Contents

1 Introduction 3
  1.1 The mutual scheme 3
  1.2 The deterministic approach 5
  1.3 Issues for further study 11

2 Payment streams and interest 13
  2.1 Basic notions of payments and interest 13
  2.2 Application to loans 18

3 Mortality 21
  3.1 Aggregate mortality 21
  3.2 Some standard mortality laws 26
  3.3 Actuarial notation 28
  3.4 Select mortality 30

4 Insurance of a single life 32
  4.1 Some standard forms of insurance 32
  4.2 The principle of equivalence 36
  4.3 Prospective reserves 38
  4.4 Thiele’s differential equation 44
  4.5 Probability distributions 47
  4.6 The stochastic process point of view 48

5 Expenses 50
  5.1 A single life insurance policy 50
  5.2 The general multi-state policy 53

6 Multi-life insurances 54
  6.1 Insurances depending on the number of survivors 54

7 Markov chains in life insurance 58
  7.1 The insurance policy as a stochastic process 58
  7.2 The time-continuous Markov chain 59
  7.3 Applications 64
  7.4 Selection phenomena 68
CONTENTS

7.5 The standard multi-state contract ......................... 70
7.6 Select mortality revisited ................................ 77
7.7 Higher order moments of present values .................. 80

8 Safety loadings and bonus .................................... 86
8.1 General considerations .................................... 86
8.2 First and second order bases ............................... 87
8.3 The technical surplus and how it emerges ............... 88
8.4 Dividends and bonus .................................... 90
8.5 Bonus prognoses ......................................... 94
8.6 Examples .................................................. 99
8.7 Including expenses ....................................... 102
8.8 Discussions ............................................... 104

9 Inference in the Markov model .............................. 108
9.1 Statistical analysis of truncated life times ................ 108
9.2 Parametric inference in the Markov model ............... 114
9.3 Confidence regions ...................................... 118
9.4 Piecewise constant intensities ............................ 120
9.5 Impact of the censoring scheme ........................... 122

A Calculus ...................................................... 1
B Indicator functions .......................................... 1
C Distribution of the number of occurring events ........... 1
D Asymptotic results from statistics .......................... 1
E The G82M mortality table ................................... 1
F List of terms .................................................. 1
Chapter 1

Introduction

1.1 The mutual scheme

A. The bank savings contract. Upon celebrating his 55th anniversary Mr. (55) – let us just call him so – decides to invest money to secure himself economically in his old age. His first plan is to deposit a capital $U_0$, say, on a savings account and draw the entire amount with earned compound interest in 15 years, on his 70th birthday. The account bears interest at rate $i = 0.045$ ( = 4.5\%) per year. In one year the capital will increase to $U_1 = (1 + i)U_0$, in two years it will increase to $U_2 = (1 + i)^2U_0$, and so on until in 15 years it will have accumulated to

\[ U_{15} = (1 + i)^{15}U_0 = 1.045^{15}U_0 = 1.935U_0. \]  

(1.1)

Now, this simple calculation takes no account of the fact that (55) will die sooner or later, maybe sooner than 15 years. Suppose he has no heirs, or he dislikes the ones he has, so that in the event of death before 70 he would consider his savings worthless. Checking population statistics he learns that about 75\% of those who are 55 will survive to 70. Thus, the relevant future prospects of the savings contract are:

- with probability 0.75 (55) survives to 70 and will then possess $U_{15}$;
- with probability 0.25 (55) dies before 70 and loses the capital.

In this perspective the expected amount at (55)’s disposal after 15 years is

\[ 0.75U_{15}. \]  

(1.2)

B. A small scale mutual fund. Having thought things over, (55) changes his mind and seeks to make the following mutual arrangement with (55)∗ and (55)∗∗, who are also 55 years old and are in exactly the same situation as (55). Each of the three deposits $U_0$ on the savings account, and those who survive to 70, if any, will then share the total accumulated capital $3U_{15}$ equally. The
prospects of this scheme are given in Table 1.1, where $S$ and $D$ signify 'survival' and 'death', respectively, $L_{70}$ is the number of survivors at age 70, and $3U_{15}/L_{70}$ is the amount at disposal per survivor (undefined if $L_{70} = 0$). There are now the following possibilities:

- with probability 0.422 (55) survives to 70 together with (55)* and (55)** and will then possess $U_{15}$;
- with probability 2·0.141 = 0.282 (55) survives to 70 together with one more survivor and will then possess 1.5$U_{15}$;
- with probability 0.047 (55) survives to 70 while both (55)* and (55)** die (may they rest in peace) and he will cash the total savings 3$U_{15}$;
- with probability 0.25 (55) dies before 70 and will get nothing.

This scheme is clearly superior to the one described in Paragraph A, with separate individual savings contracts: If (55) survives to 70, which is the only scenario of interest to him, he will cash no less than the amount $U_{15}$ he would cash under the individual scheme, and it is likely that he will get more. As compared with (1.2), the expected amount at (55)’s disposal after 15 years is now

$$0.422 \cdot U_{15} + 0.282 \cdot 1.5 \cdot U_{15} + 0.047 \cdot 3 U_{15} = 0.985 U_{15}.$$  

The point is that under the present scheme the savings of those who die are bequeathed to the survivors. The total savings are retained in the group so that nothing is left to others unless the unlikely thing happens that the whole group goes extinct within the term of the contract. This is essentially the kind of solidarity that is underlying pension funds. From the point of view of the group as a whole, the probability that all three participants will die before 70 is only 0.016, which should be compared to the probability 0.25 that (55) will die and lose everything under the individual savings program.

C. A large scale mutual scheme. Inspired by the success of the mutual fund already on the small scale of three participants, (55) starts to play with the idea of extending it to a large number of participants. Let us assume that
a total number of \( L_{55} \) persons, who are in exactly the same situation as \( (55) \), agree to join a scheme similar to the one described for the three. Then the total savings after 15 years amount to \( L_{55} U_{15} \), which yields an individual share equal to

\[
\frac{L_{55} U_{15}}{L_{70}}
\]

to each of the \( L_{70} \) (say) survivors if \( L_{70} > 0 \). By the strong law of large numbers, \( L_{70}/L_{55} \to 0.75 \) as \( L_{55} \to \infty \). Therefore, as the number of participants increases, the individual share per survivor tends to

\[
\frac{1}{0.75} U_{15},
\]

and in the limit \( (55) \) is faced with the following situation:

- with probability 0.75 he survives to 70 and gets \( \frac{1}{0.75} U_{15} \);
- with probability 0.25 he dies before 70 and gets nothing.

The expected amount at \( (55) \)'s disposal after 15 years converges to

\[
0.75 \cdot \frac{1}{0.75} U_{15} = U_{15},
\]

the same as (1.1). Thus, the bequest mechanism of the mutual scheme has raised \( (55) \)'s expectancies of future pension to what they would be with the individual savings contract if he were immortal. This is what we could expect since, in an infinitely large scheme, some will survive to 70 for sure and share the total savings. All the money will remain in the scheme and will be redistributed among its members by the lottery mechanism of death and survival.

### 1.2 The deterministic approach

**A. Life and death in the classical actuarial perspective.** Insurance mathematics is widely held to be boring. Hopefully, the present text will not sustain that prejudice. It is admittedly a fact, however, that actuaries use to cheer themselves up with jokes like: “What is the difference between an English and a Sicilian actuary? Well, the English actuary can predict fairly precisely how many English citizens will die next year. Likewise, the Sicilian actuary can predict how many Sicilians will die next year, but he can tell their names as well.” The English actuary is definitely the more typical representative of the actuarial profession since he takes a purely statistical view of mortality. Still he is able to analyse insurance problems adequately since what insurance is essentially about, is to average out the randomness associated with the individual risks.

Contemporary life insurance is based on the paradigm of the large scheme studied in Paragraph 1.1.C. Small mutual funds are no longer so common. Insurance today is dominated by insurance companies that sell insurance as a service in an open market. The typical insurance company serves tens and some even hundreds of thousands of customers, sufficiently many to ensure that
the survival rates are stable as assumed in Paragraph 1.1.C. On the basis of statistical investigations the actuary constructs a so-called decrement series, which takes as its starting point a large number \( l_0 \) (say) of new-born and, for each age \( x = 1, 2, \ldots \), specifies the number of survivors, \( l_x \).

Table 1.2: Excerpt from the mortality table G82M

<table>
<thead>
<tr>
<th>( x )</th>
<th>0</th>
<th>25</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ell_x )</td>
<td>100 000</td>
<td>98 083</td>
<td>91 119</td>
<td>82 339</td>
<td>65 024</td>
<td>37 167</td>
<td>9 783</td>
</tr>
<tr>
<td>( d_x )</td>
<td>58</td>
<td>119</td>
<td>617</td>
<td>1 275</td>
<td>2 345</td>
<td>3 111</td>
<td>1 845</td>
</tr>
<tr>
<td>( q_x )</td>
<td>0.000579</td>
<td>0.001206</td>
<td>0.006774</td>
<td>0.015484</td>
<td>0.036069</td>
<td>0.083711</td>
<td>0.188617</td>
</tr>
<tr>
<td>( p_x )</td>
<td>0.999421</td>
<td>0.998794</td>
<td>0.993226</td>
<td>0.984516</td>
<td>0.963931</td>
<td>0.916289</td>
<td>0.811383</td>
</tr>
</tbody>
</table>

Table 1.2 is an excerpt of the G82M table, shown in full in Appendix E, which is used by Danish insurers to describe the mortality of insured Danish males. The second row in the table lists some entries of the decrement series. It shows e.g. that about 65% of all new-born will celebrate their 70th anniversary. The number of survivors decreases with age; \( \ell_x \geq \ell_{x+1} \). The difference

\[ d_x = \ell_x - \ell_{x+1} \]

is the number of deaths at age \( x \) (more precisely, between the ages of \( x \) and \( x + 1 \)). These numbers are shown in the third row of the table. It is seen that the number of deaths peaks somewhere around age 80. From this it cannot be concluded that 80 is the “most dangerous age”. The actuary measures the mortality at any age \( x \) by the one-year mortality rate

\[ q_x = \frac{d_x}{\ell_x} \]

which tells how big proportion of those who survive to age \( x \) will die within one year. This rate, shown in the fourth row of the table, increases with the age. For instance, 8.4% of the 80 years old will die within a year, whereas 18.9% of the 90 years old will die within a year. The bottom row shows the one year survival rates

\[ p_x = \frac{\ell_{x+1}}{\ell_x} = 1 - q_x \]

We shall present some typical forms of products that an insurance company can offer to (55) and see how they compare with the corresponding arrangements, if any, that he can make with his bank.

### B. Bank saving versus pension insurance

Being unable to find his perfect matches (55)*, (55)**..., our hero (55) abandons the idea of creating a mutual fund and resumes discussions with his bank. The bank operates with a fixed annual interest rate \( i \), whereby a unit deposited today accumulates to \((1 + i)^t\)
in $t$ years. Accordingly, the present value today of a unit withdrawn in $t$ years is $v^t$, where

$$v = 1/(1 + i), \quad (1.4)$$
called the annual discount factor since it is what the bank would pay you today if you sell to it (discount) a default-free claim of 1 in one year.

A general savings contract over $n$ years specifies that at each time $t = 0, ..., n$ (55) is to deposit an amount $c_t$ (contribution) and withdraw an amount $b_t$ (benefit). As a matter of definition we take $c_t$ and $b_t$ to be non-negative, and $b_t = 0$ whenever $c_t > 0$ and vice versa. The net payment of deposit less withdrawal at time $t$ is denoted by $a_t = c_t - b_t$ (amount). At any time $t$ the cash balance of the account, henceforth also called the retrospective reserve, is the total of past (including present) deposits less withdrawals compounded with interest,

$$U_t = \sum_{j=0}^{t} (1 + i)^{t-j}a_j. \quad (1.5)$$

It develops in accordance with the “forward” recursive scheme

$$U_t = U_{t-1}(1 + i) + a_t, \quad (1.6)$$

commencing from $U_{-1} = 0$ (per definition). Each year (55) will receive from the bank a statement of account with the calculation (1.6), showing how the current balance emerges from the previous balance, the interest earned meanwhile, and the current deposit/withdrawal.

The balance of a savings account must always be non-negative,

$$U_t \geq 0, \quad (1.7)$$

and by time $n$, when the contract terminates and the account is closed, it must be null,

$$U_n = 0. \quad (1.8)$$

In the course of the contract the bank must maintain a so-called prospective reserve to meet its future obligations to the customer. At time $t$ the adequate reserve is

$$V_t = \sum_{j=t+1}^{n} v^{j-t}(-a_j), \quad (1.9)$$

the present value of future withdrawals less deposits. Similar to (1.6), the prospective reserve is calculated by the “backward” recursive scheme

$$V_t = v(-a_{t+1} + V_{t+1}), \quad (1.10)$$
The constraint (1.8) can be recast in terms of the prospective reserve as

$$V_{-1} = 0.$$  

(1.11)

It implies that, at any time $t$ the retrospective reserve equals the prospective reserve,

$$U_t = V_t,$$

as is easily verified: insert the defining expression (1.5) into (1.8), split the sum $\sum_{j=0}^{n} \rightarrow \sum_{t=0}^{j} + \sum_{n-j+1}^{t}$, multiply with $(1+i)^{t-n}$, and .

The bank proposes a savings contract according to which (55) saves a fixed amount $c$ annually in 15 years, at ages 55,...,69, and thereafter withdraws a fixed amount of $b = 100,000$ Danish crowns annually in 10 years, at ages 70,...,79.

The annual interest rate is $i = 0.045$. Using (1.5) (or rather repeated use of (1.6)), the bank calculates

$$U_{24} = \sum_{j=0}^{14} (1 + 1)^{24-j}c - \sum_{j=15}^{24} (1 + 1)^{24-j}b,$$

and, by (1.8), determines

$$c = \frac{\sum_{j=15}^{24} v^j}{\sum_{j=0}^{14} v^j} b = 0.381 b,$$

(1.12)

that is, 38,100 Danish crowns. The bank’s sales agent argues convincingly that, due to interest, this amount is considerably smaller than $(10/15) b = 0.667 b = 66,700$, which is what (55) would have to save per year if he should choose to tuck his money away under his mattress.

Still, to (55) 38,100 Danish crowns is a considerable expense. He believes in a life before death, and it should be blessed with the joys of high consumption. Now! He talks to an insurance agent, and is delighted to learn that, under a life annuity policy designed precisely as the savings scheme, he would have to deposit an annual amount of only 25,800 crowns.

The insurance agent explains: The calculations of the bank depend only on the amounts $a_t$ and would apply to any other customer $(x)$, say, who would enter into the same contract at age $x$. Thus, to the bank the customer is Mr. X, unknown. To the insurance company, however, he is not just Mr. X, but the significant Mr. $(x)$, who is now $x$ years old. Working under the hypothesis that $(x)$ is one of the $\ell_x$ survivors at age $x$ in the decrement series and that they all hold identical contracts, the insurer offers $(x)$ a general life annuity policy whereby each deposit or withdrawal is conditional on survival. For the entire portfolio of $\ell_x$ participants the retrospective reserve at time $t$, denoted by $U_t^p$, ...
CHAPTER 1. INTRODUCTION

develops as

\[ U_t^p = U_{t-1}^p (1 + i) + a_t \ell_{x+t} \]  
\[ = \sum_{j=0}^{t} (1 + i)^{t-j} a_j \ell_{x+j} . \]  

(1.13)  
(1.14)

The prospective portfolio reserve, denoted by \( V_t^p \), develops as

\[ V_t^p = v(-a_{t+1} \ell_{x+t+1} + V_{t+1}^p) \]  
\[ = \sum_{j=t+1}^{n} v^{j-t} (-a_j) \ell_{x+j} . \]  

(1.15)  
(1.16)

In particular, for the life annuity analogue to (55)'s savings contract, we have \( a_t = c \) for \( t = 0, \ldots, 14 \) and \( a_t = -b \) for \( t = 15, \ldots, 24 \). The annual premium comes out of \( V_{14}^p = 0 \) (or \( U_{24}^p = 0 \)) as

\[ c = \frac{\sum_{j=12}^{24} v^j \ell_{55+j}}{\sum_{j=0}^{24} v^j \ell_{55+j}} b = 0.258 b , \]  

(1.17)

i.e. 25,800 crowns.

Inspection of the expressions in (1.12) and (1.17) shows that the latter is smaller due to the fact that \( \ell_x \) is decreasing. This phenomenon is known as mortality bequest since the savings of the diseased are passed on (bequeathed) to the survivors. We shall pursue this issue in Paragraph D below.

C. A life assurance contract. Suppose, contrary to the former hypothesis, that (55) has dependents whom he very much cares for. Then he might be concerned that, if he should die within the term of the contract, the survivors in the pension scheme will be his heirs, leaving his wife and kids with nothing. He figures that, in the event of his untimely death before the age of 70, the family would need a down payment of \( b = 1000,000 \) Danish crowns to compensate the loss of their bread-winner. The insurance agent triumphantly (and correctly) asserts that the bank can not help in this matter; the benefit of \( b \) must be raised immediately since (55) could die tomorrow, and it would be meaningless to borrow the money since full repayment of the loan would be due immediately upon death. The insurance company, however, can offer (55) a so-called term life assurance policy that provides the wanted death benefit against an affordable annual premium of 0.0170 \( b = 17,000 \) crowns.

The calculation for a general life insurance written by \((x)\), with sum insured \( b_t \) payable at time \( t \) if he dies in year \( t \) (between the times \( t-1 \) and \( t \)) and premium \( c_t \) payable at time \( t \) if he is then alive, \( t = 0, \ldots, n \), goes by the retrospective scheme,

\[ U_t^p = U_{t-1}^p (1 + i) + c_t \ell_{x+t} - b_t d_{x+t-1} \]  
\[ = \sum_{j=0}^{t} (1 + i)^{t-j} (c_j \ell_{x+j} - b_j d_{x+j-1}) , \]  

(1.18)  
(1.19)
or by the prospective scheme

\[
V^P_t = v \left( b_{t+1} d_{x+t} - c_{t+1} \ell_{x+t+1} + V^P_{t+1} \right) 
\]

(1.20)

\[
= \sum_{j=t+1}^{n} v^{j-t} (b_j d_{x+j-1} - c_j \ell_{x+j}) ,
\]

(1.21)

In particular, for the policy written by (55), with level annual premium \( c_0 = \ldots = c_{14} = c \), say, and \( b_1 = \ldots = b_{15} = b = 1000,000 \), we find from \( V^P_{14} = 0 \) (or \( U^P_{15} = 0 \)) that

\[
c = \frac{\sum_{j=1}^{15} v^j d_{55+j-1}}{\sum_{j=0}^{14} v^j \ell_{55+j}} b = 0.0170 b ,
\]

(1.22)

as previously announced.

**D. Individual reserves and mortality bequest.** In the insurance schemes described above the contracts of diseased members are void, and the reserves of the portfolio are therefore to be shared equally between the survivors at any time. Thus, we introduce the individual retrospective and prospective reserves at time \( t \),

\[
U_t = U^P_t/\ell_{x+t} , \quad V_t = V^P_t/\ell_{x+t} .
\]

Since we have established that \( U_t = V_t \), we shall henceforth be referring to them as the individual reserve or just the reserve.

For the pension insurance contract in Paragraph B we get from (1.13) that the reserve develops as

\[
U_t = U_{t-1} \frac{\ell_{x+t-1}}{\ell_{x+t}} (1 + i) + a_t
\]

(1.23)

The bequest mechanism is clearly seen by comparing (1.6) to (1.23): the additional term \( U_{t-1} d_{x+t-1} (1 + i)/\ell_{x+t} \) in the latter is precisely the share per survivor of the savings left over to them by those who died during the year. Virtually, the mortality bequest acts as an increase of the interest rate.

Table 1.3 shows how the reserve develops for the bank solution and for the insurance solution. In both cases the balance reaches maximum at time 14, as could be expected. In general the insurance scheme requires a smaller reserve than the bank savings scheme. You should be able to explain why.

For the life assurance described in the previous paragraph we obtain from (1.18) that the reserve develops as

\[
U_t = U_{t-1} (1 + i) + c_t - \frac{d_{x+t-1}}{\ell_{x+t}} (b_t - U_{t-1} (1 + i)) .
\]

(1.24)
Table 1.3: reserve $U_t = V_t$ for savings account (SA) and for life annuity (LA)

<table>
<thead>
<tr>
<th>$t$</th>
<th>0</th>
<th>4</th>
<th>9</th>
<th>14</th>
<th>19</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>0.381</td>
<td>2.083</td>
<td>4.678</td>
<td>7.913</td>
<td>4.390</td>
<td>0</td>
</tr>
<tr>
<td>LA</td>
<td>0.258</td>
<td>1.448</td>
<td>3.433</td>
<td>6.376</td>
<td>3.718</td>
<td>0</td>
</tr>
</tbody>
</table>

The prospective counterpart is readily obtained from (1.20). Solving out $V_{t+1}$, we get

$$V_{t+1} = V_t (1 + i) + c_{t+1} - \frac{d_{x+t+1}}{l_{x+t+1}} (b_{t+1} - V_t (1 + i)) .$$

which is equivalent to (1.24). Another version is obtained by solving out $V_t$ and subtracting it from $V_{t+1}$:

$$V_{t+1} - V_t = v \left( V_{t+1} i + \frac{l_{x+t+1}}{l_{x+t}} c_{t+1} - \frac{d_{x+t}}{l_{x+t}} (b_{t+1} - V_t (1 + i)) \right) .$$

1.3 Issues for further study

The simple pieces of actuarial reasoning in the previous sections involve two constituents, interest and mortality, and these are to be studied separately in the two following chapters. Next we shall escalate the discussion to more complex situations. For instance, suppose (55) wants a life insurance that is paid out only if his wife survives him, or with a sum insured that depends on the number of children that are still alive at the time of his death. Or he may demand a pension payable during disability or unemployment. We need also to study the risk associated with insurance, which is due to the uncertain developments of the insurance portfolio and the investment portfolio: the deaths in a finite insurance portfolio do not follow the mortality table (1.2) exactly, and the interest earned on the investments may differ from the assumed 4.5% per year, and neither can be predicted precisely at the outset when the policies are issued.
In a scheme of the classical mutual type the problem was how to share ex-
isting money in a fair manner. A typical insurance contract of today, however,
specifies that certain benefits will be paid contingent on certain events related
only to the individual insured under the contract. An insurance company work-
ing with this concept faces a risk of insolvency, which has to be controlled in
some way. With these issues in mind, we now commence our studies of the
theory of life insurance.

The reader is advised to consult the following authoritative textbooks on the
subject: [5] (a good classic – sharpen your German!), [4], [15] (lexicographic,
treats virtually every variation of standard insurance products, and includes a
good chapter on population theory), [27] (an excellent early text based on prob-
abilistic models, placing emphasis on risk considerations), [7], [9] (an original
approach to the field – sharpen your French!), and [11] (the most recent of the
mentioned texts, still classical in its orientation).
Chapter 2

Payment streams and interest

2.1 Basic notions of payments and interest

A. Streams of payments. What is money? In lack of a precise definition you may add up the face values of the coins and notes you find in your purse and say that the total amount is your money. Now, if you do this each time you open your purse, you will realize that the development of the amount over time is important. In the context of insurance and finance the time aspect is essential since payments are usually regulated by a contract valid over some period of time. We shall give some formal mathematical structure to the notion of payment streams and, referring to Appendix A), deal only with their properties as functions of time and not venture to discuss their possible stochastic properties for the time being.

To fix ideas and terminology, consider a financial contract commencing at time 0 and terminating at a later time \( t \) (\( t \leq \infty \)), say, and denote by \( A_t \) the total amount paid in respect of the contract during the time interval \([0, t]\). The payment function \( \{ A_t \}_{t \geq 0} \) is assumed to be the difference of two non-decreasing, finite-valued functions representing incomes and outgoes, respectively, and is thus of finite variation (FV). Furthermore, the payment function is assumed to be right-continuous (RC). From a practical point of view this assumption is just a matter of convention, stating that the balance of the account changes on and after the time of any deposit or withdrawal. From a mathematical point of view it is convenient, since payment functions can then serve as integrators. In fact, we shall restrict attention to payment functions that are piecewise differentiable (PD):

\[
A_t = A_0 + \int_0^t a_\tau \, d\tau + \sum_{0 < \tau \leq t} (A_\tau - A_{\tau^-}).
\]  

(2.1)

The integral adds up payments that fall due continuously, and the sum adds up
lump sum payments. In differential form (2.1) reads
\[ dA = a_t \, dt + A_t - A_{t-}. \] (2.2)

It seems natural to count incomes as positive and outgoes as negative. Sometimes, and in particular in the context of insurance, it is convenient to work with outgoes less incomes, and to avoid ugly minus signs we introduce \( B = -A \).

Having explained what payments are, let us now see how they accumulate under the force of interest. There are monographs written especially for actuaries on the topic, see [17] and [12], but we will gather the basics of the theory in only a few lines.

B. Interest. Suppose money is currently invested on (or borrowed from) an account that bears interest. This means that a unit deposited on the account at time \( u \) gives the account holder the right to cash, at any other time \( t \), a certain amount \( v(t, u) \), typically different from 1. The function \( v \) must be strictly positive, and we shall argue that it must satisfy the functional relationship
\[ v(s, u) = v(s, t) \cdot v(t, u), \] (2.3)
implying, of course, that \( v(t, t) = 1 \) (put \( s = t = u \) and use strict positivity): If the account holder invests 1 at time \( u \), he may cash the amount on the left of (2.3) at time \( s \). If he instead cashes \( v(t, u) \) at time \( t \) and immediately reinvests this amount again, he will obtain at time \( s \) the amount on the right of (2.3). To avoid arbitrary gains, so-called arbitrage, the two strategies must give the same result.

It is easy to verify that \( v(t, u) \) satisfies (2.3) if and only if it is of the form
\[ v(t, u) = v_t^{-1} v_u \] (2.4)
for some strictly positive function \( v_t \) (allowing an abuse of notation), which can be taken to satisfy
\[ v_0 = 1. \]
Then, \( v_u \) must be the value at time 0 of a unit invested at time \( u \), and we call it the discounting function. Correspondingly, \( v_t^{-1} \) is the value at time \( t \) of a unit invested at time 0, and we call it the accumulation function.

In practical banking operations one uses
\[ v_t = e^{-\int_0^t r} \quad \text{and} \quad v_t^{-1} = e^{\int_0^t r}, \] (2.5)
where \( r_t \) is some piecewise continuous function, usually positive. (The shorthand exemplified by \( \int r = \int r \, d\tau \) will be in frequent use throughout.) Under the rule (2.5) the dynamics of accumulation and discounting are given by
\[ de^{\int_0^t r} = e^{\int_0^t r} \, r_t \, dt, \] (2.6)
\[ de^{-\int_0^t r} = -e^{-\int_0^t r} \, r_t \, dt. \] (2.7)
The relation (2.6) says that the interest earned in a small time interval is proportional to the length of the interval and to the current amount on deposit. The proportionality factor \( r_t \) is called the force of interest or the (instantaneous) interest rate at time \( t \). In integral form (2.6) and (2.7) read

\[
e^{\int_0^t r \, d\tau} = 1 + \int_0^t e^{\int_0^{t-\tau} r \, d\tau} \, d\tau ,
\]

(2.8)

\[
e^{-\int_0^t r \, d\tau} = 1 - \int_0^t e^{-\int_0^{t-\tau} r \, d\tau} \, d\tau .
\]

(2.9)

We will henceforth assume that interest is earned in accordance with (2.5) and will be working with the expressions

\[
v(t, u) = e^{-\int_u^t r \, d\tau}
\]

for the general discount factor when \( t < u \) and

\[
v(t, s) = e^{\int_t^s r \, d\tau}
\]

for the general accumulation factor when \( s < t \).

By constant interest rate \( r \) we have \( v_t = v^r \), where

\[
v = e^{-r}
\]

is the constant annual discount factor. In this case the constant annual accumulation factor is

\[
v^{-1} = e^r = 1 + i ,
\]

(2.11)

where \( i \) is the annual interest rate.

### C. Valuation of payment streams.

Suppose that the incomes/outgoes created by the payment stream \( A \) are currently deposited on/drawn from an account which bears interest at rate \( r_t \) at time \( t \). By (2.4) the value at time \( t \) of the amount \( dA_\tau \) paid in the small time interval around time \( \tau \) is \( v(t, \tau) \, dA_\tau = v_t^{-1} \, v_\tau \, dA_\tau \). Summing over all time intervals, and using (2.5), we get the value at time \( t \) of the entire payment stream,

\[
e^{\int_0^t r \, d\tau} \int_0^t e^{-\int_0^{t-\tau} r \, d\tau} \, dA_\tau = U_t - V_t ,
\]

where

\[
U_t = e^{\int_0^t r \, d\tau} \int_0^t e^{-\int_0^{t-\tau} r \, d\tau} \, dA_\tau = \int_0^t e^{\int_\tau^t r \, d\tau} \, dA_\tau
\]

(2.12)
is the accumulated value of past incomes less outgoes, and (recall the convention 
\( B = -A \))

\[
V_t = e^{\int_0^t r \, dB_\tau} \int_t^\infty e^{-\int_\tau^t r \, dB_\tau} \, dB_\tau = \int_t^\infty e^{-\int_\tau^t r \, dB_\tau} \, dB_\tau \quad (2.13)
\]

is the discounted value of future outgoes less incomes. This decomposition is 
particularly relevant for payments governed by some contract; \( U_t \) is the cash balance, that is, the amount held at the time of consideration, and \( V_t \) is the 
future liability. The difference between the two is the current value of the 
contract.

The development of the cash balance can be viewed in various ways: Application of (A.8) to (2.12), taking

\[
X_t = \exp \left( \int_0^t r_s \, ds \right) \quad \text{(continuous, with dynamics given by (2.6))}
\]

and

\[
Y_t = \int_0^t \exp \left( -\int_0^\tau r_s \, ds \right) \, dA_\tau,
\]

develops

\[
dU_t = U_t r_t \, dt + dA_t, \quad (2.14)
\]

which integrates to (note that \( U_0 = A_0 \))

\[
U_t = A_t + \int_0^t U_\tau r_\tau \, d\tau. \quad (2.15)
\]

An alternative expression,

\[
U_t = A_t + \int_0^t e^{\int_\tau^t r \, dA_\tau} \, d\tau, \quad (2.16)
\]

is derived from (2.12) upon applying the rule (A.9) of integration by parts. For instance, in the first expression in (2.12), put

\[
\int_0^t e^{-\int_\tau^t r \, dA_\tau} \, dA_\tau = A_0 + \int_0^t e^{-\int_\tau^t r \, dA_\tau} \, dA_\tau.
\]

\[
= A_0 + e^{-\int_0^t r \, dA_\tau} A_t - A_0 - \int_0^t A_\tau e^{-\int_\tau^t r \, d\tau} (-r_\tau) \, d\tau.
\]

The relations (2.14) – (2.16) show, in an easily interpretable manner, how the cash balance emerges from payments and earned interest. As a special case of (2.15) we have the trivial relationship

\[
e^{\int_0^t r \, dB_\tau} = 1 + \int_0^t e^{\int_\tau^t r \, dB_\tau} \, dB_\tau, \quad (2.17)
\]

which shows how a unit invested at time 0 accumulates with interest. Compare with (2.8).

Likewise, from (2.13) we derive

\[
dV_t = V_t r_t \, dt - dB_t, \quad (2.18)
\]
CHAPTER 2. PAYMENT STREAMS AND INTEREST

\[ V_t = B_n - B_t - \int_t^n V_{\tau} r_{\tau} d\tau, \]  
(2.19)

and

\[ V_t = B_n - B_t - \int_t^n e^{-J_{\tau} r_t} (B_n - B_{\tau}) r_{\tau} d\tau, \]  
(2.20)

the last two relationships valid for \( n = \infty \) only if \( B_\infty < \infty \). Again interpretations are easy; (2.19) and (2.20) state, in different ways, that the debt can be settled immediately at a price which is the total debt minus the present value of future interest saved by advancing the repayment.

Typically, the financial contract will lay down that incomes and outgoes be equivalent in the sense that

\[ U_n = 0 \quad \text{or} \quad V_0 = 0. \]  
(2.21)

These two relationships are equivalent and they imply that, for any \( t \),

\[ U_t = V_t. \]  
(2.22)

We anticipate here that, in the insurance context, the equivalence requirement is usually not exercised at the level of the individual policy: the very purpose of insurance is to redistribute money among the insured. Thus the principle must be applied at the level of the portfolio in some sense, which we shall discuss later. Moreover, in insurance the payments, and typically also the interest rate, are not foreseeable at the outset, so in order to establish equivalence one may have to currently adapt the payments to the development in some way or other.

D. Some standard payment functions and their values. Certain simple payment functions are so frequently used that they have been given names. An endowment of 1 at time \( n \) is defined by \( A_t = \varepsilon_n(t) \), where

\[ \varepsilon_n(t) = \begin{cases} 0, & 0 \leq t < n, \\ 1, & t \geq n. \end{cases} \]  
(2.23)

(The only payment is \( A_n - A_{n-} = 1. \) By constant interest rate \( r \) the present value at time 0 of the endowment is \( e^{-rn} \) or, recalling the notation in Chapter 1, \( v^{-n} \).

An \( n \)-year immediate annuity of 1 per year consists of a sequence of endowments of 1 at times \( t = 1, \ldots, n \), and is thus given by

\[ A_t = \sum_{j=1}^{n} \varepsilon_j(t) = [t] \wedge n. \]

By constant interest rate \( r \) its present value at time 0 is

\[ a_{nm} = \sum_{j=1}^{n} e^{-rj} = \frac{1 - e^{-rn}}{i}, \]  
(2.24)
confer (2.10) – (2.11).

An n-year annuity-due of 1 per year consists of a sequence of endowments of 1 at times \( t = 0, \ldots, n - 1 \), that is,

\[
A_t = \sum_{j=0}^{n-1} \varepsilon_j(t) = [t + 1] \wedge n.
\]

By constant interest rate its present value at time 0 is

\[
\ddot{a}_n = \sum_{j=0}^{n-1} e^{-rj} = (1 + i) \frac{1 - e^{-rn}}{i}.
\] (2.25)

An n-year continuous annuity payable at level rate 1 per year is given by

\[
A_t = t \wedge n.
\] (2.26)

For the case with constant interest rate its present value at time 0 is (recall (2.10))

\[
\ddot{a}_n = \int_0^n e^{-r\tau} d\tau = \frac{1 - e^{-rn}}{r}.
\] (2.27)

An everlasting (perpetual) annuity is called a perpetuity. Putting \( n = \infty \) in the (2.24), (2.25), and (2.27), we find the following expressions for the present values of the immediate perpetuity, the perpetuity-due, and the continuous perpetuity:

\[
\overline{a}_\infty = \frac{1}{i}, \quad \ddot{a}_\infty = \frac{1 + i}{i}, \quad \overline{\ddot{a}}_\infty = \frac{1}{r}.
\] (2.28)

An m-year deferred n-year temporary life annuity commences only after m years and is payable throughout n years thereafter. Thus it is just the difference between an \( m + n \) year annuity and an m year annuity. For the continuous version,

\[
A_t = ((t - m) \vee 0) \wedge n = (t \wedge (m + n)) - (t \wedge m).
\] (2.29)

Its present value at time 0 by constant interest is denoted \( \dot{a}_{m|n} \) and must be

\[
\dot{a}_{m|n} = \overline{a}_{m+n} - \ddot{a}_m = v^m \ddot{a}_n.
\] (2.30)

2.2 Application to loans

A. Basic features of a loan contract. Traditional loans and savings accounts in banks are among the simplest financial contracts since they are entirely deterministic. Let us consider a loan contract stipulating that at time 0, say, the bank pays to a borrower an amount \( H \), called the principal (‘first’
in Latin), and that the borrower thereafter pays back or *amortizes* the loan in accordance with a non-decreasing payment function \( \{A_t\}_{0 \leq t \leq n} \) called the amortization function. The term of the contract, \( n \), is sometimes called the duration of the loan. Without loss of generality we assume henceforth that \( H = 1 \) (the principal is proclaimed monetary unit).

The amortization function is to fulfill \( A_0 = 0 \) and \( A_n \geq 1 \). The excess of total amortizations over the principal is the total amount of interest. We denote it by \( R_n \) and have \( A_n = 1 + R_n \). General principles of book-keeping, needed e.g. for taxation purposes, prescribe that the decomposition of the amortizations into repayments and interest be extended to all \( t \in [0, n] \). Thus,

\[
A_t = F_t + R_t ,
\]

where \( F \) is a non-decreasing repayment function satisfying

\[
F_0 = 0 , \quad F_n = 1
\]

(formally a distribution function due to the convention \( H = 1 \)), and \( R \) is a non-decreasing interest payment function.

Furthermore, the contract is required to specify a nominal force of interest \( r_t , 0 \leq t \leq n \), under which the value of the amortizations should be equivalent to the value of the principal, that is,

\[
\int_0^n e^{-\int_0^\tau r} \, dA_\tau = 1 .
\]

(2.32)

There are, of course, infinitely many admissible decompositions (2.31) satisfying (2.32). A clue to constraints on \( F \) and \( R \) is offered by the relationship

\[
\int_0^n e^{-\int_0^\tau r} \, dR_\tau = \int_0^n e^{-\int_0^\tau r} (1 - F_\tau) r_\tau \, d\tau ,
\]

(2.33)

which is obtained upon inserting (2.31) into (2.32) and then using integration by parts on the term \( \int_0^n \exp(-\int_0^\tau r) \, dF_\tau = -\int_0^n \exp(-\int_0^\tau r) \, d(1 - F_\tau) \). The condition (2.33) is trivially satisfied if

\[
dR_\tau = (1 - F_\tau) r_\tau \, dt ,
\]

that is, interest is paid currently and instantaneously on the outstanding (part of the) principal, \( 1 - F_\tau \). This will be referred to as natural interest.

Under the scheme of natural interest the relation (2.31) becomes

\[
dA_t = dF_t + (1 - F_t) r_t \, dt ,
\]

(2.34)

which establishes a one-to-one correspondence between amortizations and repayments. The differential equation (2.34) is easily solved: First, integrate (2.34) over \((0, t]\) to obtain

\[
A_t = F_t + \int_0^t (1 - F_\tau) r_\tau \, d\tau ,
\]

(2.35)
which determines amortizations when repayments are given. Second, multiply (2.34) with \( \exp\left(-\int_0^t r\right) \) to obtain \( \exp\left(-\int_0^t r\right) dA_t = -d \left( \exp\left(-\int_0^t r\right) (1 - F_t) \right) \) and then integrate over \((t, n]\) to arrive at

\[
\int_t^n e^{-\int_0^t r} dA_t = 1 - F_t ,
\]

which determines (outstanding) repayments when amortizations are given. Interpretations of the relationships are obvious. For instance, since \( 1 - F_t \) is the remaining debt at time \( t \), (2.36) is the time \( t \) update of the equivalence requirement (2.32).

**B. Standard forms of loans.** We list some standard types of loans, taking now \( r \) constant. It is understood that we consider only times \( t \) in \([0, n]\).

The simplest form is the fixed loan, which is repaid in its entirety only at the term of the contract, that is, \( F_t = \varepsilon_n(t) \), the endowment defined by (2.23). The amortization function is obtained directly from (2.35): \( A_t = \varepsilon_n(t) + rt \).

A series loan has repayments of annuity form. The continuous version is given by \( F_t = t/n \), confer (2.26). The amortization plan is obtained from (2.35): \( A_t = t/n + rt(1 - t/2n) \). Thus, \( dF_t/dt = 1/n \) (fixed) and \( dR_t/dt = r(1 - t/n) \) (linearly decreasing).

An annuity loan is called so because the amortizations, which are the amounts actually paid by the borrower, are of annuity form. The continuous version is given by \( A_t = t/a_n \), confer (2.32) and (2.27). From (2.36) we easily obtain \( F_t = 1 - a_n - n/a_n \). We find \( dF_t/dt = e^{-r(n-1)/a_n} \) (exponentially increasing), and \( dR_t/dt = (1 - e^{-r(n-1)/a_n})/a_n \).

