

Seven Introductory Lectures on Actuarial Mathematics

Mogens Steffensen

5. februar 2008

Indhold

1	Compound Interest	3
1.1	Time Value of Money - Single Payments	3
1.2	Time value of money - Streams of Payments	5
1.3	Fair Payment Streams	6
1.4	Particular Payment Streams	7
1.5	Fractionated Interest and Payments	9
1.6	Continuous Interest and Payments	10
1.7	Exercises	11
2	Life Insurance Risk: Mortality	12
2.1	The newborn's lifetime	12
2.2	The x -year old's residual lifetime	13
2.3	Expected values	16
2.4	Comments	17
2.5	Exercises	18
3	Non-life Insurance Risk: Frequency and Severity	20
3.1	Compound risk models	20
3.2	Frequency risk	22
3.3	Severity risk	24
3.4	Exercises	26
4	Valuation principles	28
4.1	Properties	29
4.2	Principles	30
4.3	Exercises	34
5	Life Insurance: Products and Valuation	34
5.1	Single Payments	34
5.2	Streams of Payments	36
5.3	Fair Payment Streams	37

5.4	Particular Payment Streams	37
5.5	Exercises	40
6	Non-life Insurance: The Aggregate Claim Distribution	41
6.1	The Panjer Recursion	43
6.2	Normal Power Approximation	46
6.3	Ruin Theory	46
6.4	Exercises	47
7	Finance and Reinsurance: Products and Valuation	49

1 Compound Interest

This chapter deals with the value of a unit paid at one point in time measures at another point in time. It is the supply and demand for capital that determines this value, typically different from one unit itself. The supply and demand for capital is driven by individuals and corporates who do not necessarily spend their money at the same time point as they earn it. A paramount role of the financial market is act as intermediary between participants who earn more money than they spend and participants who earn less money than they spend at some time point.

Individuals typically earn less money than they spend in their younger years when they are under education and in their older years when they have left the labor market. On the other hand, they earn more money than they spend during their working years. The financial market provides products that enables the individuals to spread out a work life of production earnings to a whole life of consumption. Some political systems with education and social security partly paid via taxes have an embedded feature of stretching out the earned money over the full life time. These systems typically also serve to partly level out the differences in life time income between individuals. This happens e.g. when the members do not contribute to the education system by payment of taxes to the same extent as they have charged the education system themselves. In other political systems the decisions in this respect are to a larger extent left to the individuals and one does not seek the same extent of levelling out differences. The way and to which extent the individual stretches out his life time earnings over a whole life of consumption is the subject matter of *personal finance*. Of course, financial products that exchange risk in connection with an uncertain life time play an important role in personal finance. Life insurance and pension saving products are primary examples of such.

Corporates also typically spend more money than they earn at the beginning. A new-born corporate invests in products, projects and ideas and hope that it is worth it. Therefore, they go to the financial market to raise capital for their projects. However, there are differences in your demand for capital depending on what kind of corporate you are: Corporates whose business is research and development like e.g. biotech corporates, can sometimes run decades without earning any money and therefore need injection of more capital than others who produce and earn money already from day one. In what way and to which extent corporates finance their investments is the subject matter of *corporate finance* which is an important discipline of financial economics. Some economists speak of corporate cycles or business cycles of different industries with a terminology similar to that of life cycles of human beings. This idea is taken to its extreme in the financial economics discipline of *evolutionary finance*.

1.1 Time Value of Money - Single Payments

In this section we derive the present value of past and future single payments based on accrual of interest.

We consider the payment of the amount c and call the time point at which this payment takes place time 0. We assume that this payment is paid to an institution, e.g. a bank, where it earns interest. The institution therefore sees the payment as a *contribution* c which explains the choice of letter. We assume that the institution announces at the beginning of period t the interest rate $i(t)$ for that period. Given the process of interest rates $(i(1), \dots, i(n))$ we can calculate the value

of the amount at later points in time and we call these values present values at future time points. The present value at time 0 is c itself. After one period the present value is the value at time 0, c , plus the interest earned in period 1, $ci(1)$, i.e. in total $c(1+i(1))$. This value now earns interest in the second period such that the value after two periods equals $c(1+i(1))(1+i(2))$, and so on. The present value at time t can be written as

$$U(t) = c \prod_{s=1}^t (1+i(s)) = c(1+i(1))(1+i(2)) \cdots (1+i(t)).$$

The process U can be studied and it is easily seen that we can characterize this process by the initial value $U(0) = c$ and the iterative function

$$U(t) = (1+i(t))U(t-1).$$

This iterative function can also be written in terms of a so-called difference equation

$$\Delta U(t) = U(t) - U(t-1) = i(t)U(t-1),$$

that expresses the change in value as a linear function of the value itself. The value $U(t)$ is the cash balance at time t of c which has been paid into the institution at the past time point 0. The institution invests this cash balance on a financial market. It is the investment gains from this investment that allows the institution to pay interest rates.

Now we consider the payment of b and call the time point at which this payment takes place time n . We assume that this payment is paid *from* the institution. The institution therefore sees the payment as a *benefit*. Since the institution announces interest rates also in the future periods we can also speak of the present value at time $t < n$ of b . This is the amount which, after accrual of interest from time t to time n turns into b . The present value at time n is b itself. One period before, the present value is $(1+i(n))^{-1}b$ since this amount at time $n-1$ after the final period of interest payments turns into the amount itself $(1+i(n))^{-1}b$ plus the interest earned $i(n)(1+i(n))^{-1}b$, i.e. in total $(1+i(n))^{-1}b + i(n)(1+i(n))^{-1}b = (1+i(n))(1+i(n))^{-1}b = b$. Two periods before time n , the value is $(1+i(n))^{-1}(1+i(n-1))^{-1}b$, and so on. The present value at time $t < n$ can be written as

$$V(t) = b \prod_{s=t+1}^n \frac{1}{1+i(s)} = \frac{1}{1+i(n)} \frac{1}{1+i(n-1)} \cdots \frac{1}{1+i(t+1)} b.$$

The process V can be studied and it is easily seen that we can characterize this process by the terminal value $V(n) = b$ and the recursive function

$$V(t) = \frac{1}{1+i(t+1)} V(t+1).$$

This backward updating function can also be written in terms of a difference equation

$$\Delta V(t) = V(t) - V(t-1) = \frac{i(t)}{1+i(t)} V(t). \quad (1)$$

The value $V(t)$ is the present value at time t of the payment b which will be paid out from the institution at the future time point n . The financial institution is interested in knowing this amount because it expresses somehow what the institution owes to the receiver of the payment. This is the

amount that the institution has to reserve for the obligation to pay b at time n . We also call V for the reserve.

The processes U and V become particularly simple if the institution works with the same interest rate each period, i.e. $i(t) = i$. Introducing the notation

$$\begin{aligned} v &= \frac{1}{1+i}, \\ d &= 1-v = \frac{i}{1+i}, \end{aligned}$$

we can write

$$\begin{aligned} U(t) &= c(1+i)^t, \\ V(t) &= v^{n-t}b, \end{aligned}$$

and characterize these processes by their initial conditions, $U(0) = c$ and $V(n) = b$, respectively, and the iterative and recursive functions, respectively, and the corresponding difference equations,

$$\begin{aligned} U(t) &= (1+i)U(t-1) \text{ or } \Delta U(t) = iU(t-1), \\ V(t) &= vV(t+1) \text{ or } \Delta V(t) = dV(t). \end{aligned}$$

Multiplication of $(1+i)^t$ in U is called accumulation of interest. The factor $(1+i)^t$ is the *accumulation factor from 0 to t* and $(1+i)$ is the *yearly accumulation factor*. Multiplication of v^{n-t} in V is called discounting. The factor v^{n-t} is the *discount factor from n to t* and v is the *yearly discount factor*.

1.2 Time value of money - Streams of Payments

In this section we derive the present value of past and future streams of payments based on accrual of interest. Throughout we assume that the interest rate is constant.

Now we assume that at each time point $0 \leq t \leq n$, the payment $c(t) - b(t)$ is paid into the institution. The (invisible) plus in front of c reflects that c is a contribution to the institution. The minus in front of b reflects that this is a benefit going out of the institution. The payment $c(t) - b(t)$ is, so to speak, the net contribution (contribution minus benefit) at time t . We can now consider a stream of such net contributions formalized by an $n+1$ -dimensional vector, $c - b = (c(0) - b(0), c(1) - b(1), \dots, c(n) - b(n))$.

We can calculate at each time point t the present value of all payments in $c - b$ in the past by summing up all the present values corresponding to each past single net contribution,

$$U(t) = \sum_{s=0}^t (c(s) - b(s)) (1+i)^{t-s}.$$

Summing until time t in the calculation of U means that we mean past payments in a non-strict sense such that the present payment of $c(t) - b(t)$ is taken into account in the present value at time t . The process U can be characterized by the initial point $U(0) = c(0) - b(0)$ and the forward updating equation

$$U(t) = (1+i)U(t-1) + c(t) - b(t),$$

or, equivalently, the difference equation

$$\Delta U(t) = iU(t-1) + c(t) - b(t).$$

The difference equation expresses the so-called dynamics of the bank account: Interest is added, contributions are paid into the bank account, and benefits are paid out of the bank account.

Now we consider the future payments of net benefits in the payments stream $b - c$. We can calculate at each time point t the present value of all net benefits $b - c$ in the future by summing up the present value corresponding to each future single net benefit payment,

$$V(t) = \sum_{s=t+1}^n (b(s) - c(s)) v^{s-t}. \quad (2)$$

Summing from time $t + 1$ for calculation of V means that we mean future payments in a strict sense such that the present payment of $c(t) - b(t)$ is not taken into account in the present value at time t . This does not conform with the previously introduced V where the present value at time n was taken to be b although this payment does not fall due in the strict future standing at time n . In a strict future interpretation we have instead that $V(n) = 0$. In order to make the backward system work correct we would then have to specify the first step as

$$V(n-1) = vb,$$

from where the backward updating equation (1) works.

With this strict future interpretation we are now ready to characterize the process V defined in (2) by the terminal condition $V(n) = 0$ and then the backward updating equation

$$V(t) = v(b(t+1) - c(t+1) + V(t+1)),$$

or, equivalently,

$$\Delta V(t) = dV(t) + v(c(t) - b(t)).$$

1.3 Fair Payment Streams

In this section we put up constraints on the payment streams in terms of the processes U and V .

Consider a contract holder who pays the contributions c and receives the benefits b until time n . The cash balance on his account is then $U(n)$ at time n when the contract closes. If $U(n) > 0$ he would feel cheated by the institution because it cashes the balance on his account. Otherwise, if $U(n) < 0$ the institution would feel cheated by the contract holder since it has to cover a loss on the account. So, the only way that both parties will be satisfied is if

$$U(n) = 0.$$

We speak of this as a fairness criterion that we can impose on the payment stream. The fairness criterion can be written as

$$\sum_{t=0}^n (c(t) - b(t)) (1+i)^{n-t} = 0,$$

which by multiplication on both sides with $(1+i)^{-n}$ gives

$$\sum_{t=0}^n (c(t) - b(t)) (1+i)^{-t} = 0.$$

Splitting the sum on the left hand side into the term for $t = 0$ and a sum from $t = 1$ to n allows us to write this as $c(0) - b(0) - V(0)$. Thus, we can also formulate the fairness criterion in terms of the present value of future payments V ,

$$V(0) = c(0) - b(0). \quad (3)$$

This means that the institution owes a value to the contract holder which is equivalent to the net contribution at time 0.

Now, splitting up the sum in the definition of V into a sum until time t and a sum from time t , the fairness criterion can be written as

$$\sum_{s=0}^t (c(s) - b(s)) (1+i)^{-s} + \sum_{s=t+1}^n (c(s) - b(s)) (1+i)^{-s} = 0.$$

Moving the sum from time t to the right hand side and multiplying by $(1+i)^t$ gives

$$\sum_{s=0}^t (c(s) - b(s)) (1+i)^{t-s} = \sum_{s=t+1}^n (b(s) - c(s)) (1+i)^{-(s-t)}$$

which can also be written as

$$U(t) = V(t).$$

Thus, if the fairness criterion is fulfilled then the present value of past net contributions (contributions minus benefits) equals the present value of the future net benefits (benefits minus contributions) at every point in time. Thinking of the bank account it sounds very reasonable that the cash balance on the account is the same value that the institution owes to the contract holder. It is important to note that this holds also in situations where we do not know the payments and the interest in the future. In this case we can not, without any fairness criterion, calculate V because its elements are unknown. But if we impose the fairness criterion on these unknown elements, V is easily calculated since it equals U .

Consider the fairness criterion (3) for a given stream of payments. For a constant interest rate we can consider this as one equation with one unknown, i . If there exists one and only one solution to this equation we call this solution for the internal interest rate. There are a lot of situations, however, where the stream of payments allows for zero solutions or more than one solution.

1.4 Particular Payment Streams

In this section we present a series of standard payment streams which are so common that their present value have their own notation.

All payment streams in this section are so-called elementary annuities which means that they pay a series of payments of one unit. We shall make use of a small but important result on geometric sums. From the derivation

$$\begin{aligned} (1-v) \sum_{t=0}^{n-1} v^t &= \sum_{t=0}^{n-1} v^t - v \sum_{t=0}^{n-1} v^t \\ &= \sum_{t=0}^{n-1} v^t - v \sum_{t=1}^n v^{t-1} \\ &= \sum_{t=0}^{n-1} v^t - \sum_{t=1}^n v^t \\ &= 1 + v + v^2 + \dots + v^{n-1} - (v + v^2 + \dots + v^{n-1} + v^n) \\ &= 1 - v^n, \end{aligned}$$

we get by dividing on both sides with $(1-v)$ that

$$\sum_{t=0}^{n-1} v^t = \frac{1-v^n}{1-v}.$$

1. Annuity in arrear: Consider $b = (0, 1, \dots, 1) = (0, 1^{1 \times n})$ (and $c = 0$). Then

$$\begin{aligned} a_{\overline{n}|} &= V(0) \\ &= \sum_{t=1}^n v^t = v \sum_{t=1}^n v^{t-1} \\ &= v \sum_{t=0}^{n-1} v^t = v \frac{1-v^n}{1-v} = \frac{1-v^n}{i}. \end{aligned}$$

It is easily seen that

$$V(t) = a_{\overline{n-t}|}.$$

2. Annuity-due: Consider $b = (1, \dots, 1, 0) = (1^{1 \times n}, 0)$ (and $c = 0$). Then

$$\begin{aligned} \ddot{a}_{\overline{n}|} &= V(0) + 1 \\ &= \sum_{t=0}^{n-1} v^t = \frac{1-v^n}{1-v} \\ &= (1+i) a_{\overline{n}|}. \end{aligned}$$

The annuity is defined as $V(0) + 1$ such that it includes the value of the one unit paid at the present time 0. This is excluded from V which only takes into account payments in the strict future. It is easily seen that

$$V(t) + 1 = \ddot{a}_{\overline{n-t}|}.$$

3. Deferred annuity in arrear: Consider $b = (0, \dots, 0, 1, \dots, 1) = (0^{1 \times k}, 0, 1^{1 \times n})$ (and $c = 0$). Then

$$\begin{aligned} {}_k a_{\overline{n}|} &= V(0) \\ (a) &= \sum_{t=k+1}^{k+n} v^t = \sum_{t=1}^n v^{t+k} = v^k a_{\overline{n}|} \\ (b) &= \sum_{t=1}^{k+n} v^t - \sum_{t=1}^k v^t = a_{\overline{k+n}|} - a_{\overline{k}|}. \end{aligned}$$

It is easily seen that

$$V(t) = \begin{cases} {}_{k-t} a_{\overline{n}|}, & t < k, \\ a_{\overline{n-(t-k)}|}, & t \geq k. \end{cases}$$

4. Deferred annuity-due: Consider $b = (0, \dots, 0, 1, \dots, 1, 0) = (0^{1 \times k}, 1^{1 \times n}, 0)$ (and $c = 0$). Then

$$\begin{aligned} {}_k \ddot{a}_{\overline{n}|} &= V(0) \\ (a) &= \sum_{t=k}^{k+n-1} v^t = \sum_{t=0}^{n-1} v^{t+k} = v^k \ddot{a}_{\overline{n}|} \\ (b) &= \sum_{t=0}^{k+n-1} v^t - \sum_{t=0}^{k-1} v^t = \ddot{a}_{\overline{k+n}|} - \ddot{a}_{\overline{k}|}. \end{aligned}$$

It is easily seen that

$$V(t) = \begin{cases} {}_{k-t} \ddot{a}_{\overline{n}|}, & t < k, \\ \ddot{a}_{\overline{n-(t-k)}|} - 1, & t \geq k. \end{cases}$$

1.5 Fractionated Interest and Payments

In this section we see what happens if we let interest be accumulated and payments be paid more frequently than one time in every period.

