, BBB. We assume for the sake of simplicity each of the bonds and the mathematical reserve (MR) is zero coupon bonds with the corresponding term. In this case the duration equals the term, as one can easily verify. Since the volatility for interest rates and credit spread movement is stated in bps, we also need to calculate the sensitivity of the corresponding values per bp.

Since the MR is considered as a $\mathcal{Z}_{(10)}$ there is only a sensitivity with respect to the EUR 10 year risk factor and an upward movement of 1% reduces the reserve by 10%, so from a capital point of view we have an entry of +1000. For 1 bp we hence have +10 and the sensitivity factor for this liability reads as $\delta_{x_1} = (0, 10, 0, 0, 0, 0)$. In the same sense we can calculate Bond 1 (δ_{x_2}) and Bond 2 (δ_{x_3}), remarking that both of them are sensitive also with respect to credit spreads we get

$$\begin{aligned} \delta_{x_2} &= (-2.5, 0, 0, 0, 0, -2.5), \\ \delta_{x_3} &= (0, 0, -8, 0, -8, 0). \end{aligned}$$

For the share we calculate the sensitivity for an increase of 1% and hence we get:

$$\delta_{x_4} = (0, 0, 0, 20, 0, 0),$$

and therefore we get for the total sensitivity:

$$\begin{aligned} \delta_{Tot} &= \sum_{k=1}^{4} \delta_{x_k} \\ &= (-2.5, +10, -8, +20, -8, -2.5). \end{aligned}$$

In a next step we need to calculate:

$$
\begin{aligned}
s \times \delta_{Tot} &= (83.53 \times -2.5, \ldots, 52.61 \times -2.5) \\
&= (-208.83, 635.42, -487.52, 375.70, -96.03, -131.51).
\end{aligned}
$$

Now we can calculate the standard deviation of the capital, considered as a random variable by:

$$
\begin{aligned}
\sigma &= \sqrt{(s \times \delta)\rho(s \times \delta)^T} \\
&= 546.6\text{M EUR}.
\end{aligned}
$$

As a consequence of this the VaR for the 99.5% corresponds to $VaR_{99.5\%} = 2.57 \times 546.6 = 1404.7$ M EUR. If one further decomposes the VaR, one could look at pure interest rate VaR. In this case one would look at the corresponding δ:

$$
s \times \delta_{Interest} = (-208.83, 635.42, -487.52, 0, 0, 0),
$$

and we would get in the same way $VaR_{99.5\%}^{\text{Interest}} = 2.57 \times 487.9 = 1253.9$. This is the way how one determines which parts of the balance sheet contribute most to the risk. In the concrete example we get (in M EUR):

Item	**Std Deviation**	**99.5% VaR**
Bonds	487.9	1253.9
Equities	375.7	965.5
Credit	204.8	526.3
Simple Sum	1068.4	2745.7
Diversification	-521.8	-1341.0
Total	546.6	**1404.7**

Finally, we want to have a look at the accuracy of the linear approximation, which we have used. For shares there is nothing to do, since the change in value of the asset is linear. Therefore we want to look at the accuracy of the approximation for bonds. For simplicity we assume that the yield curve is flat at $i = 4\%$. In this context we have:

$$
\pi_t(\mathcal{Z}_{(n)})[i + \Delta_i] - \pi_t(\mathcal{Z}_{(n)})[i] = (1 + i + \Delta_i)^{-n} - (1 + i)^{-n}.
$$

We remark that the volatility of bonds is about 60 bps, therefore looking at the 99.5% (which is about $2.57 \times \sigma$) implies, looking at the precision of the approximation, at a shift of c 150 bps.

Term 20 yrs	True Change	Delta Approx.	Error
-300	0.3631	0.2738	-24.5%
-200	0.2165	0.1825	-15.7%
-150	0.1538	0.1369	-11.0%
-100	0.0972	0.0912	-6.1%
-50	0.0461	0.0456	-1.1%
50	-0.0417	-0.0456	9.3%
100	-0.0794	-0.0912	14.8%
150	-0.1136	-0.1369	20.4%
200	-0.1445	-0.1825	26.2%
300	-0.1979	-0.2738	38.3%

Consequently, we see that this approximation has some non negligible errors, which can be mitigated by adjusting the duration accordingly. Another source of such non-linearities are options, where a standard model is inadequate. Hence we want to have a look at a possible solution. In order to do that we need to go back to first principles, which define the model *before* approximations. In our case we assume that the risk factors are following a multi-normal distribution for $Z = \Delta R$. This means that for some of our assets $(\mathcal{A}_i)_{i \in S_A}$ or liabilities $(\mathcal{L}_i)_{i \in S_L}$ the linear approximation is inadequate. The method which we want to show here works in general and can either be applied to one or more of the underlying assets and liabilities. It works for example for plain-vanilla stock options, which can be valued using the :

The price for a *put*-option with payout $C(T, P) = \max(K - S; 0)$ at time t and strike price K and equity price S is given by:

$$
\begin{aligned}
P &= K \times e^{-r \times T} \times \Phi(-d_2) - S_0 \times \Phi(-d_1), \\
d_1 &= \frac{\log(S_0/K) + (r + \sigma^2/2) \times T}{\sigma \times \sqrt{T}}, \\
d_2 &= d_1 - \sigma \times \sqrt{T}, \\
\Phi(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(\frac{-\zeta^2}{2}) d\zeta.
\end{aligned}
$$

The risk factors which enter into the calculation of the price $\pi(\mathcal{A}) := P$ are the following:

– Share price S_t,

– Volatility of the share price σ, and

– The interest rate for the corresponding term r.

In order to understand how these options are synthetically "constructed" one needs to understand the concept of a replicating portfolio. One holds at every point in time a portfolio P_t with the aim that this portfolio matches at time T just the payout of the option mentioned above. In order to construct such portfolios one usually uses the "Greeks". These Greek letters represent the sensitivity of an option in case of a change of the underlying economic parameters such as equity price, interest rate level, etc. We have the following relationships:

$$
\begin{aligned}
\Delta_P &= \frac{\partial P}{\partial S} \\
&= \Phi(d_1) \\
\Gamma &= \frac{\partial^2 P}{\partial S^2}
\end{aligned}
$$

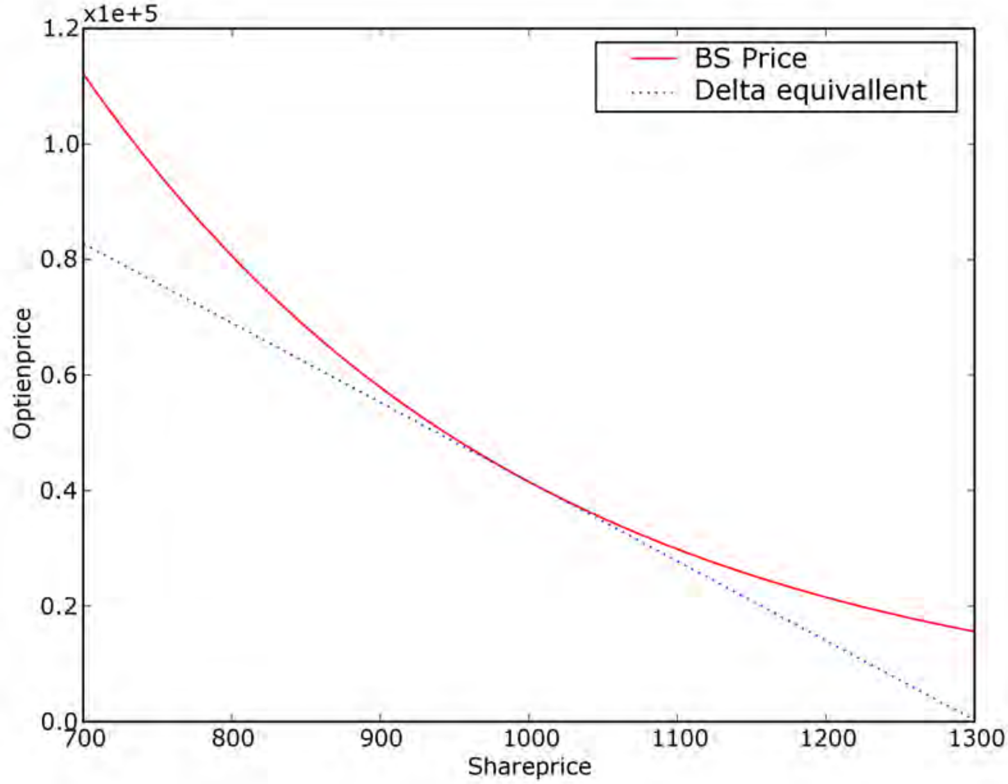

Figure 9.3. δ–Hedging

$$
\begin{aligned}
&= \frac{\Phi'(d_1)}{S \times \sigma \times \sqrt{T}} \\
\Lambda &= \frac{\partial P}{\partial \sigma} \\
&= S \times \Phi'(d_1) \times \sqrt{T-t} \\
P_P &= \frac{\partial P}{\partial r} \\
&= -(T-t) \times K \times e^{-r \times (T-t)} \times \Phi(-d_2)
\end{aligned}
$$

Based on the above partial derivatives it is now possible to define different hedging strategies, one of them being a "delta-hedge". The idea is to define at each point in time t a portfolio P_t consisting of cash and shares, which have the same value and for which the partial derivative with respect to equity price S is the same. Hence we look for a Taylor approximation of order 1 in the variable S. Figure 9.3 shows such a delta hedge. For the concrete example we have the following put option:

Interest	$r = 3.0\%$
Term	10 years
Equity Price	$S_0 = 1000$
Strike	$K = 900$
Volatility	$\sigma = 15\%$
Number of Shares	1000
Value of Put	$P = 41535.7$
Delta	$\Delta_P = -137472$

What becomes obvious is the fact that the hedge is quite good if the stock market does not move too far away during the time between the updating of the replicating portfolio, for example updating the hedge portfolio once a day.