Putting \( n = \infty \), the fixed loan and the series loan both specialize to an infinite loan without repayment. Amortizations consist only of interest, which is paid indefinitely at rate \( r \).
Chapter 3

Mortality

3.1 Aggregate mortality

A. The stochastic model. Consider an aggregate of individuals, e.g. the population of a nation, the persons covered under an insurance scheme, or a certain species of animals. The individuals need not be animate beings; for instance, in engineering applications one is often interested in studying the work-life until failure of technical components or systems. Having demographic and actuarial problems in mind, we shall, however, be speaking of persons and life lengths until death.

Due to differences in inheritance and living conditions and also due to events of a more or less purely random nature, like accidents, diseases, etc., the life lengths vary among individuals. Therefore, the life length of a randomly selected new-born can suitably be envisaged as a non-negative random variable \( T \) with a cumulative distribution function

\[
F(t) = \Pr[T \leq t].
\]  

(3.1)

In survival analysis it is convenient to work with the survival function

\[
\bar{F}(t) = \Pr[T > t] = 1 - F(t).
\]  

(3.2)

Fig. 3.1 shows \( F \) and \( \bar{F} \) for the mortality law G82M used by Danish life insurers as a basis for calculating premiums for insurances on male lives. Find the median life length and some other percentiles of this life distribution by inspection of the graphs!

We assume that \( F \) is absolutely continuous and denote the density by \( f \);

\[
f(t) = \frac{d}{dt} F(t) = -\frac{d}{dt} \bar{F}(t).
\]  

(3.3)

The density of the distribution in Fig. 3.1 is depicted in Fig. 3.2. Find the mode by inspection of the graph! Can you already at this stage figure why the median and the mode of \( F \) in Fig. 3.1 appear to exceed those of the mortality law of the Danish male population?
Figure 3.1: The G82M mortality law: $F$ broken line, $\bar{F}$ whole line.

Figure 3.2: The density $f$ for the G82M mortality law.
B. The force of mortality. When dealing with non-negative random variables representing life lengths, it is convenient to work with the derivative of $-\log \overline{F}$, which is well defined for all $t$ such that $\overline{F}(t) > 0$. For small, positive $dt$ we have

$$
\mu(t) = \int dt \frac{d}{dt} \left[ -\log \overline{F}(t) \right] = \int dt \frac{f(t)}{\overline{F}(t)},
$$

which is well defined for all $t$ such that $\overline{F}(t) > 0$. For small, positive $dt$ we have

$$
\mu(t) dt = \int dt \frac{f(t) dt}{\overline{F}(t)} = \frac{\mathbb{P}[t < T \leq t + dt]}{\mathbb{P}[T > t]} = \mathbb{P}[T \leq t + dt | T > t].
$$

(In the second equality we have neglected a term $o(dt)$ such that $o(dt)/dt \to 0$ as $dt \to 0$.) Thus, for a person aged $t$, the probability of dying within $dt$ years is (approximately) proportional to the length of the time interval, $dt$. The proportionality factor $\mu(t)$ depends on the attained age, and is called the force of mortality at age $t$. It is also called the mortality intensity or hazard rate at age $t$, the latter expression stemming from reliability theory, which is concerned with the durability of technical devices.

Fig 3.3 shows the force of mortality corresponding to $\overline{F}$ in Fig. 3.1. Assess roughly the probability that a $t$ years old person will die within one year for $t = 60, 70, 80, 90$.

Integrating (3.4) from 0 to $t$ and using $\overline{F}(0) = 1$, we obtain

$$
\overline{F}(t) = e^{-\int_0^t \mu(t)}.
$$

Relation (3.4) may be cast as

$$
f(t) = \overline{F}(t) \mu(t) = e^{-\int_0^t \mu(t)} \mu(t),
$$

which says that the probability $f(t) dt$ of dying in the age interval $(t, t + dt)$ is the product of the probability $\overline{F}(t)$ of survival to $t$ and the conditional probability $\mu(t) dt$ of then dying before age $t + dt$.

The functions $F$, $\overline{F}$, $f$, and $\mu$ are equivalent representations of the mortality law; each of them corresponds one-to-one to any one of the others.

Since $\overline{F}(\infty) = 0$, we must have $\int_0^\infty \mu = \infty$. Thus, if there is a finite highest attainable age $\omega$ such that $\overline{F}(\omega) = 0$ and $\overline{F}(t) > 0$ for $t < \omega$, then $\int_0^t \mu \not\to \infty$ as $t \not\to \omega$. If, moreover, $\mu$ is non-decreasing, we must also have $\lim_{t \not\to \omega} \mu(t) = \infty$.
C. The distribution of the remaining life length. Let $T_x$ denote the remaining life length of an individual chosen at random from the $x$ years old members of the population. Then $T_x$ is distributed as $T - x$, conditional on $T > x$, and has cumulative distribution function

$$F(t|x) = P[T \leq x + t \mid T > x] = \frac{F(x + t) - F(x)}{1 - F(x)}$$

and survival function

$$\bar{F}(t|x) = P[T > x + t \mid T > x] = \frac{F(x + t)}{F(x)}, \quad (3.7)$$

which are well defined for all $x$ such that $\bar{F}(x) > 0$. The density of this conditional distribution is

$$f(t|x) = \frac{f(x + t)}{\bar{F}(x)}. \quad (3.8)$$

Denote by $\mu(t|x)$ the force of mortality of the distribution $F(t|x)$. It is obtained by inserting $f(t|x)$ from (3.8) and $\bar{F}(t|x)$ from (3.7) in the places of $f$ and $\bar{F}$ in the definition (3.4). We find

$$\mu(t \mid x) = f(x + t) / \bar{F}(x + t) = \mu(x + t). \quad (3.9)$$

Alternatively, we could insert (3.5) into (3.7) to obtain

$$\bar{F}(t|x) = e^{-\int_{x+t}^{\infty} \mu(y) \, dy} = e^{-\int_{x}^{\infty} \mu(x+\tau) \, d\tau}, \quad (3.10)$$

which by the general relation (3.5) entails (3.9). Relation (3.9) explains why the force of mortality is particularly handy; it depends only on the attained age $x + t$, whereas the conditional density in (3.8) depends in general on $x$ and $t$ in a more complex manner. Thus, the properties of all the conditional survival distributions are summarized by one simple function of the total age only.

Figs. 3.4 – 3.6 depict the functions $F(t|70)$, $\bar{F}(t|70)$, $f(t|70)$, and $\mu(t|70) = \mu(70 + t)$ derived from the life time distribution in Fig. 3.1. Observe that the first three of these functions are obtained simply by scaling up the corresponding graphs in Figs. 3.1 – 3.2 by the factor $1 / \bar{F}(70)$ over the interval $(70, \infty)$. The force of mortality remains unchanged, however.

D. Expected values in life distributions. Let $T$ be a non-negative r.v. with distribution function $F$, not necessarily absolutely continuous, and let $G : \mathbb{R}_+ \to \mathbb{R}$ be a PD and RC function such that $E[G(T)]$ exists and is finite. Integrating by parts, we find

$$E[G(T)] = G(0) + \int_0^\infty \bar{F}(\tau) \, dG(\tau). \quad (3.11)$$
CHAPTER 3. MORTALITY

Figure 3.4: Conditional distribution of remaining life length for the G82M mortality law: $F(t|70)$ broken line, $\hat{F}(t|70)$ whole line.

Figure 3.5: Conditional density of remaining life length $f(t|70)$ for the G82M mortality law.

Figure 3.6: The force of mortality $\mu(t|70) = \mu(70 + t)$, $t > 0$, for the G82M mortality law.
CHAPTER 3. MORTALITY

Taking \( G(t) = t^k \) we get

\[
\mathbb{E}[T^k] = k \int_0^\infty t^{k-1} \bar{F}(t) \, dt, \quad (3.12)
\]

and, in particular,

\[
\mathbb{E}[T] = \int_0^\infty \bar{F}(t) \, dt, \quad (3.13)
\]

The expected remaining life time for an \( x \) years old person is

\[
\bar{e}_x = \int_0^\infty \bar{F}(t \mid x) \, dt. \quad (3.14)
\]

From (3.10) it is seen that \( \bar{F}(t \mid x) \) is a decreasing function of \( t \) for fixed \( t \) if \( \mu \) is an increasing function. Then \( \bar{e}_x \) is a decreasing function of \( x \). One can easily construct mortality laws for which \( \bar{F}(t \mid x) \) and \( \bar{e}_x \) are not decreasing functions of \( x \).

Consider the more general function

\[
G(t) = ( (t \wedge b) - (t \wedge a) )^k = \begin{cases} 
0, & 0 \leq t < a, \\
(t - a)^k, & a \leq t < b, \\
(b - a)^k, & b \leq t,
\end{cases} \quad (3.15)
\]

that is, \( dG(t) = k(t - a)^{k-1} \, dt \) for \( a < t < b \) and 0 elsewhere. It is realized that \( G(T) \) is the \( k \)th power of the number of years lived between age \( a \) and age \( b \). From (3.11) we obtain

\[
\mathbb{E}[G(T)] = k \int_a^b (t - a)^{k-1} \bar{F}(t) \, dt, \quad (3.16)
\]

In particular, the expected number of years lived between the ages of \( a \) and \( b \) is \( \int_a^b \bar{F}(t) \, dt \), which is the area between the \( t \)-axis and the survival function in the interval from \( a \) to \( b \). The formula can be motivated directly by noting that \( \bar{F}(t) \, dt \) is the expected number of years survived in the small time interval \( (t, t + dt) \) and using that the “expected value of the sum is the sum of the expected values”.

3.2 Some standard mortality laws

A. The exponential distribution. Suppose the force of mortality is \( \mu(t) = \lambda \), independent of the age. This means there are no wear-out effects; each morning when you wake up (if you wake up) life starts anew with the same prospects of longevity as for a new-born. Then the survival function (3.5) becomes

\[
\bar{F}(t) = e^{-\lambda t}, \quad (3.17)
\]
and the density (3.6) becomes

$$f(t) = \lambda e^{-\lambda t}.$$  \hspace{1cm} (3.18)

Thus, $T$ is exponentially distributed with parameter $\lambda$. The conditional survival function (3.10) becomes $\bar{F}(t \mid x) = e^{-\lambda t}$, hence

$$\bar{F}(t \mid x) = \bar{F}(t),$$  \hspace{1cm} (3.19)

the same as (3.17). The exponential distribution is a suitable model for certain technical devices like bulbs and electronic components. Unfortunately, it is not so apt for description of human lives.

One could arrive at the exponential distribution by specifying that (3.19) be valid for all $x$ and $t$, that is, the probability of surviving another $t$ years is independent of the age $x$. Then, from the general relation (3.7) we get

$$\bar{F}(x + t) = \bar{F}(x)\bar{F}(t)$$  \hspace{1cm} (3.20)

for all non-negative $x$ and $t$. It follows by induction that for each pair of positive integers $m$ and $n$, $\bar{F}(\frac{m}{n}) = \bar{F}(\frac{1}{n})^m = \bar{F}(1)^\frac{m}{n}$, hence

$$F(t) = \bar{F}(1)^t$$  \hspace{1cm} (3.21)

for all positive rational $t$. Since $\bar{F}$ is right-continuous, (3.21) must hold true for all $t > 0$. Putting $\bar{F}(1) = e^{-\lambda}$, we arrive at (3.17).

Fig. 3.7 shows the survival function and the density for two different values of $\lambda$. 

Figure 3.7: Two exponential laws with intensities $\lambda_1$ and $\lambda_2$ such that $\lambda_1 < \lambda_2$; $\bar{F}_1$ and $\bar{F}_2$ whole line, $f_1$ and $f_2$ broken line.
B. The Weibull distribution. The intensity of this distribution is of the form

\[ \mu(t) = \beta \alpha^{-\beta} t^{\beta - 1}, \tag{3.22} \]

\( \alpha, \beta > 0 \). The corresponding survival function is \( \bar{F}(t) = \exp \left( - \left( \frac{t}{\alpha} \right)^\beta \right) \).

If \( \beta > 1 \), then \( \mu(t) \) is increasing, and if \( \beta < 1 \), then \( \mu(t) \) is decreasing. If \( \beta = 1 \), the Weibull law reduces to the exponential law. Draw the graphs of \( \bar{F} \) and \( f \) for some different choices of \( \alpha \) and \( \beta \! \)!

We have \( \mu(x + t) = \beta \alpha^{-\beta} (x + t)^{\beta - 1} \), and, by virtue of (3.22), \( \bar{F}(t \mid x) \) is not a Weibull law.

C. The Gompertz-Makeham distribution. This distribution is widely used as a model for survivorship of human lives, especially in the context of life insurance. Thus, as it will be frequently referred to, we shall use the acronym G-M for this law. Its mortality intensity is of the form

\[ \mu(t) = \alpha + \beta c t, \tag{3.23} \]

\( \alpha, \beta, c > 0 \). The corresponding survival function is

\[ \bar{F}(t) = \exp \left( - \int_0^t (\alpha + \beta c s) \, ds \right) = \exp \left( - \alpha t - \beta (c^t - 1) / \ln c \right). \tag{3.24} \]

If \( \beta > 0 \) and \( c > 1 \), then \( \mu(t) \) is an increasing function of \( t \). The constant term \( \alpha \) accounts for age-independent causes of death like certain accidents and epidemic diseases, and the term \( \beta c^t \) accounts for all kinds of wear-out effects due to aging.

We have \( \mu(x + t) = \alpha + \beta c^x c^t \), and so (3.23) shows that \( \bar{F}(t \mid x) \) is also a G-M survival function with parameters \( \alpha, \beta c^x, c \). The special case \( \alpha = 0 \) is referred to as the (pure) Gompertz law.

The G82M mortality law depicted in Fig. 3.1 is the G-M law with

\[ \alpha = 5 \cdot 10^{-4}, \quad \beta = 7.5858 \cdot 10^{-5}, \quad c = 1.09144. \tag{3.25} \]

Table E.1 in Appendix E shows \( \mu(t), \bar{F}(t) \) and \( f(t) \) for integer \( t \).

3.3 Actuarial notation

A. Actuaries in all countries – unite! The International Association of Actuaries (IAA) has laid down a notational standard, which is generally accepted among actuaries all over the world. Familiarity with this notation is a must for anyone who wants to communicate in writing or reading with actuaries, and we shall henceforth adopt it in those simple situations where it is applicable.
CHAPTER 3. MORTALITY

B. A list of some standard symbols. According to the IAA standard, the quantities introduced so far are denoted as follows:

\[ t_q x = F(t | x) , \]  
\[ t_p x = \bar{F}(t | x) , \]  
\[ \mu_{x+t} = \mu(x + t) . \]  

In particular, \( t_0 = F(t) \) and \( t_0 = \bar{F}(t) \). One-year death and survival probabilities are abbreviated as

\[ q_x = 1 q x , \quad p_x = 1 p x . \]  

Frequently used is also the “\( n \)-year deferred probability of death within \( m \) years”,

\[ n|m q x = m + n q x - n q x = n p x - m + n p x = n p x m q x + n . \]  

The formulas in Section 3.1 are easily translated, e.g.

\[ t_p x = \exp(- \int_0^t \mu_{x+\tau} d\tau) , \]  
\[ f(t | x) = t_p x \mu_{x+t} , \]  
\[ \bar{e}_x = \int_0^\infty t p_x dt . \]  

Frequently actuaries work with expected numbers of survivors instead of probabilities. Consider a population of \( l_0 \) new-born who are subject to the same law of mortality given by (3.28). The expected number of survivors at age \( x \) is

\[ l_x = l_0 x p_0 . \]  

The function \( \{ l_x ; x > 0 \} \) is called the decrement function or, when considered only at integer values of \( x \), the decrement series. Expressed in terms of the decrement function we find e.g.

\[ t_p x = l_{x+t} / l_x , \]  
\[ \mu_{x+t} = - l'_{x+t} / l_x , \]  
\[ f(t | x) = - l'_{x+t} / l_x , \]  
\[ \bar{e}_x = \int_0^\infty l_{x+t} dt / l_x . \]  

The pieces of IAA notation we have shown here are quite pleasing to the eye and also space-saving; for instance, the symbol on the left of (3.27) involves three typographical entities, whereas the one on the right involves six.
3.4 Select mortality

A. The insurance portfolio consists of selected lives. Consider an individual who purchases a life insurance at age $x$. In short, he will be referred to as $(x)$ in what follows.

It is quite common in actuarial practice to assume that the force of mortality of $(x)$ depends on $x$ and $t$ in a more complex manner than the simple relationship (3.9), which rested on the assumption that $(x)$ is chosen at random from the $x$ years old individuals in the population. The fact that $(x)$ purchases insurance adds information to the mere fact that he has attained age $x$; he does not represent a purely random draw from the population, but is rather selected by some mechanisms. It is easy to think of examples of such mechanisms. For instance that poor people can not afford to buy insurance and, to the extent that longevity depends on economic situation, the mortality experience for insured people would reflect that they are wealthy enough to buy insurance (‘survival of the fattest’). Judging from textbooks on life insurance, e.g. [4] and [15] and many others, it seems that the underwriting standards of the insurer are generally held to be the predominant selective mechanism; before an insurance policy is issued, the insurer must be satisfied that the applicant meets certain requirements with regard to health, occupation, and other factors that are assumed to determine the prospects of longevity. Only first class lives are accepted as insurable at ordinary rates.

Thus there is every reason to account for selection effects by letting the force of mortality be some more general function $\mu_x(t)$ or, in other words, specify that $T_x$ follows a survival function $F_x(t)$ that is not necessarily of the form (3.7). One then speaks of select mortality.

B. More of actuarial notation. The standard actuarial notation for select mortality is

$$
\tau q\{x\}+t = P[T_x \leq t + \tau \mid T_x > t],
$$

$$
\tau p\{x\}+t = P[T_x > t + \tau \mid T_x > t],
$$

$$
\mu\{x\}+t = \lim_{\tau \downarrow 0} \frac{h q\{x\}+t}{h}.
$$

The idea is that the both the current age, $x + t$, and the age at entry, $x$, are directly visible in $[x] + t$.

From a technical point of view select mortality is just as easy as aggregate mortality; we work with the distribution function $t q\{x\}$ instead of $q_x$, and are interested in it as a function of $t$. For instance,

$$
\tau p\{x\}+t = \frac{t+\tau p\{x\}}{t p\{x\}} = \exp \left( - \int_t^{t+\tau} \mu\{x\}+s \, ds \right).
$$

C. Features of select mortality. There is ample empirical evidence to support the following facts about select mortality in life insurance populations:
For insured lives of a given age the rate of mortality usually increases with increasing duration.

- The effect of selection tends to decrease with increasing duration and becomes negligible for practical purposes when the duration exceeds a certain select period.
- The mortality among insured lives is generally lower than the mortality in the population.

There are many possible ways of building such features into the model. For instance, one could modify the aggregate G-M intensity as

$$\mu_{x+t} = \alpha(t) + \beta(t) e^{ct},$$

where $\alpha$ and $\beta$ are non-negative and non-decreasing functions bounded from above. In Section 7.6 we shall show how the selection mechanism can be explained in models that describe more aspects of the individual life histories than just survival and death.
Chapter 4

Insurance of a single life

4.1 Some standard forms of insurance

A. The single-life status. Consider a person aged $x$ with remaining life length $T_x$ as described in the previous section. In actuarial parlance this life is called the single-life status $(x)$. Referring to Appendix B, we introduce the indicator of the event of survival in $t$ years, $I_t = 1_{\{T_x > t\}}$. It is a binomial random variable with success probability $t \cdot p_x$. The indicator of the event of death within $t$ years is $1 - I_t = 1_{\{T_x \leq t\}}$, which is a binomial variable with success probability $t \cdot q_x = 1 - t \cdot p_x$. (We apologize for sometimes using technical terms where they may sound cynical.) Note that, being 0 or 1, any indicator $1_A$ satisfies $(1_A)^q = 1_A$ for $q > 0$.

The present section lists some standard forms of insurance that $(x)$ can purchase, investigates some of their properties, and presents some standard actuarial methods and formulas.

We assume that the investments of the insurance company yield interest at a fixed rate $r$, hence discounting is at annual rate $v = e^{-r}$. Standard actuarial notation pertaining to this case is employed throughout.

B. The pure endowment insurance. An $n$-year pure (life) endowment of 1 is a unit that is paid to $(x)$ at the end of $n$ years if he is then still alive. Recalling (2.23), the associated payment function is an endowment of $I_n$ at time $n$. Its present value at time 0 is

$$V^{e:n} = e^{-rn} I_n.$$  \hfill (4.1)

The expected value of $V^{e:n}$ is

$$n \cdot E_x = e^{-rn} n \cdot p_x.$$  \hfill (4.2)

For any $q > 0$ we have $(V^{e:n})^q = e^{-qrn} I_n$ (recall that $I^n_q = I_n$), and so the $q$-th non-central moment of $V^{e:n}$ may be expressed as

$$E[(V^{e:n})^q] = n \cdot E_x^{(qr)}.$$  \hfill (4.3)
where the topscript \((qr)\) signifies that discounting is made under a force of interest that is \(q\) times the standard \(r\).

In particular, the variance of \(V^{n}e\) is

\[
V[V^{n}e] = nE_x^{(2r)} - nE_x^2.
\]  

\(4.4\)

C. The life assurance. A life assurance contract specifies that a certain amount, called the sum insured, is to be paid upon the death of the insured, possibly limited to a specified period. We shall here consider only insurances payable immediately upon death, and take the sum to be 1 (just a matter of notation).

First, an \(n\)-year term insurance is payable upon death within \(n\) years. The payment function is a lump sum of \(1 - I_n\) at time \(T_x\). Its present value at time 0 is

\[
V^{ti,n} = e^{-rT_x} (1 - I_n).
\]  

\(4.5\)

The expected value of \(V^{ti,n}\) is

\[
\bar{A}_x = \int_0^n e^{-rt} t p_x \mu_{x+t} dt,
\]  

\(4.6\)

and, similar to (4.3),

\[
E[(V^{ti,n}q)'] = \bar{A}_x^{(qr)}\frac{1}{\Omega}.
\]  

\(4.7\)

In particular,

\[
V[V^{ti,n}] = \bar{A}_x^{(2r)} - \bar{A}_x^2.
\]  

\(4.8\)

An \(n\)-year endowment insurance is payable upon death if it occurs within \(n\) and otherwise at time \(n\). The payment function is a lump sum of 1 at time \(T_x \wedge n\). Its present value at time 0 is

\[
V^{ei,n} = e^{-r(T_x \wedge n)}.
\]  

\(4.9\)

The expected value of \(V^{ei,n}\) is

\[
\bar{A}_x = \int_0^n e^{-rt} t p_x \mu_{x+t} dt + e^{-rn} n p_x = \bar{A}_x + nE_x,
\]  

\(4.10\)

and

\[
E(V^{ei,n}q) = \bar{A}_x^{(qr)}\frac{1}{\Omega}.
\]  

\(4.11\)

It follows that

\[
V[V^{ei,n}] = \bar{A}_x^{(2r)} - \bar{A}_x^2.
\]  

\(4.12\)
D. The life annuity. An $n$-year temporary life annuity of 1 per year is payable as long as $(x)$ survives but limited to $n$ years. We consider here only the continuous version. Recalling (2.26), the associated payment function is an annuity of 1 in $T_x \wedge n$. Its present value at time 0 is

$$V^{a,n} = \bar{a}_{x \wedge n} = \frac{1 - e^{-r(T_x \wedge n)}}{r}.$$  \hfill (4.13)

The expected value of $V^{a,n}$ is

$$\bar{a}_x \bar{m} = \int_0^n \bar{a}_t p_x \mu_{x+t} dt + \bar{a}_n p_x = \int_0^n e^{-rt} t p_x dt = \int_0^n t E_x dt.$$  \hfill (4.14)

The last expression, which follows upon integrating by parts, displays that the annuity is a “sum of pure endowments of $dt$ in each small interval $[t, t+dt]$” up to time $n$, confer (4.2). We shall demonstrate below that

$$E[(V^{a,n})^q] = \frac{q}{r^{q-1}} \sum_{p=1}^q (-1)^{p-1} \binom{q-1}{p-1} \bar{a}_{x \wedge n},$$  \hfill (4.15)

from which we derive

$$V[V^{a,n}] = \frac{2}{r} \left( \bar{a}_x \bar{m} - \bar{a}_{x \wedge n} \right) - \bar{a}_x^2.$$  \hfill (4.16)

The endowment insurance is a combined benefit consisting of an $n$-year term insurance and an $n$-year pure endowment. By (4.9) and (4.13) it is related to the life annuity by

$$V^{a,n} = \frac{1 - V^{e,i,n}}{r} \quad \text{or} \quad V^{e,i,n} = 1 - r V^{a,n},$$  \hfill (4.17)

which just reflects the more general relationship (2.27). Taking expectation in (4.17), we get

$$\bar{A}_x \bar{m} = 1 - r \bar{a}_x \bar{m}. $$  \hfill (4.18)

Also, since $V^{t;i,n} = V^{e,i,n} - V^{e;n} = 1 - r V^{a,n} - V^{e;n}$, we have

$$\bar{A}_x \bar{m} = 1 - r \bar{a}_x \bar{m} - n E_x.$$  \hfill (4.19)

The formerly announced result (4.15) follows by operating with the $q$-th moment on the first relationship in (4.17), and then using (4.12) and (4.18) and rearranging a bit. One needs the binomial formula

$$(x+y)^q = \sum_{p=0}^q \binom{q}{p} x^{q-p} y^p$$

and the special case $\sum_{p=0}^q \binom{q}{p} (-1)^{q-p} = 0$ (for $x = -1$ and $y = 1$).

A whole-life annuity is obtained by putting $n = \infty$. Its expected present value is denoted simply by $\bar{a}_x$ and is obtained by putting $n = \infty$ in (4.14), that is

$$\bar{a}_x = \int_0^\infty e^{-rt} t p_x dt,$$  \hfill (4.20)

and the same goes for the variance in (4.16) (justify the limit operations).
E. Deferred benefits. An $m$-year deferred $n$-year temporary life annuity commences only after $m$ years, provided that $(x)$ is then still alive, and is payable throughout $n$ years thereafter as long as $(x)$ survives. The present value of the benefits is

$$V^{a;m|n} = \frac{e^{-r(T_x \wedge m) - e^{-r(T_x \wedge (m+n))}}}{r}$$

The expected present value is

$$E \left[ V^{a;m|n} \right] = E \left[ V^{a;m|n} \mid I_m \right] = m p_x E_x \bar{a}_{x+m} m$$

The last expression can be obtained also by the rule of iterated expectation, and we carry through this small exercise just to illustrate the technique:

$$V^{a;m|n} = \int_m^{m+n} e^{-rt} t p_x dt = m E_x \bar{a}_{x+m} m.$$  \hspace{1cm} (4.22)

An $m$-year deferred whole life annuity is obtained by putting $n = \infty$. The expected value is denoted by $m|\bar{a}_x$. Deferred life assurances, although less common in practice, are defined likewise. For instance, an $m$-year deferred $n$-year term assurance of 1 is payable upon death in the time interval $(m, m+n]$. Its present value at time 0 is

$$V^{t_i;m|n} = V^{t_i;m+n} - V^{t_i;m},$$

and its expected present value is

$$m|\bar{A}_x = \frac{\bar{A}_x}{m+m} - \frac{\bar{A}_x}{m} = m E_x \bar{A}_{x+m} m,$$

F. Computational aspects. Distribution functions of present values and many other functions of interest can be calculated easily; after all there is only one random variable in play, and finding expected values amounts just to forming integrals in one dimension. We shall, however, not pursue this approach because it will turn out that a different point of view is needed in more complex situations to be studied in the sequel.

Anyway, by methods to be developed later, we easily compute the three first moments of the present values considered above, and find the expected values, coefficients of variation, and skewnesses shown in Table 4.1. The reader should contemplate the results, keeping in mind that the coefficient of variation may be taken as a simple measure of “riskiness”.

We interpose that numerical techniques will be dominant in our context. Explicit formulas cannot be obtained even for trivial quantities like $\bar{a}_x m$ under the Gompertz-Makeham law (3.23); age dependence and other forms of inhomogeneity of basic entities leave little room for aesthetics in actuarial science. Also relationships like (4.18) are of limited interest; they are certainly not needed for computational purposes, but may provide some general insight.
Table 4.1: Expected value (E), coefficient of variation (CV), and skewness (SK) of the present value at time 0 of a pure endowment (PE) with sum 1, a term insurance (TI) with sum 1, an endowment insurance (EI) with sum 1, and a life annuity (LA) with level intensity 1 per year, when \( x = 30, n = 30, \mu \) is given by (3.25), and \( r = \ln(1.045) \).

<table>
<thead>
<tr>
<th></th>
<th>PE</th>
<th>TI</th>
<th>EI</th>
<th>LA</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>0.2257</td>
<td>0.06834</td>
<td>0.2940</td>
<td>16.04</td>
</tr>
<tr>
<td>CV</td>
<td>0.4280</td>
<td>2.536</td>
<td>0.3140</td>
<td>0.1308</td>
</tr>
<tr>
<td>SK</td>
<td>-1.908</td>
<td>2.664</td>
<td>4.451</td>
<td>-4.451</td>
</tr>
</tbody>
</table>

4.2 The principle of equivalence

A. A note on terminology. Like any other good or service, insurance coverages are bought at some price. And, like any other business, an insurance company must fix prices that are sufficient to defray the costs. In one respect, however, insurance is different: for obvious reasons the customer is to pay in advance. This circumstance is reflected by the insurance terminology, according to which payments made by the insured are called *premiums*. This word has the positive connotation “prize” (reward), rather antonymous to “price” (sacrifice, due), but the etymological background is, of course, that premium means “first” (French: prime).

B. The equivalence principle. The equivalence principle of insurance states that the expected present values of premiums and benefits should be equal. Then, roughly speaking, premiums and benefits will balance on the average. This idea will be made precise later. For the time being all calculations are made on an individual net basis, that is, the equivalence principle is applied to each individual policy, and without regard to expenses incurring in addition to the benefits specified by the insurance treaties. The resulting premiums are called (individual) *net premiums*.

The premium rate depends on the premium payment scheme. In the simplest case, the full premium is paid as a single amount immediately upon the inception of the policy. The resulting *net single premium* is just the expected present value of the benefits, which for standard forms of insurance is given in Section 4.1.

The net single premium may be a considerable amount and may easily exceed the liquid assets of the insured. Therefore, premiums are usually paid by a series of installments extending over some period of time. The most common solution is to let a fixed level amount fall due periodically, e.g. annually or monthly, from the inception of the agreement until a specified time \( m \) and contingent on the survival of the insured. Assume for the present that the premiums are paid continuously at a fixed level rate \( \pi \). (This is admittedly an artificial assumption, but it can serve well as an approximative description of periodical payments, which will be treated later.) Then the premiums form an \( m \)-year temporary life
annuity, payable by the insured to the insurer. Its present value is $\pi V^{a,m}$, with expected value $\pi \bar{a}_x \bar{v}$ given by (4.14). We list formulas for the net level premium rate for a selection of standard forms of insurance: For the pure endowment in Paragraph 4.1.B,

$$\pi_{n/m}^e = \frac{n E_x}{\bar{a}_x \bar{v}}.$$  \hspace{1cm} (4.25)

For the $k$-year deferred $n$-year temporary annuity in Paragraph 4.1.E,

$$\pi_{k|n/m}^a = \frac{k \bar{a}_x}{\bar{v}}.$$  \hspace{1cm} (4.26)

For the term insurance in Paragraph 4.1.C,

$$\pi_{n/m}^t = \frac{\bar{A}_1}{\bar{a}_x \bar{v}}.$$  \hspace{1cm} (4.27)

For the endowment insurance in Paragraph 4.1.C,

$$\pi_{n/m}^{e|t} = \frac{\bar{A}_1}{\bar{a}_x \bar{v}}.$$  \hspace{1cm} (4.28)

In formula (4.26) the far most common solution in practice is to let $k = m$, that is, the premium payment period coincides with the deferment period. In the other formulas $m = n$ corresponds to common practice. The case $m > n$ is of no practical interest since, basically, premiums are paid in advance. The following formulas are immediate consequences of (4.22) and (4.18):

$$\pi_{m|n/m}^a = \frac{\bar{a}_x m + n}{\bar{a}_x \bar{v}} - 1,$$  \hspace{1cm} (4.29)

$$\pi_{n/n}^{e|t} = 1 - \frac{1}{\bar{a}_x \bar{v}}.$$  \hspace{1cm} (4.30)

Usually, when working in a context with fixed contract terms, all top- and subscripts and the like will be dropped from the symbol $\pi$.

C. The net economic result for a policy. The random variables studied in Section 4.1 represent the uncertain future liabilities of the insurer. Now, unless single premiums are used, also the premium incomes are dependent on the insured’s life length and become a part of the insurer’s uncertainty. Therefore, the relevant random variable associated with an insurance policy is the present value of benefits less premiums,

$$V = V^b - \pi V^{a,m},$$  \hspace{1cm} (4.31)

where $V^b$ is the present value of the benefits, e.g. $V^{e|t,n}$ in the case of an $n$-year endowment insurance.

Stated precisely, the equivalence principle lays down that

$$\mathbb{E}[V] = 0.$$  \hspace{1cm} (4.32)
CHAPTER 4. INSURANCE OF A SINGLE LIFE

For example, with \( V^b = V^{ei,n} \) (4.32) becomes 0 = \( \bar{A}_x - \pi \bar{a}_x \), which yields (4.30) when \( m = n \).

A measure of the uncertainty associated with the economic result of the policy is the variance \( \mathbb{V}[V] \). For example, with \( V^b = V^{ei,n} \) and \( m = n \),

\[
\mathbb{V}[V] = \mathbb{V} \left[ \mu_{x+n} - \frac{1 - \mu_x^{n}}{r} \right] - (1 + \pi/r)^2 \mathbb{V}[\mu_x^{n}]
\]

\[
= \frac{2 (\bar{a}_x - \bar{a}_x^{2r})}{r \bar{a}_x^{2r}} - 1.
\]

4.3 Prospective reserves

A. The case. We shall discuss the notion of reserve in the framework of a combined insurance which comprises all standard forms of contingent payments that have been studied so far and, therefore, easily specializes to each of those. The insured is \( x \) years old upon issue of the contract, which is for a term of \( n \) years. The benefits consist of a term insurance with sum insured \( b_t \) payable upon death at time \( t \in (0, n) \) and a pure endowment with sum \( b_n \) payable upon survival at time \( n \). A lump sum premium of \( \pi_0 \) is due immediately upon issue of the policy at time 0, and thereafter premiums are payable at rate \( \pi_t \) per time unit contingent on survival at time \( t \in (0, n) \).

The expected present value at time 0 of total benefits less premiums under the contract is

\[
- \pi_0 + \int_0^n v^{n-t} \mu_{x+t} b_t - \pi_t \} d\tau + b_n v^{n-t} p_x.
\]

Under the equivalence principle this is set equal to 0, a constraint on the premium function \( \pi \).

B. Definition of the reserve. The expected value (4.34) represents, in an average sense, an assessment of the economic prospects of the policy at the outset. At any time \( t > 0 \) in the subsequent development of the policy the assessment should be updated with regard to the information currently available. If the policy has expired by death before time \( t \), there is nothing more to be done. If the policy is still in force, a renewed assessment must be based on the conditional distribution of the remaining life length. Insurance legislation lays down that at any time the insurance company must provide a reserve to meet future net liabilities on the contract, and this reserve should be precisely the expected present value at time \( t \) of total benefits less premiums in the future. Thus, if the policy is still in force at time \( t \), the reserve is

\[
V_t = \int_t^n v^{n-t} \mu_{x+t} b_t - \pi_t \} d\tau + b_n v^{n-t} p_x.
\]
CHAPTER 4. INSURANCE OF A SINGLE LIFE

Figure 4.1: The net reserve for an $n$-year pure endowment of 1 against single net premium.

More precisely, this quantity is called the prospective reserve at time $t$ since it “looks ahead”. Under the principle of equivalence it is usually called the net premium reserve. We shall here take the liberty to just speak of the reserve.

There is a retrospective formula for the net premium reserve, which is obtained upon setting the expression in (4.34) equal to 0, then splitting the integral $\int_0^n$ into $\int_0^t + \int_t^n$, and observing that the latter integral plus the last term in (4.34) is $v^t q_x V_t$. Then, solving with respect to $V_t$, we obtain

$$V_t = \frac{1}{t_p x} \left( (1 + i)^t \pi_0 + \int_0^t (1 + i)^{t-\tau} q_x \{ \pi_0 - \mu_x \} d\tau \right).$$

(4.36)

This formula expresses $V_t$ as the surplus of transactions in the past, accumulated at time $t$ with the “benefit of interest and survivorship”.

C. Some special cases. The net reserve is easily put up for the standard forms of insurance treated in Sections 4.1 and 4.2. It is assumed that premiums are based on the equivalence principle, which can be expressed as

$$- \pi_0 + V_0 = 0.$$  

(4.37)

Consider first the pure endowment introduced in Paragraph 4.1.B. If the single net premium $nE_x$ is collected at time 0, then

$$V_t = n - t E_{x+t}, \quad 0 < t \leq n.$$  

(4.38)

The graph of $V_t$ will typically look as in Fig. 4.1. At points of discontinuity a dot marks the value of the function.
Figure 4.2: The net reserve for an $n$-year pure endowment of 1 against level premium in the contract period.

Figure 4.3: The net reserve for an $m$-year deferred whole life annuity against level premium in the deferment period.
CHAPTER 4. INSURANCE OF A SINGLE LIFE

Figure 4.4: The net reserve for an $n$-year term insurance against level premium in the insurance period

Figure 4.5: The net reserve for an $n$-year endowment insurance with level premium payable in the contract period.
If premiums are payable continuously at level rate $\pi = \pi_n^c$ given by (4.25) throughout the contract period, then

$$
V_t = n-tE_x+t - \pi \bar{a}_{x+t \mid n-t} = \frac{nE_x}{\bar{a}_{x \mid n}} \bar{a}_{x+t \mid n-t}.
$$

(4.39)

A typical graph of this function is shown in Fig. 4.2.

Next, consider an $m$-year deferred whole life annuity against the level net premium $\pi = \pi_m^c \mid \infty/m$ given by (4.26) with $k = m$. The net reserve is

$$
V_t = \begin{cases} 
m-t \bar{a}_{x+t} - \pi \bar{a}_{x+t \mid m-t} & 0 < t < m, \\
\bar{a}_{x+t} - \bar{a}_{x+t \mid m-t} & t \geq m,
\end{cases}
$$

(4.40)

(with the understanding that $\bar{a}_{x+t \mid m-t} = 0$ if $t > m$). A typical graph of this function is shown in Fig. 4.3.

For the $n$-year term insurance, considered in Paragraph 4.1.C, with level net premium $\pi$ given by (4.27) and $m = n$,

$$
V_t = \bar{A}_{x+t \mid n-t} - \pi \bar{a}_{x+t \mid n-t} = 1 - r \bar{a}_{x+t \mid n-t} - \frac{nE_x}{\bar{a}_{x \mid n}} \bar{a}_{x+t \mid n-t}.
$$

(4.41)

A typical graph of this function is shown in Fig. 4.4.

Finally, for the $n$-year endowment insurance, with level net premium $\pi$ given by (4.28) payable in the insurance period,

$$
V_t = \bar{A}_{x+t \mid n-t} - \pi \bar{a}_{x+t \mid n-t} = 1 - r \bar{a}_{x+t \mid n-t} - \frac{nE_x}{\bar{a}_{x \mid n}} \bar{a}_{x+t \mid n-t}.
$$

(4.42)

A typical graph of this function is shown in Fig. 4.5.