Until now we have only considered the case where interest is accumulated with the same frequency as it is announced, namely one time in every period. However, often the institution announces e.g. a yearly interest rate which is accumulated monthly. Let us split up each period in m subperiods. Announcing a periodwise interest rate $i^{(m)}$ accumulated m times in each period means that actually the interest rate $i^{(m)}/m$ is accumulated in every subperiod. Thus, after the first subperiod a payment of 1 at time 0 has the value $(1 + i^{(m)}/m)$ and after m subperiods corresponding to one period the value is

$$\left(1 + \frac{i^{(m)}}{m}\right)^m.$$

We call $i^{(m)}$ the nominal interest rate. The yearly interest rate corresponding to the nominal interest rate $i^{(m)}$, i.e. the solution i to

$$1 + i = \left(1 + \frac{i^{(m)}}{m}\right)^m, \quad (4)$$

is called the effective interest. For a given effective interest rate we can of course calculate the nominal interest rate corresponding to m accumulation subperiods by solving the equation (4) with respect to $i^{(m)}$,

$$i^{(m)} = m \left((1 + i)^{1/m} - 1 \right). \quad (5)$$

Note that only if $m = 1$, the nominal interest and the effective interest coincide.

We can also allow for payments at the beginning of every subperiod. Let e.g. the payment stream be given by the vector of benefits

$$b = \left(b(0), b\left(\frac{1}{m}\right), b\left(\frac{2}{m}\right), \dots, b\left(\frac{mn}{m}\right) \right). \quad (6)$$

We can now define the present value of future payments by summing over payments in future subperiods,

$$V(t) = \sum_{s=mt+1}^{mn} b(s) v^{s/m-t}$$

and, in particular,

$$V(0) = \sum_{t=1}^{mn} b(t) v^{t/m}.$$

We can, in particular, introduce fractionated annuities corresponding to the annuities in the previous section. Then we replace the payment of 1 by the payment of $1/m$. E.g. the fractionated annuity in arrear corresponds to the benefit stream $b = (0, \frac{1}{m}, \dots, \frac{1}{m}) = \left(0, \left(\frac{1}{m}\right)^{1 \times mn}\right)$. In this way it is still so that during one period consisting of m subperiods, the total amount paid is m times $1/m$, i.e. 1. We then get

$$\begin{aligned} a_{\frac{n}{m}}^{(m)} &= V(0) \\ &= \frac{1}{m} \sum_{t=1}^{mn} v^{t/m} \\ &= \frac{1}{m} \sum_{t=1}^{mn} \left(v^{\frac{1}{m}}\right)^t \\ &= \frac{1}{m} \tilde{a}_{\frac{mn}{m}}, \end{aligned}$$

where $\widetilde{a}_{\overline{m}|}$ is an annuity calculated with an interest rate \widetilde{i} solving

$$\left(\frac{1}{1+i}\right)^{\frac{1}{m}} = \frac{1}{1+\widetilde{i}} \Leftrightarrow \widetilde{i} = (1+i)^{\frac{1}{m}} - 1.$$

1.6 Continuous Interest and Payments

In this section we see what happens if we let the frequency of interest rate accumulation and payment of payments go to infinity corresponding to continuous accumulation and payments.

In (5) we calculated the nominal interest rate paid in every subperiod corresponding to the effective interest rate i . A special situation arises if we let m tend to infinity. In the limit we say that interest is accumulated continuously and we can calculate the continuous interest rate corresponding to a yearly interest rate of i ,

$$\begin{aligned} \lim_{m \rightarrow \infty} i^{(m)} &= \lim_{m \rightarrow \infty} m \left((1+i)^{1/m} - 1 \right) \\ &= \lim_{m \rightarrow \infty} \frac{(1+i)^{1/m} - (1+i)^0}{1/m} \\ &= \lim_{t \rightarrow 0} \frac{(1+i)^{0+t} - (1+i)^0}{t} \\ &= \left. \frac{d}{dt} (1+i)^t \right|_{t=0} \\ &= \left. \frac{d}{dt} e^{t \log(1+i)} \right|_{t=0} \\ &= \left. \log(1+i) e^{t \log(1+i)} \right|_{t=0} \\ &= \log(1+i). \end{aligned}$$

We denote this continuous interest rate by $\delta = \log(1+i)$. We can then write the accumulation and discount factors in terms of the continuous interest rate and the exponential function,

$$\begin{aligned} (1+i)^t &= e^{\delta t}, \\ v^{n-t} &= e^{-\delta(n-t)}. \end{aligned}$$

It is also possible to work with continuous payments. If we let m go to infinity in the payment stream (6) what happens in the limit is that the total payment over small intervals is proportional to the interval length and the proportionality factor is called the payment rate. We get the following definition of U and V for a stream of payments $b - c$,

$$\begin{aligned} U(t) &= \int_0^t e^{\delta(t-s)} (c(s) - b(s)) ds, \\ V(t) &= \int_t^n e^{-\delta(s-t)} (b(s) - c(s)) ds. \end{aligned}$$

There exist also continuous versions of the elementary annuities. A continuous annuity corresponds to the benefit rate $b(t) = 1$, taking $c(t) = 0$, and its present value is given by

$$\begin{aligned} \bar{a}_{\overline{n}|} &= V(0) \\ &= \int_0^n e^{-\delta t} dt = \left[-\frac{1}{\delta} e^{-\delta t} \right]_0^n = \frac{1 - e^{-\delta n}}{\delta}. \end{aligned}$$

A deferred continuous annuity corresponds to $b(t) = 1_{(k < t < k+n)}$, taking $c(t) = 0$, and its present value is given by

$$\begin{aligned} {}_k\bar{a}_{\bar{n}|} &= V(0) \\ (a) &= \int_k^{k+n} e^{-\delta t} dt = \int_0^n e^{-\delta(k+t)} dt = e^{-\delta k} \int_0^n e^{-\delta t} dt = e^{-\delta k} \bar{a}_{\bar{n}|} \\ (b) &= \int_0^{k+n} e^{-\delta t} dt - \int_0^k e^{-\delta t} dt = \bar{a}_{\overline{k+n}|} - \bar{a}_{\bar{k}|}. \end{aligned}$$

1.7 Exercises

Exercise 1 *There also exist special notation for terminal present values. Let $c = (0, 1, \dots, 1) = (0, 1^{1 \times n})$ (and $b = 0$). Show that*

$$s_{\bar{n}|} \equiv U(n) = \frac{(1+i)^n - 1}{i} = (1+i)^n a_{\bar{n}|}.$$

Now let $c = (1, \dots, 1, 0) = (1^{1 \times n}, 0)$ (and $b = 0$). Show that

$$\ddot{s}_{\bar{n}|} \equiv U(n) = \frac{(1+i)^n - 1}{d} = (1+i) s_{\bar{n}|} = (1+i)^n \ddot{a}_{\bar{n}|}$$

Exercise 2 *We can get continuous time versions of the difference equations characterizing the U and V in the discrete time case. Show that U and V are characterized by the ordinary differential equations*

$$\begin{aligned} \frac{d}{dt} U(t) &= \delta U(t) + c(t) - b(t), U(0) = 0, \\ \frac{d}{dt} V(t) &= \delta V(t) - (b(t) - c(t)), V(n) = 0. \end{aligned}$$

Show that $U(n) = 0$ and $V(0) = 0$ are equivalent fairness criterions and show that they imply that $U(t) = V(t)$ for all t .

Exercise 3 *We now study the contribution payment stream $c = (0, \pi 1^{1 \times k}, 0^{1 \times n})$ for a constant π in combination with the benefit payment stream $b = (0, 0^{1 \times k}, 1^{1 \times n})$. Determine $V(t)$ for $t \geq k$ and for $t < k$, respectively. In particular, show that*

$$V(0) = {}_k a_{\bar{n}|} - \pi \left(\bar{a}_{\bar{k}|} \right).$$

Calculate π such that the fairness criterion is fulfilled. Calculate π for $k = 30, n = 10$ and $i = 5\%$.

The continuous time version reads $b(t) - c(t) = 1_{(k < t < k+n)} - 1_{(0 < t < k)}$. We have that

$$V(t) = \begin{cases} \bar{a}_{\overline{k+n-t}|}, & t > k, \\ {}_{k-t}\bar{a}_{\bar{n}|} - \pi \bar{a}_{\overline{k-t}|}, & t < k. \end{cases}$$

Calculate the rate π in accordance with the fairness criterion. Calculate π for $k = 10, n = 10$ and $\delta = \log(1 + 5\%)$.

2 Life Insurance Risk: Mortality

Life insurance risk is all kinds of uncertainties that affect the payment stream of a life insurance contract. In its simplest form this reduces to the uncertain lifetime such that the risk can be described by one stochastic variable containing this lifetime. In more general forms, also other events affecting significantly the lifetime income of the individual are considered as life insurance risk. The most important example is disability risk. Then more stochastic variables are in play, and in general it becomes more natural, and even necessary, to work with stochastic processes that are, essentially, time indexed stochastic variables.

At the end of the day life insurance is a betting game on life history events like death, survival, and disability. Betting in this context is very different from betting on horses or football since the benefit is paid conditional on a life history event that has already caused financial distress of the better. In Chapter 6 we look closely at the construction of these products and see how they serve to cover the financial loss that appears as a result of some financially adverse events. The nature of the betting game explains why the society encourages its members to play this betting game (partly in contrast to other betting games) under somewhat regulated conditions, though: It has a stabilizing effect on the financial position of individuals and families and thereby on the society as a whole. Therefore, also, there is a great deal of political discussion concerning individual and collective bearing of burdens in the discipline of insurance, although often these discussions are well disguised in stochastic variables and formulas.

In Chapter 6 we are going to present some standard relevant, fair, and reliable bets. But in order to discuss at all, what is a fair bet on the lifetime, we need to determine the odds, i.e. the probability of survival or death within a certain life span. These probabilities are the topic of this Chapter.

2.1 The newborn's lifetime

In this section we characterize the stochastic variable describing the lifetime of a newborn individual.

Consider a newborn individual. We introduce the stochastic variable T which denotes the uncertain lifetime of this individual. We can now characterize the distribution of the stochastic variable in terms of the distribution function F ,

$$F(t) = P(T \leq t),$$

specifying the probability that the individual dies before age t . Equivalently, we can introduce the so-called survival function \bar{F} by

$$\bar{F}(t) = 1 - F(t) = P(T > t),$$

specifying the probability that the individual survives age t . If $0 < T \leq \omega$ (possibly $\omega = \infty$) we assume that

$$\begin{aligned}\bar{F}(0) &= 1 - F(0) = 1, \\ \bar{F}(\omega) &= 1 - F(\omega) = 0,\end{aligned}$$

in words being that the individual survives age zero with probability 1 and survives age ω with probability 0. The latter means that the individual 'cannot live forever'. Alternatively one could have worked with a distribution with a positive probability mass at ω meaning that there is a positive probability that the individual never dies. We shall here only consider distribution functions which are differentiable, and we say then that the distribution is continuous and introduce the density (function) by

$$f(t) = \frac{d}{dt} F(t).$$

We can then also express the distribution function in terms of the density by the integral

$$F(t) = \int_0^t f(s) ds.$$

Probabilities are now obtained by integrating the density. E.g., the probability that the individual dies in his 60'es equals

$$P(60 < T \leq 70) = F(70) - F(60) = \int_{60}^{70} f(t) dt.$$

2.2 The x -year old's residual lifetime

In this section we characterize the stochastic variable describing the residual lifetime of an x -year old individual.

Often one knows more about the individual than just that he was alive at age zero. If we know that the individual survived age x this of course influences the probability that he will survive age $x + t$. Think of $x = 60$ and $t = 10$. Then the probability that a newborn survives age $x + t$ equals $\bar{F}(70)$. But if we know that he survived age 60 we should use this knowledge to come up with a better (higher) estimate. We introduce the so-called conditional distribution function

$$\begin{aligned} P(T \leq x + t | T > x) &= \frac{P(\{T \leq x + t\} \cap \{T > x\})}{P(T > x)} \\ &= \frac{P(x < T \leq x + t)}{P(T > x)} \\ &= \frac{F(x + t) - F(x)}{1 - F(x)} \\ &= \frac{\bar{F}(x) - \bar{F}(x + t)}{\bar{F}(x)} \\ &\equiv {}_t q_x, \end{aligned}$$

the conditional survival function

$$\begin{aligned} P(T > x + t | T > x) &= \frac{P(\{T > x + t\} \cap \{T > x\})}{P(T > x)} \\ &= \frac{P(T > x + t)}{P(T > x)} \\ &= \frac{1 - F(x + t)}{1 - F(x)} \\ &= \frac{\bar{F}(x + t)}{\bar{F}(x)} \\ &\equiv {}_t p_x, \end{aligned}$$

and, in case of a continuous distribution, the conditional density

$$f_x(t) \equiv \frac{d}{dt} P(T \leq x+t | T > x) = -\frac{1}{\bar{F}(x)} \frac{d}{dt} \bar{F}(x+t)$$

Here we have discovered also some so-called actuarial notation by introducing the abbreviations ${}_tq_x$ and ${}_tp_x$. Particularly, we use q_x and p_x for the one-year survival and death probabilities, i.e. for the special case where $t = 1$. Note that

$$\begin{aligned} {}_tq_0 &= F(t) \Rightarrow {}_tq_x = \frac{F(x+t) - F(x)}{1 - F(x)} = \frac{{}_{x+t}q_0 - {}_xq_0}{1 - {}_xq_0}, \\ {}_tp_0 &= \bar{F}(t) \Rightarrow {}_tp_x = \frac{\bar{F}(x+t)}{\bar{F}(x)} = \frac{{}_{x+t}p_0}{{}_xp_0}. \end{aligned}$$

A much more important concept than the conditional density is the so-called mortality intensity which characterizes the distribution of T in terms of the conditional probability that the individual dies within a 'small' time interval after time t having survived until then, i.e. $P(T < t + \Delta t | T > t)$. If the distribution function is continuous this probability is proportional to Δt for infinitesimally small Δt and this proportionality factor is called the mortality intensity and denoted by μ (greek m for mortality). It turns out that there is a very comfortable mathematical connection between this concept and the survival function. The mathematics are

$$\begin{aligned} \mu(t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P(T < t + \Delta t | T > t) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{F(t + \Delta t) - F(t)}{\bar{F}(t)} \\ &= \frac{1}{\bar{F}(t)} \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \\ &= \frac{1}{\bar{F}(t)} \frac{d}{dt} F(t) \\ &= -\frac{1}{\bar{F}(t)} \frac{d}{dt} \bar{F}(t) \\ &= -\frac{d}{dt} \log \bar{F}(t). \end{aligned}$$

This is actually an ordinary differential equation for $\log \bar{F}(t)$ with initial condition,

$$\begin{aligned} \frac{d}{dt} \log \bar{F}(t) &= -\mu(t) \\ \log \bar{F}(0) &= 0. \end{aligned}$$

The solution to this ordinary differential equation is $\log \bar{F}(t) = -\int_0^t \mu(s) ds$ such that

$$\bar{F}(t) = e^{-\int_0^t \mu(s) ds},$$

expressing the survival function in terms of the mortality intensity.

