Basic Life Insurance Mathematics

Ragnar Norberg

Version: September 2002

Contents

Chapter 1

Introduction

1.1 Banking versus insurance

A. The bank savings contract. Upon celebrating his 55th anniversary Mr. (55) (let us call him so) decides to invest money to secure himself economically in his old age. The first idea that occurs to him is to deposit a capital of $S_0 = 1$ (e.g. one hundred thousand pounds) on a savings account today 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 $S_1 = S_0 + S_0\, i = S_0(1+i)$, in two years it will increase to $S = S_1 + S_1\, i = S_0(1+i)^2$, and so on until in 15 years it will have accumulated to

$$S_{15} = S_0\,(1+i)^{15} = 1.045^{15} = 1.935\,. \tag{1.1}$$

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 waisted. Checking population statistics he learns that about 75% of those who are 55 will survive to 70. Thus, the relevant prospects of the contract are:

– with probability 0.75 (55) survives to 70 and will then possess S_{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

$$0.75\, S_{15}\,. \tag{1.2}$$

B. A small scale mutual fund. Having thought things over, (55) 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 $S_0 = 1$ on the savings account, and those who survive to 70, if any, will then share the total accumulated capital $3\, S_{15}$ equally.

The prospects of this scheme are given in Table 1.1, where + and − signify *survival* and *death*, respectively, L_{70} is the number of survivors at age 70, and

Table 1.1: Possible outcomes of a savings scheme with three participants.

(55)	$(55)^*$	$(55)^{**}$	L_{70}	$3\,S_{15}/L_{70}$	Probability
+	+	+	3	S_{15}	$0.75 \cdot 0.75 \cdot 0.75 = 0.422$
+	+	−	2	$1.5\,S_{15}$	$0.75 \cdot 0.75 \cdot 0.25 = 0.141$
+	−	+	2	$1.5\,S_{15}$	$0.75 \cdot 0.25 \cdot 0.75 = 0.141$
+	−	−	1	$3S_{15}$	$0.75 \cdot 0.25 \cdot 0.25 = 0.047$
−	+	+	2	$1.5\,S_{15}$	$0.25 \cdot 0.75 \cdot 0.75 = 0.141$
−	+	−	1	$3\,S_{15}$	$0.25 \cdot 0.75 \cdot 0.25 = 0.047$
−	−	+	1	$3\,S_{15}$	$0.25 \cdot 0.25 \cdot 0.75 = 0.047$
−	−	−	0	undefined	$0.25 \cdot 0.25 \cdot 0.25 = 0.016$

$3\,S_{15}/L_{70}$ 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 S_{15};
– with probability $2 \cdot 0.141 = 0.282$ (55) survives to 70 together with one more survivor and will then possess $1.5\,S_{15}$;
– 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 $3S_{15}$;
– with probability 0.25 (55) dies before 70 and will get nothing.
This scheme is 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 S_{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

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

The point is that under the present scheme the savings of those who die are bequeathed to the survivors. Thus the total savings are retained for 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 unites the members of a pension fund. 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 idea 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 $L_{55}\,S_{15}$, which yields an individual share equal to

$$\frac{L_{55}\,S_{15}}{L_{70}} \tag{1.3}$$

to each of the L_{70} survivors if $L_{70} > 0$. By the so-called law of large numbers, the proportion of survivors L_{70}/L_{55} tends to the individual survival probability 0.75 as the number of participants L_{55} tends to infinity. Therefore, as the number of participants increases, the individual share per survivor tends to

$$\frac{1}{0.75}S_{15}\,, \tag{1.4}$$

and in the limit (55) is faced with the following situation:
– with probability 0.75 he survives to 70 and gets $\frac{1}{0.75}S_{15}$;
– with probability 0.25 he dies before 70 and gets nothing.
The expected amount at (55)'s disposal after 15 years is

$$0.75\frac{1}{0.75}S_{15} = S_{15}, \tag{1.5}$$

the same as (1.1). Thus, the bequest mechanism of the mutual scheme has raised (55)'s expectations 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.

The fact that L_{70}/L_{55} tends to 0.75 as L_{55} increases, and that (1.3) thus stabilizes at (1.4), is precisely what is meant by saying that "insurance risk is diversifiable". The risk can be eliminated by increasing the size of the portfolio.

1.2 Mortality

A. Life and death in the classical actuarial perspective. Insurance mathematics is widely held to be boring. Hopefully, the present text will not support that prejudice. It must be admitted, 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 analyze 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 (diversification) studied in Paragraph 1.1C. 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 1.1C. On the basis of statistical investigations the actuary constructs a so-called *decrement series*, which takes as it starting point a large number ℓ_0 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

x:	0	25	50	60	70	80	90
ℓ_x:	100 000	98 083	91 119	82 339	65 024	37 167	9 783
d_x:	58	119	617	1 275	2 345	3 111	1 845
q_x:	.000579	.001206	.006774	.015484	.036069	.083711	.188617
p_x:	.999421	.998794	.993226	.984516	.963931	.916289	.811383

Table 1.2 is an excerpt of the table 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 age x and age $x+1$). 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 (55) can make with his bank.

1.3 Banking

A. Interest. 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 annual interest rate i_t in year $t = 1, 2, ...$ Thus, a unit $S_0 = 1$ deposited at time 0 will accumulate with compound interest as follows: In one year the capital increases to $S_1 = S_0 + S_0\, i_1 = 1 + i_1$, in two years it increases to $S_2 = S_1 + S_1\, i_2 = (1+i_1)(1+i_2)$, and in t years it increases to

$$S_t = (1 + i_1) \cdots (1 + i_t)\,, \tag{1.6}$$

called the t-year *accumulation factor*.

Accordingly, the present value at time 0 of a unit withdrawn in j years is

$$S_j^{\;-1} = \frac{1}{S_j}\,, \tag{1.7}$$

called the j-year *discount factor* since it is what the bank would pay you at time 0 if you sell to it (discount) a default-free claim of 1 at time j.

Similarly, the value at time t of a unit deposited at time $j < t$ is

$$(1 + i_{j+1}) \cdots (1 + i_t) = \frac{S_t}{S_j}\,,$$

called the accumulation factor over the time period from j to t, and the value at time t of a unit withdrawn at time $j > t$ is

$$\frac{1}{(1 + i_{t+1}) \cdots (1 + i_j)} = \frac{S_t}{S_j}\,,$$

the discount factor over the time period from t to j.

In general, the value at time t of a unit due at time j is $S_t\, S_j^{\;-1}$, an accumulation factor if $j < t$ and a discount factor if $j > t$ (and of course 1 if $j = t$).

From (1.6) it follows that $S_t = S_{t-1}(1 + i_t)$, hence

$$i_t = \frac{S_t - S_{t-1}}{S_{t-1}}\,,$$

which expresses the interest rate in year t as the relative increase of the balance in year t.

B. Saving in the bank. 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*). The net amount of deposit less withdrawal at time t is $c_t - b_t$. 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,

$$U_t = S_t \sum_{j=0}^{t} S_j{}^{-1}(c_j - b_j)\,. \tag{1.8}$$

It develops in accordance with the "forward" recursive scheme

$$U_t = U_{t-1}(1 + i_t) + c_t - b_t\,, \tag{1.9}$$

$t = 1, 2, \ldots, n$, commencing from

$$U_0 = c_0 - b_0\,.$$

Each year (55) will receive from the bank a statement of account with the calculation (1.9), showing how the current balance emerges from the previous balance, the interest earned meanwhile, and the current movement (deposit less withdrawal).

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

$$U_t \geq 0\,, \tag{1.10}$$

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

$$U_n = 0\,. \tag{1.11}$$

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

$$V_t = S_t \sum_{j=t+1}^{n} S_j{}^{-1}(b_j - c_j)\,, \tag{1.12}$$

the present value of future withdrawals less deposits . Similar to (1.9), the prospective reserve is calculated by the "backward" recursive scheme

$$V_t = (1 + i_{t+1})^{-1}(b_{t+1} - c_{t+1} + V_{t+1})\,, \tag{1.13}$$

$t = n - 1, n - 2, \ldots, 0$, starting from

$$V_n = 0\,.$$

The constraint (1.11) is equivalent to

$$\sum_{j=0}^{n} S_j{}^{-1}c_j - \sum_{j=0}^{n} S_j{}^{-1}b_j = 0\,, \tag{1.14}$$

which says that the discounted value of deposits must be equal to the discounted value of the withdrawals. 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 (1.8) with $t = n$ into (1.11), split the sum $\sum_{j=0}^{n}$ into $\sum_{j=0}^{t} + \sum_{j=t+1}^{n}$, and multiply with S_t/S_n, to arrive at $U_t - V_t = 0$.)

C. The endowment contract. 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 $b = 1$ (one hundred thousand pounds, say) at age 70. Suppose the annual interest rate is fixed and equal to $i = 0.045$, so that the accumulation factor in t years is $S_t = (1+i)^t$, the discount factor in j years is ${S_j}^{-1} = (1+i)^{-j}$, and the time t value of a unit due at time j is

$$S_t\, {S_j}^{-1} = (1+i)^{t-j} \,.$$

For the present contract the equivalence requirement (1.14) is

$$\sum_{j=0}^{14} (1+i)^{-j} c \;-\; (1+i)^{-15} 1 \;=\; 0 \,,$$

from which the bank determines

$$c \;=\; \frac{(1+i)^{-15}}{\sum_{j=0}^{14} (1+i)^{-j}} \;=\; 0.04604 \,, \tag{1.15}$$

Due to interest, this amount is considerably smaller than $1/15 = 0.06667$, which is what (55) would have to save per year if he should choose to tuck the money away under his mattress.

1.4 Insurance

A. The life endowment. Still, to (55) 0.04604 (four thousand six hundred and four pounds) is a considerable expense. He believes in a life before death, and it should be blessed with the joys that money can buy. 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 0.03743 (three thousand seven hundred and forty three pounds).

The insurance agent explains: The calculations of the bank depend only on the amounts $c_t - b_t$ and would apply to any customer (x) who would enter into the same contract at age x, say. Thus, to the bank the customer is really an unknown Mr. X. To the insurance company, however, he is not just Mr. X, but the significant Mr. (x) now x years old. Working under the hypothesis that (x) is one of the ℓ_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 the retrospective reserve at time t is

$$\begin{aligned} U_t^p &= S_t \sum_{j=0}^{t} S_j^{-1}(c_j - b_j)\,\ell_{x+j} &(1.16)\\ &= U_{t-1}^p(1+i_t) + (c_t - b_t)\,\ell_{x+t}\,. &(1.17)\end{aligned}$$

The prospective portfolio reserve at time t is

$$\begin{aligned} V_t^p &= S_t \sum_{j=t+1}^{n} S_j^{-1}(b_j - c_j)\ell_{x+j} &(1.18)\\ &= (1+i_{t+1})^{-1}((b_{t+1} - c_{t+1})\ell_{x+t+1} + V_{t+1}^p)\,. &(1.19)\end{aligned}$$

In particular, for the life endowment analogue to (55)'s savings contract, the only payments are $c_t = c$ for $t = 0, \dots, 14$ and $b_{15} = 1$. The equivalence requirement (1.14) becomes

$$\sum_{j=0}^{14}(1+i)^{-j} c\,\ell_{55+j} \,-\, (1+i)^{-15} 1\,\ell_{70} \,=\, 0\,, \tag{1.20}$$

from which the insurer determines

$$c \,=\, \frac{(1+i)^{-15}\,\ell_{70}}{\sum_{j=0}^{14}(1+i)^{-j}\ell_{55+j}} \,=\, 0.03743\,. \tag{1.21}$$

Inspection of the expressions in (1.15) and (1.21) shows that the latter is smaller due to the fact that ℓ_x is decreasing. This phenomenon is known as *mortality bequest* since the savings of the deceased are bequeathed to the survivors. We shall pursue this issue in Paragraph C below.

B. A life assurance contract. Suppose, contrary to the former hypothesis, that (55) has dependents whom he 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 $b = 1$ (one hundred thousand pounds) to compensate the loss of their bread-winner. The bank can not help in this matter; the benefit of b would have to 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 $c = 0.01701$.

The equivalence requirement (1.14) now becomes

$$\sum_{j=0}^{14}(1+i)^{-j} c\,\ell_{55+j} \,-\, \sum_{j=1}^{15}(1+i)^{-j}\,1\,d_{55+j-1} \,=\, 0\,, \tag{1.22}$$

from which the insurer determines

$$c = \frac{\sum_{j=1}^{15}(1+i)^{-j}\, d_{55+j-1}}{\sum_{j=0}^{14}(1+i)^{-j}\ell_{55+j}} = 0.01701\,. \tag{1.23}$$

C. Individual reserves and mortality bequest. In the insurance schemes described above the contracts of deceased 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,

$$U_t = U_t^p/\ell_{x+t}\,, \qquad V_t = V_t^p/\ell_{x+t}\,.$$

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 general pension insurance contract in Paragraph A we get from (1.17) that the individual reserve develops as

$$\begin{aligned} U_t &= U_{t-1}\frac{\ell_{x+t-1}}{\ell_{x+t}}(1+i_t) + (c_t - b_t) \\ &= U_{t-1}\left(1 + \frac{d_{x+t-1}}{l_{x+t}}\right)(1+i_t) + (c_t - b_t)\,. \end{aligned} \tag{1.24}$$

The bequest mechanism is clearly seen by comparing (1.24) to (1.9): the additional term $U_{t-1}(1+i_t)\, d_{x+t-1}/l_{x+t}$ 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 endowment contracts offered by the bank and the insurance company, respectively. It is seen that the insurance scheme requires a smaller reserve than the bank savings scheme.

Table 1.3: Reserve $U_t = V_t$ for bank savings account and for life endowment insurance

t :	0	4	9	14
Savings account:	0.04604	0.25188	0.56577	0.95694
Life endowment:	0.03743	0.21008	0.49812	0.92523

For the life assurance described in Paragraph B we obtain similarly that the individual reserve develops as as shown in Table 1.4.

D. Insurance risk in a finite portfolio. The perfect balance in (1.20) and (1.22) rests on the hypothesis that the decrement series ℓ_{x+t} follows the pattern of an infinitely large portfolio. In a finite portfolio, however, the factual numbers of survivors, L_{x+t}, will be subject to randomness and will be determined by

Table 1.4: Reserve $U_t = V_t$ for a term life assurance of 1 against level premium in 15 years from age 55

t :	0	4	9	14
	0.01701	0.04460	0.06010	0.03170

the survival probabilities p_{x+t} (some of which are) shown in Table 1.2. The difference between discounted premiums and discounted benefits,

$$D = \sum_{j=0}^{14} (1+i)^{-j} 0.0374\, L_{55+j} \; - \; (1+i)^{-15}\, L_{70}\,,$$

will be a random quantity. It will have expected value 0, and its standard deviation measures how much insurance risk is left due to "imperfect diversification" in a finite portfolio. An easy exercise in probability calculus shows that the standard deviation of D/L_{55} is $\frac{1}{\sqrt{L_{55}}}0.1685$. It tends to 0 as L_{55} goes to infinity.

For the term insurance contract the corresponding quantity is $\frac{1}{\sqrt{L_{55}}}0.3478$, indicating that term insurance is a more risky business than life endowment.

1.5 With-profit contracts: Surplus and bonus

A. With-profit contracts. Insurance policies are long term contracts, with time horizons wide enough to capture significant variations in interest and mortality. For simplicity we shall focus on interest rate uncertainty and assume that the mortality law remains unchanged over the term of the contract. We will discuss the issue of surplus and bonus in the framework of the life endowment contract considered in Paragraph 1.4.A.

At time 0, when the contract is written with benefits and premiums binding to both parties, the future development of the interest rates i_t is uncertain, and it is impossible to foresee what premium level c will establish the required equivalence

$$\sum_{j=0}^{14} {S_j}^{-1} c\, \ell_{55+j} = {S_{15}}^{-1} 1\, \ell_{70}\,, \tag{1.25}$$

with

$$S_j = (1+i_1)\cdots(1+i_j)\,.$$

If it should turn out that, due to adverse development of interest and mortality, premiums are insufficient to cover the benefit, then there is no way the insurance company can avoid a loss; it cannot reduce the benefit and it cannot increase the premiums since these were irrevocably set out in the contract at time 0. The only way the insurance company can prevent such a loss, is to charge a premium 'on the safe side', high enough to be adequate under all likely scenarios. Then, if everything goes well, a surplus will accumulate. This surplus belongs to the insured and is to be repaid as so-called *bonus*, e.g. as increased benefits or reduced premiums.

B. First order basis. The usual way of setting premiums to the safe side is to base the calculation of the premium level and the reserves on a provisional *first order basis*, assuming a fixed annual interest rate i^*, which represents a worst case scenario and leads to higher premium and reserves than are likely to be needed. The corresponding accumulation factor is $S_t^* = (1 + i^*)^t$. The individual reserve based on the prudent first order assumptions is called the *first order reserve*, and we denote it by V_t^* as before. The premiums are determined so as to satisfy equivalence under the first order assumption.

C. Surplus. At any time t we define the technical surplus Q_t as the difference between the retrospective reserve under the factual interest development and the retrospective reserve under the first order assumption:

$$
\begin{aligned}
Q_t &= S_t \sum_{j=0}^{t} S_j^{-1}\, c\, \ell_{55+j} \;-\; S_t^* \sum_{j=0}^{t} S_j^{*\,-1}\, c\, \ell_{55+j} \\
&= S_t \sum_{j=0}^{t-1} S_j^{-1}\, c\, \ell_{55+j} \;-\; S_t^* \sum_{j=0}^{t-1} S_j^{*\,-1}\, c\, \ell_{55+j}\,.
\end{aligned}
$$

Setting $S_t = S_{t-1}(1+i_t)$ and $S_t^* = S_{t-1}^*(1+i^*)$, writing $1+i^* = 1+i_t-(i_t-i^*)$ in the latter, and rearranging a bit, we find that Q_t develops as

$$
Q_t \;=\; Q_{t-1}(1+i_t) \;+\; V_{t-1}^*\,(i_t-i^*)\,\ell_{55+t-1}\,, \tag{1.26}
$$

commencing from

$$
Q_0 = 0\,.
$$

The contribution to the technical surplus in year t is

$$
V_{t-1}^*\,(i_t-i^*)\,\ell_{55+t-1}\,,
$$

which is easy to interpret: it is precisely the interest earned on the reserve in excess of what has been assumed under the prudent first order assumption.

The surplus is to be redistributed as *bonus*. Several bonus schemes are used in practice. One can repay currently the contribution $V_{t-1}^*\,(i_t-i^*)\,\ell_{55+t-1}$ as so-called *cash bonus* (a premium deductible), whereby each survivor at time t will receive

$$
V_{t-1}^*\,(i_t-i^*)\,\ell_{55+t-1}/\ell_{55+t}\,.
$$

Another possibility is to postpone repayment until the term of the contracts and grant so-called *terminal bonus* to the survivors (an added benefit), the amount per survivor being b^+ given by

$$S_{15} \sum_{j=1}^{15} S_j^{-1} \, V^*_{j-1} \, (i_j - i^*) \, \ell_{55+j} \; = \; \ell_{70} b^+ \, .$$

Between these two solutions there are countless other possibilities. In any case, the point is that the financial risk can be eliminated: the insurer observes the development of the factual interest and only in arrears repays the insured so as to restore equivalence on the basis of the factual interest rate development. This works well provided the first order interest rate is set on the safe side so that $i_t \geq i^*$ for all t.

There is a problem, however: Negative bonus can never be applied. Therefore the insurer will suffer a loss if the factual interest falls short of the technical interest rate. In this perspective cash bonus is the most risky solution and terminal bonus is the least risky solution.

If the financial market is sufficiently rich in assets, then the interest rate guarantee that is thus inherent in the with-profit policy can be priced, and the insured can be charged an extra premium to cover it. This would ultimately eliminate the financial risk by diversifying, not only the insurance portfolio, but also the investment portfolio.

1.6 Unit-linked insurance

A quite different way of going about the financial risk is the so-called unit-linked contract. As the name indicates, the idea is to relate payments directly to the development of the investment portfolio, i.e. the interest rate. Consider the balance equation for an endowment of b against premium c_t in year $t = 1, \ldots, 14$:

$$S_{15} \sum_{t=0}^{14} S_t^{-1} \, c_t \, \ell_{55+t} \; - \; b \, \ell_{70} \; = \; 0 \, . \tag{1.27}$$

A perfect link between payments and investments performance is obtained by letting the premiums and the benefit be inflated by the index S,

$$c_t = S_t \, c \, ,$$

and

$$b_{15} = S_{15} \, 1 \, .$$

Here c is a baseline premium, which is to be determined. Then (1.27) becomes

$$S_{15} \sum_{t=0}^{14} S_t^{-1} \, S_t \, c \, \ell_{55+t} \; - \; S_{15} \, \ell_{70} \; = \; 0 \, ,$$

which reduces to

$$\sum_{t=0}^{14} c \, \ell_{55+t} \; - \; \ell_{70} \; = \; 0 \, ,$$

and we find

$$c = \frac{\ell_{70}}{\sum_{t=0}^{14} \ell_{55+t}}\,.$$

Again financial risk has been perfectly eliminated and diversification of the insurance portfolio is sufficient to establish balance between premiums and benefits.

Perfect linking as defined here is not common in practice. Presumably, remnants of social welfare thinking have led insurers to modify the unit-linked concept in various ways, typically by introducing a guarantee on the sum insured to the effect that it cannot be less than 1 (say). Also the premium is usually not index-linked. Under such modified variations of the unit-linked policy one cannot in general obtain balance by the simple device above. However, the problem can in principle be resolved by calculating the price of the financial claim thus introduced and to charge the insured with the needed additional premium.

1.7 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 existing 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 working with this concept in a finite portfolio, with imperfect diversification of insurance risk, faces a risk of insolvency as indicated in Paragraph 1.3.D. In addition comes the financial risk, and ways of getting around that were indicated in Sections 1.5 and 1.6. The total risk 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: [6] (a good classic – sharpen your German!), [4], [29] (lexicographic, treats virtually every variation of standard insurance products, and includes a good chapter on population theory), [45] (an excellent early text based on probabilistic models, placing emphasis on risk considerations), [11], [15] (an original approach to the field – sharpen your French!), and [23] (the most recent of the mentioned texts, still classical in its orientation).

Chapter 2

Payment streams and interest

2.1 Basic definitions and relationships

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 will give precise mathematical content to the notion of payment streams and, referring to Appendix A, we deal only with their properties as functions of time and do 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 n ($\leq \infty$), 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 at 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 piece-wise differentiable (PD):

$$A_t = A_0 + \int_0^t a_\tau \, d\tau + \sum_{0<\tau\leq t} \Delta A_\tau \,, \tag{2.1}$$

where $\Delta A_\tau = A_\tau - A_{\tau-}$. The integral adds up payments that fall due continuously, and the sum adds up lump sum payments. In differential form (2.1) reads

$$dA_t = a_t\,dt + \Delta A_t\,. \tag{2.2}$$

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 [31] and [17], 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 s gives the account holder the right to cash, at any other time t, a certain amount $S(s,t)$, typically different from 1. The function S must be strictly positive, and we shall argue that it must satisfy the functional relationship

$$S(s,u) = S(s,t)\,S(t,u)\,, \tag{2.3}$$

implying, of course, that $S(t,t) = 1$ (put $s = t = u$ and use strict positivity): If the account holder invests 1 at time s, he may cash the amount on the left of (2.3) at time u. If he instead withdraws the value $S(s,t)$ at time t and immediately reinvests it again, he will obtain the amount on the right of (2.3) at time u. To avoid arbitrary gains, so-called *arbitrage*, the two strategies must give the same result.

It is easy to verify that the function $S(s,t)$ satisfies (2.3) if and only if it is of the form

$$S(s,t) = \frac{S_t}{S_s} \tag{2.4}$$

for some strictly positive function S_t (allowing an abuse of notation), which can be taken to satisfy

$$S_0 = 1\,.$$

Then S_t must be the value at time t of a unit deposited at time 0, and we call it the *accumulation function*. Correspondingly, S_t^{-1} is the value at time 0 of a unit withdrawn at time t, and we call it the *discount function*.

We will henceforth assume that S_t is of the form

$$S_t = e^{\int_0^t r}\,, \qquad S_t^{-1} = e^{-\int_0^t r}\,, \tag{2.5}$$

where r_t is some piece-wise continuous function, usually positive. (The shorthand exemplified by $\int r = \int r_\tau\,d\tau$ will be in frequent use throughout.) Accumulation factors of this form are invariably used in basic banking operations (loans and savings) and also for bonds issued by governments and corporations.

Under the rule (2.5) the dynamics of accumulation and discounting are given by

$$dS_t = S_t\, r_t\, dt\,, \tag{2.6}$$

$$dS_t^{-1} = -S_t^{-1}\, r_t\, dt\,. \tag{2.7}$$

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) reads

$$S_t = S_s + \int_s^t S_\tau\, r_\tau\, d\tau\,, \qquad s \le t\,, \tag{2.8}$$

and (2.7) reads $S_u^{-1} = S_t^{-1} - \int_t^u S_\tau^{-1}\, r_\tau\, d\tau$ or

$$S_t^{-1} = S_u^{-1} + \int_s^t S_\tau^{-1}\, r_\tau\, d\tau\,, \qquad t \le u\,. \tag{2.9}$$

We will be working with the expressions

$$S(s,t) = e^{-\int_s^t r}$$

for the general *discount factor* when $t \le s$ and

$$S(t,u) = e^{\int_t^u r}$$

for the general *accumulation factor* when $t \le u$.

By constant interest rate r we have $S_t = e^{rt}$ and and $S_t^{-1} = e^{-rt}$. Upon introducing the the *annual interest rate*

$$i = e^r - 1\,, \tag{2.10}$$

whereby the *annual accumulation factor* is $S_1 = (1+i)$, and the *annual discount factor*

$$v = e^{-r} = \frac{1}{1+i}\,, \tag{2.11}$$

we have

$$S_t = (1+i)^t\,, \qquad S_t^{-1} = v^t\,. \tag{2.12}$$

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_τ paid in the small time interval around time τ is $e^{\int_0^t r} e^{-\int_0^\tau r}\, dA_\tau$.

Summing over all time intervals we get the value at time t of the entire payment stream,

$$e^{\int_0^t r} \int_{0-}^{n} e^{-\int_0^\tau r} \, dA_\tau = U_t - V_t \, ,$$

where

$$U_t = e^{\int_0^t r} \int_{0-}^{t} e^{-\int_0^\tau r} \, dA_\tau = \int_{0-}^{t} e^{\int_\tau^t r} \, dA_\tau \tag{2.13}$$

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

$$V_t = e^{\int_0^t r} \int_{0-}^{t} e^{-\int_0^\tau r} \, dB_\tau = \int_{t}^{n} e^{-\int_t^\tau r} \, dB_\tau \tag{2.14}$$

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 (2.13), taking $X_t = e^{\int_0^t r}$ (continuous, with dynamics given by (2.6)) and $Y_t = \int_{0-}^t e^{-\int_0^\tau r} \, dA_\tau$, yields

$$dU_t = U_t r_t \, dt + dA_t \, , \tag{2.15}$$

By definition,

$$U_0 = A_0 \, . \tag{2.16}$$

Integrating (2.15) from 0 to t, using the initial condition (2.16), gives

$$U_t = A_t + \int_0^t U_\tau r_\tau \, d\tau \, . \tag{2.17}$$

An alternative expression,

$$U_t = A_t + \int_0^t e^{\int_\tau^t r} A_\tau \, r_\tau \, d\tau \, , \tag{2.18}$$

is derived from (2.13) upon applying the rule (A.9) of integration by parts:

$$\begin{aligned} \int_{0-}^{t} e^{-\int_0^\tau r} \, dA_\tau &= A_0 + \int_0^t e^{-\int_0^\tau r} \, dA_\tau \\ &= A_0 + e^{-\int_0^t r} A_t - A_0 - \int_0^t A_\tau e^{-\int_0^\tau r} (-r_\tau) \, d\tau \, . \end{aligned}$$

The relationships (2.15) – (2.18) show how the cash balance emerges from payments and earned interest. They are easy to interpret and can be read aloud in non-mathematical terms.

It follows from (2.18) that, if $r \geq 0$, then an increase of A results in an increase of U. In particular, advancing payments of a given amount produces a bigger cash balance.

Likewise, from (2.14) we derive

$$dV_t = V_t r_t \, dt - dB_t \,. \tag{2.19}$$

By definition, if $n < \infty$,

$$V_n = 0 \,, \tag{2.20}$$

or, setting $\Delta B_n = B_n - B_{n-}$,

$$V_{n-} = \Delta B_n \,. \tag{2.21}$$

Integrating (2.19) from t to n, using the ultimo condition (2.20), gives the following analogue to (2.17):

$$V_t = B_n - B_t - \int_t^n V_\tau \, r_\tau \, d\tau \,. \tag{2.22}$$

The analogue to (2.18) is

$$V_t = B_n - B_t - \int_t^n e^{-\int_t^\tau r} (B_n - B_\tau) \, r_\tau \, d\tau \,. \tag{2.23}$$

The last two relationships are valid for $n = \infty$ only if $B_\infty < \infty$. Yet another expression is

$$V_t = e^{-\int_t^n r} (B_n - B_t) + \int_t^n e^{-\int_t^\tau r} (B_\tau - B_t) \, r_\tau \, d\tau \,, \tag{2.24}$$

which is obtained upon integrating by parts in (2.23) or, simpler, multiplying $B_n - B_t$ with

$$1 = e^{-\int_t^n r} - \int_t^n e^{-\int_t^\tau r} \, r_\tau \, d\tau$$

(a twist on (2.9)) and gathering terms.

Again interpretations are easy. The relations (2.22) and (2.23) 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. The relation (2.24) states that repayment can be postponed until the term of the contract at the expense of paying interest on the outstanding amounts meanwhile. It follows from (2.24) that, if $r \geq 0$, then an increase of the outstanding payments produces an increase in the reserve. In particular, advancing payments of a given amount leads to a bigger reserve.

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

$$U_n = 0 \quad \text{ or } \quad V_{0-} = 0 \,. \tag{2.25}$$

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

$$U_t = V_t \, . \tag{2.26}$$

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 known 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

$$\varepsilon_n(t) = \begin{cases} 0, & 0 \le t < n \, , \\ 1, & t \ge n \, . \end{cases} \tag{2.27}$$

(The only payment is $\Delta A_n = 1$.) 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

$$a_{\overline{n}|} = \sum_{j=1}^{n} e^{-rj} = \frac{1 - e^{-rn}}{i} \, , \tag{2.28}$$

see (2.11) – (2.10).

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

$$\ddot{a}_{\overline{n}|} = \sum_{j=0}^{n-1} e^{-rj} = (1+i)\, a_{\overline{n}|} = (1+i) \frac{1 - e^{-rn}}{i} \, . \tag{2.29}$$

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

$$A_t = t \wedge n \, . \tag{2.30}$$

For the case with constant interest rate its present value at time 0 is (recall (2.11))

$$\bar{a}_{\overline{n}|} = \int_0^n e^{-r\tau}\,d\tau = \frac{1-e^{-rn}}{r}\,. \tag{2.31}$$

An everlasting (perpetual) annuity is called a *perpetuity.* Putting $n=\infty$ in the (2.28), (2.29), and (2.31), we find the following expressions for the present values of the immediate perpetuity, the perpetuity-due, and the continuous perpetuity:

$$a_{\overline{\infty}|} = \frac{1}{i}\,,\quad \ddot{a}_{\overline{\infty}|} = \frac{1+i}{i}\,,\quad \bar{a}_{\overline{\infty}|} = \frac{1}{r}\,. \tag{2.32}$$

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,

$$A_t = ((t-m)\vee 0)\wedge n = (t\wedge(m+n)) - (t\wedge m)\,. \tag{2.33}$$

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

$$\bar{a}_{m|n} = \bar{a}_{\overline{m+n}|} - \bar{a}_{\overline{m}|} = v^m \bar{a}_{\overline{n}|}\,. \tag{2.34}$$

2.2 Application to loans

A. Basic features of a loan contract. Loans and saving accounts in banks are particularly simple financial contracts for which interest is invariably calculated in accordance with (2.5). Let us consider a loan contract stipulating that at time 0, say, the bank pays to the 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 $\{A_t\}_{0\le t\le n}$ 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 $A_n \ge 1$. 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,

$$A_t = F_t + R_t\,, \tag{2.35}$$

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,

$$\int_0^n e^{-\int_0^\tau r}\, dA_\tau = 1\,. \tag{2.36}$$

There are, of course, infinitely many admissible decompositions (2.35) satisfying (2.36). A clue to constraints on F and R is offered by the relationship

$$\int_0^n e^{-\int_0^\tau r}\, dR_\tau = \int_0^n e^{-\int_0^\tau r}(1-F_\tau)\, r_\tau\, d\tau\,, \tag{2.37}$$

which is obtained upon inserting (2.35) into (2.36) and then using integration by parts on the term $\int_0^n \exp\left(-\int_0^\tau r\right)\, dF_\tau = -\int_0^n \exp\left(-\int_0^\tau r\right)\, d(1-F_\tau)$. The condition (2.37) is trivially satisfied if

$$dR_t = (1-F_t)\, r_t\, dt\,,$$

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

Under the scheme of natural interest the relation (2.35) becomes

$$dA_t = dF_t + (1-F_t)\, r_t\, dt\,, \tag{2.38}$$

which establishes a one-to-one correspondence between amortizations and repayments. The differential equation (2.38) is easily solved:

First, integrate (2.38) over $(0, t]$ to obtain

$$A_t = F_t + \int_0^t (1-F_\tau)\, r_\tau\, d\tau\,, \tag{2.39}$$

which determines amortizations when repayments are given.

Second, multiply (2.38) 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

$$\int_t^n e^{-\int_t^\tau r} dA_\tau = 1 - F_t\,, \tag{2.40}$$

which determines (outstanding) repayments when amortizations are given.

The relationships (2.39) and (2.40) are easy to interpret. For instance, since $1-F_t$ is the remaining debt at time t, (2.40) is the time t update of the equivalence requirement (2.36). When it comes to numerical computation, the integral expressions in (2.39) and (2.40) are not so useful, however. Whether we want to compute A for given F or the other way around, we would use the differential equation (2.38).

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 (2.27). The amortization function is obtained directly from (2.39): $A_t = \varepsilon_n(t) + rt$.

A *series loan* has repayments of annuity form. The continuous version is given by $F_t = t/n$, see (2.30). The amortization plan is obtained from (2.39): $A_t = t/n + rt(1 - t/2n)$. 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 $A_t = t/\bar{a}_{\overline{n}|}$, see (2.36) and (2.31). From (2.40) we easily obtain $F_t = 1 - \bar{a}_{\overline{n-t}|}/\bar{a}_{\overline{n}|}$. We find $dF_t/dt = e^{-r(n-t)}/\bar{a}_{\overline{n}|}$ (exponentially increasing), and $dR_t/dt = (1 - e^{-r(n-t)})/\bar{a}_{\overline{n}|}$.

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 represented by a non-negative random variable T with a cumulative distribution function

$$F(t) = \mathbb{P}[T \leq t] \,. \tag{3.1}$$

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

$$\bar{F}(t) = \mathbb{P}[T > t] = 1 - F(t) \,. \tag{3.2}$$

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;

$$f(t) = \frac{d}{dt} F(t) = -\frac{d}{dt} \bar{F}(t) \,. \tag{3.3}$$

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?

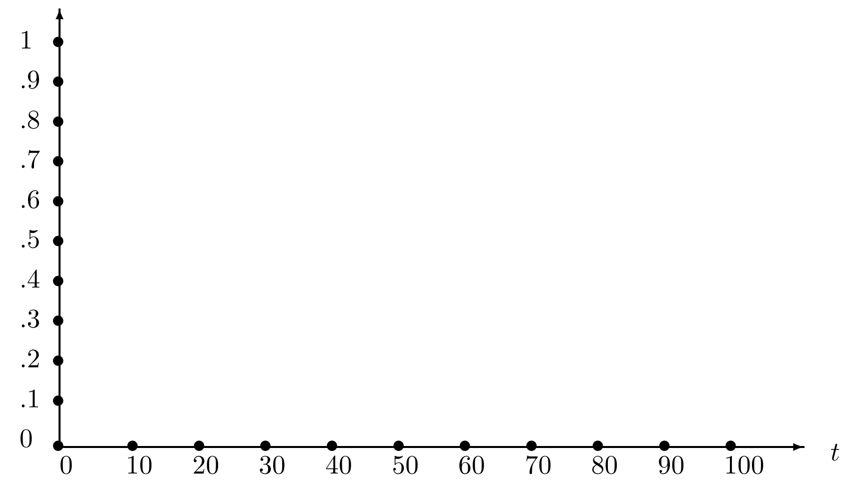


Figure 3.1: The G82M mortality law: F broken line, $\bar{F}$ whole line.

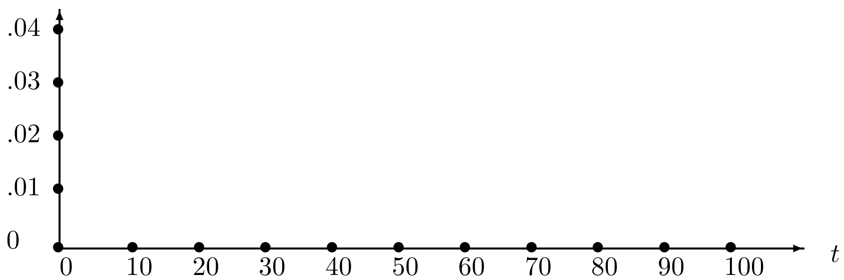


Figure 3.2: The density f for the G82M mortality law.

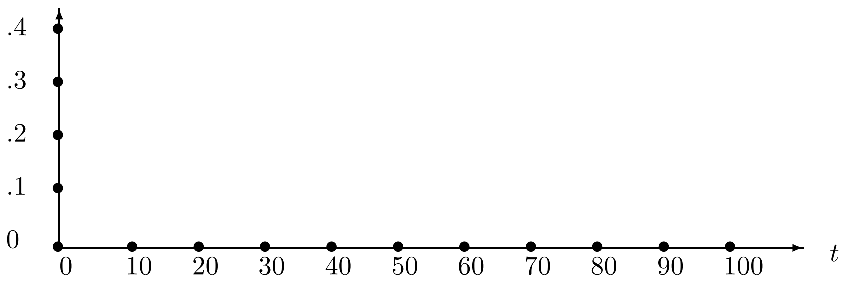


Figure 3.3: The force of mortality μ for the G82M mortality law.

B. The force of mortality. The density is the derivative of $-\bar{F}$, see (3.3). When dealing with non-negative random variables representing life lengths, it is convenient to work with the derivative of $-\ln \bar{F}$,

$$\mu(t) = \frac{d}{dt}\{-\ln \bar{F}(t)\} = \frac{f(t)}{\bar{F}(t)}\,, \tag{3.4}$$

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

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

(In the second equality we have neglected a term $o(dt)$ such that $o(dt)/dt \to 0$ as $dt \searrow 0$.) 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 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 $\bar{F}(0) = 1$, we obtain

$$\bar{F}(t) = e^{-\int_0^t \mu}. \tag{3.5}$$

Relation (3.4) may be cast as

$$f(t) = \bar{F}(t)\mu(t) = e^{-\int_0^t \mu}\mu(t)\,, \tag{3.6}$$

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

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

Since $\bar{F}(\infty) = 0$, we must have $\int_0^\infty \mu = \infty$. Thus, if there is a finite highest attainable age ω such that $\bar{F}(\omega) = 0$ and $\bar{F}(t) > 0$ for $t < \omega$, then $\int_0^t \mu \nearrow \infty$ as $t \nearrow \omega$. If, moreover, μ is non-decreasing, we must also have $\lim_{t\nearrow\omega} \mu(t) = \infty$.

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

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

and survival function

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

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

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

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

$$\mu(t \,|\, x) = f(x+t)/\bar{F}(x+t) = \mu(x+t)\,. \tag{3.9}$$

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

$$\bar{F}(t|x) = e^{-\int_x^{x+t} \mu(y)\,dy} = e^{-\int_0^t \mu(x+\tau)d\tau}\,, \tag{3.10}$$

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 random variable with absolutely continuous distribution function F, and let $G : \mathbb{R}_+ \to \mathbb{R}$ be a PD and RC function such that $\mathbb{E}[G(T)]$ exists and is finite. Integrating by parts in the defining expression

$$\mathbb{E}[G(T)] = \int_0^\infty G(\tau)\,dF(\tau)\,,$$

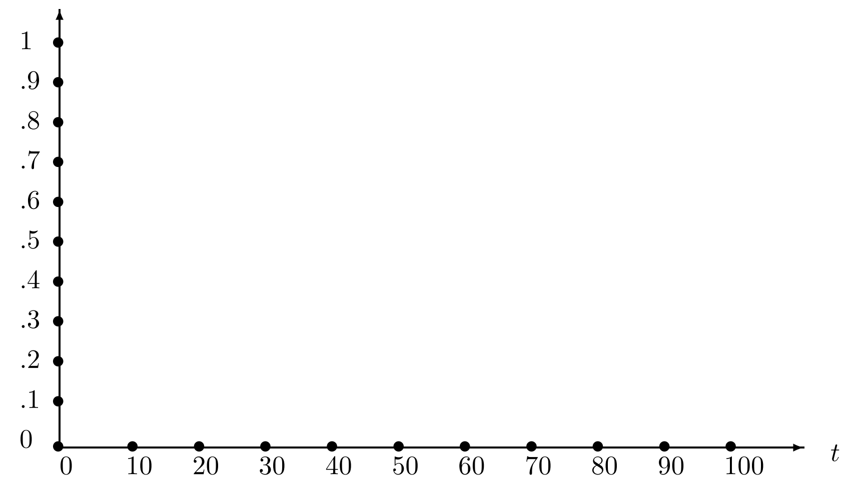


Figure 3.4: Conditional distribution of remaining life length for the G82M mortality law: $F(t|70)$ broken line, $\bar{F}(t|70)$ whole line.

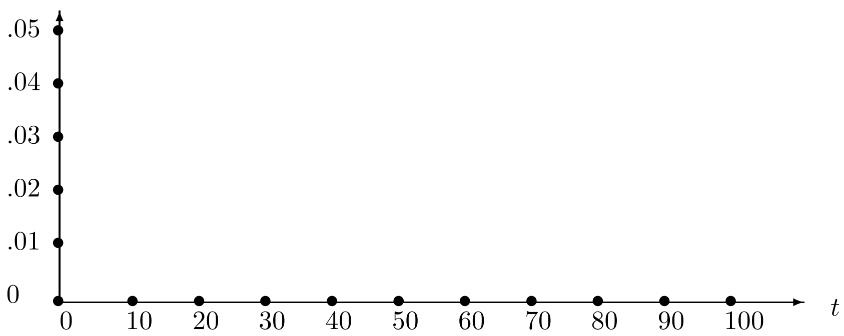


Figure 3.5: Conditional density of remaining life length $f(t|70)$ for the G82M mortality law.

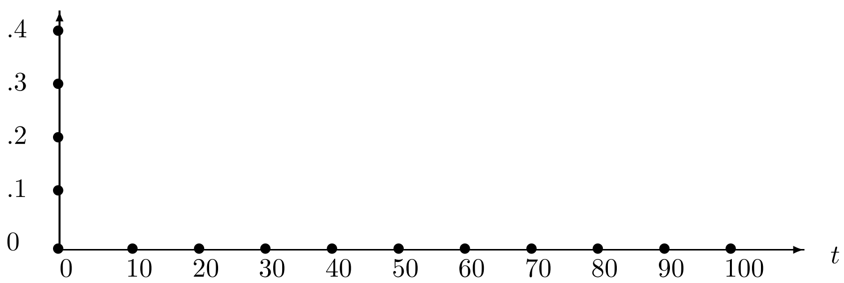


Figure 3.6: The force of mortality $\mu(t|70) = \mu(70 + t)$, $t > 0$, for the G82M mortality law.

we find

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

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

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

and, in particular,

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

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

$$\bar{e}_x = \int_0^\infty \bar{F}(t \,|\, x)\, dt\,. \tag{3.14}$$

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

Consider the more general function

$$G(t) = ((t \wedge b) - (t \wedge a))^k = \begin{cases} 0\,, & 0 \le t < a, \\ (t-a)^k\,, & a \le t < b, \\ (b-a)^k\,, & b \le t, \end{cases} \tag{3.15}$$

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 kth power of the number of years lived between age a and age b. From (3.11) we obtain

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

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

$$\bar{F}(t) = e^{-\lambda t}\,, \tag{3.17}$$

Figure 3.7: Two exponential laws with intensities λ_1 and λ_2 such that $\lambda_1 < \lambda_2$; $\bar{F}_1$ and $\bar{F}_2$ whole line, f_1 and f_2 broken line.

and the density (3.6) becomes

$$f(t) = \lambda e^{-\lambda t} \,. \tag{3.18}$$

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

$$\bar{F}(t\,|\,x) = \bar{F}(t) \,, \tag{3.19}$$

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

$$\bar{F}(x+t) = \bar{F}(x)\bar{F}(t) \tag{3.20}$$

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

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

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 λ.

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 $\bar{F}(t) = \exp\left(-(\frac{t}{\alpha})^{\beta}\right)$.

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 α and β!

We have $\mu(x+t) = \beta\alpha^{-\beta}(x+t)^{\beta-1}$, and, by virtue of (3.22), $\bar{F}(t\,|\,x)$ 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

$$\mu(t) = \alpha + \beta e^{\gamma t}\,, \tag{3.23}$$

$\alpha, \beta \geq 0$. The corresponding survival function is

$$\bar{F}(t) = \exp\left(-\int_0^t (\alpha + \beta e^{\gamma s})ds\right) = \exp\left(-\alpha t - \beta(e^{\gamma t} - 1)/\gamma\right)\,. \tag{3.24}$$

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

We have $\mu(x+t) = \alpha + \beta e^{\gamma x}e^{\gamma t}$, and so (3.23) shows that $\bar{F}(t|x)$ is also a G-M survival function with parameters α, $\beta e^{\gamma x}$, γ. 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

$$\alpha = 5 \cdot 10^{-4}\,, \qquad \beta = 7.5858 \cdot 10^{-5}\,, \qquad \gamma = \ln(1.09144)\,. \tag{3.25}$$

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 standard notation, 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.

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

$$
\begin{aligned}
{}_tq_x &= F(t\,|\,x)\,, && (3.26)\\
{}_tp_x &= \bar{F}(t\,|\,x)\,, && (3.27)\\
\mu_{x+t} &= \mu(x+t)\,. && (3.28)
\end{aligned}
$$

In particular, ${}_tq_0 = F(t)$ and ${}_tp_0 = \bar{F}(t)$. One-year death and survival probabilities are abbreviated as

$$q_x = {}_1q_x\ , \quad p_x = {}_1p_x\ . \tag{3.29}$$

Frequently used is also the "n-year deferred probability of death within m years",

$${}_{n|m}q_x = {}_{m+n}q_x - {}_nq_x = {}_np_x - {}_{m+n}p_x = {}_np_x\,{}_mq_{x+n}\ . \tag{3.30}$$

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

$$
\begin{aligned}
{}_tp_x &= \exp(-\int_0^t \mu_{x+\tau}\,d\tau), && (3.31)\\
f(t\,|\,x) &= {}_tp_x\mu_{x+t}, && (3.32)\\
\bar{e}_x &= \int_0^\infty {}_tp_x dt. && (3.33)
\end{aligned}
$$

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

$$l_x = l_0\,{}_xp_0\ . \tag{3.34}$$

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.

$$
\begin{aligned}
{}_tp_x &= l_{x+t}/l_x\ , && (3.35)\\
\mu_{x+t} &= -l'_{x+t}/l_{x+t}\ , && (3.36)\\
f(t\,|\,x) &= -l'_{x+t}/l_x\ , && (3.37)\\
\bar{e}_x &= \int_0^\infty l_{x+t}\,dt/l_x\ . && (3.38)
\end{aligned}
$$

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 [29] 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 eligible to insurance 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

$$
\begin{aligned}
{}_{\tau}q_{[x]+t} &= \mathbb{P}[T_x \le t+\tau \,|\, T_x > t]\,, && (3.39)\\
{}_{\tau}p_{[x]+t} &= \mathbb{P}[T_x > t+\tau \,|\, T_x > t]\,, && (3.40)\\
\mu_{[x]+t} &= \lim_{h\searrow 0} \frac{{}_{h}q_{[x]+t}}{h}\,. && (3.41)
\end{aligned}
$$

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 ${}_{t}q_x$, 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

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

where α and β 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, $I_t = 1[T_x > t]$. This is a binomial random variable with 'success' probability ${}_tp_x$. 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 ${}_tq_x = 1 - {}_tp_x$. (We apologize for sometimes using technical terms where they may sound misplaced.) 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 basic actuarial methods and formulas.

We assume that the investments of the insurance company yield interest at a fixed rate r so that accumulation and discounting take place in in accordance with (2.12).

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. In other words, the associated payment function is an endowment of I_n at time n. Its present value at time 0 is

$$PV^{e;n} = e^{-rn} I_n \,. \tag{4.1}$$

The expected value of $PV^{e;n}$, denoted by ${}_nE_x$, is

$${}_nE_x = e^{-rn}\, {}_np_x \,. \tag{4.2}$$

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

$$\mathbb{E}[(PV^{e;n})^q] = {}_nE_x^{(qr)} \,, \tag{4.3}$$

where the top-script (qr) signifies that discounting is made under a force of interest that is qr.

In particular, the variance of $PV^{e;n}$ is

$$\mathbb{V}[PV^{e;n}] = {}_nE_x^{(2r)} - {}_nE_x^2 \,. \tag{4.4}$$

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

$$PV^{ti;n} \quad = \quad e^{-rT_x}\,(1 - I_n)\,. \tag{4.5}$$

The expected value of $PV^{ti;n}$ is

$$\bar{A}_{\overset{1}{x}\,\overline{n}|} = \int_0^n e^{-r\tau}\,{}_\tau p_x\,\mu_{x+\tau}\,d\tau\,, \tag{4.6}$$

and, similar to (4.3),

$$\mathbb{E}[(PV^{ti;n})^q] = \bar{A}^{(qr)}_{\overset{1}{x}\,\overline{n}|}\,. \tag{4.7}$$

In particular,

$$\mathbb{V}[PV^{ti;n}] = \bar{A}^{(2r)}_{\overset{1}{x}\,\overline{n}|} - \bar{A}^{2}_{\overset{1}{x}\,\overline{n}|}\,. \tag{4.8}$$

An n-year *endowment insurance* is payable upon death if it occurs within time 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

$$PV^{ei;n} = e^{-r(T_x \wedge n)}\,. \tag{4.9}$$

The expected value of $PV^{ei;n}$ is

$$\bar{A}_{x\,\overline{n}|} = \int_0^n e^{-r\tau}\,{}_\tau p_x\,\mu_{x+\tau}\,d\tau + e^{-rn}\,{}_np_x = \bar{A}_{\overset{1}{x}\,\overline{n}|} + {}_nE_x\,, \tag{4.10}$$

and

$$\mathbb{E}(PV^{ei;n})^q = \bar{A}^{(qr)}_{x\,\overline{n}|}\,. \tag{4.11}$$

It follows that

$$\mathbb{V}[PV^{ei;n}] = \bar{A}^{(2r)}_{x\,\overline{n}|} - \bar{A}^{2}_{x\,\overline{n}|}\,. \tag{4.12}$$

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 (2.30), the associated payment function is an annuity of 1 in $T_x \wedge n$ years. Its present value at time 0 is

$$PV^{a;n} = \bar{a}_{\overline{T_x \wedge n}|} = \frac{1 - e^{-r(T_x \wedge n)}}{r} . \tag{4.13}$$

The expected value of $PV^{a;n}$ is

$$\bar{a}_{x\,\overline{n}|} = \int_0^n \bar{a}_{\overline{\tau}|}\, {}_\tau p_x\, \mu_{x+\tau}\, d\tau + \bar{a}_{\overline{n}|}\, {}_n p_x .$$

A more appealing formula is

$$\bar{a}_{x\,\overline{n}|} = \int_0^n e^{-r\tau}\, {}_\tau p_x\, d\tau , \tag{4.14}$$

which displays the life annuity as a "continuum of life endowments", $\bar{a}_{x\,\overline{n}|} = \int_0^n {}_\tau E_x\, d\tau$. There are several ways of proving (4.14). Using brute force, one can integrate by parts:

$$\begin{aligned}
\bar{a}_{\overline{n}|}\, {}_n p_x &= \bar{a}_{\overline{0}|}\, {}_0 p_x + \int_0^n \frac{d}{d\tau} \bar{a}_{\overline{\tau}|}\, {}_\tau p_x\, d\tau + \int_0^n \bar{a}_{\overline{\tau}|} \frac{d}{d\tau} {}_\tau p_x\, d\tau \\
&= \int_0^n e^{-r\tau}\, {}_\tau p_x\, d\tau - \int_0^n \bar{a}_{\overline{\tau}|}\, {}_\tau p_x\, \mu_{x+\tau}\, d\tau .
\end{aligned}$$

Using the brain instead, one realizes that the expected present value at time 0 of the payments in any small time interval $(\tau, \tau + d\tau)$ is $e^{-r\tau}\, d\tau\, {}_\tau p_x$, and summing over all time intervals one arrives at (4.14) ("the expected value of a sum is the sum of the expected values"). This kind of reasoning will be omnipresent throughout the text, and would also immediately produce formula (4.6) and (4.10). The recipe is: *Find the expected present value of the payments in each small time interval and add up.*

We shall demonstrate below that

$$\mathbb{E}[(PV^{a;n})^q] = \frac{q}{r^{q-1}} \sum_{p=1}^{q} (-1)^{p-1} \binom{q-1}{p-1} \bar{a}^{(pr)}_{x\,\overline{n}|} , \tag{4.15}$$

from which we derive

$$\mathbb{V}[PV^{a;n}] = \frac{2}{r} \left(\bar{a}_{x\,\overline{n}|} - \bar{a}^{(2r)}_{x\,\overline{n}|} \right) - \bar{a}^2_{x\,\overline{n}|}. \tag{4.16}$$

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

$$PV^{a;n} = \frac{1 - PV^{ei;n}}{r} \quad \text{or} \quad PV^{ei;n} = 1 - rPV^{a;n} , \tag{4.17}$$

which just reflects the more general relationship (2.31). Taking expectation in (4.17), we get

$$\bar{A}_{x\,\overline{n}|} = 1 - r\bar{a}_{x\,\overline{n}|}\,. \tag{4.18}$$

Also, since $PV^{ti;n} = PV^{ei;n} - PV^{e;n} = 1 - rPV^{a;n} - PV^{e;n}$, we have

$$\bar{A}_{\overset{1}{x}\,\overline{n}|} = 1 - r\bar{a}_{x\,\overline{n}|} - {}_nE_x\,. \tag{4.19}$$

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

$$\bar{a}_x = \int_0^\infty e^{-r\tau}\, {}_\tau p_x \, d\tau, \tag{4.20}$$

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

$$\begin{aligned} PV &= PV^{a;m+n} - PV^{a;m} = \bar{a}_{\overline{T_x \wedge (m+n)}|} - \bar{a}_{\overline{T_x \wedge m}|} \\ &= \frac{e^{-r(T_x\wedge m)} - e^{-r(T_x \wedge (m+n))}}{r} \end{aligned} \tag{4.21}$$

The expected present value is

$$_{m|n}\bar{a}_x = \bar{a}_{x\,\overline{m+n}|} - \bar{a}_{x\,\overline{m}|} = \int_m^{m+n} e^{-rt}\, {}_tp_x \, dt = {}_mE_x\, \bar{a}_{x+m\,\overline{n}|}\,. \tag{4.22}$$

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:

$$\begin{aligned} \mathbb{E}[PV] &= \mathbb{E}\,\mathbb{E}[PV \mid I_m] \\ &= {}_mp_x\, \mathbb{E}[PV \mid I_m = 1] + {}_mq_x\, \mathbb{E}[PV \mid I_m = 0] \\ &= {}_mp_x\, v^m\, \bar{a}_{x+m\,\overline{n}|}\,. \end{aligned}$$

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

$$PV = PV^{ti;m+n} - PV^{ti;m}\,, \tag{4.23}$$

and its expected present value is

$$_{m|n}\bar{A}_{\overset{1}{x}} = \bar{A}_{\overset{1}{x}\,\overline{m+n|}} - \bar{A}_{\overset{1}{x}\,\overline{m|}} = {}_mE_x\,\bar{A}_{\overset{1}{x+m}\,\overline{n|}}, . = \int_m^{m+n} e^{-r\tau}\,{}_\tau p_x\,\mu_{x+\tau}\,d\tau\,. \tag{4.24}$$

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.

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 $x = 30$, $n = 30$, μ is given by (3.25), and $r = \ln(1.045)$.

	PE	TI	EI	LA
E	0.2257	0.06834	0.2940	16.04
CV	0.4280	2.536	0.3140	0.1308
SK	−1.908	2.664	4.451	−4.451

Anyway, by methods to be developed later, we easily compute the three first moments of the present values considered above, and find their 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 $\bar{a}_{x\,\overline{n|}}$ 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.

4.2 The principle of equivalence

A. A note on terminology. Like any other good or service, insurance coverage is 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 basic 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 π. (This is admittedly an artificial assumption, but it can serve well as an approximate 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 $\pi PV^{a;m}$, with expected value $\pi\bar{a}_{x\,\overline{m|}}$ given by (4.14). We list formulas for the net level premium rate for a selection of basic forms of insurance: For the pure endowment (Paragraph 4.1.B) against level premium in the insurance period,

$$\pi = \frac{{}_nE_x}{\bar{a}_{x\,\overline{n|}}}\,. \tag{4.25}$$

For the m-year deferred n-year temporary annuity (Paragraph 4.1.E) against level premium in the deferred period,

$$\pi = \frac{{}_{m|n}\bar{a}_x}{\bar{a}_{x\,\overline{m|}}} = \frac{\bar{a}_{x\,\overline{m+n|}}}{\bar{a}_{x\,\overline{m|}}} - 1\,. \tag{4.26}$$

For the term insurance (Paragraph 4.1.C) against level premium in the insurance period,

$$\pi = \frac{\bar{A}^{\,1}_{x\,\overline{n|}}}{\bar{a}_{x\,\overline{n|}}}\,. \tag{4.27}$$

For the endowment insurance (Paragraph 4.1.C) against level premium in the insurance period,

$$\pi = \frac{\bar{A}_{x\,\overline{n|}}}{\bar{a}_{x\,\overline{m|}}} = \frac{1}{\bar{a}_{x\,\overline{n|}}} - r\,, \tag{4.28}$$

the last expression following from (4.18).

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,

$$PV = PV^b - \pi PV^{a;m}\,, \tag{4.29}$$

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

Stated precisely, the equivalence principle lays down that

$$\mathbb{E}[PV] = 0\,. \tag{4.30}$$

For example, with $PV^b = PV^{ei;n}$ (4.30) becomes $0 = \bar{A}_{x\,\overline{n|}} - \pi\bar{a}_{x\overline{m|}}$, which yields (4.28) when $m = n$.

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

$$\begin{aligned} \mathbb{V}[PV] &= \mathbb{V}\left[v^{T_x\wedge n} - \pi\,\frac{1 - v^{T_x\wedge n}}{r}\right] = (1 + \pi/r)^2\,\mathbb{V}[v^{T_x\wedge n}] \\ &= \frac{2\left(\bar{a}_{x\,\overline{n|}} - \bar{a}^{(2r)}_{x\,\overline{n|}}\right)}{r\bar{a}^2_{x\,\overline{n|}}} - 1\,. \end{aligned} \tag{4.31}$$

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 any 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. The premiums consist of a lump sum π_0 payable immediately upon the inception of the policy at time 0, and thereafter an annuity payable continuously at rate π_t per time unit contingent on survival at time $t \in (0, n)$. As before, assume that the interest rate is a deterministic function r_t.

The expected present value at time 0 of total benefits less premiums under the contract can be put up directly as the sum of the expected discounted payments in each small time interval:

$$- \pi_0 + \int_0^n e^{-\int_0^\tau r} {}_\tau p_x \{\mu_{x+\tau} b_\tau - \pi_\tau\} \, d\tau + b_n \, e^{-\int_0^n r} {}_n p_x \,. \tag{4.32}$$

Under the equivalence principle this is set equal to 0, a constraint on the premium function π.

B. Definition of the reserve. The expected value (4.32) 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

$$V_t = \int_t^n e^{-\int_t^\tau r} {}_{\tau-t} p_{x+t} \{\mu_{x+\tau} b_\tau - \pi_\tau\} \, d\tau + b_n \, e^{-\int_t^n r} {}_{n-t} p_{x+t} \,. \tag{4.33}$$

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 will take the liberty to just speak of the *reserve*.

Upon inserting ${}_{\tau-t} p_{x+t} = e^{-\int_t^\tau \mu_{x+s} \, ds}$, (4.33) assumes the form

$$V_t = \int_t^n e^{-\int_t^\tau (r_s + \mu_{x+s}) \, ds} \{\mu_{x+\tau} b_\tau - \pi_\tau\} \, d\tau + e^{-\int_t^\tau (r_s + \mu_{x+s}) \, ds} b_n \,. \tag{4.34}$$

Glancing behind at (2.14), we see that, formally, the expression in (4.34) is the reserve at time t for a deterministic contract with payments given by $\Delta B_0 = -\pi_0$ (comes from setting (4.32) equal to 0), $dB_t = (\mu_{x+t} \, b_t - \pi_t) \, dt$, $0 < t < n$ and $\Delta B_n = b_n$, and with interest rate $r_t + \mu_{x+t}$. We can, therefore, reuse the relationships in Chapter 2.

For instance, by (2.26) and (2.13), we have the following *retrospective formula* for the prospective premium reserve:

$$V_t = e^{\int_0^t (r_s + \mu_{x+s}) \, ds} \pi_0 + \int_0^t e^{\int_\tau^t (r_s + \mu_{x+s}) \, ds} (\pi_\tau - \mu_{x+\tau} \, b_\tau) \, d\tau \,. \tag{4.35}$$

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

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

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

$$V_0 = \pi_0 \,. \tag{4.36}$$

First, for the pure endowment against single net premium $_nE_x$ collected at time 0,

$$V_t = \, _{n-t}E_{x+t}, \quad 0 \le t < n. \tag{4.37}$$

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.

If premiums are payable continuously at level rate π given by (4.25) throughout the insurance period, then

$$\begin{aligned} V_t &= \, _{n-t}E_{x+t} - \pi \bar{a}_{x+t\,\overline{n-t}|} \\ &= \, _{n-t}E_{x+t} - \frac{_nE_x}{\bar{a}_{x\,\overline{n}|}} \bar{a}_{x+t\,\overline{n-t}|}. \end{aligned} \tag{4.38}$$

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

Next, for an m-year deferred whole life annuity against level net premium π given by (4.26),

$$\begin{aligned} V_t &= \begin{cases} _{m-t|}\bar{a}_{x+t} - \pi \bar{a}_{x+t\,\overline{m-t}|}, & 0 < t < m, \\ \bar{a}_{x+t}, & t \ge m, \end{cases} \\ &= \bar{a}_{x+t} - \bar{a}_{x+t\,\overline{m-t}|} - \frac{\bar{a}_x - \bar{a}_{x\,\overline{m}|}}{\bar{a}_{x\,\overline{m}|}} \bar{a}_{x+t\,\overline{m-t}|} \\ &= \bar{a}_{x+t} - \frac{\bar{a}_x}{\bar{a}_{x\,\overline{m}|}} \bar{a}_{x+t\,\overline{m-t}|} \end{aligned} \tag{4.39}$$

Figure 4.2: The net reserve for an n-year pure endowment of 1 against level premium in the insurance period.

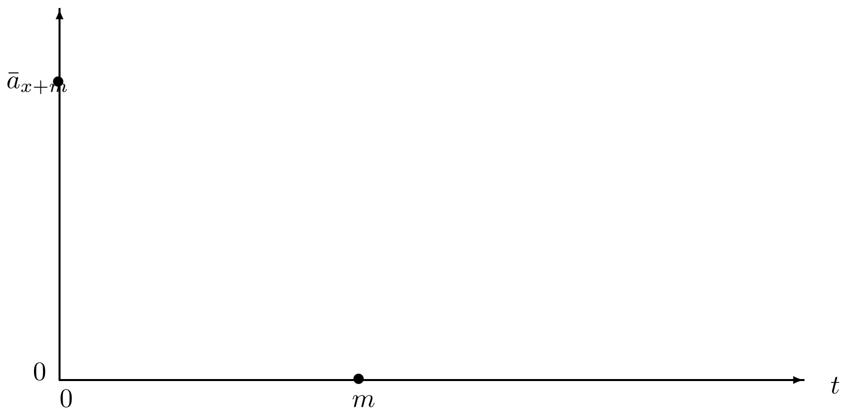


Figure 4.3: The net reserve for an m-year deferred whole life annuity against level premium in the deferred period.

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 insurance period.

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

For the n-year term insurance against level net premium π given by (4.27),

$$\begin{aligned} V_t &= \bar{A}^{\,1}_{x+t\,\overline{n-t|}} - \pi \bar{a}_{x+t\,\overline{n-t|}} \\ &= 1 - r\bar{a}_{x+t\,\overline{n-t|}} - {}_{n-t}E_{x+t} - \frac{1 - r\bar{a}_{x\,\overline{n|}} - {}_nE_x}{\bar{a}_{x\,\overline{n|}}}\bar{a}_{x+t\,\overline{n-t|}} \\ &= 1 - {}_{n-t}E_{x+t} - (1 - {}_nE_x)\frac{\bar{a}_{x+t\,\overline{n-t|}}}{\bar{a}_{x\,\overline{n|}}}. \end{aligned} \tag{4.40}$$

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

Finally, for the n-year endowment insurance against level net premium π given by (4.28),

$$\begin{aligned} V_t &= \bar{A}_{x+t\,\overline{n-t|}} - \pi \bar{a}_{x+t\,\overline{n-t|}} \\ &= 1 - r\bar{a}_{x+t\,\overline{n-t|}} - \frac{1 - r\bar{a}_{x\,\overline{n|}}}{\bar{a}_{x\,\overline{n|}}}\bar{a}_{x+t\,\overline{n-t|}} \\ &= 1 - \frac{\bar{a}_{x+t\,\overline{n-t|}}}{\bar{a}_{x\,\overline{n|}}}. \end{aligned} \tag{4.41}$$

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

The reserve in (4.41) is, of course, the sum of the reserves in (4.39) and (4.40). 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 policy-holder only to pay the premiums, and the contract can be terminated at any time the policy-holder wishes.) Therefore, it is in practice required that

$$V_t \geq 0, \quad t \geq 0. \tag{4.42}$$

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 the expected present values ${}_{n-t}E_{x+t}$, $\bar{a}_{x+t:\overline{n-t|}}$, $\bar{A}_{x+t\,\overline{n-t|}}$, and $\bar{A}^{\,1}_{x+t\,\overline{n-t|}}$ appearing in the formulas in Section 4.3.

First,

$$ {}_{n-t}E_{x+t} = e^{-\int_t^n (r+\mu_{x+s})\,ds} $$

is seen to be an increasing function of t no matter what are the interest rate and the mortality rate. The derivative is

$$ \frac{d}{dt}\,{}_{n-t}E_{x+t} = {}_{n-t}E_{x+t}\,(r+\mu_{x+t})\,. $$

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

Next consider

$$ \bar{a}_{x+t:\overline{n-t}|} = \int_t^n e^{-\int_t^\tau (r+\mu_{x+s})\,ds}\,d\tau\,. $$

The following inequalities are obvious:

$$ \bar{a}_{x+t:\overline{n-t}|} \le \frac{1}{r+\inf_{s\ge t}\mu_{x+s}} \le \frac{1}{r}\,. $$

The last expression is just the present value of a perpetuity, (2.32). If μ is an increasing function, then

$$ \bar{a}_{x+t:\overline{n-t}|} \le \frac{1}{r+\mu_{x+t}}\,. $$

We find the derivative

$$ \frac{d}{dt}\bar{a}_{x+t:\overline{n-t}|} = (r+\mu_{x+t})\bar{a}_{x+t:\overline{n-t}|} - 1\,. $$

It follows that $\bar{a}_{x+t:\overline{n-t}|}$ is a decreasing function of t if μ is increasing, which is quite natural. You can easily invent an example where $\bar{a}_{x+t:\overline{n-t}|}$ is not decreasing.

From the identity

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

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

For

$$ \bar{A}_{\overset{1}{x+t}\,\overline{n-t}|} = 1 - r\bar{a}_{x+t:\overline{n-t}|} - {}_{n-t}E_{x+t} $$

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, (4.37), is always increasing. Assume henceforth that μ 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 (4.38) 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 (4.41) of the endowment insurance. It is left to the diligent reader to show that the reserve in (4.39) is increasing throughout the deferred period and thereafter turns decreasing. This is best done by examining (4.33) and (4.35) in the cases $t \leq m$ and $t \leq m$, respectively. It follows in particular that the reserve is non-negative. The same trick serves also to show that (4.40) is first increasing and thereafter decreasing.

4.4 Thiele's differential equation

A. The differential equation. We turn back to the general case with the reserve given by (4.33) or (4.33), the latter being the more convenient since we can draw on the results in Chapter 2.

The differential form (2.19) translates to the celebrated *Thiele's differentila equation*,

$$\frac{d}{dt} V_t = \pi_t - b_t \, \mu_{x+t} + (r + \mu_{x+t}) \, V_t, \tag{4.43}$$

valid at each t where b, π, and μ are continuous. The right hand side expression in (4.43) shows how the fund per surviving policy-holder 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

$$V_{n-} \quad = \quad b_n \,, \tag{4.44}$$

the differential equation (4.43) determines V_t for fixed b and π.

If the principle of equivalence is exercised, then we must add the condition (4.36). This represents a constraint on the contractual payments b and π; typically, one first specifies the benefit b and then determines the premium rate for a given premium plan (shape of π).

Thiele's differential equation is a so-called *backward differential equation*. This term indicates that we take our stand at the beginning of the time interval we are interested in and also that the differential equation is to be solved by a backward scheme starting from the ultimo condition (2.21). The differential equation may be put up by the *direct backward construction* which goes as follows. Suppose the policy is in force at time $t \in (0, n)$. Use the rule of iterated expectation, conditioning on what happens in the small time interval $(t, t + dt]$: with probability $\mu_{x+t} \, dt + o(dt)$ the insured dies, and the conditional expected value is then just b_t; with probability $1 - \mu_{x+t} \, dt + o(dt)$ the insured survives, and the conditional expected value is then $-\pi_t dt + e^{-r \, dt} \, V_{t+dt}$. We gather

$$V_t = b_t \, \mu_{x+t} \, dt - \pi_t \, dt + (1 - \mu_{x+t} \, dt) e^{-r \, dt} \, V_{t+dt} + o(dt). \tag{4.45}$$

Subtract V_{t+dt} on both sides, divide by dt and let dt tend to 0. Observing that $(e^{-rdt} - 1)/dt \to -r$ as $dt \to 0$, one arrives at (4.43)

B. Savings premium and risk premium. Suppose the equivalence principle is in use. Rearrange (4.43) as

$$\pi_t = \frac{d}{dt}\, V_t - rV_t + (b - V_t)\mu_{x+t}. \tag{4.46}$$

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

$$\pi_t^s = \frac{d}{dt}\, V_t - rV_t, \tag{4.47}$$

and a *risk premium*,

$$\pi_t^r = (b_t - V_t)\mu_{x+t}. \tag{4.48}$$

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 (4.43) is easily obtained as

$$V_t = \begin{cases} \pi \bar{s}_{\overline{t}|}, & 0 < t < m, \\ b\bar{a}_{\overline{m+n-t}|}, & m < t < m+n, \end{cases}$$

where $\bar{s}_{\overline{t}|} = \int_0^t (1+i)^{t-\tau}\, d\tau$. 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 contract described in Section 4.4 and will refer to this as the *standard contract.* For ease of reference we fetch Thiele's differential from (4.43):

$$\frac{d}{dt}\, V_t = \pi_t - \mu_{x+t}\, b_t + (r_t + \mu_{x+t})\, V_t\,. \tag{4.49}$$

The boundary condition following from the very definition of the reserve is

$$V_{n-} = b_n \,. \tag{4.50}$$

With premiums determined by the principle of equivalence, we also have

$$V_0 = \pi_0 \,, \tag{4.51}$$

where π_0 is the lump sum premium payment collected upon the inception of the policy (it may be 0, of course).

Now consider a different model with interest r_t^* and mortality μ_{x+t}^* and a different contract with benefits b_t^* and premiums π_t^*. This will be referred to as the *special contract*. The reserve function V_t^* under this contract satisfies

$$\frac{d}{dt}\,V_t^* = \pi_t^* - \mu_{x+t}^* b_t^* + (r_t^* + \mu_{x+t}^*)\,V_t^* \,, \tag{4.52}$$

$$V_{n-}^* = b_n^* \,, \tag{4.53}$$

$$V_0^* = \pi_0^* \,. \tag{4.54}$$

Assume that

$$\pi_0^* = \pi_0 \,, \qquad b_n^* = b_n \,. \tag{4.55}$$

We are interested in the difference $V_t^* - V_t$, and a few words are in order to motivate this: The reserve is accounted as a liability on the part of the insurance company. To be on the safe side, the company should, at any time, provide a reserve in excess of what seems likely to be needed. This is usually obtained by using 'technical' elements r_t^* and μ_{x+t}^* that are different from the 'realistic' elements r_t and μ_{x+t}, and that produce a reserve V_t^* bigger than the 'realistic' V_t.

Subtract (4.49) from (4.52) to get

$$\frac{d}{dt}\,(V_t^* - V_t) = \eta_t + (r_t^* + \mu_{x+t}^*)\,(V_t^* - V_t)\,, \tag{4.56}$$

where

$$\eta_t = (\pi_t^* - \pi_t) + (\mu_{x+t} b_t - \mu_{x+t}^* b_t^*) + (r_t^* - r_t + \mu_{x+t}^* - \mu_{x+t})\,V_t \,. \tag{4.57}$$

Integrate (4.56) from 0 to t, using $V_0 = V_0^*$, to obtain

$$V_t^* - V_t = \int_0^t e^{\int_s^t (r^* + \mu^*)} \eta_s \, ds \,.$$

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

$$\eta_t \begin{smallmatrix} \leq \\ \geq \end{smallmatrix} 0 \text{ for } t \begin{smallmatrix} < \\ > \end{smallmatrix} t_0\,, \tag{4.58}$$

then $V_t^* \leq V_t$ for all t. In particular, this is the case if η_t is non-decreasing. The result remains valid if all inequalities are made strict. We can now prove the following:

(1) For a contract with level premium intensity throughout the insurance period, and with non-decreasing reserve, a uniform increase of the interest rate results in a decrease of the reserve.

Proof: Now $r_t^* - r_t = \Delta r$ is a positive constant, π^* and π are both constants, all other elements are unchanged, and V_t increasing. Then $\eta_t = (\pi^* - \pi) + \Delta r\, V_t$ is increasing.

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

Proof: Now $(\mu_{x+t}^* - \mu_{x+t})$ is positive and decreasing (non-increasing), π^* and π are constants, $b_t = b_t^* = b$ constant, V_t increasing (this is the case for the endowment insurance if μ in increasing). Then, since $V_t \leq b$, we have $\eta_t = (\pi^* - \pi) - (\mu_{x+t}^* - \mu_{x+t})(b - V_t)$ is increasing.

(3) 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 (4.58) can be used to check whether the reserve V_t is non-negative (as it should be).

Proof: Putting $\pi_t^* = \mu_{x+t} b_t$, means premiums covers current expected benefits, so there is no accumulation of reserve; $V_t^* = 0$. Now $\eta_t = -\pi_t + \mu_{x+t} b_t$, so if this is increasing, then $0 = V_t^* \leq V_t$. This is the case e.g. if π and b are constants and μ_t is increasing.

The reason why the impact on the reserve of a change in valuation and/or contract elements is a bit involved is that, under the equivalence principle, the premium is also affected by the change. However, if we require that the premium be constant as function of t, then $(\pi^* - \pi)$ appearing in the expression for η_t is constant and does not affect the monotonicity properties of η_t. Note also that, since $V_t^* - V_t$ starts and ends at 0, η_t cannot be strictly positive in some part of $(0, n)$ without being strictly negative in some other part.

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 [14] and Dhaene [16] 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 π per year as long as the contract is in force (see Paragraphs 4.1.C-D). The present value of benefits less premiums on the contract is

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

where $\bar{a}_{\overline{t}|} = \int_0^t e^{-r\tau}\, d\tau = (1 - e^{-rt})/r$ 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

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

The jump at $-\pi\bar{a}_{\overline{n}|}$ 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 insurance 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).

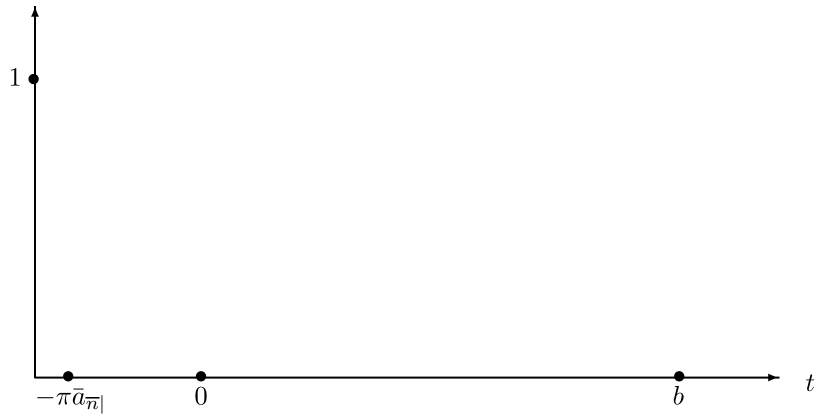


Figure 4.6: The probability distribution of the present value of a term insurance against level premium.

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 didactic purposes, however, it is worthwhile taking it already here as a rehearsal for more complicated situations where stochastic processes cannot be dispensed with.

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

$$\begin{aligned}
dB_t^{e;n} &= I_t\, d\varepsilon_n(t)\,, \\
dB_t^{ti;n} &= 1_{(0,n]}(t)\, dN_t\,, \\
dB_t^{a;n} &= I_t 1_{(0,n)}(t)\, dt\,, \\
dB_t^{ei;n} &= dB_t^{ti;n} + dB_t^{e;n}\,.
\end{aligned}$$

Their present values at time 0 are

$$\begin{aligned}
PV^{e;n} &= e^{-\int_0^n r} I_n\,, \\
PV^{ti;n} &= \int_0^n e^{-\int_0^\tau r} dN_\tau\,,
\end{aligned}$$

$$
\begin{aligned}
PV^{a;n} &= \int_0^n e^{-\int_0^\tau r} I_\tau d\tau \,, \\
PV^{ei;n} &= V^{ti;n} + V^{e;n} \,.
\end{aligned}
$$

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

$$
\begin{aligned}
\mathbb{E}\,[I_\tau] &= {}_\tau p_x \,, \\
\mathbb{E}\,[dN_\tau] &= {}_\tau p_x \, \mu_{x+\tau} \, d\tau \,.
\end{aligned}
$$

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

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

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

$$\pi = b\frac{\bar{A}_{x\,\overline{n}|}}{\bar{a}_{x\,\overline{n}|}} = b\left(\frac{1}{\bar{a}_{x\,\overline{n}|}} - r\right) , \tag{5.1}$$

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

$$V_t = b\bar{A}_{x+t\,\overline{n-t}|} - \pi\bar{a}_{x+t\,\overline{n-t}|} = b\left(1 - \frac{\bar{a}_{x+t\,\overline{n-t}|}}{\bar{a}_{x\,\overline{n}|}}\right) . \tag{5.2}$$

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 π', 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 α-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

$$\alpha' + \alpha'' b\,. \tag{5.3}$$

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

$$\beta' + \beta'' \pi' \tag{5.4}$$

throughout the premium-paying period.

Finally, in the third place there are the so-called γ-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

$$\gamma' + \gamma'' b + \gamma''' V_t' \tag{5.5}$$

at time t if the policy is then in force.

The constant terms α', β', and γ' 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 α-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 β-expenses. The term $\gamma''' V_t'$ represents expenses in connection with management of the investment portfolio, which can reasonably be divided between the policy-holders 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 π' 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 π' 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:

$$\frac{d}{dt} V_t' = \pi' - \beta' - \beta'' \pi' - \gamma' - \gamma'' b - \gamma''' V_t' - \mu_{x+t} b + r V_t' + \mu_{x+t} V_t'\,. \tag{5.6}$$

The appropriate side conditions are

$$V'_{n-} = b\,, \tag{5.7}$$

and

$$V'_0 = -(\alpha' + \alpha'' b)\,. \tag{5.8}$$

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 π' for given benefits and expense factors.

Gathering terms involving V'_t on the left of (5.6) and multiplying on both sides with $e^{\int_t^n (r-\gamma'''+\mu)}$ gives

$$\frac{d}{dt}\left(e^{\int_t^n (r-\gamma'''+\mu)} V'_t\right) = e^{\int_t^n (r-\gamma'''+\mu)} \{(1-\beta'')\pi' - \beta' - \gamma' - (\gamma'' + \mu_{x+t})b\}\,. \tag{5.9}$$

Now integrate (5.9) between t and n, using (5.7), and rearrange a bit to obtain

$$\begin{aligned} V'_t \;\; = \;\; & \int_t^n e^{-\int_t^\tau (r-\gamma'''+\mu)} \{\beta' + \gamma' + (\gamma'' + \mu_{x+\tau})b - (1-\beta'')\pi'\}\, d\tau \\ & + e^{-\int_t^n (r-\gamma'''+\mu)} b\,. \end{aligned} \tag{5.10}$$

Upon inserting $t = 0$ into (5.10) and using (5.8), we find

$$\pi' = \frac{\alpha' + \alpha'' b + \int_0^n e^{-\int_0^\tau (r-\gamma'''+\mu)} \{\beta' + \gamma' + (\gamma'' + \mu_{x+\tau})b\} d\tau + e^{-\int_0^n (r-\gamma'''+\mu)} b}{(1-\beta'') \int_0^n e^{-\int_0^\tau (r-\gamma'''+\mu)} d\tau}\,. \tag{5.11}$$

In the special case where $\gamma''' = 0$ we could determine π' 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

$$\pi' = \frac{1}{1-\beta''}\left(\pi + \frac{\alpha' + \alpha'' b}{\bar{a}_{x\,\overline{n}|}} + \beta' + \gamma' + \gamma'' b\right), \tag{5.12}$$

and

$$V'_t = V_t - \frac{\bar{a}_{x+t\,\overline{n-t}|}}{\bar{a}_{x\,\overline{n}|}}(\alpha' + \alpha'' b). \tag{5.13}$$

It is seen that $\pi' > \pi$, as was anticipated at the outset. Furthermore, $V_t' < V_t$ for $0 \le t < n$, 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 $(1-\beta'')\pi'$, see (5.12). Thus, the only expense factor that appears in the gross reserve is the non-amortized initial α-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 analysis is required as a basis for assumptions about mortality and interest, thorough cost analysis 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

$$ {}_t p_x = e^{-\int_0^t \mu_{x+\tau}\, d\tau} . \tag{6.1} $$

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,

$$ {}_n E_x \quad = \quad v^n \, {}_n p_x \,, \tag{6.2} $$

the life annuity,

$$ \bar{a}_{x\,\overline{n}|} = \int_0^n v^t \, {}_t p_x dt = \int_0^n {}_t E_x dt \,, \tag{6.3} $$

the endowment insurance,

$$ \bar{A}_{x\,\overline{n}|} = 1 - r \bar{a}_{x\,\overline{n}|} \,, \tag{6.4} $$

and the term insurance,

$$ \bar{A}^{1}_{x\,\overline{n}|} = \bar{A}_{x\,\overline{n}|} - {}_n E_x \,. \tag{6.5} $$

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 ${}_tp_x$. Once we have determined ${}_tp_x$, all other functions of interest are obtained by integration, possibly by some numerical method, and elementary algebraic operations.

B. Multi-dimensional 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 μ_{x_j+t} and, hence, has survival function

$$ {}_tp_{x_j} = e^{-\int_0^t \mu_{x_j+\tau}\, d\tau} . \tag{6.6}$$

(The function μ 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 multi-dimensional survival function

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

or, equivalently, by the density

$$\prod_{j=1}^r {}_{t_j}p_{x_j}\mu_{x_j+t_j} . \tag{6.7}$$

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

$$T_{x_1\ldots x_r} = \min\{T_1, \ldots, T_r\} . \tag{6.8}$$

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 ${}_tp_{x_1\ldots x_r}$ and is

$$ {}_tp_{x_1\ldots x_r} = \mathbb{P}\left[\cap_{j=1}^r \{T_j > t\}\right] = e^{-\int_0^t \sum_{j=1}^r \mu_{x_j+\tau}\, d\tau} . \tag{6.9}$$

From this survival function we form the present values of an endowment ${}_nE_{x_1\ldots x_r}$, a life annuity $\bar{a}_{x_1\ldots x_r\,\overline{n}|}$, an endowment insurance $\bar{A}_{x_1\ldots x_r\,\overline{n}|}$, and a term insurance, $\bar{A}^1_{x_1\ldots x_r\,\overline{n}|}$, 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,

$$\mu_{x_1\ldots x_r}(t) = \sum_{j=1}^r \mu_{x_j+t} . \tag{6.10}$$

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

$$\mu_{x_j+t} = \alpha_j + \beta_j e^{\gamma\,(x_j+t)}, \tag{6.11}$$

then (6.10) becomes

$$\mu_{x_1\ldots x_r}(t) = \alpha' + \beta' e^{\gamma\,t} \tag{6.12}$$

with

$$\alpha' = \sum_{j=1}^{r} \alpha_j\,,\;\; \beta' = \sum_{j=1}^{r} \beta_j e^{\gamma\,x_j}\,, \tag{6.13}$$

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

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

$$T_{\overline{x_1\ldots x_r}} = \max\{T_1, \ldots, T_r\}\,. \tag{6.14}$$

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 ${}_tp_{\overline{x_1\ldots x_r}}$. By the general addition rule for probabilities (Appendix C),

$$\begin{aligned} {}_tp_{\overline{x_1\ldots x_r}} &= \mathbb{P}\left[\cup_{j=1}^{r}\{T_j > t\}\right] \\ &= \sum_j {}_tp_{x_j} - \sum_{j_1<j_2} {}_tp_{x_{j_1}x_{j_2}} + \ldots + (-1)^{r-1}\,{}_tp_{x_1\ldots x_r}\,. \end{aligned} \tag{6.15}$$

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 (G). Formulas for benefits contingent on survival, obtained from (6.2) and (6.3), will reflect the structure of (G) 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 $\overset{q}{\overline{x_1\ldots x_r}}$ 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 sub-groups of lives by direct application of the theorem in Appendix C:

$$ {}_tp_{\overset{q}{\overline{x_1\ldots x_r}}} = \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{q-1} \sum_{j_1<\ldots<j_p} {}_tp_{x_{j_1}\ldots x_{j_p}}\,. \tag{6.16}$$

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

$$ {}_nE_{\overline{x_1\ldots x_r}^{\,q}} = \sum_{p=q}^{r}(-1)^{p-q}\binom{p-1}{q-1}\sum_{j_1<\ldots<j_p} {}_nE_{x_{j_1}\ldots x_{j_p}}\,, \tag{6.17} $$

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

$$ \bar{a}_{\overline{x_1\ldots x_r}^{\,q}\,\overline{n}|} = \sum_{p=q}^{r}(-1)^{p-q}\binom{p-1}{q-1}\sum_{j_1<\ldots<j_p} \bar{a}_{x_{j_1}\ldots x_{j_p}\,\overline{n}|}\,. \tag{6.18} $$

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):

$$ \bar{A}_{\overline{x_1\ldots x_r}^{\,q}\,\overline{n}|} = 1 - r\bar{a}_{\overline{x_1\ldots x_r}^{\,q}\,\overline{n}|}\,, \tag{6.19} $$

$$ \bar{A}^{1}_{\overline{x_1\ldots x_r}^{\,q}\,\overline{n}|} = \bar{A}_{\overline{x_1\ldots x_r}^{\,q}\,\overline{n}|} - {}_nE_{x_{j_1}\ldots x_{j_p}}\,. \tag{6.20} $$

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

$$ \bar{A}_{\overline{x_1\ldots x_r}^{\,q}\,\overline{n}|} = \sum_{p=q}^{r}(-1)^{p-q}\binom{p-1}{q-1}\sum_{j_1<\ldots<j_p} \bar{A}_{x_{j_1}\ldots x_{j_p}\,\overline{n}|}\,. \tag{6.21} $$

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, $\mathcal{Z} = \{0, 1, \ldots, r\}$, 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 $\mathcal{Z}$, 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 modeled 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 modeling 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 [26] as a suitable framework throughout this text. A useful basic source is [30].

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 $j_1, \ldots, j_p \in \mathcal{Z}$. Now

$$\begin{aligned} &\mathbb{P}\left[Z(t_h) = j_h,\, h = 1, \ldots, p\right] \\ &\quad = \prod_{h=1}^{p} \mathbb{P}\left[Z(t_h) = j_h \mid Z(t_g) = j_g,\, g = 0, \ldots, h-1\right], \end{aligned} \tag{7.1}$$

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 $j_1, \ldots, j_p \in \mathcal{Z}$,

$$\begin{aligned} &\mathbb{P}\left[Z(t_p) = j_p \mid Z(t_h) = j_h,\, h = 1, \ldots, p-1\right] \\ &\quad = \mathbb{P}[Z(t_p) = j_p \mid Z(t_{p-1}) = j_{p-1}], \end{aligned} \tag{7.2}$$

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

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

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

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

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 $j_1, \ldots, j_p, j, j_{p+1}, \ldots, j_{p+q}$ in $\mathcal{Z}$,

$$\begin{aligned} &\mathbb{P}\left[\, Z(t_h) = j_h,\, h = p+1, \ldots, p+q \mid Z(t) = j,\, Z(t_h) = j_h, h = 1, \ldots, p\right] \\ &\quad = \mathbb{P}\left[\, Z(t_h) = j_h,\, h = p+1, \ldots, p+q \mid Z(t) = j\right]. \end{aligned} \tag{7.5}$$

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 $\mathcal{Z}$.

From the simple transition probabilities we form the more general transition probability from j to some subset $\mathcal{K} \subset \mathcal{Z}$,

$$p_{j\mathcal{K}}(t,u) = \mathbb{P}[Z(u) \in \mathcal{K} \,|\, Z(t) = j] = \sum_{k \in \mathcal{K}} p_{jk}(t,u)\,. \tag{7.6}$$

We have, of course,

$$p_{j\mathcal{Z}}(t,u) = \sum_{k \in \mathcal{Z}} p_{jk}(t,u) = 1\,. \tag{7.7}$$

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 $p_{jk}(t,u)$, $j,k \in \mathcal{Z}$, $0 \le t \le u \le n$, such that $\sum_{k \in \mathcal{Z}} p_{jk}(t,u) = 1$ and, for any $0 \le t_1 < \cdots < t_p$ in $[0,n]$ and $\{j_1,\ldots,j_p\}$ in $\mathcal{Z}$,..."

We shall briefly outline more general definitions of the Markov property. For $\mathcal{T} \subset [0,n]$ let $\mathcal{H}_{\mathcal{T}}$ denote the class of all events generated by $\{Z(t)\}_{t \in \mathcal{T}}$. It represents everything that can be observed about Z in the time set $\mathcal{T}$. For instance, $\mathcal{H}_{\{t\}}$ is the information carried by the process at time t and consists of the elementary events $\emptyset$, Ω, 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 $\mathcal{T}$ of interval type are frequently encountered, and we abbreviate $\mathcal{H}_{\le t} = \mathcal{H}_{[0,t]}$ (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 $B \in \mathcal{H}_{>t}$,

$$\mathbb{P}[B \,|\, \mathcal{H}_{\le t}] = \mathbb{P}[B \,|\, \mathcal{H}_{\{t\}}]\,. \tag{7.8}$$

This is the general form of (7.5).

An alternative definition says that, for any $A \in \mathcal{H}_{<t}$ and $B \in \mathcal{H}_{>t}$,

$$\mathbb{P}[A \cap B \,|\, \mathcal{H}_{\{t\}}] = \mathbb{P}[A \,|\, \mathcal{H}_{\{t\}}]\, \mathbb{P}[B \,|\, \mathcal{H}_{\{t\}}]\,, \tag{7.9}$$

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 $\{Z(t) = j\}$, $j \in \mathcal{Z}$, are disjoint and their union is the almost sure event. It follows that

$$\begin{aligned}\mathbb{P}[Z(u) = k \,|\, Z(s) = i] &= \sum_{j \in \mathcal{Z}} \mathbb{P}[Z(t) = j,\; Z(u) = k \,|\, Z(s) = i] \\ &= \sum_{j \in \mathcal{Z}} \mathbb{P}[Z(t) = j \,|\, Z(s) = i]\, \mathbb{P}[Z(u) = k \,|\, Z(s) = i\,,\; Z(t) = j]\,.\end{aligned}$$

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

$$p_{ik}(s,u) = \sum_{j \in \mathcal{Z}} p_{ij}(s,t) p_{jk}(t,u)\,, \tag{7.10}$$

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 $0 = t_0 < t_1 < \cdots < t_q = n$: then we could just take the $p_{jk}(t_{q-1}, t_q)$ as any non-negative numbers satisfying $\sum_{k=0}^{r} p_{jk}(t_{p-1}, t_p) = 1$ for each $j \in \mathcal{Z}$ 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*,

$$\mu_{jk}(t) = \lim_{h \downarrow 0} \frac{p_{jk}(t, t+h)}{h} \tag{7.11}$$

exist for each $j, k \in \mathcal{Z}$, $j \neq k$, and $t \in [0, n)$ and, moreover, that they are piece-wise continuous. Another way of phrasing (7.11) is

$$p_{jk}(t, t+dt) = \mu_{jk}(t)dt + o(dt)\,, \tag{7.12}$$

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 $j \notin \mathcal{K} \subset \mathcal{Z}$, we define the intensity of transition from state j to the set of states $\mathcal{K}$ at time t as

$$\mu_{j\mathcal{K}}(t) = \lim_{u \downarrow t} \frac{p_{j\mathcal{K}}(t,u)}{u-t} = \sum_{k \in \mathcal{K}} \mu_{jk}(t)\,. \tag{7.13}$$

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

$$\mu_{j\cdot}(t) = \sum_{k; k \neq j} \mu_{jk}(t)\,. \tag{7.14}$$

From (7.7) and (7.12) we get

$$p_{jj}(t, t+dt) = 1 - \mu_{j\cdot}(t)dt + o(dt)\,. \tag{7.15}$$

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 $1 - \mu_{j\cdot}(t)\,dt$ 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

$$\begin{aligned} p_{jk}(t,u) \quad &= \quad (1 - \mu_{j\cdot}(t)\,dt)\,p_{jk}(t+dt,u) \\ &\quad + \sum_{g;g\neq j} \mu_{jg}(t)\,dt\,p_{gk}(t+dt,u) + o(dt)\,, \end{aligned} \tag{7.16}$$

Upon putting $d_t p_{jk}(t,u) = p_{jk}(t+dt,u) - p_{jk}(t,u)$ in the infinitesimal sense, we arrive at

$$d_t p_{jk}(t,u) = \mu_{j\cdot}(t)\,dt\,p_{jk}(t,u) - \sum_{g;g\neq j} \mu_{jg}(t)\,dt\,p_{gk}(t,u)\,. \tag{7.17}$$

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

$$p_{jk}(u,u) = \delta_{jk}\,. \tag{7.18}$$

Here δ_{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

$$\frac{\partial}{\partial t} p_{jk}(t,u) = \mu_{j\cdot}(t) p_{jk}(t,u) - \sum_{g;g\neq j} \mu_{jg}(t) p_{gk}(t,u)\,. \tag{7.19}$$

Since we have assumed that the intensities are piece-wise continuous, the indicated derivatives exist piece-wise. 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

$$p_{ij}(s,t+dt) = \sum_{g;g\neq j} p_{ig}(s,t)\,\mu_{gj}(t)\,dt + p_{ij}(s,t)(1-\mu_{j\cdot}(t)\,dt) + o(dt)\,,$$

hence

$$d_t p_{ij}(s,t) = \sum_{g;g\neq j} p_{ig}(s,t)\mu_{gj}(t)\,dt\; - p_{ij}(s,t)\mu_{j\cdot}(t)\,dt\,. \tag{7.20}$$

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

$$p_{ij}(s,s) = \delta_{ij}\,. \tag{7.21}$$

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 $\mathcal{H}_{\{t_1,\ldots,t_r\}}$ from the finite-dimensional distribution (7.4). In fact, with finite $\mathcal{Z}$ 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 $\mathcal{H}_{\mathcal{T}}$ with $\mathcal{T}$ 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,

$$p_{\overline{jj}}(t,u) = \mathbb{P}[Z(\tau) = j, \tau \in (t,u] \,|\, Z(t) = j]\,. \tag{7.22}$$

Obviously $p_{\overline{jj}}(t,u) = p_{\overline{jj}}(t,s)\,p_{\overline{jj}}(s,u)$ for $t < s < u$. By the "backward" construction and (7.15) we get

$$p_{\overline{jj}}(t,u) = (1-\mu_{j\cdot}(t)\,dt)\,p_{\overline{jj}}(t+dt,u) + o(dt)\,. \tag{7.23}$$

From here proceed as above, using $p_{\overline{jj}}(u,u) = 1$, to obtain

$$p_{\overline{jj}}(t,u) = e^{-\int_t^u \mu_{j\cdot}} . \tag{7.24}$$

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 $e^{\int_t^u \mu_{j\cdot}}$:

$$d_t \left(e^{\int_t^u \mu_{j\cdot}} p_{jk}(t,u) \right) = -e^{\int_t^u \mu_{j\cdot}} \sum_{g;g\neq j} \mu_{jg}(t)\, dt\, p_{gk}(t,u).$$

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

$$\delta_{jk} - e^{\int_t^u \mu_{j\cdot}} p_{jk}(t,u) = -\int_t^u e^{\int_\tau^u \mu_{j\cdot}} \sum_{g;\, g\neq j} \mu_{gj}(\tau) p_{jk}(\tau,u)\, d\tau\, .$$

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

$$p_{jk}(t,u) = \int_t^u p_{\overline{jj}}(t,\tau) \sum_{g;\, g\neq j} \mu_{jg}(\tau) p_{gk}(\tau,u)\, d\tau + \delta_{jk} p_{\overline{jj}}(t,u)\, . \tag{7.25}$$

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

$$p_{ij}(s,t) = \delta_{ij} p_{\overline{ii}}(s,t) + \sum_{g;\, g\neq j} \int_s^t p_{ig}(s,\tau)\mu_{gj}(\tau) p_{\overline{jj}}(\tau,t) d\tau\, . \tag{7.26}$$

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 modeled as a positive random variable T with survival function $\bar{F}$. There are

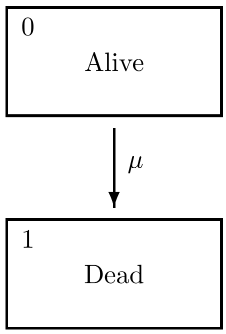


Figure 7.1: Sketch of the mortality model with one cause of death.

two states, 'alive' and 'dead'. Labeling these by 0 and 1, respectively, the state process Z is simply

$$Z(t) = 1[T \leq t]\,,\; t \in [0,n]\,,$$

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

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

The Chapman-Kolmogorov equation reduces to the trivial

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

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

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

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 μ_{0j} is superfluous and has been dropped.

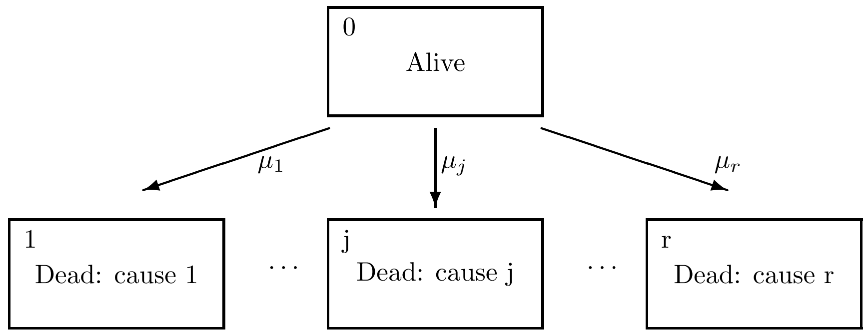


Figure 7.2: Sketch of the mortality model with r causes of death.

Relation (7.14) implies that the total mortality intensity is the sum of the intensities of death from different causes,

$$\mu(t) = \sum_{j=1}^{r} \mu_j(t)\,. \tag{7.28}$$

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). The present enriched model opens possibilities of expressing ideas about the relative importance of various causes of death and thus better motivate a specific mortality law in the aggregate. For instance, the G-M law in the simple mortality model may be motivated as resulting from two causes of death, one with intensity α independent of age (pure accident) and the other with intensity βc^t (wear-out).

The probability that a t years old will die from cause j before age u is

$$p_{0j}(t,u) = \int_t^u e^{-\int_t^\tau \mu}\, \mu_j(\tau)\, d\tau\,. \tag{7.29}$$

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 μ_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 lunge 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 re-labeling 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

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

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

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

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

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

(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 $p_{00}(s,s) = 0\,,\;\; p_{11}(s,s) = 1\,.$).

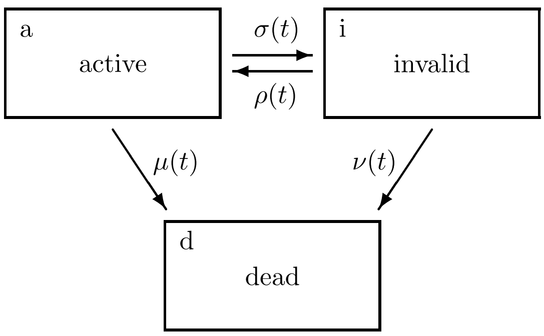


Figure 7.3: Sketch of a Markov chain model for disabilities, recoveries, and death.

When the intensities are sufficiently simple functions, one may find explicit closed expressions for the transition probabilities. Work through the case with constant intensities.

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 [27].

B. Aggregating states of a Markov chain. Let Z be the general continuous time Markov chain introduced in Section 7.2. Let $\{\mathcal{Z}_0, \ldots, \mathcal{Z}_{\tilde{r}}\}$ be a partition of $\mathcal{Z}$, that is, the $\mathcal{Z}_g$ are disjoint and their union is $\mathcal{Z}$. By convention, $0 \in \mathcal{Z}_0$. Define a stochastic process $\tilde{Z}$ on the state space $\tilde{\mathcal{Z}} = \{0, \ldots, \tilde{r}\}$ by

$$\tilde{Z}(t) = g \;\text{ iff }\; Z(t) \in \mathcal{Z}_g \,. \tag{7.34}$$

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

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

$$\mathbb{P}[\tilde{Z}(u) = h \,|\, Z(s) = i,\, \tilde{Z}(t) = g] = \frac{1}{p_{i\mathcal{Z}_g}(s,t)} \sum_{j \in \mathcal{Z}_g} p_{ij}(s,t) p_{j\mathcal{Z}_h}(t,u)\,.$$

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

$$\lim_{u \downarrow t} \frac{\mathbb{P}[\tilde{Z}(u) = h \mid Z(s) = i,\, \tilde{Z}(t) = g]}{u - t} = \frac{1}{p_{i\mathcal{Z}_g}(s,t)} \sum_{j \in \mathcal{Z}_g} p_{ij}(s,t)\mu_{j\mathcal{Z}_h}(t)\,.$$

We can not 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, $\mathcal{Z}_0 = \{0, 1\}$, and put $\mathcal{Z}_1 = \{2\}$. Thus we observe only whether the insured is alive or not. The process $\tilde{Z}$ is Markov, of course (recall the argument in Paragraph 7.3.A). The survival probability is

$$\tilde{p}_{00}(0,t) = p_{00}(0,t) + p_{01}(0,t)\,,$$

and the mortality intensity at age t is

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

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 $p_{j\mathcal{Z}_h}(t,u)$ considered as functions of j are constant on each $\mathcal{Z}_g$, that is, there exist functions $\tilde{p}_{gh}(t,u)$ such that, for each $t < u$ and $g, h \in \mathcal{Z}^*$,

$$p_{j\mathcal{Z}_h}(t,u) = \tilde{p}_{gh}(t,u)\,,\ \forall j \in \mathcal{Z}_g\,. \tag{7.35}$$

Then we shall say that the probabilities of transition between the subsets $\mathcal{Z}_g$ 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 $\mathcal{Z}_g$ are non-differential, then the process $\tilde{Z}$ defined by (7.34) is Markov with transition probabilities $\tilde{p}_{gh}(t,u)$ defined by (7.35). If Z possesses intensities μ_{jk}, then the process Z^* has intensities $\tilde{\mu}_{gh}$ given by*

$$\tilde{\mu}_{gh} = \mu_{j\mathcal{Z}_h}\,,\ j \in \mathcal{Z}_g\,. \tag{7.36}$$

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

$$\mathbb{P}[\tilde{Z}(u) = h \mid A\,, \tilde{Z}(t) = g] = \tilde{p}_{gh}(t,u)\,. \tag{7.37}$$

Using first the fact that $[\tilde{Z}(t) = g] = \cup_{j \in \mathcal{Z}_g}[Z(t) = j]$ is a union of disjoint events, then that $A \in \mathcal{H}_{<t}$ and Z is Markov, and finally assumption (7.35), we

get

$$\begin{aligned}
&\mathbb{P}[A,\,\tilde{Z}(t)=g,\,\tilde{Z}(u)=h]\\
&\quad=\sum_{j\in\mathcal{Z}_g}\mathbb{P}[A,\,Z(t)=j,\,Z(u)\in\mathcal{Z}_h]\\
&\quad=\sum_{j\in\mathcal{Z}_g}\mathbb{P}[A,\,Z(t)=j]p_{j\mathcal{Z}_h}(t,u)\,.\\
&\quad=\mathbb{P}[A,\,\tilde{Z}(t)=g]\,\tilde{p}_{gh}(t,u)\,,
\end{aligned}$$

which is equivalent to (7.37). □

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 μ_{jr}, $j\in\mathcal{H}$, are identical and equal to λ, say. Then, by Theorem 1, the survival probability is the same in all states $j\in\mathcal{H}$:

$$p_{j\mathcal{H}}(t,u)=e^{-\int_t^u\lambda}\,. \tag{7.38}$$

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

$$p_{jk\,|\,\mathcal{H}}(t,u)=\frac{p_{jk}(t,u)}{p_{j\mathcal{H}}(t,u)}=p_{jk}(t,u)e^{\int_t^u\lambda}\,. \tag{7.39}$$

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

$$\begin{aligned}
p_{jk\,|\,\mathcal{H}}(t,u)e^{-\int_t^u\lambda} \;=\;& \int_t^u e^{-\int_t^\tau\mu_{j,\mathcal{H}-\{j\}}-\int_t^\tau\lambda}\sum_{g\in\mathcal{H}-\{j\}}\mu_{jg}(\tau)p_{gk\,|\,\mathcal{H}}(\tau,u)e^{-\int_\tau^u\lambda}d\tau\\
&+\delta_{jk}e^{-\int_t^u\mu_{j,\mathcal{H}-\{j\}}-\int_t^u\lambda}\,.
\end{aligned}$$

Multiplying with $e^{\int_s^t\lambda}$, 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 $\mu_{j,k}$, $j,k\in\mathcal{H}$. 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 $N_{jk}(t) = \sharp\{\tau;\ Z(\tau-) = j,\ Z(\tau) = k,\ \tau \in (0, t]\}$ (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

$$dI_j(t) = dN_{\cdot j}(t) - dN_{j\cdot}(t)\,, \tag{7.40}$$

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

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))

$$dB(t) = \sum_k I_k(t)\, dB_k(t) + \sum_{k\neq\ell} b_{k\ell}(t)\, dN_{k\ell}(t)\,, \tag{7.41}$$

where each B_k, of form $dB_k(t) = b_k(t)\, dt + B_k(t) - B_k(t-)$, 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 ℓ (a general life assurance). When different from 0, $B_k(t) - B_k(t-)$ 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 $\mathcal{D} = \{t_0, t_1, \ldots, t_q\}$ (say).