The reserve in (4.42) is, of course, the sum of the reserves in (4.40) and (4.41). Note that the pure term insurance requires a much smaller reserve than the other insurance forms, with elements of savings in them. However, at old ages $x$ (where people typically are not covered against the risk of death since death will incur soon with certainty) also the term insurance may have a $V_t$ close to 1 in the middle of the insurance period.
D. Non-negativity of the reserve. In all the examples given here the net reserve is sketched as a non-negative function. Non-negativity of \( V_t \) is not a consequence of the definition. One may easily construct premium payment schemes that lead to negative values of \( V_t \) (just let the premiums fall due after the payment of the benefits), but such payment schemes are not used in practice. The reason is that the holder of a policy with \( V_t < 0 \) is in expected debt to the insurer and would thus have an incentive to cancel the policy and thereby get rid of the debt. (The agreement obliges the policyholder only to pay the premiums, and the contract can be terminated at any time the policyholder wishes.) Therefore, it is in practice required that

\[
V_t \geq 0, \quad t \geq 0. \tag{4.43}
\]

E. The reserve considered as a function of time. We will now take a closer look at the prospective reserve as a function of time, bearing in mind that it should be non-negative. The building blocks are expected present values \( n-tE_{x+t} \), \( \bar{a}_{x+t} \), \( \bar{A}_{x+t} \) and \( \bar{A}_{x+1} \) appearing in the formulas in Section 4.3.

First, \( n-tE_{x+t} = e^{-\int_t^n (r+\mu+s) \, ds} \) is seen to be an increasing function of \( t \) no matter what are the interest rate and the mortality rate. The derivative is

\[
\frac{d}{dt} n-tE_{x+t} = n-tE_{x+t} (r + \mu). 
\]

We interpose here that nothing is changed if \( r \) depends on time. The expressions above show that, for this pure survival benefit, \( r \) and \( \mu \) play identical parts in the expected present value. Thus, mortality bequest acts as an increase of the interest rate.

Next consider \( \bar{a}_{x+t} = \int_t^n e^{-\int_t^\tau (r+\mu+s) \, ds} \, d\tau \).

The following inequalities are obvious:

\[
\bar{a}_{x+t} \leq \frac{1}{r + \inf_{s \geq t} \mu_{x+s}} \leq \frac{1}{r}. 
\]

The last expression is just the present value of a perpetuity, (2.28). If \( \mu \) is an increasing function, then

\[
\bar{a}_{x+t} \leq \frac{1}{r + \mu_{x+t}}.
\]

We find the derivative

\[
\frac{d}{dt} \bar{a}_{x+t} = (r + \mu)\bar{a}_{x+t} - 1.
\]
It follows that $\overline{a}_{x+t,\overline{n-t}}$ is a decreasing function of $t$ if $\mu$ is increasing, which is quite natural. You can easily invent an example where $\overline{a}_{x+t,\overline{n-t}}$ is not decreasing.

From the identity

$$\overline{A}_{x+t,\overline{n-t}} = 1 - r\overline{a}_{x+t,\overline{n-t}}$$

we conclude that $\overline{A}_{x+t,\overline{n-t}}$ is an increasing function of $t$ if $\mu$ is increasing.

For

$$\overline{A}_{x+t,\overline{n-t}} = 1 - r\overline{a}_{x+t,\overline{n-t}} - n-tE_{x+t}$$

no general statement can be made as to whether it is decreasing or increasing.

Looking back at the formulas derived in Paragraph C above, we can conclude that the reserve for the pure life endowment against single premium, (4.38), is always increasing. Assume henceforth that $\mu$ is increasing, as is usually the case at ages when people are insured and certainly holds for the Gompertz-Makeham law. Then also the reserve (4.39) for the pure life endowment against level premium during the term of the contract is increasing, and the same is the case for the reserve (4.42) of the endowment insurance. It is left to the industrious reader to show that (4.40) is increasing throughout the deferment period and thereafter turns decreasing. Also (4.41) is first increasing and thereafter decreasing, as can most easily be inferred by examining the formulas (4.35) and (4.36).

### 4.4 Thiele’s differential equation

**A. The differential equation.** We turn back to the general case with the reserve given by (4.35). Suppose the policy is in force at time $t \in (0, n)$. Upon conditioning on what happens in the small time interval $(t, t + dt)$, we find

$$V_t = b_t \mu_{x+t} dt - \pi_t dt + (1 - \mu_{x+t} dt)e^{-r dt} V_{t+dt}. \quad (4.44)$$

Subtract $V_{t+dt}$ on both sides, divide by $dt$ and let $dt$ tend to 0. Observing that $(e^{-r dt} - 1)/dt \to -r$ as $dt \to 0$, one obtains *Thiele’s differential equation*,

$$\frac{d}{dt} V_t = \pi_t - b_t \mu_{x+t} + (r + \mu_{x+t}) V_t, \quad (4.45)$$

valid at each $t$ where $b$, $\pi$, and $\mu$ are continuous. The right hand side expression in (4.45) shows how the fund per surviving policyholder changes per time unit at time $t$. It is increased by the excess of premiums over benefits (which may be negative, of course), by the interest earned, $rV_t$, and by the fund inherited from those who die, $\mu_{x+t} V_t$.

When combined with the boundary condition

$$V_{n-} = b_n, \quad (4.46)$$

the differential equation (4.45) determines $V_t$ for fixed $b$ and $\pi$. 
If the principle of equivalence is exercised, then we must add the condition (4.37). This represents a constraint on the contractual payments $b$ and $\pi$; typically, one first specifies the benefit $b$ and then determines the premium rate for a given premium plan (shape of $\pi$).

**B. Savings premium and risk premium.** Suppose the equivalence principle is in use. Rearrange (4.45) as

$$
\pi_t = \frac{d}{dt} V_t - r V_t + (b - V_t) \mu_{x+t}.
$$

(4.47)

This form of the differential equation shows how the premium at any time decomposes into a savings premium,

$$
\pi^s_t = \frac{d}{dt} V_t - r V_t,
$$

(4.48)

and a risk premium,

$$
\pi^r_t = (b - V_t) \mu_{x+t}.
$$

(4.49)

The savings premium provides the amount needed in excess of the earned interest to maintain the reserve. The risk premium provides the amount needed in excess of the available reserve to cover an insurance claim.

**C. Uses of the differential equation.** In the examples given above, Thiele’s differential equation was useful primarily as a means of investigating the development of the reserve. It was not required in the construction of the premium and the reserve, which could be put up by direct prospective reasoning. In the final example to be given Thiele’s differential equation is needed as a constructive tool.

Assume that the pension treaty studied above is modified so that the reserve is paid back at the moment of death in case the insured dies during the contract period, the philosophy being that “the savings belong to the insured”. Then the scheme is supplied by an $(n + m)$-year temporary term insurance with sum $b_t = V_t$ at any time $t \in (0, m + n)$. The solution to (4.45) is easily obtained as

$$
V_t = \begin{cases} 
\pi^s_t \bar{\mu}_t, & 0 < t < m, \\
b^s\bar{u}_{m+n-t}, & m < t < m + n,
\end{cases}
$$

where $\bar{\mu}_t = \int_0^t (1 + i)^{-\tau} \, d\tau$. The reserve develops just as for ordinary savings contracts offered by banks.

**D. Dependence of the reserve on the contract elements.** A small collection of results due to Lidstone (1905) and, in the time-continuous set-up, Norberg (1985), deal with the dependence of the reserve on the contract elements, in particular mortality and interest.

The starting point in the time-continuous case is Thiele’s differential equation. For the sake of concreteness, we adopt the model assumptions and the
合同描述在第4.4节，并将此称为标准合同。

Thiele的微分方程是

\[
\frac{d}{dt} V_t = \pi_t - \mu_{x+t} b_t + (r_t + \mu_{x+t}) V_t. 
\]

(4.50)

边界条件来自储备的定义

\[
V_{t-} = b_n. 
\]

(4.51)

有保费由等价原则确定，我们也有

\[
V_0 = \pi_0, 
\]

(4.52)

\[\pi_0\]是保单成立时的单次保费（当然可能是0）。

现在考虑一个不同的模型，其中存在不同利益

\[b^*_t\]

和不同合同与利益

\[b^*_t\]和保费

\[\pi^*_0\]。这将被称为特殊合同。此合同下的储备函数

\[V^*_t\]满足

\[
\frac{d}{dt} V^*_t = \pi^*_t - \mu^*_{x+t} b^*_t + (r^*_t + \mu^*_{x+t}) V^*_t, 
\]

(4.53)

\[V^*_n = b^*_n, \]

(4.54)

\[V^*_0 = \pi^*_0. \]

(4.55)

假设

\[\pi^*_0 = \pi_0, \quad b^*_n = b_n. \]

(4.56)

我们感兴趣于差值

\[V^*_t - V_t, \]

和一些话是为了激励这个：储备被计入为保险公司的负债。为了安全起见，公司应在任何时候提供储备超过所需的储备。这通常是通过使用‘技术’元素

\[r^*_t, \mu^*_{x+t}\]

和不同的合同与利益

\[b^*_t\]和保费

\[\pi^*_0\]。这将被称作特殊合同。此合同下的储备函数

\[V^*_t\]满足

\[
\frac{d}{dt} (V^*_t - V_t) = \eta_t + (r^*_t + \mu^*_{x+t}) (V^*_t - V_t), 
\]

(4.57)

其中

\[
\eta_t = (\pi^*_t - \pi_t) + (\mu_{x+t} b_t - \mu^*_{x+t} b^*_t) + (r^*_t - r_t + \mu^*_{x+t} - \mu_{x+t}) V_t. 
\]

(4.58)

将(4.57)从0到\(t\)积分，使用\(V_0 = V_0^*\)，得到

\[
V^*_t - V_t = \int_0^t e^{\int_0^s (r^* + \mu^*)} \eta_s ds. 
\]
Similarly, integrate from \( t \) to \( n \), using \( V_{n-} = V^*_{n-} \), to obtain

\[
V^*_t - V_t = -\int_t^n e^{-\int_t^s (r^* + \mu^*)} \eta_s \, ds.
\]

From these relations conclude: If there exists a \( t_0 \in [0, n] \) such that

\[
\eta_t \leq 0 \text{ for } t < t_0,
\]

then \( V^*_t \leq V_t \) for all \( t \). In particular, this is the case if \( \eta_t \) is non-decreasing. The result remains valid if all inequalities are made strict.

(b) Use the result in (a) to prove the following: For a contract with level premium intensity throughout the contract period, and with non-decreasing reserve, a uniform increase of the interest rate results in a decrease of the reserve.

(c) Consider an endowment insurance with fixed sum insured and level premium rate throughout the contract period. Prove that a change of mortality from \( \mu \) to \( \mu^* \) such that \( \mu^*_t - \mu_t \) is positive and non-increasing, leads to a decrease of the reserve.

(d) Consider a policy with no down premium payment at time 0 and no life endowment at time \( n \). Let the special contract be the same as the standard one, except that the special contract charges so-called natural premium, \( \pi^*_t = b_t \mu_{x+t} \). Then \( V^*_t = 0 \) for all \( t \), and the result in (a) can be used to check whether the reserve \( V_t \) is non-negative (as it should be).

### 4.5 Probability distributions

**A. Motivation.** The basic paradigm being the principle of equivalence, life insurance mathematics centers on expected present values. The key tool is Thiele’s differential equation, which describes the development of such expected values and forms a basis for computing them by recursive methods. In Chapter 7 we shall obtain analogous differential equations for higher order moments, which will enable us to compute the variance, skewness, kurtosis, and so on of the present value of payments under a fairly general insurance contract.

We shall give an example of how to determine the probability distribution of a present value, which is at the base of the moments and of any other expected values of interest. Knowledge of this distribution, and in particular its upper tail, gives insight into the riskiness of the contract beyond what is provided by the mean and some higher order moments.

The task is easy for an insurance on a single life since then the model involves only one random variable (the life length of the insured). De Pril [8] and Dhaene [10] offer a number of examples. In principle the task is simple also for
insurances involving more than one life or, more generally, a finite number of random variables. In such situations the distributions of present values (and any other functions of the random variables) can be obtained by integrating the finite-dimensional distribution.

B. A simple example. Consider the single life status \((x)\) with remaining life time \(T_x\) distributed as described in Chapter 3. Suppose \((x)\) buys an \(n\) year term insurance with fixed sum \(b\) and premiums payable continuously at level rate \(\pi\) per year as long as the contract is in force (confer Paragraphs 4.1.C-D). The present value of benefits less premiums on the contract is

\[
U(T_x) = be^{-\pi T_x}1_{(0<T_x<n)} - \pi \bar{a}_{T_x\wedge n},
\]

where \(\bar{a}_t = \int_0^t e^{-r\tau} d\tau = (1 - e^{-rt})/r\) is the present value of an annuity certain payable continuously at level rate 1 per year for \(t\) years. The function \(U\) is non-increasing in \(T_x\), and we easily find the probability distribution

\[
P[U \leq u] = \begin{cases} 
0, & u < -\pi \bar{a}_n, \\
\mathbb{P}[T_x > n], & -\pi \bar{a}_n \leq u < be^{-rn} - \pi \bar{a}_n, \\
\mathbb{P}[T_x > \frac{1}{r} \ln \left( \frac{be^{-rn} - \pi \bar{a}_n}{u + \pi} \right)], & be^{-rn} - \pi \bar{a}_n \leq u < b, \\
1, & u \geq b.
\end{cases}
\tag{4.59}
\]

The jump at \(-\pi \bar{a}_n\) is due to the positive probability of survival to time \(n\). Similar effects are to be anticipated also for other insurance products with a finite contract period since, in general, there is a positive probability that the policy will remain in the current state until the contract terminates.

The probability distribution in (4.59) is depicted in Fig. 4.6 for the G82M case with \(r = \ln(1.045)\) and \(\mu(t|x) = 0.0005 + 10^{-4.12+0.038(x+t)}\) when \(x = 30, n = 30, b = 1,\) and \(\pi = 0.0042608\) (the equivalence premium).

4.6 The stochastic process point of view

A. The processes indicating survival and death. In Paragraph A of Section 4.1 we introduced the indicator of the event of survival to time \(t\), \(I_t = 1_{(T_x > t)}\), and the indicator of the complementary event of death within time \(t\), \(N_t = 1 - I_t = 1_{(T_x \leq t)}\). Viewed as functions of \(t\), they are stochastic processes. The latter counts the number of deaths of the insured as time is progresses and is thus a simple example of a counting process as defined in Paragraph D of Appendix A. This motivates the notation \(N_t\). By their very definitions, \(I_t\) and \(N_t\) are RC.

In the present context, where everything is governed by just one single random variable, \(T_x\), the process point of view is not important for practical purposes. For didactical purposes, however, it is worthwhile taking it already here as a rehearsal for more complicated situations where stochastic processes cannot be dispensed with.
Figure 4.6: The probability distribution of the present value of a term insurance against level premium.

The payment functions of the benefits considered in Section 4.1 can be recast in terms of the processes $I_t$ and $N_t$. In differential form they are

$$
\begin{align*}
    dA^c_t &= I_t d\varepsilon_n(t), \\
    dA^{i; n}_t &= (1 - \varepsilon_n(t)) dN_t, \\
    dA^a_t &= (1 - \varepsilon_n(t)) I_t dt, \\
    dA^{e; n}_t &= dA^{c; n}_t + dA^{e; n}_t. 
\end{align*}
$$

Their present values are

$$
\begin{align*}
    V^{c; n} &= e^{-\int_0^\infty r I_n}, \\
    V^{i; n} &= \int_0^n e^{-\int_0^\tau r} dN_\tau, \\
    V^{a; n} &= \int_0^n e^{-\int_0^\tau r} I_\tau d\tau, \\
    V^{e; n} &= V^{i; n} + V^{e; n}.
\end{align*}
$$

The expressions in (4.14) and (4.10) are obtained directly by taking expectation under the integral sign (Fubini), using the obvious relations

$$
\begin{align*}
    \mathbb{E}[I_t] &= \tau p_x, \\
    \mathbb{E}[dN_t] &= \tau p_x \mu_{x+\tau} d\tau.
\end{align*}
$$

The relationship (4.19) reemerges in its more basic form upon integrating by parts to obtain

$$
e^{-\int_0^\infty r I_n} = 1 + \int_0^n e^{-\int_0^\tau r (-r) I_\tau} d\tau + \int_0^n e^{-\int_0^\tau r} dI_\tau,$$

and setting $dI_t = -dN_t$ in the last integral.
Chapter 5

Expenses

5.1 A single life insurance policy

A. Three categories of expenses. Any firm has to defray expenditures in addition to the net production costs of the commodities or services it offers, and these expenses must be taken account of in the prices paid by the customers. Thus, the rate of premium charged for a given insurance contract must not merely cover the contractual net benefits, but also be sufficient to provide for all items of expenditure connected with the operations of the insurance company.

For the sake of concreteness, and also of loyalty to standard actuarial notation, we shall introduce the issue of expenses in the framework of the simple single life policy encountered in Chapter 4. To get a case that involves all main types of payments, let us consider a life \((x)\) who purchases an \(n\)-year endowment insurance with a fixed sum insured, \(b\), and premium payable continuously at level rate as long as the policy is in force.

We recall that the net premium rate determined by the principle of equivalence is

\[
\pi = b \frac{\bar{A}_x \bar{m}}{\bar{a}_x \bar{m}} = b \left( \frac{1}{\bar{a}_x \bar{m}} - r \right),
\]

and that the corresponding net premium reserve to be provided if the insured is alive at time \(t\), is

\[
V_t = b\bar{A}_{x+t} \underline{\bar{m}} = \pi \bar{a}_{x+t} \underline{\bar{m}} = b \left( 1 - \frac{\bar{a}_{x+t} \underline{\bar{m}}}{\bar{a}_x \bar{m}} \right).
\]

The term *net* means “net of administration expenses”.

When expenses are included in the accounts, one will have to charge the policy with a gross premium rate \(\pi'\), which obviously must be greater than the net premium rate, and the gross premium reserve \(V'_t\) to be provided if the policy is in force at time \(t\) will also in general differ from the net premium reserve. The precise definitions of these quantities can only be made after we
have made specific assumptions about the structure of the expenses, which we now turn to.

The expenses are usually divided into three categories. In the first place there are the so-called $\alpha$-expenses that incur in connection with the establishment of the contract. They comprise sales costs, including advertising and agent’s commission, and costs connected with health examination, issue of the policy, entering the details of the contract into the data files, etc. It is assumed that these expenses incur immediately at time 0 and that they are of the form

$$\alpha' + \alpha'' b.$$  \hspace{1cm} (5.3)

In the second place there are the so-called $\beta$-expenses that incur in connection with encashment and accounting of premiums. They are assumed to incur continuously at constant rate

$$\beta' + \beta'' \pi'$$  \hspace{1cm} (5.4)

throughout the premium-paying period.

Finally, in the third place there are the so-called $\gamma$-expenses that comprise all expenditures not included in the former two categories, such as wages to employees, rent, taxes, fees, and maintenance of the business operations in general. These expenses are assumed to incur continuously at rate

$$\gamma' + \gamma'' b + \gamma''' V'_t$$  \hspace{1cm} (5.5)

at time $t$ if the policy is then in force.

The constant terms $\alpha'$, $\beta'$, and $\gamma'$ represent costs that are the same for all policies. The terms $\alpha'' b$, $\beta'' \pi'$, and $\gamma'' b$ represent costs that are proportional to the size of the contract as measured by the amounts specified in the policy. Typically this is the case for the agent’s commission, which may be a considerable portion of the $\alpha$-expenses on individual insurances sold in an open competitive market, and also for the debt collector’s or solicitor’s commission, which in former days made up the major part of the $\beta$-expenses. The term $\gamma''' V'_t$ represents expenses in connection with management of the investment portfolio, which can reasonably be divided between the policyholders in proportion to their current reserves.

**B. The gross premium and the gross premium reserve.** Upon exercising the equivalence principle in the presence of expenses, one will determine the gross premium rate $\pi'$ and the corresponding gross premium reserve function $V'_t$. When expenses depend on the reserve, as specified in (5.5), we have to resort to the Thiele technique to construct $\pi'$ and $V'_t$. We can immediately put up the following differential equation by adding the expenses to the benefits in the set-up of Section 4.4:

$$\frac{d}{dt} V'_t = \pi' - \beta' - \beta'' \pi' - \gamma' - \gamma'' b - \gamma''' V'_t - \mu_{x+t} b + r V'_t + \mu_{x+t} V'_t.$$  \hspace{1cm} (5.6)
The appropriate side conditions are
\[ V'_{n-} = b , \]  
(5.7)
and
\[ V'_0 = -(\alpha' + \alpha''b) . \]  
(5.8)
As before, (5.7) is a matter of definition and relies only on the fact that the endowment benefit falls due upon survival at time \( n \), and (5.8) is the equivalence requirement, which determines \( \pi' \) for given benefits and expense factors.

Gathering terms involving \( V'_t \) on the left of (5.6) and multiplying on both sides with \( e^{\int_r (r - \gamma''') dt} \) gives
\[
\frac{d}{dt} \left( e^{\int_r (r - \gamma''') dt} V'_t \right) = e^{\int_r (r - \gamma''') dt} \left\{ (1 - \beta'') \pi' - \beta' - \gamma' - (\gamma'' + \mu x + \tau) b \right\} .
\]  
(5.9)
Now integrate (5.9) between \( t \) and \( n \), using (5.7), and rearrange a bit to obtain
\[
V'_t = \int_t^n e^{-\int_r (r - \gamma''') dt} (\beta' + \gamma' + (\gamma'' + \mu x + \tau) b - (1 - \beta'') \pi') d\tau + e^{-\int_0^n (r - \gamma''') dt} b .
\]  
(5.10)
Upon inserting \( t = 0 \) into (5.10) and using (5.8), we find
\[
\pi' = \frac{\alpha' + \alpha''b + \int_0^n e^{-\int_r (r - \gamma''') dt} (\beta' + \gamma' + (\gamma'' + \mu x + \tau) b) d\tau + e^{-\int_0^n (r - \gamma''') dt} b}{(1 - \beta'') \int_0^n e^{-\int_r (r - \gamma''') dt} d\tau} .
\]  
(5.11)
The practical calculations start with determining \( \pi' \) in (5.11). The easiest way is to compute the expressions in numerator and denominator by the program 'proresin.pas'. Then compute the function \( V'_t \) in (5.10) using 'proresin.pas' again.

In the special case where \( \gamma''' = 0 \) we could determine \( \pi' \) and \( V'_t \) directly from the defining relations without using the differential equation. That goes, in fact, also for the general case with \( \gamma''' \neq 0 \) by the following consideration: By inspection of the differential equation (5.6) and the side conditions, it is realized that, formally, the problem amounts to determining the “net premium rate” \( (1 - \beta'') \pi' \) and “net premium reserve” \( V'_t \) for a policy with (admittedly unrealistic) benefits consisting of a lump sum payment of \( \alpha' + \alpha''b \) at time 0, a continuous level life annuity of \( \beta' + \gamma' + \gamma''b \) per year, and an endowment insurance of \( b \), when the interest rate is \( r - \gamma''' \).

Easy calculations show that, when \( \gamma''' = 0 \), the gross and net quantities are related by
\[
\pi' = \frac{1}{1 - \beta''} \left( \pi + \frac{\alpha' + \alpha''b}{a_x m} + \beta' + \gamma' + \gamma''b \right) ,
\]  
(5.12)
and
\[ V_t' = V_t - \frac{\bar{a}_{x+t\,n-1}}{a_x} (\alpha' + \alpha'') \]  

(5.13)

It is seen that \( \pi' > \pi \), as was anticipated at the outset. Furthermore, \( V_t' < V_t \) for \( 0 \leq t < n \), which may be less obvious. The relationship (5.13) can be explained as follows: All expenses that incur at a constant rate throughout the term of the contract are compensated by an equal component in the "effective" gross premium rate \( 1 - \beta'' \pi' \), confer (5.12). Thus, the only expense factor that appears in the gross reserve is the unamortized initial \( \alpha \)-cost, which is the last term on the right of (5.13). It represents a debt on the part of the insured and is therefore to be subtracted from the net reserve.

In Paragraph 4.3.D we have advocated non-negativity of the reserve. Now, already from (5.8) it is clear that the gross premium sets out negative at the time of issue of the contract and it will remain negative for some time thereafter until a sufficient amount of premium has been collected. The only way to get around this problem would be to charge an initial lump sum premium no less than the initial expense, but this is usually not done in practice (presumably) because a substantial down payment might keep customers with liquidity problems from buying insurance.

5.2 The general multi-state policy

A. General treatment of expenses. Consider now the general multi-state insurance policy treated in Chapter 7. Expenses are easily accommodated in the theory of that chapter since they can be treated as additional benefits of annuity and assurance type. Thus, from a technical point of view expenses do not create any additional difficulties, and we can therefore suitably end this chapter here. We round off by saying that expenses are still of conceptual and great practical importance. Assumptions about the various forms of expenses are part of the technical basis, which must be verified by the insurer and is subject to approval of the supervisory authority. Thus, just as statistical and economic analyses are required as a basis for assumptions about mortality and interest, thorough cost analyses are required as a basis for assumptions about the expense factors.
Chapter 6

Multi-life insurances

6.1 Insurances depending on the number of survivors

A. The single-life status reinterpreted. In the treatment of the single life status \( (x) \) in Chapters 3–4 we were having in mind the remaining life time \( T \) of an \( x \) year old person. From a mathematical point of view this interpretation is not essential. All that matters is that \( T \) is a non-negative random variable with an absolutely continuous distribution function, so that the survival function is of the form

\[
t_p_x = e^{-\int_0^t \mu_{x+r} \, dr}.
\]  

(6.1)

The footscript \( x \) serves merely to indicate what mortality law is in play. Regardless of the nature of the status \( (x) \) and the notion of lifetime represented by \( T \), the previous results remain valid. In particular, all formulas for expected present values of payments depending on \( T \) are preserved, the basic ones being the endowment,

\[
E_x^n = v^n p_x,
\]  

(6.2)

the life annuity,

\[
\bar{a}_x = \int_0^\infty v^t p_x \, dt = \int_0^\infty t E_x \, dt,
\]  

(6.3)

the endowment insurance,

\[
\bar{A}_x \equiv 1 - r \bar{a}_x,
\]  

(6.4)

and the term insurance,

\[
\bar{A}_x^1 = \bar{A}_x - E_x.
\]  

(6.5)

These formulas demonstrate that present values of all main types of payments in life insurance — endowments, life annuities, and assurances — can be traced...
back to the present value \( tE_x \) of an endowment and, as far as the mortality law is concerned, to the survival function \( t p_x \). Once we have determined \( t p_x \), all other functions of interest are obtained by integration, possibly by some numerical method, and elementary algebraic operations.

**B. Multidimensional survival functions.** Consider a body of \( r \) individuals, the \( j \)-th of which is called \((x_j)\) and has remaining lifetime \( T_j, j = 1, \ldots, r \). For the time being we shall confine ourselves to the case with independent lives. Thus, assume that the \( T_j \) are stochastically independent, and that each \( T_j \) possesses an intensity denoted by \( \mu_{x_j} + t \) and, hence, has survival function

\[
 t p_{x_j} = e^{-\int_0^t \mu_{x_j} + \tau \, d\tau}.
\]  

(6.6)

(The function \( \mu \) need not be the same for all \( j \) as the notation suggests; we have dropped an extra index \( j \) just to save notation.) The simultaneous distribution of \( T_1, \ldots, T_r \) is given by the multidimensional survival function

\[
 P \left[ \bigcap_{j=1}^r \{ T_j > t_j \} \right] = \prod_{j=1}^r t_j p_{x_j} = e^{-\sum_{j=1}^r \int_0^t \mu_{x_j} + \tau \, d\tau}
\]  

or, equivalently, by the density

\[
 \prod_{j=1}^r t_j p_{x_j} \mu_{x_j} + t_j.
\]  

(6.7)

**C. The joint-life status.** The joint life status \((x_1 \ldots x_r)\) is defined by having remaining lifetime

\[
 T_{x_1 \ldots x_r} = \min\{T_1, \ldots, T_r\}.
\]  

(6.8)

Thus, the \( r \) lives are looked upon as a single entity, which continues to exist as long as all members survive, and terminates upon the first death. The survival function of the joint-life is denoted by \( t p_{x_1 \ldots x_r} \) and is

\[
 t p_{x_1 \ldots x_r} = P \left[ \bigcap_{j=1}^r \{ T_j > t_j \} \right] = e^{-\int_0^t \sum_{j=1}^r \mu_{x_j} + \tau \, d\tau}.
\]  

(6.9)

From this survival function we form the present values of an endowment \( nE_{x_1 \ldots x_r} \), a life annuity \( \bar{a}_{x_1 \ldots x_r} \), an endowment insurance \( \bar{A}_{x_1 \ldots x_r} \), and a term insurance, \( A^1_{x_1 \ldots x_r} \), by just putting (6.9) in the role of the survival function in (6.2) – (6.5).

By inspection of (6.9), the mortality intensity of the joint-life status is simply the sum of the component mortality intensities,

\[
 \mu_{x_1 \ldots x_r}(t) = \sum_{j=1}^r \mu_{x_j+t}.
\]  

(6.10)
In particular, if the component lives are subject to G-M mortality laws with a common value of the parameter $c$,

$$
\mu_{x_j+t} = \alpha_j + \beta_j c^{x_j+t},
$$

(6.11)

then (6.10) becomes

$$
\mu_{x_1 \ldots x_r}(t) = \alpha' + \beta' c^t
$$

(6.12)

with

$$
\alpha' = \sum_{j=1}^{r} \alpha_j, \quad \beta' = \sum_{j=1}^{r} \beta_j c^{x_j},
$$

(6.13)

again a G-M law with the same $c$ as in the component laws.

**D. The last-survivor status.** The last survivor status $x_1 \ldots x_r$ is defined by having remaining lifetime

$$
T_{x_1 \ldots x_r} = \max\{T_1, \ldots, T_r\}.
$$

(6.14)

Now the $r$ lives are looked upon as an entity that continues to exist as long as at least one member survives, and terminates upon the last death. The survival function of this status is denoted by $t p_{x_1 \ldots x_r}$. By the general addition rule for probabilities (Appendix C),

$$
t p_{x_1 \ldots x_r} = \mathbb{P}\left[ \bigcup_{j=1}^{r} \{ T_j > t \} \right] = \sum_j t p_{x_j} - \sum_{j_1 < j_2} t p_{x_{j_1} x_{j_2}} + \ldots + (-1)^{r-1} t p_{x_1 \ldots x_r}.
$$

(6.15)

This way actuarial computations for the last survivor are reduced to computations for joint lives, which are simple. As explained in Paragraph A, all main types of present values can be built from (6.15). Formulas for benefits contingent on survival, obtained from (6.2) and (6.3), will reflect the structure of (6.15) in an obvious way. Formulas for death benefits are obtained from (6.4) and (6.5). The expressions are displayed in the more general case to be treated in the next paragraph.

**E. The q survivors status.**

The $q$ survivors status $x_{q+1} \ldots x_r$ is defined by having as remaining lifetime the $(r-q+1)$-th order statistic of the sample $\{T_1, \ldots, T_r\}$. Thus the status is "alive" as long as there are at least $q$ survivors among the original $r$. The survival function can be expressed in terms of joint life survival functions of subgroups of lives by direct application of the theorem in Appendix C:

$$
t p_{x_{q+1} \ldots x_r} = \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{q-1} \sum_{j_1 < \ldots < j_p} t p_{x_{j_1} \ldots x_{j_p}}.
$$

(6.16)
Present values of standard forms of insurances for the \( q \) survivors status are now obtained along the lines described in the previous paragraph. First, combine (6.16) with (6.2) to obtain

\[
nE_{x_1 \ldots x_p} = \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{q-1} \sum_{j_1 < \ldots < j_p} nE_{x_{j_1} \ldots x_{j_p}},
\]

the notation being self-explaining. Next, combine (6.3) and (6.17) to obtain

\[
a_{x_1 \ldots x_r} \eta = \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{q-1} \sum_{j_1 < \ldots < j_p} a_{x_{j_1} \ldots x_{j_p}} \eta.
\]

Finally, present values of endowment and term insurances are obtained by inserting (6.18) and (6.17) in the general relations (6.4) and (6.5):

\[
\bar{A}_{x_1 \ldots x_r} \eta = 1 - r \bar{a}_{x_1 \ldots x_r} \eta,
\]

\[
\bar{A}^1_{x_1 \ldots x_r} \eta = \bar{A}_{x_1 \ldots x_r} \eta - nE_{x_{j_1} \ldots x_{j_p}}
\]

The following alternative to the expression in (6.19) has some aesthetic appeal as it expresses the insurance by corresponding insurances on joint lives:

\[
\bar{A}_{x_1 \ldots x_r} \eta = \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{q-1} \sum_{j_1 < \ldots < j_p} \bar{A}_{x_{j_1} \ldots x_{j_p}} \eta.
\]

It is obtained upon substituting (6.18) on the right of (6.19), then inserting (recall (6.4)) \( \bar{a}_{x_{j_1} \ldots x_{j_p}} = (1 - \bar{A}_{x_{j_1} \ldots x_{j_p}})/r \) and using (C.8) in Appendix C. A similar expression for the term insurance is obtained upon subtracting (6.17) from (6.21).
Chapter 7

Markov chains in life insurance

7.1 The insurance policy as a stochastic process

A. The basic entities. Consider an insurance policy issued at time 0 for a finite term of \( n \) years. We have in mind life or pension insurance or some other form of insurance of persons like disability or sickness coverage. In such lines of business benefits and premiums are typically contingent upon transitions of the policy between certain states specified in the contract. Thus, we assume there is a finite set of states, \( Z = \{0, 1, \ldots, r\} \), such that the policy at any time is in one and only one state, commencing in state 0 (say) at time 0. Denote the state of the policy at time \( t \) by \( Z(t) \). Regarded as a function from \([0, n]\) to \( Z \), \( Z \) is assumed to be right-continuous, with a finite number of jumps, and \( Z(0) = 0 \). To account for the random course of the policy, \( Z \) is modelled as a stochastic process on some probability space \((\Omega, \mathcal{H}, \mathbb{P})\).

B. Model deliberations; realism versus simplicity. On specifying the probability model, two concerns must be kept in mind, and they are inevitably conflicting. On the one hand, the model should reflect the essential features of (a certain piece of) reality, and this speaks for a complex model to the extent that reality itself is complex. On the other hand, the model should be mathematically tractable, and this speaks for a simple model allowing of easy computation of quantities of interest. The art of modelling is to strike the right balance between these two concerns.

Favouring simplicity in the first place, we shall be working under Markov assumptions, which allow for fairly easy computation of relevant probabilities and expected values. Later on we shall demonstrate the versatility of this model framework, showing that it is capable of representing virtually any conception one might have of the mechanisms governing the development of the policy. We shall take the Markov chain model presented in [13] as a suitable framework.
7.2 The time-continuous Markov chain

A. The Markov property. A stochastic process is essentially determined by its finite-dimensional distributions. In the present case, where $Z$ has only a finite state space, these are fully specified by the probabilities of the elementary events $\cap_{h=1}^p \{ Z(t_h) = j_h \}$, $t_1 < \cdots < t_p$ in $[0, n]$ and $j_1, \ldots, j_p \in Z$. Now

$$\mathbb{P} [ Z(t_h) = j_h, h = 1, \ldots, p ] = \prod_{h=1}^p \mathbb{P} [ Z(t_h) = j_h | Z(t_g) = j_g, g = 0, \ldots, h - 1 ] , \quad (7.1)$$

where, for convenience, we have put $t_0 = 0$ and $j_0 = 0$ so that $\{ Z(t_0) = j_0 \}$ is the trivial event with probability 1. Thus, the specification of $\mathbb{P}$ could suitably start with the conditional probabilities appearing on the right of (7.1).

A particularly simple structure is obtained by assuming that, for all $t_1 < \cdots < t_p$ in $[0, n]$ and $j, k \in Z$,

$$\mathbb{P} [ Z(t_p) = j_p | Z(t_h) = j_h, h = 1, \ldots, p - 1 ] = \mathbb{P} [ Z(t_p) = j_p | Z(t_{p-1}) = j_{p-1} ] , \quad (7.2)$$

which means that process is fully determined by the (simple) transition probabilities

$$p_{jk}(t, u) = \mathbb{P} [ Z(u) = k | Z(t) = j ] , \quad (7.3)$$

$t < u$ in $[0, n]$ and $j, k \in Z$. In fact, if (7.2) holds, then (7.1) reduces to

$$\mathbb{P} [ Z(t_h) = j_h, h = 1, \ldots, p ] = \prod_{h=1}^p p_{j_{h-1}j_h}(t_{h-1}, t_h) , \quad (7.4)$$

and one easily proves the equivalent that, for any $t_1 < \cdots < t_p < t < t_{p+1} < \cdots < t_{p+q}$ in $[0, n]$ and $j, j_p, j, j_{p+1}, \ldots, j_{p+q}$ in $Z$,

$$\mathbb{P} [ Z(t_h) = j_h, h = p + 1, \ldots, p + q | Z(t) = j, Z(t_h) = j_h, h = 1, \ldots, p ] = \mathbb{P} [ Z(t_h) = j_h, h = p + 1, \ldots, p + q | Z(t) = j ] . \quad (7.5)$$

Proclaiming $t$ “the present time”, (7.5) says that the future of the process is independent of its past when the present is known. (Fully known, that is; if the present state is only partly known, it may certainly help to add information about the past.)

The condition (7.2) is called the Markov property. We shall assume that $Z$ possesses this property and, accordingly, call it a continuous time Markov process on the state space $Z$. 

throughout this text. A useful basic source is [16].
CHAPTER 7. MARKOV CHAINS IN LIFE INSURANCE

From the simple transition probabilities we form the more general transition probability from \( j \) to some subset \( K \subset \mathbb{Z} \),
\[
p_{jK}(t, u) = \mathbb{P}[Z(u) \in K \mid Z(t) = j] = \sum_{k \in K} p_{jk}(t, u).
\] (7.6)
We have, of course,
\[
p_{j\mathbb{Z}}(t, u) = \sum_{k \in \mathbb{Z}} p_{jk}(t, u) = 1.
\] (7.7)

B. Alternative definitions of the Markov property. It is straightforward to demonstrate that (7.2), (7.4), and (7.5) are equivalent, so that any one of the three could have been taken as definition of the Markov property. Then (7.4) should be preceded by: “Assume there exist non-negative functions \( p_{jk}(t, u) \), \( j, k \in \mathbb{Z}, 0 \leq t \leq u \leq n \), such that \( \sum_{k \in \mathbb{Z}} p_{jk}(t, u) = 1 \) and, for any \( 0 \leq t_1 < \cdots < t_p \in [0, n] \) and \( \{j_1, \ldots, j_p\} \subset \mathbb{Z} \),”

We shall briefly outline more general definitions of the Markov property. For \( T \subset [0, n] \) let \( \mathcal{H}_T \) denote the class of all events generated by \( \{Z(t)\}_{t \in T} \). It represents everything that can be observed about \( Z \) in the time set \( T \). For instance, \( \mathcal{H}_{\{t\}} \) is the information carried by the process at time \( t \) and consists of the elementary events \( \emptyset, \Omega, \) and \( Z(t) = j \), \( j = 0, \ldots, r \), and all possible unions of these events. More generally, \( \mathcal{H}_{[t_1, \ldots, t_p]} \) is the information carried by the process at times \( t_1, \ldots, t_p \). Some sets \( T \) of interval type are frequently encountered, and we abbreviate \( \mathcal{H}_{\leq t} = \mathcal{H}_{[0, t]} \) (the entire history of the process by time \( t \)), \( \mathcal{H}_{< t} = \mathcal{H}_{(0, t]} \) (the strict past of the process by time \( t \)), and \( \mathcal{H}_{> t} = \mathcal{H}_{(t, n]} \) (the future of the process by time \( t \)).

The process \( Z \) is said to be a Markov process if, for any \( B \in \mathcal{H}_{> t} \),
\[
\mathbb{P}[B \mid \mathcal{H}_{\leq t}] = \mathbb{P}[B \mid \mathcal{H}_{\{t\}}].
\] (7.8)
This is the general form of (7.5).

An alternative definition says that, for any \( A \in \mathcal{H}_{< t} \) and \( B \in \mathcal{H}_{> t} \),
\[
\mathbb{P}[A \cap B \mid \mathcal{H}_{\{t\}}] = \mathbb{P}[A \mid \mathcal{H}_{\{t\}}] \mathbb{P}[B \mid \mathcal{H}_{\{t\}}],
\] (7.9)
that is, the past and the future of the process are conditionally independent, given its present state. In the case with finite state space (countability is equally simple) it is easy to prove that (7.8) and (7.9) are equivalent by working with the finite-dimensional distributions, that is, take \( A \in \mathcal{H}_{(t_1, \ldots, t_p)} \) and \( B \in \mathcal{H}_{(t_{p+1}, \ldots, t_{p+q})} \) with \( t_1 < \cdots < t_p < t < t_{p+1} < \cdots < t_{p+q} \).