One thing that makes the intensity particularly convenient to work with is the fact that condi-

tional probability are expressed in a similar way. E.g.

$$\begin{aligned}
{}_t p_x &= P(T > x + t | T > x) \\
&= \frac{\bar{F}(x + t)}{\bar{F}(x)} \\
&= \frac{e^{-\int_0^{x+t} \mu(s) ds}}{e^{-\int_0^x \mu(s) ds}} \\
&= e^{-\int_x^{x+t} \mu(s) ds} \\
&= e^{-\int_0^t \mu(x+s) ds}.
\end{aligned}$$

The last equation is obtained by a simple substitution $\tau = s - x$ and renaming τ by s .

In general one considers the mortality intensity as the fundamental function characterizing the distribution of T . There are various models of intensities which may be relevant for different types of individuals. If the 'individual' is a light bulb, an electronic system, a computer network or something similar, one typically works with a constant mortality (or failure) intensity. This means that the probability of failure within a small time interval given that the system works, is constant over time. This is an awkward assumption for biological individuals like human beings. There one would expect that the probability of dying in the next small time interval is larger for a 70 year old individual who is alive than for a 30 year old individual who is alive. Therefore one typically uses increasing mortality functions.

As always there is a trade-off between simple models which have a small number of parameters and are mathematically tractable and involved models which have a high number of parameters and which describe reality better. The most prominent compromise for modelling lifetimes of human beings is the celebrated Gompertz-Makeham intensity which is a three-dimensional (α, β, γ) model,

$$\mu(t) = \alpha + \beta e^{\gamma t}.$$

In the Danish mortality table for males used by Danish insurance companies from 1982 (G82M) the parameters are given by

$$\begin{aligned}
\alpha &= 5 \cdot 10^{-4}, \\
\beta &= 7.5858 \cdot 10^{-5}, \\
\gamma &= \log(1.09144),
\end{aligned}$$

such that

$$\mu(t) = 5 \cdot 10^{-4} + 7.5858 \cdot 10^{-5} \cdot 1.09144^t.$$

In mortality tables, one typically specifies the probabilities in terms of so-called decrement series. One considers a population of l_0 new-born individuals, with l_0 typically set to 100.000. Then l_x denotes the number of expected survivors at age x and equals $l_0 \bar{F}(x)$. Furthermore, d_x denotes the number of expected deaths in age x and equals $l_x - l_{x+1}$. Survival and death probabilities can now be expressed in terms of the decrement series by

$$\begin{aligned}
{}_t p_x &= \frac{\bar{F}(x + t)}{\bar{F}(x)} = \frac{l_{x+t}}{l_x} \\
{}_t q_x &= 1 - {}_t p_x = \frac{\bar{F}(x) - \bar{F}(x + t)}{\bar{F}(x)} = \frac{l_x - l_{x+t}}{l_x}
\end{aligned}$$

The 'beginning' and the 'end' of the mortality table G82M is illustrated below,

x	$\mu(x)$	$f(x)$	l_x	d_x	q_x
0	0.00057586	0.00057586	100000	58	0.00057911
1	0.00058279	0.00058246	99942	59	0.00058635
2	0.00059036	0.00058968	99883	59	0.00059426
3	0.00059863	0.00059758	99824	60	0.00060289
4	0.00060765	0.00060621	99764	61	0.00061231
5	0.00061749	0.00061565	99703	62	0.00062259
6	0.00062823	0.00062598	99641	63	0.00063381
7	0.00063996	0.00063726	99578	65	0.00064606
8	0.00065276	0.00064958	99513	65	0.00065942
9	0.00066672	0.00066304	99448	67	0.00067401
10	0.00068197	0.00067775	99381	69	0.00068993
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
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	0.66709109
110	1.14865351	0.00000218	0	0	0.69892078

2.3 Expected values

In this section we derive a formula for calculation of the expected value of a differentiable function of the lifetime T .

Let T be a lifetime stochastic variable with a continuous distribution such that F is differentiable. Then for a differentiable function G ,

$$\begin{aligned}
E[G(T)] &= \int_0^\omega G(t) f(t) dt \\
&= \int_0^\omega G(t) \left(\frac{d}{dt} F(t) \right) dt \\
&= - \int_0^\omega G(t) \left(\frac{d}{dt} \bar{F}(t) \right) dt \\
&= -G(\omega) \bar{F}(\omega) + G(0) \bar{F}(0) + \int_0^\omega \bar{F}(t) \left(\frac{d}{dt} G(t) \right) dt \\
&= G(0) + \int_0^\omega \bar{F}(t) \left(\frac{d}{dt} G(t) \right) dt
\end{aligned}$$

In the second last equality we use the rule of partial integration,

$$\begin{aligned} G(\omega) F(\omega) - G(0) F(0) &= \int_0^\omega \left(\frac{d}{dt} (G(t) F(t)) \right) dt \\ &= \int_0^\omega G(t) \left(\frac{d}{dt} F(t) \right) dt + \int_0^\omega F(t) \left(\frac{d}{dt} G(t) \right) dt. \end{aligned}$$

In the last equality we use that $\bar{F}(\omega) = 0$ and $\bar{F}(0) = 1$.

Consider e.g. calculation of the q th moment of T . This is obtained by setting $G(t) = t^k$, and we get

$$E[T^k] = 0^k + \int_0^\omega \left(\frac{d}{dt} (t^k) \right) \bar{F}(t) dt = k \int_0^\omega t^{k-1} \bar{F}(t) dt.$$

In particular, the expected lifetime equals

$$E[T] = \int_0^\omega \bar{F}(t) dt = \int_0^\omega e^{-\int_0^t \mu(s) ds} dt \equiv e_0.$$

The expected remaining life time of an x years old individual equals

$$\int_0^{\omega-x} {}_t p_x dt = \int_0^{\omega-x} e^{-\int_x^{x+t} \mu(s) ds} dt = \int_0^{\omega-x} e^{-\int_0^t \mu(x+s) ds} dt \equiv e_x.$$

2.4 Comments

- Demography

is the study of populations. This is simply a study of a group of individual and the decrement series can be seen as such. This can be called a closed population since the only change in the population comes from death. However, there may be features that complicates things when looking at non-closed populations. Firstly, the population reproduces itself so fertility intensities play an important role. Secondly, the population may be influenced by migration (movements in and out of the population which are not driven by birth or death). Disregarding migration, a stationary population is a population which is supported by the same number of births every year, e.g. l_0 .

- Longevity

In the presentation above we have assumed that the mortality intensity is only dependent on age. In practice, this is not the full truth. A 70-year old individual had another mortality intensity in the calendar year 2000 than in the calendar year 1900. We say that the mortality intensity is also dependent on calendar time. This can be modelled more or less sophisticated and we shall not go further into this now. It is one of the main tasks of life insurance actuaries to model and control this risk connected to future lifetimes.

- Above we presented the Danish mortality table for males. As you may have guessed there exists a similar table G82F for females. However, working with different mortality tables for males and females has been found to be sexual discriminating in some situations. Therefore the Danish insurance companies in these situations work with so-called unisex mortality tables that cover a population consisting of appropriate proportions of males and females. This shows that although there is biological evidence for differentiation between certain groups of individuals there may be political reasons to disregard this in certain situations.

- Along with the prediction of mortality comes a discussion about sociology, medicine, biology, genetics etc. Mortality effects are to be explained by genetics (some people are born with better health than others) and environmental/behavioral (some people live in areas with more pollution than others and some people take more drugs than others) causes. But every observed effect seems to be a complex of causes and it is often difficult to clarify what is the cause and what is the effect. Such studies often end up in partly political issues that are not free of cultural values. Try to figure out why mortality decreases with socioeconomic status and education.

2.5 Exercises

Exercise 4 Consider the survival function $\bar{F}(t) = \left(1 - \frac{t}{110}\right)^{\frac{3}{4}}$, $0 \leq t \leq 110$. Calculate

- the probability that a newborn will survive age 10
- the probability that a newborn will die between ages 60 and 70,
- the probability that a 20 years old will survive age 40,
- the mortality intensity at age 65.

Exercise 5 Consider the survival function $\bar{F}(t) = e^{-\beta t^\alpha}$. Calculate α and β such that $\mu(40) = 4\mu(10)$ and $\mu(25) = 0.001$.

Exercise 6 Consider the constant mortality rate 0.004. Calculate

- ${}_{10}p_{30}$,
- ${}_5q_{15}$
- the age x such that a newborn has probability 90% of surviving age x .

Exercise 7 Complete and interpret the life table below which is an extract of G82M,

x	l_x	d_x	q_x
50	91119	617	
51		665	
52			0.00796941

Exercise 8 The exponential mortality law is characterized by a constant force of mortality, $\mu(t) = \lambda$. Show that

$$\begin{aligned}\bar{F}(t) &= e^{-\lambda t}, \\ f(t) &= \frac{d}{dt}F(t) = \lambda e^{-\lambda t}.\end{aligned}$$

Introduce the conditional survival function

$${}_tp_x = \frac{\bar{F}(x+t)}{\bar{F}(x)}.$$

Show that

$${}_tp_x = {}_tp_0,$$

i.e., in particular, this is independent of x . Conclude that

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

and interpret this result.

Exercise 9 The Weibull mortality law is characterized by the intensity

$$\mu(t) = \beta \alpha^{-\beta} t^{\beta-1}, \alpha, \beta > 0.$$

Show that

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

Show that μ is increasing for $\beta > 1$, decreasing for $\beta < 1$, and that the Weibull law equals the exponential law for $\beta = 1$. Calculate ${}_t p_x$.

Exercise 10 The Gompertz-Makeham mortality law is characterized by the mortality intensity

$$\mu = \alpha + \beta e^{\gamma t}, \alpha, \beta > 0.$$

Show that

$$\overline{F}(t) = e^{(-\alpha t - \beta(e^{\gamma t} - 1)/\gamma)}.$$

Show that if $\gamma > 0$, then μ is an increasing function of t . Calculate ${}_t p_x$.

3 Non-life Insurance Risk: Frequency and Severity

3.1 Compound risk models

In this chapter, we study the accumulated claim of a non-life insurance portfolio. *Accumulated* means that all claims from different policies and possibly different branches are looked at in aggregation over some time period, e.g. a year. *Claim* means the amount that the insurance company has to pay out to its policy holders. *Non-life insurance* means obviously types of insurance that cannot be characterized as life insurance. These include most types of insurances on 'things'. *Portfolio* means that we consider an aggregate or a collective of insurance contracts. The ideas and models can also be interpreted on an individual insurance level, but for this purpose there are many alternatives which we shall not consider here. The model we consider here is known as the *collective risk model* or simply *risk model*.

The accumulated claim on the insurance company contains basically two sources of uncertainty: What is the number of claims in the portfolio? This uncertainty is modelled by probability assumptions on the *claim number distribution*. For a given claim, what is the size of this claims? This uncertainty is modelled by probability assumptions on the *claim size distribution*. It is typically assumed that each claim has the same claim size distribution. If the claim number and the claim sizes are not uncertain, it is easy to calculate the accumulated claim: If there are three claims with claim sizes 10, 20, and 25 units, then the accumulated claim is 55. But what is this sum if both the number of claims (elements in the sum) and each and every claim (element in the sum) is not known?

The claim number distribution is a discrete distribution on non-negative integers. We will not allow for -5 claims or 6.32 claims. Let N be the number of claims (N for number) occurring in a particular time interval, say, a year. Then the distribution of N is described by the probabilities

$$p(n) = P(N = n), n = 0, 1, \dots$$

Let Y_i denote the amount of the i th claim. We assume that Y_1, Y_2, \dots are positive, mutually independent, identically distributed with common distribution function F , and independent of the claim number N . All this means that we will not allow for negative claims (the policy holders do not give an amount to the company if their house burns), we do not learn anything about the size of claim Y_i just from knowing anything about the size of claim Y_j , $j \neq i$, (the severity of one fire and its following claims has nothing to do with the severity of another fire and its following claims), any to claims have the same distribution F (prior to the two fires we make the same assumption about what claims they will lead to), and we do not learn anything about the severity of claims just from knowing something about the number of claims (Just because there are many fires one year we do not expect that there fires are particularly big or small).

Let X be the accumulated claim amount, that is

$$X = \sum_{i=1}^N Y_i = Y_1 + Y_2 + \dots + Y_N.$$

There are many reasons for studying the total claim. This shall teach us something about the business such that we can take appropriate actions with respect to premiums, reserves, reinsurance, etc. Premiums are calculated before we underwrite (in connection with insurance this is really the correct translation of 'skrive under' or 'tegne') this portfolio of insurances; reserves are calculated

for the sake of accounting and includes a prediction of the outstanding liabilities; reinsurance is a way in which the insurance company can get rid of parts of its business that for some reason is undesirable.

In general, one is interested in calculating characteristics in the form $E[f(X)]$ for different choices of f . Different moments of X are obtained for specific choices of f and the identity function just gives the expectation of X . The distribution function of X is given by taking f to be the indicator function $1[x \leq u]$. We shall show a relation in the discrete claim size distribution case that is just a small example in a very wide, important and useful theory of conditional expectations and conditional distributions,

$$\begin{aligned}
E[f(X)] &= \sum_{x=0}^{\infty} f(x) P(f(X) = f(x)) \\
&= \sum_{x=0}^{\infty} f(x) \sum_{n=0}^{\infty} P(f(X) = f(x) \cap N = n) \\
&= \sum_{x=0}^{\infty} f(x) \sum_{n=0}^{\infty} P(f(X) = f(x) | N = n) P(N = n) \\
&= \sum_{n=0}^{\infty} \sum_{x=0}^{\infty} f(x) P(f(X) = f(x) | N = n) P(N = n) \\
&= \sum_{n=0}^{\infty} E[f(X) | N = n] P(N = n),
\end{aligned}$$

where $E[f(X) | N = n]$ is the expected value in the conditional distribution with probabilities $P(f(X) = f(x) | N = n)$. The result

$$E[f(X)] = \sum_{n=0}^{\infty} E[f(X) | N = n] p(n),$$

also holds if the claim size distribution is continuous. In order to go on from here, we need to say more about p , F , and f , i.e. the claim number distribution, the claim size distribution, and the characteristic f . We shall here proceed a step further for two specific choices of f and get as far as we can without specifying p and F .