In order to assess the corresponding risk, one can for example use a partial or full simulation approach. In the first case the whole distribution is simulated with a sufficiently big sample and the effective change in capital is evaluated and recorded in order to determine the value of the chosen risk measure such as the VaR or TVaR. One can also use a partial simulation approach remarking that only nonlinear instruments need to be simulated. Here the question is how to "marry" the simulated and the analytical parts. One approach is to use control variables.

In order to describe this approach let's assume that we have one asset $\mathcal{A}$, which is not linearly dependent on the risk factor and assume for sake of simplicity that we have the following:

$$\begin{aligned}\tilde{\pi}_t(x) - \pi_0(x) &= \delta^T \times \Delta(R), \text{ and} \\ \pi_t(x) - \pi_0(x) &= f(\Delta(R)).\end{aligned}$$

In the above we denote with $\tilde{\pi}_t(x)$ and $\pi_t(x)$, the approximated and the "true" change in value, respectively. The function f denotes the "true" change in value for a given $\Delta(R)$. So in order to do a partial simulation approach one needs to do the following:

1. Simulate n (say 50000) times the random variable R, resulting in a series of $(r_k)_{k=1,2,\ldots n}$.
2. Calculate the analytic value of the risk measure C_a^l for the linear approximation.
3. Calculate the simulated value of the risk measure C_r^l for the linear approximation, using the series $(r_k)_{k=1,2,\ldots n}$.
4. Calculate the simulated value of the risk measure C_r^f for the "true" value, using the series $(r_k)_{k=1,2,\ldots n}$.

As a consequence of the weak law of big numbers we have $C_a^l = \lim_{n\to\infty} C_r^l(n)$. Hence using the difference $C_a^l - C_r^l$ as a correction to C_r^f normally improves the quality of the approximation.

The table below show an example for the accuracy of the linear approximation in case of a plain vanilla put option. At time $t = 0$ we assume a stock price of 1000 and we consider a strike for the put option at 900. For this example a sample size of 50000 has been chosen and analysed for the first 500, the first 1000 samples etc.

Sample size	Linear Value	Linear Error	BS - Price Value	BS - Price Error
500	97680	+0.7%	159238	+1.5%
1000	97911	+0.9%	160029	+2.0%
2000	97908	+0.9%	160017	+2.0%
5000	96412	-0.6%	154571	-1.4%
10000	97911	+0.9%	160029	+2.0%
20000	96672	-0.4%	157896	+0.7%
50000	97038	ref	156831	ref

What can be seen from the above example is that the linear model converges much faster and in this case the value of the put-option to the company holding it is *underestimated*.

9.4 Interpreting the Results

This section provides a reporting template which can be used for financial risk management. This template risk report is subdivided into the following parts:

Summary: The aim of this section is to provide a concise summary. In order to get a high level view on the duration gap between assets and liabilities, the corresponding durations are calculated. Furthermore we see the impact of an increase in interest rates of 10 bps and an increase of 1% in equity prices, separately for assets and liabilities only and combined. After these deterministic measures we see some important key measures in terms of VaR, for both a one in ten year (1:10) and a one in 250 year (1:250) event. Here we look at combinations of risk factors. Namely we look separately at equities, bonds, surrenders and the total. This total VaR needs to be compared with the available capital (Market Value). Finally, also the Tail VaR or Expected Shortfall (ES) is shown. The figure underneath shows the required capital for different return periods (separately for assets and for the total). The two red balls represent the VaR in a 1:10 and a 1:250 year event and these numbers reconcile to the table.

Decomposition of VaR: In order to better understand where capital is consumed the total VaR is further decomposed into its components. It is possible to see which parts of the assets and liabilities consume the majority of the capital. In the concrete example we can see that most of the total required capital of 3216 M EUR is consumed by credit risk (2042 M EUR). Furthermore it becomes obvious that the pure ALM risk (in terms of interest rates) is quite small with 764 M EUR. Finally we see that equity risk and hedge funds account still with 764 M EUR and 464 M EUR respectively. At the bottom of the page we see that the surrender risk amounts to 736 M EUR.

Individual Capital Assessment (ICA): In order to be able to compare the model with the regulatory standard ICA model the corresponding results have been included in this section. It needs to be stressed that the ICA model covers more risk factors, but is not as granular for the market risk factors as the own model.

Scenarios: In this section some scenarios are shown and in particular, how the company balance sheet would look after such an event. Section 9.4.2 shows the main characteristics of the scenarios used. The balance sheet items for each of the scenarios follow a typical IFRS balance sheet. The figures underneath the table show the change in shareholders equity and the decomposition of the balance sheet post stress respectively.

Stress Tests: The section stress tests is thought to represent some additional scenarios as described above. The only difference is the fact that here the scenarios are shown in a summarised version and are based on group requirements. The scenarios currently used have been defined by FSA[1] and are quite self-explanatory.

Limits: This section aims to show the limits currently in place to limit the ALM risk. The table shows the limits currently in place: The target which is limited, the threshold, the current level and the headroom. The program has been built up in such a way that every number which is checked against a limit is either printed in green (e.g. within limit) or red (e.g. limit breach).

9.4.1 Notation

In this section the main elements in respect of notation are documented.

VaR	Value at risk, eg the Loss which occurs according to a certain probability. In the analysis a 99.6% VaR is used. This means that the loss represents a loss which in the long run is expected to occur every 250 years. It needs to be noted at this point of time, that analytical models tend to underestimate such losses since the risk factors have been modelled as normally distributed.
1:10	This symbol also relates to a VaR, in this case corresponding to the 90th quantile, e.g. once every 10 years.
Duration	The modified duration which represents the risk intrinsic to a bond portfolio
Sensi Bonds (+10 bps)	The change in value of a bond portfolio if the yield curve is shifted by 10 bps (= 0.1%).
ES 99%	The expected shortfall in a 1 in 100 year event is defined of the average loss looking at all events occurring less than once in 100 years. This measure is more sensitive in the tails than the VaR.
Intangibles	The intangibles in the balance sheet (eg goodwill etc.) In case of an impairment of participation the model reduces the intangibles in a first step.
MR	mathematical reserves for traditional business. They are moving in line with the interest rates.
UDS	Undistributed surplus.
Tax	Taxes and deferred tax assets and liabilities are not modelled.
SHE	Market Value Shareholder funds. This corresponds in principle to the corresponding MCEV.
Δ SHE	Change in SHE in case of a certain scenario.

- The figure *Distribution of Losses* shows the probability density function of the losses. The two red circles represent the VaR 1:10 and in respect of the 99.5% quantile.
- The figure *Cash Flow Profile* shows the inflows (red: bond payments and yellow: premiums) v.s. the outflows (blue: expected claims).
- The figure *Diversification* shows the diversification effect in relation to the main asset risks.
- The figure *Decomposition of required Capital* and *Credit Risk by Rating* show which risk and which credit risk absorbs most of the required capital.

[1] Financial Services Authority in the UK.

9.4.2 Scenarios

In the following section the different used scenarios are defined in some greater detail:

	Credit	Yen	Depr.	FSA	Hard Land.	Depr. ii
i-rate 2 yrs (bps)	0	-399	-399	50	90	-220
i-rate 7 yrs (bps)	0	-316	-316	50	90	-220
i-rate 10 yrs (bps)	0	-323	-323	50	90	-220
i-rate 25 yrs (bps)	0	-303	-303	50	90	-220
Shares (%)	-18	-18	-65	-34	9	-32
Properties (%)	-5	-5	-55	-19	-28	-36
AA Credit Spread (bps)	103	0	103	50	40	110

9.4.3 What can and what cannot be done with this

As indicated above, a model is not reality and hence it is of utmost importance to recognise the limitations of such a model. In this section we try to show some of the limitations of the model currently used. One of the possible risks of this model is that it is overly simplistic.

From a high level point of view the main shortcomings of the model are:

The model is linear: The different risk factors enter linearly into the calculation of the loss. Therefore for options, the corresponding delta equivalent is used. In a next step such effects should be captured better.

The model uses a standard multi-normal distribution.

Management actions: Management actions are not taken into account.

Insurance and operational risks: The model purely focuses on ALM risk.

Dynamic Lapses: Dynamic lapses are also an area where the model used needs refinement.

9.5 Reporting Example

9.5.1 Summary

Assessment and key Figures

- Globally the company has currently not enough risk capital, from a purely economic perspective, to run the corresponding ALM risks, since the margins have become tighter due to the losses in the equity and corporate bond portfolio and widening of the credit spread.
- The statutory reserve set up for GMDB at the year-end (31.12.2008) was 130m EUR whereas the more economic vision used in the MCEV calculation produced a value of 162m EUR. During the first quarter the statutory reserves increased to c230m EUR, affecting the IGD Solvency adversely by c82m EUR.

