Insurance Mathematics
A European Model

Horst Behncke
University of Osnabrück

May 2000

Contents

Why Study Insurance Mathematics?

Why should a student take up the study of insurance mathematics? This is a problem many students face when they encounter such a course among the choice of courses offered. There are in fact a number of good reasons why taking such a course or maybe even majoring in this area is a wise decision.

1. Insurance determines a good deal of our lives. An average household will in general have four or more insurance policies. The most common policies are car liability insurance, which is mandatory, fire insurance, which is also required by law for house owners, life and health insurances. But of course there is a host of other insurance types covering all aspects and disasters that may befall you. Thus a basic knowledge of the insurance process is a must for any well educated person, especially if this person has a good understanding of mathematics.

2. Since the insurance industry permeates our lives so much, it forms an important sector of our economy. The economic influence of the insurance sector is enhanced by the fact that these companies invest a great deal of their assets in other businesses, which gives them considerable economic influence and leverage. Here is a short list of premiums and assets held by insurance companies:

Germany (from the Statistical Yearbook)	1993	1994	1995	1996
Life insurance (premiums in billions of DM)	76.6	84.0	89.3	93.9
Health insurance (premiums)	25.7	28.0	32.1	34.3
Accident, Car liability	92.2	98.8	102.3	102.5
Assets	1993	1994	1995	1996
Life insurance	582.8	641.3	703.8	768.0
Total of all insurances	914.4	996.2	1,116.7	1,221.2

USA	1992	1993	1994	1995
Life ins. (premiums in billions of $)	429.9	466.4	480.8	512.2
Assets	1665	1839	1942	2144

Correspondingly, insurance companies employ a large number of staff from all walks of life. Of these, mathematicians, computer specialists and operations research specialists play an important role. In particular, insurance companies are currently looking for qualified mathematicians. The importance of life and health insurance will most certainly increase in the future due to the growing group of dependent senior citizens in all western countries, whose government insurance schemes are usually inadequate. This increase is enhanced by the fact that older people incur higher costs for nursing and health care. Thus there is a large market for life and health insurance.

3. Finally, insurance mathematics is a useful application of stochastic analysis and statistics. Although the mathematics is only demanding for a few instances, the majority of the work involves basic knowledge. A good understanding and impression can therefore be obtained from just a few courses. It should be added, however, that a basic knowledge of business, finance and law will also be required.

This course also aims to cover the European aspect of the insurance industry. This is necessary because different regulations, tax and social systems have led to different types of insurance systems in the different states. On the other hand, since 1994 insurance companies may operate freely within the European Union. Both this liberty and deregulation makes it necessary for an insurance mathematician to understand the insurance environment in other EU countries.

As stated above, this course will require some elementary probability theory and statistics. Since there are a number of good books in this field, we will not present an introduction or summary of probability and statistics theory here. Here is a short list which you may consult:

K.L. Chung	A Course in Probability Theory, Academic Press
J.E. Freund	Modern Elementary Statistics, Prentice Hall
W. Mendenhall,	Mathematical Statistics with Applications
D. Wackerly, R. Scheaffer	Boston PWS Kent Publishing 1990
D.S. Moore	The Basic Practice of Statistics, Freeman and Co. 1995

Before we commence with the development of insurance mathematics per se, a few remarks concerning the actuary profession should be made.

The Actuary

A professional insurance mathematician is called an actuary. An actuary will thus employ statistical and probabilistic techniques to problems arising with insurances or pensions. The activities are listed below.

The tasks of an actuary

1. Rate making.
2. Computing reserves and surplus.
3. Surveying the insurance market and determination of demand.
4. Developing new insurance products. In more detail this involves
   (a) fixing the parameters
   (b) exemplary computations

   (c) comparison with competing products
   (d) profit testing
   (e) model office calculations
   (f) developing test and calculation software
   (g) developing application forms and marketing material
   (h) formulation of conditions of contract

5. Setting up the balance and accounting plan.

6. Assessment of investments, bonds and securities.

7. Systematic management of assets, liabilities and risks.

Who employs actuaries?

In general, an actuary can be described as a *risk engineer*. Because of their wide range of tasks, actuaries are not only employed by insurance companies, but also by governments, pension fund organizations, banks and other financial institutions. In recent years a growing number of mathematicians have been employed by such instutions for actuarial work, and the demand is still growing. Because of the importance of such work, a considerable number of mathematians are now working as managing directors in such firms.

How can you become an actuary?

As yet no formal training exists for the actuary profession outside Great Britain. The rule was rather that mathematicians with an interest in financial business matters worked for insurance companies for some years and then passed a test administered by a professional organization of actuaries. This could apply also to business majors. Formally this may be described as the *Study period*

1. Study of mathematics (major or minor)

2. Computer science

3. Some study of business and accounting (minor or major)

4. Some knowledge of law.

All this has to be taken with a pinch of salt, however. Some of the gaps in your knowledge of the field may be filled while training on the job. But what is required most is the willingness to learn and to show interest in these matters, and of course the specific demand, e.g. the required knowledge of tax law might well depend on the section of the company you are working in.

Studying at a university is merely the first step to becoming an actuary. Parallel to working in a company, actuaries-to-be will also have to learn the basics of

1. life insurance
2. non-life insurance
3. pensions
4. health insurance
5. mathematics of finance
6. specialist knowledge in various areas of insurance, business and finance.

Candidates will be examined in these fields by the organization of actuaries in the relative country.

Besides these qualifications you should be able and willing to work in a team and you need be communicative, because new products are usually developed in a team consisting of a mathematician or actuary, a business major, marketing analysts and a legal specialist. Last but not least, international experience and fluency in at least one foreign language, usually English, will be required. If all these requirements sound overwhelming, do not despair because you would be given time to gain the knowledge and many of these qualifications are also required in other areas.

The future outlook

In recent years insurances have represented a fairly stable part in the economies of all EU countries. This is not only due to their long-range planning and their large assets, but also to the ever growing demand for insurances. This demand for life and health insurance, pension plans and other forms of life insurance will grow, because in all European countries government social insurances have exceeded their limits and capabilities. Furthermore, more people are living longer and fewer people belong to the work force. This is valid for most European countries, as can be inferred from the plot of the pensioners quota. In addition, increasing private wealth also motivates people to use the instrument of life insurance or the stock market to save money and taxes.

It is not surprising therefore that an increasing number of governments, even social democratic ones, e.g. the labour government in GB or the SPD government in Germany, stress the need for private provision for old age.

Summary: The inadequacy of government health and pension schemes causes increasingly more people to buy private insurance in these domains. Because of the relative increase of people over 60 and because of the large assets of this group, life insurance and pension schemes will grow in the future.

Thus life and health insurances will also grow. The liberalization, deregulation and increasing competition on this market will lead to a multitude of new products. All this leads to a higher demand for insurance mathematicians.

Problems:

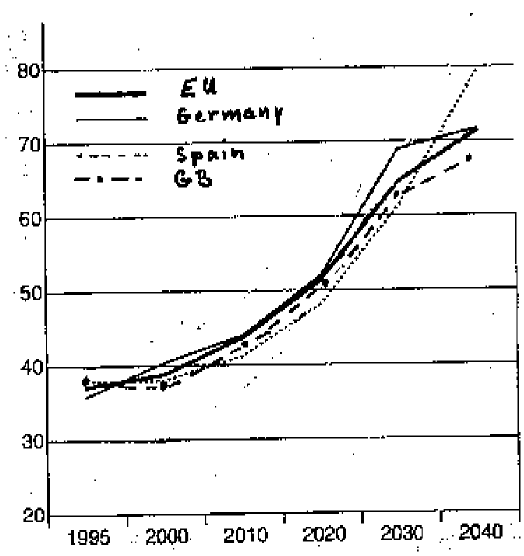
1. Which types of insurance does your family have? What protection do they offer and how much do they cost?

2. Compare two pure risk life insurances for persons of your age until the age of 60.

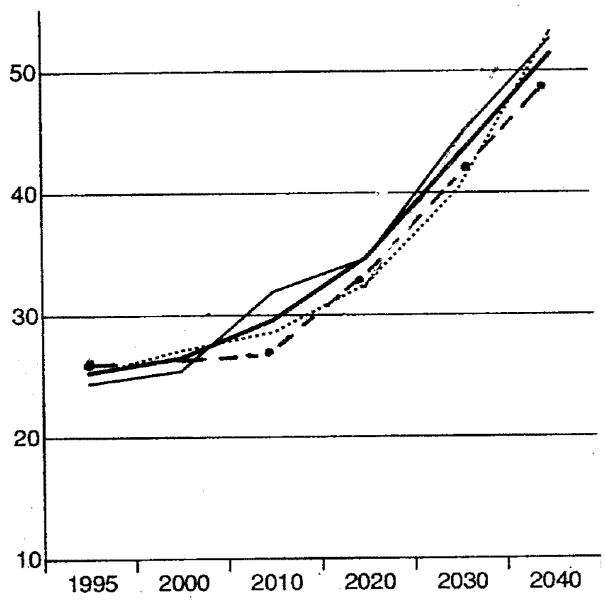
3. Consult a consumer report on life insurances. What are the pertinent points?

4. How would you go about determining the evaluation of the pensioners quota for your country? What information can you obtain from the statistical year book of your country?

5. What is the average pension of a pensioner who has worked for 40 years of his/her life? How does this compare with the cost of living in a typical old-people's home? If you are unable to obtain such data, ask in your neighborhood.



Pensioners quotient: Ratio of persons over 60 per 100 persons of age 20 to 60



Pensioners quotient: Ration of persons over 65 per 100 persons of age 20 to 65

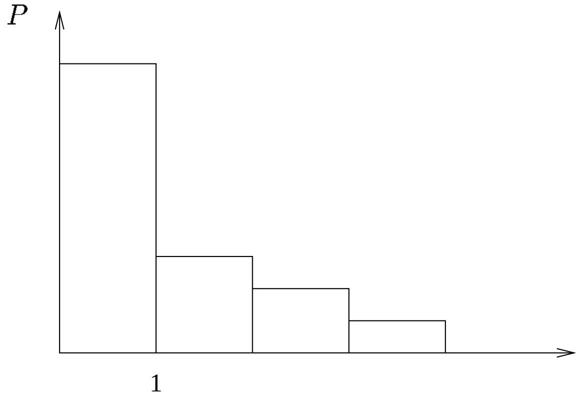
1. General Concepts in Insurance

1.1 The basic principles of insurance

Insurance is a means to protect people or groups from losses due to unforseeable events like death, accidents, sickness or liability. One crucial aspect is that these losses are accidental and that the damage can be measured in monetary terms. In this case the insurer or insurance company will compensate the insured if damage has occurred. For this the insurer demands a premium for underwriting the risk. The insured person thus exchanges the uncertain risk or loss X against the certain loss of the premium Pr. Thus we describe the loss by a random variable X, because the risks are accidental. In this context we define risk to be the uncertainty regarding losses. Thus we will also speak of the risk X. The distribution of X measures the size and probability of losses. In many cases, when losses are rounded to the next currency unit, the distribution of X is entirely determined by the list of probabilities p_x with which a loss of $X = x$ occurs.

In these notes we will not use any particular currency unit, and you may insert the pound £, franc F, peseta Pts, mark DM, Euro or any other unit you please. The numbers chosen, however, are oriented towards the Euro in order to be realistic.

The graphical representation of the p_x and the distribution of X can be an important tool in understanding the loss X.



In many cases one may also consider X as a real variable and represent $p_x = p(x)$ as a function. This is particularly useful if the function $p(x)$ can be expressed in simple terms. This, however, is rarely the case in realistic situations.

One of the central problems in the insurance industry and mathematics is to determine the list of (p_x) or the function $p(x)$. This requires some statistics and numerical mathematics. Later we will describe this process in detail for life insurances, but for the moment the following simple example for a fire insurance should suffice.

Assume you intend to enter the fire insurance industry and sell policies for family homes. For simplicity we also assume that all these homes cost the same amount H. It is known that in previous years for n insured houses n_x houses $0 \leq x \leq H$ experienced damage to x caused by fire or storm. Then common sense tells us that $\hat{p}_x = \frac{n_x}{n}$, the relative frequency of loss x, is a good estimate for p_x. This of course is also borne out by statistics. Since these estimated numbers $\hat{p}_x$ still show too much fluctuation, numerical methods are used to smoothen the data, and obtain a list of (p_x) on which premium calculations can be based. If the losses are to be described by a continuous variable, the n_x/n could also be fitted by a suitable parametrized family of functions.

For the moment we will thus assume that the distribution of X, the list of (p_x), is known. Then

$$P(x_1 \leq X \leq x_2) = \sum_{x=x_1}^{x_2} p_x \tag{1.1.1}$$

measures the probability that the damage lies between x_1 and x_2. The expectation or expected value $E(X)$ of X is then

$$E(X) = \sum_{x=0}^{L} x \cdot p_x. \tag{1.1.2}$$

Here L is the largest observed damage. $E(X)$ represents the average damage that is observed in the long run. One should note, however, that there is a distinction between sample mean, which can be observed in concrete situations and $E(X)$. For very large samples both are almost the same, however. Later we will also need the *variance* σ_X^2 of X

$$\sigma_X^2 = \sum_{x=0}^{L} (x - E(X))^2 p_x. \tag{1.1.3}$$

σ, the *standard deviation* of X, is a measure of the dispersion of the distribution of X around its centre $E(X)$.

If the distribution of X is determined by the density function $p(x)$, one similarly has

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} p(x)dx.$$

$$E(X) = \int_0^{\infty} xp(x)dx \qquad \text{and } \sigma_x^2 = \int_0^{\infty} (x - EX)^2 p(x)dx.$$

Example 1.1 For a particular class of cars it has been observed that only 10% are involved in accidents per year. Of the cars involved in accidents 80% have minor damage and only need repairs costing 2,000, while 10% of the cars involved in accidents lead to costs of 5,000 and the remainder incur costs of 10,000.

In this case the distribution is given by $p_0 = .9$, $p_{2,000} = .1 \cdot .8 = .08$, $p_{5,000} = .01$, and $p_{10,000} = .01$. Thus

$$\begin{aligned} E(X) &= .9 \cdot 0 + .08 \cdot 2,000 + .01 \cdot 5,000 + .01 \cdot 10,000 \\ &= 160 + 50 + 100 = 310 \end{aligned}$$

and

$$\sigma_X^2 = .9 \cdot (0 - 310)^2 + .08 \cdot (2,000 - 310)^2 + 0.01 \cdot (5,000 - 310)^2$$
$$.01 \cdot (10,000 - 310)^2 = 1,473,900.$$

One can show that

$$\sigma_X^2 = E(X^2) - E(X)^2 = \sum p_x x^2 - E(X)^2. \tag{1.1.4}$$

With these definitions we can now begin to answer some of the basic questions in insurance mathematics.

What premium should an insurance company demand for underwriting a risk X?

We begin by studying this problem for a very simple type of insurance – one with which the damage only has one value.

Fire Insurance

Consider fire insurance for detached family homes where the damage D is fixed, e.g. the value of the house if the house burns down. Thus $p_D = p$ is the probability of a fire and $1 - p = p_0$ is the probability that no fire occurs. To make things explicit assume $p = .001$ and $D = 100,000$. Then

$$E(X) = p \cdot D = 100$$

and

$$\sigma = D(p(1-p))^{\frac{1}{2}} = D \cdot \sqrt{.001 \cdot .999} = 3160.70,$$

because the distribution of the random variable X is essentially a binomial distribution.

Assume the insurance company has underwritten n identical risks $X_1, X_2, \ldots, X_n$ of this type. Usually n is fairly large, about 10,000 or even more. Then the insurance company will have to pay

$$Y = X_1 + X_2 + \ldots + X_n.$$

If these risks are all unrelated one may assume that the X_i are independent. This happens, for example, if the corresponding houses are far apart. If all the X_i are independent, the strong law of large numbers tells us that

$$\overline{X} = \frac{1}{n}(X_1 + \ldots + X_n) \to E(X), \tag{1.1.5}$$

i.e. for larger and larger n, $\overline{X}$, the average damage of all policies converges to $E(X)$. This is of course plausible when you consider how p_D has been determined. Remember that p_D is approximately the relative frequency with which a house burns down.

Altogether this implies that the company will have to pay $E(X)$ per policy on average. This of course holds under much more general conditions, not only for binomially distributed risks. In fact, let us assume an insurance company has insured identical risks (policies) $X_1, X_2, \dots, X_n$ with bounded variances σ^2. If these are independent

$$\frac{1}{n}(X_1 + \ldots + X_n) \to E(X), \tag{1.1.6}$$

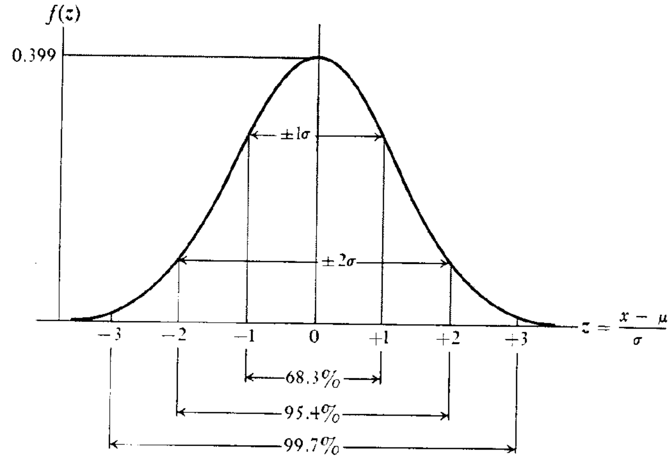
i.e. the insurance company will have to pay $E(X)$ on average. This of course requires that n is fairly large. For this reason $E(X)$ is sometimes called the *net premium* and the equivalence principle for insurances is formulated as

Expected losses = net premium.

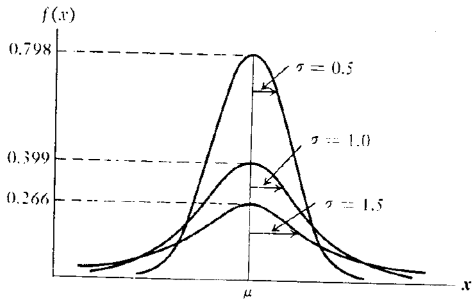
The strong law of large numbers is not good enough to give us an idea of how large n should be and what "the company pays $E(X)$ on average" means. Again we shall study this first with our fictitious example of the binomial fire insurance. To do this we need the *central limit theorem.*

Theorem 1.1 Let $X_1, X_2, \dots, X_n$ be n identically distributed independent random variables with expectated values $\mu = E(X)$ and variance σ^2. Then the distribution of the arithmetic mean $\overline{X} = \frac{1}{n}(X_1 + \ldots + X_n)$ is approximately given by a normal distribution $N(\mu, \frac{\sigma}{\sqrt{n}})$ with mean value μ and variance $\frac{\sigma^2}{n}$.
The normal distribution with mean μ and standard deviation σ has as its density the well-known bell shaped curve

$$f_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$



Standardized normal distrubution: $f(z) = (1/\sqrt{2\pi})e^{-(1/2)z^2}$



Three normal distributions with different standard deviations

which is depicted, for example, on the 10 DM note in Germany. But plots of it can also be found in almost any textbook on statistics.
Thus the probability that an $N(\mu, \sigma)$ distributed random variable has values between a and b is given by

$$P(a \leq X \leq b) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx. \tag{1.1.7}$$

The computation of these numbers can be reduced to the standard normal distribution $N(0, 1)$ and its distribution function ϕ

$$\phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx \tag{1.1.8}$$

by a simple linear transformation

$$P(a \leq X \leq b) = \phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right). \tag{1.1.9}$$

The indefinite integral for the normal distribution cannot be evaluated in closed form. However, because of its importance, ϕ is tabulated. Since the standard bell curve $f_{0,1}$ is symmetric around 0, one has

$$\phi(-x) = 1 - \phi(x) \qquad x \geq 0. \tag{1.1.10}$$

Thus ϕ is usually only tabulated for values $x \geq 0$. Let us now apply this to our fictitious fire insurance.

Example 1.2 $D = 100,000$, $p_{100,000} = p = .001$, $p_0 = .999$, $E(X) = 100$, $\sigma_X = 3160.70$. If the company has $n = 1,000$ $(10,000;\ 100,000)$ such policies the company will pay on average $n \cdot E(X)$ i.e. 100,000 (1,000,000, 10,000,000). But in 25% of all years the company may have to pay $.675 \cdot \sqrt{1,000} \cdot 3,160.70 = 67,466.32$ $(0.675 \cdot \sqrt{10,000} \cdot 3,160.70 = 213,347.25,\ .675 \cdot \sqrt{100,000}3,160.70 = 674,663.20)$

more due to fluctuations, because $1-\phi(.675)=.25$. For $n=1,000$ the fluctuations exceed two thirds of the expectation value in 25% of all years. For $n=100,000$ this excess is reduced to 2.5%. In 10% of all years the fluctuations might even exceed $1.282\cdot 3,160.70\cdot\sqrt{1,000}=128,136.04\quad(405,201.74,\ 1.281.360.41)$.

More generally, assume an insurance company has insured n identical independent risks $X_1,\ldots,X_n$. Then the damage paid by the company amounts to

$$Y=X_1+\ldots+X_n=nE(X)+\text{Fluctuations}. \tag{1.1.11}$$

If n is large and the $X_1,X_2,\ldots,X_n$ are independent, the fluctuations will exceed

$$\text{Fluctuations}\geq a\cdot\sqrt{n}\sigma \tag{1.1.12}$$

with a probability of approximately $1-\phi(a)$. Here and in the remainder, σ_X^2 or simply σ_X^2 will stand for the variance of the random variable X. Here are some typical values.

Excess prob. $1-\phi(a)$	.25	.1	.05	.01	.005	.001
Excess value a	.67449	1.2816	1.6449	2.3263	2.5758	3.0902

Thus an insurance company will have to be able to cover the risk of fluctuations. Its advantage, however, lies in the fact that the relative fluctuations get increasingly smaller as n gets larger. In addition, there are costs of administration and acquisition to cover, because the company will have to pay the insurance agent who sells you your policy, and the people who keep track of your records also want a salary. This also applies to the mathematician who computed your premium. Finally, the company wants to make a profit. A typical insurance premium can thus be broken down as follows:

$$\textit{Premium}=\textit{Expectation of Damage}+\textit{Fluctuation}+\textit{Administrative Term}+\textit{Profit}. \tag{1.1.13}$$

The administrative term is usually taken proportional to $E(X)$, mostly 5 to 10 per cent of $E(X)$. This applies likewise to the profit. Thus we shall usually take both together. From what we have seen above, the fluctuation term should be taken proportional to σ. If this is the case, one speaks of the *standard deviation principle* for the premium. In some cases it is also taken proportional to σ^2. This would be called the *variance principle* for the premium. In many cases, however, it is taken as simply proportional to $E(X)$. This would be called the *expectation value principle.* Unfortunately, this is excessive terminology, but used nonetheless.

To summarize: The premium Pr for a risk X will have the form

$$Pr=E(X)+Fl(X)+\lambda_1E(X), \tag{1.1.14}$$

and the fluctuation term can be taken as

$Fl=\lambda_2E(X)$	expectation value principle
$\lambda_2\sigma$	standard deviation principle
$\lambda_2\sigma^2$	variance principle.

In the administrative term $\lambda_1E(X)$, λ_1 usually lies between .06 and .12.

Here we have included a profit of 1 to 2 per cent of $E(X)$ with the administration cost. As yet we have not discussed how high the compensation for the fluctuation should be for the company. In fact, this is a rather complicated business. So we will only touch upon it here and take it up later in more detail.

If an insurance company has insured n identical independent policies $X_1, \ldots, X_n$ it could in principle happen that it would have to pay the maximal amount D for each insurance policy. In the example above this would be 1,000 $\times$ 100,000 or 100 million. If this were to happen, the insurance company would certainly go bankrupt. Fortunately enough, this happens extremely rarely. Thus, even though extreme losses cannot be avoided completely, a reasonable way to compute the fluctuation compensation term is to define a priori an excess probability p_e and fix $Fl(X)$ such that payments in excess of the premium income occur only with a probability of less than p_e. We will again consider our fire insurance example from above.

Example 1.3 $n = 1,000$, $p = .001$, $D = 100,000$, $E(X) = 100$, $\sigma_X = 3,160.70$. Assume the premium income of the insurance amounts to $1,000E(X)+1,000Fl(X)$, because the administrative term is already used up. With the above table we see that the chance that the total claims exceed $1,000 \cdot E(X) + 1,000\frac{\sigma_x}{\sqrt{1,000}} \cdot 2.3263 = 332,513.94$ and 408,865.83 are just $p_e = .01$ and $p_e = .001$, respectively.

More generally, this shows that the fluctuation term should be chosen as

$$Fl(X) = \frac{\sigma}{\sqrt{n}} v_e \tag{1.1.15}$$

when the insurance has n policies and where the excess value v_e can be taken from the table $\phi(v_e) = 1 - p_e$. Thus an insurance company could make the claims excess arbitrarily small if only $Fl(X)$ was chosen large enough. This of course cannot be done in general, because of competition with other companies or because of government regulations. The situation is even more complicated because the company may possess capital and reserves. We will consider again a simple example for this.

Example 1.4 Assume again that the loss X is binomial distributed with $D = 100,000$ but $p = .005$, assume the company has built up reserves in the last 10 years of 50% of $E(X)$ each year, and assume the premium Pr is given by

$$Pr = E(X) + .3 \cdot E(X) + .1E(X)$$

where the middle term is intended to cover the fluctuations. What is the chance that the claims exceed the premium income plus the reserves if the number of policies n is 1,000 (10,000)?

We have

$$\begin{aligned} E(X) &= .005 \cdot 100,000 = 500 \\ \sigma_X &= 100,000 \cdot (.005 \cdot .995)^{\frac{1}{2}} = 7,053.37. \end{aligned}$$

The total premium income is $500,000 \cdot 1.3$ $(5,000,000 \cdot 1.3)$, because the administrative part of the premium should not be counted. The terms exceeding the expectation values are thus $.3 \cdot 500,000$ $(.3 \cdot 5,000,000)$ and the reserves are $.5 \cdot 500,000$ $(.5 \cdot 5,000,000)$. In order to compute the probability that the claims exceed the reserves and fluctuation terms of $500,000$ $(.8 \times 5,000,000)$ we have to compute the variance of all claims. This is

$$n \cdot \frac{\sigma}{\sqrt{n}} = \sqrt{n} \cdot \sigma = \sqrt{n}\sqrt{.005 \cdot .995} \cdot 100,000.$$

For $n = 1,000$ $(10,000)$ this amounts to 223,040.8 and 705,336.80, respectively. The chance that the claims demand exceed even the reserves plus the premium income are thus $1 - \phi(\frac{.8 \cdot 500,000}{223,040.8}) = 1 - \phi(1.7934) = 0.0364$ and $(1 - \phi(\frac{.8 \cdot 5,000,000}{705,336.8}) = 1 - \phi(5.67) = 0.000$, respectively.

This example shows two things. Firstly, how important it is to have a sufficiently large number of policies in the portfolio: $n = 1,000$ is not enough in most cases. Secondly, it shows the importance of reserves to overcome the vagaries of bad luck, i.e. when fluctuations become too large.

More generally, assume the company has insured n identical independent risks X_1, $X_2, \ldots, X_n$, assume the company has reserves K and assume the individual premium Pr has the form

$$Pr = E(X) + \lambda_1 E(X) + \lambda_2 E(X),$$

where $\lambda_1 E(X)$ represents the fluctuation term of the premium. Then, apart from the administrative term and profit $\lambda_2 E(X)$, which is already used up, the amount $K + nE(X)\ (1 + \lambda_1)$ could be used to satisfy the claims of the insured. Because of the central limit theorem the claims

$$Y = X_1 + \ldots + X_n$$

are approximately normally distributed with parameters nEX and $n\frac{\sigma}{\sqrt{n}} = \sqrt{n}\sigma$; $Y \sim N(nEX, \sqrt{n}\sigma)$. The probability P that the fluctuations exceed $n\lambda_1 E(X) + K$ is thus approximately given by

$$P = 1 - \phi\left(\frac{n\lambda_1 EX + K}{\sqrt{n}\sigma}\right). \tag{1.1.16}$$

We can summarize these results so far as:

Insurance is a means to turn uncertainty into a reasonable amount of certainty. Formally, it is thus a contract that transfers risk from the insured to the insurance company, which in turn obtains recompense in form of the premium. The premium is defined by

$$\begin{aligned} &\text{Premium} = \text{Expected Loss} + \text{Fluctuation Term} + \text{ Admin. Term} + \text{Profit} \quad (1.1.17)\\ &Pr(X) = E(X) + Fl(X) + \lambda E(X). \end{aligned}$$

From the point of view of the insurer this will work best if

1. the policies are determined for similar risks
2. the risks are independent

3. the number of insurance policies is large.

Let us interpret these statements.

1. This means that the policies should cover only similar and related risks. Thus it is rather difficult to compare a life insurance with a fire insurance. But even fire insurances for family homes and office buildings are quite distinct, because the damages and probabilities are rather different. Thus the distributions of the various claims $X_1, \ldots, X_n$ should be reasonably similar.

2. The independence assumption is quite critical for the insurer, because without it the fluctuations in claims may be quite large. Think, for example, of a hailstorm which damages *all* cars in a town. Similarly consider the effect of an earthquake or a flood on buildings. If all these buildings are insured with the same company, it might well go bankrupt after such a disaster.

3. In the example above we involved the central limit theorem to get an idea of the loss which awaits the insurance company. For independent policies of approximately the same size the fluctuations behave like $\frac{\sigma}{\sqrt{n}}$. More generally, this can be interpreted as: The larger the number of independent policies, the smaller the relative fluctuations. These decrease roughly proportional to the inverse of the root of the number of policies.

The previous considerations have also shown that high loss, but low probability events have fairly large standard deviation, which may exeed the expectation value by far. Finally, we note that building up reserves can be a means to overcome the negative effects of fluctuations.

As a final example we want to compute the premium and short-term security for a company which has a portfolio of five different classes of binomially distributed risks. These are

	p	Claim	Number of policies	$E(X)$	$\sigma^2 = p(1-p)\text{Claim}^2$
1	.005	$100,000$	500	500	$4.975 \cdot 10^7$
2	.01	$200,000$	700	$2,000$	$3.96 \cdot 10^8$
3	.01	$50,000$	1000	500	$2.475 \cdot 10^7$
4	.015	$100,000$	400	$1,500$	$1.4775 \cdot 10^8$
5	.02	$120,000$	600	$2,400$	$2.8224 \cdot 10^8$

We will consider two different forms of premiums

$$\begin{aligned} Pr_1 &= EX + \lambda_1 E(X) + .1E(X) \\ Pr_2 &= EX + \lambda_2 \sigma(X) + .1E(X). \end{aligned}$$

In both cases $.1E(X)$ represents the administration term and profit. This part does not concern us here. Rather our interest lies with the fluctuation term, which has to be chosen such that the chance of claims exceeding the premium income is less than .001. We will also consider the case when the reserves are half of the expected yearly claims. Now let $Y = \sum_{i=1}^{5} \sum_{j=1}^{n_i} X_{ij}$ be the sum of all claims. Then

$$E(Y) = \sum n_i E(X_i) = 500 \cdot 500 + \cdots + 600 \cdot 2400 = 4.19 \cdot 10^6$$
$$\sigma^2 = \sigma^2(Y) = 500 \cdot 4.975 \cdot 10^7 + 700 \cdot 3.96 \cdot 10^8 + \cdots = 5.55269 \cdot 10^{11}$$

and, as argued before, we may assume that Y approximately has a normal distribution. Now the security level for the standard normal distribution for .001 is 3.0902. Thus the total fluctuation term of the premium must amount to $3.0902 \cdot \sigma = 2.3027 \cdot 10^6$. This, however, equals $\lambda_1 E(X)$. Thus

$$\lambda_1 = \frac{3.0902 \cdot \sigma}{4.19 \cdot 10^6} = .549571598.$$

If the reserves are taken into account, $\lambda_1 = .04957\ldots$ suffices for this security level. In reality insurance companies will take an intermediate value for λ_1, say .2 or .25. But then they also have to take into account that they want to stay in business for a long time.

Above we have shown that the fluctuation term is $3.0902 \cdot \sigma(Y)$. By construction we have

$$\begin{aligned} 3.0902 \cdot \sigma &= \lambda_2 \cdot [\sqrt{4.975 \cdot 10^7} + 700 \cdot \sqrt{3.96 \cdot 10^8} + \ldots] \\ 2.3027 \cdot 10^6 &= \lambda_2 \cdot 3.737354361 \cdot 10^7 \qquad \text{or} \\ \lambda_2 &= 0.06161. \end{aligned}$$

Summary: The fundamental principle for premiums is

$$\begin{aligned} \text{Premium} &= \text{Expected Losses} + \text{Fluctuation Term} + \text{Administrative Term} \\ Pr(X) &= E(X) + Fl(X) + \lambda E(X). \end{aligned}$$

For insurances it is important to have a large number of independent claims. Then the central limit theorem is applicable.

Problems

1. Why does one equate the net premium and $E(X)$?
2. In what sense is σ_X^2 the natural candidate for the fluctuation term?
3. Why is the independence assumption for the insured risks so important? Can you think of other cases in which it is violated?
4. Explain the role of the reserves.
5. What is the distribution of the risk for the company in examples 1.3 and 1.4?
6. Prove formula (1.1.4).

7. For an accident insurance the following values for p_x have been found.
   $p_0 = .8$, $p_{1{,}000} = .01$, $p_{2{,}000} = .01, \ldots, p_{20{,}000} = .01$. Determine $E(X)$ and σ_X^2.