Firstly, we calculate the expected value of X by letting $f(x) = x$. This gives

$$\begin{aligned}
E[X] &= \sum_{n=0}^{\infty} E[X | N = n] p(n) \\
&= \sum_{n=0}^{\infty} E[Y_1 + Y_2 + \dots + Y_n] p(n) \\
&= \sum_{n=0}^{\infty} \sum_{i=1}^n E[Y_i] p(n) \\
&= \sum_{n=0}^{\infty} n E[Y_1] p(n) \\
&= E[Y_1] \sum_{n=0}^{\infty} n p(n) \\
&= E[Y_1] E[N].
\end{aligned}$$

Recognize that in the second equality we use that the claims and the claim number are independent. Otherwise this expectation should somehow depend on n $E_n[Y_1 + Y_2 + \dots + Y_n]$ to indicate that the a particular distribution of the claim sizes should be used given that $N = n$. Recognize also that in the last equality we use that the Y_1, Y_2, \dots are identically distributed or, rather, that they have the same expectation. We have *not* used here that they are independent. The interpretation of the result is very intuitive. The expected accumulated claim equals the expected number of claims multiplied by the expected claim size.

Secondly, we calculate the second moment of X by letting $f(x) = x^2$. With the notation

$$\begin{aligned}\sum_i &= \sum_{i=1}^n, \\ \sum_{i,j} &= \sum_i \sum_j, \\ \sum_{i \neq j} &= \sum_i \sum_{j:j \neq i},\end{aligned}$$

we have that

$$\begin{aligned}E[X^2] &= \sum_{n=0}^{\infty} E[X^2 | N = n] p(n) \\ &= \sum_{n=0}^{\infty} E[(Y_1 + Y_2 + \dots + Y_n)^2] p(n) \\ &= \sum_{n=0}^{\infty} E\left[\sum_{i,j} Y_i Y_j\right] p(n) \\ &= \sum_{n=0}^{\infty} \sum_{i,j} E[Y_i Y_j] p(n) \\ &= \sum_{n=0}^{\infty} \sum_{i \neq j} E[Y_i Y_j] p(n) + \sum_{n=0}^{\infty} \sum_i \sum_{j:j=i} E[Y_i Y_j] p(n) \\ &= \sum_{n=0}^{\infty} \sum_{i \neq j} E[Y_i] E[Y_j] p(n) + \sum_{n=0}^{\infty} \sum_i E[Y_i^2] p(n) \\ &= \sum_{n=0}^{\infty} n(n-1) E^2[Y_1] p(n) + \sum_{n=0}^{\infty} n E[Y_1^2] p(n) \\ &= E^2[Y_1] \sum_{n=0}^{\infty} (n^2 - n) p(n) + E[Y_1^2] \sum_{n=0}^{\infty} n p(n) \\ &= E^2[Y_1] (E[N^2] - E[N]) + E[Y_1^2] E[N] \\ &= E^2[Y_1] E[N^2] + E[N] \text{Var}[Y_1].\end{aligned}$$

Here we have also used that the claims are identically distributed and independent of the claim number (where?). Furthermore, in the sixth equality we have used that the claims are even mutually independent, since only then we know that $E[Y_i Y_j] = E[Y_i] E[Y_j]$ for $i \neq j$.

The first and second moments makes it possible to calculate the variance of the accumulated claim. This obviously tells us something about the riskiness of the business. This riskiness may play a role in many actions on premium, reserve and reinsurance calculations. We have that

$$\begin{aligned}\text{Var}[X] &= E[X^2] - E^2[X] \\ &= E^2[Y_1] E[N^2] + E[N] \text{Var}[Y_1] - E^2[Y_1] E^2[N] \\ &= E[N] \text{Var}[Y_1] + E^2[Y_1] \text{Var}[N].\end{aligned}$$

The expected value and the variance are probably the two most important characteristic of a stochastic variable. We can now calculate these for the accumulated claim amount provided that we know the expected value and the variance of the claim number and the claim sizes, respectively. In the next subsections, we study particular claim number and claim size distribution which for various reasons play special roles.

3.2 Frequency risk

In this section, we present three important claim number distributions. All three distribution belongs to a certain so-called (a, b) -class of distributions for which

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1).$$

- The *Poisson distribution* is characterized for $\lambda > 0$ by

$$p(n) = e^{-\lambda} \frac{\lambda^n}{n!}, n = 0, 1, \dots$$

The Poisson distribution is seen to belong to the (a, b) -class by

$$p(n) = e^{-\lambda} \frac{\lambda^n}{n!} = \frac{\lambda}{n} e^{-\lambda} \frac{\lambda^{n-1}}{(n-1)!} = \frac{\lambda}{n} p(n-1),$$

such that

$$\begin{aligned} a &= 0, \\ b &= \lambda, \lambda > 0. \end{aligned}$$

- The *binomial distribution* is characterized for $0 < p < 1$ and $m = 1, 2, \dots$ by

$$p(n) = \binom{m}{n} p^n (1-p)^{m-n}, n = 0, 1, \dots, m.$$

That the binomial distribution belongs to the (a, b) -class is seen by

$$\begin{aligned} p(n) &= \binom{m}{n} p^n (1-p)^{m-n} \\ &= \frac{p(m-n+1)}{n(1-p)} \frac{m!}{(n-1)!(m-n+1)!} p^{n-1} (1-p)^{m-(n-1)} \\ &= \left(-\frac{p}{1-p} + \frac{(m+1)p}{n(1-p)} \right) p(n-1), \end{aligned}$$

such that

$$\begin{aligned} a &= -\frac{p}{1-p}, 0 < p < 1 \\ b &= (m+1) \frac{p}{1-p} = -(m+1)a. \end{aligned}$$

The fact that $a < 0$ has the interpretation that, compared with the Poisson distribution, the 'probability of more claims decreases' with the number of claims. In fact the effect is that there is only a finite number of claims. If there are as many as m claims 'probability of more claims' becomes zero. This is spoken of as negative contagion. One interpretation of this claim number process is the case where there is a finite known number m of policies and each policy gives rise to 0 or 1 claim and not more. The probability of 1 claim is p .

- The *negative binomial distribution* is characterized for $0 < q < 1$ and $\alpha > 0$ by

$$p(n) = \binom{\alpha+n-1}{n} q^n (1-q)^\alpha, n = 0, 1, \dots$$

That the negative binomial distribution belongs to the (a, b) -class is seen by

$$\begin{aligned}
p(n) &= \binom{\alpha + n - 1}{n} q^n (1 - q)^\alpha \\
&= \frac{\Gamma(\alpha + n)}{n! \Gamma(\alpha)} q^n (1 - q)^\alpha \\
&= \frac{q}{n} \frac{(\alpha + n - 1) \Gamma(\alpha + n - 1)}{(n - 1)! \Gamma(\alpha)} q^{n-1} (1 - q)^\alpha \\
&= \left(q + \frac{q(\alpha - 1)}{n} \right) \frac{\Gamma(\alpha + n - 1)}{(n - 1)! \Gamma(\alpha)} q^{n-1} (1 - q)^\alpha,
\end{aligned}$$

such that

$$\begin{aligned}
a &= q, 0 < q < 1, \\
b &= q(\alpha - 1) > -q = -a.
\end{aligned}$$

The fact that $a > 0$ has the interpretation that, compared with the Poisson distribution, the 'probability of more claims increases' with the number of claims. This is spoken of as positive contagion. One interpretation of positive contagion is that one car participating in a car accident 'increases the probability of more claims' since more cars may be involved in the accident. In this sense cars 'infect each other with claims'. If we instead consider the accident as the claim (and not the individual damaged car) we only count one claim also if there are more cars involved. Then the claim number may be Poisson distributed but with a claim size which takes into account the risk of having more cars damaged in one claim. This opposite contagion effects in the negative binomial distribution compared to the binomial distribution explains actually the term *negative binomial*.

3.3 Severity risk

In this section, we present three claim size distributions. There are many different ways of characterizing and categorizing claim size distributions depending on the purpose. In particular, one is often interested in characteristics of the riskiness. We know that the variance is one measure of risk. However, this is for some purposes and for some distributions a completely insufficient piece of information. As we shall see, there exists highly relevant claim size distributions that behave so wildly that the variance not even exists, i.e. is infinite. But still information about the mean and the variance are important pieces of information.

Another important characteristic is the so-called *mean excess loss*

$$e(y) = E[Y - y | Y > y].$$

This measures the expected claim size beyond y given that the claims size is above y . This characteristic plays different roles and has different names in different contexts. In the life context it is called the mean residual lifetime, and you can compare with the similar function introduced in Chapter 2. For biological life this function is typically assumed to be decreasing. For technological life (the light bulb) this function is typically assumed to be constant. In a financial risk management context the function is called the expected shortfall. In the non-life insurance context it is,

as mentioned, called the mean excess loss and is typically estimated to be increasing. The way it is increasing tells something about the 'danger' of a distribution. If the mean excess loss goes to infinity, the distribution is called heavy-tailed (tunghalet) corresponding to a 'dangerous' distribution. If the mean excess loss function converges to a finite number, the distribution is called light-tailed (lethalet) corresponding to a 'less dangerous' distribution.

Heavy-tailed distributions are e.g. used in motor, fire and storm with different explanations: Motor is hit by large claims due to the liability insurance; fire is hit by large claims due to industrial buildings with valuable production machinery and typically combined with a protection against loss of production in the build-up period; storm, and other catastrophes of nature, is hit by large claims due to the mercilessness of Mother Earth. Light-tailed distributions are e.g. used in insurance against theft.

- The Gamma (γ, δ) distribution is characterized for $\gamma > 0$ and $\delta > 0$ by the density

$$f(y) = \frac{\delta^\gamma}{\Gamma(\gamma)} y^{\gamma-1} e^{-\delta y}, \quad y \geq 0,$$

where actually the Gamma function Γ is determined such that this is indeed a probability distribution,

$$\int_0^\infty \frac{\delta^\gamma}{\Gamma(\gamma)} y^{\gamma-1} e^{-\delta y} dy = 1 \Leftrightarrow \Gamma(\gamma) = \int_0^\infty \delta^\gamma y^{\gamma-1} e^{-\delta y} dy$$

We can then calculate

$$\begin{aligned} E[Y^k e^{-aY}] &= \int_0^\infty y^k e^{-ay} \frac{\delta^\gamma}{\Gamma(\gamma)} y^{\gamma-1} e^{-\delta y} dy \\ &= \frac{\delta^\gamma}{\Gamma(\gamma)} \int_0^\infty y^{k+\gamma-1} e^{-(\delta+a)y} dy \\ &= \frac{\Gamma(\gamma+k)}{\Gamma(\gamma)} \frac{\delta^\gamma}{(\delta+a)^{\gamma+k}}. \end{aligned}$$

The Gamma distribution is a light-tailed distribution.

- The Pareto (λ, α) distribution is characterized for $\alpha > 0$ by the distribution function

$$F(y) = 1 - \left(\frac{\lambda}{y}\right)^\alpha, \quad y > \lambda,$$

which gives the density

$$f(y) = \frac{\alpha}{\lambda} \left(\frac{\lambda}{y}\right)^{\alpha+1}, \quad y > \lambda.$$

We can check that this is really a probability distribution by the calculation,

$$\int_\lambda^\infty \frac{\alpha}{\lambda} \left(\frac{\lambda}{y}\right)^{\alpha+1} dy = \frac{\alpha}{\lambda} \lambda^{\alpha+1} \int_\lambda^\infty y^{-\alpha-1} dy = \frac{\alpha}{\lambda} \lambda^{\alpha+1} \left[-\frac{1}{\alpha} y^{-\alpha} \right]_\lambda^\infty = 1.$$

We can then calculate for $k < \alpha$

$$\begin{aligned}
E[Y^k] &= \int_{\lambda}^{\infty} y^k \frac{\alpha}{\lambda} \left(\frac{\lambda}{y}\right)^{\alpha+1} dy \\
&= \int_{\lambda}^{\infty} \frac{\alpha}{\lambda} \frac{\lambda^{\alpha+1}}{y^{\alpha-k+1}} dy \\
&= \int_{\lambda}^{\infty} \frac{\alpha}{\alpha-k} \lambda^k \frac{\alpha-k}{\lambda} \frac{\lambda^{\alpha-k+1}}{y^{\alpha-k+1}} dy \\
&= \frac{\alpha}{\alpha-k} \lambda^k \int_{\lambda}^{\infty} \frac{\alpha-k}{\lambda} \frac{\lambda^{\alpha-k+1}}{y^{\alpha-k+1}} dy \\
&= \frac{\alpha}{\alpha-k} \lambda^k.
\end{aligned}$$

Note that the k th moments only exist up to α . In dangerous lines of business, one will often find that the estimated value of α is close to 1. Note that α may be smaller than 1 such that not even the expectation exists, i.e. it is infinite.

A very appealing and interesting property of the Pareto distribution is that if Y is Pareto (λ, α) distributed and $\alpha > 1$, then the conditional distribution of Y given $Y > y$ is again Pareto distributed with parameters (y, α) . Thus,

$$e(y) = \frac{\alpha y}{\alpha - 1} - y = \frac{y}{\alpha - 1}, \quad y > \lambda.$$

This shows that e goes to linearly to infinite and the Pareto is heavy-tailed.

- The lognormal (α, σ^2) distribution is characterized by the density

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{y} e^{-\frac{1}{2}\left(\frac{\log y - \alpha}{\sigma}\right)^2}, \quad y > 0.$$

The moments of the lognormal distribution can be calculated by

$$E[Y^k] = e^{\alpha k + \frac{\sigma^2}{2} k^2}.$$

The lognormal distribution is a heavy-tailed distribution but not as heavy-tailed as the Pareto distribution since the mean excess loss function goes to infinity, not linearly, but as $\frac{y}{\log y}$.

3.4 Exercises

Exercise 11 Assume that N is Poisson distributed with parameter λ . Show that

$$\begin{aligned}
E[N] &= \lambda, \\
E[N^2] &= \lambda^2 + \lambda, \\
\text{Var}[N] &= \lambda,
\end{aligned}$$

and use this to show that

$$\begin{aligned}
E[X] &= \lambda E[Y], \\
E[X^2] &= \lambda E[Y^2] + \lambda^2 E^2[Y], \\
\text{Var}[X] &= \lambda E[Y^2].
\end{aligned}$$

Exercise 12 Assume that N is negative binomial distributed with parameters (α, q) . Show that

$$\begin{aligned} E[N] &= q\alpha + qE[N] \\ E[N^2] &= q^2 E[N^2] + q^2 (2\alpha + 1) E[N] + q^2 (\alpha^2 + \alpha) + E[N] \end{aligned}$$

and calculate that

$$\begin{aligned} E[N] &= \frac{q\alpha}{1-q}, \\ E[N^2] &= \frac{1+q^2(2\alpha+1)}{1-q^2} E[N] + \frac{q^2}{1-q^2} (\alpha^2 + \alpha) \\ &= \frac{q\alpha}{1-q^2} \left(\frac{1+q^2(2\alpha+1)}{1-q} + q(\alpha+1) \right) \\ &= \frac{q\alpha}{1-q^2} \left(\frac{1+q^2\alpha+q\alpha+q}{1-q} \right), \\ \text{Var}[N] &= \frac{q\alpha}{1-q^2} \left(\frac{1+q^2(2\alpha+1)}{1-q} + q(\alpha+1) - \frac{(1-q^2)q}{(1-q)^2} \alpha \right) \\ &= \frac{q\alpha}{(1-q)^2}. \end{aligned}$$

Exercise 13 We introduce the mean, variance, coefficient of variation, and skewness,

$$\begin{aligned} \mu &= E[Y], \\ \sigma^2 &= E[(Y - E[Y])^2] \\ &= E[Y^2] - E^2[Y], \\ \kappa &= \frac{\sigma}{\mu}, \\ \nu &= \frac{E[(Y - E[Y])^3]}{\sigma^3} \\ &= \frac{E[Y^3]}{\sigma^3} - \frac{3}{\kappa} - \frac{1}{\kappa^3}. \end{aligned}$$

Argue why μ and σ are not independent of monetary unit, while κ and ν are.