Data Quality

- Private Equities are currently within the category hedge funds
- The current analysis is still in draft form and is based on data as of 31.12.2008.

Item	Assets	Liabilities	Total
Duration	5.81	5.12	
Sensi Bonds (+10bps)	-177.87	144.98	-32.88
Sensi Equities (+1%)	14.90	-0.21	14.69
1:10 Bonds	1505.30	1174.40	352.94
1:10 Equities	370.32	15.68	380.18
1:10 Total	1809.60	983.14	1600.50
VaR Bonds	3025.60	2360.40	709.39
VaR Sx	–	736.69	736.69
VaR Equities	744.32	31.51	764.13
VaR Total	3637.10	1976.00	3216.90
ES 99%			3098.60
Market Value	–	45343.00	**2780.00**

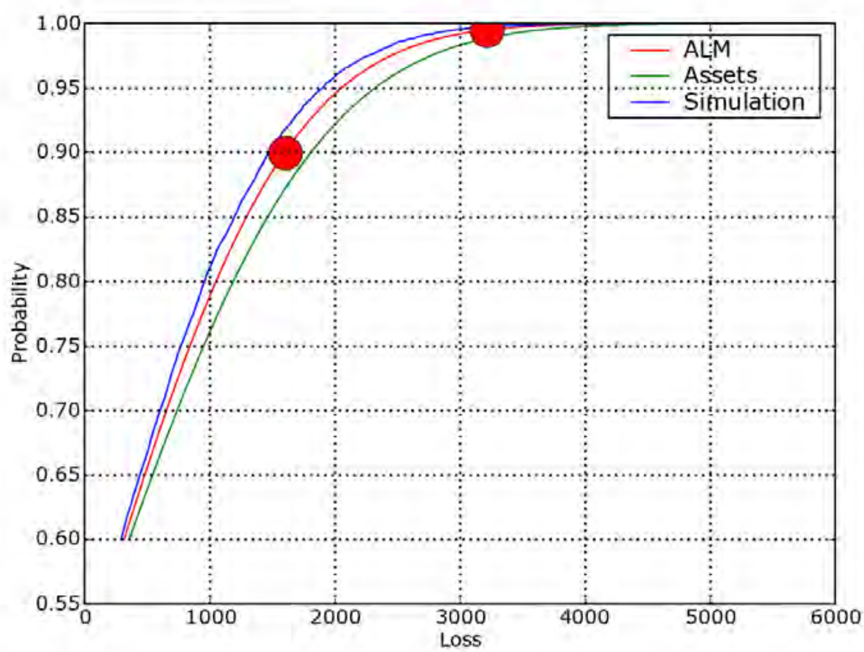

Figure 9.4. Distribution of Losses

9.5.2 Decomposition of VaR

Assessment and key Figures	Data Quality
– The total required risk capital amounts to 3.2 bn EUR (pre Tax) and to c 2.1 bn EUR (post Tax), compared with a available risk capital of c 2.7 bn EUR (after Tax). The 2.1 bn EUR compare to 2.5 bn EUR for the YE2008 ICA calculation. The biggest difference is the fact the ALM Capital does not take into account risk other than market and surrender risks. The biggest additional contribution is the expense risk capital of c 0.6 bn EUR. Adding this to the ALM Capital the two numbers get closer with a difference below 0.1 bn EUR. Overall, both metrics result in similar numbers. – The ALM mismatch consumes about $\frac{1}{2}$ and the equities et al exposure ca. $\frac{1}{2}$ of the total risk capital. This indicates the company has a rather high risk in equities, private equities, properties and hedge funds. In particular, the capital needed for alternative investments is almost 20% of the total available risk capital. – Most of the ALM mismatch stems from the long duration liabilities which are not matched with corresponding assets.	– Replicating the portfolio for one major product line under review – Available risk capital not yet calculated and the current figure is based on an estimation. – GMDB exposure reflected via δ–equivalent for equities and volatility via θ. Interest rate sensitivity not yet reflected.

Item	**Assets**	**Liabilities**	**Total**
Market Value	–	45343.00	**2780.00**
Bonds EUR ¡3	48.26	52.35	11.78
Bonds EUR 3-7	616.54	454.94	162.66
Bonds EUR 8-12	1446.40	810.34	636.14
Bonds EUR 13-24	1007.60	925.96	109.24
Bonds EUR ¿25	26.64	263.45	236.81
Bonds EUR Total	3025.60	2360.40	709.39
Div. Ben.	-119.88	-146.66	-447.24
Bonds GBP ¡3	–	–	–
Bonds GBP 3-7	–	–	–
Bonds GBP 8-12	–	–	–
Bonds GBP 13-24	–	–	–
Bonds GBP ¿25	–	–	–
Bonds GBP Total	–	–	–
Div. Ben.	–	–	–
Bonds USD ¡3	–	–	–
Bonds USD 3-7	–	–	–
Bonds USD 8-12	–	–	–
Bonds USD 13-24	–	–	–
Bonds USD ¿25	–	–	–
Bonds USD Total	–	–	–
Div. Ben.	–	–	–
Bonds CHF ¡3	–	–	–
Bonds CHF 3-7	–	–	–
Bonds CHF 8-12	–	–	–
Bonds CHF 13-24	–	–	–
Bonds CHF ¿25	–	–	–
Bonds CHF Total	–	–	–
Div Ben.	–	–	–
All Bonds	**3025.60**	**2360.40**	**709.39**
Div. Ben.	–	–	–
Credit Risk	2042.20	–	2042.20
Shares MSCIEMU	744.32	–	744.32
Shares MSCICHF	–	–	–
Shares MSCIUK	–	–	–
Shares MSCIUS	–	–	–
All Shares	**744.32**	**31.51**	**764.13**
Div. Ben.	–	–	*-11.70*
FX GBP	–	–	–
FX USD	–	–	–
FX GBP	–	–	–
FX Total	**–**	**–**	**–**
Div. Ben.	–	–	–
Real Estate	131.20	9.45	121.75
Alternatives	464.46	–	464.46
Participations	191.90	–	191.90
Total	**3637.10**	**1976.00**	**3216.90**
Div. Ben.	*-920.34*	*-425.30*	*965 29*
Surrenders	–	736.69	736.69

9.5.3 Figures

Assessment and key Figures	**Data Quality**
– The duration of the bonds with 5.6 years is considerably shorter than the ones of the liabilities with 10.6 years. In part this is due to the special characteristics of a particular insurance portfolio and corresponding analysis are under way. – From a liquidity point of view the company has considerable amounts of bonds maturing within 1 year and 2 years leading to an excess liquidity of ca 2bn EUR and 1bn EUR respectively. – The table relating the shift in asset value for a shift in credit spreads shows clearly the high credit quality of the underlying assets corresponding to a average rating of slightly above AA	– Replicating portfolio for particular product line under review – Derivatives not yet reflected in analysis