8. A certain risk is found to be exponentially distributed $p(x) = \lambda e^{-\lambda x}$. Determine $E(X)$ and σ_X^2.

9. What would the net premium be in the case of example 3 if the insurance only has to pay the damage which exceeds $a\lambda$, i.e. the claims Y are

$$Y = \begin{cases} 0 & \text{if} \quad X \leq a\lambda \\ X - a\lambda & \text{if} \quad X \geq a\lambda \end{cases}$$

10. Do the same for problem 2 if the excess limit is $1,000$.

11. A company has insured $n = 10,000(5,000,\ 2,000)$ objects. For each individual policy the damage can be 50,000 with probability .001 and 100,000 with probability .001. Compute $E(X)$ and $\sigma(X)$. How large must the fluctuations term of the premium be if the company has no reserves and wants the chance of claims exceeding income to be less than .001?

12. A company has insured n claims with the same loss profile as in problem 7. Its premium is given by $Pr = E(X) + .4 \cdot E(X) + 0.1E(X)$, with $.4E(X)$ being the fluctuation term. How large must n be so that the probability of claims exceeding revenues is less than .001? How large must n be if the reserves are $n \cdot .5E(X)$?

13. A company has insured n_1 risks with a binomial distribution $p_{50{,}000} = .001$ and n_2 policies with binomial distribution $p_{100{,}000} = .001$. Determine the expectation value and variance for the whole portfolio. Assume the premiums are determined by the rule

$$Pr(X) = E(X) + .4E(X) + .1E(X)$$

   and assume $n_1 = 2,000$ and $n_2 = 3,000$. What is the probability that the claims exceed the premium income of the company?

   (Hint: If X and Y are independent random variables $\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y)$)

14. (a) The owner of a house is considering to buy a property insurance which covers all losses up to 10,000. Assume $p_0 = .9$ and $p_{1{,}000} = p_{2{,}000} = \ldots = p_{10{,}000} = \frac{1}{100}$. Compute $E(X)$ and $\sigma(X)$. How do these values change if the owner takes a 1,000 deductible, i.e. the claim against the insurance is $\max(0, X - 1,000)$?

    (b) Assume the premium is calculated by the formula

$$Pr(X) = E(X) + .04E(X) + 200$$

    and assume that the installation of an alarm system reduces the $p_{1{,}000}$, $\ldots, p_{10{,}000}$ to half of their value. Should the owner buy the alarm system for 5,000 if it has a lifespan of 8 years?

15. Considering the last example, show that it is sensible to use the variance principle to define the premium. Now compute λ_2 in this case.

16. Derive formula (1.1.9).

1.2 Risk, a personal perspective

Risk is usually defined as adverse chance or the chance of something bad happening. But it is mainly the personal perception of risk which makes people take out insurance. This view of risk depends on

1. personal feelings regarding the risk, and gender.

2. the state of society and the person's social and cultural environment.

3. the size, form and threat of expected damage.

4. the probability with which the damage occurs.

Let us consider each of these aspects separately.

1. The psychology of risk and risk perception has been studied at great length by psychologists. Obviously, the perception of risk is much greater if you have already experienced or witnessed such a disaster. A fire that destroys your house would not only destroy 200,000 worth of property, it would also destroy the home of your family. Your illness or that of your relatives causes personal feelings of grief, helplessness and compassison. Thus the feeling of responsibility for the family causes many to take out health, accident and life insurance. An additional factor which plays a great role in risk perception is gender, i.e. women and men often perceive risks differently [P.E. Gustafson; Risk Analysis **18** 1998, 805].

2. The perception of death is quite different for married people with children than for people without such responsibilites. The perception of a risk is more vivid if it has been covered extensively by the media. In this context consider for example AIDS, cancer or nuclear reactors. In general, people only have rather imperfect and rudimentary knowledge and perception of rare hazards like floods, earthquakes or storms. This has been studied extensively in surveys. Very few people in the western world talk about malaria, even though its toll is greater than that of AIDS. But malaria is restricted almost entirely to the tropics, as is AIDS to a large extent. For this reason AIDS is hardly covered in the media any more. The assessment of damage or danger depends largely on factors like social boundary conditions such as education, position and wealth, understanding of the processes involved, the general framework in which the risk process develops and the press coverage of recent related events. Thus the perception of risk is fluid and varies along with many of these factors. In general, people judge events which lead to the death of many people to be

more dangerous than those which take a steady toll. This applies, for example, to airline accidents versus road accidents. Thus people might take out travel insurance before boarding a plane. But would you take out insurance against lightning? The attitude towards and perception of risk might even vary from nation to nation. People in France, for example, think quite differently about nuclear power than people in Germany. This is no surprise, because 70% of France's electricity is nuclear based, while Germany is planning to phase out nuclear power. More generally, the cultural background, beyond sociodemographic factors, influences the way we perceive risk [J. Brenot, S. Bonnefous, C. Harris; Risk Analysis **18** 1998, 729].

3. Single events which threaten your life or that of your family are considered usually more dangerous than a steady and permanent toll on lives and health. Thus people take out accident insurances, but continue smoking. From the above it is clear that many factors apart from purely monetary considerations induce people to look for insurance coverage. But even if we restrict our study solely to monetary consequences, it should be clear that our understanding of the value of money is not covered by the nominal amount. In mathematics this is usually described by the term *utility*. We will discuss this point shortly.

1.3 The Perception of Probability

Even if we considered only the financial aspects of a risk, its mathematical equivalent X is in general not an adequate description of our personal assessment, because our understanding of chance and probability is not sufficiently well developed to grasp the details of the distribution of X. It is known that people's understanding does not in general allow more than nine different levels of distinction.

This misjudgement of probabilities holds in particular for rare or very rare occurrences. Consider this example: For which of the following four cases would you take out insurance?

1. A loss of 100 with probability .01
2. A loss of 1,000 with probability .001
3. A loss of 10,000 with probability .0001
4. A loss of 100,000 with probability .00001

It is obvious that most people would take out insurance for cases 3. and 4. alone. Why? For one, this certainly has to do with our limited ability to judge probabilities. The second factor here is the consequences of the loss. A loss of 100 or even 1,000 can be overcome somehow. One just has to live a little more frugally, maybe by missing out on this year's cruise! A loss of 10,000 however is much more serious. It might mean delaying the purchase of the long-awaited car or discontinuing studying mathematics or economics because of a lack of funds. For most people the loss of 100,000 would be a disaster which would change their lives considerably, thus a clear case of looking for insurance protection. In all cases the net premium is of course $1 = E(X)$, but the standard deviations differ considerably. They are given by

$$\text{Loss} \cdot (p(1-p))^{1/2}.$$

In our case this is $100\sqrt{.01 \cdot .99} = \sqrt{99} = 9.95$, $1,000\sqrt{.001 \cdot .999} = \sqrt{999}$ or $\sqrt{9,999}$ and $\sqrt{99,999}$. But this of course is not what one perceives. This will in general amount to something like: 4. is riskier than 3., etc. with maybe some thoughts given to the consequences for one's own life. In section 1 we restricted insurance considerations solely to the monetary consequences of losses or damage. This was for a good reason, because non-monetary arguments are almost impossible to describe and grasp in mathematical terms. Even the consequences of the death of a person are quite difficult to evaluate in monetary terms. But one should realize that the decision to look for insurance might be determined by much more than purely mathematical considerations. This can also be seen clearly when considering advertisments for insurances (see the example: "Life insurance for your family" later in this section).

Nevertheless, we shall restrict ourselves essentially to the money aspect and describe risk in terms of the monetary consequences. But even in this simplified situation, the nominal value of money does not adequately describe its importance or usefulness. This discrepancy is usually described in the framework of *utility theory*, which was developed as an aid in decision-making. In this context we deal with the decision to buy insurance coverage.

1.4 Utility

Hence we have to talk about the *utility* of money. If x is the nominal value of money the utility function $u(x)$ measures the usefulness of this amount. It is clear that this is a very personal description. Thus the utility function varies from person to person, if not even from situation to situation. For positive values of x the function u approximately has the form presented in Figure 1.

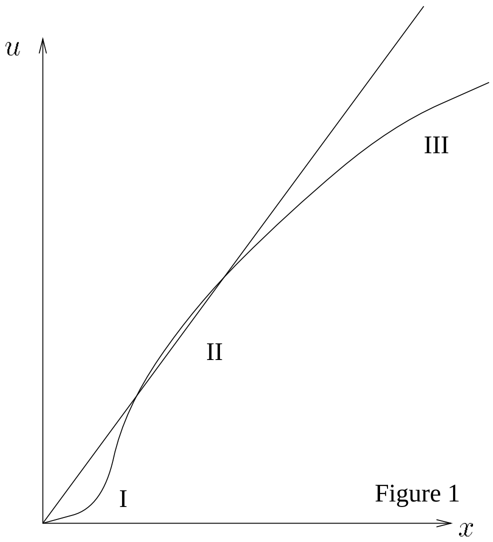


Figure 1

One can distinguish three different domains. In domain I, we attach little importance to money. This is the domain where people bet and gamble. Domain II is the domain involving sums of money used in our everyday lives. It ranges from our daily shopping to our monthly income. Most people have a fairly good and sensible feeling for the value of money in this area. Domain III could be called the area of decreasing marginal utility. Most people, for example, could satisfy all of their wishes with about 1,000,000. But beyond that? At least 2,000,000 would not make them twice as happy!

For losses, i.e. $x \leq 0$, the utility curve will look somewhat like Figure 2. Again there are three domains. In domain I the losses are not considered as serious as the actual monetary consequences. This is again the domain of betting, gambling and letting go. It is probably somewhat larger than the corresponding domains for gains. In II people usually have a quite good feeling for the losses. But if they come unexpected they are usually overvalued a little. This is also true if people are overly conscious of debts.

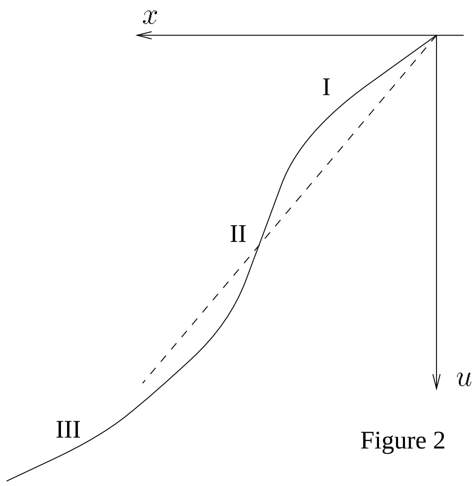


Figure 2

Larger losses in this domain, like the loss of a house, tend to be considered catastrophic, because they would induce major changes in one's life.

Again there is a domain of diminishing marginal utility. The loss of 2,000,000 is not twice as catastrophic as the loss of 1,000,000 – it is simply catastrophic.

As mentioned above, utility functions are highly personal. They depend on the individual's income, wealth, upbringing, social status and many other things.

The utility function may be quite useful in understanding people's behaviour in risky situations. Decision theory, however, has driven this concept to its limits in the form of *expected utility*. This can be summarized as follows:

Let X and Y describe two random economic prospects – X, Y are negative in the case of losses and let w be the wealth of a person and u his/her utility function. Then this person will choose prospect X if

$$E(u(w+X)) > E(u(w+Y)). \tag{1.4.18}$$

Expected utility has some basic flaws. Namely

1. Above we described utility in a rather informal way through its graph. A proper definition of utility requires a very careful determination through a series of financial thought experiments. Even then it is not clear that the values established this way are independent of other outside factors. Most likely the utility function itself will be rather fluid and situation dependent. Even though people are generally rather consistent in their reasoning, establishing a utility function puts far stricter demands on consistency than one can expect in general.

2. We have already argued above that people's perception of probability is not very well developed. But the use of expected utility requires much more. But nonetheless, expected utility holds in a rather vague sense. We will illustrate this with two examples.

Why do people play the lottery?

To take a realistic example, I will deal with the German "6 out of 49" lottery. In this game players have to select 6 numbers out of 49. For definiteness sake assume you win 2,000,000 (20,000, 200, 5) if 6 (5, 4, 3) numbers are selected correctly. Then the expectation value is .794, which is somewhat better than in reality where the expectation value is about .5. If we assume $u(2,000,000) = 1,500,000$, $u(20,000) = 20,000$, $u(200) = 200$ and $u(5) = 3$, the expected utility is .7230.

Compare this with the investment of 1, which is the stake. Hence if the utility of 1 is less than .7, people might be inclined to play the lottery. It should also be realized that many people play the lottery regularly with a stake of 5 or even 20. But the situation is even worse because for most people only the top prize of 2,000,000 counts, and for this prize alone the expected utility is only 0.1073. Hence the utility of 1 is very low indeed.

Life insurance for your family

You might have seen an advertisment for life insurance which states "Protect your family and your loved ones for just pennies a day. Now you are in the prime of your life. You have a good income, a house and family. Have you thought what would happen in the case of an accident? For just 1.50 a day you can buy your family a secure future, because XYZ insurance will pay your family 100,000 in the case of a fatal accident or untimely death."

Even though the language in such advertisments or brochures is certainly rather more ornate, let us get down to the mathematical details.

A daily premium of 1.50 means about 550 per year. The yearly mortality of a 40-year-old male is about $3 \cdot 10^{-3}$. Thus the distribution of loss X is again binomial with $p_{100,000} = 3 \cdot 10^{-3}$ and $E(X) = 300$, $\sigma(X) = 100,000 \cdot (.003 \cdot .997)^{1/2} = 5469$. If you compute everything on a day-to-day basis, you even have $p_{100,000} = \frac{3}{365,000}$ with $E(X) = 0.822$.

How does the expected utility come in here? The trick in the advertisment consists of moving the yearly premium of 550 into the low utility domain of $0 < x < 20$, the domain of gambling and spendthrifts. But now this spending does not have the negative connotation of gambling or waste, but rather that of caring. So using expected utility, you have to weigh the certain loss of 1.50, i.e. the loss of utility of $u(1.5)$ against $\frac{3}{365,000} \cdot u(100,000)$.

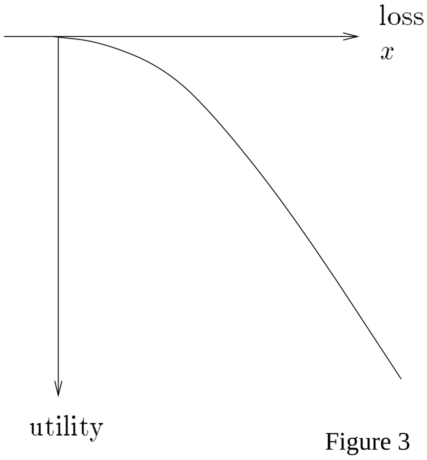


Figure 3

You will have realized by now that both these examples are close to being mirror images of each other. Similar types of advertisements can be found for travel insurances in the Tube in London. Here the rates "per day" are outrageous in my opinion, though on a day by day basis they seem low.

The following may be skipped at first reading.

1.5 Expected Utility

Many utility functions u have the properties

$$u'(x) > 0, \qquad u''(x) \leq 0, \qquad (1.5.19)$$

i.e. they are concave. This holds at least in some domain, as we have seen above. Then *Jensen's inequality* states

$$E(u(X)) \leq u(E(X)). \qquad (1.5.20)$$

For proof consider Figure 4 with $\mu = E(X)$. Then the tangent g to the graph of u in $(\mu, u(\mu))$ is given by

$$g : y = u(\mu) + u'(\mu)(x - \mu)$$

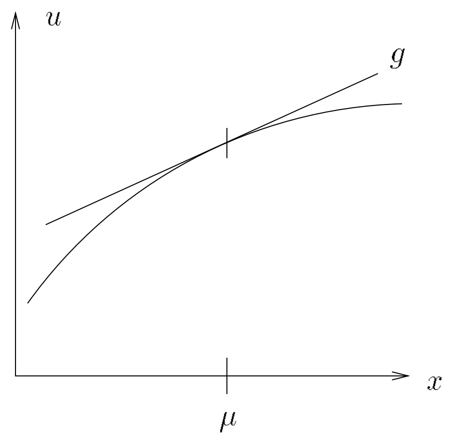


Figure 4

and u satisfies

$$u(x) \leq u(\mu) + u'(\mu)(x - \mu).$$

This can be seen directly from the graph, but you should also be able to prove this analytically. This inequality gives

$$u(X) \leq u(\mu) + u'(\mu)(X - \mu).$$

Now take expectation values and use $E(X - \mu) = 0$.

Jensen's inequality can be applied with expected utility as follows. Assume there is a person with utility function u with (1.5.19) and wealth w who faces a risk X. Let Pr be the premium he is willing to pay. Then by expected utility

$$u(w - Pr) = E(u(w - X)) \leq u(w - E(X)).$$

Since $u'(x) > 0$, u is strictly increasing and

$$w - Pr \leq w - E(X)$$

or

$$Pr \geq E(X). \qquad (1.5.21)$$

This person is thus willing to pay more than the net premium. For this reason utility functions with $u''(x) < 0$ are called *risk averse* or risk avoiding.

Thus people may have the incentive to buy insurance coverage even when the premium exceeds the expected loss. In such cases insurance companies can take advantage by pooling many related risks, because the relative fluctuations get smaller as the portfolio increases. Let us summarize the results of this part so far.

The perception of risk depends to a large extent on psychological and personal factors like status, wealth, position .... People's perception of the size and probability of a threat is more of a qualitative nature.

Utility is a useful concept to describe decision making in the case of monetary risks. People, however, only have a vague and qualitative feeling for utility, in particular as regards expected utility. Above we have seen why people might want to take out insurance coverage. But there are further reasons beyond utility which induce people to buy insurance protection.

In many situations insurance is required by law. Examples are liability insurance for car owners, fire insurance and in some cases also health insurance. The principle in these cases is usually that the possibly arising damage might exceed the financial constraints of the individual. In order to avoid this, the state has to step in and this risk is transferred to the insurance policies. A bonus in these cases is that most of these insurances are tax deductible. In fact, this is the second major reason why people take out insurance – they are tax deductible or lead to tax reduction. The main example of this is life insurance. This case will be discussed later in detail.

Since we have spent so much time with individuals' perception of risk, we should also spare a few thoughts for a working definition of risk from the insurance point of view, although the rest of these notes will be devoted to this. It is quite obvious that the utility of money for an insurance company will be the identity function, $u(x) = x$. As regards the risk, this arises mainly from the fluctuation term if the net premium is the expectation value of the damage. The protection from risk of an insurance company thus results from a large number of insured risks, the loading of the premium with the fluctuation term and from the reserves. Almost 70 years ago, Cramér, the founder of risk theory, wrote: "The object of the theory of risk is to give a mathematical analysis of the random fluctuations in an insurance business and to discuss the various means of protection against their inconvenient effects".

Summary: An important factor that makes people buy insurance coverage is the perception of risk. This depends on many individual factors.

Problems

1. Make a list of events or disasters that you find most threatening.

2. What is in your opinion the chance of being mugged. Compare it with available crime statistics.

3. In these days people consider unemployment to be a serious risk. How would you go about providing coverage against unemployment? What are the key factors? In what way would the premiums for a computer specialist and a handyman differ?

4. Make a list of factors concerning your own perception of the risks of nuclear reactors (airline accidents).

5. Show that if u is a utility function, then $\tilde{u} = au + b$, $a > 0$ is a utility function too. $\tilde{u}$ leads to the same results when used for expected utility.

6. How does personal wealth influence the utility function?

7. Suppose the utility function of a person is linear $u(x) = ax + b$ with $a > 0$. Assume X and Y are random monetary prospects. When would this person prefer X to Y on the basis of expected utility?

8. If people's decisions are guided by utility functions what would the appropriate function of a company be?

9. A person with wealth w and utility function u is faced with a loss X. Show that he would accept an insurance premium Pr against this loss if

$$u(w - Pr) \leq E(u(w - X)) \tag{1.5.22}$$

   if this person bases his/her decision on the expected utility principle.

10. Prove Jensen's inequality in all details.

11. Show that $u(x) = -\exp(-\alpha x)$ $\alpha > 0$ is a risk averse utility function, which is independent of the wealth of individuals as regards monetary decisions.

12. Estimate the probablity of

    (a) a fire in your family home

    (b) a 50-year-old male (female) dying within a year. Compare this with actual data obtained from life tables or insurance records.

13. In your opinion, what is the chance of a car driver being involved in an accident and what do you expect the average damage to be?

14. Do you consider \$ 2 for each \$ 1000 of property a fair insurance premium for earthquakes in California?

15. How would you determine the "value" of a person's life. What are the main factors?

Literature:

H.S. Deneberg et al	Risk and Insurance, Englewood Cliffs N.J. Prentice Hall 1974
H. Kunreuther	Risk. A Seminar Series. Int. Inst. for Applied Systems Analysis, Laxemburg Austria 1981
N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones, C.J. Nesbit	Actuarial Mathematics
H. Jungemann	Subjektive Wahrnehmung von Risiken (Spektrum der Wiss. Februar 1998)

2. History and Regulation of Insurance

2.1 The History of Insurance

Insurance probably began with the first appearance of large scale commerce. In bottomry contracts ships or their cargo were pledged as a security against which a merchant could borrow at elevated rates. If the ship or its cargo were lost the debt would be cancelled. Such forms were practised in Babylon, ancient Greece and Rome. In the middle ages the sharing of risk, of fire, theft, floods or life was practised by guilds for its members. The first laws relating to insurance date back to Hammurabi around 1800 BC. Until the 16th century insurance was mainly practised by merchants, who occasionally underwrote policies relating to commerce. In this informal way marine insurance started in the 17th century in the coffee house of Lloyd's in London, where merchant shipowners, brokers and underwriters met. But marine insurance was also known to other seafaring nations like Spain.

The roots of insurance as a specialized enterprise began in the latter half of the 17th century and expanded greatly in the 18th century. The fire of London in 1666 gave a tremendous impetus to fire insurance. J. Graunt (1620–1674) began a systematic study of death and church registers and thus laid the groundwork for mortality tables, which were then constructed by Halley. But they were only used systematically about a century later.

A great advance for insurance also came from the theoretical side. Orginating from the study of games of chance in 17th century France, Jakob Bernoulli (1654–1705) formulated the law of large numbers, which we used essentially in chapter 1. These studies were expanded by Laplace, Euler and K.F. Gauss (1777–1855) and finally led to the central limit theorem. Regression analysis and the systematic study of errors in observation were also first developed by Gauss. Both domains are essential for insurance mathematics. The study of limit theorems has been pursued actively since that time and it is still a flourishing field. For insurances, however, more elementary versions suffice. Since then the insurance industry has grown tremendously parallel to the growth of industry and trade. Insurance mathematics, however, was only developed in this century in the wake of probability theory. This applies in particular to risk theory and credibility.

2.2 Government Regulation and Legal Aspects

In the remaining sections of this chapter we will cover the economics of insurance, which mathematicians also need to know.

Types of insurance

Although insurance developed from the commercial sector, it meanwhile covers all domains of public life. In fact life, health and car liability are now the most dominant types of insurance. The following types can be distinguished

- Marine and aviation insurance
- Property insurance (fire, theft, car, glass breakage, ... , vandalism, earthquake, flood, weather)
- Liability insurance (cars, business, doctors, ... , private)
- Personal insurance (life, health, accident, disability, unemployment)
- Reinsurance.

Of course there are many more forms of insurance, but the basic principles are the same for all. Besides, the insurances above cover the largest part of the insurance spectrum. In addition to individual insurance one also finds *group insurance*, e.g. a firm buys a health or life insurance package for all its employees. Such insurance packages are particularly important in Japan, where people usually stay with one firm throughout their lives. Also in Europe such group insurance packages are of growing importance, because most government schemes are insufficient.

Types of insurance companies

Insurance companies may be classified according to their function (product) and their organization or method of operation.

Stock companies: These companies are organized as stock companies. Thus they earn profit for their stockholders which they obtain in the form of dividends. If the premium income does not cover all losses, the stockholders are liable. This requires efficient management to keep operating costs low. The board of directors of such a company is elected by the stockholders.

Mutual companies are corporations which are owned by the policy holders. Mutual companies are thus organized for the purpose of providing insurance for its members. Thus all operation surplus is returned to the members. A closely related form are interinsurers or reciprocal underwriters.

Lloyd's is an association of a number of persons, each of whom accepts share of a risk to be covered. A new form of this type is becoming increasingly popular in the form of insurance bonds for high risk insurances. In this case the dividends depend on the occurrence of the risk.

Government Insurance is, as the name implies, organized by a government. In the US social security is of this type. Health care, accident, unemployment insurance plans and pensions are organized in this way in many countries, as we will see later.

Since insurance first developed in the commercial sector, its form as a contract was apparent right from the beginning. Thus the first laws concerning marine and robbery insurance can be found in the code of Hammurabi (around 1800 BC). Later on it became increasingly clear that laws and regulations were needed to guard against abuse. Usually regulations apply to the following areas

1. Form and character of insurance contracts
2. Solvency requirements
3. Regulation of prices and premiums
4. Admission to an insurance
5. Entry of new firms and domains of activity.

Control is usually exercised by the central government, the local government or some assembly of insurances.

In more detail this means

1. An insurance is a contract between the insurer and the insured. Thus it must meet several requirements. These are: It must specify the purpose and interest of the insured. Both must be lawful. Moreover, the insurer and insured must have the legal capacity to contract. Finally, this contract must specify the premium, the conditions and amount of coverage. Think, for example, of suicide clauses for life insurance. In addition, it must be evident that the insured has a personal interest in the insurance. The latter condition means that one may buy an insurance only for oneself or one's family, but not for one's neighbour. Insurance is a coverage for incurred losses. Thus it should also be clear that the insured is not better off after the loss than before. This is called indemnity. This, however, does not apply to life insurance, for obvious reasons.

2. In particular in the early years of insurances, some people considered insurance as an easy-to-get-rich scheme. A fictitious example for this can be found in the novel "The Rainmaker" by J. Grisham. This amply shows the need to regulate solvency and reserves. Usually this will also specify how and to what extent money should be invested. Solvency regulations are in fact more complicated than they appear at first sight. For one, these rules must specify the level of security p_e which is considered adequate. Secondly, such regulations have to deal with the costs and benefits of such reserves. All this has to be done on a continuous level. In the US, Great Britain and some other countries the device of a guaranty fund is used. In this case other insurance firms will step in if an insurer becomes insolvent. Such a system has the advantage of offering a high degree of consumer protection and it induces self-regulatory control.

This is not yet practised on an international scale. Even though solvency requirements vary greatly from country to country, there were only very few insolvencies in recent years. It should be noted, however, that strict solvency rules usually keep small firms out of the market or prevent new ones from entering.

3. *Regulation of prices and premiums*

   If prices and premiums are too low, the insurance company might go bankrupt. If they are too high competitors may make the business. Thus a good deal of price regulation is provided by the market. This is limited, however, by the inability of the public to effectively compare insurance contracts and coverage, or by cases of monopoly. For this reason price regulation varies greatly from country to country with extremely varying benefits for the consumer. In Germany, for example, a heavily regulated country, premium schedules had to be submitted to a government bureau (*Versicherungsaufsichtsbehörde*). In some countries the role of a government agency may be taken over by the insurers' association. A somewhat simpler mechanism is the *ex post* control of profit. A price control is clearly also required if an insurance is mandatory. In some cases price control is achieved by a public insurer. Price and premium control requires some form of standardization of the insurance contracts. Altogether it seems that unregulated markets are imperfect, but regulated markets seem to be worse (The Economics of Insurance Regulation; J. Finsinger, M. Pauly, Mc Millan 1986, p. 20). The following list from the book by Finsinger, Pauly p. 21 gives a simple comparison for life insurance.

   Mean annual net premium for a 5-year term life insurance for a 30-year-old male with 100,000 benefit. The currency unit is DM.

France	326	Germany	338
Canada	259	Sweden	440[2]
US	270	Switzerland	268
Great Britain	138[1]		

   It shows that the heaviest regulation in Germany also leads to the highest premium.

4. *Admission to an insurance*

   In particular, if an insurance is compulsory, the access to insurance coverage must be regulated.

5. *Entry of new firms, domain of activity*

[1]The British system of taxing life insurance means an implicit subsidy to term life policies of approximately 15 per cent. Even after adjustment for this tax effect the British market offers the lowest premiums.

[2]Only the gross premium is available. This is the same for all firms. Depending on profits, individual firms pay dividends or increase coverage, such that the net premium is lower *ex post.* For all other countries these net premiums have been calculated.

In several European countries insurers may not underwrite both life insurance and property insurance. Also the admission of foreign insurers may be restricted.

Since 1994 the insurance market in the EU has been liberalized and deregulated. Liberalization means

1. Free flow of capital within the EU
2. Freedom to enter the insurance industry in all EU countries
3. Single licence, i.e. admission by a government agency in one country suffices for admission in all EU countries. Solvency control is undertaken only in the home country.
4. This also requires harmonizing the rules. Thus uniform and minimal constraints for all insurances hold throughout the EU.

For some heavily regulated countries like Germany, for example, this also implies deregulation, i.e. a reduction of constraints. Rules pertaining to these aspects of insurance have been introduced gradually since 1957. Since 1994 the market within the EU has been largely liberalized and deregulated. This makes it mandatory for insurance mathematicians to study the different types of products in various countries.

The aim of liberalizing the market was to create a larger and more diversified spectrum of insurance, to increase competition and improve customer orientation. So far this has goal has not been attained, mainly because insurance has evolved in all countries over a long time, thus reflecting a good deal of economic and social culture of the various countries. As a result there has been a tendency for companies to buy themselves into existing companies in other countries in order to acquire their name, know-how and infrastructure. Liberalizing the market and creating a uniform European insurance market will therefore take time.

2.3 Life insurance in the EU

The life insurance market in the EU can be characterized by the following:

1. There is a relatively large amount of people over the age of 60 (see the plots for the old age ratio by the *Statistisches Bundesamt* in Section 1).
2. Even though more governments are in debt, there is a large amount of private wealth. In Germany, e.g. the private money reserves are about $5.7 \cdot 10^{12}$DM. In earlier years people used to put their money into a savings account. Nowadays the stock market, bonds and insurances are the primary destinations for money. In 1998 people in Germany invested $1.26 \cdot 10^{12}$DM in insurance, but only $1.19 \cdot 10^{12}$DM in savings accounts.

3. A good deal of this wealth will be passed on to the next generation within the next 10 years. Therefore a market for pension schemes exists.
4. Government insurance plans in most countries are inadequate. Thus there is a considerable demand for additional private protection.
5. This is partly supported by the governments which offer tax reductions and tax saving to varying degrees.
6. These tax laws have grown and come into existence over the last 20 years or so and are now largely inadequate. What is needed is a unified treatment of all systems providing for old age, e.g. saving accounts, investment funds, stock, insurances, etc.
7. Besides harmonizing these laws nationally, a harmonization at EU level is needed. This applies in particular to the tax system.
8. Since 1994 the life insurance market has been liberalized and deregulated. This has led to increased competition among insurers. This in turn led to a larger variety of insurance products and more consumer orientation.
9. Thus insurances will become more flexible and contracts will be increasingly tailored to specific needs by modularization.
10. New forms of life insurances, e.g. insurances tied to stock or funds, Universal life, ... will spread. Life insurances with a large savings component will increase.
11. New methods of selling, direct insurers or sales by banks or via the internet will increase likewise.
12. The largest hurdle in creating a unified European market are differing tax laws and different socioeconomic parameters.

A life insurance may be bought as an individual insurance or a group insurance. This latter form can be used by firms for their employees. The economic importance of life insurance arises from

1. Stabilizing the economy by reducing the burden of care and support for the government. In most European countries government insurance schemes are generally inadequate. Thus additional insurance from the employer or private insurance is needed.
2. The savings effect of life insurance stabilizes the economy and provides funds for investment.
3. They are a means to acquire property and wealth. One should realize, however, that investment in stock would have realized greater returns in the last 20 years. The tax on gains in the stock market is generally much more severe. In recent years, however, insurance companies have offered more insurances tied to the stock market or particular investment schemes.

Summary: From its beginning insurance was tied intimately to commerce. Thus the internationalization of trade has similar consequences for insurance. On the other hand, some aspects of insurance depend on the socioeconomic factors of the corresponding country.

2.4 Types of Life Insurance

As noted above, insurance products are already rather diversified. The list below will thus only give an overview of the classical standard types of insurances.

Term Insurance: This type is a protection against disasters resulting from death during a specific period of time. This may be 1, 5, 10 years or even longer or, e.g. until the insured reaches the age of 65. The face value of the insurance is paid to the beneficiary if the insured dies within the specified term. Otherwise nothing is paid.

Sometimes the policy may be extended at a higher premium. The term insurance is pure protection against the risk of death and contains no saving element.

Whole Life Insurance: This insurance provides protection to the face amount at a level premium for the whole life time of the insured. Thus the insurance pays the face value in case of death regardless of when it occurs. In the *limited pay whole life insurance* the protection is as above, but the payment of the premium is restricted to a certain period, e.g. as long as the insured person works.

Endowment Insurance: The face amount will be paid to the beneficiary if the insured dies during the policy period. Otherwise the face value will be paid to the insured. In many cases companies will add dividends on the surplus to the face value. The claim and dividends may also be tied to a particular investment scheme.

Universal Life: Universal Life is similar to a whole life insurance, but it is more flexible regarding both premiums and coverage, e.g. payment of premiums may be interrupted and the coverage may be altered later. Moreover, it is easier to cancel the insurance or increase the reserves.

So far this type is mainly restricted to the USA.

Life insurance tied to investment funds or stocks: Provision for old age plays an important role in life insurance business. Thus whole life insurances or those with participation in the gain play an important role. On the other hand the yield of such products is lower than other forms of investment like stocks or funds. In order to increase the yield, life insurances have therefore developed a product which is tied to an investment fund. In this case the yield is generally higher, but the subscriber also carries part of the risk.