Consider the Pareto claim size distribution. Show that

$$\begin{aligned} \mu &= \frac{\alpha\lambda}{\alpha-1}, \\ \sigma^2 &= \frac{\alpha\lambda^2}{(\alpha-2)(\alpha-1)^2}, \\ \kappa &= \frac{1}{\sqrt{\alpha(\alpha-2)}}. \end{aligned}$$

Assume that N is negative binomial distributed with parameters (q, β) . Calculate $E[X]$ and $\text{Var}[X]$.

4 Valuation principles

In this chapter we discuss some valuation principles and their properties.

In Chapters 1 and 2 we have presented different types of uncertainty, the lifetime of an individual, the number of claims in a portfolio of non-life insurances, and the size of the claims in a portfolio of insurances. The idea of insurance is basically to transfer this uncertainty from the policy holder to the insurance company in the sense that he pays a certain premium and for this premium he receives a benefit linked to the stochastic phenomena. In fire insurance he pays a certain premium which e.g. covers the fire risk during a year. If a fire accidentally burns down the house, he receives a benefit sufficient to rebuild the house such that he is not financially affected by the accident. Obviously, he may still be emotionally affected but the insurance company is not supposed to do anything about that. The *premium* is not the payment of an *award or prize* but the payment of an initiation or admission fee which is paid *in advance or premiere*. We use the letter π (Greek 'p') for the premium.

The question is, for a given risk or stochastic benefit, what should naturally be required from a premium and which general premium formulas fulfill these requirements. Let us start with the easiest premium principle, namely the so-called equivalence principle stating for a stochastic benefit or variable X that

$$\pi(X) = E[X], \quad (7)$$

also called the equivalence premium. This seems very reasonable except for that no risk averse (meaning that one prefers certainty over uncertainty) insurance company would be willing to take over the risk for this premium. With this premium the insurance company takes over the risk without being rewarded for that, so why should they do it? Nevertheless, the equivalence principle plays a very important role in life insurance for reasons that go beyond the scope of this exposition.

A very fundamental principle from finance is the principle of *no-arbitrage* stating that it should not be possible to make certain gains. This principle is typically (in finance) imposed in an abstract market where all participants are able to buy and sell all risks in all fractions. Several of the properties below are related to this no-arbitrage principle.

A discussion about actuarial valuation principles is closely related to a discussion about risk measures that goes on in the academia and the industry these years. We wish to assign one value to a stochastic variable that expresses important characteristics about this variable in a pricing/management sense: This could be for calculation of premium: What does it cost to transfer the risk?; for calculating entries in a balance scheme: When the industrial companies present a snapshot of the corporate to the financial world, which information is then relevant/important/sufficient?; for calculating solvency rules: When the industrial companies defend their 'state of health' to supervisory authorities, which information is then relevant/important/sufficient?

An important risk measure called VaR which abbreviates Value-at-Risk is more or less the same as the quantile principle below. This appears to be oversimple but is a fundamental tool in financial risk management nowadays, where financial risk managers have just become aware that some financial risks are non-normal and therefore the variance does not always tell everything.

4.1 Properties

In this section we go through a series of properties that for various reasons are appropriate to require from a premium function or premium principle.

1. *Non-negative loading*

If the equivalence premium is too small since it does not compensate for the transfer of risk to a risk averse institution, one should require that

$$\pi(X) \geq E[X].$$

A more mathematical argument for this property is the following: Assume that we have n independent and identically distributed risks X_1, \dots, X_n and that $\pi(X) < E[X]$. Then the law of large number teaches us that

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{n} \sum_{i=1}^n X_i > \pi(X)\right) = 1,$$

which means that the total net loss of the insurance company $\sum_{i=1}^n X_i - n\pi(X)$ is positive with probability 1 when the portfolio grows to infinity, i.e. $n \rightarrow \infty$. This is very unsatisfying for the insurance company.

2. *Maximal loss*

If we know that X will not exceed a maximum loss m then the premium cannot be higher than m . In mathematical terms,

$$X \leq m \Rightarrow \pi(X) \leq m.$$

It is very reasonable that the policy holder would never be willing to pay more than the worst case loss. If this is not fulfilled, i.e. if $X \leq m$ **and** $\pi(X) > m$, then the insurance company obtains an arbitrage by buying X at the price $\pi(X)$ with the loss $X - \pi(X) \leq m - \pi(X) < 0$, leading to a certain gain. Although this is in principle the most simple arbitrage strategy (compared to the ones below) this maximal loss property turns out **not** to be fulfilled for very common and simple premium principles.

3. *Monotonicity*

If one coverage is better than another coverage in the sense that one risk Y pays more than another risk X , no matter what happens, then the price of Y should be larger than the price of X . In mathematical terms,

$$X \leq Y \Rightarrow \pi(X) \leq \pi(Y).$$

If this is not fulfilled, i.e. $X \leq Y$ and $\pi(X) > \pi(Y)$ then an arbitrage is obtained by selling X at the price $\pi(X)$ and buying Y at the price $\pi(Y)$. The net loss is then $X - \pi(X) + \pi(Y) - Y = \pi(Y) - \pi(X) + X - Y < 0$, which gives a certain gain.

4. Subadditivity

If one has two risks X and Y one would expect that the premium for the total risk $X + Y$ is not larger than the premium for X plus the premium for Y ,

$$\pi(X + Y) \leq \pi(X) + \pi(Y).$$

If e.g. one has an insurance which covers against fire (X) and storm (Y) then one should get a non-negative rebate for holding one contract with both coverages instead holding one contract for each coverage. Otherwise the policy holders would get a discount for splitting up risk.

One could also require additivity where, obviously, the inequality above is an equality, and superadditivity, where the inequality is turned around. If additivity is not fulfilled it is possible to make an arbitrage: Assume that $\pi(X + Y) < \pi(X) + \pi(Y)$. Then sell the risk X at the price $\pi(X)$ with the loss $X - \pi(X)$, sell the risk Y at the price $\pi(Y)$ with the loss $Y - \pi(Y)$, and buy the risk $X + Y$ at the price $\pi(X + Y)$ with the loss $\pi(X + Y) - (X + Y)$. The total loss is then $X - \pi(X) + Y - \pi(Y) + \pi(X + Y) - (X + Y) = \pi(X + Y) - \pi(X) - \pi(Y) < 0$ and we have a certain gain. This Subadditivity is **not** consistent with no-arbitrage.

4.2 Principles

In this section we go through a series of premium principles and discuss to which extent they possess the properties in the previous subsection.

- *The expected value principle*

1. This principle is based on the idea of loading the expectation of X with a proportion a of the expectation itself, i.e. for $a > 0$,

$$\pi(X) = (1 + a) E[X].$$

We can now check the different properties from the previous subsection.

1. *Non-negative loading* is obviously fulfilled.
2. *Maximal loss* is not fulfilled: Consider a constant risk such that X actually equals its maximal value, i.e. $X = m$. Then

$$\pi(m) = (1 + a) E[m] = (1 + a) m > m,$$

and the *Maximal loss* is violated.

3. *Monotonicity* is fulfilled: $X \leq Y \Leftrightarrow X - Y \leq 0 \Rightarrow E[X - Y] \leq 0 \Leftrightarrow E[X] \leq E[Y] \Leftrightarrow (1 + a) E[X] \leq (1 + a) E[Y]$.

4. *Subadditivity* is fulfilled:

$$\begin{aligned} \pi(X + Y) &= (1 + a) E[X + Y] \\ &= (1 + a) (E[X] + E[Y]) \\ &= (1 + a) E[X] + (1 + a) E[Y] \\ &= \pi(X) + \pi(Y) \end{aligned}$$

Thus, even additivity is fulfilled.

- *Standard deviation principle*

This principle is based on the idea of loading the expectation of X with a proportion of the standard deviation of X , i.e. for $a > 0$,

$$\pi(X) = E[X] + a\sqrt{Var[X]}.$$

We can now check the different properties from the previous subsection.

1. *Non-negative loading* is obviously fulfilled.
2. *Maximal loss* is not fulfilled: Consider a binomial risk such that $P(X = m) = p = 1 - P(X = 0)$. Then

$$\begin{aligned} EX &= mp \\ VarX &= pm^2 - (pm)^2 = m^2p(1-p) \end{aligned}$$

and as a function of p the premium equals

$$g(p) = \pi(X) = mp + a\sqrt{m^2p(1-p)} = m\left(p + a\sqrt{p(1-p)}\right),$$

with

$$\begin{aligned} g(0) &= 0, \\ g(1) &= m. \end{aligned}$$

The derivative of the premium with respect to p equals

$$g'(p) = m\left(1 + \frac{1}{2} \frac{a(1-2p)}{\sqrt{p(1-p)}}\right),$$

with

$$\begin{aligned} \lim_{p \rightarrow 0} g'(p) &= m\left(1 + \frac{1}{2} \frac{a}{0^+}\right) = \infty, \\ \lim_{p \rightarrow 1} g'(p) &= m\left(1 - \frac{1}{2} \frac{a}{0^+}\right) = -\infty. \end{aligned}$$

But then there must exist a $p < 1$ such that

$$g(p) = \pi(X) > m.$$

3. *Monotonicity* is not fulfilled: First note that **if** we have *monotonicity* **and** the premium for a constant risk is the constant risk itself, **then** we have *maximal loss*: If Y equals m , we have by monotonicity that $X \leq m \Rightarrow \pi(X) \leq \pi(m) = m$. But this is exactly the *maximal loss*. Second, the premium for a constant risk is the constant itself for the standard deviation principle since the standard deviation for a constant is zero. Then we know that the *maximal loss* is not fulfilled and one of the conditions is fulfilled. Then the other condition, *monotonicity* cannot be fulfilled.

4. *Subadditivity* is fulfilled: With the definition of the covariance,

$$\begin{aligned} \text{Cov}[X, Y] &= E[(X - EX)(Y - EY)] \\ &= E[XY - YEX - XEY + EXEY] \\ &= E[XY] - EXEY, \end{aligned}$$

we can calculate the variance of $X + Y$,

$$\begin{aligned} \text{Var}[X + Y] &= E(X + Y)^2 - E^2(X + Y) \\ &= EX^2 + EY^2 + 2E[XY] - (EX + EY)^2 \\ &= EX^2 - E^2X + EY^2 - E^2Y + 2E[XY] - 2EXEY \\ &= \text{Var}X + \text{Var}Y + 2\text{Cov}[X, Y]. \end{aligned}$$

Then we exploit the result

$$\begin{aligned} \text{Cov}[X, Y] &= E[(X - EX)(Y - EY)] \\ &\leq \sqrt{E(X - EX)^2} \sqrt{E(Y - EY)^2} \\ &= \sqrt{\text{Var}[X]} \sqrt{\text{Var}[Y]}, \end{aligned}$$

(where the inequality is beyond the scope of this presentation) to calculate that

$$\begin{aligned} \sqrt{\text{Var}[X + Y]} &= \sqrt{\text{Var}X + \text{Var}Y + 2\text{Cov}[X, Y]} \\ &\leq \sqrt{\text{Var}X + \text{Var}Y + 2\sqrt{\text{Var}[X]}\sqrt{\text{Var}[Y]}} \\ &= \sqrt{\left(\sqrt{\text{Var}X} + \sqrt{\text{Var}Y}\right)^2} \\ &= \sqrt{\text{Var}X} + \sqrt{\text{Var}Y} \end{aligned}$$

from which Subadditivity follows.

- *Variance principle*

This principle is based on the idea of loading the expectation of X with a proportion of the variance of X , i.e. for $a > 0$,

$$\pi(X) = E[X] + a\text{Var}[X].$$

We can now check the different properties from the previous subsection.

1. *Non-negative loading* is obviously fulfilled.
2. *Maximal loss* is not fulfilled: Consider again the binomial risk such that $P(X = m) = p = 1 - P(X = 0)$. Now we have that the premium as a function of p equals

$$g(p) = \pi(X) = mp + am^2p(1 - p) = mp(1 + am(1 - p)),$$

such that

$$g(p) > m \Leftrightarrow mp(1 + am(1 - p)) > m \Leftrightarrow pam > 1.$$

Thus, we can choose p , a , and m such that the *maximal loss* is not fulfilled.

3. *Monotonicity* is not fulfilled: The same argument as for the standard deviation principle applies since the price for a constant risk is the constant itself also for the variance principle.
4. *Subadditivity* is not fulfilled: Consider the two risks $\frac{1}{2}X$ and $\frac{1}{2}X$.

$$\begin{aligned}
\pi\left(\frac{1}{2}X\right) + \pi\left(\frac{1}{2}X\right) &= 2\left(E\left[\frac{1}{2}X\right] + a\text{Var}\left[\frac{1}{2}X\right]\right) \\
&= 2\left(\frac{1}{2}E[X] + a\frac{1}{4}\text{Var}[X]\right) \\
&= E[X] + \frac{a}{2}\text{Var}[X] \\
&< E[X] + a\text{Var}[X] \\
&= \pi(X),
\end{aligned}$$

which shows that subadditivity is violated. In general we can split up a risk in n proportions and get the relation,

$$\begin{aligned}
n\pi\left(\frac{X}{n}\right) &= E[X] + \frac{a}{n}\text{Var}[X] \\
&= \pi(X) - a\left(1 - \frac{1}{n}\right)\text{Var}[X] \\
&< \pi(X).
\end{aligned}$$

Since for $\pi(X) - a\left(1 - \frac{1}{n}\right)\text{Var}[X] \rightarrow \pi(X)$ for $n \rightarrow \infty$, we can get around the risk loading by splitting up in many small proportions. This is a very unfortunate property of the variance principle.

- *Exponential principle*

This principle is based on the idea of tilting X by an exponential (convex) function, taking the expectation, and tilting back by the logarithmic function.

$$\pi(X) = \frac{1}{a} \log E[e^{aX}].$$

- *Esscher principle*

This principle is based on the idea of tilting X with the stochastic factor $\frac{e^{aX}}{E[e^{aX}]}$, that has expectation one but is correlated with X and then take expectation,

$$\pi(X) = E\left[X \frac{e^{aX}}{E[e^{aX}]}\right] = \frac{E[Xe^{aX}]}{E[e^{aX}]}.$$

- *Quantile principle*

This principle is based on the idea that one should pay the smallest premium such that the probability of a net loss is smaller than a given tolerance level,

$$\pi(X) = \min\{m : P(X > m) \leq a\}.$$

4.3 Exercises

Exercise 14 Calculate the premium according to the variance principle, the exponential principle, and the Esscher principle in the case where X is Gamma distributed.

Exercise 15 Let X be a risk corresponding to a Pareto distributed claim happening with probability p . Calculate the premium according to the expected value principle, the variance principle, the standard deviation principle, and the quantile principle.

Exercise 16 A risk X is split into two parts kX and $(1 - k)X$. One part is transferred to an insurance company using the risk loading factor a for premium calculation. The other part is transferred to an insurance company using the risk loading factor $b > a$ for premium calculation. Which choice of k is the cheapest one in case of the expected value principle, the standard deviation principle, and the variance principle?

5 Life Insurance: Products and Valuation

5.1 Single Payments

In this chapter we return to the single payments and payment streams that we introduced in Chapter 1. The idea is to link the state of life of an individual to these payments. By linking the state of life to a payment we mean that whether a given payment actually falls due at a given time point in the future relies on whether the individual is alive or dead then.