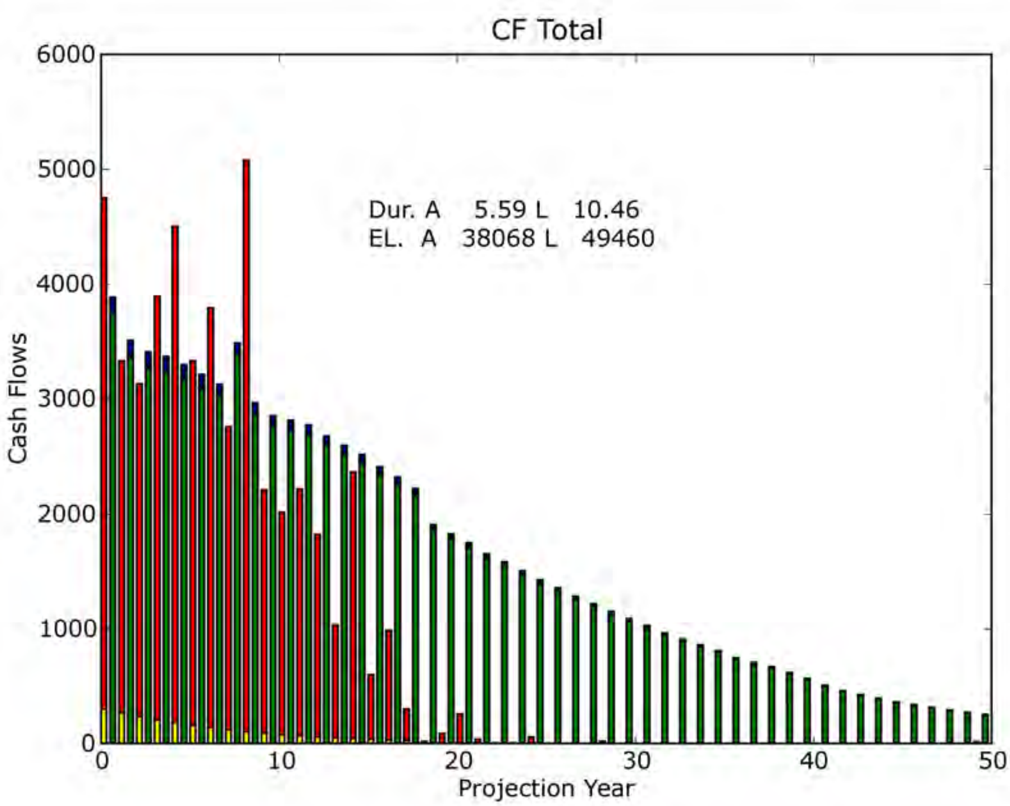

Figure 9.5. Cash Flow Profile

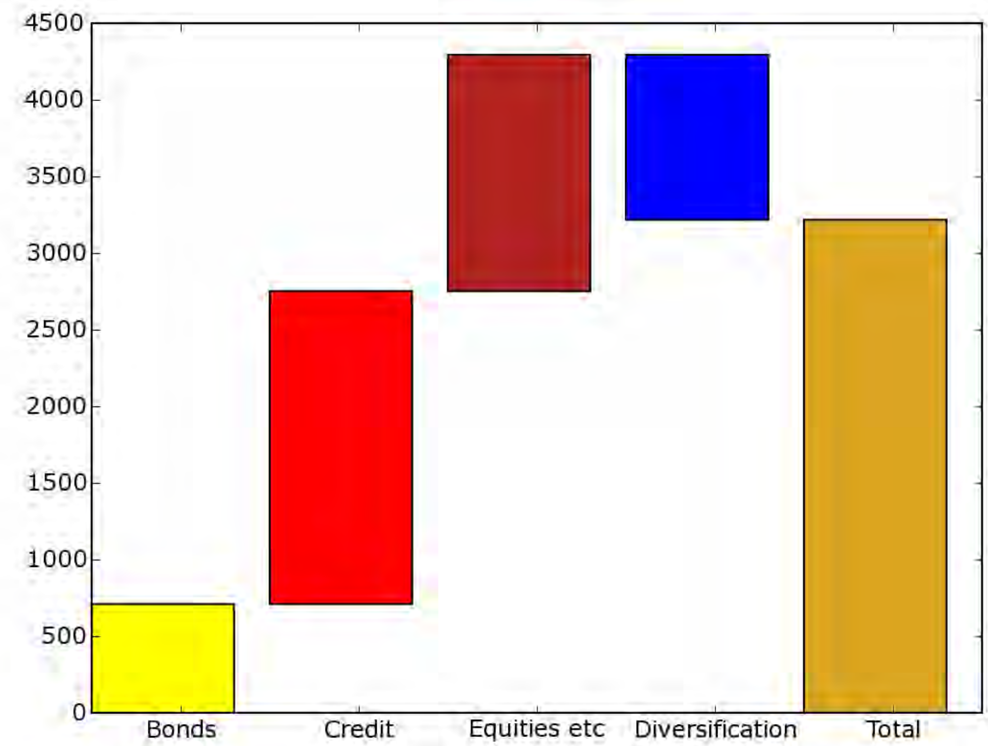

Figure 9.6. Diversification

Credit Quality	**+10 bps Spread**	**Percentile**	**Δ Profit**
EURO AAA	-89.13	1%	2905.30
EURO AA	-23.76	5%	2054.20
EURO A	-42.69	10%	1600.50
EURO BBB	-24.16	33%	538.04
USD AAA	–	66%	-537.70
USD AA	–	90%	-1600.50
USD A	–	95%	-2054.20
USD BBB	–	99%	-2905.30
Total	**-179.76**	99.5%	-3216.90

9.5.4 Scenarios

Assessment and key Figures
- The main three scenarios consist of a widening of credit spreads by a further 50%, a falling of the interest rates to YEN levels and a global severe depression.

Data Quality
- The current analysis is work in progress and is based on data as of 30.12.08

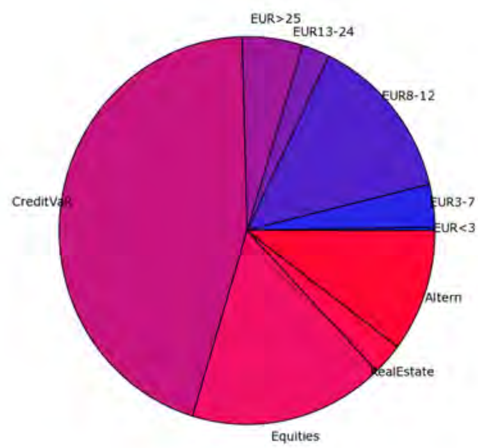

Figure 9.7. Decomposition of required Capital

Item	**Start**	**Y to D**	**Credit**	**YEN**	**Depression**
Cash	6221.10	6221.10	6221.10	6187.00	6187.00
Bonds	35734.00	35734.00	33730.00	41567.00	39564.00
Shares	4468.30	4207.50	4207.50	4207.50	3499.50
Properties	1137.10	1081.70	1081.70	1081.70	528.15
Hedge Funds	595.70	521.24	521.24	521.24	59.57
Private Equity	64.00	56.00	12.80	56.00	6.40
Loans	1452.10	1452.10	1452.10	1452.10	1452.10
Unit Linked Assets	14330.00	14330.00	14330.00	14330.00	14330.00
Other Assets	3612.00	3612.00	3612.00	3612.00	3612.00
Intangibles	204.90	152.75	167.65	152.75	-63.30
MR	47066.00	47066.00	47326.00	51773.00	51773.00
Unit Linked Liabilities	14330.00	14330.00	14330.00	14330.00	14330.00
UDS	–	–	–	–	37.60
Debt	–	–	–	–	–
Deferred Tax	127.70	127.70	127.70	127.70	127.70
Other Lia	3801.90	3797.90	3797.90	3797.90	3758.00
SHE	2493.00	2046.20	-245.49	3139.00	-851.01
Δ SHE		**-446.82**	**-2738.50**	**645.97**	**-3344.00**

Year-to-date: The interest rates decreased by c 200 bps at the short and 30 bps at the long end. At the same time credit spreads widened between 90 bps (AAA) and 380 bps (BBB). Stock markets reduced by 35+%.
Credit Scenario: An additional spread widening of 50%.
YEN Scenario: Based on the current scenario, interest rates have been lowered to levels where the Yen was at the its deepest level
Depression: YEN interest rates, a 40% credit spread widening and a cumulative reduction of 65% for shares, PE, HF and properties.

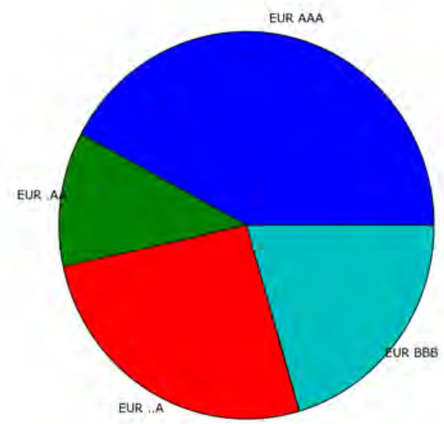

Figure 9.8. Credit Risk by Rating

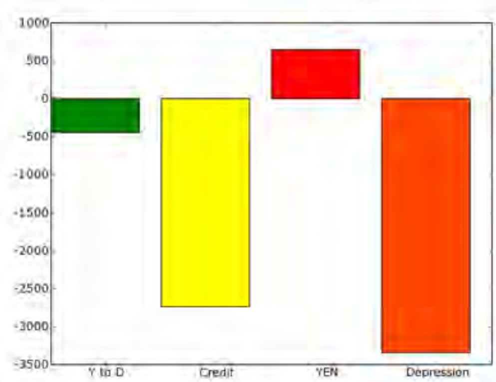

Figure 9.9. Δ SHE for the scenarios

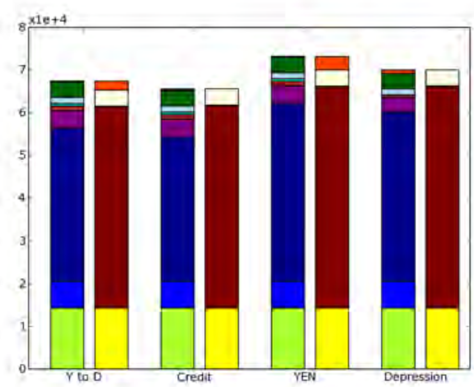

Figure 9.10. BS for Scenarios

9.5.5 ICA Capital

Assessment and key Figures

- The falls in available economic capital and the increase in capital requirements are largely driven by:
  – Falls in equity markets and increases in credit risks have lead to minimum investment guarantees "biting", with a direct burn-through impact on shareholder assets.
  – Saving products experienced significant erosion in value of future profits (VIF), with asset returns over the year close to minimum guarantees.
  – Falls in the equity market meant unit linked contracts experienced an increase in Guaranteed Minimum Death Benefit (GMDB) risk.
- Changes in YE2008 stress methodology, in particular, the "softening" of equity and credit spread tests were key to keeping the funds from going into deficit on an economic basis.
- The company has completed an equity de-risking initiative, leading to a further fall in capital requirement for equity risk.
- The company believes a significant part of the credit spread widening is linked to liquidity premium and for some products, creates an artificial and unnecessarily high capital requirement.
- A separate exercise will need to be performed to quantify the reputational risk associated to structured products that have been sold with the underlying guarantees provided by third parties. The default risk is borne by the client but a reputational risk would remain with the company (total reserves 4,844m). This is not part of the YE08 SSTEC requirements, but will be investigated given the potentially material impact.