Annuities: An annuity is a periodic payment made during a fixed period of time. A *life* annuity is paid during the lifetime of the annuitant, while a temporary annuity is only paid for a fixed time or until the annuitant dies.

Of course there are many other forms of life insurance, some of which we will discuss later in examples.

The Premiums: The standard premium is the *level premium*. This type of premium is constant throughout the contract period. A *single premium* is paid only once at the beginning of the contract period. But other forms of variable premiums may also be agreed upon.

There has been a trend recently in the insurance industry to modularize insurance in order to adapt them to the needs of customers. Thus the number and forms of policies have increased considerably. This applies in particular to the various savings and dividend schemes. An important motivation for life insurances, in particular those with profit participation, are the possibilities of *tax reduction*. Thus the specific tax laws play an important role in defining insurance schemes. Insurance mathematicians will thus have to analyse the market regularly to see what competitors offer and consider the tax laws to tailor new and efficient insurance schemes. These laws and conditions, however, are permanently being altered because governments, although aware of the need for provision for old age, keep a greedy eye on money in the insurance sector.

Below we shall list in more detail the life insurance environment and the type of policies used in France, Germany, Great Britain and Spain. These data are extracted from a study of S. Nyfeler and A. Lehmann "Lebensversicherungsprodukte im europäischen Vergleich. Ergebnisse einer internationalen Studie in 14 Ländern". These results were compiled somewhat before 1995, so there may be some changes by now.

2.5 The Insurance Environment in Various Countries

France

Social System

The French social system consists of several subsystems, in each of which a particular group of people is insured. This insurance is mandatory but can be voluntary in certain cases. The majority (75%) of working people are covered by the "Régime Général". Since the other systems for particular occupational groups are similar, we will only deal with the Régime Général. The Régime Général consists of benefits for old age, surviving dependents and insurance for sickness, invalidity and accidents. Insurance is mandatory for all employees in industry, commerce and the service sector, regardless of their income. This system is paid for by those currently employed. Special social systems exist for employees in agriculture and for self-employed people.

Unemployment benefits are organized by the government. Payment is via CNAM (Caisse national d'Assurance Maladie). The contracts are negotiated between representatives of employees and employers.

Health insurance is a family insurance which also benefits the dependents of the employees. Even though 99% of all French people benefit from it, 66% have additional insurance by mutual companies (mutuelles). Only about 4% have a health insurance with a private company. Health insurance is administrated by the CNAM, which also supports an *accident insurance*.

Provisions for old age

Pensions are provided by the government via the CNAV (Caisse nationale d'Assurance Vieillesse). For this the employer pays 12.8% of the income and the employee pays 6.8%. The pension amounts to 50% of the income of the 10 best years if the employer has worked for 37.5 years. Hence there is an additional mandatory insurance for all employees which is organized by ARRCO (Ass. des régimes de retraites complementaires). For higher level employees a similar insurance is organized by AGRIC (Ass. pour la gestion des institutions de retraites de cadres). Private insurances for old age are tax supported, because the public pension plans are insufficient. More than 60% of all savings contracts operate via such an insurance and thus allow a considerable savings of taxes. Term insurances are non-tax deductible, though annuities are added to the general income, but are taxed at a reduced rate, depending on the age of the annuitant. Dividends in endowment insurances are generally not taxed. The tax laws have been particularly favourable to the self-employed since 1994 ("la Loi Madelin"). This has led to strong competition on the annuity insurance market.

Insurance Law

There is a general insurance law (Code des Assurance) and an insurance supervisory law (la Loi Evin, 31.12.89) which is practised by a department in the Ministry of Economic Affairs (Service des Assurances du Ministere de l'Economie). Supervision is carried out by the "Commission de Contrôle des Assurances".

The French Life Insurance Market (1992)

Number of life insurances	145
Share of premiums of GNP	3.8%
Share of foreign insurance companies	2%
Life insurance contracts per capita	0.624
Premium exprenses per capita	4,700 F

	Gross Premium (10^9 FRF)	Market Share
UAP Vie	28.393	10.5 %
CNP Ass.	19.583	7.3 %
AGF Vie	19.382	7.2 %

Germany

The social system

The social system in Germany consists of a regulated unemployment, health, accident and old age insurance. These types of insurance are mandatory, with the exception of self-employed people or those with an income of less than DM 630. The premiums are reasonably proportional to the income. The premiums for civil servants, judges and the military are covered by the government.

Unemployment benefits are administered by the "Bundesanstalt für Arbeit in Nürnberg". The premiums for this are covered in equal parts of 6.5% by employers and employees alike.

Health insurance: Government health insurance covers about 87% of the population. Each insured person pays about 13% of his/her income for this. Self-employed people are usually privately insured, while civil servants generally have their own or an additional health insurance.

Accident insurance: Accident insurance is administered by occupational guilds (Berufsgenossenschaften). Many people, in particular the self-employed, are privately insured or use additional private coverage.

Provisions for old age: The government pension scheme is administered by the BfA (Bundesanstalt für Angestellte). It is financed by the premiums of those currently employed – about 20% of the income – Generationenvertrag. The increase in the group of elderly people will cause serious problem in the near future.

Development of the Number of Pensioners per 100 employed

Year	1992	2000	2033
	48	59	≥ 100

Serious problems also arise because pensions for citizens from the former East Germany and ethnic Germans from Russia have to be paid out of this system. There is also a government pension scheme for civil servants. About every other employee has an additional pension plan from his employer. These payments are tax exempt for up to DM 3,000 per year. Self-employed people beyond a certain income are usually privately insured. The tax exemption is DM 6,000 for single persons and 12,000 for couples per year for most types of life insurance, including term insurance. The tax exemption is also valid to a large extent for the proceeds, including dividends, of these insurances.

Insurance Law: The insurance industry is mainly regulated by the following law.

1. Law to establish the insurance supervisory organisation (BAG Gesetz zur Errichtung des Bundesaufsichtsamtes).

2. Insurance supervision law (VAG Versicherungsaufsichtsgesetz). The government insurance supervision organisation is a separate government institution associated to the Ministry of Finance. It controls general insurance conditions, premiums, dividends, participation in profits, investment of insurance companies and solvency.

3. Law for insurance contracts.

4. Law for terms of business.

The liberalization and deregulation since 1994 has greatly simplified the law and conditions for the German insurance market.

The German Life Insurance Market

Number of life insurances	122
Share of premiums of GNP	2.31%
Share of foreign insurance companies	3.6%
Life insurance contract per capita	.95
Premium expenses per capita	856.80

Top three life insurers

	Premiums (10^9 DM)	Market share
Allianz	8.703	12.9%
Hamburg-Mannheimer	4.051	6 %
Volksfürsorge	3.292	4.9%

Great Britain

The social system

The British social system is a general insurance system which covers the whole population. It consists of the National Insurance and the National Health Insurance. The National Insurance addresses old age support, benefits for surviving dependents, insurance benefits for invalidity, accidents at work and unemployment. The National Health Service only provides health care benefits. This social system is paid for by employees and employers, depending on income. A reduction of these payments is possible if the employee uses a pension scheme run by the employer (contracting out). The control of both systems lies with the Department of Health and Social Security.

Unemployment benefits depend on the average weekly income over the previous year. In addition, the unemployed have to demonstrate that they are actively looking for employment. The benefits are limited to 1 year.

National Health Service: There is no mandatory health insurance in Britain. Thus 90% of the population are solely insured by the NHS, while about 10% have an additional insurance. The NHS provides out- and in-patient services and services by specialists free of charge. In addition, there is a statutory sick pay and a sickness benefit.

Accident Insurance is provided by the NI. The benefits depend on the degree to which the accident was employment related.

Provisions for old age: The governmental pension scheme is administered by the NI. As regards pensions one has to distinguish between the basic pension and the State earnings related pension scheme – SERPS. All employees who have worked at least 90% of their working lives are eligible for the basic pension, which starts at the age of 65 for men and 60 for women. Employees are required to supplement their basic pension via SERPS. This is not mandatory for self-employed people. The basic pension scheme and SERPS amount on average to about 40% of their last pay. An additional private or occupational insurance is thus desirable. In fact this situation has been criticised by the charity organisation "Help the Aged". Their slogan in the autumn and winter of 1998 was "Thousands of aged people will die from cold" because they cannot pay for adequate heating. Thus the new Labour government is planning to overhaul the current system.

People can leave SERPS if they have an appropriate personal pension or a special personal plan or belong to an occupational pension scheme – contracting out. About 55% of all employers are covered by an occupational pension scheme. This is attractive, because it is tax exempt to a large extent. This tax exemption increases with age. Thus many employees also supplement this by Additional Voluntary Contributions (AVC-scheme).

About 5 million people have a personal pension plan, which also allows some tax reduction. As noted above, there are currently numerous changes in the British social system. For the latest information you should therefore consult the web site http://www.dss.gov.uk.

Insurance Law

Insurance companies in Britain are supervised by the Department of Trade and Industry (DTI). The relevant bodies of law are:

- Insurance Companies Act (ICA 1982) and
- Financial Services Act (FSA 1988).

As a whole the law is rather liberal and weakly regulated, apart from solvency control, and relies to a great extent on self-regulation. Thus life insurances have an appointed actuary. This construction has been introduced in all EU countries since 1994. Moreover, supervision and control is relegated largely from the DTI to the Securities Investment Board – SIB. The members of the SIB are laymen or come from insurance companies. But they are nominated by the DTI. In addition, insurance companies have formed their own supervisory organisations:

- LAUTRO – Life Assurance Trust Regulatory Organisation
- FIMBRA – Financial Intermediaries, Managers and Brokers Regulatory Association

- IMRA – Investment Managers Regulatory Association.

The British Life Insurance Market

Number of life insurances	130
Share of premiums of GNP	7.5%
Share of foreign life insurance comp.	24%
Premium expenses per capita (1992)	£750.43

Top three life insurers (1992)

	Premiums (£10^9)	Market share
Prudential	3.741	8.6%
Standard Life	3.112	7.2%
Legal & General Group	2.172	5 %

Spain

The Social System

The social system in Spain consists of a general system and several special systems for particular occupational groups (civil servants, farmers, self-employed people). With the exception of these groups, this general social insurance is mandatory for all employees. The premiums are 4.9% of the income paid by the employee and 24.4% paid by the employer. This system comprises a pension, health, accident and invalidity insurance. Unemployment benefits are organized by INEM (organization for labor). The premiums are 6.2% for the employees and 1.1% of income for the employers. Unemployment benefits are paid for a maximum of 2 years. The amount depends on the average income over the last 6 months. More precisely, 80% of the average income is paid for 6 months, then 70% for the next 6 months and 60% thereafter.

Health insurance is administered by the Organisation for Health – INSALUD. Self-employed people can become voluntary members. The general health insurance also covers sickness pay, which maybe supplemented by employers to about 90% of the general income.

Private health insurance is in general used as a supplementary insurance.

Accident insurance encompasses medical treatment, rehabilitation, sickness pay and pension (if necessary). Similar regulations hold for invalidity insurance.

Provision for old age is administered by the Organisation for Social Services (INSERNO). People are entitled to a pension from the age of 65 onward to 100% of their average income over the last 8 years. The system is financed by the current premiums. Because of the relative increase of old aged people this system faces serious problems in the years to come. The situation is not as grave as in Germany though. Nonetheless, a reduction in payments can be expected.

Large enterprises often offer particular pension benefits. These are usually tax deductible for the firms, but may be taxed partially on the side of the employee. Thus again private pension schemes have played an ever increasing role as supplements. Here one has to distinguish between "Plan de Pension", where the premiums are tax deductible up to 15% of the income and Seguro de Ahorro/Jubilación. For the latter the premiums are tax deductible up to 500,000 Pts, while the tax reduction for the pension lies between 30% and 70% and depends on the age of the recipient and type of pension.

Insurance Law: The fundamental insurance law is the "Ley de Ordenación del Seguro Privado" of 1984 and a number of decrees since that time. The supervision is enacted by the "Dirección General de Seguros" which is associated to the Ministry of Finance. The DGS supervises premiums, rate-fixing and solvency. Since 1984 an insurance company may not trade simultaneously in life and non-life insurances.

The insurance contract law is the "Ley de Contrato Seguro" of 1980.

The Spanish life insurance market: The insurance market is rather atomized and concentrated on small companies with intensive competition. Thus the top 10 companies have a market share of only 53%. Economic difficulties in the late 1980s have led to considerable premature cancellations of life insurance contracts. The market share of foreign companies is rather high (40%).

The Spanish Life Insurance Market

Number of life insurances	141
Share of premiums of GNP	1.46%
Share of foreign insurance companies	40%
Life insurance contracts per capita	0.124
Premium expenses per capita	22.539 Pts

Top three life insurers (1992)

	Premiums $\cdot 10^9$ Pts	Market share
Euroseguros	104.726	13.12 %
Vida Ceixa	65.971	8.27 %
Grupo Vitalico	55.994	7.02 %

Regulation Framework for Life insurances

	Germany	France	Great Britain	United States
Whole life insurance Reserves and solvency	National government sets requirements on amounts and types of assets	Reserve and solvency rules	Solvency requirements; levy on life premiums guarantees 90 per cent of benefits in event of insolvency (not funded)	Solvency and assets regulated by states, usually via minimum capital requirements
Premiums	Unregulated	Uniform and excessively high premiums: firm-specific premiums	Entirely unregulated	Unregulated, except limits on expenses in 3 states
Term life insurance Reserves and solvency	National government sets requirements	Reserve and solvency requirements	Same as for whole life	Minimum capital requirements
Premiums	Unregulated	Firm-specific regulation of markup type; firm-specific rebates	Entirely unregulated	Entirely unregulated
Taxation of whole life insurance	Some or all premiums paid are tax deductible	All benefits exempt from taxation; depending on personal characteristics some or all premiums paid are tax deductible	Premiums receive 15 per cent (inclusive) subsidy if qualifying policy; income generated by fund taxed on income minus expense basis; benefits exempt from taxation	Excess of benefits over premiums subject to income tax in period benefits are received; death benefits not subject to tax

The current trend in the insurance market in Europe can be summarized as follows:

1. The national markets and peculiarities will remain, because these have grown over the last 50 years or so and reflect the social and judicial environment and tax laws.

2. The increase in the group of people at pension age increases the trend for governments to reduce its share in the insurance sector. Thus group insurance by employers or additional private insurance will become increasingly important. In particular, insurances with a larger savings component will increase. An additional momentum for this is the growth of private wealth throughout Europe. High government deficits, however, lead governments repeatedly to attempt to get its share of the private wealth by modifying the tax laws. So these are permanently being overhauled with the aim of getting more taxes,

but so that the public are not aware of it.

3. Because of the growing importance of private insurances and the inadequacy of public schemes, tax reduction for such products will remain. It will thus be important in the development of new products.

4. Modularization and specialization in the insurance market will allow each individual to find insurance tailored to his needs. This flexibility in the market is also a consequence of the deregulation and liberalization of the insurance market.

5. Increasing competition will lead to more flexible insurance schemes and more advantageous premiums. New products will be introduced more rapidly.

6. Harmonizing the European insurance market will depend on harmonizing tax laws and social standards.

Summary: As a consequence of industrialization, governments have set up insurance schemes for health and accident insurance, invalidity and pensions. The amount of coverage and conditions differ from country to country. In most countries the coverage is inadequate and the systems have to be overhauled. In all countries changes in the social system are being introduced, but we are still far from uniform standards.

Problems

Problem 2.1 Check within your family which types of insurance they have and how provision for old age is insured.

Problem 2.2 Check with an insurance agent about new life insurance contracts.

Project: Use government statistics to project the development of the population. What does this tell you about the cost of welfare?

3. Life Insurance

3.1 General Considerations

A life insurance is a scheme to control and share the risk of premature death or invalidity. Thus the only chance element involved is that of mortality. Since mortality is usually rather low and increases with age, life insurance contracts operate on a long-term scale, usually 10 to 40 years. In addition to such life insurance, pension schemes form an important segment of life insurance. Besides risk sharing and provisions for old age, tax saving is the main motivation for such schemes. Life insurance is not a contract of indemnity, because the monetary value of human life cannot be determined precisely. It is simply impossible to put the insured or his dependents in exactly the same financial position as before the death of the insured person. Thus life insurance policies are bought at a fixed face amount.

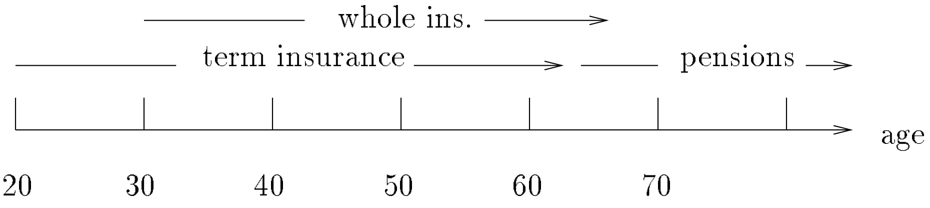
The question of *insurable interest* seldom arises in life insurance, because in most cases the insurance is bought by the insured with beneficiaries being the next of kin. A creditor may have an insurable interest in the life of his debtor. But in these cases the claim is usually confined by statute to about the amount of debt. Coverage of a life insurance is usually effective when the application has been received and the first premium been paid.

Since these notes primarily address the mathematics of life insurances, we will not analyse the legal concepts and provisions for life insurance contracts, but rather refer to the book "Life Insurance" by S.S. Huebner; K. Black; Prentice Hall Int. 1982. It should be noted, however, that current laws and insurance contracts are so complicated that they cannot be understood by laypeople. This makes an effective comparison of contracts almost impossible. One can only hope that the European harmonization also leads to a simplification of contract law and contracts.

3.2 The Use of Life Insurances

In most European countries governments provides some form of insurance scheme. But in general this is not sufficient. Thus life insurance is used to supplement government and employer's insurance schemes. For families private life insurances are determined by *need*, *means* and *taxation*. The following life cycle model for life insurances highlights need as a function of age.

Thus at an early age of up to 45 people tend to insure against the risk of death. Then at the height of their career, saving for a pension is important. Of course, in view of the tax law to alleviate the "high" tax burden.



The life cycle model

The following example should illustrate how to proceed on the basis of need. It is somewhat more detailed than the life cycle model.

Example 3.1 Mr. Breadwinner is married with two children. His wife, who formerly worked in a bank, is 35 and their children are 10 and 8 years of age. The Breadwinners own a house with a mortgage of 120,000. Their savings have just been used up for the new car. Mr. Breadwinner currently holds a life insurance of 80,000. The government insurance plan and the pension plan of his employer would guarantee his family an income of 2,000 (1,500 with one child) until the children come of age. Thus he is considering buying additional life insurance coverage. Here he is also considering that his wife could work half-time until the children are old enough and full-time thereafter, when the youngest is 18. To arrive at the required amount, he begins to put down the needs and the possible income

Needs:		*Income:*	
Mortage for the house	120,000	Government and employers' insurance benefits	1,500
Clean up fund for funeral medical costs, debts emergencies	10,000	Income of wife beginning the second year	1,000
Needs of the family until the oldest is 25	5,000/month	Income of wife in full-time work	2,000
Thereafter until the youngest is 25	4,000/month	Wife's own insurance plan at retirement	800
Then	3,000/month	Life insurance	80,000

It is obvious that the need of Mr. Breadwinner's family varies with time. More precisely we have:

Needs 1st year:		*Income 1st year*	
Mortage	120,000	Life insurance	80,000
Clean up fund	10,000	Pensions	18,000
Living expenses	60,000		
	190,000		98,000
Needs year 2–9		*Income*	
Living expenses	60,000/year	Pensions	18,000
		Wages	12,000
Needs year 10–15		*Income*	
Living expenses	60,000	Pensions	18,000
		Wages	24,000
Needs year 16– 17		*Income*	
Living expenses	48,000	Pensions	18,000
		Wages	24,000
Needs year 18–24		*Income*	
Living expenses	36,000	Pensions	18,000
		Wages	24,000
Needs thereafter		*Income*	
Living expenses	36,000	Pensions	27,600

Altogether this defines a demand of

Year 1	Years 2–9	Years 10–15	Years 16–17	Years 18–24	Thereafter
92,000	30,000	18,000	6,000	– 6,000	8,400

Depending on his means, Mr. Breadwinner can now buy a life insurance which covers all or part of these expected expenses. In doing so, one has to consider that all these sums have to be discounted, as we will see later.

Problem 3.1 Determine the face value of the life insurance policy which covers all eventualities until year 10 (16, 18, 24).

In real life the situation may be even more complex, because the government pension may depend on the number of dependent children. Along these lines you should now be able to compute finance schemes for all possible situations.

Another factor which we have not considered here is the possible death of his wife as well. All this will be taken up in the next chapter. In this chapter we will also provide the necessary mathematics to compute a life insurance for a pension plan. In all these considerations, in particular for supplementary pension schemes, the tax savings effect of premiums and the tax exempt status of benefits plays a key role.

In business a creditor will sometimes ask the debtor to buy a life insurance as a collateral. In some cases the creditor may even demand to be explicitly stated as a beneficiary in the life insurance contract. In this way the credit rating of a firm may be greatly improved. Another form of insurance might be a key person indemnification for a person whose skills are essential for the enterprise.

The largest scale use of personal insurance in industry, however, are welfare plans for employees. These include health, accident, life insurance and pension plans. As such they are an important supplement to government welfare schemes. Such plans have the advantage of binding employees closer to the enterprise and they allow tax savings. A further important element which we will take up later is the use of life insurance as an additional pension.

Summary: Buying a life insurance is determined by need, financial possibilities and tax saving opportunities. During one's life the need for a particular contract changes from pure risk considerations to an emphasis on saving.

3.3 Mathematics of Finance

Since life insurance contracts are usually long-term contracts, the interest on premiums, capital and claims plays an important role. Let p be the rate of interest, $p \in [0, 100]$. Then the *simple interest* for a capital K for n years is $n\,p/100 \cdot K$. This formula is only useful for short periods. For longer times one has to use *compound interest*. The corresponding formulas can be derived as follows: After 1 year the capital K has become

$$K_1 = K + K\frac{p}{100} = K\left(1 + \frac{p}{100}\right).$$

This is really the formula for simple interest. But if we leave the capital K_1 for another year in the bank it will become

$$K_2 = K_1\left(1 + \frac{p}{100}\right) = K\left(1 + \frac{p}{100}\right)^2.$$

Similarly, one sees that a capital K with compound interest has become

$$K_3 = K_2\left(1 + \frac{p}{100}\right) = K\left(1 + \frac{p}{100}\right)^3$$

after three years. Now the general formula is obvious:
An amount of money K with interest rate p will grow after n years in the bank to

$$K_n = K\left(1 + \frac{p}{100}\right)^n. \tag{3.3.1}$$

K is also called the *principal*, while K_n is referred to as the *accumulated amount*. The expression $i = \frac{p}{100}$ is also called the *interest rate*. In this formula we have used one year as a time unit. But the formula remains valid for other time periods as well. Then however one has to adjust the *interest rate*. If p is the yearly interest rate, it is customary to determine the interest rate for a shorter period by

$$\text{Interest rate per period} = \text{Interest per year} \cdot \frac{\text{Length of Period}}{\text{year}}. \tag{3.3.2}$$

Here are some examples:

Example 3.2 Assume the yearly interest rate is 8%. Then the monthly interest rate is 8/12% while the weekly interest rate is 8/52%.

Example 3.3 If one computes compound interest, the time period used is relevant. Assume we have a principal $P = 1,000$ and interest rate 6%. Then the accumulated amount after 3 years is

$$K_3 = 1,000\left(1+\frac{6}{100}\right)^3 = 1,191.02$$

if the interest is taken on a yearly basis. If one uses a monthly basis, one gets

$$K_3' = 1,000\left(1+\frac{6}{1,200}\right)^{36} = 1,196.68.$$

On a weekly basis one even gets

$$K_3'' = 1,000\left(1+\frac{6}{5,200}\right)^{3\cdot 52} = 1,197.09.$$

In the remainder we will mainly denote the time period by years, though this may be replaced by quarter years, months, weeks or even days, as long as the yearly interest rate p is adjusted to $p/4$, $p/12$, $p/52$ or $p/360$ respectively.

Formula (3.3.1) also shows that a principal of $P = K(1+\frac{p}{100})^{-n}$ will accumulate to

$$P\cdot\left(1+\frac{p}{100}\right)^n = K\left(1+\frac{p}{100}\right)^{-n}\cdot\left(1+\frac{p}{100}\right)^n = K$$

in n years with compound interest rate p. For this reason one calls $K(1+\frac{p}{100})^{-n}$ the *present value* of the amount K due in n years. The process $K \to K(1+\frac{p}{100})^{-n}$ is called *discounting*. Thus a payment of K in n years presently only has a value $P = K(1+P/100)^{-n}$, because one can put this amount into the bank and receive the amount $K = (1+\frac{p}{100})^n P$ after n years. In mathematics of finance it is thus not only important what amount is paid, but also when it is paid. In all computations one has to choose a fixed point in time as a reference point.

This is particularly important if several payments are made at different times. Usually the reference time is the present moment and then one speaks of the *present value*. It may also be the terminal time. For example, payments K_1 at the end of year t_1 (from now), K_2 at $t_2, \ldots, K_n$ at t_n, define a present value of

$$K = \sum_{i=1}^{n} K_i v^{t_i} \qquad \text{where} \qquad v = \left(1+\frac{p}{100}\right)^{-1}. \tag{3.3.3}$$

Such expressions are used quite often in the payments of mortgages or installment plans. An *annuity* is a series of payments made at certain specified intervals. Usually this is a constant amount and the intervals are regular.

Let us assume then an annuity A is always paid for n years at the beginning of the year. Then the present value of this payment is with $r = 1 + \frac{p}{100}$

$$K = A + Ar^{-1} + \cdots + Ar^{-(n-1)} = \sum_{i=0}^{n-1} Ar^{-i} = A\frac{1-r^{-n}}{1-r^{-1}}. \tag{3.3.4}$$

Example 3.4 You buy a house of value V. Since you have savings S, you have to take out a loan of $L = V - S$. This loan is paid in installment rates R at the end of each quarter for altogether n quarters. Then the present value of the loan L should be equal to the present value of all payments

$$\begin{aligned} L &= Rr^{-1} + Rr^{-2} + \ldots + Rr^{-n} \qquad r = \Big(1 + \frac{p}{400}\Big) \\ &= Rr^{-1}[1 + r^{-1} + \ldots + r^{-(n-1)}] = Rr^{-1}\frac{1 - r^{-n}}{1 - r^{-1}} = R\frac{r^n - 1}{r - 1}r^{-n}. \end{aligned}$$

If we choose the terminal time as the reference time, we get

$$Lr^n = Rr^{(n-1)} + Rr^{(n-2)}) + \ldots + Rr^{\circ} = R\frac{r^n - 1}{r - 1}.$$

Both equations are clearly equivalent.

Example 3.5 Sometimes banks require mortgages to be paid monthly, though the interest is only compounded yearly. What is the difference if the interest rate is $p = 8\%$, $R = 1,000$ and the payment is for 30 years? (This case actually was practised in Germany from 1899 until the high court ruled it illegal in 1986.)

With monthly compounding one has $K_m = 1,000r^{-1}\frac{1-r^{-360}}{1-r^{-1}} = 136,283.49$ where $r = 1 + \frac{8}{1,200}$. With yearly compounding one has

$$K_y = 12,000 \cdot 1.08^{-1}\frac{(1 - 1.08^{-30})}{(1 - 1.08^{-1})} = 135,093.40.$$

Thus the banks gain 1190.09.

The previous examples also show that it is necessary to specify when annuities are paid, at the beginning of the period or at the end. Premiums are usually collected at the beginning, while payments by insurances tend to be made at the end of the relevant periods.

Example 3.6 To save for their grandchild's education, the Clever family puts aside 50 Euro per month. This is put into a savings account with an annual interest rate of 5%. How much will grandchild Clever have at the age of 20 if the interest is compounded on a yearly basis?

The Clevers pay $12 \cdot 50 = 600$ per year. Thus the amount in the savings account is

$$600 \cdot 1.05^{19} + 600 \cdot 1.05^{18} + \cdots + 600 = 600 \cdot \frac{1.05^{20} - 1}{1.05 - 1} = 19,839.57.$$

Here one should observe that the interest for the payments in the first year works only for 19 years, ...

Example 3.7 Clever jr has just purchased a 150,000 Euro house and made a down payment of 30,000. He can amortize the balance in 30 years at 8%. What is his monthly payment if the interest is compounded yearly? How much interest does he have to pay and what equity does he have on his house after 20 years?

Solution: If the monthly payments are R then he pays $12 \cdot R$ per year. The present value of his payments is thus

$$K = 12 \cdot R[r^{-1} + \cdots + r^{-30}] = 12 \cdot Rr^{-1}[1 + \cdots + r^{-29}] = 12 \cdot R \cdot r^{-1}\frac{1 - r^{-30}}{1 - r^{-1}}.$$

This has to be equal to 120,000. Thus

$$120,000 = 12 \cdot R \cdot 1.08^{-1}\frac{1 - 1.08^{-30}}{1 - 1.08^{-1}} = 12R \cdot 11.25778334.$$

Thus

$$R = 888.274.$$

In these 30 years Clever has paid $30 \cdot 12 \cdot 888.274 = 319,778.64$. Hence he has paid a total of $319,778.64 - 120,000$ in interest. With 20 years of payment, the present value of his payments is

$$12 \cdot R\, 1.08^{-1}\frac{1 - 1.08^{-20}}{1 - 1.08^{-1}} = 104,654.46.$$

He therefore owes the bank 15,345.54 in present value. But in 20 years' time this has to be multiplied by $1.08^{20} = 4,660957$ or 71524.90.

Choosing a fixed reference point in time also allows us to compare payment schemes. If one has a scheme of paying K_1 at the end of year t_1, K_2 at the end of year $t_2 \ldots$, for a loan K the *effective rate of interest* is defined by (3.3.3) with p unknown.

Example 3.8 When one takes a loan from a bank, sometimes the bank will demand a *disagio*. Thus if the loan is L and the disagio $d\%$, the bank will actually pay you only $L \cdot (1 - \frac{d}{100})$. A disagio is thus the opposite of a cash discount. Assume the disagio is 2% for a loan which has to be paid back in 10 years at an interest rate of 6%. Assume R is the yearly payment. Then

$$L = R\frac{1 - 1.06^{-10}}{1 - 1.06^{-1}}.$$

But actually you have only received $L \cdot 0.98$. If p is the effective rate of interest, you get

$$.98L = R\frac{1 - (1 + \frac{p}{100})^{-10}}{1 - (1 + \frac{p}{100})^{-1}}.$$

Thus

$$.98 \cdot R \cdot \frac{1 - 1.06^{-10}}{1 - 1.06^{-1}} = R\frac{1 - (1 + \frac{p}{100})^{-10}}{1 - (1 + \frac{p}{100})^{-1}} \qquad \text{or}$$

$$7.64565843 = \frac{(1 - (1 + \frac{p}{100})^{-10})}{(1 - (1 + \frac{p}{100})^{-1})}.$$

One finds $p = 6.5365\%$ as the effective rate of interest.

Another thing which is usually not considered much in insurance mathematics is *inflation*. If the rate of inflation is $g\%$ the *purchasing power* of K in n years is only $K(1-\frac{g}{100})^n$. In this case money brought to the bank now with $p\%$ interest will have a purchasing power of $K(1+\frac{p}{100})^n(1-\frac{g}{100})^n$. With this all formulas above can be modified to cover inflation as well.

Retirement at 60

Since unemployment in the EU is rather high, some politicians in Germany have suggested that employees should already stop working at the age of 60. In order to induce many people to do this there should be no essential cut in the pension. To understand this proposal better let us compute how much of his income an employee has to put aside to finance a pension which begins x years earlier than usual. For this we will assume that this person begins his working life at the age of 20 and that people usually work until 65, when the government pension scheme sets in. We may set the yearly wages equal to 1 if we work with wage units. Let us assume this employee puts amount R aside each year and that the interest rate is 6%. For $x = 1$ this person will save R_1 each year, to be used during his 64th year. Thus we have with $r = 1,06$.

$$R_1[r^{43} + r^{42} + \cdots + r + 1] = 1$$

or

$$R_1 \frac{r^{44}-1}{r-1} = R_1 \cdot 199.758 = 1.$$

Thus $R_1 = 5.00606 \cdot 10^{-3}$ and he has to put about one half per cent of his income aside. If we assume a 1 % inflation, however, we get

$$R_1[r^{43} + \cdots + 1] = 1.01^{44}$$

or $R_1 = 7.756 \cdot 10^{-3}$. Even with a very moderate inflation rate three quarter per cent of the income will have to be saved. Now consider $x = 2$. Then as before

$$R_2[r^{42} + r^{41} + \cdots + r + 1] = 1 + \frac{1}{1,06}$$

or

$$R_2 \frac{r^{43}-1}{r-1} = R_2 \; 187.5076 = 1 + \frac{1}{1,06}.$$

This gives $R_2 = 1.03644 \cdot 10^{-2}$ or $R_2 = 1.5899 \cdot 10^{-2}$ with inflation.

You will be asked to work out the cases $x = 3$, 4 and 5.

P.S.: This proposal has been shelved meanwhile for obvious reasons.

Summary: The basic formulas for compound interest are derived. They also have many applications outside insurance.

Problems

1. 1 Euro is put into a bank with $p = 2\%(3\%)$ annually. How much is it after 2,000 (1,000) years? How much gold can you buy if the current price of gold is 7,730 Euro/kg? How big a sphere would this be and how would its radius grow at that time? (Hint.: The density of gold is 19.63 g/cm^3).