We start out by looking at a single benefit payment b paid out at time t . Recall that the present value at time 0 is given by its t -year discounted value,

$$V(0) = (1 + i)^{-t} b = v^t b = e^{-\delta t} b. \quad (8)$$

In general it may be so that the payment b is not deterministic. Let us replace the payment of b by the payment of a stochastic variable b . Now the present value is given by $e^{-\delta t} b$, which is of course also a stochastic variable but now measured at time 0. This present value is not known at time 0 where the agreement of the payment is underwritten. In order to e.g. determine a price for which one can buy at time 0 this agreement we need to assign a value to the present value $e^{-\delta t} b$. We save the notation $V(0)$ for that value. The equivalence principle (7) suggests a value defined as the *expected present value* (EPV),

$$V(0) = E[e^{-\delta t} b] = e^{-\delta t} E[b], \quad (9)$$

where we use that the discount factor is deterministic. Note that (8) is a special case of (9) since, if b is deterministic, then (8) and (9) coincide.

In connection with life insurance we shall use the indicator function

$$1_{\{\dots\}} = \begin{cases} 1 & \text{if } \{\dots\} \text{ is true,} \\ 0 & \text{if } \{\dots\} \text{ is false,} \end{cases}$$

for definition of the payment b in terms of the residual life time of an x -year old. We assume that the time (of an insurance policy) starts at this age, i.e. the individual has age x at time 0. This is

just a matter of convention. One elementary payment is

$$b = 1_{(T_x > t)}, \quad (10)$$

such that 1 unit is paid out if the individual is alive at time t where the individual is $x + t$ years old. Another elementary payment is

$$b = 1_{(s < T_x \leq t)}, \quad (11)$$

such that 1 unit is paid out if the individual dies between time s and time t , i.e. between age $x + s$ and $x + t$.

Recall the conditional probability function and the conditional density,

$$\begin{aligned} {}_tq_x &= 1 - {}_tp_x = 1 - e^{-\int_0^t \mu(x+s)ds} = 1 - e^{-\int_x^{x+t} \mu(s)ds}, \\ f_x(t) &= \frac{d}{dt} {}_tq_x = e^{-\int_0^t \mu(x+s)ds} \mu(x+t). \end{aligned}$$

We can now calculate $E[b]$ for the two choices of b above. In terms of the conditional density f_x we get for b given by (10),

$$E[1_{(T_x > t)}] = \int_0^\infty 1_{(u > t)} f_x(u) du = \int_t^\infty f_x(u) du.$$

Another more direct calculation comes from 'the expected value of an indicator function equals the probability of the event indicated'. First note that although T_x has a continuous distribution $1_{(T_x > t)}$ has a discrete distribution, since it only takes the values 0 and 1. The expression 'the expected value of an indicator function equals the probability of the event indicated' follows from (for the event $X = y$ in a discrete distribution)

$$E[1_{(X=y)}] = \sum_x 1_{(x=y)} P(X=x) = P(X=y).$$

In mathematics, (the solution for constant mortality in brackets),

$$E[1_{(T_x > t)}] = P(T_x > t) = {}_tp_x = e^{-\int_0^t \mu(x+s)ds} (= e^{-\mu t}).$$

In terms of the conditional density f_x we get for b given by (11),

$$E[1_{(s < T_x \leq t)}] = \int_0^\infty 1_{(s < u \leq t)} f_x(u) du = \int_s^t f_x(u) du.$$

The mathematics of the more direct calculation are

$$\begin{aligned} E[1_{(s < T_x \leq t)}] &= P(s < T_x \leq t) = {}_sp_x - {}_tp_x = e^{-\int_0^s \mu(x+u)du} - e^{-\int_0^t \mu(x+u)du} \\ &= e^{-\int_0^s \mu(x+u)du} \left(1 - e^{-\int_s^t \mu(x+u)du}\right) = e^{-\int_0^s \mu(x+u)du} \left(1 - e^{-\int_{x+s}^{x+t} \mu(u)du}\right) \\ &= {}_sp_x - {}_tp_x = e^{-\mu s} \left(1 - e^{-(t-s)\mu}\right). \end{aligned}$$

Returning to the value (according to the equivalence principle) $V(0)$ of the payment in (10) we have that

$$V(0) = e^{-\delta t} e^{-\int_0^t \mu(x+s)ds} = e^{-\int_0^t (\delta + \mu(x+s))ds}.$$

What would be the point in having such an agreement instead of putting the money in the bank? Why would the individual give up the chance of getting 1 unit paid out (to his dependants) in

case he is dead at time t . Well, because it saves him some money now. Should he wish a contract that pays out 1 unit at time t , no matter his state of life, this would cost him $e^{-\delta t}$. If he, for some reason, is not interested in supporting his dependants by the payment of one unit in case he is dead, this would only cost him $e^{-\int_0^t (\delta + \mu(x+s)) ds}$. Thus, he gets a discount (rabat) in accordance with the probability that the insurance company should not pay anything after all. Reasons for not supporting the dependants at time t could be that he has no dependants or that his dependants are his children and they are supposed to be able to take care of themselves by the time t . Thus, to buy this pure saving contract in the insurance company is simply cheaper than buying a saving contract (with no life insurance risk) in a bank.

5.2 Streams of Payments

As in the first chapter we can generalize the single payment to a stream of payments. We should then decorate b with a time argument such that $b(t)$ is a benefit paid out at time t . Recall from Chapter 1 that if $b = (b(0), b(1), \dots, b(n))$ is deterministic, then

$$V(0) = \sum_{t=1}^n (1+i)^{-t} b(t) = \sum_{t=1}^n v^t b(t) = \sum_{t=1}^n e^{-\delta t} b(t).$$

In case b is stochastic we can still calculate the present value above but this is itself then a stochastic variable of which we, at time 0, do not know the outcome. Therefore we cannot speak of this present value as the value. Again, in accordance with the equivalence principle, we use the notation V for the *expected* present value,

$$V(0) = E \left[\sum_{t=1}^n e^{-\delta t} b(t) \right] = \sum_{t=1}^n e^{-\delta t} E[b(t)].$$

This value $V(0)$ is then interpreted as the 'price' of b . But there may be reasons that the individual is not able to pay this price now. He wishes to pay for his benefits by paying a contribution payment stream, possibly also stochastic. Conforming with above, the definition of V for general payment processes with b and $c = (c(0), c(1), \dots, c(n))$ is

$$V(0) = E \left[\sum_{t=1}^n e^{-\delta t} (b(t) - c(t)) \right] = \sum_{t=1}^n e^{-\delta t} (E[b(t)] - E[c(t)]). \quad (12)$$

As in Chapter 1 we can also here consider a reserve at a time point t . One could here suggest the reserve formula

$$V(t) = E \left[\sum_{s=t+1}^n e^{-\delta(s-t)} (b(s) - c(s)) \right],$$

calculating the expected present value of all future payments discounted back to time t . This is similar to the prospective reserve defined in Chapter 1 except for the expectation operator. Therefore, it definitely at the same time generalizes (12) to valuation at other points in time than 0 and generalizes the Chapter 1 prospective reserve to stochastic payments. However, this is *not* how a prospective reserve is defined in life insurance!!! What seems to be natural is to take into account the information about the insured that the company has gathered over $[0, t]$. If all risk is described by T_x , this information amounts to knowing whether $\{T_x > t\}$ is true or not. If we know that the insured is alive, i.e. $\{T_x > t\}$ is true, then why not update the expectation in accordance with this knowledge, and forming a conditional expectation given $T_x > t$. On the other hand, if we know that the insured is dead, i.e. $\{T_x > t\}$ is false, then why not update the expectation in accordance with this knowledge, and forming a conditional expectation given that $T_x \leq t$: E.g. if

there are no more payments after death, and the insured is dead, then it seems natural to take the prospective reserve to be zero. Working with this 'correct' notion of prospective reserves as conditional expected values is eased by basic abstract knowledge about conditioning and we will not pursue this further here.

5.3 Fair Payment Streams

In the deterministic case we discussed the fairness criterions, respectively,

$$\begin{aligned} U(n) &= 0, \\ V(0) &= c(0) - b(0), \end{aligned}$$

and we found that these criterions were equivalent. Here, we forget about the first and maintain the second one as a generalized (to stochastic payments) equivalence principle with $V(0)$ defined in (12). Since $b(0) - c(0)$ is known at time 0, we have that $E[b(0) - c(0)] = b(0) - c(0)$ and we can write the equivalence principle as

$$V(0) + b(0) - c(0) = E \left[\sum_{t=0}^n e^{-\delta t} (b(t) - c(t)) \right] = 0.$$

Taking the contributions on the right hand side we get

$$E \left[\sum_{t=0}^n e^{-\delta t} b(t) \right] = E \left[\sum_{t=0}^n e^{-\delta t} c(t) \right], \quad (13)$$

stating that the expected present value of benefits must equal the expected present value of contributions.

The equivalence principle is one equation, so there has to be one unknown which is then to be determined in accordance. This could be a particular benefit or a particular premium. Typically it is an (unknown) *single premium* (indskudspræmie) paid at time 0 or a *level premium* (løbende præmie) paid with the same (unknown) premium amount paid in every period. This single or level premium is then determined per benefit. To see how this machinery works, we can just as well go directly to the most important life insurance payments streams.

5.4 Particular Payment Streams

- A *pure endowment insurance* (ren oplevelsesforsikring) pays one unit to the policy holder at time n if he is alive at time n . This formalized in $b = (0^{1 \times n}, 1_{(T_x > n)})$ or in more concentrated form $b(t) = 1_{(T_x > t)} 1_{(t=n)}$ (realize this!). The expected present value at time 0 of the benefit is given by

$$\begin{aligned} {}_nE_x &= E [e^{-\delta n} 1_{(T_x > n)}] \\ &= e^{-\delta n} E [1_{(T_x > n)}] \\ &= e^{-\delta n} {}_n p_x \\ &= e^{-\delta n} e^{-\int_0^n \mu(x+s) ds} \\ &= e^{-\int_0^n \delta + \mu(x+s) ds}. \end{aligned}$$

If the mortality intensity is constant we get that

$${}_nE_x = e^{-\int_0^n (\delta + \mu) ds} = e^{-(\delta + \mu)n}.$$

If this is paid for by a single premium π_0 , i.e. $c = (\pi_0, 0^{1 \times n})$, the equivalence principle reads

$$\pi_0 = e^{-\int_0^n \delta + \mu(x+s) ds},$$

so ${}_nE_x$ is also the equivalence single premium π_0 for a pure endowment.

- A *term insurance* (ren dødsfaldsforsikring) pays one unit to the policy holder's dependants at time t if he died in the t th period before time n . This is formalized by in $b = (0, 1_{(0 < T_x \leq 1)}, 1_{(1 < T_x \leq 2)}, \dots, 1_{(n-1 < T_x \leq n)})$ or in more concentrated form $b(t) = 1_{(t-1 < T_x \leq t)} 1_{(t \leq n)}$ starting with $b(0) = 0$ since $T_x > 0$. The expected present value at time 0 of the benefits is given by

$$\begin{aligned} A_{xn}^1 &= E \left[\sum_{t=1}^n e^{-\delta t} 1_{(t-1 < T_x \leq t)} \right] \\ &= \sum_{t=1}^n e^{-\delta t} E [1_{(t-1 < T_x \leq t)}] \\ &= \sum_{t=1}^n e^{-\delta t} ({}_{t-1}p_x - {}_t p_x) \\ &= \sum_{t=1}^n e^{-\delta t} {}_{t-1}p_x q_{x+t-1}, \end{aligned}$$

which is also the equivalence single premium for the term insurance.

If the mortality intensity is constant we get

$$\begin{aligned} A_{xn}^1 &= \sum_{t=1}^n e^{-\delta t} e^{-\mu(t-1)} (1 - e^{-\mu}) \\ &= (1 - e^{-\mu}) e^{\mu} \sum_{t=1}^n e^{-(\delta+\mu)t} \\ &= (1 - e^{-\mu}) e^{\mu} e^{-(\delta+\mu)} \sum_{t=1}^n e^{-(\delta+\mu)(t-1)} \\ &= (1 - e^{-\mu}) e^{-\delta} \sum_{t=0}^{n-1} e^{-(\delta+\mu)t} \\ &= (1 - e^{-\mu}) e^{-\delta} \sum_{t=0}^{n-1} e^{-(\delta+\mu)t} \\ &= (1 - e^{-\mu}) e^{-\delta} \frac{1 - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}}. \end{aligned}$$

Note that for $\delta = 0$,

$$A_{xn}^1 = E \left[\sum_{t=1}^n 1_{(t-1 < T_x \leq t)} \right] = E [1_{(T_x \leq n)}] = {}_nq_x.$$

Thus, it is the weighting of the payments by the discount factor which makes the calculations harder than just calculating the probability that the insured dies before time n .

It is also possible to have a 'never-ending' term insurance. This is called a whole life term insurance and has the expected present value

$$A_x^1 = \sum_{t=1}^{\infty} e^{-\delta t} E [1_{(t-1 < T_x \leq t)}].$$

- An *endowment insurance* (oplevelsesforsikring med dødsfaldsdækning eller livsforsikring med udbetaling) combines a pure endowment insurance and a term insurance into one insurance (and therefore it is obviously not canonical in the sense that it can be decomposed further). This is formalized in $b = (0, 1_{(0 < T_x \leq 1)}, 1_{(1 < T_x \leq 2)}, \dots, 1_{(n-1 < T_x \leq n)} + 1_{(T_x > n)})$ or in more

concentrated form $b(t) = 1_{(t-1 < T_x \leq t)} 1_{(t \leq n)} + 1_{(T_x > t)} 1_{(t=n)}$. The expected present value at time 0 is

$$\begin{aligned} A_{xn} &= \sum_{t=1}^n e^{-\delta t} E[1_{(t-1 < T_x \leq t)}] + e^{-\delta n} E[1_{(T_x > n)}] \\ &= \sum_{t=1}^n e^{-\delta t} {}_{t-1}p_x q_{x+t-1} + e^{-\delta n} {}_n p_x \\ &= {}_n E_x + A_{xn}^1. \end{aligned}$$

which is also the equivalence single premium for the term insurance.

If the mortality intensity is constant we get

$$\begin{aligned} A_{xn} &= (1 - e^{-\mu}) e^{-\delta} \frac{1 - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}} + e^{-(\delta+\mu)n} \\ &= e^{-\delta} \frac{1 - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}} - e^{-(\delta+\mu)} \frac{1 - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}} + e^{-(\delta+\mu)n} \\ &= e^{-\delta} \frac{1 - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}} - e^{-(\delta+\mu)} \frac{1 - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}} + \frac{e^{-(\delta+\mu)n} (1 - e^{-(\delta+\mu)})}{1 - e^{-(\delta+\mu)}} \\ &= e^{-\delta} \frac{1 - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}} - \frac{e^{-(\delta+\mu)} - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}}. \end{aligned}$$

Note that a whole life endowment insurance is the same as a whole life term insurance since the pure endowment part only has an effect if the insurance contract ends at some time point.