C. The Chapman-Kolmogorov equation. For a fixed \( t \in [0, n] \) the events \( \{Z(t) = j\}, j \in \mathbb{Z} \), are disjoint and their union is the almost sure event. It follows that
\[
\mathbb{P}[Z(u) = k \mid Z(s) = i] = \sum_{j \in \mathbb{Z}} \mathbb{P}[Z(t) = j, Z(u) = k \mid Z(s) = i] = \sum_{j \in \mathbb{Z}} \mathbb{P}[Z(t) = j \mid Z(s) = i] \mathbb{P}[Z(u) = k \mid Z(s) = i, Z(t) = j].
\]
CHAPTER 7. MARKOV CHAINS IN LIFE INSURANCE

If \( Z \) is Markov, and \( 0 \leq s \leq t \leq u \), this reduces to

\[
p_{ik}(s, u) = \sum_{j \in Z} p_{ij}(s, t) p_{jk}(t, u),
\]

which is known as the Chapman-Kolmogorov equation.

D. Intensities of transition. In principle, specifying the Markov model amounts to specifying the \( p_{jk}(t, u) \) in such a manner that the expressions on the right of (7.4) define probabilities in a consistent way. This would be easy if \( Z \) were a discrete time Markov chain with \( t \) ranging in a finite time set \( 0 = t_0 < t_1 < \cdots < t_q = \tau \): then we could just take the values of the \( p_{jk}(t_{p-1}, t_p) \) as any non-negative numbers satisfying \( \sum_{k=0}^{q} p_{jk}(t_{p-1}, t_p) = 1 \) for each \( j \in Z \) and \( p = 1, \ldots, q \). This simple device does not carry over without modification to the continuous time case since there are no smallest finite time intervals from which we can build all probabilities by (7.4). An obvious way of adapting the basic idea to the time-continuous case is to add smoothness assumptions that give meaning to a notion of transition probabilities in infinitesimal time intervals.

More specifically, we shall assume that the intensities of transition,

\[
\mu_{jk}(t) = \lim_{h \downarrow 0} \frac{p_{jk}(t, t + h)}{h},
\]

exist for each \( j, k \in Z, j \neq k, \) and \( t \in [0, \tau) \) and, moreover, that they are piecewise continuous. Another way of phrasing (7.11) is

\[
p_{jk}(t, t + dt) = \mu_{jk}(t) dt + o(dt),
\]

where the term \( o(dt) \) is such that \( o(dt)/dt \to 0 \) as \( dt \to 0 \). Thus, transition probabilities over a short time interval are assumed to be (approximately) proportional to the length of the interval, and the proportionality factors are just the intensities, which may depend on the time. What is “short” in this connection depends on the sizes of the intensities. For instance, if the \( \mu_{jk}(\tau) \) are approximately constant and \( < < 1 \) for all \( k \neq j \) and all \( \tau \in [t, t + 1] \), then \( \mu_{jk}(t) \) approximates the transition probability \( p_{jk}(t, t + 1) \). In general, however, the intensities may attain any positive values and should not be confused with probabilities.

For \( j \notin K \subset Z \), we define the intensity of transition from state \( j \) to the set of states \( K \) at time \( t \) as

\[
\mu_{jk}(t) = \lim_{u \uparrow t} \frac{p_{jk}(l, t)}{u - t} = \sum_{k \in K} \mu_{jk}(t).
\]

In particular, the total intensity of transition out of state \( j \) at time \( t \) is \( \mu_{j, Z - \{j\}}(t) \), which is abbreviated

\[
\mu_{j}(t) = \sum_{k, k \neq j} \mu_{jk}(t).
\]
From (7.7) and (7.12) we get
\[ p_{jj}(t, t + dt) = 1 - \mu_j(t) dt + o(dt). \] (7.15)

E. The Kolmogorov differential equations.
The transition probabilities are two-dimensional functions of time, and in non-trivial situations it is virtually impossible to specify them directly in a consistent manner or even figure how they should look on intuitive grounds. The intensities, however, are one-dimensional functions of time and, being easily interpretable, they form a natural starting point for specification of the model. Luckily, as we shall now see, they are also basic entities in the system as they determine the transition probabilities uniquely.

Suppose the process \( Z \) is in state \( j \) at time \( t \). To find the probability that the process will be in state \( k \) at a given future time \( u \), let us condition on what happens in the first small time interval \( (t, t + dt) \). In the first place \( Z \) may remain in state \( j \) with probability \( 1 - \mu_j(t) dt \) and, conditional on this event, the probability of ending up in state \( k \) at time \( u \) is \( p_{jk}(t + dt, u) \). In the second place, \( Z \) may jump to some other state \( g \) with probability \( \mu_{jg}(t) dt \) and, conditional on this event, the probability of ending up in state \( k \) at time \( u \) is \( p_{gk}(t + dt, u) \). Thus, the total probability of \( Z \) being in state \( k \) at time \( u \) is
\[ p_{jk}(t, u) = (1 - \mu_j(t) dt) p_{jk}(t + dt, u) \]
\[ + \sum_{g \neq j} \mu_{jg}(t) dt p_{gk}(t + dt, u) + o(dt), \] (7.16)

Upon putting \( \frac{d}{dt} p_{jk}(t, u) = p_{jk}(t + dt, u) - p_{jk}(t, u) \) in the infinitesimal sense, we arrive at
\[ \frac{d}{dt} p_{jk}(t, u) = \mu_j(t) dt p_{jk}(t, u) - \sum_{g \neq j} \mu_{jg}(t) dt p_{gk}(t, u). \] (7.17)

For given \( k \) and \( u \) these differential equations determine the functions \( p_{jk}(\cdot, u) \), \( j = 0, \ldots, r \), uniquely when combined with the obvious conditions
\[ p_{jk}(u, u) = \delta_{jk}. \] (7.18)

Here \( \delta_{jk} \) is the Kronecker delta defined as 1 if \( j = k \) and 0 otherwise.

The relation (7.16) could have been put up directly by use of the Chapman-Kolmogorov equation (7.10), with \( s, t, i, j \) replaced by \( t, t + dt, j, g \), but we have carried through the detailed (still informal though) argument above since it will be in use repeatedly throughout the text. It is called the backward (differential) argument since it focuses on \( t \), which in the perspective of the considered time period \( [t, u] \) is the very beginning. Accordingly, (7.17) is referred to as the Kolmogorov backward differential equations, being due to A.N. Kolmogorov.

At points of continuity of the intensities we can divide by \( dt \) in (7.17) and obtain a limit on the right as \( dt \) tends to 0. Thus, at such points we can write
(7.17) as
\[
\frac{\partial}{\partial t} p_{jk}(t, u) = \mu_j(t) p_{jk}(t, u) - \sum_{g \neq j} \mu_{jg}(t) p_{gk}(t, u).
\]  
(7.19)

Since we have assumed that the intensities are piecewise continuous, the indicated derivatives exist piecewise. We prefer, however, to work with the differential form (7.17) since it is generally valid under our assumptions and, moreover, invites algorithmic reasoning; numerical procedures for solving differential equations are based on approximation by difference equations for some fine discretization and, in fact, (7.16) is basically what one would use with some small \(dt > 0\).

As one may have guessed, there exist also Kolmogorov forward differential equations. These are obtained by focusing on what happens at the end of the time interval in consideration. Reasoning along the lines above, we have
\[
p_{ij}(s, t + dt) = \sum_{g \neq j} p_{ig}(s, t) \mu_{gj}(t) dt + p_{ij}(s, t)(1 - \mu_j(t) dt) + o(dt),
\]  
(7.20)

hence
\[
d_s p_{ij}(s, t) = \sum_{g \neq j} p_{ig}(s, t) \mu_{gj}(t) dt - p_{ij}(s, t) \mu_j(t) dt.
\]  
(7.21)

For given \(i\) and \(s\), the differential equations (7.20) determine the functions \(p_{ij}(s, \cdot)\), \(j = 0, \ldots, r\), uniquely in conjunction with the obvious conditions
\[
p_{ij}(s, s) = \delta_{ij}.
\]  
(7.22)

In some simple cases the differential equations have nice analytical solutions, but in most non-trivial cases they must be solved numerically, e.g. by the Runge-Kutta method.

Once the simple transition probabilities are determined, we may calculate the probability of any event in \(H_{\{t_1, \ldots, t_r\}}\) from the finite-dimensional distribution (7.4). In fact, with finite \(Z\) every such probability is just a finite sum of probabilities of elementary events to which we can apply (7.4).

Probabilities of more complex events that involve an infinite number of coordinates of \(Z\), e.g. events in \(H_T\) with \(T\) an interval, cannot in general be calculated from the simple transition probabilities. Often we can, however, put up differential equations for the requested probabilities and solve these by some suitable method.

Of particular interest is the probability of staying uninterruptedly in the current state for a certain period of time,
\[
p_{jj}(t, u) = P[Z(\tau) = j, \tau \in (t, u) \mid Z(t) = j].
\]  
(7.23)
From here proceed as above, using $p_{jj}(u, u) = 1$, to obtain

$$p_{jj}(t, u) = e^{-\int_t^u \mu_j \, dt}.$$  \hfill (7.24)

**F. Backward and forward integral equations.** From the backward differential equations we obtain an equivalent set of integral equations as follows. Switch the first term on the right over to the left and, to obtain a complete differential there, multiply on both sides by the integrating factor $e^{-\int_t^u \mu_j \, dt}$:

$$dt \left( e^{\int_t^u \mu_j \, dt} p_{jk}(t, u) \right) = -e^{\int_t^u \mu_j \, dt} \sum_{g : g \neq j} \mu_{gj}(t) \, dt \, p_{gk}(t, u).$$

Now integrate over $(t, u]$ and use (7.18) to obtain

$$\delta_{jk} - e^{\int_t^u \mu_j \, dt} p_{jk}(t, u) = - \int_t^u e^{\int_t^{\tau} \mu_j \, d\tau} \sum_{g : g \neq j} \mu_{gj}(\tau) p_{gk}(\tau, u) \, d\tau.$$

Finally, carry the Kronecker delta over to the right, multiply by $-e^{-\int_t^u \mu_j \, dt}$, and use (7.24) to arrive at the backward integral equations

$$p_{jk}(t, u) = \int_t^u p_{jj}(t, \tau) \sum_{g : g \neq j} \mu_{gj}(\tau) p_{gk}(\tau, u) \, d\tau + \delta_{jk} p_{jj}(t, u).$$  \hfill (7.25)

In a similar manner we obtain the following set of forward integral equations from (7.20):

$$p_{ij}(s, t) = \delta_{ij} p_{ii}(s, t) + \sum_{g : g \neq j} \int_s^t p_{ij}(s, \tau) \mu_{gj}(\tau) p_{gj}(\tau, t) \, d\tau.$$  \hfill (7.26)

The integral equations could be put up directly upon summing the probabilities of disjoint elementary events that constitute the event in question. For (7.26) the argument goes as follows. The first term on the right accounts for the possibility of ending up in state $j$ without making any intermediate transitions, which is relevant only if $i = j$. The second term accounts for the possibility of ending up in state $j$ after having made intermediate transitions and is the sum, over all states $g \neq j$ and all small time intervals $(\tau, \tau + d\tau)$ in $(s, t)$, of the probability of arriving for the last time in state $j$ from state $g$ in the time interval $(\tau, \tau + d\tau)$. In a similar manner (7.25) is obtained upon splitting up by the direction and the time of the first departure, if any, from state $j$.

We now turn to some specializations of the model pertaining to insurance of persons.

### 7.3 Applications

**A. A single life with one cause of death.** The life length of a person is modelled as a positive random variable $T$ with survival function $F$. There are
two states, 'alive' and 'dead'. Labelling these by 0 and 1, respectively, the state process $Z$ is simply

$$Z(t) = 1_{\{T \leq t\}}, \quad t \in [0, n],$$

which counts the number of deaths by time $t \geq 0$. The process $Z$ is right-continuous and is obviously Markov since in state 0 the past is trivial, and in state 1 the future is trivial. The transition probabilities are

$$p_{00}(s, t) = \frac{\bar{F}(t)}{\bar{F}(s)}.$$

The Chapman-Kolmogorov equation reduces to the trivial

$$p_{00}(s, u) = p_{00}(s, t)p_{00}(t, u)$$

or $\frac{\bar{F}(u)}{\bar{F}(s)} = \{\bar{F}(t)/\bar{F}(s)\}\{\bar{F}(u)/\bar{F}(t)\}$. The only non-null intensity is $\mu_{01}(t) = \mu(t)$, and

$$p_{00}(t, u) = e^{-\int_t^u \mu}.$$  
(7.27)

The Kolmogorov differential equations reduce to just the definition of the intensity (write out the details).

The simple two state process with state 1 absorbing is outlined in Fig. 7.1

**B. A single life with $r$ causes of death.** In the previous paragraph it was, admittedly, the process set-up that needed the example and not the other way around. The process formulation shows it power when we turn to more complex situations. Fig. 7.2 outlines a first extension of the model in the previous paragraph, whereby the single absorbing state ("dead") is replaced by $r$ absorbing states representing different causes of death, e.g. "dead in accident", "dead from heart disease", etc. The index 0 in the intensities $\mu_{0j}$ is superfluous and has been dropped.
Relation (7.14) implies that the total mortality intensity is the sum of the intensities of death from different causes,

$$\mu(t) = \sum_{j=1}^{r} \mu_j(t).$$

(7.28)

For a person aged $t$ the probability of survival to $u$ is the well-known survival probability $p_{00}(t, u)$ given by (7.27), now with a nuanced explanation in the present enriched model. For instance, the G-M law in the simple mortality model may be motivated as resulting from two causes of death, one with intensity $\alpha$ independent of age (pure accident) and the other with intensity $\beta e^{\gamma t}$ (wear-out).

The probability of a $t$ years old dying from cause $j$ before age $u$ is

$$p_{0j}(t, u) = \int_{t}^{u} e^{-\int_{t}^{\tau} \mu_j(\tau) d\tau} d\tau.$$  

(7.29)

This follows from e.g. (7.25) upon noting that $p_{rr}(t, u) = 1$, but — being totally transparent — it can be put up directly.

Inspection of (7.28) – (7.29) gives rise to a comment. An increase of one mortality intensity $\mu_k$ results in a decrease of the survival probability (evidently) and also of the probabilities of death from every other cause $j \neq k$, hence (since the probabilities sum to 1) an increase of the probability of death from cause $k$ (also evident). Thus, the increased proportions of deaths from heart diseases and cancer in our times could be sufficiently explained by the fact that medical progress has practically eliminated mortality by lung inflammation, childbed fever, and a number of other diseases.

The above discussion supports the assertion that the intensities are basic entities. They are the pure expressions of the forces acting on the policy in each given state, and the transition probabilities are resultants of the interplay between these forces.
C. A model for disabilities, recoveries, and death. Fig. 7.3 outlines a model suitable for analyzing insurances with payments depending on the state of health of the insured, e.g. sickness insurance providing an annuity benefit during periods of disability or life insurance with premium waiver during disability. Many other problems fit into the same scheme by mere relabeling of the states. For instance, in connection with a pension insurance with additional benefits to the spouse, states 0 and 1 would be "unmarried" and "married", and in connection with unemployment insurance they would be "employed” and "unemployed”.

For a person who is active at time $s$ the Kolmogorov forward differential (7.20) equations are

$$\frac{\partial}{\partial t} p_{00}(s, t) = p_{01}(s, t) \rho(t) - p_{00}(s, t)(\mu(t) + \sigma(t)),$$  \hfil (7.30)

$$\frac{\partial}{\partial t} p_{01}(s, t) = p_{00}(s, t) \sigma(t) - p_{01}(s, t)(\nu(t) + \rho(t)).$$  \hfil (7.31)

(The probability $p_{02}(s, t)$ is determined by the other two.) The initial conditions (7.21) become

$$p_{00}(s, s) = 1,$$  \hfil (7.32)

$$p_{01}(s, s) = 0.$$  \hfil (7.33)

(For a person who is disabled at time $s$ the forward differential equations are the same, only with the first subscript 0 replaced by 1 in all the probabilities, and the side conditions are $p_{00}(s, s) = 0$, $p_{11}(s, s) = 1$.)

When the intensities are sufficiently simple functions, one may find explicit closed expressions for the transition probabilities. Work through the case with constant intensities.
CHAPTER 7. MARKOV CHAINS IN LIFE INSURANCE

7.4 Selection phenomena

A. Introductory remarks. The Markov model (like any other model) may be accused of being oversimplified. For instance, in the disability model it says that the prospects of survival of a disabled person are unaffected by information about the past such as the pattern of previous disabilities and recoveries and, in particular, the duration since the last onset of disability. One could imagine that there are several types or degrees of disability, some of them light, with rather standard mortality, and some severe with heavy excess mortality. In these circumstances information about the past may be relevant: if we get to know that the onset of disability incurred a long time ago, then it is likely that one of the light forms is in play; if it incurred yesterday, it may well be one of the severe forms by which a soon death is to be expected.

Now, the kind of heterogeneity effect mentioned here can be accommodated in the Markov framework simply by extending the state space, replacing the single state “disabled” by more states corresponding to different degrees of disability. From this Markov model we deduce the model of what we can observe as sketched in Fig. 7.3 upon letting the state “disabled” be the aggregate of the disability states in the basic model. What we end up with is typically no longer a Markov model.

Generally speaking, by variation of state space and intensities, the Markov set-up is capable of representing extremely complex phenomena. In the following we shall formalize these ideas with a particular view to selection phenomena often encountered in insurance. The ideas are to a great extent taken from [14].

B. Aggregating states of a Markov chain. Let $Z$ be a continuous time Markov chain as defined in Section 7.2. Let $\{Z_0, \ldots, Z_r\}$ be a partition of $Z$, that is, the $Z_g$ are disjoint and their union is $Z$. By convention, $0 \in Z_0$. Define a stochastic process $Z^\ast$ on the state space $Z^\ast = \{0, \ldots, r^\ast\}$ by

$$Z^\ast(t) = g \text{ iff } Z(t) \in Z_g.$$  

(7.34)

The interpretation is that we can observe the process $Z^\ast$ which represents summary information about some not fully observable Markov process $Z$.

Suppose the basic process is observed to be in state $i$ at time $s$. The subsequent development of $Z^\ast$ can be projected by conditional probabilities for the process $Z$. We have, for $s < t < u$,

$$\mathbb{P}[Z^\ast(u) = h \mid Z(s) = i, Z^\ast(t) = g] = \frac{1}{p_{iZ_{g}}(s, t)} \sum_{j \in Z_{g}} p_{ij}(s, t)p_{jZ_{h}}(t, u).$$

From this expression we obtain conditional transition intensities of the aggregate process:

$$\lim_{u \uparrow t} \frac{\mathbb{P}[Z^\ast(u) = h \mid Z(s) = i, Z^\ast(t) = g]}{u - t} = \frac{1}{p_{iZ_{g}}(s, t)} \sum_{j \in Z_{g}} p_{ij}(s, t)\mu_{jZ_{h}}(t).$$
We cannot speak of the intensities since they would in general be different if more information about $Z^*$ were conditioned on.

As an example, consider the aggregate of the states 0 and 1 in the disability model in Paragraph 7.3.C, $Z_0 = \{0, 1\}$, and put $Z_1 = \{2\}$. Thus we observe only whether the insured is alive or not. The process $Z^*$ is Markov, of course (recall the argument in Paragraph 7.3.A). The survival probability is

$$p_{00}^*(0, t) = p_{00}(0, t) + p_{01}(0, t),$$

and the mortality intensity at age $t$ is

$$\mu^*(t) = \frac{p_{00}(0, t)\mu(t) + p_{01}(0, t)\nu(t)}{p_{00}(0, t) + p_{01}(0, t)},$$

a weighted average of the mortality intensities of active and disabled, the weights being the probabilities of staying in the respective states.

C. Non-differential probabilities. Suppose the transition probabilities $p_j z_h(t, u)$ considered as functions of $j$ are constant on each $Z_g$, that is, there exist functions $p_{gh}^*(t, u)$ such that, for each $t < u$ and $g, h \in Z^*$,

$$p_j z_h(t, u) = p_{gh}^*(t, u), \forall j \in Z_g.$$  (7.35)

Then we shall say that the probabilities of transition between the subsets $Z_g$ are non-differential (within the individual subsets). The following result is evident on intuitive grounds, but never the less merits emphasis.

**Theorem 1.** If the transition probabilities of the process $Z$ between the subsets $Z_g$ are non-differential, then the process $Z^*$ defined by (7.34) is Markov with transition probabilities $p_{gh}^*(t, u)$ defined by (7.35). If $Z$ possesses intensities $\mu_{jk}$, then the process $Z^*$ has intensities $\mu_{gh}^*$ given by

$$\mu_{gh}^* = \mu_j z_h, j \in Z_g.$$  (7.36)

**Proof:** By (7.8), we must show that, for any event $A$ depending only on $\{Z^*(\tau)\}_{0 \leq \tau < t}$,

$$P[Z^*(u) = h \mid A, Z^*(t) = g] = p_{gh}^*(t, u).$$  (7.37)

Using first the fact that $[Z^*(t) = g] = \cup_{j \in Z_g}[Z(t) = j]$ is a union of disjoint events, then that $A \in H_{<t}$ and $Z$ is Markov, and finally assumption (7.35), we get

$$P[A, Z^*(t) = g, Z^*(u) = h]$$

$$= \sum_{j \in Z_g} P[A, Z(t) = j, Z(u) \in Z_h]$$

$$= \sum_{j \in Z_g} P[A, Z(t) = j] p_j z_h(t, u).$$

$$= P[A, X^*(t) = g] p_{gh}^*(t, u),$$

which is equivalent to (7.37). \qed
D. Non-differential mortality. Let state \( r \) be absorbing, representing death, and let \( \mathcal{H} = \{0, \ldots, r-1\} \) be the aggregate of states where the insured is alive. Assume that the mortality is non-differential, which means that all \( \mu_{jr}, j \in \mathcal{H} \), are identical and equal to \( \lambda \), say. Then, by Theorem 1, the survival probability is the same in all states \( j \in \mathcal{H} \):

\[
p_{jH}(t, u) = e^{-\int_t^u \lambda ds}.
\]

The conditional probability of staying in state \( k \in \mathcal{H} \) at time \( t \), given survival, is

\[
p_{jk|H}(t, u) = \frac{p_{jk}(t, u)}{p_{jH}(t, u)} = p_{jk}(t, u)e^{\int_t^u \lambda ds}.
\]

Inserting \( p_{jk}(t, u) = p_{jk|H}(t, u)e^{-\int_t^u \lambda ds} \) into (7.25), we get for each \( j, k \in \mathcal{H} \) that

\[
p_{jk|H}(t, u)e^{-\int_t^u \lambda ds} = \int_t^u e^{-\int_s^u \mu_{j,\mathcal{H}-\{j\}} \lambda ds} \sum_{g \in \mathcal{H}-\{j\}} \mu_{jg}(\tau)p_{gk|H}(\tau, u)e^{-\int_\tau^u \lambda ds} d\tau \\
+ \delta_{jk}e^{-\int_t^u \mu_{j,\mathcal{H}-\{j\}} \lambda ds}.
\]

Multiplying with \( e^{\int_t^u \lambda ds} \), we see that the conditional probabilities in (7.39) satisfy the integral equations (7.25) for the transition probabilities in the so-called partial model with state space \( \mathcal{H} \) and transition intensities \( \mu_{jk}, j, k \in \mathcal{H} \). Thus, to find the transition probabilities in the full model, work first in the simple partial model for the states as alive and multiply the partial probabilities obtained there with the survival probability.

7.5 The standard multi-state contract

A. The contractual payments. We refer to the insurance policy with development as described in Paragraph 7.1.A. Taking \( Z \) to be a stochastic process with right-continuous paths and at most a finite number of jumps, the same holds also for the associated indicator processes \( I_j \) and counting processes \( N_{jk} \) defined, respectively, by \( I_j(t) = 1_{\{Z(t) = j\}} \) (1 or 0 according as the policy is in the state \( j \) or not at time \( t \)) and \( N_{jk}(t) = \#\{\tau; Z(\tau-) = j, Z(\tau) = k, \tau \in (0, t]\} \) (the number of transitions from state \( j \) to state \( k \) \( (k \neq j) \) during the time interval \( (0, t]\) ). The indicator processes \( \{I_j(t)\}_{t \geq 0} \) and the counting processes \( \{N_{jk}(t)\}_{t \geq 0} \) are related by the fact that \( I_j \) increases/decreases (by 1) upon a transition into/out of state \( j \). Thus

\[
dI_j(t) = dN_{j}(t) - dN_{j}(t),
\]

where a dot in the place of a subscript signifies summation over that subscript, e.g. \( N_j = \sum_{k \neq j} N_{jk} \).

The policy is assumed to be of standard type, which means that the payment function representing contractual benefits less premiums is of the form (recall
the device (A.15))

\[ dB(t) = \sum_k I_k(t) dB_k(t) + \sum_{\ell, \ell \neq k} b_{k\ell}(t) dN_{k\ell}(t), \]  

(7.41)

where each \( B_k \), of form \( dB_k(t) = b_k(t) dt + B_k^0 \), is a deterministic payment function specifying payments due during sojourns in state \( k \) (a general life annuity), and each \( b_{k\ell} \) is a deterministic function specifying payments due upon transitions from state \( k \) to state \( \ell \) (a general life assurance). When different from 0, \( B_k^0(t) - B_k^0(t^-) \) is an endowment at time \( t \). The functions \( b_k \) and \( b_{k\ell} \) are assumed to be finite-valued and piecewise continuous. The set of discontinuity points of any of the annuity functions \( B_k \) is \( D = \{ t_0, t_1, \ldots, t_q \} \) (say).

Positive amounts represent benefits and negative amounts represent premiums. In practice premiums are only of annuity type. At times \( t \not\in [0, n] \) all payments are null.

B. Identities revisited. Here we make an intermission to make a comment that does not depend on the probability structure to be specified below. The identity (4.18) rests on the corresponding identity (4.17) between the present values. The latter is, in its turn, a special case of the identities put up in Section 2.1, from which many identities between present values in life insurance can be derived.

Suppose the investment portfolio of the insurance company bears interest with intensity \( r(t) \) at time \( t \). The following identity, which expresses life annuities by endowments and life assurances, is easily obtained upon integrating by parts, using (7.40):

\[
\int_t^u e^{-\int_\tau^\tau_0 r} I_j(\tau) dB_j(\tau) = e^{-\int_\tau^\tau_0 r} I_j(u) B_j(u) - e^{-\int_\tau^\tau_0 r} I_j(t) B_j(t) + \int_t^u e^{-\int_\tau^\tau_0 r} I_j(\tau) B_j(\tau) r(\tau) d\tau + \int_t^u e^{-\int_\tau^\tau_0 r} B_j(\tau) \cdot (dN_j(\tau) - N_j(\tau)).
\]

C. Expected present values and prospective reserves. At any time \( t \in [0, n] \), the present value of future benefits less premiums under the contract is

\[ V(t) = \int_t^u e^{-\int_\tau^\tau_0 r} dB(\tau). \]  

(7.42)

This is a liability for which the insurer is to provide a reserve, which by statute is the expected value. Suppose the policy is in state \( j \) at time \( t \). Then the conditional expected value of \( V(t) \) is

\[ V_j(t) = \int_t^u e^{-\int_\tau^\tau_0 r} \sum_k p_{jk}(t, \tau) \left( dB_k(\tau) + \sum_{\ell, \ell \neq k} b_{k\ell}(\tau) \cdot \mu_{k\ell}(\tau) \cdot d\tau \right). \]  

(7.43)
This follows by taking expectation under the integral in (7.42), inserting $dB(\tau)$ from (7.41), and using

$$E[I_k(\tau) | Z(t) = j] = p_{jk}(t, \tau),$$

$$E[N_{k\ell}(\tau) | Z(t) = j] = p_{jk}(t, \tau)\mu_{k\ell}(\tau) d\tau.$$ 

We expound the result as follows. With probability $p_{jk}(t, \tau)$ the policy stays in state $k$ at time $\tau$, and if this happens the life annuity provides the amount $dB_k(\tau)$ during a period of length $d\tau$ around $\tau$. Thus, the expected present value at time $t$ of this contingent payment is $p_{jk}(t, \tau)e^{-\int_0^\tau r} dB_k(\tau)$. With probability $p_{jk}(t, \tau)\mu_{k\ell}(\tau) d\tau$ the policy jumps from state $k$ to state $\ell$ during a period of length $d\tau$ around $\tau$, and if this happens the assurance provides the amount $b_{k\ell}(\tau)$. Thus, the expected present value at time $t$ of this contingent payment is $p_{jk}(t, \tau)\mu_{k\ell}(\tau) d\tau e^{-\int_0^\tau r} b_{k\ell}(\tau)$. Summing over all future times and types of payments, we find the total given by (7.43).

Let $0 \leq t < u < n$. Upon separating payments in $[t, u]$ and in $(u, n]$ on the right of (7.43), and using Chapman-Kolmogorov on the latter part, we obtain

$$V_j(t) = \int_t^u e^{-\int_0^\tau r} \sum_k p_{jk}(t, \tau) \left( dB_k(\tau) + \sum_{\ell,k \neq k} b_{k\ell}(\tau)\mu_{k\ell}(\tau) d\tau \right)$$
$$+ e^{-\int_0^\tau r} \sum_k p_{jk}(t, u) V_k(u). \quad (7.44)$$

This expression is also immediately obtained upon conditioning on the state of the policy at time $u$.

Throughout the term of the policy the insurance company must currently maintain a reserve to meet future net liabilities in respect of the contract. By statute, if the policy is in state $j$ at time $t$, then the company is to provide a reserve that is precisely $V_j(t)$. Accordingly, the functions $V_j$ are called the (state-wise) prospective reserves of the policy. One may say that the principle of equivalence has been carried over to time $t$, now requiring expected balance between the amount currently reserved and the discounted future liabilities, given the information currently available. (Only the present state of the policy is relevant due to the Markov property and the simple memoryless payments under the standard contract).

D. The backward (Thiele’s) differential equations. By letting $u$ approach $t$ in (7.44), we obtain a differential form that displays the dynamics of the reserves. In fact, we are going to derive a set of backward differential equations and, therefore, take the opportunity to apply the direct backward differential argument demonstrated and announced previously in Paragraph 7.2.E.

Thus, suppose the policy is in state $j$ at time $t \notin \mathcal{D}$. Conditioning on what happens in a small time interval $[t, t + dt]$ (not intersecting $\mathcal{D}$) we write

$$V_j(t) = b_j(t) dt + \sum_{k,k \neq j} \mu_{jk}(t) dt b_{jk}(t)$$
CHAPTER 7. MARKOV CHAINS IN LIFE INSURANCE

\[ + (1 - \mu_j(t) \, dt) e^{-r(t) \, dt} V_j(t + dt) + \sum_{k, k \neq j} \mu_{jk}(t) \, dt \, e^{-r(t) \, dt} V_k(t + dt). \]

Proceeding from here along the lines of the simple case in Section 4.4, we easily arrive at the backward or Thiele’s differential equations for the state-wise prospective reserves,

\[ \frac{d}{dt} V_j(t) = (r(t) + \mu_j(t)) V_j(t) - \sum_{k, k \neq j} \mu_{jk}(t) V_k(t) - b_j(t) - \sum_{k, k \neq j} b_{jk}(t) \mu_{jk}(t). \]  

(7.45)

The differential equations are valid in the open intervals \((t_{p-1}, t_p), p = 1, \ldots, q,\) and together with the conditions

\[ V_j(t_{p-1}) = (B_j(t_p) - B_j(t_{p-1})) + V_j(t_p), \quad p = 1, \ldots, q, j \in \mathbb{Z}, \]  

(7.46)

they determine the functions \(V_j\) uniquely.

A comment is in order on the differentiability of the \(V_j.\) At points of continuity of the functions \(b_j, b_{jk}, \mu_{jk},\) and \(r\) there is no problem since there the integrand on the right of (7.43) is continuous. At possible points of discontinuity of the integrand the derivative \(\frac{d}{dt} V_j\) does not exist. However, since such discontinuities are finite in number, they will not affect the integrations involved in numerical procedures. Thus we shall throughout allow ourselves to write the differential equations on the form (7.45) instead of the generally valid differential form obtained upon putting \(dV_j(t)\) on the left and multiplying with \(dt\) on the right.

E. Solving the differential equations. Only in rare cases of no practical interest is it possible to find closed form solutions to the differential equations. In practice one must resort to numerical methods to determine the prospective reserves. As a matter of experience a fourth order Runge-Kutta procedure works reliably in virtually all situations encountered in practice.

One solves the differential equations ‘from top down’. First solve (7.45) in the upper interval \((t_{n-1}, t_n)\) subject to (7.46), which specializes to \(V_j(t_{n-1}) = B_j(t_n) - B_j(t_{n-1})\) since \(V_j(t_n) = 0\) for all \(j\) by definition. Then go to the interval below and solve (7.45) subject to \(V_j(t_{q-1}) = (B_j(t_{q-1}) - B_j(t_{q-1})) + V_j(t_{q-1}),\) where \(V_j(t_{q-1})\) was determined in the first step. Proceed in this manner downwards.

It is realized that the Kolmogorov backward equations (7.17) are a special case of the Thiele equations (7.45); the transition probability \(p_{jk}(t, u)\) is just the prospective reserve in state \(j\) at time \(t\) for the simple contract with the only payment being a lump sum payment of 1 at time \(u\) if the policy is then in state \(k,\) and with no interest. Thus a numerical procedure for computation of prospective reserves can also be used for computation of the transition probabilities.
F. The equivalence principle. If the equivalence principle is invoked, one must require that
\[ V_0(0) = -B_0(0). \] (7.47)
This condition imposes a constraint on the contractual functions \( b_j, B_j, \) and \( b_{jk}, \) viz. on the premium level for given benefits and 'design' of the premium plan. It is of a different nature than the conditions (7.46), which follow by the very definition of prospective reserves (for given contractual functions).

G. Savings premium and risk premium. The equation (7.45) can be recast as
\[ -b_j(t) dt = dV_j(t) - r(t) dt V_j(t) + \sum_{k; k \neq j} R_{jk}(t) \mu_{jk}(t) dt. \] (7.48)
where
\[ R_{jk}(t) = b_{jk}(t) + V_k(t) - V_j(t). \] (7.49)
The quantity \( R_{jk}(t) \) is called the sum at risk associated with (a possible) transition from state \( j \) to state \( k \) at time \( t \) since, upon such a transition, the insurer must immediately pay out the sum insured and also provide the appropriate reserve in the new state, but he can cash the reserve in the old state. Thus, the last term in (7.48) is the expected net payout in connection with a possible transition out of the current state \( j \) in \((t, t + dt)\), and it is called the risk premium. The two first terms on the right of (7.48) constitute the savings premium in \((t, t + dt)\), called so because it is the amount that has to be provided to maintain the reserve in the current state; the increment of the reserve less the interest earned on it. On the left of (7.48) is the premium paid in \((t, t + dt)\), and so the relation shows how the premium decomposes in a savings part and a risk part. Although helpful as an interpretation, this consideration alone cannot carry the full understanding of the differential equation since (7.48) is valid also if \( b_j(t) \) is positive (a benefit) or 0.

H. Integral equations. In (7.45) let us switch the term \((r(t) + \mu_j(t)) V_j(t)\) appearing on the right of over to the left, and multiply the equation by \(e^{-\int_0^t (r + \mu_j) dt} \) to form a complete differential on the left:
\[ \frac{d}{dt} \left( e^{-\int_0^t (r + \mu_j) dt} V_j(t) \right) = \]
\[ -e^{-\int_0^t (r + \mu_j) dt} \left( \sum_{k; k \neq j} \mu_{jk}(t) V_k(t) + b_j(t) + \sum_{k; k \neq j} b_{jk}(t) \mu_{jk}(t) \right). \]
Now integrate over an interval \((t, u)\) containing no jumps \(B_j(\tau) - B_j(\tau^-)\) and, recalling that \(e^{-\int_t^\tau \mu_j} = p_{jj}(t, \tau)\), rearrange a bit to obtain the integral equation

\[
V_j(t) = \int_t^u p_{jj}(t, \tau) e^{-\int_t^\tau r} \left( b_j(\tau) + \sum_{k, k \neq j} \mu_{jk}(\tau)(b_{jk}(\tau) + V_k(\tau)) \right) d\tau \\
+ p_{jj}(t, u) e^{-\int_t^u r} V_j(u^-).
\] (7.50)

This result generalizes the backward integral equations for the transition probabilities (7.25) and, just as in that special case, also the expression on the right hand side of (7.50) is easy to interpret; it decomposes the future payments into those that fall due before and those that fall due after the time of the first transition out of the current state in the time interval \((t, u)\) or, if no transition takes place, those that fall due before and those that fall due after time \(u\).

We shall take a direct route to the integral equation (7.50) that actually is the rigorous version of the backward technique. Suppose that the policy is in state \(j\) at time \(t\). Let us apply the rule of iterated expectations to the expected value \(V_j(t)\), conditioning on whether a transition out of state \(j\) takes place within time \(u\) or not and, in case it does, also condition on the time and the direction of the first transition. We then get

\[
V_j(t) = \int_t^u p_{jj}(t, \tau) \sum_{k, k \neq j} \mu_{jk}(\tau) d\tau \left( \int_t^\tau e^{-\int_t^\tau r} b_j(s) ds + e^{-\int_t^\tau r}(b_{jk}(\tau) + V_k(\tau)) \right) \\
+ p_{jj}(t, u) \left( \int_t^u e^{-\int_t^s r} b_j(s) ds + e^{-\int_t^u r} V_j(u^-) \right). 
\] (7.51)

To see that this is the same as (7.50), we need only to observe that

\[
\int_t^u p_{jj}(t, \tau) \mu_j(\tau) \int_t^\tau e^{-\int_t^\tau r} b_j(s) ds d\tau \\
= \int_t^u \int_t^s \frac{d}{d\tau} (-p_{jj}(t, \tau)) d\tau e^{-\int_t^\tau r} b_j(s) ds \\
= -p_{jj}(t, u) \int_t^u e^{-\int_t^s r} b_j(s) ds + \int_t^u p_{jj}(t, s) e^{-\int_t^s r} b_j(s) ds.
\]

I. Uses of the differential equations. If the contractual functions do not depend on the reserves, the defining relation (7.43) give explicit expressions for the state-wise reserves and strictly speaking the differential equations (7.45) are not needed for constructive purposes. They are, however, computationally convenient since there are good methods for numerical solution of differential equations. They also serve to give insight into the dynamics of the policy.

The situation is entirely different if the contractual functions are allowed to depend on the reserves in some way or other. The most typical examples are repayment of a part of the reserve upon withdrawal (a state "withdrawn" must then be included in the state space \(Z\)) and expenses depending partly on the...
reserve. Also the primary insurance benefits may in some cases be specified as functions of the reserve. In such situations the differential equations are an indispensable tool in the construction of the reserves and determination of the equivalence premium. We shall provide an example in the next paragraph.

J. An example: Widow’s pension. A married couple buys a combined life insurance and widow’s pension policy specifying that premiums are to be paid with level intensity $c$ as long as both husband and wife are alive, pensions are to be paid with intensity $b$ as long as the wife is widowed, and a life assurance with sum $s$ is due immediately upon the death of the husband if the wife is already dead (a benefit to their dependents). The policy terminates at time $n$. The relevant Markov model is sketched Fig. 7.4. We assume that $r$ is constant.