2. In the US bankers use the so-called 72-rule, which means that some money K doubles its value with an interest rate of $p\%$ roughly in $\frac{72}{p}$ years, when compound interest is used. Motivate this approximation. How good is it?

3. In order to avoid working with real exponents one sometimes uses the formula
$$K_t = K\left(1+\frac{p}{100}\right)^{[t]}\left(1+(t-[t])\frac{p}{100}\right),$$
where $[t]$ is the largest integer $\leq t$. How good is this?

4. Principal K is put into a bank for n years. What do you get if the interest is compounded yearly (monthly, weekly, daily, hourly, ...)?

5. How long does it take for 1,500 Euro to become 10,000 Euro at $p = 4\%$ ($p = 5\%$)?

6. What percentage of interest is used if a purchase of 1,200 Euro is paid in 13 monthly installments of 100 Euro?

7. The lifetime of a machine of value 20,000 Euro is 8 years. How much of it can be deducted yearly?

8. For 20 years you have a monthly annuity with inflation correction, i.e. the payment in the n-th month is $R(1-\frac{g}{1{,}200})^{-(n-1)}$. Compute the present value.

9. What would you get in Example 5 if the interest is compounded quarterly or monthly.

10. Clever jr. in Example 6 expects a pay rise of 2% each year. Correspondingly he is willing to let R increase by 2% each year. What is R in this case?

11. Use the same initial R as in Example 6, but assume that R grows by 2% each year, i.e. $R_k = R\cdot(1.02)^{k-1}$. When is the mortgage then amortized?

12. Finish the computations in the example "Retirement at 60" and consider also the case of a 1.5% inflation.

13. Mr. V. Bisee has the option to get a loan L either from bank A or bank B.

    The conditions for bank A are: Payment of 100,000 after 1 quarter. Then 50,000 after another quarter and 180,000 after two further quarters.

    Bank B demands 200,000 after one year, 80,000 after another half year and 90,000 after three further quarters.

    Which of the loans is better if an interest rate of 4% per quarter and quarterly compounding holds?

14. Prove the following formula: A loan L with a constant yearly annuity A is always paid at the end of the period. Then the remaining debt after n payments is
$$K_n = Lr^n - A\frac{r^n - 1}{r - 1} \qquad r = \left(1 + \frac{p}{100}\right).$$

15. Determine the effective rate of interest for a loan of 108,000 which is repaid with an annuity of 75,000 in the first and second year.

16. You want to buy a car for 50,000. If you pay cash the dealer offers a discount of 3%, and you can borrow 50,000 at 6% from your local bank.

    But the dealer also has connections to a finance company which charges only 5%. But then you do not get a discount. Which offer should you choose if the repayment is made at constant rates over 5 years?

17. Here is another clever idea of a politician. (This was actually published in a newspaper based on an interview). In order to further the formation of capital for working people one should allow for mortages which run for 100 years. To make this precise compare the repayment of 100,000 in 30 and 100 years, respectively at $p = 6$ or 7% and annual compounding of interest. What is the monthly payment in all cases? Compare. What do you think of it?

3.4 Mortality

The only element of chance in life insurance is the probability of death or mortality q. More precisely, we denote the probability that a person of age x dies within the next year by q_x. Since q_x varies smoothly and slowly with age x, it suffices to define q_x by the list $(q_0, q_1, q_2, \ldots, q_\omega) = (q_x)_0^\omega$, the *mortality table*, where x is measured in years and where ω is some general upper limit on human life, usually $\omega = 100$.

The (q_x) depend on many external factors like gender, status of health, occupation, race, social environment, education and family situation. For this reason the (q_x) are generally determined for groups where these factors agree to some extent. Otherwise the (q_x) will have to be interpreted as averaged over these properties. But it is important to note that the (q_x) only make sense with respect to particular groups of people. These groups also have to be fairly large because, in principle at least, probabilities are defined with respect to infinite entities. As groups one usually has all males (females) in the population of a country or all males (females) in the portfolio of a particular life insurance company. But if an insurance company wants to come up with a particular policy for non-smokers, it will have to consider groups of smokers and groups of non-smokers.

Mortality tables for males (females) are usually published by a government office devoted to the statistical analysis of the population. Governments need such data for all sorts of planning purposes.

Mortalities also depend on time, because medical services and living conditions improve constantly. After the reunification there was a slight increase in mortality in Germany though. In principle, mortality could be defined for a particular age cohort, e.g. for the group of all people born in 1930 (generation table). But then it would take a whole lifetime to determine the mortality. Then, however, the mortality table would be outdated. So nowadays mortalities are measured over a particular time period, which for populations is usually a year around a census. In general, however, one should distinguish between generation mortality tables and period tables. For insurance companies they could in principle be defined on a yearly basis.

Problem 3.2 List a number of domains for a government for which the mortality table of a population is needed.

The mortality tables of companies usually differ slightly from the general mortality table, because

1. Life insurance companies usually demand a health check before issuing a policy. How would this influence the (q_x)? It is known, however, that this bias wears off after 6 or 7 years.

2. People with health problems are more inclined to buy a life insurance than healthy ones. What bias does this introduce?

3. The particular life insurance might be connected to a particular occupation. What qualitative differences do you expect for the mortality of miners and clerks?

Now we need a little theory in order to relate the (q_x) to some other quantities. To do so we fix a group of people for which we want to analyse the mortality. As above, x will always be measured in years. Then let p_x denote the probability that a person in this group will become older than x years. The list $(p_x)_{x=1}^{\omega}$ is usually called the *life table*, though in general one defines $\ell_0 = 100,000$ and $\ell_x = \ell_0 \cdot p_x$ and calls $(\ell_x)_{x=0}^{\omega}$ the life table. Here ω is the terminal age for the life table. Usually ω is 100 or even 110 years. Life tables published by government statistics bureaus are usually given separately for men and women. In such a set up ℓ_x is then the number of persons of a fictitious group of ℓ_0 people who reach an age of at least x. q_x is the conditional probability that an x-year-old person dies. Thus

$$q_x = \frac{(\ell_x - \ell_{x+1})}{\ell_x} = \frac{(p_{x+1} - p_x)}{p_x} \tag{3.4.5}$$

because the denominator is the number of people who reach at least the age of x, but do not reach $x+1$. So the quotient just gives the probability. From (3.4.5) one gets

$$\ell_0 = 100,000, \qquad \ell_{x+1} = \ell_x(1 - q_x). \tag{3.4.6}$$

The second formula means: All people of age x who do not die within the next year reach $x+1$ years. (3.4.6) shows that the (ℓ_x) can be computed recursively once the (q_x) are given. Conversely if the (ℓ_x) are given, the (q_x) can be computed by (3.4.5). Since the (q_x) depend not only on the group of people for which they are defined, but also vary in time, they have to be determined from actual data. For this one proceeds with the following steps

1. Estimation of the raw mortalities.
2. Smoothing and fitting.
3. Correcting.

Appendix

Lifetime: We consider the random variable *lifetime* X or more precisely age at death. X can be the lifetime of a machine, a class of animals or of a group of people. Then the underlying probability space is $[0,\infty)$ if the lifetime is a continuous variable. If X only takes discrete values, years only, for example, the probability space is $\{1,2,3,\dots\}$. In the first case one usually assumes that X has a density f. In this case the probability that X is less than x is given by

$$P([X\leq x]) = \int\limits_0^x f(t)dt = F(x), \tag{3.4.7}$$

where $f(t)\geq 0$ and $\int_0^\infty f(t)dt = 1$. $1-F=S$ is called the survival function. In the discrete case this becomes

$$P(X\leq x) = \sum_{t=0}^{x} f_t, \tag{3.4.8}$$

where $f_t\geq 0$ and $\sum_{t=0}^\infty f_t = 1$. In the discrete case the mortality q_x was defined as

$$q_x = \frac{P(x<X\leq x+1)}{P(X>x)} = \frac{f_{x+1}}{\sum\limits_{t=x+1}^{\infty} f_t}$$

and this agrees with the above definition. In the continuous model there is no natural unit of time. Thus one defines

$$\mu(x) = q(x) = \lim_{h\to 0}\frac{P(x\leq X\leq x+h)}{hP(x\leq X)}. \tag{3.4.9}$$

q_x is then called the *force of mortality* or simply mortality. In order to relate this to f we note $F'=f$. Then (3.4.9) implies

$$q(x) = \lim\frac{F(x+h)-F(x)}{h(1-F)} = \frac{F'}{1-F}(x) = -\frac{(1-F)'(x)}{1-F} = -\frac{d}{dx}\ln(1-F(x)).$$

Thus

$$\ln(1-F(x)) = -\int\limits_0^x q(t)dt$$

or

$$1-F(x) = \exp -\int\limits_0^x q(t)dt.$$

This shows

$$F(x) = 1 - \exp\left(-\int_0^x q(t)dt\right) \tag{3.4.10}$$

and

$$f(x) = F'(x) = q(x)\exp\left(-\int_0^x q(t)dt\right). \tag{3.4.11}$$

This "continuous" description of lifetime is used in particular for the description of the lifetimes of machines to describe the latency of a sickness. In these circumstances $\mu(x) = q(x)$ is also called the *failure rate* or *hazard rate*. The continuous model has the advantage that many computations can be performed more easily if the $q(x)$ have a simple functional form. For practically oriented insurances the discrete $(q_x)_{x=0}^{\omega}$ are more important.

Problems:

1. (a) Let $q(x)$ be the force of mortality. Then

   $${}_tp_x = e^{-\int_x^{x+t}} q(s)ds$$

   is the probability that an x-year-old reaches the age $x + t$.

   (b) Then ${}_tp_x q(x + t)dt$ is the probability that an x-year-old dies between t and $t + dt$ (from now).

   (c) Show that ${}_tp_x q(x)$ describes a probability distribution on $[x, \omega)$. Express it as in 3.4.11.

2. Show

   $$\frac{d}{dt}(1 -_t p_x) =_t p_x q(x + t).$$

3. On the basis of the standard life table for the population of the US (1979–81) determine the probability that a 30-year-old will

   (a) get older than 90.

   (b) die between 60 and 70.

   (c) die before the age of 40.

The construction of life tables was introduced by J. Graunt. His table, however, was based on general observations. Below is a life table for women in England and Wales around 1841. Here $\ell_0 = 4873$. Later a procedure was found to determine the q_x on the basis of population counts.

Extract from a Life Table

Age x	Number Living ℓ_x	Number Dying before $(x+1)$ d_x	Probability of Death q_x
0	4873	646	0.133
1	4227	255	0.060
2	3972	134	0.034
3	3838	90	0.023
4	3748	66	0.018
10	3505	21	0.006
30	2986	30	0.010
50	2324	36	0.015
70	1271	77	0.061
90	66	18	0.273

Problems

1. Estimate the mortalities q_3, q_{30}, q_{50}.
2. Compare these mortalities with those from the German mortality table.
3. If one compares the mortality table from 1980 and 1990 what kind of qualitative difference does one expect?
4. Compare the mortality of men and women.
5. Compare two mortality tables from your own country which are about 10 years apart.
6. Which distribution do you get if you choose $q(x) = c = \text{const.}$ in the continuous model.
7. More generally, consider the case $q(x) = \alpha x^\beta$, which leads to the Weibull distribution.
8. Determine the mortality for Lagrange's mortality law $\ell(x) = \frac{\ell_0}{86}(86 - x)$.
9. If $(\ell_x)_0^{100}$ is a life table, how can one generate a life table for people of age $x \geq 20$.
10. Why would one expect the mortality (q_x) to grow with x? Yet, the German and American table contains a "hump" in the early twenties. What do you think causes it?
11. If you look at the US table, what is the chance of people getting older than 65? What is the chance of a 20-year-old getting older than 60?

Below is the general mortality table for Germany based on the population count from 1960. The table lists $1{,}000q_x$ for men and women

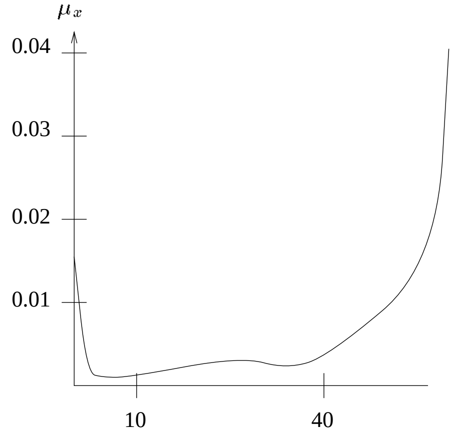
Life table 1960/62 (Germany) List of $10^3 \cdot qx$

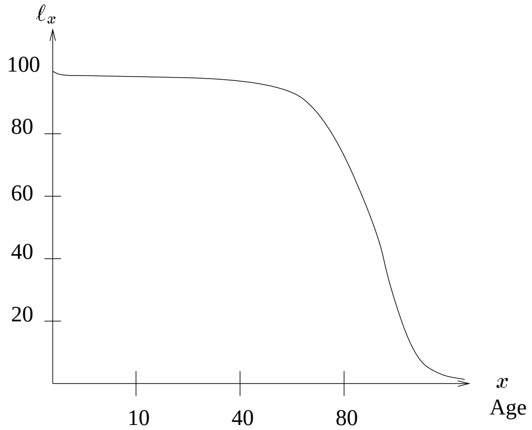
Age	Men	Women	mod.	Age	Men	Women	mod.
0	35.33	27.78	35.83	51	8.25	4.81	9.24
1	2.31	2.01	2.81	52	9.24	5.21	10.35
2	1.40	1.08	1.90	53	10.35	5.66	11.59
3	1.00	0.79	1.50	54	11.59	6.16	12.97
4	0.87	0.65	1.37	55	12.97	6.72	14.49
5	0.80	0.56	1.30	56	14.49	7.35	16.16
6	0.73	0.48	1.23	57	16.16	8.07	17.98
7	0.64	0.40	1.14	58	17.98	8.89	19.95
8	0.56	0.35	1.06	59	19.94	9.81	22.04
9	0.49	0.31	0.99	60	22.04	10.85	24.27
10	0.45	0.28	0.95	61	24.27	12.04	26.80
11	0.43	0.28	0.93	62	26.61	13.38	29.08
12	0.45	0.28	0.95	63	29.07	14.91	31.64
13	0.50	0.31	1.00	64	31.64	16.65	34.33
14	0.60	0.35	1.10	65	34.33	18.82	37.18
15	0.75	0.40	1.25	66	37.17	20.83	40.18
16	0.95	0.46	1.45	67	40.19	23.31	43.43
17	1.19	0.52	1.69	68	43.43	26.11	46.96
18	1.46	0.57	1.96	69	46.96	29.27	50.88
19	1.69	0.60	2.19	70	50.87	32.85	55.26
20	1.85	0.62	2.35	71	55.26	36.95	60.22
21	1.90	0.62	2.40	72	60.22	41.62	65.81
22	1.87	0.62	2.37	73	65.82	46.94	72.04
23	1.80	0.63	2.30	74	72.04	52.94	78.86
24	1.72	0.67	2.22	75	78.85	59.61	86.23
25	1.69	0.73	2.19	76	86.22	66.92	94.15
26	1.66	0.78	2.16	77	94.16	74.90	102.83
27	1.66	0.83	2.16	78	102.82	83.60	112.40
28	1.66	0.89	2.16	79	112.41	93.06	122.97
29	1.68	0.94	2.19	80	122.97	103.31	134.49
30	1.70	0.99	2.20	81	134.48	114.40	146.84
31	1.74	1.05	2.24	82	146.85	126.31	159.94
32	1.80	1.11	2.30	83	159.93	138.96	173.58
33	1.88	1.19	2.38	84	173.60	152.30	188.00
34	1.98	1.28	2.49	85	188.02	166.26	203.35
35	2.09	1.38	2.59	86	203.34	180.79	220.18
36	2.22	1.50	2.72	87	220.15	196.23	238.66
37	2.38	1.61	2.88	88	238.71	212.88	258.79
38	2.56	1.74	3.06	89	258.70	230.93	279.31
39	2.75	1.87	3.26	90	279.21	248.21	297.99
40	2.95	2.01	3.44	91	297.95	265.25	316.41
41	3.16	2.16	3.66	92	316.23	281.88	334.29
42	3.40	2.33	3.90	93	333.84	297.91	350.50
43	3.68	2.52	4.18	94	350.61	313.17	365.64
44	4.02	2.74	4.52	95	366.38	327.50	381.94
45	4.43	2.99	4.93	96	380.99	340.75	393.26
46	4.90	3.26	5.42	97	394.31	352.79	407.41
47	5.42	3.54	6.01	98	406.23	363.48	421.88
48	6.00	3.82	6.64	99	416.63	372.74	432.43
49	6.65	4.13	7.89	100	425.43	380.47	—
50	7.39	4.55	8.26				

The following life table of the US is based on the census around 1980

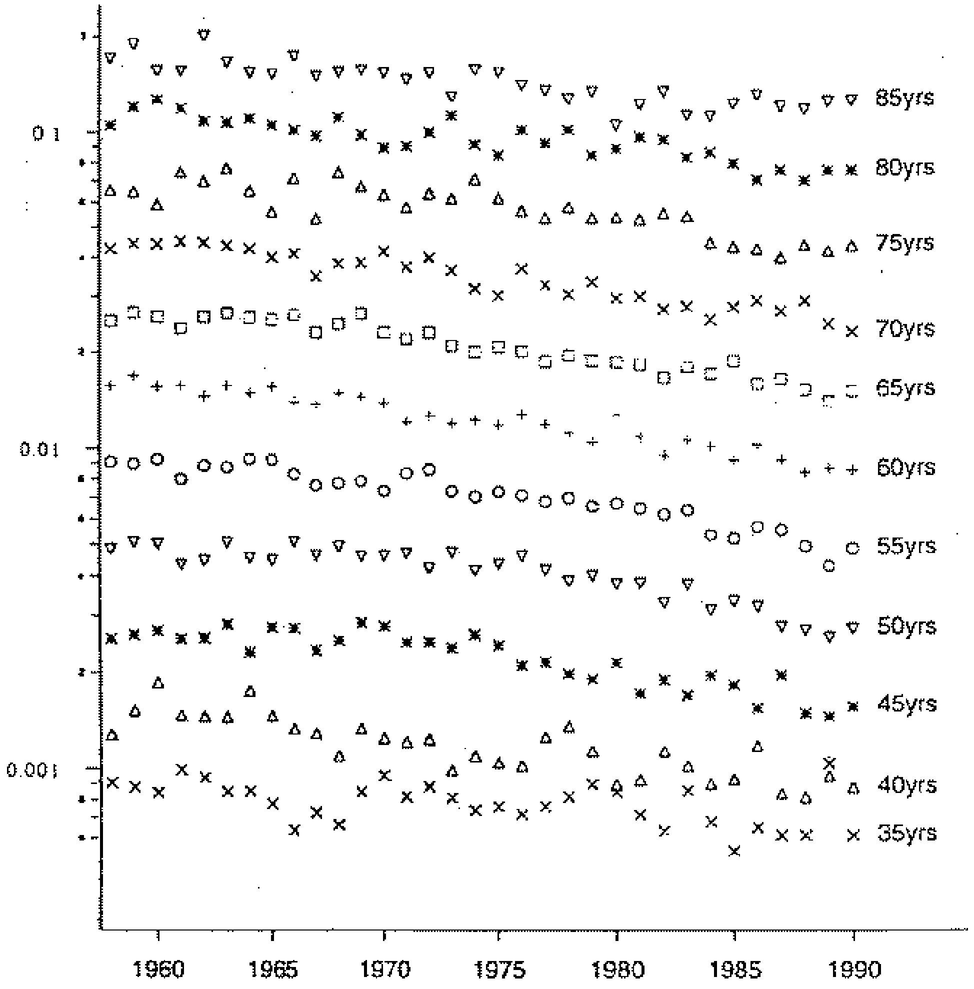
Years	q_x	ℓ_x	Years	q_x	ℓ_x
0-1	0.01260	100 000	55-56	0.00902	88 348
1-2	0.00093	98 740	56-57	0.00978	87 551
2-3	0.00065	98 648	57-58	0.01059	86 695
3-4	0.00050	98 584	58-59	0.01151	85 776
4-5	0.00040	98 535	59-60	0.01254	84789
5-6	0.00037	98 495	60-61	0.01368	83 726
6-7	0.00033	98 459	61-62	0.01493	82 581
7-8	0.00030	98 426	62-63	0.01628	81 348
8-9	0.00027	98 396	63-64	0.01767	80 024
9-10	0.00023	98 370	64-65	0.01911	78 609
10-11	0.00020	98 347	65-66	0.02059	77 107
11-12	0.00019	98 328	66-67	0.02216	75 520
12-13	0.00025	98 309	67-68	0.02389	73846
13-14	0.00037	98 285	68-69	0.02585	72 082
14-15	0.00053	98 248	69-70	0.02806	70 218
15-16	0.00069	98 196	70-71	0.03052	68 248
16-17	0.00083	98 129	71-72	0.03315	66 165
17-18	0.00095	98 047	72-73	0.03593	63 972
18-19	0.00105	97 953	73-74	0.03882	61 673
19-20	0.00112	97 851	74-75	0.04184	59 279
20-21	0.00120	97 741	75-76	0.04507	56 799
21-22	0.00127	97 623	76-77	0.04867	54 239
22-23	0.00132	97 499	77-78	0.05274	51 599
23-24	0.00134	97 370	78-79	0.05742	48 878
24-25	0.00133	97 240	79-80	0.06277	46 071
25-26	0.00132	97 110	80-81	0.06882	43 180
26-27	0.00131	96 982	81-82	0.07552	40 208
27-28	0.00130	96 856	82-83	0.08278	37 172
28-29	0.00130	96 730	83-84	0.09041	34 095
29-30	0.00131	96 604	84-85	0.09842	31 012
30-31	0.00133	96 477	85-86	0.10725	27 960
31-32	0.00134	96 350	86-87	0.11712	24 961
32-33	0.00137	96 220	87-88	0.12717	22 038
33-34	0.00142	96 088	88-89	0.13708	19 235
34-35	0.00150	95 951	89-90	0.14728	16 598
35-36	0.00159	95 808	90-91	0.15868	14 154
36-37	0.00170	95 655	91-92	0.17169	11 908
37-38	0.00183	95 492	92-93	0.18570	9 863
38-39	0.00197	95 317	93-94	0.20023	8 032

39-40	0.00213	95 129	94-95	0.21495	6 424
40-41	0.00232	94 926	95-96	0.22976	5 043
41-42	0.00254	94 706	96-97	0.24338	3 884
42-43	0.00279	94 465	97-98	0.25637	2 939
43-44	0.00306	94 201	98-99	0.26868	2 185
44-45	0.00335	93 913	99-100	0.28030	1 598
45-46	0.00366	93 599	100-101	0.29120	1 150
46-47	0.00401	93 256	101-102	0.30139	815
47-48	0.00442	92 882	102-103	0.31089	570
48-49	0.00488	92 472	103-104	0.31970	393
49-50	0.00538	92 021	104-105	0.32786	267
50-51	0.00589	91 526	105-106	0.33539	179
51-52	0.00642	90 986	106-107	0.34233	119
52-53	0.00699	90 402	107-108	0.34870	78
53-54	0.00761	89 771	108-109	0.35453	51
54-55	0.00830	89 087	109-110	0.35988	33





The problems with the mortalities are actually not as simple as they seem to appear with (3.4.5) or (3.4.6), because in reality mortality depends on time. In recent years mortalities have been decreasing steadily (see Figure) because health service and living conditions have improved. Then we should really write $q(x, t)$. The difficulty that arises now consists of the fact that the mortality of an x-year-old in n years is really $q(x + n, \underline{t + n})$ if the present moment is t. If one follows a particular cohort (generation) of people, all of the same age, through their lives, and determines the morality, one obtains a lifetable which is called a *generation life table.* If one determines the mortality as a function of age during a particular period, usually the year of a census, one gets a *period lifetable.* For practical reasons insurance companies use period life tables.



. Log crude mortality rate vs. period, duration 5+, ages 35, 40, ..., 85

(From Renshaw et al. The modelling of recent mortality trends in the United Kingdom BAJ2 1996)

So far insurance companies mainly base their calculations on time independent mortality tables, which are updated every 7 to 10 years. However, it is possible to incorporate the t-dependence by extrapolation. For this one uses several life tables determined at different times

$$(q(x,t_1))_{x=0}^{\omega},\ (q(x,t_2))_{x=0}^{\omega}, \dots ,(q(x,t_K))_{x=0}^{\omega}$$

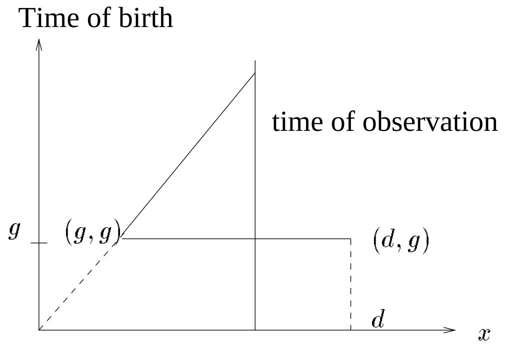
and then makes the ansatz

$$q(x,t) = q_x + h(x)g(t), \tag{3.4.12}$$

where h, g are suitable functions which can be fitted to the data. In the simplest case one uses linear interpolation and chooses $g(t) = (t - t_1)$. This is quite acceptable in most cases, because $q(x,t)$ changes only very gradually with t. It is very likely that future premiums for insurances of very long duration might be based on mortality tables of the form $(q(x,t))_{x=0}^{\omega}$. This is currently only done for pensions.

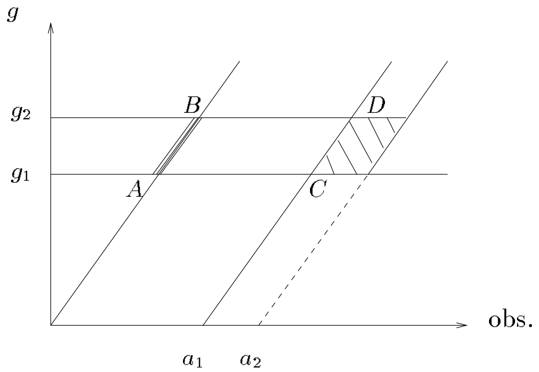
Even though the sample for the life table of a whole population is in general much larger than the stock of an insurance company, companies will in general determine their own mortality tables. Why? The following methods are employed in order to determine the life table of the population of a country.

Method of year of death



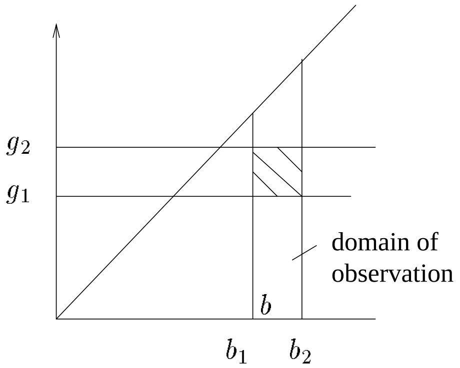
Each member of the population is represented by its lifeline, which starts at (g,g), g = time of birth and ends at (d,g), d = time of death running parallel to the time of observation axis.

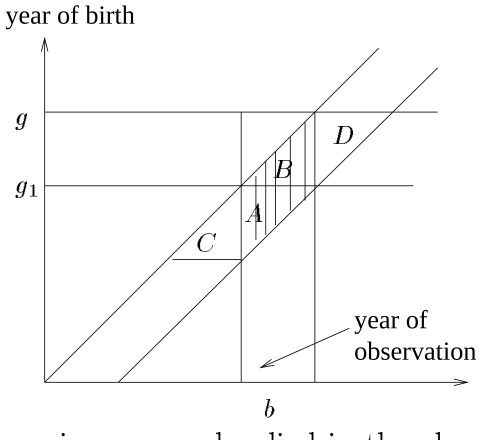
The number of all people alive at the time of observation b is then the number of lifelines intersecting the line $x = b$.



The number of all people born between g_1 and g_2 is the number of lines starting on the segment AB. Such a person gets older than a_1 years of age if his lifeline cuts the segment CD. It lies between age a_1 and a_2 if his lifeline ends in the shaded domain.

The difficulty, however, is that in general one only counts the number of people who have died within the domain of observation b.





Now let $M(g,x,b)$ = number of people born in year g who died in the observation interval b at age x. Then since $M(g-1,x,b) \approx M(g-1,x,b-1)$, $A \approx C$ and $M(g,x,b) \approx M(g,x,b+1)$, $B \approx D$ we see:
Number of people who died in b at age $x = D_x \approx \frac{1}{2}[M(g,x,b) + M(g-1,x,b) + M(g,x,b+1) + M(g-1,x,b-1)]$ and thus

$$\hat{q}_x = \frac{D_x}{\text{Average number of age } x + \frac{1}{2}D_x}. \tag{3.4.13}$$

This averaging is not as bad as it looks, since the data have yet to be smoothed over.

Another method is the method of the *year of birth*

$$\hat{q}'_x = \frac{\text{Number of people of age } x \text{ born in year } g \text{ and who die in obs. interv. } b}{\text{Number of people of age } x \text{ born in } g}. \tag{3.4.14}$$

An even simpler method to estimate the death rate is

$$\hat{q}''_x = \frac{\text{Number of people of age } x \text{ who die during the observ. time}}{\frac{1}{2}\,[\text{Number of people of age } x \text{ who are still alive at the beginning of } b + \text{number of people of age } x \text{ who are still alive at the end of } b]}. \tag{3.4.15}$$

Of course there are other methods, however, since these are the most well known ones, we will not dwell any further on measurements of $\hat{q}_x$.

Closely connected with the definition of the $\hat{q}_x$ is the age x, which is used in insurance policies. Possible definitions are

1. A person is of age x if he has completed his x-th year, but not the $(x+1)$st.
2. A person who exceeds his x-th birthday by more (less) than 6 months has age $x+1(x)$.
3. Policy age = year of the beginning of policy − year of birth.

It should be noted that in general the mortality of the portfolio of an insurance will differ slightly from that of the general population. This bias may be due to the fact the insurances screen for healthy people. This was already noted with the earliest life tables. One also calls this select mortalities. Another reason is due to the fact that more people with a questionable health status will look for an insurance. Finally, mortality depends on the country or occupation involved. Nonetheless, the mortality functions (q_x) look very much alike. From this point of view it should be possible to improve the quality of the data.

3.5 Smoothing the mortalities

Once the raw $\hat{q}_x$ have been determined, they have to be smoothed and evened out, because the fluctuation of each individual q_x is still of the order $\sqrt{q_x(1-q_x)/L_x}$, L_x = number of persons of age x. As an example consider the following situation. An insurance company has about 10^5 policies in its portfolio. This means $L_x \approx 3,000$. With $q_x \approx .001$ the error is still of the order $.001 \cdot 1/\sqrt{3}$, which is about 60% of the q_x. Fluctuations are due to chance, but also because death occurs unevenly throughout a year. Usually the death rate is higher in winter. But also an epidemic like influenza may take its toll.

Standard methods for smoothing the data are

1. Graphical interpolation: This means one draws a closely fitting smooth curve through the data points. Since this method is rather arbitrary and does not readily yield numbers, it is not used for insurance companies.

2. Sliding weighted mean

$$\hat{q}_x = \sum_{k=-r}^{r} a_k \hat{q}_{x+k}, \quad \sum_{k=-r}^{r} a_k = 1. \tag{3.5.16}$$

Here are some examples:

Wittstein: $q_x = \frac{1}{5} \sum_{K=-2}^{2} \hat{q}_{x+k}$

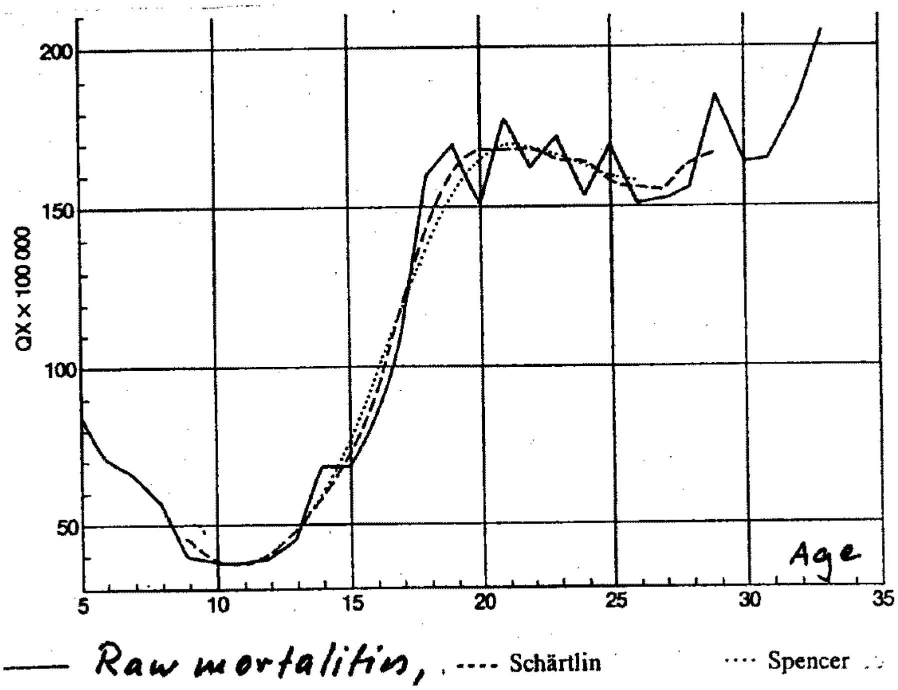
Finlaison: $q_x = \frac{1}{25}(5\hat{q}_x + 4\hat{q}_{x\pm1} + 3q_{x\pm2} + 2q_{x\pm3} + x_{\pm4})$

Schärtlin: $\hat{q}_x = \frac{1}{27}(9\hat{q}_x + 8\hat{q}_{x\pm1} + 2q_{x\pm2} - q_{x\pm4})$

Spencer: $q_x = \frac{1}{320}(74\hat{q}_x + 67\hat{q}_{x\pm1} + 46\hat{q}_{x\pm2} + 21\hat{q}_{x\pm3} + 3\hat{q}_{x\pm4}$

$q_{x\pm k}$ means $q_{x+k} + q_{x-k}$. Examples of these interpolations can be seen in the following plot

(Smoothing raw mortalities by sliding averages)



It is rather difficult to judge the quality of these approximates and a good deal of their use depends on belief and confidence. One criterion used is the exactness on simple functions, like polynomials, e.g. the Schärtlin and Spencer method are exact on cubic polynomials. In the plots you can see for yourself how these methods operate. But one should also keep in mind that there is no optimal method. Needless to say, there are many more such convolution weighted mean operators. There is also a systematic way of deriving such formulae, but this goes far beyond the need in insurance mathematics. The disadvantage with these convolution means is that they do not yield values for the q_x for small and large x. The values at the end of the table, however, may be determined by extrapolation. More on the weighted average can be found in B. Benjamin, F.H. Pollard: The Analysis of Mortality and Other Actuarial Statistics, Butterworth, Heinemann 1992.