Positive amounts represent benefits and negative amounts represent premiums. In practice premiums are only of annuity type. At times $t \notin [0, n]$ 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):

$$\begin{aligned}\int_t^u e^{-\int_0^\tau r} I_j(\tau)\, dB_j(\tau) \;&=\; e^{-\int_0^u r} I_j(u)\, B_j(u) \;-\; e^{-\int_0^t r} I_j(t) B_j(t) + \int_t^u e^{-\int_0^\tau r} I_j(\tau) B_j(\tau) r(\tau)\, d\tau \\ &\quad + \int_t^u e^{-\int_0^\tau r} B_j(\tau_-)\, d(N_{j\cdot}(\tau) - N_{\cdot j}(\tau))\,.\end{aligned}$$

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

$$V(t) = \int_t^n e^{-\int_t^\tau r} dB(\tau) \,. \tag{7.42}$$

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

$$V_j(t) = \int_t^n e^{-\int_t^\tau r} \sum_k p_{jk}(t,\tau) \left(dB_k(\tau) + \sum_{\ell;\ell\neq k} b_{k\ell}(\tau)\mu_{k\ell}(\tau)\, d\tau \right) . \tag{7.43}$$

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

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

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

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

Let $0 \le t < u < n$. 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

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

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

$$\begin{aligned} V_j(t) &= b_j(t)\,dt + \sum_{k;k\neq j} \mu_{jk}(t)\,dt\,b_{jk}(t) \\ &\quad +(1-\mu_{j\cdot}(t)\,dt)e^{-r(t)\,dt}\,V_j(t+dt) + \sum_{k;k\neq j} \mu_{jk}(t)\,dt\,e^{-r(t)\,dt}\,V_k(t+dt)\,. \end{aligned}$$

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,

$$\begin{aligned} \frac{d}{dt}V_j(t) &= (r(t)+\mu_{j\cdot}(t))\,V_j(t) - \sum_{k;k\neq j}\mu_{jk}(t)\,V_k(t) \\ &\quad -b_j(t) - \sum_{k;k\neq j} b_{jk}(t)\mu_{jk}(t)\,. \end{aligned} \tag{7.45}$$

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

$$V_j(t_p-) = (B_j(t_p) - B_j(t_p-)) + V_j(t_p)\,, \quad p = 1, \ldots, q,\ j \in \mathcal{Z}, \tag{7.46}$$

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}, μ_{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_{q-1}, n) subject to (7.46), which specializes to $V_j(n-) = B_j(n) - B_j(n-)$ since $V_j(n) = 0$ 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

$$V_0(0) = -B_0(0)\,. \tag{7.47}$$

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

$$-b_j(t)\,dt = dV_j(t) - r(t)\,dt\,V_j(t) + \sum_{k;k\neq j} R_{jk}(t)\mu_{jk}(t)\,dt\,. \tag{7.48}$$

where

$$R_{jk}(t) = b_{jk}(t) + V_k(t) - V_j(t)\,. \tag{7.49}$$

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 outlay 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 $(r(t)+\mu_{j\cdot}(t))\,V_j(t)$ appearing on the right of over to the left, and multiply the equation by $e^{-\int_0^t(r+\mu_{j\cdot})}$ to form a complete differential on the left:

$$\begin{aligned}
&\frac{d}{dt}\left(e^{-\int_0^t(r+\mu_{j\cdot})}V_j(t)\right) = \\
&\quad -e^{-\int_0^t(r+\mu_{j\cdot})}\left(\sum_{k;k\neq j}\mu_{jk}(t)\,V_k(t)+b_j(t)+\sum_{k;k\neq j}b_{jk}(t)\mu_{jk}(t)\right).
\end{aligned}$$

Now integrate over an interval (t,u) containing no jumps $B_j(\tau)-B_j(\tau-)$ and, recalling that $e^{-\int_t^\tau \mu_{j\cdot}}=p_{\overline{jj}}(t,\tau)$, rearrange a bit to obtain the integral equation

$$\begin{aligned}
V_j(t) \;=\; & \int_t^u p_{\overline{jj}}(t,\tau)\,e^{-\int_t^\tau r}\left(b_j(\tau)+\sum_{k;\,k\neq j}\mu_{jk}(\tau)(b_{jk}(\tau)+V_k(\tau))\right)d\tau \\
& +p_{\overline{jj}}(t,u)\,e^{-\int_t^u r}\,V_j(u-)\,. \qquad (7.50)
\end{aligned}$$

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

$$\begin{aligned}
V_j(t) \;=\; & \int_t^u p_{\overline{jj}}(t,\tau)\sum_{k;\,k\neq j}\mu_{jk}(\tau)\,d\tau\left(\int_t^\tau e^{-\int_t^s r}b_j(s)\,ds+e^{-\int_t^\tau r}(b_{jk}(\tau)+V_k(\tau))\right) \\
& +p_{\overline{jj}}(t,u)\left(\int_t^u e^{-\int_t^s r}b_j(s)\,ds+e^{-\int_t^u r}\,V_j(u-)\right). \qquad (7.51)
\end{aligned}$$

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

$$\begin{aligned}
&\int_t^u p_{\overline{jj}}(t,\tau)\,\mu_{j\cdot}(\tau)\int_t^\tau e^{-\int_t^s r}b_j(s)\,ds\,d\tau \\
&\quad = \int_t^u\int_s^u \frac{d}{d\tau}(-p_{\overline{jj}}(t,\tau))\,d\tau e^{-\int_t^s r}b_j(s)\,ds \\
&\quad = -p_{\overline{jj}}(t,u)\int_t^u e^{-\int_t^s r}b_j(s)\,ds+\int_t^u p_{\overline{jj}}(t,s)e^{-\int_t^s r}b_j(s)\,ds.
\end{aligned}$$

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 $\mathcal{Z}$) 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 payable at level rate c as long as both husband and wife are alive, widow's pension is payable at level rate b (as long as the wife survives the husband), 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. F.4. We assume that r is constant.

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

$$\begin{aligned} \frac{d}{dt}V_0(t) &= (r + \mu(t) + \nu(t))\, V_0(t) \\ &\quad -\mu(t)V_1(t) - \nu_{02}(t)V_2(t) \, + c\,, \qquad (7.52) \\ \frac{d}{dt}V_1(t) &= (r + \nu'(t))\, V_1(t) - b\,, \qquad (7.53) \\ \frac{d}{dt}V_2(t) &= (r + \mu'(t))\, V_2(t) - \mu'(t)\, s\,. \qquad (7.54) \end{aligned}$$

Consider a modified contract, by which 50% of the reserve is to be paid back to the husband in case he becomes a widower 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 $0.5V_0(t)\nu(t)$ must be subtracted on the right of (7.52), which then changes to

$$\begin{aligned} \frac{d}{dt}V_0(t) &= (r + \mu(t) + 0.5\nu(t))\, V_0(t) + c \\ &\quad -\mu(t)V_1(t) - \nu(t)V_2(t)\,, \qquad (7.55) \end{aligned}$$

Together with the conditions $V_j(n) = 0,\ j = 0, 1, 2$, these equations are easily solved.

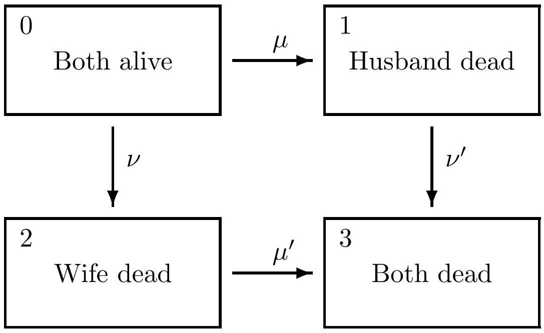


Figure 7.4: Sketch of a model for two lives.

As a second case the widow's pension shall be analyzed 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 $a\,V_j(t)$ 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 categories 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. Non-insurability occurs upon onset of disability or serious illness or other intervening circumstances that entail excess mortality. Hence it is assumed that

$$\lambda_x > \kappa_x; \;\; x > 0. \tag{7.56}$$

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

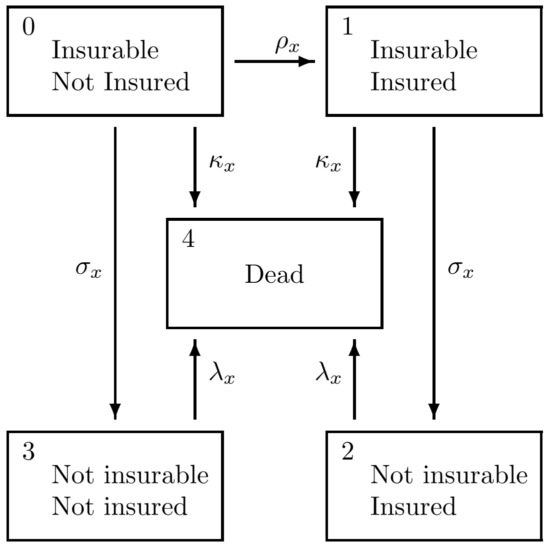


Figure 7.5: A Markov model for occurrences of non-insurability, purchase of insurance, and death.

formulas can be put up directly:

$$p_{11}(x, x+t) = \exp\{-\int_x^{x+t} (\sigma + \kappa)\}, \tag{7.57}$$

$$p_{12}(x, x+t) = \int_x^{x+t} \exp\{-\int_x^u (\sigma + \kappa)\} \, \sigma_u \exp\{-\int_u^{x+t} \lambda\} \, du, \tag{7.58}$$

$$p_{00}(0, x) = \exp\{-\int_0^x (\sigma + \kappa + \rho)\}\,. \tag{7.59}$$

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 non-insurable. 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

$${}_tp_{[x]} = p_{11}(x, x+t) + p_{12}(x, x+t)\,, \tag{7.60}$$

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

$$\mu_{[x]+t} = \frac{\kappa_{x+t}\, p_{11}(x, x+t) + \lambda_{x+t}\, p_{12}(x, x+t)}{p_{11}(x, x+t) + p_{12}(x, x+t)}. \tag{7.61}$$

(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 an decreasing function of the age at entry. Formula (7.61) can be recast as

$$\mu_{[x]+y-x} = \kappa_y + \zeta(x, y)(\lambda_y - \kappa_y)\,, \tag{7.62}$$

where

$$\begin{aligned} \zeta(x, y) &= \frac{p_{12}(x, y)}{p_{11}(x, y) + p_{12}(x, y)} && (7.63)\\ &= \frac{1}{1 + p_{11}(x, y)/p_{12}(x, y)}\,. && (7.64) \end{aligned}$$

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 κ_y and λ_y, the weights being (of course) the conditional probabilities of being insurable and non-insurable, respectively. The weight attached to λ_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 non-insurable in the period since entry; the longer the duration, the larger the proportion of non-insurable lives. In particular, those who have just entered, are known to be insurable, that is, $\mu_{[x]} = \kappa_x$.

D. Comparison with the mortality in the population. Let $\bar{\mu}_x$ denote the force of mortality of a randomly chosen life of age x from the population. A formula for $\bar{\mu}_x$ is easily obtained starting from the survival function ${}_x\bar{p}_0 = \sum_{i=0}^{3} p_{0i}(0, x)$. 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. $\bar{\mu}_y = \mu_{[0]+y}$. Then, since $\mu_{[x]+y-x}$ is a decreasing function of x and $\bar{\mu}_y$ corresponds to $x = 0$, it follows that $\bar{\mu}_y > \mu_{[x]+y-x}$ for all $x < y$.

Again the explanation is trivial; due to the underwriting standards, the proportion of non-insurable 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 $\mathcal{D} = \{t_0, t_1 \dots, t_m\}$ (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 of $V(t)$. By the Markov property, we need only the state-wise conditional moments

$$V_j^{(q)}(t) = \mathbb{E}[V(t)^q | Z(t) = j]\,,$$

$q = 1, 2, \dots$

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

$$\begin{aligned} \frac{d}{dt} V_j^{(q)}(t) &= (qr(t) + \mu_{j\cdot}(t)) V_j^{(q)}(t) - q b_j(t) V_j^{(q-1)}(t) \\ &\quad - \sum_{k;\,k \neq j} \mu_{jk}(t) \sum_{p=0}^{q} \binom{q}{p} (b_{jk}(t))^p V_k^{(q-p)}(t)\,, \end{aligned}$$

valid on $(0,n) \backslash \mathcal{D}$ *and subject to the conditions*

$$V_j^{(q)}(t-) = \sum_{p=0}^{q} \binom{q}{p} (B_j(t) - B_j(t-))^p V_j^{(q-p)}(t)\,, \tag{7.65}$$

$t \in \mathcal{D}$. $\square$

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

$$V(t) = V(t,u) + e^{-\int_t^u r} V(u)\,, \tag{7.66}$$

For any $q = 1, 2, \dots$ we have by the binomial formula

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

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

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

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

$$\binom{q}{p} (1 - \mu_{j\cdot}(t)\,dt)\,(b_j(t)\,dt)^p\, e^{-(q-p)r(t)\,dt}\, V_j^{(q-p)}(t+dt) \quad (7.69)$$

$$+ \binom{q}{p} \sum_{k;\, k \neq j} \mu_{jk}(t)\,dt\,(b_j(t)\,dt + b_{jk}(t))^p e^{-(q-p)r(t)\,dt} V_k^{(q-p)}(t+dt) . \quad (7.70)$$

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

$$(1 - \mu_{j\cdot}(t)\,dt) e^{-qr(t)\,dt}\, V_j^{(q)}(t+dt) ,$$

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

$$\begin{aligned} V_j^{(q)}(t) \quad = \quad & (1 - \mu_{j\cdot}(t)\,dt) e^{-qr(t)\,dt}\, V_j^{(q)}(t+dt) \\ & + q\, b_j(t)\,dt\, e^{-(q-1)r(t)\,dt}\, V_j^{(q-1)}(t+dt) \\ & + \sum_{p=0}^{q} \binom{q}{p} \sum_{k;\, k \neq j} \mu_{jk}(t)\,dt\,(b_{jk}(t))^p e^{-(q-p)r(t)\,dt} V_k^{(q-p)}(t+dt) . \end{aligned}$$

Now subtract $V_j^{(q)}(t+dt)$ on both sides, divide by dt, let dt tend to 0, and use $\lim_{t \downarrow 0} \left(e^{-qr(t)\,dt} - 1 \right) / dt = -qr(t)$ 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. □

A rigorous proof is given in [38].

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

$$m_j^{(1)}(t) = V_j^{(1)}(t)\,, \tag{7.71}$$

$$m_j^{(q)}(t) = \sum_{p=0}^{q}(-1)^{q-p}\binom{q}{p}V_j^{(p)}(t)\left(V_j^{(1)}(t)\right)^{q-p}\,. \tag{7.72}$$

B. Computations. The computation goes as follows. First solve the differential equations in the upper interval (t_{m-1}, n), where the side conditions (7.65) are just

$$V_j^{(q)}(n-) = (B_j(n) - B_j(n-))^q \tag{7.73}$$

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

$$\mu_x = \nu_x = 0.0005 + 0.000075858 \cdot 10^{0.038x}\,,$$
$$\sigma_x = 0.0004 + 0.0000034674 \cdot 10^{0.06x}\,,$$
$$\rho_x = 0.005\,.$$

The intensities μ, ν, and σ 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 $\mu_{02}(t) = \mu_{12}(t) = \mu_{30+t}$, $\mu_{01}(t) = \sigma_{30+t}$, $\mu_{10}(t) = \rho_{30+t}$, $0 < t < 30$ $(= n)$. The central moments $m_t^{(q)j}$ defined in (7.71) – (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.

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 ε-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 ε-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

Time t	0	6	12	18	24	30
$m_t^{(1)0} = m_t^{(1)1}$:	0.0683	0.0771	0.0828	0.0801	0.0592	0
$m_t^{(2)0} = m_t^{(2)1}$:	0.0300	0.0389	0.0484	0.0549	0.0484	0
$m_t^{(3)0} = m_t^{(3)1}$:	0.0139	0.0191	0.0262	0.0343	0.0369	0

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

Time t	0	6	12	18	24	30
$m_t^{(1)0}$:	15.763	13.921	11.606	8.698	4.995	0
$m_t^{(1)1}$:	0.863	0.648	0.431	0.230	0.070	0
$m_t^{(2)0}$:	5.885	5.665	4.740	2.950	0.833	0
$m_t^{(2)1}$:	7.795	5.372	3.104	1.290	0.234	0
$m_t^{(3)0}$:	−51.550	−44.570	−32.020	−15.650	−2.737	0
$m_t^{(3)1}$:	78.888	49.950	25.099	8.143	0.876	0

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

Time t	0	6	12	18	24	30
$m_t^{(1)0}$:	0.277	0.293	0.289	0.239	0.119	0
$m_t^{(1)1}$:	15.176	13.566	11.464	8.708	5.044	0
$m_t^{(2)0}$:	1.750	1.791	1.646	1.147	0.364	0
$m_t^{(2)1}$:	11.502	8.987	6.111	3.107	0.716	0
$m_t^{(3)0}$:	15.960	14.835	11.929	6.601	1.277	0
$m_t^{(3)1}$:	−101.500	−71.990	−42.500	−17.160	−2.452	0

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:

Time t	0	6	12	18	24	30
$m_t^{(1)0}$:	0.0000	0.0410	0.0751	0.0858	0.0533	0
$m_t^{(1)1}$:	7.6451	6.8519	5.8091	4.4312	2.5803	0
$m_t^{(2)0}$:	0.4869	0.5046	0.4746	0.3514	0.1430	0
$m_t^{(2)1}$:	2.7010	2.0164	1.2764	0.5704	0.0974	0
$m_t^{(3)0}$:	2.1047	1.9440	1.5563	0.8686	0.1956	0
$m_t^{(3)1}$:	-12.1200	-8.1340	-4.3960	-1.5100	-0.1430	0

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, $y_{1-\varepsilon} - m^{(1)}$. A possible measure of the riskiness of the portfolio is the ratio $R = \left(y_{1-\varepsilon} - m^{(1)}\right)/P$, 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 $N = 100000$.

7.8 A Markov chain interest model

7.8.1 The Markov model

A. The force of interest process.
The economy (or rather the part of the economy that governs the interest) is a homogeneous time-continuous Markov chain Y on a finite state space $\mathcal{J}^Y = \{1, \ldots, J^Y\}$, with intensities of transition λ_{ef}, $e, f \in \mathcal{J}^Y$, $e \neq f$. The force of interest is r_e when the economy is in state e, that is,

$$r(t) = \sum_e I_e^Y(t) r_e \,, \tag{7.74}$$

where $I_e^Y(t) = 1[Y(t) = e]$ is the indicator of the event that Y is in state e at time t.

Figure F.5 shows a flow-chart of a simple Markov chain interest rate model with three states, 0.02, 0.05, 0.08. Direct transition can only be made to a neighbouring state, and the total intensity of transition out of any state is 0.5, that is, the interest rate changes every two years on the average. By symmetry, the long run average interest rate is 0.05.

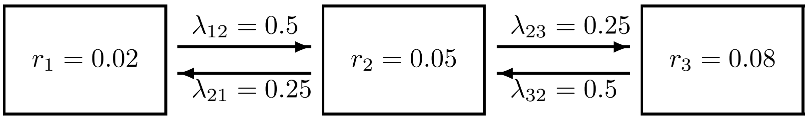


Figure 7.6: Sketch of a simple Markov chain interest model.

B. The payment process.
We adopt the standard Markov chain model of a life insurance policy in Section 7.2 and equip the associated indicator and counting processes with topscript Z to distinguish them from the corresponding entities for the Markov chain governing the interest. We assume the payment stream is of the standard type considered in Section 7.5.

C. The full Markov model.
We assume that the processes Y and Z are independent. Then (Y, Z) is a Markov chain on $\mathcal{J}^Y \times \mathcal{J}^Z$ with intensities

$$\kappa_{ej,fk}(t) = \begin{cases} \lambda_{ef}(t)\,, & e \neq f,\, j = k\,, \\ \mu_{jk}(t)\,, & e = f,\, j \neq k\,, \\ 0\,, & e \neq f,\, j \neq k\,. \end{cases}$$

7.8.2 Differential equations for moments of present values

A. The main result.
For the purpose of assessing the contractual liability we are interested in aspects of its conditional distribution, given the available information at time t. We focus here on determining the conditional moments. By the Markov assumption, the functions in quest are the state-wise conditional moments

$$V_{ej}^{(q)}(t) = \mathrm{E}\left[\left(\frac{1}{v(t)}\int_t^n v\,dB\right)^q \,\middle|\, Y(t) = e, Z(t) = j\right].$$

Copying the proof in Section 7.7, we find that the functions $V_{ej}^{(q)}(\cdot)$ are determined by the differential equations

$$\begin{aligned} \frac{d}{dt}V_{ej}^{(q)}(t) &= (qr_e + \mu_{j\cdot}(t) + \lambda_{e\cdot})V_{ej}^{(q)}(t) - qb_j(t)V_{ej}^{(q-1)}(t) \\ &- \sum_{k;k\neq j}\mu_{jk}(t)\sum_{p=0}^{q}\binom{q}{p}(b_{jk}(t))^p V_{ek}^{(q-p)}(t) - \sum_{f;f\neq e}\lambda_{ef}V_{fj}^{(q)}(t)\,, \end{aligned} \tag{7.75}$$

valid on $(0,n)\backslash\mathcal{D}$ and subject to the conditions

$$V_{ej}^{(q)}(t-) = \sum_{p=0}^{q}\binom{q}{p}(\Delta B_j(t))^p V_{ej}^{(q-p)}(t)\,,\;\; t\in\mathcal{D}.\ \Box \tag{7.76}$$

For $q = 2, 3, \ldots$, denote by $m_{ej}^{(q)}(t)$ the q-th central moment corresponding to $V_{ej}^{(q)}(t)$, and define $m_{ej}^{(1)}(t) = V_{ej}^{(1)}(t)$. Having computed the non-central moments, we obtain the central moments of orders $q > 1$ from

$$m_{ej}^{(q)}(t) = \sum_{p=0}^{q} \binom{q}{p} (-1)^{q-p} V_{ej}^{(p)}(t) \left(V_{ej}^{(1)}(t) \right)^{q-p} .$$

B. Numerical results for a combined insurance policy.
Consider a combined life insurance and disability pension policy issued at time 0 to a person who is then aged x, say. The relevant states of the policy are 1 = *active*, 2 = *disabled*, and 3 = *dead*. At time t, when the insured is $x + t$ years old, transitions between these states take place with intensities

$$\begin{aligned}
&\mu_{13}(t) = \mu_{23}(t) = 0.0005 + 0.000075858 \cdot 10^{0.038(x+t)} , \\
&\mu_{12}(t) = 0.0004 + 0.0000034674 \cdot 10^{0.06(x+t)} , \\
&\mu_{21}(t) = 0.005 .
\end{aligned}$$

We extend the model by assuming that the force of interest may assume three values, $r_1 = \ln(1.00) = 0$ (*low* – in fact no interest), $r_2 = \ln(1.045) = 0.04402$ (*medium*), and $r_3 = \ln(1.09) = 0.08618$ (*high*), and that the transitions between these states are governed by a Markov chain with infinitesimal matrix of the form

$$\Lambda = \lambda \begin{pmatrix} -1 & 1 & 0 \\ 0.5 & -1 & 0.5 \\ 0 & 1 & -1 \end{pmatrix} . \tag{7.77}$$

The scalar λ can be interpreted as the expected number of transitions per time unit and is thus a measure of interest volatility.

Table 1 displays the first three central moments of the present value at time 0 for the following case, henceforth referred to as *the combined policy* for short: the age at entry is $x = 30$, the term of the policy is $n = 30$, the benefits are a life assurance with sum 1 ($= b_{13} = b_{23}$) and a disability annuity with level intensity 0.5 ($= b_2$), and premiums are payable in active state continuously at level rate π ($= -b_1$), which is taken to be the net premium rate in state (2,1) (i.e. the rate that establishes expected balance between discounted premiums and benefits when the insured is active and the interest is at medium level at time 0).

The first three rows in the body of the table form a benchmark; $\lambda = 0$ means no interest fluctuation, and we therefore obtain the results for three cases of fixed interest. It is seen that the second and third order moments of the present value are strongly dependent on the (fixed) force of interest and, in fact, their absolute values decrease when the force of interest increases (as could be expected since increasing interest means decreasing discount factors and, hence, decreasing present values of future amounts).

Table 7.5: Central moments $m_{ej}^{(q)}(0)$ of orders $q = 1, 2, 3$ of the present value of future benefits less premiums for *the combined policy* in interest state e and policy state j at time 0, for some different values of the rate of interest changes, λ. Second column gives the net premium π of a policy starting from interest state 2 (medium) and policy state 1 (active).

	$e, j:$		$1,1$	$1,2$	$2,1$	$2,2$	$3,1$	$3,2$
λ	π	q						
		1	0.15	13.39	0.00	7.65	−0.39	5.03
0	.0131	2	2.55	12.50	0.49	2.70	0.13	0.80
		3	20.45	−99.02	2.11	−12.12	0.37	−2.38
		1	0.06	11.31	0.00	7.90	−0.03	5.78
.05	.0137	2	1.61	12.26	0.62	5.41	0.25	2.43
		3	11.94	−42.87	3.20	−4.33	0.94	−0.08
		1	0.02	8.43	0.00	7.81	−0.02	7.24
.5	.0134	2	0.65	4.90	0.55	4.15	0.46	3.52
		3	3.34	−13.35	2.59	−10.13	2.02	−7.74
		1	0.00	7.77	0.00	7.70	0.00	7.64
5	.0132	2	0.51	2.86	0.50	2.91	0.49	2.86
		3	2.26	−12.51	2.20	−12.19	2.14	−11.88
		1	0.00	7.69	0.00	7.69	0.00	7.69
∞	.0132	2	0.50	2.74	0.50	2.74	0.50	2.74
		3	2.15	−12.37	2.15	−12.37	2.15	−12.37

It is seen that, as λ increases, the differences across the three pairs of columns get smaller and in the end they vanish completely. The obvious interpretation is that the initial interest level is of little importance if the interest changes rapidly.

The overall impression from the two central columns corresponding to medium interest is that, as λ increases from 0, the variance of the present value will first increase to a maximum and then decrease again and stabilize. This observation supports the following piece of intuition: the introduction of moderate interest fluctuation adds uncertainty to the final result of the contract, but if the interest changes sufficiently rapidly, it will behave like fixed interest at the mean level. Presumably, the values of the net premium in the second column reflect the same effect.

7.8.3 Complement on Markov chains

A. Time-continuous Markov chains.
Let $X = \{X(t)\}_{t\geq 0}$ be a time-continuous Markov chain on the finite state space $\mathcal{J} = \{1, \ldots, J\}$. Denote by $P(t, u)$ the $J \times J$ matrix whose j, k-element is the transition probability $p_{jk}(t, u) = P[X(u) = k \mid X(t) = j]$. The Markov property implies the Chapman-Kolmogorov equation

$$P(s, u) = P(s, t)P(t, u)\,, \tag{7.78}$$

valid for $0 \leq s \leq t \leq u$. In particular

$$P(t, t) = I^{J\times J}\,, \tag{7.79}$$

the $J \times J$ identity matrix. The intensity of transition from state j to state k ($\neq j$) at time t is defined as $\kappa_{jk}(t) = \lim_{dt\downarrow 0} p_{jk}(t, t+dt)/dt$ or, equivalently, by

$$p_{jk}(t, t + dt) = \kappa_{jk}(t)\, dt + o(dt)\,, \tag{7.80}$$

when the limit exists. Then, obviously,

$$p_{jj}(t, t + dt) = 1 - \kappa_{j\cdot}(t)dt + o(dt)\,, \tag{7.81}$$

where $\kappa_{j\cdot}(t) = \sum_{k;\, k\neq j} \kappa_{jk}(t)$ can appropriately be termed the total intensity of transition out of state j at time t. The infinitesimal matrix $M(t)$ is the $J \times J$ matrix with $\kappa_{jk}(t)$ in row j and column k, defining $\kappa_{jj}(t) = -\kappa_{j\cdot}(t)$. With this notation (7.80) – (7.81) can be assembled in

$$P(t, t + dt) = I + M(t)dt\,. \tag{7.82}$$

The probabilities determine the intensities. Conversely, the probabilities are determined by the intensities through Kolmogorov's differential equations, which are readily obtained upon combining (7.78) and (7.82). There is a forward equation,

$$\frac{\partial}{\partial t}P(s, t) = P(s, t)M(t)\,, \tag{7.83}$$

and a backward equation,

$$\frac{\partial}{\partial t}P(t,u) = -M(t)P(t,u)\,, \tag{7.84}$$

each of which determine $P(t,u)$ when combined with the condition (7.79).

B. Stationary Markov chains.
When $M(t) = M$, a constant, then (as is obvious from the Kolmogorov equations) $P(s,t) = P(0,t-s)$ depends on s and t only through $t-s$. In this case we write $P(t) = P(0,t)$, allowing a slight abuse of notation. The equations (7.83) – (7.84) now reduce to

$$\frac{d}{dt}P(t) = P(t)M = MP(t)\,. \tag{7.85}$$

The limit $\Pi = \lim_{t\to\infty} P(t)$ exists, and the j-th row of Π is the limiting (stationary) distribution of the state of the process, given that it starts from state j. We shall assume throughout that all states communicate with each other. Then the stationary distribution $\pi' = (\pi_1,\ldots,\pi_J)$, say, is independent of the initial state, and so

$$\Pi = 1^{J\times 1}\pi'\,, \tag{7.86}$$

where $1^{J\times 1}$ is the J-dimensional column vector with all entries equal to 1.

Letting $t \to \infty$ in (7.85) and using (7.86), we get $1^{J\times 1}\pi' M = M1^{J\times 1}\pi' = 0^{J\times J}$ (a matrix of the indicated dimension with all elements equal to 0), that is,

$$\pi' M = 0^{1\times J}\,,\ M1^{J\times 1} = 0^{J\times 1}\,. \tag{7.87}$$

Thus, 0 is an eigenvalue of M, and π' and $1^{J\times 1}$ are corresponding left and right eigenvectors, respectively.

From Paragraph 4.3 of Karlin and Taylor (1975) we gather the following useful representation result. Let ρ_j, $j = 1,\ldots,J$, be the eigenvalues of M and, for each j, let ψ_j' and ϕ_j be the corresponding left and right eigenvectors, respectively. Let Φ be the $J\times J$ matrix whose j-th column is ϕ_j. Then the j-th row of Φ^{-1} is just ψ_j', and introducing $R(t) = \mathrm{diag}(e^{\rho_j t})$, the transition matrix $P(t)$ can be expressed as

$$P(t) = \Phi R(t)\Phi^{-1} = \sum_{j=1}^{J} e^{\rho_j t}\phi_j\psi_j'\,, \tag{7.88}$$

which is computationally convenient. We can take ρ_1 to be 0 and $\phi_1 = 1^{J\times 1}$. Then $\psi_1' = \pi'$, and we obtain

$$P(t) = 1^{J\times 1}\pi' + \sum_{j=2}^{J} e^{\rho_j t}\phi_j\psi_j'\,. \tag{7.89}$$

All the ρ_j, $j = 2, \ldots, J$ are strictly negative, and so the representation shows that the transition probabilities converge exponentially to the stationary distribution.

At this point we need to make precise that in (7.85) the $\frac{d}{dt}$ is to be thought of as an operator, to be distinguished from the matrix $P^{(1)}(t)$ of derivatives it produces when applied to $P(t)$. Now, for $\lambda > 0$, define

$$P_\lambda(t) = P(\lambda t)\,. \tag{7.90}$$

Upon differentiating this relationship and using (7.85), we obtain

$$\frac{d}{dt}P_\lambda(t) = \frac{d}{dt}P(\lambda t) = P^{(1)}(\lambda t)\lambda = P_\lambda(t)\lambda M\,,$$

which shows that $P_\lambda(t)$, which is certainly a matrix of transition probabilities, has infinitesimal matrix

$$M_\lambda = \lambda M\,. \tag{7.91}$$

Thus, doubling (say) the intensities of transition affects the transition probabilities the same way as a doubling of the time period.

7.9 Dependent lives

7.9.1 Introduction

Actuarial tables for multi-life statuses are invariably based on the assumption of mutual independence between the remaining lengths of the individual component lives. The independence hypothesis is computationally convenient or, rather, was so in those days when tables had to be constructed. In the present era of scientific computing such concerns are not so important.

Let S and T be real-valued random variables defined on some probability space. Being mainly interested in survival analysis related to life insurance, we shall let S and T represent the remaining life lengths of two individuals insured under the same policy, let us say husband and wife, respectively. Thus, we make the convenient (but not essential) assumptions that S and T are strictly positive with probability 1 and that they possess a joint density.

The variables S and T are stochastically independent if

$$\mathbb{P}[S > s\,,\, T > t] \;=\; \mathbb{P}[S > s]\,\mathbb{P}[T > t] \text{ for all } s \text{ and } t.$$

In particular, stochastic independence implies that $\mathbb{C}(g(S), h(T)) = 0$ for all functions g and h such that the covariance is well defined. (We let $\mathbb{C}$ and $\mathbb{V}$ denote covariance and variance, respectively.)

Mortality statistics suggest that life lengths of husband and wife are dependent and, moreover, that they are positively correlated. It is easy to think of possible explanations to this empirical fact. For instance, that people who marry do so because they have something in common ('birds of a feather fly together'), or that married people share lifestyle and living conditions and therefore also hazards of diseases and accidents, or that death of the spouse impairs the living conditions for the survivor ('a grief effect', or maybe the husband just does not know where the kitchen is and starves to death shortly after the loss of the spouse).

Correlation is a rather special measure of dependence – essentially it measures linear dependence between random variables – and it is not sufficiently refined for our purposes.

7.9.2 Notions of positive dependence

There are various notions of positive dependence between pairs of random variables, and we will introduce three of them here. A comprehensive reference text is [5].

Definition PQD: S and T are *positively quadrant dependent*, written $\mathrm{PQD}(S,T)$, if

$$\mathbb{P}[S>s\,,\,T>t]\;\geq\;\mathbb{P}[S>s]\,\mathbb{P}[T>t]\text{ for all } s \text{ and } t. \tag{7.92}$$

This definition is symmetric in the two variables, so $\mathrm{PQD}(S,T)$ is the same as $\mathrm{PQD}(T,S)$. The defining inequality (7.92) is equivalent to

$$\mathbb{P}[S>s\,|\,T>t]\;\geq\;\mathbb{P}[S>s]\,, \tag{7.93}$$

which is easy to interpret: knowing e.g. that the wife will survive at least s years improves the survival prospects of the husband.

Definition AS: S and T are *associated*, written $\mathrm{AS}(S,T)$, if

$$\mathbb{C}(g(S,T),h(S,T))\geq 0 \tag{7.94}$$

for all real-valued functions g and h that are increasing in both arguments (and for which the covariance exists).

Also the definition of AS is symmetric in the two variables, so $\mathrm{AS}(S,T)$ is the same as $\mathrm{AS}(T,S)$.

Definition RTI: S is *right tail increasing in* T, written $\mathrm{RTI}(S|T)$, if

$$\mathbb{P}[S>s\,|\,T>t]\text{ is an increasing function of } t \text{ for each fixed } s. \tag{7.95}$$

The definition of RTI is not symmetric in the two variables.

To each notion of positive dependence there is a corresponding notion of negative dependence. We can reasonably say that S and T are negatively quadrant dependent if the inequality (7.92) is reversed. This is the same as $\mathrm{PQD}(-S,T)$, see Exercise 21. We can say that S and T are negatively associated ('dissociated' does not have the right connotation) if the inequality (7.94) is reversed. This is the same as $\mathrm{AS}(-S, T)$, see Exercise 21. We say that S is *right tail decreasing in* T, written $\mathrm{RTD}(S|T)$ if $\mathbb{P}[S > s \,|\, T > t]$ is a decreasing function of t for each fixed s. This is the same as $\mathrm{RTD}(-S|T)$, see Exercise 21. Since results on positive dependence thus translate into results on negative dependence, we will henceforth focus on the former.

Theorem 1: $\mathrm{RTI}(S|T) \Rightarrow \mathrm{AS}(S,T) \Rightarrow \mathrm{PQD}(S,T)$.

Proof (incomplete): The first implication, $\mathrm{RTI}(S|T) \Rightarrow \mathrm{AS}(S,T)$, is the hard part. The proof is long and technical and can be found in [20].

The second implication $\mathrm{AS}(S,T) \Rightarrow \mathrm{PQD}(S,T)$ is easy. For $g(S,T) = 1_{(s,\infty)}(S) = 1[S > s]$ and $h(S,T) = 1_{(t,\infty)}(T) = 1[T > t]$, (7.94) reduces to

$$\mathbb{C}(1[S > s], 1[T > t]) \geq 0\,, \tag{7.96}$$

which is just a reformulation of the defining inequality (7.92).

As a partial compensation for the absence of proof of the first implication, let us prove the shortcut implication $\mathrm{RTI}(S|T) \Rightarrow \mathrm{PQD}(S,T)$: if $\mathrm{RTI}(S|T)$, then $\mathbb{P}[S > s \,|\, T > t] \geq \mathbb{P}[S > s \,|\, T > 0]$ for $t > 0$, which is the same as (7.93). □

The following result is a partial converse to the second implication in Theorem 1. It could be formulated by saying that positive quadrant dependence is equivalent to "marginal association".

Lemma 1: $\mathrm{PQD}(S,T) \Leftrightarrow \mathbb{C}(g(S), h(T)) \geq 0$ for increasing functions g and h.

Proof: By (7.96) the result holds for increasing indicator functions. Then it holds for increasing simple functions, $g(S) = g_0 + \sum_{i=1}^{m} g_i 1[S > s_i]$ and $h(T) = h_0 + \sum_{j=1}^{n} h_j 1[T > t_j]$ (with constant coefficients g_i and h_j and $g_i > 0$, $i = 1, \ldots, m$ and $h_j > 0$, $j = 1, \ldots, n$), as is seen from

$$\mathbb{C}(g(S), h(T)) = \sum_{i=1}^{m} \sum_{j=1}^{n} g_i h_j \mathbb{C}\left(1[S > s_i], 1[T > t_j]\right) \geq 0\,.$$

Then it holds for all increasing functions $g(S)$ and $h(T)$ since since any increasing function from $\mathbb{R}$ to $\mathbb{R}$ can be written as the limit of a sequence of increasing simple functions (monotone convergence). □

In the definitions of PQD, AS, and RTI we could equally reasonably have entered the events $S \geq s$ and $T \geq t$. Due to the continuity property of probability measures, it does not matter which inequalities we use, $>$ or $\geq$. See Exercise 21. for $A_1 \subseteq A_2 \subseteq \ldots \lim \mathbb{P}[A_n] = \mathbb{P}[\cup_n A_n]$ for $A_1 \supseteq A_2 \supseteq \ldots \lim \mathbb{P}[A_n] = \mathbb{P}[\cap_n A_n]$

7.9.3 Dependencies between present values

The following table lists formulas for the life lengths and the survival functions of the four statuses husband, wife, their joint life, and their last survivor.

Status (z)	Life length U	Survival function $\mathbb{P}[U > \tau]$
Husband (x)	S	$\mathbb{P}[S > \tau]$
Wife (y)	T	$\mathbb{P}[T > \tau]$
Joint life (x, y)	$S \wedge T$	$\mathbb{P}[S > \tau\,,\, T > \tau]$
Last survivor $\overline{x, y}$	$S \vee T$	$\mathbb{P}[S > \tau] + \mathbb{P}[T > \tau] - \mathbb{P}[S > \tau\,,\, T > \tau]$

The next table recapitulate the formulas for present values and their expected values for the most basic insurance benefits to a status (z) with remaining life length U.

Payment scheme	Present value	Expected present value
Pure endowment	$e^{-rn}\, 1[U > n]$	${}_nE_z = e^{-rn}\, \mathbb{P}[U > n]$
Life annuity	$\int_0^n e^{-r\tau}\, 1[U > \tau]\, d\tau$	$\bar{a}_{z\,\overline{n}\rceil} = \int_0^n e^{-r\tau}\, \mathbb{P}[U > \tau]\, d\tau$
Term insurance	$e^{-rU} 1[U < n]$	$\bar{A}_{\overset{1}{x}\,\overline{n}\rceil} = 1 - {}_nE_z - \bar{a}_{z\,\overline{n}\rceil}$

The life lengths of the statuses listed in the first table are increasing functions of both S and T. From the second table we see that, for a general status with remaining life length U, the present value of a pure survival benefit (life endowment or life annuity) is an increasing functions of U, whereas the present value of the pure death benefit is a decreasing function of U (we assume the interest rate r is positive.)

Combining these observations and Theorem 1, we can infer the following (and many other things): If PQD(S, T), then pure survival benefits on any two statuses are positively dependent, pure death benefits on any two statuses are positively dependent, and any pure survival benefit and any death benefit are negatively dependent.

We can also draw conclusions about the bias introduced in equivalence premiums by erroneously adopting the independence hypothesis. For instance, if PQD(S, T) and we work under the independence hypothesis, then the present value of a survival benefit on the joint life will be underestimated, whereas the present value of a survival benefit on the last survivor will be overestimated. For the death benefit it is the other way around. Combining these things we may conclude e.g. that, for a death benefit on the joint life against level premium during joint survival, the equivalence premium will be overestimated.

7.9.4 A Markov chain model for two lives

It is not easy to create a given form of dependence between the life lengths S and T by direct specification of their joint distribution. However, the process point of view, which is a powerful one, quite naturally allows us to express various ideas about dependencies between life lengths of a couple. A suitable framework is the Markov model sketched in Figure F.4.

The following formulas are obvious:

$$\begin{aligned}
p_{00}(s,t) &= e^{-\int_s^t \mu+\nu}\,, \\
p_{01}(s,t) &= \int_s^t e^{-\int_s^\tau \mu+\nu}\mu_\tau e^{-\int_\tau^t \nu'}\,d\tau\,, \\
p_{02}(s,t) &= \int_s^t e^{-\int_s^\tau \mu+\nu}\nu_\tau e^{-\int_\tau^t \mu'}\,d\tau\,.
\end{aligned}$$

The joint survival function of S and T is

$$\begin{aligned}
\mathbb{P}[S>s\,,\,T>t] &= \begin{cases} p_{00}(0,t)+p_{00}(0,s)p_{01}(s,t)\,, s\le t\,, \\ p_{00}(0,s)+p_{00}(0,t)p_{02}(s,t)\,, s>t\,, \end{cases} \\
&= \begin{cases} e^{-\int_0^t \mu+\nu}+\int_s^t e^{-\int_0^\tau \mu+\nu}\mu_\tau e^{-\int_\tau^t \nu'}\,, s\le t\,, \\ e^{-\int_0^s \mu+\nu}+\int_t^s e^{-\int_0^\tau \mu+\nu}\nu_\tau e^{-\int_\tau^s \mu'}\,, s>t\,. \end{cases} \qquad (7.97)
\end{aligned}$$

The marginal survival function of T is (put $s=0$ in (7.97))

$$\begin{aligned}
\mathbb{P}[T>t] &= p_{00}(0,t)+p_{01}(0,t) \\
&= e^{-\int_0^t \mu+\nu}+\int_0^t e^{-\int_0^\tau \mu+\nu}\mu_\tau e^{-\int_\tau^t \nu'}\,, t\ge 0\,. \qquad (7.98)
\end{aligned}$$

It is intuitively obvious that S and T are independent if $\mu'_\tau=\mu_\tau$ and $\nu'_\tau=\nu_\tau$ for all τ, and this will follow from Theorem 2 below (see also Exercise 1). It is also intuitively clear that S and T will become dependent if we let the mortality rates depend on marital status. Let us see what happens if the mortality rate increases upon the loss of the spouse.

Theorem 2: If $\mu'_\tau\ge\mu_\tau$ and $\nu'_\tau\ge\nu_\tau$ for all τ, then S and T are positively dependent in the sense RTI($S|T$) (hence AS(S,T) and PQD(S,T)).

If $\mu'_\tau\le\mu_\tau$ and $\nu'_\tau\le\nu_\tau$ for all τ, then S and T are negatively dependent in the sense RTD($S|T$) (hence AS($-S,T$) and PQD($-S,T$)).

If $\mu'_\tau=\mu_\tau$ and $\nu'_\tau=\nu_\tau$ for all τ, then S and T are independent.

Proof: Consider first the case $s\le t$. From (7.97) and (7.98) we get

$$\begin{aligned}
\mathbb{P}[S>s\,|\,T>t] &= \frac{e^{-\int_0^t \mu+\nu}+\int_s^t e^{-\int_0^\tau \mu+\nu}\,\mu_\tau\,e^{-\int_\tau^t \nu'}\,d\tau}{e^{-\int_0^t \mu+\nu}+\int_0^t e^{-\int_0^\tau \mu+\nu}\,\mu_\tau\,e^{-\int_\tau^t \nu'}\,d\tau} \\
&= 1-\frac{\int_0^s e^{-\int_0^\tau \mu+\nu-\nu'}\,\mu_\tau\,d\tau}{e^{-\int_0^t \mu+\nu-\nu'}+\int_0^t e^{-\int_0^\tau \mu+\nu-\nu'}\,\mu_\tau\,d\tau}\,.
\end{aligned}$$

Now we need only to study the denominator in the second term as a function of t. Its derivative is

$$e^{-\int_0^s \mu+\nu-\nu'}(\nu'_t-\nu_t)\,.$$

It follows that $\mathbb{P}[S > s \,|\, T > t]$ is an increasing function function of t if $\nu'_t \geq \nu_t$ and a decreasing function function of t if $\nu'_t \leq \nu_t$.

Next consider the case $s > t$. This is a bit more complicated. From (7.97) and (7.98) we get

$$\mathbb{P}[S > s \,|\, T > t] \quad = \quad \frac{e^{-\int_0^s \mu+\nu} + \int_t^s e^{-\int_0^\tau \mu+\nu} \nu_\tau \, e^{-\int_\tau^s \mu'} \, d\tau}{e^{-\int_0^t \mu+\nu} + \int_0^t e^{-\int_0^\tau \mu+\nu} \mu_\tau \, e^{-\int_\tau^t \nu'} \, d\tau} \, .$$

By the rule $d(u/v) = (v\,du - u\,dv)/v^2$, the sign of $\frac{\partial}{\partial t}\mathbb{P}[S > s \,|\, T > t]$ is the same as that of

$$\begin{aligned}
&\left(e^{-\int_0^t \mu+\nu} + \int_0^t e^{-\int_0^\tau \mu+\nu} \mu_\tau \, e^{-\int_\tau^t \nu'} \, d\tau\right) \left(-e^{-\int_0^t \mu+\nu} \nu_t \, e^{-\int_t^s \mu'}\right) \\
&- \left(e^{-\int_0^s \mu+\nu} + \int_t^s e^{-\int_0^\tau \mu+\nu} \nu_\tau \, e^{-\int_\tau^s \mu'} \, d\tau\right) \times \\
&\left(e^{-\int_0^t \mu+\nu}(-\mu_t - \nu_t) + e^{-\int_0^t \mu+\nu}\mu_t + \int_0^t e^{-\int_0^\tau \mu+\nu} \mu_\tau \, e^{-\int_\tau^t \nu'} \, d\tau \, (-\nu'_t)\right) \, .
\end{aligned}$$

In this expression two terms cancel in the last parenthesis. Further, to get rid of some common factors, let us multiply with $e^{\int_0^s \mu+\nu} \, e^{\int_0^t \mu+\nu}$, which preserves the sign and turns the expression into

$$\begin{aligned}
&- \left(1 + \int_0^t e^{\int_\tau^t \mu+\nu-\nu'} \mu_\tau \, d\tau\right) e^{\int_t^s \mu-\mu'+\nu} \, \nu_t \\
&+ \left(1 + \int_t^s e^{\int_\tau^s \mu-\mu'+\nu} \nu_\tau \, d\tau\right) \left(\nu_t + \int_0^t e^{\int_\tau^t \mu+\nu-\nu'} \mu_\tau \, d\tau \, \nu'_t\right) \, .
\end{aligned}$$

Substituting

$$\begin{aligned}
\int_t^s e^{\int_\tau^s \mu-\mu'+\nu} \nu_\tau \, d\tau \quad &= \quad \int_t^s e^{\int_\tau^s \mu-\mu'+\nu} (\mu_\tau - \mu'_\tau + \nu_\tau) \, d\tau \\
&\quad + \int_t^s e^{\int_\tau^s \mu-\mu'+\nu} (\mu'_\tau - \mu_\tau) \, d\tau \\
&= \quad e^{\int_t^s \mu-\mu'+\nu} - 1 + \int_t^s e^{\int_\tau^s \mu-\mu'+\nu} (\mu'_\tau - \mu_\tau) \, d\tau
\end{aligned}$$

and rearranging a bit, we arrive at

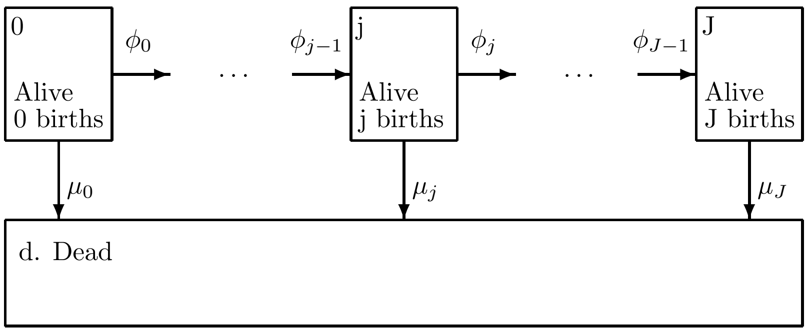
$$\begin{aligned}
&\left(\nu_t + \nu'_t \int_0^t e^{\int_\tau^t \mu+\nu-\nu'} \mu_\tau \, d\tau\right) \int_t^s e^{\int_\tau^s \mu-\mu'+\nu} (\mu'_\tau - \mu_\tau) \, d\tau \\
&+ \int_0^t e^{\int_\tau^s \mu+\nu-\nu'} \mu_\tau \, d\tau \, (\nu'_t - \nu_t) \, .
\end{aligned}$$

It follows that $\mathbb{P}[S > s \,|\, T > t]$ is an increasing function function of t if $\mu'_t \geq \mu_t$ and $\nu'_t \geq \nu_t$ and a decreasing function function of t if $\mu'_t \leq \mu_t$ and $\nu'_t \leq \nu_t$. $\square$

7.10 Conditional Markov chains

7.10.1 Retrospective fertility analysis

In connection with a pension insurance scheme there is an additional benefit which is a sum insured payable to possible dependent children less than 18 years old at the time of death of the insured. In the technical basis we therefore need to make assumptions about births. We have to distinguish by sex, and in the following we consider female insured only. The Figure below shows a flowchart for possible life histories with death and births (at most J). To keep things simple, we assume that the insured enters the scheme in state 0 at age 0 and that the process is Markov: for a t year old who has given birth to j children, the mortality rate $\mu_j(t)$ and the fertility rate $\phi_j(t)$ are functions of t and j only.



Assume now that the past history of births and death is observed only upon death of the insured, when the additional benefit to the possible dependents is due. Suppose that the statistical data comprise only those who are dead at the time of consideration and that for each of those there is a complete record of the times of possible births and of death. In these data the observed life history of a woman, who entered the scheme u years ago, is governed by a Markov process as described above, but with intensities

$$\mu_j^*(t) = \mu_j(t)\frac{1}{p_{jd}(t,u)}\,, \tag{7.99}$$

$$\phi_j^*(t) = \phi_j(t)\frac{p_{j+1,d}(t,u)}{p_{jd}(t,u)}\,. \tag{7.100}$$

We see that $\mu_j^*(t) \geq \mu_j(t)$, which is easy to explain (we are looking at the mortality, given death).

We are going to prove a more interesting result: If mortality increases with the number of births, that is,

$$\mu_j(t) \leq \mu_{j+1}(t)\,,\ j = 0,\ldots,J-1,\ t > 0\,, \tag{7.101}$$

then

$$\phi_j^*(t) \geq \phi_j(t)\,,\ j = 0, \ldots, J-1\,,\ t > 0\,. \tag{7.102}$$

We need to prove that $p_{j+1,d}(t,u) \geq p_{jd}(t,u)$, $j = 1, \ldots, J-1$. It is convenient to work with

$$p_j(t,u) = 1 - p_{jd}(t,u) = \sum_{k=j}^{J} p_{jk}(t,u)\,, \tag{7.103}$$

the probability that a t year old with j births will survive to age u, and to prove the hypothesis

$$H_j:\ p_k(t,u) \leq p_j(t,u)\,, \qquad k = j+1, \ldots, J, \tag{7.104}$$

for $j = 0, \ldots, J-1$. The proof goes by induction 'downwards', proving that H_{j+1} implies H_j. Thus assume H_{j+1} is true.

By direct reasoning (or an easy calculation) the mortality intensity at age u ($> t$) associated with the survival function (F.49) is

$$\mu_j(t,u) = \frac{\sum_{k\geq j} p_{jk}(t,u)\mu_k(u)}{\sum_{k\geq j} p_{jk}(t,u)}\,, \tag{7.105}$$

hence

$$p_j(t,u) = e^{-\int_t^u \mu_j(t,s)ds}\,. \tag{7.106}$$

Two more expressions for $p_j(t,u)$, both obvious, are

$$p_j(t,u) = \sum_{k\geq j} p_{jk}(t,\tau)\, p_k(\tau,u)\,, \tag{7.107}$$

$t \leq \tau \leq u$, and

$$p_j(t,u) = e^{-\int_t^u (\phi_j+\mu_j)} + \int_t^u e^{-\int_t^\tau (\phi_j+\mu_j)}\phi_j(\tau)\, p_{j+1}(\tau,u)\, d\tau\,. \tag{7.108}$$

By (F.47) and (7.105) we have

$$\mu_j(u) \leq \mu_{j+1}(t,u)\,, \tag{7.109}$$

hence

$$e^{-\int_t^u \mu_j} \geq p_{j+1}(t,u)\,.$$

Therefore, from (F.53) we get

$$\begin{aligned} p_j(t,u) \quad &\geq \quad e^{-\int_t^u \phi_j}\, p_{j+1}(t,u) \\ &\quad + \int_t^u e^{-\int_t^\tau \phi_j}\phi_j(t,\tau)\, p_{j+1}(t,\tau)\, p_{j+1}(\tau,u)\, d\tau\,. \end{aligned} \tag{7.110}$$

Focusing on the last two factors under the integral, use in succession (F.49), the induction hypothesis (7.104), and (F.50), to deduce

$$\begin{aligned} p_{j+1}(t,\tau)p_{j+1}(\tau,u) &= \sum_{k=j+1}^{J} p_{j+1,k}(t,\tau)p_{j+1}(\tau,u) \\ &\geq \sum_{k=j+1}^{J} p_{j+1,k}(t,\tau)p_k(\tau,u) \\ &= p_{j+1}(t,u)\,. \end{aligned}$$

Putting this into (7.110), we obtain

$$p_j(t,u) \geq \left(e^{-\int_t^u \phi_j} + \int_t^u e^{-\int_t^\tau \phi_j}\phi_j(t,\tau)\,d\tau\right) p_{j+1}(t,u) \;=\; p_{j+1}(t,u)\,.$$

It follows that H_j is true. Since H_{J-1} is obviously true, we are done. □

Comment: The inequality (F.48) means that the fertility rates will be overestimated if one uses the estimators for the ϕ_j^* based on diseased participants in the scheme. If the inequalities (F.47) are reversed, then also the inequality (F.48) will be reversed, and the estimators the ϕ_j^* will underestimate the fertility. In particular it follows that, under the hypothesis of non-differential mortality, the fertility rates will be unbiasedly estimated from the selected material of diseased participants.

Chapter 8

Probability distributions of present values

Abstract: A system of integral equations is obtained for the statewise probability distributions of the present value of future payments on a multistate life insurance policy under Markov assumptions. They are brought on a differential form convenient for computation and applied to some cases.

Key words: Multistate life insurance, Markov counting process, optional sampling, stochastic interest.

8.1 Introduction

A. Background and motive of the present study.
By tradition, life insurance mathematics centers on conditional expected values of discounted cashflows. A key tool of the theory are Thiele's differential equations, which describe the development of such expected values for a multistate policy driven by a Markov process. In two recent papers one of the authors (Norberg, 1994, 1995) obtains differential equations for higher order moments of present values and offers examples of their potential uses in solvency assessments and in construction of untraditional insurance products. In continuance of those results, we undertake to determine the probability distributions that are at the base of the moments and of any other expected values of interest. Knowledge of the distribution of the present value, and in particular its upper tail, gives insight into the riskiness of the contract beyond what is provided by the mean and the higher order moments.

B. Contents of the paper.
By way of introduction, Section 2 deals briefly with models involving only 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. This approach comes down overagainst more complex situations where stochastic processes have to be employed. In Section 3 we consider a general multistate insurance policy with payments of assurance and annuity types and with state-dependent force of interest. Assuming that the state process is Markov, we derive in Section 4 a set of integral equations for the conditional probability distribution function of the present value of future benefits less premiums, given the state at the time of consideration. The equations are converted to a differential form that forms the basis of a computational scheme. Examples of applications, including numerical illustrations, are given in Section 5.

8.2 Calculation of probability distributions of present values by elementary methods

A. A simple example involving only one life length.
De Pril (1989) and Dhaene (1990) compile lists of distributions of present values of standard single-life insurance benefits ; pure endowment, term assurance, endowment assurance, and life annuity. To offer an example along this line, consider a life insurance policy specifying that the sum assured b is to be paid out immediately upon the (possible) death of the insured within n years after the date of issue of the policy and that premiums are payable continuously at level rate c per year as long as the contract is in force. Suppose interest accumulates with constant intensity δ so that the τ years discount factor is $v^\tau = e^{-\delta\tau}$. Denoting the remaining life time of the insured by T, the present value of benefits less premiums on the contract is $U(T) = bv^T 1_{(0,n]}(T) - c\bar{a}_{\overline{T\wedge n}|}$, where $\bar{a}_{\overline{t}|} = \int_0^t v^\tau d\tau = (1 - e^{-\delta t})/\delta$ is the present value of an annuity certain payable continuously at level rate 1 per year for t years. The function U is nonincreasing in T and, letting T be a random variable, we easily find the probability distribution (4.59).

The jump at $-c\bar{a}_{\overline{n}|}$ is due to the positive probability of survival to time n. Similar effects are to be anticipated also for more complex finite term insurance products since, in general, there is a positive probability that the policy will remain in the current state until the contract terminates.

Fig. 1 shows the graph of the function in (4.59) for the case where $\delta = \ln(1.045) = 0.044017$, $\mathrm{P}[T > t] = e^{-\int_0^t \mu(\tau)d\tau}$ with $\mu(t) = 0.0005 + 0.000075858 \cdot 10^{0.038(30+t)}$, $n = 30$, $b = 1$, and $c = 0.0042608$; interest and mortality are according to the first order technical basis currently used by most Danish insurers, supposing the insured is a male who is 30 years old at the time when the policy is issued, and the given value of c is the net premium rate. This contract will henceforth be referred to as *the term insurance policy.*

Figure 1 about here.

Fig. 1: Probability distribution of the present value at time 0 of the term in-

surance policy.

B. Models involving a finite number of random variables.
The analysis above was straightforward because the present value is a function of only one random variable. The approach works whenever only a finite number of random variables, $T_1, \ldots, T_r$, are involved, e.g. for a multilife insurance depending on r lifetimes. The probability distribution of a present value $U(T_1, \ldots, T_r)$, say, is obtained by integrating the joint probability function of the T_i-s over the sets $\{(t_1, \ldots, t_r);\, U(t_1, \ldots, t_r) \leq u\}$ for (in principle) all u. When applicable at all, this procedure would usually be far more complicated than the one we are going to demonstrate. Thus, with no further ado, we now turn to the principal message of the paper.

8.3 The general Markov multistate policy

A. Payments, interest, and present values.
Consider a stream of payments in respect of a contract issued at time 0 and terminating at time n, say, and denote by $B(t)$ the total amount paid in the time interval $[0, t]$. The payment function $\{B(t)\}_{t \geq 0}$ is assumed to be right-continuous and of bounded variation. Money is currently invested in (or borrowed from) a fund that at any time t yields $\delta(t)$ in return per unit of time and unit amount on deposit, that is, $\delta(t)$ is the force of interest at time t. Then the discounted value at time t of a unit due at time τ is $e^{-\int_t^\tau \delta}$, and so the present value at time t of the future payments under the contract is

$$\int_t^n e^{-\int_t^\tau \delta}\, dB(\tau). \tag{8.1}$$

The short-hand exemplified by $\int \delta = \int \delta(s) ds$ will be in frequent use throughout. By convention, $\int_a^b$ means $\int_{(a,b]}$ if $b < \infty$ and $\int_{(a,\infty)}$ if $b = \infty$.

B. The multi-state insurance policy.
We adopt the standard set-up of life insurance mathematics as presented in Norberg (1994). There is a set of states, $\mathcal{J} = \{1, \ldots, J\}$, such that at any time $t \in [0, n]$ the policy is in one an only one state. Denote by $X(t)$ the state of the policy at time t. Considered as a function of t, $\{X\}_{t \geq 0}$ is taken to be right-continuous, and $X(0) = 1$ implying (as a convention) that the policy commences in state 1. Introduce $I_j(t) = 1[X(t) = j]$, the indicator of the event that the policy is in state j at time t, and $N_{jk}(t) = \sharp\{\tau; \tau \in (0, t], X_{\tau-} = j, X(\tau) = k\}$, the total number of transitions of X from state j to state k $(\neq j)$ by time t. The payment function B is assumed to be of the form (7.41), where each B_j is a deterministic payment function specifying payments due during sojourns in state j (a general life annuity) and each b_{jk} is a deterministic function specifying payments due upon transitions from state j to state k (a general life assurance). The left-limit in $I_j(t-)$ means that the state j annuity is effective at time t if

the policy is in state j just prior to (but not necessarily at) time t. Consistently we define $I_j(0-) = 1$.

We shall allow the force of interest to depend on the current state, that as in (7.74.

C. The time-continuous Markov model.
It is assumed that $\{X(t)\}_{t\geq 0}$ is a (continuous-time) Markov chain. Denote the transition probabilities by

$$p_{jk}(t,u) = \mathrm{P}[X(u) = k | X(t) = j].$$

The transition intensities

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

are assumed to exist for all $j, k \in \mathcal{J}$, $j \neq k$. The total intensity of transition out of state j is $\mu_{j\cdot}(t) = \sum_{k;k\neq j} \mu_{jk}(t)$. The probability of staying uninterruptedly in state j during the time interval from t to u is $e^{-\int_t^u \mu_{j\cdot}}$.

8.4 Differential equations for statewise distributions

A. The statewise probability distributions.
The problem is to determine the conditional probability distribution of the liability in (8.1), given the information available at time t. Since the Markov assumption implies conditional independence between past and future for fixed present state of the policy, the relevant functions are *the statewise probability distributions* defined by

$$P_j(t,u) = \mathrm{P}\left[\int_t^n e^{-\int_t^\tau \delta} dB(\tau) \leq u \,\middle|\, I_j(t) = 1\right], \tag{8.1}$$

$t \in [0,n]$, $u \in \mathcal{R}$, $j \in \mathcal{J}$.

B. A system of integral equations.
A simple heuristic argument will establish that the probabilities in (8.1) satisfy the integral equations

$$\begin{aligned} P_j(t,u) \quad = & \quad \sum_{k;k\neq j} \int_t^n e^{-\int_t^s \mu_{j\cdot}} \mu_{jk}(s)\, ds \\ & \cdot P_k\left(s, e^{\delta_j(s-t)}u - \int_t^s e^{\delta_j(s-\tau)} dB_j(\tau) - b_{jk}(s)\right) \\ & + e^{-\int_t^n \mu_{j\cdot}} 1\left[\int_t^n e^{-\delta_j(\tau - t)} dB_j(\tau) \leq u\right]. \end{aligned} \tag{8.2}$$

In the first terms on the right here the factor $e^{-\int_t^s \mu_{j\cdot}}\mu_{jk}(s)ds$ is the probability that the policy stays in state until time $s\,(< n)$ and then makes a transfer to state $k\,(\neq j)$ in the small time interval $[s, s+ds)$. In this case the annuity B_j is in force during the time interval $(t, s]$, the lump sum $b_{jk}(s)$ falls due at time s, and the interest rate during this time interval is δ_j, and so the event

$$\int_t^n e^{-\int_t^\tau \delta} dB(\tau) \leq u \tag{8.3}$$

takes place if

$$\int_t^s e^{-\delta_j(\tau-t)} dB_j(\tau) + e^{-\delta_j(s-t)} b_{jk}(s) + e^{-\delta_j(s-t)} \int_s^n e^{-\int_s^\tau \delta} dB(\tau) \leq u$$

or, equivalently,

$$\int_s^n e^{-\int_s^\tau \delta} dB(\tau) \leq e^{\delta_j(s-t)} \left(u - \int_t^s e^{-\delta_j(\tau-t)} dB_j(\tau) \right) - b_{jk}(s).$$

Thus, the corresponding conditional probability is

$$P_k \left(s, e^{\delta_j(s-t)} u - \int_t^s e^{\delta_j(s-\tau)} dB_j(\tau) - b_{jk}(s) \right).$$

Summing over all times s and states k, we obtain the first terms on the right of (8.2), which thus is the part of the total probability that pertains to exit from state j before time n.

Likewise it is realized that the last term on the right of (8.2) is the remaining part of the probability, pertaining to the case of no transition out of state j before time n.

This heuristic argument is made rigorous by applying Doob's optional sampling theorem to the martingale generated by the indicator of the event in (8.3) and the stopping time defined as the the minimum of n and the time of the first transition after t from the current state j.

C. A system of differential equations.
Already (8.2) might serve as a basis for computation of the statewise probability functions, but is not convenient since the integrand on the right depends on t. Introduce the auxiliary functions Q_j defined by

$$Q_j(t, u) = P_j \left(t, e^{\delta_j t} \left(u - \int_0^t e^{-\delta_j \tau} dB_j(\tau) \right) \right) \tag{8.4}$$

or

$$P_j(t,u) = Q_j\left(t, e^{-\delta_j t}u + \int_0^t e^{-\delta_j \tau} dB_j(\tau)\right). \tag{8.5}$$

Multiply by $e^{-\int_0^t \mu_{j\cdot}}$ in (8.2), insert $e^{\delta_j t}\left(u - \int_0^t e^{-\delta_j \tau} dB_j(\tau)\right)$ in the place of u, and rearrange a bit to obtain

$$\begin{aligned} e^{-\int_0^t \mu_{j\cdot}} Q_j(t,u) &= \int_t^n e^{-\int_0^s \mu_{j\cdot}} \sum_{k;k\neq j} \mu_{jk}(s)\, ds \\ &\quad \cdot Q_k\Bigg(s, e^{(\delta_j - \delta_k)s}u + \int_0^s e^{-\delta_k \tau} dB_k(\tau) \\ &\quad - e^{(\delta_j-\delta_k)s}\int_0^s e^{-\delta_j\tau} dB_j(\tau) - e^{-\delta_k s} b_{jk}(s)\Bigg) \\ &\quad + e^{-\int_0^n \mu_{j\cdot}}\, 1\left[\int_0^n e^{-\delta_j\tau} dB_j(\tau) \le u\right]. \end{aligned} \tag{8.6}$$

Now the integrand on the right does not contain t, and we are allowed to differentiate along t on the right hand side by simply substituting t for s in minus the integrand. Performing this and cancelling the common factor $e^{-\int_0^t \mu_{j\cdot}}$, we arrive at the following main result, where the side condition (7.76) comes directly out of (8.6) by letting $t \uparrow n$:

Theorem. *The functions Q_j in* (8.4) *are the unique solutions to the differential equations*

$$\begin{aligned} d_t Q_j(t,u) &= \mu_{j\cdot}(t) dt\, Q_j(t,u) - \sum_{k;k\neq j} \mu_{jk}(t)\, dt \\ &\quad \cdot Q_k\Bigg(t, e^{(\delta_j-\delta_k)t}u + \int_0^t e^{-\delta_k\tau} dB_k(\tau) \\ &\quad - e^{(\delta_j-\delta_k)t}\int_0^t e^{-\delta_j\tau} dB_j(\tau) - e^{-\delta_k t} b_{jk}(t)\Bigg), \end{aligned} \tag{8.7}$$

$0 \le t \le n$, subject to the conditions

$$Q_j(n,u) = 1\left[\int_0^n e^{-\delta_j\tau} dB_j(\tau) \le u\right]. \tag{8.8}$$

Having determined the auxiliary Q_j, we obtain the P_j from (8.5).

Remark: The differentiation is in the Stieltje's sense for functions of bounded variation and does not require differentiability or any other smoothness properties of the functions involved.

D. Computational scheme.
A simple numerical procedure consists in approximating the functions Q_j by the functions Q^*_j obtained from the finite difference version of (8.7):

$$\begin{aligned} Q^*_j(t-h,u) &= (1-\mu_{j\cdot}(t)h)Q^*_j(t,u) + h\sum_{k;k\neq j}\mu_{jk}(t) \\ &\quad \cdot Q^*_k\left(t, e^{(\delta_j-\delta_k)t}u + \int_0^t e^{-\delta_k\tau}dB_k(\tau)\right. \\ &\quad \left. -e^{(\delta_j-\delta_k)t}\int_0^t e^{-\delta_j\tau}dB_j(\tau) - e^{-\delta_k t}b_{jk}(t)\right). \end{aligned} \tag{8.9}$$

Starting from (8.8) (with Q^*_j in the place of Q_j), one calculates first the functions $Q^*_j(n-h,\cdot)$ by (8.9) and continues recursively until the $Q^*_j(0,\cdot)$ have been calculated in the final step.

The $Q^*_j(t,u)$ are defined for $t \in \{0, h, 2h, \ldots, n\}$ and $u \in \{a, a+h', a+2h', \ldots, b\}$, say, where the steplengths h and h' must be sufficiently small and a and b must be chosen such that the supports of the $Q_j(t,\cdot)$ are sufficiently well covered by $[a,b]$. What is "sufficient" must be decided on in each individual case by judgement and by trial and error. If two states j and k are intercommunicating, then $Q_j(t,\cdot)$ and $Q_k(t,\cdot)$ have the same support. The supports are finite and usually easy to determine in situations where the number of assurance payments has a non-random upper bound. The u-arguments in the functions Q^*_k on the right of (8.9) must, of course, be rounded to the nearest point in $\{a, a+h', a+2h', \ldots, b\}$.

E. Comments on the method.
Before turning to applications, we pause in this paragraph to offer some motivation and discussion of our approach. It is not needed in the sequel and may be skipped on the first reading.

The equation (8.7) is just the differential form of the integral equation (8.6). It does not require that the derivatives $\frac{\partial}{\partial t}Q_j(t,u)$ exist, which they do not in general for the obvious reason that the statewise annuity functions on the right of (8.7) may have jumps.

If one should attempt to construct differential equations along the lines of Norberg (1994), the starting point would be the martingale M defined by

$$M(t) = \mathrm{P}\left[\int_0^n e^{-\int_0^\tau \delta} dB(\tau) \leq u \mid \mathcal{F}_t\right],$$

where $\mathcal{F}_t = \sigma\{X(\tau); \tau \leq t\}$ is the information generated by the state process up to time t. Using the Markov property of conditional independence between past and future, given the present, we find

$$M(t) = \sum_j I_j(t)\, P_j\left(t, e^{\int_0^t \delta}\left(u - \int_0^t e^{-\int_0^\tau \delta} dB(\tau)\right)\right).$$

Now, the recipe would be to apply the change of variable formula to the expression on the right and then to identify the martingale component that is predictable (and of bounded variation) and hence constant. Accomplishing this without caring about justification, would lead to the first order partial differential equations

$$\begin{aligned}&\frac{\partial}{\partial t}P_j(t,u) + \frac{\partial}{\partial u}P_j(t,u)(\delta_j u - b_j(t)) \\ &\quad + \sum_{k;k\neq j} \mu_{jk}(t)\left(P_k(t, u - b_{jk}(t)) - P_j(t,u)\right) = 0,\end{aligned}$$

valid between jumps of the contractual annuity functions B_j, and

$$P_j(t-,u) = P_j(t, u - \Delta B_j(t)),$$

valid at jumps of the B_j, and subject to the condition that the $P_j(n,u)$ are 0 and 1 according as $u < 0$ or $u \geq 0$.

The approach requires that the functions P_j possess first order derivatives in both directions. As we have seen already in the introductory example of Section 2 they generally do not. This difficulty might possibly be circumvented by conditioning on the ultimate state at time n, but proving differentiability of the conditional probabilities would require additional assumptions and would not be straightforward.

These remarks serve to show that the method developed in Section 4 is not just one among several candidate approaches to the problem; it is the only mathematically sound solution we are able to offer. One general conclusion we can extract is that there is no single general technique for solving the bulk of problems of the kind considered here; the method will have to be designed for each individual problem at hand and will depend on the model assumptions and the functional of interest.

8.5 Applications

A. The Poisson distribution.

In continuance of the example in Paragraph 6B of Norberg (1994), consider the special case with two states, $\mathcal{J} = \{1, 2\}$, no interest, $\delta_1 = \delta_2 = 0$, and the only payments being an assurance of 1 payable upon each transition, $b_{12} = b_{21} = 1$.

Then, taking $n = 1$, the present value in (8.1) is just the number of transition in the time interval $(t, 1]$, $N_{12}(1) + N_{21}(1) - N_{12}(t) - N_{21}(t)$. Furthermore, take $\mu_{12} = \mu_{21} = \mu$, a constant ($> 0$). Then it is seen from the defining relation (8.4) that the functions P_j and Q_j are all the same. Denoting this function by P, we can work with (8.2), which becomes

$$P(t, u) = \int_t^1 e^{-\mu(s-t)} \mu P(s, u-1) ds + e^{-\mu(1-t)} 1[0 \leq u]. \tag{8.1}$$

Using that $P(t, u) = 0$ for $u < 0$, we readily obtain from (8.1) that $P(t, u) = e^{-\mu(1-t)}$ for $0 \leq u < 1$. Then, for $1 \leq u < 2$, it follows from (8.1) that $P(t, u) = \mu(1-t)e^{-\mu(1-t)} + e^{-\mu(1-t)}$. Proceeding by induction we obtain, for each $u \geq 0$, that

$$P(t, u) = \sum_{i=0}^{[u]} \frac{(\mu(1-t))^i}{i!} e^{-\mu(1-t)},$$

which is the Poisson distribution with parameter $(1-t)\mu$, of course.

As a check on the accuracy of the numerical method, we list the computed values of $P(0, u)$, $u = 0, 1, \ldots, 8$, for $t = 0$ and $\mu = 1$ together with the exact values of the Poisson probabilities (in parantheses): 0.3670 (0.3679), 0.7358 (0.7358), 0.9202 (0.9197), 0.9813 (0.9810), 0.9965 (0.9864), 0.9994 (0.9994), 0.9999 (0.9999), 0.9999 (0.9999), 1.0000 (1.0000). These results were obtained with $h = 1/200$, $h' = 1/100$, $a = -0.5$, and (truncating the infinite support) $b = 9.5$.

B. The term insurance policy.

To analyse the term insurance policy in Paragraph 2A, take $\mathcal{J} = \{1, 2\}$, $n = 30$, $\mu_{12}(t) = 0.0005 + 0.000075858 \cdot 10^{0.038(30+t)}$, $\delta_1 = \ln(1.045)$, $b_1 = -0.0042608$, $b_{12} = 1$, and all other intensities and payments null.

Again, as a check on the accuracy of the numerical method, we list the computed values of $P(0, u)$ together with exact values (in parantheses): $P(0, u) = 0$ for $u < -0.0705$ (0 for $u < -0.0709$), $P(0, u) = 0.8453$ for $u \in [-0.0705, 0.1965)$ (0.8452 for $u \in [-0.0709, 0.1961)$), $P(0, 0.2) = 0.8491$ (0.8490), $P(0, 0.4) = 0.9465$ (0.9467), $P(0, 0.6) = 0.9774$ (0.9776), $P(0, 0.8) = 0.9918$ (0.9919), $P(0, 1) = 1.0000$ (1.0000). These results are based on $h = 1/1000$, $h' = 1/2000$, $a = -0.1$, and $b = 1.1$.

C. A combined insurance policy.

In our final numerical example we consider what will be referred to as *the combined policy*, which is the same as the one in Paragraph B, but with a disability pension added. More specifically, 1 is payable upon death, an annuity with level intensity 0.5 is payable during disability, and premium is payable with level intensity in active state. The relevant model entities are $\mathcal{J} = \{1, 2, 3\}$,

$n = 30$, $b_{13} = b_{23} = 1$, $b_1 = -0.013108$ (net premium when the intensities are as specified below), $b_2 = 0.5$, and, adoping the standard Danish technical basis except for the recovery intensity, $\delta = \ln(1.045)$ (independent of state), and

$$\mu_{13}(t) = \mu_{23}(t) = 0.0005 + 0.000075858 \cdot 10^{0.038x},$$
$$\mu_{12}(t) = 0.0004 + 0.0000034674 \cdot 10^{0.06(30+t)},$$
$$\mu_{21}(t) = 0.005,$$

all other payments and intensities being null.

Figure 2 about here

Fig. 2: Probability distribution of the present value of the combined insurance policy in (a) state 1 and (b) state 2.

This example is a follow-up of Paragraphs 4B-C in Norberg (1994), where the first three moments of the present value are calculated for the combined policy.

D. Numerical evaluation of multiple integrals.
Numerical integration in higher dimensions is in general complicated, and there exists no technique held to be universally superior. The technique developed here can be used to evaluate integrals that, possibly after a reinterpretation, can be recognized as a probability related to a present value for a suitably specified policy. Just to illustrate the idea, suppose T_1 and T_2 are independent positive random variables with cumulative distribution functions F_1 and F_2 with densities f_1 and f_2, respectively, and that we seek $\mathrm{P}[(T_2 \wedge 1) - (T_1 \wedge 1) \leq u]$. It is realized that this probability is found as $P(0, u)$ for the policy with $\mathcal{J} = \{1, 2, 3, 4\}$, $\mu_{12}(t) = \mu_{34}(t) = f_1(t)/(1 - F_1(t))$, $\mu_{13}(t) = \mu_{24}(t) = f_2(t)/(1 - F_2(t))$, $b_1(t) = -1$, $b_2(t) = 1$, $n = 1$, and no interest. Countless examples of this kind can be constructed.

Chapter 9

Reserves

Prospective and retrospective reserves are defined as conditional expected values, given some information available at the time of consideration. Each specification of the information invoked gives rise to a corresponding pair of reserves. Relationships between reserves are established in the general set-up. For the prospective reserve the present definition conforms with, and generalizes, the traditional one. For the retrospective reserve it appears to be novel. Special attention is given to the continuous time Markov chain model frequently used in the context of life and pension insurance. Thiele's differential equation for the prospective reserve is shown to have a retrospective counterpart. It is pointed out that the prospective and retrospective differential equations have, respectively, the Kolmogorov backward and forward differential equations as special cases. Practical uses of the differential equations are demonstrated by examples.

9.1 Introduction

A. Sketch of the idea. The concept of prospective reserve is no matter of dispute in life insurance mathematics. It is defined as the conditional expected present value of future benefits less premiums on the policy, given its present state. A straightforward generalization is obtained by conditioning on some other piece of information, e.g. on the policy's staying in some subset of the state space. It is proposed here to define the retrospective reserve analogously as the conditional expected present value of past premiums less benefits.

B. An example: insurance of a single life. A person aged x buys a life insurance policy specifying that the sum assured, b, is payable immediately upon death before age $x+n$ and that premiums are to be contributed continuously with level intensity c throughout the insurance period. Let T_x denote the person's remaining life length after the policy is issued at time 0, say. Assume that the survival function ${}_tp_x = P\{T_x > t\}$ is of the form ${}_tp_x = e^{-\int_0^t \mu_{x+s}\, ds}$, with continuous force of mortality, μ. Finally, assume that interest is earned with a constant, nonrandom intensity δ so that $v = e^{-\delta}$ is the annual discount factor and $i = e^{\delta} - 1$ is the annual rate of interest ($1 + i = v^{-1}$ is the annual interest factor).

At any time $t \in [0, n]$ the policy is either in state 0 = "alive" or in state 1 = "dead". The prospective reserves in the two states, indicated by subscripts 0 and 1, are

$$V_0^+(t) = \int_t^n v^{\tau - t} \; {}_{\tau - t}p_{x+t} \left\{ b\mu_{x+\tau} - c \right\} d\tau \tag{9.1}$$

(by the usual heuristic argument, the sum of expected discounted benefits minus premiums in small time intervals $(\tau, \tau + d\tau)$, $0 < \tau < t$, and, of course,

$$V_1^+(t) = 0. \tag{9.2}$$

The statewise retrospective reserves as defined above are

$$V_0^-(t) = c \int_0^t (1+i)^{t-\tau} \, d\tau \tag{9.3}$$

(trivial) and

$$V_1^-(t) = \frac{1}{1 - \; {}_tp_x} \int_0^t (1+i)^{t-\tau} \{ c \, ({}_\tau p_x - \; {}_tp_x) - b \, {}_\tau p_x \, \mu_{x+\tau} \} \, d\tau \tag{9.4}$$

(use the same kind of argument as in (9.1) noting that, conditional on death within time t, the probability of survival to τ is $({}_\tau p_x - \; {}_tp_x)/(1 - \; {}_tp_x)$ and the probability of death in $(\tau, \tau + d\tau)$ is ${}_\tau p_x \mu_{x+\tau} d\tau/(1 - \; {}_tp_x)$, $0 < \tau < t$) .

The state at time t is $X(t) = 1[T_x \leq t]$, the "number of deaths" of the person within time t. This is the information on which the reserves in (9.1) – (9.4) are based.

Now, suppose the complete prehistory of the policy is currently recorded, so that it is known at any time if the person is alive or dead and, in the latter case, when he died. The information available at time t is the pair $(X(t), min(T_x, t))$. Denote the reserves correspondingly by a double subscript. The reserves in state 0 remain as above, $V_{0,t}^{\pm}(t) = V_0^{\pm}(t)$, and so does the prospective reserve in state 1, of course, $V_{1,T_x}^+(t) = V_1^+(t) = 0$. Only the retrospective reserve in state 1 is affected by the additional information on the exact time of death. It now becomes simply the value at time t of past premiums less the benefit payment,

$$V_{1,T_x}^-(t) = c \int_0^{T_x} (1+i)^{t-\tau} d\tau - b(1+i)^{t-T_x}. \tag{9.5}$$

The quantity in (9.1) is what traditionally is referred to as the prospective reserve. The notion of retrospective reserve launched here differs from the traditional one, which in the present example is

$$ {}^{-}V_0(t) = \frac{1}{{}_tp_x} \int_0^t (1+i)^{t-\tau}\, {}_\tau p_x\, (c - b\mu_{x+\tau})d\tau. \tag{9.6}$$

This quantity emerges from the principle of equivalence, which requires that benefits and premiums should balance in the mean at the outset:

$$\int_0^n v^\tau\, {}_\tau p_x (b\mu_{x+\tau} - c)d\tau = 0. \tag{9.7}$$

Splitting $\int_0^n$ into $\int_0^t + \int_t^n$ in (9.7) and substituting from (9.1) and (9.6), yields

$${}^{-}V_0(t) = V_0^+(t)\,. \tag{9.8}$$

Thus, the traditional concept of "retrospective reserve" is rather a retrospective formula for the prospective reserve, valid for b and c satisfying the equivalence principle. For general b and c the quantity ${}^{-}V_0(t)$ is not an expected value, and it has no probabilistic interpretation in the present model involving one single policy. In an extended (artificial) model, with m independent replicates of the policy issued at time 0, it may be interpreted as the almost sure limit of the total accumulated surplus per survivor by time t as m tends to infinity.

As compared with (9.8), the retrospective reserve introduced here is related to the prospective reserve under the equivalence principle by the identity

$${}_tp_x V_0^-(t) + (1 - {}_tp_x)V_1^-(t) = {}_tp_x V_0^+(t)\,. \tag{9.9}$$

C. Outline of the paper. In Section 2 the present notions of reserves are defined for quite general stochastic payment streams and discounting rules, and certain relationships between them are established. No particular reference to the insurance context is made at this stage. In Sections 3 and 4 the framework of the further discussions is presented: payments of the life annuity and life insurance types in a continuous time Markov chain model. A useful auxiliary result is that a Markov process behaves like a composition of mutually independent Markov processes in disjoint intervals when its values at the dividing points between the intervals are fixed. This together with standard results for Markov chains is used in Section 5 to investigate the properties of reserves in the Markov chain case. The prospective reserve, given the state at the time of consideration, is the traditional one, which satisfies the well-known generalized Thiele's differential equations (see e.g. Hoem, 1969a), here also referred to as the *prospective differential equations*. The statewise retrospective reserves turn out to satisfy a set of *retrospective differential equations*, different from the prospective ones. It is pointed out the differential equations for the reserves have the Kolmogorov differential equations for the transition probabilities as special cases. Surprisingly, maybe, it is the retrospective equations that generalize the Kolmogorov forward equations, while the prospective equations generalize the Kolmogorov backward equations. In Section 6 some more examples are supplied.

9.2 General definitions of reserves and statement of some relationships between them

A. Payment streams and their discounted values. First some basic definitions and results are quoted from Norberg (1990).

Consider a stream of payments commencing at time 0. It is defined by a finite-valued *payment function* A, which for each time $t \geq 0$ specifies the total amount $A(t)$ paid in $[0, t]$. Negative payments are allowed for; it is only required that A be of bounded variation in finite intervals and, by convention, right-continuous. This means that $A = B - C$, where B and C are non-negative, nondecreasing, finite-valued, and right-continuous functions representing the outgoes and incomes, respectively, of some business. In the context of insurance B represents benefits and C represents contributed premiums on an insurance policy (or a portfolio of policies). The payment function extends in a unique way to a payment measure on the Borel sets, which is also denoted by A.

Assuming that payments are valuated by a piecewise monotone and continuous discount function v, the present value of A at time t is

$$V(t, A) = \frac{1}{v(t)} \int_{[0,\infty)} v(\tau)\, dA(\tau) \tag{9.1}$$

(the sum of all payments in small intervals discounted at time 0, multiplied by the interest factor $1/v(t)$ for the interval from 0 to t). Just to obtain transparent formulas, it will be assumed throughout that v is of the form

$$v(t) = e^{-\int_0^t \delta}, \tag{9.2}$$

with piecewise continuous interest intensity δ. (The shorthand exemplified by $\int_0^t \delta = \int_0^t \delta(\tau)\, d\tau$ will be in frequent use throughout.)

B. Definitions of retrospective and prospective reserves. The restriction of A to a (measurable) time set $\mathcal{T}$ is the measure $A_{\mathcal{T}}$ counting only those A-payments that fall due in $\mathcal{T}$; $A_{\mathcal{T}}\{\mathcal{S}\} = A\{\mathcal{S} \cap \mathcal{T}\}$.

At any time $t \geq 0$ the payment stream splits into payments after time t and payments up to and including time t; $A = A_{(t,\infty)} + A_{[0,t]}$. The present value in (9.1) splits correspondingly into

$$V(t,A) = V^+(t,A) - V^-(t,A), \tag{9.3}$$

where

$$V^+(t,A) = V(t, A_{(t,\infty)}) = \frac{1}{v(t)} \int_{(t,\infty)} v(\tau)\, dA(\tau) \tag{9.4}$$

is the discounted value of future net outgoes (in the insurance context benefits less premiums), and

$$V^-(t,A) = V(t, -A_{[0,t]}) = \frac{1}{v(t)} \int_{[0,t]} v(\tau)\, d(-A)(\tau) \tag{9.5}$$

is the value, with accumulation of interest, of past net incomes (in the insurance context premiums less benefits). The quantity $V^-(t,A)$ is observable by time t and can suitable be called the *individual retrospective reserve* of the policy at time t. If the future development of (v,A) were known, then $V^+(t,A)$ would be the appropriate amount to set aside to cover future excess of benefits over premiums on the individual policy. However, if the future course of (v,A) is uncertain, it is not possible to provide $V^+(t,A)$ as a prospective reserve on an individual basis.

Assume now that A and, possibly, also v are stochastic processes on some probability space $(\Omega, \mathcal{F}, P)$. An operational definition of the prospective reserve must depend solely on information that is at hand at the moment when the reserve is to be provided. Let $\mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}$ be a family of sub-sigmaalgebras of $\mathcal{F}$, $\mathcal{F}_t$ representing some piece of information available at time t. The family $\mathbf{F}$ may be increasing, that is, $\mathcal{F}_s \subset \mathcal{F}_t,\ s < t$, but this is not required in general. Reserves are defined as conditional expected values, given the information provided by $\mathbf{F}$. At time t *the prospective* $\mathbf{F}$*-reserve* is

$$V^+_{\mathbf{F}}(t,A) = E_{\mathcal{F}_t} V^+(t,A), \tag{9.6}$$

and *the retrospective* $\mathbf{F}$*-reserve* is

$$V^-_{\mathbf{F}}(t,A) = E_{\mathcal{F}_t} V^-(t,A), \tag{9.7}$$

where the subscript on the expectation sign signifies conditioning. The prospective $\mathbf{F}$-reserve meets the operationality requirement formulated above as it is determined by the current information. Even though the retrospective individual reserve in (9.5) is observable by time t, it may be judged relevant to calculate retrospective reserves with respect to some more summary information $\mathbf{F}$. For a given realization of $\{v(\tau)\}_{0\leq\tau\leq t}$, $\mathcal{F}_t$ may be thought of as a classification of the policies, whereby all policies with the same characteristics as specified by $\mathcal{F}_t$ are grouped together. Forming the mean, conditional on $\mathcal{F}_t$, means averaging over all policies in the same group, roughly speaking.

The reserves are conditional means of the present values $V^{\pm}(t, A)$. Other features of the conditional distributions of these random variables may be of interest. In particular, as measures of variability, introduce the variances

$$V_{\mathbf{F}}^{\pm(2)}(t, A) = Var_{\mathcal{F}_t} V^{\pm}(t, A). \tag{9.8}$$

C. Relationships between reserves. When only one payment stream is considered, notation can be saved by dropping the symbol A from $V(t, A)$. Thus, abbreviations like $V(t)$ and $V_{\mathbf{F}}^{\pm}(t)$ will be frequently used in the sequel.

By (9.1) – (9.3), the value of A at time 0 is related to its value at any time $t \geq 0$ by

$$\begin{aligned} V(0) &= v(t)V(t) \\ &= v(t)\{V^{+}(t) - V^{-}(t)\}. \end{aligned}$$

Taking expectation gives

$$E\,V(0) = E\,\{v(t)V(t)\} \tag{9.9}$$

$$= E\,\{v(t)(V^{+}(t) - V^{-}(t))\}. \tag{9.10}$$

The *equivalence principle* of insurance states that premiums and benefits should balance on the average as seen at the outset, that is,

$$E\,V(0) = 0. \tag{9.11}$$

It does not imply $EV(t) = 0$ for $t > 0$ unless v is a deterministic function, confer (9.9). Taking iterated expectations in (9.10), the equivalence requirement can be cast as

$$E\,\{v(t)V_{\mathbf{F}}^{-}(t)\} = E\,\{v(t)V_{\mathbf{F}}^{+}(t)\} \tag{9.12}$$

if $v(t)$ is determined by $\mathcal{F}_t$, and

$$E\,V_{\mathbf{F}}^{-}(t) = E\,V_{\mathbf{F}}^{+}(t) \tag{9.13}$$

if v is deterministic. Relation (9.9) is a special case of (9.13).

Let $\mathbf{F}' = \{\mathcal{F}'_t\}_{t\geq 0}$ be some sub-sigmaalgebra representing more summary information than $\mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}$ in the sense that $\mathcal{F}'_t \subset \mathcal{F}_t,\ t \geq 0$. The rule of iterated expextations yields the following relationship between reserves on different levels of information:

$$V_{\mathbf{F}'}^{\pm}(t) = E_{\mathcal{F}_t'} V_{\mathbf{F}}^{\pm}(t). \tag{9.14}$$

By the general rule $Var X = E\, Var_Y X + Var\, E_Y X$, variances denoted as in (9.8) are related by

$$V_{\mathbf{F}'}^{\pm(2)}(t) = E_{\mathcal{F}_t'}\{V_{\mathbf{F}}^{\pm(2)}(t) + (V_{\mathbf{F}}^{\pm}(t))^2\} - (V_{\mathbf{F}'}^{\pm}(t))^2. \tag{9.15}$$

There exist also useful relationships between reserves at different times for a fixed source of information, $\mathbf{F}$. The discounted values of future net outgoes at two different points of time, $t < u$, are related by

$$v(t)V^+(t) = \int_{(t,u]} v(\tau)\, dA(\tau) + v(u)V^+(u).$$

Similarly, for $s < t$,

$$v(t)V^-(t) = v(s)V^-(s) + \int_{(s,t]} v(\tau)\, d(-A)(\tau).$$

Taking expectations in these two identities, yields a pair of basic relationships under the assumption that v is deterministic. If $\{A(\tau)\}_{\tau \geq u}$ depends stochastically only on $\mathcal{F}_u$ for given $\mathcal{F}_t$ and $\mathcal{F}_u$, then

$$v(t)V_{\mathbf{F}}^+(t) = \int_{(t,u]} v(\tau)\, d\, E_{\mathcal{F}_t} A(\tau) + v(u) E_{\mathcal{F}_t} V_{\mathbf{F}}^+(u). \tag{9.16}$$

Likewise, if $\{A(\tau)\}_{\tau \leq s}$ depends stochastically only on $\mathcal{F}_s$ for given $\mathcal{F}_s$ and $\mathcal{F}_t$, then

$$v(t)V_{\mathbf{F}}^-(t) = v(s) E_{\mathcal{F}_t} V_{\mathbf{F}}^-(s) + \int_{(s,t]} v(\tau)\, d(-E_{\mathcal{F}_t} A)(\tau). \tag{9.17}$$

D. Right-continuity of the reserve processes. As defined by (9.6) and (9.7) the reserves are right-continuous stochastic processes. They could alternatively be made left-continuous by letting the integrals in (9.4) and (9.5) extend over $[t, \infty)$ and $[0, t)$, respectively. This would be in keeping with tradition, but the right-continuous versions are chosen here since they fit into the general apparatus of stochastic integrals and differential equations and thus are the more convenient quantities to deal with in anticipated applications of the theory to complex models. Anyway, the right-continuous and the left-continuous versions differ only at points of time where non-null amounts fall due with positive probability.

9.3 Description of payment streams appearing in life and pension insurance

A. Specification of the insurance treaty terms. Having insurances of persons in mind, consider insurance treaties specifying terms of the following general form. There is a set $\mathcal{J} = \{0, \ldots, J\}$ of possible states of the policy. The policy is issued at time 0, say. At any time $t \geq 0$ it is in one and only one of the states in $\mathcal{J}$, commencing in state 0. Payments are of two kinds: general life annuities by which the amount $A_g^\circ(t) - A_g^\circ(s)$ is paid during a sojourn in state g throughout the time interval $(s, t]$, and general life insurances by which an amount $a_{gh}^\circ(t)$ is paid immediately upon a transition from state g to state h at time t. They comprise benefits, administration expenses of all kinds, and premiums (negative). The *contractual functions* A_g° and a_{gh}° are, respectively, a payment function and a finite-valued, right-continuous function.

B. The form of the payment function. Let $X(t)$ be the state of the policy at time t. The development of the policy is given by $\{X(t)\}_{t\geq 0}$. This process, regarded as a function from $[0, \infty)$ to $\mathcal{J}$, is assumed to be right-continuous, with a finite number of jumps in any finite time interval. Let N_{gh} be the process counting the transitions from state g to state h, that is, $N_{gh}(t) = \#\{\tau \in [0, t];\ X(\tau-) = g, X(\tau) = h\}$. The stream of net payments is of the form

$$A = \sum_g \{A_g + \sum_{h; h\neq g} A_{gh}\},$$

with

$$\begin{aligned} dA_g(t) &= 1[X(t) = g]\, dA_g^\circ(t), && (9.1)\\ dA_{gh}(t) &= a_{gh}^\circ(t)\, dN_{gh}(t). && (9.2) \end{aligned}$$

The behaviour of the retrospective and prospective reserves is now to be studied for some specifications of $\mathbf{F}$ that are of relevance in insurance. The model framework will be the traditional Markov chain, which yields lucid results.

9.4 The Markov chain model

A. Model assumptions and basic relationships. The process $\{X(t)\}_{t\geq 0}$ is assumed to be a continuous time Markov chain on the state space $\mathcal{J}$. The transition probabilities are denoted

$$p_{jk}(t, u) = P\{X(u) = k \mid X(t) = j\}.$$

The transition intensities

$$\mu_{jk}(t) = \lim_{u \downarrow t} \frac{p_{jk}(t,u)}{u-t}$$

are assumed to exist for all t and $j \neq k$. To simplify matters, the functions μ_{jk} are furthermore assumed to be piecewise continuous. The total transition intensity from state j is

$$\mu_j = \sum_{k;\, k \neq j} \mu_{jk}\,.$$

From the Chapman-Kolmogorov equations,

$$p_{jk}(t,u) = \sum_{g \in \mathcal{J}} p_{jg}(t,\tau)\, p_{gk}(\tau,u),$$

valid for $t \leq \tau \leq u$, one obtains *Kolmogorov's differential equations*, the *forward*,

$$\frac{\partial}{\partial t} p_{ij}(s,t) = \sum_{g;\, g \neq j} p_{ig}(s,t)\, \mu_{gj}(t) - p_{ij}(s,t)\mu_j(t), \tag{9.1}$$

and the *backward*,

$$\frac{\partial}{\partial t} p_{jk}(t,u) = \mu_j(t) p_{jk}(t,u) - \sum_{g;\, g \neq j} \mu_{jg}(t) p_{gk}(t,u). \tag{9.2}$$

Together with the initial conditions $p_{jk}(t,t) = \delta_{jk},\ j,k \in \mathcal{J}$, (the Kronecker delta) they determine the transition probabilities uniquely. The Kolmogorov equations are the major tools for constructing the transition probabilities from the intensities, which are the basic entities in the system; they are functions of one argument only and, being readily interpretable, they form the natural starting point for specification of the model.

The conditional probability of staying uninterruptedly in state j throughout the time interval $[t,u]$, given that $X(t) = j$, is (solve a simple differential equation)

$$p_{\overline{jj}}(t,u) = e^{-\int_t^u \mu_j}. \tag{9.3}$$

B. Moments of present values. Throughout the balance of the paper it will be assumed that the interest intensity is nonstochastic. Consider the general annuities and insurances defined by (9.1) and (9.2). The present values at time t of their contributions in some time interval $\mathcal{T}$ are

$$
\begin{aligned}
V(t, A_{g\mathcal{T}}) &= \frac{1}{v(t)} \int_{\mathcal{T}} v(\tau) 1[X(\tau) = g] dA_g^{\circ}(\tau), \\
V(t, A_{gh\mathcal{T}}) &= \frac{1}{v(t)} \int_{\mathcal{T}} v(\tau) a_{gh}^{\circ}(\tau)\, dN_{gh}(\tau).
\end{aligned}
$$

Here follows a list of formulas for the first and second order moments of such present values, conditional on $X(s)$, with $\mathcal{T} \subset [s, \infty)$. The mean values are

$$
E_{X(s)=i} V(t, A_{g\mathcal{T}}) = \frac{1}{v(t)} \int_{\mathcal{T}} v(\tau) p_{ig}(s, \tau)\, dA_g^{\circ}(\tau), \tag{9.4}
$$

$$
E_{X(s)=i} V(t, A_{gh\mathcal{T}}) = \frac{1}{v(t)} \int_{\mathcal{T}} v(\tau) a_{gh}^{\circ}(\tau) p_{ig}(s, \tau) \mu_{gh}(\tau)\, d\tau. \tag{9.5}
$$

The noncentral second order moments, for intervals $\mathcal{S}, \mathcal{T} \subset [s, \infty)$, are

$$
\begin{aligned}
&E_{X(s)=i}\{V(t, A_{e\mathcal{S}}) V(t, A_{g\mathcal{T}})\} \\
&\quad = \frac{1}{v^2(t)} \iint_{\mathcal{S}\times\mathcal{T}} v(\vartheta) v(\tau) \{1[\vartheta \leq \tau] p_{ie}(s, \vartheta) p_{eg}(\vartheta, \tau) \\
&\qquad + 1[\vartheta > \tau] p_{ig}(s, \tau) p_{ge}(\tau, \vartheta)\}\, dA_e^{\circ}(\vartheta)\, dA_g^{\circ}(\tau),
\end{aligned} \tag{9.6}
$$

$$
\begin{aligned}
&E_{X(s)=i}\{V(t,A_{ef\mathcal{S}})V(t,A_{gh\mathcal{T}})\}\\
&\quad=\frac{1}{v^2(t)}\,[\int\!\!\int_{\mathcal{S}\times\mathcal{T}} v(\vartheta)v(\tau)\{1[\vartheta<\tau]p_{ie}(s,\vartheta)p_{fg}(\vartheta,\tau)\\
&\qquad+1[\vartheta>\tau]p_{ig}(s,\tau)p_{he}(\tau,\vartheta)\}\mu_{ef}(\vartheta)\mu_{gh}(\tau)a^{\circ}_{ef}(\vartheta)a^{\circ}_{gh}(\tau)\,d\vartheta\,d\tau\\
&\qquad+\delta_{ef,gh}\int_{\mathcal{S}\cap\mathcal{T}} v^2(\tau)p_{ig}(s,\tau)\mu_{gh}(\tau)a^{\circ\,2}_{gh}(\tau)\,d\tau\,],
\end{aligned}
\tag{9.7}
$$

$$
\begin{aligned}
&E_{X(s)=i}\{V(t,A_{e\mathcal{S}})V(t,A_{gh\mathcal{T}})\}\\
&\quad=\frac{1}{v^2(t)}\int\!\!\int_{\mathcal{S}\times\mathcal{T}} v(\vartheta)v(\tau)\{1[\vartheta\le\tau]p_{ie}(s,\vartheta)p_{eg}(\vartheta,\tau)\\
&\qquad+1[\vartheta>\tau]p_{ig}(s,\tau)p_{he}(\tau,\vartheta)\}dA^{\circ}_{e}(\vartheta)\,\mu_{gh}(\tau)a^{\circ}_{gh}(\tau)\,d\tau.
\end{aligned}
\tag{9.8}
$$

Variances and covariances are now formed by the rule $Cov(X,Y)=E(XY)-EXEY$.

The formulas are easily established by a heuristic type of argument. To motivate e.g. (9.8), put

$$
\begin{aligned}
&E_{X(s)=i}\,\{\int_{\mathcal{S}} v(\vartheta)1[X(\vartheta)=e]\,dA^{\circ}_{e}(\vartheta)\int_{\mathcal{T}} v(\tau)a^{\circ}_{gh}(\tau)\,dN_{gh}(\tau)\,\}\\
&\quad=\int\!\!\int_{\mathcal{S}\times\mathcal{T}} v(\vartheta)v(\tau)E_{X(s)=i}\{1[X(\vartheta)=e]\,dN_{gh}(\tau)\}a^{\circ}_{gh}(\tau)\,dA^{\circ}_{e}(\vartheta),
\end{aligned}
$$

and calculate the expected value in the integrand for $\vartheta\le\tau$ and for $\vartheta>\tau$. Rigorous proofs will not be given here. Some of the results in (9.4) – (9.8) are proved for absolutely continuous A°_{g} by Hoem (1969a) and Hoem & Aalen (1978).

C. Conditional Markov chains. Viewing reserves as conditional expected values, one may be concerned with various sub-sigmaalgebras of the basic sigmaalgebra $\mathcal{F}$. They will typically be of the form $\mathcal{F}_{\mathcal{T}}=\sigma\{X(\tau);\ \tau\in\mathcal{T}\}$, the sigmaalgebra representing all information provided by the process X in the time set $\mathcal{T}$. Of particular interest are $\mathcal{F}_{[0,t]}$ and $\mathcal{F}_{\{t_0,\dots,t_q\}}$.

The Markov property means precisely that for $\mathcal{B}\in\mathcal{F}_{(t,\infty)}$,

$$
P\{\mathcal{B}\mid\mathcal{F}_{[0,t]}\}=P\{\mathcal{B}\mid\mathcal{F}_{\{t\}}\},
\tag{9.9}
$$

that is, the future development of the process depends on its past and present states only through the present. An equivalent formulation is that, for $\mathcal{A}\in\mathcal{F}_{(0,t)}$ and $\mathcal{B}\in\mathcal{F}_{(t,\infty)}$,

$$P\{\mathcal{A} \cap \mathcal{B} \mid \mathcal{F}_{\{t\}}\} = P\{\mathcal{A} \mid \mathcal{F}_{\{t\}}\}P\{\mathcal{B} \mid \mathcal{F}_{\{t\}}\}, \tag{9.10}$$

which says that for a fixed present state of the process its future and past are conditionally independent. The equivalence of (9.9) and (9.10) is easily established in the present situation with finite state space. By induction (9.10) extends to the following: for $0 = t_0 < t_1 < \ldots < t_q < t_{q+1} = \infty$ and $\mathcal{A}_p \in \mathcal{F}_{(t_{p-1},t_p)}$, $p = 1, \ldots, q+1$,

$$P\{\cap_{p=1}^{q+1}\mathcal{A}_p \mid \mathcal{F}_{\{t_0,\ldots,t_q\}}\} = \prod_{p=1}^{q+1} P\{\mathcal{A}_p \mid \mathcal{F}_{\{t_{p-1},t_p\}}\},$$

where $\mathcal{F}_{\{t_q,\infty\}} = \mathcal{F}_{\{t_q\}}$. Thus, conditional on $X(0), X(t_1), \ldots, X(t_q)$, the developments of the process in the intervals (t_{p-1}, t_p) are mutually independent. So, to study the conditional process, given its values at fixed points, interest can be focused on the behaviour of the segment $\{X(\tau)\}_{s<\tau<t}$ for fixed $X(s)$ and $X(t)$, say. For $s < \tau < \vartheta < t$ and $\mathcal{A} \in \mathcal{F}_{(s,\tau)}$,

$$\begin{aligned}
&P\{X(\vartheta) = h \mid \mathcal{A}, X(\tau) = g,\ X(s) = i,\ X(t) = j\} \\
&\quad = \frac{P\{X(s) = i,\ \mathcal{A},\ X(\tau) = g,\ X(\vartheta) = h,\ X(t) = j\}}{P\{X(s) = i,\ \mathcal{A},\ X(\tau) = g,\ X(t) = j\}} \\
&\quad = \frac{P\{X(s) = i,\ \mathcal{A},\ X(\tau) = g\}\ p_{gh}(\tau,\vartheta)p_{hj}(\vartheta,t)}{P\{X(s) = i,\ \mathcal{A},\ X(\tau) = g\}p_{gj}(\tau,t)}\,.
\end{aligned}$$

Thus, the conditional process is Markov with transition probabilities

$$p_{gh|ij}(\tau,\vartheta \mid s,t) = \frac{p_{gh}(\tau,\vartheta)p_{hj}(\vartheta,t)}{p_{gj}(\tau,t)} \tag{9.11}$$

and intensities

$$\mu_{gh|ij}(\tau \mid s,t) = \mu_{gh}(\tau)\,\frac{p_{hj}(\tau,t)}{p_{gj}(\tau,t)}, \tag{9.12}$$

both independent of s and i. As $\tau \uparrow t$, a point of continuity of all the intensities, the expression in (9.12) tends to 0 if $g = j$, to $\mu_{gh}(t)\mu_{hj}(t)/\mu_{gj}(t)$ if $g, h \neq j$, and to $+\infty$ if $h = j$, which reflects that the conditional process is forced to end up in state j at time t. The conditional process $\{X(\tau)\}_{0\leq\tau\leq t}$, given $X(t) = j$, was studied by Hoem (1969b) in connection with statistical analysis of selected samples in demography.

9.5 Reserves in the Markov chain model

A. Individual reserves. There are as many definitions of the reserves in (9.6) and (9.7) as there are possible specifications of **F**. Each **F** represents a way of classifying the policies, and the **F**-reserves at time t are classwise averages of discounted excess of benefits over premiums in the future or accumulated excess of premiums over benefits in the past.

The simplest concepts of reserves are the *individual reserves* obtained by conditioning on the full history of the individual policy, that is, take $\mathcal{F}_t = \mathcal{F}_{[0,t]}$. The individual retrospective reserve is simply the cash value at time t of the net gains so far, defined by (9.5) and exemplified by (9.5):

$$V_{\mathbf{F}}^{-}(t) = V^{-}(t). \tag{9.1}$$

By the Markov property, the individual prospective reserve is

$$V_{\mathbf{F}}^{+}(t) = E_{X(t)}V^{+}(t).$$

Thus, in the Markov case it suffices to study the prospective reserve for given current state.