The differential equations (7.45) now specialize to the following (we omit the trivial equation for $V_3(t) = 0$):

$$
\frac{d}{dt} V_0(t) = (r + \mu_{01}(t) + \mu_{02}(t)) V_0(t) - \mu_{01}(t) V_1(t) - \mu_{02}(t) V_2(t) + c, \quad (7.52)
$$

$$
\frac{d}{dt} V_1(t) = (r + \mu_{13}(t)) V_1(t) - b, \quad (7.53)
$$

$$
\frac{d}{dt} V_2(t) = (r + \mu_{23}(t)) V_2(t) - \mu_{23}(t) s. \quad (7.54)
$$

Consider a modified contract, by which 50% of the reserve is to be paid back to the husband in case he is widowed before time $n$, the philosophy being that couples receiving no pensions should have some of their savings back. Now
the differential equations are really needed. Under the modified contract the
equations above remain unchanged except that the term \(0.5V_0(t)\mu_{02}(t)\) must be
subtracted on the right of (7.52), which then changes to
\[
\frac{d}{dt} V_0(t) = (r + \mu_{01}(t) + 0.5\mu_{02}(t)) V_0(t) + c
- \mu_{01}(t)V_1(t) - \mu_{02}(t)V_2(t),
\]
(7.55)
Together with the conditions \(V_j(n) = 0, j = 0, 1, 2\), these equations are easily
solved.

As a second case the widow’s pension shall be analysed in the presence of
administration expenses that depend partly on the reserve. Consider again
the policy terms described in the introduction of this paragraph, but assume
that administration expenses incur with an intensity that is \(a\) times the current
reserve throughout the entire period \([0, n]\).

The differential equations for the reserves remain as in (7.52)–(7.54), except
that for each \(j\) the term \(aV_j(t)\) is to be subtracted on the right of the differential
equation for \(V_j\). Thus, the administration costs related to the reserve has the
same effect as a decrease of the interest intensity \(r\) by \(a\).

7.6 Select mortality revisited

A. A simple Markov chain model. Referring to Section 3.4 we shall present
a simple Markov model that offers an explanation of the selection phenomenon.

The Markov model sketched in Fig. 7.5 is designed for studies of selection
effects due to underwriting standards. The population is grouped into four cat-
egories or states by the criteria insurable/uninsurable and insured/not insured.
In addition there is a category comprising the dead. It is assumed that each
person enters state 0 as new-born and thereafter changes states in accordance
with a time-continuous Markov chain with age-dependent forces of transition as
indicated in the figure. Uninsurability occurs upon onset of disability or serious
illness or other intervening circumstances that entail excess mortality. Hence it
is assumed that
\[
\lambda_x > \kappa_x; \quad x > 0.
\]
(7.56)

Let \(Z(x)\) be the state at age \(x\) for a randomly chosen new-born, and denote
the transition probabilities of the Markov process \(\{Z(x); x > 0\}\). The following
formulas can be put up directly:
\[
p_{11}(x, x + t) = \exp\{-\int_x^{x+t} (\sigma + \kappa)\}, \tag{7.57}
\]
\[
p_{12}(x, x + t) = \int_x^{x+t} \exp\{-\int_x^u (\sigma + \kappa)\} \sigma_u \exp\{-\int_u^{x+t} \lambda\} du, \tag{7.58}
\]
\[
p_{00}(0, x) = \exp\{-\int_0^{x} (\sigma + \kappa + \rho)\}. \tag{7.59}
\]
CHAPTER 7. MARKOV CHAINS IN LIFE INSURANCE

Figure 7.5: A Markov model describing occurrences of uninsurability, purchase of life insurance, and death.
B. Select mortality among insured lives. The insured lives are in either state 1 or state 2. Those who are in state 2 reached to buy insurance before they turned uninsurable. However, the insurance company does not observe transitions from state 1 to state 2; the only available information are $x$ and $x + t$. Thus, the relevant survival function is

$$\mathbb{P}[x] = p_{11}(x, x + t) + p_{12}(x, x + t),$$

(7.60)

the probability that a person who entered state 1 at age $x$, will attain age $x + t$. The symbol on the left of (7.60) is chosen in accordance with standard actuarial notation, see Sections 3.3 – 3.4.

The force of mortality corresponding to (7.60) is

$$\mu[x + t] = \kappa y + \lambda y - \kappa y,$$

(7.61)

(Self-evident by conditioning on $Z(x)$.) In general, the expression on the right of (7.61) depends effectively on both $x$ and $t$, that is, mortality is select.

We can now actually establish that under the present model the select mortality intensity behaves as stated in Paragraph 3.4.C. It is suitable in the following to fix $x + t = y$, say, as we are interested in how the mortality at a certain age depends on the age of entry.

C. The select force of mortality is a decreasing function of the age at entry. Formula (7.61) can be recast as

$$\mu[x + y - x] = \kappa y + \zeta(x, y)(\lambda y - \kappa y),$$

(7.62)

where

$$\zeta(x, y) = \frac{p_{12}(x, y)}{p_{11}(x, y) + p_{12}(x, y)}$$

(7.63)

$$= \frac{1}{1 + p_{11}(x, y)/p_{12}(x, y)}.$$ 

(7.64)

We easily find that

$$p_{12}(x, y)/p_{11}(x, y) = \int_x^y \sigma_u \exp\{ \int_u^y (\sigma + \kappa - \lambda) \} \, du,$$

which is a decreasing function of $x$. It follows that $\mu[x + y - x]$ is a decreasing function of $x$ as asserted in the heading of this paragraph.

The explanation is simple. Formula (7.61) expresses $\mu[x + y - x]$ as a weighted average of $\kappa y$ and $\lambda y$, the weights being (of course) the conditional probabilities of being insurable and uninsurable, respectively. The weight attached to $\lambda y$, the larger of the two rates, decreases as $x$ increases. Or, put in terms of everyday speech: in a body of insured lives of the same age $x$ and duration $t = y - x$, some will have turned uninsurable in the period since entry; the longer the duration, the larger the proportion of uninsurable lives. In particular, those who have just entered, are known to be insurable, that is, $\mu[x] = \kappa x$. 

CHAPTER 7. MARKOV CHAINS IN LIFE INSURANCE

D. Comparison with the mortality in the population. Let $\mu_x$ denote the force of mortality of a randomly chosen life of age $x$ from the population.

A formula for $\mu_x$ is easily obtained starting from the survival function $\bar{p}_0 = \sum_{i=0}^{\infty} p_{0i}(0, x)$. It can, however, also be picked directly from the results of the previous paragraph by noting that the pattern of mortality must be the same in the population as among lives insured as newly-born, i.e. $\mu_y = \mu_x + y \bar{p}_0$. Then, since $\mu_x + y - x$ is a decreasing function of $x$ and $\mu_y$ corresponds to $x = 0$, it follows that $\mu_y > \mu_x + y - x$ for all $x < y$.

Again the explanation is trivial; due to the underwriting standards, the proportion of uninsurable lives will be less among insured people than in the population as a whole.

7.7 Higher order moments of present values

A. Differential equations for moments of present values. Our framework is the Markov model and the standard insurance contract. The set of time points with possible lump sum annuity payments is $D = \{t_0, t_1, \ldots, t_m\}$ (with $t_0 = 0$ and $t_m = n$).

Denote by $V(t, u)$ the present value at time $t$ of the payments under the contract during the time interval $(t, u]$ and abbreviate $V(t) = V(t, n)$ (the present value at time $t$ of all future payments). We want to determine higher order moments $V^{(q)}(t)$.

For $q = 1, 2, \ldots$

Theorem 2. The functions $V^{(q)}(t)$ are determined by the differential equations

$$\frac{d}{dt} V^{(q)}(t) = (qr(t) + \mu_j(t))V^{(q)}(t) - qB_j(t) V^{(q-1)}(t)$$

$$- \sum_{k \neq j} \mu_{jk}(t) \sum_{p=0}^{q-1} \binom{q}{p} (B_{jk}(t) - B_j(t-))^{q-p} V^{(q-p)}(t),$$

valid on $(0, n) \setminus D$ and subject to the conditions

$$V_j^{(q)}(t) = \sum_{p=0}^{q} \binom{q}{p} (B_j(t) - B_j(t-))^p V_j^{(q-p)}(t),$$

$t \in D$. □

Proof: Obviously, for $t < u < n$,

$$V(t) = V(t, u) + e^{-\int_t^u r(t')} V(u),$$
For any \( q = 1, 2, \ldots \) we have by the binomial formula

\[
V^q(t) = \sum_{p=0}^{q} \binom{q}{p} V(t, u)^p \left( e^{-\int_t^u r V(u) \, du} \right)^{q-p}.
\] (7.67)

Consider first a small time interval \((t, t + dt)\) without any lump sum annuity payment. Putting \( u = t + dt \) in (7.67) and taking conditional expectation, given \( Z(t) = j \), we get

\[
V_j^q(t) = \sum_{p=0}^{q} \binom{q}{p} \mathbb{E} \left[ V(t, t + dt)^p \left( e^{-\int_t^u r V(u) \, du} \right)^{q-p} \bigg| Z(t) = j \right].
\] (7.68)

By use of iterated expectations, conditioning on what happens in the small interval \((t, t + dt)\), the \( p \)-th term on the right of (7.68) becomes

\[
\binom{q}{p} (1 - \mu_j(t) \, dt) (b_j(t) \, dt)^p e^{-(q-p) \, r(t) \, dt} V_j^{(q-p)}(t + dt)
\] (7.69)

\[
+ \binom{q}{p} \sum_{k, k \neq j} \mu_{jk}(t) \, dt (b_j(t) \, dt + b_{jk}(t))^p e^{-(q-p) \, r(t) \, dt} V_{k}^{(q-p)}(t + dt).
\] (7.70)

Let us identify the significant parts of this expression, disregarding terms of order \( o(dt) \). First look at (7.69); for \( p = 0 \) it is

\[
(1 - \mu_j(t) \, dt) e^{-q \, r(t) \, dt} V_j^{(q)}(t + dt),
\]

for \( p = 1 \) it is

\[
q b_j(t) \, dt e^{-(q-1) \, r(t) \, dt} V_j^{(q-1)}(t + dt),
\]

and for \( p > 1 \) is \( o(dt) \). Next look at (7.70); the factor

\[
dt (b_j(t) \, dt + b_{jk}(t))^p = dt \sum_{r=0}^{p} \binom{p}{r} (b_j(t) \, dt)^r (b_{jk}(t))^{p-r}
\]

reduces to \( dt (b_{jk}(t))^p \) so that (7.70) reduces to

\[
\binom{q}{p} \sum_{k, k \neq j} \mu_{jk}(t) \, dt (b_{jk}(t))^p e^{-(q-p) \, r(t) \, dt} V_k^{(q-p)}(t + dt).
\]

Thus, we gather

\[
V_j^{(q)}(t) = (1 - \mu_j(t) \, dt) e^{-q \, r(t) \, dt} V_j^{(q)}(t + dt)
\]

\[
+ q b_j(t) \, dt e^{-(q-1) \, r(t) \, dt} V_j^{(q-1)}(t + dt)
\]

\[
+ \sum_{p=0}^{q} \binom{q}{p} \sum_{k, k \neq j} \mu_{jk}(t) \, dt (b_{jk}(t))^p e^{-(q-p) \, r(t) \, dt} V_k^{(q-p)}(t + dt).
\]
Now subtract $V_j^{(q)}(t + dt)$ on both sides, divide by $dt$, let $dt$ tend to 0, and use $\lim_{t \to 0} (e^{-qr(t)dt} - 1)/dt = -qr(t)$ to obtain the differential equation (7.65).

The condition (7.65) follows easily by putting $t - dt$ and $t$ in the roles of $t$ and $u$ in (7.67) and letting $dt$ tend to 0. \(\square\)

A rigorous proof is given in [22].

Central moments are easier to interpret and therefore more useful than the non-central moments. Letting $m_j^{(q)}$ denote the $q$-th central moment corresponding to the non-central $V_j^{(q)}$, we have

\[
m_j^{(1)}(t) = V_j^{(1)}(t),
\]

\[
m_j^{(q)}(t) = \sum_{p=0}^{q} (-1)^{q-p} \binom{q}{p} V_j^{(p)}(t) \left(V_j^{(1)}(t)\right)^{q-p}.
\]

**B. Computations.** The computational procedure goes by the program ‘proresin.pas’ as follows. First solve the differential equations in the upper interval $(t_{m-1}, n)$, where the side conditions (7.65) are just

\[
V_j^{(q)}(n-) = (B_j(n) - B_j(n-))^q
\]

since $V_j^{(q)}(n) = \delta_{q0}$ (the Kronecker delta). Then, if $m > 1$, solve the differential equations in the interval $(t_{m-2}, t_{m-1})$ subject to (7.65) with $t = t_{m-1}$, and proceed in this manner downwards.

**C. Numerical examples.** We shall calculate the first three moments for some standard forms of insurance related to the ‘disability model’ in Paragraph 7.3.C. We assume that the interest rate is constant and 4.5% per year, $r = \ln(1.045) = 0.044017$, and that the intensities of transitions between the states depend only on the age $x$ of the insured and are

\[
\begin{align*}
\mu_x &= \nu_x = 0.0005 + 0.000075858 \cdot 10^{0.038x}, \\
\sigma_x &= 0.0004 + 0.0000034674 \cdot 10^{0.062}, \\
\rho_x &= 0.005.
\end{align*}
\]

The intensities $\mu$, $\nu$, and $\sigma$ are those specified in the G82M technical basis. (That basis does not allow for recoveries and uses $\rho = 0$).

Consider a male insured at age 30 for a period of 30 years, hence use $\mu_{02}(t) = \mu_{12}(t) = \mu_{30+0}, \mu_{10}(t) = \sigma_{30+0}, \mu_{10}(t) = \rho_{30+0}, 0 < t < 30 \ (= n)$. The central moments $m_j^{(q)}$ defined in (7.71) \(\sim\) (7.72) have been computed for the states 0 and 1 (state 2 is uninteresting) at times $t = 0, 6, 12, 18, 24$, and are shown – in Table 7.1 for a term insurance with sum 1 \(\ (= b_{02} = b_{12})\).
in Table 7.2 for an annuity payable in active state with level intensity 1 ($= b_0$);
in Table 7.3 for an annuity payable in disabled state with level intensity 1 ($= b_1$);
in Table 7.4 for a combined policy providing a term insurance with sum 1 ($= b_{02} = b_{12}$) and a disability annuity with level intensity 0.5 ($= b_1$) against level net premium 0.013108 ($= -b_0$) payable in active state.
You should try to interpret the results.

\textbf{D. Solvency margins in life insurance – an illustration.} Let $Y$ be the present value of all future net liabilities in respect of an insurance portfolio. Denote the $q$-th central moment of $Y$ by $m^{(q)}$. The so-called normal power approximation of the upper $\varepsilon$-fractile of the distribution of $Y$, which we denote by $y_{1-\varepsilon}$, is based on the first three moments and is

$$y_{1-\varepsilon} \approx m^{(1)} + c_{1-\varepsilon} \sqrt{m^{(2)}} + \frac{c_{1-\varepsilon}^2 - 1}{6} \frac{m^{(3)}}{m^{(2)}},$$

where $c_{1-\varepsilon}$ is the upper $\varepsilon$-fractile of the standard normal distribution. Adopting the so-called break-up criterion in solvency control, $y_{1-\varepsilon}$ can be taken as a
### Table 7.1: Moments for a life assurance with sum 1

<table>
<thead>
<tr>
<th>Time</th>
<th>0</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_t^{(1)}$</td>
<td>0.0683</td>
<td>0.0771</td>
<td>0.0828</td>
<td>0.0801</td>
<td>0.0592</td>
<td>0</td>
</tr>
<tr>
<td>$m_t^{(2)}$</td>
<td>0.0300</td>
<td>0.0389</td>
<td>0.0484</td>
<td>0.0549</td>
<td>0.0484</td>
<td>0</td>
</tr>
<tr>
<td>$m_t^{(3)}$</td>
<td>0.0139</td>
<td>0.0191</td>
<td>0.0262</td>
<td>0.0343</td>
<td>0.0369</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 7.2: Moments for an annuity of 1 per year while active:

<table>
<thead>
<tr>
<th>Time</th>
<th>0</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_t^{(1)}$ : 15.763</td>
<td>13.921</td>
<td>11.606</td>
<td>8.698</td>
<td>4.995</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(2)}$ : 5.885</td>
<td>5.665</td>
<td>4.740</td>
<td>2.950</td>
<td>0.833</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(3)}$ : 7.795</td>
<td>5.372</td>
<td>3.104</td>
<td>1.290</td>
<td>0.234</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(4)}$ : -51.550</td>
<td>-44.570</td>
<td>-32.020</td>
<td>-15.650</td>
<td>-2.737</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(5)}$ : 78.888</td>
<td>49.950</td>
<td>25.099</td>
<td>8.143</td>
<td>0.876</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

### Table 7.3: Moments for an annuity of 1 per year while disabled:

<table>
<thead>
<tr>
<th>Time</th>
<th>0</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_t^{(1)}$ : 0.277</td>
<td>0.293</td>
<td>0.289</td>
<td>0.239</td>
<td>0.119</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(2)}$ : 15.176</td>
<td>13.566</td>
<td>11.464</td>
<td>8.708</td>
<td>5.044</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(3)}$ : 1.750</td>
<td>1.791</td>
<td>1.646</td>
<td>1.147</td>
<td>0.364</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(4)}$ : 11.502</td>
<td>8.987</td>
<td>6.111</td>
<td>3.107</td>
<td>0.716</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(5)}$ : 15.960</td>
<td>14.835</td>
<td>11.929</td>
<td>6.601</td>
<td>1.277</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$m_t^{(6)}$ : -101.500</td>
<td>-71.990</td>
<td>-42.500</td>
<td>-17.160</td>
<td>-2.452</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Table 7.4: Moments for a life assurance of 1 plus a disability annuity of 0.5 per year against net premium of 0.013108 per year while active:

<table>
<thead>
<tr>
<th>Time $t$</th>
<th>0</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_t^{(1)0}$</td>
<td>0.0000</td>
<td>0.0410</td>
<td>0.0751</td>
<td>0.0858</td>
<td>0.0533</td>
<td>0</td>
</tr>
<tr>
<td>$m_t^{(1)1}$</td>
<td>7.6451</td>
<td>6.8519</td>
<td>5.8091</td>
<td>4.4312</td>
<td>2.5803</td>
<td>0</td>
</tr>
<tr>
<td>$m_t^{(2)0}$</td>
<td>0.4869</td>
<td>0.5046</td>
<td>0.4746</td>
<td>0.3514</td>
<td>0.1430</td>
<td>0</td>
</tr>
<tr>
<td>$m_t^{(2)1}$</td>
<td>2.7010</td>
<td>2.0164</td>
<td>1.2764</td>
<td>0.5704</td>
<td>0.0974</td>
<td>0</td>
</tr>
<tr>
<td>$m_t^{(3)0}$</td>
<td>2.1047</td>
<td>1.9440</td>
<td>1.5563</td>
<td>0.8686</td>
<td>0.1956</td>
<td>0</td>
</tr>
<tr>
<td>$m_t^{(3)1}$</td>
<td>−12.1200</td>
<td>−8.1340</td>
<td>−4.3960</td>
<td>−1.5100</td>
<td>−0.1430</td>
<td>0</td>
</tr>
</tbody>
</table>

minimum requirement on the technical reserve at the time of consideration. It decomposes into the premium reserve, $m^{(1)}$, and what can be termed the fluctuation reserve, $y_1 - \epsilon - m^{(1)}$. A possible measure of the riskiness of the portfolio is the ratio $R = (y_1 - \epsilon - m^{(1)}) / P$, where $P$ is some suitable measure of the size of the portfolio at the time of consideration. By way of illustration, consider a portfolio of $N$ independent policies, all identical to the one described in connection with Table 7.4 and issued at the same time. Taking as $P$ the total premium income per year, the value of $R$ at the time of issue is 48.61 for $N = 10$, 12.00 for $N = 100$, 3.46 for $N = 1000$, 1.06 for $N = 10000$, and 0.332 for $N = 100000$. 
Chapter 8

Safety loadings and bonus

8.1 General considerations

A. Bonus – what it is. The word bonus is Latin and means 'good'. In insurance terminology it denotes various forms of repayments to the policyholders of that part of the company’s surplus that stems from good performance of the insurance portfolio, a sub-portfolio, or the individual policy. We shall here concentrate on the special form it takes in traditional life insurance.

The issue of bonus presents itself in connection with every standard life insurance contract, characteristic of which is its specification of nominal contingent payments that are binding to both parties throughout the term of the contract. All contracts discussed so far are of this type, and a concrete example is the combined policy described in 7.4: upon inception of the contract the parties agree on a death benefit of 1 and a disability benefit of 0.5 per year against a level premium of 0.013108 per year, regardless of future developments of the intensities of mortality, disability, and interest. Now, life insurance policies like this one are typically long term contracts, with time horizons wide enough to capture significant variations in intensities, expenses, and other relevant economic-demographic conditions. The uncertain development of such conditions subjects every supplier of standard insurance products to a risk that is non-diversifiable, that is, independent of the size of the portfolio; an adverse development can not be countered by raising premiums or reducing benefits, and also not by cancelling contracts (the right of withdrawal remains one-sidedly with the insured). The only way the insurer can safeguard against this kind of risk is to build into the contractual premium a safety loading that makes it cover, on the average in the portfolio, the contractual benefits under any likely economic-demographic development. Such a safety loading will typically create a systematic surplus, which by statute is the property of the insured and has to be repaid in the form of bonus.
B. Sketch of the usual technique. The approach commonly used in practice is the following. At the outset the contractual benefits are valued, and the premium is set accordingly, on a first order (technical) basis, which is a set of hypothetical assumptions about interest, intensities of transition between policy-states, costs, and possibly other relevant technical elements. The first order model is a means of prudent calculation of premiums and reserves, and its elements are therefore placed to the safe side in a sense that will be made precise later. As time passes reality reveals true elements that ultimately set the realistic scenario for the entire term of the policy and constitute what is called the second order (experience) basis. Upon comparing elements of first and second order, one can identify the safety loadings built into those of first order and design schemes for repayment of the systematic surplus they have created. We will now make these things precise.

To save notation, we disregard administration expenses for the time being and discuss them separately in Section 8.7 below.

8.2 First and second order bases

A. The second order model. The policy-state process $Z$ is assumed to be a time-continuous Markov chain as described in Section 7.2. In the present context we need to equip the indicator processes and counting processes related to the process $Z$ with a topscript, calling them $I^Z_j$ and $N^Z_{jk}$. The probability measure and expectation operator induced by the transition intensities are denoted by $\mathbb{P}$ and $\mathbb{E}$, respectively.

The investment portfolio of the insurance company bears interest with intensity $r(t)$ at time $t$.

The intensities $r$ and $\mu_{jk}$ constitute the experience basis, also called the second order basis, representing the true mechanisms governing the insurance business. At any time its past history is known, whereas its future is unknown.

We extend the set-up by viewing the second order basis as stochastic, whereby the uncertainty associated with it becomes quantifiable in probabilistic terms. In particular, prediction of its future development becomes a matter of model-based forecasting. Thus, let us consider the set-up above as the conditional model, given the second order basis, and place a distribution on the latter, whereby $r$ and the $\mu_{jk}$ become stochastic processes. Let $G_t$ denote their complete history up to, and including, time $t$ and, accordingly, let $\mathbb{E}[\cdot|G_t]$ denote conditional expectation, given this information.

For the time being we will work only in the conditional model and need not specify any particular marginal distribution of the second order elements.

B. The first order model. We let the first order model be of the same type as the conditional model of second order. Thus, the first order basis is viewed as deterministic, and we denote its elements by $r^*$ and $\mu^*_{jk}$ and the corresponding probability measure and expectation operator by $\mathbb{P}^*$ and $\mathbb{E}^*$, respectively. The first order basis represents a prudent initial assessment of the development of
the second order basis, and its elements are placed on the safe side in a sense that will be made precise later.

By statute, the insurer must currently provide a reserve to meet future liabilities in respect of the contract, and these liabilities are to be valuated on the first order basis. The first order reserve at time $t$, given that the policy is then in state $j$, is

$$V^*_j(t) = E^* \left[ \int_t^\infty e^{-\int_\tau^t r^*} dB(\tau) \, \Big| \, Z(t) = j \right]$$

$$= \int_t^\infty e^{-\int_\tau^t r^*} \sum g p^*_j(t, \tau) \left( dB_g(\tau) + \sum_{k, h \neq g} b_{gh}(\tau) \mu^*_{gh}(\tau) d\tau \right) . \quad (8.1)$$

We need Thiele’s differential equations

$$dV^*_j(t) = r^*(t)V^*_j(t) dt - dB_j(t) - \sum_{k, k \neq j} R^*_{jk}(t) \mu^*_{jk}(t) dt , \quad (8.2)$$

where

$$R^*_{jk}(t) = b_{jk}(t) + V^*_j(t) - V^*_k(t) \quad (8.3)$$

is the sum at risk associated with a possible transition from state $j$ to state $k$ at time $t$.

The premiums are based on the principle of equivalence exercised on the first order valuation basis,

$$E^* \left[ \int_0^t e^{-\int_\tau^t r^*} dB(\tau) \right] = 0 , \quad (8.4)$$

or, equivalently,

$$V^*_0(0) = -\Delta B_0(0) . \quad (8.5)$$

### 8.3 The technical surplus and how it emerges

**A. Definition of the mean portfolio surplus.** With premiums determined by the principle of equivalence (8.4) based on prudent first order assumptions, the portfolio will create a systematic technical surplus if everything goes well. Quite naturally, the surplus is some average of past net incomes valuated on the factual second order basis less future net outgoes valuated on the conservative first order basis. The portfolio-wide mean surplus thus construed is

$$S(t) = E \left[ \int_0^t e^{\int_\tau^t r^*} d(-B)(\tau) \, \big| \, \mathcal{G}_t \right] - \sum_j p_{0j}(0, t) V^*_j(t)$$

$$= -e^{\int_0^t r^*} \int_0^t e^{-\int_\tau^t r^*} \sum_j p_{0j}(0, \tau) \left( dB_j(\tau) + \sum_{k, k \neq j} b_{jk}(\tau) \mu_{jk}(\tau) d\tau \right)$$

$$- \sum_j p_{0j}(0, t) V^*_j(t) . \quad (8.6)$$
CHAPTER 8. SAFETY LOADINGS AND BONUS

The definition conforms with basic principles of insurance accountancy; at any time the balance is the difference between, on the debit, the factual income in the past and, on the credit, the reserve that by statute is to be provided in respect of future liabilities. In particular, due to (8.5),

$$S(0) = 0$$  \hspace{1cm} (8.7)

and, due to $V_j^*(n) = 0$,

$$S(n) = E \left[ \int_0^n e^{\int_\tau^r d(-B)(\tau)} \left| G_n \right| \right],$$  \hspace{1cm} (8.8)

as it ought to be.

Note that the expression in (8.6) involves only the past history of the second order basis, which is currently known.

B. The contributions to the surplus. Differentiating (8.6), applying the Kolmogorov forward equation (7.20) and the Thiele backward equation (8.2) to the last term on the right, leads to

$$dS(t) = -e^{\int_0^r r(t) dt} \int_0^t e^{-\int_\tau^r \{r(t) - r^*(t)\} d\tau} \left| G_n \right| d\tau,$$

with

$$c_j(t) = \{r(t) - r^*(t)\} V_j^*(t) + \sum_{k \neq j} R_{jk}^*(t) \{ \mu_{jk}(t) - \mu_{jk}(t) \}.$$  \hspace{1cm} (8.9)

Finally, integrating up and using (8.7), we arrive at

$$S(t) = \int_0^t e^{\int_\tau^r \sum_j p_{0j}(0, \tau) c_j(\tau) d\tau},$$  \hspace{1cm} (8.10)
which expresses the technical surplus at any time as the sum of past contributions compounded with second order interest.

One may arrive at the definition of the contributions \((8.9)\) by another route, starting from the individual surplus defined, quite naturally, as

\[
S_{\text{ind}}(t) = e^{\int_0^t r} \int_{0^-}^t e^{-\int_0^{\tau} r} d( - B)(\tau) - \sum_j I_j^r(t) V_j^*(t). 
\]  

(8.11)

Upon differentiating this expression, and proceeding along the same lines as above, one finds that \(S_{\text{ind}}(t)\) consists of a purely erratic term and a systematic term. The latter is 

\[
\int_{0^-}^t e^{\int_0^{\tau} r} \sum_j I_j^r(\tau)c_j(\tau) d\tau, 
\]

which is the individual counterpart of \((8.10)\), showing how the contributions emerge at the level of the individual policy. They form a random payment function \(C\) defined by

\[
dC(t) = \sum_j I_j^r(t) c_j(t) dt .
\]

(8.12)

With this definition, we can recast \((8.10)\) as

\[
S(t) = \mathbb{E} \left[ \int_0^t e^{\int_0^{\tau} r} dC(\tau) \bigg| \mathcal{G}_0 \right].
\]

(8.13)

C. Safety margins. The expression on the right of \((8.9)\) displays how the contributions arise from safety margins in the first order force of interest (the first term) and in the transition intensities (the second term). The purpose of the first order basis is to create a non-negative technical surplus. This is certainly fulfilled if

\[
r(t) \geq r^*(t) \quad \text{(8.14)}
\]

(assuming that all \(V_j^*(t)\) are non-negative as they should be) and

\[
\text{sign} \{ \mu_{jk}^*(t) - \mu_{jk}(t) \} = \text{sign} R_{jk}^*(t) .
\]

(8.15)

8.4 Dividends and bonus

A. The dividend process. Legislation lays down that the technical surplus belongs to the insured and has to be repaid in its entirety. Therefore, to the contractual payments \(B\) there must be added dividends, henceforth denoted by \(D\). The dividends are currently adapted to the development of the second order basis and, as explained in Paragraph 8.1.A, they can not be negative. The purpose of the dividends is to establish, ultimately, equivalence on the true second order basis:

\[
\mathbb{E} \left[ \int_0^n e^{-\int_0^{\tau} r} d\{B + D\}(\tau) \bigg| \mathcal{G}_n \right] = 0 .
\]

(8.16)
We can state (8.16) equivalently as
\[
\mathbb{E} \left[ \int_0^n e^{\int_0^\tau r} d[B + D](\tau) \bigg| G_n \right] = 0.
\] (8.17)

The value at time \( t \) of past individual contributions less dividends, compounded with interest, is
\[
U^d(t) = \int_0^t e^{\int_0^\tau r} d[C - D](\tau).
\] (8.18)

This amount is an outstanding account of the insured against the insurer, and we shall call it the dividend reserve at time \( t \).

By virtue of (8.8) and (8.13) we can recast the equivalence requirement (8.17) in the appealing form
\[
\mathbb{E}[U^d(n) | G_n] = 0.
\] (8.19)

From a solvency point of view it would make sense to strengthen (8.19) by requiring that compounded dividends must never exceed compounded contributions:
\[
\mathbb{E}[U^d(t) | G_t] \geq 0,
\] (8.20)

\( t \in [0, n] \). At this point some explanation is in order. Although the ultimate balance requirement is enforced by law, the dividends do not represent a contractual obligation on the part of the insurer; the dividends must be adapted to the second order development up to time \( n \) and can, therefore, not be stipulated in the terms of the contract at time 0. On the other hand, at any time, dividends allotted in the past have irrevocably been credited to the insured’s account. These regulatory facts are reflected in (8.20).

If we adopt the view that “the technical surplus belongs to those who created it”, we should sharpen (8.19) by imposing the stronger requirement
\[
U^d(n) = 0.
\] (8.21)

This means that no transfer of redistributions across policies is allowed. The solvency requirement conforming with this point of view, and sharpening (8.20), is
\[
U^d(t) \geq 0,
\] (8.22)

\( t \in [0, n] \).

The constraints imposed on \( D \) in this paragraph are of a general nature and leave a certain latitude for various designs of dividend schemes. We shall list some possibilities motivated by practice.

**B. Special dividend schemes.** The so-called contribution scheme is defined by \( D = C \), that is, all contributions are currently and immediately credited to the account of the insured. No dividend reserve will accrue and, consequently, the only instrument on the part of the insurer in case of adverse second order
experience is to cease crediting dividends. In some countries the contribution principle is enforced by law. This means that insurers are compelled to operate with minimal protection against adverse second order developments.

By *terminal dividend* is meant that all contributions are currently invested and their compounded total is credited to the insured as a lump sum dividend payment only upon the termination of the contract at some time $T$ after which no more contributions are generated. Typically $T$ would be the time of transition to an absorbing state (death or withdrawal), truncated at $n$. If compounding is at second order rate of interest, then

$$D(t) = 1[t \geq T] \int_0^T e^{\int_\tau^T r} dC(\tau).$$

Contribution dividends and terminal dividends represent opposite extremes in the set of conceivable dividend schemes, which are countless. One class of intermediate solutions are those that yield dividends only at certain times $T_1 < \cdots < T_K \leq n$, e.g. annually or at times of transition between certain states. At each time $T_i$ the amount $\Delta D(T_i) = \int_{T_{i-1}}^{T_i} e^{\int_{\tau}^{T_i} r} dC(\tau)$ (with $T_0 = 0$) is entered to the insured’s credit.

C. Allocation of dividends; bonus. Once they have been allotted, dividends belong to the insured. They may, however, be disposed of in various ways and need not be paid out currently as they fall due. The actual payouts of dividends are termed *bonus* in the sequel, and the corresponding payment function is denoted by $B^b$.

The compounded value of credited dividends less paid bonuses at time $t$ is

$$U^b(t) = \int_0^t e^{\int_\tau^t r} d\{D - B^b\}(\tau). \quad (8.23)$$

This is a debt owed by the insurer to the insured, and we shall call it the *bonus reserve* at time $t$. Bonuses may not be advanced, so $B^b$ must satisfy

$$U^b(t) \geq 0 \quad (8.24)$$

for all $t \in [0, n]$. In particular, since $D(0) = 0$, one has $B^b(0) = 0$. Moreover, since all dividends must eventually be paid out, we must have

$$U^b(n) = 0. \quad (8.25)$$

We have introduced three notions of reserves that all appear on the debit side of the insurer’s balance sheet. First, the premium reserve $V^*$ is provided to meet net outgoes in respect of future events; second, the dividend reserve $U^d$ is provided to settle the excess of past contributions over past dividends; third, the bonus reserve $U^b$ is provided to settle the unpaid part of dividends credited in the past. The premium reserve is of prospective type and is a predicted
amount, whereas the dividend and bonus reserves are of retrospective type and are indeed known amounts summing up to

\[ U^d(t) + U^b(t) = \int_0^t e^{\int_\tau^t r \, d\{C - B^b\}(\tau)} \, dC - B^b \]

the compounded total of past contributions not yet paid back to the insured.

D. Some commonly used bonus schemes. The term *cash bonus* is, quite naturally, used for the scheme \( B^b = D \). Under this scheme the bonus reserve is always null, of course.

By *terminal bonus*, also called *reversionary bonus*, is meant that all dividends, with accumulation of interest, are paid out as a lump sum upon the termination of the contract at some time \( T \), that is,

\[ B^b(t) = 1[t \geq T] \int_0^T e^{\int_\tau^T r \, dD(\tau)} \, dB^b \]

Here we could replace the integrator \( D \) by \( C \) since terminal bonus obviously does not depend on the dividend scheme; all contributions are to be repaid with accumulation of interest.

Assume now, what is common in practice, that dividends are currently used to purchase additional insurance coverage of the same type as in the primary policy. It seems natural to let the *additional benefits* be proportional to those stipulated in the primary policy since they represent the desired profile of the product. Thus, the dividends \( dD(s) \) in any time interval \( [s, s + ds) \) are used as a single premium for an insurance with payment function of the form

\[ dQ(s)\{B^+(\tau) - B^+(s)\}, \]

\( \tau \in (s, n] \), where the topscript “+” signifies, in an obvious sense, that only positive payments (benefits) are counted.

Supposing that additional insurances are written on first order basis, the proportionality factor \( dQ(s) \) is determined by

\[ dD(s) = dQ(s)\{V^+_{Z(s)}(s)\} \]

where

\[ V^+_{Z(s)}(s) = E^* \left[ \int_s^n e^{-\int_s^\tau r \, dB^+(\tau)} Z(s) \right] \]

is the single premium at time \( s \) for the future benefits under the policy.

Now the bonus payments \( B^b \) are of the form

\[ dB^b(t) = Q(t)dB^+(t) \]

Being written on first order basis, also the additional insurances create technical surplus. The total contributions under this scheme develop as

\[ dC(t) + Q(t)dB^+(t) \]
where the first term on the right stems from the primary policy and the second term stems from the \( Q(t) \) units of additional insurances purchased in the past, each of which has payment function \( B^+ \) producing contributions \( C^+ \) of the form

\[
dC^+ (t) = \sum_j I_j^+ (t) c^+_j (t) \, dt, \quad \text{with}
\]

\[
c^+_j (t) = \{ r(t) - r^* (t) \} V^+_j (t) + \sum_{k, k \neq j} R^+_{jk} (t) \{ \mu^+_{jk} (t) - \mu_{jk} (t) \},
\]

\[
R^+_{jk} (t) = b^+_{jk} (t) + V^+_k (t) - V^+_j (t).
\]

The present situation is more involved than those encountered previously since, not only are dividends driven by the contractual payments, but it is also the other way around. To keep things relatively simple, suppose that the contribution principle is adopted so that the dividends in (8.27) are set equal to the contributions in (8.29). Then the system is governed by the dynamics

\[
dC (t) + Q(t)dC^+ (t) = dQ(t) V^+_Z (t)
\]

or, realizing that \( V^+_Z (t) \) is strictly positive whenever \( dC (t) \) and \( dC^+ (t) \) are,

\[
dQ(t) - Q(t) dG(t) = dH(t), \quad (8.30)
\]

where \( G \) and \( H \) are defined by

\[
dG(t) = \frac{1}{V^+_Z (t)} dC^+ (t), \quad (8.31)
\]

\[
dH(t) = \frac{1}{V^+_Z (t)} dC (t). \quad (8.32)
\]

Multiplying with \( \exp(-G(t)) \) to form a complete differential on the left and then integrating from 0 to \( t \), using \( Q(0) = 0 \), we obtain

\[
Q(t) = \int_0^t e^{G(\tau) - G(t)} dH(\tau). \quad (8.33)
\]

### 8.5 Bonus prognoses

**A. A Markov chain environment.** We shall adopt a simple Markov chain description of the uncertainty associated with the development of the second order basis. Let \( Y(t), 0 \leq t \leq n \), be a time-continuous Markov chain with finite state space \( Y = \{1, \ldots, q\} \) and constant intensities of transition, \( \lambda_{ef} \). Denote the associated indicator processes by \( I^Y_e \). The process \( Y \) represents the “economic-demographic environment”, and we let the second order elements depend on the current \( Y \)-state:

\[
r(t) = \sum_e I^Y_e (t) r_e = r_Y (t),
\]

\[
\mu_{jk} (t) = \sum_e I^Y_e (t) \mu^e_{jk} (t) = \mu_{Y(t)jk} (t).
\]
The $r_e$ are constants and the $\mu_{e,j,k}(t)$ are intensity functions, all deterministic.

With this specification of the full two-stage model it is realized that the pair $X = (Y, Z)$ is a Markov chain on the state space $\mathcal{X} = \mathcal{Y} \times \mathcal{Z}$, and its intensities of transition, which we denote by $\kappa_{e,j,k}(t)$ for $(e,j), (f,k) \in \mathcal{X}$, $(e,j) \neq (f,k)$, are

$$\kappa_{e,j,f}(t) = \lambda_{ef}, \quad e \neq f; \tag{8.34}$$
$$\kappa_{e,j,k}(t) = \mu_{e,j,k}(t), \quad j \neq k, \tag{8.35}$$

and null for all other transitions.