3. The method of Whittaker Henderson

   This method uses the idea that the q_x should depend smoothly on x. For this reason one defines the difference operator $\triangle^s$ of order s by

$$\triangle^s \hat{q}_{x_0+k} = \sum_0^s (-1)^v \binom{s}{v} \hat{q}_{x_0+k+v}.$$

   Then the q_x are chosen such that the expression

$$\sum_{k=0}^{n} w_k(\hat{q}_{x_0+k} - q_{x_0+k})^2 + g \sum_{k=0}^{n-s} (\triangle^s q_{x_0+k})^2 \to \min$$

   is minimized. Here the w_k are weight factors with $\sum w_k = 1$ and g is a measure for the required smoothness. g and the w_k can be chosen freely. [F.Y. Chan, et al., Properties and modification of Whittaker, Henderson graduation, Scand. Actuarial J. 1982].

4. The method of splines

   This technique is usually dealt with in mathematics courses on numerical analysis. Thus we will only give a rough idea of this method, which goes back to Reinsch [Smoothing by Spline functions I, II Num. Math. **10** 1967, 177–183; **16** 1970, 451–454]. Given $(\hat{q}_x)_{x=0}^{\omega}$ and the corresponding standard deviations σ_x. ω is here the largest considered age, e.g. $\omega = 100$. Find a smooth function g, which is twice continuously differentiable such that

$$\sum_{x=0}^{\omega} \left(\frac{g(x) - \hat{q}_x}{\sigma_x} \right)^2 \le S \qquad \text{with}$$

$$\int_0^w g''(x)^2 dx \to \min .$$

   The corresponding Euler-Lagrange equation shows that

$$g(x) = a_i + b_i(x-i) + c(x-i)^2 + d_i(x-i)^3$$

   is a cubic spline. The coefficients satisfy $c_0 = c_\omega = 0$ and for $i < \omega$ $d_i = (c_{i+1} - c_i)/3$ $b_i = a_{i+1} - a_i - c_i - d_i$, but the (a_i) and (c_i) still have to be determined from the $\hat{q}_x$.

5. The methods of Woolhouse and Karup.

This method consists of two steps. In the first step one finds for each x a cubic polynomial p_x through the points $(x-5,\ \hat{q}_{x-5})$, $(x,\ \hat{q}_x)$, $(x+5,\ \hat{q}_{x+5})$ and $(x+10,\ \hat{q}_{x+10})$. In the second step one sets

$$q_x = \frac{1}{5}(p_{x\pm 2}(x) + p_{x\pm 1}(x) + p_x(x)).$$

As you can imagine, there are a host of further interpolation and smoothing methods along these lines. Their credibility is quite difficult to assess, however, in particular since the "real" q_x are unknown. As a final method we will consider

6. Mortality laws

   In this case one starts with a class of functions parametrized by $a_1, a_2, \dots, a_k$. If $f(x,a)$ is such a function, the problem is: Choose the parameters $(a_1, \dots, a_k) = a$ such that $f(x,a)$ is a "good fit" for (q_x). Here of course it is still unclear what a "good fit" means. One mathematical rather convenient method is to choose the parameters $a_1, \dots, a_k$ such that

   $$\sum_{x=0}^{w} \frac{(f(x,a) - \hat{q}_x)^2}{\sigma_x^2} \to \min$$

   is minimal. This type of approach is also encountered in statistics under the heading of "nonlinear regression". There are also other methods of defining a good fit, e.g. the method of moments. But we will leave it at that.

   The simplest, but not the most clever way is to choose f, a polynomial, and optimize the fit over suitably small intervals. Quadratic or cubic polynomials over intervals of length 5 or 6 are common. In this way the whole domain $0 \le x \le w$ is made up of patches where the $\hat{q}_x$ are smoothed out.

   A more clever way to smoothen the $\hat{q}_x$ is to choose functions which are better designed to allow a fit over larger domains. These are the mortality laws, which will be dealt with in the next section.

   With the above techniques the methods of smoothing the $\hat{q}_x$ are by far not exhausted, and numerical mathematics and statistics provide further methods for smoothing. In addition, the above methods can be applied to suitable transforms of the q_x. In this connection $\ln q_x$ and (ℓ_x) are most commonly used.

Once one has determined the $(q_x)_{x=0}^{\omega}$, ω the terminal age, from the raw mortalities $(\hat{q}_x)_{x=0}^{\omega}$, one still has to assess the validity of these values and judge the fit.
The obvious requirement

$$\hat{q}_x - c\frac{\sqrt{\hat{q}_x(1-\hat{q}_x)}}{\sqrt{L_x}} \le q_x \le \hat{q}_x + c\frac{\sqrt{\hat{q}_x(1-\hat{q}_x)}}{\sqrt{L_x}} \tag{3.5.17}$$

is in general not very stringent, even for large portfolios.

The *smoothness* of the q_x can be judged by considering

$$\sum_{x=2}^{\omega-2}(\triangle^2 q_x)^2,$$

where $\triangle^2$ is the second difference operator. Roughly speaking this expression is a measure for

$$\int_0^{\omega} q''(x)dx.$$

The goodness of fit can be estimated by expressions

$$\sum_{x=0}^{\omega}(\hat{q}_x - q_x)^2 = Q \tag{3.5.18}$$

or weighted versions of this, like

$$\sum_{x=0}^{\omega}(\hat{q}_x - q_x)^2 w_x = Q_w. \tag{3.5.19}$$

The χ^2 test in statistics leads to a similar expression with $w_x = L_x/q_x$ where, for example, the weight factor w_x is proportional to L_x, the number of persons in the class of x-year-olds.

In the method of *moments* one checks whether the first m moments of $(\hat{q}_x)$ and (q_x) agree. Here the k-th moment of (q_x) is defined by

$$M_k = \sum_{x=0}^{\omega} x^k q_x. \tag{3.5.20}$$

One could also apply standard statistical tests to $(q_x - \hat{q}_x)$ to test for trend or randomness. A simple example for this would be the sign test, where one considers

$$n_+ = \text{ the number of } x \text{ with } q_x > \hat{q}_x.$$

Then one would expect n_+ to be approximately $\omega/2$. Thus if $|n_+ - \omega/2|$ is too large the fit would not be acceptable.

Despite the large numbers of smoothing methods and the multitude of methods to assess the quality of fit, there seems to be no method in the literature based on actual work in insurance. By this we mean that criteria to assess the goodness of fit should be based on criteria derived from applications of the mortality laws. Here are some simple proposals for this:

1. If the majority of all life insurance policies are in the age bracket [25, 70] the accuracy of fit outside this range is of lesser importance.

2. If all life insurance policies are of a certain type, e.g. term insurances from 25 to 65, the set-up of these policies should lead to an assessment criterion.

3. Since the majority of all contracts are for a duration of more than 20 years, the tests should be based on averaging over such intervals and consider ${}_tq_x$ for $t \geq 20$, where $1 -_t q_x = \prod_0^t (1 - q_{x+v})$ intervals.

4. A standard method to assess the fit is to apply it to model contracts. This could still be combined with bootstrapping.

Problems

1. Below is the raw and normalized life table of the Californian population [Chin Long Chong, The Life Table and its Applications Florida 1984]. Using $\hat{\ell}_0 = 100,000$ and $\ell_{x+1} = \ell_x(1 - q_x)$ show how one arrives at the normalized table by computing the first 10 ℓ_x.

Raw Mortalities for the Californian Population

x	l_x	D_x	9_x	x	l_x	D_x	9_x
0	340483	6234	.01801	36	224302	475	.00212
1	326154	368	.00113	37	228652	519	.00227
2	313699	269	.00086	38	226727	549	.00242
3	323441	237	.00073	39	235980	606	.00256
4	338904	175	.00052	40	249027	665	.00267
5	362161	179	.00049	41	232893	719	.00308
6	379642	171	.00045	42	239747	863	.00359
7	386980	131	.00034	43	238783	874	.00365
8	391610	121	.00031	44	248100	993	.00399
9	397724	121	.00030	45	253828	1140	.00448
10	406118	126	.00031	46	249857	1268	.00506
11	388927	127	.00033	47	247955	1362	.00548
12	395025	138	.00035	48	252137	1422	.00562
13	388526	158	.00041	49	242126	1530	.00630
14	385085	186	.00048	50	243799	1594	.00652
15	377127	235	.00062	51	220599	1710	.00772
16	368156	344	.00093	52	213448	1793	.00837
17	366198	385	.00105	53	302618	1870	.00914
18	354932	506	.00142	54	202388	1981	.00974
19	350966	584	.00166	55	201750	2217	.01093
20	359833	583	.00162	56	193828	2333	.01196
21	349557	562	.00161	57	187257	2483	.01317
22	365839	572	.00156	58	178602	2392	.01330
23	370548	564	.00152	59	171807	2517	.01454
24	295189	421	.00143	60	174613	2733	.01553
25	304013	416	.00137	61	157734	2743	.01724
26	305558	391	.00128	62	154174	2911	.01870
27	310554	461	.00148	63	144149	2968	.02038
28	275897	411	.00149	64	140100	2954	.02086
29	261592	392	.00150	65	135857	3391	.02465
30	264083	399	.00151	66	129386	3278	.02502
31	247777	378	.00152	67	123925	3352	.02669
32	241726	388	.00160	68	112574	3331	.02916
33	232025	365	.00157	69	119063	3736	.03089
34	233778	434	.00187	70	114066	3846	.03316
35	234338	439	.00187				

Extract from the normalized Mortality Table for the Calif. Popul.

x	9_x	l_x	d_x	x	9_x	l_x	d_x
0	.01801	100000	1801	44	.00399	92492	369
1	.00113	98199	111	45	.00448	92123	413
2	.00086	98088	84	46	.00506	91719	464
3	.00073	98004	72	47	.00548	91246	500
4	.00052	97932	51	48	.00562	90746	510
5	.00049	97881	48	49	.00630	90236	568
6	.00045	97833	44	50	.00652	89668	858
7	.00034	97789	33	51	.00772	89083	688
8	.00031	97756	30	52	.00837	88395	740
9	.00030	97726	29	53	.00914	87655	801
10	.00031	97697	30	54	.00974	86854	846
11	.00033	97667	32	55	.01093	86008	940
12	.00035	97635	34	56	.01196	85068	1017
13	.00041	97601	40	57	.01317	84051	1107
14	.00048	96561	47	58	.01330	82944	1103
15	.00062	97514	60	59	.01454	81841	1190
16	.00093	97454	91	60	.01553	80651	1253
17	.00105	97363	102	61	.01724	79398	1369
18	.00142	97261	138	62	.01870	78029	1459
19	.00156	97123	161	63	.02038	76570	1560
20	.00162	96962	157	64	.02086	75010	1565
21	.00161	96805	156	65	.02465	73445	1810
22	.00156	96649	151	66	.02502	71635	1792
23	.00152	96498	147	67	.02669	69843	1864
24	.00143	96351	138	68	.02916	67979	1982
25	.00137	96213	132	69	.03089	65997	2039
26	.00128	96081	123	70	.03316	63958	2121
27	.00148	95958	142	71	.03609	61837	2232
28	.00149	95816	143	72	.03906	59605	2328
29	.00150	95673	144	73	.04167	57277	2387
30	.00151	95529	144	74	.04586	54890	2517
31	.00152	95385	145	75	.05142	52373	2693
32	.00160	95240	152	76	.05787	49680	2875
33	.00157	95088	149	77	.05863	46805	2744
34	.00185	94939	176	78	.06668	44061	2938
35	.00187	94763	177	79	.07223	41123	2970
36	.00212	94586	201	80	.07391	38153	2820
37	.00227	94385	214	81	.08501	35333	3004
38	.00242	94171	228	82	.09613	32329	3108
39	.00256	93943	240	83	.10334	29221	3020
40	.00267	93703	250	84	.11171	26201	2927
41	.00308	93453	288				
42	.00359	93165	334	85+	1.00000	23274	23274
43	.00365	92831	339				

2. Plot $\hat{q}_x$ for $x = 0, \dots 25$ and determine q_x by the method of Wittstein.

3. Apply the same technique to $\ln \hat{q}_x$ and compare.

4. Fit $\hat{q}_x$ $x = 0, \dots, 10$ by a cubic polynomial.

The mortalities (q_x) obtained by smoothing and interpolation from the raw mortalities $(\hat{q}_x)$ are not considered adequate by many insurances for use as a life table to compute premiums and reserves. In particular with German insurance companies there has been a tendency to build the "fluctuation term" into the life table by increasing the (q_x) systematically (see, e.g. life table 1960/62 mod) in that case $q_{x\,\mathrm{mod}} = \max(q_{x+1}, q_x + 0.0005)$. Previously this has been done by using older life tables, whose (q_x) are larger than the present (q_x). Thus a life table from 1962 has been in use for more than 20 years. Also in France, Spain and Great Britain the life tables are renewed only every 10 years or so. The Canadian Institute of Actuaries has used $q_x + c_x$ instead of q_x, where the c_x are inversely proportional to the life expectancy of x-year-olds. The practice recommended by the German Actuarial Society replaces the q_x by $\tilde{q}_x$, where

$$\tilde{q}_x = r_x(q_x + s_x^\alpha).$$

The factor r_x is considered as a safeguard against an increase in mortality. This is absurd even though there might have been a slight increase in mortality due to immigrants from the east and due to the reunification. They use

$$r_x = \begin{cases} 1.2 & x \leq 20 \\ 1.2 - .01(x - 20) & 21 \leq x \leq 33 \\ 1.07 & x \geq 33. \end{cases}$$

The term s_x^α is supposed to be a safeguard against the fluctuaion s at a safety level α, i.e. the probability that the number of deaths exceeds

$$\sum (q_x + s_x^\alpha) L_x$$

is less than α. Here L_x is the number of persons aged x in the portfolio. Assuming a total portfolio of $\sum L_x = 300,000$ and $\alpha = .01$ the s_x^α can be computed given the age distribution (L_x). In this way a new life table based on the q_x from the census in 1986 was derived. Though insurances have not been bound since 1994 to use particular life tables, this is now the recommended life table in Germany for whole and term insurances.

Life tables for pensions should not be single period tables. One should rather use two-dimensional tables $(q_{x,t})$. The form recommended by the German Actuarial Society uses

$$\ln q(x, t) = \ln q^B(x) - F(x)t,$$

where the base value $q^B(x)$ for the year 2000 are determined as above in the form

$$q_x^B = f_x(q_x - s_x^\alpha)$$

with

$$f_x = \begin{cases} .9 & x \leq 20 \\ .9 - .01(x-20) & 21 \leq x \leq 29 \\ .8 & 30 \leq x \leq 59 \\ .6 & 60 \leq x \leq 65 \\ .6 + .015(x-65) & 66 \leq x \leq 74 \\ .75 & 75 \leq x \leq 110. \end{cases}$$

Again the s_x^α are fluctuation safeguards.

Problem 3.3 Why would an insurance company increase (decrease) the "actual" (q_x) for term or whole insurances (pensions)?

3.6 Mortality Laws

We have already mentioned the use of particular classes of functions to describe mortality. This does not mean that mortality obeys some form of a law, as people thought earlier, but rather that particular functions allow a simple description of mortality. The oldest mortality law is that of de Moivre (1724):

$$\ell_x = \ell_0 \left(1 - \frac{x}{86}\right). \tag{3.6.21}$$

Lambert improved this in 1765 to

$$\ell_x = \ell_0 \left\{ \left(1 - \frac{x}{96}\right)^2 - .6176 \left(\exp\frac{x}{13.68} - \exp\frac{x}{2.43}\right) \right\}. \tag{3.6.22}$$

Also the law

$$\ell_x = \ell_0 - a x^{1/4} - b x^{9/4} - c x^{17/4} \tag{3.6.23}$$

by Moser (1839) shows that mortality laws were investigated intensively in the 19th century. While these laws are mainly of historical interest, the Gompertz-Makeham law, which was postulated by Gompertz in 1824 and extended by Makeham in 1860,

$$\ell_x = k s^x g^{c^x} \tag{3.6.24}$$

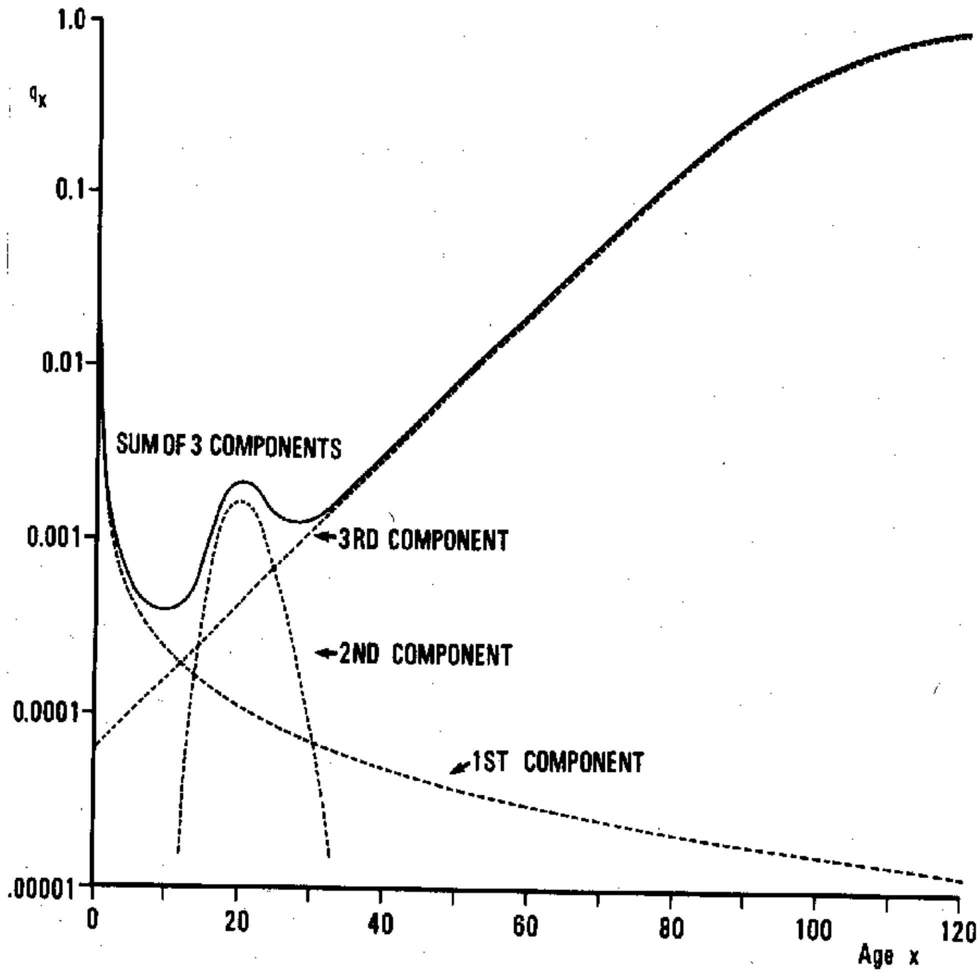
achieved wider fame and has been used extensively for smoothing data on subdomains of the life table. Quiquet (1893) generalized the approach of Gompertz Makeham to

$$\ln \ell_x = A + B_x + \sum_1^n e^{r_i x} h_i(x). \tag{3.6.25}$$

In the early 20th century actuaries realized that mortality depends on too many factors to allow a simple description. So nowadays mortality laws are rather considered a convenient description of a life table. Then, however, one is torn between simplicity and accuracy. It follows from general principles that the $(\hat{q}_x)_{x=0}^w$ can be fitted exactly by a polynomial. This polynomial, however, has degree ω, so nothing is gained by this. Thus the idea is to fit the (q_x) by a smooth function which depends on a few parameters, say 4 to 8, but allows a "good" fit of the 100 data values (q_x). Perks (1932) used μ_x or $q_x = (A + Bc^x)/(1 + Dc^x)$. This formula was refined by Barnett to $q_x = (a - Hx + Bc^x)/(1 + A - Hx + Bc^x)$. One of the best formulas in this context is the one by Heligman and Pollard [The Age Pattern of Mortality, J. of Inst. of Actuaries **107**, 1980, p. 49–80]. In this paper the authors study Australian mortality for the periods 1946–48, 1960–62 and 1971 using the formula

$$\frac{q_x}{p_x} = A^{(x+B)^c} + De^{-E(\ln x - \ln F)^2} + GH^x, \tag{3.6.26}$$

where $p_x = (1 - q_x)$. The parameters $A, \dots, H$ have to be estimated from the data. The Heligman-Polland model



The formula contains three terms, each representing distinct components of mortality. The first summand, a rapidly declining exponential, reflects the rapid decrease of mortality in early childhood. Here C is a measure for the rapidity of that decline. B is generally rather small, $0.01 \le B \le .03$. The second summand describes the "accident hump" in the early twenties, while the third summand, the well-known Gompertz exponential, reflects the near geometric rise in the mortality of adults.

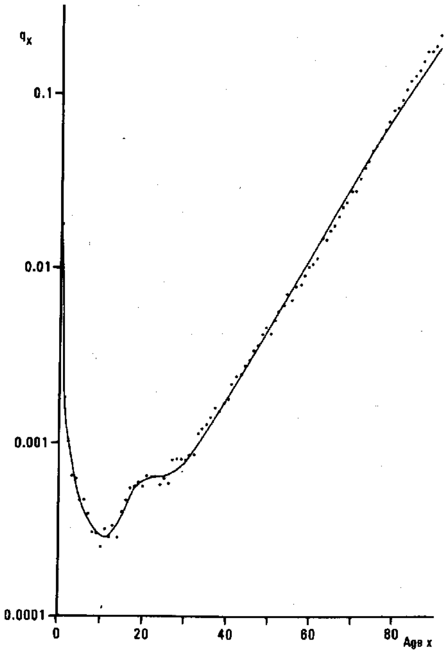
Parameters for model (1). Australian national mortality 1946-48, 1960-62, and 1970-

72.

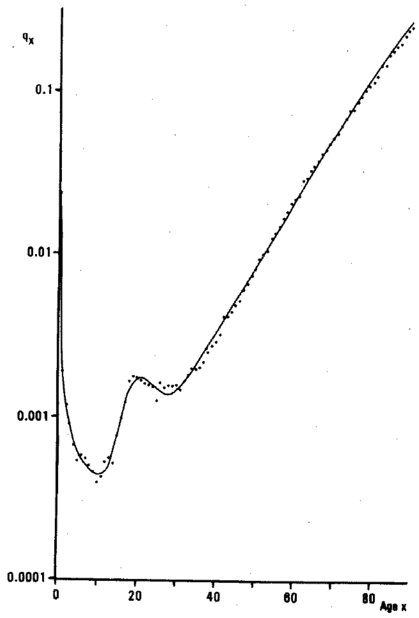
Parameter	Males 1946-48	Males 1960-62	Males 1970-72	Females 1946-48	Females 1960-62	Females 1970-72
A	.00341	.00184	.00163	.00293	.00177	.00137
B	.0208	.0189	.0144	.0336	.0304	.0251
C	.1284	.1189	.1182	.1339	.1309	.1249
D	.00094	.00110	.00164	.00156	.00025	.00039
E	$9 \cdot 49$	$13 \cdot 55$	$18 \cdot 49$	$1 \cdot 29$	$8 \cdot 83$	$16 \cdot 80$
F	$20 \cdot 22$	$20 \cdot 43$	$19 \cdot 88$	$53 \cdot 17$	$20 \cdot 37$	$18 \cdot 58$
G	.0000862	.0000711	.0000643	.0000196	.0000353	.0000383
(x_0)	$(101 \cdot 1)$	$(101 \cdot 0)$	$(100 \cdot 0)$	$(101 \cdot 7)$	$(105 \cdot 4)$	$(106 \cdot 0)$
H	$1 \cdot 0970$	$1 \cdot 0992$	$1 \cdot 1013$	$1 \cdot 1136$	$1 \cdot 1022$	$1 \cdot 1007$

Graduation of Australian Mortality

Females 1960/62



Males 1960/62



The close fit of these curves as well as the small number of parameters suggests that this model would be an ideal candidate for smoothing data. It should be added, however, that the numerics to compute the parameters from actual data is not quite simple. However, if these values are approximately known, Newton's method will readily give the best fit with a few iterations. The Heligman Pollard model also has the advantage that mortality is fitted by a class of functions which approximates the general form of mortality fairly well. Thus only a few parameters (1.1.8) are needed. Since life insurances are mainly interested in mortalities beyond 20 years, one essentially needs only four parameters. Thus this model is ideal for smoothing or fitting data. A slight improvement of the Heligman Pollard model has be proposed by Kostaki [Math. Popul. Studies 1992 **3** 277–288]. However, it improves only the fit of the accident hump, so it is of limited value for insurances. One question in this context is:

Does fitting a mortality law to the $(\hat{q}_x)$ allow avoiding smoothing by other forms?

The general advantage of such mortality laws results from the fact that the data $(\hat{q}_x)_{x=0}^{\omega}$ are compressed to a few parameters, which thus have a considerably better stability and accuracy. In this respect one statistical problem is the distribution of these parameters. So far this has not been investigated at all, though bootstrapping should give some answers.

We have already seen that the mortality in the portfolio of an insurance company will in general deviate from the mortality in the population. Similarly, one would expect that the mortality of certain subgroups in a portfolio will deviate from the general mortality. The insurance company might consider for example to introduce a special policy for nonsmokers or people who have already suffered from a heart attack. This has actually been done recently. The problem in this case is then the following:

The mortality table $(q_x)_{x=0}^{\omega}$ of the general group is given. Determine the mortality of a subgroup or related group $(\tilde{q}_x)$ under the assumption that

1. The (q_x) and $(\tilde{q}_x)$ behave similarly and
2. $q_x - \tilde{q}_x$ is not too large.

We will discuss three methods for determining the $\tilde{q}_x$. These are

The parameter method

If the (q_x) are determined by a mortality law with not too many parameters, this law can be fitted to the

$$\hat{q}_x = \frac{D_x}{L_x}.$$

Here D_x is the number of x-year-olds in the subgroup who have died, while L_x is the number of persons in the subgroup. In the case of people with heart attacks and the Heligman Pollard model, it might only be necessary to vary the parameters in the third summand. One might also consider fitting

$$\hat{q}_x - q_x$$

by a simple function.

The multiplicative method

In this case one will want to fit the $\hat{q}_x$ by an expression of the form mq_x, where m is a fixed factor. One way to determine m would then be to minimize the weighted sum of squares

$$Q(m) = \sum_x (mq_x - \hat{q}_x)^2 \frac{L_x}{L} \to \min,$$

where $L - \sum L_x$. Another possibility might be the maximum likelihood method. The likelihood function in this case is

$$L(m) = \prod_x \binom{L_x}{D_x} (mq_x)^{D_x} (1 - mq_x)^{L_x - D_x}$$

so that it suffices to minimize L or $\ln L$ with respect to m.

Problem 3.4 Compute $d/dm \ln L(m)$. What equation do you obtain for m?

The additive method

This method is rather like the multiplicative method, but now one considers

$$\hat{q}_x = (q_x + a)$$

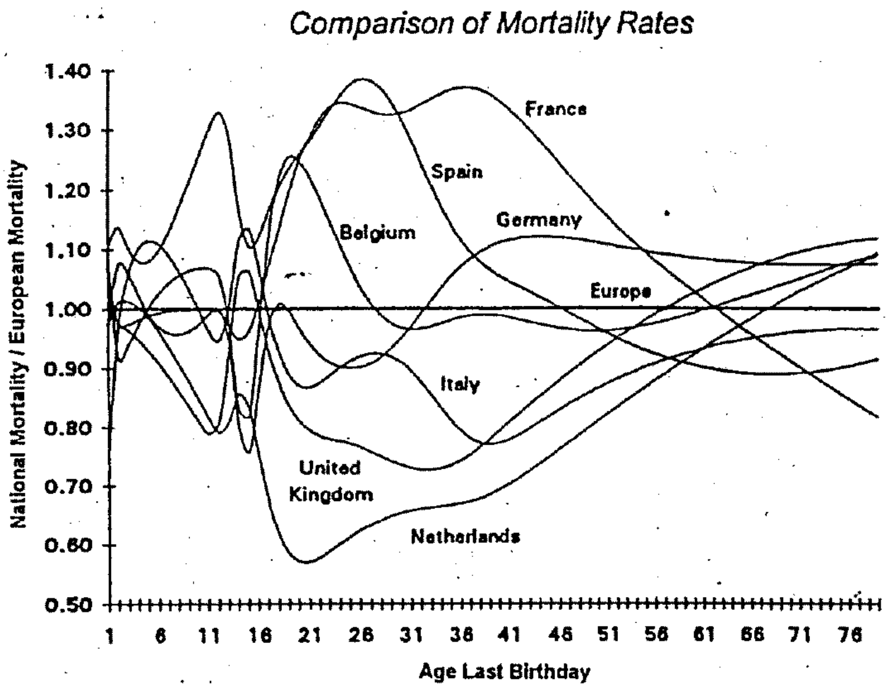
and determines the additive term similarly to the methods above.
It may also be possible to choose different factors or summands a in different regions.

3.7 A Comparison of European Mortality

As a consequence of European unification there will eventually be a single European life insurance market. This makes it necessary to understand the currently existing differences in insurance law, taxation, the social systems, economies and all other factors relating to life insurance. The social and economic conditions have been discussed above. Thus there remain the comparative analysis of mortality and rate making. In the analysis of mortality we refer to the paper of M. Kastenholz: "A Comparison of European Mortality 1989–1991" Blätter der Deutschen Gesellschaft für Versicherungsmathematik 323–357. Such analysis is necessary for the following reasons. Would a single European insurance market allow a single life table? If not, where are the differences in mortality and what causes these differences?
The study cited above is based on raw population mortality data published in Eurostat. Demographic Statistics 1991–1993. These data date back to 1989–1991. They were all analysed by a modified Heligman Pollard law

$$\frac{q_x}{p_x} = A^{(x+B)^c} + D/x^L \exp[-E^{2K}((\ln x)^J - (\ln F)^J)^{2K}]. \tag{3.7.27}$$

The following diagram shows the ratios of graduated national mortality rates versus the European average.



This plot illustrates that there are considerable differences between national population mortalities. Unfortunately it is not clear to what extent different statistics or measures of accounting are responsible for the observed differences. This plot also shows that there are three knots where the mortalities are rather close. These are in the domain 3–5, 14–15 and 57–63.

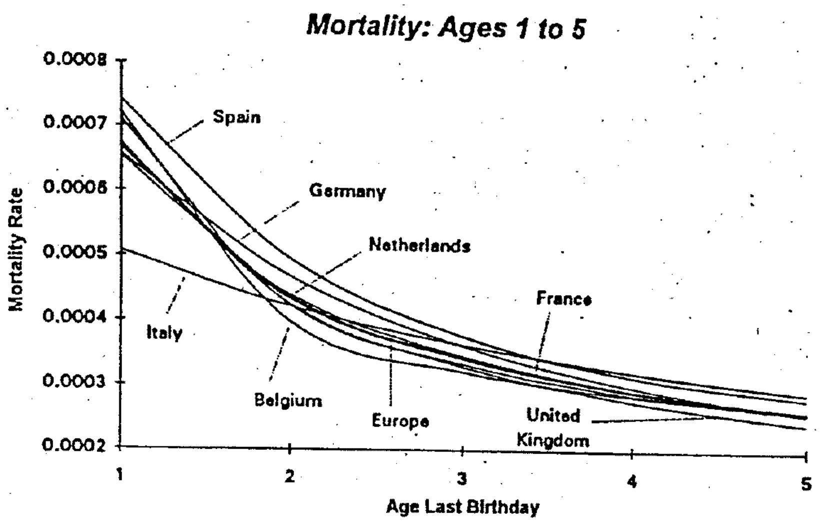
In his analysis Kastenholz uses two indices. The first, $S_1(x, y)$, is the expected number of deaths per 1000 for a standard population between age x and y. $S_2(x, y)$ gives this as a ratio versus the European standard. The indices $S_1(1, 79)$ and $S_2(1, 79)$ may serve as a rough comparison.