B. Reserves in single states. Now, let $\mathbf{F} = \{\mathcal{F}_{\{t\}}\}_{t\geq 0}$ so that reserves at time t are formed by averaging over policies in each single state at that time. In short, write $V_j^{\pm}(t)$ for the reserves $V_{\mathbf{F}}^{\pm}(t)$ when $X(t) = j$. The prospective reserves are the same as in Paragraph A. The explicit formula is gathered from (9.4) – (9.5):

$$\begin{aligned} V_j^{+}(t) &= \frac{1}{v(t)} \int_{(t,\infty)} v(\tau) \sum_{g} p_{jg}(t,\tau)\{dA_g^{\circ}(\tau) \\ &\quad + \sum_{h;\,h\neq g} a_{gh}^{\circ}(\tau)\mu_{gh}(\tau)\,d\tau\}, \end{aligned} \tag{9.2}$$

defined on $[0,\infty)$ for $j = 0$ and on $(0,\infty)$ for $j \neq 0$.

The retrospective reserve in state j is obtained by combining the results in Paragraphs 4B – C. Conditional on $X(t) = j$ for $t > 0$ (and $X(0) = 0$), $\{X(\tau)\}_{0\leq\tau\leq t}$ is a Markov chain starting in state 0, with transition probabilities and intensities given by (9.11) and (9.12). Noting that

$$p_{0g|0j}(0,\tau \mid 0,t)\mu_{gh|0j}(\tau \mid 0,t) = \frac{p_{0g}(0,\tau)\mu_{gh}(\tau)p_{hj}(\tau,t)}{p_{0j}(0,t)},$$

the general formulas (9.4) and (9.5) give

$$
\begin{aligned}
V_j^-(t) &= -\frac{1}{v(t)p_{0j}(0,t)}\int_{[0,t]} v(\tau)\sum_g p_{0g}(0,\tau)\{dA_g^\circ(\tau)p_{gj}(\tau,t) \\
&\quad + \sum_{h;\,h\neq g} a_{gh}^\circ(\tau)\mu_{gh}(\tau)p_{hj}(\tau,t)\,d\tau\}. \qquad (9.3)
\end{aligned}
$$

This expression is valid if $p_{0j}(0,t) > 0$. In particular, $V_0^-(0) = -A_0^\circ(0)$, of course, whereas $V_j^-(0)$ is not defined for $j \neq 0$. Formula (9.3) is easy to interpret by direct heuristic reasoning: conditionally, given $X(0) = 0$ and $X(t) = j$, the probability of staying in state g at time τ is $p_{0g}(0,\tau)p_{gj}(\tau,t)/p_{0j}(0,t)$, and the probability of a transfer from state g to state h in the time interval $(\tau, \tau + d\tau)$ is $p_{0g}(0,\tau)\mu_{gh}(\tau)d\tau p_{hj}(\tau,t)/p_{0j}(0,t)$.

The conditional variances,

$$V_j^{\pm(2)}(t) = Var_{X(t)=j}V^\pm(t),$$

are composed from (9.6) – (9.8) upon inserting the relevant conditional transition probabilities and intensities.

Finally, note that the future and the past developments of the process are independent, hence uncorrelated, for fixed present $X(t) = j$.

C. More general reserves. Referring to Paragraph 2C, let $\mathbf{F}$ be as in Paragraph B and let $\mathbf{F}' = \{\mathcal{F}_t'\}_{t\geq 0}$ be the more summary information with $\mathcal{F}_t'$ generated by the events $X(t) \in \mathcal{J}_l,\ l = 1,\ldots,m$, where $\{\mathcal{J}_1,\ldots,\mathcal{J}_m\}$ is some partitioning of $\mathcal{J}$. Thus, $\mathbf{F}'$-reserves at time t are now conditional on $X(t) \in \mathcal{J}_l,\ l \in \{1,\ldots,m\}$. For any $\mathcal{K} \subset \mathcal{J}$ and $t \leq u$ put

$$p_{j\mathcal{K}}(t,u) = P\{X(u) \in \mathcal{K} \mid X(t) = j\} = \sum_{k\in\mathcal{K}} p_{jk}(t,u).$$

Applying (9.14) and (9.15), one obtains (with obvious notation)

$$V_{\mathcal{J}_l}^\pm(t) = \frac{1}{p_{0\mathcal{J}_l}(0,t)}\sum_{j\in\mathcal{J}_l} p_{0j}(0,t)V_j^\pm(t)$$

and

$$V_{\mathcal{J}_l}^{\pm(2)}(t) = \frac{1}{p_{0\mathcal{J}_l}(0,t)}\sum_{j\in\mathcal{J}_l} p_{0j}(0,t)\{V_j^{\pm(2)}(t) + V_j^\pm(t)^2\} - V_{\mathcal{J}_l}^\pm(t)^2,$$

hence $\mathbf{F}'$-reserves and $\mathbf{F}'$-variances can be composed from the corresponding single-state quantities.

A variety of reserves based on different choices of **F** can be analysed by the results in Section 4. For instance, taking $\mathcal{F}_t = \mathcal{F}_{\{t_1,t_2,\ldots\}\cap[0,t]}$ means that reserves are formed by averaging over policies that are known to have visited certain states at given epochs in the past. The reserves and variances are found for each of the intervals and, by conditional independence, simply added to give the required total reserve and variance.

D. Differential equations for prospective reserves in single states. In the present Markov chain set-up (9.16) specializes to

$$\begin{aligned} W_j^+(t) &= \int_{(t,u]} v(\tau) \sum_g p_{jg}(t,\tau)\{dA_g^\circ(\tau) + \sum_{h;h\neq g} a_{gh}^\circ(\tau)\mu_{gh}(\tau)d\tau\} \\ &\quad + \sum_g p_{jg}(t,u) W_g^+(u), \end{aligned} \tag{9.4}$$

where, for each $j \in \mathcal{J}$,

$$W_j^+(t) = v(t) V_j^+(t). \tag{9.5}$$

Consider the case where all payments are restricted to a finite interval $[0,n]$. Assume first that the functions A_g° are absolutely continuous between 0 and n, that is, $dA_g^\circ(t) = a_g^\circ(t)\,dt$, $t \in (0,n)$. Then all W_g^+ are continuous in (0,n). Assume furthermore that all the functions a_g°, a_{gh}°, μ_{gh}, and δ are continuous in $(0,n)$. Then, for $u = t + dt$ (9.4) becomes

$$\begin{aligned} W_j^+(t) &= v(t)\{a_j^\circ(t) + \sum_{g;\,g\neq j} \mu_{jg}(t) a_{jg}^\circ(t)\}\,dt \\ &\quad + (1 - \mu_j(t)\,dt) W_j^+(t+dt) + \sum_{g;\,g\neq j} \mu_{jg}(t) dt W_g^+(t+dt) + o(dt), \end{aligned}$$

where $o(dt)/dt \to 0$ as $dt \to 0$. Now, subtract $W_j^+(t+dt)$ and divide by dt on both sides, and let $dt \downarrow 0$ to obtain the *prospective differential equations*, which are the multistate generalization of Thiele's classical equation:

$$\begin{aligned} \frac{\partial}{\partial t} W_j^+(t) &= -\,v(t)\{a_j^\circ(t) + \sum_{g;\,g\neq j} \mu_{jg}(t) a_{jg}^\circ(t)\} \\ &\quad + \mu_j(t) W_j^+(t) - \sum_{g;\,g\neq j} \mu_{jg}(t) W_g^+(t). \end{aligned} \tag{9.6}$$

For fixed contractual functions the V_j^+ are uniquely determined by the equations (9.6) together with the conditions

$$V_j^+(n-) = A_j^\circ(n) - A_j^\circ(n-). \tag{9.7}$$

If the principle of equivalence (9.11) is invoked, then the additional condition

$$V_0^+(0) = -A_0^\circ(0) \tag{9.8}$$

must be imposed — a constraint on the contractual functions.

The procedure is to be modified only slightly if A_j° has jumps at points $t_{j1}, \ldots, t_{j,q_j-1}$ in $(0,n)$. Put $t_{j0} = 0$ and $t_{jq_j} = n$. The solution of (9.6) has to be obtained in each of the intervals $(t_{j,p-1}, t_{jp})$, $p = 1, \ldots, q_j$, and the piecewise solutions in adjacent intervals must be hooked together by the conditions

$$V_j^+(t_{jp}-) = V_j^+(t_{jp}) + A_j^\circ(t_{jp}) - A_j^\circ(t_{jp}-), \tag{9.9}$$

$p = 1, \ldots, q_j$, which comprise (9.7).

A comment is needed also on possible discontinuities of the functions a_{jg}°, μ_{jg}, and δ. At such points and at the points t_{gp} (if any) where A_g° jumps and W_g^+ is discontinuous for some $g \neq j$, the derivative $\frac{\partial}{\partial t} W_j^+$ does not exist. The left and right derivatives exist, however, and if the discontinuities of the kind mentioned are finite in number, they will cause no technical problem as they will not affect the integrations that have to be performed to obtain the solution of the differential equations.

Pursuing the previous remark, switch the term $\mu_j(t) W_j^+(t)$ appearing on the right of (9.6) over to the left, multiply by the integrating factor $e^{-\int_0^t \mu_j}$ to form the complete differential $\frac{\partial}{\partial t}\{e^{-\int_0^t \mu_j} W_j^+(t)\}$ on the left, and finally integrate over an interval (t,u) containing no jumps of A_j° to obtain an integral equation. Insert (9.5) and solve (recall (9.2) and (9.3))

$$\begin{aligned} V_j^+(t) &= \frac{1}{v(t)} \Big[\int_t^u v(\tau) p_{\overline{jj}}(t,\tau) \{ a_j^\circ(\tau) + \sum_{g;\, g \neq j} \mu_{jg}(\tau) (a_{jg}^\circ(\tau) + V_g^+(\tau)) \} d\tau \\ &\quad + v(u) p_{\overline{jj}}(t,u) V_j^+(u-) \Big]. \end{aligned} \tag{9.10}$$

This expression is easy to interpret: it decomposes the future payments into those that fall due before and those that fall due after time u or the time of the first transition out of the current state, whichever occurs first. By fixed contractual functions the integral equations in conjunction with the conditions (9.9) determine the V_j^+.

E. Differential equations for retrospective reserves in single states. The starting point is (9.17), which now specializes to

$$
\begin{aligned}
v(t)V_j^-(t) &= v(s)\sum_g \frac{p_{0g}(0,s)p_{gj}(s,t)}{p_{0j}(0,t)}V_g^-(s) \\
&\quad -\int_{(s,t]} v(\tau)\sum_g \frac{p_{0g}(0,\tau)}{p_{0j}(0,t)}\{dA_g^\circ(\tau)p_{gj}(\tau,t) \\
&\quad +\sum_{h;\,h\neq g} a_{gh}^\circ(\tau)\mu_{gh}(\tau)p_{hj}(\tau,t)\,d\tau\}. \qquad (9.11)
\end{aligned}
$$

Upon multiplying by $p_{0j}(0,t)$ and introducing

$$
W_j^-(t) = v(t)p_{0j}(0,t)V_j^-(t), \qquad (9.12)
$$

$j \in \mathcal{J}$, (9.11) can be reshaped as

$$
\begin{aligned}
W_j^-(t) &= \sum_g W_g^-(s)p_{gj}(s,t) \\
&\quad -\int_{(s,t]} v(\tau)\sum_g p_{0g}(0,\tau)\{dA_g^\circ(\tau)p_{gj}(\tau,t) \\
&\quad +\sum_{h;\,h\neq g} a_{gh}^\circ(\tau)\mu_{gh}(\tau)p_{hj}(\tau,t)d\tau\}. \qquad (9.13)
\end{aligned}
$$

Again, consider the case where all payments are restricted to a finite interval $[0,n]$, with $dA_g^\circ(t) = a_g^\circ(t)\,dt$, $t \in (0,n)$, and a_g°, a_{gh}°, μ_{gh}, and δ all continuous in $(0,n)$. With t and $t+dt$ in the places of s and t, (9.13) becomes

$$
\begin{aligned}
W_j^-(t+dt) &= \sum_{g;g\neq j} W_g^-(t)\mu_{gj}(t)dt + W_j^-(t)(1-\mu_j(t)dt) \\
&\quad -v(t)\sum_g p_{0g}(0,t)\{a_g^\circ(t)dt\,p_{gj}(t,t+dt) \\
&\quad +\sum_{h;\,h\neq g} a_{gh}^\circ(t)\mu_{gh}(t)p_{hj}(t,t+dt)dt\} + o(dt).
\end{aligned}
$$

As $p_{ij}(t,t+dt)$ is $\mu_{ij}(t)dt + o(dt)$ for $i \neq j$ and $1-\mu_j(t)dt + o(dt)$ for $i=j$, all terms $p_{ij}(t,t+dt)$ on the right can be replaced by δ_{ij} (the difference is absorbed in $o(dt)$). Then proceed in the same manner as in the previous paragraph to obtain the *retrospective differential equations*

$$
\begin{aligned}
\frac{\partial}{\partial t}W_j^-(t) &= \sum_{g;g\neq j} W_g^-(t)\mu_{gj}(t) - W_j^-(t)\mu_j(t) \\
&\quad -v(t)\{p_{0j}(0,t)a_j^\circ(t) + \sum_{g;\,g\neq j} p_{0g}(0,t)a_{gj}^\circ(t)\mu_{gj}(t)\}. \qquad (9.14)
\end{aligned}
$$

For fixed contractual functions the retrospective reserves are determined uniquely by (9.14) together with the conditions

$$W_j^-(0) = -\delta_{0j} A_0^\circ(0). \tag{9.15}$$

The equivalence condition (9.11) can be cast in terms of the retrospective reserves as

$$\sum_{j \in \mathcal{J}} p_{0j}(0,n) V_j^-(n) = 0. \tag{9.16}$$

Possible discontinuities are accounted for in the same manner as for the prospective reserves, the retrospective counterpart of (9.9) being

$$V_j^-(t_{jp}) = V_j^-(t_{jp}-) - (A_j^\circ(t_{jp}) - A_j^\circ(t_{jp}-)), \tag{9.17}$$

$p = 1, \ldots, q_j$.

Copying essentially the steps leading to (9.10), one obtains from (9.14) the integral equation

$$\begin{aligned} V_j^-(t-) \;=\; & \frac{1}{v(t) p_{0j}(0,t)} \,[\, v(s) p_{0j}(0,s) p_{\overline{jj}}(s,t) V_j^-(s) \\ & - \int_s^t v(\tau) \{ p_{0j}(0,\tau) p_{\overline{jj}}(\tau,t) a_j^\circ(\tau) \\ & + \sum_{g;\, g \neq j} p_{0g}(0,\tau) \mu_{gj}(\tau) p_{\overline{jj}}(\tau,t) (a_{gj}^\circ(\tau) - V_g^+(\tau)) \} d\tau \,], \end{aligned} \tag{9.18}$$

valid in intervals (s,t) containing no jumps of A_j°. It decomposes the past payments into those that fall due before and those that fall due after time s or the time of the last transition into the current state, whichever is the latter.

F. The prospective and retrospective differential equations are generalizations of Kolmogorov's backward and forward differential equations, respectively. This result is established by considering a simple endowment in the special case where $v(t) \equiv 1$ (no interest). First, assume that the only payment provided is a unit benefit payable at time u contingent upon $X(u) = k$, that is, $A_k^\circ = \varepsilon_u$, the measure with a unit mass at u, and all other A_g° and all a_{gh}° are null. Then, for $t < u$, $W_j^+(t)$ in (9.5) reduces to $p_{jk}(t,u)$, and the prospective differential equations in (9.6) specialize to the Kolmogorov backward equations in (9.2). Second, the Kolmogorov forward differential equations in (9.1) are obtained as a specialization of the retrospective differential equations in (9.14) by letting the only payment be a unit premium at time s, contingent upon $X(s) = i$, whereby $W_j^-(t)$ in (9.12) reduces to $p_{0i}(0,s)p_{ij}(s,t)$ for $t > s$.

Likewise, integral equations for the transition probabilities come out as special cases. From (9.10) one gets

$$p_{jk}(t,u) = \int_t^u \; p_{\overline{jj}}(t,\tau) \sum_{g;\, g \neq j} \mu_{jg}(\tau) p_{gk}(\tau,u) d\tau \; + \; \delta_{jk} p_{\overline{jj}}(t,u), \qquad (9.19)$$

and from (9.18)

$$p_{ij}(s,t) = \delta_{ij} p_{\overline{ii}}(s,t) + \sum_{g;\, g \neq j} \int_s^t \; p_{ig}(s,\tau) \mu_{gj}(\tau) p_{\overline{jj}}(\tau,t) d\tau. \qquad (9.20)$$

These equations are easy to interpret.

In a similar manner one may also find differential and integral equations for expected sojourn times. For $v(t) \equiv 1$, $dA_k^\circ(t) = dt$, and all other contractual functions null, $V_j^+(t)$ and $V_j^-(t)$ are just the expected future and past total sojourn times in state k, conditional on $X(t) = j$.

G. Uses of the differential equations. In the case where the contractual functions do not depend on the reserves, the defining relations (9.2) and (9.3) are explicit expressions for the reserves. The differential equations (9.6) and (9.14) are not needed for constructive purposes — they serve only 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 $\mathcal{J}$) 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 indispensable tools in the construction of the reserves and determinaton of the equivalence premium. The simplest case is when the payments are linear functions of the reserve, with coefficients possibly depending on time. Then the operations leading to (9.10) and (9.18) can basically be reproduced after collecting all terms involving the reserve in state j on the left hand side and multiplying by the appropriate integrating factor. Such techniques are standard and frequently used for the traditional prospective reserve. They carry over to the retrospective reserve, as will be illustrated by examples in the final section.

Apart from some very simple situations, like the one encountered in Paragraph 1B, the computation of the reserves will usually require numerical solution of the set of differential equations (or the equivalent integral equations). There is, however, an important class of situations which allow for more direct computation by iterated numerical integrations, and a comment shall be rendered on those. When returns to formerly visited states are impossible, $\mu_{jg} = 0$ for

$g < j$, say, the equations (9.10) form the basis of a recursive computational procedure. The sum over g on the right of (9.10) extends only over $g > j$ (void if $j = J$), and so one can suitably start by determining V_J^+, which is easy since the equation for V_J^+ involves no other V_g^+ (typically the policy is no longer in force in state J and V_J^+ is identically 0). Then proceed downwards through the state space: having determined $V_{j+1}^+, \ldots, V_J^+$, use (9.10) with $u = t_{jp}$ and (9.9) in each interval $(t_{j,p-1}, t_{jp})$, starting from time $t_{jq_j} = n$. Similarly, the equations (9.18) are solved recursively starting with the equation for V_0^-, which involves no other V_g^-: having determined $V_0^+, \ldots, V_{j-1}^+$, use (9.18) with $s = t_{j,p-1}$ and (9.17) in each interval $(t_{j,p-1}, t_{jp})$, starting from time $t_{j0} = 0$. Again, the examples in the following section are referred to.

H. Behaviour of the retrospective reserves in the vicinity of 0. In establishing the retrospective differential equations, the auxiliary functions W_j^- defined in (9.12) are more convenient to work with than the reserves themselves. The V_j^- are defined only for those t where $p_{0j}(0,t) > 0$, whereas the W_j^- are well defined in all of $[0, \infty)$. The W_j^- are right-continuous, and their values in 0, given by (9.15), are simple initial conditions for the differential equations. The reserves V_j^- are also right-continuous, but in 0 only V_0^- is defined $(-A_0^\circ(0))$. For $j \neq 0$ the limit $V_j^-(0+)$ must be obtained by letting $t \downarrow 0$ in (9.3). Leaving details aside, only the intuitively appealing result is reported. A state $j \neq 0$ is said to be immediately accessible from state 0 in q steps via the directed path $\mathbf{g} = (g_1 g_2 \ldots g_{q-1})$ if $\mu_{\mathbf{g}}(0) = \mu_{0g_1}(0)\mu_{g_1 g_2}(0) \ldots \mu_{g_{q-1}j}(0) > 0$. Let q_j be the minimum of all q for which this property holds for some path, and denote the set of such minimal paths by $\mathcal{P}_j$. Then

$$V_j^-(0+) = -A_0^\circ(0) - \frac{\sum_{\mathbf{g} \in \mathcal{P}_j} \mu_{\mathbf{g}}(0)\{a_{0g_1}^\circ(0) + a_{g_1 g_2}^\circ(0) + \ldots + a_{g_{q_j-1}j}^\circ(0)\}}{\sum_{\mathbf{g} \in \mathcal{P}_j} \mu_{\mathbf{g}}(0)}.$$

I. A different notion of retrospective reserve. In a recent paper Wolthuis & Hoem (1990) have launched a notion of retrospective reserve quite different from the one introduced here. Working in the Markov chain model, they require that the statewise retrospective reserves should satisfy

$$E_{X(0)=i} V(0) = v(t) \sum_{j \in \mathcal{J}} p_{ij}(0,t)\{V_j^+(t) - V_j^-(t)\} \tag{9.21}$$

for $i = 0$, which conforms with (9.10). Then, imagining that the policy might start from any state different from 0, they require that (9.21) be valid for all i, whereby some hypothetical values must be chosen for the $E_{X(0)=i}V(0)$, $i \neq 0$. There exists no such choice that can produce the retrospective reserves (9.3), and so the approach is incompatible with the one taken here. The same is the case for the approach proposed by Hoem (1988), where (9.21) is required for $i = 0$ and $V_j^-(t)$ is put equal to $V_j^+(t)$ for $j \neq 0$.

9.6 Some examples

A. Life insurance of a single life (continued from Paragraph 1B). In this simple situation the prospective differential equations (9.6) are

$$\frac{\partial}{\partial t}W_0^+(t) = v^t\,(c - \mu_{x+t}b) + \mu_{x+t}W_0^+(t) - \mu_{x+t}W_1^+(t), \tag{9.1}$$

$$\frac{\partial}{\partial t}W_1^+(t) = 0. \tag{9.2}$$

With the conditions

$$W_0^+(n-) = W_1^+(n-) = 0 \tag{9.3}$$

they lead to (9.1) and (9.2), which could be put up by direct prospective reasoning. The equivalence premium is

$$c = b\,\frac{\int_0^n v^\tau\ {}_\tau p_x \mu_{x+\tau} d\tau}{\int_0^n v^\tau\ {}_\tau p_x d\tau}.$$

Suppose the treaty is modified so that the prospective reserve is paid out as an additional benefit upon death before time n; $a_{01}^\circ(t) = b + V_0^+(t)$. Then the reserve and the equivalence premium cannot be determined directly by prospective reasoning, and it is necessary to employ the differential equations (9.6). Now, instead of (9.1) one gets

$$\begin{aligned}\frac{\partial}{\partial t}W_0^+(t) &= v^t\,\{c - \mu_{x+t}(b + V_0^+(t))\} + \mu_{x+t}W_0^+(t) - \mu_{x+t}W_1^+(t)\\ &= v^t\,\{c - \mu_{x+t}b\} - \mu_{x+t}W_1^+(t).\end{aligned} \tag{9.4}$$

Equation (F.38) and the conditions (F.39) remain unchanged. One finds $V_1^+(t) = 0$ as before, and

$$V_0^+(t) = \int_t^n v^{\tau - t}(\mu_{x+\tau}b - c)\,d\tau. \tag{9.5}$$

The equivalence premium is determined upon inserting $t = 0$ in (9.5) and equating to 0:

$$c = b\,\frac{\int_0^n v^\tau \mu_{x+\tau} d\tau}{\int_0^n v^\tau d\tau}. \tag{9.6}$$

These techniques and results are classical, and are referred here for the sake of comparison with what now follows.

For the standard contract, with no repayment of the reserve, the differential equations (9.14) become

$$\begin{aligned}\frac{\partial}{\partial t}W_0^-(t) &= -W_0^-(t)\mu_{x+t} + v^t\,{}_tp_x c, &(9.7)\\ \frac{\partial}{\partial t}W_1^-(t) &= W_0^-(t)\mu_{x+t} - v^t\,{}_tp_x\mu_{x+t}b. &(9.8)\end{aligned}$$

With the conditions

$$W_0^-(0) = W_1^-(0) = 0 \tag{9.9}$$

they lead to (9.3) and (9.4), which could be put up directly.

Suppose now that the retrospective reserve is to be paid out as an additional benefit upon death; $a_{01}^{\circ}(t) = b + V_0^-(t)$. Then the retrospective differential equations must be employed. Instead of (9.8) one gets

$$\begin{aligned}\frac{\partial}{\partial t}W_1^-(t) &= W_0^-(t)\mu_{x+t} - v^t\,{}_tp_x\mu_{x+t}(b + V_0^-(t))\\ &= -v^t\,{}_tp_x\mu_{x+t}b,\end{aligned}$$

whereas equation (9.7) and the conditions (9.9) remain unchanged. One arrives at the same expression for V_0^- as before, of course, and

$$V_1^-(t) = -\frac{b}{1-{}_tp_x}\int_0^t (1+i)^{t-\tau}\,{}_\tau p_x\mu_{x+\tau}d\tau.$$

The equivalence premium c is determined by (9.16):

$$c = b\,\frac{\int_0^n v^\tau\,{}_\tau p_x\mu_{x+\tau}d\tau}{{}_np_x\int_0^n v^\tau d\tau}.$$

B. Widow's pension. A married couple buys a widow's pension policy specifying that premiums are to be paid with intensity c as long as both husband and wife are alive, and pensions are to be paid with intensity b as long as the wife is widowed. The policy terminates at time n or upon the death of the wife, whichever occurs first. The relevant Markov model is sketched in Fig. reffig:two-lives. Expressions for the transition probabilities are easily obtained by direct reasoning or by use of (9.19). The reserves can be picked directly from (9.2) and (9.3):

$$\begin{aligned}V_0^+(t) &= \frac{1}{v(t)}\int_t^n v(\tau)\{b\,p_{01}(t,\tau) - c\,p_{00}(t,\tau)\}d\tau,\\ V_1^+(t) &= \frac{b}{v(t)}\int_t^n v(\tau)p_{11}(t,\tau)d\tau,\\ V_2^+(t) &= V_3^+(t) = 0,\end{aligned}$$

$$
\begin{aligned}
V_0^-(t) &= \frac{c}{v(t)}\int_0^t v(\tau)\,d\tau, &(9.10)\\
V_1^-(t) &= \frac{1}{v(t)\,p_{01}(0,t)}\int_0^t v(\tau)\,\{c\,p_{00}(0,\tau)p_{01}(\tau,t)\\
&\quad -b\,p_{01}(0,\tau)p_{11}(\tau,t)\}d\tau, &(9.11)\\
V_2^-(t) &= \frac{c}{v(t)\,p_{02}(0,t)}\int_0^t v(\tau)\,p_{00}(0,\tau)p_{02}(\tau,t)\,d\tau, &(9.12)\\
V_3^-(t) &= \frac{1}{v(t)\,p_{03}(0,t)}\int_0^t v(\tau)\,\{c\,p_{00}(0,\tau)p_{03}(\tau,t)\\
&\quad -b\,p_{01}(0,\tau)p_{13}(\tau,t)\}d\tau. &(9.13)
\end{aligned}
$$

The differential equations are not needed to construct these formulas. The retrospective ones are listed for ease of reference:

$$
\begin{aligned}
\frac{\partial}{\partial t}W_0^-(t) &= -W_0^-(t)(\mu_{01}(t)+\mu_{02}(t))+v(t)p_{00}(0,t)\,c, &(9.14)\\
\frac{\partial}{\partial t}W_1^-(t) &= W_0^-(t)\mu_{01}(t)-W_1^-(t)\mu_{13}(t)-v(t)p_{01}(0,t)\,b, &(9.15)\\
\frac{\partial}{\partial t}W_2^-(t) &= W_0^-(t)\mu_{02}(t)-W_2^-(t)\mu_{23}(t), &(9.16)\\
\frac{\partial}{\partial t}W_3^-(t) &= W_1^-(t)\mu_{13}(t)+W_2^-(t)\mu_{23}(t). &(9.17)
\end{aligned}
$$

Consider a modified policy, by which the retrospective reserve is to be paid back to the husband in case he is widowered before time n, the philosophy being that couples receiving no pensions should have their savings back. Now the retrospective differential equations are needed. The equations above remain unchanged except that the term $v(t)p_{00}(0,t)V_0^-(t)\mu_{02}(t) = W_0^-(t)\mu_{02}(t)$ must be subtracted on the right of (9.16), which then changes to

$$
\frac{\partial}{\partial t}W_2^-(t) = -W_2^-(t)\mu_{23}(t). \tag{9.18}
$$

Together with the conditions $W_j^-(0) = 0$, $j = 0,1,2,3$, these equations are easily solved. Obviously, the expressions for $V_0^-(t)$ and $V_1^-(t)$ remain the same as in (9.10) and (9.11). From (9.18) follows $V_2^-(t) = 0$, which is also obvious. Finally, (9.17) gives

$$
\begin{aligned}
V_3^-(t) &= \frac{1}{v(t)\,p_{03}(0,t)}\int_0^t v(\tau)\,\{c\,p_{00}(0,\tau)p_{013}(\tau,-,t)\\
&\quad -b\,p_{01}(0,\tau)p_{13}(\tau,t)\}\,d\tau, &(9.19)
\end{aligned}
$$

where

$$p_{013}(\tau, -, t) = \int_{\tau}^{t} p_{01}(\tau, \vartheta)\mu_{13}(\vartheta)d\vartheta$$

is the probability of passing from state 0 to state 3 via state 1 in the time interval $[\tau, t]$, given that $X(\tau) = 0$.

As a final example 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 as follows. At time t expenses fall due with intensity $e'_0(t) + e''_0(t)c$ in state 0 (the latter term represents encashment commision) and with intensity $e'_1(t)$ in state 1. In addition, expenses related to maintenance of the reserve fall due with intensity $e(t)\times$(current retrospective reserve) throughout the entire period $[0, n]$.

Instead of (9.14) – (9.17) one now gets

$$\begin{aligned}
\frac{\partial}{\partial t}W_0^-(t) &= -W_0^-(t)(\mu_{01}(t)+\mu_{02}(t)) \\
&\quad - v(t)p_{00}(0,t)\{e_0'(t)+e_0''(t)c+e(t)V_0^-(t)-c\}, && (9.20)\\
\frac{\partial}{\partial t}W_1^-(t) &= W_0^-(t)\mu_{01}(t)-W_1^-(t)\mu_{13}(t) \\
&\quad - v(t)p_{01}(0,t)\{b+e_1'(t)+e(t)V_1^-(t)\}, && (9.21)\\
\frac{\partial}{\partial t}W_2^-(t) &= W_0^-(t)\mu_{02}(t)-W_2^-(t)\mu_{23}(t) \\
&\quad -v(t)p_{02}(0,t)e(t)V_2^-(t), && (9.22)\\
\frac{\partial}{\partial t}W_3^-(t) &= W_1^-(t)\mu_{13}(t)+W_2^-(t)\mu_{23}(t) \\
&\quad -v(t)p_{03}(0,t)e(t)V_3^-(t), && (9.23)
\end{aligned}$$

and the conditions (9.15) become $W_j^-(0)=0,\ j=0,1,2,3$. Now a small trick. In each of the equations (9.20) – (9.23), say the one for $\frac{\partial}{\partial t}W_j^-(t)$, there appears a term $-v(t)p_{0j}(0,t)e(t)V_j^-(t) = -e(t)W_j^-(t)$ on the right hand side. Switch this over to the left and multiply on both sides by $e^{\int_0^t e}$. Form a complete differential $\frac{\partial}{\partial t}\{e^{\int_0^t e}W_j^-(t)\}$ on the left hand side and absorb everywhere the factor $e^{\int_0^t e}$ into $v(t)=e^{-\int_0^t \delta}$ (remember $v(t)$ is a factor in $W_j^-(t)$). What remains are retrospective differential equations for the the same situation as the original one, modified to the effect that the administration costs related to the reserve have vanished and the interest intensity δ has been decreased by e. Thus, one can apply the explicit expression (9.3) for this modified case, and arrive at the following formulas, where $v^*(t)=e^{-\int_0^t(\delta-e)}$:

$$\begin{aligned}
V_0^-(t) &= \frac{1}{v^*(t)}\int_0^t v^*(\tau)\{c(1-e_0''(\tau))-e_0'(\tau)\}\,d\tau,\\
V_1^-(t) &= \frac{1}{v^*(t)\,p_{01}(0,t)}\int_0^t v^*(\tau)\,[\,p_{00}(0,\tau)p_{01}(\tau,t)\{c(1-e_0''(\tau))-e_0'(\tau)\}\\
&\quad -p_{01}(0,\tau)p_{11}(\tau,t)(b+e_1'(\tau))\,]\,d\tau,\\
V_2^-(t) &= \frac{1}{v^*(t)\,p_{02}(0,t)}\int_0^t v^*(\tau)p_{00}(0,\tau)p_{02}(\tau,t)\{c(1-e_0''(\tau))-e_0'(\tau)\}d\tau,\\
V_3^-(t) &= \frac{1}{v^*(t)\,p_{03}(0,t)}\int_0^t v^*(\tau)\,[\,p_{00}(0,\tau)p_{03}(\tau,t)\{c(1-e_0''(\tau))-e_0'(\tau)\}\\
&\quad -p_{01}(0,\tau)p_{13}(\tau,t)(b+e_1'(\tau))\,]\,d\tau.
\end{aligned}$$

Using the equivalence principle in the form (9.16), one obtains the equivalence premium

$$c=\frac{\int_0^n v^*(\tau)\{p_{00}(0,\tau)e_0'(\tau)+p_{01}(0,\tau)(b+e_1'(\tau))\}\,d\tau}{\int_0^n v^*(\tau)p_{00}(0,\tau)(1-e_0''(\tau))\,d\tau}. \tag{9.24}$$

If the cost element $e(t)V_j^-(t)$ is replaced by $e(t)V_j^+(t)$, the prospective differential equations must be used. The procedure above can essentially be repeated, and just as for the retrospective reserves it turns out that the prospective reserves are those correspending to $e(t) = 0$ and discount function v^*. The equivalence premium remains as in (9.24).

Chapter 10

Safety loadings and bonus

10.1 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 valuated, 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 10.7 below.

10.2 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_j^Z and N_{jk}^Z. 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 μ_{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 μ_{jk} become stochastic processes. Let $\mathcal{G}_t$ denote their complete history up to, and including, time t and, accordingly, let $\mathbb{E}[\,\cdot\,|\,\mathcal{G}_t]$ 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 μ_{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

$$\begin{aligned} V_j^*(t) &= \mathbb{E}^*\left[\int_t^n e^{-\int_t^\tau r^*} dB(\tau) \,\middle|\, Z(t)=j\right] \\ &= \int_t^n e^{-\int_t^\tau r^*} \sum_g p_{jg}^*(t,\tau)\left(dB_g(\tau)+\sum_{h;h\neq g} b_{gh}(\tau)\mu_{gh}^*(\tau)\,d\tau\right). \end{aligned} \tag{10.1}$$

We need Thiele's differential equations

$$dV_j^*(t) = r^*(t)V_j^*(t)\,dt - dB_j(t) - \sum_{k;\,k\neq j} R_{jk}^*(t)\,\mu_{jk}^*(t)\,dt\,, \tag{10.2}$$

where

$$R_{jk}^*(t) = b_{jk}(t) + V_k^*(t) - V_j^*(t) \tag{10.3}$$

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,

$$\mathbb{E}^*\left[\int_{0-}^n e^{-\int_0^\tau r^*}\,dB(\tau)\right] = 0\,, \tag{10.4}$$

or, equivalently,

$$V_0^*(0) = -\Delta B_0(0)\,. \tag{10.5}$$

10.3 The technical surplus and how it emerges

A. Definition of the mean portfolio surplus. With premiums determined by the principle of equivalence (10.4) 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

$$\begin{aligned} S(t) &= \mathbb{E}\left[\int_{0-}^t e^{\int_\tau^t r}\,d(-B)(\tau)\,\middle|\,\mathcal{G}_t\right] - \sum_j p_{0j}(0,t)\,V_j^*(t) \\ &= -e^{\int_0^t r}\int_{0-}^t e^{-\int_0^\tau r}\sum_j p_{0j}(0,\tau)\left(dB_j(\tau)+\sum_{k;k\neq j} b_{jk}(\tau)\mu_{jk}(\tau)\,d\tau\right) \\ &\quad - \sum_j p_{0j}(0,t)\,V_j^*(t)\,. \end{aligned} \tag{10.6}$$

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 (10.5),

$$S(0) = 0 \tag{10.7}$$

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

$$S(n) = \mathbb{E}\left[\int_{0-}^{n} e^{\int_\tau^n r}\, d(-B)(\tau) \,\middle|\, \mathcal{G}_n \right] , \tag{10.8}$$

as it ought to be.

Note that the expression in (10.6) involves only the past history of the second order basis, which is currently known.

B. The contributions to the surplus. Differentiating (10.6), applying the Kolmogorov forward equation (7.20) and the Thiele backward equation (10.2) to the last term on the right, leads to

$$\begin{aligned} dS(t) \quad = \quad & -e^{\int_0^t r}\, r(t)\, dt \int_{0-}^{t} e^{-\int_0^\tau r} \sum_j p_{0j}(0,\tau) \left(dB_j(\tau) + \sum_{k;k\neq j} b_{jk}(\tau)\mu_{jk}(\tau)\, d\tau \right) \\ & - \sum_j p_{0j}(0,t) \left(dB_j(t) + \sum_{k;k\neq j} b_{jk}(t)\mu_{jk}(t)\, dt \right) \\ & - \sum_j \left(\sum_{g;\, g\neq j} p_{0g}(0,t)\, \mu_{gj}(t)\, dt \;-\; p_{0j}(0,t)\, \mu_{j\cdot}(t)\, dt\,. \right) V_j^*(t) \\ & - \sum_j p_{0j}(0,t) \left(r^*(t) V_j^*(t)\, dt - dB_j(t) - \sum_{k;\, k\neq j} R_{jk}^*(t)\, \mu_{jk}^*(t)\, dt \right) . \end{aligned}$$

Reusing the relation (10.6) in the first line here and gathering terms, we obtain

$$dS(t) = r(t)\, dt\, S(t) \;+\; \sum_j p_{0j}(0,t) c_j(t)\, dt\,,$$

with

$$c_j(t) = \{r(t) - r^*(t)\}\, V_j^*(t) + \sum_{k;\, k\neq j} R_{jk}^*(t) \{\mu_{jk}^*(t) - \mu_{jk}(t)\}\,. \tag{10.9}$$

Finally, integrating up and using (10.7), we arrive at

$$S(t) = \int_0^t e^{\int_\tau^t r} \sum_j p_{0j}(0,\tau) c_j(\tau)\, d\tau\,, \tag{10.10}$$

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 (10.9) by another route, starting from the *individual surplus* defined, quite naturally, as

$$S_{ind}(t) = e^{\int_0^t r} \int_{0-}^{t} e^{-\int_0^\tau r}\, d(-B)(\tau) - \sum_j I_j^Z(t) V_j^*(t). \tag{10.11}$$

Upon differentiating this expression, and proceeding along the same lines as above, one finds that $S_{ind}(t)$ consists of a purely erratic term and a systematic term. The latter is $\int_0^t e^{\int_\tau^t r} \sum_j I_j^Z(\tau) c_j(\tau)\, d\tau$, which is the individual counterpart of (10.10), showing how the contributions emerge at the level of the individual policy. They form a random payment function C defined by

$$dC(t) = \sum_j I_j^Z(t)\, c_j(t)\, dt\,. \tag{10.12}$$

With this definition, we can recast (10.10) as

$$S(t) = \mathbb{E}\left[\int_0^t e^{\int_\tau^t r} dC(\tau) \,\middle|\, \mathcal{G}_t \right]. \tag{10.13}$$

C. Safety margins. The expression on the right of (10.9) 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

$$r(t) \geq r^*(t) \tag{10.14}$$

(assuming that all $V_j^*(t)$ are non-negative as they should be) and

$$\operatorname{sign}\{\mu_{jk}^*(t) - \mu_{jk}(t)\} = \operatorname{sign} R_{jk}^*(t)\,. \tag{10.15}$$

10.4 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 10.1.A, they can not be negative. The purpose of the dividends is to establish, ultimately, equivalence on the true second order basis:

$$\mathbb{E}\left[\int_{0-}^{n} e^{-\int_0^\tau r}\, d\{B + D\}(\tau) \,\middle|\, \mathcal{G}_n \right] = 0\,. \tag{10.16}$$

We can state (10.16) equivalently as

$$\mathbb{E}\left[\int_{0-}^{n} e^{\int_\tau^n r}\, d\{B+D\}(\tau)\,\middle|\, \mathcal{G}_n\right] = 0\,. \tag{10.17}$$

The value at time t of past individual contributions less dividends, compounded with interest, is

$$U^d(t) = \int_{0-}^{t} e^{\int_\tau^t r}\, d\{C-D\}(\tau)\,. \tag{10.18}$$

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 (10.8) and (10.13) we can recast the equivalence requirement (10.17) in the appealing form

$$\mathbb{E}[U^d(n)\,|\,\mathcal{G}_n] = 0\,. \tag{10.19}$$

From a solvency point of view it would make sense to strengthen (10.19) by requiring that compounded dividends must never exceed compounded contributions:

$$\mathbb{E}[U^d(t)\,|\,\mathcal{G}_t] \geq 0\,, \tag{10.20}$$

$t \in [0, n]$. 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 (10.20).

If we adopt the view that "the technical surplus belongs to those who created it", we should sharpen (10.19) by imposing the stronger requirement

$$U^d(n) = 0\,. \tag{10.21}$$

This means that no transfer of redistributions across policies is allowed. The solvency requirement conforming with this point of view, and sharpening (10.20), is

$$U^d(t) \geq 0\,, \tag{10.22}$$

$t \in [0, n]$.

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

$$U^b(t) = \int_0^t e^{\int_\tau^t r} \, d\{D - B^b\}(\tau) \, . \tag{10.23}$$

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

$$U^b(t) \geq 0 \tag{10.24}$$

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

$$U^b(n) = 0 \, . \tag{10.25}$$

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

$$U^d(t) + U^b(t) = \int_0^t e^{\int_\tau^t r}\, d\{C - B^b\}(\tau)\,, \tag{10.26}$$

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,

$$B^b(t) = 1[t \geq T] \int_0^T e^{\int_\tau^T r} dD(\tau)\,.$$

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

$$dD(s) = dQ(s) V^{*+}_{Z(s)}(s), \tag{10.27}$$

where

$$V^{*+}_{Z(s)}(s) = \mathbb{E}^*\left[\int_s^n e^{-\int_s^\tau r^*} dB^+(\tau) \,\middle|\, Z(s)\right]$$

is the single premium at time s for the future benefits under the policy.

Now the bonus payments B^b are of the form

$$dB^b(t) = Q(t)dB^+(t)\,. \tag{10.28}$$

Being written on first order basis, also the additional insurances create technical surplus. The total contributions under this scheme develop as

$$dC(t) + Q(t)dC^+(t)\,, \tag{10.29}$$

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 $dC^+(t) = \sum_j I_j^Z(t)\, c_j^+(t)\, dt$, 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 (10.27) are set equal to the contributions in (10.29). Then the system is governed by the dynamics

$$dC(t) + Q(t)dC^+(t) = dQ(t)V_{Z(t)}^{*+}(t)$$

or, realizing that $V_{Z(t)}^{*+}(t)$ is strictly positive whenever $dC(t)$ and $dC^+(t)$ are,

$$dQ(t) - Q(t)\, dG(t) = dH(t)\,, \tag{10.30}$$

where G and H are defined by

$$dG(t) = \frac{1}{V_{Z(t)}^{*+}(t)}\, dC^+(t)\,, \tag{10.31}$$

$$dH(t) = \frac{1}{V_{Z(t)}^{*+}(t)}\, dC(t)\,. \tag{10.32}$$

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

$$Q(t) = \int_0^t e^{G(t)-G(\tau)}\, dH(\tau)\,. \tag{10.33}$$

10.5 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 $\mathcal{Y} = \{1, \ldots, q\}$ and constant intensities of transition, λ_{ef}. Denote the associated indicator processes by I_e^Y. The process Y represents the "economic-demographic environment", and we let the second order elements depend on the current Y-state:

$$r(t) = \sum_e I_e^Y(t)\, r_e = r_{Y(t)}\,,$$

$$\mu_{jk}(t) = \sum_e I_e^Y(t)\, \mu_{e;jk}(t) = \mu_{Y(t);jk}(t)\,.$$

The r_e are constants and the $\mu_{e;jk}(t)$ 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 $\kappa_{ej,fk}(t)$ for (e, j), $(f, k) \in \mathcal{X}$, $(e, j) \neq (f, k)$, are

$$\kappa_{ej,fj}(t) = \lambda_{ef}\,, \qquad e \neq f, \tag{10.34}$$

$$\kappa_{ej,ek}(t) = \mu_{e;jk}(t)\,, \quad j \neq k, \tag{10.35}$$

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

$$dC(t) = c(t)\,dt = \sum_{e,j} I_e^Y(t) I_j^Z(t) c_{ej}(t)\,dt,$$

where

$$c_{ej}(t) = \{r_e - r^*(t)\}\, V_j^*(t) + \sum_{k;\, k\neq j} R_{jk}^*(t)\{\mu_{jk}^*(t) - \mu_{e;jk}(t)\}\,. \tag{10.36}$$

Under the scheme of additional benefits described in Paragraph 10.4.D a similar convention goes for C^+ and c^+ and, accordingly, (10.31) and (10.32) become

$$dG(t) = g(t)\,dt = \sum_{e,j} I_e^Y(t) I_j^Z(t) g_{ej}(t)\,dt\,, \tag{10.37}$$

$$g_{ej}(t) = \frac{c_{ej}^+(t)}{V_j^{*+}(t)}\,, \tag{10.38}$$

$$dH(t) = h(t)\,dt = \sum_{e,j} I_e^Y(t) I_j^Z(t) h_{ej}(t)\,dt\,, \tag{10.39}$$

$$h_{ej}(t) = \frac{c_{ej}(t)}{V_j^{*+}(t)}\,. \tag{10.40}$$

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 10.4.

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),

$$W_{ej}(t) = \mathbb{E}\left[\int_t^n e^{-\int_t^\tau r}\, c(\tau)\, d\tau \,\middle|\, X(t) = (e,j) \right] .$$

They are determined by the appropriate version of Thiele's differential equation,

$$\begin{aligned}\frac{d}{dt} W_{ej}(t) &= r_e\, W_{ej}(t) - c_{ej}(t) - \sum_{f;f\neq e} \lambda_{ef}\left(W_{fj}(t) - W_{ej}(t)\right) \\ &\quad - \sum_{k;k\neq j} \mu_{e;jk}(t)\, \left(W_{ek}(t) - W_{ej}(t)\right) , \end{aligned} \tag{10.41}$$

subject to

$$W_{ej}(n-) = 0\,, \qquad \forall e, j\,. \tag{10.42}$$

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,

$$W = \int_0^T e^{\int_\tau^T r} c(\tau)\, d\tau = \int_0^t e^{\int_\tau^t r}\, c(\tau)\, d\tau\, W'(t) + W''(t)\,, \tag{10.43}$$

where

$$\begin{aligned} W'(t) &= e^{\int_t^T r}, \\ W''(t) &= \int_t^T e^{\int_\tau^T r} c(\tau)\, d\tau\,. \end{aligned}$$

We need the state-wise expected values

$$\begin{aligned} W'_e(t) &= \mathbb{E}[W'(t) \,|\, Y(t) = e], \\ W''_{ej}(t) &= \mathbb{E}[W''(t) \,|\, X(t) = (e,j)]\,, \end{aligned}$$

to find the state-wise predictors of W in (10.43),

$$W_{ej}(t) = \int_0^t e^{\int_\tau^t r}\, c(\tau)\, d\tau\, W'_e(t) + W''_{ej}(t)\,.$$

We shall find these functions by the backward construction, starting from

$$\begin{aligned} W'(t) &= e^{r\,dt}\, W'(t+dt), \\ W''(t) &= c(t)\,dt\, W'(t) + W''(t+dt)\,. \end{aligned}$$

Conditioning on what happens in the small time interval $(t, t+dt]$, we get

$$W'_e(t) = e^{r_e\,dt} \left((1 - \lambda_{e\cdot}\,dt)\, W'_e(t+dt) + \sum_{f;\, f\neq e} \lambda_{ef}(t)\,dt\, W'_f(t+dt) \right) ,$$

and

$$\begin{aligned} W''_{ej}(t) &= c_{ej}(t)\,dt\, W'_e(t) + (1 - (\lambda_{e\cdot} + \mu_{e;j\cdot}(t))\,dt)\, W''_{ej}(t+dt) \\ &\quad + \sum_{f;\, f\neq e} \lambda_{ef}(t)\,dt\, W''_{fj}(t+dt) \\ &\quad + \sum_{k;\, k\neq j} \mu_{e;jk}(t)\,dt\, W''_{ek}(t+dt)\,. \end{aligned}$$

From these relationships we easily obtain the differential equations

$$\frac{d}{dt} W'_e(t) = -r_e\, W'_e(t) - \sum_{f;\, f\neq e} \lambda_{ef}\, \left(W'_f(t) - W'_e(t) \right) , \tag{10.44}$$

$$\begin{aligned} \frac{d}{dt} W''_{ej}(t) &= -\,c_{ej}(t)\, W'_e(t) - \sum_{f;\, f\neq e} \lambda_{ef}\, \left(W''_{fj}(t) - W''_{ej}(t) \right) \\ &\quad - \sum_{k;\, k\neq j} \mu_{e;jk}(t)\, \left(W''_{ek}(t) - W''_{ej}(t) \right) , \end{aligned} \tag{10.45}$$

which are to be solved subject to

$$W'_e(n-) = 1\,, \qquad W''_{ej}(n-) = 0\,, \qquad \forall e, j\,. \tag{10.46}$$

E. Additional benefits. Suppose we want to predict the total future bonuses discounted with respect to second order interest,

$$W(t) = \int_t^n e^{-\int_t^\tau r}\, Q(\tau)\, dB^+(\tau)\,,$$

with Q defined by (10.33). Recalling (10.37)–(10.40), we reshape $W(t)$ as

$$\begin{aligned} W(t) &= \int_t^n e^{-\int_t^\tau r} \int_0^\tau e^{\int_r^\tau g}\, h(r)\, dr\, dB^+(\tau) \\ &= \int_t^n e^{-\int_t^\tau r} \left(\int_0^t e^{\int_r^t g}\, h(r)\, dr\, e^{\int_t^\tau g} + \int_t^\tau e^{\int_r^\tau g}\, h(r)\, dr \right) dB^+(\tau) \\ &= \int_0^t e^{\int_r^t g}\, h(r)\, dr\, W'(t) + W''(t)\,, \end{aligned} \tag{10.47}$$

with

$$
\begin{aligned}
W'(t) &= \int_t^n e^{\int_t^\tau (g-r)}\, dB^+(\tau), \\
W''(t) &= \int_t^n e^{-\int_t^\tau r}\, W'(\tau)\, h(\tau)\, d\tau\,.
\end{aligned}
$$

Thus, we need the state-wise expected values

$$
\begin{aligned}
W'_{ej}(t) &= \mathbb{E}[W'(t)\,|\,X(t)=(e,j)]\,, \\
W''_{ej}(t) &= \mathbb{E}[W''(t)\,|\,X(t)=(e,j)]\,,
\end{aligned}
$$

in order to find the state-wise predictors of $W(t)$ in (10.47),

$$
W_{ej}(t) = \int_0^t e^{\int_r^t g}\, h(r)\, dr\, W'_{ej}(t) + W''_{ej}(t)\,.
$$

The backward equations start from

$$
\begin{aligned}
W'(t) &= dB^+(t) + e^{(g(t)-r(t))\,dt}\, W'(t+dt)\,, \\
W''(t) &= W'(t)\, h(t)\, dt + e^{-r(t)\,dt}\, W''(t+dt)\,,
\end{aligned}
$$

from which we proceed in the same way as in the previous paragraph to obtain

$$
\begin{aligned}
dW'_{ej}(t) &= -\,dB^+_j(t) + (r_e - g_{ej}(t))\, dt\, W'_{ej}(t) \\
&\quad - \sum_{f;\,f\neq e} \lambda_{ef}\, dt\, \left(W'_{fj}(t) - W'_{ej}(t)\right) \\
&\quad - \sum_{k;k\neq j} \mu_{e;jk}(t)\, dt\, \left(b^+_{jk}(t) + W'_{ek}(t) - W'_{ej}(t)\right)\,, \qquad (10.48) \\
dW''_{ej}(t) &= -\,W'_{ej}(t) h_{ej}(t)\, dt + r_e\, dt\, W''_{ej}(t) \\
&\quad - \sum_{f;\,f\neq e} \lambda_{ef}\, dt\, \left(W''_{fj}(t) - W''_{ej}(t)\right) \\
&\quad - \sum_{k;\,k\neq j} \mu_{e;jk}(t)\, dt\, \left(W''_{ek}(t) - W''_{ej}(t)\right)\,. \qquad (10.49)
\end{aligned}
$$

The appropriate side conditions are

$$
W'_{ej}(n-) = \Delta B^+_j(n)\,, \qquad W''_{ej}(n-) = 0\,, \qquad \forall e,j\,. \qquad (10.50)
$$

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

$$\mathbb{E}[c_{Y(s)i}(s) \,|\, Y(t) = e] = \sum_f p^Y_{ef}(t,s) c_{fi}(s)\,.$$

10.6 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 $\mathcal{Z} = \{a, d\}$ ('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:

$$\begin{aligned} r^* &= \ln(1.045)\,, \\ \mu^*_{ad}(t) &= \mu^*(t) = 0.0005 + 0.000075858 \cdot 10^{0.038(30+t)}\,. \end{aligned}$$

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 λ_i between the first order rate r^* and a better rate $\epsilon_i r^*$ ($\epsilon_i > 1$) and, similarly, μ switches with a constant intensity λ_m between the first order rate μ^* and a better rate $\epsilon_m \mu^*$ ($\epsilon_m < 1$). (We choose to express ourselves this way although (10.15) 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.)

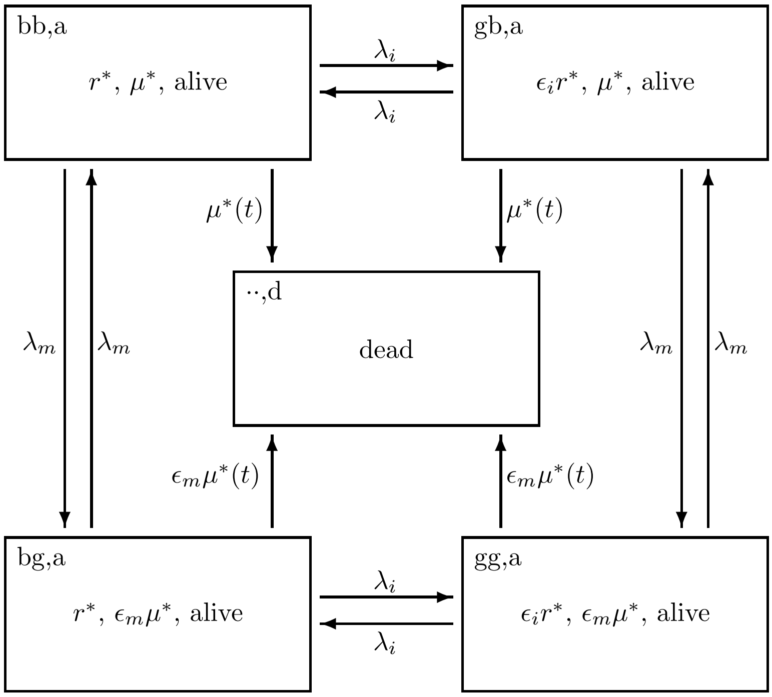


Figure 10.1: The Markov process $X = (Y, Z)$ for a single life insurance in an environment with two interest states and two mortality states.

The situation fits into the framework of Paragraph 10.5.A; Y has states $\mathcal{Y} = \{bb, gb, bg, gg\}$ representing all combinations of "bad" (b) and "good" (g) interest and mortality, and the non-null intensities are

$$\lambda_{bb,gb} = \lambda_{gb,bb} = \lambda_{bg,gg} = \lambda_{gg,bg} = \lambda_i\,,$$

$$\lambda_{bb,bg} = \lambda_{bg,bb} = \lambda_{gb,gg} = \lambda_{gg,gb} = \lambda_m\,.$$

The first order basis is just the worst-scenario bb.

Adopting the device (10.34)–(10.35), 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 10.1.

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 10.1 displays, for each of the three policies, the state-wise expected values of discounted contributions obtained by solving (10.41)–(10.42). 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, see (10.15). 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, see 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 10.1.

Table 10.2 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 10.5 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.

Table 10.1: 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).

	bb	gb	bg	gg
TI :	.00851	.00854	.01061	.01059
PE :	.01613	.01823	.01595	.01807
EI :	.02463	.02677	.02656	.02865

10.7 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,

$$dA(t) = \sum_j I_j^Z(t)\, dA_j(t) + \sum_{j \neq k} a_{jk}(t)\, dN_{jk}^Z(t)\,. \tag{10.51}$$

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 (10.5) now turns into

$$V_0^*(0) = -\Delta A_0^*(0) - \Delta B_0(0)\,, \tag{10.52}$$

the surplus at time 0 becomes

$$S(0) = \Delta A_0^*(0) - \Delta A_0(0)\,, \tag{10.53}$$

Table 10.2: 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).

			bb	gb	bg	gg
TI:	E	C :	.02153	.02222	.02436	.02505
	E	TB :	.03693	.03916	.04600	.04847
	E	AB :	.02949	.03096	.03545	.03706
PE:	E	C :	.04342	.04818	.04314	.04791
	E	TB :	.07337	.08687	.07264	.08615
	E	AB :	.07337	.08687	.07264	.08615
EI:	E	C :	.06495	.07040	.06750	.07296
	E	TB :	.11030	.12603	.11864	.13462
	E	AB :	.10723	.12199	.11501	.13003

and the contributions consist of a jump

$$\Delta C(0) = \Delta A_0^*(0) - \Delta A_0(0) \tag{10.54}$$

at time 0 and thereafter a continuous part, which is defined upon replacing (10.9) with

$$\begin{aligned} c_j(t) &= \{r(t) - r^*(t)\}\, V_j^*(t) + \{a_j^*(t) - a_j(t)\} \\ &\quad + \sum_{k;\, k \neq j} \{a_{jk}^*(t) - a_{jk}(t)\}\mu_{jk}(t) \\ &\quad + \sum_{k;\, k \neq j} R_{jk}^*(t)\{\mu_{jk}^*(t) - \mu_{jk}(t)\}\,, \end{aligned} \tag{10.55}$$

where now

$$R_{jk}^*(t) = a_{jk}^*(t) + b_{jk}(t) + V_k^*(t) - V_j^*(t)\,. \tag{10.56}$$

Referring to the discussion in Paragraph 10.3.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 $\Delta A_0(0)$, $a_j(t)$, and $a_{jk}(t)$ are deterministic. Then the methods in Section 10.5 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 10.5.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,

$$\begin{aligned}
\Delta A_0(0) &= \sum_e I_e^Y(t)\,\Delta A_{e0}(0)\,,\\
a_j(t) &= \sum_e I_e^Y(t)\,a_{ej}(t)\,,\\
a_{jk}(t) &= \sum_e I_e^Y(t)\,a_{ejk}(t)\,,
\end{aligned}$$

with deterministic ΔA_{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 $a_j(t) = e^{\int_0^t \gamma} a_j^0(t)$ and $a_{jk}(t) = e^{\int_0^t \gamma} a_{jk}^0(t)$ 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 γ. More realistically one should let γ 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 10.5.

10.8 Discussions

A. The principle of equivalence. This principle, as formulated in (10.4), 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 individual 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 premiums 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 10.7.)

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 (10.16). 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 (10.16) 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 10.6. Table 10.3 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 10.3: 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 $r = \epsilon_i r^*$ and mortality $\mu = \epsilon_m \mu^*$ for various choices of ϵ_i and ϵ_m.

			TI			LA		
		ϵ_m :	1.5	1.0	0.5	1.5	1.0	0.5
ϵ_i :	0.5	$\mathbb{E}$:	.14636	.10119	.05250	20.545	20.996	21.467
		SD :	.27902	.24104	.18041	03.691	03.101	02.257
	1.0	E :	.09927	.06834	.03531	15.750	16.039	16.340
		SD :	.20245	.17330	.12857	02.505	02.097	01.521
	1.5	E :	.06976	.04782	.02460	12.466	12.655	12.852
		SD :	.15858	.13437	.09868	01.759	01.468	01.061

Chapter 11

Statistical inference in the Markov chain model

Think of the insurance company as a car:
At the steering wheel sits the managing
director trying to keep the vehicle steadily
on the road. In the front passenger seat
sits the sales manager pushing the speed
pedal. At the rear sits the actuary peeping
out the back window and giving the directions.

11.1 Estimating a mortality law from fully observed life lengths

A. Completely observed life lengths. Let the life length of an individual be represented by a non-negative random variable T with cumulative distribution function of the form

$$F(t;\theta) = 1 - e^{-\int_0^t \mu(s;\theta)\, ds} . \tag{11.1}$$

The mortality intensity μ is assumed to be continuous (at least piece-wise) so that the density

$$f(t) = \mu(t)(1 - F(t)) \tag{11.2}$$

exists.

A. Right-censored life times. Let the total life length of an individual be represented by a non-negative random variable T with cumulative distribution function of the form

$$F(t) = 1 - e^{-\int_0^t \mu(s)\, ds} . \tag{11.3}$$

The mortality intensity μ is assumed to be continuous (at least piece-wise) so that the density

$$f(t) = \mu(t)(1 - F(t)) , \tag{11.4}$$

exists.

Suppose that the individual is observed continually in z years from its birth so that only the truncated life length $W = T \wedge z$ is observed. A technical term for this kind of observational plan is *right-censoring at time z*. The cumulative distribution of W is

$$\mathbb{P}[W \le t] = \begin{cases} F(t), & 0 < t < z , \\ 1, & t \ge z . \end{cases}$$

and the density (with respect to Lebesgue measure on $(0, z)$ and the unit mass at z) is (recall (11.4))

$$g(t) = \begin{cases} \mu(t)(1 - F(t)), & 0 < t < z\,, \\ 1 - F(z), & t = z\,. \end{cases}$$

Introduce

$$d(t) = 1_{(0,z)}(t) = \begin{cases} 1\,, & 0 < t < z\,, \\ 0\,, & t \geq z\,, \end{cases} \tag{11.5}$$

to obtain the closed expression

$$g(t) = \mu(t)^{d(t)}(1 - F(t))\ ,\ 0 < t \leq z\,. \tag{11.6}$$

Denote the indicator function of survival to age $u > 0$ by $I(u) = 1[U > u]$. The indicator of death before age z is

$$D = d(T) = 1 - I(z-) = 1 - I(z)\,, \tag{11.7}$$

where the last equality holds with probability 1.

We will need the following formulas, valid whenever the displayed moments exist:

$$\mathbb{E}[D^k] = F(z)\ ,\ k = 1, 2, \ldots, \tag{11.8}$$

$$\mathbb{E}[T^k] = k \int_0^z t^{k-1}(1 - F(t))dt\ ,\ k = 1, 2, \ldots, \tag{11.9}$$

$$\mathbb{E}[DT] = \mathbb{E}[T] - z(1 - F(z))\,. \tag{11.10}$$

To verify (11.10), use (11.7) 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

$$F(u; \mu) = 1 - e^{-\mu u}\ ,\ u > 0\,, \tag{11.11}$$

that is, μ is constant, independent of age. The expected life length is

$$\nu = \int_0^\infty (1 - F(u; \mu))du = \frac{1}{\mu}\,. \tag{11.12}$$

Using (11.8) – (11.11) one easily calculates

$$\mathbb{E}[D] = 1 - e^{-\mu z}\,, \tag{11.13}$$

$$\mathbb{V}[D] = e^{-\mu z}(1 - e^{-\mu z})\,, \tag{11.14}$$

$$\mathbb{E}[T] = \frac{1 - e^{-\mu z}}{\mu}\,, \tag{11.15}$$

$$\mathbb{V}[T] = \frac{1 - 2\mu z e^{-\mu z} - e^{-2\mu z}}{\mu^2}\,, \tag{11.16}$$

$$\mathbb{E}[D - \mu T] = 0\,, \tag{11.17}$$

$$\mathbb{V}[D - \mu T] = 1 - e^{-\mu z}\,. \tag{11.18}$$

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 μ 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 (11.7), put $N_i = 1[T_i < z_i]$, $i = 1, \ldots, n$. By (11.6) and (11.11), the likelihood of the observables is

$$\Lambda = \Pi_{i=1}^{n} \mu^{N_i} e^{-\mu T_i} = \mu^{N} e^{-\mu W} = e^{\ln \mu \, N - \mu W}, \tag{11.19}$$

where

$$\begin{aligned} N &= \sum_{i=1}^{n} N_i, \text{ the total number of deaths occurred,} \\ W &= \sum_{i=1}^{n} T_i, \text{ the total time exposed to risk of death.} \end{aligned}$$

Clearly, (N, W) is a sufficient statistic. Take the logarithm,

$$\ln \Lambda = \ln \mu \, N - \mu W \, ,$$

and form the derivatives

$$\frac{\partial}{\partial \mu} \ln \Lambda = \frac{N}{\mu} - W \, , \tag{11.20}$$

$$\frac{\partial^2}{\partial \mu^2} \ln \Lambda = -\frac{N}{\mu^2} \, . \tag{11.21}$$

Putting the expression in (11.20) equal to 0 and noting that the second derivative is non-positive, we find that the maximum likelihood estimator (MLE) of μ is the so-called *occurrence-exposure rate*, (OE rate)

$$\hat{\mu} = \frac{N}{W} \, , \tag{11.22}$$

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 (11.12) is

$$\hat{\nu} = \frac{W}{N} \, , \tag{11.23}$$

defined as $+\infty$ when $N = 0$. This estimator has no mean (and no higher order moments).

The expressions for the likelihood in (11.19) and the MLE in (11.22) do not appear to depend on the censoring mechanoism. The censoring is, however, decisive of the probability distribution of $\hat{\mu}$ and, hence, of its performance as an estimator of μ. 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 N_i are 1, $N = n$, $W = \sum_{i=1}^{n} T_i$, and the likelihood in (11.19) becomes

$$\Lambda = e^{\ln \mu \, n - \mu W} \, . \tag{11.24}$$

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

$$\frac{\mu^n}{\Gamma(n)} w^{n-1} e^{-\mu w} \; , \, w > 0 \, .$$

One finds (perform the easy calculations) for $k > -n$ that

$$\mathbb{E}[W^k] = \frac{\Gamma(n+k)}{\Gamma(n) \mu^k} \, ,$$

hence

$$\mathbb{E}[\hat{\mu}] = \frac{n\mu}{n-1} \;,\; n > 1 \,,$$
$$\mathbb{V}[\hat{\mu}] = \frac{n^2\mu^2}{(n-1)^2(n-2)} \;,\; n > 2 \,,$$

and

$$\mathbb{E}[\hat{\nu}] = \nu \;,\; n \geq 1 \,,$$
$$\mathbb{V}[\hat{\nu}] = \frac{\nu^2}{n} \;,\; n \geq 1 \,.$$

The estimator $\hat{\mu}$ is biased and, on the average, overestimates μ by $\mu/(n-1)$, which is negligible for large n. An unbiased estimator of μ is $\breve{\mu} = (n-1)/W$. Its variance is $\mu^2/(n-2)$. The estimator $\hat{\nu}$ is now just the observed average life length, the empirical counterpart of ν. It is unbiased, of course. In fact, $\breve{\mu}$ and $\hat{\nu}$ are UMVUE (uniformly minimum variance unbiased estimators) since they are based on W, which is a sufficient and complete statistic: by (11.24), the distribution belongs to an exponential family with canonical parameter μ varying in the open set $(0, \infty)$.

E. Asymptotic results by uniform censoring. Suppose all z_i are equal to z, say. Writing

$$\hat{\mu} = \frac{\frac{1}{n}\sum_{i=1}^n N_i}{\frac{1}{n}\sum_{i=1}^n T_i} \,,$$

and noting that $\mathbb{E}[N_i] = \mu\mathbb{E}T_i$ by (11.13) and (11.15), it follows by the strong law of large numbers that the estimator is strongly consistent,

$$\hat{\mu} \overset{\text{a.s.}}{\to} \mu \,.$$

To investigate its asymptotic distribution, look at

$$\sqrt{n}(\hat{\mu} - \mu) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n (N_i - \mu T_i)}{\frac{1}{n}\sum_{i=1}^n T_i} \,.$$

The denominator of this fraction converges a.s. to $\mathbb{E}[T_i]$ given by (11.15). By the central limit theorem, the limiting distribution of the numerator is normal with mean 0 (recall (11.17)) and variance given by (11.18). It follows that

$$\hat{\mu} \sim_{\text{as}} \mathrm{N}\left(\mu, \frac{\mu^2}{n(1 - e^{-\mu z})}\right) \,. \tag{11.25}$$

Copying the arguments above (or using (D.6) in Appendix D), it can also be concluded that $\hat{\nu}$ defined by (11.23) is strongly consistent and that

$$\hat{\nu} \sim_{\text{as}} \mathrm{N}\left(\nu, \frac{1}{n\mu^2(1 - e^{-\mu z})}\right) \,. \tag{11.26}$$

No strong conclusions as to optimality can be drawn in parallel to those in the previous paragraph. The reason is seen from (11.19): 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 (11.15) is equivalent to

$$\sum_{i=1}^{n}(1-e^{-\mu z_i}) \to \infty\,, \tag{11.27}$$

that is, the expected number of deaths grows to infinity. Thus assume that (11.27) is satisfied. In the following the relationships (11.13) – (11.18) will be used frequently without explicit mentioning.

First, to prove consistency, use (11.15) and (11.18) to write

$$\begin{aligned}\hat{\mu}-\mu &= \frac{\sum_{i=1}^{n}(N_i-\mu T_i)}{\sum_{i=1}^{n}T_i}\,. \\ &= \frac{\sum_{i=1}^{n}(N_i-\mu T_i)}{\sum_{i=1}^{n}\mathbb{V}[N_i-\mu T_i]}\left(\frac{\sum_{i=1}^{n}T_i}{\mu\sum_{i=1}^{n}\mathbb{E}[T_i]}\right)^{-1}\,. \end{aligned} \tag{11.28}$$

The first factor in (11.28) has expected value 0 and variance

$$\frac{1}{\sum_{i=1}^{n}\mathbb{V}[N_i-\mu T_i]} = \frac{1}{\sum_{i=1}^{n}(1-e^{-\mu z_i})}\,,$$

which tends to 0 as n increases. Therefore, this factor tends to 0 in probability. The second factor in (11.28) 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

$$\begin{aligned}\frac{\sum_{i=1}^{n}\mathbb{V}[T_i]}{\left(\sum_{i=1}^{n}(1-e^{-\mu z_i})\right)^2} &= \frac{\sum_{i=1}^{n}(1-2\mu z_i e^{-\mu z_i}-e^{-2\mu z_i})}{\left(\sum_{i=1}^{n}(1-e^{-\mu z_i})\right)^2} \\ &= \frac{\sum_{i=1}^{n}a(\mu z_i)(1-e^{-\mu z_i})}{\sum_{i=1}^{n}(1-e^{-\mu z_i})}\,\frac{1}{\sum_{i=1}^{n}(1-e^{-\mu z_i})}\,, \end{aligned} \tag{11.29}$$

where a is defined as

$$a(t) = \frac{1-2te^{-t}-e^{-2t}}{1-e^{-t}}\ ,\ t \geq 0.$$

The function a is bounded since it is continuous and tends to 0 as $t \searrow 0$ (use l'Hospital's rule) and to 1 as $t \nearrow \infty$. The first factor in (11.29) 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 (11.29) tends to 0 and, consequently, that the second factor in (11.28) converges in probability to μ. It can be concluded that the expression in (11.28) converges in probability to 0, so that $\hat{\mu}$ is weakly consistent;

$$\hat{\mu} \xrightarrow{\mathrm{p}} \mu\,.$$

Next, to prove asymptotic normality, look at

$$\sqrt{\sum_{i=1}^{n}(1-e^{-\mu z_i})}(\hat{\mu}-\mu) = \frac{\sum_{i=1}^{n}(N_i-\mu T_i)}{\sqrt{\sum_{i=1}^{n}(1-e^{-\mu z_i})}}\left(\frac{\sum_{i=1}^{n}T_i}{\mu\sum_{i=1}^{n}\mathbb{E}[T_i]}\right)^{-1}\,.$$

In the presence of (11.27) 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 μ. It follows that

$$\hat{\mu} \sim_{\mathrm{as}} \mathrm{N}\left(\mu, \frac{\mu^2}{\sum_{i=1}^{n}(1-e^{-\mu z_i})}\right)\,. \tag{11.30}$$

Likewise, it also holds that $\hat{\nu}$ is weakly consistent, and

$$\hat{\nu} \sim_{\mathrm{as}} \mathrm{N}\left(\nu, \frac{1}{\mu^2\sum_{i=1}^{n}(1-e^{-\mu z_i})}\right)\,. \tag{11.31}$$

G. Random censoring. In Paragraph E the censoring time was assumed to be the same for all individuals. Thereby the pairs (N_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 (11.27), 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 augmented model where the 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 μ. This way the triplets (N_i, T_i, Z_i), $i = 1, 2, \ldots$, become stochastic replicates, and the i.i.d. situation is restored with all its powers.

The likelihood of the observations now becomes

$$\Lambda = \Pi_{i=1}^{n} \mu^{N_i} e^{-\mu T_i}\, h(Z_i) = e^{\ln \mu\, N - \mu W}\, \Pi_{i=1}^{n} h(Z_i)\,. \tag{11.32}$$

Maximizing (11.32) with respect to μ is equivalent to maximizing the likelihood (11.19) 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 $Z \sim H$.

11.2 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 represented by a continuous time Markov process Z on a finite state space $\mathcal{Z} = \{0, 1, \ldots, J\}$. 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 μ_{gh} are assumed to exist, and to be piecewise continuous.

Suppose the policy is observed continuously throughout the time period $[\underline{t}, \bar{t}\,]$, commencing in state g_0 at time $\underline{t}$. One then speaks of *left-censoring* and *right-censoring* at times $\underline{t}$ and $\bar{t}$, respectively, and the triplet $z = (\underline{t}, \bar{t}, g_0)$ will be referred to as *observational design* or *censoring scheme* of the policy.

Consider a specific realization of the observed part of the process:

$$X(\tau) = \begin{cases} g_0 & , \underline{t} < \tau < t_1, \\ g_1 & , t_1 + dt_1 < \tau < t_2, \\ & \ldots \\ g_{q-2} & , t_{q-2} + dt_{q-2} < \tau < t_{q-1}, \\ g_{q-1} & , t_{q-1} + dt_{q-1} < \tau < \bar{t}. \end{cases}$$

By the given censoring, the probability of this realization is as follows, where $t_0 = \underline{t}$, $t_q = \bar{t}$, and $\mu_g = \sum_{h \neq g} \mu_{gh}$ denotes the total intensity of transition out of state g:

$$\begin{aligned} & \exp\left(-\int_{t_0}^{t_1} \mu_{g_0}\right) \mu_{g_0 g_1}(t_1) dt_1 \exp\left(-\int_{t_1}^{t_2} \mu_{g_1}\right) \mu_{g_1 g_2}(t_2) dt_2 \ldots \\ & \exp\left(-\int_{t_{q-2}}^{t_{q-1}} \mu_{g_{q-2}}\right) \mu_{g_{q-2} g_{q-1}}(t_{q-1}) dt_{q-1} \exp\left(-\int_{t_{q-1}}^{t_q} \mu_{g_{q-1}}\right) \\ = & \prod_{p=1}^{q-1} \mu_{g_{p-1} g_p}(t_p) dt_p \;\; \exp\left(-\sum_{1}^{q} \int_{t_{p-1}}^{t_p} \mu_{g_{p-1}}\right) \end{aligned}$$

$$= \exp\left(\sum_{p=1}^{q-1} \ln \mu_{g_{p-1}g_p}(t_p) - \sum_1^q \int_{t_{p-1}}^{t_p} \mu_{g_{p-1}}\right) dt_1 \ldots dt_{q-1}\,.$$

It follows that the likelihood of the observables is

$$\begin{aligned}\Lambda &= \exp\left(\sum_{g\neq h} \int_{\underline{t}}^{\bar{t}} \ln \mu_{gh}(\tau) dN_{gh}(\tau) - \sum_g \int_{\underline{t}}^{\bar{t}} \mu_g(\tau) I_g(\tau) d\tau\right) \\ &= \exp\left(\sum_{g\neq h} \int_{\underline{t}}^{\bar{t}} \{\ln \mu_{gh}(\tau) dN_{gh}(\tau) - \mu_{gh}(\tau) I_g(\tau) d\tau\}\right)\,. \qquad (11.33)\end{aligned}$$

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 θ.

Suppose that inference is to be made about the intensities or, equivalently, the parameter θ 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 Z 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 (11.33),

$$\ln \Lambda = \sum_{g\neq h} \int \{\ln \mu_{gh}(\tau, \theta) dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau\}\,, \qquad (11.34)$$

with

$$N_{gh} = \sum_{m=1}^n N_{gh}^{(m)}\,,\; I_g = \sum_{m=1}^n I_g^{(m)}\,.$$

The censoring schemes are not visualized in (11.34), and they need not be if, as a matter of definition, $dN_{gh}^{(m)}(t)$ and $I_g^{(m)}(t)$ are taken as 0 for $t \notin [\underline{t}^{(m)}, \bar{t}^{(m)}]$. Likewise, introduce

$$p_g^{(m)}(t) = p_{g_0^{(m)}g}(\underline{t}^{(m)}, t) 1_{[\underline{t}^{(m)}, \bar{t}^{(m)}]}(t)\,,$$

the probability that the censored process $Z^{(m)}$ stays in g at time t, by definition taken as 0 for $t \notin [\underline{t}^{(m)}, \bar{t}^{(m)}]$.

In the MLE construction we need the derivatives of (11.34), of first order (an s-vector),

$$\frac{\partial}{\partial\theta} \ln \Lambda = \sum_{g\neq h} \int \frac{\partial}{\partial\theta} \ln \mu_{gh}(\tau, \theta)\{dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau\}\,, \qquad (11.35)$$

and of second order (an $s \times s$ matrix),

$$\begin{aligned}\frac{\partial^2}{\partial\theta\partial\theta'} \ln \Lambda &= \sum_{g\neq h} \int \Big\{ \frac{\partial^2}{\partial\theta\partial\theta'} \ln \mu_{gh}(\tau, \theta)\{dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau\} \\ &\quad - \frac{\partial}{\partial\theta} \ln \mu_{gh}(\tau, \theta) \frac{\partial}{\partial\theta'} \mu_{gh}(\tau, \theta) I_g(\tau)\, d\tau \Big\} \qquad (11.36)\end{aligned}$$

By (11.35) the MLE $\hat{\theta}$ is the solution to

$$\sum_{g\neq h} \int \frac{\partial}{\partial\theta} \ln \mu_{gh}(\tau, \theta)\{dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau\} \mid_{\theta=\hat{\theta}} = 0^{sx1}\,. \qquad (11.37)$$

A comment on the form of the likelihood (9.31): For each type of transition $g \to h$ introduce N_{gh}, the number of transitions of that type, and (if $N_{gh} > 0$) $T_{gh}^{(1)}, \ldots, T_{gh}^{(N_{gh})}$, the times when such transitions occurred. In terms of these quantities the log likelihood in (9.32) assumes the form

$$\sum_{g \neq h} \sum_{j=1}^{N_{gh}} \ln \mu_{gh}(T_{gh}^{(j)}, \hat{\theta}) - \sum_{g \neq h} \int \mu_{gh}(\tau, \hat{\theta}) \, I_g(\tau) d\tau \,,$$

and the ML equations (9.35) become

$$\sum_{g \neq h} \sum_{j=1}^{N_{gh}} \frac{\frac{\partial}{\partial \theta_i} \mu_{gh}(T_{gh}^{(j)}, \hat{\theta})}{\mu_{gh}(T_{gh}^{(j)}, \hat{\theta})} = \sum_{g \neq h} \int \frac{\partial}{\partial \theta_i} \mu_{gh}(\tau, \hat{\theta}) \, I_g(\tau) d\tau \,, \tag{11.38}$$

$i = 1, \ldots, s$. The form (11.38) is explicit and is, of course, the one we will work with when it comes to numerical computation of the MLE: The good thing about the form (9.32) is that it, by use of the counting processes, writes the sum on the left as a sum of contributions from all small time intervals. This is particularly useful in the derivation of the statistical properties of the MLE. (A similar remark could be made about the benefit of using the counting processes to define the payment stream for a general insurance policy in Section 7.5.)

Referring to Appendix D, the large sample distribution properties of the MLE are given by

$$\hat{\theta} \sim_{\mathrm{as}} \mathrm{N}(\theta, \Sigma(\theta)) \,, \tag{11.39}$$

where $\Sigma(\theta)$ is given by its inverse, the so-called *information matrix*,

$$\Sigma(\theta)^{-1} = -\mathbb{E}\left[\frac{\partial}{\partial \theta \partial \theta'} \ln \Lambda\right] . \tag{11.40}$$

Taking expectation in (11.36), noting that the terms

$$dN_{gh}(\tau) - \mu_{gh}(\tau, \theta) I_g(\tau) d\tau$$

have zero means, we obtain

$$\Sigma(\theta)^{-1} = \sum_{g \neq h} \int \frac{1}{\mu_{gh}(\tau, \theta)} \left(\frac{\partial}{\partial \theta} \mu_{gh}(\tau, \theta) \frac{\partial}{\partial \theta'} \mu_{gh}(\tau, \theta) \right) \sum_{m=1}^{n} p_g^{(m)}(\tau, \theta) d\tau \,. \tag{11.41}$$

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.

C. Estimating the parameters of a Gompertz-Makeham mortality law.

The actuarial office in a life insurance company is to estimate the mortality law governing the company's portfolio of term insurance policies. (The mortality law for the portfolio of life annuities may be different since people who (believe they) are in good health would probably find a pure survival benefit more useful and profitable than a pure death benefit. Thus, it seems appropriate to perform a separate mortality investigation for each line of life insurance business. Moreover, since mortality also depends on sex, the study would typically include only males or only females.)

Suppose the statistical data comprises n individuals who have been insured under the scheme during a certain period of time. For each individual No. m $(= 1, \ldots, n)$ there is a policy record with the following pieces of information:

– x_m, the age on entry into the study;
– y_m, the age on exit from the study;
– N_m, the number of deaths during the study (0 or 1).

Here x_m would typically be the age at issue of the policy. If $N_m = 1$, then y_m is the age at death, and if $N_m = 0$, then y_m is the age at the time of censoring, either at the term of the contract or upon termination of the study, whichever occurred first. In any case $y_m - x_m$ is the time spent under observation as alive during the study. With these definitions x_m takes the role of $\underline{t}^{(m)}$ in the general set-up and, for $N_m = 0$, y_m takes the role of $\bar{t}^{(m)}$.

The state space is now just $\mathcal{Z} = \{0, 1\}$ ("alive" and "dead"). Assume the mortality law is Gompertz-Makeham so that the mortality intensity at age t is

$$\mu(\tau, \theta) = \alpha + \beta e^{\gamma\tau},$$

with

$$\theta = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

We need the derivatives of the intensity w.r.t. all three parameters,

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\mu(\tau,\theta) &= 1\,,\\
\frac{\partial}{\partial\beta}\mu(\tau,\theta) &= e^{\gamma\tau}\,,\\
\frac{\partial}{\partial\gamma}\mu(\tau,\theta) &= \beta\, e^{\gamma\tau}\,\tau\,,
\end{aligned}$$

and their integrals with respect to time,

$$\begin{aligned}
\int_x^y \frac{\partial}{\partial\alpha}\mu(\tau,\theta)\,d\tau &= y - x\,,\\
\int_x^y \frac{\partial}{\partial\beta}\mu(\tau,\theta)\,d\tau &= \frac{e^{\gamma y} - e^{\gamma x}}{\gamma}\,,\\
\int_x^y \frac{\partial}{\partial\gamma}\mu(\tau,\theta)\,d\tau &= \beta\left(\frac{e^{\gamma y}y - e^{\gamma x}x}{\gamma} - \frac{e^{\gamma y} - e^{\gamma x}}{\gamma^2}\right).
\end{aligned}$$

The MLE equations (9.35) specialize to

$$\begin{aligned}
\sum_{m;\,N_m=1} \frac{1}{\hat\alpha + \hat\beta e^{\hat\gamma y_m}} &= \sum_m (y_m - x_m)\,,\\
\sum_{m;\,N_m=1} \frac{e^{\hat\gamma y_m}}{\hat\alpha + \hat\beta e^{\hat\gamma y_m}} &= \sum_m \frac{e^{\hat\gamma y_m} - e^{\hat\gamma x_m}}{\hat\gamma}\,,\\
\sum_{m;\,N_m=1} \frac{e^{\hat\gamma y_m} y_m}{\hat\alpha + \hat\beta e^{\hat\gamma y_m}} &= \sum_m \left(\frac{e^{\gamma y_m} y_m - e^{\gamma x_m} x_m}{\gamma} - \frac{e^{\gamma y_m} - e^{\gamma x_m}}{\gamma^2}\right).
\end{aligned}$$

To find the information matrix (9.38) we need the matrix with the products of the derivatives,

$$\frac{\partial}{\partial\theta}\mu(\tau,\theta)\frac{\partial}{\partial\theta'}\mu(\tau,\theta) = \begin{pmatrix} 1 & e^{\gamma\tau} & \beta e^{\gamma\tau}\tau \\ \cdot & e^{2\gamma\tau} & \beta e^{2\gamma\tau}\tau \\ \cdot & \cdot & \beta^2 e^{2\gamma\tau}\tau^2 \end{pmatrix} \tag{11.42}$$

(symmetric) and the probabilities $p_0^{(m)}(\tau,\theta)$, which are

$$\begin{aligned}
p_0^{(m)}(\tau,\theta) &= \exp\left(-\int_{x_m}^{\tau} (\alpha + \beta e^{\gamma\, s})\,ds\right)\\
&= \exp\left(-\alpha(\tau - x_m) - \beta\frac{e^{\gamma\tau} - e^{\gamma x_m}}{\gamma}\right), \qquad (11.43)
\end{aligned}$$

for $\tau \in [x_m, y_m]$ (the survival probability) and 0 otherwise. (There is only one kind of transition, from 0 to 1, and the summation over g, h in the information matrix (9.38) 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 θ.

11.3 Confidence regions

A. An asymptotic confidence ellipsoid. From the asymptotic normality of the MLE it follows that

$$(\hat{\theta}-\theta)'\Sigma^{-1}(\theta)(\hat{\theta}-\theta) \sim_{\mathrm{as}} \chi^2_s\,, \tag{11.44}$$

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 θ satisfying

$$(\theta-\hat{\theta})'\Sigma^{-1}(\theta)(\theta-\hat{\theta}) \leq \chi^2_{s,\,1-\varepsilon}\,. \tag{11.45}$$

The expression on the left here will typically be a complicated function of θ, and it is in general not easy to find the values of θ 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

$$\hat{\Sigma}\Sigma^{-1}(\theta) \to I\,. \tag{11.46}$$

Then it is easy to show that also the relation

$$(\theta-\hat{\theta})'\hat{\Sigma}^{-1}(\theta-\hat{\theta}) \leq \chi^2_{s,\,1-\varepsilon} \tag{11.47}$$

determines an asymptotic $1-\varepsilon$ confidence region. The relation (11.47) 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 θ 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 θ. Unfortunately, this is the case only in certain simple situations, typically when the state space $\mathcal{Z}$ 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 Σ, 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 $I_g^{(m)}(\tau)$ and put

$$\sum_{m=1}^{n} p_g^{(m)}(\tau) \approx I_g(\tau)\,. \tag{11.48}$$

B. Simultaneous confidence intervals. The confidence ellipsoid (11.47) can be resolved in simultaneous confidence intervals for all linear functions of θ in the following way. The Schwarz inequality says that for all vectors a and x in $\mathcal{R}^s$,

$$|a'x| \leq \sqrt{a'a\,x'x}\,,$$

with equality for $a = cx$. Thus, noting that the quadratic form on the left of (11.47) is $(\hat{\Sigma}^{-1/2}(\theta-\hat{\theta}))'\,(\hat{\Sigma}^{-1/2}(\theta-\hat{\theta}))$, the confidence statement can be cast equivalently as

$$|a'\hat{\Sigma}^{-1/2}(\theta-\hat{\theta})| \leq \sqrt{a'a\,\chi^2_{s,\,1-\varepsilon}}\,,\ \forall a\,. \tag{11.49}$$

Since $\hat{\Sigma}$ is of full rank, the vector $\hat{\Sigma}^{-1/2}a$ ranges through all of $\mathcal{R}^s$ as a ranges in $\mathcal{R}^s$. Thus, writing $a'a = (\hat{\Sigma}^{-1/2}a)'\hat{\Sigma}(\hat{\Sigma}^{-1/2}a)$, (11.49) is equivalent to

$$|a'(\theta-\hat{\theta})| \leq \sqrt{a'\hat{\Sigma}a\,\chi^2_{s,\,1-\varepsilon}}\,,\ \forall a\,,$$

that is,

$$a'\theta \in [a'\hat{\theta} - \sqrt{\chi^2_{s,\,1-\varepsilon}a'\hat{\Sigma}a}\,,\ a'\hat{\theta} + \sqrt{\chi^2_{s,\,1-\varepsilon}a'\hat{\Sigma}a}]\,,\ \forall a\,. \tag{11.50}$$

The intervals in (11.50) are (asymptotic) simultaneous confidence intervals for all linear functions of θ in the sense that the probability is at least $1-\varepsilon$ that they all hold true.

C. Confidence band for the G-M mortality intensity. Returning to the mortality study example in Paragraph 11.2.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 (11.42) and (11.42), and the appropriate variance matrix Σ is obtained by inverting the upper left 2×2 block in the information matrix defined by (11.41), (11.42), and (11.43).

From (11.50) we obtain simultaneous confidence intervals for all $\mu(\tau) = \alpha + \beta e^{\gamma\tau}$, constituting a confidence band in the space of mortality intensity functions;

$$\mu(\tau) \in [\hat{\alpha} + \hat{\beta} e^{\gamma\tau} - \sqrt{\chi^2_{2,\,1-\varepsilon}\hat{\sigma}_\tau}\,,\, \hat{\alpha} + \hat{\beta} e^{\gamma\tau} - \sqrt{\chi^2_{2,\,1-\varepsilon}\hat{\sigma}_\tau}]\,,\, \forall \tau > 0\,, \tag{11.51}$$

where

$$\hat{\sigma}_\tau = (1,\, e^{\gamma\tau})\hat{\Sigma} \begin{pmatrix} 1 \\ e^{\gamma\tau} \end{pmatrix}.$$

11.4 More on simultaneous confidence intervals

A. Simultaneous confidence intervals based on a confidence region.
Let $X \in \mathcal{X}$ be some observation(s) with distribution P_θ, $\theta \in \Theta$, some s-dimensional set. A $1-\varepsilon$ confidence region for θ is a function C from $\mathcal{X}$ to the set of subsets of Θ such that the random set $\mathcal{C} = C(X)$ satisfies

$$P_\theta\{\theta \in \mathcal{C}\} = 1 - \varepsilon, \;\; \forall \theta \in \Theta\,.$$

From a confidence region we readily obtain confidence intervals for the values $g(\theta)$ of all real-valued functions $g : \Theta \to \mathcal{R}$ as follows. For a fixed g, define the random variables $\underline{g} = \inf_{\theta \in \mathcal{C}} g(\theta)$ and $\overline{g} = \sup_{\theta \in \mathcal{C}} g(\theta)$. Clearly, $\theta \in \mathcal{C}$ implies $\underline{g} \le g(\theta) \le \overline{g}$ for all g, hence

$$P_\theta\{\underline{g} \le g(\theta) \le \overline{g}\,,\; \forall g\} \ge 1 - \varepsilon\,,\;\; \forall \theta \in \Theta\,.$$

Thus, by stating that

$$g(\theta) \in [\,\underline{g},\, \overline{g}\,]\,,\;\; \forall g\,, \tag{11.52}$$

all statements hold true simultaneously with a probability no less than $1-\varepsilon$. In this sense the intervals in (11.52) are simultaneous confidence intervals with (simultaneous) confidence level $1-\varepsilon$. A practical consequence is that we are allowed to snoop around in the data set, looking for possible interesting effects (within the model), and still keep control of the probability of making false statements.

B. Confidence ellipsoid for a normal mean; Scheffé intervals.
Let $\hat{\theta}$ is an estimator of an s-dimensional parameter vector θ and assume that

$$\hat{\theta} \sim N(\theta, \Sigma)\,,$$

with Σ known. A $1-\varepsilon$ confidence region of θ is the ellipsoid defined by (3.2):

$$\mathcal{C} = \{\theta\,;(\theta - \hat{\theta})'\Sigma^{-1}(\theta - \hat{\theta}) \le \chi^2_{s,\,1-\varepsilon}\}\,.$$

To construct the confidence interval (11.52) for a specific function $g(\theta)$, we are left with the mathematical problem of finding the extrema of g over the ellipsoid, which may be a difficult task. For linear functions it is simple, however, and it goes by the the technique in Paragraph 3B, which is due to Scheffé. The simultaneous confidence intervals for linear functions $a'\theta = \sum_{p=1}^{s} a_p\theta_p$ are, with a bit sloppy notation,

$$a'\theta \in a'\hat{\theta} \pm \sigma_a \sqrt{\chi^2_{s,\,1-\varepsilon}}\,,\, \forall a \in \mathcal{R}^s\,,$$

where

$$\sigma_a^2 = a'\,\Sigma\,a$$

is the variance of the point estimator $a'\hat{\theta}$.

C. Narrowing the confidence intervals.
Generally speaking, in the presence of uncertainty, the price we have to pay for making many safe statements is that each individual statement has to be vague. In our situation this general truth takes a very manifest form: for a fixed confidence level $1-\varepsilon$ the lengths of the confidence intervals increase with the dimension of θ since $\chi^2_{s,\,1-\varepsilon}$ is an increasing function of s (why?).

We can gain precision in terms of lengths of the intervals by reducing the number of statements we want to make. Suppose we are only interested in drawing inferences about linear combinations of r linearly independent functions $b_j'\theta$, $j = 1, \ldots, r$, $r < s$. Thus, putting $B = (b_1, \ldots, b_r)$, an $s \times r$ matrix, we are only interested in linear functions $a'\theta$ with $a = Bc$ for some r-vector c, that is, $a \in \mathcal{R}(B)$, the r-dimensional linear space spanned by the b_j. Then, start from

$$B'\hat{\theta} \sim N(B'\theta, B'\Sigma B)\,,$$

and apply the results above to obtain that simultaneous confidence intervals for all linear functions of $B'\,\theta$ are given by

$$c'B'\theta \in c'B'\hat{\theta} \pm \sqrt{c'B'\Sigma Bc\,\chi^2_{r,\,1-\varepsilon}}\,,\ \forall\, c \in \mathcal{R}^r\,, \tag{11.53}$$

or, equivalently,

$$a'\theta \in a'\hat{\theta} \pm \sigma_a\,\sqrt{\chi^2_{r,\,1-\varepsilon}}\,,\ \forall a \in \mathcal{R}(B)\,.$$

It is seen that, by reducing the "dimension of our statements" from s to r, we have gained a reduction of the lengths of the confidence intervals by a factor $\sqrt{\chi^2_{r,\,1-\varepsilon}/\chi^2_{s,\,1-\varepsilon}}$.

B. Bonferroni intervals for a finite number of parameter functions.
Let $g_j(\theta)$, $j = 1, \ldots, r$, be a finite number of real-valued parameter functions, and assume that for each $g_j(\theta)$ we have constructed an individual confidence interval $[\,\underline{g}_j\,, \overline{g}_j\,]$ with level $1-\varepsilon_j$. Thus, denoting the event $g_j(\theta) \in [\underline{g}_j\,, \overline{g}_j]$ by A_j, we have $P_\theta(A_j) \geq 1-\varepsilon_j$ for each j. The probability that all A_j hold true at the same time is

$$\begin{aligned} P_\theta\left(\cap_j A_j\right) &= 1 - P_\theta\left(\cup_j A_j^c\right) \\ &\geq 1 - \sum_{j=1}^r P_\theta(A_j^c) \\ &\geq 1 - \sum_{j=1}^r \varepsilon_j\,. \end{aligned}$$

It follows that the intervals taken together are simultaneous confidence intervals with simultaneous confidence level no less than $1-\varepsilon$, where $\varepsilon = \sum_{j=1}^r \varepsilon_j$.

This simple device, due to Bonferroni, represents an attractive alternative to the approach in Paragraphs A – C in situations where we take interest only in a finite number of parameter functions. It turns out that the Bonferroni intervals often are shorter than the Scheffé intervals, which aim at an infinite number of parameter functions. Bonferroni intervals with simultaneous confidence level $1-\varepsilon$ for q linear functions $a_j'\theta$, $j = 1, \ldots, q$, are

$$a_j'\theta \in a_j'\hat{\theta} \pm \sigma_{a_j}\,\sqrt{\chi^2_{1,\,1-\varepsilon/q}}\,, \tag{11.54}$$

(Note that $\sqrt{\chi^2_{1,\,1-\varepsilon/q}}$ is just the $(1-\varepsilon/2q)$-fractile of the standard normal distribution, so we recognize (11.54) as the traditional individual $1-\varepsilon/q$ confidence interval for a normal mean.) Let r $(\leq s)$ be the dimension of the space spanned by the coefficient vectors a_j. The corresponding Scheffé intervals based on (11.53) are

$$a_j'\theta \in a_j'\hat{\theta} \pm \sigma_{a_j}\,\sqrt{\chi^2_{r,\,1-\varepsilon}}\,.$$

The ratio of the length of the intervals by the two constructions is $B/S(q, r, \varepsilon) = \sqrt{\chi^2_{1,\,1-\varepsilon/q}/\chi^2_{r,\,1-\varepsilon}}$. Clearly, the ratio decreases with r. It increases with q, and (for $r > 1$) it starts from $q = 1$

with a value smaller than 1 and tends to infinity as q grows. There will be some value $q(r, \varepsilon)$ such that the ratio is ≤ 1 for $q \leq q(r, \varepsilon)$. Inspection of a table of the chi-square fractiles shows e.g. that $q(2, 0.025) = 5$, $q(4, 0.1) = 20$, $q(6, 0.25) = 50$ (approximately).

E. Confidence intervals based on consistent and asymptotically normal point estimators.
The results and considerations in the previous paragraphs carry over to the situation in Section 3, where (3.4) formed the basis for simultaneous inference.

Suppose we are interested in functions of a set of parameter functions $f_j(\theta)$, $j = 1, \dots, r$, $r < s$. Put $f = (f_1, \dots, f_r)'$. If f is continuously differentiable, first order Taylor expansion gives

$$f(\hat{\theta}) \sim_{\text{as}} N\left(f(\theta), \, \Sigma_f(\theta)\right) \, ,$$

with

$$\Sigma_f(\theta) = Df(\theta)' \, \Sigma(\theta) \, Df(\theta) \, ,$$

and

$$Df(\theta) = \frac{\partial}{\partial \theta} f(\theta) \, ,$$

an $s \times r$ matrix. We obtain the asymptotic confidence ellipsoid

$$\mathcal{C}_f = \{ f \, ; (f - f(\hat{\theta}))' \hat{\Sigma}_f^{-1} (f - f(\hat{\theta})) \leq \chi^2_{r,\,1-\varepsilon} \} \, ,$$

where $\hat{\Sigma}_f$ is some consistent estimator of Σ_f in the sense of (3.3), e.g. $\hat{\Sigma}_f = \Sigma_f(\hat{\theta})$. Asymptotic Scheffé intervals for all functions $g(\theta) = h(f(\theta))$, with h real-valued and continuously differentiable, are

$$g(\theta) \in g(\hat{\theta}) \pm \hat{\sigma}_g \sqrt{\chi^2_{r,\,1-\varepsilon}} \, ,$$

where

$$\hat{\sigma}_g = \frac{\partial}{\partial \theta'} g(\theta) \, \Sigma(\theta) \, \frac{\partial}{\partial \theta} g(\theta) \, |_{\theta = \hat{\theta}} \, .$$

Asymptotic Bonferroni intervals for a finite collection of functions g_j, $j = 1, \dots, q$, are obtained upon replacing r and ε with 1 and ε / q.

F. The G-M mortality intensity revisited.
The confidence intervals (3.8) are infinitely many, so Bonferroni ideas cannot help here. If also c is to be estimated, we obtain asymptotic confidence intervals by the device above. The same goes for any function of actuarial relevance, like the reserve of a life insurance or a portfolio of insurances. Think of examples.
Returning to the mortality study example in Paragraph 2C, let γ 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 eqnarrays (11.42) and (11.42), and the appropriate variance matrix Σ is obtained by inverting the upper left 2×2 block in the information matrix defined by (11.41), (11.42), and (11.43).

From (11.50) we obtain simultaneous confidence intervals for all $\mu(\tau) = \alpha + \beta e^{\gamma \tau}$, constituting a confidence band in the space of mortality intensity functions;

$$\mu(\tau) \in [\hat{\alpha} + \hat{\beta} e^{\gamma \tau} - \sqrt{\chi^2_{2,\,1-\varepsilon} \hat{\sigma}_\tau} \, , \, \hat{\alpha} + \hat{\beta} e^{\gamma \tau} - \sqrt{\chi^2_{2,\,1-\varepsilon} \hat{\sigma}_\tau}] \, , \, \forall \tau > 0 \, ,$$

where

$$\hat{\sigma}_\tau = (1, \, e^{\gamma \tau}) \hat{\Sigma} \begin{pmatrix} 1 \\ e^{\gamma \tau} \end{pmatrix} \, .$$

11.5 Piecewise constant intensities

A. Piecewise constant intensities. Let $0 = t_0 < t_1 < \cdots < \cdots < t_r = \bar{t}$ be some finite partition of the time interval $[0, \bar{t}]$, and assume that the intensities are step functions of the form

$$\begin{aligned} \mu_{gh}(\tau) &= \mu_{gh,q}\,,\ \tau \in [t_{q-1}, t_q)\ ,\ q = 1, \ldots, r\,, \\ &= \sum_{q=1}^{r} 1_{[t_{q-1},t_q)}(\tau)\mu_{gh,q}\,, \end{aligned} \tag{11.55}$$

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 (11.34) now becomes

$$\ln \Lambda = \sum_{g \neq h} \sum_{q=1}^{r} \{\ln \mu_{gh,q} N_{gh,q} - \mu_{gh,q} W_{g,q}\}\,, \tag{11.56}$$

where

$$N_{gh,q} = \int_{t_{q-1}}^{t_q} dN_{gh}(\tau)\,, \tag{11.57}$$

$$W_{g,q} = \int_{t_{q-1}}^{t_q} I_g(\tau) d\tau\,, \tag{11.58}$$

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:

$$\frac{\partial}{\partial \mu_{gh,q}} \ln \Lambda = \frac{1}{\mu_{gh,q}} N_{gh,q} - W_{g,q}\,, \tag{11.59}$$

$$\frac{\partial^2}{\partial \mu_{gh,q} \partial \mu_{g'h',q'}} \ln \Lambda = -\delta_{ghq,g'h'q'} \frac{1}{\mu_{gh,q}^2} N_{gh,q}\,. \tag{11.60}$$

It follows from (11.59) that the MLE is

$$\hat{\mu}_{gh,q} = \frac{N_{gh,q}}{W_{g,q}}\,, \tag{11.61}$$

an O-E rate of the same kind as in the simple model of 11.2.B. Noting that, by (11.57),

$$\mathbb{E}[N_{gh,q}] = \mu_{gh,q} \int_{t_{q-1}}^{t_q} \sum_{m=1}^{n} p_g^{(m)}(\tau) d\tau\,,$$

we obtain from (11.60) that the asymptotic variance matrix becomes

$$\Sigma(\theta) = \mathrm{Diag}\left(\ldots, \frac{\mu_{gh,q}}{\int_{t_{q-1}}^{t_q} \sum_{m=1}^{n} p_g^{(m)}(\tau) d\tau}, \ldots\right)\,, \tag{11.62}$$

a diagonal matrix, implying that the estimators of the $\mu_{gh,q}$ are asymptotically independent.

An estimator of Σ is obtained upon replacing the parameter functions appearing on the right of (11.62) by their straightforward estimators: put $\mu_{gh,q} \approx \hat{\mu}_{gh,q}$ defined by (11.61) and, by the device (11.48),

$$\int_{t_{q-1}}^{t_q} \sum_{m=1}^{n} p_g^{(m)}(\tau) d\tau \approx \int_{t_{q-1}}^{t_q} I_g(\tau) d\tau = W_{g,q}\,,$$

to obtain

$$\hat{\Sigma} = \mathrm{Diag}\left(\ldots, \frac{N_{gh,q}}{W_{g,q}^2}, \ldots\right)\,. \tag{11.63}$$

C. Smoothing O-E rates. The MLE of the intensity function is obtained upon inserting the estimators (11.61) in (11.55). 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 11.2 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 τ_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 θ by θ^* minimizing

$$(\mu(\theta) - \hat{\mu})' A(\mu(\theta) - \hat{\mu}) = \sum_{pq} a_{pq}(\mu(\tau_p, \theta) - \hat{\mu}_p)(\mu(\tau_q, \theta) - \hat{\mu}_q)\,. \tag{11.64}$$

If the intensity is a linear function of θ (like in the G-M study with known γ),

$$\mu(\theta) = Y(\tau)\theta\,, \tag{11.65}$$

then

$$\hat{\theta} = (Y'AY)^{-1} Y'A\hat{\mu}\,. \tag{11.66}$$

The asymptotic variance of $\hat{\theta}$ is $(Y'AY)^{-1}Y'A\Sigma(\theta)AY(Y'AY)^{-1}$. By the Gauss-Markov theorem it is minimized by taking $A = \Sigma(\theta)^{-1}$, and the minimum is $(Y'\Sigma(\theta)^{-1}Y)^{-1}$. Thus, asymptotically the best choice of A is $\hat{\Sigma}^{-1}$, where $\hat{\Sigma}$ is some estimate of Σ satisfying (11.46). Write $\Sigma = \Sigma(\theta)$. The symmetric, pd matrix Σ has a symmetric pd square root W such that $\Sigma = W^2$.

$$\begin{aligned}
\Delta &= (Y'AY)^{-1}Y'A\Sigma AY(Y'AY)^{-1} - (Y'\Sigma^{-1}Y)^{-1} \\
&= (Y'AY)^{-1}Y'AW\left[I - WY((WY)'(WY))^{-1}(WY)'\right] WAY(Y'AY)^{-1} \\
&= (Y'AY)^{-1}Y'AW\,H\,WAY(Y'AY)^{-1}\,.
\end{aligned}$$

where

$$H = I - P(P'P)^{-1}P'\,.$$

The matrix $H = I - P(P'P)^{-1}P'$ is symmetric, $H = H'$, and idempotent, $H^2 = H$, Thus, $H = H'H$ and

$$\Delta = \left(HWAY(Y'AY)^{-1}\right)' \left(HWAY(Y'AY)^{-1}\right)$$

which is indeed pd.

More on analytic graduation - the G-M example: Let us focus on one intensity that is to be graduated and, to fix ideas, assume it is the mortality intensity in the simple model with two states 'alive' and 'dead'..

The first step is to assume that the intensity is piece-wise constant:

$$\mu(t) = \mu_q\,, \qquad q - 1 \le t < q\,, j = 1, 2, \ldots$$

The log likelihood (9.53) is

$$\ln \Lambda = \sum_q (\ln \mu_q N_q - \mu_q W_q)\,,$$

where N_q and W_q are, respectively, the number of deaths and the total time spent alive in the age interval $[q-1, q)$.

Each μ_q is a parameter which is functionally unrelated to all the others, so there are many parameters in this model! For instance, if we are interested in mortality up to age 100 and have data in the age range from 0 to 100, there are 100 parameters, which is quite a lot. Remember, however, that this model is just a first step in a two-stage procedure where the second step is to graduate (smooth) the ML estimators resulting from the present naive model with piece-wise constant intensity.