In this extended set-up the contributions, whose dependence on the second order elements was not visualized earlier, can appropriately be represented as

$$dC(t) = c(t) \, dt = \sum_{e,j} I_e^Y(t) I_j^Z(t) c_{e,j}(t) \, dt,$$

where

$$c_{e,j}(t) = \{r_e - r^*(t)\} V^*_j(t) + \sum_{k, k \neq j} R^*_j(t) \{\mu^*_j(t) - \mu_{e,j,k}(t)\}. \tag{8.36}$$

Under the scheme of additional benefits described in Paragraph 8.4.D a similar convention goes for $C^+$ and $c^+$ and, accordingly, (8.31) and (8.32) become

$$dG(t) = g(t) \, dt = \sum_{e,j} I_e^Y(t) I_j^Z(t) g_{e,j}(t) \, dt, \tag{8.37}$$
$$g_{e,j}(t) = \frac{c^+_{e,j}(t)}{V^*_{j^+} (t)}, \tag{8.38}$$
$$dH(t) = h(t) \, dt = \sum_{e,j} I_e^Y(t) I_j^Z(t) h_{e,j}(t) \, dt, \tag{8.39}$$
$$h_{e,j}(t) = \frac{c^+_{e,j}(t)}{V^*_{j^+} (t)}. \tag{8.40}$$

B. Preparatory remarks on the issue of bonus prognoses. There is no single functional of the future bonus stream that presents itself as the relevant quantity to prognosticate. One could e.g. take the total bonuses discounted by some suitable inflation rate, or the undiscounted total bonuses, or the rate at which bonus will be paid at certain times, and one could apply any of these possibilities to the random development of the policy or to some representative fixed development. We shall focus on the expected value, and in the simplest cases also higher order moments, of the future bonuses discounted by the stochastic second order interest. From this we can easily deduce predictors for a number of other relevant quantities. We turn now to the analysis of some of the schemes described in Section 8.4.
C. Contribution dividends and cash bonus. This case, where $B^b = C = D$, is particularly simple since the bonus payments at any time depend only on the current state of the process. We can then employ the appropriate version of Thiele’s differential equation to calculate the state-wise expected discounted future bonuses ($= $ contributions),

$$W_{ej}(t) = E \left[ \int_t^0 e^{-r(\tau - t)} c(\tau) d\tau \mid X(t) = (e, j) \right].$$

They are determined by the appropriate version of Thiele’s differential equation,

$$\frac{d}{dt} W_{ej}(t) = r_e W_{ej}(t) - c_{ej}(t) - \sum_{j,j \neq e} \lambda_{ej} (W_{fj}(t) - W_{ej}(t))$$

$$- \sum_{k,k \neq j} \mu_{e;jk}(t) (W_{ek}(t) - W_{ej}(t)),$$

subject to

$$W_{ej}(n-) = 0, \quad \forall e, j.$$  \hspace{1cm} (8.42)

D. Terminal dividend and/or bonus. Under the terminal bonus scheme dividends and bonuses are the same, of course. The problem of predicting the total bonus payments discounted with respect to second order interest is basically the same as in the previous paragraph since it amounts to adding the total amount of compounded past contributions, which is known, and the state-wise predictor of discounted future contributions.

Suppose instead that at time $t$, the policy still being in force, it is decided to predict the undiscounted value of the terminal bonus amount,

$$W = \int_0^T e^{\int_\tau^T r(\tau) d\tau} d\tau = \int_0^t e^{\int_\tau^T r(\tau) d\tau} d\tau W'(t) + W''(t),$$  \hspace{1cm} (8.43)

where

$$W'(t) = e^{\int_t^T r(\tau) d\tau},$$

$$W''(t) = \int_t^T e^{\int_\tau^T r(\tau) d\tau} d\tau.$$

We need the state-wise expected values

$$W'_e(t) = E[W'(t) \mid Y(t) = e],$$

$$W''_e(t) = E[W''(t) \mid X(t) = (e, j)],$$

to find the state-wise predictors of $W$ in (8.43),

$$W_{ej}(t) = \int_0^t e^{\int_\tau^T r(\tau) d\tau} d\tau W'_e(t) + W''_e(t).$$
We shall find these functions by the backward construction, starting from

\[ W'(t) = e^{r \, dt} W'(t + dt), \]
\[ W''(t) = c(t) \, dt \, W'(t) + W''(t + dt). \]

Conditioning on what happens in the small time interval \((t, t + dt]\), we get

\[
W'_e(t) = e^{r \, dt} 
\left( (1 - \lambda_e \, dt) \, W'_e(t + dt) + \sum_{f, f \neq e} \lambda_{ef} (t) \, dt \, W'_f(t + dt) \right),
\]

and

\[
W''_{ej}(t) = c_{ej}(t) \, dt \, W'(t) + (1 - (\lambda_e + \mu_{e;j} (t)) \, dt) \, W''_{ej}(t + dt) + \sum_{f, f \neq e} \lambda_{ef} (t) \, dt \, W''_{fj}(t + dt) + \sum_{k; k \neq j} \mu_{e;jk}(t) \, dt \, W''_{ek}(t + dt).
\]

From these relationships we easily obtain the differential equations

\[
\frac{d}{dt} W'_e(t) = -r_e W'_e(t) - \sum_{f; f \neq e} \lambda_{ef} (W'_f(t) - W'_e(t)), \tag{8.44}
\]
\[
\frac{d}{dt} W''_{ej}(t) = -c_{ej}(t) W'_e(t) - \sum_{f; f \neq e} \lambda_{ef} (W''_{fj}(t) - W''_{ej}(t)) - \sum_{k; k \neq j} \mu_{e;jk}(t) (W''_{ek}(t) - W''_{ej}(t)), \tag{8.45}
\]

which are to be solved subject to

\[ W'_e(n-) = 1, \quad W''_{ej}(n-) = 0, \quad \forall e, j. \tag{8.46} \]

E. Additional benefits. Suppose we want to predict the total future bonuses discounted with respect to second order interest,

\[ W(t) = \int_t^n e^{-L_r \, r} Q(r) \, dB^+(r), \]

with \(Q\) defined by (8.33). Recalling (8.37)–(8.40), we reshape \(W(t)\) as

\[
W(t) = \int_t^n e^{-L_r \, r} \int_0^r e^{L_r \, g \, h(r) \, dr} dB^+(r) = \int_t^n e^{-L_r \, r} \left( \int_0^r e^{L_r \, g \, h(r) \, dr} e^{L_r \, g} + \int_t^r e^{L_r \, g \, h(r) \, dr} \right) dB^+(r) = \int_0^t e^{L_r \, g \, h(r) \, dr} W'(t) + W''(t), \tag{8.47}
\]
with
\[
W'(t) = \int_t^n e^\int_{\tau}^{t} (g-r) \, dB^+(\tau),
\]
\[
W''(t) = \int_t^n e^{-\int_{\tau}^{t} r} \, W'(\tau) h(\tau) \, d\tau.
\]
Thus, we need the state-wise expected values
\[
W'_e(t) = \mathbb{E}[W'(t) \mid X(t) = (e, j)],
\]
\[
W''_e(t) = \mathbb{E}[W''(t) \mid X(t) = (e, j)],
\]
in order to find the state-wise predictors of \(W(t)\) in (8.47),
\[
W_e(t) = \int_t^0 e^{\int_{\tau}^{t} g(r) \, dr} W'_e(t) + W''_e(t).
\]
The backward equations start from
\[
W'(t) = dB^+(t) + e^{(g(t)-r(t)) \, dt} W'(t+dt),
\]
\[
W''(t) = W'(t) h(t) \, dt + e^{-r(t) \, dt} W''(t+dt),
\]
from which we proceed in the same way as in the previous paragraph to obtain
\[
dW'_e(t) = -dB_e^+(t) + (r_e - g_e(t)) \, dt \, W'_e(t) \]
\[\quad - \sum_{f,j \neq e} \lambda_{ef} \, dt \, (W'_f(t) - W'_e(t)) \]
\[\quad - \sum_{k,j \neq e} \mu_{ejk} \, dt \, (h_{jk}^+(t) + W'_k(t) - W'_e(t)), \tag{8.48}
\]
\[
dW''_e(t) = -W'_e(t) h_e(t) \, dt + r_e \, dt \, W''_e(t) \]
\[\quad - \sum_{f,j \neq e} \lambda_{ef} \, dt \, (W''_f(t) - W''_e(t)) \]
\[\quad - \sum_{k,j \neq e} \mu_{ejk} \, dt \, (W''_k(t) - W''_e(t)). \tag{8.49}
\]
The appropriate side conditions are
\[
W'_e(n) = \Delta B^+_e(n), \quad W''_e(n) = 0, \quad \forall e, j. \tag{8.50}
\]
**F. Predicting undiscounted amounts.** If the undiscounted total contributions or additional benefits is what one wants to predict, one can just apply the formulas with all \(r_e\) replaced by 0.
G. Predicting bonuses for a given policy path. Yet another form of prognosis, which may be considered more informative than the two mentioned above, would be to predict bonus payments for some possible fixed pursuits of a policy instead of averaging over all possibilities. Such prognoses are obtained from those described above upon keeping the realized path $Z(\tau)$ for $\tau \in [0,t]$, where $t$ is the time of consideration, and putting $Z(\tau) = z(\tau)$ for $\tau \in (t,n]$, where $z(\cdot)$ is some fixed path with $z(t) = Z(t)$. The relevant predictors then become essentially functions only of the current $Y$-state and are simple special cases of the results above.

As an example of an even simpler type of prognosis for a policy in state $j$ at time $t$, the insurer could present the expected bonus payment per time unit at a future time $s$, given that the policy is then in state $i$, and do this for some representative selections of $s$ and $i$. If $Y(t) = e$, then the relevant prediction is

$$\mathbb{E}[c_Y(s) | Y(t) = e] = \sum_f p^Y_{ef}(t,s)c_{fi}(s).$$

8.6 Examples

A. The case. For our purpose, which is to illustrate the role of the stochastic environment in model-based prognoses, it suffices to consider simple insurance products for which the relevant policy states are $Z = \{a, d\}$ (‘alive’ and ‘dead’).

We will consider a single life insured at age 30 for a period of $n = 30$ years, and let the first order elements be those of the Danish technical basis G82M for males:

$$r^* = \ln(1.045),$$
$$\mu^*_{ad}(t) = \mu^*(t) = 0.0005 + 0.000075858 \cdot 10^{0.038(30+t)}.$$

Three different forms of insurance benefits will be considered, and in each case we assume that premiums are payable continuously at level rate as long as the policy is in force. First, a term insurance (TI) of $1 = b_{ad}(t)$ with first order premium rate $0.0042608 = -b_a(t)$. Second, a pure endowment (PE) of $1 = \Delta B_a(30)$ with first order premium rate $0.0140690 = -b_a(t)$. Third, an endowment insurance (EI), which is just the combination of the former two; $1 = b_{ad}(t) = \Delta B_a(30), 0.0183298 = -b_a(t)$.

Just as an illustration, let the second order model be the simple one where interest and mortality are governed by independent time-continuous Markov chains and, more specifically, that $r$ switches with a constant intensity $\lambda_i$ between the first order rate $r^*$ and a better rate $\epsilon_i r^*$ ($\epsilon_i > 1$) and, similarly, $\mu$ switches with a constant intensity $\lambda_m$ between the first order rate $\mu^*$ and a better rate $\epsilon_m \mu^*$ ($\epsilon_m < 1$). (We choose to express ourselves this way although (8.15) shows that, for insurance forms with negative sum at risk, e.g. pure endowment insurance, it is actually a higher second order mortality that is “better” in the sense of creating positive contributions.)
The situation fits into the framework of Paragraph 8.5.A; \( Y \) has states \( Y = \{bb, gb, bg, gg\} \) representing all combinations of “bad” (b) and “good” (g) interest and mortality, and the non-null intensities are
\[
\lambda_{bb,gb} = \lambda_{gb,bb} = \lambda_{bg,gg} = \lambda_{gg,bg} = \lambda_i, \\
\lambda_{bb,bg} = \lambda_{gg,bb} = \lambda_{gb,gg} = \lambda_{gg,bg} = \lambda_m.
\]
The first order basis is just the worst-scenario \( bb \).

Adopting the device (8.34)–(8.35), we consider the Markov chain \( X = (Y, Z) \) with states \( (bb, a), (gb, a), \) etc. It is realized that all death states can be merged into one, so it suffices to work with the simple Markov model with five states sketched in Figure 8.1.

**B. Results.** We shall report some numerical results for the case where \( \epsilon_i = 1.25, \epsilon_m = 0.75, \) and \( \lambda_i = \lambda_m = 0.1 \). Prognoses are made at the time of issue of the policy. Computations were performed by the fourth order Runge-Kutta method, which turns out to work with high precision in the present class of situations.
Table 8.1 displays, for each of the three policies, the state-wise expected values of discounted contributions obtained by solving (8.41)–(8.42). We shall be content here to point out two features: First, for the term insurance the mortality margin is far more important than the interest margin, whereas for the pure endowment it is the other way around (the latter has the larger reserve). Note that the sum at risk is negative for the pure endowment, so that the first order assumption of excess mortality is really not to the safe side, confer (8.15). Second, high interest produces large contributions, but, since high initial interest also induces severe discounting, it is not necessarily true that good initial interest will produce a high value of the expected discounted contributions, confer the two last entries in the row TI.

The latter remark suggests the use of a discounting function different from the one based on the second order interest, e.g. some exogenous deflator reflecting the likely development of the price index or the discounting function corresponding to first order interest. In particular, one can simply drop discounting and prognosticate the total amounts paid. We shall do this in the following, noting that the expected value of bonuses discounted by second order interest must in fact be the same for all bonus schemes, and are already shown in Table 8.1.

Table 8.2 shows state-wise expected values of undiscounted bonuses for three different schemes; contribution dividends and cash bonus ($C$), the same as total undiscounted contributions), terminal bonus ($TB$), and additional benefits ($AB$).

We first note that, now, any improvement of initial second order conditions helps to increase prospective contributions and bonuses. Furthermore, expected bonuses are generally smaller for $C$ than for $TB$ and $AB$ since bonuses under $C$ are paid earlier. Differences between $TB$ and $AB$ must be due to a similar effect. Thus, we can infer that $AB$ must on the average fall due earlier than $TB$, except for the pure endowment policy, of course. One might expect that the bonuses for the term insurance and the pure endowment policies add up to the bonuses for the combined endowment insurance policy, as is the case for $C$ and $TB$. However, for $AB$ it is seen that the sum of the bonuses for the two component policies is generally smaller than the bonuses for the combined policy. The explanation must be that additional death benefits and additional survival benefits are not purchased in the same proportions under the two policy strategies. The observed difference indicates that, on the average, the additional benefits fall due later under the combined policy, which therefore must have the smaller proportion of additional death benefits.

C. Assessment of prognostication error. Bonus prognoses based on the present model may be equipped with quantitative measures of the prognostication error. By the technique of proof shown in Section 8.5 we may derive differential equations for higher order moments of any of the predictands considered and calculate e.g. the coefficient of variation, the skewness, and the kurtosis.
CHAPTER 8. SAFETY LOADINGS AND BONUS

Table 8.1: Conditional expected present value at time 0 of total contributions for term insurance policy (TI), pure endowment policy (PE), and endowment insurance policy (EI), given initial second order states of interest and mortality (b or g).

<table>
<thead>
<tr>
<th></th>
<th>bb</th>
<th>gb</th>
<th>bg</th>
<th>gg</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI</td>
<td>.00851</td>
<td>.00854</td>
<td>.01061</td>
<td>.01059</td>
</tr>
<tr>
<td>PE</td>
<td>.01613</td>
<td>.01823</td>
<td>.01595</td>
<td>.01807</td>
</tr>
<tr>
<td>EI</td>
<td>.02463</td>
<td>.02677</td>
<td>.02656</td>
<td>.02865</td>
</tr>
</tbody>
</table>

8.7 Including expenses

A. The form of the expenses. Expenses are assumed to incur in accordance with a non-decreasing payment function \( A \) of the same type as the contractual payments, that is,

\[
dA(t) = \sum_j I_j^Z(t) dA_j(t) + \sum_{j \neq k} a_{jk}(t) dN_{jk}^Z(t) .
\] (8.51)

It is common in practice to assume, furthermore, that expenses of annuity type incur with a lump sum of initial costs at time 0 and thereafter continuously at a rate that depends on the current state, that is, \( \Delta A_0(0) > 0 \) and \( dA_j(t) = a_j(t) dt \) for \( t > 0 \). The transition costs \( a_{jk}(t) \) are not explicitly taken into account in practice, but we include them here since they add realism without adding mathematical complexity.

B. First order assumptions. The elements \( \Delta A_0(0) \), \( a_j(t) \), and \( a_{jk}(t) \) will in general depend on the second order development, and the first order basis must, therefore, specify prudent estimates \( \Delta A_0^*(0) \), \( a_j^*(t) \), and \( a_{jk}^*(t) \). Denote the corresponding payment function by \( A^* \).

C. Surplus and contributions in the presence of expenses. The introduction of expenses adds a new feature to the previous set-up in that also the payments become dependent on the second order development. However, the essential parts of the analyses in the previous sections carry over with merely notational modifications; all it takes is to replace everywhere the contractual payment function \( B \) with \( A + B \) in the past and \( A^* + B \) in the future. One finds, in particular, that the first order equivalence relation (8.5) now turns into

\[
V_0^*(0) = -\Delta A_0^*(0) - \Delta B_0(0) ,
\] (8.52)

the surplus at time 0 becomes

\[
S(0) = \Delta A_0^*(0) - \Delta A_0(0) ,
\] (8.53)
Table 8.2: Conditional expected value (E) of undiscounted total contributions (C), terminal bonus (TB), and total additional benefits (AB) for term insurance policy (TI), pure endowment policy (PE), and endowment insurance policy (EI), given initial second order states of interest and mortality (b or g).

<table>
<thead>
<tr>
<th></th>
<th>bb</th>
<th>gb</th>
<th>bg</th>
<th>gg</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI:</td>
<td>.02153</td>
<td>.02222</td>
<td>.02436</td>
<td>.02505</td>
</tr>
<tr>
<td></td>
<td>.03693</td>
<td>.03916</td>
<td>.04600</td>
<td>.04847</td>
</tr>
<tr>
<td></td>
<td>.02949</td>
<td>.03096</td>
<td>.03545</td>
<td>.03706</td>
</tr>
<tr>
<td>PE:</td>
<td>.04342</td>
<td>.04818</td>
<td>.04314</td>
<td>.04791</td>
</tr>
<tr>
<td></td>
<td>.07337</td>
<td>.08687</td>
<td>.07264</td>
<td>.08615</td>
</tr>
<tr>
<td></td>
<td>.07337</td>
<td>.08687</td>
<td>.07264</td>
<td>.08615</td>
</tr>
<tr>
<td>EI:</td>
<td>.06495</td>
<td>.07040</td>
<td>.06750</td>
<td>.07296</td>
</tr>
<tr>
<td></td>
<td>.11030</td>
<td>.12603</td>
<td>.11864</td>
<td>.13462</td>
</tr>
<tr>
<td></td>
<td>.10723</td>
<td>.12199</td>
<td>.11501</td>
<td>.13003</td>
</tr>
</tbody>
</table>

and the contributions consist of a jump

$$\Delta C(0) = \Delta A^*_0(0) - \Delta A_0(0)$$  \hspace{1cm} (8.54)

at time 0 and thereafter a continuous part, which is defined upon replacing (8.9) with

$$c_j(t) = \{ r(t) - r^*(t) \} V^*_j(t) + \{ a^*_j(t) - a_j(t) \}$$

$$+ \sum_{k, k \neq j} (a^*_jk(t) - a_{jk}(t)) \mu_{jk}(t)$$

$$+ \sum_{k, k \neq j} R^*_jk(t) \{ \mu^*_jk(t) - \mu_{jk}(t) \},$$  \hspace{1cm} (8.55)

where now

$$R^*_jk(t) = a^*_jk(t) + b_{jk}(t) + V^*_k(t) - V^*_j(t).$$  \hspace{1cm} (8.56)

Referring to the discussion in Paragraph 8.3.C, we see that the contributions emerge from safety margins in all first order elements, r*, μ^*_jk, and A*.

**D. Prediction in the presence of expenses.** The complexity of the prediction problem depends heavily on the assumptions made about the second order expenses, and at this point some new problems may arise.

Just to get started, suppose first that the expense elements ∆A_0(0), a_j(t), and a_{jk}(t) are deterministic. Then the methods in Section 8.5 carry over with only trivial modifications. Presumably, this simplistic model is at the base of the frequently encountered claim that “administration expenses can be regarded
as additional benefits”. Unfortunately, real life expenses are of a different, and
typically less pleasant, nature. An exhaustive discussion of this issue could easily
exhaust the reader, so we shall be content with just outlining some tentative
ideas.

The problem is that expenses are made up of wages, commissions, rent,
taxes and other items that are governed by the economic development. In
the framework of the Markov model in Paragraph 8.5.A, one simple way of
accounting for such effects is to make the second order expenses dependent on
the current state of \( Y \), that is,

\[
\begin{align*}
\Delta A_0(0) &= \sum_e I^Y_e(t) \Delta A_{e0}(0), \\
a_j(t) &= \sum_e I^Y_e(t) a_{ej}(t), \\
a_{jk}(t) &= \sum_e I^Y_e(t) a_{ejk}(t),
\end{align*}
\]

with deterministic \( \Delta A_{e0}, a_{ej}, \) and \( a_{ejk} \). By enriching sufficiently the state space
of \( Y \), one can in principle create a fairly realistic model.

Perhaps the most reasonable point of view is that expenses are inflated by
some time-dependent rate \( \gamma(t) \) so that we should put \( a_j(t) = e^{\int_0^t \gamma(t') \, dt'} a_j^0(t) \) and
\( a_{jk}(t) = e^{\int_0^t \gamma(t') \, dt'} a_{jk}^0(t) \) with \( a_j^0 \) and \( a_{jk}^0 \) deterministic. One possibility is to put
the second order force of interest \( r \) in the role of \( \gamma \). More realistically one should
let \( \gamma \) be something else, but still related to \( r \) through joint dependence on a
suitably specified \( Y \). We shall not pursue this idea any further here, but note,
by way of warning, that prognostication in this kind of inflation model will
present problems in addition to those solved in Section 8.5.

8.8 Discussions

A. The principle of equivalence. This principle, as formulated in (8.4), is
basic in life insurance. The expected value represents averaging over a large
(really infinite) portfolio of policies, the philosophy being that, even if the in-
dividual policy creates a (possibly large) loss or gain, there will be balance on
the average between outgoes and incomes in the portfolio as a whole if the pre-
miums are set by equivalence. The deviation from perfect balance, which is
inevitable in a finite world with finite portfolios, represents profit or loss on the
part of the insurer and has to be settled by an adjustment of the equity capital.
(The possibility of loss, about as likely and about as large as the possible profit,
might seem unacceptable to an industry that needs to attract investors, but it
should be kept in mind that salaries to employees and dividends to owners are
accounted as part of the expenses discussed in Section 8.7.)

B. On the notion of second order basis. The definition of the second
order basis as the true one is slightly at variance with practical usage (which is
not uniform anyway). The various amendments made to our idealized definition in practice are due to administrative and procedural bottlenecks: The factual development of interest, mortality, etc. has to be verified by the insurer and then approved by the supervisory authority. Since this can not be a continuous operation, any regulatory definition of the second order basis must to some extent involve realistic, still typically conservative, short term forecasts of the future development. However, our definition can certainly be agreed upon as the intended one.

C. Model deliberations. The Markov chain model is mathematically tractable since state-wise expected values are determined by solving (in most cases simple) systems of first order ordinary differential equations. At the same time, when equipped with a sufficiently rich state space and appropriate intensities of transition, it is able to picture virtually any conceivable notion of the real object of the model.

The Markov chain model is particularly apt to describe the development of life insurance policy since the paths of \( Z \) are of the same kind as the true ones. When used to describe the development of the second order basis, however, the approximative nature of the Markov chain is obvious, and it will surface immediately as e.g. the experienced force of interest takes values outside of the finite set allowed by the model. This is not a serious objection, however, and the next paragraph explains why.

D. The role of the stochastic environment model. A paramount concern is that of establishing equivalence conditional on the factual second order history in the sense of (8.16). Now, in this conditional expectation the marginal distribution of the second order elements does not appear and is, in this respect, irrelevant. Also the contributions and, hence, the dividends are functions only of the realized experience basis and do not involve the distribution of its elements.

Then, what remains the purpose of placing a distribution on the second order elements is to form a basis for prognostication of bonus. Subsidiary as it is, this role is still an important part of the play; although a prognosis does not commit the insurer to pay the forecasted amounts, it should as much as possible be a reliable piece of information to the insured. Therefore, the distribution placed on the second order elements should set a reasonable scenario for the course of events, but it need not be perfectly true. This is comforting since any view of the mechanisms governing the economic-demographic development is to some extent guess-work. When the accounts are eventually made up, every speculative element must be absent, and that is precisely what the principle (8.16) lays down.

E. A digression: Which is more important, interest or mortality? Actuarial wisdom says it is interest. This is, of course, an empirical statement based on the fact that, in the era of contemporary insurance, mortality rates
have been smaller and more stable than interest rates. Our model can add some other kind of insight. We shall again be content with a simple illustration related to the single life described in Section 8.6. Table 8.3 displays expected values and standard deviations of the present values at time 0 of a term life insurance and a life annuity under various scenarios with fixed interest and mortality, that is, conditional on fixed Y-state throughout the term of the policy. The impact of interest variation is seen by reading column-wise, and the impact of mortality variation is seen by reading row-wise. The overall impression is that mortality is the more important element by term insurance, whereas interest is the (by far) more important by life annuity insurance.
Table 8.3: Expected value (E) and standard deviation (SD) of present values of a term life insurance (TI) with sum 1 and a life annuity (LA) with level intensity 1 per year, with interest $r = \epsilon_i \cdot r^*$ and mortality $\mu = \epsilon_m \mu^*$ for various choices of $\epsilon_i$ and $\epsilon_m$.

<table>
<thead>
<tr>
<th>$\epsilon_i$</th>
<th>1.5</th>
<th>1.0</th>
<th>0.5</th>
<th>1.5</th>
<th>1.0</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon_m$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>.14636</td>
<td>.10119</td>
<td>.05250</td>
<td>20.545</td>
<td>20.996</td>
<td>21.467</td>
</tr>
<tr>
<td>SD</td>
<td>.27902</td>
<td>.24104</td>
<td>.18041</td>
<td>.03691</td>
<td>.03101</td>
<td>.02257</td>
</tr>
<tr>
<td>LA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>.09927</td>
<td>.06834</td>
<td>.03531</td>
<td>15.750</td>
<td>16.039</td>
<td>16.340</td>
</tr>
<tr>
<td>SD</td>
<td>.20245</td>
<td>.17330</td>
<td>.12857</td>
<td>.02505</td>
<td>.02097</td>
<td>.01521</td>
</tr>
<tr>
<td>E</td>
<td>.06976</td>
<td>.04782</td>
<td>.02460</td>
<td>12.466</td>
<td>12.655</td>
<td>12.852</td>
</tr>
<tr>
<td>SD</td>
<td>.15858</td>
<td>.13437</td>
<td>.09868</td>
<td>.01759</td>
<td>.01468</td>
<td>.01061</td>
</tr>
</tbody>
</table>
Chapter 9

Inference in the Markov model

9.1 Statistical analysis of truncated life times

A. Right-censored life times. Let the total life length of some item be represented by a non-negative random variable $U$ with cumulative distribution function $F$ of the form

$$ F(u) = 1 - e^{-\int_0^u \mu(s)\,ds}. \quad (9.1) $$

The mortality intensity $\mu$ is assumed to be piecewise continuous so that the density,

$$ f(u) = \mu(u)(1 - F(u)), \quad (9.2) $$

exists almost everywhere.

Suppose now that the period of observation is delimited to $z$ so that only the truncated life length $T = U \wedge z$ is observed. A technical term for this kind of observational plan is right-censoring at time $z$. The cumulative distribution of $T$ is

$$ G(t) = \begin{cases} F(t), & 0 < t < z, \\ 1, & t \geq z, \end{cases} $$

and the density (with respect to Lebesgue measure on $(0, z)$ and the unit mass at $z$) is (recall (9.2))

$$ g(t) = \begin{cases} \mu(t)(1 - F(t)), & 0 < t < z, \\ 1 - F(z), & t = z. \end{cases} $$

Introduce

$$ d(t) = 1_{(0,z)}(t) = \begin{cases} 1, & 0 < t < z, \\ 0, & t \geq z, \end{cases} \quad (9.3) $$

to obtain the closed expression

$$ g(t) = \mu(t)^{d(t)}(1 - F(t)), \quad 0 < t \leq z. \quad (9.4) $$
CHAPTER 9. INFERENCE IN THE MARKOV MODEL

Denote the indicator function of survival to age \( u > 0 \) by \( I(u) = 1_{\{U > u\}} \).

The indicator of death before age \( z \) is

\[
D = d(T) = 1 - I(z-) = 1 - I(z),
\]

where the last equality holds with probability 1.

In the sequel we shall need the following formulas, valid whenever the displayed moments exist:

\[
E[D^k] = F(z), \quad k = 1, 2, \ldots, \quad (9.6)
\]

\[
E[T^k] = k \int_0^z t^{k-1}(1 - F(t))dt, \quad k = 1, 2, \ldots, \quad (9.7)
\]

\[
E[DT] = E[T] - z(1 - F(z)). \quad (9.8)
\]

To verify (9.8), use (9.5) and \( T = \int_0^z I(t)dt \) to write \( DT = \int_0^z I(t)dt - zI(z) \), and take expectation.

B. The truncated exponential distribution. We set out by analyzing the simple case with constant mortality intensity, partly as a motivating example, but also because the techniques are at the base of an important class of procedures in actuarial life history analysis.

Thus, assume that \( U \) is exponentially distributed with cumulative distribution function

\[
F(u; \mu) = 1 - e^{-\mu u}, \quad u > 0, \quad (9.9)
\]

that is, \( \mu \) is constant, independent of age. The expected life length is

\[
\nu = \int_0^\infty (1 - F(u; \mu))du = \frac{1}{\mu}. \quad (9.10)
\]

From (9.6) – (9.9) one easily calculates

\[
E[D] = 1 - e^{-\mu z}, \quad (9.11)
\]

\[
\mathbb{V}[D] = e^{-\mu z}(1 - e^{-\mu z}), \quad (9.12)
\]

\[
E[T] = \frac{1 - e^{-\mu z}}{\mu}, \quad (9.13)
\]

\[
\mathbb{V}[T] = \frac{1 - 2\mu ze^{-\mu z} - e^{-2\mu z}}{\mu^2}, \quad (9.14)
\]

\[
E[D - \mu T] = 0, \quad (9.15)
\]

\[
\mathbb{V}[D - \mu T] = 1 - e^{-\mu z}. \quad (9.16)
\]
C. Maximum likelihood estimators based on censored exponential variates. Let \( U_i, i = 1, 2, \ldots \), be independent replicates of \( U \). Consider the problem of estimating \( \mu \) from a sample of \( n \) censored life lengths, \( T_i = U_i \wedge z_i, i = 1, \ldots, n \). The interpretation is that a mortality study is carried out in a population during a certain period of time terminating at time \( \bar{t} \), say, the sample being \( n \) individuals born during the period at times \( \bar{t} - z_i, i = 1, \ldots, n \).

Referring to (9.5), put \( D_i = 1_{\{T_i < z_i\}}, i = 1, \ldots, n \). By (9.4) and (9.9), the likelihood of the observables is
\[
\Lambda = \prod_{i=1}^{n} \mu^{D_i} e^{-\mu T_i} = \mu^N e^{-\mu W} = e^{\ln \mu N - \mu W},
\]
(9.17)
where
\[
N = \sum_{i=1}^{n} D_i, \text{ the total number of deaths occurred},
\]
\[
W = \sum_{i=1}^{n} T_i, \text{ the total time exposed to risk of death}.
\]

Clearly, \((N, W)\) is a sufficient statistic. Take the logarithm,
\[
\ln \Lambda = \ln \mu N - \mu W,
\]
and form the derivatives
\[
\frac{\partial}{\partial \mu} \ln \Lambda = - \frac{N}{\mu} - W, \quad (9.18)
\]
\[
\frac{\partial^2}{\partial \mu^2} \ln \Lambda = - \frac{N}{\mu^2}. \quad (9.19)
\]

Putting the expression in (9.18) equal to 0 and noting that the second derivative is non-positive, we find that the MLE (maximum likelihood estimator) of \( \mu \) is the so-called O-E (occurrence-exposure) rate,
\[
\hat{\mu} = \frac{N}{W}, \quad (9.20)
\]
the number of deaths occurred per unit of time exposed to risk of death in the sample. It is the empirical counterpart of the mortality intensity, which is the expected number of deaths per time unit, roughly speaking.

The MLE of the expected life length in (9.10) is
\[
\hat{\nu} = \frac{W}{N}, \quad (9.21)
\]
defined as \(+\infty\) when \( N = 0 \). This estimator has no mean (and no higher order moments).

The forms of the likelihood in (9.17) and the MLE in (9.20) are independent of the censoring pattern. The censoring is, however, decisive of the probability distribution of \( \hat{\mu} \) and, hence, of its performance as an approximation to \( \mu \).
Unfortunately, this probability distribution is not easy to calculate in general, and we shall therefore have to add assumptions about the censoring mechanism, ranging from the special case of no censoring, where everything is simple and a lot of powerful results can be proved, to weak conditions under which only certain asymptotic properties are in reach.

D. The special case with no censoring. Suppose now that the $n$ lives are completely observed without censoring, that is, $z_i = \infty$ and $T_i = U_i$, $i = 1, \ldots, n$. Then all $D_i$ are 1, $N = n$, $W = \sum_{i=1}^{n} U_i$, and the likelihood in (9.17) becomes

$$\Lambda = e^{|n\mu - \mu W|}. \quad (9.22)$$

In this simple situation it is easy to investigate the small sample properties of the estimators. The sum of the life lengths, $W$, is now a sufficient statistic. It has a gamma distribution with shape parameter $n$ and scale parameter $\nu = 1/\mu$, whose density is

$$\frac{\mu^n}{\Gamma(n)} w^{n-1} \exp(-\mu w), \quad w > 0.$$ 

One finds (perform the easy calculations) for $k > -n$ that

$$\mathbb{E}[W^k] = \frac{\Gamma(n + k)}{\Gamma(n)\mu^k},$$

hence

$$\mathbb{E}[\hat{\mu}] = \frac{n\mu}{n-1}, \quad n > 1,$$

$$\mathbb{V}[\hat{\mu}] = \frac{n^2\mu^2}{(n-1)(n-2)}, \quad n > 2,$$

and

$$\mathbb{E}[\hat{\nu}] = \nu, \quad n \geq 1,$$

$$\mathbb{V}[\hat{\nu}] = \frac{\nu^2}{n}, \quad n \geq 1.$$

The estimator $\hat{\mu}$ is biased and, on the average, overestimates $\mu$ by $\mu/(n-1)$ (negligible for large $n$). An unbiased estimator of $\mu$ is $\bar{\mu} = (n-1)/W$. Its variance is $\mu^2/(n-2)$. The estimator $\hat{\nu}$ is now just the observed average life length, the straightforward empirical counterpart of $\nu$. It is unbiased, of course. In fact, $\bar{\mu}$ and $\bar{\nu}$ are UMVUE (uniformly minimum variance unbiased estimators) since they are based on $W$, which is a sufficient and complete statistic: by (9.22), the distribution belongs to an exponential family with canonical parameter $\mu$ varying in the open set $(0, \infty)$. 

E. Asymptotic results by uniform censoring. Suppose all \( z_i \) are equal to \( z \), say. Writing
\[
\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} D_i / \sum_{i=1}^{n} T_i,
\]
and noting that \( \mathbb{E}[D_i] = \mu \mathbb{E}T_i \) by (9.11) and (9.13), it follows by the strong law of large numbers that the estimator is strongly consistent,
\[
\hat{\mu} \overset{a.s.}{\to} \mu.
\]
To investigate its asymptotic distribution, look at
\[
\sqrt{n}(\hat{\mu} - \mu) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( D_i - \mu T_i \right) / \frac{1}{n} \sum_{i=1}^{n} T_i.
\]
The denominator of this fraction converges a.s. to \( \mathbb{E}[T_i] \) given by (9.13). By the central limit theorem, the limiting distribution of the numerator is normal with mean 0 (recall (9.15)) and variance given by (9.16). It follows that
\[
\hat{\mu} \sim_{a.s.} \mathcal{N}\left(\mu, \mu^2 n (1 - e^{-\mu z}) \right).
\] (9.23)

Copying the arguments above (or using (D.6) in Appendix D), it can also be concluded that \( \hat{\nu} \) defined by (9.21) is strongly consistent and that
\[
\hat{\nu} \sim_{a.s.} \mathcal{N}\left(\nu, \frac{1}{n \mu^2 (1 - e^{-\mu z})} \right).
\] (9.24)

No strong conclusions as to optimality can be drawn in parallel to those in the previous paragraph. The reason is seen from (9.17): the distribution belongs to a general exponential family with canonical parameter \((\ln \mu, \mu)\), which does not vary in an open (two-dimensional) set. Therefore, the sufficient statistic \((N, W)\) cannot be proved to be complete (not the usual way at least), and standard theory for inference in regular exponential families of distributions cannot be employed.

F. Asymptotic results by fairly general censoring. Consider now the general situation in Paragraph C with censoring varying among the individuals. A bit more effort must now be put into the study of the asymptotic properties of the MLE. It turns out that a sufficient condition for consistency and asymptotic normality is that the expected exposure grows to infinity in the sense that
\[
\sum_{i=1}^{n} \mathbb{E}[T_i] \to \infty,
\]
which by (9.13) is equivalent to
\[
\sum_{i=1}^{n} (1 - e^{-\mu z_i}) \to \infty, \quad (9.25)
\]
that is, the expected number of deaths grows to infinity. Thus assume that 
(9.25) is satisfied. In the following the relationships (9.11) – (9.16) will be used 
frequently without explicit mentioning.

First, to prove consistency, use (9.13) and (9.16) to write

\[
\hat{\mu} - \mu = \frac{\sum_{i=1}^{n} (D_i - \mu T_i)}{\sum_{i=1}^{n} T_i}.
\]

(9.26)

The first factor in (9.26) has expected value 0 and variance

\[
\frac{1}{\sum_{i=1}^{n} \mathbb{V}[D_i - \mu T_i]} = \frac{1}{\sum_{i=1}^{n} (1 - e^{-\mu z_i})},
\]

which tends to 0 as \( n \) increases. Therefore, this factor tends to 0 in probability.
The second factor in (9.26) is the inverse of \( \sum_{i=1}^{n} T_i / \mu \sum_{i=1}^{n} \mathbb{E}[T_i] \), which has 
expected value \( 1/\mu \) and variance equal to \( 1/\mu^2 \) times

\[
\frac{\sum_{i=1}^{n} \mathbb{V}[T_i]}{(\sum_{i=1}^{n} (1 - e^{-\mu z_i}))^2} = \frac{\sum_{i=1}^{n} (1 - 2\mu z_i e^{-\mu z_i} - e^{-2\mu z_i})}{(\sum_{i=1}^{n} (1 - e^{-\mu z_i}))^2} = \frac{\sum_{i=1}^{n} a(\mu z_i)(1 - e^{-\mu z_i})}{\sum_{i=1}^{n} (1 - e^{-\mu z_i})} = \frac{1}{\sum_{i=1}^{n} (1 - e^{-\mu z_i})},
\]

(9.27)

where \( a \) is defined as

\[
a(t) = \frac{1 - 2te^{-t} - e^{-2t}}{1 - e^{-t}}, \quad t \geq 0.
\]

The function \( a \) is bounded since it is continuous and tends to 0 as \( t \downarrow 0 \) and to
1 as \( t \uparrow \infty \). The first factor in (9.27) is bounded since it is a weighted average 
of values of \( a \), and the second factor tends to 0 by assumption. It follows that 
the expression in (9.27) tends to 0 and, consequently, that the second factor 
in (9.26) converges in probability to \( \mu \). It can be concluded that the whole 
expression in (9.26) converges in probability to 0, so that \( \hat{\mu} \) is weakly consistent;

\[
\hat{\mu} \xrightarrow{p} \mu.
\]

Next, to prove asymptotic normality, look at

\[
\sqrt{\sum_{i=1}^{n} (1 - e^{-\mu z_i})(\hat{\mu} - \mu)} = \frac{\sum_{i=1}^{n} (D_i - \mu T_i)}{\sqrt{\sum_{i=1}^{n} (1 - e^{-\mu z_i})}} \left( \frac{\sum_{i=1}^{n} T_i}{\mu \sum_{i=1}^{n} \mathbb{E}[T_i]} \right)^{-1}.
\]

In the presence of (9.25) the first factor on the right converges in distribution 
to a standard normal variate. (Verify that the Lindeberg condition is satisfied, 
see Appendix D.)
The second factor has just been proved to converge in probability to $\mu$. It follows that
\[ \hat{\mu} \sim_{\text{as}} N \left( \mu, \frac{\mu^2}{\sum_{i=1}^{n}(1 - e^{-\mu z_i})} \right). \]  
(9.28)
Likewise, it also holds that $\hat{\nu}$ is weakly consistent, and
\[ \hat{\nu} \sim_{\text{as}} N \left( \nu, \frac{1}{\mu^2} \sum_{i=1}^{n}(1 - e^{-\mu z_i}) \right). \]  
(9.29)

G. Random censoring. In Paragraph E the censoring time was assumed to be the same for all individuals. Thereby the pairs $(D_i, T_i)$ became stochastic replicates, and we could invoke simple asymptotic theory for i.i.d. variates to prove strong consistency and asymptotic normality of MLE-s. In Paragraph F the censoring was allowed to vary among the individuals, but it turned out that the asymptotic results essentially remained true, although only weak consistency could be achieved. All we required was (9.25), which says that the censoring must not turn too severe so that information deteriorates in the end: there must be a certain stability in the censoring pattern so that individuals with sufficient exposure time enter the study sufficiently frequently in the long run.