Mortality of Children aged 1–5

Neither the diagram nor the indices indicate any dramatic differences

Single-Figure Indices Ages 1 to 5

Country	SMR1 (1, 5)	SMR2 (1, 5)
Belgium	0.40	99.4 %
Germany	0.41	102.2 %
Spain	0.44	110.5 %
France	0.40	100.7 %
Italy	0.38	95.3 %
Netherlands	0.40	99.5 %
United Kingdom	0.39	97.4 %

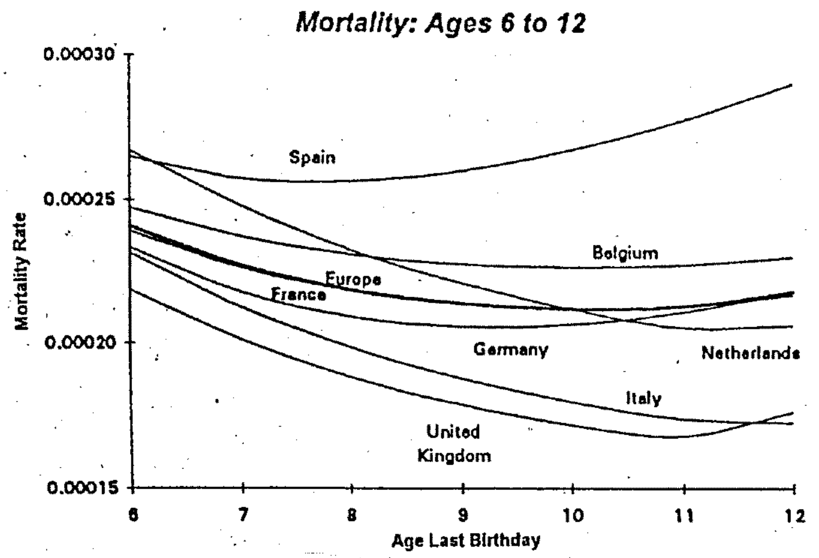


Mortality of Children aged 6–12

Both the diagram and the indices indicate a considerable difference. At this age children are mainly exposed to the risks of accidental death or death from infectious diseases. It remains to be studied whether the type of schools and medical services are partly responsible for these differences.

Single-Figure Indices Ages 6 to 12

Country	SMR1 (6, 12)	SMR2 (6, 12)
Belgium	0.23	105.4 %
Germany	0.21	97.3 %
Spain	0.27	121.3 %
France	0.22	99.7 %
Italy	0.23	103.0 %
Netherlands	0.19	87.7 %
United Kingdom	0.19	84.3 %

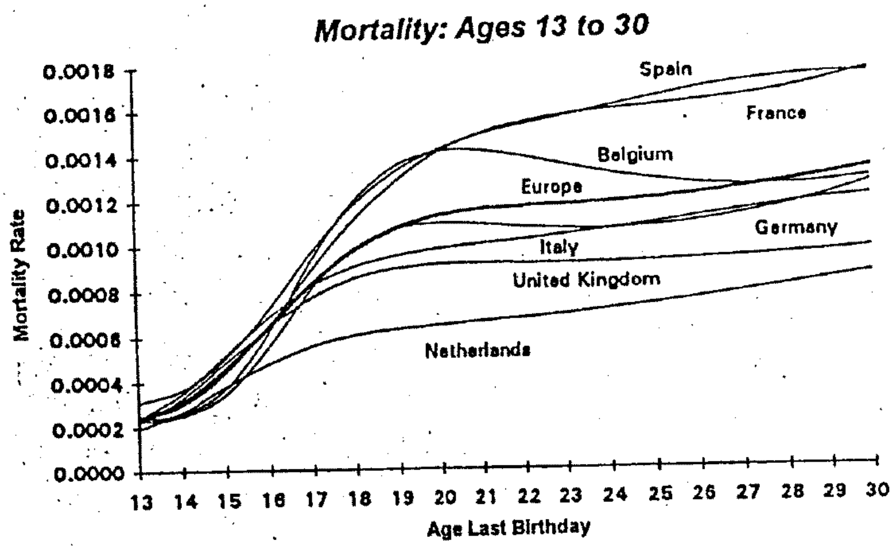


Mortality: Ages 13 to 30

In this age range accidental mortality, homicide, suicide and drug abuse are the major causes of mortality. It still remains to be analyzed which true factors are responsible for this discrepancy. In this age range one also finds the "accident hump". However, the relatively high mortality in Spain does not permit us to relate this to the general traffic density.

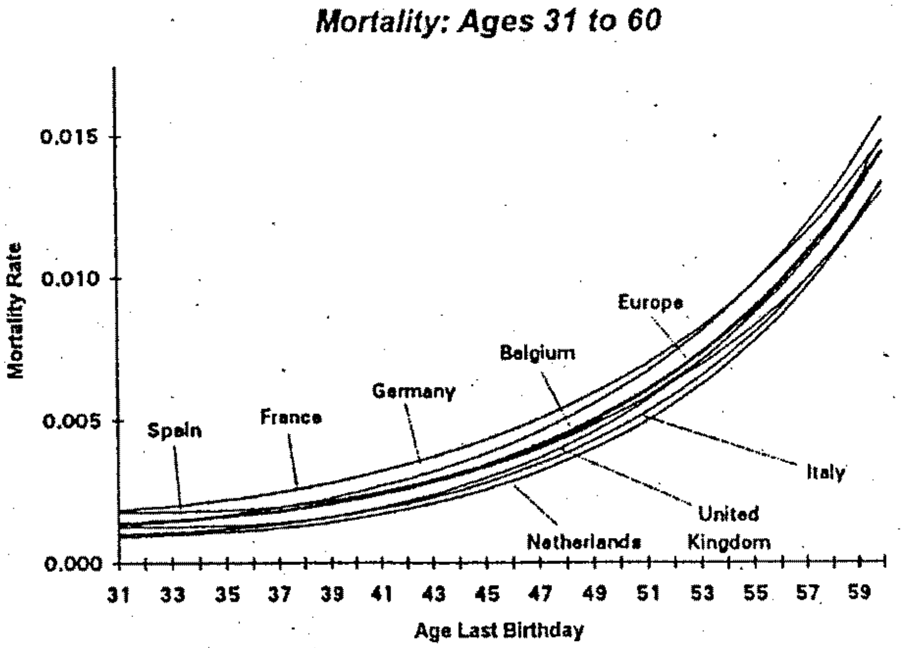
Single-Figure Indices Ages 13 to 30

Country	SMR1 (13, 30)	SMR2 (13, 30)
Belgium	1.11	108.9 %
Germany	0.95	92.6 %
Spain	1.33	130.3 %
France	1.29	126.6 %
Italy	0.94	92.0 %
Netherlands	0.64	62.8 %
United Kingdom	0.82	80.5 %



Mortality: Ages 31–60

In this range the mortality exhibits the same trend, but still there are considerable differences, which can readily be recognized from the indices. The major factor for mortality here should be lifestyle. This range is also responsible for a large part of life insurances. So we can expect differences in premiums of up to 40%.

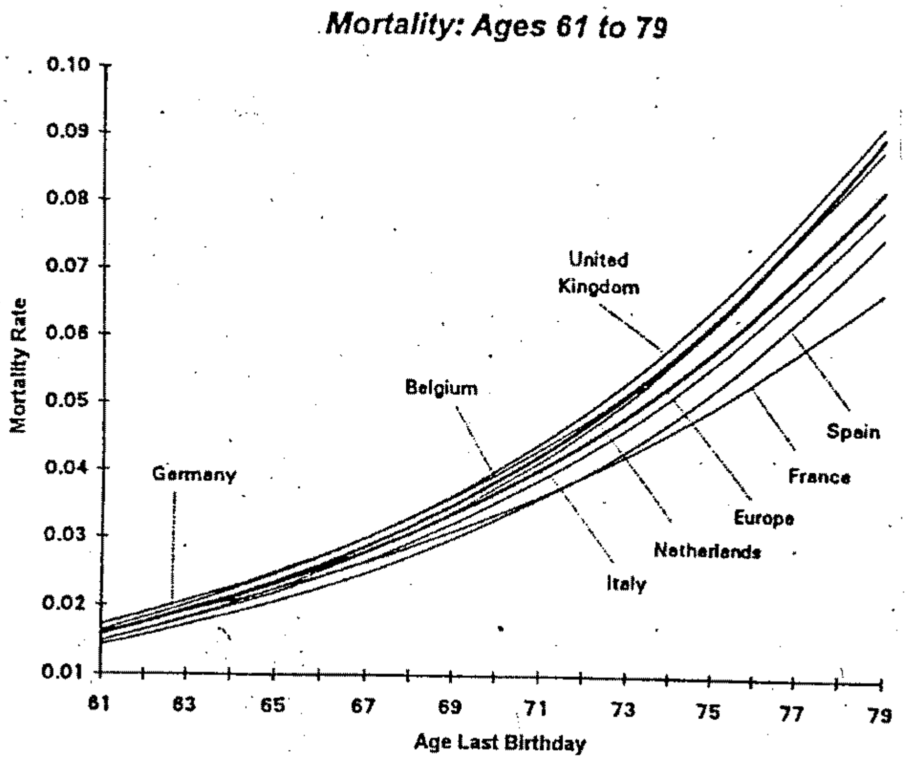


The Accident Humps

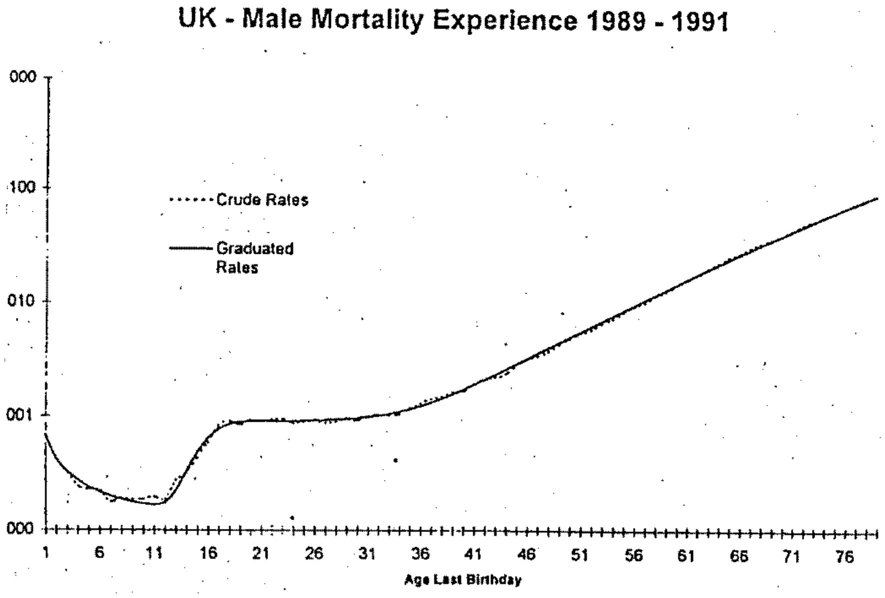
Country	Rate (per 1000) at Age			Deviation from European Average		
	20	25	30	20	25	30
Belgium	1.42	1.29	1.29	+25 %	+7 %	-3 %
Germany	1.10	1.08	1.27	-3	-10 %	-5 %
Spain	1.42	1.65	1.76	+26 %	+38 %	+32 %
France	1.41	1.61	1.76	+25 %	+35 %	+32 %
Italy	0.99	1.09	1.21	-13 %	-9 %	-9 %
Netherlands	0.65	0.74	0.87	-43 %	-39 %	-35 %
United Kingdom	0.92	0.93	0.98	-19 %	-23 %	-27 %
Average	1.13	1.20	1.33			

Mortality: Ages 61–79

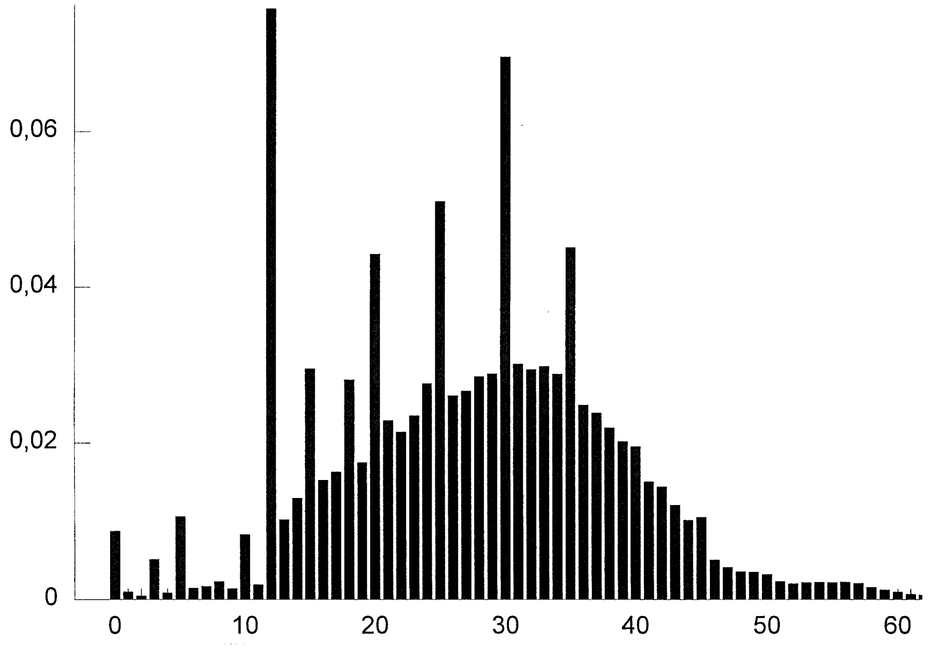
Again all curves reflect the steady rise in mortality. However, the slopes for France and Italy are well below the European average. Is this the “red wine effect”?



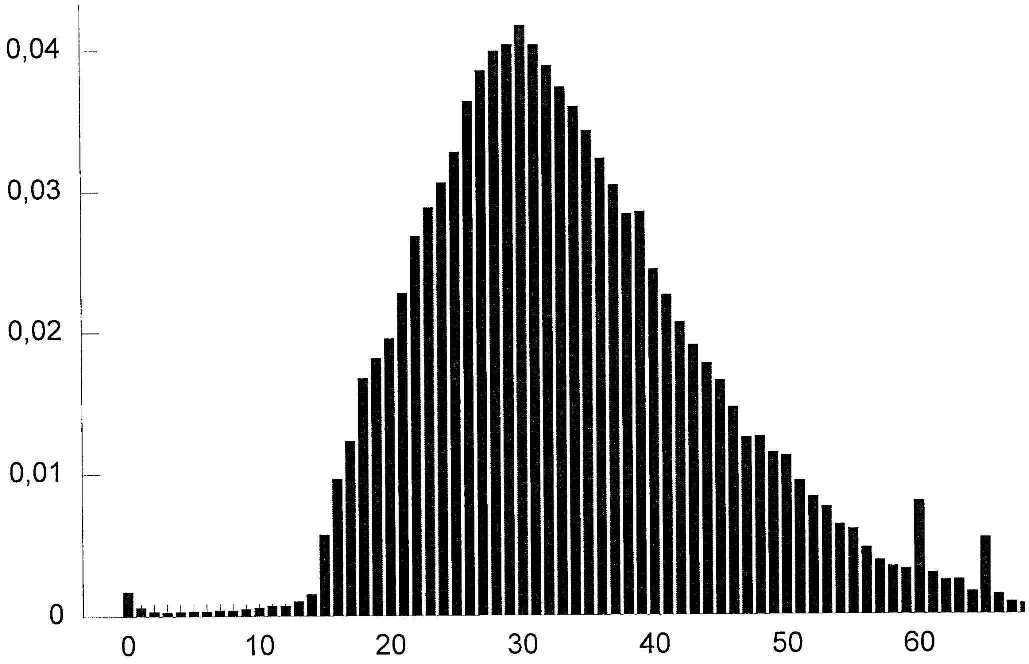
As a final example we show the fit of this Heligman Pollard model to the UK crude mortalities. A similar good fit can be observed for the other countries of this study.



Summary: Mortality is the key stochastic variable in life insurance. While governments usually use the method of "year of death" or "year of birth" in conjunction with a census, insurances use the direct method of estimation $\hat{q}_x = Dx/Lx$. The raw mortalities are then smoothed. For this there are many methods available whose merit is difficult to judge. So far there are no methods available which are directly related to insurance to assess the fit and validity of mortalities. In some cases the fluctuation term is built into the mortality table if elevated values are used.



Duration of life insurances



Age at the beginning of life insurance

4. Life Insurance Premiums

4.1 The Equivalence Principle

In the first chapter we saw that in general the insurance premium is made up of three parts

$$\text{Premium} = \text{Net premium} + \text{Fluctuation Term} + \text{Administr. costs.} \tag{4.1.1}$$

For the net premium we derived on the basis of general statistical considerations

$$\text{Net premium} = \text{Expected Losses.} \tag{4.1.2}$$

These principles are still correct, but they have to be augmented for life insurances because

1. Life insurance contracts usually cover a long time span of 20 to 40 years, so compound interest has to be considered.
2. Premiums are usually paid over a longer time interval. Thus also the payment of the premiums has a stochastic element, because the insured might die.

Taking both these points into account we are led to the general *Equivalence Principle for Life Insurances*

$$\text{Current value of expected net premiums} = \text{Current value of expected losses.} \tag{4.1.3}$$

In order to actually compute insurance premiums, a few remarks are necessary. The first is that insurance companies tend to work with a fixed interest rate p, even though p actually varies. Since p determines the premiums and reserves decisively, governments usually have some rules on how to fix p. In Germany p used to be fixed at $p = 3\%$ until 1987 and 3.5% until 1991. Since then p has to be chosen adequately. Only the interest rate used for computing the reserves is fixed at 4%. Recently the rule in Germany was that p has to be about 60% of the average interest paid on government bonds. So in 2000, 2001 p will be 3.5%. This is now the general rule throughout the EU. The adequate rate is usually determined by the organization of insurance companies or one uses 60% of a weighted average over the last 5 or 10 years. In France, $p = 4.5\%$ was used for life insurances with a large savings component, otherwise p was oriented towards the interest rates of banks. In Spain, p was only limited to a maximum of 6%. But even then there were a few exceptions.

Roughly speaking we can thus expect insurance companies to use an interest rate of between 3 and 5 per cent.

Administrative Costs

We will now comment on the term administrative costs. As noted above, these tend to be proportional to the claim. The administrative costs consist of: cost of acquisition, cost of collecting, cost of settlement and cost of administration.

Cost of acquisition comprises all costs necessary to gain new customers. This includes costs for marketing, costs for establishing branch offices and networks of insurance agents. It may also include costs for medical examinations. The largest part of the cost of acquisition, however, are the commissions for insurance agents. In general, these range from 1% to 3.5% of the face amount. The precise amount depends on the company and the status of the agents. In most cases the commission is proportional to the face value, the yearly premium or pension. It may also depend on the duration of the contract. Sometimes, when the agent has a fixed salary, the bonus will depend on the face value. It is a problem, however, if agents are more commission-oriented than willing to offer good advice. In Great Britain, commissions have been regulated somewhat by LAUTRO. In recent years there has been a trend towards direct insurance or of banks moving into the insurance business. The main reason for this is of course to reduce the cost of acqusition.

Cost of collecting For most life insurances the premiums are paid regularly. The cost of transferring the money is the cost of collecting. In the age of computers and automatic collecting, these costs are now minimal. This also holds for the

Cost of administration which is mainly keeping the accounts and correspondence with customers. For insurances tied to funds or stocks a little more work might be involved. Thus in Great Britain insurances will demand between .5 and 1% in fund management charges per year. In France the administrative costs are not fixed and vary between 0.5 and 1.5%. To the cost of administration we may also add research and the work in preparing new policies.

Cost of settlement comprises claim investigation, legal expenses and cost of disbursing.

Altogether, one can say that commissions, the cost of collecting and administration in general make up between 5 and 10 per cent of the face value of the insurance.

In the next section we will treat the problem of defining the fluctuation terms. For this reason we will study net premiums first. By the equivalence principle the net premiums are also the single premiums, paid at the very beginning of the contract.

Assume we have an insurance policy (X_i) which promises to pay X_i in year i from now. Here X_i may be a random variable. Then (4.1.2) implies

$$\text{Net premium} = \sum_{j=1}^{n} E(X_j)\left(1+\frac{p}{100}\right)^{-j} = \sum_{j=1}^{n} E(X_j)v^j. \tag{4.1.4}$$

The right-hand side is clearly based on the assumption that the insurance benefits are paid at the end of the year. If the payment is made immediately and spread out approximately evenly throughout each year, one could write approximately

$$\text{Net premium} = \sum_{j=1}^{n} E(X_j) v^{j-1/2}. \qquad (4.1.5)$$

In life insurances the random variable X_j will have the form

$$X_j = \begin{cases} S_j & \text{if the person survives year } j \\ D_j & \text{if the person dies in year } j. \end{cases}$$

The probabilities of survival or death can be computed from the life table. Thus assume $(q_x)_{x=0}^{\omega-1}$ with $q_{\omega-1} = 1$ is a given life table. From it compute the life table $(\ell_x)_{x=0}^{\omega}$ by

$$\ell_{x+1} = \ell_x(1 - q_x) \qquad \ell_0 = 100,000. \qquad (4.1.6)$$

Then the probability that a person of age x dies in year $[j-1, j)$ from now is

$$w_j = \frac{\ell_{x+j-1}}{\ell_x} q_{x+j-1} =_{j-1} p_x \; q_{x+j-1}. \qquad (4.1.7)$$

In this formula ℓ_{x+j-1}/ℓ_x is the conditional probability that a person of age x now reaches the age $x+j-1$, while the second factor gives the probability of dying in year $x+j$ after having reached the age $x+j-1$.

Problem 4.1 Prove that formula (4.1.7) defines a probability measure on $\{0, 1, \dots, \omega - x\}$. What is the meaning of the expected value with respect to this measure?

Term Insurance

The simplest type of life insurance is the term insurance. This is a contract against the risk of death which usually runs for 5, 10,20 ... years. In this case the policy owner is paid the face value F if he dies during the given period of n years. In this case each X_j, $1 \le j \le n$ is binomially distributed

$$X_j = \begin{cases} F & \text{with prob. } w_j \quad j = 1, \dots, n \\ 0 & \text{with prob. } 1 - w_j. \end{cases}$$

Thus (4.1.5) and (4.1.6) give with $v = (1 + p/100)^{-1}$

$$P_{x,n\rceil} = F \cdot \sum_{s=0}^{n-1} {}_s p_x q_{x+s} v^{s+1} = F \cdot A^1_{x,n\rceil}. \qquad (4.1.8)$$

If the term insurance is not limited, (4.1.7) is still valid but we have to sum until $\omega - x$. Thus

$$P_x = F \sum_{s=0}^{\omega - x} {}_s p_x q_{x+s} v^{s+1} = F \cdot A_x. \qquad (4.1.9)$$

At this point two observations are due. First we see that the net premium is proportional to F, the face value. Thus it suffices to list or compute only the values A_x or A_x. Thus $A_{x,n\rceil}$ is the net premium for a term insurance for n years with face value 1. In the remainder we shall keep to this type of normalisation. The second observation concerns the mortality values (q_x). Larger (q_x) imply higher premiums. Thus insurance companies can use this mechanism to include the fluctuation term into the net premium by choosing larger (q_x). This has already been done by German life insurance companies. As an example consider the German mortality table 60/62 which was based on the census of 1961. It was modified according to the rule

$$q_x \operatorname{mod} = \max(q_x + .0005,\ q_{x+1}).$$

It is also used in other European countries whenever older, outdated mortality tables are used. It follows readily from (4.1.5) and (4.1.8) that

$$A_{x,n\rceil} = A_x - A_{x+n}. \tag{4.1.10}$$

Thus it suffices just to compute the A_x. But these expressions can still be simplified. We have

$$A_x = \sum_{s=0}^{\omega-x} {}_s p_x q_{x+s} v^{s+1} = \sum_{s=0}^{\omega-x} \frac{\ell_{x+s}}{\ell_s} \frac{d_{x+s}}{\ell_{x+s}} v^{s+1},$$

where d_{x+s} is the number of deceased persons in year $x+s$; $d_x + s = \ell_{x+s} - \ell_{x+s+1}$. Thus

$$A_x = \sum_{s=0}^{\omega-x} \frac{d_{x+s} v^{x+s+1}}{\ell_x v^x} = \sum_{s=0}^{\omega-x} \frac{C_{x+s}}{D_x} = \frac{M_x}{D_x}. \tag{4.1.11}$$

Here $C_x = d_x v^{x+1}$ is called the *discounted number of the dead* while $D_x = \ell_x v^x$ is the *discounted number of the living.* M_x is the sum of the discounted number of the deceased. These numbers C_x, D_x and M_x are called *commutation numbers.* They were used extensively to aid the computations of these sums. (4.1.10) and (4.1.11) imply

$$A^1_{x,n\rceil} = \frac{(M_x - M_{x+n})}{D_x}. \tag{4.1.12}$$

We have

$$C_x = d_x v^{x+1} = (\ell_x - \ell_{x+1}) v^{x+1} = v D_x - D_{x+1}.$$

Thus

$$\begin{aligned} M_x &= \sum_{s=0}^{w-x} C_{x+s} = \sum_{s=0}^{w-x} v D_{x+s} - D_{x+s+1} = v N_x - (N_x - D_x) \\ &= D_x - (1-v) N_x. \end{aligned}$$

This shows $A_x = 1 - (1-v)\ddot{a}_x$.

Problems

1. What is the influence of the interest rate on the premiums of term insurances?
2. Show that $C_x = vD_x - D_{x+1}$.
3. Compute C_x, D_x for $x = 25$, 30, 35 and $p = 3\%$ based on the above modified mortality table 1960/62 for men.
4. Based on the above modified mortality table for men show one of the results

p%	3	3.5	7
A_{40}	.418291	.366857	.163251

5. As in Probem 1. show one of the results

p	3	3.5	7
$A^1_{40,25\rceil}$	.177449	.164330	0.99949

6. The following is from an advertisment in a German weekly journal "Cheaper, faster, more direct, XXX Insurance .... Example: Term insurance with face value DM 100,000 for a woman aged 20 for 10 years, just 3.99 per month." Any comments? Use the life table and commutation numbers in the appendix and table of commutation numbers.

Endowment Insurance

A *pure endowment contract* promises to pay the face value if the insured survives a specified period. To be more specific, assume the person is of age x and the pure endowment contract is for n years. Then ℓ_{x+n}/ℓ_x is the probability that this person survives the n years. Thus the net premium for a face value 1 in this case is

$$A_{x,n\rceil}^{\;\;1} = \frac{\ell_{x+n}}{\ell_x}v^n = \frac{\ell_{x+n}v^{x+n}}{\ell_x v^x} = \frac{D_{x+n}}{D_x}. \tag{4.1.13}$$

The usual endowment insurance combines the effect of an n-year term insurance and pure endowment insurance. Such an insurance promises to pay the face value whenever the policy holder dies within the period of n years and it promises to pay the face value at the end of n years if the policy holder survives. In this case we clearly have for the net premium

$$A_{x,n\rceil} = A^1_{x,n\rceil} + A_{x,n\rceil}^{\;\;1} = \frac{(M_x - M_{x+n} + D_{x+n})}{D_x}. \tag{4.1.14}$$

$$\begin{aligned} A_{x,n\rceil} &= \frac{M_x}{D_x} - \frac{M_{x+n}}{D_{x+n}}\frac{D_{x+n}}{D_x} + \frac{D_{x+n}}{D_x} \\ &= A_x + \frac{D_{x+n}}{D_x}(1 - A_{x+n}) = (1 - d\ddot{a}_x) + \frac{D_{x+n}}{D_x}(1 - (1 - d\ddot{a}_{x+n})) \\ &= 1 - d\ddot{a}_x + d\frac{D_{x+n}}{D_x}\ddot{a}_{x+n} = 1 - d\ddot{a}_{x,n\rceil}, \end{aligned}$$

where $d = (1 - v)$.

Deferred insurance:

An m-year deferred insurance provides for a benefit only if the insured dies at least m years following policy issue. To be specific, consider an m-year deferred whole life insurance for an x-year-old person with face value 1.

$$ {}_{m|}A_x = \sum_{s=m}^{\omega-x} {}_sp_x q_{x+s} v^{s+1} = \sum_{s=0}^{\omega-x-m} {}_{s+m}p_x q_{x+s+m} v^{s+m+1}. \tag{4.1.15}$$

Problem 4.2 Derive the corresponding formulas for the other types of deferred insurances.

Annuities

An annuity is a periodic payment – usually once a year or once a month – made during a fixed period contingent on the life of the annuitant. A whole life annuity is thus generally a yearly payment, as long as the annuitant lives. Here one also has to make the distinction when the payments are made. Payments at the beginning of the period are called due.

With this is mind the net premium for a whole life annuity of value 1 due is

$$\begin{aligned} \ddot{a}_x = \sum_{s=0}^{\omega-x} {}_sp_x v^x \quad &= \sum_{s=0}^{\omega-x} \frac{\ell_{x+s}}{\ell_x} = \sum_{s=0}^{\omega-x} \frac{\ell_{x+s} v^{s+x}}{\ell_x s^x} \\ = \sum_{s=0}^{\omega-x} \frac{D_{x+s}}{D_x} &= \frac{\sum_{s=0}^{\omega-x} D_{x+s}}{D_x} = \frac{N_x}{D_x}. \end{aligned} \tag{4.1.16}$$

Here again N_x is a commutation number.

You may have wondered by now about the strange letters and notations employed for these types of insurances. These are actually international standard actuarial notations, the list of which is stated in the appendix to this chapter. In this case $\ddot{}$ always indicates payments at the beginning of the period.

Problem 4.3 Now it should be easy for you to derive the formulas for

1. an n-year temporary life annuity due $\ddot{a}_{x,n]}$ or
2. an m-year deferred whole life annuity ${}_{m|}\ddot{a}_x$.

Problem 4.4 Show $\ddot{a}_{x,n]} = (N_x - N_{x+n})/D_x$.

If the annuity is paid at the end of the period, the term *immediate* is used. Thus the present value of a whole life annuity immediate is

$$a_x = \sum_{s=0}^{\infty} {}_sp_x v^{s+1}. \tag{4.1.17}$$

Problem 4.5 Show $a_x = \ddot{a} - 1$ and $A_x = v\ddot{a}_x - a_x$.

Problem 4.6 Show that an n-year deferred whole life annuity immediate is given by ${}_{n|}a_x = N_{x+n+1}/D_x$.

Problem 4.7 What is the effect of inflation on

1. whole life insurances?

2. annuities?

Let us now consider the situation that the insured person wants to counter the effect of inflation by increasing the face value each year by a factor of $(1-g/100)^{-1} = w$, where g is the rate of inflation. The net premium for an inflation-adjusted whole life insurance is then

$$P_x = \sum_{s=0}^{\omega-x} {}_sp_x\, q_{x+s} v^{s+1} w^{s+1}. \tag{4.1.18}$$

It is clear that this expression is again of the form (4.1.9), if one replaces v by $v \cdot w$.

Problem 4.8 Derive the corresponding formula for inflation-adjusted annuities.

Premiums: So far we have only computed the formulas for the net premium, more precisely a single net premium payable at the beginning of the insurance contract. While this actually occurs, e.g. if the insured has inherited, the rule is that the premiums are paid at a constant rate, usually monthly. Such level premiums are thus more or less the rule. But a level premium is nothing but an annuity, paid by the insured. Since premiums are paid in advance, we have from (4.1.16)

$$\text{Present value of premium } Pr = Pr \cdot \ddot{a}_x. \tag{4.1.19}$$

If the premiums are supposed to be paid only for n years we get

$$\text{Present value } = Pr \cdot \ddot{a}_{x,\overline{n}|}. \tag{4.1.20}$$

With these expressions and (4.1.5), we can, at least in principle, compute level premiums.

Example 4.1 M. Clever is 30 and needs a term insurance with face value $F =$ 100,000 until he is 65, when he expects to retire. Compute the net level premium.

As we have seen above the present value is $F \cdot A^1_{x,\overline{n}|}$. If Pr is the level premium, we thus have

$$FA^1_{x,\overline{35}|} = Pr \cdot \ddot{a}_{x,\overline{35}|}.$$

Example 4.2 One of the standard life insurances is the endowment insurance. For a face value F and time n for a person of age x, its present value is $A_{x,\overline{n}|}F = (M_x - M_{x+n} + D_{x+n})\, F/D_x$. A level premium Pr for such an insurance is thus

$$Pr = \frac{(M_x - M_{x+n} + D_{x+n})F}{D_x \cdot \ddot{a}_{x,\overline{n}|}} = \frac{(M_x - M_{x+n} + D_{x+n})F}{(N_x - N_{x+n})}.$$

Example 4.3 Since M. Breadwinner considers the government pension system inadequate, he wants to buy himself a life annuity of $F = 1,000$, starting with his retirement at the age of 65. He is currently 30 years old.

From the insurance side, we are talking about a deferred life annuity due. Its present value is

$$1000 \cdot_{35|} \ddot{a}_{30} = 1000 \cdot \sum_{s=35}^{\omega-65} {}_s p_x v^v.$$

On the premium side with a level premium we have a net value of $Pr\ddot{a}_{30,\,35\rceil}$. Thus

$$Pr = \frac{(1000 \cdot_{35|} \ddot{a}_{30})}{\ddot{a}_{30,\,35\rceil}}.$$

Problem 4.9 Since M. Breadwinner expects a yearly increase in his salary of $g\%$, he is willing to let his premiums grow by that rate. Formulate the corresponding problem.