The ML estimators are the occurrence-exposure rates,

$$\hat{\mu}_q = \frac{N_q}{W_q}\,,$$

which are well defined for all q such that $W_q > 0$ (i.e. in age intervals where there were survivors exposed to risk of death). The $\hat{\mu}_q$ are asymptotically (as n increases) normally distributed, mutually independent, unbiased, and with variances given by

$$\sigma_q^2 = \mathrm{as.}\mathbb{V}[\hat{\mu}_q] = \frac{\mu_q}{\mathbb{E}[W_q]}\,, \tag{11.67}$$

where the expected exposure is

$$\mathbb{E}[W_q] = \sum_{m=1}^{n} \int_{q-1}^{q} p^{(m)}(\tau)\, d\tau\,,$$

$p^{(m)}(\tau)$ being the probability that individual No. m is alive and under observation at time τ.

The variance σ_q^2 is inversely proportional to the corresponding expected exposure. In the present simple model, with only one intensity of transition from the state 'alive' to the absorbing state 'dead', we find explicit expressions for the expected exposure.

For instance, suppose we have observed each individual life from birth until death or until attained age 100, whichever occurs first (i.e. censoring at age 100). Then, for $\tau \in [q-1, q)$ with $q = 1, \ldots, 100$, we have

$$\begin{aligned} p^{(m)}(\tau) &= \exp\left(-\int_0^{\tau} \mu(s)\, ds\right) \\ &= \exp\left(-\sum_{p=1}^{q-1} \mu_p - (\tau - (q-1))\,\mu_q\right)\,, \end{aligned} \tag{11.68}$$

hence

$$\begin{aligned} \mathbb{E}[W_q] &= n \int_{q-1}^{q} \exp\left(-\sum_{p=1}^{q-1} \mu_p - (\tau - (q-1))\,\mu_q\right) d\tau \\ &= n \exp\left(-\sum_{p=1}^{q-1} \mu_p\right) \frac{1 - \exp(-\mu_q)}{\mu_q}\,, \end{aligned}$$

and

$$\sigma_q^2 = \frac{1}{n}\, \frac{\mu_q}{\exp\left(-\sum_{p=1}^{q-1} \mu_p\right)\,(1 - \exp(-\mu_q))}\,. \tag{11.69}$$

You should look at other censoring schemes and discuss the impact of censoring on the variance. Take e.g. the case where person No m enters at age z_m and is observed until death or age 100, whichever occurs first (all z_m less than 100).

Estimators $\hat{\sigma}_q^2$ of the variances are obtained upon replacing the μ_j in (G.55) by the estimators $\hat{\mu}_j$. Simpler estimators are obtained by just replacing μ_q and $\mathbb{E}[W_q]$ in (G.53) with their straightforward empirical counterparts: $\hat{\sigma}_q^2 = \hat{\mu}_q / W_q = N_q / W_q^2$.

Now to graduation. The occurrence-exposure rates will usually have a ragged appearance. Assuming that the real underlying mortality intensity is a smooth function, we will therefore fit a suitable function to the occurrence-exposure rates. Suppose we assume that the true mortality rate is a Gompertz Makeham function, $\mu(t) = \alpha + \beta e^{\gamma t}$. Then, take some representative age τ_q (typically $\tau_q = q - 0.5$) in each age interval and fit the parameters α, β, γ by minimizing a weighted sum of squared errors

$$Q = \sum_q a_q \left(\hat{\mu}_q - \alpha - \beta e^{\gamma \tau_q}\right)^2 .$$

This is a matter of non-linear regression. The optimal weights a_q are the inverse of the variances, but since these are unknown, we plug in the estimators and use $a_q = 1/\hat{\sigma}_q^2$.

The minimizing values α^*, β^*, and γ^* are obtained by differentiating Q with respect to each of the three parameters and setting the derivatives equal to 0. The derivatives are:

$$\begin{aligned}
\frac{\partial}{\partial \alpha} Q &= \sum_q a_q 2 \left(\hat{\mu}_q - \alpha - \beta e^{\gamma \tau_q}\right)(-1)\,, \\
\frac{\partial}{\partial \beta} Q &= \sum_q a_q 2 \left(\hat{\mu}_q - \alpha - \beta e^{\gamma \tau_q}\right)(-e^{\gamma \tau_q})\,, \\
\frac{\partial}{\partial \gamma} Q &= \sum_q a_q 2 \left(\hat{\mu}_q - \alpha - \beta e^{\gamma \tau_q}\right)(-\beta e^{\gamma \tau_q} \tau_q)\,.
\end{aligned}$$

Thus α^*, β^*, and γ^* are the solution to the equations

$$\begin{aligned}
\sum_q a_q \left(\hat{\mu}_q - \alpha^* - \beta^* e^{\gamma^* \tau_q}\right) &= 0\,, \\
\sum_q a_q \left(\hat{\mu}_q - \alpha^* - \beta^* e^{\gamma^* \tau_q}\right) e^{\gamma^* \tau_q} &= 0\,, \\
\sum_q a_q \left(\hat{\mu}_q - \alpha^* - \beta^* e^{\gamma^* \tau_q}\right) e^{\gamma^* \tau_q} \tau_q &= 0\,.
\end{aligned}$$

This is in general a set of non-linear equations that does not allow of an explicit solution. Actually these equations are just as involved as the maximum likelihood equations in Paragraph 9.2C above, which is disappointing since the two-stage procedure considered here was supposed to be simpler. (Occurrence-exposure rates are easy to find and they are asymptotically independent, which makes it easy to find their asymptotic variances. The graduation is, however, messy.)

Suppose now that γ is taken to be known. Then only the two first equations above are relevant and they reduce to

$$\begin{aligned}
\sum_q a_q \, \alpha^* + \sum_q a_q e^{\gamma \tau_q} \, \beta^* &= \sum_q a_q \hat{\mu}_q \,, \\
\sum_q a_q e^{\gamma^* \tau_q} \, \alpha^* - \sum_q a_q e^{2\gamma^* \tau_q} \, \beta^* &= \sum_q a_q \hat{\mu}_q e^{\gamma^* \tau_q} \,.
\end{aligned}$$

This is a linear system of equations with an explicit solution, which the industrious reader certainly will find.

11.6 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 11.2 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 τ. (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 11.2.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

$$\sum_{m=1}^{n} p_g^{(m)}(\tau) = n\, p_g^{(1)}(\tau)\,,\ g = 0, \ldots, J,$$

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 11.2.C is conducted cross-sectionally throughout a calendar period of duration $\bar{t}$, and that it comprises n individuals at ages $\underline{t}^{(m)}$, $m = 1, \ldots, n$, at the beginning of the study. In this case the factor depending on the design in the information matrix is

$$\sum_{m=1}^{n} p^{(m)}(\tau) = \sum_{m=1}^{n} 1_{\,[\,\underline{t}^{(m)}, \underline{t}^{(m)} + \bar{t}\,]}(\tau) \exp\left(-\alpha(\tau - \underline{t}^{(m)}) - \beta \frac{e^{\gamma \underline{t}^{(m)}}(e^{\gamma\tau} - 1)}{\gamma}\right).$$

Chapter 12

Heterogeneity models

12.1 The notion of heterogeneity – a two-stage model

The life length T of an individual depends on a number of factors like biological inheritance (some are strong and healthy, others are weak and sickly), occupation (mining and forestry have higher accident rates than office work), leisure activities (mountain climbing is more wholesome but also more dangerous than philately), nutrition (see the weekly magazines for current wisdom), smoking and drinking habits. The list might be extended endlessly. Let all such individual characteristics be represented by a parameter θ, which may be quite complex, possibly of large dimension comprising numbers and strings of letters. The dependence of T on these characteristics is accounted for by letting the probability law of T depend on θ. Thus, write $F_\theta(t)$ for the probability that a person with characteristics θ dies within age t and $\bar{F}_\theta(t)$ for the corresponding survival probability. Assuming that F_θ possesses a density f_θ, the force of mortality is $\mu_\theta(t) = f_\theta(t)/\bar{F}_\theta(t)$, and $\bar{F}_\theta(t) = \exp\left(-\int_0^t \mu_\theta(s)\,ds\right)$, $\bar{F}_\theta(t\,|\,x) = \exp\left(-\int_x^{x+t} \mu_\theta(s)\,ds\right)$, and so on.

Apparently the aggregate mortality law treated in Section *** conflicts with the myopic viewpoint taken here, but this is really not so: the aggregate law describes the mortality pattern in the population as a whole and represents the prospects of longevity for a person when nothing is known as to his/her personal characteristics. To make this precise we let the individual characteristics of a randomly chosen newly born be a random element Θ with a distribution G, and the conditional distribution of the life length for fixed $\Theta = \theta$ is F_θ. If the value of Θ is observed to be θ, then F_θ is the relevant basis for making predictions about T. If Θ is not observed, then predictions about T can only be based on the unconditional distribution function

$$F(t) = \mathbb{E}[F_\Theta(t)] = \int F_\theta(t)\,dG(\theta) \tag{12.1}$$

or any of the equivalent functions in (3.2) – (3.4), which now assume the forms

$$\begin{aligned}
\bar{F}(t) &= \int \bar{F}_\theta(t)\,dG(\theta), && (12.2)\\
f(t) &= \int f_\theta(t)\,dG(\theta), && (12.3)\\
\mu(t) &= f(t)/\bar{F}(t) = \frac{\int f_\theta(t)\,dG(\theta)}{\int \bar{F}_\theta(t)\,dG(\theta)} = \frac{\int \mu_\theta(t)\bar{F}_\theta(t)\,dG(\theta)}{\int \bar{F}_\theta(t)\,dG(\theta)} && (12.4)\\
&= \mathbb{E}[\mu_\Theta(t)\,|\,T>t]\,. && (12.5)
\end{aligned}$$

Formula (12.5) calls for a comment. The joint distribution of T and Θ is given by

$$\mathbb{P}\,[T \in A, \Theta \in B] = \int_B \int_A f_\theta(t)\,dt\,dG(\theta) = \int_B F_\theta[A]\,dG(\theta).$$

The marginal distribution of T is given by

$$\mathbb{P}[T \in A] = F[A] = \int F_\theta[A]\,dG(\theta)\,,$$

where the integral sign without indication of the area signifies integration over the entire range of the variable. The conditional distribution of Θ, given $T \in A$, is

$$G[B \,|\, T \in A] = \frac{\int_B F_\theta[A]\,dG(\theta)}{\int F_\theta[A]\,dG(\theta)} \tag{12.6}$$

or, in terms of the generalized density,

$$dG(\theta \,|\, T \in A) = \frac{F_\theta[A]\,dG(\theta)}{\int F_{\theta'}[A]\,dG(\theta')}\,. \tag{12.7}$$

In particular, the conditional distribution of Θ, given $T > t$, is given by

$$dG(\theta \,|\, T > t) = \frac{\bar{F}_\theta(t)\,dG(\theta)}{\int \bar{F}_{\theta'}(t)\,dG(\theta')}\,, \tag{12.8}$$

and the last expression in (12.4) is recognized as the expected value of $\mu_\Theta(t)$ with respect to this distribution. Inserting $A = [t, t + dt]$ and $F_\theta[A] = f_\theta\,dt$ in (12.7) gives (somewhat informally)

$$dG(\theta \,|\, T = t) = \frac{f_\theta\,dG(\theta)}{\int f_{\theta'}(t)\,dG(\theta')}\,. \tag{12.9}$$

The probabilities in (12.7) have a straightforward interpretation as proportions in a cohort of individuals born at the same time. They cannot in general be interpreted as proportions in a given population, and the reason for this is that no assumptions have been made as to the birth rates that would govern the development of the age composition of the population over time. It is certainly true that (12.8) is the distribution of Θ amongst those who, at a given moment, are exactly t years old in the population. It is not the distribution of Θ amongst those who, at a given moment, are t years or older (as the conditioning on $T > t$ might suggest). Likewise, (12.9) is the distribution of Θ amongst those who are known to have died at age t in the past.

The conditional probability $\mathbb{P}[T \in A' \,|\, T \in A]$ may be expressed as

$$\begin{aligned} F[A' \,|\, A] &= \frac{F[A' \cap A]}{F[A]} \\ &= \frac{\int F_\theta[A' \cap S]\,dG(\theta)}{F[A]} \\ &= \int \frac{F_\theta[A' \in A]}{F_\theta[A]}\,\frac{F_\theta[A]\,dG(\theta)}{F[A]} \\ &= \int F_\theta[A' \,|\, A]\,dG(\theta \,|\, T \in A). \end{aligned} \tag{12.10}$$

Formula (12.10) says that, when $T \in A$ is given, the probability of $T \in A'$ is to be formed in the usual way, by taking the average of the probability of $T \in A'$ for fixed Θ over the distribution of Θ, all distributions being conditional on $T \in A$. In particular,

$$\bar{F}(t \,|\, x) = \int \bar{F}_\theta(t \,|\, x)\,dG(\theta \,|\, T > x), \tag{12.11}$$

which resembles (12.2), only that the distributions of T and Θ are updated in regard of the information that the person has survived to age x.

Those who adhere to a deterministic world picture would presumably fancy the idea that T could be exactly determined if only the individual and its surroundings could be sufficiently accurately described – down to the atoms. Then T would be just a function $T(\Theta)$ of Θ, and F_θ would reduce to a one-point distribution in $T(\theta)$ which need no longer be explicated in the model since the mortality law would simply be GT^{-1}. The model formulation (12.1) with F_θ a non-degenerate distribution, expresses the point of view that something remains unexplained beyond Θ. Two interpretations are possible. Either that there exists such a thing as pure randomness in the world, or that not all explanatory factors are included in Θ. Now, leaving such speculations to the philosophers, let us pursue the chosen approach.

12.2 The proportional hazard model

The perhaps simplest means of describing mortality variations is the so-called proportional hazard model, which specifies that Θ is a positive random variable, and

$$\mu_\theta(t) = \theta\mu^\circ(t), \tag{12.12}$$

where $\mu^\circ(t)$ is some "baseline" force of mortality. According to this assumption, the mortality pattern is basically the same for all people, only that some die "faster" than others at all ages. It does not allow for the possibility that e.g. some people have mortality above the average in the youth and below the average in the old age.

Introduce the *accumulated baseline intensity* at age t,

$$W(x) = \int_0^x \mu^\circ(s)\, ds. \tag{12.13}$$

Under the assumption (12.11), the conditional survival function by fixed $\Theta = \theta$ is

$$\bar{F}_\theta(t) = e^{-\theta W(t)}, \tag{12.14}$$

and the conditional density is

$$f_\theta(t) = \theta\mu^\circ(t)\, e^{-\theta W(t)}. \tag{12.15}$$

The functions in (12.2) – (12.5) become

$$\begin{aligned}
\bar{F}(t) &= \int e^{-\theta W(t)}\, dG(\theta), & (12.16)\\
f(t) &= \mu^\circ(t) \int \theta\, e^{-\theta W(t)}\, dG(\theta), & (12.17)\\
\mu(t) &= \mu^\circ(t) \frac{\int \theta\, e^{-\theta W(t)}\, dG(\theta)}{\int e^{-\theta W(t)}\, dG(\theta)} = \mu^\circ(t)\, E(\Theta \,|\, T > t). & (12.18)
\end{aligned}$$

Although the unconditional distribution F is the only thing that matters when Θ is unobservable, the present two-stage model is of interest also in such circumstances. In the first place it is of theoretical interest due to its explanatory import, and in the second place it is of practical value since it produces candidates for suitable aggregate laws. Assume now that G is $G_{\gamma,\delta}$, the gamma distribution with shape parameter γ and scale parameter δ^{-1}, whose density is

$$g_{\gamma,\delta}(\theta) = \frac{\delta^\gamma}{\Gamma(\gamma)}\, \theta^{\gamma-1}\, e^{-\theta\delta}, \theta > 0, \tag{12.19}$$

and γ, $\delta > 0$. Here Γ is the gamma function

$$\Gamma(\gamma) = \int_0^\infty t^{\gamma-1}\, e^{-t}\, dt\,,$$

which satisfies

$$\Gamma(\gamma) = (\gamma - 1)\,\Gamma(\gamma - 1), \gamma > 1,$$

(easy to prove by integration by parts) and $\Gamma(1) = 1$. It is immediately seen that

$$\begin{aligned}
\mathbb{E}\left(\Theta^r\, e^{-\Theta s}\right) &= \frac{\delta^\gamma}{\Gamma(\gamma)}\, \frac{\Gamma(r+\gamma)}{(s+\delta)^{r+\gamma}}, r > -\gamma, s > -\delta,\\
&= \frac{\delta^\gamma (r+\gamma-1)^{(r)}}{(s+\delta)^{r+\gamma}}, r = 0, 1, \cdots, s > -\delta. & (12.20)
\end{aligned}$$

In particular, all moments of Θ exist, and

$$\mathbb{E}\,\Theta = \gamma/\delta, \tag{12.21}$$

$$\mathbb{V}\,\Theta = \gamma/\delta^2. \tag{12.22}$$

The role of the scale parameter becomes clear by substituting $\theta' = \theta\delta$ in the integral

$$G_{\gamma,\delta}(\theta) = \int_0^\theta \frac{\delta^\gamma}{\Gamma(\gamma)}\,\theta^{\gamma-1}\,e^{-\theta\delta}\,d\theta = \int_0^{\theta\delta} \frac{1}{\Gamma(\gamma)}\,\theta'^{-\gamma-1}\,e^{-\theta'}\,d\theta' = G_{\gamma,1}(\theta\delta), \tag{12.23}$$

which is an increasing function of δ.

By use of (12.20), it is readily seen that the present case (12.16)—(12.18) specialize to

$$\bar{F}(t) = \left(\frac{\delta}{W(t)+\delta}\right)^\gamma, \tag{12.24}$$

$$f(t) = \mu^\circ(t)\,\frac{\delta^\gamma\,\gamma}{(W(t)+\delta)^{\gamma+1}}, \tag{12.25}$$

$$\mu(t) = \mu^\circ(t)\,\frac{\gamma}{W(t)+\delta}. \tag{12.26}$$

Observe that

$$P[W(t) > w] = \bar{F}(W^{-1}(w)) = \left(\frac{\delta}{w+\delta}\right)^\gamma,$$

that is, $W(t)+\delta$ is Pareto-distributed with parameters (δ,γ). Observe also that $\mu(t)$ is $\mu^\circ(t)$ times a factor that decreases with t. Notice that μ° is not a force of mortality in the unconditional law.

The survival function for an x year old is obtained from (12.24) as

$$\bar{F}(t\,|\,x) = \bar{F}(x+t)/\bar{F}(x) = \left(\frac{W(x)+\delta}{W(x+t)+\delta}\right)^\gamma, \tag{12.27}$$

which can be written as

$$\bar{F}(t\,|\,x) = \left(\frac{\delta(x)}{W(t\,|\,x)+\delta(x)}\right)^\gamma \tag{12.28}$$

with

$$W(t\,|\,x) = W(t+x) - W(x) = \int_0^t \mu^\circ(x+\tau)\,d\tau,$$

$$\delta(x) = W(x) + \delta.$$

Thus, the survival distribution of an x-year old is of the same form as that of a newly born, see (12.24), but with updated value of the scale parameter.

Despite what has been said about Θ and G as elements of a mere explanatory background of the "surface" entities T and F, it may be of interest to study the conditional distribution (12.6) for some given aspect $T \in A$ of the life length. As an example, take the general proportional hazard model and focus attention at (see (12.6)

$$\bar{G}(\theta\,|\,T > t) = P[\Theta > \theta\,|\,T > t] = \frac{\int_\theta^\infty e^{-\theta' W(t)}\,dG(\theta')}{\int_0^\infty e^{-\theta' W(t)}\,dG(\theta')}.$$

To study the dependence of this function on t, differentiate:

$$\frac{\partial}{\partial t}\bar{G}(\theta\,|\,T > t) = \frac{\int_0^\infty e^{-\theta' W(t)}\,dG(\theta')\,\int_\theta^\infty e^{-\theta' W(t)}(-\theta'\,\mu^\circ(t))\,dG(\theta') - \int_0^\infty e^{-\theta' W(t)}(-\theta'\,\mu^\circ(t))\,dG(\theta')\,\int_\theta^\infty e^{-\theta' W(t)}\,dG(\theta')}{\left\{\int_0^\infty e^{-\theta' W(t)}\,dG(\theta')\right\}^2} \tag{12.29}$$

By introduction of the probability measure H given by

$$dH(\theta) = \frac{e^{-\theta W(x)}\, dG(\theta)}{\int_0^\infty e^{-\theta' W(t)}\, dG(\theta')}\,,$$

(12.29) can be cast as

$$\mu^\circ(t)\, \{-E_H(\Theta\, 1[\Theta > \theta]) + E_H \Theta\, E_H 1[\Theta > \theta]\} = -\mu^\circ(t) Cov_H(\Theta, 1[\Theta > \theta]), \qquad (12.30)$$

where subscript H signifies that the operators are formed by the distribution H. Now, Θ and $1[\Theta > \theta]$ are both increasing functions of Θ, hence associated and positively correlated. It follows that the expression in (12.30) is negative, hence that the probability $G(\theta \,|\, T > t)$ is a decreasing function of t and so $RTD(\Theta \,|\, T)$. Thus, old people are likely to have a small value of θ.

Specialize now to the gamma case (12.19) again. By (12.8), the conditional density of Θ, given $X > x$, is

$$g_{\gamma,\delta}(\theta \,|\, T > t) = c\theta^{\gamma-1}\, e^{-\theta(W(t)+\delta)}, \qquad (12.31)$$

where c does not depend on θ. The constant c need not be calculated since, by inspection of (12.19) and (12.31) it follows that

$$g_{\gamma,\delta}(\theta \,|\, T > t) = g_{\gamma, W(t)+\delta}(\theta)\,. \qquad (12.32)$$

Formula (12.28) is also obtained from (12.11) upon inserting

$$F_\theta(t \,|\, x) = \exp[-\theta\{W(x+t) - W(x)\}]$$

and, from (12.32), $dG(\theta \,|\, X > x) = g_{\gamma, W(x)+\delta}(\theta)\, d\theta$ and using (12.20). From a computational point of view this is a detour, but it uncovers the role of the hidden Θ in the play.

Chapter 13

Group life insurance

13.1 Basic characteristics of group insurance

.
A group insurance treaty is an arrangement whereby a group of persons is covered by a single contract with an insurer. In its broadest context, group insurance would include a variety of coverages, including life insurance, accident insurance, health insurance, annuities, civil property insurance, and others. The major types of groups eligible for group insurance are *individual employer groups* (the employees of a firm are covered by a contract between an insurer and the employer, typically as a part of a labour–management negotiated employee welfare and security plan), *multiple employer groups* (the same as individual employer groups, except that the individual firm is extended to two or more firms/employers, e.g. a trade association or an entire industry), *labour union groups* (the members of a labour union are covered by a contract issued directly to the union), and *creditor – debtor groups* (life and/or health insurance is provided for debtors through a contract issued to the creditor, e.g. a commercial bank; if the borrower dies or is disabled, benefits are paid to the lender to cancel the insured part of the debt). A great variety of groups beyond the foregoing classifications are covered by group insurance. Among such *miscellaneous groups* are associations of public and private employees, professional organizations, fraternal societies, and many others.

When group insurance is contrasted with individual insurance, a number of characteristic features are evident, In the first place, the coverage is offered to all members of the group, usually without medical examination or other evidence of individual insurability. Thus, the criteriae by which individuals are recruited to a group are considered to provide a sufficient guarantee against adverse selection of high-risk individuals, so-called antiselection against the insurer. For instance, it is to be expected that the staff of an engineering workshop or publishing house or the membership of an association of lawyers or teachers has only a small infusion of impaired lives. To further preclude the possibility of antiselection, there are usually some requirements pertaining to the minimum number of persons needed to constitute a group and to the minimum proportion covered in the entire group.

Another feature characteristic of group insurance is that the persons insured under a contract are not parties to the contract, since legally the contract is between the insurer and the policyholder (usually an employer or an organization).

A third characteristic of group insurance is that it is essentially low-cost, mass protection. Marketing and administration costs are far below the level typical of individual insurance.

As a fourth characteristic, it should be pointed out that group insurance contracts are of a continuing nature, in that the contract and the plan may last long beyond the lifetime, or membership in the plan, of any one individual. New persons are added to the group from time to time, and others terminate their coverage. The contract is renewed regularly, typically annually. Therefore, both contract terms and premium rate can be currently adjusted in accordance with the observed development of costs, risk conditions and other circumstances influencing the economic result. This feature sets a difference of great principal and practical importance between group contracts and individual contracts in life insurance. An individual life insurance policy usually specifies premiums and benefits that remain unalterned throughout the contract period, which may extend over several decades. Therefore, a substantial safety loading is usually built into the individual premium to meet possible unfavourable future developments of mortality and expenses. In group life insurance there is no need for this kind of safety loading since the premium rate can be currently adjusted in accordance with the experiences.

The last remark points directly to the final feature of group insurance to be mentioned here. To save expenses, usually only very summary characteristics of the groups are observed and used as a basis for the rating of premiums at the outset. Those risk characteristics that are not observed may vary considerably between the groups and give rise to substantial risk differentials between them, despite that they "appear to be similar". As an example, in group life insurance one may choose to observe only the number of persons insured under the plan of a group and leave other characteristics such as occupation and age composition unobserved. If the groups differ considerably with respect to these unobserved risk characteristics, they will have different "true underlying risk premiums". These differentials will be reflected by the risk experiences of the individual groups as time passes and claims statistics accrues. Thus, the individual claims record of a group provides some information on its "risk profile", which could be taken into account in the current adjustment of the premium. When the premium is regulated this way for each group in regard of its claims experience, one speaks of *experience rating*.

There is yet another point of difference between individual and group insurance which ought to be mentioned because it explains why experience rating is widely used in group life insurance. If you should ask holders of individual life insurance policies if they find the premiums reasonable, the answers would typically be "I guess so" or "I don't know". They don't know and don't haggle over the price, simply because they have no access to statistics from which they could judge the fairness of the premiums. In group life insurance this is different. Each master contract is managed by a policyholder who can compare premium payments with received benefits in the long run. Those policy holders who find that premiums exceed by far the benefits, will sooner or later call for a discount (the others will remain silent). Therefore, a competitive market will tend to enforce experience rating of groups life contracts.

13.2 A proportional hazard model for complete individual policy and claim records

Consider a group life portfolio for which statistical records have been maintained during the period (τ', τ''), where τ'' is the present moment. The portfolio comprises I master contracts, labeled by $i = 1, \ldots, I$. Let (τ_i', τ_i'') be the period during which contract i has been in force ($\tau_i'' < \tau''$ if the contract has been terminated in its entirety). Let J_i be the number of persons currently or formerly insured under the plan of contract i. They are labeled by (i, j), $j = 1, \ldots, J_i$. For each individual (i, j) introduce the following quantities, which are observable by time τ'' :

τ_{ij}',	the time of entry into the group,
x_{ij},	the age at enty,
T_{ij},	the time exposed to risk as insured before time τ_i''
K_{ij},	the number of times the coverage has been terminated on an individual basis before time τ_i'',
M_{ij},	the number of deaths as insured before time τ_i''.

The pairs (K_{ij}, M_{ij}) can only assume the values (0,0), (1,0), (0,1) (implying that participation in the group will not be resumed once it has been terminated). The hidden mortality characteristics of group i are represented by a latent quantity Θ_i. The T_{ij}, K_{ij}, M_{ij}, and Θ_i are viewed as random variables, and the following assumptions are made.

(i) Variables belonging to different groups are stochastically independent and $\Theta_1, \ldots, \Theta_I$ are iid (independent and identically distributed).

(ii) Variables belonging to different persons within one and the same group i are conditionally independent, given Θ_i.

(iii) All persons in one and the same group follow the same pattern of mortality and termination. More specifically, it is assumed that the Θ_i are positive and that, conditional on $\Theta_i = \theta_i$, a person who entered group i at age x and is still a member of the group at age $x + t$ $(x, t > 0)$, then has a force of termination

$$\kappa_i(x, t) \tag{13.1}$$

and a force of mortality of the form

$$\theta_i \mu(x, t). \tag{13.2}$$

Assumption (i) corresponds to the idea that the groups are independent random selections from a population of groups that are comparable, but not entirely similar. It is this assumption, in conjunction with (15.2), that establishes a relationship between the groups and forms the rationale of combining portfoliowide mortality experience with the mortality experience of a given group in an assessment of the mortality in that group. The "proportional hazard" assumption (15.2) represents, perhaps, the simplest possible way of modelling mortality variations between groups. It states that the risk characteristics specific of a group act on the force of mortality only through a multiplicative factor, implying that the mortality pattern is basically the same for all groups. Such an assumption is not apt for describing more complex mortality differences, e.g. that a group may have a mortality below the average at early ages and above the average towards the end of life. For example, it is thinkable that such hazardous occupations as blast furnace operation and mining attract only physically fit and healthy applicants and that those who are employed quickly get worn out by the severe working conditions.

The statistical data presently available from group i are the individual entrance times, τ'_{ij}, entrance ages, x_{ij}, and histories as insurees,

$$\mathcal{O}_i = \{(K_{ij}, M_{ij}, T_{ij});\ \ j = 1, \ldots, J_i\}.$$

The conditional distribution of (K_{ij}, M_{ij}, T_{ij}), when $\Theta_i = \theta_i$, is given by

$$\begin{aligned} &P\{K_{ij} = 0,\ M_{ij} = 1,\ t_{ij} < T_{ij} t_{ij} + dt_{ij} \,|\, \Theta_i = \theta_i\} \\ &\quad = \ \theta_i \mu(x_{ij}, t_{ij})\, dt_{ij}\ \exp\left[-\int_0^{t_{ij}} \{\kappa_i(x_{ij}, t) +\ \theta_i \mu(x_{ij}, t)\}\, dt \right], \\ &\qquad\qquad\qquad\qquad\qquad\qquad 0 < t_{ij} < \tau''_i - \tau'_{ij}, \end{aligned}$$

$$\begin{aligned} &P\{K_{ij} = 1,\ M_{ij} = 0,\ t_{ij} < T_{ij} t_{ij} + dt_{ij} \,|\, \Theta_i = \theta_i\} \\ &\quad = \ \kappa_i(x_{ij}, t_{ij})\, dt_{ij} \exp\left[-\int_0^{t_{ij}} \{\kappa_i(x_{ij}, t) + \theta_i \mu(x_{ij}, t)\}\, dt \right], \\ &\qquad\qquad\qquad\qquad\qquad\qquad 0 < t_{ij} < \tau''_i - \tau'_{ij}, \end{aligned}$$

$$
\begin{aligned}
&P\{K_{ij} = M_{ij} = 0, T_{ij} = t_{ij} \,|\, \Theta_i = \theta_i\} \\
&= \quad \exp\left[- \int_0^{t_{ij}} \{\kappa_i(x_{ij}, t) + \theta_i \mu(x_{ij}, t)\}\, dt \right],
\end{aligned}
$$

$$t_{ij} = \tau_i'' - \tau_{ij}'.$$

From these expressions we gather the following formula for the conditional likelihood of $\mathcal{O}_i$:

$$
\begin{aligned}
&\prod_{j=1}^{J_i} (\kappa_i(x_{ij}, T_{ij})^{K_{ij}} \{\theta_i \mu(x_{ij}, T_{ij})\}^{M_{ij}} \\
&\qquad \exp[- \int_0^{T_{ij}} \{\kappa_i(x_{ij}, t) + \theta_i \mu(x_{ij}, t)\}\, dt]),
\end{aligned} \tag{13.3}
$$

$(K_{ij}, M_{ij}) \in \{(0,1),(1,0)\}$ and $0 < T_{ij} < \tau_i'' - \tau_{ij}'$ or $(K_{ij}, M_{ij}) = (0,0)$ and $T_{ij} = \tau_i'' - \tau_{ij}'$, $j = 1, \ldots, J_i$. For each person (i,j) introduce the cumulative basic force of mortality

$$W_{ij} = \int_0^{T_{ij}} \mu(x_{ij}, t)\, dt. \tag{13.4}$$

It is seen from (15.3) that a set of sufficient statistics for group i are

$$M_i = \sum_{j=1}^{J_i} M_{ij}, \quad \text{the total number of deaths,} \tag{13.5}$$

$$W_i = \sum_{j=1}^{J_i} W_{ij}, \quad \text{the sum of cumulative basic intensities,} \tag{13.6}$$

and that the conditional likelihood, considered as a function of θ_i, is proportional to

$$\theta_i^{M_i}\, e^{-\theta_i W_i}. \tag{13.7}$$

The expression in (15.7) is gamma shaped. Motivated by the convenience of the analysis in the previous chapter, it is assumed that the common distribution of the latent Θ_i is the gamma distribution with density

$$g_{\gamma,\delta}(\theta_i) = \frac{\delta^\gamma}{\Gamma(\gamma)}\, \theta_i^{\gamma-1}\, e^{-\theta_i \delta}, \ \ \theta_i > 0. \tag{13.8}$$

The conditional density of Θ_i, given $\mathcal{O}$, is proportional to the product of the expressions in (15.7) and (15.8), hence

$$g_{\gamma,\delta}(\theta_i \,|\, \mathcal{O}_i) = g_{M_i+\gamma, W_i+\delta}(\theta_i). \tag{13.9}$$

By use of (14.25)–(14.27), it follows that

$$E(\Theta_i \,|\, \mathcal{O}_i) \quad = \quad \frac{M_i + \gamma}{W_i + \delta}, \tag{13.10}$$

$$Var(\Theta_i \,|\, \mathcal{O}_i) \quad = \quad \frac{M_i + \gamma}{(W_i + \delta)^2}, \tag{13.11}$$

$$E(e^{-w\Theta_i} \,|\, \mathcal{O}_i) = \left(\frac{W_i + \delta}{w + W_i + \delta} \right)^{M_i+\gamma}, \ \ w > -(W_i + \delta). \tag{13.12}$$

The conditional mean in (15.10) is the Bayes estimator $\tilde{\Theta}_i$ (say) of Θ_i with respect to squared loss. It can be cast as

$$\tilde{\Theta}_i = \zeta_i \hat{\Theta}_i + (1-\zeta_i)\gamma/\delta, \tag{13.13}$$

where

$$\hat{\Theta}_i = M_i/W_i \tag{13.14}$$

is the maximum likelihood estimator of θ_i in the conditional model, given $\Theta_i = \theta_i$, and

$$\zeta_i = W_i/(W_i+\delta). \tag{13.15}$$

The expression in (15.13) is a weighted mean of the sample estimator $\hat{\Theta}_i$ and the unconditional mean, $E\Theta_i = \gamma/\delta$. The weight ζ_i attached to the experience of the group, is an increasing function of the exposure times T_{ij}, confer (15.4).

13.3 Experience rated net premiums

The set of master contracts in force at the present moment is

$$\mathcal{I} = \{i : \tau_i'' = \tau''\},$$

and for each group $i \in \mathcal{I}$ the set of persons presently covered under the plan of the group is

$$\mathcal{J}_i = \{j : \tau_{ij}' + T_{ij} = \tau''\}.$$

For each person (i,j) presently insured let M_{ij}'' and S_{ij}'' denote, respectively, the number of deaths and the sum payable by death in the next year, $(\tau'', \tau''+1)$. To prevent technicalities from obscuring the main points, disregard interest and assume that all the S_{ij}'' will remain constant throughout the year.

For a group $i \in \mathcal{I}$ the net annual premium based on the available information $\mathcal{O}_i$ is

$$P_i^A = \sum_{j\in\mathcal{J}_i} S_{ij}''\, E(M_{ij}''\,|\,\mathcal{O}_i). \tag{13.16}$$

The expected values appearing in (15.16) are

$$\begin{aligned} E(M_{ij}''\,|\,\mathcal{O}_i) &= E\{E(M_{ij}''\,|\,\Theta_i, \mathcal{O}_i)\,|\,\mathcal{O}_i\} \\ &= E[1-\exp\left\{-\Theta_i \int_0^1 \mu(x_{ij}, T_{ij}+t)\; dt\right\}|\,\mathcal{O}_i], \end{aligned}$$

which can be calculated by formula (15.12). Defining

$$w_{ij}'' = \int_0^1 \mu(x_{ij}, T_{ij}+t)\, dt \tag{13.17}$$

and

$$Q_{i,w} = 1 - \left(\frac{W_i+\delta}{w+W_i+\delta}\right)^{M_i+\gamma}, \tag{13.18}$$

one finds

$$E(M_{ij}''\,|\,\mathcal{O}_i) = Q_{i,w_{ij}''}. \tag{13.19}$$

Substituting (15.19) into (15.16) yields

$$P_i^A = \sum_{j \in \mathcal{J}_i} S''_{ij}\, Q_{i,w''_{ij}}, \tag{13.20}$$

with $Q_{i,w''_{ij}}$ defined by (15.17) and (15.18).

As an alternative to the premium (15.20), which is exact on an annual basis, one could use the "instantaneous net premium" per time unit at time τ'',

$$P_i^I = \lim_{\Delta\tau \downarrow 0} E\left\{ \sum_{j \in \mathcal{J}_i} S''_{ij}\, M''_{ij}(\Delta\tau) \,|\, \mathcal{O}_i \right\} /\Delta\tau, \tag{13.21}$$

where $M''_{ij}(\Delta\tau)$ is the number of deaths of person (i,j) in the time interval $(\tau'', \tau'' + \Delta\tau)$. Now,

$$\begin{aligned} E\{M''_{ij}(\Delta\tau) \,|\, \mathcal{O}_i\} &= E[E\{M''_{ij}(\Delta\tau)) \,|\, \Theta_i, \mathcal{O}_i\} \,|\, \mathcal{O}_i] \\ &= E\{\Theta_i \mu(x_{ij}, T_{ij})\Delta\tau + o(\Delta t) \,|\, M_i, W_i\} \\ &= \mu(x_{ij}, T_{ij})\Delta t \tilde{\Theta}_i + o(\Delta t), \end{aligned} \tag{13.22}$$

the last passage being a consequence of (15.10) and (15.13). Combine (15.21) and (15.22), to obtain

$$P_i^I = \sum_{j \in \mathcal{J}_i} S''_{ij}\, \mu(x_{ij}, T_{ij})\tilde{\Theta}_i. \tag{13.23}$$

To see that P_i^I is an approximation to P_i^A, apply the first order Taylor expansion $(1+x)^{-\alpha} \approx 1 - \alpha x$ to the second term on the right of (15.18) and then approximate w''_{ij} in (15.17) by $\mu(x_{ij}, T_{ij})$, which gives

$$\begin{aligned} Q_{i,w''_{ij}} &= 1 - \{1 + w''_{ij}/(W_i + \delta)\}^{-(M_i+\gamma)} \\ &\approx 1 - \{1 - (M_i + \gamma)w''_{ij}/(W_i + \delta)\} \\ &= w''_{ij}\tilde{\Theta}_i \\ &\approx \mu(x_{ij}, T_{ij})\tilde{\Theta}_i. \end{aligned} \tag{13.24}$$

Using (15.25) in (15.20), yields $P_i^A \approx P_i^I$. The approximation is good if the w''_{ij} are $<< W_i$, which is the case for groups with a reasonably large risk exposure in the past, and if the $\mu(x_{ij}, T_{ij} + t)$ are nearly constant for $0 < t < 1$.

13.4 The fluctuation reserve

A measure of the uncertainty associated with the annual result for group i is the conditional variance,

$$\begin{aligned} V_i^A &= Var\left(\sum_{j \in \mathcal{J}_i} S''_{ij}\, M''_{ij} \,|\, \mathcal{O}_i) \right) \\ &= E\left\{ \sum_{j,k \in \mathcal{J}_i} S''_{ij} S''_{ik} M''_{ij} M''_{ik} \,|\, \mathcal{O}_i \right\} - (P_i^A)^2 \\ &= \sum_{j \in \mathcal{J}_i} S''^2_{ij} E(M''_{ij} \,|\, \mathcal{O}_i) \\ &\quad +2 \sum_{j,k \in \mathcal{J}_i; j<k} S''_{ij} S''_{ik} E(M''_{ij} M''_{ik} \,|\, \mathcal{O}_i) - (P_i^A)^2 \end{aligned} \tag{13.25}$$

(M''_{ij} is equal to its square since it is 0 or 1). The expected values appearing in the second sum in (15.26) are

$$\begin{aligned} E(M''_{ij}M''_{ik}\,|\,\mathcal{O}_i) &= E\{E(M''_{ij}M''_{ik}\,|\,\Theta_i,\mathcal{O}_i)\,|\,\mathcal{O}_i\} \\ &= E\left[\{1-\exp(-\Theta_i w''_{ij})\}\{1-\exp(-\Theta_i w''_{ik})\}\,|\,\mathcal{O}_i\right] \\ &= Q_{i,w''_{ij}} + Q_{i,w''_{ik}} - Q_{i,w''_{ij}+w''_{ik}}, \end{aligned} \tag{13.26}$$

confer (15.12) and (15.18). Now, insert the expressions (15.19) and (15.26) into (15.25) to obtain

$$\begin{aligned} V_i^A &= \sum_{j\in\mathcal{J}_i} S''^2_{ij} Q_{i,w''_{ij}} + 2\sum_{j,k\in\mathcal{J}_i;\,j<k} S''_{ij}S''_{ik}(Q_{i,w''_{ij}} + Q_{i,w''_{ik}} - Q_{i,w''_{ij}+w''_{ik}}) - (P_i^A)^2 \\ &= 2\sum_{j\in\mathcal{J}_i} S''_{ij}P_i^A - \sum_{j\in\mathcal{J}_i} S''^2_{ij}Q_{i,w''_{ij}} \\ &\quad -2\sum_{j,k\in\mathcal{J}_i;\,j<k} S''_{ij}S''_{ik}Q_{i,w''_{ij}+w''_{ik}} - (P_i^A)^2, \end{aligned} \tag{13.27}$$

where $Q_{i,w''_{ij}}$ and P_i^A are given by (15.17), (15.18), and (15.20). By use of the approximation (15.24),

$$V_i^A \approx \sum_{j\in\mathcal{J}_i} S''^2_{ij} w''_{ij}\tilde{\Theta}_i - (P_i^A)^2. \tag{13.28}$$

Charging each group i its net premium P_i^A would only secure expected equivalence of premium incomes and benefit payments for the portfolio as a whole. (At any time groups with low mortality will subsidize those with high mortality, but as time passes and risk experience accrues, these transfers will diminish: eventually each group will be charged its true risk premium.) To meet unfavourable random fluctuations in the results, the company should provide a reserve for the entire portfolio. By approximation to the normal distribution, which is reasonable for a portfolio of some size, a fluctuation reserve given by

$$F^A = 2.33\left(\sum_{i\in\mathcal{I}} V_i^A\right)^{\frac{1}{2}}, \tag{13.29}$$

with V_i^A defined by (15.27) or (15.28), will be sufficient to cover claim expenses in excess of the total net premium, $\sum_{i\in\mathcal{I}} P_i^A$, with 99% probability. To establish the reserve in (15.29), it may be necessary to charge each group an initial loading in addition to the net premium. Thereafter the reserve can be maintained by transfer of surplus in years with favourable results for the whole portfolio. (Charging the insurees a total premium loading equal to F^A each year is, of course, not necessary: that would create an unlawful profit on the part of the insurer.) The loadings can be determined in several ways. One reasonable possibility is to let each group i contribute to F^A by an amount F_i^A proportional to the standard deviation $(V_i^A)^{\frac{1}{2}}$, that is,

$$F_i^A = F^A(V_i^A)^{\frac{1}{2}}/\sum_{k\in\mathcal{I}}(V_k^A)^{\frac{1}{2}}. \tag{13.30}$$

Upon termination of a master contract, the group should be credited with the amount (15.30).

13.5 Estimation of parameters

At time τ'' the observations that can be utilized in parameter estimation are $\mathcal{O}_i,\ i = 1, \ldots, I$. It is, of course, only in this connection that the data from terminated master contracts come into play.

Assume now that the basic mortality law is of Gompertz-Makeham type and is aggregate, that is, $\mu(x, t) = \mu(x + t)$, where

$$\mu(y) = \alpha + \beta c^y, \quad y > 0. \tag{13.31}$$

(In group life insurance there is actually no reason to expect selectional effects since eligibility is not made conditional on the insuree's health or other individual risk characteristics.)

From (15.3), (15.8) and (15.31) one gathers the following expression for the unconditional likelihood of the data:

$$\prod_{j=1}^{j_i} \left[\int_0^\infty \left\{ \prod_{j=1}^{J_i} \kappa_i(x_{ij}, T_{ij})^{K_{ij}} \right\} \theta_i^{M_i} \left\{ \prod_{j=1}^{J_i} (\alpha + \beta c^{x_{ij}+T_{ij}})^{M_{ij}} \right\} \right.$$
$$\exp\left\{ -\sum_{j=1}^{J_i} \int_0^{T_{ij}} \kappa_i(x_{ij}, t)\, dt - \ \theta_i \sum_{j=1}^{J_i} \int_0^{T_{ij}} (\alpha + \beta c^{x_{ij}+t})\, dt \right\}$$
$$\left. \frac{\delta^\gamma}{\Gamma(\gamma)}\, \theta_i^{\gamma-1}\, e^{-\delta\theta_i}\, d\theta_i \right].$$

The forces of termination do not appear in any of the expressions for premiums and reserves, and so one can concentrate on the estimation of $\gamma, \delta, \alpha, \beta, c$. The essential part of the likelihood is

$$\left\{ \prod_{i=1}^{I} \prod_{j=1}^{J_i} (\alpha + \beta c^{x_{ij}+T_{ij}})^{M_{ij}} \right\} \left(\frac{\delta^\gamma}{\Gamma(\gamma)} \right)^I$$
$$\prod_{i=1}^{I} \frac{\Gamma(M_i + \gamma)}{\left\{ \alpha \sum_{j=1}^{J_i} T_{ij} + \beta \sum_{j=1}^{J_i} c^{x_{ij}} (c^{T_{ij}} - 1)/\ln c + \delta \right\}^{M_i+\gamma}}.$$

The maximum likelihood estimators $\gamma^*, \delta^*, \alpha^*, \beta^*, c^*$ have to be determined by numerical methods, e.g. steepest ascent or Newton-Raphson techniques. Note that the number of parameters is essentially only four since a scale parameter in θ_i can be absorbed in μ in (15.2). On should, therefore, put $\gamma = \delta$ or $\alpha = 1$ or $\beta = 1$.

Chapter 14

Hattendorff and Thiele

Hattendorff's classical result on zero means and uncorrelatedness of the losses created in disjoint time intervals by a life insurance policy is an immediate consequence of the very definition of the concept of loss. Thus, the result is formulated and proved here in a setting with quite general payments, discount function, and time intervals, all stochastic. A general representation is given for the variances of the losses. They are easy to compute when sufficient structure is added to the model. The traditional continuous time Markov chain model is given special consideration. A stochastic Thiele's differential equation is obtained in a fairly general counting process framework.

14.1 Introduction

A. Short review
Hattendorff's (1868) theorem states that the losses in different years on a life insurance policy have zero means and are uncorrelated, hence the variance of the total loss is the sum of the variances of the per year losses. The loss in a year is defined as the net outgoes (insurance benefits less premiums) during the year, plus the reserve that has to be provided at the end of the year, and minus the reserve released at the beginning of the year, all quantities discounted at time 0. This classical result has had its recent revival with the advent of modern life insurance mathematics based on the theory of stochastic processes, notable references being Gerber (1979, 1986), Papatriandafylou and Waters (1984), Wolthuis (1987), and Ramlau-Hansen (1988). The first three of these papers deal with losses in fixed time intervals but, apart from that, the proofs do not really rest on any specific model assumptions. The two last-mentioned papers deal with stochastic periods, namely the total sojourn times in different states in the framework of the continuous time Markov chain model, with deterministic contractual benefits and constant interest rate. The proofs, hence apparently also the results, depend on the particular structures of the model.

B. Outline of the present paper
It is shown here that Hattendorff's results are valid for virtually any payment stream and rule for accumulation of interest, no matter what stochastic mechanisms govern them, and for any natural time periods, random or not. The crux of the matter is the very definition of the losses: as pointed out by Gerber (1979), they are the increments of a martingale. The rest follows from the optional sampling theorem. This general result is established in Section 2.

It is only when it comes to formulas for the variances of the losses that specific model assumptions are crucial. Section 3 deals with the situation where the events upon which payments are contingent are governed by a continuous time stochastic process with finite state space, essentially a multivariate counting process. A representation theorem for martingales adapted to a counting process is used to identify the variance process that generates the variances of the losses. As a byproduct, Thiele's differential equation, well-known from the case with Markov counting process and benefits depending only on the present state, turns out to be a special case of a stochastic differential equation valid under far more general assumptions.

The Markov case with nonrandom benefits and discount function, is simple enough to allow for easy computation of the variances. As an example of what can be gained by the general formulation of Hattendorff's theorem, it is applied to the periods of time spent upon the n-th visit to a certain state. Benefits depending on the past development of the process can also be dealt with without great difficulties. Relaxing the Markov assumption is what makes computation of variances cumbersome.

For ease of reference, some elements from the theory of martingales are gathered in the final Section 4. They are taken as prerequisites throughout.

14.2 The general Hattendorff theorem

A. Payment streams and discounted values

A suitable framework for a general treatment of properties of payment streams is set in two recent papers by the author (Norberg, 1990, 1991). It is adopted here with sufficient explanation to make the presentation selfcontained.

Consider a stream of payments commencing at time 0. It is defined by the payment function A, which to each time $t \geq 0$ specifies the total amount $A(t)$ paid in $[0,t]$. Negative payments are allowed for; it is only required that A be of bounded variation in finite intervals and, by convention, right-continuous. This essentially means that A is a difference of two non-decreasing, finite-valued, and right-continuous payment functions representing outgoes and incomes, respectively. When the discount function v is used, the present value at time 0 of the payments is

$$V = \int v \, dA. \tag{14.1}$$

To allow for random development of payments and interest, the functions A and v are assumed to be stochastic processes defined on some probability space $(\Omega, \mathcal{F}, P)$. They are adapted to a right-continuous filtration $\mathbf{F} = \{\mathcal{F}_t\}_{t \geq 0}$, each $\mathcal{F}_t$ comprising the events that govern the development of payments and interest up to and including time t.

B. The notions of reserve and loss

At any time t the insurer must provide a reserve to meet future net liabilities on the policy. An adequate reserve would be the cash value at time t of future outgoes minus incomes, but this quantity is in general unobservable by time t. What can be entered in the accounts, is its expected value given the information available by time t, the so-called (*prospective*) $\mathbf{F}$-*reserve*,

$$V_{\mathbf{F}}(t) = \frac{1}{v(t)} \mathrm{E} \left(\int_{(t,\infty)} v \, dA \mid \mathcal{F}_t \right). \tag{14.2}$$

The $\mathbf{F}$-reserve plays a central role in insurance practice since it is taken as a factual liability and, by statute, is to be accounted as a debt in the balance sheet of the insurer at time t. In accordance with this accountancy convention the insurer's *loss in the period* $(s,t]$ is defined as

$$L_{(s,t]} = \int_{(s,t]} v \, dA + v(t) V_{\mathbf{F}}(t) - v(s) V_{\mathbf{F}}(s), \tag{14.3}$$

the expression on the right being construed as follows: at the end of the period the net outgoes throughout the period have been covered (the first term), the new reserve must be provided (the second term), and the reserve set aside prior to the period can be cashed (the third term).

C. Hattendorff's theorem generally stated.

The basic circumstance underlying Hattendorff's theorem and its generalizations is the fact that the loss as defined by (14.3) is the increment over $(s,t]$ of the martingale generated by the value V in (14.1). More precisely, assuming that $\mathrm{E}|V| < \infty$, define for each $t \geq 0$

$$\begin{aligned} M(t) &= \mathrm{E}\,(V|\mathcal{F}_t) \\ &= \int_{[0,t]} v\,dA + v(t)V_{\mathbf{F}}(t), \end{aligned} \tag{14.4}$$

the second equality due to **F**-adaptedness of A and v and the definition (14.2). By inspection of (14.3) and (14.4),

$$L_{(s,t]} = M(t) - M(s). \tag{14.5}$$

Clearly, $\{M(t)\}_{t\geq 0}$ is an **F**-martingale converging to V. It can always be taken to be right-continuous (see e.g Protter, 1990, p.8) so that the optional sampling property and its consequences apply, confer Section 4.

A general Hattendorff theorem is now obtained by merely spelling out some basic properties of martingales and stochastic integrals quoted in Section 4, in particular (F.38), (F.39), and (14.10), with stopping times in the roles of the t_i. Thus, let $0 = T_0 < T_1 < \cdots$ be a nondecreasing sequence of **F**-stopping times. This covers the simple case where the T_i are fixed (typically $T_i = i$, the end of year No. i) and all practically relevant cases where they are random. Abbreviate

$$L_i = L_{(T_{i-1},T_i]}, \qquad i = 1, 2, \ldots, \tag{14.6}$$

and define consistently the loss at time 0 as

$$L_0 = A(0) + V_{\mathbf{F}}(0) - \mathrm{E}\,V. \tag{14.7}$$

Theorem 1. *Assume that the value V in* (14.1) *has finite variance. Then the losses L_i defined in* (14.6) – (14.7) *have zero means and are uncorrelated, and this is also true conditionally: for $i < j < k$*

$$\mathrm{E}\,(L_j|\mathcal{F}_{T_i}) = 0, \tag{14.8}$$

$$\mathrm{Cov}\,(L_j, L_k \mid \mathcal{F}_{T_i}) = 0, \tag{14.9}$$

hence

$$\mathrm{Var}\left(\sum_{k=j}^{\infty} L_k \mid \mathcal{F}_{T_i}\right) = \sum_{k=j}^{\infty} \mathrm{Var}\,(L_k \mid \mathcal{F}_{T_i}). \tag{14.10}$$

The variances in (14.10) *are of the form*

$$\begin{aligned} \mathrm{Var}\,(L_j \mid \mathcal{F}_{T_i}) &= \mathrm{E}\,(\langle M\rangle(T_j) - \langle M\rangle(T_{j-1}) \mid \mathcal{F}_{T_i}) && (14.11) \\ &= \mathrm{E}\left(\int_{(T_{j-1},T_j]} d\langle M\rangle \mid \mathcal{F}_{T_i}\right), && (14.12) \end{aligned}$$

where $\langle M\rangle$ is the variance process defined by (14.9) *in Section 4.*

The results listed in the theorem are quite independent of the nature of the processes involved. They are rooted in the very definition of the concept of loss, not in any particular assumptions as to what mechanisms they stem from. There are Hattendorff results for virtually any payment stream and rule for accumulation of interest, in insurance, life as well as non-life, and in finance in general. The periods may be any time intervals delimited by stopping times.

The qualitative part of the theorem, (14.8) – (14.10), is done with once and for all. The quantitative part concerning the form of the variances, (14.11) – (14.12) and (14.9), calls for further examination of special models to find computable expressions. The next section is devoted to the case where payments are generated by an insurance policy modelled by a multivariate counting process.

14.3 Application to life insurance

A. The general life insurance policy
A life or pension insurance treaty is typically of the following general form. There is a set $\mathcal{J} = \{0, \ldots, J\}$ of possible states of the policy. At any time t it is in one and only one of the states in $\mathcal{J}$, commencing in state 0 at time 0, say. Let $X(t)$ be the state of the policy at time t. The development of the policy, $\{X(t)\}_{t\geq 0}$, is a stochastic process. Regarded as a function from $[0,\infty)$ to $\mathcal{J}$, it is assumed to be right-continuous, with a finite number of jumps in any finite time interval, and $X(0) = 0$. For each $j \in \mathcal{J}$, let $I_j(t) = 1[X(t) = j]$ be the indicator of the event that the process is in state j at time $t \geq 0$, and for each j, k with $j \neq k$ let $N_{jk}(t) = \sharp\{\tau \in (0,t];\, X(\tau-) = j, X(\tau) = k\}$ be the number of transitions from state j to state k up to and including time $t > 0$. Define $N_{jk}(0) = 0$ for all $j \neq k$. The processes X, $(I_j)_{j\in\mathcal{J}}$, and $(N_{jk})_{j\neq k,\, j,k\in\mathcal{J}}$ correspond mutually one-to-one and carry the same information, which is the filtration $\mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}$ generated by $\{X(t)\}_{t\geq 0}$. In particular, since the process is in state j at a given time if and only if the number of entries into j exceeds (necessarily by 1) the number of departures from j by that time,

$$I_j(t) = \delta_{0j} + \sum_{k;\, k\neq j} (N_{kj}(t) - N_{jk}(t)) \,. \tag{14.1}$$

It is assumed that the process possesses intensities. The intensity of transition from state j to state k, $j \neq k$, is denoted by $\lambda_{jk}(t)$ and is of the form

$$\lambda_{jk}(t) = I_j(t)\mu_{jk}(t) \,, \tag{14.2}$$

with μ_{jk} some $\mathbf{F}$-adapted process.

Insurance benefits are of two kinds. In the first place, the general life annuity provides the amount $A^\circ_j(t) - A^\circ_j(s)$ during a sojourn in state j throughout the time interval $(s,t]$. In the second place, the general assurance provides the lump sum $a^\circ_{jk}(t)$ immediately upon a transition from state j to state k at time t. Insurance premiums are counted as negative benefits. The *contractual functions* A°_j and a°_{jk} are, respectively, a payment function as defined in Paragraph 2A and a left-continuous, finite-valued function, both predictable with respect to $\mathbf{F}$, which means that benefits at any time may depend on the past development of the policy. The total stream of payments A is of the form

$$dA(t) = \sum_j I_j(t)\, dA^\circ_j(t) + \sum_{j\neq k} a^\circ_{jk}(t)\, dN_{jk}(t) \,.$$

It is assumed throughout that the discount function v is continuous and deterministic. In general the reserve $V_{\mathbf{F}}(t)$ depends on the entire history $\mathcal{F}_t$, but it is convenient to suppress this in the notation and visualize only the dependence on the present state of the process. Thus, henceforth write $V_j(t)$ for the reserve when $X(t) = j$.

B. The variance process
In the present model the martingale in (14.4) takes the form

$$\begin{aligned} M(t) \quad &= \quad A^\circ_0(0) + \int_{(0,t]} v(\tau) \left(\sum_j I_j(\tau) dA^\circ_j(\tau) + \sum_{j\neq k} a^\circ_{jk}(\tau)\, dN_{jk}(\tau) \right) \\ &\quad + \sum_j I_j(t)\, v(t) V_j(t) \end{aligned} \tag{14.3}$$

Assume that $\mathrm{E}(\int v\, dA)^2 < \infty$ so that $\{M(t)\}_{t\geq 0}$ is square integrable, that is, $\sup_{t\geq 0} \mathrm{E}\, M^2(t) < \infty$ for all $t \geq 0$. Then a general representation theorem (see Bremaud, 1981, Section III.3) says that M is of the form

$$M(t) \quad = \quad M(0) + \int_{(0,t]} \sum_{j \neq k} H_{jk}(\tau) \left(dN_{jk}(\tau) - \lambda_{jk}(\tau) d\tau \right) , \tag{14.4}$$

where the H_{jk} are some predictable processes satisfying

$$\mathrm{E} \sum_{j \neq k} \int_{(0,t]} H_{jk}^2(\tau)\, \lambda_{jk}(\tau) d\tau < \infty \,. \tag{14.5}$$

Moreover, the variance process is given by

$$d\langle M \rangle(t) = \sum_{j \neq k} H_{jk}^2(t)\, \lambda_{jk}(t) dt \,. \tag{14.6}$$

Thus, to find the variance process, it suffices to identify those functions in (14.3) that take the roles of the H_{jk} in (14.4).

To simplify notation, introduce

$$\tilde{a}_{jk}^{\circ}(t) \quad = \quad v(t)\, a_{jk}^{\circ}(t) \,, \tag{14.7}$$

$$\tilde{A}_{j}^{\circ}(t) \quad = \quad \int_{(0,t]} v(\tau)\, dA_{j}^{\circ}(\tau) \,, \tag{14.8}$$

$$W_j(t) \quad = \quad v(t) V_j(t) \,, \tag{14.9}$$

the last two being right-continuous processes with bounded variation (define $W_j(0) = V_j(0+)$ for $j \neq 0$). Note, in passing, that the function $\tilde{A}_j^{\circ} + W_j$ is continuous since

$$\begin{aligned} \tilde{A}_j^{\circ}(t-) + W_j(t-) &= \tilde{A}_j^{\circ}(t-) + (\tilde{A}_j^{\circ}(t) - \tilde{A}_j^{\circ}(t-)) + W_j(t) \\ &= \tilde{A}_j^{\circ}(t) + W_j(t) \,. \end{aligned}$$

In terms of the functions in (14.7) – (14.9) the relation (14.3) is

$$\begin{aligned} M(t) \quad &= \quad A_0^{\circ}(0) + \int_{(0,t]} \left(\sum_j I_j(\tau) d\tilde{A}_j^{\circ}(\tau) + \sum_{j \neq k} \tilde{a}_{jk}^{\circ}(\tau)\, dN_{jk}(\tau) \right) \\ &\quad + \sum_j I_j(t)\, W_j(t) \,. \end{aligned} \tag{14.10}$$

The last term on the right of (14.10) can be reshaped as follows. Integration by parts gives

$$I_j(t)\, W_j(t) = I_j(0)\, W_j(0) + \int_{(0,t]} I_j(\tau)\, dW_j(\tau) + \int_{(0,t]} W_j(\tau-) dI_j(\tau) \,.$$

In the last term here use (14.1) to write

$$dI_j(t) = \sum_{k;\, k \neq j} d(N_{kj}(t) - N_{jk}(t))$$

and obtain

$$\begin{aligned} \sum_j I_j(t)\, W_j(t) \quad &= \quad \sum_j I_j(0)\, W_j(0) + \sum_j \int_{(0,t]} I_j(\tau)\, dW_j(\tau) \\ &\quad + \sum_j \int_{(0,t]} W_j(\tau-) \sum_{k;\, k \neq j} d(N_{kj}(\tau) - N_{jk}(\tau)) \\ &= \quad W_0(0) + \int_{(0,t]} \sum_j I_j(\tau)\, dW_j(\tau) \\ &\quad + \int_{(0,t]} \sum_{j \neq k} (W_k(\tau-) - W_j(\tau-)) dN_{jk}(\tau) \,. \end{aligned}$$

Upon inserting this, (14.10) becomes

$$\begin{aligned} M(t) &= A_0^\circ(0) + W_0(0) + \int_0^t \sum_j I_j(\tau) d(\tilde{A}_j^\circ(\tau) + W_j(\tau)) \\ &\quad + \int_{(0,t]} \sum_{j \neq k} (\tilde{a}_{jk}^\circ(\tau) + W_k(\tau-) - W_j(\tau-))\, dN_{jk}(\tau)\,. \end{aligned} \tag{14.11}$$

Now, identify the discontinuous parts on the right of (14.4) and (14.11), to obtain the following result.

Theorem 2. *For any continuous discount function and any predictable contractual functions such that* $\mathrm{E}(\int v\, dA)^2 < \infty$*, the variance process* (14.6) *is given by*

$$H_{jk}(t) \quad = \quad \tilde{a}_{jk}^\circ(t) + W_k(t-) - W_j(t-)\,, \tag{14.12}$$

with the elements on the right defined by (14.7) *and* (14.9).

The functions H_{jk} in (14.12) can be expressed as

$$H_{jk}(t) = v(t) R_{jk}(t)\,, \tag{14.13}$$

where

$$R_{jk}(t) = a_{jk}^\circ(t) + V_k(t-) - V_j(t-) \tag{14.14}$$

is the so-called *sum at risk* in respect of transition from state j to state k at time t.

C. A stochastic Thiele's differential equation.
Upon identifying the continuous parts of the expressions on the right of (14.4) and (14.11) and recalling (14.2), the following result is obtained.

Theorem 3. *For any continuous discount function and any predictable contractual functions such that* $\mathrm{E}(\int v\, dA)^2 < \infty$, the identity

$$I_j(t)\, d(\tilde{A}_j^\circ(t) + W_j(t)) + \sum_{k;\, k \neq j} H_{jk}(t)\, \lambda_{jk}(t) dt = 0 \tag{14.15}$$

holds almost surely, the elements being defined by (14.7) – (14.9) *and* (14.12).

This is a generalization of the classical Thiele's differential equation, well-known from the case with Markov counting process and deterministic contractual functions. The stochastic differential equation (14.15) is valid for any counting process possessing intensities, and for any predictable benefit functions, including lump sum survival benefits.

For technical purposes the compact form (14.15) is expedient. For ease of interpretation and comparison with the traditional Thiele's equation, an alternative form is suitable. Assume that the discount function is of the form $v(t) = \exp(-\int_0^t \delta(\tau) d\tau)$, so that $\delta(t)$ is the interest intensity at time t. Insert (14.8), (14.9), and (14.14) in (14.15), put $dW_j(t) = dv(t)V_j(t) + v(t)dV_j(t) = -v(t)\delta(t)dtV_j(t) + v(t)dV_j(t)$, divide by $v(t)$, and rearrange a bit to obtain

$$\begin{aligned} I_j(t)\, d(-A_j^\circ)(t) &= I_j(t)\, (dV_j(t) - V_j(t)\delta(t)dt) \\ &\quad + \sum_{k;\, k \neq j} R_{jk}(t)\, \lambda_{jk}(t) dt\,. \end{aligned} \tag{14.16}$$

On the left of (14.16) is the premium (possibly negative) paid in state j during a small time interval around t. The expression on the right shows how it decomposes into a *savings premium* (the first term), which provides the amount needed for maintenance of the reserve in excess of the interest it creates, and a *risk premium* (the second term), which covers the outgo related to transitions from the current state (sum insured plus new reserve minus old reserve).

D. Losses in a given state
In the present set-up it is meaningful to speak of the total loss $L^{(j)}$ in a certain state j. Let $S_n^{(j)}$ and $T_n^{(j)}$ denote the (stopping) times of arrival and departure, respectively, by the n-th visit of the policy in state j (they are ∞ if no such visit takes place). Then

$$\begin{aligned} L^{(j)} &= \sum_{n=1}^{\infty}\left(M(T_n^{(j)}) - M(S_n^{(j)})\right) && (14.17)\\ &= \sum_{n=1}^{\infty}\int_{(S_n^{(j)},T_n^{(j)}]} dM(\tau) && (14.18)\\ &= \int I_j(\tau-)dM(\tau), \end{aligned}$$

the last equality due to the right-continuity of X, by which

$$\{X(t-)=j\} = \bigcup_{n=1}^{\infty}\{S_n^{(j)} < t \le T_n^{(j)}\}.$$

The function $I_j(t-)$ is left-continuous, hence predictable. Then, by a general martingale result ((14.8) in Section 4), $\mathrm{Var}\, L^{(j)} = \mathrm{E}\int_{(0,\infty)} I_j(\tau-)\, d\langle M\rangle(\tau)$. Inserting (14.6) and recalling (14.2), gives

$$\mathrm{Var}\, L^{(j)} = \int_0^{\infty} \mathrm{E}\left(I_j(\tau-)\sum_{k;\, k\neq j}\mu_{jk}(\tau)\, H_{jk}^2(\tau)\right) d\tau, \tag{14.19}$$

E. The Markov chain model revisited
To obtain feasible expressions for the variance in (14.19), more structure must be placed on the model. Thus, assume now that the process $\{X(t)\}_{t\geq 0}$ is a time-continuous Markov chain, and denote the transition probabilities by

$$p_{jk}(t,u) = P\{X(u)=k \mid X(t)=j\},$$

$0 \le t \le u$, $j,k \in \mathcal{J}$. The μ_{jk} appearing in (14.2) are now deterministic functions given by

$$\mu_{jk}(t) = \lim_{u\downarrow t}\frac{p_{jk}(t,u)}{u-t}.$$

For the time being the contractual functions A_j° and a_{jk}° are taken to be deterministic, that is, the benefits depend only on the current state. Then the discounted reserves $W_j(t)$ are deterministic functions given by the integral expressions

$$W_j(t) = \int_{(t,\infty)} v(\tau)\sum_g p_{jg}(t,\tau)\left(dA_g^{\circ}(\tau) + \sum_{h;\, h\neq g} a_{gh}^{\circ}(\tau)\,\mu_{gh}(\tau)\, d\tau\right), \tag{14.20}$$

which are easy to compute.

It follows that (14.19) reduces to

$$\mathrm{Var}\, L^{(j)} = \int_0^{\infty} p_{0j}(0,\tau)\sum_{k;\, k\neq j}\mu_{jk}(\tau)\, H_{jk}^2(\tau)\, d\tau. \tag{14.21}$$

The expression (14.21) was obtained by Ramlau-Hansen (1988) by use of martingale theorems applied to the compensated multivariate Markov counting process in combination with the classical Thiele's differential equation, implying that the life annuity benefits are absolutely continuous. As for the qualitative part of the Hattendorff theorem, an examination of the proof indicates that the special model assumptions are redundant: at the stage where Thiele's differential equation is invoked, all terms pertaining to intermediate transitions between visits to state j vanish.

It follows from Theorem 1 that not only are the losses created in different states uncorrelated, but so are also the losses created in respect of different visits to one and the same state. The loss in connection with the n-th visit to state j is

$$L_n^{(j)} = M(T_n^{(j)}) - M(S_n^{(j)}) = \int_{(S_n^{(j)},T_n^{(j)}]} dM(t),$$

the n-th term on the right of (14.17) or (14.18). (If there are fewer than n visits to state j, then $S_n^{(j)} = T_n^{(j)} = \infty$ by definition and $L_n^{(j)} = 0$ since $\mathrm{E}|V| < \infty$ implies $\lim_{t\to\infty} \int_t^\infty v\, dA = 0$ almost surely.) Repeating the argument above, replacing $I_j(t-)$ with $1[S_n^{(j)} < t \leq T_n^{(j)}]$ (also a left-continuous function), one arrives at

$$\operatorname{Var} L_n^{(j)} = \int_0^\infty p_{0j}^{(n)}(0,\tau) \sum_{k;\, k\neq j} \mu_{jk}(\tau)\, H_{jk}^2(\tau)\, d\tau, \tag{14.22}$$

where

$$p_{0j}^{(n)}(0,t) = P\{S_n^{(j)} < t \leq T_n^{(j)}\}$$

is the probability of sojourning in state j for the n-th time just before time t. Summing (14.22) over n gives (14.21), a consequence of the general theorem.

The results above carry over to visits within a fixed time interval $(t,u]$ since for any stopping time T the truncated times $(T \vee t) \wedge u$ and $t \vee (T \wedge u)$ are also stopping times ($\vee$ and $\wedge$ form maximum and minimum, respectively); just replace $\int_0^\infty$ by $\int_t^u$. For any stopping time T the $\mathcal{F}_T$-conditional variances of losses in state j after time T are obtained upon replacing $\int_0^\infty$ by $\int_T^\infty$ and $p_{0j}^{(n)}(0,\tau)$ by the appropriate conditional probability.

F. Computational problems in more complex models

An immediate extension of the simple set-up in the previous paragraph, which does not destroy the computability of the variance of the loss, consists in letting the contractual functions A_j° and a_{jk}° depend on the past development of the policy. The expressions may become messy, however.

An interesting issue is to study computational problems under non-Markov assumptions, e.g. when the transition intensities are allowed to depend on the duration of the period that has elapsed since the policy entered the current state. This would complicate matters immensely since integrations would have to be performed over the times of transitions. In principle a numerical procedure can always be arranged. Another possibility would be to let the counting process be doubly stochastic.

Finally, it should be mentioned that the discount function and the contractual functions could be made stochastic and independent of the development of the policy. Then computations would still be feasible by the rule of iterated expectations.

All the extensions indicated here present computational problems that do not pertain particularly to Hattendorff's and Thiele's results, and this is not the right place to pursue such technical issues.

14.4 Excerpts from martingale theory

A. Definition of martingale

Let $(\Omega, \mathcal{F}, P)$ be some probability space. A family $\mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}$ of sub-sigma-algebras of $\mathcal{F}$ is called a *filtration* if it is non-decreasing, that is, $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$. The index t represents time, and $\mathcal{F}_t$ is some collection of $\mathcal{F}$-events whose occurrence or non-occurrence is settled by time t. Thus, $\mathbf{F}$ represents a description of the evolution of the internal history of the phenomena encountered. It is assumed thet $\mathbf{F}$ is right-continuous, which means that $\mathcal{F}_s = \cap_{t;t>s}\mathcal{F}_t$.

A stochastic process $\{M(t)\}_{t\geq 0}$ is said to be $\mathbf{F}$-*adapted* if $M(t)$ is measurable with respect to $\mathcal{F}_t$ for each t. It is an $\mathbf{F}$-*martingale* if it is $\mathbf{F}$-adapted, $\mathrm{E}|M(t)|$ exists for all t, and

$$\mathrm{E}\,(M(t) - M(s)|\mathcal{F}_s) = 0 \tag{14.1}$$

for $s \leq t$. (Relations involving random variables are understood to hold almost surely.)

B. Basic properties of square integrable martingales
In what follows $\{M(t)\}_{t\geq 0}$ is assumed to be a square integrable martingale, that is, (14.1) holds and $\sup_{t\geq 0}\mathrm{E}\,M^2(t) < \infty$. Let $0 = t_0 < t_1 < \cdots$ be some partitioning of $[0,\infty)$, and denote the increments of M over the intervals by

$$\Delta_j M = M(t_j) - M(t_{j-1}). \tag{14.2}$$

As straightforward consequences of the martingale property, the increments have conditional zero means and covariances: for $i < j < k$,

$$\begin{aligned}
\mathrm{E}\,(\Delta_j M \mid \mathcal{F}_{t_i}) &= \mathrm{E}\left(\mathrm{E}\,(\Delta_j M \mid \mathcal{F}_{t_{j-1}}) \mid \mathcal{F}_{t_i}\right) \\
&= 0, && (14.3)\\
\mathrm{Cov}\,(\Delta_j M, \Delta_k M \mid \mathcal{F}_{t_i}) &= \mathrm{E}\,(\Delta_j M\,\Delta_k M \mid \mathcal{F}_{t_i}) \\
&= \mathrm{E}\left(\Delta_j M\, E(\Delta_k M \mid \mathcal{F}_{t_{k-1}}) \mid \mathcal{F}_{t_i}\right) \\
&= 0. && (14.4)
\end{aligned}$$

It follows that, for $i < j < l \leq \infty$ and for H any square integrable $\mathbf{F}$-adapted process,

$$\begin{aligned}
&\mathrm{E}\left(\sum_{k=j}^{l} H(t_{k-1})\Delta_k M \mid \mathcal{F}_{t_i}\right) \\
&\quad = \mathrm{E}\left(\sum_{k=j}^{l} H(t_{k-1})\mathrm{E}(\Delta_k M \mid \mathcal{F}_{t_{k-1}}) \mid \mathcal{F}_{t_i}\right) \\
&\quad = 0, && (14.5)
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{k=j}^{l} H(t_{k-1})\Delta_k M \mid \mathcal{F}_{t_i}\right) \\
&\quad = \mathrm{E}\left(\sum_{k=j}^{l}\sum_{k'=j}^{l} H(t_{k-1})H(t_{k'-1})\mathrm{E}\,(\Delta_k M \Delta_{k'} M \mid \mathcal{F}_{t_{k-1}\vee t_{k'-1}}) \mid \mathcal{F}_{t_i}\right) \\
&\quad = \mathrm{E}\left(\sum_{k=j}^{l} H^2(t_{k-1})\mathrm{E}\left((\Delta_k M)^2 \mid \mathcal{F}_{t_{k-1}}\right) \mid \mathcal{F}_{t_i}\right). && (14.6)
\end{aligned}$$

Integral analogues of (14.5) and (14.6), obtained by passing to the limit, are valid with certain qualifications. Denote by $\mathcal{F}_{t-}$ the history as specified by $\mathbf{F}$ up to, but not including, time t. It is a sub-sigma-algebra of $\mathcal{F}_t$. Assume that the process H is *predictable* in the sense that $H(t)$ is $\mathcal{F}_{t-}$-measurable for each t. Left-continuity certainly implies predictability. Now, let $s = t_i < t_{i+1} < \cdots < t_l = t$ be a partitioning of $[s,t]$. Refine it indefinitely to obtain from (14.5) that

$$\mathrm{E}\left(\int_{(s,t]} H\,dM \mid \mathcal{F}_s\right) = 0, \tag{14.7}$$

and from (14.6) that

$$\mathrm{Var}\left(\int_{(s,t]} H\,dM \mid \mathcal{F}_s\right) = \mathrm{E}\left(\int_{(s,t]} H^2\,d\langle M\rangle \mid \mathcal{F}_s\right), \tag{14.8}$$

where $\langle M\rangle$ is the so-called *variance process* of M defined (informally) by

$$d\langle M\rangle(t) = \mathrm{E}\left((dM(t))^2 \mid \mathcal{F}_{t-}\right) . \tag{14.9}$$

As a special case, take $H(t) = 1_{(t_{j-1},t_j]}(t)$, a left-continuous function of t. Then (14.8) reduces to

$$\begin{aligned} \mathrm{Var}\,(\Delta_j M \mid \mathcal{F}_{t_i}) &= \mathrm{E}\,(\langle M\rangle(t_j) - \langle M\rangle(t_{j-1}) \mid \mathcal{F}_{t_i}) \\ &= \mathrm{E}\left(\int_{(t_{j-1},t_j]} d\langle M\rangle \mid \mathcal{F}_{t_i}\right). \end{aligned} \tag{14.10}$$

C. Stopping times and optional sampling
So far the time points t_i have been taken as fixed. It turns out that all results stated above carry over to certain random times T_i, now to be defined. A nonnegative random variable T is an **F**-*stopping time* if $\{T \le t\} \in \mathcal{F}_t, \forall t \ge 0$, which means that the information provided by the description of the history by **F** suffices to ascertain at any time whether or not T has occurred. The collection of all events that are observable by time T is denoted by $\mathcal{F}_T$. Formally, these are the events $F \in \mathcal{F}$ such that $F \cap \{T \le t\} \in \mathcal{F}_t, \forall t$. Obviously they form a sigma-algebra.

Doob's fundamental optional sampling theorem states that (14.1) remains valid if s and t are replaced by **F**-stopping times S and T. More precisely, if $\{M(t)\}_{t\ge 0}$ is a right-continuous **F**-martingale, which converges to a random variable $M(\infty)$ with finite mean, then

$$\mathrm{E}\,(M(T) - M(S)|\mathcal{F}_S) = 0 \tag{14.11}$$

for S and T **F**-stopping times such that $S \le T$. Now, since all the results above rest on the martingale property (14.1), they remain valid if the t_i are replaced by **F**-stopping times T_i in the presence of (14.11).

By necessity the present brief survey is imprecise at some points. Recommendable rigorous expositions of martingale theory are e.g. the readable text by Gihman and Skorohod (1979) or the recent one by Protter (1990).

Acknowledgement

My thanks are due to Christian Max Mller for many useful discussions during the progress of this work.

References

Bremaud, P. (1981). *Point processes and queues.* Springer-Verlag, New York, Heidelberg, Berlin.

Gerber, H.U. (1979). *An Introduction to Mathematical Risk Theory.* Huebner Foundation, Univ. of Pennsylvania, Philadelphia.

Gerber, H.U. (1986). *Lebensversicherungsmathematik.* Springer-Verlag, Berlin, Heidelberg, New York.

Gihman, I.L. and Skorohod, A.V. (1979). *The Theory of Stochastic Processes Vol. 3.* Springer-Verlag, Heidelberg, New York.

Hattendorff, K. (1868). Das Risiko bei der Lebensversicherung. *Masius Rundschau der Versicherungen* **18**, 169-183.

Norberg, R. (1990). Payment measures, interest, and discounting – an axiomatic approach with applications to insurance. *Scand. Actuarial J.* **1990**, 14-33.

Norberg, R. (1991). Reserves in life and pension insurance. *Scand. Actuarial J.* **1991**, 1-22.

Papatriandafylou, A. and Waters, H.R. (1984). Martingales in life insurance. *Scand. Actuarial J.* **1984**, 210-230.

Protter, P. (1990). *Stochastic Integration and Differential Equations.* Springer-Verlag, Berlin, New York.

Ramlau-Hansen, H. (1988). Hattendorff's theorem: a Markov chain and counting process approach. *Scand. Actuarial J.* **1988**, 143-156.

Wolthuis, H. (1987). Hattendorff's theorem for a continuous-time Markov model. *Scand. Actuarial J.* **1987**, 157-175.

Addendum to Hattendorff's Theorem and Thiele's Differential Equation generalized

RAGNAR NORBERG

A. Purpose of this note

Any specific application of the theory in Section 3 of the paper would demand that the statewise reserves V_j be precisely defined. There is some latitude at this point, however, and it turns out that Theorem 3 as stated may require that an appropriate definition be used. Paragraph B of the present note adds rigour on this issue. Paragraph C offers some guidance as to how to construct and compute the reserves in nontrivial cases. Some technical lemmas are placed in the final Paragraph D.

Notation and results from the background paper are used without further explanation (formula numbers in that paper indicate section and formula within the section separated by decimal point, in contrast to the integer formula numbers internal to the present note). We shall feel free also to employ certain basic concepts and results in the theory of point processes and their martingales as presented in Paragraphs II.1-4 of Andersen & al. A number of statements about paths of stochastic processes are tacitly understood to be with the qualification 'almost surely'. We shall also frequently use the fact that we are dealing only with processes of bounded variation without explicitly saying so.

B. Supplement to Section 3

The history of the state process throughout the time interval $(0, t]$ is completely described by the present state, $X(t)$, the time of last entry into the present state, $S(t)$, and the "strict past" made up by the times of (possible) previous transitions and the state arrived at upon each of those transitions (define the first transition to take place at time 0, taking the process into state 0). Therefore, fairly generally, the reserve (2.2) is of the form

$$V_{\mathbf{F}}(t) = f(S(t), t, U(t), X(t)), \tag{14.12}$$

where U(t) is some (finite-dimensional) predictable semimartingale. Now, (14.12) may be cast as

$$V_{\mathbf{F}}(t) = \sum_j I_j(t) f(S(t), t, U(t), j) = \sum_j I_j(t) f(S_j(t), t, U(t), j), \tag{14.13}$$

where

$$S_j(t) = I_j(t)S(t) + (1 - I_j(t))t \tag{14.14}$$

is $S(t)$ during sojourns in state j and t otherwise. The statewise reserves introduced at the end of Paragraph 3A may conveniently be defined as

$$V_j(t) = f(S_j(t), t, U(t), j), \tag{14.15}$$

and, accordingly, $W_j(t) = v(t)V_j(t)$.

Obviously, the definition of $S_j(t)$ at times t when $I_j(t) = 0$ is immaterial. The choice (14.14) reflects the fact that, even if we cannot predict times of transition of the policy, we know just prior to any transition what $S(t)$ will be and hence what new value the reserve will assume upon the transition. The semimartingale $S_j(t)$ is continuous at (the random) times of transition into state j and also everywhere else except at times t of transition out of state j, where it jumps from $S(t)$ $(< t)$ to t.

We are now in a position to explicate the arguments following (3.11) and leading to Theorems 2 and 3. First, noting that $\sum_k I_k \equiv 1$, put

$$\begin{aligned} I_j(t)dW_j(t) &= \sum_{k;k\neq j} I_k(t-)I_j(t)(W_j(t) - W_j(t-)) + I_j(t-)I_j(t)dW_j(t) \\ &= \sum_{k;k\neq j} dN_{kj}(t)v(t)(f(t,t,U(t),j) - f(t-,t-,U(t-),j)) \quad (14.16) \\ &\quad + I_j(t-)I_j(t)dW_j(t). \quad (14.17) \end{aligned}$$

where we have used that, as a matter of definition, $S_j(\tau) = \tau$ for all τ in the non-degenerate interval $[S(t-), t]$ if the policy enters state j at time $t > 0$. By Lemma 1 below (in Paragraph D), the predictable process $f(t, t, U(t), j)$ has no jumps coinciding with jumps of the counting processes, and so the sum in (14.16) is 0. By Lemma 3 below, the term in (14.17) is predictable (or, more presicely, the increment of a predictable process). It follows that $I_j(t)dW_j(t)$ is predictable. Finally, by Lemma 2 below, $I_j(t)d\tilde{A}_j^\circ(t)$ is predictable since $d\tilde{A}_j^\circ(t)$ is. Comparison of the predictable (bounded variation) parts of (3.4) and (3.11) leads to Theorems 2 and 3.

With the present definition of the statewise reserves it makes no difference if we replace $W_k(t-)$ with $W_k(t)$ in (3.12) (and $V_k(t)$ with $V_k(t-)$ in (3.14)). This observation settles the apparent disagreement with the result (3.5) in Møller (1993). He considers the case where the contractual payment functions, the transition intensities and, hence, the reserve depend only on t and $S(t)$ or, equivalently, on t and $t - S(t)$ (he uses $U(t)$ to denote the latter).

C. Construction of the reserve

Usually Thiele's differential equations are mobilized when the reserve cannot be put up by a direct prospective argument, typically when the payments depend on the current reserve. However, as long as the payments and the transition intensities do not depend on the past, the statewise reserves remain functions of t only and can, therefore, be determined by a set of simple differential equations. Real difficulties arise only when payments and transition intensities are allowed to depend on the past in a more or less complex manner; then the reserves will be functions of several variables and it is rather obvious that they cannot be determined by just a set of first order ordinary differential equations.

Let us look briefly at the case mentioned at the balance of the previous paragraph, where payments, intensities, and hence the reserve depend on $S(t)$ (and t). We seek the two-dimensional statewise discounted reserve-functions $W_j(s,t)$ (say), $0 \leq s \leq t \leq n$, $j \in \mathcal{J}$, where n is the time of expiry of the contract. Conditioning on whether or not there is a transition out of state j within time n and, in case there is, the time and the direction of the first such transition, we obtain the integral equation

$$\begin{aligned} W_j(s,t) &= \int_t^n e^{-\int_t^\tau \mu_{j\cdot}(s,u)du} \sum_{k;k\neq j} \mu_{jk}(s,\tau)d\tau \\ &\quad \cdot(\tilde{A}_j^\circ(s,\tau) - \tilde{A}_j^\circ(s,t) + \tilde{a}_{jk}^\circ(s,\tau) + W_k(\tau,\tau)) \\ &\quad + e^{-\int_t^n \mu_{j\cdot}(s,u)du}(\tilde{A}_j^\circ(s,n) - \tilde{A}_j^\circ(s,t)). \end{aligned} \tag{14.18}$$

Again, we see that $W_k(\tau,\tau)$ can be replaced with $W_k(\tau-,\tau-)$ since the integration with respect to $d\tau$ annihilates the at most countable number of differences between them. To determine the reserve functions, solve first the $W_j(t,t)$ from the integral equations with $s = t$, and then solve the $W_j(s,t)$ from the general equations.

To obtain a differential form of the integral equation (14.18), add $\tilde{A}_j^\circ(s,t)$ and multiply by $e^{-\int_0^t \mu_{j\cdot}(s,u)du}$ on both sides, then differentiate with respect to t, and finally divide by the common factor $e^{-\int_0^t \mu_{j\cdot}(s,u)du}$, which gives the following equivalent of (3.15):

$$d_t(W_j(s,t) + \tilde{A}_j^\circ(s,t)) = -\sum_{k;k\neq j} \mu_{jk}(s,t)dt(\tilde{a}_{jk}^\circ(s,t) + W_k(t,t) - W_j(s,t)). \tag{14.19}$$

D. Some auxiliary results

Lemma 1. *If H is a predictable process, then*

$$\int_0^t H(\tau-)dN_{jk}(\tau) = \int_0^t H(\tau)dN_{jk}(\tau). \tag{14.20}$$

Consequently, H has no jumps in common with the counting processes.

Proof: Let M_{jk} be the martingale defined by $dM_{jk}(t) = dN_{jk}(t) - \lambda_{jk}(t)dt$. The difference between the two integrals in (14.20) is

$$\begin{aligned}\int_0^t (H(\tau) - H(\tau-))dN_{jk}(\tau) &= \int_0^t (H(\tau) - H(\tau-))\lambda_{jk}(\tau)d\tau \\ &\quad + \int_0^t (H(\tau) - H(\tau-))dM_{jk}(\tau).\end{aligned}$$

The first term on the right is zero because the integrand differs from 0 at most at a countable set of points. Likewise, the second term is also 0 because its variance is

$$\mathrm{E}\int_0^t (H(\tau) - H(\tau-))^2\lambda_{jk}(\tau)d\tau.$$

The last assertion is an obvious implication of the argument above. □

Lemma 2. *If A is a predictable semimartingale, then so is also the process H defined by $dH(t) = I_j(t)dA(t)$.*

Proof: By the rule of integration by parts, we have the identities

$$dH(t) = dI_j(t)A(t-) + I_j(t)dA(t) = dI_j(t)A(t) + I_j(t-)dA(t).$$

Forming the difference between the last two expressions and using (3.1), gives

$$I_j(t)dA(t) = I_j(t-)dA(t) + \sum_{k;k\neq j}(A(t) - A(t-))(dN_{kj}(t) - dN_{jk}(t)).$$

The first term on the right is predictable and the second term is 0 by Lemma 1. □

Lemma 3. *If G is a semimartingale, then the process H defined by $dH(t) = I_j(t-)I_j(t)dG(t)$ is predictable.*

Proof: Inserting the Doob-Meyer decomposition

$$dG(t) = dG_0(t) + \sum_{i\neq k} G_{ik}(t)(dN_{ik}(t) - \lambda_{ik}(t)dt),$$

with $G_0(t)$ and the $G_{ik}(t)$ predictable, and noting that $I_j(t-)I_j(t)dN_{ik}(t) = 0$, we get

$$dH(t) = I_j(t-)I_j(t)dG_0(t) - \sum_{i\neq k} G_{ik}(t)\lambda_{ik}(t)dt.$$

Here the first term is predictable by Lemma 2, and the second term is certainly predictable as it is continuous. □

Chapter 15

Financial mathematics in insurance

15.1 Finance in insurance

Finance was always an essential part of insurance. Trivially, one might say, because any business has to attend to its money affairs. However, for at least two reasons, insurance is not just any business. In the first place insurance products are not physical goods or services, but financial contracts with obligations related to uncertain future events. Therefore, pricing is not just a piece of accountancy involving the four basic arithmetical operations, but requires assessment and management of risk by sophisticated mathematical models and methods. In the second place, insurance contracts are more or less long term (in life insurance for up to several decades), and they are typically paid in advance (hence the term 'premium' for 'price' derived from French 'prime' for 'first'). Therefore, the insurance industry is a major accumulator of capital, and insurance companies especially pension funds are major institutional investors in today's society. It follows that the financial operations (investment strategy) of an insurance company may be as decisive of its revenues as its insurance operations (design of products, risk management, premium rating, procedures of claims assessment, and the pure randomness in the claims process). Accordingly one speaks of *assets risk* or *financial risk* and *liability risk* or *insurance risk*. We anticipate here that financial risk may well be the more severe: Insurance risk created by random deviations of individual claims from their expected is *diversifiable* in the sense that, by the law of large numbers, it can be eliminated in a sufficiently large insurance portfolio. This notion of diversification does not apply to catastrophe coverage and it does not account for the risk associated with long term contracts

insurance

ic insurance risk

Financial risk created random economic events are

by booms, recessions and, rare but disastrous, crashes in the market as a whole is not diversifiable; it is the uncertain events are part of our the world history non-replicable averaged out in any meaningful sense. Financial risk created by the day-to-day rises and falls of individual stocks, is not *diversifiable* on large random ups and downs is held to be *indiversifiable* since the entire portfolio is affected by the development of the economy.

On this background one may ask why insurance mathematics traditionally centers on measurement and control of the insurance risk. The answer may partly be found in institutional circumstances: The insurance industry used to be heavily regulated, solvency being the primary concern of the regulatory authority. Possible adverse developments of economic factors (e.g. inflation, weak returns on investment, low interest rates, etc.) would be safeguarded against by placing premiums on the safe side. The comfortable surpluses, which would typically accumulate under this regime, were redistributed as bonuses (dividends) to the policyholders only in arrears, after interest and other financial parameters had been observed. Furthermore, the insurance industry used to be separated from other forms of business and protected from competition within itself, and severe restrictions were placed on its investment operations. In these circumstances financial matters appeared to be something the traditional actuary did not need to worry about. Another reason why insurance mathematics used to be void of financial considerations was, of course, the absence of a well developed theory for description and control of financial risk.

All this has changed. National and institutional borders have been downsized or eliminated and regulations have been liberalized: Mergers between insurance companies and banks are now commonplace, new insurance products are being created and put on the market virtually every day, by insurance companies and other financial institutions as well, and without prior licencing by the supervisory authority. The insurance companies of today find themselves placed on a fiercely competitive market. Many new products are directly linked to economic indices, like unit-linked life insurance and catastrophe derivatives. By so-called securitization also insurance risk can be put on the market and thus open new possibilities of inviting investors from outside to participate in risk that previously had to be shared solely between the participants in the insurance insurance schemes. These developments in practical insurance coincide with the advent of modern financial mathematics, which has equipped the actuaries with a well developed theory within which financial risk and insurance risk can be analyzed, quantified and controlled.

A new order of the day is thus set for the actuarial profession. The purpose of this chapter is to give a glimpse into some basic ideas and results in modern financial mathematics and to indicate by examples how they may be applied to actuarial problems involving management of financial risk.

15.2 Prerequisites

A. Probability and expectation.

Taking basic measure theoretic probability as a prerequisite, we represent the relevant part of the world and its uncertainties by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Here Ω is the set of possible outcomes ω, $\mathcal{F}$ is a sigmaalgebra of subsets of Ω representing the events to which we want to assign probabilities, and $\mathbb{P} : \mathcal{F} \mapsto [0, 1]$ is a probability measure.

A set $A \in \mathcal{F}$ such that $\mathbb{P}[A] = 0$ is called a *nullset*, and a property that takes place in all of Ω, except possibly on a nullset, is said to hold *almost surely* (a.s.). If more than one probability measure are in play, we write "nullset ($\mathbb{P}$)" and "a.s. ($\mathbb{P}$)" whenever emphasis is needed. Two probability measures $\tilde{\mathbb{P}}$ and $\mathbb{P}$ are said to be *equivalent*, written $\tilde{\mathbb{P}} \sim \mathbb{P}$, if they are defined on the same $\mathcal{F}$ and have the same nullsets.

Let $\mathcal{G}$ be some sub-sigmaalgebra of $\mathcal{F}$. We denote the restriction of $\mathbb{P}$ to $\mathcal{G}$ by $\mathbb{P}_{\mathcal{G}}$; $\mathbb{P}_{\mathcal{G}}[A] = \mathbb{P}[A]$, $A \in \mathcal{G}$. Note that also $(\Omega, \mathcal{G}, \mathbb{P}_{\mathcal{G}})$ is a probability space.

A $\mathcal{G}$-measurable random variable (r.v.) is a function $X : \Omega \mapsto \mathbb{R}$ such that $X^{-1}(B) \in \mathcal{G}$ for all $B \in \mathcal{R}$, the Borel sets in $\mathbb{R}$. We write $X \in \mathcal{G}$ in short.

The expected value of a r.v. X is the probability-weighted average $\mathbb{E}[X] = \int X \, d\mathbb{P} = \int_\Omega X(\omega) \, d\mathbb{P}(\omega)$, provided this integral is well defined.

The conditional expected value of X, given $\mathcal{G}$, is the r.v. $\mathrm{E}[X|\mathcal{G}] \in \mathcal{G}$ satisfying

$$\mathbb{E}\{\mathbb{E}[X|\mathcal{G}] \, Y\} = \mathbb{E}[XY] \tag{15.1}$$

for each $Y \in \mathcal{G}$ such that the expected value on the right exists. It is unique up to nullsets ($\mathbb{P}$). To motivate (15.1), consider the special case when $\mathcal{G} = \sigma\{B_1, B_2, \ldots\}$, the sigma-algebra generated by the $\mathcal{F}$-measurable sets $B_1, B_2, \ldots$, which form a partition of Ω. Being $\mathcal{G}$-measurable, $\mathbb{E}[X|\mathcal{G}]$ must be of the form $\sum_k b_k 1[B_k]$. Putting this together with $Y = 1[B_j]$ into the relationship (15.1) we arrive at

$$\mathbb{E}[X|\mathcal{G}] = \sum_j 1[B_j] \frac{\int_{B_j} X \, d\mathbb{P}}{\mathbb{P}[B_j]} \, ,$$

as it ought to be. In particular, taking $X = 1[A]$, we find the conditional probability $\mathbb{P}[A|B] = \mathbb{P}[A \cap B]/\mathbb{P}[B]$.

One easily verifies the *rule of iterated expectations*, which states that, for $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$,

$$\mathbb{E}\left\{ \mathbb{E}[X|\mathcal{G}] \middle| \, \mathcal{H} \right\} = \mathbb{E}[X|\mathcal{H}] \, . \tag{15.2}$$

F. Change of measure.

If L is a r.v. such that $L \geq 0$ a.s. ($\mathbb{P}$) and $\mathbb{E}[L] = 1$, we can define a probability measure $\tilde{\mathbb{P}}$ on $\mathcal{F}$ by

$$\tilde{\mathbb{P}}[A] = \int_A L \, d\mathbb{P} = \mathbb{E}[1[A]L] \, . \tag{15.3}$$

If $L > 0$ a.s. ($\mathbb{P}$), then $\tilde{\mathbb{P}} \sim \mathbb{P}$.

The expected value of X w.r.t. $\tilde{\mathbb{P}}$ is

$$\tilde{\mathbb{E}}[X] = \mathbb{E}[XL] \tag{15.4}$$

if this integral exists; by the definition (15.3), the relation (15.4) is true for indicators, hence for simple functions and, by passing to limits, it holds for measurable functions. Spelling out (15.4) as $\int X \, d\tilde{\mathbb{P}} = \int XL \, d\mathbb{P}$ suggests the notation $d\tilde{\mathbb{P}} = L \, d\mathbb{P}$ or

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = L \, . \tag{15.5}$$

The function L is called the *Radon-Nikodym derivative* of $\tilde{\mathbb{P}}$ w.r.t. $\mathbb{P}$.

Conditional expectation under $\tilde{\mathbb{P}}$ is formed by the rule

$$\tilde{\mathbb{E}}[X|\mathcal{G}] = \frac{\mathbb{E}[XL|\mathcal{G}]}{\mathbb{E}[L|\mathcal{G}]} \, . \tag{15.6}$$

To see this, observe that, by definition,

$$\tilde{\mathbb{E}}\{\tilde{\mathbb{E}}[X|\mathcal{G}] \, Y\} = \tilde{\mathbb{E}}[XY] \tag{15.7}$$

for all $Y \in \mathcal{G}$. The expression on the left of (15.7) can be reshaped as

$$\mathbb{E}\{\tilde{\mathbb{E}}[X|\mathcal{G}]\,YL\} = \mathbb{E}\{\tilde{\mathbb{E}}[X|\mathcal{G}]\,\mathbb{E}[L|\mathcal{G}]\,Y\}\,.$$

The expression on the right of (15.7) is

$$\mathbb{E}[XYL] = \mathbb{E}\{\mathbb{E}[XL|\mathcal{G}]\,Y\}\,.$$

It follows that (15.7) is true for all $Y \in \mathcal{G}$ if and only if

$$\tilde{\mathbb{E}}[X|\mathcal{G}]\,\mathbb{E}[L|\mathcal{G}] = \mathbb{E}[XL|\mathcal{G}]\,,$$

which is the same as (15.6).

For $X \in \mathcal{G}$ we have

$$\tilde{\mathbb{E}}_{\mathcal{G}}[X] = \tilde{\mathbb{E}}[X] = \mathbb{E}[XL] = \mathbb{E}\left\{X\,\mathbb{E}[L|\mathcal{G}]\right\} = \mathbb{E}_{\mathcal{G}}\left\{X\,\mathbb{E}[L|\mathcal{G}]\right\}\,, \tag{15.8}$$

showing that

$$\frac{d\tilde{\mathbb{P}}_{\mathcal{G}}}{d\mathbb{P}_{\mathcal{G}}} = \mathbb{E}[L|\mathcal{G}]\,. \tag{15.9}$$

B. Stochastic processes.

To describe the evolution of random phenomena over some time interval $[0, T]$, we introduce a family $\mathbf{F} = \{\mathcal{F}_t\}_{0\leq t\leq T}$ of sub-sigmaalgebras of $\mathcal{F}$, where $\mathcal{F}_t$ represents the information available at time t. More precisely, $\mathcal{F}_t$ is the set of events whose occurrence or non-occurrence can be ascertained by time t. If no information is ever sacrificed, we have $\mathcal{F}_s \subset \mathcal{F}_t$ for $s < t$. We then say that $\mathbf{F}$ is a *filtration*, and $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ is called a *filtered probability space*.

A stochastic process is a family of r.v.-s, $\{X_t\}_{0\leq t\leq T}$. It is said to be *adapted to* the filtration $\mathbf{F}$ if $X_t \in \mathcal{F}_t$ for each $t \in [0, T]$, that is, at any time the current state (and also the past history) of the process is fully known if we are currently provided with the information $\mathbf{F}$. An adapted process is said to be *predictable* if its value at any time is entirely determined by its history in the strict past, loosely speaking. For our purposes it is sufficient to think of predictable processes as being either left-continuous or deterministic.

C. Martingales.

An adapted process X with finite expectation is a *martingale* if

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s$$

for $s < t$. The martingale property depends both on the filtration and on the probability measure, and when these need emphasis, we shall say that X is martingale $(\mathbf{F}, \mathbb{P})$. The definition says that, "on the average", a martingale is always expected to remain on its current level. One easily verifies that, conditional on the present information, a martingale has uncorrelated future increments. Any integrable r.v. Y induces a martingale $\{X_t\}_{t\geq 0}$ defined by $X_t = \mathbb{E}[Y|\mathcal{F}_t]$, a consequence of (15.2).

Abbreviate $\mathbb{P}_t = \mathbb{P}_{\mathcal{F}_t}$, introduce

$$L_t = \frac{d\tilde{\mathbb{P}}_t}{d\mathbb{P}_t}\,,$$

and put $L = L_T$. By (15.9) we have

$$L_t = \mathbb{E}[L|\mathcal{F}_t]\,, \tag{15.10}$$

which is a martingale $(\mathbf{F}, \mathbb{P})$.

D. Counting processes.
As the name suggests, a counting process is a stochastic process $N = \{N_t\}_{0 \le t \le T}$ that commences from zero ($N_0 = 0$) and thereafter increases by isolated jumps of size 1 only. The *natural filtration* of N is $\mathbf{F}^N = \{\mathcal{F}_t^N\}_{0 \le t \le T}$, where $\mathcal{F}_t^N = \sigma\{N_s; s \le t\}$ is the history of N by time t. This is the smallest filtration to which N is adapted. The strict past history of N at time t is denoted by $\mathcal{F}_{t-}^N$.

An $\mathbf{F}^N$-predictable process $\{\Lambda_t\}_{0 \le t \le T}$ is called a *compensator* of N if the process M defined by

$$M_t = N_t - \Lambda_t \tag{15.11}$$

is a zero mean $\mathbf{F}^N$-martingale. If Λ is absolutely continuous, that is, of the form

$$\Lambda_t = \int_0^t \lambda_s \, ds\,,$$

then the process λ is called the *intensity* of N. We may also define the intensity informally by

$$\lambda_t \, dt = \mathbb{P}\left[dN_t = 1 \,|\, \mathcal{F}_{t-}\right] = \mathbb{E}\left[dN_t \,|\, \mathcal{F}_{t-}\right]\,,$$

and we sometimes write the associated martingale (15.11) in differential form,

$$dM_t = dN_t - \lambda_t \, dt\,. \tag{15.12}$$

A stochastic integral w.r.t. the martingale M is an $\mathbf{F}^N$-adapted process H of the form

$$H_t = H_0 + \int_0^t h_s \, dM_s\,, \tag{15.13}$$

where H_0 is $\mathcal{F}_0^N$-measurable and h is an $\mathbf{F}^N$-predictable process such that H is integrable. The stochastic integral is also a martingale.

A fundamental representation result states that every $\mathbf{F}^N$ martingale is a stochastic integral w.r.t. M. It follows that every integrable $\mathcal{F}_t^N$ measurable r.v. is of the form (15.13).

If $H_t^{(1)} = H_0^{(1)} + \int_0^t h_s^{(1)} \, dM_s$ and $H_t^{(2)} = H_0^{(2)} + \int_0^t h_s^{(2)} \, dM_s$ are stochastic integrals with finite variance, then an easy heuristic calculation shows that

$$\mathrm{Cov}[H_T^{(1)}, H_T^{(2)}|\mathcal{F}_t] = \mathbb{E}\left[\int_t^T h_s^{(1)} h_s^{(2)} \lambda_s \, ds \,|\, \mathcal{F}_t\right]\,, \tag{15.14}$$

and, in particular,

$$\mathrm{Var}[H_T|\mathcal{F}_t] = \mathbb{E}\left[\int_t^T h_s^2 \lambda_s \, ds \,|\, \mathcal{F}_t\right]\,.$$

$H^{(1)}$ and $H^{(2)}$ are said to be orthogonal if they have conditionally uncorrelated increments, that is, the covariance in (15.14) is null. This is equivalent to saying that $H^{(1)}\,H^{(2)}$ is a martingale.

The intensity is also called the infinitesimal characteristic if the counting process since it entirely determines it probabilistic properties. If λ_t is deterministic, then N_t is a Poisson process. If λ depends only on N_{t-}, then N_t is a Markov process.

A comprehensive textbook on counting processes in life history analysis is [3].

E. The Girsanov transform.

Girsanov's theorem is a celebrated one in stochastics, and it is basic in mathematical finance. We formulate and prove the counting process variation:

Theorem *(Girsanov). Let N_t be a counting process with $(\mathbf{F}, \mathbb{P})$-intensity λ_t, and let $\tilde{\lambda}_t$ be a given non-negative $\mathbf{F}$-adapted process such that $\tilde{\lambda}_t = 0$ if and only if $\lambda_t = 0$. Then there exists a probability measure $\tilde{\mathbb{P}}$ such that $\tilde{\mathbb{P}} \sim \mathbb{P}$ and N has $(\mathbf{F}, \tilde{\mathbb{P}})$-intensity $\tilde{\lambda}_t$. The likelihood process (15.10) is*

$$L_t = \exp\left(\int_0^t (\ln \tilde{\lambda}_s - \ln \lambda_s)\, dN_s \; + \int_0^t (\lambda_s - \tilde{\lambda}_s)\, ds\right) .$$

Proof: We shall give a constructive proof, starting from a guessed L in (15.5). Since L must be strictly positive a.e. $(\mathbb{P})$, a candidate would be $L = L_T$, where

$$L_t = \exp\left(\int_0^t \phi_s\, dN_s + \int_0^t \psi_s\, ds\right)$$

with ϕ predictable and ψ adapted.

In the first place, L_t should be a martingale $(\mathbf{F}, \mathbb{P})$. By Itô's formula,

$$\begin{aligned} dL_t &= L_t\, \psi_t\, dt + L_{t-}(e^{\phi_t} - 1)\, dN_t \\ &= L_t\left(\psi_t + \left(e^{\phi_t} - 1\right)\lambda_t\right) dt + L_{t-}\left(e^{\phi_t} - 1\right) dM_t\,. \end{aligned}$$

The representation result (15.13) tells us that to make L a martingale, we must make the drift term vanish, that is,

$$\psi_t = \left(1 - e^{\phi_t}\right)\lambda_t\,, \tag{15.15}$$

whereby

$$dL_t = L_{t-}\left(e^{\phi_t} - 1\right) dM_t\,,$$

In the second place, we want to determine ϕ_t such that the process $\tilde{M}$ given by

$$d\tilde{M}_t = dN_t - \tilde{\lambda}_t\, dt \tag{15.16}$$

is a martingale $(\mathbf{F}, \tilde{\mathbb{P}})$. Thus, we should have $\tilde{\mathbb{E}}[\tilde{M}_t|\mathcal{F}_s] = \tilde{M}_s$ or, by (15.6),

$$\frac{\mathbb{E}\left[\tilde{M}_t L|\mathcal{F}_s\right]}{\mathbb{E}\left[L|\mathcal{F}_s\right]} = \tilde{M}_s\,.$$

Using the martingale property (15.10) of L_t, this is the same as

$$\mathbb{E}\left[\tilde{M}_t L_t|\mathcal{F}_s\right] = \tilde{M}_s L_s$$

i.e. $\tilde{M}_t L_t$ should be a martingale $(\mathbf{F}, \mathbb{P})$. Since

$$\begin{aligned} d(\tilde{M}_t L_t) &= (-\tilde{\lambda}_t\, dt) L_t + \tilde{M}_t (e^{\phi_t} - 1) L_t (-\lambda_t\, dt) \\ &\quad + \Big((\tilde{M}_{t-} + 1) L_{t-} e^{\phi_t} - \tilde{M}_{t-} L_{t-} \Big)\, dN_t \\ &= L_t\, dt \Big(-\tilde{\lambda}_t + e^{\phi_t} \lambda_t \Big) + \Big((\tilde{M}_{t-} + 1) L_{t-} e^{\phi_t} - \tilde{M}_{t-} L_{t-} \Big)\, dM_t\,. \end{aligned}$$

we conclude that the martingale property is obtained by choosing $\phi_t = \ln \tilde{\lambda}_t - \ln \lambda_t$.