- A *temporary life annuity(-due)* (ophørende (forudbetalt) livrente) is an annuity that runs as long as the insurance is alive but only until time t . Usually we write that the annuity runs until death or termination whatever occurs first. This is formalized in $b = (1_{(T_x > 0)}, 1_{(T_x > 1)}, \dots, 1_{(T_x > n-1)}, 0)$ or in more concentrated form $b(t) = 1_{(T_x > t)} 1_{(t < n)}$. The expected present value at time 0 equals

$$\begin{aligned} \ddot{a}_{x\overline{n}|} &= \sum_{t=0}^{n-1} e^{-\delta t} E[1_{(T_x > t)}] \\ &= \sum_{t=0}^{n-1} e^{-\delta t} {}_t p_x. \end{aligned}$$

If the mortality intensity is constant we get

$$\begin{aligned} \ddot{a}_{x\overline{n}|} &= \sum_{t=0}^{n-1} e^{-\delta t} e^{-\mu t} \\ &= \sum_{t=0}^{n-1} e^{-(\delta+\mu)t} \\ &= \frac{1 - e^{-(\delta+\mu)n}}{1 - e^{-(\delta+\mu)}}. \end{aligned}$$

This also exists in a non-temporary form, i.e. as a whole life annuity. The expected present value is

$$\ddot{a}_x = \sum_{t=0}^{\infty} e^{-\delta t} {}_t p_x.$$

The temporary life annuity is presented above as a benefit payment stream. However, more often it represents the premium payment profile in the sense that a premium π is paid at the beginning of each period until death or termination whatever occurs first. This is formalized in $c = (\pi 1_{(T_x > 0)}, \pi 1_{(T_x > 1)}, \dots, \pi 1_{(T_x > n-1)}, 0)$ or in more concentrated form $c(t) = \pi 1_{(T_x > t)} 1_{(t < n)}$. A

policy holder is e.g. interested in such a premium payment profile if he cannot afford the payment of the single premium at time 0. With this premium profile we can now, for a given benefit payment stream, calculate the equivalence level premium π in accordance with (13).

Consider first the pure endowment insurance. For this contract with $b = (0^{1 \times n}, 1_{(T_x > n)})$ and $c = (\pi 1_{(T_x > 0)}, \pi 1_{(T_x > 1)}, \dots, \pi 1_{(T_x > n-1)}, 0)$, the equivalence relation can be written

$$e^{-\delta n} E [1_{(T_x > n)}] = \pi \sum_{t=0}^{n-1} e^{-\delta t} E [1_{(T_x > t)}],$$

or, in terms of survival probabilities,

$$e^{-\delta n} {}_n p_x = \pi \sum_{t=0}^{n-1} e^{-\delta t} {}_t p_x,$$

or, in terms of actuarial notation,

$${}_n E_x = \pi \ddot{a}_{x:\overline{n}|}.$$

This is one equation with one unknown and the solution is

$$\pi = \frac{{}_n E_x}{\ddot{a}_{x:\overline{n}|}}.$$

This premium paid at the beginning of every period sees to it that the insurance company at time zero where the contract is issued, faces the same expected present value as income and outgo.

Similar calculations for the term insurance benefit gives that the equivalence level premium solves the equivalence relation $A_{xn}^1 = \pi \ddot{a}_{xn}$ such that

$$\pi = \frac{A_{xn}^1}{\ddot{a}_{x:\overline{n}|}}.$$

For the endowment insurance we get the equivalence premium

$$\pi = \frac{A_{xn}}{\ddot{a}_{x:\overline{n}|}} = \frac{{}_n E_x + A_{xn}^1}{\ddot{a}_{x:\overline{n}|}},$$

i.e. just the sum of the premium for the pure endowment and the premium for the term insurance.

All premium calculated above are elementary premiums in the sense that they form the 'price' of *one unit* of all the different benefits. But it is very easy to calculate the premium of multiple units. If e.g. the pure endowment pays out k instead of 1 upon survival until time n , then the expected present value of benefits equals $k {}_n E_x$, which is also the equivalence single premium, and the equivalence level premium equals

$$\pi = k \frac{{}_n E_x}{\ddot{a}_{x:\overline{n}|}}.$$

Thus, the equivalence premium for k benefits is just equal to k times the equivalence premium for 1 benefit.

5.5 Exercises

Exercise 17 We consider a so-called deferred temporary life annuity(-due) (*opsat ophørende forudbetalt livrente*) where

$$b(t) = 1_{(T_x > t, k \leq t < k+n)} = 1_{(T_x > t)} 1_{(k \leq t < k+n)}$$

with present value

$$\sum_{t=0}^{k+n} b(t) e^{-\delta t}.$$

a) Explain this product in words and argue why a pension saver may be interested in such a product in practice.

b) Calculate $E \left[\sum_{t=0}^{k+n} b(t) e^{-\delta t} \right]$ in general in terms of ${}_t p_x$ and show that

$$E \left[\sum_{t=0}^{k+n} b(t) e^{-\delta t} \right] = {}_k E_x \ddot{a}_{x+k|\overline{n}|}.$$

c) Calculate $E \left[\sum_{t=0}^{k+n} b(t) e^{-\delta t} \right]$ in the special cases where μ is constant, where $n = \infty$, and the combination where μ is constant and $n = \infty$.

d) Suggest an actuarial notation for the deferred temporary life annuity and for the deferred life annuity ($n = \infty$), inspired by the notation you already know (look in Chapter 1 for the deferment notation and in Chapter 5 for the 'everlasting' notation).

e) Assume that the annuity benefit is paid for by a level premium π in k periods (løbende præmie) with expected present value $\ddot{a}_{x:\overline{k}|}$. Calculate the premium according to the equivalence principle.

f) Calculate the premium in the special cases where μ is constant, where $n = \infty$, and the combination where μ is constant and $n = \infty$.

Exercise 18 Consider an insurance contract on a couple consisting of two spouses (ægtefæller) with initial ages x and y . We assume that the insurance company uses a unisex mortality table, i.e. the same age dependent mortality rate μ is used for males and females. We also assume that the lifetimes of the two spouses are independent (!!!). Consider an insurance contract where y receives a unit at time n if at that time, x is dead and y is alive.

a) Formalize the claim $b(n)$ in terms of the two stochastic variables T_x and T_y .

b) What is the expected present value at time 0 of this claim?

c) Assume that the premium is paid annually as long as x is alive. Determine the equivalence premium.

Exercise 19 Assume a linear mortality intensity, i.e. $\mu(t) = \mu t$. Show that

$$e^{-\int_s^t \mu(x+u) du} = e^{-\mu(x(t-s) + \frac{1}{2}(t^2-s^2))}.$$

Show that

$$\begin{aligned} {}_n E_x &= e^{-(\mu x + \delta)n - \mu \frac{1}{2} n^2}, \\ A_{x:\overline{n}|}^1 &= \sum_{t=1}^n e^{-\delta t} e^{-\mu(x(t-1) + \frac{1}{2}(t-1)^2)} \left(1 - e^{-\mu(x+t-\frac{1}{2})}\right), \\ A_{x:\overline{n}|} &= e^{-\delta n} e^{-\mu(xn + \frac{1}{2} n^2)} + \sum_{t=1}^n e^{-\delta t} e^{-\mu(x(t-1) + \frac{1}{2}(t-1)^2)} \left(1 - e^{-\mu(x+t-\frac{1}{2})}\right), \\ \ddot{a}_{x:\overline{n}|} &= \sum_{t=0}^{n-1} e^{-(\mu x + \delta)t - \mu \frac{1}{2} t^2} \end{aligned}$$

6 Non-life Insurance: The Aggregate Claim Distribution

We shall here continue the study of the distribution of the total claim and expected values of functions of the total claim, i.e. $X = \sum_{i=1}^N Y_i$ and $E[f(X)]$. We wish to calculate the probability

distribution $P(X \leq x)$, such that we can calculate so-called fixed time ruin probabilities. Assume that the insurance company starts out with an initial capital c and receives the premiums π . Then the so-called ruin probability $P(X > c + \pi) = 1 - P(X \leq c + \pi)$ specifies the probability that the initial capital plus the premium is inadequate to cover the total losses. There are also other motivations for calculating the total claim distribution.

In the study of stochastic variables, one often studies certain functions namely the probability generating function, the moment generating function, the characteristic function and the Laplace transform

$$\begin{aligned} P_X(s) &= E[s^X], \\ M_X(s) &= E[e^{sX}], \\ \Phi_X(s) &= E[e^{isX}], \\ L_X(s) &= E[e^{-sX}]. \end{aligned}$$

We shall here work with the probability generating function $P_X(s)$. We have the following probability generating functions for the claim number and claim size distributions,

$$\begin{aligned} P_N(s) &= E[s^N] = \sum_{n=0}^{\infty} p(n) s^n, \\ P_Y(s) &= E[s^Y]. \end{aligned}$$

With these functions we get

$$\begin{aligned} P_X(s) &= E[s^X] \left(= \sum_{n=0}^{\infty} E[s^X | N = n] p(n) \right) \\ &= \sum_{n=0}^{\infty} E \left[s^{\sum_{i=1}^n Y_i} \right] p(n) \\ &= \sum_{n=0}^{\infty} E[s^{Y_1} s^{Y_2} \dots s^{Y_n}] p(n) \\ &= \sum_{n=0}^{\infty} E[s^{Y_1}] E[s^{Y_2}] \dots E[s^{Y_n}] p(n) \\ &= \sum_{n=0}^{\infty} (E[s^Y])^n p(n) \\ &= P_N(P_Y(s)). \end{aligned} \tag{14}$$

Now, assume that the severities are positive integer-valued variables with probabilities

$$f(y) = P(Y_1 = y), \quad y = 1, 2, \dots$$

The distribution of X is then given by

$$\begin{aligned} g(x) &= P(X = x) \\ &= \sum_{n=0}^{\infty} P(X = x, N = n) \\ &= \sum_{n=0}^{\infty} p(n) P(X = x | N = n) \\ &= \sum_{n=0}^{\infty} p(n) P(Y_1 + \dots + Y_n = x) \\ &= \sum_{n=0}^{\infty} p(n) f^{*n}(x), \end{aligned}$$

where

$$\begin{aligned}
f^{*n}(x) &= P(Y_1 + \dots + Y_n = x) \\
&= \sum_{y=1}^x P(Y_1 + \dots + Y_n = x, Y_n = y) \\
&= \sum_{y=1}^x f(y) P(Y_1 + \dots + Y_n = x | Y_n = y) \\
&= \sum_{y=1}^x f(y) P(Y_1 + \dots + Y_{n-1} = x - y) \\
&= \sum_{y=1}^x f(y) f^{*(n-1)}(x - y).
\end{aligned} \tag{15}$$

Unfortunately this calculation is slowly and inefficient (One can realize that the number of terms involved in calculation of $g(x)$ is of order $O(x^3)$). In the following two subsections we shall study two alternatives. One is a precise recursion formula which works for certain claim number distributions, the other is a moment-based approximation formula.

Example Let N be Poisson distributed with $\lambda = 1$. Let Y be distributed such that $f(1) = 1/4, f(2) = 1/2, f(3) = 1/4$. Note that $f^{*1}(x) = f(x)$. Then

$$\begin{aligned}
f^{*2}(2) &= f(1)f^{*1}(1) = \frac{1}{4} \frac{1}{4} = \frac{1}{16}, \\
f^{*2}(3) &= f(1)f^{*1}(2) + f(2)f^{*1}(1) = \frac{1}{4} \frac{1}{2} + \frac{1}{2} \frac{1}{4} = \frac{1}{4}, \\
f^{*2}(4) &= f(1)f^{*1}(3) + f(2)f^{*1}(2) + f(3)f^{*1}(1) = \frac{1}{4} \frac{1}{4} + \frac{1}{2} \frac{1}{2} + \frac{1}{4} \frac{1}{4} = \frac{3}{8}, \\
&\vdots \\
f^{*3}(3) &= f(1)f^{*2}(2) = \frac{1}{4} \frac{1}{16} = \frac{1}{64}, \\
&\vdots
\end{aligned}$$

$$\begin{aligned}
g(0) &= p(0) = e^{-1}, \\
g(1) &= p(1)f^{*1}(1) = \frac{1}{4}e^{-1}, \\
g(2) &= p(1)f^{*1}(2) + p(2)f^{*2}(2) = \frac{1}{2}e^{-1} + \frac{1}{16} \frac{1}{2}e^{-1} = \left(\frac{1}{2} + \frac{1}{32}\right)e^{-1}, \\
g(3) &= p(1)f^{*1}(3) + p(2)f^{*2}(3) + p(3)f^{*3}(3) \\
&= \frac{1}{4}e^{-1} + \frac{1}{4} \frac{1}{2}e^{-1} + \frac{1}{64} \frac{1}{6}e^{-1} = \left(\frac{1}{4} + \frac{1}{8} + \frac{1}{384}\right)e^{-1}. \square
\end{aligned}$$

6.1 The Panjer Recursion

Let N belong to the (a, b) -class. We shall first derive a differential equation for the probability generating function for the claim number process,

$$\begin{aligned}
P'_N(s) &= \sum_{n=0}^{\infty} np(n) s^{n-1} \\
&= \sum_{n=1}^{\infty} n \left(a + \frac{b}{n} \right) p(n-1) s^{n-1} \\
&= \sum_{n=1}^{\infty} (na + b) p(n-1) s^{n-1} \\
&= \sum_{n=0}^{\infty} ((n+1)a + b) p(n) s^n \\
&= \sum_{n=0}^{\infty} (a+b) p(n) s^n + as \sum_{n=0}^{\infty} np(n) s^{n-1} \\
&= (a+b) P_N(s) + as P'_N(s),
\end{aligned}$$

which leads to the differential equation

$$P'_N(s) = \frac{a+b}{1-as} P_N(s).$$

This differential equation has the side condition

$$P_N(1) = \sum_{n=0}^{\infty} p(n) = 1$$

If $a = 0$ (the Poisson distribution), the solution to this equation is

$$P_N(s) = e^{b(s-1)}$$

For $a \neq 0$ (the binomial and negative binomial distributions), the solution is

$$P_N(s) = \left(\frac{1-a}{1-as} \right)^{\frac{a+b}{a}}.$$

By differentiation of (14) we get

$$\begin{aligned}
P'_X(s) &= P'_N(P_Y(s)) P'_Y(s) \\
&= \frac{a+b}{1-aP_Y(s)} P_N(P_Y(s)) P'_Y(s) \\
&= \frac{a+b}{1-aP_Y(s)} P_X(s) P'_Y(s),
\end{aligned}$$

such that

$$P'_X(s) - aP'_X(s) P_Y(s) = (a+b) P_X(s) P'_Y(s).$$

Plugging in $P_X(s) = \sum_{x=0}^{\infty} g(x) s^x$, $P'_X(s) = \sum_{x=0}^{\infty} xg(x) s^{x-1}$, $P_Y(s) = \sum_{y=1}^{\infty} f(y) s^y$, and

$P'_Y(s) = \sum_{y=1}^{\infty} y f(y) s^{y-1}$, we get

$$\begin{aligned}
& \sum_{x=0}^{\infty} x g(x) s^{x-1} - a \sum_{y=1}^{\infty} f(y) s^y \sum_{x=0}^{\infty} x g(x) s^{x-1} \\
= & (a+b) \sum_{x=0}^{\infty} g(x) s^x \sum_{y=1}^{\infty} y f(y) s^{y-1} \Leftrightarrow \\
& \sum_{x=1}^{\infty} x g(x) s^{x-1} - a \sum_{y=1}^{\infty} \sum_{x=0}^{\infty} x f(y) g(x) s^{x+y-1} \\
= & (a+b) \sum_{y=1}^{\infty} \sum_{x=0}^{\infty} y g(x) f(y) s^{x+y-1} \Leftrightarrow (z=x+y) \\
& \sum_{x=1}^{\infty} x g(x) s^{x-1} - a \sum_{y=1}^{\infty} \sum_{z=y}^{\infty} (z-y) f(y) g(z-y) s^{z-1} \\
= & (a+b) \sum_{y=1}^{\infty} \sum_{z=y}^{\infty} y g(z-y) f(y) s^{z-1} \Leftrightarrow (\text{switch summation}) \\
& \sum_{x=1}^{\infty} x g(x) s^{x-1} - a \sum_{z=1}^{\infty} \sum_{y=1}^z (z-y) f(y) g(z-y) s^{z-1} \\
= & (a+b) \sum_{z=1}^{\infty} \sum_{y=1}^z y g(z-y) f(y) s^{z-1} \Leftrightarrow (\text{rename } x=z) \\
& \sum_{x=1}^{\infty} x g(x) s^{x-1} - a \sum_{x=1}^{\infty} \sum_{y=1}^x (x-y) f(y) g(x-y) s^{x-1} \\
= & (a+b) \sum_{x=1}^{\infty} \sum_{y=1}^x y g(x-y) f(y) s^{x-1}
\end{aligned}$$

from which we collect terms and summations into

$$\sum_{x=1}^{\infty} \left(x g(x) - \sum_{y=1}^x (ax + by) g(x-y) f(y) \right) s^{x-1} = 0$$

If this true for all s , then

$$\begin{aligned}
x g(x) - \sum_{y=1}^x (ax + by) g(x-y) f(y) &= 0 \Leftrightarrow \\
g(x) &= \sum_{y=1}^x \left(a + b \frac{y}{x} \right) g(x-y) f(y)
\end{aligned}$$

This recursive formula is considerably more efficient and faster than (15), (One can realize that the number of terms involved in calculation of $g(x)$ is of order $O(x^2)$).