in M EUR	**YE 2007**	**YE 2008**	**1Q2009**	**4Q2009**
Available Economic Capital	3657	2780		
Reg. Capital Required	2083	2508		
Cover	176%	111%		
Diversification Benefit	39%	41%		
MV Assets	59390	56390		
MV Liabilities	54674	52558		

9.5.6 Stress Tests

Nr	**Name**	**Equity**	**Δ Assets**	**Δ Lia**	**Δ Equity**
	B/S	2493.00			
	Y to D	2046.20	-450.81	-3.98	-446.82
1	Equities -10%	1763.00	-603.96	126.01	-729.97
2	Equities -20%	1479.90	-757.11	256.01	-1013.10
3	Equities -40%	1043.60	-1063.40	386.01	-1449.40
4	Equity Vola + 10%	2044.10	-450.81	-1.88	-448.92
5	Property -7.5%	1908.00	-529.70	55.33	-585.03
6	Property -15%	1769.80	-608.58	114.65	-446.82
7	Property -30%	1558.40	-766.35	168.28	-934.64
8	I rates -50 bps	2210.60	438.52	720.94	-282.42
9	I rates -100 bps	2375.00	1327.80	1445.90	-118.01
10	I rates -200 bps	2703.80	3106.50	2895.70	210.79
11	I rates Twist: long down	1988.70	834.64	1339.00	-504.31
12	I rates Twist: long up	2102.20	-1750.40	-1359.60	-390.84
13	Cred spread +50 bps	1147.40	-1349.60	-3.98	-1345.60
14	Cred spread +100 bps	248.60	-2248.40	-3.98	-2244.40
15	Cred spread +200 bps	-1549.00	-4046.00	-3.98	-4042.00
16	FSA	530.28	-2703.00	-740.28	-1962.70
17	FSA Hard Landing	1287.10	-2533.00	-1327.00	-1205.90
18	FSA Depression 2010	206.99	875.04	3161.10	-2286.00

9.5.7 Limits

In this section the various limits are checked:

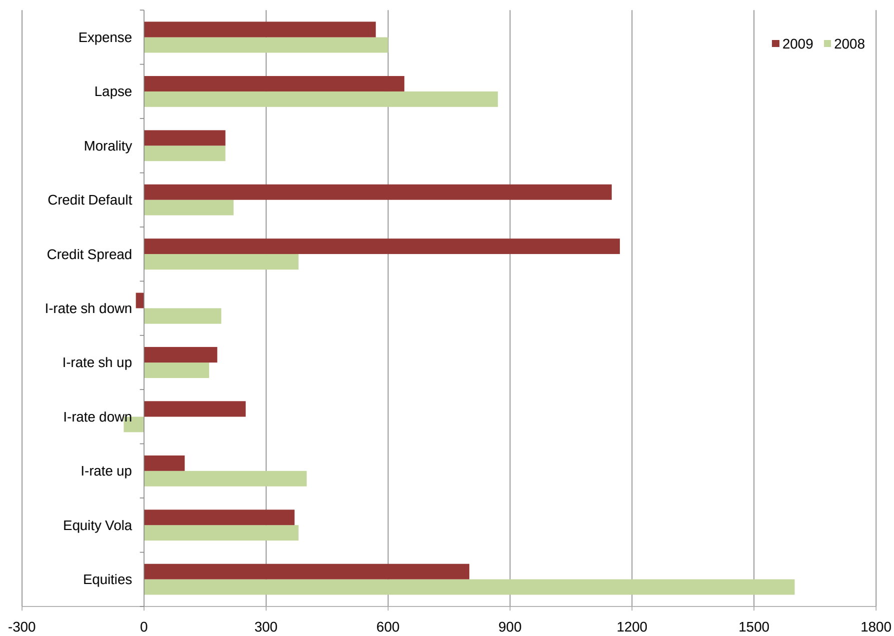

Figure 9.11. Required ICA Capital by Risk

Limit	Threshold	Actual	Headroom
VaR Equity Asset (20%) ◇	409.24	744.32	**-335.08**
Total VaR (80%) ◇	1636.90	3216.90	**-1580.00**
Base Point Sensitivity	87.98	-32.88	55.10
Alternatives VaR (10%) ◇	204.62	464.46	**-259.84**
Credit VaR (20%) ◇	409.24	2042.20	**-1633.00**
Credit Scenario SHE > 0	0.00	3139.00	3139.00
Yen Scenario SHE > 0	0.00	-245.49	245.49
Combined 2 Scenario SHE > 0	0.00	1287.10	1287.10
Depression Scenario SHE > 0	0.00	-851.01	851.01
Properties VaR (10%)	204.62	131.20	73.42

9.6 Summary Reporting Example

Figure 9.12 provides an example of a summary on a page for the financial risk a company is facing. The aim is to be concise and also action oriented. Hence the table envisages the following entries:

Name The name of the risk is indicated together with a measure for its size, such as the amount of assets affected, a risk measure etc.

Risk Category The risk category aims to indicate which type of risk is described, such as credit risk, liquidity risk, ...

Risk The risk is described in a concise manner in order that a knowledgeable third party can understand, what the risk is.

Actions This one is the most important column, since the mitigation actions performed and planned are described. This helps to see the development with respect to the corresponding risk

Remarks Here additional information needed to better understand the issue is documented.

The principle for the writing of such reportings must be *relevant, concise and action oriented*.

Name	Risk Category	Risk	Actions	Remarks
Credit Defaults	Credit Risk	The company had and has substantial credit exposure, in particular towards the banking sector and might suffer credit defaults	• Reduction in holdings in Italian and Greek government bonds • Reduction (€ 100 M) of counterparty exposure to X • Regular exposure control and reporting	The counterparty exposure vis-à-vis our strategic partners is a particular concern due to its strategic rationale.
Hedge Funds	Market Risk	An unclear asset investment rationale and risk controlling was found at the beginning of this year. As a consequence the company faces a variety of risks – exposure € c500M	• Decision to exit hedge funds • So far € 70M have been sold	Due to the introduction of gates and side pockets within the hedge funds the reduction of hedge fund investments will take some time
Equity and GMDB	Market Risk	As a consequence of the credit crises in 2008, the market value of equities dropped considerably together with a spike in equity-volatility. As a consequence the company was and still is vunerable against equity market movements.	• Reduction of € >3bn so far • Hedging of Equity exposure with derivatives • Mitigation actions with respect to GMDB exposure examined • Continuous monitoring	• Equity exposure (€ c900M) protected with put option strategy maturing Mar 2010 • GMDB reserve € 160M exposed towards volatility increase
Interest Rate	Market Risk	Since many of the bond portfolios covering insurance liabilities are not perfectly matched, the company is vunerable against interest rate movements. Most liability portfolios are of a shorter duration than the assets.	• Cash flow analysis performed for major blocks of business. As a consequence the corresponding risk for life portfolios is adequate • GI portfolios were short duration in liabilities and a reduction in asset duration is under way	Regular ALM risk controlling to be built up
Lapse Risk	Insurance Risk	As a consequence of the financial crisis, which now also effects the "real" economy, a lot of our policyholders lapsed their policies, with a potentially adverse impact on our franchise value	• Actions are under way to safeguard the franchise value of the company • The distribution channels have been asked to report back on the steps undertaken so far • Regular reporting of lapse experience	A side effect of the financial crisis is the fact that some of our clients switch their unit linked policies into traditional ones, which are less profitable

Figure 9.12. Financial Risk Reporting

A. Notes on stochastic integration

This appendix summarises definitions and results in the area of stochastic integration and martingales. For obvious reasons we will not present proofs for all results. In fact we will only give a survey of various results and then refer to the corresponding literature. Foremost we refer to the monographs [Pro90] and [IW81].

A.1 Stochastic processes and martingales

Definition A.1.1. *A probability space $(\Omega, \mathcal{A}, P)$ is called* filtered*, if there exists a family of σ-algebras $F = (\mathcal{F}_t)_{t\geq 0}$ such that*

1. $\mathcal{F}_0 \supset \{A \in \mathcal{A} | P(A) = 0\}$,
2. $\mathcal{F}_s \subset \mathcal{F}_t$ *for* $s \leq t$.

The filtration is called right continuous, if $\mathcal{F}_t = \bigcap_{t'>t} \mathcal{F}_{t'}, \forall t \geq 0$.

Definition A.1.2. *A random variable $T : \Omega \to [0, \infty]$ is a* stopping time*, if $\{T \leq t\} \in \mathcal{F}_t$ for all $t \in \mathbb{R}_+$.*

Theorem A.1.3. *T is a stopping time if and only if $\{T < t\} \in \mathcal{F}_t$ for all $t \in \mathbb{R}_+$. ([Pro90] Thm. 1.1.1.)*

Definition A.1.4. *Let X, Y be stochastic processes. X is called* modification of Y*, if*

$$X_t = Y_t \qquad\qquad P\text{–almost surely for all } t.$$

X and Y are indistinguishable*, if*

$$X_t = Y_t, \forall t \qquad\qquad P\text{–almost surely.}$$

Definition A.1.5. 1. *A stochastic process is called* càdlàg *(continue à droite, limites à gauche), if its trajectories are right continuous with left limits.*

2. *A stochastic process is called* càglàd *(continue à gauche, limites à droite), if its trajectories are left continuous with right limits.*
3. *A stochastic process is* adapted*, if $X_t \in \mathcal{F}_t$ (X_t is $\mathcal{F}_t$-measurable).*

Theorem A.1.6. 1. *Let Λ be an open set and X be an adapted càdlàg process. Then $T := \inf\{t \in \mathbb{R}_+ : X_t \in \Lambda\}$ is a stopping time.*

2. *Let S, T be stopping times and $\alpha > 1$. Then the following random variables are also stopping times:* $\min(S,T)$, $\max(S,T)$, $S+T$, $\alpha \cdot T$.