The multivariate case goes in the same way; just replace by vector-valued processes.

15.3 A Markov chain financial market - Introduction

A. Motivation.
The theory of diffusion processes, with its wealth of powerful theorems and model variations, is an indispensable toolkit in modern financial mathematics. The seminal papers of Black and Scholes [10] and Merton [32] were crafted with Brownian motion, and so were most of the almost countless papers on arbitrage pricing theory and its bifurcations that followed over the past quarter of a century.

A main course of current research, initiated by the martingale approach to arbitrage pricing ([24] and [25]), aims at generalization and unification. Today the core of the matter is well understood in a general semimartingale setting, see e.g. [13]. Another course of research investigates special models, in particular various Levy motion alternatives to the Brownian driving process, see e.g. [18] and [40]. Pure jump processes have been widely used in finance, ranging from plain Poisson processes introduced in [33] to quite general marked point processes, see e.g. [8]. And, as a pedagogical exercise, the market driven by a binomial process has been intensively studied since it was launched in [12].

The present paper undertakes to study a financial market driven by a continuous time homogeneous Markov chain. The idea was launched in [39] and reappeared in [19], the context being limited to modelling of the spot rate of interest. The purpose of the present study is two-fold: In the first place, it is instructive to see how well established theory turns out in the framework of a general Markov chain market. In the second place, it is worthwhile investigating the feasibility of the model from a theoretical as well as from a practical point of view. Poisson driven markets are accommodated as special cases.

B. Preliminaries: Notation and some useful results.
Vectors and matrices are denoted by in bold letters, lower and upper case, respectively. They may be equipped with topscripts indicating dimensions, e.g. $\mathbf{A}^{n\times m}$ has n rows and m columns. We may write $\mathbf{A} = (a^{jk})_{j\in\mathcal{J}}^{k\in\mathcal{K}}$ to emphasize the ranges of the row index j and the column index k. The transpose of $\mathbf{A}$ is denoted by $\mathbf{A}'$. Vectors are invariably taken to be of column type, hence row vectors appear as transposed. The identity matrix is denoted by $\mathbf{I}$, the vector with all entries equal to 1 is denoted by $\mathbf{1}$, and the vector with all entries equal to 0 is denoted by $\mathbf{0}$. By $\mathbf{D}_{j=1,\dots,n}(a^j)$, or just $\mathbf{D}(\boldsymbol{a})$, is meant the diagonal matrix with the entries of $\boldsymbol{a} = (a^1, \dots, a^n)'$ down the principal diagonal. The n-dimensional Euclidean space is denoted by $\mathbb{R}^n$, and the linear subspace spanned by the columns of $\mathbf{A}^{n\times m}$ is denoted by $\mathbb{R}(\mathbf{A})$.

A diagonalizable square matrix $\mathbf{A}^{n\times n}$ can be represented as

$$\mathbf{A} = \boldsymbol{\Phi}\,\mathbf{D}_{j=1,\ldots,n}(\rho_j)\,\boldsymbol{\Phi}^{-1} = \sum_{j=1}^{n} \rho_j \boldsymbol{\phi}_j \boldsymbol{\psi}_j'\,, \tag{15.1}$$

where the $\boldsymbol{\phi}_j$ are the columns of $\boldsymbol{\Phi}^{n\times n}$ and the $\boldsymbol{\psi}_j'$ are the rows of $\boldsymbol{\Phi}^{-1}$. The ρ_j are the eigenvalues of $\mathbf{A}$, and $\boldsymbol{\phi}_j$ and $\boldsymbol{\psi}_j'$ are the corresponding right and left eigenvectors, respectively. Eigenvectors (right or left) corresponding to eigenvalues that are distinguishable and non-null are mutually orthogonal. These results can be looked up in e.g. [30].

The exponential function of $\mathbf{A}^{n\times n}$ is the $n \times n$ matrix defined by

$$\exp(\mathbf{A}) = \sum_{p=0}^{\infty} \frac{1}{p!}\mathbf{A}^p = \boldsymbol{\Phi}\,\mathbf{D}_{j=1,\ldots,n}(e^{\rho_j})\,\boldsymbol{\Phi}^{-1} = \sum_{j=1}^{n} e^{\rho_j} \boldsymbol{\phi}_j \boldsymbol{\psi}_j'\,, \tag{15.2}$$

where the last two expressions follow from (15.1). The matrix $\exp(\mathbf{A})$ has full rank.

If $\boldsymbol{\Lambda}^{n\times n}$ is positive definite symmetric, then $\langle \boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2\rangle_{\boldsymbol{\Lambda}} = \boldsymbol{\zeta}_1'\boldsymbol{\Lambda}\boldsymbol{\zeta}_2$ defines an inner product on $\mathbb{R}^n$. The corresponding norm is given by $\|\boldsymbol{\zeta}\|_{\boldsymbol{\Lambda}} = \langle \boldsymbol{\zeta}, \boldsymbol{\zeta}\rangle_{\boldsymbol{\Lambda}}^{1/2}$. If $\mathbf{F}^{n\times m}$ has full rank m $(\le n)$, then the $\langle\,\cdot\,,\,\cdot\,\rangle_{\boldsymbol{\Lambda}}$-projection of $\boldsymbol{\zeta}$ onto $\mathbb{R}(\mathbf{F})$ is

$$\boldsymbol{\zeta}_{\mathbf{F}} = \mathbf{P}_{\mathbf{F}}\boldsymbol{\zeta}\,, \tag{15.3}$$

where the projection matrix (or projector) $\mathbf{P}_{\mathbf{F}}$ is

$$\mathbf{P}_{\mathbf{F}} = \mathbf{F}(\mathbf{F}'\boldsymbol{\Lambda}\mathbf{F})^{-1}\mathbf{F}'\boldsymbol{\Lambda}\,. \tag{15.4}$$

The projection of $\boldsymbol{\zeta}$ onto the orthogonal complement $\mathbb{R}(\mathbf{F})^{\perp}$ is

$$\boldsymbol{\zeta}_{\mathbf{F}^\perp} = \boldsymbol{\zeta} - \boldsymbol{\zeta}_{\mathbf{F}} = (\mathbf{I} - \mathbf{P}_{\mathbf{F}})\boldsymbol{\zeta}\,.$$

Its squared length, which is the squared $\langle\,\cdot\,,\,\cdot\,\rangle_{\boldsymbol{\Lambda}}$-distance from $\boldsymbol{\zeta}$ to $\mathbb{R}(\mathbf{F})$, is

$$\|\boldsymbol{\zeta}_{\mathbf{F}^\perp}\|_{\boldsymbol{\Lambda}}^2 = \|\boldsymbol{\zeta}\|_{\boldsymbol{\Lambda}}^2 - \|\boldsymbol{\zeta}_{\mathbf{F}}\|_{\boldsymbol{\Lambda}}^2 = \boldsymbol{\zeta}'\boldsymbol{\Lambda}(\mathbf{I} - \mathbf{P}_{\mathbf{F}})\boldsymbol{\zeta}\,. \tag{15.5}$$

The cardinality of a set $\mathcal{Y}$ is denoted by $|\mathcal{Y}|$. For a finite set it is just its number of elements.

15.4 The Markov chain market

A. The continuous time Markov chain.
At the base of everything (although slumbering in the background) is some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $\{Y_t\}_{t\ge 0}$ be a continuous time Markov chain with finite state space $\mathcal{Y} = \{1,\ldots,n\}$. We assume that it is time homogeneous so that the transition probabilities

$$p_t^{jk} = \mathbb{P}[Y_{s+t} = k \mid Y_s = j]$$

depend only on the length of the transition period. This implies that the transition intensities

$$\lambda^{jk} = \lim_{t\searrow 0} \frac{p_t^{jk}}{t}\,, \tag{15.1}$$

$j \neq k$, exist and are constant. To avoid repetitious reminders of the type "$j,k \in \mathcal{Y}$", we reserve the indices j and k for states in $\mathcal{Y}$ throughout. We will frequently refer to

$$\mathcal{Y}^j = \{k; \lambda^{jk} > 0\} \,,$$

the set of states that are directly accessible from state j, and denote the number of such states by

$$n^j = |\mathcal{Y}^j| \,.$$

Put

$$\lambda^{jj} = -\lambda^{j\cdot} = -\sum_{k;k\in\mathcal{Y}^j} \lambda^{jk}$$

(minus the total intensity of transition out of state j). We assume that all states intercommunicate so that $p_t^{jk} > 0$ for all j,k (and $t > 0$). This implies that $n^j > 0$ for all j (no absorbing states). The matrix of transition probabilities,

$$\mathbf{P}_t = (p_t^{jk}) \,,$$

and the infinitesimal matrix,

$$\Lambda = (\lambda^{jk}) \,,$$

are related by (F.41), which in matrix form reads $\Lambda = \lim_{t\searrow 0} \frac{1}{t}(\mathbf{P}_t - \mathbf{I})$, and by the backward and forward Kolmogorov differential equations,

$$\frac{d}{dt}\mathbf{P}_t = \mathbf{P}_t\Lambda = \Lambda\mathbf{P}_t \,. \tag{15.2}$$

Under the side condition $\mathbf{P}_0 = \mathbf{I}$, (15.2) integrates to

$$\mathbf{P}_t = \exp(\Lambda t) \,. \tag{15.3}$$

In the representation (15.2),

$$\mathbf{P}_t = \mathbf{\Phi}\,\mathbf{D}_{j=1,\ldots,n}(e^{\rho_j t})\,\mathbf{\Phi}^{-1} = \sum_{j=1}^{n} e^{\rho_j t}\boldsymbol{\phi}_j\boldsymbol{\psi}_j' \,, \tag{15.4}$$

the first (say) eigenvalue is $\rho_1 = 0$, and corresponding eigenvectors are $\boldsymbol{\phi}_1 = \mathbf{1}$ and $\boldsymbol{\psi}_1' = (p^1,\ldots,p^n) = \lim_{t\nearrow\infty}(p_t^{j1},\ldots,p_t^{jn})$, the stationary distribution of Y. The remaining eigenvalues, $\rho_2,\ldots,\rho_n$, are all strictly negative so that, by (15.4), the transition probabilities converge exponentially to the stationary distribution as t increases.

Introduce

$$I_t^j = 1[Y_t = j] \,, \tag{15.5}$$

the indicator of the event that Y is in state j at time t, and

$$N_t^{jk} = |\{s;\, 0 < s \leq t,\, Y_{s-} = j \,,\, Y_s = k\}| \,, \tag{15.6}$$

the number of direct transitions of Y from state j to state $k \in \mathcal{Y}^j$ in the time interval $(0,t]$. For $k \notin \mathcal{Y}^j$ we define $N_t^{jk} \equiv 0$. Taking Y to be right-continuous, the same goes for the indicator processes I^j and the counting processes N^{jk}. As is seen from (15.5), (15.6), and the obvious relationships

$$Y_t = \sum_j jI_t^j \,, \qquad I_t^j = I_0^j + \sum_{k;k\neq j} (N_t^{kj} - N_t^{jk}) \,,$$

the state process, the indicator processes, and the counting processes carry the same information, which at any time t is represented by the sigma-algebra $\mathcal{F}_t^Y = \sigma\{Y_s; 0 \leq s \leq t\}$. The corresponding filtration, denoted by $\mathbf{F}^Y = \{\mathcal{F}_t^Y\}_{t\geq 0}$, is taken to satisfy the usual conditions of right-continuity ($\mathcal{F}_t = \cap_{u>t}\mathcal{F}_u$) and completeness ($\mathcal{F}_0$ contains all subsets of $\mathbb{P}$-nullsets), and $\mathcal{F}_0$ is assumed to be the trivial $(\emptyset, \Omega)$. This means, essentially, that Y is right-continuous (hence the same goes for the I^j and the N^{jk}) and that Y_0 deterministic.

The compensated counting processes M^{jk}, $j \neq k$, defined by

$$dM_t^{jk} = dN_t^{jk} - I_t^j \lambda^{jk}\, dt \tag{15.7}$$

and $M_0^{jk} = 0$, are zero mean, square integrable, mutually orthogonal martingales w.r.t. $(\mathbf{F}^Y, \mathbb{P})$.

We now turn to the subject matter of our study and, referring to introductory texts like [9] and [43], take basic notions and results from arbitrage pricing theory as prerequisites.

B. The continuous time Markov chain market.

We consider a financial market driven by the Markov chain described above. Thus, Y_t represents the state of the economy at time t, $\mathcal{F}_t^Y$ represents the information available about the economic history by time t, and $\mathbf{F}^Y$ represents the flow of such information over time.

In the market there are $m+1$ basic assets, which can be traded freely and frictionlessly (short sales are allowed, and there are no transaction costs). A special role is played by asset No. 0, which is a "locally risk-free" *bank account* with state-dependent interest rate

$$r_t = r^{Y_t} = \sum_j I_t^j r^j\,,$$

where the state-wise interest rates r^j, $j = 1, \ldots, n$, are constants. Thus, its price process is

$$B_t = \exp\left(\int_0^t r_s\, ds\right) = \exp\left(\sum_j r^j \int_0^t I_s^j\, ds\right),$$

with dynamics

$$dB_t = B_t\, r_t\, dt = B_t \sum_j r^j I_t^j\, dt\,.$$

(Setting $B_0 = 1$ a just a matter of convention.)

The remaining m assets, henceforth referred to as *stocks*, are risky, with price processes of the form

$$S_t^i = \exp\left(\sum_j \alpha^{ij} \int_0^t I_s^j\, ds + \sum_j \sum_{k\in\mathcal{Y}^j} \beta^{ijk} N_t^{jk}\right), \tag{15.8}$$

$i = 1, \ldots, m$, where the α^{ij} and β^{ijk} are constants and, for each i, at least one of the β^{ijk} is non-null. Thus, in addition to yielding state-dependent returns of the same form as the bank account, stock No. i makes a price jump of relative size

$$\gamma^{ijk} = \exp\left(\beta^{ijk}\right) - 1$$

upon any transition of the economy from state j to state k. By the general Itô's formula, its dynamics is given by

$$dS_t^i = S_{t-}^i \left(\sum_j \alpha^{ij} I_t^j \, dt + \sum_j \sum_{k \in \mathcal{Y}^j} \gamma^{ijk} dN_t^{jk} \right) . \tag{15.9}$$

Taking the bank account as numeraire, we introduce the discounted stock prices $\tilde{S}_t^i = S_t^i / B_t$, $i = 0, \ldots, m$. (The discounted price of the bank account is $\tilde{B}_t \equiv 1$, which is certainly a martingale under any measure). The discounted stock prices are

$$\tilde{S}_t^i = \exp\left(\sum_j (\alpha^{ij} - r^j) \int_0^t I_s^j \, ds + \sum_j \sum_{k \in \mathcal{Y}^j} \beta^{ijk} N_t^{jk} \right) , \tag{15.10}$$

with dynamics

$$d\tilde{S}_t^i = \tilde{S}_{t-}^i \left(\sum_j (\alpha^{ij} - r^j) I_t^j \, dt + \sum_j \sum_{k \in \mathcal{Y}^j} \gamma^{ijk} dN_t^{jk} \right) , \tag{15.11}$$

$i = 1, \ldots, m$.

We stress that the theory we are going to develop does not aim at explaining how the prices of the basic assets emerge from supply and demand, business cycles, investment climate, or whatever; they are exogenously given basic entities. (And God said "let there be light", and there was light, and he said "let there also be these prices".) The purpose of the theory is to derive principles for consistent pricing of financial contracts, derivatives, or claims in a given market.

C. Portfolios.

A dynamic *portfolio* or *investment strategy* is an $m+1$-dimensional stochastic process

$$\boldsymbol{\theta}_t' = (\eta_t, \boldsymbol{\xi}_t') ,$$

where η_t represents the number of units of the bank account held at time t, and the i-th entry in

$$\boldsymbol{\xi}_t = (\xi_t^1, \ldots, \xi_t^m)'$$

represents the number of units of stock No. i held at time t. As it will turn out, the bank account and the stocks will appear to play different parts in the show, the latter being the more visible. It is, therefore, convenient to costume the two types of assets and their corresponding portfolio entries accordingly. To save notation, however, it is useful also to work with double notation

$$\boldsymbol{\theta}_t = (\theta_t^0, \ldots, \theta_t^m)' ,$$

with $\theta_t^0 = \eta_t$, $\theta_t^i = \xi_t^i$, $i = 1, \ldots, m$, and work with

$$\mathbf{S}_t^* = (S_t^0, \ldots, S_t^m)' , \qquad S_t^0 = B_t .$$

The portfolio $\boldsymbol{\theta}$ is adapted to $\mathbf{F}^Y$ (the investor cannot see into the future), and the shares of stocks, $\boldsymbol{\xi}$, must also be $\mathbf{F}^Y$-predictable (the investor cannot, e.g. upon a sudden crash of the stock market, escape losses by selling stocks at prices quoted just before and hurry the money over to the locally risk-free bank account.)

The *value* of the portfolio at time t is

$$V_t^{\theta} = \eta_t B_t + \sum_{i=0}^{m} \xi_t^i S_t^i = \eta_t B_t + \boldsymbol{\xi}_t' \mathbf{S}_t = \boldsymbol{\theta}_t' \mathbf{S}_t^*$$

Henceforth we will mainly work with discounted prices and values and, in accordance with (15.10), equip their symbols with a tilde. The discounted value of the portfolio at time t is

$$\tilde{V}_t^{\theta} = \eta_t + \boldsymbol{\xi}_t' \, \tilde{\mathbf{S}}_t = \boldsymbol{\theta}_t' \, \tilde{\mathbf{S}}_t^* \,. \tag{15.12}$$

The strategy $\boldsymbol{\theta}$ is *self-financing* (SF) if $dV_t^{\theta} = \boldsymbol{\theta}_t' \, d\mathbf{S}_t^*$ or, equivalently,

$$d\tilde{V}_t^{\theta} = \boldsymbol{\theta}_t' \, d\tilde{\mathbf{S}}_t^* = \sum_{i=1}^{m} \xi_t^i \, d\tilde{S}_t^i \,. \tag{15.13}$$

We explain the last step: Put $Y_t = B_t^{-1}$, a continuous process. The dynamics of the discounted prices $\tilde{\mathbf{S}}_t^* = Y_t \mathbf{S}_t^*$ is then $d\tilde{\mathbf{S}}_t^* = dY_t \mathbf{S}_t^* + Y_t \, d\mathbf{S}_t^*$. Thus, for $\tilde{V}_t^{\theta} = Y_t V_t^{\theta}$, we have

$$d\tilde{V}_t^{\theta} = dY_t \, V_t^{\theta} + Y_t \, dV_t^{\theta} = dY_t \, \boldsymbol{\theta}_t' \, \mathbf{S}_t^* + Y_t \boldsymbol{\theta}_t' \, d\mathbf{S}_t^* = \boldsymbol{\theta}_t' \, (dY_t \mathbf{S}_t^* + Y_t \, d\mathbf{S}_t^*) = \boldsymbol{\theta}_t' \, d\tilde{\mathbf{S}}_t^* \,,$$

hence the property of being self-financing is preserved under discounting.

The SF property says that, after the initial investment of V_0^{θ}, no further investment inflow or dividend outflow is allowed. In integral form:

$$\tilde{V}_t^{\theta} = \tilde{V}_0^{\theta} + \int_0^t \boldsymbol{\theta}_s' \, d\tilde{\mathbf{S}}_s^* = \tilde{V}_0^{\theta} + \int_0^t \boldsymbol{\xi}_s' \, d\tilde{\mathbf{S}}_s \,. \tag{15.14}$$

Obviously, a constant portfolio $\boldsymbol{\theta}$ is SF; its discounted value process is $\tilde{V}_t^{\theta} = \boldsymbol{\theta}' \, \tilde{\mathbf{S}}_t^*$, hence (15.13) is satisfied. More generally, for a continuous portfolio $\boldsymbol{\theta}$ we would have $d\tilde{V}_t(\theta) = d\boldsymbol{\theta}_t' \, \tilde{\mathbf{S}}_t^* + \boldsymbol{\theta}_t' \, d\tilde{\mathbf{S}}_t^*$, and the self-financing condition would be equivalent to the a budget constraint $d\boldsymbol{\theta}_t' \, \tilde{\mathbf{S}}_t^* = 0$, which says that any purchase of assets must be financed by a sale of some other assets. We urge to say that we shall typically be dealing with portfolios that are not continuous and, in fact, not even right-continuous so that "$d\boldsymbol{\theta}_t$" is meaningless (integrals with respect to the process $\boldsymbol{\theta}$ are not well defined).

D. Absence of arbitrage.

An SF portfolio $\boldsymbol{\theta}$ is called an *arbitrage* if, for some $t > 0$,

$$V_0^{\theta} < 0 \;\text{ and }\; V_t^{\theta} \geq 0 \quad \text{a.s. } \mathbb{P} \,,$$

or, equivalently,

$$\tilde{V}_0^{\theta} < 0 \;\text{ and }\; \tilde{V}_t^{\theta} \geq 0 \quad \text{a.s. } \mathbb{P} \,.$$

A basic requirement on a well-functioning market is the absence of arbitrage. The assumption of no arbitrage, which appears very modest, has surprisingly far-reaching consequences as we shall see.

A *martingale measure* is any probability measure $\tilde{\mathbb{P}}$ that is equivalent to $\mathbb{P}$ and such that the discounted asset prices $\tilde{\mathbf{S}}_t^*$ are martingales $(\mathbf{F}, \tilde{\mathbb{P}})$. The fundamental theorem of arbitrage pricing says: If there exists a martingale measure, then there is no arbitrage. This result follows from easy calculations starting from (15.14): Forming expectation $\tilde{\mathbb{E}}$ under $\tilde{\mathbb{P}}$ and using the martingale property of $\tilde{\mathbf{S}}^*$ under $\tilde{\mathbb{P}}$, we find

$$\mathbb{E}[\tilde{V}_t^\theta] = \tilde{V}_0^\theta + \mathbb{E}[\int_0^t \boldsymbol{\xi}_s' \, d\tilde{\mathbf{S}}_s] = \tilde{V}_0^\theta$$

(the stochastic integral is a martingale). It is seen that arbitrage is impossible.

We return now to our special Markov chain driven market. Let

$$\tilde{\Lambda} = (\tilde{\lambda}^{jk})$$

be an infinitesimal matrix that is equivalent to Λ in the sense that $\tilde{\lambda}^{jk} = 0$ if and only if $\lambda^{jk} = 0$. By Girsanov's theorem, there exists a measure $\tilde{\mathbb{P}}$, equivalent to $\mathbb{P}$, under which Y is a Markov chain with infinitesimal matrix $\tilde{\Lambda}$. Consequently, the processes $\tilde{M}^{jk}$, $j = 1, \ldots, n$, $k \in \mathcal{Y}^j$, defined by

$$d\tilde{M}_t^{jk} = dN_t^{jk} - I_t^j \tilde{\lambda}^{jk} \, dt \,, \tag{15.15}$$

and $\tilde{M}_0^{jk} = 0$, are zero mean, mutually orthogonal martingales w.r.t. $(\mathbf{F}^Y, \tilde{\mathbb{P}})$. Rewrite (15.11) as

$$d\tilde{S}_t^i = \tilde{S}_{t-}^i \left[\sum_j \left(\alpha^{ij} - r^j + \sum_{k \in \mathcal{Y}^j} \gamma^{ijk} \tilde{\lambda}^{jk} \right) I_t^j \, dt + \sum_j \sum_{k \in \mathcal{Y}^j} \gamma^{ijk} d\tilde{M}_t^{jk} \right] , \tag{15.16}$$

$i = 1, \ldots, m$. The discounted stock prices are martingales w.r.t. $(\mathbf{F}^Y, \tilde{\mathbb{P}})$ if and only if the drift terms on the right vanish, that is,

$$\alpha^{ij} - r^j + \sum_{k \in \mathcal{Y}^j} \gamma^{ijk} \tilde{\lambda}^{jk} = 0 \,, \tag{15.17}$$

$j = 1, \ldots, n$, $i = 1, \ldots, m$. From general theory it is known that the existence of such an equivalent martingale measure $\tilde{\mathbb{P}}$ implies absence of arbitrage.

The relation (15.17) can be cast in matrix form as

$$r^j \mathbf{1} - \boldsymbol{\alpha}^j = \boldsymbol{\Gamma}^j \tilde{\boldsymbol{\lambda}}^j \,, \tag{15.18}$$

$j = 1, \ldots, n$, where $\mathbf{1}$ is $m \times 1$ and

$$\boldsymbol{\alpha}^j = \left(\alpha^{ij} \right)_{i=1,\ldots,m} , \quad \boldsymbol{\Gamma}^j = \left(\gamma^{ijk} \right)_{i=1,\ldots,m}^{k \in \mathcal{Y}^j} , \quad \tilde{\boldsymbol{\lambda}}^j = \left(\tilde{\lambda}^{jk} \right)_{k \in \mathcal{Y}^j} .$$

The existence of an equivalent martingale measure is equivalent to the existence of a solution $\tilde{\boldsymbol{\lambda}}^j$ to (15.18) with all entries strictly positive. Thus, the market is arbitrage-free if (and we can show only if) for each j, $r^j \mathbf{1} - \boldsymbol{\alpha}^j$ is in the interior of the convex cone of the columns of $\boldsymbol{\Gamma}^j$.

Assume henceforth that the market is arbitrage-free so that (15.16) reduces to

$$d\tilde{S}_t^i \quad = \quad \tilde{S}_{t-}^i \sum_j \sum_{k \in \mathcal{Y}^j} \gamma^{ijk} d\tilde{M}_t^{jk} \,, \tag{15.19}$$

where the $\tilde{M}^{jk}$ defined by (15.15) are martingales w.r.t. $(\mathbf{F}^Y, \tilde{\mathbb{P}})$ for some measure $\tilde{\mathbb{P}}$ that is equivalent to $\mathbb{P}$.

Inserting (15.19) into (15.13), we find that $\boldsymbol{\theta}$ is SF if and only if

$$d\tilde{V}_t^{\theta} = \sum_j \sum_{k \in \mathcal{Y}^j} \sum_{i=1}^{m} \xi_t^i \tilde{S}_{t-}^i \gamma^{ijk} d\tilde{M}_t^{jk} \,, \tag{15.20}$$

implying that $\tilde{V}^{\theta}$ is a martingale w.r.t. $(\mathbf{F}^Y, \tilde{\mathbb{P}})$ and, in particular,

$$\tilde{V}_t^{\theta} = \tilde{\mathbb{E}}[\tilde{V}_T^{\theta} \,|\, \mathcal{F}_t] \,. \tag{15.21}$$

Here $\tilde{\mathbb{E}}$ denotes expectation under $\tilde{\mathbb{P}}$. (Note that the tilde, which in the first place was introduced to distinguish discounted values from the nominal ones, is also attached to the equivalent martingale measure and certain related entities. This usage is motivated by the fact that the martingale measure arises from the discounted basic price processes, roughly speaking.)

E. Attainability.
A T-claim is a contractual payment due at time T. Formally, it is an $\mathcal{F}_T^Y$-measurable random variable H with finite expected value. The claim is *attainable* if it can be perfectly duplicated by some SF portfolio $\boldsymbol{\theta}$, that is,

$$\tilde{V}_T^{\theta} = \tilde{H} \,. \tag{15.22}$$

If an attainable claim should be traded in the market, then its price must at any time be equal to the value of the duplicating portfolio in order to avoid arbitrage. Thus, denoting the price process by π_t and, recalling (15.21) and (15.22), we have

$$\tilde{\pi}_t = \tilde{V}_t^{\theta} = \tilde{\mathbb{E}}[\tilde{H} \,|\, \mathcal{F}_t] \,, \tag{15.23}$$

or

$$\pi_t = \tilde{\mathbb{E}}\left[e^{-\int_t^T r} H \,\middle|\, \mathcal{F}_t \right] . \tag{15.24}$$

By (15.23) and (15.20), the dynamics of the discounted price process of an attainable claim is

$$d\tilde{\pi}_t = \sum_j \sum_{k \in \mathcal{Y}^j} \sum_{i=1}^{m} \xi_t^i \tilde{S}_{t-}^i \gamma^{ijk} d\tilde{M}_t^{jk} \,. \tag{15.25}$$

F. Completeness.
Any T-claim H as defined above can be represented as

$$\tilde{H} = \tilde{\mathbb{E}}[\tilde{H}] + \int_0^T \sum_j \sum_{k \in \mathcal{Y}^j} \zeta_t^{jk} d\tilde{M}_t^{jk} \,, \tag{15.26}$$

where the ζ_t^{jk} are $\mathbf{F}^Y$-predictable and integrable processes. Conversely, any random variable of the form (15.26) is, of course, a T-claim. By virtue of (15.22), and (15.20), attainability of H means that

$$\begin{aligned} \tilde{H} &= \tilde{V}_0^{\theta} + \int_0^T d\tilde{V}_t^{\theta} \\ &= \tilde{V}_0^{\theta} + \int_0^T \sum_j \sum_{k \in \mathcal{Y}^j} \sum_i \xi_t^i \tilde{S}_{t-}^i \gamma^{ijk} d\tilde{M}_t^{jk} \,. \end{aligned} \tag{15.27}$$

Comparing (15.26) and (15.27), we see that H is attainable iff there exist predictable processes $\xi_t^1, \ldots, \xi_t^m$ such that

$$\sum_{i=1}^{m} \xi_t^i \tilde{S}_{t-}^i \gamma^{ijk} = \zeta_t^{jk} \,,$$

for all j and $k \in \mathcal{Y}^j$. This means that the n^j-vector

$$\boldsymbol{\zeta}_t^j = (\zeta_t^{jk})_{k \in \mathcal{Y}^j}$$

is in $\mathbb{R}(\boldsymbol{\Gamma}^{j\prime})$.

The market is *complete* if every T-claim is attainable, that is, if every n^j-vector is in $\mathbb{R}(\boldsymbol{\Gamma}^{j\prime})$. This is the case if and only if $\text{rank}(\boldsymbol{\Gamma}^j) = n^j$, which can be fulfilled for each j only if $m \geq \max_j n_j$.

15.5 Arbitrage-pricing of derivatives in a complete market

A. Differential equations for the arbitrage-free price.
Assume that the market is arbitrage-free and complete so that prices of T-claims are uniquely given by (15.23) or (15.24).

Let us for the time being consider a T-claim of the form

$$H = h(Y_T, S_T^\ell) \,. \tag{15.1}$$

Examples are a European call option on stock No. ℓ defined by $H = (S_T^\ell - K)^+$, a caplet defined by $H = (r_T - g)^+ = (r^{Y_T} - g)^+$, and a zero coupon T-bond defined by $H = 1$.

For any claim of the form (15.1) the relevant state variables involved in the conditional expectation (15.24) are t, Y_t, S_t^ℓ, hence π_t is of the form

$$\pi_t = \sum_{j=1}^{n} I_t^j f^j(t, S_t^\ell) \,, \tag{15.2}$$

where the

$$f^j(t,s) = \tilde{\mathbb{E}}\left[e^{-\int_t^T r} H \,\middle|\, Y_t = j,\, S_t^\ell = s \right] \tag{15.3}$$

are the state-wise price functions.

The discounted price (15.23) is a martingale w.r.t. $(\mathbf{F}^Y, \tilde{\mathbb{P}})$. Assume that the functions $f^j(t,s)$ are continuously diferentiable. Using Itô on

$$\tilde{\pi}_t = e^{-\int_0^t r} \sum_{j=1}^{n} I_t^j f^j(t, S_t^\ell) \,, \tag{15.4}$$

we find

$$\begin{aligned} d\tilde{\pi}_t \;\; &= \;\; e^{-\int_0^t r} \sum_j I_t^j \left(-r^j\, f^j(t, S_t^\ell) + \frac{\partial}{\partial t} f^j(t, S_t^\ell) + \frac{\partial}{\partial s} f^j(t, S_t^\ell) S_t^\ell \alpha^{\ell j} \right) dt \\ & \quad + e^{-\int_0^t r} \sum_j \sum_{k \in \mathcal{Y}^j} \left(f^k(t, S_{t-}^\ell(1 + \gamma^{\ell jk})) - f^j(t, S_{t-}^\ell) \right) dN_t^{jk} \end{aligned}$$

$$
\begin{aligned}
= \quad & e^{-\int_0^t r} \sum_j I_t^j \Big(-r^j f^j(t, S_t^\ell) + \frac{\partial}{\partial t} f^j(t, S_t^\ell) + \frac{\partial}{\partial s} f^j(t, S_t^\ell) S_t^\ell \alpha^{\ell j} \\
& + \sum_{k \in \mathcal{Y}^j} \{ f^k(t, S_{t-}^\ell(1+\gamma^{\ell jk})) - f^j(t, S_{t-}^\ell) \} \tilde{\lambda}^{jk} \Big) \, dt \\
& + e^{-\int_0^t r} \sum_j \sum_{k \in \mathcal{Y}^j} \Big(f^k(t, S_{t-}^\ell(1+\gamma^{\ell jk})) - f^j(t, S_{t-}^\ell) \Big) \, d\tilde{M}_t^{jk} \, . \qquad (15.5)
\end{aligned}
$$

By the martingale property, the drift term must vanish, and we arrive at the non-stochastic partial differential equations

$$
\begin{aligned}
& -r^j f^j(t,s) + \frac{\partial}{\partial t} f^j(t,s) + \frac{\partial}{\partial s} f^j(t,s) s \alpha^{\ell j} \\
& \quad + \sum_{k \in \mathcal{Y}^j} \Big(f^k(t, s(1+\gamma^{\ell jk})) - f^j(t,s) \Big) \tilde{\lambda}^{jk} = 0 \, , \qquad (15.6)
\end{aligned}
$$

$j = 1, \ldots, n$, which are to be solved subject to the side conditions

$$
f^j(T, s) = h(j, s) \, , \qquad (15.7)
$$

$j = 1, \ldots, n$.

In matrix form, with

$$
\mathbf{R} = \mathbf{D}_{j=1,\ldots,n}(r^j) \, , \qquad \mathbf{A}^\ell = \mathbf{D}_{j=1,\ldots,n}(\alpha^{\ell j}) \, ,
$$

and other symbols (hopefully) self-explaining, the differential equations and the side conditions are

$$
-\mathbf{R}\mathbf{f}(t,s) + \frac{\partial}{\partial t}\mathbf{f}(t,s) + s\mathbf{A}^\ell \frac{\partial}{\partial s}\mathbf{f}(t,s) + \tilde{\Lambda}\mathbf{f}(t, s(1+\gamma)) = \mathbf{0} \, , \qquad (15.8)
$$

$$
\mathbf{f}(T,s) = \mathbf{h}(s) \, . \qquad (15.9)
$$

B. Identifying the strategy.
Once we have determined the solution $f^j(t,s)$, $j = 1, \ldots, n$, the price process is known and given by (15.2).

The duplicating SF strategy can be obtained as follows. Setting the drift term to 0 in (15.5), we find the dynamics of the discounted price;

$$
d\tilde{\pi}_t = e^{-\int_0^t r} \sum_j \sum_{k \in \mathcal{Y}^j} \Big(f^k(t, S_{t-}^\ell(1+\gamma^{\ell jk})) - f^j(t, S_{t-}^\ell) \Big) \, d\tilde{M}_t^{jk} \, . \qquad (15.10)
$$

Identifying the coefficients in (15.10) with those in (15.25), we obtain, for each state j, the equations

$$
\sum_{i=1}^m \xi_t^i S_{t-}^i \gamma^{ijk} = f^k(t, S_{t-}^\ell(1+\gamma^{\ell jk})) - f^j(t, S_{t-}^\ell) \, , \qquad (15.11)
$$

$k \in \mathcal{Y}^j$. The solution $\boldsymbol{\xi}_t^j = (\xi_t^{i,j})'_{i=1,\ldots,m}$ (say) certainly exists since $\text{rank}(\boldsymbol{\Gamma}^j) \le m$, and it is unique iff $\text{rank}(\boldsymbol{\Gamma}^j) = m$. Furthermore, it is a function of t and $\mathbf{S}_{t-}$ and is thus predictable. This simplistic argument works on the open intervals between the jumps of the process Y, where $d\tilde{M}_t^{jk} = -I_t^j \tilde{\lambda}^{jk} \, dt$. For the dynamics (15.10) and (15.25) to be the same also at jump times, the coefficients must clearly be left-continuous. We conclude that

$$\boldsymbol{\xi}_t = \sum_{j=1}^n I_{t-}^j \boldsymbol{\xi}_t\,,$$

which is predictable.

Finally, η is determined upon combining (15.12), (15.23), and (15.4):

$$\eta_t = e^{-\int_0^t r} \sum_{j=1}^n \left(I_t^j f^j(t, S_t^\ell) - I_{t-}^j \sum_{i=1}^m \xi_t^{i,j} S_t^i \right) .$$

C. The Asian option.

As an example of a path-dependent claim, let us consider an Asian option, which essentially is a T-claim of the form $H = \left(\int_0^T S_\tau^\ell\, d\tau - K\right)^+$, where $K \geq 0$. The price process is

$$\begin{aligned} \pi_t &= \tilde{\mathbb{E}}\left[e^{-\int_t^T r} \left(\int_0^T S_\tau^\ell\, d\tau - K \right)^+ \middle| \mathcal{F}_t^Y \right] \\ &= \sum_{j=1}^n I_t^j f^j \left(t, S_t^\ell, \int_0^t S_\tau^\ell\, d\tau \right) , \end{aligned}$$

where

$$f^j(t,s,u) \quad = \quad \tilde{\mathbb{E}}\left[e^{-\int_t^T r} \left(\int_t^T S_\tau^\ell + u - K \right)^+ \middle| Y_t = j,\, S_t^\ell = s \right] .$$

The discounted price process is

$$\tilde{\pi}_t = e^{-\int_0^t r} \sum_{j=1}^n I_t^j\, f^j \left(t, S_t^\ell, \int_0^t S_s^\ell \right) .$$

We obtain partial differential equations in three variables.

The special case $K = 0$ is simpler, with only two state variables.

D. Interest rate derivatives.

A particularly simple, but still important, class of claims are those of the form $H = h(Y_T)$. Interest rate derivatives of the form $H = h(r_T)$ are included since $r_T = r^{Y_T}$. For such claims the only relevant state variables are t and Y_t, so that the function in (15.3) depends only on t and j. The equation (15.6) reduces to

$$\frac{d}{dt} f_t^j = r^j f_t^j - \sum_{k \in \mathcal{Y}^j} (f_t^k - f_t^j) \tilde{\lambda}^{jk}\,, \tag{15.12}$$

and the side condition is (put $h(j) = h^j$)

$$f_T^j = h^j\,. \tag{15.13}$$

In matrix form,

$$\frac{d}{dt} \mathbf{f}_t = (\tilde{\mathbf{R}} - \tilde{\boldsymbol{\Lambda}}) \mathbf{f}_t\,,$$

subject to

$$\mathbf{f}_T = \mathbf{h}\,.$$

The solution is

$$\mathbf{f}_t = \exp\{(\tilde{\Lambda} - \mathbf{R})(T-t)\}\mathbf{h}\,. \tag{15.14}$$

It depends on t and T only through $T-t$.

In particular, the zero coupon bond with maturity T corresponds to $\mathbf{h} = \mathbf{1}$. We will henceforth refer to it as the T-bond in short and denote its price process by $p(t,T)$ and its state-wise price functions by $\mathbf{p}(t,T) = (p^j(t,T))_{j=1,\ldots,n}$;

$$\mathbf{p}(t,T) = \exp\{(\tilde{\Lambda} - \mathbf{R})(T-t)\}\mathbf{1}\,. \tag{15.15}$$

For a call option on a U-bond, exercised at time T ($< U$) with price K, $\mathbf{h}$ has entries $h^j = (p^j(T,U) - K)^+$.

In (15.14) – (15.15) it may be useful to employ the representation shown in (15.2),

$$\exp\{(\tilde{\Lambda} - \mathbf{R})(T-t)\} = \tilde{\mathbf{\Phi}}\,\mathbf{D}_{j=1,\ldots,n}(e^{\tilde{\rho}_j(T-t)})\,\tilde{\mathbf{\Phi}}^{-1}\,, \tag{15.16}$$

say.

15.6 Numerical procedures

A. Simulation.
The homogeneous Markov process $\{Y_t\}_{t\in[0,T]}$ is simulated as follows: Let K be the number of transitions between states in $[0,T]$, and let $T_1,\ldots,T_K$ be the successive times of transition. The sequence $\{(T_n, Y_{T_n})\}_{n=0,\ldots,K}$ is generated recursively, starting from the initial state Y_0 at time $T_0 = 0$, as follows. Having arrived at T_n and Y_{T_n}, generate the next waiting time $T_{n+1} - T_n$ as an exponential variate with parameter $\lambda_{Y_n\cdot}$ (e.g. $-\ln(U_n)/\lambda_{Y_n\cdot}$, where U_n has a uniform distribution over $[0,1]$), and let the new state $Y_{T_{n+1}}$ be k with probability $\lambda_{Y_n k}/\lambda_{Y_n\cdot}$. Continue in this manner $K+1$ times until $T_K < T \le T_{K+1}$.

B. Numerical solution of differential equations.
Alternatively, the differential equations must be solved numerically. For interest rate derivatives, which involve only ordinary first order differential equations, a Runge Kutta will do. For stock derivatives, which involve partial first order differential equations, one must employ a suitable finite difference method, see e.g. [46].

15.7 Risk minimization in incomplete markets

A. Incompleteness.
The notion of incompleteness pertains to situations where a contingent claim cannot be duplicated by an SF portfolio and, consequently, does not receive a unique price from the no arbitrage postulate alone.

In Paragraph 15.4F we were dealing implicitly with incompleteness arising from a scarcity of traded assets, that is, the discounted basic price processes are incapable of spanning the space of all martingales w.r.t. $(\mathbf{F}^Y, \tilde{\mathbb{P}})$ and, in particular, reproducing the value (15.26) of every financial derivative (function of the basic asset prices).

Incompleteness also arises when the contingent claim is not a purely financial derivative, that is, its value depends also on circumstances external to the financial market. We have in mind insurance claims that are caused by events like death or fire and whose claim amounts are e.g. inflation adjusted or linked to the value of some investment portfolio.

In the latter case we need to work in an extended model specifying a basic probability space with a filtration $\mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}$ containing $\mathbf{F}^Y$ and satisfying the usual conditions. Typically it will be the natural filtration of Y and some other process that generates the insurance events. The definitions and conditions laid down in Paragraphs 15.4C-E are modified accordingly, so that adaptedness of η and predictability of $\boldsymbol{\xi}$ are taken to be w.r.t. $(\mathbf{F}, \mathbb{P})$ (keeping the symbol $\mathbb{P}$ for the basic probability measure), a T-claim H is $\mathcal{F}_T$ measurable, etc.

B. Risk minimization.
Throughout the remainder of the paper we will mainly be working with discounted prices and values without any other mention than the notational tilde. The reason is that the theory of risk minimization rests on certain martingale representation results that apply to discounted prices under a martingale measure. We will be content to give just a sketchy review of some main concepts and results from the seminal paper of Föllmer and Sondermann [21].

Let $\tilde{H}$ be a T-claim that is not attainable. This means that an *admissible* portfolio $\boldsymbol{\theta}$ satisfying

$$\tilde{V}_T^{\theta} = \tilde{H}$$

cannot be SF. The *cost*, $\tilde{C}_t^{\theta}$, of the portfolio by time t is defined as that part of the value that has not been gained from trading:

$$\tilde{C}_t^{\theta} = \tilde{V}_t^{\theta} - \int_0^t \boldsymbol{\xi}_\tau' d\tilde{\mathbf{S}}_\tau \, .$$

The *risk* at time t is defined as the mean squared outstanding cost,

$$\tilde{R}_t = \tilde{\mathbb{E}}\left[(\tilde{C}_T^{\theta} - \tilde{C}_t^{\theta})^2 \,\middle|\, \mathcal{F}_t \right] . \tag{15.1}$$

By definition, the risk of an admissible portfolio $\boldsymbol{\theta}$ is

$$\tilde{R}_t^{\theta} = \tilde{\mathbb{E}}\left[(\tilde{H} - \tilde{V}_t^{\theta} - \int_t^T \boldsymbol{\xi}_\tau' d\tilde{\mathbf{S}}_\tau)^2 \,\middle|\, \mathcal{F}_t \right] ,$$

which is a measure of how well the current value of the portfolio plus future trading gains approximates the claim. The theory of risk minimization takes this entity as its object function and proves the existence of an optimal admissible portfolio that minimizes the risk (15.1) for all $t \in [0, T]$. The proof is constructive and provides a recipe for how to actually determine the optimal portfolio.

One sets out by defining the *intrinsic value* of $\tilde{H}$ at time t as

$$\tilde{V}_t^H = \tilde{\mathbb{E}}\left[\tilde{H} \,|\, \mathcal{F}_t \right] .$$

Thus, the intrinsic value process is the martingale that represents the natural current forecast of the claim under the chosen martingale measure. By the Galchouk-Kunita-Watanabe representation, it decomposes uniquely as

$$\tilde{V}_t^H \;=\; \tilde{\mathbb{E}}[\tilde{H}] + \int_0^t \boldsymbol{\xi}_t^{H'} d\tilde{\mathbf{S}}_t + L_t^H \, ,$$

where L^H is a martingale w.r.t. $(\mathbf{F}, \tilde{\mathbb{P}})$ which is orthogonal to $\tilde{S}$. The portfolio $\boldsymbol{\theta}^H$ defined by this decomposition minimizes the risk process among all admissible strategies. The minimum risk is

$$\tilde{R}_t^H = \tilde{\mathbb{E}}\left[\int_t^T d\langle L^H \rangle_\tau \,\middle|\, \mathcal{F}_t \right].$$

C. Unit-linked insurance.
As the name suggests, a life insurance product is said to be *unit-linked* if the benefit is a certain predetermined number of units of an asset (or portfolio) into which the premiums are currently invested. If the contract stipulates a minimum value of the benefit, disconnected from the asset price, then one speaks of *unit-linked insurance with guarantee.* A risk minimization approach to pricing and hedging of unit-linked insurance claims was first taken by Møller [34], who worked with the Black-Scholes-Merton financial market. We will here sketch how the analysis goes in our Markov chain market, which conforms well with the life history process in that they both are intensity-driven.

Let T_x be the remaining life time of an x years old who purchases an insurance at time 0, say. The conditional probability of survival to age $x+u$, given survival to age $x+t$ $(0 \le t < u)$, is

$$ {}_{u-t}p_{x+t} = \mathbb{P}[T_x > u \,|\, T_x > t] = e^{-\int_t^u \mu_{x+s}\, ds}, \tag{15.2}$$

where μ_y is the mortality intensity at age y. We have

$$d\, {}_{u-t}p_{x+t} = {}_{u-t}p_{x+t}\, \mu_{x+t}\, dt. \tag{15.3}$$

Introduce the indicator of survival to age $x+t$,

$$I_t = 1[T_x > t],$$

and the indicator of death before time t,

$$N_t = 1[T_x \le t] = 1 - I_t.$$

The process N_t is a (very simple) counting process with intensity $I_t\, \mu_{x+t}$, that is, M given by

$$dM_t = dN_t - I_t\, \mu_{x+t}\, dt \tag{15.4}$$

is a martingale w.r.t. $(\mathbf{F}, \mathbb{P})$. Assume that the life time T_x is independent of the economy Y. We will work with the martingale measure $\tilde{\mathbb{P}}$ obtained by replacing the intensity matrix Λ of Y with the martingalizing $\tilde{\Lambda}$ and leaving the rest of the model unaltered.

Consider a unit-linked pure endowment benefit payable at a fixed time T, contingent on survival of the insured, with sum insured equal to one unit of stock No. ℓ, but guaranteed no less than a fixed amount g. This benefit is a contingent T-claim,

$$H = (S_T^\ell \vee g)\, I_T.$$

The single premium payable as a lump sum at time 0 is to be determined.

Let us assume that the financial market is complete so that every purely financial derivative has a unique price process. Then the intrinsic value of H at time t is

$$\tilde{V}_t^H = \tilde{\pi}_t\, I_t\, {}_{T-t}p_{x+t},$$

where $\tilde{\pi}_t$ is the discounted price process of the derivative $S_T^\ell \vee g$.

Using Itô and inserting (15.4), we find

$$
\begin{aligned}
d\tilde{V}_t^H &= d\tilde{\pi}_t \, I_{t-} \; {}_{T-t}p_{x+t} + \tilde{\pi}_t \, I_{t-} \; {}_{T-t}p_{x+t} \, \mu_{x+t} \, dt + (0 - \tilde{\pi}_{t-} \; {}_{T-t}p_{x+t}) \, dN_t \\
&= d\tilde{\pi}_t \, I_{t-} \; {}_{T-t}p_{x+t} - \tilde{\pi}_{t-} \; {}_{T-t}p_{x+t} \, dM_t \, .
\end{aligned}
$$

It is seen that the optimal trading strategy is that of the price process of the sum insured multiplied with the conditional probability that the sum will be paid out, and that

$$
dL_t^H = -{}_{T-t}p_{x+t} \, \tilde{\pi}_{t-} \, dM_t \, .
$$

Consequently,

$$
\begin{aligned}
\tilde{R}_t^H &= \int_t^T {}_{T-s}p_{x+s}^2 \, \tilde{\mathbb{E}}\left[\tilde{\pi}_s^2 \middle| \mathcal{F}_t \right] {}_{s-t}p_{x+t} \, \mu_{x+s} \, ds \\
&= {}_{T-t}p_{x+t} \int_t^T \tilde{\mathbb{E}}\left[\tilde{\pi}_s^2 \middle| \mathcal{F}_t \right] {}_{T-s}p_{x+s} \, \mu_{x+s} \, ds \, . \qquad (15.5)
\end{aligned}
$$

15.8 Trading with bonds: How much can be hedged?

A. A finite zero coupon bond market.

Suppose an agent faces a contingent T-claim and is allowed to invest only in the bank account and a finite number m of zero coupon bonds with maturities T_i, $i = 1, \ldots, m$, all post time T. For instance, regulatory constraints may be imposed on the investment strategies of an insurance company. The question is, to what extent can the claim be hedged by self-financed trading in these available assets?

An allowed SF portfolio has discounted value process $\tilde{V}_t^\theta$ of the form

$$
d\tilde{V}_t^\theta = \sum_{i=1}^m \xi_t^i \sum_j \sum_{k \in \mathcal{Y}^j} (\tilde{p}^k(t, T_i) - \tilde{p}^j(t, T_i)) d\tilde{M}_t^{jk} = \sum_j d(\tilde{\mathbf{M}}_t^j)' \mathbf{F}_t^j \boldsymbol{\xi}_t \, ,
$$

where $\boldsymbol{\xi}$ is predictable, $\tilde{\mathbf{M}}_t^{j\,\prime} = (\tilde{M}_t^{jk})^{k \in \mathcal{Y}^j}$ is the n^j-dimensional row vector comprising the non-null entries in the j-th row of $\tilde{\mathbf{M}}_t = (\tilde{M}_t^{jk})$, and

$$
\mathbf{F}_t^j = \mathbf{Y}^j \mathbf{F}_t
$$

where

$$
\mathbf{F}_t = (\tilde{p}^j(t, T_i))_{j=1,\ldots,n}^{i=1,\ldots,m} = (\tilde{\mathbf{p}}(t, T_1), \cdots, \tilde{\mathbf{p}}(t, T_m)) \, , \qquad (15.1)
$$

and $\mathbf{Y}^j$ is the $n^j \times n$ matrix which maps $\mathbf{F}_t$ to $(\tilde{p}^k(t, T_i) - \tilde{p}^j(t, T_i))_{k \in \mathcal{Y}^j}^{i=1,\ldots,m}$. If e.g. $\mathcal{Y}^n = \{1, \ldots, p\}$, then $\mathbf{Y}^n = (\mathbf{I}^{p \times p} \, , \, \mathbf{0}^{p \times (n-p-1)} \, , \, -\mathbf{1}^{p \times 1})$.

The sub-market consisting of the bank account and the m zero coupon bonds is complete in respect of T-claims iff the discounted bond prices span the space of all martingales w.r.t. $(\mathbf{F}^Y, \tilde{\mathbb{P}})$ over the time interval $[0, T]$. This is the case iff, for each j, $\mathrm{rank}(\mathbf{F}_t^j) = n^j$. Now, since $\mathbf{Y}^j$ obviously has full rank n^j, the rank of $\mathbf{F}_t^j$ is determined by the rank of $\mathbf{F}_t$ in (15.1). We will argue that, typically, $\mathbf{F}_t$ has full rank. Thus, suppose $\mathbf{c} = (c_1, \ldots, c_m)'$ is such that

$$
\mathbf{F}_t \mathbf{c} = \mathbf{0}^{n \times 1} \, .
$$

Recalling (15.15), this is the same as

$$\sum_{i=1}^{m} c_i \exp\{(\tilde{\Lambda} - \mathbf{R})T_i\}\mathbf{1} = \mathbf{0}\,,$$

or, by (15.16) and since $\tilde{\mathbf{\Phi}}$ has full rank,

$$\mathbf{D}_{j=1,\ldots,n}(\sum_{i=1}^{m} c_i e^{\tilde{\rho}^j T_i})\tilde{\mathbf{\Phi}}^{-1}\mathbf{1} = \mathbf{0}\,. \tag{15.2}$$

Since $\tilde{\mathbf{\Phi}}^{-1}$ has full rank, the entries of the vector $\tilde{\mathbf{\Phi}}^{-1}\mathbf{1}$ cannot be all null. Typically all entries are non-null, and we assume this is the case. Then (15.2) is equivalent to

$$\sum_{i=1}^{m} c_i e^{\tilde{\rho}^j T_i} = 0\,, \qquad j = 1, \ldots, n. \tag{15.3}$$

Using the fact that the generalized Vandermonde matrix has full rank, we know that (15.3) has a non-null solution $\mathbf{c}$ if and only if the number of distinct eigenvalues $\tilde{\rho}^j$ is less than m, see Section 15.9 below.

In the case where $\mathrm{rank}(\mathbf{F}^j_t) < n^j$ for some j we would like to determine the Galchouk-Kunita-Watanabe decomposition for a given $\mathcal{F}^Y_T$-claim. The intrinsic value process has dynamics

$$d\tilde{H}_t = \sum_j \sum_{k\in\mathcal{Y}^j} \zeta^{jk}_t d\tilde{M}^{jk}_t = \sum_j d(\tilde{\mathbf{M}}^j_t)'\boldsymbol{\zeta}^j_t\,. \tag{15.4}$$

We seek a decomposition of the form

$$\begin{aligned} d\tilde{V}_t &= \sum_i \xi^i_t \, d\tilde{p}(t,T_i) + \sum_j \sum_{k\in\mathcal{Y}^j} \psi^{jk}_t \, d\tilde{M}^{jk}_t \\ &= \sum_j \sum_{j\in\mathcal{Y}^j} \sum_i \xi^i_t \, (\tilde{p}^k(t,T_i) - \tilde{p}^j(t,T_i))\, d\tilde{M}^{jk}_t + \sum_j \sum_{k\in\mathcal{Y}^j} \psi^{jk}_t d\tilde{M}^{jk}_t \\ &= \sum_j d(\tilde{\mathbf{M}}^j_t)'\mathbf{F}^j_t\boldsymbol{\xi}^j_t + \sum_j d(\tilde{\mathbf{M}}^j_t)'\boldsymbol{\psi}^j_t\,, \end{aligned}$$

such that the two martingales on the right hand side are orthogonal, that is,

$$\sum_j I^j_{t-} \sum_{k\in\mathcal{Y}^j} (\mathbf{F}^j_t\boldsymbol{\xi}^j_t)'\tilde{\mathbf{\Lambda}}^j\boldsymbol{\psi}^j_t = 0\,,$$

where $\tilde{\mathbf{\Lambda}}^j = \mathbf{D}(\tilde{\boldsymbol{\lambda}}^j)$. This means that, for each j, the vector $\boldsymbol{\zeta}^j_t$ in (15.4) is to be decomposed into its $\langle\,,\,\rangle_{\tilde{\mathbf{\Lambda}}^j}$ projections onto $\mathbb{R}(\mathbf{F}^j_t)$ and its orthocomplement. From (15.3) and (15.4) we obtain

$$\mathbf{F}^j_t\boldsymbol{\xi}^j_t = \mathbf{P}^j_t\boldsymbol{\zeta}^j_t\,,$$

where

$$\mathbf{P}^j_t = \mathbf{F}^j_t(\mathbf{F}^{j\prime}_t\tilde{\mathbf{\Lambda}}^j\mathbf{F}^j_t)^{-1}\mathbf{F}^{j\prime}_t\tilde{\mathbf{\Lambda}}^j\,,$$

hence

$$\boldsymbol{\xi}^j_t = (\mathbf{F}^{j\prime}_t\tilde{\mathbf{\Lambda}}^j\mathbf{F}^j_t)^{-1}\mathbf{F}^{j\prime}_t\tilde{\mathbf{\Lambda}}^j\boldsymbol{\zeta}^j_t\,. \tag{15.5}$$

Furthermore,

$$\boldsymbol{\psi}_t^j = (\mathbf{I} - \mathbf{P}_t^j)\boldsymbol{\zeta}_t^j \,, \tag{15.6}$$

and the risk is

$$\int_t^T \sum_j p_{s-t}^{Y_t j} \sum_{k \in \mathcal{Y}^j} \lambda^{jk} (\psi_s^{jk})^2 \, ds \,. \tag{15.7}$$

The computation goes as follows: The coefficients ζ^{jk} involved in the intrinsic value process (15.4) and the state-wise prices $p^j(t, T_i)$ of the T_i-bonds are obtained by simultaneously solving (15.6) and (15.12), starting from (15.9) and (15.12), respectively, and at each step computing the optimal trading strategy $\boldsymbol{\xi}$ by (15.5) and the $\boldsymbol{\psi}$ from (15.6), and adding the step-wise contribution to the variance (15.7) (the step-length times the current value of the integrand).

B. First example: The floorlet.
For a simple example, consider a 'floorlet' $H = (r^* - r_T)^+$, where $T < \min_i T_i$. The motivation could be that at time T the insurance company will ascribe interest to the insured's account at current interest rate, but not less than a prefixed guaranteed rate r^*. Then H is the amount that must be provided per unit on deposit and per time unit at time T.

Computation goes by the scheme described above, with the $\zeta_t^{jk} = \tilde{f}_t^k - \tilde{f}_t^j$ obtained from (15.12) subject to (15.13) with $h^j = (r^* - r^j)^+$.

C. Second example: The interest guarantee in insurance.
A more practically relevant example is an interest rate guarantee on a life insurance policy. Premiums and reserves are calculated on the basis of a prudent so-called *first order* assumption, stating that the interest rate will be at some fixed (low) level r^* throughout the term of the insurance contract. Denote the corresponding first order reserve at time t by V_t^*. The (portfolio-wide) mean surplus created by the first order assumption in the time interval $[t, t+dt)$ is $(r^* - r_t)^+ \, {}_tp_x^* V_t^* \, dt$. This surplus is currently credited to the account of the insured as *dividend*, and the total amount of dividends is paid out to the insured at the term of the contracts at time T. Negative dividends are not permitted, however, so at time T the insurer must cover

$$H \;=\; \int_0^T e^{\int_s^T r} (r^* - r_s)^+ \, {}_sp_x^* V_s^* \, ds \,.$$

The intrinsic value of this claim is

$$\begin{aligned}
\tilde{H}_t &= \tilde{\mathbb{E}}\left[\int_0^T e^{-\int_0^s r} (r^* - r_s)^+ \, {}_sp_x^* V_s^* \, ds \,\middle|\, \mathcal{F}_t \right] \\
&= \int_0^t e^{-\int_0^s r} (r^* - r_s)^+ \, {}_sp_x^* V_s^* \, ds + e^{-\int_0^t r} \sum_j I_t^j f_t^j \,,
\end{aligned}$$

where the f_t^j are the state-wise expected values of future guarantees, discounted at time t,

$$f_t^j = \tilde{\mathbb{E}}\left[\int_t^T e^{-\int_t^s r} (r^* - r_s)^+ \, {}_sp_x^* V_s^* \, ds \,\middle|\, Y_t = j \right] .$$

Working along the lines of Section 15.5, we determine the f_t^j by solving

$$\frac{d}{dt} f_t^j = -(r^* - r^j)^+ \, {}_t p_x^* V_t^* + r^j f_t^j - \sum_{k \in \mathcal{Y}^j} (f_t^k - f_t^j) \tilde{\lambda}^{jk} \,,$$

subject to

$$f_T^j = 0 \,. \tag{15.8}$$

The intrinsic value has dynamics (15.4) with $\zeta_t^{jk} = \tilde{f}_t^k - \tilde{f}_t^j$.

From here we proceed as described in Paragraph A.

D. Computing the risk.
Constructive differential equations may be put up for the risk. As a simple example, for an interest rate derivative the state-wise risk is

$$\tilde{R}_t^j = \int_t^T \sum_g p_{\tau - t}^{jg} \sum_{k; k \neq g} \lambda^{gk} \left(\psi_\tau^{gk} \right)^2 d\tau \,.$$

Differentiating this equation, we find

$$\frac{d}{dt} \tilde{R}_t^j = - \sum_{k; k \neq j} \lambda^{jk} \left(\psi_t^{jk} \right)^2 + \int_t^T \sum_g \frac{d}{dt} p_{\tau - t}^{jg} \sum_{k; k \neq g} \left(\psi_\tau^{gk} \right)^2 d\tau \,,$$

and, using the backward version of (15.2),

$$\frac{d}{dt} p_{s-t}^{jg} = - \sum_{h; h \neq j} \lambda^{jh} p_{s-t}^{hg} + \lambda^{j\cdot} p_{s-t}^{jg} \,,$$

we arrive at

$$\frac{d}{dt} \tilde{R}_t^j = - \sum_{k; k \neq j} \lambda^{jk} \left(\psi_t^{jk} \right)^2 - \sum_{k; k \neq j} \lambda^{jk} \tilde{R}_t^k + \lambda^{j\cdot} \tilde{R}_t^j \,.$$

15.9 The Vandermonde matrix in finance

A. The Vandermonde matrix.
Let $\mathbf{A}_n$ denote the generic $n \times n$ matrix of the form

$$\mathbf{A}_n = \left(e^{\alpha_i \beta_j} \right)_{i=1,\ldots,n}^{j=1,\ldots,n} \,, \tag{15.1}$$

where $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$ are reals. This is a classic in matrix theory, known as the generalized *Vandermonde matrix* (usually its elements are written in the form $x_i^{\beta_j}$ with $x_i > 0$). It is well known that it is non-singular iff all α_i are different and all β_j are different, see Gantmacher [22] p. 87.

B. Purpose of the study.
The matrix $\mathbf{A}_n$ in (15.1) and its close relative

$$\mathbf{A}_n - \mathbf{1}_n \mathbf{1}_n' = \left(e^{\alpha_i \beta_j} - 1 \right)_{i=1,\ldots,n}^{j=1,\ldots,n} \,, \tag{15.2}$$

arise naturally in zero coupon bond prices based on spot interest rates driven by certain homogeneous Markov processes. It turns out that, in such bond markets, the issue of completeness is closely related to the rank of the two archetype matrices. Roughly speaking, non-singularity of matrices of types (15.1) or (15.2) ensures that any simple T-claim can be duplicated by a portfolio consisting of the risk-free bank account and a sufficiently large number of zero coupon bonds. The non-singularity results are proved in Section 15.10, and applications to bond markets are presented in Section 15.11.

15.10 Two properties of the Vandermonde matrix

A. The main result.
We take the opportunity here to provide a short proof of the quoted result on non-singularity of the Vandermonde matrix in (15.1), and will supply a similar result about its relative defined in (15.2).

Theorem
(i) *If the α_i are all different and the β_j are all different, then $\mathbf{A}_n$ is non-singular.*
(ii) *If, furthermore, the α_i and the β_j are all different from 0, then $\mathbf{A}_n - \mathbf{1}_n\mathbf{1}_n'$ is non-singular.*

Proof: The proof goes by induction. Let H_n be the hypothesis stated in the two items of the lemma. Trivially, H_1 is true. Assuming that H_{n-1} is true, we need to prove H_n.

Addressing first item (i) of the the hypothesis, it suffices to prove that $\det(\mathbf{A}_n) \neq 0$. Recast this determinant as

$$\begin{aligned}
\det(\mathbf{A}_n) &= \left(\prod_{j=1}^{n} e^{\alpha_n\beta_j}\right) \det \begin{pmatrix} e^{(\alpha_1-\alpha_n)\beta_1} & \cdot\ \cdot\ \cdot & e^{(\alpha_1-\alpha_n)\beta_n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ e^{(\alpha_{n-1}-\alpha_n)\beta_1} & \cdot\ \cdot\ \cdot & e^{(\alpha_{n-1}-\alpha_n)\beta_n} \\ 1 & \cdot\ \cdot\ \cdot & 1 \end{pmatrix} \\
&= \left(\prod_{j=1}^{n} e^{\alpha_n\beta_j} \prod_{i=1}^{n-1} e^{(\alpha_i-\alpha_n)\beta_n}\right) \det \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{1}_{n-1} \\ \mathbf{1}_{n-1}' & 1 \end{pmatrix} \qquad (15.1)
\end{aligned}$$

where

$$\mathbf{A}_{n-1} = \left(e^{(\alpha_i-\alpha_n)(\beta_j-\beta_n)}\right)_{i=1,\ldots,n-1}^{j=1,\ldots,n-1}. \qquad (15.2)$$

The determinant appearing in (15.1) remains unchanged upon subtracting the n-th row of the matrix from all other rows, which gives

$$\begin{aligned}
\det \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{1}_{n-1} \\ \mathbf{1}_{n-1}' & 1 \end{pmatrix} &= \det \begin{pmatrix} \mathbf{A}_{n-1} - \mathbf{1}_{n-1}\mathbf{1}_{n-1}' & \mathbf{0}_{n-1} \\ \mathbf{1}_{n-1}' & 1 \end{pmatrix} \\
&= \det\left(\mathbf{A}_{n-1} - \mathbf{1}_{n-1}\mathbf{1}_{n-1}'\right). \qquad (15.3)
\end{aligned}$$

Now, since the α_i are all different and also the β_j are all different, the matrix $\mathbf{A}_{n-1}$ in (15.2) is of the form required in item (ii) of the lemma and so, by the assumed hypothesis H_{n-1}, $\det(\mathbf{A}_{n-1} - \mathbf{1}_{n-1}\mathbf{1}'_{n-1}) \neq 0$. It follows from (15.1) and (15.3) that $\det(\mathbf{A}_n) \neq 0$, hence item (i) of H_n holds true.

Next, we turn to item (ii) of H_n. Preparing for an ad absurdum argument, assume that $\mathbf{A}_n$ is as specified in item (ii) of the lemma and that $\mathbf{A}_n - \mathbf{1}_n\mathbf{1}'_n$ is singular. Then there exists a vector $\mathbf{c} = (c_1, \ldots, c_n)' \neq \mathbf{0}_n$ such that

$$\mathbf{A}_n\mathbf{c} = \mathbf{1}_n\mathbf{1}'_n\mathbf{c}\,. \tag{15.4}$$

Introducing the function

$$f(\alpha) = \sum_{j=1}^{n} c_j e^{\alpha\beta_j}\,,$$

and putting $\alpha_0 = 0$, we can spell out (15.4) as

$$f(\alpha_0) = f(\alpha_1) = \cdots = f(\alpha_n)\,, \tag{15.5}$$

that is, f assumes the same value at $n+1$ distinct values of α. Since f is continuously differentiable, Rolle's theorem implies that the derivative f' of f is 0 at n distinct values $\alpha_1^*, \ldots, \alpha_n^*$ (say) of α. Now,

$$f'(\alpha) = \sum_{j=1}^{n} c_j\beta_j e^{\alpha\beta_j}\,,$$

and since some c_j are different from 0 and all β_j are different from 0, it follows that the matrix $\mathbf{A}_n^* = \left(e^{\alpha_i^*\beta_j}\right)_{i=1,\ldots,n}^{j=1,\ldots,n}$ should be singular. This contradicts the previously established item (i) under H_n, showing that the assumed singularity of $\mathbf{A}_n - \mathbf{1}_n\mathbf{1}'_n$ is absurd. We conclude that also item (ii) of H_n holds true. □

B. Remarks.
In fact, if $\alpha_1 < \cdots < \alpha_n$ and $\beta_1 < \cdots < \beta_n$, then $\det(\mathbf{A}_n) > 0$ (see [22]). If we take this fact for granted, (15.1) and (15.3) show that also $\det(\mathbf{A}_n - \mathbf{1}_n\mathbf{1}'_n) > 0$, implying that the latter is non-singular under the hypothesis of item (ii) in the theorem. The sign of a general Vandermonde determinant is, of course, the product of the signs of the row and column permutations needed to order the α_i and the β_j by their size.

15.11 Applications to finance

A. Zero coupon bond prices.
A zero coupon bond with maturity T, or just T-bond in short, is the simple contingent claim of 1 at time T. Taking an arbitrage-free financial market for granted, the price process $\{p(t,T)\}_{t\in[0,T]}$ of the T-bond is

$$p(t,T) = \tilde{\mathbb{E}}\left[e^{-\int_t^T r_u\,du} \middle| \mathcal{F}_t \right]\,, \tag{15.1}$$

where $\tilde{\mathbb{E}}$ denotes expectation under some martingale measure, and $\mathcal{F}_t$ is the information available at time t.

We will provide some examples where the results in Section 15.10 are instrumental for establishing linear independence of price processes of bonds with different maturities. The issue is non-trivial only in cases where the bond prices are governed by more than one source of randomness, of course, so we have to look into cases where the spot rate of interest is driven by more than one martingale.

B. Markov chain interest rate.
Referring to Chapter 7, let us model the spot rate of interest $\{r_t\}_{t\geq 0}$ as a continuous time, homogeneous, recurrent Markov chain with finite state space $\{r^1, \ldots, r^n\}$.

We are working under some martingale measure given by an infinitesimal matrix $\tilde{\Lambda} = (\tilde{\lambda}^{jk})$ of the Markov chain, that is, the transition intensities are $\tilde{\lambda}^{jk}$, $j \neq k$, and $\tilde{\lambda}^{jj} = -\sum_{k;k\neq j} \tilde{\lambda}^{jk}$. The price at time $t \leq T$ of a zero coupon bond with maturity T is

$$p(t,T) = \sum_{j=1}^{n} I_t^j p^j(t,T)\,,$$

where $I_t^j = 1[r_t = r^j]$ and

$$p^j(t,T) = \tilde{\mathbb{E}}\left[e^{-\int_t^T r_u\, du} \middle| r_t = r^j \right] .$$

The vector of state-wise prices,

$$\mathbf{p}(t,T) = (p^j(t,T))_{j=1,\ldots,n}\,,$$

is given by (15.15),

$$\mathbf{p}(t,T) = \exp\{(\tilde{\Lambda} - \mathbf{R})(T-t)\}\mathbf{1} = \mathbf{\Phi}\,\mathrm{Diag}(e^{\rho_j(T-t)})\,\mathbf{\Psi}\mathbf{1}\,,$$

where $\mathbf{R} = \mathrm{Diag}(r^j)$ is the $n \times n$ diagonal matrix with the entries r^j down the principal diagonal, $\mathbf{1}$ is the n-vector with all entries equal to 1, $\tilde{\rho}^j$, $j = 1, \ldots, n$, are the eigenvalues of $\tilde{\Lambda} - \mathbf{R}$, and $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are the $n \times n$ matrices formed by the right and left eigenvectors, respectively.

The price processes of m zero coupon bonds with maturities $T_1 < \cdots < T_m$ are linearly independent only if the matrix

$$(e^{\tilde{\rho}^j T_i})_{i=1,\ldots,m}^{j=1,\ldots,n}$$

has rank m. From item (i) in the theorem in Paragraph 15.10A we conclude that this is the case if there are at least m distinct eigenvalues $\tilde{\rho}^j$. It also follows that the market consisting of the bank account with price process $\exp\left(\int_0^t r_s\, ds\right)$ and the m zero coupon bonds is complete for the class of all $\mathcal{F}_{T_1}^r$-claims only if both the number of distinct eigenvalues and the number of bonds are no less than the maximum number of states that can be directly accessed from any single state of the Markov chain.

C. Mixed Vaciček interest rate.
The Vasiček model takes the spot rate of interest to be an Ornstein-Uhlenbeck process given by

$$dr_t = \alpha\,(\rho - r_t)\,dt + \sigma\, d\tilde{W}_t\,. \tag{15.2}$$

Here ρ is the stationary mean of the process, α is a positive mean reversion parameter, σ is a positive volatility parameter, and $\tilde{W}$ is a standard Brownian motion under a martingale measure. The dynamics of the discounted T-bond price,

$$\tilde{p}(t,T) = e^{-\int_0^t r_u\,du} p(t,T)\,, \tag{15.3}$$

is

$$d\tilde{p}(t,T) \;=\; \tilde{p}(t,T)\frac{\sigma}{\alpha}\left(e^{-\alpha\,(T-t)} - 1\right) d\tilde{W}_t\,, \tag{15.4}$$

see e.g. [9]. Obviously, any $\mathcal{F}_T^{\tilde{W}}$ claim can be duplicated by a self-financing portfolio in the T-bond and the bank account, and so the completeness issue is trivial in this model.

To create an example where one bond is not sufficient to complete the market, let us concoct a mixed Vasiček model by putting

$$r_t = \sum_{j=1}^{n} r_t^j\,,$$

where the r^j are independent Ornstein-Uhlenbeck processes,

$$dr_t^j = \alpha^j\,(\rho^j - r_t^j)\,dt + \sigma^j\,d\tilde{W}_t^j\,,$$

$j = 1, \ldots, n$, and the $\tilde{W}^j$ are independent standard Brownian motions. We assume that the α^j are all distinct (otherwise we could gather all processes r^j with coinciding mean reversion parameter into one Ornstein-Uhlenbeck process). The mixed Vasiček process is not mean-reverting in the same simple sense as the traditional Vasiček process. It is stationary, however, and is apt to describe interest that is subject to several random phenomena, each of mean-reverting type.

By the assumed independence, the price of the T-bond is just

$$p(t,T) = \prod_{j=1}^{n} p^j(t,T)\,,$$

where $p^j(t,T) = \tilde{\mathbb{E}}\left[e^{-\int_t^T r_u\,du}\middle|\, r_t^j\right]$, and the discounted price is

$$\tilde{p}(t,T) = \prod_{j=1}^{n} \tilde{p}^j(t,T)\,,$$

where $\tilde{p}^j(t,T)$ is the "j-analogue" to (15.3). By virtue of (15.4), we conclude that the discounted T-bond price has dynamics

$$d\tilde{p}(t,T) = \tilde{p}(t,T)\sum_{j=1}^{n}\frac{\sigma^j}{\alpha^j}\left(e^{-\alpha^j\,(T-t)} - 1\right) d\tilde{W}_t^j\,. \tag{15.5}$$

Now, consider the market consisting of the bank account and m zero coupon bonds with maturities $T_1 < \cdots < T_m$. From (15.5) it is seen that this market is complete for the class of $\mathcal{F}_{T_1}^{\tilde{W}_1,\ldots,\tilde{W}_n}$-claims if and only if the matrix

$$\left(e^{-\alpha^j\,(T_i-t)} - 1\right)_{j=1,\ldots,n}^{i=1,\ldots,m} \tag{15.6}$$

has rank n. By virtue of item (ii) in the theorem in Paragraph 15.10A, we conclude that this is the case if $m \geq n$.

D. Mixed Poisson-driven Ornstein-Uhlenbeck interest rate.
Referring to [40], let us replace the Brownian motions in Paragraph C above with independent compensated Poisson processes, that is,

$$d\tilde{W}_t^j = dN_t^j - \lambda^j \, dt \, ,$$

where each N^j is a Poisson process with intensity λ^j. Instead of (15.5) we obtain

$$d\tilde{p}(t,T) = \tilde{p}(t-,T) \sum_{j=1}^{n} \left(\exp\left\{ \frac{\sigma^j}{\alpha^j} \left(e^{-\alpha^j \, (T-t)} - 1 \right) \right\} - 1 \right) d\tilde{W}_t^j \, . \tag{15.7}$$

It is seen from (15.7) that the market consisting of the bank account and m zero coupon bonds with maturities $T_1 < \cdots < T_m$ is complete for the class of $\mathcal{F}_{T_1}^{\tilde{N}_1,\ldots,\tilde{N}_n}$-claims if and only if the matrix

$$\left(\exp\left\{ \frac{\sigma^j}{\alpha^j} \left(e^{-\alpha^j \, (T-t)} - 1 \right) \right\} - 1 \right)_{j=1,\ldots,n}^{i=1,\ldots,m}$$

has rank n. By item (ii) in the theorem in Paragraph 15.10A, we know that the matrix (15.6) has full rank. Thus, completeness of a market consisting of the bank account and at least n bonds would be established – and we would be done – if we could prove that the $n \times m$ matrix $(e^{\gamma_{ji}} - 1)$ has full rank whenever (γ_{ji}) has full rank. With this conjecture our study of these problems will have to halt for the time being.

15.12 Martingale methods

A. Preliminaries. Referring to Sections 5.4 and 7.3 we shall see examples of how martingale methods can be used to prove results that throughout the text have been obtained by the direct backward argument.

The policy was described as a time-continuous Markov chain Z with finite state space and transition intensities μ_{jk}, $j \neq k$. We introduced

$$I_j(t) = 1[Z(t) = j] \, ,$$

the indicator of the event that Z is in state j at time t, and

$$N_{jk}(t) = \sharp\{s;\, 0 < s \leq t,\, Z(s-) = j \, , \, Z(s) = k\} \, ,$$

the number of direct transitions of Z from state j to state k $(\neq j)$ in the time interval $(0,t]$. Taking Z to be right-continuous, the same goes for the indicator processes I_j and the counting processes N_{jk}.

Let $\mathcal{F}_t^Z = \sigma\{Z_s;\, 0 \leq s \leq t\}$, $t \geq 0$, be the filtration generated by Z. The compensated counting processes M_{jk}, $j \neq k$, defined by

$$dM_{jk}(t) = dN_{jk}(t) \, - \, I_j(t) \, \mu_{jk} \, dt \tag{15.8}$$

and $M_{jk}(0) = 0$, are zero mean, square integrable, mutually orthogonal martingales w.r.t. the filtration.

We need a couple of general results:
1. Let X be a real-valued random variable such that $\mathbb{E}|X| < \infty$. Then the process M defined by

$$M(t) = \mathbb{E}[X|\mathcal{F}_t]$$

is a martingale. This follows by the rule of iterated expectation and the filtration property, $\mathcal{F}_s \subset \mathcal{F}_t$ for $s < t$:

$$\mathbb{E}[M(t)\,|\,\mathcal{F}_s] \;=\; \mathbb{E}\{\mathbb{E}[X|\mathcal{F}_t]\,|\,\mathcal{F}_s\} \;=\; \mathbb{E}[X\,|\,\mathcal{F}_s] \;=\; M(s)\,.$$

2. A martingale M with paths that are (almost surely) continuous and of finite variation in every finite interval is constant as a function of time; $M(t) = M(0)$ for all t. This is seen as follows. Since M has finite variation, it obeys the rules of ordinary calculus and, in particular,

$$M^2(t) \;=\; M^2(0) \,+\, 2\int_0^t M(s)\,dM(s)\,.$$

Since M is continuous, it is also predictable so that the integral $\int_0^t 2M(s)\,dM(s)$ is a martingale. It follows that

$$\mathbb{E}\left[M^2(t)\right] \;=\; M^2(0)\,.$$

Since $\mathbb{E}[M(t)] \;=\; M(0)$, we conclude that

$$\mathbb{V}\mathrm{ar}[M(t)] \;=\; 0\,,$$

hence M is constant.

Now to the martingale technique:

B. First example: Thiele's differential equation. Consider the standard multi-state insurance policy defined in Paragraph 5.4.A and recall the form of the payment function,

$$dB(t) \;=\; \sum_j I_j(t)\,dB_j(t) \,+\, \sum_{j\neq k} b_{jk}(t)\,dN_{jk}(t)\,, \tag{15.9}$$

where the B_j and the b_{jk} are deterministic functions. Motivated by Section 5.6, we allow the interest rate to depend on the current state of the Markov process:

$$r(t) \;=\; r_{Z(t)} \;=\; \sum_j I_j(t)\,r_j\,. \tag{15.10}$$

(This does not complicate matters.)

Define the martingale

$$\begin{aligned} M(t) \;&=\; \mathbb{E}\left[\int_{0-}^n e^{-\int_0^\tau r} dB(\tau)\bigg|\,\mathcal{F}_t\right] \\ &=\; \int_{0-}^t e^{-\int_0^\tau r} dB(\tau) \,+\, e^{-\int_0^t r}\mathbb{E}\left[\int_t^n e^{-\int_t^\tau r} dB(\tau)\bigg|\,\mathcal{F}_t\right] \\ &=\; \int_{0-}^t e^{-\int_0^\tau r} dB(\tau) \,+\, e^{-\int_0^t r}\sum_j I_j(t)V_j(t)\,. \end{aligned} \tag{15.11}$$

(Recall the short-hand $\int_0^\tau r = \int_0^\tau r(s)\,ds$.) The last step follows from the Markov property of Z and the fact that the payments at any time are (functionally) independent of the past:

$$\begin{aligned}
\mathbb{E}\left[\int_t^n e^{-\int_t^\tau r} dB(\tau)\middle|\mathcal{F}_t\right] &= \mathbb{E}\left[\int_t^n e^{-\int_t^\tau r} dB(\tau)\middle| Z_t\right] = V_{Z(t)}(t) \\
&= \sum_j I_j(t) V_j(t)\,.
\end{aligned}$$

Now apply Itŏ's formula to (15.11):

$$\begin{aligned}
dM(t) &= e^{-\int_0^t r} dB(t) + e^{-\int_0^t r}(-r(t)\,dt) \sum_j I_j(t)\,V_j(t) \\
&\quad + e^{-\int_0^t r} \sum_j I_j(t)\,dV_j(t) + e^{-\int_0^t r} \sum_{j\neq k} dN_{jk}(t)\,(V_k(t) - V_j(t-))\,.
\end{aligned}$$

The last term on the right takes care of the jumps of the Markov process: upon a jump from state j to state k the last term in (15.11) changes immediately from the discounted value of the reserve in state j just before the jump to the value of the reserve in state k at the time of the jump. Since the state-wise reserves are deterministic functions with finite variation, they have at most a countable number of discontinuities at fixed times. The probability that the Markov process jumps at any such time is 0. Therefore, we need not worry about possible common points of discontinuity of the $V_j(t)$ and the $I_j(t)$. For the same reason we can also disregard the left limit in $V_j(t-)$ in the last term.

We proceed by inserting the expressions (15.9) for $dB(t)$, (15.10) for $r(t)$, and the expression $dN_{jk}(t) = dM_{jk}(t) - I_j(t)\,\mu_{jk}(t)\,dt$ obtained from (15.8), and gather

$$\begin{aligned}
dM(t) &= e^{-\int_0^t r} \sum_j I_j(t) \left(dB_j(t) - r_j\,V_j(t)\,dt + dV_j(t) + \sum_{k;\,k\neq j} \mu_{jk}\,dt\,R_{jk}(t) \right) \\
&\quad + e^{-\int_0^t r} \sum_{j\neq k} R_{jk}(t)\,dM_{jk}(t)\,,
\end{aligned} \tag{15.12}$$

where

$$R_{jk}(t) = b_{jk}(t) + V_k(t) - V_j(t)$$

is recognized as the sum at risk in respect of transition from j to k at time t.

Since the last term on the right of (15.12) is the increment of a martingale, the first term of the right is the difference between the increments of two martingales and is thus itself the increment of a martingale. This martingale has finite variation and, as will be explained below, is also continuous, and must therefore be constant. For this to be true for all realizations of the indicator functions I_j, we must have

$$dB_j(t) - r_j\,V_j(t)\,dt + dV_j(t) + \sum_{k;\,k\neq j} \mu_{jk}\,dt\,R_{jk}(t) = 0\,. \tag{15.13}$$

This is nothing but Thiele's differential equation. We also obtain that

$$dM(t) = e^{-\int_0^t r} \sum_{j\neq k} R_{jk}(t)\,dM_{jk}(t)\,,$$

which displays the dynamics of the martingale M.

Finally, we explain why the (15.13) is the increment at t of a continuous function. The dt terms are continuous increments, of course. Outside jump times of the B_j both the B_j themselves and the V_j are continuous. At any time t where there is a jump in some B_j the reserve V_j jumps by the same amount in the opposite direction since $V_j(t-) = \Delta B_j(t) + V_j(t)$. Thus, $B_j + V_j$ is indeed continuous.

C. Second example: Revisiting Exercise 56, the geometric Poisson price process. Martingale methods are certainly not needed in this case since the Poisson process is sufficiently well structured to allow of direct computation of all functions asked for in the exercise. But just for the sake of the example:

Introduce

$$V(t) \,=\, \mathbb{E}[S(t)]\,.$$

Let $\mathcal{F}_t$, $t \geq 0$, be the filtration generated by the Poisson process N. Fix a time $T > 0$ and introduce the martingale

$$\begin{aligned} M(t) &= \mathbb{E}\left[\, S(T) \mid \mathcal{F}_t \right] \\ &= e^{\alpha t + \beta N(t)} \mathbb{E}\left[e^{\alpha(T-t)+\beta(N(T)-N(t)}\Big|\, \mathcal{F}_t \right] \\ &= S(t)\, V(T-t)\,. \end{aligned}$$

Here we have made use of the fact that the Poisson process has stationary and independent increments.

Write $V' = \frac{d}{dt}V$ and apply Itŏ:

$$dM(t) \,=\, S(t)\,\alpha\, dt\, V(T-t) \,+\, S(t)\, V'(T-t)(-dt) \,+\, (S(t) \,-\, S(t-))\, V(T-t), \tag{15.14}$$

the last term accounting of jumps. Rewrite

$$\begin{aligned} S(t) \,-\, S(t-) &= e^{\alpha t + \beta N(t)} - e^{\alpha t + \beta N(t-)} \\ &= e^{\alpha t + \beta(N(t-)+\Delta N(t))} - e^{\alpha t + \beta N(t-)} \\ &= S(t-)\left(e^{\beta \Delta N(t)} - 1\right) \\ &= S(t-)\left(e^{\beta} - 1\right) dN(t)\,, \end{aligned}$$

where the last two step is due to the zero-or-one nature of the increments of counting process N. Upon inserting this into (15.14) and introducing the martingale

$$dM_N(t) \,=\, dN(t) \,-\, \lambda\, dt\,,$$

we get

$$\begin{aligned} dM(t) &= S(t)\left[\alpha\, V(T-t) \,-\, V'(T-t) \,+\, \left(e^{\beta} - 1\right) \lambda\, V(T-t)\right] dt \\ &\quad + S(t-)\left(e^{\beta} - 1\right) dM_N(t)\,, \end{aligned}$$

Here we have used the fact that $S(t-)\, dt \,=\, S(t)\, dt$ which is to be understood in integral form: $\int f(t-)\, dt \,=\, \int f(t)\, dt$ since the integral is not affected by a change of the integrand at an at most countable set of points. Arguing as in the previous example, the drift term in the dynamics of M must vanish, and we arrive at the differential equation

$$V'(t) \,=\, \left[\alpha \,+\, \left(e^{\beta} - 1\right) \lambda\right] V(t)\,,$$

The solution, subject to the obvious condition

$$V(0) = 1\,,$$

is

$$V(t) \;=\; \exp\left[\left(\alpha + \left(e^{\beta} - 1\right)\lambda\right) t\right].$$

To rehearse the technique, you should use it to solve Item (c) in exercise 56. Start from the martingale

$$M(t) \,=\, \mathbb{E}\left[\int_0^T S^{-1}(\tau)\,d\tau \,\middle|\, \mathcal{F}_t\right] \,=\, \int_0^t S^{-1}(\tau)\,d\tau \,+\, e^{-\alpha t-\beta N(t)}\,V(T-t)\,,$$

where

$$V(t) \,=\, \mathbb{E}\left[\int_0^t S^{-1}(\tau)\,d\tau\right].$$

D. Third example: Revisiting pricing of the unit-linked term insurance with guarantee, pages 166-167. We need to determine

$$\pi \,=\, \mathbb{E}\left[\int_0^n \left(1 \vee e^{-\int_0^\tau r} g\right) f(\tau)\,d\tau\right].$$

Let $\mathcal{F}_t = \sigma\{Y_s;\, 0 \le s \le t\}$, $t \ge 0$, be the filtration generated by the 'economy process' Y. Start from the martingale

$$\begin{aligned}
M(t) \;&=\; \mathbb{E}\left[\left(1 \vee e^{-\int_0^\tau r} g\right) f(\tau)\,d\tau \,\middle|\, \mathcal{F}_t\right] \\
&=\; \int_0^t \left(1 \vee e^{-\int_0^\tau r} g\right) f(\tau)\,d\tau \,+\, e^{-\int_0^t r}\,\mathbb{E}\left[\int_t^n \left(e^{\int_0^t r} \vee e^{-\int_t^\tau r} g\right) f(\tau)\,d\tau \,\middle|\, \mathcal{F}_t\right] \\
&=\; \int_0^t \left(1 \vee e^{-\int_0^\tau r} g\right) f(\tau)\,d\tau \,+\, e^{-\int_0^t r} \sum_e I_e(t)\, W_e(t, e^{\int_0^\tau r})\,,
\end{aligned}$$

where

$$W_e(t,u) = \mathbb{E}\left[\int_t^n \left(u \vee e^{-\int_t^\tau r} g\right) f(\tau)\,d\tau \,\middle|\, Y(t) = e\right].$$

Assuming that the partial derivatives $\frac{\partial}{\partial t}W_e(t,u)$ and $\frac{\partial}{\partial u}W_e(t,u)$ exist, apply Itŏ:

$$\begin{aligned}
dM(t) \;&=\; \left(1 \vee e^{-\int_0^t r} g\right) f(t)\,dt \,+\, e^{-\int_0^t r}(-r(t)\,dt) \sum_e I_e(t)\, W_e(t, e^{\int_0^\tau r}) \\
&\quad + e^{-\int_0^t r} \sum_e I_e(t) \left(\frac{\partial}{\partial t} W_e(t, e^{\int_0^\tau r})\,dt \,+\, \frac{\partial}{\partial u} W_e(t, e^{\int_0^\tau r})\, e^{\int_0^\tau r} r(t)\,dt\right) \\
&\quad + e^{-\int_0^t r} \sum_{e \ne f} dN_{ef}(t)\,(W_f(t, e^{\int_0^\tau r}) \,-\, W_e(t, e^{\int_0^\tau r}))\,.
\end{aligned}$$

Substitute

$$dN_{ef}(t) \,=\, dM_{ef}(t) \,+\, \lambda_{ef}\,dt\,,$$

where the M_{ef} are martingales, and put $U(t) = e^{\int_0^\tau r}$. Arguing along the lines of Paragraph B, we conclude that

$$\begin{aligned}
&\left(1 \vee U(t)^{-1} g\right) f(t) \,-\, U(t)^{-1}\, r_e\, W_e(t, U(t)) \\
&+ U(t)^{-1} \left(\frac{\partial}{\partial t} W_e(t, U(t)) \,+\, \frac{\partial}{\partial u} W_e(t, U(t))\, U(t)\, r_e(t) \right) \\
&+ U(t)^{-1} \sum_{f;\, f \neq e} \lambda_{ef} \left(W_f(t, U(t)) \,-\, W_e(t, U(t))\right) = 0\,.
\end{aligned}$$

for all realizations of Y. We end up with the partial differential equation

$$\begin{aligned}
&(u \vee g)\, f(t) \,-\, r_e\, W_e(t,u) \,+\, \left(\frac{\partial}{\partial t} W_e(t,u) \,+\, \frac{\partial}{\partial u} W_e(t,u)\, u\, r_e(t) \right) \\
&+ \sum_{f;\, f \neq e} \lambda_{ef} \left(W_f(t,u) \,-\, W_e(t,u)\right) = 0\,,
\end{aligned}$$

which are to be solved subject to the obvious conditions

$$W_e(n,u) \,=\, 0\,.$$

E. Remark on the technique. In all three examples we needed to determine the expected value of some random variable W that depends on the development of a stochastic process. Here is an outline of the method: Start from the martingale $M(t) \,=\, \mathbb{E}\,[W \mid \mathcal{F}_t]$. Inspect W and try to write it in the form

$$W = W(T(t), U(t))\,,$$

where $T(t)$ depends only on the future development of the process and $U(t)$ depends only on the past history $\mathcal{F}_t$. The random variable $U(t)$ is the called the 'state variable(s)' (it may be multi-dimensional). How to proceed from here depends on the properties of the driving stochastic process. In our situations (let us keep the example in Paragraph D in mind) we use the Markov property of Y to conclude that the conditional expected value $M(t)$ must be a function only of the current time, state, and value of the state variable:

$$M(t) = F_{Y(t)}(t, U(t)) = \sum_e I_e(t)\, F_e(t, U(t))\,.$$

Assuming continuous differentibility of the functions F_e with respect to t and u, use Itŏ to form the dynamics of M:

$$\begin{aligned}
dM(t) \quad &= \quad \sum_e I_e(t) \left(\frac{\partial}{\partial t} F_e(t, U(t))\, dt \,+\, \frac{\partial}{\partial u} F_e(t, U(t))\, dU^c(t) \right) \\
&\qquad + \sum_e \left[I_e(t)\, F_e(t, U(t)) \,-\, I_e(t-)\, F_e(t-, U(t-)) \right]\,.
\end{aligned}$$

Here U^c denotes the continuous part of U. The jump part may have contributions from possible jumps of U outside jump times for Y, but these will always cancel out and vanish in the end, however, see the example in Paragraph B above. In any case the jump part will consist of the following terms due to jumps of Y:

$$\sum_{f \neq e} dN_{ef}(t)\, [F_f(t, U(t)) - F_e(t, U(t-))]\,.$$

Now, insert $dN_{ef}(t) = dM_{ef}(t) + \lambda_{ef}\, dt$ and identify the martingale part and the drift part with the factor dt in the expression for $dM(t)$. The drift part must vanish, and we arrive at a set of constructive non-stochastic differential equations from which the functions $F_e(t, u)$ can be solved (at least numerically).

Bibliography

[1] Aase, K.K. and Persson, S.-A. (1994). Pricing of unit-linked life insurance policies. *Scand. Actuarial J.*, **1994**, 26-52.

[2] Aczel J. *Lectures on Functional Equations and their Applications*. Academic Press, 1966.

[3] Andersen, P.K., Borgan, Ø., Gill, R.D., Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York, Berlin, Heidelberg.

[4] Anderson, J.L. and Dow, J.B. (1948). *Actuarial Statistics, Vol. II: Constructions of Mortality and other Tables.* Cambridge University Press.

[5] Barlow, R.E. and Proschan, F (1981): *Statistical Theory of Reliability and Life Testing*, Holt, Reinhart and Winston Inc.

[6] Berger, A. (1939): *Mathematik der Lebensversicherung*. Verlag von Julius Springer, Vienna.

[7] Bibby J.M., Mardia, K.V., and Kent J.T. *Multivariate Analysis*. Academic Press, 1979.

[8] Björk, T., Kabanov, Y., Runggaldier, W. (1997): Bond market structures in the presence of marked point processes. *Mathematical Finance*, **7**, 211-239.

[9] Björk, T. (1998): *Arbitrage Theory in Continuous Time*, Oxford University Press.

[10] Black, F., Scholes, M. (1973): The pricing of options and corporate liabilities. *J. Polit. Economy*, **81**, 637-654.

[11] Bowers, N.L. Jr., Gerber, H.U., Hickman, J.C., and Nesbitt, C.J. (1986). Actuarial Mathematics. The Society of Actuaries. Itasca, Illinois.

[12] Cox, J., Ross, S., Rubinstein, M. (1979): Option pricing: A simplified approach. *J. of Financial Economics*, **7**, 229-263.

[13] Delbaen, F., Schachermayer, W. (1994): A general version of the fundamental theorem on asset pricing. *Mathematische Annalen*, **300**, 463-520.

[14] De Pril, N. (1989). The distributions of actuarial functions. *Mitteil. Ver. Schweiz. Vers.math.*, **89**, 173-183.

[15] De Vylder, F. and Jaumain, C. (1976). *Exposé moderne de la théorie mathématique des opérations viagère*. Office des Assureurs de Belgique, Bruxelles.

[16] Dhaene, J. (1990). Distributions in life insurance. *ASTIN Bull.* **20**, 81-92.

[17] Donald D.W.A. *Compound Interest and Annuities Certain.* Heinemann, London, 1970.

[18] Eberlein, E., Raible, S. (1999): Term structure models driven by general Lévy processes. *Mathematical Finance*, **9**, 31-53.

[19] Elliott, R.J., Kopp, P.E. (1998): *Mathematics of financial markets*, Springer-Verlag.

[20] Esary, J.D. and Proschan, F (1972): Relationships among some notions of bivariate dependence, *Annals of Mathematical Statistics*, **43**, 651-655.

[21] Föllmer, H., Sondermann, D. (1986): Hedging of non-redundant claims. In *Contributions to Mathematical Economics in Honor of Gerard Debreu*, 205-223, eds. Hildebrand, W., Mas-Collel, A., North-Holland.

[22] Gantmacher, F.R. (1959): *Matrizenrechnung II*, VEB Deutscher Verlag der Wissenschaften, Berlin.

[23] Gerber, H.U. (1995). *Life Insurance Mathematics*, 2nd edn. Springer-Verlag.

[24] Harrison, J.M., Kreps, D.M. (1979): Martingales and arbitrage in multiperiod securities markets. *J. Economic Theory*, **20**, 1979, 381-408.

[25] Harrison, J.M., Pliska, S. (1981): Martingales and stochastic integrals in the theory of continuous trading. *J. Stoch. Proc. and Appl.*, **11**, 215-260.

[26] Hoem, J.M. (1969): Markov chain models in life insurance. *Blätter Deutsch. Gesellschaft Vers.math.*, **9**, 91–107.

[27] Hoem, J.M. (1969): Purged and partial Markov chains. *Skandinavisk Aktuarietidskrift*, **52**, 147–155.

[28] Hoem, J.M. and Aalen, O.O. (1978). Actuarial values of payment streams. *Scand. Actuarial J.* **1978**, 38-47.

[29] Jordan, C.W. (1967). *Life Contingencies.* The Society of Actuaries, Chicago.

[30] Karlin, S., Taylor, H. (1975): *A first Course in Stochastic Processes*, 2nd. ed., Academic Press.

[31] Kellison S. (1970). *The theory of interest.* Richard D. Irwin, Inc. Homewood, Illinois.

[32] Merton, R.C. (1973): The theory of rational option pricing. *Bell Journal of Economics and Management Science*, **4**, 141-183.

[33] Merton, R.C. (1976): Option pricing when underlying stock returns are discontinuous. *J. Financial Economics*, **3**, 125-144.

[34] Møller, T. (1998): Risk minimizing hedging strategies for unit-linked life insurance. *ASTIN Bull.*, **28**, 17-47.

[35] Norberg, R. (1985). Lidstone in the continuous case. *Scand. Actuarial J.* **1985**, 27-32.

[36] Norberg, R. (1991). Reserves in life and pension insurance. *Scand. Actuarial J.* **1991**, 1-22.

[37] Norberg R. (1993). Identities for present values of life insurance benefits. *Scand. Actuarial J.*, 1993, p. 100–106.

[38] Norberg, R. (1995): Differential equations for moments of present values in life insurance. *Insurance: Math. & Econ.*, **17**, 171-180.

[39] Norberg, R. (1995): A time-continuous Markov chain interest model with applications to insurance. *J. Appl. Stoch. Models and Data Anal.*, 245-256.

[40] Norberg, R. (1998): Vasiček beyond the normal. *Working paper No. 152*, Laboratory of Actuarial Math., Univ. Copenhagen.

[41] Norberg, R. (1999): A theory of bonus in life insurance. *Finance and Stochastics.*, **3**, 373-390

[42] Norberg, R. (2001): On bonus and bonus prognoses in life insurance. *Scand. Actuarial J.*

[43] Pliska, S.R. (1997): *Introduction to Mathematical Finance*, Blackwell Publishers.

[44] Ramlau-Hansen, H. (1991): Distribution of surplus in life insurance. *ASTIN Bull.*, **21**, 57-71.

[45] Sverdrup, E. (1969): *Noen forsikringsmatematiske emner.* Stat. Memo. No. 1, Inst. of Math., Univ. of Oslo. (In Norwegian.)

[46] Thomas, J.W. (1995): *Numerical Partial Differential Equations: Finite Difference Methods*, Springer-Verlag.

[47] Width, E. (1986): A note on bonus theory. *Scand. Actuarial J.* **1986**, 121-126.

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 $\mathcal{D}(X)$ 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

$$X_t = X_0 + \int_0^t x_\tau \, d\tau + \sum_{0 < \tau \leq t} (X_\tau - X_{\tau-}) \,. \tag{A.1}$$

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

$t \notin \mathcal{D} = \mathcal{D}(X) \cup \mathcal{D}(x)$ we have $\frac{d}{dt} X_t = x_t$, that is, the function X grows (or decays) continuously at rate x_t.

As a convenient notational device we shall frequently write (A.1) in differential form as

$$dX_t = x_t\, dt \,+\, X_t - X_{t-}\,. \tag{A.2}$$

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

$$\int_s^t Y_\tau\, dX_\tau = \int_s^t Y_\tau x_\tau\, d\tau + \sum_{s<\tau\le t} Y_\tau (X_\tau - X_{\tau-})\,, \tag{A.3}$$

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

$$X_t = X_s + \int_s^t dX_\tau\,, \tag{A.4}$$

saying that the value of X at time t is its value at time s plus all its increments in $(s, t]$.

By definition,

$$\int_s^{t-} Y_\tau\, dX_\tau = \lim_{r \nearrow t} \int_s^r Y_\tau\, dX_\tau = \int_s^t Y_\tau\, dX_\tau - Y_t(X_t - X_{t-}) = \int_{(s,t)} Y_\tau\, dX_\tau\,,$$

a left-continuous function of t. Likewise,

$$\int_{s-}^t Y_\tau\, dX_\tau = \lim_{r \nearrow s} \int_r^t Y_\tau\, dX_\tau = \int_s^t Y_\tau\, dX_\tau + Y_s(X_s - X_{s-}) = \int_{[s,t]} Y_\tau\, dX_\tau\,,$$

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 $f : \mathcal{R}^m \mapsto \mathcal{R}$ 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

$$df(X_t) = \sum_{i=1}^m \frac{\partial f}{\partial x^i}(X_t)\; x_t^i\, dt + f(X_t) - f(X_{t-})\,, \tag{A.5}$$

or, in integral form,

$$f(X_t) = f(X_s) + \int_s^t \sum_{i=1}^m \frac{\partial}{\partial x^i} f(X_\tau)\, x_\tau^i \, d\tau + \sum_{s<\tau\leq t} \{f(X_\tau) - f(X_{\tau-})\} \,. \tag{A.6}$$

Obviously, $f(X_t)$ is PD and RC.

A frequently used special case is (check the formulas!)

$$\begin{aligned} d(X_t Y_t) &= X_t \, y_t \, dt + Y_t \, x_t \, dt + X_t Y_t - X_{t-} Y_{t-} \\ &= X_{t-} \, dY_t + Y_{t-} \, dX_t + (X_t - X_{t-})(Y_t - Y_{t-}) \\ &= X_{t-} \, dY_t + Y_t \, dX_t \,. \end{aligned} \tag{A.7}$$

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

$$d(X_t Y_t) = X_t \, dY_t + Y_t \, dX_t \,. \tag{A.8}$$

The integral form of (A.7) is the so-called rule of *integration by parts*:

$$\int_s^t Y_\tau dX_\tau = Y_t X_t - Y_s X_s - \int_s^t X_{\tau-} \, dY_\tau \,. \tag{A.9}$$

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

$$Y_t X_t = Y_0 X_0 + \int_0^t Y_\tau \, dX_\tau + \int_0^t X_{\tau-} \, dY_\tau \,, \tag{A.10}$$

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. To keep notation simple assume $X_t = \sum_{j=0}^{[t]} x_j$ and $Y_t = \sum_{j=0}^{[t]} y_j$. Then

$$\begin{aligned} X_t Y_t &= \sum_{i=0}^{[t]} x_i \sum_{j=0}^{[t]} y_j \\ &= x_0 y_0 + \sum_{i=1}^{[t]} \sum_{j=0}^{i} y_j \, x_i + \sum_{j=1}^{[t]} \sum_{i=0}^{j-1} x_i \, y_j \\ &= X_0 Y_0 + \sum_{i=1}^{[t]} Y_i \, x_i + \sum_{j=1}^{[t]} X_{j-1} \, y_j \\ &= X_0 \, Y_0 + \int_0^t Y_\tau \, dX_\tau + \int_0^t X_{\tau-} \, dY_\tau \,, \end{aligned}$$

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,

$$X_t = X_0 + \int_0^t x_\tau \, d\tau \, , Y_t = Y_0 + \int_0^t y_\tau \, d\tau \, .$$

Take $X_0 = Y_0 = 0$ for the time being. Then

$$\begin{aligned} X_t Y_t &= \int_0^t x_\sigma \, d\sigma \int_0^t y_\tau \, d\tau \\ &= \int\int_{0<\tau\le\sigma\le t} y_\tau \, d\tau \, x_\sigma \, d\sigma + \int\int_{0<\sigma<\tau\le t} x_\sigma \, d\sigma \, y_\tau \, d\tau \\ &= \int_0^t \int_0^\sigma y_\tau \, d\tau \, x_\sigma \, d\sigma + \int_0^t \int_0^{\tau-} x_\sigma \, d\sigma \, y_\tau \, d\tau \\ &= \int_0^t Y_\sigma x_\sigma \, d\sigma + \int_0^t X_{\tau-} \, y_\tau \, d\tau \\ &= \int_0^t Y_\tau \, dX_\tau + \int_0^t X_\sigma \, dY_\sigma \, , \end{aligned}$$

which also conforms with (A.10): the left limit in the next to last line disappeared since an integral with respect to $d\tau$ 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

$$\hat{y}_0 = Y_0 \, , \qquad \hat{y}_j = \int_{j-1}^j y_\tau \, d\tau \, , \; j = 1, \ldots, [t].$$

We have $X_t = X_{[t]}$ and

$$\begin{aligned} X_t Y_t &= X_{[t]} Y_{[t]} + X_{[t]}(Y_t - Y_{[t]}) \\ &= \sum_{i=0}^{[t]} x_i \sum_{j=0}^{[t]} \hat{y}_j + X_{[t]} \int_{[t]}^t y_\tau \, d\tau \, . \end{aligned} \tag{A.11}$$

Upon applying our first result for two discrete functions, the first term in (A.11) becomes

$$\begin{aligned} &X_0 Y_0 + \sum_{i=1}^{[t]} \sum_{j=0}^{i} \hat{y}_j \, x_i + \sum_{j=1}^{[t]} \sum_{i=0}^{j-1} x_i \, \hat{y}_j \\ &= X_0 Y_0 + \sum_{i=1}^{[t]} Y_j \, x_i + \sum_{j=1}^{[t]} X_{j-1} \int_{j-1}^j y_\tau \, d\tau \\ &= X_0 Y_0 + \int_0^t Y_\tau \, dX_\tau + \int_0^{[t]} X_{\tau-} dY_\tau \end{aligned}$$

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 $dY_\tau = y_\tau \, d\tau$.