One way of securing such stability is to regard the censoring times as outcomes of i.i.d. random variables. Such an assumption seems particularly apt in a non-experimental context like insurance. The censoring is not subject to planning, and the censoring times are just as random in their nature as anything else observed about the individuals.

Thus, we henceforth work with an enriched model, where the distributional assumptions in Paragraph F constitute the conditional model for given censoring times $Z_i = z_i, i = 1, 2, \ldots$, and the $Z_i$ are independent selections from some distribution function $H$ with (generalized) density $h$ independent of $\mu$. This way the triplets $(D_i, T_i, Z_i), i = 1, 2, \ldots$, become stochastic replicates, and the i.i.d. situation is restored with all its conveniences.

The likelihood of the observations now becomes
\[ \Lambda = \prod_{i=1}^{n} \mu^{D_i} e^{-\mu T_i} h(Z_i) = e^{\ln \mu N - \mu W} \prod_{i=1}^{n} h(Z_i). \]  
(9.30)
Maximization of (9.30) with respect to $\mu$ is equivalent to maximization of the likelihood (9.17) in the conditional model for fixed censoring, hence the MLE remains the same as before. Its distribution is affected by the structure now added to the model, however. It is easy to prove that the results in Paragraph E carry over to the present case, only that the expression $1 - e^{-\mu Z}$ is everywhere to be replaced by $1 - \mathbb{E} \left[ e^{-\mu Z} \right]$, where $Z \sim H$.

9.2 Parametric inference in the Markov model

A. The likelihood of a time-continuous Markov process. Consider now the general set-up, whereby the development of an insurance policy is
CHAPTER 9. INFERENCE IN THE MARKOV MODEL

represented by a continuous time Markov process $X$ on a finite state space $\mathcal{J} = \{0, 1, \ldots, J\}$. As usual, let $I_g(t)$ and $N_{gh}(t)$ denote, respectively, the indicator of the event that the process is staying in state $g$ at time $t \geq 0$, and the number of transitions from state $g$ to state $h$ in the time interval $(0, t]$. The transition intensities $\mu_{gh}$ are assumed to exist, and to be piecewise continuous.

Suppose the policy is observed continuously throughout the time period $[L, l]$, commencing in state $g_0$ at time $L$. One then speaks of left-censoring and right-censoring at times $L$ and $l$, respectively, and the triplet $z = (L, l, g_0)$ will be referred to as the censoring scheme of the policy.

Consider a specific realization of the observed part of the process:

$$X(\tau) = \begin{cases} g_0, & \frac{L}{2} < \tau < t_1, \\ g_1, & t_1 + dt_1 < \tau < t_2, \\ \vdots \\ g_{q-2}, & t_{q-2} + dt_{q-2} < \tau < t_{q-1}, \\ g_{q-1}, & t_{q-1} + dt_{q-1} < \tau < \bar{t}. \end{cases}$$

By the given censoring, the probability of this realization is as follows, where $t_0 = L$, $t_q = l$, and $\mu_g = \sum_{h \neq g} \mu_{gh}$ denotes the total intensity of transition out of state $g$:

$$\exp \left( - \int_{t_0}^{t_1} \mu_{g_0}(t) \, dt_1 \exp \left( - \int_{t_1}^{t_2} \mu_{g_1}(t) \, dt_2 \right) \right) \exp \left( - \int_{t_{q-2}}^{t_{q-1}} \mu_{g_{q-2}}(t) \, dt_{q-2} \exp \left( - \int_{t_{q-1}}^{t_q} \mu_{g_{q-1}}(t) \, dt_{q-1} \right) \right)$$

$$= \prod_{p=1}^{q-1} \mu_{g_{p-1}g_p}(t_p) \exp \left( - \sum_{p=1}^{q-1} \int_{t_{p-1}}^{t_p} \mu_{g_{p-1}}(t) \, dt_{p-1} \right)$$

$$= \exp \left( \sum_{p=1}^{q-1} \ln \mu_{g_{p-1}g_p}(t_p) - \sum_{p=1}^{q-1} \int_{t_{p-1}}^{t_p} \mu_{g_{p-1}}(t) \, dt_{p-1} \right) \, dt_1 \ldots dt_{q-1}.$$ 

It follows that the likelihood of the observables is

$$\Lambda = \exp \left( \sum_{g \neq h} \int_{L}^{l} \ln \mu_{gh}(\tau) \, dN_{gh}(\tau) - \sum_{g} \int_{L}^{l} \mu_{g}(\tau) I_g(\tau) \, d\tau \right)$$

$$= \exp \left( \sum_{g \neq h} \int_{L}^{l} \{\ln \mu_{gh}(\tau) \, dN_{gh}(\tau) - \mu_{gh}(\tau) I_g(\tau) \, d\tau \} \right). \quad (9.31)$$

B. ML estimation of parametric intensities. Now consider a parametric model where the intensities are of the form $\mu_{gh}(t, \theta)$, with $\theta = (\theta_1, \ldots, \theta_s)'$ varying in an open set in the $s$-dimensional euclidean space, $s < \infty$. We assume they are twice continuously differentiable functions of $\theta$. 
CHAPTER 9. INFERENCE IN THE MARKOV MODEL

Suppose that inference is to be made about the intensities or, equivalently, the parameter $\theta$ on the basis of data from a sample of $n$ similar policies. Equip all quantities related to the $m$-th policy by topscript $(m)$. The processes $X^{(m)}$ are assumed to be stochastically independent replicates of the process $X$ described above, but their censoring schemes $z^{(m)}$ may be different.

By independence, the likelihood of the whole data set is the product of the individual likelihoods: $\Lambda = \prod_{m=1}^{n} \Lambda^{(m)}$. Thus, by (9.31),

$$
\ln \Lambda = \sum_{g \neq h} \int \{ \ln \mu_{gh}(\tau, \theta) dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau \}, 
$$

(9.32)

with

$$
N_{gh} = \sum_{m=1}^{n} N_{gh}^{(m)}, \quad I_g = \sum_{m=1}^{n} I^{(m)}_g.
$$

The censoring schemes are not visualized in (9.32), and they need not be if, as a matter of definition, $dN_{gh}^{(m)}(t)$ and $I^{(m)}_g(t)$ are taken as 0 for $t \notin [\bar{t}_g^{(m)}, \bar{t}_g^{(m)}]$. Likewise, introduce

$$
p^{(m)}_g(t) = p_{g0}^{(m)}(t|\xi^{(m)}, t)1_{[\xi^{(m)}, \bar{t}^{(m)}]}(t),
$$

the probability that the censored process $X^{(m)}$ stays in $g$ at time $t$, by definition taken as 0 for $t \notin [\bar{t}_g^{(m)}, \bar{t}_g^{(m)}]$.

In the MLE construction we need the derivatives of (9.32), of first order (an $s$-vector),

$$
\frac{\partial}{\partial \theta} \ln \Lambda = \sum_{g \neq h} \int \frac{\partial}{\partial \theta} \ln \mu_{gh}(\tau, \theta) \{ dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau \},
$$

(9.33)

and of second order (an $s \times s$ matrix),

$$
\frac{\partial^2}{\partial \theta \partial \theta'} \ln \Lambda = \sum_{g \neq h} \int \left\{ \frac{\partial^2}{\partial \theta \partial \theta'} \ln \mu_{gh}(\tau, \theta) \{ dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau \} 

- \frac{\partial}{\partial \theta} \ln \mu_{gh}(\tau, \theta) \frac{\partial}{\partial \theta'} \mu_{gh}(\tau, \theta) I_g(\tau) \right\}
$$

(9.34)

By (9.33) the MLE $\hat{\theta}$ is the solution of

$$
\sum_{g \neq h} \int \frac{\partial}{\partial \theta} \ln \mu_{gh}(\tau, \theta) \{ dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau \} \bigg|_{\theta = \hat{\theta}} = 0^{s \times 1}.
$$

(9.35)

Referring to Appendix D, the large sample distribution properties of the MLE are given by

$$
\hat{\theta} \sim_{as} N(\theta, \Sigma(\theta)),
$$

(9.36)
where $\Sigma(\theta)$ is given by its inverse, the so-called information matrix,

$$
\Sigma(\theta)^{-1} = -\mathbb{E} \left[ \frac{\partial}{\partial \theta \partial \theta'} \ln \Lambda \right].
$$

Taking expectation in (9.34), noting that the terms $dN_{gh}(\tau) - \mu_{gh}(\tau, \theta)I_g(\tau)d\tau$ have zero means, we obtain

$$
\Sigma(\theta)^{-1} = \sum_{g \neq h} \int \frac{1}{\mu_{gh}(\tau, \theta)} \left( \frac{\partial}{\partial \theta} \mu_{gh}(\tau, \theta) \cdot \frac{\partial}{\partial \theta} \mu_{gh}(\tau, \theta) \right) \sum_{m=1}^{n} p_g^{(m)}(\tau, \theta)d\tau.
$$

The expression in parentheses under the integral sign is an $s \times s$ matrix and all other quantities are scalar.

It is seen that the information matrix tends to infinity, hence the variance matrix of the MLE tends to 0, if the terms $\sum_{m=1}^{n} p_g^{(m)}(\tau, \theta)$ grow to infinity as $n$ increases, roughly speaking, which means that the expected number of individuals exposed to risk in different states gets unlimited.


In a mortality study the relevant state space is $J = \{0, 1\}$ (“alive” and “dead”). Assume the mortality intensity is of G-M form,

$$
\mu(\tau, \theta) = \alpha + \beta \tau,
$$

with

$$
\theta = (\alpha, \beta, c)'.
$$

To estimate $\theta$, form the derivatives of first order,

$$
\frac{\partial}{\partial \theta} \mu(\tau, \theta) = \begin{pmatrix}
\frac{\partial}{\partial \alpha} \mu(\tau, \theta) \\
\frac{\partial}{\partial \beta} \mu(\tau, \theta) \\
\frac{\partial}{\partial c} \mu(\tau, \theta)
\end{pmatrix} = \begin{pmatrix}
1 \\
\beta \tau \\
\beta c \tau^{-1}
\end{pmatrix}.
$$

For each individual $i$ participating in the study, denote by $D_i$ the number of deaths during the study (0 or 1), and by $T_i$ the age at departure from the study (the age at death if $D_i = 1$). The MLE is the solution of the set of equations

$$
\sum_{i, D_i = 1} \frac{1}{\hat{\alpha} + \hat{\beta} \hat{c} T_i} = \sum_i T_i, 
$$

$$
\sum_{i, D_i = 1} \hat{c}^{T_i} = \sum_i \hat{c}^{T_i} - 1 \ln \hat{c},
$$

$$
\sum_{i, D_i = 1} \frac{\hat{c}^{T_i}}{\hat{\alpha} + \hat{\beta} \hat{c} T_i} = \sum_i \left( \frac{\hat{c}^{T_i} T_i}{\ln \hat{c}} - \frac{\hat{c}^{T_i} - 1}{(\ln \hat{c})^2} \right).
$$

(9.39)  (9.40)  (9.41)
To find the information matrix (9.38) we need
\[
\frac{\partial}{\partial \theta} \mu(\tau, \theta) \frac{\partial}{\partial \theta} \mu(\tau, \theta) = \begin{pmatrix}
1 & c\tau & \beta \tau c^{\tau - 1} \\
\cdot & c^2\tau & \beta \tau c^{\tau - 1} \\
\cdot & \cdot & \beta^2 \tau c^{\tau(\tau - 1)}
\end{pmatrix}
\]
(9.42)
symmetric) and the probabilities \(p_0^{(m)}(\tau, \theta)\). The latter depend on the observational scheme. Consider the simple case where the individuals are observed from birth until age \(z\) or death, whichever occurs first. Then
\[
p_0^{(m)}(\tau, \theta) = \exp \left( -\int_0^\tau (\alpha + \beta c^s) ds \right)
= \exp \left( -\alpha \tau - \beta c^\tau - \frac{1}{\ln c} \right),
\]
(9.43)
(the survival probability) for \(\tau \leq z\) and 0 for \(\tau > z\). (There is only one kind of transition, from 0 to 1, and the summation over \(g, h\) in the information matrix can be dropped.)

We see that all ingredients in the asymptotic variance matrix are given by explicit formulas, and it remains only to perform a numerical integration to find its value for given \(\theta\). □

9.3 Confidence regions

A. An asymptotic confidence ellipsoid. From the asymptotic normality of the MLE it follows that
\[
(\hat{\theta} - \theta)^\top \Sigma^{-1}(\theta)(\hat{\theta} - \theta) \sim_{\text{as}} \chi^2_s,
\]
the chi-squared distribution with \(s\) degrees of freedom. Therefore, denoting the \((1 - \varepsilon)\)-fractile of this distribution by \(\chi^2_s,1-\varepsilon\), an asymptotic \(1 - \varepsilon\) confidence region is the set of all \(\theta\) satisfying
\[
(\theta - \hat{\theta})^\top \Sigma^{-1}(\theta)(\theta - \hat{\theta}) \leq \chi^2_s,1-\varepsilon.
\]
(9.45)
The expression on the left here will typically be a complicated function of \(\theta\), and it is in general not easy to find the values of \(\theta\) that satisfy the inequality and constitute a confidence region. Now, suppose \(\Sigma(\theta)\) can be estimated by some function of the data, \(\hat{\Sigma}\), and that the estimator is consistent in the sense that
\[
\hat{\Sigma} \Sigma^{-1}(\theta) \to I.
\]
(9.46)
Then it is easy to show that also the relation
\[
(\theta - \hat{\theta})^\top \hat{\Sigma}^{-1}(\theta - \hat{\theta}) \leq \chi^2_s,1-\varepsilon
\]
(9.47)
determines an asymptotic \(1 - \varepsilon\) confidence region. The relation (9.47) defines an ellipsoid, which is a fairly simple geometric figure and, as we shall see in the
following paragraph, a convenient basis for deriving other confidence statements of interest.

A straightforward way of constructing \( \hat{\Sigma} \) would be to replace \( \theta \) in \( \Sigma(\theta) \) by the consistent estimator \( \hat{\theta} \), that is, put

\[
\hat{\Sigma} = \Sigma(\hat{\theta}).
\]

This works well if the entries in \( \Sigma(\theta) \) are closed expressions in \( \theta \). Unfortunately, this is the case only in certain simple situations, typically when the state space \( J \) is small and the pattern of transitions is hierarchical. One example is the mortality study with parametric mortality law, e.g. of G-M type. In more complex situations we cannot in general find closed formulas for the probabilities \( p_g^{(m)}(\tau) \) involved in \( \Sigma \), even if the intensities themselves are simple parametric functions. Then a different construction is required. A simple device is to replace the \( p_g^{(m)}(\tau) \) by their empirical counterparts \( \hat{I}_g^{(m)}(\tau) \) and put

\[
\sum_{m=1}^{n} p_g^{(m)}(\tau) \approx \hat{I}_g(\tau). \quad (9.48)
\]

**B. Simultaneous confidence intervals.** The confidence ellipsoid (9.47) can be resolved in simultaneous confidence intervals for all linear functions of \( \theta \) in the following way. The Schwarz inequality says that for all vectors \( a \) and \( x \) in \( \mathbb{R}^s \),

\[
|a'x| \leq \sqrt{a'a \chi^2_s},
\]

with equality for \( a = cx \). Thus, noting that the quadratic form on the left of (9.47) is \( (\Sigma^{-1/2}(\theta - \hat{\theta}))' (\Sigma^{-1/2}(\theta - \hat{\theta})) \), the confidence statement can be cast equivalently as

\[
|a'\Sigma^{-1/2}(\theta - \hat{\theta})| \leq \sqrt{a'a \chi^2_{s,1-\epsilon}}, \forall a. \quad (9.49)
\]

Since \( \hat{\Sigma} \) is of full rank, the vector \( \Sigma^{-1/2}a \) ranges through all of \( \mathbb{R}^s \) as \( a \) ranges in \( \mathbb{R}^s \). Thus, writing \( a'a = (\Sigma^{-1/2}a)'(\Sigma^{-1/2}a) \), (9.49) is equivalent to

\[
|a'(\theta - \hat{\theta})| \leq \sqrt{a'a \chi^2_{s,1-\epsilon}}, \forall a,
\]

that is,

\[
a'\theta \in [a'\hat{\theta} - \sqrt{\chi^2_{s,1-\epsilon}a'\hat{\Sigma}a}, a'\hat{\theta} + \sqrt{\chi^2_{s,1-\epsilon}a'\hat{\Sigma}a}], \forall a. \quad (9.50)
\]

The intervals in (9.50) are (asymptotic) simultaneous confidence intervals for all linear functions of \( \theta \) in the sense that the probability is at least \( 1 - \epsilon \) that they all hold true.

**C. Confidence band for the G-M mortality intensity.** Returning to the mortality study example in Paragraph 9.2.C, let \( c \) be taken as known so that the mortality intensity is a linear function of the unknown parameter \( \theta = (\alpha, \beta)' \). The MLE is obtained by solving the equations (9.39) and (9.40), and
the appropriate variance matrix $\Sigma$ is obtained by inverting the upper left $2 \times 2$ block in the information matrix defined by (9.38), (9.42), and (9.43).

From (9.50) we obtain simultaneous confidence intervals for all $\mu(\tau) = \alpha + \beta c^\tau$, constituting a confidence band in the space of mortality intensity functions;

$$\mu(\tau) \in [\hat{\alpha} + \hat{\beta} c^\tau - \sqrt{\chi^2_{2(1-\epsilon)} \hat{\sigma}^2}, \hat{\alpha} + \hat{\beta} c^\tau - \sqrt{\chi^2_{2(1-\epsilon)} \hat{\sigma}^2}], \forall \tau > 0,$$

(9.51)

where

$$\hat{\sigma} = (1, c^\tau) \hat{\Sigma} \left( \begin{array}{c} 1 \\ c^\tau \end{array} \right).$$

### 9.4 Piecewise constant intensities

#### A. Piecewise constant intensities.

Let $0 = t_0 < t_1 < \cdots < t_r = \bar{\tau}$ be some finite partition of the time interval $[0, \bar{\tau}]$, and assume that the intensities are step functions of the form

$$\mu_{gh}(\tau) = \mu_{gh,q}, \tau \in [t_{q-1}, t_q), q = 1, \ldots, r,$$

$$= \sum_{q=1}^{r} 1_{(t_{q-1}, t_q)}(\tau) \mu_{gh,q},$$

(9.52)

where the $\mu_{gh,q}$ take values in $(0, \infty)$, with no relationships between them. The situation fits into the general framework with $\theta = (\ldots, \mu_{gh,q}, \ldots)'$, a vector of (typically high) dimension $J \times J \times r$.

#### B. The MLE estimators are O-E rates.

The log likelihood in (9.32) now becomes

$$\ln \Lambda = \sum_{g \neq h} \sum_{q=1}^{r} \left\{ \ln \mu_{gh,q} N_{gh,q} - \mu_{gh,q} W_{g,q} \right\},$$

(9.53)

where

$$N_{gh,q} = \int_{t_{q-1}}^{t_q} dN_{gh}(\tau),$$

(9.54)

$$W_{g,q} = \int_{t_{q-1}}^{t_q} I_g(\tau) d\tau,$$

(9.55)

are, respectively, the total number of transitions from state $g$ to state $h$ and the total time spent in state $g$ during the age interval $[t_{q-1}, t_q)$.

Since the $\mu_{gh,q}$ are functionally unrelated, the log likelihood decomposes into terms that depend on one and only one of the basic parameters, and finding maximum amounts to maximizing each term. The derivatives involved in the ML construction now become particularly simple:

$$\frac{\partial}{\partial \mu_{gh,q}} \ln \Lambda = \frac{1}{\mu_{gh,q}} N_{gh,q} - W_{g,q},$$

(9.56)
\[
\frac{\partial^2}{\partial \mu_{gh,q} \partial \mu_{g'h',q'}} \ln \Lambda = -\delta_{gh,q,g'h',q'} \frac{1}{\mu_{gh,q}} N_{gh,q} .
\]

It follows from (9.56) that the MLE is
\[
\hat{\mu}_{gh,q} = \frac{N_{gh,q}}{W_{g,q}} ,
\]
an O-E rate of the same kind as in the simple model of 9.2.B. Noting that, by (9.54),
\[
\mathbb{E}[N_{gh,q}] = \mu_{gh,q} \int_{t_{q-1}}^{t_q} \sum_{m=1}^{n} p_g^{(m)}(\tau)d\tau ,
\]
we obtain from (9.57) that the asymptotic variance matrix becomes
\[
\Sigma(\theta) = \text{Diag} \left( \ldots , \frac{\mu_{gh,q}}{\int_{t_{q-1}}^{t_q} \sum_{m=1}^{n} p_g^{(m)}(\tau)d\tau} , \ldots \right) ,
\]
a diagonal matrix, implying that the estimators of the \( \mu_{gh,q} \) are asymptotically independent.

An estimator of \( \Sigma \) is obtained upon replacing the parameter functions appearing on the right of (9.59) by their straightforward estimators: put \( \mu_{gh,q} \approx \hat{\mu}_{gh,q} \) defined by (9.58) and, by the device (9.48),
\[
\int_{t_{q-1}}^{t_q} \sum_{m=1}^{n} p_g^{(m)}(\tau)d\tau \approx \int_{t_{q-1}}^{t_q} I_g(\tau)d\tau = W_{g,q} ,
\]
to obtain
\[
\hat{\Sigma} = \text{Diag} \left( \ldots , \frac{N_{gh,q}}{W_{g,q}^2} , \ldots \right) .
\]

**C. Smoothing O-E rates.** The MLE of the intensity function is obtained upon inserting the estimators (9.58) in (9.52). The resulting function will typically have a ragged appearance due to the estimation error in a finite sample. This is unsatisfactory since the intensities are expected to be smooth functions: for instance, there are a priori reasons to assume that the mortality intensity is a continuous and non-decreasing function of the age. Now, the very assumption of piecewise constant intensities is artificial, of course, and the estimates obtained under this assumption cannot serve as an ultimate answer in practice. In fact, they represent only the first step in a two-stage procedure, where the second step is to fit some smooth functions to the raw estimates delivered by the O-E rates. The functions used for fitting constitute the model we have in mind. It may be objected that the two-stage procedure is a detour since, if the intensities are assumed to be functions of a smaller set of parameters, one could follow the prescription in Section 9.2 and maximize the likelihood directly. There are two reasons why the two-stage procedure never the less merits special treatment: in
the first place, the O-E rates and their asymptotic variance matrix are easy to construct; in the second place, a comparative plot of the fitted functions and the O-E rates makes it possible to detect systematic deviations between model assumptions and facts.

A commonly used fitting technique is the so-called generalized least squares method, which amounts to minimizing a positive definite quadratic form in the deviations between the raw estimates and the fitting functions. In the following brief outline of the procedure we focus on one given intensity and drop the subscripts \(g, h\).

For each interval \([t_{q-1}, t_q)\) choose a "representative" point \(\tau_q\), e.g. the interval midpoint. Put \(\hat{\mu} = (\ldots, \hat{\mu}_q, \ldots)'\), the vector of O-E rates, and (with a bit sloppy notation) \(\mu(\theta) = (\ldots, \mu(\tau_q, \theta), \ldots)'\), the vector of true values. Let \(A = (a_{pq})\) be some positive definite matrix of order \(r \times r\). Estimate \(\theta\) by \(\theta^*\) minimizing

\[
(\mu(\theta) - \hat{\mu})' A (\mu(\theta) - \hat{\mu}) = \sum_{pq} a_{pq} (\mu(\tau_p, \theta) - \hat{\mu}_p)(\mu(\tau_q, \theta) - \hat{\mu}_q).
\]

If the intensity is a linear function of \(\theta\) (like in the G-M study with known \(c\)),

\[
\mu(\theta) = Y(\tau) \theta,
\]

then

\[
\hat{\theta} = (Y' AY)^{-1} Y' A \hat{\mu}.
\]

The asymptotic variance of \(\hat{\theta}\) is \((Y' AY)^{-1} Y' A \Sigma(\theta) A Y (Y' A Y)^{-1}\). By the Gauss-Markov theorem it is minimized by taking \(A = \Sigma(\theta)^{-1}\), and the minimum is \((Y' A \Sigma(\theta)^{-1} Y)^{-1}\). Thus, asymptotically the best choice of \(A\) is \(\hat{\Sigma}^{-1}\), where \(\hat{\Sigma}\) is some estimate of \(\Sigma\) satisfying (9.46).

### 9.5 Impact of the censoring scheme

#### A. The precision of the estimation.

The precision of the MLE depends on the amount of information provided by the censoring scheme of the study. Asymptotically it is the variance matrix \(\Sigma(\theta)\) that determines everything, and in Section 9.2 it was pointed out that the size of this matrix depends on the censoring scheme only through the functions \(\sum_{m=1}^n p_g^{(m)}(\tau)\), \(g = 0, \ldots, J\), the expected numbers of individuals staying in each state \(g\) at time \(\tau\). (It depends also on the parametric structure of the intensities, of course.) We shall look at two censoring schemes frequently encountered in practice.

#### B. Longitudinal observation (cohort studies).

The term cohort stems from Latin and originally signified a unit division in an ancient Roman legion. In demography it means a class of individuals born in a particular year or more general period of time (a "generation"), and a cohort study is one where a cohort is observed over a certain period, possibly until it is extinct. This was the situation in Paragraph 9.2.C.
Thus, let the \( n \) Markov processes in the general set-up be stochastic replicates, all commencing in state 0 at time 0 and thereafter observed continuously throughout the time interval \([0, \bar{t}]\). In this case

\[
\sum_{m=1}^{n} p_{g}^{(m)}(\tau) = n p_{g}^{(1)}(\tau), \quad g = 0, \ldots, J,
\]

and

\[
\Sigma(\theta) = \frac{1}{n} \left( \sum_{g \neq h} \int_{0}^{\bar{t}} \frac{1}{\mu(\tau, \theta)} \frac{\partial}{\partial \theta} \mu(\tau, \theta) \frac{\partial}{\partial \theta'} \mu(\tau, \theta) p_{g}^{(1)}(\tau) d\tau \right)^{-1}.
\]

This matrix tends to 0 as \( n \) increases if the inverse matrix indicated exists.

C. Cross-sectional observation. In a cross-sectional study a population is observed over a certain period of time. As an example, suppose the G-M mortality study in Paragraph 9.2.C is conducted cross-sectionally throughout a calendar period of duration \( \bar{t} \), and that it comprises \( n \) individuals at ages \( t(m) \), \( m = 1, \ldots, n \), at the beginning of the study. In this case the factor depending on the design in the information matrix is

\[
\sum_{m=1}^{n} p^{(m)}(\tau) = \sum_{m=1}^{n} 1_{[t(m), t(m)+\bar{t}]}(\tau) \exp \left( -\alpha(\tau - t(m)) - \beta \frac{t(m)(c^{\tau} - 1)}{\ln c} \right).
\]
Bibliography


Appendix A

Calculus

A. Piecewise differentiable functions. Being concerned with operations in time, commencing at some initial date, we will consider functions defined on the positive real line \([0, \infty)\). Thus, let us consider a generic function \(X = \{X_t\}_{t \geq 0}\) and think of \(X_t\) as the state or value of some process at time \(t\). For the time being we take \(X\) to be real-valued.

In the present text we will work exclusively in the space of so-called piecewise differentiable functions. From a mathematical point of view this space is tiny since only elementary calculus is needed to move about in it. From a practical point of view it is huge since it comfortably accommodates any idea, however sophisticated, that an actuary may wish to express and analyse. It is convenient to enter this space from the outside, starting from a wider class of functions.

We first take \(X\) to be of finite variation (FV), which means that it is the difference between two non-decreasing, finite-valued functions. Then the left-limit \(X_t^- = \lim_{s \uparrow t} X_s\) and the right-limit \(X_t^+ = \lim_{s \downarrow t} X_s\) exist for all \(t\), and they differ on at most a countable set \(D(X)\) of discontinuity points of \(X\).

We are particularly interested in FV functions \(X\) that are right-continuous (RC), that is, \(X_t = \lim_{s \downarrow t} X_s\) for all \(t\). Any probability distribution function is of this type, and any stream of payments accounted as incomes or outgoes, can reasonably be taken to be FV and, as a convention, RC. If \(X\) is RC, then \(\Delta X_t = X_t - X_{t^-}\), when different from 0, is the jump made by \(X\) at time \(t\).

For our purposes it suffices to let \(X\) be of the form

\[
X_t = X_0 + \int_0^t x_\tau \, d\tau + \sum_{0 < \tau \leq t} (X_\tau - X_{\tau^-}).
\]  

(A.1)

The integral, which may be taken to be of Riemann type, adds up the continuous increments/decrements, and the sum, which is understood to range over discontinuity times, adds up increments/decrements by jumps.

We assume, furthermore, that \(X\) is piecewise differentiable (PD); A property holds piecewise if it takes place everywhere except, possibly, at a finite number of points in every finite interval. In other words, the set of exceptional points,
if not empty, must be of the form \{t_0, t_1, \ldots\}, with \( t_0 < t_1 < \cdots \), and, in case it is infinite, \( \lim_{j \to \infty} t_j = \infty \). Obviously, \( X \) is PD if both \( X \) and \( x \) are piecewise continuous. At any point \( t \not\in D = D(X) \cup D(x) \) we have \( \frac{d}{dt} X_t = x_t \), that is, the function \( X \) grows (or decreases) continuously at rate \( x_t \).

As a convenient notational device we shall frequently write (A.1) in differential form as

\[
dX_t = x_t \, dt + X_t - X_{t-}.
\] (A.2)

A left-continuous PD function may be defined by letting the sum in (A.1) range only over the half-open interval \([0, t)\). Of course, a PD function may be neither right-continuous nor left-continuous, but such cases are of no interest to us.

B. The integral with respect to a function. Let \( X \) and \( Y \) both be PD and, moreover, let \( X \) be RC and given by (A.2). The integral over \((s, t]\) of \( Y \) with respect to \( X \) is defined as

\[
\int_{s}^{t} Y_{\tau} \, dX_{\tau} = \int_{s}^{t} Y_{\tau} x_{\tau} \, d\tau + \sum_{s < \tau \leq t} Y_{\tau}(X_{\tau} - X_{\tau-}),
\] (A.3)

provided that the individual terms on the right and also their sum are well defined. Considered as a function of \( t \) the integral is itself PD and RC with continuous increments \( Y_t x_t \, dt \) and jumps \( Y_t(X_t - X_{t-}) \). One may think of the integral as the weighted sum of the \( Y \)-values, with the increments of \( X \) as weights, or vice versa. In particular, (A.1) can be written simply as

\[
X_t = X_s + \int_{s}^{t} dX_{\tau},
\] (A.4)

saying that the value of \( X \) at time \( t \) is its value at time \( s \) plus all its increments in \((s, t]\).

By definition,

\[
\int_{s}^{t-} Y_{\tau} \, dX_{\tau} = \lim_{r \to t} \int_{s}^{r} Y_{\tau} \, dX_{\tau} = \int_{s}^{t} Y_{\tau} \, dX_{\tau} - Y_{t}(X_{t} - X_{t-}) = \int_{(s,t]} Y_{\tau} \, dX_{\tau},
\]

a left-continuous function of \( t \). Likewise,

\[
\int_{s-}^{t} Y_{\tau} \, dX_{\tau} = \lim_{r \to s} \int_{r}^{t} Y_{\tau} \, dX_{\tau} = \int_{s}^{t} Y_{\tau} \, dX_{\tau} + Y_{s}(X_{s} - X_{s-}) = \int_{[s,t]} Y_{\tau} \, dX_{\tau},
\]

a left-continuous function of \( s \).

C. The chain rule (Itô’s formula). Let \( X_t = (X_t^1, \ldots, X_t^m) \) be an \( m \)-variate function with PD and RC components given by \( dX_t^i = x_t^i \, dt + (X_t^i - X_{t-}^i) \). Let \( f : \mathbb{R}^m \mapsto \mathbb{R} \) have continuous partial derivatives, and form the composed
function $f(X_t)$. On the open intervals where there are neither discontinuities in the $x^i$ nor jumps of the $X^i$, the function $f(X_t)$ develops in accordance with the well-known chain rule for scalar fields along rectifiable curves. At the exceptional points $f(X_t)$ may change (only) due to jumps of the $X^i$, and at any such point $t$ it jumps by $f(X_t) - f(X_{t-})$. Thus, we gather the so-called change of variable rule or Itô’s formula, which in our simple function space reads

$$df(X_t) = \sum_{i=1}^m \frac{\partial f}{\partial x^i}(X_t) x^i_t \, dt + f(X_t) - f(X_{t-}), \quad (A.5)$$

or, in integral form,

$$f(X_t) = f(X_s) + \int_s^t \sum_{i=1}^m \frac{\partial f}{\partial x^i}(X_{\tau}) x^i_{\tau} \, d\tau + \sum_{s < \tau \leq t} \{f(X_{\tau}) - f(X_{\tau-})\}. \quad (A.6)$$

Obviously, $f(X_t)$ is PD and RC.

A frequently used special case is (check the formulas!)

$$d(X_t Y_t) = X_t y_t \, dt + Y_t x_t \, dt + X_t Y_t - X_{t-} Y_{t-}$$

$$= X_{t-} \, dY_t + Y_{t-} \, dX_t + (X_t - X_{t-})(Y_t - Y_{t-})$$

$$= X_{t-} \, dY_t + Y_t \, dX_t. \quad (A.7)$$

If $X$ and $Y$ have no common jumps, as is certainly the case if one of them is continuous, then (A.7) reduces to the familiar

$$d(X_t Y_t) = X_t \, dY_t + Y_t \, dX_t. \quad (A.8)$$

The integral form of (A.7) is the so-called rule of integration by parts:

$$\int_s^t Y_{\tau} \, dX_{\tau} = Y_t X_t - Y_s X_s - \int_s^t X_{\tau-} \, dY_{\tau}. \quad (A.9)$$

Let us consider three special cases for which (A.9) can be obtained by direct calculation and specialises to well-known formulas. Setting $s = 0$ (just a matter of notation), (A.9) can be cast as

$$Y_t X_t = Y_0 X_0 + \int_0^t Y_{\tau} \, dX_{\tau} + \int_0^t X_{\tau-} \, dY_{\tau}, \quad (A.10)$$

which shows how the product of $X$ and $Y$ at time $t$ emerges from its initial value at time 0 plus all its increments in the interval $(0, t]$.

Assume first that $X$ and $Y$ are both discrete. For notational simplicity assume $X_t = \sum_{j=0}^{[t]} x_j$ and $Y_t = \sum_{j=0}^{[t]} y_j$. Then

$$X_t Y_t = \sum_{i=0}^{[t]} x_i \sum_{j=0}^{[t]} y_j$$
\[ \sum_{i=0}^{[t]} \sum_{j=0}^{i} y_j x_i + \sum_{j=1}^{[t]} \sum_{i=0}^{j-1} x_i y_j = x_0 y_0 + \sum_{i=0}^{[t]} \sum_{j=0}^{i} y_j x_i + \sum_{j=1}^{[t]} \sum_{i=0}^{j-1} x_i y_j = X_0 Y_0 + \sum_{i=1}^{[t]} Y_i x_i + \sum_{j=1}^{[t]} X_{j-1} y_j = X_0 Y_0 + \int_0^t Y_\tau \, dX_\tau + \int_0^t Y_\tau \, dX_\tau, \]

which is (A.10). We see here that the left limit on the right of (A.10) is essential. This case is basically nothing but the rule of changing the order of summation in a double sum, the only new thing being that we formally consider the sums \( X \) and \( Y \) as functions of a continuous time index; only the values at integer times matter, however.

Assume next that \( X \) and \( Y \) are both continuous, that is,

\[ X_t = X_0 + \int_0^t x_\tau \, d\tau, \quad Y_t = Y_0 + \int_0^t y_\tau \, d\tau. \]

Take \( X_0 = Y_0 = 0 \) for the time being. Then

\[
X_t Y_t = \int_0^t x_\sigma \, d\sigma \int_0^t y_\tau \, d\tau \\
= \int \int_0^{\sigma < \tau \leq t} y_\tau \, d\tau \, x_\sigma \, d\sigma + \int \int_0^{0 < \sigma < \tau \leq t} x_\sigma \, d\sigma \, y_\tau \, d\tau \\
= \int_0^t \int_0^\tau y_\tau \, d\tau \, x_\sigma \, d\sigma + \int_0^t \int_0^{\tau -} x_\sigma \, d\sigma \, y_\tau \, d\tau \\
= \int_0^t Y_\sigma \, x_\sigma \, d\sigma + \int_0^t X_{\tau -} \, y_\tau \, d\tau \\
= \int_0^t Y_\tau \, dX_\tau + \int_0^t X_\sigma \, dY_\sigma, 
\]

which also conforms with (A.10): the left limit in the next to last line disappeared since an integral with respect to \( dt \) remains unchanged if we change the integrand at a countable set of points. The result for general \( X_0 \) and \( Y_0 \) is obtained by applying the formula above to \( X_t - X_0 \) and \( Y_t - Y_0 \).

Finally, let one function be discrete and the other continuous, e.g. \( X_t = \sum_{j=0}^{[t]} x_j \) and \( Y_t = Y_0 + \int_0^t y_\tau \, d\tau \). Introduce

\[ \hat{y}_0 = Y_0, \quad \hat{y}_j = \int_{j-1}^j y_\tau \, d\tau, \quad j = 1, \ldots, [t]. \]
We have $X_t = X_{[t]}$ and

$$X_t Y_t = X_{[t]} Y_{[t]} + X_{[t]} (Y_t - Y_{[t]}) = \sum_{i=0}^{[t]} x_i \sum_{j=0}^{[t]} \tilde{y}_j + X_{[t]} \int_{[t]}^t y_\tau \, d\tau . \quad (A.11)$$

Upon applying our first result for two discrete functions, the first term in (A.11) becomes

$$X_0 Y_0 + \sum_{i=1}^{[t]} \sum_{j=0}^{[t]} \tilde{y}_j x_i + \sum_{j=1}^{[t]} \sum_{i=0}^{j-1} x_i \tilde{y}_j = X_0 Y_0 + \sum_{i=1}^{[t]} Y_j x_i + \sum_{j=1}^{[t]} X_{j-1} \int_{j-1}^j y_\tau \, d\tau$$

$$= X_0 Y_0 + \int_0^t Y_\tau \, dX_\tau + \int_0^t X_{\tau-} \, dY_\tau$$

The second term in (A.11) is $\int_{[t]}^t X_{\tau-} \, dY_\tau$. Thus, also in this case we arrive at (A.10). Again the left-limit is irrelevant since $dY_t = y_t \, dt$.

The general formula now follows from these three special cases by the fact that the the integral is a linear operator with respect to the integrand and the integrator.