It is now possible to formulate the general life insurance problem.
A person aged x years wants an insurance policy that pays him D_i if he dies in his $x + i$th year and pays him S_i, if he survives the $(x + i)$th year. For this he pays the premium Pr_i at the beginning of his $(x + i)$th year. Such an insurance premium system we abbreviate by $(x, (D_i), (S_i), (Pr_i))$. Then we have
Present value in case of death

$$= \sum_{s=0}^{\omega-x} {}_s p_x q_{x+s} D_s v^{s+1}$$

Present value for survival

$$= \sum_{s=0}^{\omega-x} {}_s p_x v^{s+1} S_s.$$

Finally, the present value of the premiums amounts to

$$\sum_{s=0}^{\omega-x} {}_s p_x v^s Pr_s.$$

By the equivalence principle, we should have

$$\sum_{s=0}^{\omega-x} {}_s p_x v^s Pr_s = \sum_{s=0}^{\omega-x} {}_s p_x q_{x+s} D_v v^{s+1} + \sum_{s=0}^{\omega-x} {}_v p_x v^{s+1} S_s. \tag{4.1.21}$$

Problem 4.10 With a fixed term insurance the face value will be paid after n years, but the premiums are only paid as long as the insured lives. Compute the present value and derive a formula for a level premium.

The above formulas also allow us to compute the present value of pensions. This is necessary because sometimes people want to invest money in a pension scheme or to get the cash equivalent for a pension. A pension is of course an annuity paid in advance. So if P_s is the amount paid in s years the present value for a person of age x is

$$\sum {}_sp_x v^s P_s.$$

Government pensions are usually inflation adjusted, so in most cases one will have $P_s = R \cdot (1-g)^{-s}$.

Problem 4.11 Mr. Active, aged 70, has just sold his firm and wants to invest 500,000 in a pension scheme with pension guarantee for 10 years, which is to begin immediately. What is his monthly pension? What is his monthly pension if the payments are inflation-adjusted with $g = 2\%$.

Problem 4.12 Mr. Active, aged 50, has received a loan of 1,000,000 from the bank at 6%, which he intends to pay in constant installments over 5 years, always at the end of the year. For this the bank demands a term insurance which covers the remaining debt. Compute the level premium for this insurance.

Sometimes an accidental death-benefit clause or double indemnity is added to a life insurance. In this case the insurance provides a double or even triple amount of the face value if the death of the insured is caused by an accident. Let us consider this case for a term insurance. The extension to other insurances is then quite obvious. Now q_x, the probability of an x-year-old dying within the next year has to be written as

$$q_x = q_{x1} + q_{x2}, \tag{4.1.22}$$

where q_{x1} and q_{x2} are the chance of dying from normal and accidental causes, respectively. With this the present value of a double indemnity n-year term insurance is

$$\begin{aligned} & F \cdot \sum_{s=0}^{n-1} {}_vp_x q_{x+s,1} v^{s+1} + 2F \sum_{s=0}^{n-1} {}_vp_x q_{x+s,2} v^{s+1} \\ &= \sum_{0}^{n-1} {}_vp_x q_{x+s} v^{s+1} + F \sum_{v=0}^{n-1} {}_vp_x q_{x+s,2} v^{s+1} \\ &= F \sum_{v=0}^{n-1} {}_vp_x (q_{x+s} + q_{x+s,2}) v^{s+1} \end{aligned} \tag{4.1.23}$$

Problem 4.13 It is sometimes advantageous to define actuarial functions recursively. In this context prove

$$A_x = q_x v + p_x v A_{x+1}.$$

For these insurances we can also compute the higher moments. These are needed, for example, to compute or estimate the fluctuation term. To begin with, let us recall the well-known formula for the variance σ_X^2 of a random variable X

$$\sigma_X^2 = E(X^2) - E(X)^2.$$

Now consider a term insurance of face value 1 for n years. Then the k-th moment is given by

$$^k A^1_{x,\overline{n}|} = \sum_{s=0}^{n-1} v_s^{k(s+1)} p_x q_{x+s} \tag{4.1.24}$$

because the probability distribution on $\{0, 1, \ldots, w - x\}$ is given by ${}_s p_x q_{x+s}$, $s = 0, 1, \ldots, w - x$.
The pure endowment insurance pays the face value if the insured survives n years. The probability of this is ${}_n p_s$. For a face value 1 the distribution is thus essentially a binomial distribution with factor v^n. The variance is therefore given by

$$v^{2n}(1 -_n p_x)_n p_x = v_n^{2n} q_x \, {}_n p_x.$$

An endowment insurance actually consist of two parts $X + Y$. The first part X pays the face amount if the insured dies within the n years. Extending (4.1.8) somewhat, its k-th moment is for a face value 1 is given by (4.1.24) The Y part describes the pure endowment, whose k-th moment is

$$^k A_{x,\overline{n^1}|} =_n p_x v^{kn}. \tag{4.1.25}$$

In order to compute the higher moments of $X + Y$, we use $X \cdot Y = 0$, so that

$$E((X + Y)^k) = E(X^k) + E(Y)^k =^k A^1_{x,\overline{n}|} +^k A^{\,1}_{x,\overline{n}|}. \tag{4.1.26}$$

Problem 4.14 Derive similar formulas for the corresponding deferred insurances.

Finally, let us consider annuities. Again, we will assume that the face value is 1. If the annuitant dies in the ℓ-th year, the total amount of

$$\frac{1 - v^{\ell+1}}{1 - v}$$

will have been paid. The probability of this is ${}_\ell p_x q_{x+\ell}$. Thus the k-th moment for this is

$$^k \ddot{a}_x = \sum_{s=0}^{w-x} {}_s p_x q_{x+s} \left(\frac{1 - v^{\ell+1}}{1 - v}\right)^k. \tag{4.1.27}$$

Its variance is thus

$$\sigma^2 = \sigma^2 \left(\frac{1 - v^{X+1}}{1 - v}\right) = \left(\frac{1}{1 - v}\right)^2 \sigma^2(v^{T+1}), \tag{4.1.28}$$

where T is the random variable life time. The latter expression also appears for the term insurance. In fact,

Problem 4.15 Use (4.1.27) and (4.1.28) to derive

$$\begin{aligned} \ddot{a}_x &= \frac{1}{d}(1 - A_x) \qquad d = 1 - v \quad \text{and} \\ 1 &= d\ddot{a}_x + A_x \end{aligned} \tag{4.1.29}$$

$$\sigma^2 = d^{-2}({}^2A_x - A_x^2). \tag{4.1.30}$$

Problem 4.16 Similarly, show:

$$\ddot{a}_{x,n\rceil} = \frac{1}{d}(1 - A_{x,n\rceil}) \tag{4.1.31}$$

$$\sigma^2 = \frac{1}{d^2}({}^2A_{x,n} - A^2_{x,n\rceil}. \tag{4.1.32}$$

Problem 4.17 The pension fund of a firm has n_i members of age $65 + i$, $i = 0, 1, \dots, 35$. Give a formula for the necessary reserves for this and derive a formula for the variance, assuming the lives to be independent.

Summary The basic principle for life insurances is:

Current value of expected net premium = current value of expected payments by the insurance company

and

Gross premium = net premium + fluctuation term + administrative term

The current value of expected losses can be expressed in terms of commutation numbers.

Appendix I

Constants for Actuarial Functions

L.I. = Life Insurance, E.I. = Endowment Insurance, L.A. = Life Annuity

Discrete Case

Name	Notation
whole L.I.	A_x
n-year term L.I.	$A^{1}_{xn\rceil}$
m-year deferred L.I. n-year term L.I.	${}_{m\|}\overline{A}^{1}_{xn\rceil}$
n-year pure E.I.	$A^{1}_{xn\rceil}$
n-year E.I.	$A_{xn\rceil}$
m-year deferred n-year E.I.	${}_{m\|}A_{xn\rceil}$
whole L.A. due	$\ddot{a}_x$
whole L.A. immediate	a_x
n-year temporary L.A. due	$\ddot{a}_{xn\rceil}$
n-year temporary L.A. immediate	$a_{xn\rceil}$
m-year deferred whole L.A. due	${}_{m\|}\ddot{a}_x$
m-year deferred whole L.A. immediate	${}_{m\|}a_x$
m-year deferred n-year temporary L.A. due	${}_{m\|}\ddot{a}_{xn\rceil}$
m-year deferred n-year temporary L.A. immediate	${}_{m\|}a_{xn\rceil}$

Continuous Case

Name	Notation
whole L.I.	$\overline{A}_x$
n-year term L.I.	$\overline{A}^1_{xn\rceil}$
m-year deferred L.I.	$_{m\mid}\overline{A}_x$
m-year deferred n-year term L.I.	$_{m\mid}\overline{A}^1_{xn\rceil}$
n-year pure E.I.	$\overline{A}_{x\ \overset{1}{n\rceil}}$
n-year E.I.	$\overline{A}_{xn\rceil}$
m-year deferred n-year E.I.	$_{m}\lvert\overline{A}_{xn\rceil}$
whole L.A.	$\overline{a}_x$
n-year temporary L.A.	$\overline{a}_{xn\rceil}$
m-year deferred whole L.A.	$_{m\mid}\overline{a}_x$
m-year deferred	$_{m\mid}\overline{a}_x$
m-year deferred n-year temporary L.A.	$_{m\mid}\overline{a}_{xn\rceil}$

Commutation Rates

$$
\begin{array}{rcll}
C_x &=& d_x V^{x+1} = (\ell_x - \ell_{x+1})v^{x+1} & \text{discounted number of the dead} \\
M_x &=& \sum_{s=0}^{\omega-x} C_{x+s} & \\
R_x &=& \sum_{s=0}^{\omega-x} M_{x+s} & \\
D_x &=& \ell_x v^x & \text{discounted number of living} \\
N_x &=& \sum_{s=0}^{\omega-x} D_{x+s} & \\
S_x &=& \sum_{s=0}^{\omega-x} N_{x+s} & \\
D_{xy} &=& v^{1/2(x+y)}\ell_{xy} & \text{Comm. rates for several lives} \\
N_{xy} &=& D_{xy} + D_{x+1,y+1} + \dots & \\
C_{xy} &=& v^{1/2(x+y)+1} d_{xy} & d_{xy} = \ell_{xy} - \ell_{x+1,y+1} \\
M_{xy} &=& C_{xy} + C_{x+1,y+1} + \dots &
\end{array}
$$

Appendix II

Life table man p = 3%, 1994 T Germany

Alter x	**vx**	**qx94T**	**lx**	**Dx**	**Nx**	**Sx**	**Cx**
0	1.00000000	0.011687	100,000	100,000	2,933,983	72,022,036	1,134.95
1	0.97087379	0.001008	98,831	95,952	2,833,983	69,088,053	94.26
2	0.94259591	0.000728	98,731	93,063	2,738,031	66,254,070	65.89
3	0.91514166	0.000542	98,659	90,287	2,644,968	63,516,039	47.09
4	0.88848705	0.000473	98,606	87,610	2,554,681	60,871,071	40.54
5	0.86260879	0.000452	98,559	85,018	2,467,071	58,316,390	37.69
6	0.83748426	0.000433	98,514	82,504	2,382,053	55,849,319	34.96
7	0.81309151	0.000408	98,471	80,066	2,299,549	53,467,266	31.58
8	0.78940923	0.000379	98,431	77,702	2,219,483	51,167,717	28.36
9	0.76641673	0.000352	98,394	75,411	2,141,781	48,948,234	26.04
10	0.74409391	0.000334	98,359	73,188	2,066,370	46,806,453	23.84
11	0.72242127	0.000331	98,326	71,033	1,993,182	44,740,083	23.15
12	0.70137987	0.000340	98,293	68,941	1,922,149	42,746,901	22.47
13	0.68095133	0.000371	98,260	66,910	1,853,208	40,824,752	23.80
14	0.66111780	0.000451	98,224	64,938	1,786,298	38,971,544	28.24
15	0.64186194	0.000593	98,180	63,018	1,721,360	37,185,246	36.14
16	0.62316693	0.000792	98,122	61,146	1,658,342	35,463,886	47.19
17	0.60501644	0.001040	98,044	59,318	1,597,196	33,805,544	59.91
18	0.58739460	0.001298	97,942	57,531	1,537,878	32,208,348	72.43
19	0.57028602	0.001437	97,815	55,783	1,480,347	30,670,470	78.07
20	0.55367575	0.001476	97,674	54,080	1,424,564	29,190,123	77.41
21	0.53754927	0.001476	97,530	52,427	1,370,484	27,765,559	75.15
22	0.52189250	0.001476	97,386	50,825	1,318,057	26,395,075	72.96
23	0.50669175	0.001476	97,242	49,272	1,267,232	25,077,018	70.84
24	0.49193374	0.001476	97,098	47,766	1,217,960	23,809,786	68.30
25	0.47760557	0.001476	96,955	46,306	1,170,194	22,591,826	66.31
26	0.46369473	0.001476	96,812	44,891	1,123,888	21,421,632	64.38
27	0.45018906	0.001476	96,669	43,519	1,078,997	20,297,744	62.50
28	0.43707676	0.001476	96,526	42,189	1,035,478	19,218,747	60.26
29	0.42434637	0.001476	96,384	40,900	993,289	18,183,269	58.50
30	0.41198677	0.001476	96,242	39,650	952,389	17,189,980	56.80
31	0.39998716	0.001476	96,100	38,439	912,739	16,237,591	55.14
32	0.38833705	0.001489	95,958	37,264	874,300	15,324,852	53.91
33	0.37702626	0.001551	95,815	36,125	837,036	14,450,552	54.54
34	0.36604491	0.001641	95,666	35,018	800,911	13,613,516	55.80
35	0.35538341	0.001747	95,509	33,942	765,893	12,812,605	57.62

Alter x	vx	qx94T	lx	Dx	Nx	Sx	Cx
36	0.34503244	0.001869	95,342	32,896	731,951	12,046,712	59.63
37	0.33498295	0.002007	95,164	31,878	699,055	11,314,761	62.12
38	0.32522617	0.002167	94,973	30,888	667,177	10,615,706	65.05
39	0.31575356	0.002354	94,767	29,923	636,289	9,948,529	68.36
40	0.30655685	0.002569	94,544	28,983	606,366	9,312,240	72.32
41	0.29762801	0.002823	94,301	28,067	577,383	8,705,874	76.86
42	0.28895923	0.003087	94,035	27,172	549,316	8,128,491	81.36
43	0.28054294	0.003387	93,745	26,299	522,144	7,579,175	86.61
44	0.27237179	0.003726	93,427	25,447	495,845	7,057,031	92.02
45	0.26443863	0.004100	93,079	24,614	470,398	6,561,186	98.07
46	0.25673653	0.004522	92,697	23,799	445,784	6,090,788	104.44
47	0.24925877	0.004983	92,278	23,001	421,985	5,645,004	111.32
48	0.24199881	0.005508	91,818	22,220	398,984	5,223,019	118.88
49	0.23495030	0.006094	91,312	21,454	376,764	4,824,035	126.83
50	0.22810709	0.006751	90,756	20,702	355,310	4,447,271	135.76
51	0.22146319	0.007485	90,143	19,963	334,608	4,091,961	145.13
52	0.21501281	0.008302	89,468	19,237	314,645	3,757,353	155.10
53	0.20875030	0.009215	88,725	18,521	295,408	3,442,708	165.78
54	0.20267019	0.010195	87,907	17,816	276,887	3,147,300	176.30
55	0.19676717	0.011236	87,011	17,121	259,071	2,870,413	186.83
56	0.19103609	0.012340	86,033	16,435	241,950	2,611,342	196.97
57	0.18547193	0.013519	84,971	15,760	225,515	2,369,392	206.90
58	0.18006983	0.014784	83,822	15,094	209,755	2,143,877	216.61
59	0.17482508	0.016150	82,583	14,438	194,661	1,934,122	226.42
60	0.16973309	0.017625	81,249	13,791	180,223	1,739,461	235.98
61	0.16478941	0.019223	79,817	13,153	166,432	1,559,238	245.42
62	0.15998972	0.020956	78,283	12,524	153,279	1,392,806	254.74
63	0.15532983	0.022833	76,643	11,905	140,755	1,239,527	263.91
64	0.15080566	0.024858	74,893	11,294	128,850	1,098,772	272.62
65	0.14641326	0.027073	73,031	10,693	117,556	969,922	281.03
66	0.14214880	0.029552	71,054	10,100	106,863	852,366	289.82
67	0.13800854	0.032350	68,954	9,516	96,763	745,503	298.93
68	0.13398887	0.035632	66,723	8,940	87,247	648,740	309.22
69	0.13008628	0.039224	64,346	8,371	78,307	561,493	318.77
70	0.12629736	0.043127	61,822	7,808	69,936	483,186	326.90
71	0.12261880	0.047400	59,156	7,254	62,128	413,250	333.81
72	0.11904738	0.052110	56,352	6,709	54,874	351,122	339.46
73	0.11557998	0.057472	53,415	6,174	48,165	296,248	344.50
74	0.11221357	0.063440	50,345	5,649	41,991	248,083	347.97

Alter x	vx	qx94T	lx	Dx	Nx	Sx	Cx
75	0.10894521	0.070039	47,151	5,137	36,342	206,092	349.26
76	0.10577205	0.077248	43,849	4,638	31,205	169,750	347.82
77	0.10269131	0.085073	40,462	4,155	26,567	138,545	343.17
78	0.09970030	0.093534	37,020	3,691	22,412	111,978	335.21
79	0.09679641	0.102662	33,557	3,248	18,721	89,566	323.75
80	0.09397710	0.112477	30,112	2,830	15,473	70,845	309.03
81	0.09123990	0.122995	26,725	2,438	12,643	55,372	291.17
82	0.08858243	0.134231	23,438	2,076	10,205	42,729	270.56
83	0.08600236	0.146212	20,292	1,745	8,129	32,524	247.74
84	0.08349744	0.158964	17,325	1,447	6,384	24,395	223.25
85	0.08106548	0.172512	14,571	1,181	4,937	18,011	197.86
86	0.07870435	0.186896	12,057	949	3,756	13,074	172.16
87	0.07641199	0.202185	9,804	749	2,807	9,318	147.04
88	0.07418640	0.218413	7,822	580	2,058	6,511	123.02
89	0.07202563	0.235597	6,114	440	1,478	4,453	100.70
90	0.06992780	0.253691	4,674	327	1,038	2,975	80.52
91	0.06789107	0.272891	3,488	237	711	1,937	62.75
92	0.06591366	0.293142	2,536	167	474	1,226	47.55
93	0.06399384	0.314638	1,793	115	307	752	35.04
94	0.06212994	0.337739	1,229	76	192	445	25.03
95	0.06032033	0.362060	814	49	116	253	17.28
96	0.05856343	0.388732	519	30	67	137	11.49
97	0.05685770	0.419166	317	18	37	70	7.34
98	0.05520165	0.452008	184	10	19	33	4.45
99	0.05359383	0.486400	101	5	9	14	2.55
100	0.05203284	0.527137	52	3	4	5	1.36
101	0.05051732	1.000000	25	1	1	1	1.23

Life table man p = 4%, 1994 T Germany

Alter x	vx	qx94T	lx	Dx	Nx	Sx	Cx
0	1.00000000	0.011687	100,000	100,000	2,375,516	50,492,905	1,124.04
1	0.96153846	0.001008	98,831	95,030	2,275,516	48,117,389	92.46
2	0.92455621	0.000728	98,731	91,282	2,180,486	45,841,873	64.01
3	0.88899636	0.000542	98,659	87,707	2,089,204	43,661,387	45.30
4	0.85480419	0.000473	98,606	84,289	2,001,497	41,572,183	38.63
5	0.82192711	0.000452	98,559	81,008	1,917,208	39,570,686	35.56
6	0.79031453	0.000433	98,514	77,857	1,836,200	37,653,478	32.68
7	0.75991782	0.000408	98,471	74,830	1,758,343	35,817,278	29.23
8	0.73069021	0.000379	98,431	71,923	1,683,513	34,058,935	26.00
9	0.70258674	0.000352	98,394	69,130	1,611,590	32,375,422	23.64
10	0.67556417	0.000334	98,359	66,448	1,542,460	30,763,832	21.44
11	0.64958093	0.000331	98,326	63,871	1,476,012	29,221,372	20.61
12	0.62459705	0.000340	98,293	61,394	1,412,141	27,745,360	19.82
13	0.60057409	0.000371	98,260	59,012	1,350,747	26,333,219	20.79
14	0.57747509	0.000451	98,224	56,722	1,291,735	24,982,472	24.43
15	0.55526451	0.000593	98,180	54,516	1,235,013	23,690,737	30.97
16	0.53390818	0.000792	98,122	52,388	1,180,497	22,455,724	40.04
17	0.51337325	0.001040	98,044	50,333	1,128,109	21,275,227	50.35
18	0.49362813	0.001298	97,942	48,347	1,077,776	20,147,118	60.28
19	0.47464243	0.001437	97,815	46,427	1,029,429	19,069,342	64.35
20	0.45638695	0.001476	97,674	44,577	983,002	18,039,913	63.19
21	0.43883361	0.001476	97,530	42,799	938,425	17,056,911	60.76
22	0.42195539	0.001476	97,386	41,093	895,626	16,118,486	58.42
23	0.40572634	0.001476	97,242	39,454	854,533	15,222,860	56.18
24	0.39012148	0.001476	97,098	37,880	815,079	14,368,327	53.64
25	0.37511681	0.001476	96,955	36,369	777,199	13,553,248	51.58
26	0.36068924	0.001476	96,812	34,919	740,830	12,776,049	49.59
27	0.34681658	0.001476	96,669	33,526	705,911	12,035,219	47.69
28	0.33347748	0.001476	96,526	32,189	672,385	11,329,308	45.53
29	0.32065142	0.001476	96,384	30,906	640,196	10,656,923	43.78
30	0.30831867	0.001476	96,242	29,673	609,290	10,016,727	42.10
31	0.29646026	0.001476	96,100	28,490	579,617	9,407,437	40.48
32	0.28505794	0.001489	95,958	27,354	551,127	8,827,820	39.20
33	0.27409417	0.001551	95,815	26,262	523,773	8,276,693	39.27
34	0.26355209	0.001641	95,666	25,213	497,511	7,752,920	39.79
35	0.25341547	0.001747	95,509	24,203	472,298	7,255,409	40.69

Alter x	vx	qx94T	lx	Dx	Nx	Sx	Cx
36	0.24366872	0.001869	95,342	23,232	448,095	6,783,111	41.70
37	0.23429685	0.002007	95,164	22,297	424,863	6,335,016	43.03
38	0.22528543	0.002167	94,973	21,396	402,566	5,910,153	44.62
39	0.21662061	0.002354	94,767	20,528	381,170	5,507,587	46.45
40	0.20828905	0.002569	94,544	19,692	360,642	5,126,417	48.67
41	0.20027793	0.002823	94,301	18,886	340,950	4,765,775	51.22
42	0.19257493	0.003087	94,035	18,109	322,064	4,424,825	53.70
43	0.18516820	0.003387	93,745	17,359	303,955	4,102,761	56.62
44	0.17804635	0.003726	93,427	16,634	286,596	3,798,806	59.58
45	0.17119841	0.004100	93,079	15,935	269,962	3,512,210	62.88
46	0.16461386	0.004522	92,697	15,259	254,027	3,242,248	66.32
47	0.15828256	0.004983	92,278	14,606	238,768	2,988,221	70.01
48	0.15219477	0.005508	91,818	13,974	224,162	2,749,453	74.05
49	0.14634113	0.006094	91,312	13,363	210,188	2,525,291	78.24
50	0.14071263	0.006751	90,756	12,771	196,825	2,315,103	82.94
51	0.13530061	0.007485	90,143	12,196	184,054	2,118,278	87.82
52	0.13009674	0.008302	89,468	11,639	171,858	1,934,224	92.94
53	0.12509302	0.009215	88,725	11,099	160,219	1,762,366	98.39
54	0.12028175	0.010195	87,907	10,574	149,120	1,602,147	103.63
55	0.11565553	0.011236	87,011	10,063	138,546	1,453,027	108.76
56	0.11120724	0.012340	86,033	9,567	128,483	1,314,481	113.56
57	0.10693004	0.013519	84,971	9,086	118,916	1,185,998	118.14
58	0.10281735	0.014784	83,822	8,618	109,830	1,067,082	122.49
59	0.09886284	0.016150	82,583	8,164	101,212	957,252	126.81
60	0.09506042	0.017625	81,249	7,724	93,048	856,040	130.89
61	0.09140425	0.019223	79,817	7,296	85,324	762,992	134.82
62	0.08788870	0.020956	78,283	6,880	78,028	677,668	138.59
63	0.08450837	0.022833	76,643	6,477	71,148	599,640	142.20
64	0.08125805	0.024858	74,893	6,086	64,671	528,492	145.48
65	0.07813274	0.027073	73,031	5,706	58,585	463,821	148.53
66	0.07512763	0.029552	71,054	5,338	52,879	405,236	151.70
67	0.07223811	0.032350	68,954	4,981	47,541	352,357	154.96
68	0.06945972	0.035632	66,723	4,635	42,560	304,816	158.76
69	0.06678819	0.039224	64,346	4,298	37,925	262,256	162.09
70	0.06421941	0.043127	61,822	3,970	33,627	224,331	164.62
71	0.06174943	0.047400	59,156	3,653	29,657	190,704	166.49
72	0.05937445	0.052110	56,352	3,346	26,004	161,047	167.68
73	0.05709082	0.057472	53,415	3,050	22,658	135,043	168.53
74	0.05489502	0.063440	50,345	2,764	19,608	112,385	168.59

Alter x	**vx**	**qx94T**	**lx**	**Dx**	**Nx**	**Sx**	**Cx**
75	0.05278367	0.070039	47,151	2,489	16,844	92,777	167.59
76	0.05075353	0.077248	43,849	2,225	14,355	75,933	165.29
77	0.04880147	0.085073	40,462	1,975	12,130	61,578	161.51
78	0.04692449	0.093534	37,020	1,737	10,155	49,448	156.25
79	0.04511970	0.102662	33,557	1,514	8,418	39,293	149.46
80	0.04338433	0.112477	30,112	1,306	6,904	30,875	141.29
81	0.04171570	0.122995	26,725	1,115	5,598	23,971	131.85
82	0.04011125	0.134231	23,438	940	4,483	18,373	121.34
83	0.03856851	0.146212	20,292	783	3,543	13,890	110.03
84	0.03708511	0.158964	17,325	642	2,760	10,347	98.20
85	0.03565876	0.172512	14,571	520	2,118	7,587	86.20
86	0.03428727	0.186896	12,057	413	1,598	5,469	74.28
87	0.03296853	0.202185	9,804	323	1,185	3,871	62.83
88	0.03170051	0.218413	7,822	248	862	2,686	52.06
89	0.03048126	0.235597	6,114	186	614	1,824	42.20
90	0.02930890	0.253691	4,674	137	428	1,210	33.42
91	0.02818163	0.272891	3,488	98	291	782	25.80
92	0.02709772	0.293142	2,536	69	193	491	19.36
93	0.02605550	0.314638	1,793	47	124	298	14.13
94	0.02505337	0.337739	1,229	31	77	174	10.00
95	0.02408978	0.362060	814	20	46	97	6.83
96	0.02316325	0.388732	519	12	26	51	4.50
97	0.02227236	0.419166	317	7	14	25	2.85
98	0.02141573	0.452008	184	4	7	11	1.71
99	0.02059205	0.486400	101	2	3	4	0.97
100	0.01980005	0.527137	52	1	1	1	0.51
101	0.01903851	1.000000	25	0	0	0	0.46

Life table man p = 4%

Alter x	vx	qx94T	lx	Dx	Nx	Sx	Cx
0	1.00000000	0.011687	100,000	100,000	1,983,347	36,771,267	1,113.33
1	0.95238095	0.001008	98,831	94,125	1,883,347	34,787,920	90.70
2	0.90702948	0.000728	98,731	89,552	1,789,222	32,904,573	62.20
3	0.86383760	0.000542	98,659	85,225	1,699,670	31,115,351	43.60
4	0.82270248	0.000473	98,606	81,123	1,614,445	29,415,681	36.83
5	0.78352617	0.000452	98,559	77,224	1,533,322	27,801,236	33.58
6	0.74621540	0.000433	98,514	73,513	1,456,098	26,267,914	30.56
7	0.71068133	0.000408	98,471	69,982	1,382,585	24,811,816	27.07
8	0.67683936	0.000379	98,431	66,622	1,312,603	23,429,231	23.85
9	0.64460891	0.000352	98,394	63,426	1,245,981	22,116,628	21.49
10	0.61391325	0.000334	98,359	60,384	1,182,555	20,870,647	19.29
11	0.58467929	0.000331	98,326	57,489	1,122,171	19,688,092	18.38
12	0.55683742	0.000340	98,293	54,733	1,064,682	18,565,921	17.50
13	0.53032135	0.000371	98,260	52,109	1,009,949	17,501,239	18.18
14	0.50506795	0.000451	98,224	49,610	957,840	16,491,290	21.16
15	0.48101710	0.000593	98,180	47,226	908,230	15,533,450	26.57
16	0.45811152	0.000792	98,122	44,951	861,004	14,625,220	34.03
17	0.43629669	0.001040	98,044	42,776	816,053	13,764,216	42.38
18	0.41552066	0.001298	97,942	40,697	773,277	12,948,163	50.26
19	0.39573396	0.001437	97,815	38,709	732,580	12,174,886	53.14
20	0.37688949	0.001476	97,674	36,812	693,871	11,442,306	51.69
21	0.35894237	0.001476	97,530	35,008	657,059	10,748,435	49.23
22	0.34184988	0.001476	97,386	33,291	622,051	10,091,376	46.88
23	0.32557131	0.001476	97,242	31,659	588,760	9,469,325	44.65
24	0.31006791	0.001476	97,098	30,107	557,101	8,880,565	42.23
25	0.29530277	0.001476	96,955	28,631	526,994	8,323,464	40.22
26	0.28124073	0.001476	96,812	27,227	498,363	7,796,470	38.30
27	0.26784831	0.001476	96,669	25,893	471,136	7,298,107	36.48
28	0.25509363	0.001476	96,526	24,623	445,243	6,826,971	34.50
29	0.24294631	0.001476	96,384	23,416	420,620	6,381,728	32.86
30	0.23137744	0.001476	96,242	22,268	397,204	5,961,108	31.29
31	0.22035947	0.001476	96,100	21,177	374,936	5,563,904	29.80
32	0.20986616	0.001489	95,958	20,138	353,759	5,188,968	28.58
33	0.19987253	0.001551	95,815	19,151	333,621	4,835,209	28.36
34	0.19035479	0.001641	95,666	18,210	314,470	4,501,588	28.46
35	0.18129028	0.001747	95,509	17,315	296,260	4,187,118	28.83

Alter x	**vx**	**qx94T**	**lx**	**Dx**	**Nx**	**Sx**	**Cx**
36	0.17265741	0.001869	95,342	16,462	278,945	3,890,858	29.27
37	0.16443563	0.002007	95,164	15,648	262,483	3,611,913	29.91
38	0.15660536	0.002167	94,973	14,873	246,835	3,349,430	30.72
39	0.14914796	0.002354	94,767	14,134	231,962	3,102,595	31.68
40	0.14204568	0.002569	94,544	13,430	217,828	2,870,633	32.87
41	0.13528160	0.002823	94,301	12,757	204,398	2,652,805	34.27
42	0.12883962	0.003087	94,035	12,115	191,641	2,448,407	35.58
43	0.12270440	0.003387	93,745	11,503	179,526	2,256,766	37.16
44	0.11686133	0.003726	93,427	10,918	168,023	2,077,240	38.73
45	0.11129650	0.004100	93,079	10,359	157,105	1,909,217	40.49
46	0.10599667	0.004522	92,697	9,826	146,746	1,752,112	42.30
47	0.10094921	0.004983	92,278	9,315	136,920	1,605,366	44.23
48	0.09614210	0.005508	91,818	8,828	127,605	1,468,446	46.33
49	0.09156390	0.006094	91,312	8,361	118,777	1,340,841	48.49
50	0.08720371	0.006751	90,756	7,914	110,416	1,222,064	50.91
51	0.08305115	0.007485	90,143	7,486	102,502	1,111,648	53.39
52	0.07909633	0.008302	89,468	7,077	95,016	1,009,146	55.97
53	0.07532984	0.009215	88,725	6,684	87,939	914,130	58.69
54	0.07174270	0.010195	87,907	6,307	81,255	826,191	61.22
55	0.06832638	0.011236	87,011	5,945	74,948	744,936	63.64
56	0.06507274	0.012340	86,033	5,598	69,003	669,988	65.82
57	0.06197404	0.013519	84,971	5,266	63,405	600,985	67.82
58	0.05902290	0.014784	83,822	4,947	58,139	537,580	69.65
59	0.05621229	0.016150	82,583	4,642	53,192	479,441	71.42
60	0.05353551	0.017625	81,249	4,350	48,550	426,249	73.01
61	0.05098620	0.019223	79,817	4,070	44,200	377,699	74.49
62	0.04855829	0.020956	78,283	3,801	40,130	333,499	75.84
63	0.04624599	0.022833	76,643	3,544	36,329	293,369	77.08
64	0.04404380	0.024858	74,893	3,299	32,785	257,040	78.10
65	0.04194648	0.027073	73,031	3,063	29,486	224,255	78.98
66	0.03994903	0.029552	71,054	2,839	26,423	194,769	79.90
67	0.03804670	0.032350	68,954	2,623	23,584	168,346	80.84
68	0.03623495	0.035632	66,723	2,418	20,961	144,762	82.03
69	0.03450948	0.039224	64,346	2,221	18,543	123,801	82.95
70	0.03286617	0.043127	61,822	2,032	16,322	105,258	83.45
71	0.03130111	0.047400	59,156	1,852	14,290	88,936	83.59
72	0.02981058	0.052110	56,352	1,680	12,438	74,646	83.38
73	0.02839103	0.057472	53,415	1,517	10,758	62,208	83.01
74	0.02703908	0.063440	50,345	1,361	9,241	51,450	82.25