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 $N_t = \sharp\{j\,;\,t_j \le t\} = \sum_j 1_{[t_j,\infty)}(t)$ or, putting $t_0 = 0$, $N_t = j$ for $t_j \le t < t_{j+1}$. The function $N = \{N_t\}_{t\ge 0}$ 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 $f : \mathcal{R} \mapsto \mathcal{R}$ and for N defined above,

$$
\begin{aligned}
f(N_t) &= f(N_s) + \sum_{s<\tau\le t} \{f(N_\tau) - f(N_{\tau-})\} && \text{(A.12)}\\
&= f(N_s) + \sum_{s<\tau\le t} \{f(N_{\tau-}+1) - f(N_{\tau-})\}(N_\tau - N_{\tau-}) && \text{(A.13)}\\
&= f(N_s) + \int_s^t \{f(N_{\tau-}+1) - f(N_{\tau-})\}\,dN_\tau\,. && \text{(A.14)}
\end{aligned}
$$

Basically, what these expressions state, is just

$$
f(j) = f(0) + \sum_{i=1}^{j} \{f(i) - f(i-1)\}\,.
$$

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 $N_t = \sharp\{\tau \in (0,t];\, X_\tau \ne X_{\tau-}\}$, the number of discontinuities of X within time t. Equipped with our notion of integral, we can now express X as

$$
dX_t = x_t^c\,dt + x_t^d\,dN_t\,, \tag{A.15}
$$

where $x_t^c = x_t$ is the instantaneous rate of continuous change and $x_t^d = X_t - X_{t-}$ is the size of the jump, if any, at t. Generalizing (A.14), we have

$$
f(X_t) = f(X_s) + \int_s^t \frac{d}{dx} f(X_\tau)\,x_\tau^c\,d\tau + \int_s^t \{f(X_{\tau-} + x_\tau^d) - f(X_{\tau-})\}\,dN_\tau\,. \tag{A.16}
$$

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

$$I_A(\omega) = \begin{cases} 1 & \text{if} \quad \omega \in A\,, \\ 0 & \text{if} \quad \omega \in A^c\,, \end{cases}$$

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

$$I_{A^c} = 1 - I_A\,, \tag{B.1}$$

where $A^c = \Omega \backslash A$ is the complement of A.

For any two sets A and B (subsets of Ω),

$$I_{A\cap B} = I_A I_B \tag{B.2}$$

and

$$I_{A\cup B} = I_A + I_B - I_A I_B\,. \tag{B.3}$$

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 $\{A_1, \ldots, A_r\}$:

$$I_{\cap_{j=1}^r A_j} = \prod_{j=1}^r I_{A_j}\,, \tag{B.4}$$

$$I_{\cup_{j=1}^r A_j} = \sum_j I_{A_j} - \sum_{j_1<j_2} I_{A_{j_1}} I_{A_{j_2}} + \ldots + (-1)^{r-1} I_{A_1} \cdots I_{A_r}\,. \tag{B.5}$$

The relation (B.4) is obvious. To demonstrate (B.5), we need the identity

$$\prod_{j=1}^r (a_j + b_j) \quad = \quad \sum_{p=0}^r \sum_{r\backslash p} a_{j_1} \cdots a_{j_p} b_{j_{p+1}} \cdots b_{j_r}\,, \tag{B.6}$$

where $r \backslash p$ signifies that the sum ranges over all $\binom{r}{p}$ 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 α 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 ω 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 ω belongs to A_k). By $\liminf A_j$ is meant the set of points ω 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 ω 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

$$I_{\cup_j \cap_{k \geq j} A_k} = \sup_j \inf_{k \geq j} I_{A_k} \,,$$

obtained upon combining (B.8) and (B.9).

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;

$$I_A \sim \mathrm{Bin}(1, \mathbb{P}[A])\,.$$

It follows that

$$\mathbb{E}\,[I_A] = \mathbb{P}[A]\,, \quad \mathbb{V}\,[I_A] = \mathbb{P}[A](1 - \mathbb{P}[A])\,. \tag{B.10}$$

Since we often will need to equip indicator functions with subscripts, we will use the notation $1[A]$ and I_A interchangeably.

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

$$Z_p = \sum_{j_1 < \ldots < j_p} \mathbb{P}\left[A_{j_1} \cap \cdots \cap A_{j_p}\right] \ , \ \ p = 1, \ldots, r, \tag{C.1}$$

and define in particular $Z_0 = 1$.

Theorem
The probability distribution of Q *can be expressed by the* Z_p *in* (C.1) as

$$\mathbb{P}[Q = q] \quad = \quad \sum_{p=q}^{r} (-1)^{p-q} \binom{p}{p-q} Z_p \, , \qquad q = 0, \ldots, r \, , \tag{C.2}$$

$$\mathbb{P}[Q \geq q] \quad = \quad \sum_{p=q}^{r} (-1)^{p-q} \binom{p-1}{p-q} Z_p \, , \qquad q = 1, \ldots, r \, . \tag{C.3}$$

Proof: Obviously,

$$\{Q = q\} = \bigcup_{r \setminus q} A_{j_1} \cap \cdots \cap A_{j_q} \cap A^c_{j_{q+1}} \cap \cdots \cap A^c_{j_r} \, .$$

The elements in the union are mutually disjoint, and so

$$I_{\{Q=q\}} = \sum_{r \setminus q} I_{j_1} \cdots I_{j_q} (1 - I_{j_{q+1}}) \cdots (1 - I_{j_r}) \, .$$

Starting from this expression, the generating function of the sequence $\{I_{\{Q=p\}}\}_{p=0,\dots,r}$ can be shaped as follows by repeated use (B.6):

$$\begin{aligned}\sum_{p=0}^{r} s^p I_{\{Q=p\}} &= \sum_{p=0}^{r} s^p \sum_{r\backslash p} I_{j_1}\cdots I_{j_p}(1-I_{j_{p+1}})\cdots(1-I_{j_r}) \\ &= \sum_{p=0}^{r} \sum_{r\backslash p} sI_{j_1}\cdots sI_{j_p}(1-I_{j_{p+1}})\cdots(1-I_{j_r})\,. \\ &= \prod_{j=1}^{r} (sI_j + 1 - I_j) \\ &= \prod_{j=1}^{r} ((s-1)I_j + 1) \\ &= \sum_{p=0}^{r} \sum_{r\backslash p} (s-1)I_{j_1}\cdots(s-1)I_{j_p}\,1^{r-p}\,,\end{aligned}$$

where the first term corresponding to $p = 0$ is to be interpreted as 1. Thus,

$$\sum_{p=0}^{r} s^p I_{\{Q=p\}} = \sum_{p=0}^{r} (s-1)^p Y_p\,, \tag{C.4}$$

where

$$Y_p = \sum_{j_1<\ldots<j_p} I_{j_1}\cdots I_{j_p}\,,\ \ p = 1,\dots,r, \tag{C.5}$$

and $Y_0 = 1$. Upon differentiating (C.4) q times with respect to s and putting $s = 0$, we get

$$q!I_{\{Q=q\}} = \sum_{p=q}^{r} p^{(q)}(-1)^{p-q}Y_p\,,$$

hence, noting that $p^{(q)}/q! = \binom{p}{q} = \binom{p}{p-q}$,

$$I_{\{Q=q\}} = \sum_{p=q}^{r} (-1)^{p-q}\binom{p}{p-q} Y_p\,. \tag{C.6}$$

Taking expectation, we arrive at (C.2).

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:

$$\begin{aligned}\sum_{p=0}^{r} s^p \left(I_{\{Q\geq p\}} - I_{\{Q\geq p+1\}}\right) &= \sum_{p=0}^{r} s^p I_{\{Q\geq p\}} - \sum_{p=1}^{r} s^{p-1} I_{\{Q\geq p\}} \\ &= 1 + \sum_{p=1}^{r} (s-1)s^{p-1} I_{\{Q\geq p\}}\,.\end{aligned}$$

Thus, recalling that $Y_0 = 1$, (C.4) is equivalent to

$$\sum_{p=1}^{r} s^{p-1} I_{\{Q\geq p\}} = \sum_{p=1}^{r} (s-1)^{p-1} Y_p\,.$$

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

$$I_{\{Q\geq q\}} = \sum_{p=q}^{r}(-1)^{p-q}\binom{p-1}{p-q}Y_p\,, \tag{C.7}$$

which implies (C.3). □

B. Comments and examples. Setting all A_j equal to the sure event Ω, all the indicators I_j become identically 1. Thus Y_p defined by (C.5) becomes $\binom{r}{p}$, and (C.7) specializes to

$$1 = \sum_{p=q}^{r}(-1)^{p-q}\binom{p-1}{p-q}\binom{r}{p}\,. \tag{C.8}$$

For $q=r$, the theorem reduces to the trivial result

$$P[Q=r] = \mathbb{P}[Q\geq r] = \mathbb{P}\left[A_1\cap\cdots\cap A_r\right]\,.$$

Putting $q=1$ in (C.3) and noting that $[Q\geq 1] = \cup_j A_j$, yields

$$\begin{aligned}\mathbb{P}\left[\cup_j A_j\right] \quad &= \quad \sum_j \mathbb{P}[A_j] - \sum_{j_1<j_2}\mathbb{P}[A_{j_1}\cap A_{j_2}]\\ &\quad + \ldots + (-1)^{r-1}\mathbb{P}[A_1\cap\cdots\cap A_r]\,,\end{aligned}$$

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 $\mathbb{P}[A\cup B] = \mathbb{P}[A]+\mathbb{P}[B]-\mathbb{P}[A\cap B]$ (see (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

$$\mathbb{P}[Q=2] = \mathbb{P}[A_1\cap A_2]+\mathbb{P}[A_1\cap A_3]+\mathbb{P}[A_2\cap A_3]-3\mathbb{P}[A_1\cap A_2\cap A_3]\,.$$

From (C.3) we obtain the probability of at least two occurrences,

$$\mathbb{P}[Q\geq 2] = \mathbb{P}[A_1\cap A_2]+\mathbb{P}[A_1\cap A_3]+\mathbb{P}[A_1\cap A_3]-2\mathbb{P}[A_1\cap A_2\cap A_3]\,.$$

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 $\mathbb{P}\left[\cap_{i=1}^{p}A_{j_i}\right] = \prod_{i=1}^{p}\mathbb{P}[A_{j_i}]$, and the expressions in (C.1) – (C.3) can be computed from the $\mathbb{P}[A_j]$ 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,

$$\sigma_i^2 = \mathbb{V}[X_i]\,,\ i = 1, 2, \ldots \tag{D.1}$$

Define

$$b_n^2 = \sum_{i=1}^{n} \sigma_i^2\ ,\ n = 1, 2, \ldots \tag{D.2}$$

The celebrated Lindeberg/Feller central limit theorem says that if

$$\frac{\sum_{i=1}^{n} \mathbb{E}\left[X_i^2 1[X_i^2 > \varepsilon b_n^2]\right]}{b_n^2} \to 0\ \ ,\forall \varepsilon > 0, \tag{D.3}$$

then

$$\frac{\sum_{i=1}^{n} X_i}{b_n} \xrightarrow{d} \mathrm{N}(0, 1)\,. \tag{D.4}$$

B. Asymptotic properties of MLE estimators The asymptotic distributions derived in Chapter 11 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 θ 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 θ 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

$$\Sigma(\theta) = \left(-\mathbb{E}\left[\frac{\partial^2}{\partial \theta \partial \theta'} \ln \Lambda_n\right]\right)^{-1} \tag{D.5}$$

tends to $0^{s\times s}$ as $n \to \infty$, the MLE is asymptotically normally distributed,

$$\hat{\theta} \sim_{\text{as}} \mathrm{N}(\theta, \Sigma(\theta))\,.$$

C. The delta method Assume that $\hat{\theta}$ is a consistent estimator of $\theta \in \Theta$, an open set in $\mathcal{R}^s$, and that $\hat{\theta} \sim_{\text{as}} \mathrm{N}(\theta, \Sigma)$. If $g : \mathcal{R}^s \to \mathcal{R}^r$ is a twice continuously differentiable function of θ, then

$$g(\hat{\theta}) \sim_{\text{as}} \mathrm{N}\left(g(\theta),\ \frac{\partial}{\partial\theta'} g(\theta) \Sigma \frac{\partial}{\partial\theta} g(\theta)\right)\,. \tag{D.6}$$

The result follows easily by inspection of the first order Taylor expansion of $g(\hat{\theta})$ around θ.

Appendix E

The G82M mortality table

Table E.1: The mortality table G82M

x	μ_x	$f(x)$	$\ell_x = 10^5 \bar{F}(x)$	d_x	q_x
0	0.00057586	0.00057586	100 000	58	0.00057911
1	0.00058279	0.00058246	99 942	59	0.00058635
2	0.00059036	0.00058968	99 883	59	0.00059426
3	0.00059863	0.00059758	99 824	60	0.00060289
4	0.00060765	0.00060621	99 764	61	0.00061231
5	0.00061749	0.00061565	99 703	62	0.00062259
6	0.00062823	0.00062598	99 641	63	0.00063381
7	0.00063996	0.00063726	99 578	65	0.00064606
8	0.00065276	0.00064958	99 513	65	0.00065942
9	0.00066672	0.00066304	99 448	67	0.00067401
10	0.00068197	0.00067775	99 381	69	0.00068993
11	0.00069861	0.00069380	99 312	70	0.00070731
12	0.00071677	0.00071134	99 242	72	0.00072627
13	0.00073659	0.00073048	99 170	74	0.00074697
14	0.00075823	0.00075137	99 096	77	0.00076956
15	0.00078184	0.00077417	99 019	78	0.00079422
16	0.00080761	0.00079906	98 941	81	0.00082113
17	0.00083574	0.00082621	98 860	85	0.00085050
18	0.00086644	0.00085583	98 775	87	0.00088256
19	0.00089994	0.00088814	98 688	90	0.00091754

x	μ_x	$f(x)$	$\ell_x = 10^8 \bar{F}(x)$	d_x	q_x
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	98 195	112	0.00114669
25	0.00117608	0.00115353	98 083	119	0.00120582
26	0.00123790	0.00121270	97 964	124	0.00127034
27	0.00130538	0.00127718	97 840	131	0.00134076
28	0.00137902	0.00134743	97 709	139	0.00141762
29	0.00145940	0.00142394	97 570	146	0.00150150
30	0.00154713	0.00150727	97 424	155	0.00159304
31	0.00164288	0.00159800	97 269	165	0.00169293
32	0.00174738	0.00169678	97 104	175	0.00180196
33	0.00186144	0.00180428	96 929	186	0.00192094
34	0.00198594	0.00192125	96 743	199	0.00205078
35	0.00212181	0.00204849	96 544	211	0.00219247
36	0.00227011	0.00218686	96 333	226	0.00234710
37	0.00243197	0.00233728	96 107	242	0.00251584
38	0.00260863	0.00250075	95 865	259	0.00269998
39	0.00280144	0.00267834	95 606	277	0.00290091
40	0.00301189	0.00287119	95 329	298	0.00312018
41	0.00324157	0.00308050	95 031	319	0.00335944
42	0.00349226	0.00330759	94 712	343	0.00362051
43	0.00376588	0.00355382	94 369	369	0.00390537
44	0.00406451	0.00382066	94 000	396	0.00421619
45	0.00439045	0.00410964	93 604	426	0.00455532
46	0.00474620	0.00442239	93 178	459	0.00492533
47	0.00513447	0.00476062	92 719	494	0.00532902
48	0.00555825	0.00512607	92 225	532	0.00576943
49	0.00602077	0.00552060	91 693	574	0.00624989
50	0.00652560	0.00594609	91 119	617	0.00677402
51	0.00707658	0.00640446	90 502	665	0.00734576
52	0.00767794	0.00689767	89 837	716	0.00796941
53	0.00833430	0.00742765	89 121	770	0.00864963
54	0.00905067	0.00799632	88 351	830	0.00939153
55	0.00983254	0.00860553	87 521	893	0.01020062
56	0.01068591	0.00925700	86 628	960	0.01108295
57	0.01161732	0.00995232	85 668	1 032	0.01204506
58	0.01263389	0.01069283	84 636	1 108	0.01309408
59	0.01374341	0.01147959	83 528	1 189	0.01423775

x	μ_x	$f(x)$	$\ell_x = 10^8 \bar{F}(x)$	d_x	q_x
60	0.01495440	0.01231325	82 339	1 275	0.01548448
61	0.01627611	0.01319402	81 064	1 366	0.01684342
62	0.01771868	0.01412149	79 698	1 460	0.01832447
63	0.01929317	0.01509456	78 238	1 560	0.01993841
64	0.02101162	0.01611127	76 678	1 664	0.02169690
65	0.02288721	0.01716867	75 014	1 771	0.02361259
66	0.02493430	0.01826263	73 243	1 882	0.02569916
67	0.02716858	0.01938769	71 361	1 996	0.02797145
68	0.02960717	0.02053690	69 365	2 112	0.03044546
69	0.03226874	0.02170163	67 253	2 229	0.03313851
70	0.03517368	0.02287138	65 024	2 345	0.03606928
71	0.03834425	0.02403370	62 679	2 461	0.03925790
72	0.04180475	0.02517403	60 218	2 573	0.04272605
73	0.04558167	0.02627566	57 645	2 680	0.04649704
74	0.04970395	0.02731973	54 965	2 781	0.05059590
75	0.05420317	0.02828534	52 184	2 873	0.05504946
76	0.05911381	0.02914974	49 311	2 953	0.05988641
77	0.06447348	0.02988871	46 358	3 020	0.06513740
78	0.07032323	0.03047703	43 338	3 069	0.07083506
79	0.07670789	0.03088920	40 269	3 102	0.07701411
80	0.08367637	0.03110030	37 167	3 111	0.08371127
81	0.09128204	0.03108704	34 056	3 098	0.09096537
82	0.09958318	0.03082907	30 958	3 059	0.09881727
83	0.10864338	0.03031033	27 899	2 994	0.10730975
84	0.11853205	0.02952051	24 905	2 901	0.11648747
85	0.12932494	0.02845661	22 004	2 781	0.12639675
86	0.14110474	0.02712418	19 223	2 635	0.13708534
87	0.15396168	0.02553851	16 588	2 465	0.14860208
88	0.16799427	0.02372520	14 123	2 274	0.16099656
89	0.18331000	0.02172028	11 849	2 066	0.17431855
90	0.20002621	0.01956945	9 783	1 845	0.18861740
91	0.21827096	0.01732660	7 938	1 619	0.20394126
92	0.23818401	0.01505134	6 319	1 392	0.22033622
93	0.25991791	0.01280578	4 927	1 172	0.23784520
94	0.28363917	0.01065073	3 755	963	0.25650675
95	0.30952951	0.00864156	2 792	772	0.27635358
96	0.33778728	0.00682433	2 020	601	0.29741104
97	0.36862894	0.00523248	1 419	453	0.31969526
98	0.40229077	0.00388474	966	332	0.34321126
99	0.43903065	0.00278447	634	233	0.36795078

x	μ_x	$f(x)$	$\ell_x = 10^8 \bar{F}(x)$	d_x	q_x
100	0.47913004	0.00192066	401	158	0.39389013
101	0.52289614	0.00127047	243	102	0.42098791
102	0.57066422	0.00080282	141	64	0.44918271
103	0.62280022	0.00048261	77	37	0.47839101
104	0.67970357	0.00027473	40	20	0.50850528
105	0.74181017	0.00014737	20	11	0.53939240
106	0.80959582	0.00007408	9	5	0.57089277
107	0.88357981	0.00003469	4	2	0.60281998
108	0.96432893	0.00001504	2	1	0.63496163
109	1.05246177	0.00000599	1	1	0.66708109
110	1.14865351	0.00000218	0	0	0.69892078

Appendix F

Exercises

Exercise 1
Acquaint yourself with the software Turbo and the program 'ode-1.pas'. Use the program to solve the following simple problems:

(a) Compute the integrals $\int_0^1 t^2\,dt$ and $\int_0^1 t^{-1/2}\,dt$ and check the answers.

(b) Produce a table of values of the standard normal distribution function $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-\frac{t^2}{2})\,dt$. (Note that $\Phi(-x) = 1 - \Phi(x)$ so you only need to compute the integral $\int_0^x$.)

(c) Consider the payment function $A_t = t$, $t \in [0, 10]$, and let the rate of interest be constant $r = 0.05$. Compute the cash balance U_t by a forward scheme and the reserve V_t (which is here negative of course) by a backward scheme for $t = 0, 1, \ldots, 10$.

Exercise 2
Draw a sketch of the payment function A_t and the cash balance U_t for the following deterministic payment streams, assuming that the interest rate is 4.5% p.a. ($r = \ln(1.045)$):

(a) A single payment (endowment) of 1 at time 5.

(b) An annuity-due with payments of 1 at times $0, 1, \ldots, 9$.

(c) An immediate annuity with payments of 1 at times $1, \ldots, 10$.

(d) An annuity paid continuously at rate 1 per time unit in the time interval $[0, 10]$. Compare the graphs in (b) – (d).

(e) Now consider the random payment function generated by premiums less benefits on a temporary term assurance with sum 1, payable immediately upon death within 10 years, against premium payable continuously at level rate 0.00212 during the insurance period. (This is the equivalence premium if the insured is 30 years old upon issue of the contract, and we use the Danish mortality law.) Draw a sketch of the payment function and the cash balance in the case where the insured dies at time $t = 5$ and in the case where the insured survives throughout the term of 10 years.

Exercise 3
Draw the graphs in Figures 3.1 - 3.7 in the lecture notes (Basic Life Insurance Mathematics). Use the program 'ode-1.pas' to do the computations.

Exercise 4

Find expressions for $f(t)$, $\bar{F}(t)$, $\mu(t)$, $f(t|x)$, and $\bar{F}(t|x)$ and sketch graphs of these functions for the following mortality laws, all of which have finite ω:

(a) De Moivre's *uniform mortality law*,

$$F(t) = \frac{t \wedge \omega}{\omega} .$$

(b) The more general law given by

$$F(t) = 1 - \left(1 - \frac{t \wedge \omega}{\omega}\right)^{\alpha} ,$$

$\alpha > 0$. Look in particular at the case $\alpha = 1/2$.

(c) The law given by

$$F(t) = \frac{\ln(1 + (t \wedge \omega))}{\ln(1 + \omega)} .$$

Exercise 5

Population statistics shows that more than half of all new-born are boys (e.g. in Norway some 51.2% at the present). It also shows that the mortality law depends on the sex. Denote by s_0^m the probability that a new-born is male and by $s_0^f = 1 - s_0^m$ the probability that a new-born is female. Let ${}_tp_0^m$ and μ_t^m be the survival function and mortality intensity, respectively, for males and let ${}_tp_0^f$ and μ_t^f be the corresponding functions for females.

(a) Find the probability s_t^m that a randomly chosen t years old person is a male.

(b) Find the survival function ${}_tp_0$ and the mortality intensity μ_t in the population (of males and females). If μ_t^m and μ_t^f are both of Gompertz-Makeham form, does it then follow that also μ_t is Gompertz-Makeham?

(c) Assume that female mortality is lower than male mortality at all ages. (This is typically the case, and a common way of modeling this difference is by age reduction, e.g. take $\mu_t^f = \mu_{t-5}^m$.) Show that under this assumption s_t^m is a decreasing function. Give a sufficient condition for s_t^m to tend to 0 as t tends to ∞.

(d) Express the following quantities pertaining to the population in terms of s_x^m, s_x^f, and the corresponding quantities specific for males and females, respectively: ${}_tp_x$, ${}_{m|n}q_x$, ${}_nE_x$, and, in general the net premium reserve at any time during the term of an insurance contract.

Exercise 6

A survival function of polynomial form can be fitted arbitrarily well to a given (finite) set of (non-contradictory) constraints if we choose the degree of the polynomial high enough. Obviously, the maximum attainable age ω must be finite for a polynomial survival function.

Consider henceforth a trinomial survival function,

$$\bar{F}(t) = a_0 + a_1\,t + a_2\,t^2 + a_3\,t^3\,,$$

$t \in [0, \omega]$.

(a) Find the general expressions for the density $f(t)$ and the mortality intensity $\mu(t)$. Verify directly, by inspection of the expression, that $\lim_{t \nearrow \omega} \mu(t) = \infty$. Observe that $\bar{F}(t|x)$, considered as function of t, is also trinomial.

(b) Having four parameters, the coefficients $(a_0, \ldots, a_3)$, a trinomial survival functions can be fitted to four independent constraints. It should certainly be required that $\bar{F}(0) = 1$ and it is quite natural to fix an ω, hence require that $\bar{F}(\omega) = 0$. Two more constraints should be added, e.g. $f(\omega) = 0$ and that the expected life length be equal to a given age $e_0 \in (0, \omega)$. Spell out the equations that the coefficients must satisfy. One must, of course, check that the solution produces a $\bar{F}$ that is decreasing and thus is a genuine survival function.

Exercise 7

We adopt the Danish technical basis in (3.25) and consider a policy issued to a life at age $x = 30$ and with term $n = 30$. Find and draw a sketch of the probability distribution of the present value at time 0 of the following benefits: (a) A pure endowment benefit of 1; (b) A term insurance with sum 1; (c) An endowment insurance with sum 1; (d) A life annuity of 1 per year.

Exercise 8

(a) Prove and explain the following relationships:

$$\begin{aligned}
{}_{m+n}E_x &= {}_mE_x\,{}_nE_{x+m}\,, \\
{}_{m|n}\bar{a}_x &= {}_mE_x\,\bar{a}_{x+m\,\overline{n}|}\,, \\
{}_{m|n}\bar{A}_x &= {}_mE_x\,\bar{A}_{x+m\,\overline{n}|}\,.
\end{aligned}$$

(b) A more general rule for expected present values of deferred benefits can be formulated in terms of the prospective reserve: Verify that

$$V_t = V_{(t,u]} + {}_{u-t}E_{x+t}\,V_u\,,$$

where $V_{(t,u]}$ is the expected present value at time t of benefits less premiums payable in the time interval $(t, u]$.

(c) Find formulas (analogous to those derived in Section 4.1) for the variances of the present values with expected values listed in item (a) above.

Exercise 9

Let $T > 0$ be the time of occurrence of some event and let $N_t = 1_{[T,\infty)}(t)$ be the (very simple) counting function that, for each time $t \geq 0$, counts how many times the event has occurred by time t (none before time T and once at and after time T).

Application of the rule of integration by parts to the product $N_t^2 = N_t\, N_t$, gives

$$N_t^2 = N_0^2 + \int_0^t N_\tau\, dN_\tau + \int_0^t N_{\tau-}\, dN_\tau\,. \tag{F.1}$$

Check that this relation is true for all $t \geq 0$ by finding the value of each individual term for $t < T$ (trivial) and for $t \geq T$ (almost trivial). The left-limit in the last integral on the right hand side of (F.1) is now essential since the two factors in the product certainly have a common jump. What happens if you ignore the left-limit?

Explain that the first integral on the right hand side of (F.1) is the same as $\int_0^t (N_{\tau-} + 1)\, dN_\tau$ and that $\int_0^t N_{\tau-}\, dN_\tau = 0$ for all t. Thus (F.1) boils down to the trivial fact that $N_t = \int_0^t dN_\tau$.

Exercise 10

Using the program 'ode-1.pas', do Items (a), (b), and (c) below; for each item run the program and print the output file.

(a) Compute the survival function ${}_tp_0$ for the mortality law in (3.25) in BL and output values for $t = 0, 10, ..., 100$.

(b) Modify the program so that it computes ${}_tp_{70}$ and outputs its values for t = 0,1,...,30. (Insert the following statements in the appropriate places: x := 70; term := 30; outp := 30;)

(c) Now consider the expected life length in t years for a new-born,

$$\bar{e}_{0:\overline{t}|} = \int_0^t {}_\tau p_0\, d\tau\,.$$

Make the program compute and output values for $t = 0, 10, ..., 100$.

Exercise 11

Study the program 'thiele1.pas', which solves Thiele's differential equation numerically for a fairly general single-life policy.

(a) Check the accuracy of the program by computing some functions that possess closed formulas (e.g. $\bar{a}_{\overline{30}|} = (1 - e^{-30r})/r$ by setting $n = 30$, $\mu = 0$, $ba = 1$, and all other payments null).

In the following use the Danish technical basis given in the program.

(b) Compute ${}_{30-t}E_{30+t}$, $\bar{a}_{30+t\,\overline{30-t}|}$, $\bar{A}^{\;1}_{30+t\,\overline{30-t}|}$, $\bar{A}_{30+t\,\overline{30-t}|}$, ${}_{(15-t)_+\,|\,15-(t-15)_+}\bar{a}_{30+t}$ for $t = 0, 5, \ldots, 30$.

(c) Compute the quantities listed in (b), now with $r = 0$, to find ${}_{30-t}p_{30+t}$, $\bar{e}_{30+t\,\overline{30-t}|}$, ${}_{30-t}q_{30+t}$, 1 (!), and ${}_{(15-t)_+\,|\,15-(t-15)_+}\bar{e}_{30+t}$. Compare with the results in (b) to get an impression of the impact of interest.

(d) Setting instead the mortality rate μ equal to 0, compute the quantities in (b) again (some of them are now trivial), identify their proper names and symbols in the world of deterministic payments, and compare with the quantities obtained in (b) to get an impression of the impact of mortality.

(e) Do the same exercise again, now assuming that both r and μ are 0, which means that you just have to write down the total (not discounted)payments under some simple deterministic contracts. Certainly you do not need the program to find the answers. Compare the results with those obtained in the previous items.

(f) Going back to the full model with mortality and interest, compute the net premium level π and net premium reserve V_t at times $t = 0, 5, 10, 15, 20, 25, 30$ for the standard forms of insurance treated in Chapter 3, first for the (rare) case with a single premium at time 0 (has already been done in Item (b) above), and then for the case with continuous premium at level rate. For the deferred annuity, take $m = n = 15$. Draw the missing graphs in Figures 4.1 – 4.5.

(g) Compute the variances of the random variables whose expected values are listed in (b).

Exercise 12

Revisit Exercise 11 Item (b), which lists the reserves for four contracts with standard benefits against single premium at time 0. Compute the corresponding reserves under the modified scheme where $k\,V_t$ is paid back to the insured (i.e. his or her dependents) upon death during the term of the contract. Use $k = 0.5$. (You have already the numbers for the cases $k = 0$ and $k = 1$, and you should be able to explain why that is so.)

Exercise 13

Prove that, if the mortality intensity is an increasing function of age, then the net premium reserve for a deferred level benefit life annuity against level equivalence premium in the deferred period (both continuously paid) is an increasing function of time throughout the deferred period, and thereafter is a decreasing function.

Exercise 14

Consider the endowment insurance against continuously payable premium, which comprises the three main types of contingent payments: a death benefit of b_t at time $t \in (0, n)$, a life endowment of b_n at time n, and a life annuity with intensity $-\pi_t$ at time $t \in (0, n)$. Thiele's differential equation for the reserve V_t is

$$\frac{d}{dt} V_t = \pi_t - b_t \, \mu_{x+t} + (r + \mu_{x+t}) \, V_t \, , \tag{F.2}$$

$t \in (0, t)$, with side condition

$$V_{n-} = b_n \, .$$

We remind of the direct construction of the differential equation at the beginning of Section 4.4 and add here some further details: Denote by $PV_{(t,u]}$ the present value at time t of benefits less premiums in the time interval $(t, u]$. By definition, $V_t = \mathbb{E}[PV_{(t,n]} \, | \, T_x > t]$. We have

$$PV_{(t,n]} = PV_{(t,t+dt]} + e^{-r\,dt} PV_{(t+dt,n]} \, . \tag{F.3}$$

Take expectation in (F.3), given that the policy is in force at time t. On the left we get V_t. On the right we use the rule of iterated expectation, conditioning on what happens in the small time interval $(t, t+dt]$: with probability $\mu_{x+t} \, dt$ the insured dies, and the conditional expected value is then just b_t; with probability $1 - \mu_{x+t} \, dt$ the insured survives, and the conditional expected value is then $-\pi_t dt + e^{-r\,dt} \, V_{t+dt}$. We gather

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

Subtracting V_{t+dt} on both sides, dividing by dt, and letting dt tend to 0, we arrive at (F.2).

(a) Use the same technique to obtain a differential equation for the conditional second order moment, $V_t^{(2)} = \mathbb{E}[(PV_{(t,n]})^2 \, | \, T_x > t]$. Start by squaring (F.3),

$$(PV_{(t,n]})^2 = (PV_{(t,t+dt]})^2 + 2 \, PV_{(t,t+dt]} \, e^{-r\,dt} PV_{(t+dt,n]} + e^{-2r\,dt} (PV_{(t+dt,n]})^2 \, ,$$

and work along the lines above. Having obtained a differential equation for $V_t^{(2)}$, you find also one for the conditional variance $M_t^{(2)} = V_t^{(2)} - V_t^2$ upon differentiating:

$$\frac{d}{dt} M_t^{(2)} = \frac{d}{dt} V_t^{(2)} - 2 \, V_t \, \frac{d}{dt} V_t \, .$$

(b) Explain how you can compute the variance function $M_t^{(2)}$ by extending the program 'thiele1.pas'.

Exercise 15

Poisson processes, which are totally memoryless, can of course be generated from continuous time Markov chains, which are more general. For instance, let $\{Z(t)\}_{t\geq 0}$ be a Markov chain on the state space $\{1, 2\}$ with intensities of transition $\mu_{12}(t) = \mu_{21}(t) = \mu$. Then $\{N(t)\}_{t\geq 0}$ defined by $N(t) = N_{12}(t) + N_{21}(t)$ (the total number of transitions in $(0, t]$) is a Poisson process with intensity μ; transitions counted by N occur with intensity μ at any time regardless of the past history of the process. Two independent Poisson processes, $\{N_1(t)\}_{t\geq 0}$ with intensity μ_1 and $\{N_2(t)\}_{t\geq 0}$ with intensity μ_2, can be generated by letting $\{Z(t)\}_{t\geq 0}$ be a Markov chain on the state space $\{1, 2, 3, 4\}$ with intensities $\mu_{12}(t) = \mu_{21}(t) = \mu_{34}(t) = \mu_{43}(t) = \mu_1$, $\mu_{13}(t) = \mu_{31}(t) = \mu_{24}(t) = \mu_{42}(t) = \mu_2$, $\mu_{14}(t) = \mu_{41}(t) = \mu_{23}(t) = \mu_{32}(t) = 0$, and defining $N_1(t) = N_{12}(t)+N_{21}(t)+N_{34}(t)+N_{43}(t)$ and $N_2(t) = N_{13}(t)+N_{31}(t)+N_{24}(t)+N_{42}(t)$. Three independent Poisson processes can be generated from a Markov chain with 8 states (work out the details), and, in general, k independent Poisson processes can be generated from a Markov chain with 2^k states.

In the following let $\{Z(t)\}_{t\geq 0}$ be a Markov chain on the state space $\{1, 2\}$ and take $\mu_{12}(t) = \mu_{21}(t) = 1$. The total time spent by Z in state 1 during the time interval $(t, n]$ is

$$T_1(t, n] = \int_t^n I_1(\tau)\, d\tau \,,$$

and the total number of transitions made from state 1 to state 2 in that interval is

$$N_{12}(t, n] = N_{12}(n) - N_{12}(t) \,.$$

(These quantities can be viewed as present values of benefits of annuity type and assurance type, respectively, for a two-state policy with no interest.)

(a) Assume $Z(0) = 1$. What are the interpretations of the random variables $T_1(t, u]$ and $N_{12}(t, u]$ in terms of the Poisson process $N(t) = N_{12}(t) + N_{21}(t)$?

(b) Find, by solving the relevant differential equations analytically, explicit expressions for the first two state-wise conditional moments

$$V_j^{(q)}(t) = \mathbb{E}[T_1^q(t, n] \mid Z(t) = j] \,,$$

$$W_j^{(q)}(t) = \mathbb{E}[N_{12}^q(t, n] \mid Z(t) = j] \,,$$

$j = 1, 2$, $q = 1, 2$, and find also the corresponding variances. (You should obtain e.g.

$$\mathbb{E}[T_1(t, n] \mid Z(t) = 1] = \frac{1}{4\mu}\left(1 + 2\mu - e^{-2\mu}\right) \,,$$

$$\mathbb{V}\mathrm{ar}[T_1(t, n] \mid Z(t) = 1] = \frac{1}{16\mu^2}\left(1 + 4\mu - (2 - e^{-2\mu})^2\right) \,.$$

(c) Using the program prores1.pas, solve the differential equations also numerically and compare the results with the exact solutions obtained in (b).

Exercise 16

Consider the Markov chain model for death, sickness, and recovery sketched in Figure 7.3, which is apt to describe insurances with payments that depend on the state of health of the insured (e.g. disability pension or waiver of premium during disability). Apart from the names of the states and the symbols for the intensities, we adopt standard notation for the transition probabilities $p_{jk}(t,u)$ and the occupancy probabilities $p_{\overline{jj}}(t,u)$.

(a) Prove, both by a forward and a backward argument, that

$$p_{\overline{aa}}(t,u) = \exp\left(-\int_t^u (\sigma(s)+\mu(s))\,ds\right).$$

(b) Given start in state a at time 0, write up the probability that the process remains in a during the time interval $[0,t_1)$, then jumps to state i in $[t_1, t_1+dt_1)$, then remains in i during the time interval $[t_1+dt_1, t_2)$, then jumps to state a in $[t_2, t_2+dt_2)$, then remains in a during the time interval $[t_2+dt_2, t_3)$, and finally jumps to state d in the time interval $[t_3, t_3+dt_3)$. This is the probability of one particular full specification of the history of the process.

(c) Write down the forward Kolmogorov differential equations for $p_{aa}(s,t)$ and $p_{ai}(s,t)$ and their appropriate initial conditions at time $t=s$. (Since $p_{ad}(s,t) = 1-p_{aa}(s,t)-p_{ai}(s,t)$, there is no need to work with the differential equation for $p_{ad}(s,t)$. This is a useful property of the forward equations.)

(d) Find an explicit solution to the differential equations in (c) in the case where the intensities are constant. Since the transition probabilities in this case depend only on $t-s$, we can set $s=0$ and simplify notation to $p_{jk}(0,t) = p_{aa}(t)$. (Hint: Differentiate the differential equation for e.g. $p_{aa}(t)$, and substitute expressions for $p_{ai}(t)$ and its first order derivative to obtain a second order differential equation involving only $p_{aa}(t)$. This equation can be solved by a standard techniques.)

Exercise 17

Continuing Exercise 16: We consider an insurance policy purchased by an x years old person in active state. We are now interested in the life history after age x, conditional on the state then being a. Therefore, to visualize the insured's age at entry and the time elapsed since issue of the policy, we put $\sigma(t)=\sigma_{x+t}$, $\rho(t)=\rho_{x+t}$, $\mu(t)=\mu_{x+t}$, $\nu(t)=\nu_{x+t}$ and ${}_tp_x^{aa} = p_{aa}(x,x+t)$ etc.

From Exercise 16 (c) we fetch the (forward) differential equations for ${}_tp_x^{aa} = p_{aa}(0,t)$ and ${}_tp_x^{ai} = p_{ai}(0,t)$ and their side conditions.

(a) Find closed form expressions for ${}_tp_x^{aa}$, ${}_tp_x^{ai}$, and ${}_tp_x^{ii}$ in the case without recovery ($\rho_{x+t}\equiv 0$). You can either solve the differential equations or put up the expressions by direct reasoning, see Chapter 7.

(b) The remaining lifetime T_x of (x) has survival function ${}_tp_{[x]} = {}_tp_x^{aa} + p_x^{ai}$. What are the conditional probabilities, ${}_t\tilde{p}_x^{aa}$ and ${}_t\tilde{p}_x^{ai}$, of being active and invalid, respectively, at time t, given survival to time t (and start as active at time 0)? Find the mortality intensity $\mu_{[x]+t}$ of the survival function ${}_tp_{[x]}$. Discuss the expression for $\mu_{[x]+t}$, observe that it is select in general, and give a verbal explanation of the selective mechanism which is at work here.

(c) Assume now that $\mu_{x+t} = \nu_{x+t}$, usually referred to as the case of *non-differential mortality*. In this case $\mu_{[x]+t} = \mu_{x+t}$, of course. Verify that the conditional probabilities ${}_t\tilde{p}_x^{aa}$ and ${}_t\tilde{p}_x^{ai}$ are the corresponding transition probabilities in the *partial model* sketched in Figure F.1. (Check that they are solutions to the forward differential equations for the transition probabilities in the partial model.)

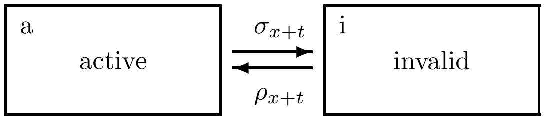


Figure F.1: A partial Markov chain model for sickness and recovery, no death.

(d) Consider the following case: Age at entry is $x = 30$, term of contract is $n = 30$, the benefit is a life insurance of 1, premium is payable at level rate while active (waiver of premium during disability), interest rate and transition intensities are as given in the standard version of the program 'prores1'. Fill in appropriate statements in the program 'prores1' to make it compute the state-wise reserves at times $t = 0, 1, \ldots, 30$.

Exercise 18

Let $Z(t)$, $t \geq 0$ be a time-continuous Markov chain on a finite state space $\mathcal{J} = 1, \ldots, J$, starting in state 1 at time 0; $Z(0) = 1$. Assume that it possesses transition intensities, and adopt standard notation for basic entities: intensities of transition $\mu_{jk}(t)$, transition probabilities $p_{jk}(t, u)$, and probabilities of uninterrupted sojourns $p_{\overline{jj}}(t, u)$.

(a) Let $t_0 \leq t_1 \leq \cdots \leq t_r \leq t_{r+1}$ be times in $[0, \infty)$ and let $j_0, j_1, \ldots, j_r, j_{r+1}$ be states in $\mathcal{J}$. Express the following probabilities in terms of the basic entities:

$$\mathbb{P}\left[\bigcap_{i=0}^{r+1} Z(t_i) = j_i\right] \; ;$$

$$\mathbb{P}\left[\bigcap_{i=1}^{r} Z(t_i) = j_i \,\middle|\, Z(t_0) = j_0,\; Z(t_{r+1}) = j_{r+1}\right] . \tag{F.4}$$

(b) Let s and t be fixed times such that $s < t$, and let i and j be fixed states. Use the result in (F.4) (if you got it right) to show that, conditional on $Z(s) = i$ and $Z(t) = j$, the process $Z(\tau)$, $\tau \in [s, t]$, is a Markov chain.

(c) Suppressing the dependence on s, t and i, j from the notation, denote the conditional Markov chain by $\tilde{Z}(\tau)$, $\tau \in [s, t]$, and denote its transition probabilities and intensities by $\tilde{p}_{gh}(\tau, \vartheta)$ and $\tilde{\mu}_{gh}(\tau)$, respectively. Determine these probabilities and intensities.

(d) What is the limit of $\tilde{\mu}_{gh}(\tau)$ as τ tends to t? Distinguish between the cases where (i) g and h are both different from the "destination state" j, (ii) $g \neq j$ and $h = j$, (iii) $g = j$ and $h \neq j$. Give a verbal explanation of the results.

(e) Now specialize to the disability model in Figure 7.3, and condition on $Z(0) = a$ and $Z(t) = i$. Write out the expression for $\tilde{\sigma}(\tau)$ and the value of $\tilde{\mu}(\tau)$ (trivial). Find an explicit expression for $\tilde{\sigma}(\tau)$ in the case with constant intensities and no recovery, and discuss it with particular attention to the limiting value as τ tends to t.

Exercise 19

(a) Describe in words the single life insurance policy for which Thiele's differential equation is

$$\frac{d}{dt}V_t \,=\, r\,V_t \,+\, c \,-\, \mu_{x+t}(1 - V_t)\,, \quad 0 < t < n\,,$$

subject to the condition

$$V_{n-} = 1\,.$$

Integrate the equation, using the technique with integrating factor, to obtain the direct prospective expression for the reserve V_t. Express the reserve in standard actuarial notation. Explain how to determine c so as satisfy the equivalence requirement $V_0 = 0$. $\square$

(b) Prove that if μ_{x+t} is an increasing function of $t \in [0, n]$, then V_t is an increasing function of $t \in [0, n]$. Construct an example where V_t is not an increasing function of t for all $t \in (0, n)$. $\square$

Exercise 20

(a) Let T be a random lifetime with continuous survival function ${}_tp_0$, and let G be a non-decreasing real-valued function such that $\mathbb{E}[G(T)]$ is finite. Prove the formula

$$\mathbb{E}[G(T)] \,=\, G(0) \,+\, \int_0^\infty {}_tp_0\, dG(t)\,. \;\square \tag{F.5}$$

Let $w(t)$, defined on $[0, \infty)$, be a continuous and strictly positive function such that $\int_0^\infty w(s)\, ds \,=\, \infty$. Define

$$W(t) = \int_0^t w(s)\, ds$$

and

$${}_tp_0 = \left(\frac{\delta}{W(t) + \delta}\right)^\gamma \,=\, (1 + W(t)/\delta)^{-\gamma}\,. \tag{F.6}$$

(b) Show that ${}_tp_0$ is a survival function, and find the corresponding mortality intensity μ_t. Find the conditional survival function ${}_tp_x$ and observe that it is of the same form as ${}_tp_0$, only with different W and δ. Prove that $W(T) + \delta$ is Pareto-distributed with parameters (γ, δ). $\square$

(c) Consider the special case where $w(t) = 2t$, and $\gamma > 1$. Use (F.5) to find an explicit expression for $\mathbb{E}[T^2]$. □

(d) The parameters γ and δ are to be estimated from data on n independent lives observed from birth until attained age z or death, whichever occurs first. Thus, if T_m is the life length of person No. m, the observations are $T_m \wedge z$, $m = 1, \ldots, n$. Write up the log likelihood function, derive the likelihood equations that determine the ML estimators $\hat{\gamma}$ and $\hat{\delta}$, and explain without doing calculations how to find their asymptotic covariance matrix. □

(e) Assume that δ is known. Then there is an explicit expression for $\hat{\gamma}$. Find this expression and use (F.5) to show that $\hat{\gamma}$ converges to γ with probability 1. □

(f) Assume now that the data available are the occurrence-exposure rates in years $j = 1, \ldots, z$. Explain how to estimate γ and δ by the technique of analytic graduation by weighted least squares. □

Exercise 21

An x years old person buys a permanent disability pension policy specifying that premium is paid continuously at fixed rate c per time unit in active state and pension is payable continuously at fixed rate b per time unit in disabled state. The policy terminates z years after issue. The relevant Markov model is sketched in Fig. F.2. Assume that the interest rate r is constant.

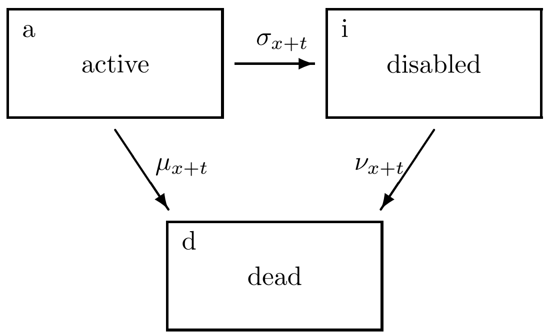


Figure F.2: A Markov chain model for permanent disability insurance.

(a) Derive the differential equations and their side conditions for the reserves in the states a and i. □

(b) Suppose the reserve is paid back upon the insured's death as active before time z. How does this change the differential equations in Item (a)? □

(c) Describe in words the product for which the reserves satisfy the differential equations

$$\begin{aligned}\frac{d}{dt}V_a(t) &= rV_a(t) + c - \sigma_{x+t}(V_i(t) - V_a(t)) - \mu_{x+t}(1 - V_a(t))\,, \\ \frac{d}{dt}V_i(t) &= r\,V_i(t) - b - \nu_{x+t}(1 - V_i(t))\,,\end{aligned}$$

with side conditions $V_a(z-) = 1$ and $V_i(z-) = 1$. $\square$

(d) Write down, without proofs, explicit expressions for the probabilities $p_{aa}(0,t)$, $p_{ii}(0,t)$, and $p_{ai}(0,t)$. $\square$

(e) Assume that n independent policies, identical to the one described above, are observed in z years, and that the intensities are of the form:

$$\begin{aligned}\mu_{x+t} &= \alpha + \beta\, e^{\gamma\,(x+t)}\,, \\ \nu_{x+t} &= \alpha + \beta\, e^{\gamma'\,(x+t)}\,, \\ \sigma_{y+t} &= \sigma \ \ (\text{constant})\,.\end{aligned}$$

Explain how to estimate the parameters α, β, γ, γ', and σ from the data. Write up the log likelihood function and derive the likelihood equations that determine the ML estimator. $\square$

Exercise 22
Use the program 'actuary.pas' to study how the (fixed) premium rate and the reserve per survivor depend on the age of the insured, the length of the premium payment period, the term of the contract, and the interest rate for various forms of insurance. In the case of life annuities, compare with the corresponding bank savings solution.

Exercise 23

Prove that an FV function has at most a countable number of discontinuities. (Hint: Consider a non-decreasing and finite-valued function X. Its set of discontinuities is

$$\mathcal{D}(X) = \cup_{m=1}^{\infty} \cup_{n=1}^{\infty} \mathcal{D}_{m,n}\,,$$

where $\mathcal{D}_{m,n} = \{t\,;\, n-1 \leq t < n\,,\, X_{t+} - X_{t-} \geq \frac{1}{m}\}$. Show that $\mathcal{D}_{m,n}$ is finite, possibly empty. Thus, $\mathcal{D}(X)$ is a countable union of finite sets, and is thus at most countable.)

Exercise 24
Define $\varepsilon_u(t) = 1_{[u,\infty)}(t)$, which for fixed u is a very simple RC and PD function of t. Apply (A.7) to the very transparent case where $X_t = \varepsilon_{t_1}(t)$,and $Y_t = \varepsilon_{t_2}(t)$, $0 < t_1 \leq t_2$, and compare the results of the formal calculations to what you can put up by direct reasoning.

Exercise 25

Let X be an RC and PD function and let N count its discontinuities. Prove that

$$dX_t^q = qX_t^{q-1}\, x_t^c \, dt + \sum_{p=0}^{q-1} \binom{q}{p} X_{t-}^p \, (x_t^d)^{q-p} \, dN_t \,, \tag{F.7}$$

which generalizes the well-known rule $dX_t^q = qX_t^{q-1}\, dX_t$ for continuous X.

Exercise 26

Let N be a counting process and consider a process S defined by

$$S_t = \exp(at + bN_t) \,,$$

where a and b are constants. Use Itô's formula to show that the dynamics of S is given by

$$dS_t = S_{t-} \left(a \, dt + (e^b - 1) dN_t \right) \,.$$

In integral form

$$S_t = 1 + \int_0^t S_{\tau-} \left(a \, d\tau + (e^b - 1) dN_\tau \right) \,.$$

Assume now that N is a Poisson process with intensity λ. By the independent increments property of N, we can heuristically conclude that $X_{\tau-}$ and dN_τ are independent, so

$$\mathbb{E}[S_{\tau-} dN_\tau] = \mathbb{E}[S_{\tau-}] \mathbb{E}[dN_\tau] \,.$$

Use this together with $\mathbb{E}[S_{\tau-}] = \mathbb{E}[S_\tau]$ to obtain an integral equation for $\mathbb{E}[S_t]$. Solve it to find

$$\mathbb{E}[S_t] = \exp\left(at + \left(e^b - 1 \right) \lambda t \right) \,.$$

Verify the result by direct calculation using that N_t has a Poisson distribution with parameter λt.

Exercise 27

Verify (2.4). (A recommended text on functional equations is [2].)

Exercise 28

(a) Apply (A.8) to (2.14), taking $X_t = \exp\left(\int_0^t r_s \, ds \right)$ and $Y_t = \int_t^n \exp\left(- \int_0^\tau r_s \, ds \right) \, dB_\tau$, to obtain (2.19) and integrate over $(t, n]$ to obtain (2.22).

(b) Prove the relationship (2.24). by applying integration by parts (A.9) to $\int_t^n e^{-\int_0^\tau r} \, dB_\tau$ (it is convenient to put $dB_\tau = d(B_\tau - B_t)$). Give a verbal interpretation of (2.24). Then prove (2.23) by using (2.9).

(c) Observe that if B consists only of a unit payable at time n, then, essentially, all the relationships (2.22), (2.23), and (2.24) reduce to (2.9).

Exercise 29

(a) Let $A = -B$ and $A' = -B'$ be two payment functions with retrospective and prospective reserves denoted by U and V and U' and V', respectively. Assuming that the interest rate is always positive, verify the following rather obvious assertions: If $A_t \leq A'_t$ for all t, then $U_t \leq U'_t$ for all t. If, moreover, $A_t < A'_t$ for some t, then $U_\tau \leq U'_\tau$ for all $\tau \geq t$. Similarly, if B_n is finite and $B_n - B_t \leq B'_n - B'_t$ for all t, then $V_t \leq V'_t$ for all t.

(b) In particular, any advancement of deposits will produce an increase of the retrospective reserve if the interest rate is positive. This is the general circumstance underlying results like the following about ordering of the present values in (2.28), (2.29), and (2.31): $a_{\overline{n}|} < \bar{a}_{\overline{n}|} < \ddot{a}_{\overline{n}|}$.

Exercise 30

Let the payment stream A represent deposits less withdrawals on an n years savings account that bears interest with strictly positive interest rate. It is required that $U_t \geq 0$ for all t, with strict inequality for some t, and that $U_n = 0$, see (1.10) and (1.11). Prove that $A_n < 0$, and explain this result.

Exercise 31

We refer to the theory of loans in Subsection 2.2. In practice amortizations are typically due at whole years after the loan is paid out and, accordingly, interest is calculated on an annual basis. Thus, the payment functions A, F, and R are pure jump functions with jumps a_t, f_t, and ρ_t, respectively, at times $t = 1, \ldots, n$. (The symbol ρ_t is chosen since r_t is reserved for the instantaneous interest rate, which is something different.) Denote the annual interest rate in year t by i_t and the corresponding discount factor by $v_t = (1 + i_t)^{-1}$. They are related to the instantaneous interest rate r_t by $1 + i_t = \exp\left(\int_{t-1}^{t} r_\tau \, d\tau\right)$.

(a) Rewrite the relationships appearing in Subsection 2.2 in terms of the A_t, a_t, F_t, f_t, R_t, ρ_t, i_t, and v_t, restricting to discrete times $t = 0, \ldots, n$. In particular, dynamical relations like (2.38) are to be written out as recursive equations in the style of Chapter 1.

(b) Use the program 'actuary.pas' on L:\KUFML (set mortality equal to 0) to study numerically the properties of various forms of loans; fixed loan ($f_t = 0$ for $t = 1, \ldots, n-1$ and $f_n = 1$), series loan ($f_t = f$ for $t = 1, \ldots, n$), and annuity loan ($a_t = a$ for $t = 1, \ldots, n$).

Exercise 32

By inspection of Fig. 3.1, estimate roughly the expected number of years spent in the age intervals $(0, 10]$, $(10, 20], \ldots$, $(90, 100]$ for a new-born. Compute the exact figures.

Exercise 33

Starting from a given aggregate mortality intensity μ_x, there are uncountably many ways of inventing a select mortality law $\mu_{[x],t}$ with select period of a given length s and ultimate mortality law μ_{x+t}, $t \geq s$. We shall propose four possibilities:

(a) A duration dependent weighted mortality intensity,

$$\mu_{[x],t} = \frac{w(t)}{w(0)}\,\nu_{x+t} + \left(1 - \frac{w(t)}{w(0)}\right)\,\mu_{x+t}\,,$$

where ν is some mortality intensity such that $\nu_x < \mu_x$ for all $x > 0$, and w is some non-increasing function such that $w(0) > 0$ and $w(t) = 0$ for $t \geq s$. For instance, one could take $\nu_x = \mu_{x-a}$ for some $a > 0$ (assuming μ is increasing) and

$$w(t) = (1 - t/s)^q \vee 0$$

with $q > 0$. If $q > 1$, then the derivative of w at $t = s$ is 0, making μ "smooth" there. Suggest another choice of w.

(b) A mortality intensity obtained by reducing age in the select period,

$$\mu_{[x],t} = \mu_{x+t-w(t)}$$

with w as in item (a) above.

(c) A simple Cox regression model,

$$\mu_{[x],t} = \exp(\beta_1 w(t) + \beta_2 w^2(t))\mu_{x+t}$$

with w as in item (a) above.

(d) A piecewise constant intensity,

$$\mu_{[x],t} = \lambda_j + \kappa_k\,,\ j-1 \leq x+t < j\,,\ k-1 \leq t < k\,,$$

$j = 1, 2, \ldots$, $k = 1, 2, \ldots, s$.

(e) Discuss in general terms how the select mortality laws in items (a) - (d) can be estimated from data. Matters may simplify in (a) - (c) if you assume that the ultimate mortality law is known. Carry out a more formal analysis for the case (d).

Exercise 34

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 $b = 1$ anually in 10 years, at ages 70,...,79. The annual interest rate is $i = 0.045$. The balance equation is

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

which gives

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

Find the annual premium for the corrresponding life annuity with benefits payable contingent on survival.

Exercise 50

So-called *Loss of profits insurance* covers losses suffered by firms during cessation of production caused by fire or other accidental events. Suppose that a firm at any time is either in state "a = active" (normal production) or "i = inactive" (no production). The insurance can be purchased when the firm is in active state. The indemnification (benefit) is paid continuously at constant rate 1 per time unit in inactive state, and premium is paid as a single amount upon issue of the policy at time 0 (say). Denote the term of the contract by n. Let $Z(t)$ denote the state of the firm at time t. Under these conventions the total indemnification in respect of the policy is the time spent in inactive state during the term of the contract,

$$T_n = \int_0^n I_{\{Z(\tau)=i\}}\, d\tau\,.$$

Assume that Z is a time-homogeneous Markov chain as sketched in Figure F.3; the intensities σ and ρ are constant. Put

$$p_{jk}(t,u) = \mathrm{P}[Z(u) = k \mid Z(t) = j]\,,$$

$0 \le t \le u \le n$ and $j,\ k\ \in \{a, i\}$.

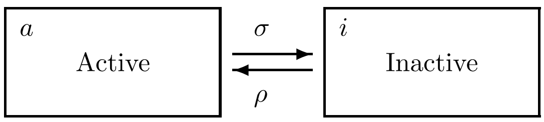


Figure F.3: A Markov chain model for loss of profit insurance.

(a) Show that

$$p_{ai}(t,u) = \frac{\sigma}{\sigma+\rho}\left(1 - e^{-(\sigma+\rho)(u-t)}\right)\,,$$

and $p_{aa}(t,u) = 1 - p_{ai}(t,u)$. The probabilities $p_{ia}(t,u)$ and $p_{ii}(t,u)$ are given by the same expressions when σ and ρ change roles. Explain why the probabilities $p_{jk}(t,u)$ depend only on the length of the time interval, $u - t$. Notice that

$$\lim_{u\to\infty} p_{ai}(t,u) = \frac{\sigma}{\sigma+\rho}\,,$$

and explain why the result is reasonable.

(b) Find the single premium $\pi_n = \mathrm{E}[T_n]$ for the loss of profit insurance described above; Employ the principle of equivalence assuming no interest. Show that the premium per year, π_n/n, is an increasing function of n, with upper limit $\sigma/(\sigma+\rho)$ (compare with the last formula in (a)).

(c) The intensities σ and ρ are to be estimated from statistical data from m independent, identical firms, which have been insured and observed over the same period of time, $[0, n]$. Assume first that for each firm there is a complete record of transitions between states in the observation period. Write up the likelihood function, explain how the maximum likelihood (ML) estimators of the parameters are derived, and find their asymptotic variances. Examine these variances as functions of n for fixed $mn = w$ (total time under observation), and comment.

(d) item For $0 \leq s_1 \leq \cdots \leq s_r \leq t \leq u \leq n$ and $j_1, \ldots, j_r \in \{a, i\}$, find the conditional probability

$$\tilde{p}_{jk}(t,u) = \mathrm{P}[Z(u) = k \,|\, Z(s_1) = j_1, \ldots, Z(s_r) = j_r, Z(t) = j, Z(n) = i]\,,$$

a function of j, k, t, and u. Conclude that, conditional on $Z(n) = i$, the process Z behaves like a Markov chain $\tilde{Z}$ with transition probabilities $\tilde{p}_{jk}(t,u)$. Find the transition intensities $\tilde{\sigma}(t)$ and $\tilde{\rho}(t)$ for $\tilde{Z}$.

(e) Assume now that $\rho = 0$ and that the data available are occurrence-exposure rates in years $j = 1, \ldots, n$ for those firms that are inactive at time n and no information about other firms in the portfolio. Using the results in (a) and (d), explain how to estimate σ by the technique of analytic graduation. (Do not try to find an explicit expression for the estimate of σ.)

Exercise 51

(a) Consider the Markov model for death, sickness, and recovery sketched in Figure 7.3. Apart from the special symbols for the intensities, we adopt standard notation and assumptions: $Z(t)$ is the state at time $t \geq 0$, the process starts from active state, $Z(0) = a$,

$$p_{jk}(t,u) = P[Z(u) = k \mid Z(t) = j]\,, \tag{F.8}$$

and

$$p_{\overline{jj}}(t,u) = P[Z(\tau) = j\,,\, \tau \in [t,u] \mid Z(t) = j]\,. \tag{F.9}$$

(a) item Derive forward differential equations for the probability $p_{ai}^{(1)}(0,t)$ of being disabled for the first time at time t and for the probability $p_{aa}^{(1)}(0,t)$ of being active at time t after having been disabled once. Proceed in the same manner to derive forward differential equations for the probabilities $p_{ai}^{(k)}(0,t)$ and $p_{aa}^{(k)}(0,t)$ of being disabled and active, respectively, at time t after exactly k onsets of disability, for $k = 2, 3, \ldots$.

What is $\sum_{k=1}^{\infty} p_{ai}^{(k)}(0,t)$?

(b) Express the probability

$$P[Z(\tau) = i; \tau \in [t-q, t] \mid Z(0) = a]$$

in terms of the basic entities defined in (F.40) and (F.9).

(c) At time 0 an active person buys an insurance policy which specifies that a pension benefit is to be paid continuously with fixed intensity 1 as long as the insured is disabled and has been so uninterruptedly for at least q years. (The term q is called the *qualifying period.*) The policy expires upon death or, at the latest, at time n ($> q$). A single premium π is paid at time 0.

Determine the premium π by the principle of equivalence, assuming that the interest rate r is constant. Find the reserve at time $t < n - q$ for an insured who is disabled and is currently receiving the disability benefit.

(d) Consider the Markov chain model in Figure 7.3. Fill in appropriate statements in the enclosed program 'prores1' to make it compute the transition probabilities $p_{ai}(t, 10)$ and $p_{ii}(t, 10)$ for $t = 0, 1, \ldots, 10$ in the case where all the intensities are constant: $\mu = \nu = \sigma = 0.01$ and $\rho = 0$.

Exercise 52

(a) We adopt the usual notation and assumptions of the theory of multi-life insurance policies and consider two independent lives (x) and (y) with remaining life lengths T_x and T_y, respectively. Assume that the benefit is an assurance of 1 payable at time T_y if $T_x + n/2 < T_y < n$ and that premium is payable at constant rate π until time $\min(T_x, T_y, n/2)$, where n is the term of the contract (fixed). Determine the equivalence premium π.

Exercise 53

The employees of a firm are automatically members of a pension scheme with salary dependent premiums and benefits defined as follows. Consider an employee (x), who enters the scheme x years old at time 0 (say), retires at pensionable age 65 at time $m = 65 - x$, and earns salary at rate $S(t)$ per time unit at time $t \in [0, m]$. Contingent on survival, premium is payable continuously at rate $\pi S(t)$ at time $t \in [0, m]$, and pension is received as an endowment of $5S(m)$ upon retirement at time m.

Assume that the economy is governed by a continuous time Markov chain $Y(t)$, $t \geq 0$, with state space $\mathcal{J} = \{1, \ldots, J\}$, constant intensities of transition λ_{jk}, $j \neq k$, and initial state $Y(0) = i$, say. At any time $t \geq 0$ the accumulation factor $U(t)$ of the investment portfolio is given by

$$U(t) = \exp\left(\int_0^t r(s)\, ds\right), \qquad r(s) = \sum_j I_j(s) r_j\,,$$

and the salary rate is given by

$$S(t) = \exp\left(\int_0^t a(s)\, ds\right), \qquad a(s) = \sum_j I_j(s) a_j\,.$$

Here, $I_j(t) = 1[Y(t) = j]$, and the r_j and the a_j are known, fixed numbers; r_j is the interest rate and a_j is the rate at which salary increases per time unit and per unit of salary when the economy is in state j.

(a) Demonstrate how to determine the premium rate π that makes expected discounted benefits less premiums equal to 0 at time 0; Construct differential equations for benefits and for premiums and specify appropriate side conditions.

(b) Explain that, if $a_j = r_j$ for all states j so that salary is perfectly linked to investments, then equivalence can be attained in a large (in principle infinitely large) portfolio of identical policies.

Exercise 54

(a) Describe briefly the main characteristics of with-profit insurance and unit-linked insurance.

(b) Consider a pure endowment benefit at time n (the term of the contract) against a single premium at time 0. Assume that the investment portfolio of the insurance company bears interest at rate $r(t)$ at time t, where r is driven by a Markov chain as described in Question 6. Show how to determine the single premium for a unit-linked contract with sum insured $\max(U(n), g)$, where $U(n)$ is the index of the investment portfolio given by $U(t)$ as defined in Question 6, and g is a guaranteed minimum sum insured.

Exercise 55

We adopt the usual notation and assumptions of the theory of multi-life insurance policies and consider two independent lives (x) and (y) with remaining life lengths T_x and T_y, respectively.

(a) Assume that the benefit is an assurance of 1 payable at time T_y if $2T_x < T_y < n$ and that premium is payable at constant rate π until time $\min(T_x, T_y, n/2)$, where n is the term of the contract (fixed). Determine the equivalence premium π.

(b) Propose a method for computing the premium numerically. (Hint: One possibility is to treat ${}_{t/2}p_x$ as a survival function ${}_t\tilde{p}_x$ with intensity $\tilde{\mu}_{x+t}$, which you would need to express in terms of μ, and then solve a Thiele differential equation numerically.)

(c) Determine the reserve at any time t, assuming that the insurer currently knows the complete past history of the two lives. You need to distinguish between various cases, whether (y) is alive or dead, whether t is before or after time $n/2$, and whether x is alive or dead and, if dead, when. Is the reserve always non-negative?

(d) What is the variance of the present value of the benefit?

SURPLUS, BONUS, WITH PROFIT, GUARANTEES, UNIT-LINKED

A. Preliminaries. We are going to restate the theory of surplus and bonus and related problems in the framework of the simple, still fairly general, single life contract treated in Section 4.4 and add material on interest guarantees and unit-linked insurance.

The terms of the contract are set out in the expression for the prospective reserve,

$$V_t = \int_t^n e^{-\int_t^\tau (r_u+\mu_{x+u})\,du}(\mu_{x+\tau}\,b_\tau - \pi_\tau)\,d\tau \;+\; e^{-\int_t^n (r_u+\mu_{x+u})\,du}\,b_n\,,$$

and the equivalence relation

$$V_0 = \pi_0\,, \tag{F.10}$$

where π_0 is the lump sum premium payment collected upon the inception of the policy (it may be 0, of course). The equivalence relation (F.10) can be cast as

$$\pi_0 \;+\; \int_0^t e^{-\int_0^\tau (r_u+\mu_{x+u})\,du}(\pi_\tau - \mu_{x+\tau}\,b_\tau)\,d\tau - e^{-\int_0^t (r_u+\mu_{x+u})\,du}V_t = 0\,. \tag{F.11}$$

The first two terms in (F.11) are the expected present value at time 0 of premiums less benefits up to and including time t. The second term is minus the expected present value at time 0 of benefits less premiums after time t. Thus, the equivalence principle ensures that, at any time, net incomes in the past provide precisely the amount needed to meet net liabilities in the future.

Thiele's differential equation is

$$\frac{d}{dt}\,V_t = \pi_t - \mu_{x+t}\,b_t + (r_t + \mu_{x+t})\,V_t\,, \tag{F.12}$$

with the boundary condition

$$V_{t-} = b_n\,. \tag{F.13}$$

B. With profit contracts (participating policies); Surplus and Bonus. Insurance policies are long term contracts, with time horizons wide enough to capture significant variations in interest and mortality. Therefore, at time 0 when the contract is written with benefits and premiums binding to both parties, the future development of (r_t, μ_{x+t}), $t > 0$, is uncertain, and it is impossible to foresee which premium level will satisfy (F.11) and establish equivalence in the end. If it should turn out that, due to adverse development of interest and mortality, premiums are insufficient to cover benefits, then there is no way the insurance company can avoid a loss; it cannot reduce the benefits and it cannot increase the premiums since these were irrevocably set out in the contract at time 0. The only way the insurance company can prevent such a loss, is to charge a premium 'on the safe side', high enough to be adequate under all likely scenarios. Then, if everything goes well, a surplus will accumulate. This surplus belongs to the insured and is to be repaid as so-called *bonus*, e.g. as increased benefits or reduced premiums.

The usual way of setting premiums to the safe side is to base the calculation of the premium level and the reserves on a provisional *first order basis*, (r_t^*, μ_{x+t}^*), $t > 0$, which represents a worst case scenario and leads to higher premium and reserves than are likely to be needed. We follow common practice and take the first order interest rate to be constant, r^*. (From a mathematical point of view this is just a matter of notation.) The reserve based on the prudent first order assumptions is called the *first order reserve*, and we denote it by V_t^*. It satisfies Thiele's differential equation

$$\frac{d}{dt}\,V_t^* = \pi_t + r^*\,V_t^* - \mu_{x+t}^*(b_t - V_t^*)\,, \tag{F.14}$$

subject to the natural side condition $V_{n-}^* = b_n$. The premiums are determined so as to satisfy the first order equivalence relation $V_0^* = \pi_0$.

Taking our stand at a given time t after the inception of the policy, the development of interest and mortality in the past, $(r_\tau, \mu_{x+\tau})$, $\tau \le t$, is now known and can be invoked in an updated calculation of the net incomes up to time t. The future development of interest and mortality, $(r_\tau, \mu_{x+\tau})$, $\tau > t$, remains uncertain, however, so for the assessment of future liabilities one must stick to the conservative first order basis, $(r^*, \mu^*_{x+\tau})$, $\tau > t$. Thus, instead of the balance equation (F.11), which cannot be set up since it involves unknown future rates of interest and mortality, we have the following expression for the *mean surplus per policy* at time t:

$$S_t = \pi_0 + \int_0^t e^{-\int_0^\tau (r_u + \mu_{x+u})\, du} (\pi_\tau - \mu_{x+\tau}\, b_\tau)\, d\tau - e^{-\int_0^t (r_u+\mu_{x+u})\, du} V^*_t\,. \quad \text{(F.15)}$$

If the factual interest and mortality in the past were more favourable than the pessimistic first order basis, then this surplus is positive.

C. Emergence of surplus. To see how the surplus is emerges, we need to study the dynamics of S_t. Differentiating (F.15), using (F.14), and rearranging terms, we obtain

$$\frac{d}{dt} S_t = e^{-\int_0^t (r_u+\mu_{x+u})\, du}\, c_t,$$

where

$$c_t = (r_t - r^*)\, V^*_t + (\mu^*_{x+t} - \mu_{x+t})(b_t - V^*_t)\,. \quad \text{(F.16)}$$

Obviously, c_t is the rate at which surplus emerges per survivor and per time unit at time t. Interpret the two terms on the right hand side of (F.16) as surplus emerging from safety margins in the interest rate and in the mortality rate, respectively. We can reasonably say that a first order element is set on the safe side if the corresponding contribution to the surplus is positive. By inspection of the first term on the right of (F.16), we see that r^* is on the safe side as long as it is less than the true r_t (provided that the first order reserve is positive, as it should be for any meaningful contract). By inspection of the second term on the right of (F.16), we see that the sign of the sum at risk by death, $b_t - V^*_t$, determines how to set first order mortality to the safe side: If the sum at risk is positive (e.g. term assurance or endowment assurance), then μ^*_{x+t} is on the safe side if it is bigger than μ_{x+t}. If the sum at risk is negative (as is the case for e.g. a pure endowment, a deferred annuity, or some other savings insurance with $b_t = 0$), then μ^*_{x+t} is on the safe side if it is less than μ_{x+t}.

D. Redistribution of surplus as bonus. 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. In the present context of life insurance it denotes the repayments of surplus stemming from favourable development of interest and mortality. Let us denote such repayments by $\tilde{b}$ in general. For the sake of concreteness, suppose bonuses are paid back continuously at rate $\tilde{b}_t$ per survivor for $0 < t < n$ and possibly with a lump sum $\tilde{b}_n$ per survivor at time $t = n$. By statute, surplus is to be repaid in its entirety, which means that equivalence is to be re-established on basis of the true interest and mortality conditions when these are ultimately known at the term of the contract:

$$\begin{aligned} \int_0^n e^{-\int_0^\tau (r_u+\mu_{x+u})\, du} c_\tau\, d\tau \quad &= \quad \int_0^n e^{-\int_0^\tau (r_u+\mu_{x+u})\, du}\, \tilde{b}_\tau\, d\tau \\ &\quad + e^{-\int_0^n (r_u+\mu_{x+u})\, du} \tilde{b}_n\,. \end{aligned} \quad \text{(F.17)}$$

In the following Paragraphs E - G we will study some commonly used bonus schemes.

E. Cash Bonus. This means that surplus is being repaid continually as it emerges, i.e. $\tilde{b}_t = c_t$, $0 < t < n$, and $\tilde{b}_n = 0$. It may for instance take the form of a premium deductible payable at rate c_t as long as the insured is alive during the contract period. In the case of a term assurance contract it could reasonably take the form of an additional payment $\hat{b}_t$ upon death at time $t \in (0, n)$, and a natural choice is $\hat{b}_t = c_t/\mu_{x+t}$.

Exercise 1-12
Verify that the two cash bonus schemes described above comply with the ultimate equivalence requirement (G.36). Construct a scheme that is a combination of the two proposed here.

F. Terminal Bonus. This means that surplus is repaid as a lump sum $\tilde{b}_n$ to survivors at the end of the term, and $\tilde{b}_t = 0$, $0 < t < n$.

Exercise 1-13
Determine $\tilde{b}_n$ by (G.36).

G. Purchase of Additional Insurance. Under this bonus scheme the surplus is spent on purchase of additional insurance. Additional insurance is written on the first order basis and will therefore also generate surplus, which in its turn will be used for further purchase of additional insurance, and so on. The scheme is non-trivial and requires a bit of theoretical reasoning:

For a policy in force at time t let V_t^{*+} denote the expected present value, on the first order basis, of future benefits only;

$$V_t^{*+} = \int_t^n e^{-\int_t^\tau (r^* + \mu^*_{x+u})\, du} \mu_{x+\tau}\, b_\tau\, d\tau \;+\; e^{-\int_t^n (r^* + \mu^*_{x+u})\, du}\, b_n\,.$$

It satisfies the Thiele's differential equation

$$\frac{d}{dt}\, V_t^{*+} = r^*\, V_t^{*+} - \mu^*_{x+t}(b_t - V_t^{*+})\,, \tag{F.18}$$

with natural side condition $V_{n-}^{*+} = b_n$.

The quantity V_t^{*+} is the single premium payable at time t if the insured then were to purchase an additional insurance for the balance of the term, with the same benefits as in the original contract. Spending the surplus $c_t\, dt$ generated in $[t, t+dt)$ as a single premium for additional benefits of the form specified in the original contract, will buy the insured a fraction $q_t\, dt$ of future benefits given by

$$c_t = q_t\, V_t^{*+}\,. \tag{F.19}$$

At any time $t \in (0, n)$ the death benefits from the original contract and from the additional benefits purchased during $(0, t]$ total

$$(1 + Q_t)\, b_t\,, \tag{F.20}$$

where

$$Q_t = \int_0^t q_\tau\, d\tau\,. \tag{F.21}$$

Likewise, the total endowment benefit at the term of the contract is

$$(1 + Q_n)\, b_n\,. \tag{F.22}$$

At time t the total surpluses from the original contract and the additional benefits purchased during $(0, t]$ emerge at rate

$$c_t = (r_t - r^*)\,(V_t^* + Q_t V_t^{*+}) \,+\, (\mu^*_{x+t} - \mu_{x+t})\,\big((1 + Q_t) b_t - V_t^* - Q_t V_t^{*+}\big)\ . \tag{F.23}$$

Now all elements needed are in place, and a dynamic computation will deliver the solution. First, at the time of the inception of the contract, the functions V_t^* and V_t^{*+} and the equivalence premium π_t are determined by use of the program 'prores1.pas' (or 'prores2.pas'). The computation goes backwards starting from the side conditions $V^*_{t-} = b_n$ and $V^{*+}_{t-} = b_n$. Then, as time goes by and surpluses are being observed and disposed of, one computes simultaneously the functions V_t^* and V_t^{*+} (again) and the random function Q_t as solutions to the differential equations (F.14), (F.18), and (rewrite (F.19))

$$\frac{d}{dt} Q_t = \frac{1}{V_t^{*+}} c_t\,, \tag{F.24}$$

with c_t given by (F.23). The computation goes forwards, starting from time $t = 0$ with the initial conditions

$$V_0^* = 0\,,$$
$$V_0^{*+} = V_0^{*+}$$

(picked from the first computation), and

$$Q_0 = 0\,.$$

Having determined Q, the benefits under this bonus scheme are now given by (F.20) and (F.22).

H. Prognostication of bonus. At regular times (typically annually) the customer receives a statement of his policy account, informing about bonus earned from surplus in the past and also predicting future bonuses based on a qualified guess as to the future development of interest and mortality.

Exercise 1-14
Outline such a statement with these pieces of information for the standard contract considered so far, including the relevant formulas, and basing the prognostication of future surplus on the assumption that $r_\tau = r^* + \Delta r$ and $\mu_{x+\tau} = \mu^*_{x+\tau} - \Delta\mu$ for some given positive Δr and $\Delta\mu$.

Exercise 1-15
Apply the present theory to a pension insurance policy for which benefits are an m year deferred life annuity payable at level rate 1 per year in n years, and premiums are payable at level rate during the deferred period. Write out all relations and formulas that differ from the corresponding ones above. Will surplus emerge also in the benefit period $[m, m + n)$?

Exercise 1-16

Extend the theory so as to include expenses. Consider an endowment insurance with constant sum insured b and constant gross premium rate π' (no down payment π'_0 at time 0). Assume that true expenses incur with a lump sum $\alpha' + \alpha'' b$ at time 0 and thereafter continuously at rate $\beta'_t + \beta''_t \pi' + \gamma'_t + \gamma''_t b + \gamma'''_t V_t^{*'}$ at time $t \in (0, n)$ as long as the policy is in force. Here $V_t^{*'}$ denotes the gross premium reserve on first order basis. First order assumptions specify that expenses incur with a lump sum $\alpha^* b$ at time 0 and thereafter continuously at constant rate $\beta^* \pi' + \gamma^* b$. Discuss how the first order elements can be set on the safe side. Observe that there is a lump sum contribution to surplus at time 0.

I. Stochastic interest. The uncertain development of the second order elements can be built into the model by describing the interest rate and the (parameters of the) mortality rate as stochastic processes. To keep things simple, we will focus on interest, which is the more important of the two, and assume that the mortality is perfectly predicted by the first order basis: $\mu^*_{x+t} = \mu_{x+t}$. This means that contributions to surplus stem only from interest gains, so that

$$c_t = (r_t - r^*) V_t^* \,. \tag{F.25}$$

As a simple, but flexible, model for stochastic interest, we will assume that $\{r_t\}_{t\geq 0}$ is generated by the Markov chain model in Section 7.8 of BL, see also Exercise 2. To save space, we will write Y_t and r_t instead of $Y(t)$ and $r(t)$ and, since subscripts are now used for the time variable t, denote the state-wise interest rate by r^e.

The statement of account, which is regularly sent to the insured, usually comes with a prognosis of future bonuses on the insurance. Such a prognosis must be based on a qualified guess about the future development of the factual valuation basis – in our simplified situation about r. This guess may be exogenous to the model, e.g. based on combined opinions of experts in the finance department of the company. Having adopted a stochastic model for r, the insurer can make an endogenous, model-based forecast of future bonus payments. Thus, consider a policy which is still in force at time t, and suppose the insurer wants to inform the insured about the conditional expected value of future bonuses, given that the current interest rate is $r_t = r^e$ (which means $Y_t = e$, assuming that all r^e are different). We will consider a few examples.

J. Cash bonus: The rate at which bonus will be paid at some fixed future time u, provided the insured is then alive, is

$$W = (r_u - r^*) V_u^* \,.$$

At time $t < u$, given $r_t = r^e$, W is predicted by its conditional expected value

$$W_e(t) = \mathbb{E}[W \mid r_t = r^e] \,.$$

Exercise 1-17

Show that the functions $W_e(t)$ are the solution to the differential equations

$$\frac{d}{dt} W_e(t) = \sum_{\ell; f \neq e} \lambda_{ef} (W_e(t) - W_f(t)) \,,$$

subject to the conditions

$$W_e(u) = (r^e - r^*)V_u^* \,,$$

$e = 1, \ldots, J^Y$.

K. Terminal bonus: Bonus payable as a lump sum at the term of the contract n, provided the insured is then alive, is

$$\begin{aligned} W &= \int_0^n e^{\int_\tau^n r_s\,ds}(r_\tau - r^*)V_\tau^*\,d\tau \\ &= W_t' \int_0^t e^{\int_\tau^t r_s\,ds}(r_\tau - r^*)V_\tau^*\,d\tau + W_t'' \,, \end{aligned}$$

where

$$\begin{aligned} W_t' &= e^{\int_t^n r_s\,ds} \,, \\ W_t'' &= \int_t^n e^{\int_\tau^n r_s\,ds}(r_\tau - r^*)V_\tau^*\,d\tau \,. \end{aligned}$$

The random variables W_t' and W_t'', which are unknown at time t, are predicted by

$$\begin{aligned} W_e'(t) &= \mathbb{E}[W_t' \mid r_t = r^e] \,, \\ W_e''(t) &= \mathbb{E}[W_t'' \mid r_t = r^e] \,. \end{aligned}$$

Exercise 1-18
Writing

$$\begin{aligned} W_t' &= e^{r_t\,dt}\,W_{t+dt}' \,, \\ W_t'' &= W_t'\,(r_t - r^*)\,V_t^*\,dt \,+\, W_{t+dt}'' \,, \end{aligned}$$

show that the functions $W_e'(t)$ and $W_e''(t)$ are the solution to the differential equations

$$\begin{aligned} \frac{d}{dt}W_e'(t) &= -r_e\,W_e'(t) + \sum_{f;f\neq e} \lambda_{ef}(W_e'(t) - W_f'(t)) \,, \\ \frac{d}{dt}W_e''(t) &= -W_e'(t)(r^e - r^*)V_t^* + \sum_{f;f\neq e} \lambda_{ef}(W_e''(t) - W_f''(t)) \,, \end{aligned}$$

subject to the conditions

$$\begin{aligned} W_e'(n-) = 1 \,, \\ W_e''(n-) = 0 \,, \end{aligned}$$

$e = 1, \ldots, J^Y$.

L. Additional benefits: At a fixed future time u bonus is paid as a multiple Q_u of the contractual benefits provided the insured is then alive. At time t we decompose Q_u into Q_t, which is known, and $Q_u - Q_t$, which is unknown, and we need to predict the latter. Recalling (F.19) and (F.25), start from the differential equation

$$\frac{d}{dt}Q_t = (r_t - r^*)\,(\frac{V_t^*}{V_t^{*+}} + Q_t)$$

and use the technique with integrating factor to obtain

$$Q_u = W_t' \, Q_t + W_t'' ,$$

where

$$\begin{aligned} W_t' &= e^{\int_t^u (r_s - r^*)\, ds} , \\ W_t'' &= \int_t^u e^{\int_\tau^u (r_s - r^*)\, ds} (r_\tau - r^*) \frac{V_\tau^*}{V_\tau^{*+}} \, d\tau . \end{aligned}$$

Exercise 1-19
Derive differential equations for the state-wise predictions $W_e'(t)$ and $W_e''(t)$ of W_t' and W_t''.

Exercise 1-20
(a) Predict discounted future cash bonuses given survival to n,

$$\int_t^n e^{-\int_t^\tau r_s\, ds} (r_\tau - r^*) V_\tau^* \, d\tau .$$

(b) Predict discounted future cash bonuses,

$$\int_t^n e^{-\int_t^\tau (r_s + \mu_{x+s})\, ds} (r_\tau - r^*) V_\tau^* \, d\tau .$$

(c) Find differential equations for the conditional variance, given $r_t = r^e$, of the future cash bonuses. You may try your hand also on the conditional variance of future terminal bonus.

As we have said before, the Markov model proposed here can hardly be 100 per cent realistic. Now, the usefulness of a model depends on its purpose. The sole purpose of the interest rate model is to provide the insured with a reasonable guess as to his future prospects of bonus, and for that purpose a rough model can certainly be adequate. Anyway, at the end of the day the bonus payments will be determined entirely by the factual interest rate and will not depend on the assumptions in our model.

M. Guaranteed interest. Recall the basic rules of the 'with profit' insurance contract: On the one hand, any surplus is to be redistributed to the insured. On the other hand, benefits and premiums set out in the contract cannot be altered to the insured's disadvantage. This means that negative surplus, should it occur, cannot result in negative bonus. Thus, the with profit policy comes with an interest rate guarantee to the effect that bonus is to be paid as if factual interest were no less than first order interest, roughly speaking. For instance, cash bonus is to be paid at rate

$$(r_t - r^*)_+ V_t^*$$

per survivor at time t, hence the insurer has to cover

$$(r^* - r_t)_+ V_t^* . \tag{F.26}$$

Similarly, terminal bonus (typical for e.g. a pure endowment benefit) is to be paid as a lump sum

$$\left(\int_0^n e^{\int_\tau^n r_s \, ds} (r_\tau - r^*) V_\tau^* \, d\tau \right)_+$$

per survivor at time n, hence the insurer has to cover

$$\left(\int_0^n e^{\int_\tau^n r_s \, ds} (r^* - r_\tau) V_\tau^* \, d\tau \right)_+ . \tag{F.27}$$

(We write $a_+ = \max(a, 0) = a \vee 0$.)

An interest guarantee of this kind represents a liability on the part of insurer. It cannot be offered for free, of course, but has to be compensated by a premium. This can certainly be done without violating the rules of game for the participating policy, which lay down that premiums and benefits be set out in the contract at time 0. Thus, for simplicity, suppose a single premium is to be collected at time 0 for the guarantee. The question is, how much should it be?

Being brought up with the principle of equivalence, we might think that the expected discounted value of the liability is an agreeable candidate for the premium. However, the rationale of the principle of equivalence, which was to make premiums and benefits balance on the average in an infinitely large portfolio, does not apply to financial risk. Interest rate variations cannot be eliminated by increasing the size of the portfolio; all policy-holders are faring together in one and the same boat on their once-in-a-lifetime voyage through the troubled waters of their chapter of economic history. This risk cannot be averaged out in the same way as the risk associated with the lengths of the individual lives. None the less, in lack of anything better, let us find the expected discounted value of the interest guarantee, and just anticipate here that this actually would be the correct premium in an extended model specifying a so-called complete financial market. Those who are familiar with basic arbitrage theory know what this means. Those who are not should just imagine that, in addition to the bank account with the interest rate r_t, there are some other investment opportunities, and that any future financial claim can be duplicated perfectly by investing a certain amount at time 0 and thereafter just selling and buying available assets without any further infusion of capital. The initial amount required to perform this duplicating investment strategy is, quite naturally, the price of the claim. It turns out that this price is precisely the expected discounted value of the claim, only under a different probability measure than the one we have specified in our physical model. With these reassuring phrases, let us proceed to find the expected discounted value of the interest guarantee.

(a) Cash bonus with gurantee given by (F.26): Given that $r_0 = r^e$ (say), the price of the total claims under the guarantee, averaged over an infinitely large portfolio, is

$$\mathbb{E}\left[\int_0^n e^{-\int_0^\tau r} (r^* - r_\tau)_+ V_\tau^* \, {}_\tau p_x \, d\tau \,\middle|\, r_0 = r^e \right] . \tag{F.28}$$

A natural starting point for creating some useful differential equations by the backward construction is the 'price of future claims under the guarantee' in state e at time t,

$$W_e(t) = \mathbb{E}\left[\int_t^n e^{-\int_t^\tau r} (r^* - r_\tau)_+ V_\tau^* \, {}_\tau p_x \, d\tau \,\middle|\, r_t = r^e \right] , \tag{F.29}$$

$e = 1, \ldots, J^Y$, $0 \le t \le n$. The price in (F.28) is precisely $W_e(0)$.

Conditioning on what happens in the time interval $(t, t+dt]$ and neglecting terms of order $o(dt)$ that will disappear in the end anyway, we find

$$W_e(t) = (1 - \lambda_{e\cdot}\, dt)\Big((r^* - r^e)_+ V_t^*\, {}_tp_x\, dt + e^{-r^e dt} W_e(t+dt)\Big) + \sum_{f;\, f\neq e} \lambda_{ef}\, dt\, W_f(t)\,.$$

From here we easily arrive at the differential equations

$$\frac{d}{dt} W_e(t) = -(r^* - r^e)_+ V_t^*\, {}_tp_x + r^e W_e(t) - \sum_{f;\, f\neq e} \lambda_{ef}\, dt\, (W_f(t) - W_e(t))\,, \qquad \text{(F.30)}$$

which are to be solved subject to the conditions

$$W_e(n-) = 0\,. \qquad \text{(F.31)}$$

(b) Terminal bonus at time n given by (F.27): Given $r_0 = r^e$, the price of the claim under the guarantee, averaged over an infinitely large portfolio, is

$$\begin{aligned} &\mathbb{E}\left[e^{-\int_0^n r_s\, ds} \left(\int_0^n e^{\int_\tau^n r_s\, ds} (r^* - r_\tau) V_\tau^*\, d\tau \right)_+ {}_np_x \,\middle|\, r_0 = r^e \right] \\ &= \mathbb{E}\left[\left(\int_0^n e^{-\int_0^\tau r} (r^* - r_\tau) V_\tau^*\, d\tau \right)_+ \,\middle|\, r_0 = r^e \right] {}_np_x\,. \end{aligned} \qquad \text{(F.32)}$$

Let us try and copy the method of Item (a) and look at the 'price of the claim at time t', which should be the conditional expected discounted value of the claim, given what we know at the time:

$$\begin{aligned} &\mathbb{E}\left[e^{-\int_t^n r_s\, ds} \left(\int_0^n e^{\int_\tau^n r_s\, ds} (r^* - r_\tau) V_\tau^*\, d\tau \right)_+ {}_np_x \,\middle|\, r_\tau;\, 0 \le \tau \le t \right] \\ &= \mathbb{E}\left[\left(U_t + \int_t^n e^{-\int_t^\tau r} (r^* - r_\tau) V_\tau^*\, d\tau \right)_+ \,\middle|\, r_\tau;\, 0 \le \tau \le t \right] {}_np_x\,. \end{aligned} \qquad \text{(F.33)}$$

where

$$U_t = \int_0^t e^{\int_\tau^t r} (r^* - r_\tau) V_\tau^*\, d\tau\,.$$

The quantity in (F.33) is more involved than the one in (F.29) since it depends effectively on the past history of interest rate through U_t. We can, therefore, not hope to end up with the same simple type of problem as in Item (a) above and in all other situations encountered so far, where we essentially had to determine the conditional expected value of some function depending only on the future course of the interest rate. Which was easy since, by the Markov property, we could look at state-wise conditional expected values $W_e(t)$, $e = 1, \ldots, J^Y$, say. These are deterministic functions of the time t only and can be determined by solving ordinary differential equations.

Let us proceed and see what happens. Due to the Markov property (conditional independence between past and future, given the present) the expression in (F.33) is a function of t, r_t and U_t. Dropping the uninteresting factor ${}_np_x$, consider its value for given $U_t = u$ and $r_t = r_e$,

$$W_e(t,u) = \mathbb{E}\left[\left(u + \int_t^n e^{-\int_t^\tau r} (r^* - r_\tau) V_\tau^*\, d\tau \right)_+ \,\middle|\, r_t = r^e \right]\,.$$

Use the backward construction:

$$
\begin{aligned}
&W_e(t,u) = \\
&(1-\lambda_{e\cdot}dt)\mathbb{E}\left[\left(u+(r^*-r^e)V_t^*\,dt+e^{-r^e\,dt}\int_{t+dt}^{n} e^{-\int_{t+dt}^{\tau} r}(r^*-r_\tau)V_\tau^*\,d\tau\right)_+ \middle|\, r_{t+dt}=r^e\right] \\
&+\sum_{f;\,f\neq e}\lambda_{ef}\,dt\,W_f(t,u) = \\
&(1-\lambda_{e\cdot}dt)e^{-r^e\,dt}W_e(t+dt,e^{r^e\,dt}u+(r^*-r^e)V_t^*\,dt)+\sum_{f;\,f\neq e}\lambda_{ef}\,dt\,W_f(t,u)\,.
\end{aligned}
$$

Insert here $e^{\pm r^e\,dt}=1\pm r^e dt+o(dt)$,

$$
\begin{aligned}
&W_e(t+dt,e^{r^e\,dt}u+(r^*-r^e)V_t^*\,dt) = \\
&W_e(t,u)+\frac{\partial}{\partial t}W_e(t,u)\,dt+\frac{\partial}{\partial u}W_e(t,u)(u\,r^e+(r^*-r^e)V_t^*)\,dt+o(dt)\,,
\end{aligned}
$$

and proceed in the usual manner to arrive at the partial differential equations

$$
\frac{\partial}{\partial t}W_e(t,u)+(ur^e+(r^*-r^e)V_t^*)\frac{\partial}{\partial u}W_e(t,u)-r^eW_e(t,u)+\sum_{f;\,f\neq j}\lambda_{ef}\,(W_f(t,u)-W_e(t,u))=0\,.
$$

These are to be solved subject to the conditions

$$
W_e(n-,u)=u_+\,,
$$

$e=1,\ldots,J^Y$.

Since the functions we are interested in involved both t and U_t, we are lead to state-wise functions in two arguments and, therefore, quite naturally end up with partial differential equations for those.

N. Unit linked insurance. We have been discussing the participating (or with profit) policy, characteristic of which is that benefits and premiums are set out in nominal amounts in the contract at time 0. Thus, For the fairly general contract described in the introduction to this note, the functions b_t and π_t would be deterministic, not dependent on the development of the interest rate over the term of the contract. Introduce

$$
U_t=e^{\int_0^t r_u\,du}\,,
$$

which is the value at time t of a unit deposited in the investment portfolio at time 0. We may call it the price index of the investment portfolio. Recast the equivalence relation (F.11) as

$$
-\pi_0\;+\;\int_0^n U_\tau^{-1}\,{}_\tau p_x(\mu_{x+\tau}\,b_\tau-\pi_\tau)\,d\tau\;+\;U_n^{-1}\,{}_n p_x\,b_n=0\,. \tag{F.34}
$$

With b_t and π_t fixed at time 0 there is no way one can make them fulfill (F.34) for all possible future courses of the interest rate process. Depending on the economic development there will be inequality in the one or the other direction. The financial risk thus introduced is hedged (hopefully perfectly) by setting premiums on a prudent first order basis, i.e. replacing the unknown r_t in (F.34) by some r^* set to the 'safe side'.

An alternative scheme for management of financial risk in life insurance is known as *unit linked insurance* (also called *variable life insurance*). The idea of this concept is to link benefits and premiums to the performance of the investment portfolio, that is, let contractual payments be inflated by the index U instead of being fixed nominal amounts.

Under a *perfect unit linked* contract we would have $b_t = U_t b_t^\circ$ and $\pi_t = U_t\,\pi_t^\circ$ for some 'baseline' benefits b_t° and premiums π_t°, $t \in [0, n]$, determined at time 0. Inserting this into (F.34) gives

$$-\pi_0 \,+\, \int_0^n {}_\tau p_x(\mu_{x+\tau}\, b_\tau^\circ - \pi_\tau^\circ)\, d\tau \,+\, {}_n p_x\, b_n^\circ = 0\,. \tag{F.35}$$

We see that, for a given baseline benefit function b_t°, the equivalence relation can be fulfilled by a suitable choice of baseline premium rate π_t°. The future course of the interest rate process has disappeared from the relation upon discounting the indexed payments and, thus, the problem with financial risk has been resolved by the perfect unit linked device.

However, in practice unit-linked contracts are usually not perfect in the sense described above. Typically, only the benefits are linked to the investment index, whereas premiums are not. Furthermore, the unit linked contract is typically equipped with a guarantee specifying that the benefit cannot fall below a certain pre-specified nominal minimum. Such modifications to the perfect linking re-introduce financial risk, of course.

Before we return to mathematics, we dare to suggest that guarantees, whether they apply to benefits under unit linked contracts or to interest under with profit contracts, are remains of the social security concern that traditionally was paramount in life insurance. They introduce a discrimination between various forms of saving; unlike those who invest in stocks, bonds, or real estate, those who invest in life or pension insurance are granted the privilege of gaining from booms without loosing from recessions. However, parity can be restored by letting the insured pay for the guarantee. Thus we proceed to determine its right price.

For an example, let us try and determine the single premium payable at time 0 for a term insurance with sum $b_t = (U_t \vee g)$ at time $t \in (0, n)$, where g is the guaranteed minimum sum insured specified at time 0. The premium is

$$\pi = \mathbb{E}\left[\int_0^n e^{-\int_0^\tau r}\left(e^{\int_0^\tau r} \vee g\right) {}_\tau p_x \mu_{x+\tau}\, d\tau\right] = \mathbb{E}\left[\int_0^n \left(1 \vee e^{-\int_0^\tau r} g\right) f_\tau\, d\tau\right]\,,$$

where we have abbreviated

$$f_t = {}_t p_x \mu_{x+t}\,.$$

Following the recipe in Item (b) of Paragraph M, consider the 'price of future claims at time t',

$$\begin{aligned}&\mathbb{E}\left[\int_t^n e^{-\int_t^\tau r}\left(e^{\int_0^\tau r} \vee g\right) f_\tau\, d\tau \,\middle|\, r_\tau;\, 0 \le \tau \le t\right]\\ &= \mathbb{E}\left[\int_t^n \left(U_t \vee e^{-\int_t^\tau r} g\right) f_\tau\, d\tau \,\middle|\, r_\tau;\, 0 \le \tau \le t\right]\,.\end{aligned} \tag{F.36}$$

Arguing as before, the expression in (F.36) is a function of t, r_t and U_t. Consider its value at time t for given $U_t = u$, and $r_t = r^e$,

$$W_e(t, u) = \mathbb{E}\left[\int_t^n \left(u \vee e^{-\int_t^\tau r} g\right) f_\tau\, d\tau \,\middle|\, r_t = r^e\right]\,.$$

When $r_0 = r^e$, the premium we seek is $W_e(t, 1)$

Now use the backward construction, this time leaving details aside:

$$\begin{aligned}
&W_e(t,u) = \\
&(1-\lambda_{e\cdot}dt)\mathbb{E}\left[(u\vee g)\, f_t\, dt + e^{-r^e dt}\int_{t+dt}^{n}\left(e^{r^e dt}u \vee e^{-\int_{t+dt}^{\tau} r}g\right) f_\tau\, d\tau\,\middle|\, r_{t+dt} = r^e\right] \\
&+ \sum_{f;\, f\neq e} \lambda_{ef}\, dt\, W_f(t,u) \\
&= (1-\lambda_{e\cdot}dt)\left((u\vee g)\, f_t\, dt + e^{-r^e dt} W_e(t+dt, e^{r^e dt}u)\right) + \sum_{f;\, f\neq e}\lambda_{ef} dt\, W_f(t,u)\,.
\end{aligned}$$

Insert $e^{\pm r^e\, dt} = 1 \pm r^e dt + o(dt)$ and

$$\begin{aligned}
W_e(t+dt, e^{r^e dt}u) &= W_e(t+dt, u + ur^e dt) + o(dt) \\
&= W_e(t,u) + \frac{\partial}{\partial t}W_e(t,u)\, dt + \frac{\partial}{\partial u}W_e(t,u)\, u\, r^e\, dt + o(dt)\,,
\end{aligned}$$

and fill in some details to arrive at the partial differential equations

$$(u\vee g) f_t - r^e W_e(t,u) + \frac{\partial}{\partial t}W_e(t,u) + \frac{\partial}{\partial u}W_e(t,u)\, u\, r_e + \sum_{f;\, f\neq e}\lambda_{ef}\left(W_f(t,u) - W_e(t,u)\right) = 0\,.$$

These are to be solved subject to the conditions

$$W_e(n-, u) = 0\,,$$

$e = 1, \ldots, J^Y$.

O. Salary dependent premiums and benefits. The employees of a firm are enrolled in a pension scheme with salary dependent premiums and benefits. Consider an employee (x), who enters the scheme x years old at time 0 (say), retires at pensionable age 65 at time $m = 65 - x$, earns salary at rate $S(t)$ per time unit at any time $t < m$, and will receive pension continuously at level rate Q (yet to be determined) for n years after retirement. Let us first work under the assumption that the interest rate r is constant and known for the entire term of the contract up to time $m + n$.

(a) We will first consider a 'defined contributions' scheme under which a fixed proportion of the salary is used as premium for additional pension benefits. It will turn out that equivalence is automatically attained regardless of the development of the salary.

In any small time interval $[t, t+dt)$, $t < m$, the insured earns $S(t)\, dt$. A fixed proportion $\pi S(t)\, dt$, $0 < \pi < 1$, of this salary is used as a single premium for a pension of $q(t)\, dt$ per time unit in the time interval $[m, m+n]$. By the principle of equivalence, q_t is given by

$$\pi\, S(t)\, dt = q(t)\, dt\; {}_{m-t|n}\bar{a}_{x+t}\,,$$

that is,

$$q(t) = \pi\,\frac{S(t)}{{}_{m-t|n}\bar{a}_{x+t}} = \pi\,\frac{S(t)}{{}_{m-t}E_{x+t}\;\bar{a}_{x+m\,\overline{n|}}}\,.$$

The total rate of pension per time unit purchased by a survivor at time m is

$$Q = \int_0^m q(\tau)\, d\tau = \pi\int_0^m \frac{S(\tau)}{{}_{m-\tau}E_{x+\tau}\;\bar{a}_{x+m\,\overline{n|}}}\, d\tau\,.$$

It should be fairly obvious that the equivalence requirement is fulfilled by this scheme since, no matter how much or little salary the insured will earn and no matter if it can be predicted or not at the outset, the benefits are entirely determined by the salary-dependent contributions. Let us, however, just check: For any given salary function, S, the expected discounted premiums are

$$\pi \int_0^m e^{-\int_0^\tau (r+\mu_{x+s})\,ds}\, S(\tau)\, d\tau = \pi \int_0^m {}_\tau E_x\, S(\tau)\, d\tau\,,$$

and the expected discounted benefits are

$$Q\; {}_{m|n}\bar{a}_x = \pi \int_0^m \frac{S(\tau)}{{}_{m-\tau}E_{x+\tau}\; \bar{a}_{x+m\,\overline{n}|}}\, d\tau\; {}_mE_x\; \bar{a}_{x+m\overline{n}|} = \pi \int_0^m {}_\tau E_x\, S(\tau)\, d\tau\,,$$

where we have used the well-known identity ${}_mE_x = {}_\tau E_x\; {}_{m-\tau}E_{x+\tau}$.

(b) Suppose now that, instead of letting the contributions determine the benefits, the benefits are linked to the salary whereas the premiums are not. More specifically, suppose pension is payable continuously at rate $0.75S(m)$ (i.e. 75% of the salary rate at the time of retirement) for n years after retirement, and that premium is payable at a prefixed level rate π while active (both contingent on survival, of course). To determine the premium level π at time 0, we now need to make assumptions about the future development of the salary. Let us also abandon the unrealistic assumption that the future development of the interest rate is known:

Assume that the economy is governed by a continuous time Markov chain $Y(t)$, $t \geq 0$, with state space $\mathcal{J} = \{1, \ldots, J\}$, constant intensities of transition λ_{jk}, $j \neq k$, and initial state $Y(0) = i$, say. At any time $t \geq 0$ the accumulation factor $U(t)$ of the investment portfolio is given by

$$U(t) = \exp\left(\int_0^t r(s)\, ds\right)\,, \qquad r(s) = \sum_j I_j(s) r_j\,, \tag{F.37}$$

and the salary rate is given by

$$S(t) = \exp\left(\int_0^t a(s)\, ds\right)\,, \qquad a(s) = \sum_j I_j(s) a_j\,.$$

Here, $I_j(t) = 1[Y(t) = j]$, and the r_j and the a_j are known, fixed numbers; r_j is the interest rate and a_j is the rate at which salary increases per time unit and per unit of salary when the economy is in state j.

Exercise 1-21

Determine the premium rate π that makes expected discounted benefits less premiums equal to 0 at time 0; Construct differential equations for benefits and for premiums and specify appropriate side conditions.

Exercise 1-22

Consider the Markov chain interest model outlined in Figure F.5. Fill in appropriate statements in the program 'prores2' to make it compute the state-wise expected discount factors $\mathbb{E}\left[\exp\left(-\int_0^5 r(s)\,ds\right)\Big|\,Y(0)=j\right]$, $j=1,2,3$.

Exercise 1-22

Let the benefit be an n years unit-linked pure life endowment with guaranteed sum insured $\max(U(n),g)$, and let premium be payable continuously at level rate π during the insurance period. Show how to determine π.

Exercise 56

Let $\{N(t)\}_{t\geq 0}$ be Poisson process with intensity λ. This is a counting process of even simpler type than the counting processes associated with a Markov chain; N is not only Markov, but also has independent increments. Thus, in any small time interval $[t,t+dt)$ the process N makes a jump of 1 with probability $\lambda\,dt$ regardless of the past history of the process in $[0,t)$.

Let the price $S(t)$ of a share of stock at time t be modelled as a so-called geometric Poisson process with drift,

$$S(t)=\exp\left(\alpha t+\beta N(t)\right)\,,$$

$t\geq 0$. If $\beta=0$, then $S(t)$ is just the accumulation factor for a bank account with fixed interest rate. The Poisson term in the exponent adds jumps at random times, and a jump at time t makes the stock price jump from $S(t-)$ to $S(t)=S(t-)\,e^{\beta}$. Thus, $\gamma=e^{\beta}-1$ is the relative change $(S(t)-S(t-))/S(t-)$ in the stock price at the jump time. Between the jumps the stock price increases at fixed "rate of interest" α.

(a) Find the expected value $\mathbb{E}[S(t)]$ at time 0 of the stock price at time t, and do this in two ways: First, work directly with the Poisson distribution of $N(t)$ and, second, solve a differential equation obtained by the direct backward construction (condition on "what happens in the small time interval $[0,dt)$"). Explain that, having determined the expected value, higher order moments are easily obtained.

(b) Find the dynamics $dS(t)$ of the stock price by applying the change of variable rule, see Appendix A.

(c) Using the direct backward construction, show that the expected present value of a perpetuity (an everlasting annuity), is

$$\mathbb{E}\left[\int_0^\infty S^{-1}(\tau)\,d\tau\right]=(\alpha+\lambda(1-e^{-\beta}))^{-1}\,.$$

(d) Let $\{N_1(t)\}_{t\geq 0}$ and $\{N_2(t)\}_{t\geq 0}$ be independent Poisson processes with intensities λ_1 and λ_2, respectively. Let the price $S(t)$ of a share of stock at time t be

$$S(t)=\exp\left(\alpha t+\beta_1 N_1(t)+\beta_2 N_2(t)\right)\,,$$

$t\geq 0$. Letting β_1 and β_2 have opposite signs, we have created a stock price which may increase or decrease with instantaneous jumps. Write out the straightforward analogs of the formulas in Items (a) - (c) for this more general model.

Exercise 57
(a) Let A and A' be two payment functions with retrospective reserves denoted by U and U', respectively. Assuming that the interest rate is always positive, verify the following rather obvious assertion: If $A_t \leq A'_t$ for all t, then $U_t \leq U'_t$ for all t. (In particular, any advancement of deposits will produce an increase of the retrospective reserve if the interest rate is positive. This is the general circumstance underlying results like the following about ordering of expected present values: $a_{\overline{n|}} \leq \bar{a}_{\overline{n|}} \leq \ddot{a}_{\overline{n|}}$.)

(b) Let the payment stream A represent deposits less withdrawals on an n years savings account that bears interest with strictly positive interest rate. It is required that $U_t \geq 0$ for all t, with strict inequality for some t, and that $U_n = 0$. Prove that $A_n < 0$, and explain this result.

Exercise 58

(a) Show that, for $t \leq s \leq u$, $p_{\overline{jj}}(t,u) = p_{\overline{jj}}(t,s)\, p_{\overline{jj}}(s,u)$, which is obvious.

(b) Given start in state a at time 0, write up the probability that the process remains in a during the time interval $[0, t_1)$, then jumps to state i in $[t_1, t_1 + dt_1)$, then remains in i during the time interval $[t_1 + dt_1, t_2)$, then jumps to state a in $[t_2, t_2 + dt_2)$, then remains in a during the time interval $[t_2 + dt_2, t_3)$, and finally jumps to state d in the time interval $[t_3, t_3 + dt_3)$. This is the probability of one particular full specification of the history of the process.