Proof. [Pro90] Thm. 1.1.3 and Thm. 1.1.5.

Definition A.1.7. *Let $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, P)$ be a filtered probability space. A stochastic process X is called* martingale, *if*

– $X_t \in L^1(\Omega, \mathcal{A}, P)$, *i.e.* $E[|X_t|] < \infty$,

– $E[X_t|\mathcal{F}_s] = X_s$ *holds for* $s < t$.

Remark A.1.8. If one replaces in the previous equation "=" by "≤" ("≥"), the process X is called supermartingale (submartingale).

Theorem A.1.9. *Let X be a supermartingale. Then the following conditions are equivalent:*

1. *The mapping $T \to \mathbb{R}, t \mapsto E[X_t]$ is right continuous.*
2. *There exists a unique modification Y of X which is càdlàg.*

Proof. [Pro90] Thm. 1.2.9.

Theorem A.1.10. *Let X be a martingale. Then there exists a unique càdlàg modification Y of X.*

Theorem A.1.11 (Doob's stopping theorem). *Let X be a right continuous martingale with closure X_∞, i.e. $X_t = E[X_\infty|\mathcal{F}_t]$. Moreover, let S and T be stopping times such that $S \leq T$ P–almost surely. Then the following statements hold:*

1. $X_S,\ X_T \in L^1(\Omega, \mathcal{A}, P)$,
2. $X_S = E[X_T|\mathcal{F}_S]$.

Proof. [Pro90] Thm. 1.2.16.

Definition A.1.12. *Let X be a stochastic process and T be a stopping time. The stopped stochastic process $(X_t^T)_{t\geq 0}$ is defined by $X_t^T = X_{\min(t,T)}$ for all $t \geq 0$.*

Theorem A.1.13 (Jensen's inequality). *Let $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function and $X \in L^1(\Omega, \mathcal{A}, P)$ with $\phi(X) \in L^1(\Omega, \mathcal{A}, P)$. Furthermore let $\mathcal{G}$ be a σ-algebra. Then the following inequality holds:*

$$\phi \circ E[X|\mathcal{G}] \leq E[\phi(X)|\mathcal{G}].$$

Proof. [Pro90] Thm. 1.2.19.

A.2 Stochastic integrals

In this section we present a short introduction to the theory of stochastic integration. We follow the approach of [Pro90].

Essentially one can understand a stochastic integral with respect to a semimartingale as a pathwise Stieltjes integral, and the latter should be known from lectures in analysis. The main idea for this type of integrals is to consider the limit of sums of the form

$$\sum f(T_k)\ (T_{k+1} - T_k)$$

for partitions with decreasing mesh size. In the following we assume that a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, P)$, which satisfies the common regularity conditions, is given.

Definition A.2.1. *1. A stochatic process H is called* simple predictable*, if it can be represented as*

$$H_t = H_0 \cdot \chi_{\{0\}}(t) + \sum_{i=1}^{n} H_i \cdot \chi_{]T_i, T_{i+1}]}(t),$$

where

$$0 = T_1 \leq \ldots \leq T_{n+1} < \infty$$

is a finite family of stopping times and the $H_i \in \mathcal{F}_t$, $(H_i)_{i=0,\ldots,n}$ are finite P-almost everywhere.

The set of simple predictable processes will be denoted by $\mathbb{S}$. Furthermore, $\mathbb{S}_u$ denotes the set $\mathbb{S}$ equipped with the topology of uniformly convergence in (t,ω) on $\mathbb{R} \times L^\infty(\Omega, \mathcal{A}, P)$.

2. The vector space of finite, real valued random variables equipped with the convergence in probability is denoted by $\mathbb{L}^0$.

Next, we want to define the expression $\int H\, dX$ for certain processes $(X_t)_{t\in\mathbb{R}}$ and $(H_t)_{t\in\mathbb{R}}$. In order to be able to call such an operator, denoted by I_X, integral, it should at least be linear and a theorem like the Lebesgue convergence theorem should hold.

For the convergence theorem we assume the following continuity: If H^n converges uniformly to H, then $I_X(H^n)$ should converge in probability to $I_X(H)$.

Given a process X. We define $I_X : \mathbb{S} \to \mathbb{L}^0$ by

$$I_X(H) = H_0\, X_0 + \sum_{i=1}^{n} H_i \left(X^{T_i} - X^{T_{i+1}}\right),$$

where

$$H_t = H_0 \cdot \chi_{\{0\}}(t) + \sum_{i=1}^{n} H_i \cdot \chi_{]T_i, T_{i+1}]}(t).$$

This definition of $I_X(H)$ is independent of the representation of H.

Definition A.2.2 (Total semimartingale). *A stochastic process $(X_t)_{t\geq 0}$ is called* total semimartingale*, if*

1. *X is càdlàg and*
2. *I_X is a continuous mapping from $\mathbb{S}_u$ to $\mathbb{L}^0$.*

Definition A.2.3 (Semimartingale). *A stochastic process $(X_t)_{t\geq 0}$ is a* semimartingale*, if X^t (compare with Definition A.1.12) is a total semimartingale for all $t \in [0,\infty[$.*

Remark A.2.4. Thus semimartingles are defined as well behaving integrators.

The following theorem summarizes the most important properties of the operator I_X:

Theorem A.2.5. *1. The set of all semimartingales is a vector space.*

2. Let Q be a measure which is absolutely continuous with respect to P. Then every P-semimartingale is also a Q-semimartingale.

3. Let $(P_n)_{n\in\mathbb{N}}$ be a sequence of probability measures such that $(X_t)_{t\geq 0}$ is a P_n-semimartingale for each n. Then $(X_t)_{t\geq 0}$ is an R-semimartingale, for $R = \sum_{n\in\mathbb{N}} \lambda_n P_n$, where $\sum_{n\in\mathbb{N}} \lambda_n = 1$.

4. (Stricker's Theorem) Let X be a semimartingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$, and $(\mathcal{G}_t)_{t\geq 0}$ be a sub-filtration of $(\mathcal{F}_t)_{t\geq 0}$ such that X is adapted to $(\mathcal{G}_t)_{t\geq 0}$. Then X is a $\mathcal{G}$-semimartingale.

Proof. The statements follow from the definition of a semimartingale. The proofs, which are recommended as an exercise to the reader, can be found in [Pro90] Chapter II.2.

Now we want to characterise the class of semimartingales.

Theorem A.2.6. *Every adapted process with càdlàg paths and finite variation on compact sets is a semimartingale.*

Proof. This theorem is based on the fact that

$$|I_X(H)| \leq ||H||_u \int_0^\infty |dX_s|,$$

where $\int_0^\infty |dX_s|$ denotes the total variation.

Theorem A.2.7. *Every square integrable martingale with càdlàg paths is a semimartingale.*

Proof. Let X be a square integrable martingale with $X_0 = 0$, $H \in \mathbb{S}$. The continuity of the operators I_X is a consequence of the following inequality:

$$\begin{aligned}
E\left[(I_X(H))^2\right] &= E\left[\left(\sum_{i=0}^n H_i \left(X^{T_i} - X^{T_{i+1}}\right)\right)^2\right] \\
&= E\left[\sum_{i=0}^n H_i^2 \left(X^{T_i} - X^{T_{i+1}}\right)^2\right] \\
&\leq ||H||_u^2 \, E\left[\sum_{i=0}^n \left(X^{T_i} - X^{T_{i+1}}\right)^2\right] \\
&= ||H||_u^2 \, E\left[\sum_{i=0}^n \left({X^{T_i}}^2 - {X^{T_{i+1}}}^2\right)\right] \\
&= ||H||_u^2 \, E\left[{X_{T^{n+1}}}^2\right] \\
&\leq ||H||_u^2 \, E\left[{X_{T^\infty}}^2\right].
\end{aligned}$$

Example A.2.8. Brownian motion is a semimartingale.

Now, after defining semimartingales, we want to enlarge the class of integrands. A class which is very well suited for this purpose is the set of càglàd processes. We will use this class, since for it the proofs remain relatively simple.

Definition A.2.9. *The set of adapted càdlàg (càglàd) processes is denoted by* $\mathbb{D}$ *(*$\mathbb{L}$*, respectively). Further,* $b\mathbb{L}$ *denotes the set of processes* $X \in \mathbb{L}$ *with bounded paths.*

Up to now we are familiar with the topology of the uniform convergence (on $\mathbb{S}_u$) and the topology of the convergence in probability on $\mathbb{L}^0$. We define a further notion of convergence.

Definition A.2.10. *Let* $t \geq 0$ *and* H *be a stochastic process. Then we set*

$$H_t^* = \sup_{0 \leq s \leq t} |H_s|.$$

A sequence $(H^n)_{n\in\mathbb{N}}$ *converges uniformly on compact sets in probability (short: convergence in ucp-topology) to* H*, if*

$$(H^n - H)_t^* \to 0$$

in probability for $n \to \infty$ *and all* $t \geq 0$.