**D. Counting processes.** Let $t_1 < t_2 < \cdots$ be a sequence in $(0, \infty)$, either finite or, if infinite, such that $\lim_{j \to \infty} t_j = \infty$. Think of $t_j$ as the $j$-th time of occurrence of a certain event. The number of events occurring within a given time $t$ is $N_t = \# \{ j; t_j \leq t \}$ or, putting $t_0 = 0$, $N_t = j$ for $t_j \leq t < t_{j+1}$. The function $N = \{ N_t \}_{t \geq 0}$ thus defined is called a *counting function* since it currently counts the number of occurred events. It is a particularly simple PD and RC function commencing from $N_0 = 0$ and thereafter increasing only by jumps of size 1 at the epochs $t_j$, $j = 1, 2, \ldots$

The change of variable rule (A.6) becomes particularly simple when $X$ is a counting function. In fact, for $f : \mathbb{R} \to \mathbb{R}$ and for $N$ defined above,

$$f(N_t) = f(N_s) + \sum_{s < \tau \leq t} \{ f(N_\tau) - f(N_{\tau-}) \} \quad (A.12)$$

$$= f(N_s) + \sum_{s < \tau \leq t} \{ f(N_{\tau-} + 1) - f(N_{\tau-}) \} (N_\tau - N_{\tau-}) \quad (A.13)$$

$$= f(N_s) + \int_s^t \{ f(N_{\tau-} + 1) - f(N_{\tau-}) \} \, dN_\tau . \quad (A.14)$$

Basically, what these expressions state, is just the fact that

$$f(j) = f(0) + \sum_{i=1}^j (f(i) - f(i-1)) .$$
Still they will prove to be useful representations when we come to stochastic counting processes.

Going back to the general PD and RC function $X$ in (A.1), we can associate with it a counting function $N$ defined by $N_t = \sharp \{ \tau \in (0, t]; X_\tau \neq X_{\tau-} \}$, the number of discontinuities of $X$ within time $t$. Equipped with our notion of integral, we can now express $X$ as

$$dX_t = x^c_t \, dt + x^d_t \, dN_t, \quad (A.15)$$

where $x^c_t = x_t$ is the instantaneous rate of continuous change and $x^d_t = X_t - X_{t-}$ is the size of the jump, if any, at $t$. Generalizing (A.14), we have

$$f(X_t) = f(X_s) + \int_s^t \frac{d}{dx} f(X_\tau) \, x^c_\tau \, d\tau + \int_s^t \{ f(X_{\tau-} + x^d_\tau) - f(X_{\tau-}) \} \, dN_\tau. \quad (A.16)$$
Appendix B

Indicator functions

A. Indicator functions in general spaces. Let Ω be some space with generic point ω, and let A be some subset of Ω. The function $I_A : \Omega \to \{0,1\}$ defined by

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \in A^c, \end{cases}$$

is called the indicator function or just the indicator of A since it indicates by the value 1 precisely those points ω that belong to A.

Since $I_A$ assumes only the values 0 and 1, $(I_A)^p = I_A$ for any $p > 0$. Clearly, $I_\emptyset = 0$, $I_\Omega = 1$, and

$$I_{A^c} = 1 - I_A,$$

where $A^c = \Omega \setminus A$ is the complement of A.

For any two sets A and B (subsets of Ω),

$$I_{A \cap B} = I_A I_B$$

and

$$I_{A \cup B} = I_A + I_B - I_A I_B.$$  \hspace{1cm} (B.2)  \hspace{1cm} (B.3)

The last two statements are displayed here only for ease of reference. They are special cases of the following results, valid for any finite collection of sets \{A_1, \ldots, A_r\}:

$$I_{\bigcap_{j=1}^r A_j} = \prod_{j=1}^r I_{A_j},$$  \hspace{1cm} (B.4)

$$I_{\bigcup_{j=1}^r A_j} = \sum_j I_{A_j} - \sum_{j_1 < j_2} I_{A_{j_1}} I_{A_{j_2}} + \ldots + (-1)^{r-1} I_{A_1} \cdots I_{A_r}.$$  \hspace{1cm} (B.5)

The relation (B.4) is obvious. To demonstrate (B.5), we need the identity

$$\prod_{j=1}^r (a_j + b_j) = \sum_{p=0}^r \sum_{j \in p} a_{j_1} \cdots a_{j_p} b_{j_{p+1}} \cdots b_{j_r},$$  \hspace{1cm} (B.6)
where $r \backslash p$ signifies that the sum ranges over all \(^{r \backslash p}\) different ways of dividing \(\{1, \ldots, r\}\) into two disjoint subsets \(\{j_1, \ldots, j_p\}\) (\(\emptyset\) when $p = 0$) and \(\{j_{p+1}, \ldots, j_r\}\) (\(\emptyset\) when $p = r$). Combining the general relation

\[ \{\cup_\alpha A_\alpha\}^c = \cap_\alpha A_\alpha^c, \tag{B.7} \]

with (B.1) and (B.4), we find

\[ I_{\cup_{j=1}^r A_j} = 1 - I_{\cap_{j=1}^r A_j^c} = 1 - \prod_{j=1}^r (1 - I_{A_j}), \]

and arrive at (B.5) by use of (B.6).

**B. Further aspects of indicators.** The algebraic expressions in (B.4) and (B.5) apply only to the finite case. For any collection \(\{A_\alpha\}\) of sets indexed by $\alpha$ ranging in an arbitrary space, possibly uncountable,

\[ I_{\cap_\alpha A_\alpha} = \inf_\alpha I_{A_\alpha} \tag{B.8} \]

and

\[ I_{\cup_\alpha A_\alpha} = \sup_\alpha I_{A_\alpha}. \tag{B.9} \]

In fact, inf and sup are attained here, so we can write min and max. In accordance with the latter two results one may define $\sup_\alpha A_\alpha = \cup_\alpha A_\alpha$ and $\inf_\alpha A_\alpha = \cap_\alpha A_\alpha$.

The relation (B.7) rests on elementary logical operations, but also follows from $1 - \sup_\alpha I_{A_\alpha} = \inf_\alpha (1 - I_{A_\alpha})$.

The representation of sets by indicators supports the understanding of some conventions and definitions in set theory. For instance, if \(\{A_j\}_{j=1, 2, \ldots}\) is a disjoint sequence of sets, some authors write \(\sum_j A_j\) instead of \(\cup_j A_j\). This is motivated by

\[ I_{\cup_j A_j} = \sum_j I_{A_j}, \]

valid for disjoint sets.

For any sequence \(\{A_j\}\) of sets one writes \(\limsup A_j\) for the set of points $\omega$ that belong to infinitely many of the $A_j$, that is,

\[ \limsup A_j = \cap_j \cup_{k \geq j} A_k \]

(for all $j$ there exists some $k \geq j$ such that $\omega$ belongs to $A_k$). By \(\liminf A_j\) is meant the set of points $\omega$ that belong to all but possibly finitely many of the $A_j$, that is,

\[ \liminf A_j = \cup_j \cap_{k \geq j} A_k \]

(there exists a $j$ such that for all $k \geq j$ the point $\omega$ belongs to $A_k$). This usage is in accordance with

\[ I_{\cap_j \cup_{k \geq j} A_k} = \inf_j \sup_{k \geq j} I_{A_k} \]
and
\[ I_{\bigcup_{j \geq 1} A_k} = \sup_{j} \inf_{k \geq j} I_{A_k}, \]
obtained upon combining (B.8) and (B.9).

\section*{C. Indicators of events.}
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be some probability space. The indicator \(I_A\) of an event \(A \in \mathcal{F}\) is a simple binomial random variable;
\[ I_A \sim \text{Bin}(1, \mathbb{P}[A]). \]

It follows that
\[ \mathbb{E}[I_A] = \mathbb{P}[A], \quad \mathbb{V}[I_A] = \mathbb{P}[A](1 - \mathbb{P}[A]). \quad (B.10) \]
Appendix C

Distribution of the number of occurring events

A. The main result. Let \( \{A_1, \ldots, A_r\} \) be a finite assembly of events, not necessarily disjoint. Introduce the short-hand \( I_j = I_{A_j} \). We seek the probability distribution of the number of events that occur out of the total of \( r \) events,

\[
Q = \sum_{j=1}^{r} I_j.
\]

It turns out that this distribution can be expressed in terms of the probabilities of intersections of selections from the assembly of sets. Introduce

\[
Z_p = \sum_{j_1 < \ldots < j_p} ^{r} P[A_{j_1} \cap \cdots \cap A_{j_p}], \quad p = 1, \ldots, r, \tag{C.1}
\]

and define in particular \( Z_0 = 1 \).

**Theorem**

The probability distribution of \( Q \) can be expressed by the \( Z_p \) in (C.1) as

\[
P[Q = q] = \sum_{p=q}^{r} (-1)^{p-q} \binom{p}{p-q} Z_p, \quad q = 0, \ldots, r, \tag{C.2}
\]

\[
P[Q \geq q] = \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{p-q} Z_p, \quad q = 1, \ldots, r. \tag{C.3}
\]

**Proof**: Obviously,

\[
\{Q = q\} = \bigcup_{r \setminus q} A_{j_1} \cap \cdots \cap A_{j_q} \cap A_{j_{q+1}}^c \cap \cdots \cap A_{j_r}^c.
\]
The elements in the union are mutually disjoint, and so
\[ I_{\{Q=q\}} = \sum_{r \leq q} I_{j_1} \cdots I_{j_p} (1 - I_{j_{p+1}}) \cdots (1 - I_{j_r}). \]

Starting from this expression, the generating function of the sequence \( \{I_{\{Q=p\}}\}_{p=0,\ldots,r} \) can be shaped as follows by repeated use (B.6):

\[
\sum_{p=0}^{r} s^p I_{\{Q=p\}} = \sum_{p=0}^{r} s^p \sum_{r \leq p} I_{j_1} \cdots I_{j_p} (1 - I_{j_{p+1}}) \cdots (1 - I_{j_r}) \\
= \sum_{p=0}^{r} \sum_{r \leq p} s I_{j_1} \cdots s I_{j_p} (1 - I_{j_{p+1}}) \cdots (1 - I_{j_r}) \\
= \prod_{j=1}^{r} (s I_j + 1 - I_j) \\
= \prod_{j=1}^{r} ((s-1)I_j + 1) \\
= \sum_{p=0}^{r} \sum_{r \leq p} (s-1)I_{j_1} \cdots (s-1)I_{j_p} 1^{r-p},
\]

where the first term corresponding to \( p = 0 \) is to be interpreted as 1. Thus,

\[
\sum_{p=0}^{r} s^p I_{\{Q=p\}} = \sum_{p=0}^{r} (s-1)^p Y_p, \quad (C.4)
\]

where

\[
Y_p = \sum_{j_1 < \cdots < j_p} I_{j_1} \cdots I_{j_p}, \quad p = 1, \ldots, r, \quad (C.5)
\]

and \( Y_0 = 1 \). Upon differentiating (C.4) \( q \) times with respect to \( s \) and putting \( s = 0 \), we get

\[
q! I_{\{Q=q\}} = \sum_{p=0}^{r} p^{(q)} (-1)^{p-q} Y_p, \quad (C.6)
\]

hence, noting that \( p^{(q)}/q! = \binom{p}{q} = \binom{p}{p-q} \),

\[
I_{\{Q=q\}} = \sum_{p=q}^{r} (-1)^{p-q} \binom{p}{p-q} Y_p. \quad (C.6)
\]

Taking expectation, we arrive at (C.2).
APPENDIX C. DISTRIBUTION OF THE NUMBER OF OCCURRING EVENTS

To prove (C.3), insert \( I\{Q=p\} = I\{Q\geq p\} - I\{Q\geq p+1\} \) on the left of (C.4) and rearrange as follows:

\[
\sum_{p=0}^{r} s^p (I\{Q\geq p\} - I\{Q\geq p+1\}) = \sum_{p=0}^{r} s^p I\{Q\geq p\} - \sum_{p=1}^{r} s^{p-1} I\{Q\geq p\}
\]

\[
= 1 + \sum_{p=1}^{r} (s-1)s^{p-1} I\{Q\geq p\}.
\]

Thus, recalling that \( Y_0 = 1 \), (C.4) is equivalent to

\[
\sum_{p=1}^{r} s^{p-1} I\{Q\geq p\} = \sum_{p=1}^{r} (s-1)s^{p-1} Y_p.
\]

Differentiating here \( q-1 \) times with respect to \( s \) and putting \( s = 0 \), gives

\[
(q-1)! I\{Q\geq q\} = \sum_{p=q}^{r} (p-1)^{(q-1)}(-1)^{p-q} Y_p,
\]

hence

\[
I\{Q\geq q\} = \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{p-q} Y_p,
\]

which implies (C.3). \( \square \)

**B. Comments and examples.** Setting all \( A_j \) equal to the sure event \( \Omega \), all the indicators \( I_j \) become identically 1. Thus \( Y_p \) defined by (C.5) becomes \( \binom{r}{p} \), and (C.7) specializes to

\[
1 = \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{p-q} \binom{r}{p}.
\]

For \( q = r \), the theorem reduces to the trivial result

\[
P[Q = r] = P[Q \geq r] = P[A_1 \cap \cdots \cap A_r].
\]

Putting \( q = 1 \) in (C.3) and noting that \([Q \geq 1] = \bigcup_j A_j\), yields

\[
P[\bigcup_j A_j] = \sum_j P[A_j] - \sum_{j_1 < j_2} P[A_{j_1} \cap A_{j_2}]
\]

\[+ \ldots + (-1)^{r-1} P[A_1 \cap \cdots \cap A_r],
\]

which also results upon taking expectation in (B.5). This is the well-known general addition rule for probabilities, called so because it generalizes the elementary rule \( P[A \cup B] = P[A] + P[B] - P[A \cap B] \) (confer (B.3)). The theorem states the most general results of this type.
As a non-standard example, let us find the probability of exactly two occurrences among three events, $A_1, A_2, A_3$. Putting $r = 3$ and $q = 2$ in (C.2), gives

$$P[Q = 2] = P[A_1 \cap A_2] + P[A_1 \cap A_3] + P[A_2 \cap A_3] - 3P[A_1 \cap A_2 \cap A_3].$$

From (C.3) we obtain the probability of at least two occurrences,

$$P[Q \geq 2] = P[A_1 \cap A_2] + P[A_1 \cap A_3] + P[A_1 \cap A_3] - 2P[A_1 \cap A_2 \cap A_3].$$

The usefulness of the theorem is due to the decomposition of complex events into more elementary ones. The observant reader may have asked why intersections rather than unions are taken as the elementary events. The reason for doing so is apparent in the case of independent events, since then $P[\cap_{i=1}^p A_j] = \prod_{i=1}^p P[A_j]$, and the expressions in (C.1) – (C.3) can be computed from the $P[A_j]$ by elementary algebraic operations.

Note that the results in the theorem are independent of the probability measure involved; they rest entirely on the set-relations (C.6) and (C.7).
Appendix D

Asymptotic results from statistics

A. The central limit theorem  Let $X_1, X_2, \ldots$ be a sequence of random variables with zero means, $\mathbb{E}[X_i] = 0$, and finite variances,

$$\sigma_i^2 = \mathbb{V}[X_i], \ i = 1, 2, \ldots$$  \hspace{1cm} (D.1)

Define

$$b_n^2 = \sum_{i=1}^{n} \sigma_i^2, \ n = 1, 2, \ldots$$  \hspace{1cm} (D.2)

The celebrated Lindeberg/Feller central limit theorem says that if

$$\sum_{i=1}^{n} \mathbb{E} \left[ X_i^2 \mathbb{1}_{\{X_i^2 > \varepsilon b_n^2\}} \right] \Rightarrow 0, \ \forall \varepsilon > 0,$$  \hspace{1cm} (D.3)

then

$$\frac{\sum_{i=1}^{n} X_i}{b_n} \xrightarrow{d} \mathcal{N}(0, 1).$$  \hspace{1cm} (D.4)

B. Asymptotic properties of MLE estimators  The asymptotic distributions derived in Chapter 9 could be obtained directly from the following standard result, which is cited here without proof.

Let $X_1, X_2, \ldots$ be a sequence of random elements with joint distribution depending on a parameter $\theta$ that varies in an open set in the $s$-dimensional euclidean space. Assume that the likelihood function of $X_1, X_2, \ldots, X_n$, denoted by $\Lambda_n(X_1, X_2, \ldots, X_n, \theta)$, is twice continuously differentiable with respect to $\theta$ and that the equation

$$\frac{\partial}{\partial \theta} \ln \Lambda_n(X_1, X_2, \ldots, X_n, \theta) = 0^{s \times 1}$$
has a unique solution \( \hat{\theta}_n(X_1, X_2, \ldots, X_n) \), called the MLE (maximum likelihood estimator). Then, if the matrix

\[
\Sigma(\theta) = \left( -\mathbb{E} \left[ \frac{\partial^2}{\partial \theta \partial \theta'} \ln \Lambda_n \right] \right)^{-1}
\]  

(D.5)
tends to \( 0^{s \times s} \) as \( n \to \infty \), the MLE is asymptotically normally distributed,

\[
\hat{\theta} \sim_{as} N(\theta, \Sigma(\theta)).
\]

C. The delta method  Assume that \( \hat{\theta} \) is a consistent estimator of \( \theta \in \Theta \), an open set in \( \mathbb{R}^s \), and that \( \hat{\theta} \sim_{as} N(\theta, \Sigma) \). If \( g: \mathbb{R}^s \to \mathbb{R}^r \) is a twice continuously differentiable function of \( \theta \), then

\[
g(\hat{\theta}) \sim_{as} N \left( g(\theta), \frac{\partial}{\partial \theta} g(\theta) \Sigma \frac{\partial}{\partial \theta} g(\theta) \right). \]  

(D.6)

The result follows easily by inspection of the first order Taylor expansion of \( g(\hat{\theta}) \) around \( \theta \).
## Appendix E

### The G82M mortality table

Table E.1: The mortality table G82M

<table>
<thead>
<tr>
<th>x</th>
<th>$\mu_x$</th>
<th>$f(x)$</th>
<th>$\ell_x = 10^8 \bar{F}(x)$</th>
<th>$d_x$</th>
<th>$q_x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.00057586</td>
<td>0.00057586</td>
<td>100000</td>
<td>58</td>
<td>0.00057911</td>
</tr>
<tr>
<td>1</td>
<td>0.00058279</td>
<td>0.00058246</td>
<td>99942</td>
<td>59</td>
<td>0.00058635</td>
</tr>
<tr>
<td>2</td>
<td>0.00059036</td>
<td>0.00058968</td>
<td>99883</td>
<td>59</td>
<td>0.00059426</td>
</tr>
<tr>
<td>3</td>
<td>0.00059863</td>
<td>0.00059758</td>
<td>99824</td>
<td>60</td>
<td>0.00060289</td>
</tr>
<tr>
<td>4</td>
<td>0.00060765</td>
<td>0.00060621</td>
<td>99764</td>
<td>61</td>
<td>0.00061231</td>
</tr>
<tr>
<td>5</td>
<td>0.00061749</td>
<td>0.00061565</td>
<td>99703</td>
<td>62</td>
<td>0.00062259</td>
</tr>
<tr>
<td>6</td>
<td>0.00062823</td>
<td>0.00062598</td>
<td>99641</td>
<td>63</td>
<td>0.00063381</td>
</tr>
<tr>
<td>7</td>
<td>0.00063996</td>
<td>0.00063726</td>
<td>99578</td>
<td>65</td>
<td>0.00064606</td>
</tr>
<tr>
<td>8</td>
<td>0.00065276</td>
<td>0.00064958</td>
<td>99513</td>
<td>65</td>
<td>0.00065942</td>
</tr>
<tr>
<td>9</td>
<td>0.00066672</td>
<td>0.00066304</td>
<td>99448</td>
<td>67</td>
<td>0.00067401</td>
</tr>
<tr>
<td>10</td>
<td>0.00068197</td>
<td>0.00067775</td>
<td>99381</td>
<td>69</td>
<td>0.00068993</td>
</tr>
<tr>
<td>11</td>
<td>0.00069861</td>
<td>0.00069380</td>
<td>99312</td>
<td>70</td>
<td>0.00070731</td>
</tr>
<tr>
<td>12</td>
<td>0.00071677</td>
<td>0.00071134</td>
<td>99242</td>
<td>72</td>
<td>0.00072627</td>
</tr>
<tr>
<td>13</td>
<td>0.00073659</td>
<td>0.00073048</td>
<td>99170</td>
<td>74</td>
<td>0.00074697</td>
</tr>
<tr>
<td>14</td>
<td>0.00075823</td>
<td>0.00075137</td>
<td>99096</td>
<td>77</td>
<td>0.00076956</td>
</tr>
<tr>
<td>15</td>
<td>0.00078184</td>
<td>0.00077417</td>
<td>99019</td>
<td>78</td>
<td>0.00079422</td>
</tr>
<tr>
<td>16</td>
<td>0.00080761</td>
<td>0.00079906</td>
<td>98941</td>
<td>81</td>
<td>0.00082113</td>
</tr>
<tr>
<td>17</td>
<td>0.00083574</td>
<td>0.00082621</td>
<td>98860</td>
<td>85</td>
<td>0.00085050</td>
</tr>
<tr>
<td>18</td>
<td>0.00086644</td>
<td>0.00085583</td>
<td>98775</td>
<td>87</td>
<td>0.00088256</td>
</tr>
<tr>
<td>19</td>
<td>0.00089994</td>
<td>0.00088814</td>
<td>98688</td>
<td>90</td>
<td>0.00091754</td>
</tr>
</tbody>
</table>
## APPENDIX E. THE G82M MORTALITY TABLE

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
x & \mu_x & f(x) & \ell_x = 10^8 \bar{F}(x) & d_x & q_x \\
\hline
20 & 0.00093652 & 0.00092338 & 98.598 & 94 & 0.00095573 \\
21 & 0.00097643 & 0.00096182 & 98.504 & 99 & 0.00099740 \\
22 & 0.00102000 & 0.00100373 & 98.405 & 102 & 0.00104288 \\
23 & 0.00106754 & 0.00104942 & 98.303 & 108 & 0.00109252 \\
24 & 0.00111944 & 0.00109924 & 98.201 & 112 & 0.00116190 \\
25 & 0.00117608 & 0.00115353 & 98.100 & 119 & 0.00123082 \\
26 & 0.00123790 & 0.00121270 & 97.999 & 124 & 0.00130434 \\
27 & 0.00130538 & 0.00127718 & 97.898 & 131 & 0.00138076 \\
28 & 0.00137902 & 0.00134743 & 97.796 & 139 & 0.00146072 \\
29 & 0.00145940 & 0.00142394 & 97.694 & 146 & 0.00154510 \\
30 & 0.00154713 & 0.00150727 & 97.592 & 155 & 0.00163304 \\
31 & 0.00164288 & 0.00159800 & 97.490 & 165 & 0.00172593 \\
32 & 0.00174738 & 0.00169678 & 97.388 & 175 & 0.00182389 \\
33 & 0.00186144 & 0.00180428 & 97.286 & 186 & 0.00192894 \\
34 & 0.00198594 & 0.00192125 & 97.184 & 199 & 0.00204078 \\
35 & 0.00212181 & 0.00204849 & 97.082 & 211 & 0.00216247 \\
36 & 0.00227011 & 0.00218686 & 96.980 & 226 & 0.00234710 \\
37 & 0.00243197 & 0.00233728 & 96.878 & 242 & 0.00251584 \\
38 & 0.00260863 & 0.00250075 & 96.776 & 259 & 0.00269998 \\
39 & 0.00280144 & 0.00267834 & 96.674 & 277 & 0.00290091 \\
40 & 0.00301189 & 0.00287119 & 96.572 & 298 & 0.00312018 \\
41 & 0.00324157 & 0.00308050 & 96.470 & 319 & 0.00335944 \\
42 & 0.00349226 & 0.00330759 & 96.368 & 343 & 0.00362051 \\
43 & 0.00376588 & 0.00355382 & 96.266 & 369 & 0.00390537 \\
44 & 0.00406451 & 0.00382066 & 96.164 & 396 & 0.00421619 \\
45 & 0.00439045 & 0.00410964 & 96.062 & 426 & 0.00455532 \\
46 & 0.00474620 & 0.00442239 & 95.960 & 459 & 0.00492533 \\
47 & 0.00513447 & 0.00476062 & 95.858 & 494 & 0.00532902 \\
48 & 0.00555825 & 0.00512607 & 95.756 & 532 & 0.00576943 \\
49 & 0.00602077 & 0.00552060 & 95.655 & 574 & 0.00624889 \\
50 & 0.00652560 & 0.00594609 & 95.553 & 617 & 0.00677402 \\
51 & 0.00707658 & 0.00640446 & 95.451 & 665 & 0.00734576 \\
52 & 0.00767794 & 0.00689767 & 95.349 & 716 & 0.00796941 \\
53 & 0.00833430 & 0.00742765 & 95.247 & 770 & 0.00864963 \\
54 & 0.00905067 & 0.00799632 & 95.145 & 830 & 0.00939153 \\
55 & 0.00983254 & 0.00860553 & 95.043 & 893 & 0.01020062 \\
56 & 0.01068591 & 0.00925700 & 94.941 & 960 & 0.01108295 \\
57 & 0.01161732 & 0.00995232 & 94.839 & 1032 & 0.01204506 \\
58 & 0.01263389 & 0.01069283 & 94.737 & 1108 & 0.01304908 \\
59 & 0.01374341 & 0.01147959 & 94.636 & 1189 & 0.01423775 \\
\hline
\end{array}
\]
### APPENDIX E. THE G82M MORTALITY TABLE

<table>
<thead>
<tr>
<th>$x$</th>
<th>$\mu_x$</th>
<th>$f(x)$</th>
<th>$\ell_x = 10^8\bar{F}(x)$</th>
<th>$d_x$</th>
<th>$q_x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>0.01495440</td>
<td>0.01231325</td>
<td>82339</td>
<td>1275</td>
<td>0.01548448</td>
</tr>
<tr>
<td>61</td>
<td>0.01627611</td>
<td>0.01319402</td>
<td>81664</td>
<td>1366</td>
<td>0.01684342</td>
</tr>
<tr>
<td>62</td>
<td>0.01771868</td>
<td>0.01412149</td>
<td>79698</td>
<td>1460</td>
<td>0.01832447</td>
</tr>
<tr>
<td>63</td>
<td>0.01929317</td>
<td>0.01509456</td>
<td>78238</td>
<td>1560</td>
<td>0.01993841</td>
</tr>
<tr>
<td>64</td>
<td>0.02101162</td>
<td>0.01611127</td>
<td>76678</td>
<td>1664</td>
<td>0.02169690</td>
</tr>
<tr>
<td>65</td>
<td>0.02288721</td>
<td>0.01716867</td>
<td>75014</td>
<td>1771</td>
<td>0.02361259</td>
</tr>
<tr>
<td>66</td>
<td>0.02493430</td>
<td>0.01826263</td>
<td>73243</td>
<td>1882</td>
<td>0.02569916</td>
</tr>
<tr>
<td>67</td>
<td>0.02716858</td>
<td>0.01938769</td>
<td>71361</td>
<td>1996</td>
<td>0.02797145</td>
</tr>
<tr>
<td>68</td>
<td>0.02960717</td>
<td>0.02053690</td>
<td>69365</td>
<td>2112</td>
<td>0.03044546</td>
</tr>
<tr>
<td>69</td>
<td>0.03226874</td>
<td>0.02170163</td>
<td>67253</td>
<td>2229</td>
<td>0.03313851</td>
</tr>
<tr>
<td>70</td>
<td>0.03517368</td>
<td>0.02287138</td>
<td>65024</td>
<td>2345</td>
<td>0.03606928</td>
</tr>
<tr>
<td>71</td>
<td>0.03834425</td>
<td>0.02403370</td>
<td>62679</td>
<td>2461</td>
<td>0.03925790</td>
</tr>
<tr>
<td>72</td>
<td>0.04180475</td>
<td>0.02517403</td>
<td>60218</td>
<td>2573</td>
<td>0.04272605</td>
</tr>
<tr>
<td>73</td>
<td>0.04558167</td>
<td>0.02627566</td>
<td>57645</td>
<td>2680</td>
<td>0.04649704</td>
</tr>
<tr>
<td>74</td>
<td>0.04970395</td>
<td>0.02731973</td>
<td>54965</td>
<td>2781</td>
<td>0.05059590</td>
</tr>
<tr>
<td>75</td>
<td>0.05420317</td>
<td>0.02828534</td>
<td>52184</td>
<td>2873</td>
<td>0.05504946</td>
</tr>
<tr>
<td>76</td>
<td>0.05911381</td>
<td>0.02914974</td>
<td>49311</td>
<td>2953</td>
<td>0.05988641</td>
</tr>
<tr>
<td>77</td>
<td>0.06447348</td>
<td>0.02988871</td>
<td>46358</td>
<td>3020</td>
<td>0.06513740</td>
</tr>
<tr>
<td>78</td>
<td>0.07032323</td>
<td>0.03047703</td>
<td>43338</td>
<td>3069</td>
<td>0.07083506</td>
</tr>
<tr>
<td>79</td>
<td>0.07670789</td>
<td>0.03088920</td>
<td>40269</td>
<td>3102</td>
<td>0.07701411</td>
</tr>
<tr>
<td>80</td>
<td>0.08367637</td>
<td>0.03110030</td>
<td>37767</td>
<td>3111</td>
<td>0.08371127</td>
</tr>
<tr>
<td>81</td>
<td>0.09128204</td>
<td>0.03108704</td>
<td>34056</td>
<td>3098</td>
<td>0.09096537</td>
</tr>
<tr>
<td>82</td>
<td>0.09958318</td>
<td>0.03082907</td>
<td>30958</td>
<td>3059</td>
<td>0.09881727</td>
</tr>
<tr>
<td>83</td>
<td>0.10864338</td>
<td>0.03031033</td>
<td>27899</td>
<td>2994</td>
<td>0.10730975</td>
</tr>
<tr>
<td>84</td>
<td>0.11853205</td>
<td>0.02952051</td>
<td>24905</td>
<td>2901</td>
<td>0.11648747</td>
</tr>
<tr>
<td>85</td>
<td>0.12932494</td>
<td>0.02845661</td>
<td>22004</td>
<td>2781</td>
<td>0.12639675</td>
</tr>
<tr>
<td>86</td>
<td>0.14110474</td>
<td>0.02712418</td>
<td>19223</td>
<td>2635</td>
<td>0.13708534</td>
</tr>
<tr>
<td>87</td>
<td>0.15396168</td>
<td>0.02553851</td>
<td>16588</td>
<td>2465</td>
<td>0.14860208</td>
</tr>
<tr>
<td>88</td>
<td>0.16799427</td>
<td>0.02372520</td>
<td>14123</td>
<td>2274</td>
<td>0.16099656</td>
</tr>
<tr>
<td>89</td>
<td>0.18331000</td>
<td>0.02172028</td>
<td>11849</td>
<td>2066</td>
<td>0.17431855</td>
</tr>
<tr>
<td>90</td>
<td>0.20002621</td>
<td>0.01956945</td>
<td>9783</td>
<td>1845</td>
<td>0.18861740</td>
</tr>
<tr>
<td>91</td>
<td>0.21827096</td>
<td>0.01732660</td>
<td>7938</td>
<td>1619</td>
<td>0.20394126</td>
</tr>
<tr>
<td>92</td>
<td>0.23814011</td>
<td>0.01505134</td>
<td>6319</td>
<td>1392</td>
<td>0.22033622</td>
</tr>
<tr>
<td>93</td>
<td>0.25991791</td>
<td>0.01280578</td>
<td>4927</td>
<td>1172</td>
<td>0.23784520</td>
</tr>
<tr>
<td>94</td>
<td>0.28363917</td>
<td>0.01065073</td>
<td>3755</td>
<td>963</td>
<td>0.25650675</td>
</tr>
<tr>
<td>95</td>
<td>0.30952951</td>
<td>0.00864156</td>
<td>2792</td>
<td>772</td>
<td>0.27635358</td>
</tr>
<tr>
<td>96</td>
<td>0.33778728</td>
<td>0.00682433</td>
<td>2020</td>
<td>601</td>
<td>0.29741104</td>
</tr>
<tr>
<td>97</td>
<td>0.36862894</td>
<td>0.00523248</td>
<td>1419</td>
<td>453</td>
<td>0.31969526</td>
</tr>
<tr>
<td>98</td>
<td>0.40229977</td>
<td>0.00388920</td>
<td>966</td>
<td>332</td>
<td>0.34321126</td>
</tr>
<tr>
<td>99</td>
<td>0.43903065</td>
<td>0.00278447</td>
<td>634</td>
<td>233</td>
<td>0.36795078</td>
</tr>
</tbody>
</table>
## APPENDIX E. THE G82M MORTALITY TABLE

<table>
<thead>
<tr>
<th>$x$</th>
<th>$\mu_x$</th>
<th>$f(x)$</th>
<th>$\ell_x = 10^8\hat{F}(x)$</th>
<th>$d_x$</th>
<th>$q_x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.47913004</td>
<td>0.00192066</td>
<td>401</td>
<td>158</td>
<td>0.39389013</td>
</tr>
<tr>
<td>101</td>
<td>0.52289614</td>
<td>0.00127047</td>
<td>243</td>
<td>102</td>
<td>0.42098791</td>
</tr>
<tr>
<td>102</td>
<td>0.57066422</td>
<td>0.00080282</td>
<td>141</td>
<td>64</td>
<td>0.44918271</td>
</tr>
<tr>
<td>103</td>
<td>0.62280022</td>
<td>0.00048261</td>
<td>77</td>
<td>37</td>
<td>0.47839101</td>
</tr>
<tr>
<td>104</td>
<td>0.67970357</td>
<td>0.00027473</td>
<td>40</td>
<td>20</td>
<td>0.50850528</td>
</tr>
<tr>
<td>105</td>
<td>0.74181017</td>
<td>0.00014737</td>
<td>20</td>
<td>11</td>
<td>0.53939240</td>
</tr>
<tr>
<td>106</td>
<td>0.80959582</td>
<td>0.00007408</td>
<td>9</td>
<td>5</td>
<td>0.57089277</td>
</tr>
<tr>
<td>107</td>
<td>0.88357981</td>
<td>0.00003469</td>
<td>4</td>
<td>2</td>
<td>0.60281998</td>
</tr>
<tr>
<td>108</td>
<td>0.96432893</td>
<td>0.00001504</td>
<td>2</td>
<td>1</td>
<td>0.63496163</td>
</tr>
<tr>
<td>109</td>
<td>1.05246177</td>
<td>0.00000599</td>
<td>1</td>
<td>1</td>
<td>0.66708109</td>
</tr>
<tr>
<td>110</td>
<td>1.14865351</td>
<td>0.00000218</td>
<td>0</td>
<td>0</td>
<td>0.69892078</td>
</tr>
</tbody>
</table>
Appendix F

List of terms

Forkortelser:
adj = adjective (tillægsord)
adv = adverb (biord)
n = noun (navneord)
iy = intransitive verb (intransitivt verbum)
tv = transitive verb (transitivt verbum)

administration cost/expense administrationsomkostning
aggregate aggregat
amortization amortisation
annual adj, ~ly adv årlig
annuity annuitet
  decreasing ~ aftagende annuitet
  ~due forskudsvis annuitet
  immediate ~ efterskudsvis annuitet
  increasing ~ voksende annuitet
  life ~ livrente
  whole life ~ livsvarig livrente
reversionary ~ adj eventuel livrente (e.g. enkepension)
basis n grundlag
  first order (technical) ~ første ordens (teknisk) grundlag
  second order (empirical) ~ anden ordens grundlag
benefit n gode, forsinkring: (forsikrings)ydelse
bequeath n testamentere, overdrage (jura)
bequest n arv ved testamente (jura)
bonus n bonus (af Latin bonus n, adj = ‘gode’, ‘god’)
cash ~ kontantbonus
reversionary ~ slutbonus
terminal ~ slutbonus
censoring censurering
compound interest rente og rentes rente
compounding factor rentefaktor, oprentningsfaktor

contribution bidrag
  ~ plan kontributionsplan

debit debet (skylde, gæld)
decrement $n$ afgang, nedgang
  ~ function dekrementfunktion

cause of $\sim$ afgangsårsag

deposit $n$ indskud (på konto), $tv$ indskyde

defer opsætte
  $\sim$ red life annuity opsat livrente

disability invaliditet, uforhåd
  ~ pension invalidepension

discount $tv$ diskontere (realisere værdi af fordring)
  ~ing factor/function diskonteringsfaktor/-funktion

dividend udbytte, dividende

duration varighed

down $tv$ udstyre, give

donation $tv$ udstyr, gave
  ~ insurance livs forsikring med udbetaling (af forsikringssum ved
toverlevelse af forsikringstiden)
  pure $\sim$ udstyrss forsikring

equity capital egenkapital

equivalence ækvivalens
  principle of $\sim$ ækvivalensprincippet

expiry udløb (af kontrakt)
exposure eksponering

force of transition overgangsintensitet

graduation udglatning

gross $adj$ brutto
  ~ premium brutto præmie
  ~ premium reserve brutto præmiereserve

impair forringe
  ~ed life mindre godt liv (med overdødelighed p.g.a. sygdom,
risikofylt erhverv e.l.)

heir $n$ arving, naturlig (født) efterfølger

hereditary $adj$ arvelig, specielt i biologisk forstand

incur $iv$ indtræffe (med konsekvenser), påløbe (om omkostninger)
in indicator function indikatorfunktion

inherit $tv$ arve; overtage ejendom ved testament eller titel
  som naturlig efterfølger (jura); modtage en arvelig genetisk egenskab (biol.)
inheritance $n$ arv

installment aftag

interest rente
  force of $\sim$ rente intensitet (rente pr. krone pr. tidsenhed; grænseværdi
  når tidsintervallet går mod null)
  ~ rate rentefod (rente pr. krone pr. tidsenhed, kan benyttes om
APPENDIX F. LIST OF TERMS

ændeligt tidsinterval, e.g. rente pr. år

issue $n$, adj udstedelse, udstede

joint life $n$, adj samliv (korteste liv), samlivs-

level $n$ niveau

level adj fast, konstant

life history analysis levetidsanalyse, forløbsanalyse

longest life længste liv, længstlivs-

longevity det at leve længe

lump klump, stykke

~ sum beløb forskelligt fra 0, der forfalder på et givet tidspunkt

mortality dødelighed

~ law dødelighedslov

~ table dødelighedstavle

aggregate ~ adj aggregat dødelighed

excess ~ overdødelighed

force of ~ dødsintensitet

nondifferential ~ ikke-differentiel dødelighed

multi-life adj flerlivs-

mutual adj gensidig

net netto

~ premium netto præmie

~ premium reserve netto premierserve

occurrence indtræffelse

part-payment afdrag

perpetuity (= perpetual annuity) perpetuel (evigvarende) annuitet

piecewise stykkevis

premium præmie

~ rate præmierate

single ~ engangspræmie

present value kapitalværdi, kontantværdi

policy police

portfolio bestand, portefølje (af fransk portfeuille = taske)

principal $n$ finans: hovedstol, lånesum

remaining ~ restgæld

qualification period karenstid (i invalidepensionsforsikring)

prospective adj prospektiv, 'fremadskuende'

redistribution tilbageforsel

reserve reserve

retrospective adj retrospektiv, 'bagudskuende'

revenue $n$ indtægt

safety loading sikkerhedstilkæg

safety margin sikkerhedsmargen

select $tv$, adj udvælge, baseret på selektion

~ mortality select dødelighed

~ period selektionsperiode

single life et liv, etlivs
APPENDIX F. LIST OF TERMS

sum at risk risikosum
surplus $n$ overskud
surrender $iv$ overgive (sig), forsikring: genkøbe
    ~ value genkøbsværdi
survival function overlevelsfunktion
temporary life annuity temporær livrente
term (1) $n$ periode
    ~ insurance korttidsforsikring, ren dødsfaldsforsikring
    ~ of contract kontraktsperiode
term (2) $n$ betingelse
contract ~s kontraktsbetingelser
transition probability overgangssandsynlighed
waiver of premium præmiefritagelse
withdraw $iv, tv$ (1) trække (sig) tilbage, tilbagekalde, forsikring: opsige
    forsikringsaftale; (2) udtrække, hæve (penge fra en konto)
    ~al $n$ (1) forsikring: opsigelse; (2) udtræk (fra en konto)
whole life insurance simpel livsforsikring