Exercise 59

Read Sections 9.1 and 9.2 in 'Basic Life Insurance Mathematics'. Suppose n independent lives, which follow the same Gompertz-Makeham mortality law with intensity $\mu(t) = \alpha + \beta \exp(\gamma t)$, are observed from birth until death. Find the equations for the Maximum likelihood estimators for the parameters α, β, and γ, and find the asymptotic properties of the estimators.

Exercise 60

In the situation of Paragraphs 9.1.B-E, consider the problem of estimating μ from the D_i alone, the interpretation being that it is only observed whether survival to z takes place or not. Show that the likelihood based on $D_i,\ i = 1, \ldots, n$, is

$$q^N (1-q)^{n-N} ,$$

with $q = 1 - e^{-\mu z}$, the probability of death before z. (Trivial: it is a binomial situation.)

Note that N is now sufficient, and that the class of distributions is a regular exponential class. The MLE of q is

$$q^* = \frac{N}{n}$$

with the first two moments

$$Eq^* = q, \;\; \mathrm{Var}q^* = \frac{q(1-q)}{n} \,.$$

It is UMVUE in the class of estimators based on the D_i.

The MLE of $\mu = -\ln(1-q)/z$ is $\mu^* = -\ln(1-q^*)/z$. Apply standard asymptotic results about MLE to show that

$$\mu^* \sim_{\mathrm{as}} \mathrm{N}\left(\mu, \frac{q}{nz^2(1-q)}\right) \,.$$

The asymptotic efficiency of $\hat{\mu}$ relative to μ^* is

$$\frac{\mathrm{asVar}\mu^*}{\mathrm{asVar}\hat{\mu}} = \left(\frac{e^{\frac{\mu z}{2}} - e^{-\frac{\mu z}{2}}}{\mu z}\right)^2 = \left(\frac{\sinh(\mu z/2)}{\mu z/2}\right)^2$$

(sinh is the hyperbolic sine function defined by $\sinh(x) = (e^x - e^{-x})/2)$. This function measures the loss of information suffered by observing only death/survival by age z as compared to inference based on complete observation throughout the time interval $(0, z)$. It is ≥ 1 and increases from 1 to ∞ as μz increases from 0 to ∞. Thus, for small μz, the number of deaths is all that matters, whereas for large μz, the life lengths are all that matters. Reflect over these findings.

Exercise 61

We refer to the disability model.

(a) Consider an x years old insured who enters an insurance scheme at time 0. The probability $p_{\overline{aa}}(0,t) = \exp(-\int_0^t (\mu_{x+s} + \sigma_{x+s})\,ds)$ can be viewed as the probability $p_{aa}^{(0)}(0,t)$ of being active at time t after having been disabled 0 times. Derive forward differential equations for the probability $p_{ai}^{(1)}(0,t)$ of being disabled for the first time at time t and for the probability $p_{aa}^{(1)}(0,t)$ of being active at time t after having been disabled once.

(b) Find the probability of being disabled for the first time at time t and that the disability has lasted for at least q years.

(c) At time 0 an active person aged x buys a disability pension insurance with the following terms: The benefit is a pension payable at level rate 1 during the first disability, but only after it has lasted for at least q years (the *qualifying period*). Premium is payable at level rate π as long as the insured is active and has not yet been disabled, but not after time $n - q$, where n is the contract period ($n > q$). Determine the premium π by the principle of equivalence, assuming that the interest rate r is constant. Find the reserve at time $t < n - q$ for an insured who is disabled for the first time and is currently receiving the disability benefit.

Exercise 62

The two-state Markov chain Z sketched in Fig. F.1 can be given many interpretations; it could describe transitions of a person into and out of the work-force ('active' means employed, 'invalid' means unemployed, and mortality is disregarded), or the transitions of a person between marital states ('active' means single, 'invalid' means married), or the transitions of a machine or mechanical device between states of functioning ('active' means intact, 'invalid' means out of order), and so on.

Let us take the model as a description of a lamp, which is always switched on (it is installed in a submarine), and which is 'active' when the bulb is intact and 'invalid' or 'inactive' when the bulb is burnt-out. One can assume that the life-time of a bulb is exponentially distributed, and it also seems reasonable to assume that the lapse of time from a bulb burns out until the failure is discovered and the bulb is replaced, is exponentially distributed. Then the intensities σ and ρ are constant (σ is the 'mortality' intensity of the life length of a bulb or the expected number of 'deaths' per time unit for an burning bulb, and ρ is the expected number of maintenance inspections per time unit.).

We follow the lamp from time 0 when it is active. Let $N_{ai}(t)$ and $N_{ia}(t)$ denote the number of failures and replacements of bulb, respectively, in the time interval $[0, t]$, and let $I_a(t)$ and $I_i(t)$ be the indicators of the events 'active at time t' and 'inactive at time t', respectively. Of course, $I_i(t)+I_a(t) = 1$, and $N_{ia}(t)-N_{ai}(t)$ is either 0 or 1.

(a) Find $\mathbb{E}[I_i(t)] = p_{ai}(0, t)$ by solving Kolmogorov's forward equations, using $p_{aa}(0, t) = 1 - p_{ai}(0, t)$.

(b) Find the expected time as inactive in $[0, t]$, $\mathbb{E}[\int_0^t I_i(\tau)\, d\tau]$. You may just integrate the function $p_{ai}(0, \tau)$ in (a). Explain why.

(c) Find $\mathbb{E}[N_{ai}(t)]$. You may also find the answer by just integrating the function $p_{aa}(0, \tau)$ in (a) and multiplying with σ. Explain why.

(d) Divide the expected values in (b) and (c) by t and find the limit as $t \to \infty$. Discuss the expressions as functions of σ and ρ.

(e) Find the failure intensity $\tilde{\sigma}(t)$, $0 < t < n$, for the Markov chain Z, conditional on $Z(n) = i$.

(f) If $\sigma = \rho$, then $N_{ai}(t) + N_{ia}(t)$ is a Poisson process with intensity σ. Explain why. What is then $N_{ai}(t)$ in terms of the Poisson process?

Exercise 63

The situation is as described in Exercise 62. Suppose data are available for m independent lamps observed over the same time interval $[0, n]$, all active at time 0.

(a) Assume complete histories have been recorded for each lamp. Find the MLE for σ and find its asymptotic distribution.

Exercise 64

(b) The parameter ρ is subject to control since it is the frequency with which the lamps are being checked by maintenance personnel. Show that the asymptotic variance of the MLE in (a) is a decreasing function of ρ.

(c) Suppose that, at time n, data are available only for lamps that are inactive at the time. Then it is the conditional process in 7.8e which is the relevant stochastic model for the individual histories. Find the MLE in this situation and find also its asymptotic variance.

Exercise 65

Go through Paragraphs 9.1 A-E and G in 'BL' and add all details in proofs.

Exercise 66

At your disposal are data from m independent lives insured under one and the same scheme. You know for each individual the time of entry into the scheme, the age upon entry, the length of the observation period, whether he/she died during the observation period, and - if died - the age at death.

Find the ML estimators under the assumption of piece-wise constant intensities and explain how they can be fitted by analytic graduation by a Gompertz Makeham function.

Exercise 67

Consider the continuous time version of an n-year endowment insurance of 1 (say) against level premium π during the insurance period. Assume that the interest rate r is constant. Prove that the premium rate π is a decreasing function of n and that the premium reserve V_t for fixed t is a decreasing function of n for $n > t$. Note that the results are valid for all specifications of interest rate and mortality rate.

Exercise 68

We refer to Basic Life Insurance Mathematics (BL), Section 7.5, Paragraph J, Widow's pension. The situation is depicted in Figure F.4. At time $t = 0$ (say) the husband (x) is x years old and the wife (y) is y years old. Their remaining life lengths after time 0 are denoted by S and T, respectively. The joint distribution of S and T is determined by the mortality rates named in the figure: for instance, at any time $t \geq 0$, (x) has mortality rate $\mu(t)$ if (y) is still alive and $\mu'(t)$ if (y) is dead.

(a) Find integral expressions for the marginal survival probabilities $\mathbb{P}[S > s]$, $\mathbb{P}[T > t]$, and for the joint survival probability $\mathbb{P}[S > s\,,\, T > t]$.

(b) Prove that, if the mortality rates are independent of marital status, $\mu = \mu'$ and $\nu = \nu'$, then S and T are stochastically independent:

$$\mathbb{P}[S > s\,,\, T > t] = \mathbb{P}[S > s]\,\mathbb{P}[T > t]\,.$$

Adopt the independence assumption in (b) and assume that $x = y = 30$ and that (x) and (y) are subject to the same law of mortality, G82M, irrespective of marital status. Thus, at any time $t \geq 0$,

$$\mu(t) = \nu(t) = \mu'(t) = \nu'(t) = 0.0005 + 0.000075858 \times 1.09144^{30+t}\,. \tag{F.38}$$

The couple buys a combined life insurance and widow's pension policy specifying that a pension is to be paid with intensity $b = 0.5$ as long as the wife is alive and husband is dead, a life assurance with sum $s = 1$ is due immediately upon the death of the husband if the wife is already dead (a benefit to their dependents), and the equivalence premium is to be paid with level intensity c as long as both husband and wife are alive. The policy terminates at time $n = 30$. Interest is earned on the reserve at constant rate $r = \ln(1.045)$.

(c) Using the direct backward argument, derive the differential equation for the non-central moments of first and second order of the discounted future benefits in state 0. (The first order moment is just the reserve.)

(d) Using the program 'prores1.pas', compute the three first moments of discounted benefits less premiums at times 0,5,...,30. Do the same for a modified contract, by which 50% of the current reserve (in state 0) is to be paid back immediately to the husband in case the wife dies before him during contract period.

(e) Let us now drop the independence assumption in (b) and instead assume that, for each party, the mortality rate increases upon the death of the spouse (a 'grief effect'):

$$\mu' > \mu \,, \qquad \nu' > \nu \,. \tag{F.39}$$

In a forthcoming note 'Dependent lives' it will be proved that, under the assumption (F.39), S and T are positively dependent in the sense that

$$\mathbb{P}[S > s \,,\, T > t] > \mathbb{P}[S > s]\, \mathbb{P}[T > t] \,.$$

Give some numerical evidence in support of this result by e.g. computing $\mathbb{P}[S > 30 \,,\, T > 30]$, $\mathbb{P}[S > 30]$, and $\mathbb{P}[T > 30]$ in the case where μ and ν are as in (F.38) and $\mu'(t) = \mu(t) + 0.0005$ and $\nu'(t) = \nu(t) + 0.0005$.

(f) Under the assumptions of Item (e) find the covariance between the present values at time 0 of a term insurance of 1 in 30 years on (x) and a similar insurance on (y). (The relationship $\mathbb{C}\mathrm{ov}(X+Y) = \frac{1}{2}[\mathbb{V}\mathrm{ar}(X+Y) - \mathbb{V}\mathrm{ar}(X) - \mathbb{V}\mathrm{ar}(Y)]$ may be useful.)

Exercise 69

The uncertain development of interest can be accounted for by letting the interest rate $r(t)$ at any time t depend on the "state of the economy" $Y(t)$ modeled as a stochastic process. We will assume that $\{Y(t)\}_{t \geq 0}$ is a continuous time Markov chain with finite state space $\mathcal{J}^Y = \{1, \ldots, J^Y\}$. The probability of transition from state e to state f in the time interval from t to u is denoted by

$$p_{ef}(t,u) = \mathbb{P}[Y(u) = f \mid Y(t) = e] \,. \tag{F.40}$$

We assume that the process is homogeneous, which means that the transition probabilities $p_{ef}(t,u)$ depend on t and u only through $u - t$, the length of the time interval. This implies that the process has constant intensities of transition,

$$\lambda_{ef} = \lim_{dt \downarrow 0} \frac{p_{ef}(t, t+dt)}{dt} \,. \tag{F.41}$$

Just like the mortality intensity, λ_{ef} is a "probability of transition per time unit". The total intensity of transition out of state e at time t is denoted by

$$\lambda_{e\cdot} = \sum_{f;\, f \neq e} \lambda_{ef} \,.$$

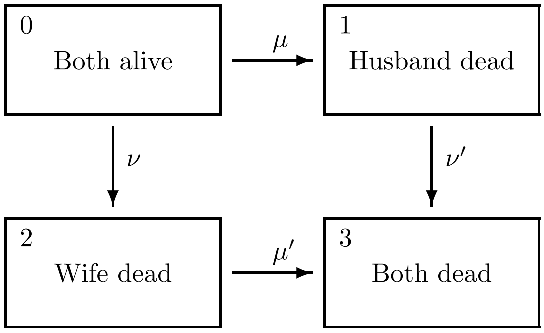


Figure F.4: Sketch of a Markov model for two lives.

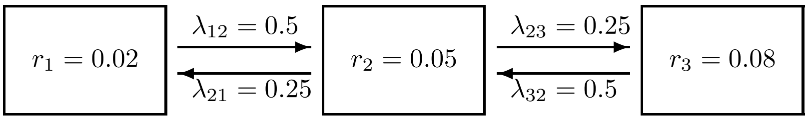


Figure F.5: Sketch of a simple Markov chain interest model.

(a) Use the "direct backward argument" to derive the so-called backward Kolmogorov differential equations for the transition probabilities:

$$\frac{\partial}{\partial t} p_{ef}(t,u) = \sum_{g;g\neq e} \lambda_{eg}(p_{ef}(t,u) - p_{gf}(t,u)) \,.$$

The side conditions are

$$p_{ef}(u,u) = \delta_{ef} \,,$$

i.e. 1 if $e = f$ and 0 otherwise (the Kroenecker delta). These differential equations can easily be solved numerically with a suitable variation of the program 'prores1.pas' or 'prores2.pas'.

Stochastic interest is now modeled by letting the interest rate assume J^Y possible values $r_1, \ldots, r_{J^Y}$ and, at any time t, the interest rate is $r(t) = r_{Y(t)}$ (dependent on the state of the economy).

Figure F.5 shows a flow-chart of a simple Markov chain interest rate model with three states, 0.02, 0.05, 0.08. Direct transition can only be made to a neighbouring state, and the total intensity of transition out of any state is 0.5, that is, the interest rate changes once in two years on the average. By symmetry, the long run average interest rate is 0.05.

(b) Use the direct backward argument to derive differential equations for the state-wise expected discount factors

$$W_e(t) = \mathbb{E}[e^{-\int_t^n r(s)\,ds} \mid Y(t) = e]\,,$$

$t \in [0, n)$, What are the side conditions?

Remark: It has been said that "all models are wrong, but some are useful". The Markov model proposed here is, of course, unable to mimic perfectly the development of the true interest rate. We can, however, by judicious choice of the state space (sufficiently big) and the transition intensities, make it catch the basic features of the real world interest fairly well.

(c) Consider the model in Fig. F.5. Fill in appropriate statements in the program 'prores2.pas' to make it compute the state-wise expected discount factors $\mathbb{E}\left[\exp\left(-\int_0^5 r(s)\,ds\right)\middle| Y(0) = e\right]$, $e = 1, 2, 3$.

Exercise 70
Poisson processes, which are totally memoryless, can of course be generated from continuous time Markov chains, which are more general. For instance, let $\{Y(t)\}_{t\geq 0}$ be a Markov chain on the state space $\{1, 2\}$ with intensities of transition $\mu_{12}(t) = \mu_{21}(t) = \lambda$. Then $\{N(t)\}_{t\geq 0}$ defined by $N(t) = N_{12}(t) + N_{21}(t)$ (the total number of transitions in $(0, t]$) is a Poisson process with intensity λ; transitions counted by N occur with intensity λ at any time regardless of the past history of the process. Two independent Poisson processes, $\{N_1(t)\}_{t\geq 0}$ with intensity λ_1 and $\{N_2(t)\}_{t\geq 0}$ with intensity λ_2, can be generated by letting $\{Y(t)\}_{t\geq 0}$ be a Markov chain on the state space $\{1, 2, 3, 4\}$ with intensities $\mu_{12}(t) = \mu_{21}(t) = \mu_{34}(t) = \mu_{43}(t) = \lambda_1$, $\mu_{13}(t) = \mu_{31}(t) = \mu_{24}(t) = \mu_{42}(t) = \lambda_2$, $\mu_{14}(t) = \mu_{41}(t) = \mu_{23}(t) = \mu_{32}(t) = 0$, and defining $N_1(t) = N_{12}(t)+N_{21}(t)+N_{34}(t)+N_{43}(t)$ and $N_2(t) = N_{13}(t)+N_{31}(t)+N_{24}(t)+N_{42}(t)$. Three independent Poisson processes can be generated from a Markov chain with 8 states (work out the details), and, in general, k independent Poisson processes can be generated from a Markov chain with 2^k states.

(a) Let $\{Y(t)\}_{t\geq 0}$ be a Markov chain on the state space $\{1, 2\}$, and take $\mu_{12}(t) = \mu_{21}(t) = 1$. Denote the corresponding indicator processes and counting processes be $I_e(t)$ and $N_{ef}(t)$. The total time spent by Y in state 1 during the time interval $(t, n]$ is

$$T_1(t, n] = \int_t^n I_1(\tau)\,d\tau\,,$$

and the total number of transitions made from state 1 to state 2 in that interval is

$$N_{12}(t, n] = N_{12}(n) - N_{12}(t)\,.$$

(These quantities can be viewed as present values of benefits of annuity type and assurance type, respectively, for a two-state policy with no interest.)

(a) Assume $Y(0) = 1$. What are the interpretations of the random variables $T_1(t, u]$ and $N_{12}(t, u]$ in terms of the Poisson process $N(t) = N_{12}(t) + N_{21}(t)$?

(b) Find, by solving the relevant differential equations analytically, explicit expressions for the first two state-wise conditional moments

$$V_e^{(q)}(t) = \mathbb{E}[T_1^q(t,n] \mid Y(t) = e] \,,$$

$$W_e^{(q)}(t) = \mathbb{E}[N_{12}^q(t,n] \mid Y(t) = e] \,,$$

$e = 1, 2$, $q = 1, 2$, and find also the corresponding variances. (You should obtain e.g.

$$\mathbb{E}[T_1(0,1] \mid Y(0) = 1] = \frac{1}{4\mu}\left(1 + 2\mu - e^{-2\mu}\right) \,,$$

$$\mathbb{V}\mathrm{ar}[T_1(0,1] \mid Y(0) = 1] = \frac{1}{16\mu^2}\left(1 + 4\mu - (2 - e^{-2\mu})^2\right) \,.$$

(c) Using the program prores1.pas (or prores2.pas), solve the differential equations also numerically and compare the results with the exact solutions obtained in (b).

Exercise 71
Let $\{N(t)\}_{t\geq 0}$ be Poisson process with intensity λ. This is a counting process of even simpler type than the counting processes associated with a Markov chain; N is not only Markov, but also has independent increments. Thus, in any small time interval $[t, t+dt)$ the process N makes a jump of 1 with probability $\lambda\, dt$ regardless of the past history of the process in $[0, t)$.

Let the price $S(t)$ of a share of stock at time t be modelled as a so-called geometric Poisson process with drift,

$$S(t) = \exp\left(\alpha t + \beta N(t)\right) \,,$$

$t \geq 0$. If $\beta = 0$, then $S(t)$ is just the accumulation factor for a bank account with fixed interest rate. The Poisson term in the exponent adds jumps at random times, and a jump at time t makes the stock price jump from $S(t-)$ to $S(t) = S(t-)\, e^{\beta}$. Thus, $\gamma = e^{\beta} - 1$ is the relative change $(S(t) - S(t-))/S(t-)$ in the stock price at the jump time. Between the jumps the stock price increases at fixed "rate of interest" α.

(a) Find the expected value $\mathbb{E}[S(t)]$ at time 0 of the stock price at time t, and do this in two ways: First, work directly with the Poisson distribution of $N(t)$ and, second, solve a differential equation obtained by the direct backward construction (condition on "what happens in the small time interval $[0, dt)$"). Explain that, having determined the expected value, higher order moments are easily obtained.

(b) Find the dynamics $dS(t)$ of the stock price by applying the change of variable rule, see Appendix A.

(c) Using the direct backward construction, show that the expected present value of a perpetuity (an everlasting annuity), is

$$\mathbb{E}\left[\int_0^\infty S^{-1}(\tau)\, d\tau\right] = (a + (1 - e^{-\beta}))^{-1} \,.$$

(d) Let $\{N_1(t)\}_{t\geq 0}$ and $\{N_2(t)\}_{t\geq 0}$ be independent Poisson processes with intensities λ_1 and λ_2, respectively. Let the price $S(t)$ of a share of stock at time t be

$$S(t) = \exp\left(\alpha t + \beta_1 N_1(t) + \beta_2 N_2(t)\right) \,,$$

$t \geq 0$. Letting β_1 and β_2 have opposite signs, we have created a stock price which may increase or decrease with instantaneous jumps. Write out the straightforward analogs of the formulas in Items (a) - (c) for this more general model.

Exercise 72
(a) Using your Turbo Pascal program with the Danish basis, compute the net premium rate π and the net premium reserve V_t (at selected times t) for an endowment insurance with age at entry $x = 30$, term $n = 30$, sum insured $b_t = b_n = b = 1$, and premium payable continuously at constant rate throughout the contract period.

(b) Compute the gross premium rate π' and the gross premium reserve V_t' assuming that administration expenses consist of a lump sum cost of $0.003 + 0.001\, b$ at time 0, costs incurring continuously at rate $0.0001 + 0.01\, \pi' + 0.005\, V_t'$ at any time t in the insurance period, a cost of 0.002 due immediately upon possible payment of the death benefit, and a cost of 0.0001 due at time n upon possible payment of the endowment benefit. Compare with the quantities net of expenses found in Item (a).

Exercise 73
In connection with a pension insurance there is an additional benefit which is a sum insured to possible dependent children less than 18 years old at the time of death of the insured. I the technical basis we therefore need to make assumptions about births. We have to distinguish by sex, and in the following we are going to consider insured women only. The Figure below shows a flowchart describing a life history with births (at most J). To keep things simple, we will assume that the process is Markov, that all participants enter the insurance scheme in state 0 at age 0, and that the individual life histories are independent replicates of the process. Assume furthermore that the mortality intensity $\mu_j(t)$ and the birth intensity $\phi_j(t)$ for a t year old woman, who has given birth to j children, are given by

$$\mu_j(t) = \alpha_j + \beta c^t \, , \; j = 0, \ldots, J, \tag{F.42}$$

$$\phi_j(t) = \eta_j f(t) \, , \; j = 0, \ldots, J-1, \tag{F.43}$$

where c and the function f are known and the parameters α_j, β, and η_j are unknown.

(a) Assume for the time being that, at the time of consideration, we have complete information about the past history for all individuals that have previously been or currently are insured under the scheme (i.e. we know the exact times of possible births and death). Put up the equations for determining the ML (maximum likelihood) estimators for the unknown parameters and, to the extent possible, find explicit expressions for the estimators. Explain also how to determine the asymptotic covariance matrix.

(b) In our model there is non-differential mortality if

$$\alpha_0 = \cdots = \alpha_J \, . \tag{F.44}$$

Derive the likelihood ratio test for the null hypothesis (F.44) and determine the rejection limit such that the asymptotic level is 10%.

Assume now that the past history of births and death is being observed only upon death of the insured, when the additional benefit to the possible dependents is due. Suppose that the statistical data comprise only those who are dead at the time of consideration and that for each of those there is a complete record of the times of possible births and of death. In these data the observed life history of a woman, who entered the scheme u years ago, is governed by a Markov process as described above, but with intensities

$$\mu_j^*(t) = \mu_j(t)\frac{1}{p_{jd}(t,u)}, \tag{F.45}$$

$$\phi_j^*(t) = \phi_j(t)\frac{p_{j+1,d}(t,u)}{p_{jd}(t,u)}. \tag{F.46}$$

It is to be proved that, if the mortality increases with the number of births, that is,

$$\mu_j(t) \leq \mu_{j+1}(t)\,,\ t > 0,\ j = 0,\ldots,J-1, \tag{F.47}$$

then

$$\phi_j^*(t) \geq \phi_j(t)\,,\ t > 0,\ j = 0,\ldots,J-1\,. \tag{F.48}$$

Introduce

$$p_j(t,u) = 1 - p_{jd}(t,u)\,, \tag{F.49}$$

the probability that a t year old with j births will survive to age u. Build the proof as follows:

(c) Prove that, for $t \leq \tau \leq u$,

$$p_j(t,u) = \sum_{k\geq j} p_{jk}(t,u) = \sum_{k\geq j} p_{jk}(t,\tau)p_k(\tau,u)\,. \tag{F.50}$$

(d) Prove that

$$p_j(t,u) = e^{-\int_t^u \mu_j(t,\vartheta)d\vartheta}\,, \tag{F.51}$$

where

$$\mu_j(t,u) = \frac{\sum_{k\geq j} p_{jk}(t,u)\mu_k(u)}{\sum_{k\geq j} p_{jk}(t,u)}\,, \tag{F.52}$$

the mortality intensity at age u for a woman who is in state j at time $t < u$).

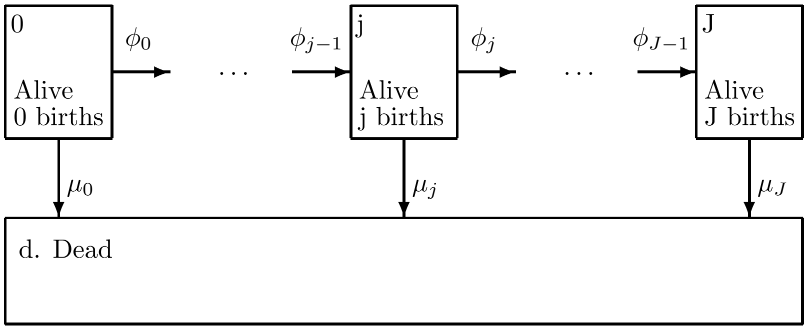
(e) Show (preferably by direct reasoning), that

$$p_j(t,u) = e^{-\int_t^u(\phi_j+\mu_j)} + \int_t^u e^{-\int_t^\tau(\phi_j+\mu_j)}\phi_j(\tau)p_{j+1}(\tau,u)\,d\tau\,. \tag{F.53}$$

(d) Use the results (F.50) – (F.53) to show that (F.47) implies

$$p_{j+1}(t,u) \leq p_j(t,u)\,,\ j = 0,\ldots,J-1\,, \tag{F.54}$$

which by (F.46) and (F.49) implies (F.48). The relationship (F.54) is easy to prove for $j = J-1$, and the result then follows by induction.



Commment: The inequality (F.48) means that the fertility rates will be overestimated if one uses the estimators for the ϕ_j^* based on diseased participants in the scheme. If the inequalities (F.47) are reversed, then also the inequality (F.48) will be reversed, and the estimators the ϕ_j^* will underestimate the fertility. In particular it follows that, under the hypothesis of non-differential mortality, the fertility rates will be unbiasedly estimated from the selected material of diseased participants.

Exercise 74
We adopt the usual notation and assumptions of the theory of multi-life insurance policies and consider two independent lives (x) and (y) with remaining life lengths T_x and T_y, respectively.

(a) Assume that the benefit is an assurance of 1 payable at time T_y if $2T_x < T_y < n$ and that premium is payable at constant rate π until time $\min(T_x, T_y, n/2)$, where n is the term of the contract (fixed). Determine the equivalence premium π.

(b) Propose a method for computing the premium numerically. (Hint: One possibility is to treat ${}_{t/2}p_x$ as a survival function ${}_t\tilde{p}_x$ with intensity $\tilde{\mu}_{x+t}$, which you would need to express in terms of μ, and then solve a Thiele differential equation numerically.)

(c) Determine the reserve at any time t, assuming that the insurer currently knows the complete past history of the two lives. You need to distinguish between various cases, whether (y) is alive or dead, whether t is before or after time $n/2$, and whether x is alive or dead and, if dead, when. Is the reserve always non-negative?

(d) What is the variance of the present value of the benefit?

Exercise 75
Find an expression for the mortality intensity of the life length of the last-survivor status $\overline{x_1 \ldots x_r}$. Do this by direct reasoning and also the hard way by calculating

$$\mu_{\overline{x_1 \ldots x_r}}(t) = -\frac{d}{dt} \ln {}_t p_{\overline{x_1 \ldots x_r}}\,.$$

Exercise 76
An example of an insurance policy with benefits that are 'path-dependent', that is, dependent on the past history of the policy:

(a) Consider two independent lives (x) and (y). Find the expected present value at time 0 of a life annuity payable continuously at rate 1 from time T_{xy} until time $\max(T_{xy} + 20\,,\, T_{\overline{xy}} + 10)$. In words, payments start at the time of the first death and continues thereafter for a term of 20 years or until 10 years after the death of the survivor, whichever is the longer period.

(b) Suppose premium is payable continuously at level rate π from time 0 until time T_{xy}. What is the reserve for this policy? This question is difficult and will be addressed later, but you can start thinking about it.

Exercise 77
Another example of 'path-dependent' payments: Consider three independent lives (x), (y), and (z). Find the expected present value of a sum insured of 1 payable at time T_x if $T_x < \min(T_z, T_y + 20)$. State in words what this contractual benefit is.

Exercise 78
Use the program 'prores1.pas' to compute

$$\pi = \frac{\bar{A}_{\overline{20,20,\overline{n}|}}}{\bar{a}_{20,20,\overline{n}|}}\,,$$

assuming that both lives follow the mortality law G82M and that the interest rate is $r = \ln(1.05)$. What are the terms of the policy for which π is the level equivalence premium rate?
Exercise 80
We refer here to Section 7.9 'Dependent lives' in BL.

(a) Prove the rather obvious statements PQD(T, T), AS(T, T), and RTI$(T|T)$.

(b) Prove that the Definitions PQD – RTI are equivalent to the modified definitions obtained upon replacing the strict inequalities $S > s$ and $T > t$ with the non-strict inequalities $S \geq s$ and $T \geq t$. For instance, for PQD prove that (7.92) is equivalent to

$$\mathbb{P}[S \geq s\,,\, T \geq t] \;\geq\; \mathbb{P}[S \geq s]\,\mathbb{P}[T \geq t] \text{ for all } s \text{ and } t.$$

(c) Negative dependence in the PQD sense: Prove that PQD$(-S, T)$ is equivalent to

$$\mathbb{P}[S > s\,,\, T > t] \;\leq\; \mathbb{P}[S > s]\,\mathbb{P}[T > t] \text{ for all } s \text{ and } t.$$

(d) Negative dependence in the AS sense: Prove that AS$(-S, T)$ is equivalent to $\mathrm{Cov}(g(S, T), h(S, T)) \leq 0$ for all real-valued functions g and h that are decreasing in S and increasing in T (and for which the covariance exists).

(e) Negative dependence in the RTI sense: Prove that RTI$(-S|T)$ is equivalent to RTD$(S|T)$.

(f) Consider the model in Figure 7.6, and let $\mu = \mu'$ and $\nu = \nu'$. We have shown in Exercise 1 that S and T are then independent.

Now, add a cause of simultaneous death (due to 'catastrophe') with intensity $\mu_{03} = \kappa$. Does it follow that RTI$(S|T)$?

Assume instead, maybe more reasonably, that catastrophe risk is present independently of the state of the marriage: $\mu_{01}(t) = \mu(t)$, $\mu_{02}(t) = \nu(t)$, $\mu_{03}(t) = \kappa(t)$, $\mu_{13}(t) = \nu(t) + \kappa(t)$, $\mu_{23}(t) = \mu(t) + \kappa(t)$. Prove that RTI$(S|T)$, hence PQD$(S,T)$.

Exercise 81
We refer to Chapter 8 of BL and Exercise paper No. 6. Note the following correction to Exercise paper No. 6, page 8, line 2 from top: "Recalling (10) and (14),...".

A 30 years old buys a pure life endowment of 1 in $n = 30$ years against premium payable continuously at level rate as long as the policy is in force. The first order basis specifies G82M mortality and interest at instantaneous rate 0.02. The second order basis specifies the same mortality G82M and stochastic interest as given in Fig. 1 of Exercise 2.

(a) See Exercise 20a. Suppose surpluses are to be repaid immediately as cash bonus. Compute the state-wise expected present values (under second order interest) at time 0 of future bonuses, given that the insured survives 30 years. Use the program 'prores2.pas'. It may be a good idea to define two 'alive' states which do not communicate, one for computation of the first order reserve, the other for computation of the expected present values of bonuses.

(b) See Exercise 19. Suppose instead that that surpluses are currently spent on purchase of additional benefits. Compute the state-wise expected values at time 0 of the additional benefit Q_{30}.

OK TO HERE

(b) Give an alternative proof of (3.16) along the following lines: Write

$$(T \wedge b) - (T \wedge a) = \int_a^b I_t \, dt \,,$$

where $I_t = I_{\{T>t\}}$. By Itô's formula,

$$\left(\int_a^b I_t \, dt \right)^k = \int_a^b \left(\int_a^t I_s \, ds \right)^{k-1} I_t \, dt \,.$$

Use $I_s I_t = I_t$ for $s < t$ and $I_t^k = I_t$, and take expectation.

Exercise 2
Find the q-th non-central moment of $PV^{a;m|n}$ in (4.21). Start from

$$(V^{a;m|n})^q = \frac{1}{r^q} \sum_{p=0}^{q} \binom{q}{p} v^{p(T_x \wedge m)} v^{(q-p)(T \wedge (m+n))} \,.$$

Exercise 4
Dependence of expected present values on age, duration, and technical basis.

Commence with the pure endowment treated in Paragraph A. The expected present value in (4.2) can be written as

$$ {}_tE_x = e^{-rt-\int_0^t \mu_{x+s}\,ds}. \tag{F.55}$$

The dependence of ${}_tE_x$ on the deferred period t, the age x, and the technical basis r and μ is to be discussed. Some qualitative aspects are obvious by inspection of (F.55) and also on intuitive grounds: ${}_tE_x$ decreases with increasing t; ${}_tE_x$ decreases with increasing x if μ is an increasing function; ${}_tE_x$ decreases with increasing r (or interest rate i); ${}_tE_x$ decreases by a general increase of μ.

A closer study, also of the quantitative aspects, can be based on the derivatives

$$\begin{aligned} \frac{\partial}{\partial t}\,{}_tE_x &= -\,{}_tE_x\,(r+\mu_{x+t}), && \text{(F.56)}\\ \frac{\partial}{\partial x}\,{}_tE_x &= -\,{}_tE_x\,(\mu_{x+t}-\mu_x) && \text{(F.57)} \end{aligned}$$

(the derivative of $\int_0^t \mu_{x+s}\,ds = \int_0^{x+t}\mu_y\,dy - \int_0^x \mu_y\,dy$ with respect to x is $\mu_{x+t}-\mu_x$),

$$\frac{\partial}{\partial r}\,{}_tE_x = -{}_tE_x t, \tag{F.58}$$

and, if $\mu_x = \mu(x,\alpha)$ is a differentiable function of some finite-dimensional parameter $\alpha = (\alpha_1,\ldots,\alpha_r)$,

$$\frac{\partial}{\partial \alpha_j}\,{}_tE_x = -\,{}_tE_x\,\frac{\partial}{\partial \alpha_j}\int_0^t \mu(x+s,\alpha)\,ds. \tag{F.59}$$

In particular, if the force of mortality is of G-M type,

$$\mu_x = \alpha + \beta c^x,$$

then

$$\begin{aligned} \int_0^t \mu_{x+s}\,ds &= \alpha t + \beta c^x\,(c^t-1)/\ln c\\ &= \alpha t + (\mu_{x+t}-\mu_x)/\ln c. \end{aligned}$$

One easily finds (check the details)

$$\begin{aligned} \frac{\partial}{\partial \alpha}\int_0^t \mu_{x+s}\,ds &= t,\\ \frac{\partial}{\partial \beta}\int_0^t \mu_{x+s}\,ds &= \frac{c^x(c^t-1)}{\ln c}\\ &= \frac{\mu_{x+t}-\mu_x}{\beta\ln c}, \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial c}\int_0^t \mu_{x+s}\,ds &= \beta c^{x-1}\left\{\frac{(x+t)c^t-x}{\ln c} - \frac{c^t-1}{\ln^2 c}\right\}\\ &= \frac{1}{c\ln c}\left\{\left(x-\frac{1}{\ln c}\right)(\mu_{x+t}-\mu_x) + t(\mu_{x+t}-\alpha)\right\}. \end{aligned}$$

Thus, in the G-M case (F.59) specializes to

$$\frac{\partial}{\partial\alpha}\,{}_tE_x = -\,{}_tE_x\,t, \tag{F.60}$$

$$\frac{\partial}{\partial\beta}\,{}_tE_x = -\,{}_tE_x\,\frac{1}{\beta\ln c}\,(\mu_{x+t}-\mu_x), \tag{F.61}$$

$$\frac{\partial}{\partial c}\,{}_tE_x = -\,{}_tE_x\,\frac{1}{c\ln c}\left\{\left(x-\frac{1}{\ln c}\right)(\mu_{x+t}-\mu_x)+t(\mu_{x+t}-\alpha)\right\}. \tag{F.62}$$

The expressions in (F.59) and (F.58) are the same, of course, since ${}_tE_x$ depends on r and α only through their sum $r+\alpha$. In general, a constant change in the force of mortality is equivalent to a change in the force of interest.

A comparison of (F.57) and (F.61) reveals that a change of β is essentially the same as a change of x. In fact, replacing β by $\beta' = \beta c^h$ (say) is equivalent to a shift from x to $x+h$.

The expression in (F.62) is not so transparent, but one may say that a change of c to $c' = c^h$ (say) amounts to an expansion or contraction in age by a factor h.

All right hand side expressions in (F.56)–(F.62) include $-\,{}_tE_x$ as a factor. Division by $-\,{}_tE_x$ gives the derivative of $-\ln\,{}_tE_x$, which is the relative decrease of ${}_tE_x$ (in units of ${}_tE_x$) by a unit increase in the argument. The factors multiplying $-\,{}_tE_x$ form a basis for comparing the arguments with respect to the import of changes. Roughly speaking, for common values of the arguments, they can be ordered as follows according to their impacts on the value of ${}_tE_x$, starting with the less important: x, t and α, β, and c. In fact, a change in t is generally of greater importance than a similar change in x. This result is immediately meaningful since t and x are measured in the same units.

Other comparisons that can be made are not so lucid since the arguments are not "commensurable". For instance, even though changes in β are more important than changes in x in the sense that $\frac{\partial}{\partial\beta}\,{}_tE_x > \frac{\partial}{\partial x}\,{}_tE_x$ (for common values of the arguments, that is), it must be kept in mind that the relevant changes in these two arguments are of quite different order of magnitude. Usually β is confined to values below 10^{-3}, whereas x typically ranges between 20 and 70. To account for this aspect, one should rather compare the two differentials $\frac{\partial}{\partial\beta}\,{}_tE_x\,db$ and $\frac{\partial}{\partial x}\,{}_tE_x\,dx$ for representative changes db and dx, and (presumably) conclude that x produces the greater changes in ${}_tE_x$. Here is another comment along the same line: Despite the fact that r and α play equivalent roles in ${}_tE_x$ from a purely mathematical point of view, the interest rate is widely held to the most important element in the technical basis. The reason is, of course, that the interest rate nowadays varies by percentages whereas the mortality is fairly stable and varies by less than per milles.

Now, consider an n-year temporary level life annuity with expected present value given by (4.14). Taking derivatives under the integral sign, one obtains results about $\bar{a}_{x\,\overline{n|}}$ by just inserting the expressions from (F.56)–(F.62). The dependence on n is obtained immediately from (4.14). The results are

$$\frac{\partial}{\partial n}\,\bar{a}_{x\,\overline{n|}} = {}_nE_x, \tag{F.63}$$

$$\begin{aligned}\frac{\partial}{\partial x}\,\bar{a}_{x\,\overline{n|}} &= -\bar{A}^{\;1}_{x\,\overline{n|}} + \mu_x\bar{a}_{x\,\overline{n|}} \\ &= -\left\{1-(r+\mu_x)\,\bar{a}_{x\,\overline{n|}} - {}_nE_x\right\},\end{aligned} \tag{F.64}$$

$$\frac{\partial}{\partial r}\,\bar{a}_{x\,\overline{n|}} = \frac{\partial}{\partial\alpha}\,\bar{a}_{x\,\overline{n|}} = -(\bar{I}\,\bar{a})_{x\,\overline{n|}}, \tag{F.65}$$

$$\frac{\partial}{\partial i}\,\bar{a}_{x\,\overline{n|}} = -(\bar{I}\,\bar{a})_{x\,\overline{n|}}\,v, \tag{F.66}$$

$$\frac{\partial}{\partial \beta}\,\bar{a}_{x\,\overline{n|}} = -\frac{1}{\beta\,\ln c}\left\{1-(r+\mu_x)\,\bar{a}_{x\,\overline{n|}} - {}_nE_x\right\}, \tag{F.67}$$

$$\begin{aligned}\frac{\partial}{\partial c}\,\bar{a}_{x\,\overline{n|}} &= -\frac{1}{c\ln c}\left\{\left(x-\frac{1}{\ln c}\right)\left(\bar{A}_{\overset{1}{x}\,\overline{n|}} - \mu_x\,\bar{a}_{x\,\overline{n|}}\right) + \left(\bar{I}\,\bar{A}\right)^{x}_{x\,\overline{n|}} - \alpha\,\left(\bar{I}\,\bar{a}\right)_{x\,\overline{n|}}\right\} \\ &= -\frac{1}{c\ln c}\Big[\left(x-\frac{1}{\ln c}\right)\{1-(r+\mu_x)\bar{a}_{x\,\overline{n|}} - {}_nE_x\} \\ &\quad + \bar{a}_{x\,\overline{n|}} - (r+\alpha)(\bar{I}\bar{a})_{x\,\overline{n|}} - n\,{}_nE_x\Big].\end{aligned} \tag{F.68}$$

Finally, the dependence of $\bar{A}_{\overset{1}{x}\,\overline{n|}}$ (and $\bar{A}_{x\,\overline{n|}}$) on the arguments is obtained from the results above upon putting $\bar{A}_{\overset{1}{x}\,\overline{n|}} = 1 - r\bar{a}_{x\,\overline{n|}} - {}_nE_x$. Work out the details.

Clearly, since the expressions in (F.56)–(F.62) are all negative when μ is increasing, so are also the expressions in (F.63)–(F.68). As a by-product one obtains that

$$1 > (r+\mu_x)\bar{a}_{x\,\overline{n|}} + {}_nE_x,$$

which can easily be established by direct calculation.

Exercise 2

In the situation of the present paragraph consider the problem of estimating μ from the D_i alone, the interpretation being that it is only observed whether survival to z takes place or not. Show that the likelihood based on $D_i,\ i=1,\ldots,n$, is

$$q^N(1-q)^{n-N},$$

with $q = 1-e^{-\mu z}$, the probability of death before z. (Trivial: it is a binomial situation.)

Note that N is now sufficient, and that the class of distributions is a regular exponential class. The MLE of q is

$$q^* = \frac{N}{n}$$

with the first two moments

$$Eq^* = q,\ \ \mathrm{Var}q^* = \frac{q(1-q)}{n}.$$

It is UMVUE in the class of estimators based on the D_i.

The MLE of $\mu = -\ln(1-q)/z$ is $\mu^* = -\ln(1-q^*)/z$. Apply (D.6) in Appendix D to show that

$$\mu^* \sim_{\mathrm{as}} \mathrm{N}\left(\mu, \frac{q}{nz^2(1-q)}\right).$$

The asymptotic efficiency of $\hat{\mu}$ relative to μ^* is

$$\frac{\mathrm{asVar}\mu^*}{\mathrm{asVar}\hat{\mu}} = \left(\frac{e^{\frac{\mu z}{2}} - e^{-\frac{\mu z}{2}}}{\mu z}\right)^2 = \left(\frac{\sinh(\mu z/2)}{\mu z/2}\right)^2$$

(sinh is the hyperbolic sine function defined by $\sinh(x) = (e^x - e^{-x})/2$). This function measures the loss of information suffered by observing only death/survival by age z as compared to inference based on complete observation throughout the time interval

$(0, z)$. It is ≥ 1 and increases from 1 to ∞ as μz increases from 0 to ∞. Thus, for small μz, the number of deaths is all that matters, whereas for large μz, the life lengths are all that matters. Reflect over these findings.

Exercise 3
Use the general theory of the present section to prove the special results in Section 11.1.

Exercise 4
Work out the details leading to (11.42) – (11.42).

Exercise 5
An insurance company is to carry out a mortality study based on complete records for n life insurance policies with unlimited term period. Policy No. i was issued z_i years ago to a person who was then aged x_i. The actuary sets out to maximize the likelihood

$$\prod_{i=1}^{n} \mu(x_i + T_i, \theta)^{D_i} \exp\left(\int_{x_i}^{x_i+T_i} \mu(s, \theta) ds\right) ,$$

where the notation is selfexplaining.

One employee in the department objects that the method represents a neglect of information; it is known that the insured have survived, not only the period they were insured, but also the period from birth until entry into the scheme. Thus, he claims, the appropriate likelihood is rather

$$\prod_{i=1}^{n} \mu(x_i + T_i, \theta)^{D_i} \exp\left(\int_{0}^{x_i+T_i} \mu(s, \theta) ds\right) .$$

Settle this apparent paradox. (Hint: A suitable framework for discussing the problem is an enriched model with three states, "uninsured", "insured", and "dead".)

Exercise 6

(a) Modify the formulas to the situation where person No. i entered the study z_i years ago at age x_i.

(b) Find explicit expressions for the entries of the asymptotic covariance matrix of the MLE.

Exercise 1
Prove (C.6) in the theorem by induction: Verify that it is true for $r = 1$ and, assuming it is true for a given r, prove that it is true also for $r + 1$.

Exercise 2
Derive the binomial distribution by applying the theorem to the situation where $A_1, \ldots, A_r$ are independent and equally probable, $\mathbb{P}[A_j] = p$, $j = 1, \ldots, r$.

Exercise 3
Use the theorem to find $\mathbb{E}[Q]$ and $\mathbb{V}[Q]$ expressed in terms of the Z_p.

Exercise 4
Find the probability that at least 3 out of 4 events occur.

Let the price at time t of a stock be

$$S(t) \,=\, e^{\alpha t + \beta N(t)}\,,$$

where $N(t)$ is a Poisson process with intensity λ. The money market account bears interest at spot rate r.

At time 0 our hero (x) purchases an n-year unit linked pure life endowment with sum insured $S(n) \vee g$ against a single premium π. Here g is the guaranteed minimum sum insured introduced to protect the insured against poor performance of the stock; if β is negative (in which case α should certainly be positive), then a Poisson event at time t represents a sudden drop in the stock price from $S(t-)$ to $S(t) = S(t-)e^{\beta}$ (a crash in the stock market if the absolute value of β is big). Combining basic principles in finance (no arbitrage) and insurance (equivalence), π should be the expected discounted value of the claim under a suitable probability measure (equivalent martingale measure for the market and physical measure for the life length):

$$\pi \,=\, \mathbb{E}\left[e^{-rn}\,(S(n)\vee g)\,1[T_x > n]\right] \,=\, \mathbb{E}\left[S(n)\vee g\right]\,e^{-rn}\;{}_np_x\,. \tag{F.69}$$

We need to find the expected value appearing in the last expression.

Start as usual from the conditional expected value of the random variable $S(n)\vee g$, given everything that is known by time t:

$$\begin{aligned}
&\mathbb{E}[S(n)\vee g \,|\, N(\tau);\, 0\le\tau\le t]\\
&= \mathbb{E}\left[e^{\alpha n+\beta N(n)}\vee g\,\middle|\, N(\tau);\,0\le\tau\le t\right]\\
&= \mathbb{E}\left[e^{\alpha t+\beta N(t)}e^{\alpha(n-t)+\beta(N(n)-N(t))}\vee g\,\middle|\, N(\tau);\,0\le\tau\le t\right]\,.
\end{aligned}$$

Here we have separated out what pertains to the past (known under the conditioning) and what pertains to the future (remains random under the conditioning), and it is seen that we can work with the function

$$W(t,u) \,=\, \mathbb{E}\left[u\,e^{\alpha(n-t)+\beta(N(n)-N(t))}\vee g\right]\,.$$

Preparing for a backward construction, write

$$W(t,u) \;=\;\; = \;\; \mathbb{E}\left[u\,e^{\alpha dt+\beta(N(t+dt)-N(t))}e^{\alpha(n-t-dt)+\beta(N(n)-N(t+dt))}\vee g\right]\,,$$

and proceed as usual, conditioning on what happens in $(t, t+dt)$:

$$\begin{aligned}
W(t,u) \;&=\; (1\,-\,\lambda\,dt)\,W\left(t+dt, ue^{\alpha dt}\right) \,+\, \lambda\,dt\,W\left(t+dt, ue^{\alpha dt+\beta}\right) \,+\, o(dt)\\
&=\; W\left(t+dt, ue^{\alpha dt}\right) \,-\, \lambda\,dt\,W(t,u) \,+\, \lambda\,dt\,W\left(t, ue^{\beta}\right) \,+\, o(dt)\\
&=\; W(t+dt, u\,+\,u\,\alpha\,dt) \,-\, \lambda\,dt\,W(t,u) \,+\, \lambda\,dt\,W\left(t, ue^{\beta}\right) \,+\, o(dt)\\
&=\; W(t,u) \,+\, \frac{\partial}{\partial t}W(t,u)\,dt \,+\, \frac{\partial}{\partial u}W(t,u)\,u\,\alpha\,dt\\
&\quad\; -\lambda\,dt\,W(t,u) \,+\, \lambda\,dt\,W\left(t, ue^{\beta}\right) \,+\, o(dt)\,.
\end{aligned}$$

We arrive at

$$\frac{\partial}{\partial t}W(t,u) \,+\, \frac{\partial}{\partial u}W(t,u)\,u\,\alpha \,-\, \lambda\,W(t,u) \,+\, \lambda\,W\left(t, ue^{\beta}\right) \,=\, 0\,,$$

which is to be solved subject to the condition

$$W(n,u) \,=\, u \vee g\,.$$

Remark: We could have written (F.69) as

$$\pi \,=\, \mathbb{E}\left[e^{\beta N(n)} \vee ge^{-\alpha\, n}\right]\, e^{(\alpha-r)\,n}\ {}_n p_x\,,$$

and, redefining $W(t,u)$ accordingly, essentially get rid of $e^{-\alpha\, t}$. We have chosen the present approach since it gives us an opportunity to see the different roles of the (non-stochastic) smooth function $e^{-\alpha\, t}$ and the (stochastic) jump process $N(t)$.

Appendix G

Solutions to exercises

Exercise 1
(a) Exact value $\int_0^1 t^2\,dt = 1/3 = 0.3333...$. Can be computed by solving numerically the ordinary differential equation

$$v'(t) \,=\, a(t) + b(t)\,v(t) \tag{G.1}$$

by 'Ode-1.pas'. Take

$$v(t) \,=\, \int_0^t s^2\,ds\,, \qquad v'(t) \,=\, t^2\,,$$

hence $a(t) = t^2$ and $b(t) = 0$. Compute forwards ('F') starting from $v(0) = 0$. Insert the following statements in the program:

```
(*SPECIFY DIMENSION OF v !*)
dim := 1;

(*SPECIFY DIRECTION OF COMPUTATION - FORWARD ('F') OR BACKWARD
('B'):!*)
BF := 'F';

(*SPECIFY THE TIME INTERVAL [0,T] BY INSERTING T IN "term" AND - IF
NEEDED - THE AGE OF THE LIFE AT TIME 0 ! *)
term := 1;

(*SPECIFY BOUNDARY CONDITION(S) AT TIME t = 0 IF FORWARD (BF =
'F') AND AT t = T if BACKWARD (BF = 'B'): !*)
v[1] := 0;

(*SPECIFY THOSE COEFFICIENTS a AND B AND DERIVATIVES a' and B'
THAT ARE NOT 0 !*)
a[1] := t*t;
a1[1] := 2*t;
```

Similar for $\int_0^1 t^{-1/2}\,dt = 0.5$. Take

$$v(t) = \int_0^t s^{-1/2}\,ds\,, \qquad v'(t) = t^{-1/2}\,,$$

hence a(t) = $1/\sqrt{t}$ and b(t) = 0. Statements as above, except

(*SPECIFY THOSE COEFFICIENTS a AND B AND DERIVATIVES a' and B' THAT ARE NOT 0 !*)
a[1] := 1/sqrt{t};
a1[1] := -0.5*a[1]/t;

(b) Since $\Phi(-x) = 1 - \Phi(x)$, we need only compute $\Phi(x)$ for positive x. Moreover, $\Phi(0) = 0.5$, hence

$$\Phi(x) = 0.5 + \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{t^2}{2}\right)\,dt\,.$$

Put

$$v(t) = \Phi(t)\,, \qquad v'(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)\,,$$

and use 'Ode-1.pas' for (G.1) with $a(t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})$ and $b(t) = 0$, working forwards ('F') starting from $v(0) = 0.5$. Statements as above, except

(*SPECIFY THOSE COEFFICIENTS a AND B AND DERIVATIVES a' and B' THAT ARE NOT 0 !*)
a[1] := (1/sqrt{2*pi})*exp(-t*t/2);
a1[1] := a[1]*(-t);

(c) U_t: Use 'Ode-1.pas' forwards for the differential equation $U_t' = rU_t + 1$.
V_t: Use 'Ode-1.pas' backwards for the differential equation $V_t' = rV_t - 1$.

Exercise 3
Fig 3.1: The functions we are interested in are

$$\begin{aligned} v_1(t) &= \bar{F}(t) = \exp\left(-\int_0^t \mu(s)\,ds\right) = \exp\left(-\alpha\,t - \beta\frac{e^{\gamma\,t} - 1}{\gamma}\right)\,, \\ v_2(t) &= f(t) = \bar{F}(t)\,\mu(t) = v_1(t)\,v_3(t)\,, \\ v_3(t) &= \mu(t) = \alpha + \beta\,e^{\gamma\,t}\,. \end{aligned}$$

They can be computed directly since they are given by explicit formulas. Thus, the program 'Ode-1.pas' is not really needed, but it is still useful since it can produce a nicely arranged output. One could drop everything that has to do with the difference scheme and just put in the following statements beginning from line 27:

dim := 3;

(*auxiliary quantities from Danish life table :*)
alpha := 0.0005; (*Gompertz-Makeham parameters*)
beta := 0.00007585775;

```
gamma := 0.038*ln(10);

for l := 0 to 100 do (*Do not confuse the letter l with the number 1*)
begin
writeln; writeln(odeout); (*Line shift*)
t := l;
v[3] := alpha + beta*exp(gamma*t);
v[1] := exp ( - alpha*t - beta*( exp(gamma*t) - 1)/gamma );
v[2] := v[1]*v[3];
write(t:4,' '); write(odeout,t:4,' ');
for j := 1 to dim do
begin
write(' ',v[j]); write(odeout,' ',v[j]);
end;
end;
close(odeout);
end.
```

Alternatively, one could use the program as it is to compute $\bar{F}$ and add statements to compute f and μ. We have

$$v'_1(t) = b_{11}(t)\, v_1(t)\,,$$

where

$$b_{11}(t) = -\alpha - \beta\, exp(\gamma\, t))\,.$$

Having computed $b_{11}(t)$ and $v_1(t)$, we compute

$$\begin{aligned} v_3(t) &= -b_{11}(t)\,, \\ v_2(t) &= v_1(t)\, v_3(t)\,. \end{aligned}$$

Statements:

```
(*SPECIFY DIMENSION OF v !*)
dim := 1;

(* SPECIFY DIRECTION OF COMPUTATION - FORWARD ('F') OR BACK-
WARD ('B'): !*)
BF := 'F';

(*auxiliary quantities from Danish life table :*)
alpha := 0.0005; (*Gompertz-Makeham parameters*)
beta := 0.00007585775;
gamma := 0.038*ln(10);

(*SPECIFY THE TIME INTERVAL [0,T] BY INSERTING T IN "term" AND - IF
NEEDED - THE AGE OF THE LIFE AT TIME 0 ! *)
x:= 0; term:= 100;

(*SPECIFY STEPS IN DIFFERENCE SCHEME AND OUTPUT !: *)
steps := 10000; (*number of steps in difference method*)
```

```
outp := 100; (*number of values in output*)
h := term/steps; (*steplength*)
count := steps/outp - 0.0005;

(* SPECIFY BOUNDARY CONDITION(S) AT TIME t = 0 IF FORWARD (BF =
'F') AND AT t = term if BACKWARD (BF = 'B'): !*)
v[1] := 1;

(*SPECIFY THOSE COEFFICIENTS a AND B AND DERIVATIVES a' and B'
THAT ARE NOT 0 !*)
B[1,1] := - alpha - beta*exp(gamma*(x+t));
B1[1,1] := - beta*exp(gamma*(x+t))*gamma;
```

(Then comes the difference scheme, and we do not need to do anything until we come to the output at times t = 0,1,... :)

```
if countout > count then
begin
v[3] := - B[1,1]; v[2] := v[1]*v[3]; (*this comes extra*)
countout := 0;
writeln; writeln(odeout);
write(t:4,' '); write(odeout,t:4,' ');
for j := 1 to 3 do (*NB! 3 instead of dim*)
begin
write(' ',v[j]); write(odeout,' ',v[j]);
end;
end;
end; (*difference scheme*)
close(odeout);
end.
```

Exercise 4

(a) A special case of (b) - just put $\alpha = 1$.

(b) Write out the defining expression in each interval where there is one analytic expression:

$$F(t) = \begin{cases} 0 & , \quad t \le 0\,, \\ 1 - \left(1 - \frac{t}{\omega}\right)^{\alpha} & , \quad 0 < t < \omega\,, \\ 1 & , \quad t \ge \omega\,. \end{cases}$$

$$\bar{F}(t) = \begin{cases} 1 & , \quad t \le 0\,, \\ \left(\frac{\omega - t}{\omega}\right)^{\alpha} & , \quad 0 < t < \omega\,, \\ 0 & , \quad t \ge \omega\,. \end{cases}$$

$$f(t) = \begin{cases} -\alpha\left(\frac{\omega - t}{\omega}\right)^{\alpha-1}\left(-\frac{1}{\omega}\right) = \frac{\alpha}{\omega}\left(\frac{\omega - t}{\omega}\right)^{\alpha-1} & , \quad 0 < t < \omega\,, \\ 0 & \quad \text{else.} \end{cases}$$

The mortality intensity is $\mu(t) = f(t)/(1 - F(t))$, defined for all t such that the denominator is positive:

$$\mu(t) = \begin{cases} \frac{\alpha}{\omega}\left(\frac{\omega - t}{\omega}\right)^{-1} = \frac{\alpha}{\omega - t} & , \quad 0 < t < \omega\,, \\ \text{undefined} & , \quad t \ge \omega\,. \end{cases}$$

Note that $\mu(t) \nearrow +\infty$ as $t \nearrow \omega$ (as is always the case if $\omega < \infty$ and μ is non-decreasing).

For x and $x+t$ both in $(0,\omega)$,

$$\bar{F}(t|x) = \frac{\bar{F}(x+t)}{\bar{F}(x)} = \left(\frac{\omega - x - t}{\omega - x}\right)^{\alpha},$$

the same type of distribution as $\bar{F}$, only with ω replaced by $\omega - x$. Thus we find $f(t|x)$ and $\mu(t|x)$ by just replacing ω with $\omega - x$ in the formulas above.

(c) Do as above.

Exercise 5
(a)

$${}_tp_0 = \mathbb{P}[T > t] = \mathbb{P}[M]\,\mathbb{P}[T > t|M] + \mathbb{P}[F]\,\mathbb{P}[T > t|F] = s_0^m\,{}_tp_0^m + s_0^f\,{}_tp_0^f\,.$$

$$s_t^m = \mathbb{P}[M|T > t] = \frac{\mathbb{P}[M \cap (T > t)]}{\mathbb{P}[T > t]} = \frac{\mathbb{P}[M]\,\mathbb{P}[T > t|M]}{\mathbb{P}[T > t]} = \frac{s_0^m\,{}_tp_0^m}{s_0^m\,{}_tp_0^m + s_0^f\,{}_tp_0^f}\,.$$

(b)

$$\begin{aligned}\mu_t &= -\frac{\frac{d}{dt}\,{}_tp_0}{{}_tp_0} = -\frac{s_0^m\frac{d}{dt}\,{}_tp_0^m + s_0^f\frac{d}{dt}\,{}_tp_0^f}{s_0^m\,{}_tp_0^m + s_0^f\,{}_tp_0^f} = \frac{s_0^m\,{}_tp_0^m\,\mu_t^m + s_0^f\,{}_tp_0^f\,\mu_t^f}{s_0^m\,{}_tp_0^m + s_0^f\,{}_tp_0^f} \\ &= s_t^m\,\mu_t^m + s_t^f\,\mu_t^f\,,\end{aligned}$$

a weighted average of the mortalities at age t for males and females, the weights being the conditional probabilities of being male and female, respectively, given survival to age t.

(c)

$$s_t^m = \frac{s_0^m}{s_0^m + s_0^f\,{}_tp_0^f/{}_tp_0^m} = \frac{s_0^m}{s_0^m + s_0^f\exp\left[\int_0^t(\mu_s^m - \mu_s^f)\,ds\right]}\,,$$

a decreasing function of t if $\mu_t^m - \mu_t^f > 0$ for all $t > 0$.
If $\int_0^\infty(\mu_s^m - \mu_s^f)\,ds = \infty$, then $s_t^m \to 0$ as $t \to \infty$.

(d)

$${}_tp_x = \frac{{}_{x+t}p_0}{{}_xp_0} = \frac{s_0^m\,{}_{x+t}p_0^m + s_0^f\,{}_{x+t}p_0^f}{s_0^m\,{}_xp_0^m + s_0^f\,{}_xp_0^f} = s_x^m\,{}_tp_x^m + s_x^f\,{}_tp_x^f\,.$$

(From this expression we could have derived the mortality intensity μ_x upon forming $\lim_{t\searrow 0}(1 - {}_tp_x)/t$, which would give the result in (b).)
Similarly

$${}_{m|n}q_x = s_x^m\,{}_{m|n}q_x^m + s_x^f\,{}_{m|n}q_x^f\,,$$
$${}_nE_x = s_x^m\,{}_nE_x^m + s_x^f\,{}_nE_x^f\,.$$

Exercise 6

$$\bar{F}(t) = a_0 + a_1\,t + a_2\,t^2 + a_3\,t^3\,, \qquad t \in [0,\omega]\,, \qquad a_3 \neq 0,.$$

(a)

$$f(t) \;=\; -\bar{F}'(t) \;=\; -a_1 - 2a_2\,t - 3a_3\,t^2\,, \qquad t \in [0,\omega]\,.$$

$$\mu(t) \;=\; \frac{f(t)}{\bar{F}(t)} \;=\; -\frac{a_1 + 2a_2\,t + 3a_3\,t^2}{a_0 + a_1\,t + a_2\,t^2 + a_3\,t^3}\,, \qquad t \in [0,\omega]\,.$$

Proof of

$$\lim_{t\nearrow\omega} \mu(t) \;=\; \lim_{t\nearrow\omega} \frac{f(t)}{\bar{F}(t)} \;=\; \infty \;: \tag{G.2}$$

We know that $\lim_{t\nearrow\omega} \bar{F}(t) = 0$. If $\lim_{t\nearrow\omega} f(t) \neq 0$, then (G.2) holds. If $\lim_{t\nearrow\omega} f(t) = 0$, then we have a $0/0$ expression in the limit and use l'Hospital's rule:

$$\lim_{t\nearrow\omega} \mu(t) = -\lim_{t\nearrow\omega} \frac{f'(t)}{f(t)}\,.$$

If $\lim_{t\nearrow\omega} f'(t) \neq 0$, then (G.2) holds. If $\lim_{t\nearrow\omega} f'(t) = 0$, use l'Hospital again:

$$\lim_{t\nearrow\omega} \mu(t) = -\lim_{t\nearrow\omega} \frac{f''(t)}{f'(t)}\,.$$

But $f''(t) = -6a_3 \neq 0$, and so (G.2) holds.

Observe that $\bar{F}(t|x)$, considered as function of t, is also trinomial.

(b)

$$\begin{aligned}
\bar{F}(0) = 1 \;:& \qquad a_0 \;=\; 1\,,\\
\bar{F}(\omega) = 0 \;:& \qquad a_0 + a_1\,\omega + a_2\,\omega^2 + a_3\,\omega^3 \;=\; 0\,,\\
f(\omega) = 0 \;:& \qquad a_1 + 2a_2\omega + 3a_3\,\omega^2 \;=\; 0\,,\\
\int_0^\omega \bar{F}(t)\,dt \;=\; e_0 \;:& \qquad a_0\,\omega + \frac{a_1}{2}\,\omega^2 + \frac{a_2}{3}\,\omega^3 + \frac{a_3}{4}\,\omega^4 \;=\; e_0\,.
\end{aligned}$$

Exercise 7
The probability distribution of the random variable T_x is given by $\mathbb{P}[T_x \leq t] = {}_tp_x$, $t \geq 0$. We are interested in the probability distribution of a present value $PV(T_x)$ which is just a real-valued function of the random variable T_x: $\mathbb{P}[PV(T_x) \leq u]$, $u \in (-\infty, \infty)$. Thus for each value of u we need to determine the set of values of T_x that make $PV(T_x) \leq u$ and determine its probability. It may be helpful to draw a graph of the function, with T_x on the horizontal axis and $PV(T_x)$ on the vertical axis, and for each given u on the vertical axis determine the set on the horizontal axis where the graph is under u. We summarize the results:

(a) Pure endowment benefit of 1:

$$PV(T_x) \;=\; e^{-rn}\,1[T_x > n] \;=\; \begin{cases} 0 & , \quad T_x \leq n\,,\\ e^{-rn} & , \quad T_x > n\,. \end{cases}$$

A very simple non-decreasing function with just two values.

$$\mathbb{P}[PV(T_x) \le u] \;=\; \begin{cases} 0 & , \quad u < 0\,, \\ 1 - {}_n p_x & , \quad 0 \le u < e^{-rn}\,, \\ 1 & , \quad e^{-rn} \le u\,. \end{cases}$$

(b) Term insurance with sum 1:

$$PV(T_x) \;=\; e^{-r\,T_x}\, 1[T_x \le n] \;=\; \begin{cases} e^{-r\,T_x} & , \quad T_x \le n\,, \\ 0 & , \quad T_x > n\,. \end{cases}$$

A simple non-increasing function starting from its maximum value 1 at $T_x = 0$, decreasing exponentially for $T_x \in (0, n]$, dropping to 0 at $T_x = n$ and remaining 0 thereafter.

$$\mathbb{P}[PV(T_x) \le u] \;=\; \begin{cases} 0 & , \quad u < 0\,, \\ {}_n p_x & , \quad 0 \le u < e^{-rn}\,, \\ {}_{-\ln u\,/\,r} p_x & , \quad e^{-rn} \le u < 1\,, \\ 1 & , \quad 1 \le u\,. \end{cases}$$

The third line on the right is the only one that takes a small bit of calculation: For $e^{-rn} \le u < 1$ solve $e^{-rt} = u$ to find $t = -\ln u\,/\,r$, and conclude that $PV(T_x) \le u$ is equivalent to $T_x \ge t$.

(c) Endowment insurance with sum 1: Same problem as (b), only simpler:

$$\mathbb{P}[PV(T_x) \le u] \;=\; \begin{cases} 0 & , \quad u < e^{-rn}\,, \\ {}_{-\ln u\,/\,r} p_x & , \quad e^{-rn} \le u < 1\,, \\ 1 & , \quad 1 \le u\,. \end{cases}$$

(d) Life annuity of 1 per year:

$$\mathbb{P}[PV(T_x) \le u] \;=\; \begin{cases} 0 & , \quad u < 0\,, \\ 1 - {}_{-\ln(1-ru)\,/\,r} p_x & , \quad 0 \le u < \bar{a}_{\overline{\infty}|}\,, \\ 1 & , \quad \bar{a}_{\overline{\infty}|} \le u\,. \end{cases}$$

Exercise 8

(a) There are several ways of proving these things. We will sketch two:

First method works directly on the expressions for the expected values:
Life endowment:

$${}_{m+n}E_x = e^{-\int_0^{m+n}(r+\mu_{x+u})\,du} = e^{-\int_0^{m}(r+\mu_{x+u})\,du}\, e^{-\int_m^{m+n}(r+\mu_{x+u})\,du}$$

or just

$${}_{m+n}E_x = v^{m+n}\;{}_{m+n}p_x = v^m\;{}_m p_x\, v^n\;{}_n p_{x+m}\,.$$

Deferred life annuity:

$${}_{m|n}\bar{a}_x = \int_m^{m+n} v^t\;{}_t p_x\,dt = v^m\;{}_m p_x \int_m^{m+n} v^{t-m}\;{}_{t-m}p_{x+m}\,dt = {}_m E_x\,\bar{a}_{x+m\,\overline{n}|}\,,$$

the last step by substituting $\tau = t - m$ in the integral.

Second method works with the indicator functions $I_t = 1[T_x > t]$ and the rule of iterated expectation $\mathbb{E}[X] = \mathbb{E}\mathbb{E}[X \,|\, Y]$. (Seems like shooting sparrows with cannons in these simple examples, but serves to illustrate a technique that may be useful):
Life endowment:

$$_{m+n}E_x = \mathbb{E}\left[v^{m+n} I_{m+n}\right] = v^m\, v^n\, \mathbb{E}\left[I_{m+n}\right] \,,$$

and (just to illustrate)

$$\begin{aligned} \mathbb{E}\left[I_{m+n}\right] &= \mathbb{E}\left[\mathbb{E}[I_{m+n} \,|\, I_m]\right] = {_m}p_x\, \mathbb{E}[I_{m+n} \,|\, I_m = 1] + {_m}q_x\, \mathbb{E}[I_{m+n} \,|\, I_m = 0] \\ &= {_m}p_x\, {_n}p_{x+m} \,. \end{aligned}$$

Deferred annuity:

$$\begin{aligned} {_{m|n}}\bar{a}_x &= \mathbb{E}\left[\int_m^{m+n} v^t\, I_t\, dt\right] \\ &= v^m\, \mathbb{E}\left[\mathbb{E}\int_m^{m+n} v^{(t-m)}\, I_t\, dt \,\middle|\, I_m\right] \\ &= v^m\, {_m}p_x \mathbb{E}\left[\mathbb{E}\int_m^{m+n} v^{(t-m)}\, I_t\, dt \,\middle|\, I_m = 1\right] \\ &\quad + v^m\, {_m}q_x \mathbb{E}\left[\mathbb{E}\int_m^{m+n} v^{(t-m)}\, I_t\, dt \,\middle|\, I_m = 0\right] \\ &= v^m\, {_m}p_x\, \bar{a}_{x+m\,\overline{n}|} \,. \end{aligned}$$

Similar for deferred endowment insurance.

(b) Here we use the second method, which is general. Let $PV_{(t,u]}$ denote the random present value at time t of benefits less premiums in $(t,u]$, and abbreviate $PV_t = PV_{(t,\infty]}$. We have

$$PV_t = PV_{(t,u]} + v^{u-t}\, PV_u \,. \tag{G.3}$$

Then

$$V_t = \mathbb{E}[PV_t \,|\, I_t = 1] = \mathbb{E}[PV_{(t,u]} \,|\, I_t = 1] + v^{u-t}\, \mathbb{E}[PV_u \,|\, I_t = 1] \,.$$

Here $\mathbb{E}[PV_{(t,u]} \,|\, I_t = 1] = V_{(t,u]}$, and

$$\mathbb{E}[PV_u \,|\, I_t = 1] = \mathbb{E}\left[\mathbb{E}[PV_u \,|\, I_t = 1,\, I_u] \,|\, I_t = 1\right] = {_{u-t}}p_{x+t}\, V_u + {_{u-t}}q_{x+t} \cdot 0 \,.$$

From this you gather the stated result.

(c) Second method is the superior one. We start from (G.3) for payments deferred in m years:

$$PV_0 = v^m\, PV_m \,.$$

Denote variance by $\mathbb{V}$. We have

$$\mathbb{V}[PV_0] = v^{2m}\mathbb{V}[PV_m]$$

and

$$\begin{aligned} \mathbb{V}[PV_m] &= \mathbb{V}\left[\mathbb{E}[PV_m \mid I_m]\right] + \mathbb{E}\left[\mathbb{V}[PV_m \mid I_m]\right] \\ &= \mathbb{V}\left[I_m\, \mathbb{E}[PV_m \mid I_m = 1]\right] + \mathbb{E}\left[I_m\, \mathbb{V}[PV_m \mid I_m = 1]\right] \\ &= \mathbb{V}[I_m]\, \left(\mathbb{E}\left[PV_m \,|\, I_m = 1\right]\right)^2 + \mathbb{E}[I_m]\, \mathbb{V}\left[PV_m \,|\, I_m = 1\right] \,, \end{aligned}$$

hence

$$\begin{aligned}\mathbb{V}[PV_0] &= v^{2m}\left\{{}_mp_x\,{}_mq_x\;(\mathbb{E}\,[PV_m\,|\,I_m=1])^2 + {}_mp_x\,\mathbb{V}\,[PV_m\,|\,I_m=1]\right\} \\ &= \left({}_mE_x^{(2r)} - {}_mE_x^2\right)\,(\mathbb{E}\,[PV_m\,|\,I_m=1])^2 + {}_mE_x^{(2r)}\,\mathbb{V}\,[PV_m\,|\,I_m=1]\;. \qquad \text{(G.4)}\end{aligned}$$

For instance, for an m year deferred n year annuity (G.4) gives the variance (see (4.16) in BL)

$${}_mE_x^{(2r)}\,\frac{2}{r}\left(\bar{a}_{x+m\,\overline{n}|} - \bar{a}^{(2r)}_{x+m\,\overline{n}|}\right) - {}_mE_x^2\,\bar{a}^2_{x+m\,\overline{n}|}\,.$$

Apply the result to the pure life endowment (a deferred benefit by its very definition) and to a deferred life insurance. (You could put up the expressions immediately using the trick shown in Chapter 4.)

Exercise 9

$$N_t^2 = N_0^2 \,+\, \int_0^t N_\tau\,dN_\tau \,+\, \int_0^t N_{\tau-}\,dN_\tau\,.$$

For $t < T$: All terms on both sides are 0, so the equation holds.
For $t \geq T$: On the left $N_t^2 = 1$. On the right $N_0^2 = 0$, $\int_0^t N_\tau\,dN_\tau = 1$, and $\int_0^t N_{\tau-}\,dN_\tau = 0$, so the equation holds. Ignoring the left limit gives 2 on the right when $t > T$ (the jump of N at T) has been "counted twice").

The relationship

$$\int_0^t N_\tau\,dN_\tau = \int_0^t (N_{\tau-}+1)\,dN_\tau$$

is true for any counting process, not only the simple one considered here. By definition

$$\int_0^t N_\tau\,dN_\tau = \sum_{\tau\le t} N(\tau)(N(\tau)-N(\tau-))$$

The sum on the right ranges effectively over the (finite number of) time points τ where N jumps, and at such a time τ it jumps (by 1) from the value it had just before the jump, $N(\tau-)$, to the value at the jump time, $N(\tau) = N(\tau-)+1$.

In particular, if N has only one jump, then $N(\tau-) = 0$ at the time jump time τ, and the integral is therefore 0.

Exercise 10

(c) Set $v_1(t) = {}_tp_0$, $v_2(t) = \bar{e}_{0:\overline{t}|}$. They satisfy the differential equations

$$v'_1(t) = -v_1(t)\,\mu_t\,,$$

$$v'_2(t) = -v_2(t)\,v_1(t)\,,$$

with side conditions $v_1(0) = 1$, $v_2(0) = 0$.
Statements:

(*SPECIFY DIMENSION OF v !*)
dim := 2;

(* SPECIFY DIRECTION OF COMPUTATION - FORWARD ('F') OR BACKWARD ('B'): !*)

```
BF := 'F';

(*auxiliary quantities from Danish life table :*)
alpha := 0.0005; (*Gompertz-Makeham parameters*)
beta := 0.00007585775;
gamma := 0.038*ln(10);

(*SPECIFY THE TIME INTERVAL [0,T] BY INSERTING T IN "term" AND - IF
NEEDED - THE AGE OF THE LIFE AT TIME 0 ! *)
x:= 0; term:= 100;

(*SPECIFY STEPS IN DIFFERENCE SCHEME AND OUTPUT !: *)
steps := 10000; (*number of steps in difference method*)
outp := 100; (*number of values in output*)
h := term/steps; (*steplength*)
count := steps/outp - 0.0005;

(* SPECIFY BOUNDARY CONDITION(S) AT TIME t = 0 IF FORWARD (BF =
'F') AND AT t = term if BACKWARD (BF = 'B'): !*)
v[1] := 1; v[2] := 0;

(*SPECIFY THOSE COEFFICIENTS a AND B AND DERIVATIVES a' and B'
THAT ARE NOT 0 !*)
B[1,1] := - alpha - beta*exp(gamma*(x+t));
B1[1,1] := - beta*exp(gamma*(x+t))*gamma;
B[2,1] := 1;
```

Exercise 13
Premium payable at rate π in the deferred period $t < m$ and pension payable at rate 1 in the pension period $m < t < m+n$, say. Use Thiele's differential equation: Inspect it directly, or integrate it from 0 to t and from t to n to obtain the retrospective formula for t in the deferred period and the prospective formula in the benefit payment period:

$$V_t = \begin{cases} \pi \int_0^t \exp\left(\int_\tau^t (r+\mu_{x+u})\,du\right)\,d\tau\,, & 0 \le t < m\,, \\ \int_t^{m+n} \exp\left(-\int_t^\tau (r+\mu_{x+u})\,du\right)\,d\tau\,, & m \le t < m+n\,. \end{cases}$$

It is seen that V_t is an increasing function for $t < m$, no matter if μ is increasing or not; increasing t gives a bigger integrand integrated over a longer interval. We know from the theory in Chapter 4 of BL that, for $t \ge m$, $V_t = \bar{a}_{x+t\,\overline{m+n-t}|}$ is decreasing if μ is increasing.

Exercise 14
(a)

$$\begin{aligned} V_t &= \mathbb{E}\left[PV_{(t,n]}\,\middle|\,T_x > t\right] \\ &= \mathbb{E}\left[PV_{(t,t+dt]} + e^{-r\,dt}PV_{(t+dt,n]}\,\middle|\,T_x > t\right] \\ &= (1-\mu_{x+t}\,dt)\,\mathbb{E}\left[PV_{(t,t+dt]} + e^{-r\,dt}PV_{(t+dt,n]}\,\middle|\,T_x > t+dt\right] \end{aligned}$$

$$
\begin{aligned}
& + \mu_{x+t}\, dt\, \mathbb{E}\left[PV_{(t,t+dt]} + e^{-r\,dt} PV_{(t+dt,n]} \,\middle|\, t < T_x < t+dt \right] \\
= \quad & (1-\mu_{x+t}\,dt)\left[-\pi_t\,dt + e^{-r\,dt} V_{t+dt} \right] + \mu_{x+t}\,dt\,\left[b_t + e^{-r\,dt}\cdot 0 \right] \,+\, o(dt) \\
= \quad & -\pi_t\,dt + (1-\mu_{x+t}\,dt)(1-r\,dt)(V_t + \frac{d}{dt}V_t\,dt) + \mu_{x+t}\,dt\,b_t \,+\, o(dt) \\
= \quad & -\,\pi_t\,dt + V_t - (\mu_{x+t}+r)\,dt\,V_t + \frac{d}{dt}V_t\,dt + \mu_{x+t}\,dt\,b_t \,+\, o(dt)
\end{aligned}
$$

Cancel V_t, divide by dt, and let dt go to 0, to arrive at Thiele's diff. eq.

(b)

$$
\begin{aligned}
V_t^{(2)} \quad = \quad & \mathbb{E}\left[PV_{(t,n]}^2 \,\middle|\, T_x > t \right] \\
= \quad & \mathbb{E}\left[(PV_{(t,t+dt]})^2 + 2\,PV_{(t,t+dt]}\,e^{-r\,dt} PV_{(t+dt,n]} + e^{-2r\,dt}(PV_{(t+dt,n]})^2 \,\middle|\, T_x > t \right] \\
= \quad & (1-\mu_{x+t}\,dt)\,\mathbb{E}\left[(PV_{(t,t+dt]})^2 + 2\,PV_{(t,t+dt]}\,e^{-r\,dt} PV_{(t+dt,n]} + e^{-2r\,dt}(PV_{(t+dt,n]})^2 \,\middle|\, T_x > t+dt \right] \\
& + \mu_{x+t}\,dt\,\mathbb{E}\left[(PV_{(t,t+dt]})^2 + 2\,PV_{(t,t+dt]}\,e^{-r\,dt} PV_{(t+dt,n]} + e^{-2r\,dt}(PV_{(t+dt,n]})^2 \,\middle|\, t < T_x < t+dt \right] \\
= \quad & (1-\mu_{x+t}\,dt)\,\left[(-\pi_t\,dt)^2 + 2(-\pi_t\,dt)e^{-r\,dt}V_{t+dt} + e^{-2r\,dt}V_{t+dt}^{(2)} \right] \\
& + \mu_{x+t}\,dt\,\left[b_t^2 + 2b_t e^{-r\,dt}\cdot 0 + e^{-2r\,dt}\cdot 0^2 \right] \,+\, o(dt) \\
= \quad & (1-\mu_{x+t}\,dt)\,2\,(-\pi_t\,dt)\,(1-r\,dt)\left(V_t + \frac{d}{dt}V_t\,dt \right) + (1-\mu_{x+t}\,dt)(1-2r\,dt)\left(V_t^{(2)} \,+\, \frac{d}{dt}V_t^{(2)}\,dt \right) \\
& + \mu_{x+t}\,dt\,b_t^2 \,+\, o(dt) \\
= \quad & -2\pi_t\,dt\,V_t + (1-(\mu_{x+t} \,+\, 2r)\,dt)V_t^{(2)} \,+\, \frac{d}{dt}V_t^{(2)}\,dt \,+\, \mu_{x+t}\,dt\,b_t^2 \,+\, o(dt)
\end{aligned}
$$

Cancel $V_t^{(2)}$, divide by dt, let dt go to 0, and rearrange a bit to arrive at

$$
\frac{d}{dt}V_t^{(2)} = 2\pi_t\,V_t \,+\, (\mu_{x+t} \,+ 2r)V_t^{(2)} \,-\, \mu_{x+t}\,b_t^2\,.
$$

Exercise 15
(a) Let $\tau_1, \tau_2, \ldots$ denote the times of transition of Z listed in chronological order. The process $N_t = \sum_i 1[\tau_i \le t]$, $t \ge 0$, which counts the total number of transitions, is a homogeneous Poisson process with intensity μ. (Write μ instead of λ to make the formulas given in item (b) meaningful.) The times $\tau_1, \tau_3, \ldots$ are the times where N takes odd values, and $\tau_2, \tau_4, \ldots$ are the times where N takes even values. Thus $N_{12}(t,u]$ is the number odd numbered occurrences of Poisson events between time t and time u. In particular, $N_{12}(t) = \left[\frac{N(t)+1}{2}\right]$. Likewise, $N_{21}(t) = \left[\frac{N(t)}{2}\right]$. $T_1(t,u]$ is the total time in $(t,u]$ with odd value of N.

(b) We could derive the backward differential for the functions $V_j^{(q)}(t)$ and $W_j^{(q)}(t)$, but they are really special cases of Thiele's differential equation for the reserve and its generalization to higher order moments: $T_1(t,n]$ is the present value at time t of an n-year annuity of 1 per time unit running in state 1 in the present simple Markov model, and with no interest. Similarly, $N_{12}(t,n]$ is the present value at time t of an n-year insurance 1 payable upon every transition from state 1 to state 2.

Thiele for the annuity:

$$\frac{d}{dt}V_1^{(1)}(t) = -1 - \mu\,(V_2^{(1)}(t) - V_1^{(1)}(t))\,, \tag{G.5}$$

$$\frac{d}{dt}V_2^{(1)}(t) = -\mu\,(V_1^{(1)}(t) - V_2^{(1)}(t))\,, \tag{G.6}$$

subject to

$$V_1^{(1)}(n) = V_2^{(1)}(n) = 0\,. \tag{G.7}$$

There are many ways of solving these equations and we mention a few:
(1) Differentiate (G.5) and substitute expressions for the first order derivatives from (G.5) and (G.6) to obtain a second order differential equation in $V_1^{(1)}(t)$ subject to conditions $V_1^{(1)}(n) = 0$ and $\frac{d}{dt}V_1^{(1)}(n) = -1$, the latter obtained from $(G.5)$. There are standard methods for this problem, see solution to Exercise No 15 in Exercise paper No 5.
(2) Add (G.5) and (G.6) to obtain

$$\frac{d}{dt}(V_1^{(1)} + V_2^{(1)})(t) = -1\,,$$

subject to $(V_1^{(1)} + V_2^{(1)})(n) = 0$. Solution:

$$V_1^{(1)} + V_2^{(1)}(t) = n - t\,. \tag{G.8}$$

Subtract (G.6) from (G.5) to obtain

$$\frac{d}{dt}(V_1^{(1)} - V_2^{(1)})(t) = -1 - 2\,\mu\,(V_1^{(1)} - V_2^{(1)})(t)\,,$$

subject to $(V_1^{(1)} - V_2^{(1)})(n) = 0$. Solution (we recognize the differential equation and side condition for the reserve on a deterministic annuity of 1 at interest rate $2\,\mu$):

$$V_1^{(1)}(t) - V_2^{(1)}(t) = \int_t^n e^{-2\,\mu(\tau-t)\,d\tau}\,d\tau = \frac{1 - e^{-2\,\mu\,(n-t)}}{2\,\mu}\,.$$

We find

$$V_1^{(1)}(t) = \frac{n-t}{2} + \frac{1 - e^{-2\mu(n-t)}}{4\mu}, \tag{G.9}$$

$$V_2^{(1)}(t) = \frac{n-t}{2} - \frac{1 - e^{-2\mu(n-t)}}{4\mu}.$$

(3) Already before we arrived at relation (G.8) we ought to have realized the following: Given $Z(t) = 1$, $T_1(t,n]$ is the time that remains to spend in the current state. Given $Z(t) = 2$, $T_1(t,n]$ is the time that remains to spend in the other state. Given $Z(t) = 1$, $T_2(t,n] = (n-t) - T_1(t,n]$ is the time that remains to spend in the other state. Due to symmetry, the two last mentioned cases are probabilistically identical, so we can conclude that

$$V_2^{(1)}(t) = (n-t) - V_1^{(1)}(t),$$

which is (G.8). Substituting this into (G.5), we get

$$\frac{d}{dt}V_1^{(1)}(t) = -1 + \mu\left((n-t) - 2\,V_1^{(1)}(t)\right),$$

which is to be solved subject to $V_1^{(1)}(n) = 0$. We easily arrive at the solution above.

Now to non-central second order moments, and we proceed with method (3) using the general differential equation, which specializes to

$$\frac{d}{dt}V_1^{(2)}(t) = \mu\,V_1^{(2)}(t) - 2\,V_1^{(1)}(t) - \mu\,V_2^{(2)}(t),$$

subject to

$$V_1^{(2)}(n) = 0.$$

Using the symmetry again, we realize that

$$V_2^{(2)}(t) = (n-t)^2 - 2\,(n-t)\,V_1^{(1)}(t) + V_1^{(2)}(t).$$

Substituting this and reorganizing a bit, the differential equation above becomes

$$\frac{d}{dt}V_1^{(2)}(t) = 2\left(\mu\,(n-t) - 1\right)V_1^{(1)}(t) - \mu\,(n-t)^2.$$

Inserting the expression (G.9), we get

$$V_1^{(2)}(t) = -\int_t^n \left(2\left(\mu\,(n-\tau) - 1\right)\left(\frac{n-\tau}{2} + \frac{1 - e^{-2\mu(n-\tau)}}{4\mu}\right) - \mu\,(n-\tau)^2\right) d\tau.$$

Substituting $n - \tau$, we are left with the simple task of integrating some standard functions. One should finally arrive at

$$\mathbb{V}\mathrm{ar}[T_1(t,n] \mid Z(t) = 1] = V_1^{(2)}(t) - (V_1^{(1)}(t))^2 = \frac{1}{4\mu}\left(n - t + \frac{1 - \left(2 - e^{-2\mu(n-t)}\right)^2}{4\mu}\right).$$

Exercise 16

(a) Obviously, for $t < s < u$,

$$p_{\overline{aa}}(t,u) = p_{\overline{aa}}(t,s)p_{\overline{aa}}(s,u).$$

Details:

$$\begin{aligned}&\mathbb{P}[Z(\tau) = a,\, \tau \in (t,u] \,|\, Z(t) = a]\\&= \mathbb{P}[Z(\tau) = a\,, \tau \in (t,s] \,|\, Z(t) = a]\mathbb{P}[Z(\tau) = a\,, \tau \in (s,u] \,|\, Z(t) = a\,, Z(\tau) = a\,, \tau \in (t,s]]\,;\end{aligned}$$

The first factor here is $p_{\overline{aa}}(t,s)$, and the second factor is (due to the Markov property) $p_{\overline{aa}}(s,u)$.

For instance forward argument (put u and $u + du$ in the roles of s and u):

$$p_{\overline{aa}}(t,u+du) = p_{\overline{aa}}(t,u)p_{\overline{aa}}(u,u+du) = p_{\overline{aa}}(t,u)(1 - (\mu(u) + \sigma(u))du)\,.$$

From this we get

$$\frac{\partial}{\partial u}p_{\overline{aa}}(t,u) = -p_{\overline{aa}}(t,u)(\mu(u) + \sigma(u))\,.$$

Integrating, using the side condition $p_{\overline{aa}}(t,t) = 1$, we arrive at the answer.

(b) Using the Markov property:

$$p_{aa}(0,t_1)\,\sigma(t_1)\,dt_1\,p_{ii}(t_1+dt_1,t_2)\,\rho(t_2)\,dt_2\,p_{aa}(t_2+dt_2,t_3)\,\mu(t_3)\,dt_3$$

(plus $o(dt_1) + o(dt_2) + o(dt_3)$, strictly speaking). Due to the factors dt_i we can replace the arguments $t_i + dt_i$ appearing here by t_i.

(d) The Kolmogorov forward equations with constant intensities, letting differentiation w.r.t. t be denoted by primes:

$$p'_{aa}(t) = -p_{aa}(t)(\mu + \sigma) + p_{ai}(t)\rho\,, \tag{G.10}$$
$$p'_{ai}(t) = p_{aa}(t)\sigma - p_{ai}(t)(\nu + \rho)\,. \tag{G.11}$$

Side conditions:

$$p_{aa}(0) = 1\,, \qquad p_{ai}(0) = 0\,. \tag{G.12}$$

Differentiate (G.10):

$$p''_{aa}(t) = -p'_{aa}(t)(\mu + \sigma) + p'_{ai}(t)\rho\,. \tag{G.13}$$

Substitute here $p'_{ai}(t)$ from (G.11):

$$p''_{aa}(t) = -p'_{aa}(t)(\mu + \sigma) + (p_{aa}(t)\sigma - p_{ai}(t)(\nu + \rho))\rho\,. \tag{G.14}$$

Now solve $p_{ai}(t)$ from (G.10) and substitute into (G.14) and rearrange a bit to arrive at

$$p''_{aa}(t) + p'_{aa}(t)((\mu + \sigma) + (\nu + \rho)) + p_{aa}(t)((\mu + \sigma)(\nu + \rho) - \sigma\rho) = 0\,. \tag{G.15}$$

This is a simple homogeneous second order ordinary differential equation, which is to be solved subject to the conditions

$$p_{aa}(0) = 1\,, \qquad p'_{aa}(0) = -(\mu + \sigma)\,, \tag{G.16}$$

the latter obtained by setting $t = s$ in (G.10). The general solution to (G.15) - (G.16) is

$$p_{aa}(t) = c_1e^{r_1t} + c_2e^{r_2t}\,,$$

when r_1 and r_2 are distinct solutions to

$$r^2 + r((\mu+\sigma) + (\nu+\rho)) + ((\mu+\sigma)(\nu+\rho) - \sigma\rho) = 0\,.$$

Here c_1 and c_2 are constants that are to be determined so as to match the side conditions (G.16):

$$\begin{aligned} c_1 + c_2 &= 1\,, \\ c_1 r_1 + c_2 r_2 &= -(\mu+\sigma)\,. \end{aligned}$$

If the roots r_1 and r_2 coincide, $r_1 = r_2 = r$ (say), then the general solution is of the form $(c_1 + c_2 t)e^{rt}$. You find

$$\left.\begin{array}{c} r_1 \\ r_2 \end{array}\right\} = -\frac{(\mu+\sigma) + (\nu+\rho) \pm \sqrt{((\mu+\sigma) - (\nu+\rho))^2 + 4\sigma\rho}}{2}$$

Fill in the details yourself.

Exercise 17
Recall the forward equations:

$$\frac{\partial}{\partial t} p_{aa}(0,t) = -p_{aa}(0,t)(\mu_{x+t} + \sigma_{x+t}) + p_{ai}(0,t)\rho_{x+t}\,, \tag{G.17}$$

$$\frac{\partial}{\partial t} p_{ai}(0,t) = p_{aa}(0,t)\sigma_{x+t} - p_{ai}(0,t)(\nu_{x+t} + \rho_{x+t})\,. \tag{G.18}$$

Side conditions

$$p_{aa}(0,0) = 1\,, \qquad p_{ai}(0,0) = 0\,. \tag{G.19}$$

Note the following: The forward differential equations for the transition probabilities are usually easier to work with than the backward differential equations. Under the forward construction we work with the functions $p_{jk}(t,\,\cdot\,)$, $k = 1, 2, ..., J$, for fixed j and t. These functions sum to one and therefore we need only solve $J - 1$ equations. Under the backward construction we work with the functions $p_{jk}(\,\cdot\,, u)$, $j = 1, 2, ..., J$, for fixed k and u, which do not sum to 1 or anything else that could be helpful. If one should never the less want to use the backward equations in the present situation, one would face a quasi-difficulty which is due to the notation: e.g. in ${}_tp_x^{aa}$, which means $p_{aa}(x, x+t)$, x appears in both time variables. To apply the backward construction one must parametrize time as in the general theory, letting starting time and ending time be functionally unrelated.

(a) When $\rho_{x+t} = 0$, (G.17) reduces to

$$\frac{\partial}{\partial t} p_{aa}(0,t) = -p_{aa}(0,t)(\mu_{x+t} + \sigma_{x+t})\,,$$

which subject to the first condition in (G.19) integrates to

$$p_{aa}(0,t) = \exp\left(-\int_0^t (\mu_{x+s} + \sigma_{x+s})\,ds\right)\,,$$

the same as the occupancy probability $p_{\overline{aa}}(0,t)$, of course. The equation (G.18) reduces to

$$\frac{\partial}{\partial t} p_{ai}(0,t) = p_{aa}(0,t)\sigma_{x+t} - p_{ai}(0,t)\,\nu_{x+t}\,,$$

or

$$\frac{\partial}{\partial t} p_{ai}(0,t) + p_{ai}(0,t)\,\nu_{x+t} = p_{aa}(0,t)\sigma_{x+t}\,,$$

where the function on the right is now known. Multiply by integrating factor $\exp\left(\int_0^t \nu_{x+s}\,ds\right)$, integrate from 0 to t, use the second condition in (G.19), and arrive at

$$p_{ai}(0,t) = \int_0^t \exp\left(-\int_0^\tau (\mu_{x+s} + \sigma_{x+s})\,ds\right)\,\sigma_{x+\tau}\,\exp\left(-\int_\tau^t \nu_{x+s}\,ds\right)\,d\tau\,,$$

or

$$p_{ai}(0,t) = \int_0^t {}_\tau p_x^{\overline{aa}}\,\sigma_{x+\tau}\;{}_{t-\tau}p_{x+\tau}^{\overline{ii}}\,d\tau\,.$$

This expression can be read aloud; Under the integral sign is the probability that the policy stays in a until an intermediate time τ, then makes the transfer to state i in the time interval $(\tau, \tau + d\tau]$, and thereafter stays in state i until time t, and the integral sums up the probabilities of these mutually exclusive events. See Section 7.F of BL. The reason why we get explicit expressions here is that there is no return to a state once the policy has left it. It should be made clear, however, that computation of the probabilities goes by numerical solution to the differential equations, which is just as easy with recovery as without. The explicit expressions are useful mainly because they can be directly understood and also because they give us a possibility of discussing how the transition probabilities depend on the intensities.

In the same manner as for $p_{aa}(0,t)$ we find in the case without recovery that

$$p_{ii}(0,t) = \exp\left(-\int_0^t \nu_{x+s}\,ds\right)\,.$$

(b) The conditional probability of being active at time t, given alive at time t (and start as active at time 0), is

$$\begin{aligned}
{}_t\tilde{p}_x^{aa} &= \mathbb{P}[Z(t) = a \,|\, Z(t) \in \{a,i\}] = \frac{\mathbb{P}[Z(t) = a \cap Z(t) \in \{a,i\}]}{\mathbb{P}[Z(t) \in \{a,i\}]} \\
&= \frac{p_{aa}(0,t)}{p_{aa}(0,t) + p_{ai}(0,t)}\,.
\end{aligned}$$

Likewise,

$${}_t\tilde{p}_x^{ai} = \frac{p_{ai}(0,t)}{p_{aa}(0,t) + p_{ai}(0,t)}\,.$$

Differentiating ${}_tp_{[x]} = p_{aa}(0,t) + p_{ai}(0,t)$ and using (G.17) and (G.18), one finds

$$\begin{aligned}
\mu_{[x]+t} &= -\frac{\frac{\partial}{\partial t}\,{}_tp_{[x]}}{{}_tp_{[x]}} \\
&= -\frac{\frac{\partial}{\partial t}p_{aa}(0,t) + \frac{\partial}{\partial t}p_{ai}(0,t)}{p_{aa}(0,t) + p_{ai}(0,t)} \\
&= -\frac{-p_{aa}(0,t)(\mu_{x+t} + \sigma_{x+t}) + p_{ai}(0,t)\rho_{x+t} + p_{aa}(0,t)\sigma_{x+t} - p_{ai}(0,t)(\nu_{x+t} + \rho_{x+t})}{p_{aa}(0,t) + p_{ai}(0,t)} \\
&= \frac{p_{aa}(0,t)\mu_{x+t} + p_{ai}(0,t)\nu_{x+t}}{p_{aa}(0,t) + p_{ai}(0,t)} \\
&= {}_t\tilde{p}_x^{aa}\,\mu_{x+t} + {}_t\tilde{p}_x^{ai}\,\nu_{x+t}\,,
\end{aligned}$$

a weighted average of mortality for active and mortality for invalid, the weights being the conditional probabilities of being active and invalid, respectively, given survival.

The select mechanism is in this case due to the fact that the insured is known to be, not just alive and x years old at time 0, but also in active state.

(c) If $\mu_{x+t} = \nu_{x+t}$ for all $t \geq 0$, then $\mu_{[x]+t} = \mu_{x+t}$, of course, and

$$ {}_tp_{[x]} = \exp\left(-\int_0^t \mu_{x+s}\,ds\right) . $$

We have

$$ {}_t\tilde{p}_x^{aj} = {}_tp_x^{aj} / {}_tp_{[x]} = {}_tp_x^{aj} \exp\left(\int_0^t \mu_{x+s}\,ds\right) , $$

$j = a, i$. Differentiating w.r.t. t, and rearranging a bit, you will see that the ${}_t\tilde{p}_x^{aj}$, $j = a, i$, satisfy the differential equations (G.17) - (G.18) with $\rho = 0$, which are the Kolmogorov forward equations for in the partial model. They also satisfy the side conditions (G.19). Thus, in the case of non-differential mortality, the transition probabilities can be obtained by first determining the transition probabilities in the simpler partial model with only two states, and then multiplying them with the survival probability.

(d) Statements:

```
(* SPECIFY NON-NULL PAYMENTS AT TIME t ! *)
bi[1,3] := 1; bi[2,3] := 1;
ca[1] := 1;

(* SPECIFY MAXIMUM ORDER OF MOMENTS AND NUMBER OF STATES !
*)
q := 1; (*moments*)
JZ := 3; (*number of states of the policy*)

(* SPECIFY TRANSITION INTENSITIES FOR POLICY Z ! Here Danish basis ex-
tended with recovery; States 1 = active, 2 = disabled, 3 = dead:*)
alpha[1,3] := 0.0005;
beta[1,3] := 0.00007585775;
gamma[1,3] := 0.038*ln(10);
alpha[2,3] := 0.0005;
beta[2,3] := 0.00007585775;
gamma[2,3] := 0.038*ln(10);
alpha[1,2] := 0.0004;
beta[1,2] := 0.000003467368;
gamma[1,2] := 0.06*ln(10);
alpha[2,1] := 0.005;

(*SPECIFY AGE x, TERM t, INTEREST RATES AND NON-NULL LIFE ENDOW-
MENTS ! *)
x:= 30; (*age*)
t := 30; (*term*)
r := ln(1+0.045); (*interest rate*)
```

be[1] := 0; be[2] := 0; (*endowments at term of contract*)

(*SPECIFY LUMP SUM PREMIUM AT TIME 0: PUT c0 := 1 IF ALL OTHER PREMIUMS ARE 0 AND ONLY MOMENTS OF BENEFITS ARE WANTED ! *)
c0 := 0; b0 := 0;

(*SPECIFY NUMBER OF STEPS IN RUNGE-KUTTA (OPTIONAL) ! *)
steps := 3000; (*number of steps*)
h := t/steps; (*steplength*)
count := 30; (*number of output times*)
count := steps/count - 0.5;

Result:

state:	1	2
time:	30.00	
NC 1:	0.000000	0.000000
time:	29.00	
NC 1:	0.009740	0.013915
time:	28.00	
NC 1:	0.017861	0.025918
time:	27.00	
NC 1:	0.024547	0.036236
.....	.....	.....
time:	15.00	
NC 1:	0.038582	0.081542
.....	.....	.....
time:	1.00	
NC 1:	0.002763	0.066309
time:	0.00	
NC 1:	0.000000	0.064598

pi = 0.004336

Exercise 18
The process must start somewhere, so let us say $Z(0) = 1$.

(a)

$$\mathbb{P}\left[\bigcap_{i=0}^{r+1} Z(t_i) = j_i\right] \quad = \quad \mathbb{P}\left[\bigcap_{i=0}^{r} Z(t_i) = j_i\right] \mathbb{P}\left[Z(t_{r+1}) = j_{r+1} \,\middle|\, \bigcap_{i=0}^{r} Z(t_i) = j_i\right]$$

$$= \quad \mathbb{P}\left[\bigcap_{i=0}^{r} Z(t_i) = j_i\right] p_{j_r,j_{r+1}}(t_r, t_{r+1})\,.$$

We have here used the Markov property of the process. Repeating this, we obtain

$$\mathbb{P}\left[\bigcap_{i=0}^{r+1} Z(t_i) = j_i\right] \quad = \quad p_{1j_0}(0, t_0) \prod_{i=1}^{r+1} p_{j_{i-1},j_i}(t_{i-1}, t_i)\,. \tag{G.20}$$

Next, using (G.20),

$$\begin{aligned}
\mathbb{P}\left[\bigcap_{i=1}^{r} Z(t_i) = j_i \,\middle|\, Z(t_0) = j_0,\ Z(t_{r+1}) = j_{r+1}\right] &= \frac{\mathbb{P}\left[\bigcap_{i=0}^{r+1} Z(t_i) = j_i\right]}{\mathbb{P}\left[Z(t_0) = j_0\,,\ Z(t_{r+1}) = j_{r+1}\right]} \\
&= \frac{p_{1j_0}(0, t_0) \prod_{i=1}^{r+1} p_{j_{i-1},j_i}(t_{i-1}, t_i)}{p_{1j_0}(0, t_0)\, p_{j_0,j_{r+1}}(t_0, t_{r+1})} \\
&= \frac{\prod_{i=1}^{r+1} p_{j_{i-1},j_i}(t_{i-1}, t_i)}{p_{j_0,j_{r+1}}(t_0, t_{r+1})}\,.
\end{aligned} \tag{G.21}$$

(b) For $s = t_0 < t_1 < \cdots < t_r < t_{r+1} = t$:

$$\begin{aligned}
&\mathbb{P}\left[Z(t_r) = j_r \,\middle|\, \bigcap_{i=1}^{r-1} Z(t_i) = j_i\,, Z(s) = i,\ Z(t) = j\right] \\
&= \mathbb{P}\left[Z(t_r) = j_r \,\middle|\, \bigcap_{i=1}^{r-1} Z(t_i) = j_i\,, Z(t_0) = j_0,\ Z(t_{r+1}) = j_{r+1}\right] \\
&= \frac{\mathbb{P}\left[\bigcap_{i=0}^{r+1} Z(t_i) = j_i\right]}{\mathbb{P}\left[\bigcap_{i=0,\ldots,r-1,r+1} Z(t_i) = j_i\right]} \\
&= \frac{p_{1j_0}(0, t_0) \prod_{i=1}^{r+1} p_{j_{i-1},j_i}(t_{i-1}, t_i)}{p_{1j_0}(0, t_0) \prod_{i=1,\ldots,r-1} p_{j_{i-1},j_i}(t_{i-1}, t_i)\, p_{j_{r-1},j_{r+1}}(t_{r-1}, t_{r+1})} \\
&= \frac{p_{j_{r-1},j_r}(t_{r-1}, t_r)\, p_{j_r,j}(t_r, t)}{p_{j_{r-1},j}(t_{r-1}, t)}\,.
\end{aligned} \tag{G.22}$$

For given $t_r = t$, $j_{r+1} = j$ (and $t_0 = s$, $j_0 = i$) this is just a function $\tilde{p}_{j_{r-1},j_r}(t_{r-1}, t_r)$ (say) of j_{r-1}, j_r, t_{r-1}, and t_r, showing that the conditional Markov chain is itself Markov.

The intensities of the conditional Markov chain are

$$\begin{aligned}
\tilde{\mu}_{gh}(\tau) &= \lim_{d\tau \searrow 0} \frac{\tilde{p}_{gh}(\tau, \tau + d\tau)}{d\tau} \\
&= \lim_{d\tau \searrow 0} \frac{p_{gh}(\tau, \tau + d\tau)\, p_{hj}(\tau + d\tau, t)}{d\tau\, p_{gj}(\tau + d\tau, t)} \\
&= \mu_{gh}(\tau)\, \frac{p_{hj}(\tau, t)}{p_{gj}(\tau, t)}\,.
\end{aligned} \tag{G.23}$$

$$\lim_{\tau \nearrow t} \tilde{\mu}_{gh}(\tau) = \begin{cases} \mu_{gh}(t)\, \frac{\mu_{hj}(t)}{\mu_{gj}(t)}\,, & g \neq j,\ h \neq j\,, \\ \infty\,, & g \neq j,\ h = j\,, \\ 0\,, & g = j,\ h \neq j\,. \end{cases}$$

The expression in the case $g \neq j$, $h \neq j$ is obtained upon writing

$$\frac{p_{hj}(\tau,t)}{p_{gj}(\tau,t)} = \frac{p_{hj}(\tau,t)/(t-\tau)}{p_{gj}(\tau,t)/(t-\tau)}$$

before taking the limit.

Think about these results - they are quite natural.

(e)

$$\tilde{\sigma}(\tau) = \sigma(\tau)\frac{p_{ii}(\tau,t)}{p_{ai}(\tau,t)}\,,$$

$$\tilde{\mu}(\tau) = \mu(\tau)\frac{p_{di}(\tau,t)}{p_{ai}(\tau,t)} = 0\,.$$

Constant intensities and no recovery:

$$\begin{aligned}\tilde{\sigma}(\tau) &= \sigma\frac{e^{-\nu\,(t-\tau)}}{\int_\tau^t e^{-(\mu+\sigma)(s-\tau)}\,\sigma\, e^{-\nu\,(t-s)}\,ds} = \frac{1}{\int_\tau^t e^{(\nu-\mu-\sigma)(s-\tau)}\,ds} \\ &= \begin{cases}\frac{\nu-\mu-\sigma}{e^{(\nu-\mu-\sigma)(t-\tau)}-1}\,, & \nu-\mu-\sigma \neq 0\,, \\ \frac{1}{t-\tau}\,, & \nu-\mu-\sigma = 0\,.\end{cases}\end{aligned}$$

Exercise 19

(a) An n-year endowment insurance with sum insured 1 against level premium c per time unit, continuous time, age of insured upon issue of contract is x.