$\mathbb{D}_{ucp}$*,* $\mathbb{L}_{ucp}$ *and* $\mathbb{S}_{ucp}$ *denote the corresponding sets equipped with the ucp-topology.*

Remark A.2.11. 1. The ucp-topology is metrizable. An equivalent metric is for example:

$$d(X, Y) = \sum_{i=1}^{\infty} \frac{1}{2^n} E\left[\min(1, (X - Y)_n^*\right].$$

2. $\mathbb{D}_{ucp}$ is a complete metric space.

The following theorem is essential for the extension of the integral I_X.

Theorem A.2.12. *The vectorspace* $\mathbb{S}$ *is dense in* $\mathbb{L}$ *with respect to the ucp-topology.*

Proof. [Pro90] Thm. 2.4.10.

Note that, I_X can be extended if we can show that I_X is continuous. To show the continuity we start with a definition.

Definition A.2.13. *Let* $H \in \mathbb{S}$ *and* X *be a semimartingale. Then we define* $J_X : \mathbb{S} \to \mathbb{D}$ *by*

$$J_X(H) = H_0\, X_0 + \sum_{i=0}^{n} H_i \left(X^{T_i} - X^{T_{i+1}}\right),$$

where

$$H_t = H_0 \cdot \chi_{\{0\}}(t) + \sum_{i=1}^{n} H_i \cdot \chi_{]T_i, T_{i+1}]}(t)$$

for $H_i \in \mathcal{F}_{T_i}$ *and stopping times* $0 = T_1 \leq \ldots \leq T_{n+1} < \infty$.

Definition A.2.14 (Stochastic integral). *Let $H \in \mathbb{S}$ and X be a càdlàg process. Then $J_X(H)$ is called the stochastic integral of H with respect to X and we use the notations:*

$$H \cdot X := \int H_s \, dX_s := J_X(H).$$

Now we have defined the stochastic integral on $\mathbb{S}$, and we want to extend it onto $\mathbb{L}$. For the extension we need the following theorem.

Theorem A.2.15. *Let X be a semimartingale. Then the mapping $J_X : \mathbb{S}_{ucp} \to \mathbb{D}_{ucp}$ is continuous. Also the extension of J_X onto $\mathbb{S}_{ucp}$ will be called stochastic integral, and the notations of Definition A.2.14 will be used also for the extension.*

Proof. [Pro90] Thm. 2.4.11.

Remark A.2.16. To extend J_X onto $\mathbb{D}$ we use the fact that $\mathbb{D}_{ucp}$ is a *complete* metric space.

The process $J_X(H) = \int H_s \, dX_s$ evaluated at the time $t \geq 0$ will be denoted by

$$H \cdot X_{\ t} := \int_0^t H_s \, dX_s := \int_{[0,t]} H_s \, dX_s.$$

A.3 Properties of the stochastic integral

After defining the stochastic integral we will now summarise its properties. We will concentrate on the statements without giving proofs.

Theorem A.3.1. *1. Let T be a stopping time. Then $(H \cdot X)^T = H \cdot \chi_{[0,T]} \cdot X = H \cdot X^T$.*

2. Let $G, H \in \mathbb{L}$ and X be a semimartingale. Then also $Y := H \cdot X$ is a semartingale. Furthermore, we have

$$G \cdot Y = G \cdot (H \cdot X) = (G \cdot H) \cdot X.$$

Proof. [Pro90] Thm. 2.5.12 and 2.5.19.

Definition A.3.2. *Let X be a càdlàg process. Then we define*

$$\begin{aligned} X_-(t) &= \lim_{s \uparrow t} X(s), \\ \Delta X(t) &= X(t) - X_-(t). \end{aligned}$$

Definition A.3.3. *A random partition σ of $\mathbb{R}$ is a finite sequence of stopping times such that*

$$0 = T_0 \leq T_1 \leq \ldots \leq T_n < \infty.$$

A sequence $(\sigma_n)_{n \in \mathbb{N}}$ of random partitions of $\mathbb{R}$ converges to the identity, if the following conditions hold:

1. $\lim\limits_{n \to \infty} (\sup_k T_k^n) = \infty$ *P-almost surely,*
2. $||\sigma_n|| := \sup_k |T_{k+1}^n - T_k^n|$ *converges P-almost surely to* 0.

Let Y be a process and σ be a random partition. Then we define

$$Y^{\sigma} := Y_0 \cdot \chi_{\{0\}} + \sum_k Y_{T_k} \cdot \chi_{]T_k, T_{k+1}]}.$$

Remark A.3.4. It is easy to show that

$$\int Y_s^{\sigma}\, dX_s = Y_0\, X_0 + \sum_k Y_{T_k} \left(X^{T_{k+1}} - X^{T_k} \right)$$

for all semimartingales X and all Y in $\mathbb{S}$, $\mathbb{D}$ and $\mathbb{L}$.

Using random partitions one calculate the stochastic integral by the following theorem.

Theorem A.3.5. *Let X be a semimartingale, $Y \in \mathbb{D}$ and $(\sigma_n)_{n\in\mathbb{N}}$ be a sequence of random partitions which converges to the identity. Then*

$$\int_{0+} Y_s^{\sigma_n}\, dX_s = \sum_k Y_{T_k^n} \left(X^{T_{k+1}^n} - X^{T_k^n} \right)$$

converges in ucp-topology towards the stochastic integral $\int (Y_-)\, dX$.

Proof. [Pro90] Thm. 2.5.21.

Definition A.3.6. *Let X and Y be semimartingales. Then we define*

$$\begin{aligned} [X,X] &= ([X,X]_t)_{t\geq 0} \text{ the quadratic variation process by} \\ [X,X] &:= X^2 - 2\int X_- dX, \end{aligned}$$

and accordingly

$$[X,Y] := XY - \int X_-\, dY - \int Y_-\, dX$$

is the covariation process.

Theorem A.3.7. *Let X be a semimartingale. Then the following statements hold:*

1. *$[X,X]$ is càdlàg, monotone increasing and adapted.*
2. *$[X,X]_0 = X_0^2$ and $\Delta[X,X] = (\Delta X)^2$.*
3. *If a sequence $(\sigma_n)_{n\in\mathbb{N}}$ of random partitions converges to 1, then*

$$X_0^2 + \sum_i (X^{T_{i+1}^n} - X^{T_i^n})^2 \longrightarrow [X,X] \text{ in ucp-topology for } n \to \infty.$$

4. *Let T be a stopping time. Then $[X^T, X] = [X, X^T] = [X^T, X^T] = [X,X]^T$.*

Proof. [Pro90] Thm. 2.6.22.

Remark A.3.8. – The mapping $(X, Y) \mapsto [X, Y]$ is bilinear and symmetric.

– The polarization identity holds:

$$[X, Y] = \frac{1}{2}\left([X+Y, X+Y] - [X, X] - [Y, Y]\right).$$

Theorem A.3.9. *The* bracket process $[X, Y]$ *of two semimartingales* X *and* Y *has paths of bounded variation on compact sets and it is a semimartingale.*

Proof. [Pro90] Cor. 2.6.1.

Theorem A.3.10 (Partial integration).

$$d(XY) = X_- \, dY + Y_- \, dX + d[X, Y].$$

Proof. [Pro90] Cor. 2.6.2.

Theorem A.3.11. *Let* M *be a local martigale. Then the first and second of the following statements are equivalent and the third is a consequence of the first two.*

1. M *is martingale with* $E[M_t^2] \leq \infty \; \forall t \geq 0$*,*
2. $E\left[[M, M]_t\right] < \infty \; \forall t \geq 0$*,*
3. $E[M_t^2] = E\left[[M, M]_t\right] \forall t \geq 0$*.*

Proof. [Pro90] Cor. 2.6.4.

Theorem A.3.12. *Let* X, Y *be semimartingales and* $H, K \in \mathbb{L}$*. Then the following statements hold:*

1. $[H \cdot X, K \cdot Y]_t = \int_0^t H_s K_s d[X, Y]_s \; \forall t \geq 0$*,*
2. $[H \cdot X, H \cdot X]_t = \int_0^t H_s^2 d[X, X]_s \; \forall t \geq 0$*.*

Proof. [Pro90] Thm. 2.6.29.

Theorem A.3.13 (Itô-formula). *Let* X *be a semimartingale and* $f \in C^2(\mathbb{R})$*. Then*

$$\begin{aligned} f(X_t) - f(X_0) &= \int_{0+}^t f'(X_s^-) dX_s + \frac{1}{2}\int_{0+}^t f''(X_s^-) d[X, X]_s^{cont} \\ &+ \sum_{0<s\leq t} \left\{ f(X_s) - f(X_s^-) - f'(X_s^-) \Delta X_s \right\}. \end{aligned}$$

Proof. [Pro90] Thm. 2.7.32.

Remark A.3.14. A function $f \in C^2(\mathbb{R})$ has the deterministic integral representation

$$f(t) - f(0) = \int_0^t f'(s) ds.$$

For a stochastic integral two further terms appear. The term

$$\frac{1}{2}\int_{0+}^{t} f''(X_s^-)d[X,X]_s^{cont}$$

is due to the quadratic variation of the process and the term

$$\sum_{0<s\le t}\left\{f(X_s)-f(X_s^-)-f'(X_s^-)\Delta X_s\right\}$$

is due to the jumps of the process.