Example Consider again the example where N is Poisson distributed with $\lambda = 1$, i.e. $a = 0, b = 1$, and Y is distributed such that $f(1) = 1/4, f(2) = 1/2, f(3) = 1/4$. Then

$$\begin{aligned}
g(0) &= p(0) = e^{-1}, \\
g(1) &= g(0) f(1) = \frac{1}{4} e^{-1}, \\
g(2) &= \frac{1}{2} g(1) f(1) + g(0) f(2) = \frac{1}{2} \frac{1}{4} e^{-1} \frac{1}{4} + e^{-1} \frac{1}{2} = \left(\frac{1}{32} + \frac{1}{2} \right) e^{-1}, \\
g(3) &= \frac{1}{3} g(2) f(1) + \frac{2}{3} g(1) f(2) + g(0) f(3) \\
&= \frac{1}{3} \left(\frac{1}{32} + \frac{1}{2} \right) e^{-1} \frac{1}{4} + \frac{2}{3} \frac{1}{4} e^{-1} \frac{1}{2} + e^{-1} \frac{1}{4} \\
&= \left(\left(\frac{1}{32} + \frac{1}{2} \right) \frac{1}{12} + \frac{2}{24} + \frac{1}{4} \right) e^{-1} \\
&= \left(\frac{1}{384} + \frac{3}{24} + \frac{1}{4} \right) e^{-1} \\
&= \left(\frac{1}{4} + \frac{1}{8} + \frac{1}{384} \right) e^{-1}. \square
\end{aligned}$$

6.2 Normal Power Approximation

In certain situation one may be willing to give up part of the accuracy in order to speed up the calculation. We shall consider a moment-based approximation method called the Normal Power approximation. Let $u_{1-\varepsilon}$ denote the $1 - \varepsilon$ fractile of G , i.e. $u_{1-\varepsilon}$ is the smallest number u such that $1 - G(u) \leq \varepsilon$, and let $z_{1-\varepsilon}$ denote the $1 - \varepsilon$ fractile of the standard normal distribution. NP approximation says that approximately

$$u_{1-\varepsilon} = E[X] + z_{1-\varepsilon} \sqrt{Var[X]} + \frac{1}{6} (z_{1-\varepsilon}^2 - 1) \frac{E[(X - E[X])^3]}{Var[X]}.$$

The total claim distribution and the normal power approximation may be used in the quantile premium principle for valuation of a claim X . Recall the premium principle

$$\pi(X) = \min \{m : P(X > m) \leq a\}.$$

If we approximate the aggregate claim distribution by the normal power we get exactly that

$$\pi(X) = u_{1-a}.$$

This is typically not thought of as a valuation principle but rather as a solvency criterion principle. A solvency rule should protect the policy holders such that the insurance company is able to pay the losses with a high probability. The solvency requirement is a minimum requirement on capital. One natural solvency rule is to require from the insurance company that it has sufficiently capital such that the probability that the aggregate losses exceed the capital is smaller than a given tolerance level ε . An approximated solvency rule is that the company should at least hold the capital $u_{1-\varepsilon}$.

6.3 Ruin Theory

Studying the probability distribution of X corresponds to studying the so-called fixed time ruin probability. Other types of ruin probabilities are also of interest and belongs to very large discipline of non-life insurance called risk theory or ruin theory.

Above we have been working with N as a claim number stochastic variable. This corresponds to measuring at a fixed point in time the total number of claims considered as a function of time and measured by a so-called counting process $N(t)$ counting the number of claims over $[0, t]$. If at any fixed point in time the claim number is Poisson distributed, the process $N(t)$ is called a Poisson process. Based on a counting process one can introduce the process of total claims or briefly, the risk process,

$$X(t) = \sum_{i=1}^{N(t)} Y_i.$$

If N is a Poisson process then X is called a compound Poisson process. Assume now that the insurance company over $[0, t]$ receives the premiums πt . Then for a fixed time horizon t^* , the probability

$$P(X(t^*) > c + \pi t^*)$$

corresponds to the probability which we shall study here. The insurance company is, however, also interested in probabilities in the forms

$$\begin{aligned} P(X(t) > c + \pi t, t < t^*), \\ P(X(t) > x + \pi t, \forall t), \end{aligned}$$

i.e. the so-called finite time ruin probability that the insurance company will not be ruined until time t^* and the infinite time ruin probability that the insurance company will never be ruined.

6.4 Exercises

Exercise 20 a) For N belong to the (a, b) -class, show that

$$E[N^k] = aE[(N+1)^k] + bE[(N+1)^{k-1}].$$

b) Find simple expressions for $E[N]$ and $\text{Var}[N]$ in terms of a and b , and show that

$$\frac{\text{Var}[N]}{E[N]} = \frac{1}{1-a}.$$

Exercise 21 The special case of the negative binomial distribution where $\alpha = 1$, i.e. $p(n) = q^n(1-q)$ is called the geometric distribution. Assume that N is geometrically distributed. Show that the distribution function G of X , $G(x) = P(X \leq x)$, satisfies the recursion formula

$$G(x) = (1-q) + q \sum_{y=1}^x f(y) G(x-y).$$

Exercise 22 Make sure that you follow all the steps in the following derivation of the Panjer

recursion for the moments of the total claim distribution,

$$\begin{aligned}
E[X^k] &= \sum_{x=0}^{\infty} x^k g(x) \\
&= \sum_{x=1}^{\infty} x^k g(x) \\
&= \sum_{x=1}^{\infty} x^{k-1} x g(x) \\
&= \sum_{x=1}^{\infty} x^{k-1} \sum_{y=1}^x (ax + by) g(x-y) f(y) \\
&= \sum_{x=1}^{\infty} \sum_{y=1}^x (ax^k + byx^{k-1}) g(x-y) f(y) \\
&= \sum_{y=1}^{\infty} \sum_{x=y}^{\infty} (ax^k + byx^{k-1}) g(x-y) f(y) \\
&= \sum_{y=1}^{\infty} \sum_{z=0}^{\infty} (a(y+z)^k + by(y+z)^{k-1}) g(z) f(y) \\
&= a \sum_{y=1}^{\infty} \sum_{z=0}^{\infty} \sum_{l=0}^k \binom{k}{l} z^l y^{k-l} g(z) f(y) \\
&\quad + b \sum_{y=1}^{\infty} \sum_{z=0}^{\infty} y \sum_{l=0}^{k-1} \binom{k-1}{l} z^l y^{k-1-l} g(z) f(y) \\
&= a \sum_{y=1}^{\infty} \sum_{z=0}^{\infty} \sum_{l=0}^{k-1} \binom{k}{l} z^l y^{k-l} g(z) f(y) + a \sum_{y=1}^{\infty} \sum_{z=0}^{\infty} z^k g(z) f(y) \\
&\quad + b \sum_{y=1}^{\infty} \sum_{z=0}^{\infty} \sum_{l=0}^{k-1} \binom{k-1}{l} z^l y^{k-l} g(z) f(y) \\
&= a \sum_{l=0}^{k-1} \binom{k}{l} \sum_{y=1}^{\infty} y^{k-l} f(y) \sum_{z=0}^{\infty} z^l g(z) + a \sum_{z=0}^{\infty} z^k g(z) \sum_{y=1}^{\infty} f(y) \\
&\quad + b \sum_{l=0}^{k-1} \binom{k-1}{l} \sum_{y=1}^{\infty} y^{k-l} f(y) \sum_{z=0}^{\infty} z^l g(z) \\
&= a \sum_{l=0}^{k-1} \binom{k}{l} E[Y^{k-l}] E[X^l] + a E[X^k] \\
&\quad + b \sum_{l=0}^{k-1} \binom{k-1}{l} E[Y^{k-l}] E[X^l] \\
&= a E[X^k] + \sum_{l=0}^{k-1} \left(a \binom{k}{l} + b \binom{k-1}{l} \right) E[Y^{k-l}] E[X^l]
\end{aligned}$$

and conclude that

$$E[X^k] = \frac{1}{1-a} \sum_{l=0}^{k-1} \left(a \binom{k}{l} + b \binom{k-1}{l} \right) E[Y^{k-l}] E[X^l].$$

Exercise 23 Consider the Normal Power approximation in case of a Poisson λ claim distribution. Show that

$$E[X^k] = \lambda \sum_{l=0}^{k-1} \binom{k-1}{l} E[Y^{k-l}] E[X^l].$$

The first two moments are known from a previous exercise. Verify these results and show that

$$E[X^3] = \lambda E[Y^3] + 3\lambda^2 E[Y^2] E[Y] + \lambda^3 E^3[Y].$$

Show that

$$E \left[(X - EX)^3 \right] = \lambda E \left[Y^3 \right],$$

and conclude the following NP approximation formula

$$u_{1-\varepsilon} = \lambda E[Y] + z_{1-\varepsilon} \sqrt{\lambda E[Y^2]} + \frac{1}{6} (z_{1-\varepsilon}^2 - 1) \frac{E[Y^3]}{E[Y^2]}.$$

Exercise 24 Verify, by means of,

$$\begin{aligned} E[N] &= \frac{a+b}{1-a}, \\ \text{Var}[N] &= \frac{a+b}{(1-a)^2}, \end{aligned}$$

and the recursion formula for moments, that the simple general relations

$$\begin{aligned} E[X] &= E[Y] E[N], \\ \text{Var}[X] &= E[N] \text{Var}[Y] + E^2[Y] \text{Var}[N], \end{aligned}$$

also, of course, holds when N belongs to the (a, b) -class.

7 Finance and Reinsurance: Products and Valuation

Risk free investment - corresponds to chapter 1

$$\begin{aligned} S^0(0) &= 1, \\ S^0(1) &= (1+r) S^0(0) = 1+r \end{aligned}$$

This is new

$$\begin{aligned} S^1(0) &= s, \\ S^1(1) &= s(1+Z), \\ Z &= \begin{cases} u \text{ with probability } p_u, \\ d \text{ with probability } p_d. \end{cases} \end{aligned}$$

but it is not new to have such a two point distribution. Recall e.g.

$$1_{(T_x > 1)}$$

which also has a two point distribution.

claim, benefit, amount that you receive

$$\Phi(S^1(1))$$

depends on the outcome of $S^1(1)$. Compare with the benefit $1_{(T_x > 1)}$ which also depends on (is even equal to) the outcome of $1_{(T_x > 1)}$.

Example: $r = 0.1$, $s = 1$, $u = 0.4$, $d = 0.9$, $p_u = 60\%$, $p_d = 40\%$.

Consider the claim

$$\max(S^1(1), 1)$$

What should be the price for this claim: Calculate the expected present value

$$\begin{aligned}
 \pi &= E[PV(0)] + \text{risk loading} \\
 &= \frac{1}{1+r} E[\max(S^1(1), 1)] + \text{risk loading} \\
 &= \frac{1}{1.1} (0.6 * 1.4 + 0.4 * 1) + \text{risk loading} \\
 &= 1.127273 + \text{risk loading}
 \end{aligned}$$

Consider the situation where I can invest the premium paid in S^1 . Then we can put up two equations with two unknowns

$$\begin{aligned}
 1.4x + (\pi - x) 1.1 &= 1.4 \\
 0.9x + (\pi - x) 1.1 &= 1
 \end{aligned}$$

$$\begin{aligned}
 0.5x &= 0.4 \Rightarrow x = 0.8 \\
 0.3 * 0.8 + 1.1\pi &= 1.4 \Rightarrow \pi = \frac{1.4 - 0.8 * 0.3}{1.1} = 1.054545
 \end{aligned}$$

$$\begin{aligned}
 \pi &= E^*[PV(0)] \\
 &= \frac{1}{1+r} E^*[\max(S^1(1), 1)] \\
 &= \frac{1}{1.1} \left(\frac{1.1 - 0.9}{1.4 - 0.9} * 1.4 + \frac{1.4 - 1.1}{1.4 - 0.9} * 1 \right) \\
 &= \frac{1}{1.1} (0.6 * 1.4 + 0.4 * 1)
 \end{aligned}$$

In general it seems very restrictive to work with only two outcomes, but if we have more outcomes we can compensate by more trading dates. The continuous (lognormally distributed) outcome combined with continuous trading was solved in 1973 and was in 1997 awarded the Nobel prize in economics.

Optioner:

$$\Phi = \max(S^1(1) - K, 0) \Rightarrow \text{determine } \pi \text{ for given } K$$

Futures:

$$\Phi = S^1(1) - K : \text{determine } k \text{ such that } \pi = 0.$$

The insurance version:

$$\begin{aligned}
 S^0(0) &= 1, \\
 S^0(1) &= (1+r) S^0(0) = 1+r
 \end{aligned}$$

This is new

$$\begin{aligned}
 S^1(0) &= \pi^r, \\
 S^1(1) &= 1_{(T_x \leq 1)}.
 \end{aligned}$$

Claim

$$\Phi = 1_{(T_x > 1)}.$$

Example $r = 0.1, p_x = 0.9, q_x = 0.1, \pi^r = 0.1$. Now put up two equations with two unknowns

$$1 * x + (\pi - 0.1x) 1.1 = 0$$

$$0 * x + (\pi - 0.1x) 1.1 = 1$$

$$x = -1$$

$$-1(1 - 0.1 * 1.1) + \pi 1.1 = 0 \Rightarrow \pi = \frac{1 - 0.1 * 1.1}{1.1} = 0.8090909$$

$$\begin{aligned} \pi &= \frac{1}{1+r} E^* [1_{(T_x > 1)}] \\ &= \frac{1}{1.1} \left(\frac{0.1 * 1.1 - 0}{1 - 0} * 0 + \frac{1 - 0.1 * 1.1}{1 - 0} * 1 \right) \\ &= \frac{1 - 0.1 * 1.1}{1.1} = 0.8090909 \end{aligned}$$

and not

$$\pi = \frac{1}{1.1} p_x = \frac{0.9}{1.1} = 0.8181818$$