(b) Integration is book-work, see lecture notes. One obtains

$$\begin{aligned}V_t &= \int_t^n e^{-\int_t^\tau (r+\mu_{x+s})\,ds}(\mu_{x+\tau}-c)\,d\tau \,+\, e^{-\int_t^n (r+\mu_{x+s})\,ds} \\ &= \bar{A}_{x+t\,\overline{n-t|}} \,-\, c\,\bar{a}_{x+t\,\overline{n-t|}}\,.\end{aligned}$$

Determine c by equivalence requirement $V_0 = 0$ (no down payment at time0):

$$c \,=\, \frac{\bar{A}_{x\,\overline{n|}}}{\bar{a}_{x\,\overline{n|}}}\,.$$

(b) One must first prove

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

which is book-work (there are several ways). See lecture notes.

Then write

$$V_t = 1 \,-\, r\,\bar{a}_{x+t\,\overline{n-t|}} \,-\, c\,\bar{a}_{x+t\,\overline{n-t|}} \,=\, 1 \,-\, (r\,+\,c)\,\bar{a}_{x+t\,\overline{n-t|}}\,.$$

Thus, the problem reduces to proving that if μ_{x+t} is an increasing function of t, then $\bar{a}_{x+t\,\overline{n-t|}}$ is a decreasing function of t. This is book-work, see lecture notes.

To construct an example where V_t is not an increasing function, i.e $\bar{a}_{x+t\,\overline{n-t|}}$ is not a decreasing function, we must obviously take a μ_{x+t} that is not increasing. Fix $t < n$ and write

$$\begin{aligned}\bar{a}_{x\,\overline{n|}} &= \int_0^n e^{-\int_0^\tau (r+\mu_{x+s})\,ds}\,d\tau \\ &= \int_0^t e^{-\int_0^\tau (r+\mu_{x+s})\,ds}\,d\tau \;+\; e^{-\int_0^t (r+\mu_{x+s})\,ds}\int_t^n e^{-\int_t^\tau (r+\mu_{x+s})\,ds}\,d\tau \\ &= \bar{a}_{x\,\overline{t|}} \;+\; {}_tE_x \bar{a}_{x+t\,\overline{n-t|}}\,.\end{aligned}$$

Keep μ_{x+s} fixed for $t \le s \le n$. By increasing μ_{x+s} for $0 \le s \le t$, we can obviously make $\bar{a}_{x\,\overline{t|}}$ and ${}_tE_x$ arbitrarily small. In particular we can arrange that $\bar{a}_{x\,\overline{t|}} \;<\; 0.5\,\bar{a}_{x+t\,\overline{n-t|}}$ and ${}_tE_x < 0.5$, hence $\bar{a}_{x\,\overline{n|}} < \bar{a}_{x+t\,\overline{n-t|}}$.

Exercise 20

(a) Book-work. See lecture notes.

(b) ${}_tp_0$ is a survival function since it is decreasing and ${}_0p_0 = 1$, ${}_\infty p_0 = 0$. The mortality intensity is

$$\mu_t = -\frac{\frac{d}{dt}\,{}_tp_0}{{}_tp_0} = \cdots = \frac{\gamma w(t)}{\delta + W(t)}\,. \tag{G.24}$$

$${}_tp_x = \frac{{}_{x+t}p_0}{{}_xp_0} = \left(\frac{\delta + W(x)}{\delta + W(x+t)}\right)^\gamma = \left(\frac{\delta + W(x)}{(\delta + W(x)) + (W(x+t) - W(x))}\right)^\gamma\,.$$

This is of the same form as ${}_tp_0$, only with $W(t)$ and δ replaced by $W(x+t) - W(x)$ and $\delta + W(x)$, respectively.

$$\mathbb{P}[W(T) + \delta > x] = \mathbb{P}[T > W^{-1}(x-\delta)] = \left(\frac{\delta}{W(W^{-1}(x-\delta)) + \delta}\right)^\gamma = \left(\frac{\delta}{x}\right)^\gamma\,,$$

$x > \delta$, a Pareto distribution with shape parameter γ and truncation parameter δ.

(c) Use (F.5) with $G(t) = t^2$, hence $dG(t) = 2\,t\,dt$

$$\mathbb{E}[T^2] = \int_0^\infty 2\,t\,\frac{1}{(1+t^2/\delta)^\gamma}\,dt\,.$$

Substitute $u = 1 + t^2/\delta$. We have $du = 2\,t/\delta\,dt$ and u is an increasing function of t, $u = 1$ for $t = 0$, and $u = \infty$ for $t = \infty$. Thus

$$\mathbb{E}[T^2] = \delta\int_1^\infty u^{-\gamma}\,du \;=\; \frac{\delta}{\gamma - 1}\,.$$

One could also observe directly that

$$\begin{aligned}\int_0^\infty 2\,t\,\left(1+t^2/\delta\right)^{-\gamma}\,dt &= \frac{\delta}{-\gamma+1}\int_0^\infty d\,\left(1+t^2/\delta\right)^{-\gamma+1} \\ &= \frac{\delta}{-\gamma+1}\left[\left(1+\infty^2/\delta\right)^{-\gamma+1} \;-\; \left(1+0^2/\delta\right)^{-\gamma+1}\right] \\ &= \frac{\delta}{1-\gamma}\,.\end{aligned}$$

(d) Recall (G.24). We need

$$\ln \mu_t = \ln \gamma + \ln w(t) - \ln(\delta + W(t))$$

and

$$\int_0^{T\wedge z} \mu_t \, dt = \int_0^{T\wedge z} \gamma \frac{w(t)}{\delta + W(t)} \, dt = \gamma \left(\ln(\delta + W(T \wedge z)) - \ln \delta\right).$$

Log likelihood is

$$\begin{aligned} \ln \Lambda &= \sum_{m;\, T_m < z_m} [\ln \gamma + \ln w(T_m) - \ln(\delta + W(T_m))] \\ &\quad - \gamma \sum_m [\ln(\delta + W(T_m \wedge z)) - \ln \delta]. \end{aligned}$$

First order derivatives:

$$\frac{\partial}{\partial \gamma} \ln \Lambda = \sum_{m;\, T_m < z_m} \gamma^{-1} - \sum_m [\ln(\delta + W(T_m \wedge z)) - \ln \delta],$$

$$\frac{\partial}{\partial \delta} \ln \Lambda = - \sum_{m;\, T_m < z_m} \frac{1}{\delta + W(T_m)} - \gamma \sum_m \left(\frac{1}{\delta + W(T_m \wedge z)} - \frac{1}{\delta} \right).$$

Put $N(z) = \sharp\{m;\, T_m < z\}$, the number of deaths within age z. The ML equations are

$$\hat{\gamma} = \frac{N(z)}{\sum_m \ln(1 + W(T_m \wedge z)/\hat{\delta})}, \tag{G.25}$$

$$\sum_{m;\, T_m < z_m} \frac{1}{\hat{\delta} + W(T_m)} + \hat{\gamma} \sum_m \left(\frac{1}{\hat{\delta} + W(T_m \wedge z)} - \frac{1}{\hat{\delta}} \right) = 0. \tag{G.26}$$

These equations have to be solved numerically. Insert the expression (G.25) for $\hat{\gamma}$ into (G.26) and solve the latter w.r.t. $\hat{\delta}$. Substitute the solution into (G.25) to find $\hat{\gamma}$.

To find the asymptotic variance matrix, form the second derivatives $\frac{\partial^2}{\partial \gamma^2} \ln \Lambda$, $\frac{\partial^2}{\partial \delta^2} \ln \Lambda$, $\frac{\partial^2}{\partial \gamma \partial \delta} \ln \Lambda$, find their expected values, change sign, and invert the information matrix.

(e) If δ is known, then (G.25) with $\hat{\delta}$ replaced by δ is an explicit expression for $\hat{\gamma}$. Divide by n in denominator and numerator, let n go to ∞, and use the law of large numbers to conclude that

$$\hat{\gamma} \to \frac{\mathbb{P}[T_1 \leq z]}{\mathbb{E}[\ln(1 + W(T_1 \wedge z)/\delta)]}. \tag{G.27}$$

From (F.6) we have

$$\mathbb{P}[T_1 \leq z] = 1 - {}_t p_0 = 1 - \left(\frac{\delta}{W(z) + \delta} \right)^{\gamma}.$$

Use (F.5) with ${}_t p_0$ given by (F.6) and $G(t) = \ln(1 + W(t \wedge z)/\delta)$, for which $G(0) = 0$ and

$$dG(t) = \begin{cases} \frac{w(t)}{\delta + W(t)} \, dt, & 0 < t < z, \\ 0 & t > z. \end{cases}$$

We find

$$\begin{aligned}
\mathbb{E}[\ln(1+W(T_1\wedge z)/\delta)] &= \int_0^z \left(\frac{\delta}{W(t)+\delta}\right)^{\gamma}\frac{w(t)}{\delta+W(t)}\,dt \\
&= \delta^{\gamma}\int_0^z \frac{w(t)}{(W(t)+\delta)^{\gamma+1}}\,dt \\
&= \delta^{\gamma}\gamma\left(\frac{1}{(W(0)+\delta)^{\gamma}} - \frac{1}{(W(z)+\delta)^{\gamma}}\right) \\
&= \gamma\left(1-\left(\frac{\delta}{W(z)+\delta}\right)^{\gamma}\right)
\end{aligned}$$

Inserting these expressions in (G.27), we find that the limit is γ.

(f) OE rate in year j is

$$\hat{\mu} = \frac{N_j}{W_j}\,,$$

where $N_j = \sharp\{m;\, j-1 < T_m \le j\}$ and $W_j = \sum_m ((T_m\wedge j)-(j-1))\vee 0$. Choose representative age $\tau_j \in (j-1, j]$ and weights w_j and minimize

$$Q = \sum_{j=1}^n w_j\,(\mu(\tau_j;\,\gamma,\delta) - \hat{\mu}_j)^2 = \sum_{j=1}^n \left(\frac{\gamma w(\tau_j)}{\delta+W(\tau_j)} - \hat{\mu}_j\right)^2.$$

Minimize by forming the partial derivatives,

$$\frac{\partial}{\partial\gamma}Q = \sum_{j=1}^n w_j\,2\left(\frac{\gamma w(\tau_j)}{\delta+W(\tau_j)} - \hat{\mu}_j\right)\frac{w(\tau_j)}{\delta+W(\tau_j)}\,,$$

$$\frac{\partial}{\partial\delta}Q = \sum_{j=1}^n w_j\,2\left(\frac{\gamma w(\tau_j)}{\delta+W(\tau_j)} - -\hat{\mu}_j\right)\frac{-\gamma\,w(\tau_j)}{(\delta+W(\tau_j))^2}\,,$$

setting them equal to zero and solving for γ^* and δ^*

Exercise 21

(a) Direct backward argument, conditioning on what happens in the small time interval $(t, t+dt)$ and splitting payments in $(t, z]$ into payments in $(t, t+dt]$ and payments in $(t+dt, z]$:

$$\begin{aligned}
V_a(t) \;=\;& (1-\mu_{x+t}\,dt - \sigma_{x+t}\,dt)(-c\,dt + e^{-r\,dt}\,V_a(t+dt)) \\
&+ \sigma_{x+t}\,dt\,(O(dt) + e^{-r\,dt}\,V_i(t+dt)) + \mu_{x+t}\,dt\,O(dt) + o(dt)\,,
\end{aligned}$$

where $O(dt)$ is of order dt and $o(dt)$ is of order less than dt (i.e $o(dt)/dt \to 0$ as $dt \to 0$). Insert (Taylor expansion to first order)

$$e^{-r\,dt} = 1 - r\,dt + o(dt)\,,$$

$$V_j(t+dt) = V_j(t) + \frac{d}{dt}V_j(t)\,dt + o(dt)\,,\;\; j = a, i,$$

and collect all terms of order dt^2 or less in $o(dt)$:

$$V_a(t) \;=\; -\,c\,dt + (1-\mu_{x+t}\,dt-\sigma_{x+t}\,dt - r\,dt)\,V_a(t) + \frac{d}{dt}V_a(t)\,dt + \sigma_{x+t}\,dt\,V_i(t) + o(dt)\,.$$

Subtract $V_a(t)$ on both sides, divide by dt and let dt go to zero, to arrive at

$$\frac{d}{dt}V_a(t) \;=\; rV_a(t) + c - \sigma_{x+t}(V_i(t) - V_a(t)) + \mu_{x+t}V_a(t)\,.$$

Similarly

$$\frac{d}{dt}V_i(t) \;=\; r\,V_i(t) - b + \nu_{x+t}V_i(t)\,.$$

Side conditions $V_a(z) = V_i(z) = 0$.

(b) The differential equations now become

$$\frac{d}{dt}V_a(t) \;=\; rV_a(t) + c - \sigma_{x+t}(V_i(t) - V_a(t))\,.$$

$$\frac{d}{dt}V_i(t) \;=\; r\,V_i(t) - b + \nu_{x+t}\,V_i(t))\,.$$

Side conditions remain the same.

(c) Same product as in (a) plus an endowment insurance with sum 1.

(d)

$$\begin{aligned}
p_{aa}(s,t) &= e^{-\int_s^t(\mu_{x+u}+\sigma_{x+u})\,du}\,,\\
p_{ii}(s,t) &= e^{-\int_s^t \nu_{x+u}\,du}\,,\\
p_{ai}(s,t) &= \int_s^t e^{-\int_s^\tau(\mu_{x+u}+\sigma_{x+u})\,du}\sigma_{x+\tau}e^{-\int_\tau^t \nu_{x+u}\,du}\,d\tau\,,\\
p_{ad}(s,t) &= 1 \,-\, p_{aa}(s,t) \,-\, p_{ai}(s,t)\,.
\end{aligned}$$

(e) As usual, let $I_j(t)$ and $N_{jk}(d)$ denote, respectively, the number of policies in state j at time t and the total number of transitions $j \to k$ up to and including time t. The log likelihood function is

$$\begin{aligned}
\ln\Lambda \;=\; & \int_0^z \ln\sigma\,dN_{ai}(t) + \int_0^z \ln(\alpha+\beta e^{\gamma(x+t)})\,dN_{ad}(t) + \int_0^z \ln(\alpha+\beta e^{\gamma'(x+t)})\,dN_{id}(t)\\
& - \int_0^z (\alpha+\beta e^{\gamma(x+t)}+\sigma)\,I_a(t)\,dt - \int_0^z (\alpha+\beta e^{\gamma'(x+t)})\,I_i(t)\,dt\,.
\end{aligned}$$

1st derivatives:

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\ln\Lambda &= \int_0^z \frac{1}{\alpha+\beta e^{\gamma(x+t)}}\,dN_{ad}(t) + \int_0^z \frac{1}{\alpha+\beta e^{\gamma'(x+t)}}\,dN_{id}(t)\\
&\quad - \int_0^z (I_a(t)+I_i(t))\,dt\,,\\
\frac{\partial}{\partial\beta}\ln\Lambda &= \int_0^z \frac{e^{\gamma(x+t)}}{\alpha+\beta e^{\gamma(x+t)}}\,dN_{ad}(t) + \int_0^z \frac{e^{\gamma'(x+t)}}{\alpha+\beta e^{\gamma'(x+t)}}\,dN_{id}(t)\\
&\quad - \int_0^z (e^{\gamma(x+t)}I_a(t) + e^{\gamma'(x+t)}I_i(t))\,dt\,,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial\gamma}\ln\Lambda &= \int_0^z \frac{\beta e^{\gamma(x+t)}(x+t)}{\alpha+\beta e^{\gamma(x+t)}}\,dN_{ad}(t) - \int_0^z \beta e^{\gamma(x+t)}(x+t)I_a(t)\,dt\,,\\
\frac{\partial}{\partial\gamma'}\ln\Lambda &= \int_0^z \frac{\beta e^{\gamma'(x+t)}(x+t)}{\alpha+\beta e^{\gamma'(x+t)}}\,dN_{id}(t) - \int_0^z \beta e^{\gamma'(x+t)}(x+t)I_i(t)\,dt\,,\\
\frac{\partial}{\partial\sigma}\ln\Lambda &= \int_0^z \frac{1}{\sigma}\,dN_{ai}(t) - \int_0^z I_a(t)\,dt\,.
\end{aligned}$$

Set equal to 0 and solve for the ML estimators.

Exercise 23

Assume that X is non-decreasing. Note that

$$\mathcal{D}(X) = \cup_{m=1}^{\infty}\cup_{n=1}^{\infty}\mathcal{D}_{m,n}\,,$$

where $\mathcal{D}_{m,n} = \{t\,;\,t\le n\,,\,X_{t+}-X_{t-}\ge\frac{1}{m}\}$. Since

$$\infty > X_n \ge X_0 + \sum_{t\in\mathcal{D}_{m,n}}(X_{t+}-X_{t-}) > X_0 + \frac{1}{m}\,\sharp\mathcal{D}_{m,n}\,,$$

we conclude that $\sharp\mathcal{D}_{m,n}$ is finite. Thus, being a countable union of finite sets, $\mathcal{D}(X)$ is (at most) countable.

Exercise 25

$$dX_t^q = qX_t^{q-1}\,x_t^c\,dt + \left((X_{t-}+x_t^d)^q - X_{t-}^q\right)\,dN_t\,.$$

Exercise 26

(See also Exercise 56. Here notation is changed: a and b are called α and β)

$$\begin{aligned}
dS(t) &= e^{\alpha t+\beta N(t)}\alpha dt + dN(t)\left(e^{\alpha t+\beta(N(t-)+1)} - e^{\alpha t+\beta N(t-)}\right)\\
&= e^{\alpha t+\beta N(t-)}\alpha dt + dN(t)e^{\alpha t+\beta N(t-)}\left(e^{\beta}-1\right)\\
&= S(t-)\left(\alpha\,dt + \left(e^{\beta}-1\right)dN(t)\right)\\
&= S(t-)\left(\alpha+\lambda\left(e^{\beta}-1\right)\right)\,dt + S(t-)\left(e^{\beta}-1\right)dM(t)\,,
\end{aligned}$$

where $M(t) = N(t)-\lambda t$, a so-called martingale (a process with conditionally zero mean and uncorrelated increments, here actually independent increments).

The rest is straightforward .

Exercise 34

Balance equation

$$\sum_{j=0}^{14}(1+i)^{24-j}c\,\ell_{55+j} - \sum_{j=15}^{24}(1+i)^{24-j}b\,\ell_{55+j} = 0$$

gives

$$c = \frac{\sum_{j=15}^{24}(1+i)^{-j}\,\ell_{55+j}}{\sum_{j=0}^{14}(1+i)^{-j}\,\ell_{55+j}} = 0.381\,.$$

Exercise 50

(a) Kolmogorov forward, using the obvious $p_{aa}(t,u) = 1 - p_{ai}(t,u)$:

$$p_{ai}(t,u+du) = p_{ai}(t,u)(1-\rho\,du) + (1-p_{ai}(t,u))\sigma\,du,$$

$$\frac{\partial}{\partial u}p_{ai}(t,u) + p_{ai}(t,u)(\rho+\sigma) = \sigma$$

$$\frac{\partial}{\partial u}\left(p_{ai}(t,u)e^{(\rho+\sigma)u}\right) = \sigma e^{(\rho+\sigma)u}$$

Integrate from t to u using $p_{ai}(u,u) = 0$ to find

$$p_{ai}(t,u)e^{(\rho+\sigma)u} = \frac{\sigma}{\rho+\sigma}\left(e^{(\rho+\sigma)u} - e^{(\rho+\sigma)t}\right)\,,$$

hence the claimed formula. It depends only on $u-t$ due to homogeneity.

(b) Either direct integration

$$\pi_n = \int_0^n p_{ai}(0,\tau)\,d\tau = \frac{\sigma}{\rho+\sigma}\left(n - \frac{1-e^{-(\rho+\sigma)n}}{\rho+\sigma}\right)\,,$$

or use Thiele's differential equations.

Single premium per unit of time insured is

$$\frac{\pi_n}{n} = \frac{\sigma}{\rho+\sigma}\left(1 - \frac{1-e^{-(\rho+\sigma)n}}{(\rho+\sigma)n}\right)\,.$$

The function

$$g(x) = \frac{1-e^{-x}}{x} \tag{G.28}$$

has derivative

$$g'(x) = \frac{e^{-x}}{x} - \frac{1-e^{-x}}{x^2} = e^{-x}\frac{1+x-e^x}{x^2}\,,$$

which is < 0. It follows that, as $n \nearrow +\infty$,

$$\frac{\pi_n}{n} \nearrow \frac{\sigma}{\rho+\sigma}\,.$$

Reasonable: Increasing function of σ, decreasing function of ρ.

(c) Likelihood

$$\Lambda = \sigma^{N_{ai}}\rho^{N_{ia}}e^{-\sigma W_a - \rho W_i}\,,$$

where $W_a = \sum_{\ell=1}^m T_n^{(m)}$, the total time spent in active state (and $W_i = nm - W_a$ the total time spent in inactive state).

$$\frac{\partial}{\partial\sigma}\ln\Lambda = \frac{N_{ai}}{\sigma} - W_a\,,$$

$$\frac{\partial^2}{\partial\sigma^2}\ln\Lambda = -\frac{N_{ai}}{\sigma^2}\,,$$

plus similar expressions for derivatives w.r.t. ρ, and

$$\frac{\partial^2}{\partial\sigma\partial\rho}\ln\Lambda = 0\,.$$

Thus, the ML estimators are the occurrence-exposure rates

$$\hat{\sigma} = \frac{N_{ai}}{W_a}\,, \qquad \hat{\rho} = \frac{N_{ia}}{W_i}\,,$$

which are asymptotically independent, normally distributed and unbiased, and

$$\text{as.Var}\hat{\sigma} = \frac{\sigma^2}{\text{E}N_{ai}} = \frac{\sigma}{mn(1-\pi_n/n)}\,,$$

where we have used

$$\text{E}N_{ai} = m\int_0^n p_{aa}(0,\tau)\sigma\,d\tau = m\sigma\int_0^n (1-p_{ai}(0,\tau))\,d\tau = mn\sigma(1-\frac{\pi_n}{n})\,,$$

see item (b).

For fixed $w = mn$ (i.e. $m = w/n$)

$$\text{as.Var}\hat{\sigma} = \frac{\sigma}{w(1-\frac{\pi_n}{n})}\,,$$

which is an increasing function of the time span n by the results in item (b).

We also find

$$\text{as.Var}\hat{\rho} = \frac{\rho^2}{\text{E}N_{ia}} = \frac{\rho}{w\frac{\pi_n}{n}}\,,$$

where we have used

$$\text{E}N_{ia} = m\int_0^n p_{ai}(0,\tau)\rho\,d\tau = mn\rho\frac{\pi_n}{n}\,.$$

We see that as.Var$\hat{\rho}$ is a decreasing function of the time span n for fixed w.

Comment: The asymptotic variance of an intensity estimator is better the longer the total expected time spent in the state from which the the relevant transition is made. All policies start from state a at time 0. For fixed total exposure the estimation of σ will be good for many policies observed over a short time (when they are likely to remain active), and estimation of ρ will be good for few policies observed over a long time when they can make it to inactive state.

(d) We find

$$\begin{aligned}
\tilde{p}_{jk}(t,u) &= \frac{\text{P}[Z(s_1)=j_1,\ldots,Z(s_r)=j_r,\,Z(t)=j,\,Z(u)=k,\,Z(n)=i]}{\text{P}[Z(s_1)=j_1,\ldots,Z(s_r)=j_r,\,Z(t)=j,\,Z(n)=i]} \\
&= \frac{p_{aj_1}(0,s_1)\cdots p_{j_{r-1}j_r}(s_{r-1},s_r)p_{j_rj}(s_r,t)p_{jk}(t,u)p_{ki}(u,n)}{p_{aj_1}(0,s_1)\cdots p_{j_{r-1}j_r}(s_{r-1},s_r)p_{j_rj}(s_r,t)p_{ji}(t,n)} \\
&= \frac{p_{jk}(t,u)p_{ki}(u,n)}{p_{ji}(t,n)}\,.
\end{aligned}$$

Corresponding intensities:

$$\tilde{\sigma}(t) = \lim_{u \searrow t} \frac{p_{ai}(t,u)}{u-t} \frac{p_{ii}(u,n)}{p_{ai}(t,n)} = \sigma \frac{p_{ii}(t,n)}{p_{ai}(t,n)}$$

and, similarly,

$$\tilde{\rho}(t) = \rho \frac{p_{ai}(t,n)}{p_{ii}(t,n)} \,,$$

with $p_{ai}(t,n)$ and $p_{ii}(t,n)$ given by expressions in item (a). Full expression for $\tilde{\sigma}(t)$ is useful later:

$$\tilde{\sigma}(t) = \sigma \frac{1 - \frac{\rho}{\sigma+\rho}\left(1 - e^{-(\sigma+\rho)(n-t)}\right)}{\frac{\sigma}{\sigma+\rho}\left(1 - e^{-(\sigma+\rho)(n-t)}\right)} = \frac{\sigma + \rho e^{-(\sigma+\rho)(n-t)}}{1 - e^{-(\sigma+\rho)(n-t)}} \,.$$

(e) When $\rho = 0$,

$$\tilde{\sigma}(t) = \frac{\sigma}{1 - e^{-\sigma(n-t)}} \,.$$

Let the occurrence-exposure rate in year $j = 1, \ldots, n$ be

$$\hat{\sigma}_j = \frac{N_{ai;j}}{W_{a;j}} \,,$$

where $N_{ai;j}$ and $W_{a;j}$ are the number of claims and the total time at risk in year j (time interval $[j-1, j)$). Graduation means estimating σ by minimizing the weighted squares differences between the observed rates $\hat{\sigma}_j$ and the theoretical intensities $\tilde{\sigma}(j-0.5)$, with weights equal to the inverse estimated variances $\widehat{\mathrm{Var}}\hat{\sigma}_j = N_{ai;j}/W^2_{a;j}$:

$$\sum_{j=1}^{n} \frac{W^2_{a;j}}{N_{ai;j}} \left(\hat{\sigma}_j - \frac{\sigma}{1 - e^{-\sigma(n-j+0.5)}} \right)^2 \,.$$

Exercise 51

(a)

$$p^{(1)}_{ai}(0, t+dt) = p^{(1)}_{ai}(0,t)(1 - (\nu(t) + \rho(t))\,dt) \;+\; p_{\overline{aa}}(0,t)\,\sigma(t)\,dt$$

leads to

$$\frac{d}{dt} p^{(1)}_{ai}(0,t) = -p^{(1)}_{ai}(0,t)(\nu(t) + \rho(t)) \;+\; p_{\overline{aa}}(0,t)\,\sigma(t) \,,$$

with side condition

$$p^{(1)}_{ai}(0,0) = 0 \,.$$

(Integrating gives the following integral expression, which could be put up by direct reasoning:

$$p^{(1)}_{ai}(0,t) = \int_0^t \exp(-\int_0^s (\mu+\sigma))\sigma(s)\,ds\,\exp(-\int_s^t (\nu+\rho))\,.)$$

Next,

$$p^{(1)}_{aa}(0, t+dt) = p^{(1)}_{aa}(0,t)(1 - (\mu(t) + \sigma(t))\,dt) \;+\; p^{(1)}_{ai}(0,t)\,\rho(t)\,dt$$

leads to

$$\frac{d}{dt} p^{(1)}_{aa}(0,t) = -p^{(1)}_{aa}(0,t)(\mu(t) + \sigma(t)) \;+\; p^{(1)}_{ai}(0,t)\,\rho(t) \,,$$

with side condition

$$p_{aa}^{(1)}(0,0) = 0, .$$

(Integral expression, which could be put up by direct reasoning:

$$p_{aa}^{(1)}(0,t) = \int_0^t p_{ai}^{(1)}(0,s)\,\rho(s)\,ds\;\exp(-\int_s^t(\mu+\sigma))\,.)$$

Repeating the argument for $k = 2, 3, ...$:

$$\frac{d}{dt}p_{ai}^{(k)}(0,t) = -p_{ai}^{(k)}(0,t)(\nu(t)+\rho(t)) \;+\; p_{aa}^{(k-1)}(0,t)\,\sigma(t)\,,$$

$$p_{ai}^{(k)}(0,0) = 0\,,$$

and

$$\frac{d}{dt}p_{aa}^{(k)}(0,t) = -p_{aa}^{(k)}(0,t)(\mu(t)+\sigma(t)) \;+\; p_{ai}^{(k)}(0,t)\,\rho(t)\,,$$

$$p_{aa}^{(k)}(0,0) = 0\,.$$

Introducing $p_{aa}^{(0)}(0,t) = p_{\overline{aa}}(0,t)$ would save work.

Obviously,

$$\sum_{k=1}^{\infty} p_{ai}^{(k)}(0,t) = p_{ai}(0,t)\,.$$

(b)

$$P[Z(\tau) = i; \tau \in [t-q,t] \mid Z(0) = a] = p_{ai}(0,t-q)p_{\overline{ii}}(t-q,t)\,.$$

(c)

$$\pi = \int_q^n e^{-r\tau}p_{ai}(0,\tau-q)p_{\overline{ii}}(\tau-q,\tau)\,d\tau\,.$$

Reserve: $\int_t^{t+q} e^{-r(\tau-t)}p_{\overline{ii}}(t,\tau)\,d\tau \;+\; \int_{t+q}^n e^{-r(\tau-t)}p_{ii}(t,\tau-q)p_{\overline{ii}}(\tau-q,\tau)\,d\tau\,.$

Exercise 52

Expected PV at time 0 of benefits is

$$\int_{n/2}^n e^{-r\tau}(1 - {}_{\tau-n/2}p_x)\,{}_\tau p_y \mu_{y+\tau}\,d\tau\,.$$

Expected PV at time 0 of premiums is π times

$$\int_0^{n/2} e^{-r\tau}\,{}_\tau p_x\,{}_\tau p_y\,d\tau\,.$$

Equivalence premium π is the ratio between these expressions.

Exercise 53

(a) Expected present value at time 0 of benefits is 5 times

$$W^b = \mathbb{E}\left[e^{\int_0^m (-r(s)+a(s)-\mu(x+s))\,ds} \middle| Y(0) = i \right] . \tag{G.29}$$

Consider the functions

$$W_j^b(t) = \mathbb{E}\left[e^{\int_t^m (-r(s)+a(s)-\mu(x+s))\,ds} \middle| Y(t) = j \right] ,$$

$t \in [0, m]$, $j = 1, \ldots, J$. We need to determine $W^b = W_i^b(0)$. Backward construction:

$$W_j^b(t) = (1 - \lambda_{j\cdot} dt) e^{(-r_j + a_j - \mu(x+t))\,dt} W_j^b(t + dt) + \sum_{k;k \neq j} \lambda_{jk} dt W_k^b(t) + o(dt)$$

leads to

$$\frac{d}{dt} W_j^b(t) = (\lambda_{j\cdot} + r_j - a_j + \mu(x+t)) W_j^b(t) - \sum_{k;\, k \neq j} \lambda_{jk}\, W_k^b(t) \,.$$

Solve backwards subject to conditions

$$W_j^b(m) = 1 \,,$$

$j = 1, \ldots, J$.

Expected present value at time 0 of premiums is π times

$$W^c = \mathbb{E}\left[\int_0^m e^{\int_0^\tau (-r(s)+a(s)-\mu(x+s))\,ds}\, d\tau \middle| Y(0) = i \right] . \tag{G.30}$$

Consider the functions

$$W_j^c(t) = \mathbb{E}\left[\int_t^m e^{\int_t^\tau (-r(s)+a(s)-\mu(x+s))\,ds}\, d\tau \middle| Y(t) = j \right] .$$

$t \in [0, m]$, $j = 1, \ldots, J$. We need to determine $W^c = W_i^c(0)$. Backward construction:

$$W_j^c(t) = (1 - \lambda_{j\cdot} dt) \left(dt + e^{(-r_j + a_j - \mu(x+t))\,dt} W_j^c(t + dt) \right) + \sum_{k;k \neq j} \lambda_{jk} dt W_k^c(t) + o(dt)$$

leads to

$$\frac{d}{dt} W_j^c(t) = -1 + (\lambda_{j\cdot} + r_j - a_j + \mu(x+t)) W_j^c(t) - \sum_{k;\, k \neq j} \lambda_{jk}\, W_k^c(t) \,.$$

Solve backwards subject to conditions

$$W_j^c(m) = 0 \,,$$

$j = 1, \ldots, J$.

Solve both systems numerically by e.g. 'prores2' and determine $\pi = W^b / W^c$.

(b) If $a_j = r_j$ for all j, then $a(t) = r(t)$ for all t and they cancel out of the expressions for the present values when we drop the expectation. Thus, equivalence in the sense defined can be attained.

Exercise 54

(a) With profit contract: Stipulating benefits and premiums in nominal values, binding to both parties. Charge a premium 'on the safe side', typically by using conservative technical (first order) valuation basis. If everything goes well, surplus will accumulate. This surplus belongs to the insured and is to be repaid as so-called *bonus*, e.g. as increased benefits or reduced premiums.

Unit linked: Unit linked contract: Stipulating benefits and (possibly) premiums in units of a share of the investment portfolio, that is, let contractual payments be inflated by the 'price index' of the investment portfolio rather than being fixed nominal amounts. With all payments perfectly linked, the financial risk will be eliminated in a large portfolio.

(b) Single premium, given $Y(0) = i$, is

$$\pi = \mathbb{E}\left[e^{-\int_0^n r}(U(n) \vee g) \Big| Y(0) = i \right] {}_n p_x \,.$$

Disregarding the uninteresting term ${}_n p_x$ henceforth, we need to determine

$$W = \mathbb{E}\left[(1 \vee e^{-\int_0^n r} g) \Big| Y(0) = i \right] .$$

The starting point is the 'price' of the claim at time t, given the current information about the past,

$$\mathbb{E}\left[e^{-\int_t^n r}\,(U(n) \vee g) \,\Big|\, Y(\tau);\, 0 \le \tau \le t \right] = \mathbb{E}\left[\left(U(t) \vee e^{-\int_t^n r} g \right) \Big| Y(\tau);\, 0 \le \tau \le t \right] .$$

Due to the Markov property this expression is a function of t, $Y(t)$ and $U(t)$. Consider its value at time t for given $U(t) = u$, and $Y(t) = j$,

$$W_j(t,u) = \mathbb{E}\left[\left(u \vee e^{-\int_t^n r} g \right) \Big| Y(t) = j \right] .$$

The premium we seek is $W_i(0,1)$

Now use the backward construction, disregarding terms of order $o(dt)$:

$$\begin{aligned}
W_j(t,u) &= (1 - \lambda_{j.}dt)\mathbb{E}\left[\left(u \vee\, e^{-r_j dt} e^{-\int_{t+dt}^{\tau} r} g \right) \Big| Y(t+dt) = j \right] + \sum_{k;\, k \neq j} \lambda_{jk} dt\, W_k(t,u) \\
&= (1 - \lambda_{j.}dt)\, e^{-r_j dt}\mathbb{E}\left[\left(u\, e^{r_j dt} \vee e^{-\int_{t+dt}^{\tau} r} g \right) \Big| Y(t+dt) = j \right] + \sum_{k;\, k \neq j} \lambda_{jk} dt\, W_k(t,u) \\
&= (1 - \lambda_{j.}\, dt)\, e^{-r_j\, dt} W_j(t+dt, u e^{r_j\, dt}) + \sum_{k;\, k \neq j} \lambda_{jk} dt\, W_k(t,u)
\end{aligned}$$

Insert $e^{\pm r_j\, dt} = 1 \pm r_j dt + o(dt)$ and

$$\begin{aligned}
W_j(t+dt, e^{r_j dt} u) &= W_j(t+dt, u + u r_j dt) + o(dt) \\
&= W_j(t,u) + \frac{\partial}{\partial t} W_j(t,u)\, dt + \frac{\partial}{\partial u} W_j(t,u)\, u\, r_j\, dt + o(dt)\,,
\end{aligned}$$

and fill in some details to arrive at the partial differential equations

$$-r_j W_j(t,u) + \frac{\partial}{\partial t} W_j(t,u) + \frac{\partial}{\partial u} W_j(t,u)\, u\, r_j + \sum_{k;\, k \neq j} \lambda_{jk} \left(W_k(t,u) - W_j(t,u) \right) = 0\,.$$

These are to be solved subject to the conditions

$$W_j(n,u) = (u \vee g)\,,$$

$j = 1, \ldots, J$.

Exercise 55

(a) Expected PV at time 0 of benefits:

$$\int_0^n e^{-r\tau}(1 - {}_{\tau/2}p_x)\,{}_\tau p_y \mu_{y+\tau}\, d\tau\,.$$

Expected PV at time 0 of premiums is π times

$$\int_0^{n/2} e^{-r\tau}\,{}_\tau p_x\,{}_\tau p_y\, d\tau\,.$$

Equivalence premium π is the ratio between these expressions.

(b) (Do not spend too much time on this.) A straightforward method for computation is to define $v_1(t) = 1 - {}_{t/2}p_x$, $v_2(t) = {}_t p_y$,

$$v_3(t) = \int_0^t e^{-r\tau}(1 - {}_{\tau/2}p_x)\,{}_\tau p_y \mu_{y+\tau}\, d\tau\,.$$

and solve numerically the system of differential equations

$$v_1'(t) = \mu_{x+t/2}\,(1/2)\,v_1(t)\,,$$
$$v_2'(t) = -\mu_{y+t}\,v_2(t)\,,$$
$$v_3'(t) = e^{-rt}v_1(t)\,v_2(t)\,\mu_{y+t}\,,$$

by a forward difference scheme starting from the conditions $v_1(0) = 0$, $v_2(0) = 1$, $v_3(0) = 0$.

A more sophisticated method hinted at in the problem: Observe that

$${}_{t/2}p_x = \exp\left(-\int_0^{t/2} \mu(x+s)\, ds\right) = \exp\left(-\int_0^t \frac{1}{2}\mu(x+\frac{1}{2}s)\, ds\right)\,,$$

formally a survival function with intensity $\tilde{\mu}(t) = \mu(x + t/2)/2$. Then the single premium is the difference between the single premiums of two well-known simple products, which may be computed by solving their Thiele differential equations numerically. Or compute by e.g. the program 'prores1' the expected discounted value of an assurance of 1 payable upon transition from state 1 to state 3 in a four states Markov model on $\{0,1,2,3\}$, starting from state 0, with transition intensities $\mu_{01}(t) = \mu_{23}(t) = \mu(x+t/2)/2$, $\mu_{02}(t) = \mu_{13}(t) = \mu(y+t)$, and all other intensities 0.

(c) Reserve V_t at time $t \in [0,n]$ depends on what is currently known about (x) and (y):
(y) dead: $V_t = 0$.
(y) alive, $t \geq n/2$, $T_x > n/2$: $V_t = 0$.

(y) alive, $t \geq n/2$, $T_x \leq n/2$: $V_t = \int_{2T_x \vee t}^{n} e^{-r(\tau-t)}\,{}_{\tau-t}p_{y+t}\mu_{y+\tau}\,d\tau$.
(y) alive, $t < n/2$, $T_x \leq t$: $V_t = \int_{2T_x}^{n} e^{-r(\tau-t)}\,{}_{\tau-t}p_{y+t}\,\mu_{y+\tau}\,d\tau$.
(y) alive, $t < n/2$, $T_x > t$: $V_t = \int_{2t}^{n} e^{-r(\tau-t)}(1-{}_{\tau/2-t}p_{x+t})\,{}_{\tau-t}p_{y+t}\mu_{y+\tau}\,d\tau - \pi\int_{t}^{n/2} e^{-r(\tau-t)}\,{}_{\tau-t}p_{x+t}\,{}_{\tau-t}p_{y+t}\,d\tau$.

(d) In general, for a unit due at some random time, the non-central 2nd moment of present value is the same as the expected value, only with $2r$ instead of r.

Exercise 21

(b) The expected present value at time 0 of the benefits is 0.75 times

$$\begin{aligned} W^b &= \mathbb{E}\left[S(m)\int_m^{m+n} U(\tau)^{-1}\,{}_{\tau}p_x\,d\tau \middle| Y(0)=i\right] \\ &= \mathbb{E}\left[e^{\int_0^m a(s)\,ds}\int_m^{m+n} e^{-\int_0^\tau (r(s)+\mu(x+s))\,ds}\,d\tau \middle| Y(0)=i\right]. \end{aligned}$$

Proceed along the lines of Problems 4.7 - 4.11. The conditional expected value of the discounted benefits, given what is known at time $t \in [0, m+n]$, is 0.75 times

$$W^b(t) = \mathbb{E}\left[e^{\int_0^m a(s)\,ds}\int_m^{m+n} e^{-\int_0^\tau (r(s)+\mu(x+s))\,ds}\,d\tau \middle| Y(\tau);\, 0\leq\tau\leq t\right].$$

The gist of the matter in the backward construction is to separate out what relates only to the past and what relates to the future here, and to work with the conditional expected values of the latter, given the relevant pieces of current information.

Firstly, for $m \leq t \leq m+n$,

$$\begin{aligned} W^b(t) &= e^{\int_0^m a(s)\,ds}\int_m^{t} e^{-\int_0^\tau (r(s)+\mu(x+s))\,ds}\,d\tau \\ &\quad + e^{\int_0^m a(s)\,ds - \int_0^t (r(s)+\mu(x+s))\,ds}\,\mathbb{E}\left[\tilde{W}^b(t)\,|\,Y(\tau);\,0\leq\tau\leq t\right], \end{aligned}$$

where

$$\tilde{W}^b(t) = \int_t^{m+n} e^{-\int_t^\tau (r(s)+\mu(x+s))\,ds}\,d\tau.$$

In the expression for $W^b(t)$ we need to determine the conditional expected value of $\tilde{W}^b(t)$ in the last term. Due to the Markov property (conditional independence between past and future, given the present), this conditional expected value depends only on the time t and the current state of the economy, $Y(t)$. Therefore, we can restrict attention to the state-wise conditional expected values,

$$\tilde{W}^b_j(t) = \mathbb{E}\left[\tilde{W}^b(t)\middle|\, Y(t)=j\right],$$

$m \leq t \leq m+n$, $j = 1, \ldots, J$. These are the solution to the differential equations

$$\frac{d}{dt}\tilde{W}^b_j(t) = (r_j + \mu(x+t))\,\tilde{W}^b_j(t) - 1 - \sum_{k;\,k\neq j}\lambda_{jk}\,(\tilde{W}^b_k(t) - \tilde{W}^b_j(t)) = 0\,, \quad \text{(G.31)}$$

subject to the conditions

$$\tilde{W}^b_j(m+n) = 0\,, \quad \text{(G.32)}$$

$j = 1, \ldots, J$. Outline of details (which you shouldn't spell out unless you are explicitly asked to do so):

$$\begin{aligned}
\tilde{W}^b(t) &= \int_t^{t+dt} e^{-\int_t^\tau (r(s)+\mu(x+s))\,ds}\,d\tau \\
&\quad + e^{-\int_t^{t+dt}(r(s)+\mu(x+s))\,ds} \int_{t+dt}^{m+n} e^{-\int_{t+dt}^\tau (r(s)+\mu(x+s))\,ds}\,d\tau \\
&= e^{O(dt)}\,dt + e^{-(r(t)+\mu(x+t))\,dt}\,\tilde{W}^b(t+dt) \,+\, o(dt)\,,
\end{aligned}$$

where $O(dt)$ is of order dt and $o(dt)$ is of order $(dt)^2$, hence

$$\tilde{W}^b(t) = dt + (1 - (r(t) + \mu(x+t))\,dt)\,\tilde{W}^b(t+dt) \,+\, o(dt)\,.$$

(We have lumped negligible terms into $o(dt)$: For instance, by Taylor expansion, $dt\,\exp(O(dt)) = dt\,(1 + O(dt)) = dt + o(dt)$ and $dt\,\tilde{W}_k^b(t+dt) = dt\,(\tilde{W}_k^b(t) + O(dt)) = dt\,\tilde{W}_k^b(t) + o(dt)$.) Now, condition on what happens in the small time interval $(t, t+dt)$:

$$\begin{aligned}
\tilde{W}_j^b(t) &= (1 - \lambda_{j\cdot}\,dt)\,\mathbb{E}\left[dt + (1 - (r(t) + \mu(x+t))\,dt)\,\tilde{W}^b(t+dt) \,+\, o(dt)\,\middle|\, Y(\tau) = j,\, \tau \in [t, t+dt]\right] \\
&\quad + \sum_{k;\, k\neq j} \lambda_{jk} dt\,\mathbb{E}\left[dt + (1 - (r(t) + \mu(x+t))\,dt)\,\tilde{W}^b(t+dt) \,+\, o(dt)\,\middle|\, Y(t) = j,\, Y(t+dt) = k\right] \\
&= (1 - \lambda_{j\cdot}\,dt)\,[\,dt + (1 - (r_j + \mu(x+t))\,dt)\,\tilde{W}_j^b(t+dt)] + \sum_{k;\, k\neq j} \lambda_{jk} dt\,\tilde{W}_k^b(t) \,+\, o(dt) \\
&= dt + \tilde{W}_j^b(t+dt) - (r_j + \mu(x+t) + \lambda_{j\cdot})\,dt\,\tilde{W}_j^b(t) + \sum_{k;\, k\neq j} \lambda_{jk} dt\,\tilde{W}_k^b(t) \,+\, o(dt)\,.
\end{aligned}$$

(The first step in the above derivation is made here just to make sure you understand the argument - it need not be given in an answer to the Problem.) Subtracting $\tilde{W}_j^b(t+dt)$ on both sides, dividing by dt, and letting dt go to 0, we arrive at (G.31). Another way to do it is to set

$$\tilde{W}_j^b(t+dt) = \tilde{W}_j^b(t) + \frac{d}{dt}\tilde{W}_j^b(t)\,dt + o(dt)\,,$$

cancel $\tilde{W}_j^b(t)$ on both sides, and finally divide by dt.

Secondly, for $0 \le t \le m$,

$$W^b(t) = e^{\int_0^t (a(s)-r(s)-\mu(x+s))\,ds}\,\mathbb{E}\left[\tilde{W}^b(t) |\, Y(\tau);\, 0 \le \tau \le t\right]\,,$$

where now

$$\tilde{W}^b(t) = e^{\int_t^m (a(s)-r(s)-\mu(x+s))\,ds} \int_m^{m+n} e^{-\int_m^\tau (r(s)+\mu(x+s))\,ds}\,d\tau\,.$$

The state-wise conditional expected values of this function are

$$\tilde{W}_j^b(t) = \mathbb{E}\left[\tilde{W}^b(t)\,\middle|\, Y(t) = j\right]\,,$$

$0 \le t \le m$, $j = 1, \ldots, J$. These are the solution to the differential equations

$$\frac{d}{dt}\tilde{W}_j^b(t) = (r_j + \mu(x+t) - a_j)\,\tilde{W}_j^b(t) - \sum_{k;\, k\neq j} \lambda_{jk}\,(\tilde{W}_k^b(t) - \tilde{W}_j^b(t)) = 0\,. \quad \text{(G.33)}$$

At time $t = m$ the defining expressions for $\tilde{W}^b(t)$ in $[0, m]$ and $[m, m+n]$ coincide, so we just proceed backwards in $[0, m]$ from the values $\tilde{W}^b_j(m)$ obtained after having completed the backward scheme in $[m, m+n]$.

Brief outline of details of the derivation of the differential equations:

$$\begin{aligned}\tilde{W}^b(t) &= e^{\int_t^{t+dt}(a(s)-r(s)-\mu(x+s))\,ds}\,\tilde{W}^b(t+dt)\\ &= (1+(a(t)-r(t)-\mu(x+t))\,dt)\,\tilde{W}^b(t+dt)\,;\end{aligned}$$

$$\begin{aligned}\tilde{W}^b_j(t) &= (1-\lambda_j\,.\,dt)\,(1+(a_j-r_j-\mu(x+t))\,dt)\tilde{W}^b_j(t+dt)\\ &\quad+\sum_{k;\,k\neq j}\lambda_{jk}dt\,\tilde{W}^b_k(t+dt)\;+\;o(dt)\,,\end{aligned}$$

and the rest is trivial.

The expected present value of future benefits at time 0, when the economy starts in state i, is

$$W^b = \tilde{W}^b_i(0)\,.$$

The expected present value at time 0 of the premiums is π times

$$\begin{aligned}W^c &= \mathbb{E}\left[\int_0^m U(\tau)^{-1}\,{}_\tau p_x\,d\tau\,\middle|\,Y(0)=i\right]\\ &= \mathbb{E}\left[\int_0^m e^{-\int_0^\tau(r(s)+\mu(x+s))\,ds}\,d\tau\,\middle|\,Y(0)=i\right].\end{aligned}$$

The situation is now simpler than in the case of the benefits. For $t \in [0, m]$ it suffices to look at the state-wise conditional expected values

$$W^c_j(t) = \mathbb{E}\left[\int_t^m e^{-\int_t^\tau(r(s)+\mu(x+s))\,ds}\,d\tau\,\middle|\,Y(t)=j\right],$$

$0 \le t \le m$, $j = 1, \ldots, J$. These we determine as the solution to the differential equations (copy arguments for the case $t \in [m, m+n]$ above)

$$\frac{d}{dt}W^c_j(t) = (r_j+\mu(x+t))\,W^c_j(t) - 1 - \sum_{k;\,k\neq j}\lambda_{jk}\,(W^c_k(t)-W^c_j(t)) = 0\,, \tag{G.34}$$

subject to the conditions

$$W^c_j(m) = 0\,, \tag{G.35}$$

$j = 1, \ldots, J$.

The expected present value at time 0, when the economy starts in state i, is

$$W^c = W^c_i(0)\,.$$

Finally, determine π from the equivalence relation $\pi\,W^c_i(0) = 0.75\,W^b_i(0)$:

$$\pi = 0.75\,W^b_i(0)/W^c_i(0)\,.$$

This solution was filled with details that would not make any good in an exam; you should just write the essential steps. To check your understanding of these things, try your hand on the modified contract where the benefit is a lump sum of $4\,S(m)$ payable upon retirement at time m. (A simpler problem, of course.)

Problem 4.15
This is a small perturbation of Problem 4.11, and actually it is simpler. We treat only the benefits (for premiums we do as in Problem 4.13).

Suppose $Y(0) = i$. The expected discounted value of the benefits is ${}_np_x$ times

$$W = \mathbb{E}\left[e^{-\int_0^n r}(U(n) \vee g)\Big| Y(0) = i\right] = \mathbb{E}\left[(1 \vee e^{-\int_0^n r}\, g)\Big| Y(0) = i\right] .$$

A suitable starting point is the 'price' of the claim at time t, given the current information about the past,

$$\mathbb{E}\left[e^{-\int_t^n r}\,(U(n) \vee g)\,\Big|\, Y(\tau);\, 0 \le \tau \le t\right] = \mathbb{E}\left[\left(U(t) \vee e^{-\int_t^n r} g\right)\Big| Y(\tau);\, 0 \le \tau \le t\right] .$$

Arguing as in Problem 4.10 of the Exercises, we realize that this expression is a function of t, $Y(t)$ and $U(t)$. Consider its value at time t for given $U(t) = u$, and $Y(t) = j$,

$$W_j(t, u) = \mathbb{E}\left[\left(u \vee e^{-\int_t^n r} g\right)\Big| Y(t) = j\right] .$$

The value W we seek is $W_i(0, 1)$.

Now use the backward construction, disregarding terms of order $o(dt)$:

$$\begin{aligned} W_j(t,u) &= (1 - \lambda_{j\cdot}\,dt)\mathbb{E}\left[\left(u \vee\, e^{-r_j\,dt} e^{-\int_{t+dt}^{\tau} r} g\right)\Big| Y(t+dt) = j\right] + \sum_{k;\, k\neq j} \lambda_{jk} dt\, W_k(t,u) \\ &= (1 - \lambda_{j\cdot}\,dt)\, e^{-r_j\,dt}\mathbb{E}\left[\left(u\, e^{r_j\,dt} \vee e^{-\int_{t+dt}^{\tau} r} g\right)\Big| Y(t+dt) = j\right] + \sum_{k;\, k\neq j} \lambda_{jk} dt\, W_k(t,u) \\ &= (1 - \lambda_{j\cdot}\; dt)\, e^{-r_j\; dt} W_j(t + dt, u e^{r_j\; dt}) + \sum_{k;\, k\neq j} \lambda_{jk} dt\, W_k(t,u)\,. \end{aligned}$$

Now do as in 4.11 and fill in some details to arrive at the partial differential equations

$$-r_j W_j(t,u) + \frac{\partial}{\partial t} W_j(t,u) + \frac{\partial}{\partial u} W_j(t,u)\, u\, r_j + \sum_{k;\, k\neq j} \lambda_{jk}\left(W_k(t,u) - W_j(t,u)\right) = 0\,.$$

These are to be solved subject to the conditions

$$W_j(n,u) = (u \vee g)\,,$$

$j = 1, \ldots, J$.

The expected discounted premiums are found as in Problem 4.13.

Exercise 57
(a) Direct from the Poisson distribution:

$$\begin{aligned} \mathbb{E}[S(t)] &= e^{\alpha t}\mathbb{E}\left[e^{\beta N(t)}\right] \\ &= e^{\alpha t} \sum_{n=0}^{\infty} e^{\beta n} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \end{aligned}$$

$$
\begin{aligned}
&= e^{\alpha t} e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(e^{\beta}\lambda t)^n}{n!} \\
&= e^{\alpha t} e^{-\lambda t} \exp\left(e^{\beta}\lambda t\right) \\
&= \exp\left(\alpha t + (e^{\beta}-1)\lambda t\right) .
\end{aligned}
$$

Backward construction, particularly simple here since $\alpha t + bN(t)$ has stationary and independent increments:

$$
\begin{aligned}
V(t) &= \mathbb{E}\left[e^{\alpha t+\beta N(t)}\right] \\
&= (1-\lambda dt)e^{\alpha dt}V(t-dt) + \lambda\, dt\, e^{\alpha dt+\beta}V(t-dt) + o(dt) \\
&= e^{\alpha dt}V(t-dt) - \lambda\, dt\, e^{\alpha dt}V(t-dt) + \lambda\, dt\, e^{\alpha dt+\beta}V(t-dt) + o(dt) \\
&= (1+\alpha\, dt)(V(t) - \frac{d}{dt}V(t)\, dt) \\
&\quad -\lambda\, dt\, e^{\alpha dt}V(t-dt) + \lambda\, dt\, e^{\alpha\, dt+\beta}V(t-dt) + o(dt) .
\end{aligned}
$$

Cancel $V(t)$, divide by dt and let dt go to 0 to obtain a simple differential equation, to be solved subject to $V(0) = 1$.

For non-central q-th moments, just replace α and β with $q\alpha$ and $q\beta$.

(b)

$$
\begin{aligned}
dS(t) &= e^{\alpha t+\beta N(t)}\alpha dt + dN(t)\left(e^{\alpha t+\beta(N(t-)+1)} - e^{\alpha t+\beta N(t-)}\right) \\
&= e^{\alpha t+\beta N(t-)}\alpha dt + dN(t)e^{\alpha t+\beta N(t-)}\left(e^{\beta}-1\right) \\
&= S(t-)\left(\alpha t + \lambda\left(e^{\beta}-1\right)\right)\, dt \;+\; S(t-)\left(e^{\beta}-1\right) dM(t) ,
\end{aligned}
$$

where $M(t) = N(t) - \lambda t$, a so-called martingale (a process with conditionally zero mean and uncorrelated increments, here actually independent increments).

(c) Replace α and β in (a) with $-\alpha$ and $-\beta$, and integrate the expression for the expected value, to obtain the claimed answer (misprint: a λ is missing).

(d) This is now trivial.

BEGIN From 305exsol

Exercise 55

(a) Expected PV at time 0 of benefits:

$$
\int_0^n e^{-r\tau}(1 - {}_{\tau/2}p_x)\, {}_{\tau}p_y \mu_{y+\tau}\, d\tau .
$$

Expected PV at time 0 of premiums is π times

$$
\int_0^{n/2} e^{-r\tau}\, {}_{\tau}p_x\, {}_{\tau}p_y\, d\tau .
$$

Equivalence premium π is the ratio between these expressions.

(b) (Do not spend too much time on this.) A straightforward method for computation is to define $v_1(t) = 1 - {}_{t/2}p_x$, $v_2(t) = {}_tp_y$,

$$v_3(t) = \int_0^t e^{-r\tau}(1 - {}_{\tau/2}p_x)\,{}_\tau p_y \mu_{y+\tau}\, d\tau\,.$$

and solve numerically the system of differential equations

$$v_1'(t) = \mu_{x+t/2}\,(1/2)\,v_1(t)\,,$$
$$v_2'(t) = -\mu_{y+t}\,v_2(t)\,,$$
$$v_3'(t) = e^{-rt}v_1(t)\,v_2(t)\,\mu_{y+t}\,,$$

by a forward difference scheme starting from the conditions $v_1(0) = 0$, $v_2(0) = 1$, $v_3(0) = 0$.

A more sophisticated method hinted at in the problem: Observe that

$${}_{t/2}p_x = \exp\left(-\int_0^{t/2}\mu(x+s)\,ds\right) = \exp\left(-\int_0^t \frac{1}{2}\mu(x+\frac{1}{2}s)\,ds\right)\,,$$

formally a survival function with intensity $\tilde{\mu}(t) = \mu(x+t/2)/2$. Then the single premium is the difference between the single premiums of two well-known simple products, which may be computed by solving their Thiele differential equations numerically. Or compute by e.g. the program 'prores1' the expected discounted value of an assurance of 1 payable upon transition from state 1 to state 3 in a four states Markov model on $\{0,1,2,3\}$, starting from state 0, with transition intensities $\mu_{01}(t) = \mu_{23}(t) = \mu(x+t/2)/2$, $\mu_{02}(t) = \mu_{13}(t) = \mu(y+t)$, and all other intensities 0.

(c) Reserve V_t at time $t \in [0,n]$ depends on what is currently known about (x) and (y):
(y) dead: $V_t = 0$.
(y) alive, $t \geq n/2$, $T_x > n/2$: $V_t = 0$.
(y) alive, $t \geq n/2$, $T_x \leq n/2$: $V_t = \int_{2T_x \vee t}^n e^{-r(\tau-t)}\,{}_{\tau-t}p_{y+t}\mu_{y+\tau}\,d\tau$.
(y) alive, $t < n/2$, $T_x \leq t$: $V_t = \int_{2T_x}^n e^{-r(\tau-t)}\,{}_{\tau-t}p_{y+t}\,\mu_{y+\tau}\,d\tau$.
(y) alive, $t < n/2$, $T_x > t$: $V_t = \int_{2t}^n e^{-r(\tau-t)}(1-{}_{\tau/2-t}p_{x+t})\,{}_{\tau-t}p_{y+t}\mu_{y+\tau}\,d\tau - \pi\int_t^{n/2} e^{-r(\tau-t)}\,{}_{\tau-t}p_{x+t}\,{}_{\tau-t}p_{y+t}\,d\tau$.

(d) In general, for a unit due at some random time, the non-central 2nd moment of present value is the same as the expected value, only with $2r$ instead of r.

Exercise 75

We work under the independence hypothesis (should have been stated in the exercise text.)

First the brute force method:

$${}_tp_{\overline{x_1\ldots x_r}} = 1 - \prod_j (1 - {}_tp_{x_j})\,.$$

Using the rule for differentiating a product (special case of Itǒ),

$$\frac{d}{dt}{}_tp_{\overline{x_1 \ldots x_r}} = -\sum_k -\frac{d}{dt}{}_tp_{x_k} \prod_{j;\, j\neq k} (1 - {}_tp_{x_j}) = -\sum_k {}_tp_{x_k}\, \mu_{x_k+t} \prod_{j;\, j\neq k} (1 - {}_tp_{x_j}) \, .$$

It follows that

$$\mu_{\overline{x_1 \ldots x_r}}(t) = \frac{\sum_k {}_tp_{x_k}\, \mu_{x_k+t} \prod_{j;\, j\neq k} (1 - {}_tp_{x_j})}{1 - \prod_j (1 - {}_tp_{x_j})} \, .$$

Second, direct reasoning: The conditional probability that the last survivor dies in $(t, t+dt)$, given that there are survivors at time t, is dt times the expression above.

Exercise 76

(a) Apology: Problems of this kind are one of the favorite sports of classical actuaries, not because they are so common in practice (market share in the per mille range), but rather because they can entertain and stimulate the brains of actuaries. The proposed product is not totally inconceivable, however: it might be useful for a couple that needs to secure economically the last survivor and also their children after the possible early death of the last survivor. It is also of some theoretical interest beyond that of mere parlor games as it is an example of a product where payments are dependent on the past history of the driving process. This is seen clearly if the problem is formulated in the set-up of the Markov chain models for two lives. Now to work:

As is almost always the case, the best method is to find the expected value of the discounted payment in each small time interval $(\tau, \tau+d\tau)$ and then sum over all times. For $\tau \leq 20$ the benefit is running if (and only if) $T_{xy} < \tau$. For $\tau > 20$ the benefit is running if $\tau - 20 < T_{xy} < \tau$ or if $T_{xy} < \tau - 20$ and $\tau - 10 < T_{\overline{xy}} < \tau$. We gather the following expected value of future discounted payments at time 0:

$$\begin{aligned} &\int_0^{20} e^{-r\tau}\, (1 - {}_\tau p_x\, {}_\tau p_y)\, d\tau \\ &+ \int_{20}^{\infty} e^{-r\tau} \left[{}_{\tau-20}p_x\, {}_{\tau-20}p_y - {}_\tau p_x\, {}_\tau p_y\right] d\tau \\ &+ \int_{20}^{\infty} e^{-r\tau} [(1 - {}_{\tau-20}p_x)\, ({}_{\tau-10}p_y - {}_\tau p_y) + (1 - {}_{\tau-20}p_y)\, ({}_{\tau-10}p_x - {}_\tau p_x)]\, d\tau \, . \end{aligned}$$

(b) The premium rate π is the ratio between the expected present value in item (a) and the expected present value

$$\int_0^{\infty} e^{-r\tau}\, {}_\tau p_x\, {}_\tau p_y\, d\tau \, .$$

The reserve is a long and tedious story. One must, at each time of consideration t, distinguish between all possible past histories of the two lives along. For instance, if $T_{xy} > t$, then the reserve is simply the first expression above minus π times the second expression above, with x and y replaced by $x+t$ and $y+t$.

Exercise 77

$$\int_0^{20} e^{-r\tau}\, {}_\tau p_x\, \mu_{x+\tau}\, {}_\tau p_z\, d\tau + \int_{20}^{\infty} e^{-r\tau}\, {}_\tau p_x\, \mu_{x+\tau}\, {}_{\tau-20}p_y\, {}_\tau p_z\, d\tau \, .$$

A benefit of 1 payable immediately upon the death of (x) if (z) is then still alive and (y) was alive 20 years ago.

Exercise 78
Let us say n = 30. Use the set-up of the four states Markov chain for two lives. Set age $x = 0$ (we could have taken $x = 20$ since, accidentally, the two start at same age, but the x in the program, which refers to a single life, doesn't really apply to the general case). Account of starting age of the two by writing $\alpha + \beta e^{\gamma\,(20+t)} = \alpha + \beta^* e^{\gamma\,t}$ with $\beta^* = \beta\, e^{\gamma\,20}$:

```
(* SPECIFY NON-NULL PAYMENTS AT TIME t ! *)
bi[2,4] := 1; bi[3,4] := 1;
ca[1] := 1;

(* SPECIFY MAXIMUM ORDER OF MOMENTS AND NUMBER OF STATES !
*)
q := 1; (*moments*)
JZ := 4; (*number of states of the policy*)

(* SPECIFY TRANSITION INTENSITIES FOR POLICY Z ! *)
alpha[1,2] := 0.0005;
gamma[1,2] := 0.038*ln(10);
beta[1,2] := 0.00007585775*exp(gamma[1,2]*20);
alpha[1,3] := 0.0005;
gamma[1,3] := 0.038*ln(10);
beta[1,3] := 0.00007585775*exp(gamma[1,3]*20);
alpha[2,4] := 0.0005;
gamma[2,4] := 0.038*ln(10);
beta[2,4] := 0.00007585775*exp(gamma[2,4]*20);
alpha[3,4] := 0.0005;
gamma[3,4] := 0.038*ln(10);
beta[3,4] := 0.00007585775*exp(gamma[3,4]*20);

(*SPECIFY AGE x, TERM t, INTEREST RATES AND NON-NULL LIFE ENDOW-
MENTS ! *)
x:= 0; (*age*)
t := 30; (*term*)
r := ln(1+0.045); (*interest rate*)
be[1] := 0; be[2] := 0; (*endowments at term of contract*)
(*SPECIFY LUMP SUM PREMIUM AT TIME 0: PUT c0 := 1 IF ALL
OTHER PREMIUMS ARE 0 AND ONLY MOMENTS OF BENEFITS ARE
WANTED ! *)
c0 := 0; b0 := 0;
```

Exercise 1-12
Pure verification - just insert the appropriate expressions on the right hand side. Any combination of cash bonus at rate $\tilde{b}_t = \alpha_t\, c_t$ and additional death benefit of $\hat{b}_t = (1 - \alpha_t)\, c_t/\mu_{x+t}$, $0 \leq \alpha \leq 1$, produces a right hand side equal to the left hand

side.

Exercise 1-13
$\tilde{b}_n$ must solve

$$\int_0^n e^{-\int_0^\tau (r_u+\mu_{x+u})\,du} c_\tau \, d\tau \; = \; e^{-\int_0^n (r_u+\mu_{x+u})\,du} \tilde{b}_n \,, \tag{G.36}$$

hence

$$\tilde{b}_n \; = \; \int_0^n e^{\int_\tau^n (r_u+\mu_{x+u})\,du} c_\tau \, d\tau \,. \tag{G.37}$$

Exercise 1-14
Relation (7) becomes

$$c_t = \Delta r \, V_t^* \; + \; \Delta\mu (b_t - V_t^*) \,.$$

Here are some examples of time t prognosis of future bonuses, assuming that the insured will survive the term of the contract:
1. Rate of cash bonus payments at time $u \in (t, n)$ is just c_u defined above.
2. Present value of future cash bonuses:

$$\int_t^n e^{-(r^*+\Delta r)(\tau - t)} \, c_\tau \, d\tau \,,$$

3. Value of terminal bonus (not discounted):

$$\int_0^n e^{\int_\tau^n (\mu^*_{x+s} \, + \, \Delta\mu \, + \, r^* \, + \, \Delta r)\, ds} \, c_\tau \, d\tau \,.$$

Exercise 1-15
All that is needed is to put $m+n$ in the role of n and work with the general formulas. Here V_t^* is given by

$$\begin{aligned} \frac{d}{dt} V_t^* \quad &= \quad (\mu^*_{x+t} \, + \, r^*) \, V_t^* \, + \, \pi \,, \qquad 0 < t < m \,, \\ \frac{d}{dt} V_t^* \quad &= \quad (\mu^*_{x+t} \, + \, r^*) \, V_t^* \, - \, 1 \,, \qquad m < t < m+n \,. \end{aligned}$$

This is the only place where the particulars of the contract matter: Thiele's differential equation is needed for the computation of V_t^* alongside that of c_t.

Since $V_t^* > 0$ for all $t \in (m, m+n)$, also $c_t > 0$ throughout this time interval.

Exercise 1-16
Expenses can be treated as benefits in addition to those specified in the contract (see Chapter 5). We need the differential equation for the first order gross reserve,

$$\frac{d}{dt} V_t^{*'} \; = \; V_t^{*'} \, r^* \, + \, \pi \, - \, \beta^* \pi \, - \, \gamma^* b \, - \, \mu^*_{x+t} (b \, - \, V_t^{*'}) \tag{G.38}$$

(with side condition $V_{n-}^{*'} \; = \; b$), and the equivalence relationship,

$$V_0^{*'} \; = \; -\alpha^* b \,,$$

which determines π. The discounted mean surplus per policy at time t is now

$$\begin{aligned} S_t &= -(\alpha' + \alpha'' b) \\ &\quad + \int_0^t e^{-\int_0^\tau (r_s + \mu_{x+s})\, ds} (\pi - \mu_{x+\tau}\, b - \beta'_\tau - \beta''_\tau \pi' - \gamma'_\tau - \gamma''_\tau b - \gamma'''_\tau V^{*'}_\tau)\, d\tau \\ &\quad - e^{-\int_0^t (r_s + \mu_{x+s})\, ds} V^{*'}_t \,. \end{aligned}$$

It is seen that

$$S_0 \,=\, -(\alpha' + \alpha'' b) \,-\, V^{*'}_0 \,=\, \alpha^* b \,-\, (\alpha' + \alpha'' b)\,,$$

which is the surplus arising immediately upon issue of the contract due to prudent first order assumptions about the initial cost. It is positive (and indeed prudent) if

$$\alpha^* b \,>\, (\alpha' + \alpha'' b)\,,$$

which means that the first order initial cost is set on the safe side. (This cannot be achieved for all $b > 0$ if $\alpha' > 0$; one then has to assume that b is greater than a certain minimum, which is certainly the case in practice.)

The dynamics of the surplus is

$$\begin{aligned} dS_t &= e^{-\int_0^t (r_s + \mu_{x+s})\, ds} (\pi - \mu_{x+t}\, b - \beta'_t - \beta''_t \pi' - \gamma'_t - \gamma''_t b - \gamma'''_t V^{*'}_t)\, dt \\ &\quad + e^{-\int_0^t (r_u + \mu_{x+u})\, du} (r_t + \mu_{x+t})\, V^{*'}_t \,-\, e^{-\int_0^t (r_u + \mu_{x+u})\, du}\, dV^{*'}_t \,. \end{aligned}$$

Inserting $dV^{*'}_t \,=\, \frac{d}{dt} V^{*'}_t\, dt$ from (G.38), we gather

$$dS_t = e^{-\int_0^t (r_s + \mu_{x+s})\, ds}\, c_t\, dt\,,$$

where

$$\begin{aligned} c_t &= (r_t - r^*) V^{*'}_t \,+\, (\beta^* \pi' - \beta'_t - \beta''_t \pi') \\ &\quad + (\gamma^* V^{*'}_t - \gamma'_t - \gamma''_t b - \gamma'''_t V^{*'}_t) \,+\, (\mu^*_{x+t} - \mu_{x+t})(b - V^{*'}_t) \end{aligned}$$

is the mean contribution to surplus per survivor at time t. This contribution decomposes into gains stemming from safety loadings in the various first order elements – interest, expenses of β type, expenses of γ type, and mortality – and how these elements can be set on the safe side. Just as for the initial cost, there is a problem with the safety loading on expenses of β and γ type: if e.g. $\gamma'_t > 0$, then there will inevitably be a loss on the γ expenses for small t since the gross reserve starts from a negative value. This loss has to be compensated by setting other first order elements sufficiently to the safe side to make c_t (or at least S_t) non-negative for all $t \in (0, n)$.

Exercise 1-17

This is a trivial one, and the same goes for the conditional expected value of any random variable that depends only on the state of Y at some fixed future time. Starting from

$$W_e(t) \,=\, (1 - \lambda_{e\cdot}\, dt) W_e(t + dt) \,+\, \sum_{\ell; f \neq e} \lambda_{ef}\, dt\, W_f(t + dt) \,+\, o(dt)\,,$$

we get

$$W_e(t) - W_e(t + dt) \,=\, -\lambda_{e\cdot}\, W_e(t + dt) \,+\, \sum_{\ell; f \neq e} \lambda_{ef}\, dt\, W_f(t + dt) \,+\, o(dt)\,,$$

and, dividing by dt and letting $dt \searrow 0$, we arrive at the answer. The side conditions are obvious (as always).

Exercise 1-18
We start with W'_e and supply details (to be precise, add a term $o(dt)$ on the right of the two expressions given for W'_t and W''_t in the exercise):

$$\begin{aligned} W'_e(t) &= (1-\lambda_e\,.dt)\mathbb{E}\left[e^{r_t\,dt}W'_{t+dt}\,\middle|\,Y_\tau = e\,,\, t \le \tau \le \tau + d\tau\right] \\ &\quad + \sum_{f\,;f\neq e} \lambda_{ef}\,dt\,\mathbb{E}\left[e^{r_t\,dt}W'_{t+dt}\,\middle|\,Y_t = e\,,\, Y_{t+dt} = f\right] \,+\, o(dt) \\ &= (1-\lambda_e\,.dt)e^{r^e\,dt}W'_e(t+dt) + \sum_{f\,;f\neq e} \lambda_{ef}\,dt\,e^{O(dt)}W'_f(t+dt) \,+\, o(dt)\,, \end{aligned}$$

where $O(dt)$ signifies a term of order dt (i.e. such that $O(dt)/dt$ is bounded as $dt \searrow 0$). Inserting the Taylor expansions

$$\begin{aligned} e^{r^e\,dt} &= 1 + r^e dt + o(dt)\,, \\ W'_e(t+dt) &= W'_e(t) + \frac{d}{dt}W'_e(t)\,dt + o(dt)\,, \\ e^{O(dt)} &= 1 + O(dt)\,, \\ W'_f(t+dt) &= W'_f(t) + O(dt)\,, \end{aligned}$$

multiplying out, gathering all $o(dt)$ terms and rearranging a bit, one obtains the differential equations for the functions W'_e.

Next, the W''_e, a bit more sketchy and gathering $o(dt)$ terms currently as they arise without further mentioning:

$$\begin{aligned} W''_e(t) &= (1-\lambda_e\,.dt)\left(W'_e(t)\,(r^e - r^*)\,V^*_t\,dt \,+\, W''_e(t+dt)\right) \\ &\quad + \sum_{f\,;f\neq e} \lambda_{ef}\,dt\,W''_e(t) \,+\, o(dt)\,. \end{aligned}$$

Proceeding as above, we obtain the differential equations for the functions W''_e.

The side conditions are obvious.

Exercise 1-19
Goes along the lines of Exercise 18.

Exercise 1-20
(a) Same thing again. Introduce

$$W_t \,=\, \int_t^n e^{-\int_t^\tau r_s\,ds}(r_\tau - r^*)V^*_\tau\,d\tau\,,$$

and write

$$W_t \,=\, (r_t - r^*)V^*_t\,dt \,+\, e^{-r_t\,dt}W_{t+dt} \,+\, o(dt)\,.$$

Apply the direct backward construction to $W_e(t) = \mathbb{E}[W_t \,|\, Y_t = e]$. Start from

$$\begin{aligned} W_e(t) &= (1 - \lambda_{e\cdot}dt)\Big((r^e - r^*)V_t^*\,dt \,+\, e^{-r^e\,dt}W_e(t+dt)\Big) \\ &\quad + \sum_{f;f\neq e} \lambda_{ef}\,dt\,W_f(t) \,+\, o(dt)\,, \end{aligned}$$

and do a small piece of paper-work to arrive at

$$\frac{d}{dt}W_e(t) \,=\, W_e(t)\,r^e \,-\, (r^e - r^*)\,V_t^* \,-\, \sum_{f;f\neq e} \lambda_{ef}(W_f(t) - W_e(t))\,.$$

Side conditions are: $W_e(n-) = 0$, $e = 1, \ldots, J^Y$.

(b) Basically the same exercise as (a).

(c) For discounted cash bonuses use the general formula for higher order moments of present values of payment streams with state-dependent payment intensity and interest intensity.

Forget about the variance of terminal bonus (too cumbersome).

Exercise 80

Items (a) - (e) are rather theoretical and not typical exam questions in ST305. Anyway, we offer something for students who accept only statements that have been firmly proved.

(a) RTI$(T|T)$ is easy to prove:

$$\mathbb{P}[T > s \,|\, T > t] \,=\, \frac{\mathbb{P}[T > \max(s,t)]}{\mathbb{P}[T > t]} \,=\, \begin{cases} \frac{\mathbb{P}[T>s]}{\mathbb{P}[T>t]} &, \quad t < s\,, \\ 1 &, \quad t \geq s\,. \end{cases}$$

This is obviously an increasing (non-decreasing) function of t for fixed s. PQD(T,T) and AS(T,T) then follow.

We could also prove PQD(T,T) directly:

$$\mathbb{P}[T > s\,,\, T > t] = \mathbb{P}[T > \max(s,t)] \geq \mathbb{P}[T > s]\,\mathbb{P}[T > t]\,.$$

A direct proof of AS(T,T) goes as follows: Let $g(s,t)$ and $h(s,t)$ be increasing functions in both arguments. Then $\tilde{g}(t) = g(t,t)$ and $\tilde{h}(t) = h(t,t)$ are increasing functions in t. Thus, marginal association is enough, see notes 'depend-l.pdf'.

(b) Proof of the continuity property of probabilities: Let A_n, $n = 1, 2, \ldots$ be an increasing sequence of events, that is, $A_n \subseteq A_{n+1}$ for all n. Write

$$\cup_{j=1}^{\infty} A_j \,=\, A_1 \,\cup\, \bigcup_{j=2}^{\infty}(A_j \cap A_{j-1}^c)\,,$$

which is a union of mutually exclusive sets. (Here A^c denotes the complement of the event A.) Then

$$\mathbb{P}\left[\cup_{j=1}^{\infty} A_j\right] \quad = \quad \mathbb{P}[A_1] \,+\, \sum_{j=2}^{\infty}(\mathbb{P}[A_j] - \mathbb{P}[A_{j-1}])$$

$$= \lim_{n\to\infty} \left(\mathbb{P}[A_1] + \sum_{j=2}^{n} (\mathbb{P}(A_j) - \mathbb{P}[A_{j-1}]) \right) ,$$

hence

$$\mathbb{P}\left[\cup_{j=1}^{\infty} A_j\right] = \lim_{n\to\infty} \mathbb{P}[A_n] .$$

Let A_n, $n = 1, 2, \ldots$, be a decreasing sequence of events, that is, $A_n \supseteq A_{n+1}$ for all n. Then

$$\mathbb{P}\left[\cap_{j=1}^{\infty} A_j\right] = \lim_{n\to\infty} \mathbb{P}[A_n] .$$

The two statements are equivalent. For instance, the latter follows by applying the former to the increasing sequence A_n^c:

$$\begin{aligned} \mathbb{P}\left[\cap_{j=1}^{\infty} A_j\right] &= 1 - \mathbb{P}\left[\left(\cap_{j=1}^{\infty} A_j\right)^c\right] = 1 - \mathbb{P}\left[\cup_{j=1}^{\infty} A_j^c\right] = 1 - \lim_{n\to\infty} \mathbb{P}\left[A_n^c\right] \\ &= 1 - \lim_{n\to\infty} (1 - \mathbb{P}[A_n]) = \lim_{n\to\infty} \mathbb{P}[A_n] . \end{aligned}$$

Now,

$$\mathbb{P}[S > s\,,\, T > t] = \mathbb{P}\left[\bigcup_n \left(S \ge s + \frac{1}{n}\,,\, T \ge t + \frac{1}{n}\right)\right] = \lim_{n\to\infty} \mathbb{P}\left[S \ge s + \frac{1}{n}\,,\, T \ge t + \frac{1}{n}\right] ,$$

hence

$$\mathbb{P}[S > s]\mathbb{P}[T > t] = \lim_{n\to\infty} \mathbb{P}\left[S \ge s + \frac{1}{n}\right] \mathbb{P}\left[T \ge t + \frac{1}{n}\right] ,$$

and

$$\mathbb{P}[S \ge s\,,\, T \ge t] = \mathbb{P}\left[\bigcap_n \left(S \ge s - \frac{1}{n}\,,\, T \ge t - \frac{1}{n}\right)\right] = \lim_{n\to\infty} \mathbb{P}\left[S \ge s - \frac{1}{n}\,,\, T \ge t - \frac{1}{n}\right] ,$$

hence

$$\mathbb{P}[S \ge s]\mathbb{P}[T \ge t] = \lim_{n\to\infty} \mathbb{P}\left[S > s - \frac{1}{n}\right] \mathbb{P}\left[T > t - \frac{1}{n}\right] .$$

It follows that the strict inequalities $>$ in the definition of PQD and RTI can be replaced by $\ge$. (The definition of AS is no issue here.)

(c) PQD$(-S|T)$ means

$$\mathbb{P}[-S > -s\,,\, T > t] \ge \mathbb{P}[-S > -s]\, \mathbb{P}[T > t]$$

for all s (or all $-s$, which is the same, of course) and all t. This is the same as

$$\mathbb{P}[S < s\,,\, T > t] \ge \mathbb{P}[S < s]\, \mathbb{P}[T > t] ,$$

which is the same as

$$\mathbb{P}[T > t] - \mathbb{P}[S \ge s\,,\, T > t] \ge (1 - \mathbb{P}[S \ge s])\, \mathbb{P}[T > t] ,$$

which is the same as

$$\mathbb{P}[S \ge s\,,\, T > t] \le \mathbb{P}[S \ge s]\, \mathbb{P}[T > t] .$$

Due to the result in (b), this is the same as the asserted result.

(d) Misprint in the exercise text: $\leq$ should be $\geq$. Now, $\mathrm{AS}(-S,T)$ means that

$$\mathbb{C}(g(-S,T),h(-S,T)) \geq 0$$

for all g and h that are increasing in both arguments. But this is equivalent to the asserted result.

(e) $\mathrm{RTI}(-S|T)$ means that $\mathbb{P}[-S > -s \,|\, T > t]$ is increasing in t for fixed s. Rewriting

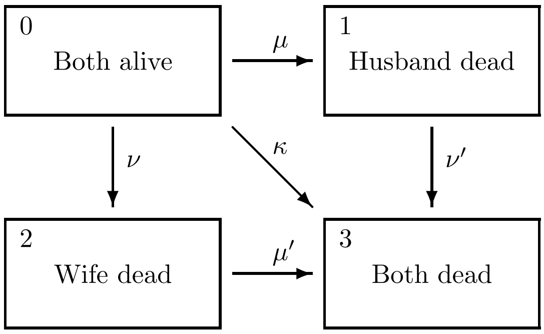
$$\mathbb{P}[-S > -s \,|\, T > t] \;=\; 1 \,-\, \mathbb{P}[S \geq s \,|\, T > t]$$

and recalling the result in (b), we arrive at the asserted result.

(f) The Markov model is sketched in the figure below. Only small amendments are needed in the calculations made in the theory ('depend-l.pdf'), but the results are a bit surprising. We will discuss the matter under the more general assumption that $\mu'_t \geq \mu_t$ and $\nu'_t \geq \nu_t$.

First the case $s \leq t$:

$$\begin{aligned}
\mathbb{P}[S > s \,|\, T > t] &= \frac{e^{-\int_0^t \mu+\nu+\kappa} + \int_s^t e^{-\int_0^\tau \mu+\nu+\kappa}\,\mu_\tau\, e^{-\int_\tau^t \nu'}\,d\tau}{e^{-\int_0^t \mu+\nu+\kappa} + \int_0^t e^{-\int_0^\tau \mu+\nu+\kappa}\,\mu_\tau\, e^{-\int_\tau^t \nu'}\,d\tau} \\
&= 1 - \frac{\int_0^s e^{-\int_0^\tau \mu+\nu+\kappa-\nu'}\,\mu_\tau\,d\tau}{e^{-\int_0^t \mu+\nu+\kappa-\nu'} + \int_0^t e^{-\int_0^\tau \mu+\nu+\kappa-\nu'}\,\mu_\tau\,d\tau}\,.
\end{aligned}$$



We need to discuss this expression as a function of t, which appears only in the denominator. The derivative of the denominator is

$$e^{-\int_0^s \mu+\nu+\kappa-\nu'}\,(\nu'_t - \nu_t - \kappa_t)\,.$$

It follows that, in the presence of a positive κ_t, $\mathbb{P}[S > s \,|\, T > t]$ is not in general an increasing function function of t if $\nu'_t \geq \nu_t$. For $\nu'_t = \nu_t$ it is actually decreasing. We have thus already answered the question and need not look into the case $s < t$.

The second part of the question is now easily sorted out in the case $s < t$ by setting $\mu'_t = \mu_t + \kappa_t$ and $\nu'_t = \nu_t + \kappa_t$ in the result above. We find that $\mathbb{P}[S > s \,|\, T > t]$ is independent of t for $s < t$. Therefore, the RTI issue is so far unsettled and we need to investigate the case $s > t$:

$$\begin{aligned}\mathbb{P}[S > s \,|\, T > t] &= \frac{e^{-\int_0^s \mu+\nu+\kappa} + \int_t^s e^{-\int_0^\tau \mu+\nu+\kappa}\,\nu_\tau\, e^{-\int_\tau^s \mu+\kappa}\,d\tau}{e^{-\int_0^t \mu+\nu+\kappa} + \int_0^t e^{-\int_0^\tau \mu+\nu+\kappa}\,\mu_\tau\, e^{-\int_\tau^t \nu+\kappa}\,d\tau} \\ &= e^{-\int_t^s \kappa}\,\mathbb{P}^*[S > s \,|\, T > t]\,,\end{aligned}$$

where $\mathbb{P}^*$ denotes probability under the independence hypothesis $\mu'_t = \mu_t$ and $\nu'_t = \nu_t$ (see expression in 'depend-l.pdf'). This is an increasing function of t, and we have proved RTI($S|T$).

Exercise 81
See 'EXERC22A.pas' on public folder.

END From 305exsol

Exercise 57
(a) Inspect (2.16) in BL. It should be quite obvious that the cash balance at any time will get bigger if at any time a bigger amount has been deposited on the account (when interest is positive). In particular this is true if a given amount of deposits is being advanced, i.e. payed earlier.

(b) If $U_t > 0$ for some $t \in (0, n)$, then, by right-continuity of U, $U_\tau > 0$ for $\tau \in [t, t+\epsilon)$, some non-degenerate interval, hence $\int_t^{t+dt} U_\tau\, r_\tau\, d\tau > 0$ if r is strictly positive. If, moreover, $U_t \geq 0$ for all $t \in (0, n)$, it follows that $\int_0^n U_\tau\, r_\tau\, d\tau > 0$. Then, if $U_n = 0$, it follows from (2.15) that

$$0 = A_n + \int_0^n U_\tau\, r_\tau\, d\tau\,,$$

and so $A_n < 0$.

Think of a savings account: Deposits are made first, interest is earned on these, and at the end one can withdraw more than the total deposited. For instance, a unit deposited at time 0 grows with interest to $\exp(\int_0^n r)$ in n years, which can then be withdrawn to make the balance nil at time n. In this case $A_t = 1$ for $0 \leq t < n$ and $A_n = 1 - \exp(\int_0^n r)$, which is negative if interest is positive.

Exercise 58
(b)

$$p_{\overline{aa}}(0, t_1)\,\sigma(t_1)\,dt_1\, p_{\overline{ii}}(t_1, t_2)\,\rho(t_2)\,dt_2 p_{\overline{aa}}(t_2, t_3)\,\mu(t_3)\,dt_3 + o(dt)\,,$$

where $dt = max(dt_1, dt_2, dt_3)$. We have used the differentiability of the transition probabilities to write e.g. $p_{\overline{ii}}(t_1 + dt_1, t_2) = p_{\overline{ii}}(t_1, t_2) + o(dt_1)$.

Exercise 61

(a)

$$p_{ai}^{(1)}(0, t + dt) = p_{ai}^{(1)}(0, t)(1 - (\nu(t) + \rho(t))\,dt) \;+\; p_{\overline{aa}}(0, t)\,\sigma(t)\,dt$$

leads to

$$\frac{d}{dt} p_{ai}^{(1)}(0, t) = -p_{ai}^{(1)}(0, t)(\nu(t) + \rho(t)) \;+\; p_{\overline{aa}}(0, t)\,\sigma(t)\,,$$

with side condition

$$p_{ai}^{(1)}(0,0) = 0\,.$$

Integrating gives the following integral expression, which could be put up by direct reasoning:

$$p_{ai}^{(1)}(0,t) = \int_0^t \exp(-\int_0^s (\mu+\sigma))\sigma(s)\,ds\,\exp(-\int_s^t (\nu+\rho))\,.$$

Next,

$$p_{aa}^{(1)}(0,t+dt) = p_{aa}^{(1)}(0,t)(1-(\mu(t)+\sigma(t))\,dt)\;+\;p_{ai}^{(1)}(0,t)\,\rho(t)\,dt$$

leads to

$$\frac{d}{dt}p_{aa}^{(1)}(0,t) = -p_{aa}^{(1)}(0,t)(\mu(t)+\sigma(t))\;+\;p_{ai}^{(1)}(0,t)\,\rho(t)\,,$$

with side condition

$$p_{aa}^{(1)}(0,0) = 0,\,.$$

Integral expression, which could be put up by direct reasoning:

$$p_{aa}^{(1)}(0,t) = \int_0^t p_{ai}^{(1)}(0,s)\,\rho(s)\,ds\,\exp(-\int_s^t (\mu+\sigma))\,.$$

As an exercise, repeat the argument for $k = 2, 3, ...$ to find differential equations for the probability $p_{ai}^{(k)}(0,t)$ of being disabled for the k-th time and the probability $p_{aa}^{(k)}(0,t)$ of being active after having been disabled k times.

One could attack this problem by redefining the state-space of the process as indicated in Figure G.1, where the notation speaks for itself:

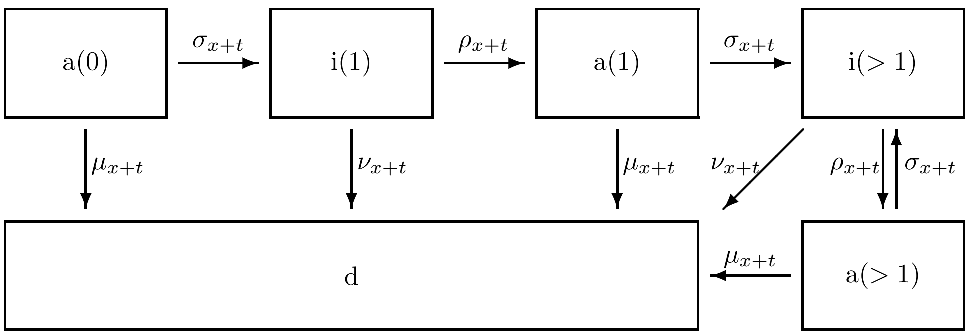


Figure G.1: Problem 7.7

(b)

$$p_{ai}^{(1)}(0,t-q)p_{\overline{ii}}(t-q,t) = \int_0^{t-q} p_{\overline{aa}}(0,\tau)\,\sigma_{x+\tau}\,p_{\overline{ii}}(\tau,t)\,d\tau\,.$$

You should be able to interpret the integral expression as a sum of probabilities of mutually exclusive favourable events.

(c)

$$\pi = \frac{\int_q^n e^{-r\tau} p_{ai}^{(1)}(0,\tau - q) p_{\overline{ii}}(\tau - q, \tau)\, d\tau}{\int_0^n e^{-r\tau} p_{\overline{aa}}(0,\tau)\, d\tau}\,.$$

Comment: The premium plan is unacceptable in practice since it will produce a negative reserve if the insured is in premium paying state after time $n - q$ (then, for sure, no benefits will be received, but premium will still be paid). Therefore, the premiums should be paid only over a shorter period and certainly not after time $n - q$.
Reserve: $\int_t^{t+q} e^{-r(\tau - t)} p_{\overline{ii}}(t,\tau)\, d\tau \; + \; \int_{t+q}^n e^{-r(\tau-t)} p_{ii}(t, \tau - q) p_{\overline{ii}}(\tau - q, \tau)\, d\tau\,.$

Exercise 62

(a) Use Kolmogorov forward, taking advantage of the obvious relationship $p_{aa}(t,u) = 1 - p_{ai}(t,u)$:

$$p_{ai}(t, u + du) = p_{ai}(t,u)(1 - \rho\, du) + (1 - p_{ai}(t,u))\sigma\, du\,.$$

You obtain a (very simple) differential equation for $p_{ai}(t,\,\cdot\,)$, which you integrate from t to u using $p_{ai}(u,u) = 0$, and find

$$p_{ai}(t,u) = \frac{\sigma}{\rho+\sigma} - \frac{\sigma}{\rho+\sigma} e^{-(\rho+\sigma)(u-t)}\,. \tag{G.39}$$

Then calculate $p_{aa}(t,u) = 1 - p_{ai}(t,u)$:

$$p_{aa}(t,u) = \frac{\rho}{\rho+\sigma} + \frac{\sigma}{\rho+\sigma} e^{-(\rho+\sigma)(u-t)}\,. \tag{G.40}$$

The transition probabilities depend only on $u - t$ due to homogeneity. Discuss the probabilities as functions of $u - t$ and look at the limits as $u - t$ tends to $+\infty$.

(b) Thiele is actually not a good idea (it is doable, of course, but being a backward equation it does not make use of the fact that $p_{aa}(t,u) + p_{ai}(t,u) = 1$). Since we have an explicit expression for $p_{ai}(t,u)$, the easiest way is to calculate

$$\mathbb{E}\left[\int_0^t I_i(\tau)\, d\tau\right] = \int_0^t \mathbb{E}[I_i(\tau)]\, d\tau = \int_0^t p_{ai}(0,\tau)\, d\tau = \frac{\sigma}{\rho+\sigma}\, t - \frac{\sigma}{(\rho+\sigma)^2}\left(1 - e^{-(\rho+\sigma)t}\right) \tag{G.41}$$

(c)

$$\mathbb{E}\left[N_{ai}(t)\right] = \mathbb{E}\left[\int_0^t dN_{ai}(t)\right] = \int_0^t p_{aa}(0,\tau)\,\sigma\, d\tau = \frac{\rho\,\sigma}{\rho+\sigma}\, t \, + \, \frac{\sigma^2}{(\rho+\sigma)^2}\left(1 - e^{-(\rho+\sigma)t}\right) \tag{G.42}$$

(d) From (G.41) the expected proportion of inactive time in t years is

$$\frac{\sigma}{\rho+\sigma} - \frac{\sigma}{(\rho+\sigma)^2} \frac{1 - e^{-(\rho+\sigma)t}}{t}\,. \tag{G.43}$$

The function

$$g(x) = \frac{1 - e^{-x}}{x} \tag{G.44}$$

is 1 for $x = 0$ (l'Hospital). It has derivative

$$g'(x) = \frac{e^{-x}}{x} - \frac{1-e^{-x}}{x^2} = -e^{-x}\frac{e^x - 1 - x}{x^2},$$

which is < 0 for $x > 0$ (Taylor expansion), hence g is decreasing. Thus, as $t \to +\infty$, the proportion in (G.43) is increasing from 0 to

$$\frac{\sigma}{\rho + \sigma}.$$

Reasonable: Increasing function of σ, decreasing function of ρ.

From (G.42) the expected number of onsets of invalidity per time unit in t years is

$$\frac{\rho\,\sigma}{\rho+\sigma} + \frac{\sigma^2}{(\rho+\sigma)^2}\frac{1-e^{-(\rho+\sigma)t}}{t}.$$

As t increases from 0 to $+\infty$ the value of this expression decreases from σ (of course!) to $(\rho\,\sigma)/(\rho+\sigma)$. It is interesting to note that the limiting expression here is symmetric in ρ and σ. You should try and figure why.

(e) If ρ were 0, we could pick the result from Item 7.2 (e) above, setting $\mu = \nu = 0$:

$$\tilde{\sigma}(t) = \frac{\sigma}{1 - e^{-\sigma\,(n-t)}}, \qquad 0 < t < n. \tag{G.45}$$

In general we have

$$\tilde{\sigma}(t) = \sigma\frac{p_{ii}(t,n)}{p_{ai}(t,n)} = \frac{\sigma + \rho e^{-(\rho+\sigma)(n-t)}}{1 - e^{-(\rho+\sigma)(n-t)}}, \qquad 0 < t < n. \tag{G.46}$$

Here we have used (with a view to (G.40), just switch roles of σ and ρ)

$$p_{ii}(t,u) = \frac{\sigma}{\rho+\sigma} + \frac{\rho}{\rho+\sigma}e^{-(\rho+\sigma)(u-t)}.$$

(f) Constant intensity of transition (whether $a \to i$ or $i \to a$) means that the transitions are generated by a Poisson process. N_{ai} and N_{ia} count every second transition: N_{ai} counts transition No. 1,3,... and so on, i.e. $N_{ai}(t)$ is distributed as $[N(t)+1]/2$, where $N(t)$ is a Poisson variate with parameter $\sigma\,t$. N_{ia} counts transition No. 2,4,... and so on, i.e. $N_{ia}(t)$ is distributed as $[N(t)]/2$, where $N(t)$ is a Poisson variate with parameter $\sigma\,t$.

Exercise 64

Problem 9.3

(a) To simplify notation, put $N_{ai} = N_{ai}(n)$ and $N_{ia} = N_{ia}(n)$. The likelihood is

$$\Lambda = \sigma^{N_{ai}}\rho^{N_{ia}}e^{-\sigma W_a - \rho W_i},$$

where $W_a = \sum_{\ell=1}^{m} T_n^{(m)}$, the total time spent in active state (and $W_i = nm - W_a$ the total time spent in inactive state).

$$\frac{\partial}{\partial\sigma}\ln\Lambda = \frac{N_{ai}}{\sigma} - W_a,$$

$$\frac{\partial^2}{\partial\sigma^2}\ln\Lambda = -\frac{N_{ai}}{\sigma^2}\,,$$

plus similar expressions for derivatives w.r.t. ρ, and

$$\frac{\partial^2}{\partial\sigma\partial\rho}\ln\Lambda = 0\,.$$

Thus, the ML estimators are the occurrence-exposure rates

$$\hat{\sigma} = \frac{N_{ai}}{W_a}\,, \qquad \hat{\rho} = \frac{N_{ia}}{W_i}\,,$$

which are asymptotically independent, unbiased, and normally distributed with asymptotic variances

$$\text{as.Var}\hat{\sigma} = \frac{\sigma^2}{\mathrm{E}N_{ai}}\,, \qquad \text{as.Var}\hat{\rho} = \frac{\rho^2}{\mathrm{E}N_{ia}}\,. \tag{G.47}$$

By (G.40),

$$\begin{aligned} \mathrm{E}N_{ai} &= m\int_0^n p_{aa}(0,\tau)\,\sigma\,d\tau \qquad (G.48)\\ &= m\,\sigma\int_0^n \left(\frac{\rho}{\rho+\sigma} + \frac{\sigma}{\rho+\sigma}e^{-(\rho+\sigma)\tau}\right)d\tau\\ &= m\,\sigma\left(\frac{\rho}{\rho+\sigma}\,n + \frac{\sigma}{(\rho+\sigma)^2}\left(1-e^{-(\rho+\sigma)n}\right)\right)\,, \qquad (G.49) \end{aligned}$$

hence

$$\text{as.Var}\hat{\sigma} = \frac{\sigma\,(\rho+\sigma)}{m\,\left(\rho\,n + \frac{\sigma}{\rho+\sigma}\,(1-e^{-(\rho+\sigma)n})\right)}\,. \tag{G.50}$$

Similarly

$$\text{as.Var}\hat{\rho} = \frac{\rho\,(\rho+\sigma)}{m\,\sigma\,\left(n - \frac{1}{\rho+\sigma}\,(1-e^{-(\rho+\sigma)n})\right)}\,. \tag{G.51}$$

(b) One can discuss the explicit expression (G.51), but there is an easier way: By (G.47) and (G.48)

$$\text{as.Var}\hat{\sigma} = \frac{\sigma}{m\int_0^n p_{aa}(0,\tau)\,d\tau}\,, \tag{G.52}$$

so it suffices to show that $p_{aa}(0,\tau)$ is an increasing function of ρ or, what is the same, an increasing function of $\rho+\sigma$. By (G.40)

$$p_{aa}(0,\tau) = \frac{\rho}{\rho+\sigma} + \frac{\sigma}{\rho+\sigma}e^{-(\rho+\sigma)\,\tau} = 1 - \sigma\,\tau\,\frac{1-e^{-(\rho+\sigma)\,\tau}}{(\rho+\sigma)\,\tau}\,,$$

so you need only recall that the function g in (G.44) is decreasing.

(c) General comment: The theory of conditional Markov chains worked out in Problem 7.2 is often needed in statistical analysis of insurance data where one does not have access to the complete 'life histories' of the policies. It is often the case that one must work with data that are selected somehow. For instance, suppose data on disabilities can be seen only from the claims records of those that are currently disabled. Then the relevant probabilities and intensities are the conditional one, given that the process is now in disabled state.

In the present situation you must therefore work with the intensities in 7.8 (e).

Exercise 66

To conform with the notation in 'Basic Life Insurance Mathematics', let us call the number of lives n instead of m. We will throughout refer to formulas in the general theory in 'Basic Life Insurance Mathematics'.

Assume piece-wise constant mortality intensity, see (9.58):

$$\mu(t) = \mu_q\,, \qquad q-1 \leq t < q\,, j = 1, 2, \ldots$$

The log likelihood (9.53) is

$$\ln \Lambda = \sum_q (\ln \mu_q N_q - \mu_q W_q)\,,$$

where N_q and W_q are, respectively, the number of deaths and the total time spent alive in the age interval $[q-1, q)$.

Each μ_q is a parameter which is functionally unrelated to all the others, so there are many parameters in this model! For instance, if we are interested in mortality up to age 100 and have data in the age range from 0 to 100, there are 100 parameters, which is quite a lot. Remember, however, that this model is just a first step in a two-stage procedure where the second step is to graduate (smoothen) the ML estimators resulting from the present naive model with piece-wise constant intensity.

The ML estimators are the occurrence-exposure rates

$$\hat{\mu}_q = \frac{N_q}{W_q}\,,$$

which are well defined for all q such that $W_q > 0$ (i.e. in age intervals where there were survivors exposed to risk of death). The $\hat{\mu}_q$ are asymptotically (as n increases) normally distributed, mutually independent, unbiased, and with variances given by

$$\sigma_q^2 = \text{as.}\mathbb{V}[\hat{\mu}_q] = \frac{\mu_q}{\mathbb{E}[W_q]}\,, \tag{G.53}$$

where the expected exposure is

$$\mathbb{E}[W_q] = \sum_{m=1}^{n} \int_{q-1}^{q} p^{(m)}(\tau)\, d\tau\,,$$

$p^{(m)}(\tau)$ being the probability that individual No. m is alive and under observation at time τ.

The variance σ_q^2 is inversely proportional to the corresponding expected exposure. In the present simple model, with only one intensity of transition from the state 'alive' to the absorbing state 'dead', we find explicit expressions for the expected exposure.

For instance, suppose we have observed each individual life from birth until death or until attained age 100, whichever occurs first (i.e. censoring at age 100). Then, for $\tau \in [q-1, q)$ with $q = 1, \ldots, 100$, we have

$$\begin{aligned} p^{(m)}(\tau) &= \exp\left(-\int_0^\tau \mu(s)\, ds\right) \\ &= \exp\left(-\sum_{p=1}^{q-1} \mu_p - (\tau - (q-1))\, \mu_q\right), \end{aligned} \tag{G.54}$$

hence

$$\begin{aligned} \mathbb{E}[W_q] &= n \int_{q-1}^{q} \exp\left(-\sum_{p=1}^{q-1} \mu_p - (\tau - (q-1))\, \mu_q\right) d\tau \\ &= n \exp\left(-\sum_{p=1}^{q-1} \mu_p\right) \frac{1 - \exp(-\mu_q)}{\mu_q}, \end{aligned}$$

and

$$\sigma_q^2 = \frac{1}{n} \frac{\mu_q^2}{\exp\left(-\sum_{p=1}^{q-1} \mu_p\right) (1 - \exp(-\mu_q))}, \tag{G.55}$$

You should look at other censoring schemes and discuss the impact of censoring on the variance. Take e.g. the case where person No m enters at age $z^{(m)}$ and is observed until death or age 100, whichever occurs first (all $z^{(m)}$ less than 100).

Estimators $\hat{\sigma}_q^2$ of the variances are obtained upon replacing the μ_j in (G.55) by the estimators $\hat{\mu}_j$. Simpler estimators are obtained by just replacing μ_q and $\mathbb{E}[W_q]$ in (G.53) with their straightforward empirical counterparts: $\hat{\sigma}_q^2 = \hat{\mu}_q / W_q = N_q / W_q^2$.

Now to graduation. The occurrence-exposure rates will usually have a ragged appearance. Assuming that the real underlying mortality intensity is a smooth function, we therefore will fit a suitable function to the occurrence-exposure rates. Suppose we assume that the true mortality rate is a Gompertz Makeham function, $\mu(t) = \alpha + \beta e^{\gamma t}$. Then, take some representative age ξ_q (typically $\xi_q = q - 0.5$) in each age interval and fit the parameters α, β, γ by minimizing a weighted sum of squared errors

$$Q = \sum_q a_q \left(\hat{\mu}_q - \alpha - \beta e^{\gamma t}\right)^2.$$

This is a matter of non-linear regression. Optimal weights a_q are the inverse of the variances, but since these are unknown, we plug in the estimators and use $a_q = 1/\hat{\sigma}_q^2$.

Belongs to what?:
For fixed $w = mn$ we have $n = w/m$ and

$$\text{as.Var}\hat{\sigma} = \frac{\sigma}{w(1 - \frac{\pi_{w/m}}{w/m})},$$

which is a decreasing function of m by the results in item (b).

We also find

$$\text{as.Var}\hat{\rho} = \frac{\rho^2}{\text{E}N_{ia}} = \frac{\rho}{w \frac{\pi_{w/m}}{w/m}},$$

where we have used

$$\mathrm{E}N_{ia} = m\int_0^n p_{ai}(0,\tau)\rho\, d\tau = mn\rho\frac{\pi_n}{n}\,.$$

We see that as.Var$\hat{\rho}$ is an increasing function of m for fixed w.

Comment: The asymptotic variance of an intensity estimator is better the longer the total expected time spent in the state from which the the relevant transition is made. All policies start from state a at time 0. For fixed total exposure the estimation of σ will be good for many policies observed in a short time (when they are likely to remain active), and estimation of ρ will be good for few policies observed over a long time when they can make it to inactive state.

Exercise 67

If r is constant, then

$$\pi = \frac{1}{\bar{a}_{x\,\overline{n}|}} - r\,,$$

which is a decreasing function of n.

For $t < n$ the premium reserve is

$$V_t = 1 - \frac{\bar{a}_{x+t\,\overline{n-t}|}}{\bar{a}_{x\,\overline{n}|}} = 1 - \frac{1}{{}_tE_x}\frac{{}_{x+t\,|\,n-t}\bar{a}_x}{\bar{a}_{x\,\overline{t}|} + {}_{x+t\,|\,n-t}\bar{a}_x}\,.$$

For fixed t this is a decreasing function of n.

In particular, a whole life insurance has smaller premium than an n-year temporary endowment insurance, and it has also smaller reserve for $t < n$.

Exercise 3

Follows from () and the fact that U is right-continuous, hence strictly positive in some interval. The result says that total withdrawals exceed total deposits, which is due to earned interest.

Exercise 3

(a) Prove (3.11) along the following lines: By definition,

$$\mathbb{E}[G(T)] = G(0)F(0) + \int_0^\infty G(t)\, dF(t)\,. \tag{G.56}$$

Observe that

$$\int_0^\infty G(t)\, dF(t) = -\int_0^\infty G(t) d\bar{F}(\tau)\,. \tag{G.57}$$

Integrate by parts to obtain

$$\int_0^n G(t)\, d\bar{F}(t) = G(n)\bar{F}(n) - G(0)\bar{F}(0) - \int_0^n \bar{F}(t-)\, dG(t)\,. \tag{G.58}$$

Assuming first that G is non-negative and non-decreasing, we have

$$G(n)\bar{F}(n) = G(n)\int_n^\infty dF(t) \le \int_n^\infty G(t)\, dF(t) \to 0\,,$$

hence $G(n)\bar{F}(n) \to 0$, as $n \to \infty$. The same holds true for an integrable G of finite variation, which is the difference between two integrable non-decreasing functions. Thus, letting $n \to \infty$ in (G.58), we obtain

$$\int_0^\infty G(t)\, d\bar{F}(t) = -G(0)\bar{F}(0) - \int_0^\infty \bar{F}(t-)\, dG(t)\,. \tag{G.59}$$

Now combine (G.56) – (G.59) to arrive at (3.11).

Exercise 1

Assume that X is non-decreasing. Note that

$$\mathcal{D}(X) = \cup_{m=1}^\infty \cup_{n=1}^\infty \mathcal{D}_{m,n}\,,$$

where $\mathcal{D}_{m,n} = \{t\,;\, t \le n\,,\, X_{t+} - X_{t-} \ge \frac{1}{m}\}$. Since

$$\infty > X_n \ge X_0 + \sum_{t \in \mathcal{D}_{m,n}} (X_{t+} - X_{t-}) > X_0 + \frac{1}{m}\, \sharp\mathcal{D}_{m,n}\,,$$

we conclude that $\sharp\mathcal{D}_{m,n}$ is finite. Thus, being a countable union of finite sets, $\mathcal{D}(X)$ is (at most) countable.

Exercise 2

$$dX_t^q = qX_t^{q-1}\, x_t^c\, dt + \Big((X_{t-} + x_t^d)^q - X_{t-}^q\Big)\, dN_t\,.$$