Theorem A.3.15 (Transformation theorem). *Let V be a stochastic process with right continuous paths of bounded variation. Furthermore, let $f \in C^1(\mathbb{R})$. Then $(f(V_t))_{t\ge 0}$ is a process of bounded variation and*

$$f(V_t)-f(V_0)=\int_{0+}^{t} f'(V_{s_-})\,dV_s+\sum_{0<s\le t}\big(f(V_s)-f(V_{s_-})-f'(V_{s_-})\Delta V_s\big)\,.$$

Theorem A.3.16 (ItÃ´-formula). *Let X be a continuous semimartingale and $f \in C^2(\mathbb{R})$. Then also $f(X)$ is a semimartingale and it satisfies*

$$f(X_t)-f(X_0)=\int_{0+}^{t} f'(X_s)dX_s+\frac{1}{2}\int_{0+}^{t} f''(X_s)d[X,X]_s.$$

References

[Bau91] H. Bauer. *Wahrscheinlichkeitstheorie.* De Gruyter, 1991.

[Bau92] H. Bauer. *Mass- und Integrationstheorie.* De Gruyter, 1992.

[BGH+86] N. L. Bowers, H.U. Gerber, J.C. Hickman, D.C. Jones, and C. J. Nesbitt. *Actuarial Mathematics.* Society of Actuaries, 1986.

[BP80] Bellhouse and Panjer. Stochastic modelling of interest rates with applications to life contingencies. *Journal of Risk and Insurance*, 47:91–110, 1980.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[Büh92] H. Bühlmann. Stochastic discounting. *Insurance Mathematics and Economics*, 11/2:113–127, 1992.

[Büh95] H. Bühlmann. Life insurance with stochastic interest rates. In G. Ottaviani, editor, *Financial Risk in Insurance*, pages 1–24. Springer, 1995.

[CHB89] J. Cox, C. Huang, and Bhattachary. Option pricing theory and its applications. In Constantinides, editor, *Theory of Valuation*, pages 272–288. Rowman and Littlefield Publishers, 1989.

[CFO08] CFO Forum. *Market Consistent Embedded Value.* http://www.cfoforum.nl/embedded_value.html.

[CIR85] J. Cox, J. Ingersoll, and S. Ross. A theory of term structure of interest rates. *Econometrica*, 53:385–408, 1985.

[CW90] K. L. Chung and R. J. Williams. *Introduction to stochastic Integration.* Birkhäuser, 2 edition, 1990.

[DAV09] DAV Unterarbeitsgruppen "Rechnungsgrundlagen der Pflegeversicherung" und "Todesfallrisiko". *Herleitung der Rechnungsgrundlagen DAV 2008 P für die Pflegerenten(zusatz)versicherung*, und *Raucher- und Nichtrauchersterbetafeln für Lebensversicherungen mit Todesfallcharakter*, und *Herleitung der Sterbetafel DAV 2008 T für Lebensversicherungen mit Todesfallcharakter. Blätter der DGVFM*, 30/1:31-140, 141-187, 189-224, 2009.

[Doo53] J. L. Doob. *Stochastic Processes.* Wiley, 1953.

[Dot90] M. Dothan. *Prices in Financial Markets.* Oxford University Press, 1990.

[DS57] N. Dunford and J. T. Schwartz. *Linear Operators Part 1: General Theory.* Wiley - Interscience, 1957.

[Duf88] D. Duffie. *Security markets: Stochastic Models.* Academic Press, 1988.

[Duf92] D. Duffie. *Dynamic Asset Pricing Theory.* Prinston University Press, 1992.

[DVJ88] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes.* Springer, 1988.

[Fel50] W. Feller. *An Introduction to probability theory and its applications.* Wiley, 1950.

[Fis78] M. Fisz. *Wahrscheinlichkeitsrechnung und mathematische Statistik.* Deutscher Verlag der Wissenschaften, 1978.

[Ger95] H. U. Gerber. *Life Insurance Mathematics.* Springer, 2 edition, 1995.

[HK79] J. M. Harrison and D. Kreps. Martingales and multiperiod security markets. *Journal of Economic Theory*, 20:381–401, 1979.

[HN96] O. Hesselager and R. Norberg. On probability distributions of present values in life insurance. *J. Insurance Math. Econom.*, 18/1:135–142, 1996.

[Hoe69] J. M. Hoem. Markov chain models in life insurance. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, 9:91–107, 1969.

[HP81] J. M. Harrison and S. R. Pliska. Martingales, stochastic integrals and continuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981.

[Hua91] C. Huang. Lecture notes on advanced financial econometrics. Technical report, Sloan School of Management, MIT, Massachusetts, 1991.

[Hul97] J. C. Hull. *Options, Futures and other Derivatives.* Prentice Hall, 1997.

[IW81] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* North-Holland, 1981.

[KP92] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Applications of Mathematics.* Springer, 1992.

[KS88] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, 1988.

[Mol92] C. M. Moller. Numerical evaluation of markov transition probabilities based on the discretized product integral. *Scand. Actuarial J.*, pages 76–87, 1992.

[Mol95] C. M. Moller. A counting process approach to stochastic interest. *Insurance Mathematice and Economics*, 17:181–192, 1995.

[MS07] T. Moller and M.Steffenson. *Market-Valuation Methods in Life and Pension Insurance* Cambridge University Press, 2007

[NM96] R. Norberg and C. M. Moller. Thiele's differential equation by stochastic interest of diffusion type. *Scand. Actuarial J.*, 1996/1:37–49, 1996.

[Nor90] R. Norberg. Payment measures, interest and discounting. *Scand. Actuarial J.*, pages 14–33, 1990.

[Nor91] R. Norberg. Reserves in life and pension insurance. *Scand. Actuarial J.*, pages 3–24, 1991.

[Nor92] R. Norberg. Hattendorff's theorem and thiele's differential equation generalized. *Scand. Actuarial J.*, pages 2–14, 1992.

[Nor94] R. Norberg. Differential equations for higher order moments of present values in life insurance. *Insurance: Mathematics and Economics*, pages 171–180, 1994.

[Nor95a] R. Norberg. Stochastic calculus in actuarial science. Working paper, Laboratory of Actuarial Mathematics University of Copenhagen, 1995.

[Nor95b] R. Norberg. A time-continuous markov chain interest model with applications to insurance. *J. Appl. Stoch. Models and Data Anal.*, 11:245–256, 1995.

[Nor96a] R. Norberg. Addendum to hattendorff's theorem and thiele's differential equation generalized. *Scand. Actuarial J.*, pages 2–14, 1996.

[Nor96b] R. Norberg. Bonus in life insurance: Principles and prognoses in a stochastic environment. Working paper, Laboratory of Actuarial Mathematics University of Copenhagen, 1996.

[Nor98] R. Norberg. Vasicek beyond the normal. Working paper, Laboratory of Actuarial Mathematics University of Copenhagen, 1998.

[Par94a] G. Parker. Limiting distributions of the present value of a portfolio. *ASTIN Bulletin*, 24/1:47–60, 1994.

[Par94b] G. Parker. Stochastic analysis of a portfolio of endowment insurance policies. *Scand. Actuarial J.*, 2:119–130, 1994.

[Par94c] G. Parker. Two stochastic approaches for discounting actuarial functions. *ASTIN Bulletin*, 24/2:167–181, 1994.

[Per94] S. A. Persson. *Pricing Life Insurance Contracts under Financial Uncertainty.* PhD thesis, Norwegian School of Economics and Business Administration, Bergen, 1994.

[Pli97] S. R. Pliska. *Introduction to Mathematical Finance. Discrete Time Models.* Blackwell Publishers, 1997.

[Pro90] P. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Applications of Mathematics.* Springer, 1990.

[PT93] H. Peter and J. R. Trippel. Auswertungen und Vergleich der Sterblichkeit bei den Einzelkapitalversicherungen der Schweizerischen Lebensversicherungs- und Rentenanstalt in den Jahren 1981 - 1990. *Mitteilungen der Schweiz. Vereinigung der Versicherungsmathematiker*, pages 23–44, 1993.

[RH90] H. Ramlau-Hansen. Thiele's differential equation as a tool in product developpement in life insurance. *Scand. Actuarial J.*, pages 97–104, 1990.

[Rog97] L. C. G. Rogers. The potential approach to the term structure of interest rates and foreign exchange rates. *Mathematical Finance*, pages 157–176, 1997.

[Vas77] O. Vasicek. An equilibrium characterisation of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[WH86] H. Wolthuis and I. Van Hoek. Stochastic models for life contingencies. *Insurance Mathematics and Economics*, 5:217–254, 1986.

[Wil86] A. D. Wilkie. Some applications of stochastic interest models. *Journal of the Institute of Actuaries Student Soc.*, 29:25–52, 1986.

[Wil95] A. D. Wilkie. More on a stocastic asset model for actuarial use. *British Actuarial Journal*, 1, 1995.

[Wol85] H. Wolthuis. Hattendorf's theorem for a continuous-time markov model. *Scand. Actuarial J.*, 70, 1985.

[Wol88] H. Wolthuis. *Savings and Risk Processes in Life Contingencies.* PhD thesis, University of Amsterdam, 1988.

Index