Alter x	vx	qx94T	lx	Dx	Nx	Sx	Cx
75	0.02575150	0.070039	47,151	1,214	7,880	42,209	80.98
76	0.02452524	0.077248	43,849	1,075	6,666	34,329	79.11
77	0.02335737	0.085073	40,462	945	5,591	27,663	76.57
78	0.02224511	0.093534	37,020	824	4,646	22,072	73.37
79	0.02118582	0.102662	33,557	711	3,822	17,426	69.51
80	0.02017697	0.112477	30,112	608	3,111	13,604	65.09
81	0.01921616	0.122995	26,725	514	2,503	10,493	60.16
82	0.01830110	0.134231	23,438	429	1,989	7,990	54.83
83	0.01742962	0.146212	20,292	354	1,560	6,001	49.25
84	0.01659964	0.158964	17,325	288	1,206	4,441	43.54
85	0.01580918	0.172512	14,571	230	918	3,235	37.85
86	0.01505636	0.186896	12,057	182	688	2,317	32.31
87	0.01433939	0.202185	9,804	141	506	1,629	27.07
88	0.01365656	0.218413	7,822	107	365	1,123	22.21
89	0.01300625	0.235597	6,114	80	258	758	17.84
90	0.01238690	0.253691	4,674	58	178	500	13.99
91	0.01179705	0.272891	3,488	41	120	322	10.70
92	0.01123529	0.293142	2,536	28	79	202	7.95
93	0.01070028	0.314638	1,793	19	51	123	5.75
94	0.01019074	0.337739	1,229	13	32	72	4.03
95	0.00970547	0.362060	814	8	19	40	2.73
96	0.00924330	0.388732	519	5	11	21	1.78
97	0.00880314	0.419166	317	3	6	10	1.12
98	0.00838394	0.452008	184	2	3	4	0.66
99	0.00798470	0.486400	101	1	1	1	0.37
100	0.00760448	0.527137	52	0	0	0	0.20
101	0.00724236	1.000000	25	0	0	0	0.17

4.2 A Continuous Presentation of Life Insurance

can be omitted at first reading

This part is more of a theoretical nature and may be omitted at first reading. So far we have developed the mathematics of life insurance essentially for the discrete model, which is conceptually easier. However, the continuous description allows the use of calculus, which may lead to some simplifications. On the other hand, it should also be clear that realistic data can only be handled numerically. Even if one uses a "realistic law", like the Makeham Gompertz law or the Heligman Pollard expression, the integrals in general can not be evaluated in closed form. In the continuous version the life table $(\ell_x)_{x=1}^{\omega}$ is replaced by a real valued function ℓ on $[0, \omega]$. Thus $\ell(x)/\ell(0)$ is the probability of reaching age x. Then

$$ {}_tq_x = \frac{\ell(x) - \ell(x+t)}{\ell(x)} \tag{4.2.33} $$

is the probability of an x-year-old dying in $[x, x+t]$. Similarly, we let

$$ {}_tp_x = \frac{\ell(x+t)}{\ell(x)} \tag{4.2.34} $$

be the probability that an x-year-old reaches $x+t$ years.
Previously, we defined the force of mortality as

$$ \mu(x) = \lim_{t\to 0_+} \frac{\ell(x) - \ell(x+t)}{t\ \ell(x)} = \frac{\ell'(x)}{\ell(x)} = -(\ln \ell(x))'. \tag{4.2.35} $$

Remark. If one has a discrete representation $(\ell_x)_{x=1}^{\omega}$ and wants to apply the continuous version, one should either use an appropriate mortality law or determine $\ell(x)$ by interpolation. From (4.2.35) we get

$$ \ln\left(\frac{\ell(x)}{\ell(x+t)}\right) = -\ln \ell(x+t) + \ln \ell(x) = \int\limits_x^{x+t} \mu(s)ds $$

or

$$ {}_tp_x = \frac{\ell(x+t)}{\ell(x)} = \exp\left(-\int\limits_x^{x+t} \mu(s)ds\right) \tag{4.2.36} $$

and

$$ {}_tq_x = 1 - \exp\left(-\int\limits_x^{x+t} \mu(s)ds\right). \tag{4.2.37} $$

Now $t \to_t q_x$ is the distribution function of the random variable M_x, which describes the mortality of an x-year-old person. If M_x has a continuous distribution its density is given by

$$ f_x(t) = \frac{d}{dt}{}_tq_x = \mu(x+t)e^{-\int_x^{x+t}\mu(s)ds} = \mu(x+t){}_tp_x. \tag{4.2.38} $$

Assume the insurance pays $D(t)$ if the insured dies at $x+t$. Then the discounted value is $D(t)v^t$ so that the expectation of the present value becomes

$$ E(D(t)v^t) = \int\limits_0^{\infty} \mu(x+t){}_tp_x v^t D(t)dt. \tag{4.2.39} $$

The variance of this can be computed with the well-known expression

$$ \sigma_x^2 = E(X^2) - E(X)^2. $$

Clearly

$$ E(D(t)v^t)^2) = \int\limits_0^{\infty} \mu(x+t){}_tp_x v^{2t} D(t)^2 dt. $$

These expressions may be slightly simplified if one writes $v^t = e^{-\delta t} = \exp(\int_0^t \delta ds) = \exp(-\int_x^{x+t} \delta ds)$. This latter expression suggests interpreting δ as the "mortality" of capital. With this (4.2.39) becomes

$$E(D(t)v^t) = \int_0^\infty \mu(x+t) e^{-\int_x^{x+t}(\mu(s)+\delta)ds} D(t)dt. \tag{4.2.40}$$

In general, the integral will, of course, only extend to $\omega - x$ because then ${}_tp_x = 0$. If $D(t)$ is not too complicated this may actually be computed.

Example 4.4 Consider $D(t) = 1$ and $\mu(x) = \dfrac{1}{86-x}$, which is Moivre's mortality law. Moreover, assume $x \geq 20$, because this law is not valid for small x

$$\begin{aligned}
E(v^t) &= \int_0^{86-x} \frac{1}{86-x-t} \left(\exp - \int_x^{x+t} \left(\delta + \frac{1}{86-s}\right) ds\right) dt \\
&= \int_0^{86-x} \frac{e^{-t\delta}}{86-x-t} exp\left(-\ln\left(\frac{86-x}{86-x-t}\right)\right| dt \\
&= \int_0^{86-x} \frac{e^{-t\delta}}{(86-x-t)} \frac{86-x-t}{86-x} dt \\
&= \frac{1}{86-x} \int_0^{86-x} e^{-t\delta} dt = \frac{1}{86-x}\frac{1}{\delta}\left(1 - e^{-\delta(86-x)}\right) \\
&= \overline{A}_x.
\end{aligned}$$

Expressions based on a continuous model are usually indicated by an overhead bar. As a particular case consider $x = 40$, $p = 4\%$ and $D = 1$. Then $v = 1.04^{-1} = e^{-\delta}$ or $e^\delta = 1.04$ or $\delta = \ln 1.04 = 0.039221$

$$E(Dv^t) = .463035.$$

Then

$$E((Dv^t)^2) = \int_0^{86-x} \frac{1}{86-x-t} \left(e^{\int_x^{x+t}(\mu(s)+2\delta)ds}\right) D^2 dt \tag{4.2.41}$$

and as above we are led to

$$E(D^2 v^{2t}) = \frac{1}{86-x} \int_0^{86-x} e^{2\delta t} dt = \frac{1}{86-x}\frac{1}{2\delta}\left(1 - e^{2\delta(86-x)}\right).$$

In our special case this is

$$E(D^2 v^{2t}) = .269629.$$

Thus in this case the variance is .0552271 and the standard deviation .235.

Problem 4.18 With the same mortality law as above compute the current value of the term insurance for $x = 30$, $n = 35$ and $p = 4\%$.

Problem 4.19 Derive a formula for the k-th moment of the above whole life insurance.

Example 4.5 Assume the force of mortality to be constant and assume p to be constant. Then a deferred, whole or term insurance is independent of the age of the insured. To be explicit, consider an m-year deferred term insurance for n years, which pays the amount 1, in the case of death. Then the density (4.2.37) is given by

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

which is independent of x as expected. This distribution is an exponential distribution and analogously to (4.2.37) we get for $m < n$ with $v^t = e^{-\delta t}$

$$_{m|}\overline{A}_{n\rceil} = \int_m^n \mu e^{-\mu t} \cdot e^{-\delta t} dt = \frac{\mu}{\mu+\delta}\left(e^{-(\mu+\delta)m} - e^{-(\mu+\delta)n}\right).$$

The k-th moment of this variable becomes

$$_{m|}{}^{k}\overline{A}_{n\rceil} = \int_m^n \mu e^{-\mu t} e^{-k\delta t} dt = \frac{\mu}{\mu+k\delta}\left(e^{-(\mu+k\delta)m} - e^{-(\mu+k\delta)n}\right).$$

The variance of the corresponding whole insurance $n \to \infty$, is therefore

$$\frac{\mu}{\mu+2\delta} e^{-(\mu+2\delta)m} - \left(\frac{\mu}{\mu+\delta}\right)^2 e^{-2(\mu+\delta)m}.$$

Problem 4.20 Compute this insurance with increasing benefits $e^{\gamma t}$, where $e^{\gamma t} = (1 - g/100)^{-t}$ is considered as a balance to inflation.

Let us extend this now to the case of an endowment insurance for a person of age x for a time of n years. Such insurance is actually the sum of two insurances, one of which is payable in the case of death within $[x, x+n]$, while the other is payable in the case that the person survives n years.

Let us write

$$X = D + S$$

for the obligation of the insurance company, where D refers to death and S to survival. Then

$$E(D) = \int_0^n \mu(x+t)_t p_x v^t D(t) dt \tag{4.2.42}$$

and

$$E(S) = {}_n p_x v^n S(n). \tag{4.2.43}$$

Here $D(t)$ is due if death occurs at $x+t$, while $S(n)$ is due if he is still alive in the year $x+n$.

In order to compute the variance, we use the rule

$$\begin{aligned}\sigma^2_{X+Y} =& E((X+Y-EX-EY)^2) = E((X-EX)^2) + E((Y-EY)^2)\\ &+ 2E((X-EX)(Y-EY))\\ &= \sigma^2_X + \sigma^2_Y + 2\sigma_{XY},\end{aligned}$$

i.e. Variance $(X+Y)$ = Variance X + Variance Y + 2 Covariance (X,Y) which holds for all random variables. In addition, we need

$$E((X-EX)(Y-EY)) = E(X(Y-EY)) = E(XY) - EX\ EY.$$

This formula is useful, because here $S \cdot D = 0$. Thus

$$\begin{aligned}\sigma^2_{S+D} =& \sigma^2_X + \sigma^2_D - 2E(D)\cdot E(S)\\ &= \int_0^n \mu(x+t)_t p_x (v^t D(t))^2 dt - \left[\int_0^n \mu(x+t)_t p_x v^t D(t)\right]^2\\ &+ v^{2n} S(n)^2_n p_x (1 -_n p_x) - 2v^n_n p_x S(n) E(D).\end{aligned}$$

Problem 4.21 Formulate the continuous version of a deferred insurance of the type considered above.

Continuous Annuities

The present value for an annuity 1 due for n years at interest rate p is

$$1 + v + \cdots + v^{n-1} = \frac{1-v^n}{1-v} \qquad v = \left(1 + \frac{p}{100}\right)^{-1}.$$

Now assume the sum $1/k$ is paid k times a year and interest is compounded every year$/k$. Then the present value becomes

$$\frac{1}{k}[1 + v_k + \cdots + v_k^{nk-1}] = \frac{1-v_k^{nk}}{k(1-v_k)}$$

with $v_k = v^{1/k}$. It should be noted that this differs slightly from our previous definition of interest rates per year$/k$. If $v = e^{-\delta}$, $v_k = e^{-\delta/k}$ so that the limit for this becomes

$$\overline{a}_n = \frac{(1-v^n)}{\delta} \tag{4.2.44}$$

because

$$\lim_{k\to\infty} k(1-e^{-\delta/k}) = \lim_k k\left(1 - 1 + \frac{\delta}{k} - \frac{\delta^2}{2k^2} + \dots\right) = \delta.$$

We could also have obtained this formula by integrating the discounted flow $e^{-\delta t}$ from 0 to n.

Thus if an x-year-old dies at time $x+t$, the present value $\overline{a}_t = 1 - v^t/\delta$ in annuities at rate 1 at age $x+t$ will be $\mu(x+t)_t p_x$. Thus we get

$$\overline{a}_x = \int_0^\infty \overline{a}_t \mu(x+t)_t p_x dt \ \overline{a}_t = \frac{(1-v^t)}{\delta} \tag{4.2.45}$$

for the present value of a whole life annuity of an x-year-old.

Problem 4.22 Assume $\ell(t) = \ell_0(100-t)/100$ for $0 \le t \le 100$. Derive the formulas for a term, whole and deferred insurance similar to the example above.

This expression can be simplified somewhat by using a simple result from probability theory.

Lemma 4.1 Let X be a random variable which has only non-negative values. Let F be its distribution function and let h be differentiable. Then

$$E(h(X)) = h(0) + \int_0^\infty h'(t)(1-F(t))dt \tag{4.2.46}$$

provided the expectation value and the integral exist.

Proof.

$$\begin{aligned} E(h(X)) &= \lim_{b\to\infty} \int_0^b h(x)d(1-F(x)) \\ &= \lim_{b\to\infty} h(x)(1-F(x)) \mid_0^b + \int_0^b h'(t)(1-F(t))dt \\ &= h(0) + \int_0^\infty h'(t)(1-F(t))dt \end{aligned}$$

by partial integration.

Here we apply this Lemma to $F(t) =_t q_x$, $1 - F(t) =_t p_x$ and $h(t) = \overline{a}_t = (1-v^t)/\delta$. Then

$$\overline{a}_x = \int_0^\infty v^t_t p_x dt \tag{4.2.47}$$

For the variance we get similarly

$$\begin{aligned} \text{Variance} &= \sigma^2\left(\frac{1-v^T}{\delta}\right) = \sigma^2\left(\frac{v^T}{\delta}\right) = \delta^{-2}\left[\int_0^\infty v^{2t}\mu(x+t)_t p_x dt - \left(\int_0^\infty v^t\mu(x+t)_t p_x\right)^2\right] \\ &= ({}^2\overline{A}_x - \overline{A}_x^2)\delta^{-2}. \end{aligned}$$

For a whole life insurance with face value 1 we get similarly

$$\overline{A}_x = \int_0^\infty v^t \mu(x+t)_t p_x dt = \int_0^\infty e^{-\delta t} \mu(x+t)_t p_x dt \tag{4.2.48}$$

$$= e^{-\delta \cdot 0} - \delta \int_0^\infty e^{-\delta t} \ {}_t p_x dt = 1 - \delta \overline{a}_x.$$

Problem 4.23 Show $\overline{A}_{x,n} = 1 - \delta \overline{a}_{x,n}$ for the term insurance by the same method.

These results of course apply immediately to the payment of premiums, which after all is nothing but an annuity paid by the insured to the company. Thus if a person of age x has bought an insurance whose present expected value is K, his level lifelong premium Pr will be

$$K = Pr \overline{a}_x$$

or

$$K = Pr \overline{a}_{x,n\rceil}$$

if he has to pay until his death or at most for n years.

Summary: It is possible to develop the theory of life insurance for the continuous situation in parallel to the discrete case. This, however, is of limited practical importance, because in realistic situations the required computations can only be done numerically.

4.3 Insurance with respect to several lives

So far we have considered insurance for one life. However, it is also possible to consider insurances for several lives. Examples are pensions for widows, dowry or education funds or annuities. For simplicity we will only consider an insurance for two lives.

Mortality The two persons will be denoted by x and y, and x and y also gives their ages. Even though the mortalities for a couple x, y will hardly be independent we will assume this and also use the same life table for both. Then

$${}_t p_{xy} =_t p_x \cdot_t p_y \tag{4.3.49}$$

is the probability that a couple (x, y) is still alive after t years. Similarly

$${}_t q_{xy} = 1 -_t p_{xy} \tag{4.3.50}$$

is the probability that at least one of the couple has died in this time. The probability that both die in this time is

$${}_t q_{\overline{xy}} =_t q_x \, {}_t q_y. \tag{4.3.51}$$

Problem 4.24 1. Show that the probability that at least one of (x, y) is alive after t years is

$$_tp_{\overline{xy}} = 1 -_t q_{\overline{xy}} =_t p_x =_t p_x +_t q_x \, _tp_y. \tag{4.3.52}$$

2. The probability that after t years exactly one of x or y is alive is

$$_tp^1_{xy} =_t p_x +_t p_y - 2_tp_{xy}. \tag{4.3.53}$$

Problem 4.25 1. Show that the probability $_t(qp)_{xy}$ that x dies in the $(t+1)$-st year and that y is alive at the end of this year is given by

$$_t(qp)_{xy} =_t p_{xy} q_{x+t} p_{y+t}. \tag{4.3.54}$$

2. What does $_{t|}q_{xy} =_t p_{xy} -_{t+1} p_{xy}$ stand for?

3. Find an expression for $_t(pp)_{xy}$, the probability that neither x nor y dies in the year $t+1$.

Once the combined mortalities have been defined, one can define premiums for all sorts of insurances for such couples. The guiding principle is again the equivalence principle for the net premium, while the gross premiums are determined as above. Since the underlying principles are always the same we shall only consider a number of examples.

Endowment Insurance The insurance pays (xy) the face amount 1 after n years, if both are still alive by then. The present value for this is

$$\begin{aligned} _nE_{xy} =_n p_{xy} v^n &= \frac{\ell_{x+n}\ell_{y+n}}{\ell_x \ell_y} v^n = \frac{v^{1/2(x+y+2n)}\ell_{x+n}\ell_{y+n}}{v^{1/2(x+y)}\ell_x\ell_y} \\ &= \frac{D_{x+n,y+n}}{D_{xy}}. \end{aligned} \tag{4.3.55}$$

Here $D_{xy} = v^{1/2(x+y)}\ell_{xy}$ is again a suitable commutation number.

Annuity due for n years The company pays an annual pension at the beginning of each year as long as both (xy) are still alive, but at most for n years. This can be considered a superposition of the previous examples. Thus the present value is

$$\begin{aligned} \ddot{a}_{xy,\overline{n}|} &= \sum_{v=0}^{n-1} {}_vp_{xy} v^v = D_{xy}^{-1}(D_{xy} + D_{x+1,y+1} + \ldots + D_{x+n-1,y+n-1} \\ &= (N_{xy} - N_{x+n,y+n})D_{xy}^{-1} \end{aligned} \tag{4.3.56}$$

where $N_{xy} = D_{xy} + D_{x+1,y+1} + D_{x+2,y+2} + \ldots$

If the annuity D_{xy} is payable until the first of (x, y) dies one gets

$$\ddot{a}_{xy} = \frac{N_{xy}}{D_{xy}}. \tag{4.3.57}$$

Problem 4.26 Derive an expression for a deferred annuity.

Annuity on the last life The annuity 1 is paid due as long as one person of (x, y) is still alive

$$\begin{aligned}\ddot{a}_{\overline{xy}} &= \sum_{v=0}^{m} {}_v p_{\overline{xy}} v^v = \sum_{v=0}^{m} ({}_v p_x +_v p_y -_v p_{xy}) v^v \\ &= \ddot{a}_x + \ddot{a}_y - \ddot{a}_{xy} \\ \text{where } m &= \max(\omega - x, w - y).\end{aligned} \tag{4.3.58}$$

Annuity with reduction The annuity 1 is paid due as long as both live. After one of (x, y) has died a reduced annuity with face value r is paid due to the survivor. The present value of this insurance is

$$\ddot{a}_{\overline{xy}} \cdot r + (1-r)\ddot{a}_{xy} = r(\ddot{a}_x + \ddot{a}_y - \ddot{a}_{xy}) + (r - r)\ddot{a}_{xy} \tag{4.3.59}$$

Problem 4.27 Prove (4.3.59).

Annuity for the survivor After the death of x the surviving person y gets an annuity 1 due. The present value for this is

$$\ddot{a}_{x|y} = \ddot{a}_y - \ddot{a}_{xy}. \tag{4.3.60}$$

Term insurance for n years The company pays the face value 1 when one of (x, y) dies, provided this occurs within n years. The present value for this insurance is

$$\begin{aligned}{}_{|n}A_{xy} &= \sum_{v=0}^{n-1} {}_v q_{xy} v^{v+1} = \sum_{v=0}^{n-1} ({}_v p_{xy} -_{v+1} p_{xy}) v^{v+1} \\ &= v\ddot{a}_{xy,n\rceil} - a_{xy,n\rceil}.\end{aligned} \tag{4.3.61}$$

Problem 4.28 Show

$${}_{|n}A_{xy} = D_{xy}^{-1}(C_{xy} + \ldots + C_{x+n-1,y+n-1}) = D_{xy}^{-1}(M_{xy} - M_{x+n,y+n}) \tag{4.3.62}$$

where

$$\begin{aligned}C_{xy} &= v^{1/2(x+y)+1}(\ell_{xy} - \ell_{x+1,y+1}) \text{ and} \\ M_{xy} &= C_{xy} + C_{x+1,y+1} + \ldots\end{aligned}$$

Problem 4.29

Whole insurance for n years The insurance company pays the face value 1 after n years if both (x, y) are still alive. If one dies in this time the value 1 is paid in that year. Show that the present value of this insurance is given by

$$\begin{aligned}A_{xy,n\rceil} &= v\ddot{a}_{xy,n\rceil} - a_{xy,n\rceil} + \frac{D_{x+n,y+n}}{D_{xy}} \\ &= v\ddot{a}_{xy,n\rceil} + 1 - \ddot{a}_{xy,n\rceil}\end{aligned} \tag{4.3.63}$$

As a final example we consider an insurance which is paid upon the death of the second person. The present value of this is denoted by $A_{\overline{xy}}$ and

$$A_{\overline{xy}} = A_x + A_x - A_{xy} = 1 - d\ddot{a}_{\overline{xy}}. \tag{4.3.64}$$

Along these lines it is in principle possible to compute all sorts of insurances for pairs (x, y). The main problem that arises in this context is, however, the very large number of commutation numbers. Instead of about 100 values for D_x, for example, one now has to consider 10,000 values D_{xy} for all combinations of x and y. In order to overcome this difficulty one writes

$$D_{xy} \approx D_{\overline{x}\overline{x}} \tag{4.3.65}$$

where the central age $\overline{x}$ somehow has to be computed. That this approximation is not unreasonable follows from the fact that usually the x and y's are close together and that the mortalities change slowly with age. The simplest method would be $\overline{x} = \frac{1}{2}(x+y)$. A somewhat more complicated method could be based on a mortality law in the following fashion. Assume the ℓ_x are determined by a law $v^{x/2}\ell_x = f(x)$. Then we have to solve the equation

$$D_{xy} = f(x)f(y) = f(\overline{x})f(\overline{x}). \tag{4.3.66}$$

If the mortality law is not too complicated, this equation may be solved in closed form. In general it seems more convenient to expand (4.3.66) around the arithmetic mean $z = \frac{1}{2}(x+y)$ to second order $x = z + \zeta$, $y = z - \zeta$. Then $\overline{x} = z + y$ leads to

$$\begin{gathered}\left[f(z) + f'(z)\xi + \frac{1}{2}f''(z)\xi^2\right]\left[f(z) - f'(z)\xi + \frac{1}{2}f''(z)\xi^2\right] \\ \left[f(z) + f'(z)\eta + \frac{1}{2}f''(z)\eta^2\right]^2 \approx f(z)^2 + 2f(z)f'(z)\eta\end{gathered}$$

because η will be small in general. For the commutation numbers one will have to proceed similarly.

Such exercises in numerical approximation should not be overvalued, however, because the use of the same mortality table for x and y is a much more serious approximation. Also the assumed independence is a serious simplification.

Problem 4.30 In most case the insurance for two lives is for a man x and a women y. How would the formulas above (4.3.49) to (4.3.55) change if we use different life tables for men and women?

Problem 4.31 With married couples the women is usually between 2–4 years younger than her husband. How could one use this information to derive a more realistic approach to insurances for couples? How would you derive a "central" age in this case and a list of commutation numbers?

Summary: With the assumed independence of the two lives the life insurance for two lives can be developed. The usual age difference and the lack of independence, however, should lead to a slight reformulation of this theory.

Project: Develop a method to deal with dependence and age difference for insurances for two lives.

4.4 Model portfolios

In this section we will define two model portfolios for a life insurance contract

I: The small portfolio

	Face Value in 10^4				
q	1	2	3	4	5
0.002	1	3	0	2	0
0.003	0	2	3	0	3
0.004	3	1	2	10	0
0.005	7	2	0	1	0
0.006	10	12	20	4	5

(a) Compute the expectation value and variance.

(b) What is the probability that the loss will exceed 120,000?

(c) What is the probability that the loss will exceed the expected value by more than 10%?

(d) With a reserve of 100,000 determine the fluctuation term, based on the expected value principle, so that the chance of bankrupcy is less than .1%.

In the following somewhat more realistic example we base our computations on the .DAV Life table from 1994 T for men

II. The big portfolio

Age	30–40	41–50	51–60
Number	10,000 each year	8,000 each year	6,000 each year

Insurance Sum.	20,000	40,000	100,000
Rel. distribution in %	30	40	30

(a) Compute the expected losses and variances in the case of a 1-year term insurance.

(b) Complete the same as above for a 5-(10-)year term insurance.

(c) Compute the reserves in this case.

(d) The above DAV mortality table overestimates mortality by 10 to 20 per cent. Thus assume an overestimation of the mortality by 10 (20) per cent. Redo the above calculations.

(e) Assume 10% of the mortality is accident based. Redo the above calculations with a double indemnity is case of an accident.

(f) Compute level premiums for the corresponding whole insurance.

5. Reserves

5.1 The Role of Reserves

We have seen in the first chapter that reserves help insurance companies to cope with fluctuations. But with life and health insurances reserves have an additional function. Let us consider an example of this.

Example: Clever jr., aged 40, at the peak of his career, has bought a term insurance of face value 200,000 until his expected retirement at the age of 65, so that his wife can pay off the mortgage on the house and get the children through school.

Based on the life table ADST 60/62 mod his net premium

in the first year would be	$3.44 \cdot 10^{-3} \cdot 200,000$	$=$	688
after 10 (15, 20, 30) years would be	$8.26 \cdot 10^{-3} \cdot 200,000$	$=$	1,652

(2,898, 4,854, 6,866). But Clever jr. wants to pay a level premium which is 2,076,40 for $p = 3.5\%$. Leaving aside the administrative costs and fluctuation term, he overpays the company in the first year by almost 1,400. Even after 10 years he is still overpaying. The amount he overpaid, however, will be needed in the second half of his contract, because from year 55 onwards he will be paying less than the actual risk.

In order to cover the elevated risk in his later insurance period the company will thus have to save whatever he overpaid initially. These initial overpayments are called *reserves*. Such reserves are thus savings for future claims. They cannot therefore be treated as profit and, in fact, present a commitment for the future. Reserves are thus a liability. Reserves are able to stabilize a portfolio of life insurance. The reserves are determined to a large extent by the assumed interest rate, because the interest rate determines to what extent the future is discounted. A large p therefore means fewer reserves, while a smaller p increases the reserves. The interest rate is therefore much more important for the reserves than for the premiums. For these reasons governments have restricted the admissible p somewhat. In Germany it used to be about 3 or 3.5%. In Spain p was limited to 6%, while in France $p = 4.5\%$ was used in most cases. Since 1994 the requirements are not as stringent, although an interest rate between 4 and 5% is acceptable. In Germany the interest rate for reserves will be 3.5% from 2000 on.

We have just noted that reserves cannot be considered as profit. However, insurance companies will invest these reserves in stocks, property, bonds or funds. In these domains interest rates of 7–8% could be realized, thus contributing to the profit or surplus. It is obvious that reserves do not arise if the net premiums are given exactly by the expected claims in that year. In this case one speaks of natural premiums. Such premiums, however, would have to change from year to year, in fact they would rise dramatically as the example shows.

The explanation of reserves given above is still too general. To be more specific we define the *reserves prospectively* as the difference at any particular point in time between the present value of future benefits and the present value of future net premiums.

Retrospectively, we could also say that the reserves are the difference between the net premiums received and claims paid.

Both definitions allow us to describe reserves mathematically and, as we will see, both descriptions are equivalent because of the equivalence principle.

Consider a general insurance $(x, (D_i), (S_i), (Pr_i))$. For this the present value of the claims is

$$\mathcal{C}(0) = \sum_{s=0}^{\omega-x} ({}_{s+1}p_x S_{s+1} +_s p_x q_{s+x} D_{s+1}) v^{s+1} \tag{5.1.1}$$

while the present value of the premiums amounts to

$$Pr(0) = \sum_{s=0}^{\omega-x} {}_s p_x Pr_{s+1} v^s. \tag{5.1.2}$$

By the equivalence principle we have

$$Pr(0) = \mathcal{C}(0). \tag{5.1.3}$$

This equality, however, only holds at the beginning of the insurance contract, as the example above shows. After m years the prospective claims are

$$\mathcal{C}(m) = \sum_{s=0}^{\omega-x-m} ({}_{s+1}p_{x+m} S_{s+m+1} +_s p_{x+m} q_{x+m+s} D_{s+m+1} v^{s+1} \tag{5.1.4}$$

and

$$Pr(m) = \sum_{s=0}^{\omega-x-m} {}_s p_{x+m} Pr_{s+m+1} v^s. \tag{5.1.5}$$

One can now show by induction on m that $\mathcal{C}(m) = Pr(m)$ only if

$$Pr_{s+1} = p_{x+s} S_{s+1} v + q_{x+s} D_{s+1}.$$

This, however, would imply time varying premiums. Thus, in general, we have $\mathcal{C}(m) \neq Pr(m)$. If $\mathcal{C}(m) < Pr(m)$ the insured has to expect fewer benefits than s/he is paying for. In this case it is advantageous for him/her to cancel the contract. In general, this is possible. Thus we will have to demand

$$\mathcal{C}(m) \geq Pr(m). \tag{5.1.6}$$

The expression

$$V_m^{pro} = \mathcal{C}(m) - Pr(m) \tag{5.1.7}$$

represents *prospective reserves*. These reserves V_m^{pro} are thus the part of future claims not covered by future premiums.
We have

$$\begin{aligned}\mathcal{C}(0) &= \sum_{s=0}^{\omega-x}({}_{s+1}p_x S_{s+1} +_s p_x q_{s+x} D_{s+1})v^{s+1}\\ &= \sum_{s=0}^{m-1} + \sum_{s=m}^{\omega-x} = v^m\ {}_m p_x \left[\sum_{s=0}^{m-1}({}_{s+1}p_x S_{s+1} +_s p_x q_{s+x} D_{s+1})\frac{r^{m-s-1}}{{}_m p_x} + \mathcal{C}(m)\right]\end{aligned}$$

with $r = v^{-1}$ and similarly

$$Pr(0) = \sum_{s=0}^{m-1}\ {}_s p_x Pr_{s+1} v^s + v^m\ {}_m p_x Pr(m).$$

Since $\mathcal{C}(0) = Pr(0)$ we get

$$v^m\ {}_m p_x \left[\sum_{s=0}^{m-1}\ {}_s p_x Pr_{s+1}\frac{r^{m-s}}{{}_m p_x} + Pr(m)\right] = v^m\ {}_m p_x \left[\sum_{s=0}^{m-1}(\ )\frac{r^{m-s-1}}{{}_s p_x} + \mathcal{C}(m)\right]$$

or

$$\sum_{s=0}^{m-1}\ {}_s p_x Pr_{s+1}\frac{r^{m-s}}{{}_m p_x} - \sum_{s=0}^{m-1}({}_{s+1}p_x S_{s+1} +_s p_x q_{s+x} D_{s+1})\frac{r^{m-s-1}}{{}_m p_x} = V_m^{pro} = V_m^{re}. \quad (5.1.8)$$

The term on the left-hand side is called the retrospective reserve V_m^{re}, because the contingent on the life of the insured is the present value at time m of the overpayment of the insured. The derivation of (5.1.8) only expresses the fact that these overpayments will be used in the future. But the future net demand is just V_m^{pro}. Thus

$$V_m^{re} = V_m^{pr} \quad (5.1.9)$$

is just the equivalence principle from the beginning of the period of contract, viewed at time m. For this reason we shall write V_m for the reserves in the remainder. It should be noted that $V_m^{pro} = V_m^{re}$ is only correct if the same p is used in the periods before and after m.

The above formulas are, in general, too complicated if one wants to determine the development of V_m as a function of time m. In this case a recursion relation is more adequate.

Let us assume again that we are given a general life insurance $(x, (S_i), (D_i), (Pr_i))$. Then

$$(V_{m-1} + Pr_m)\left(1 + \frac{p}{100}\right) = p_{x+m-1}V_m + p_{x+m-1}S_m + q_{x+m-1}D_m. \quad (5.1.10)$$

In order to interpret this formula, assume that the insurance company keeps a separate account for each policy. The premiums are then considered as deposits, while the claims correspond to withdrawals. The interest rate applied to this account is of course p. With this interpretation V_m represents the balance for this account in the year m. Then (5.1.10) expresses the fact that at the beginning of year m the balance is V_{m-1} – from last year – plus the deposit Pr_m. On this amount interest is paid so that at the end of year m the credit is $(1+p/100)(V_{m-1}+Pr_m)$. If the policy holder survives the year m, this credit is used to form the reserves V_m and to S_m. If the insured dies, the amount D_m is due. (5.1.10) can be rewritten as

$$V_m = \frac{[(V_{m-1}+Pr_m)r]}{p_{x+m-1}} - S_m - \frac{q_{x+m-1}}{p_{x+m-1}} D_m, \tag{5.1.11}$$

which is the derived recursion relation. Together with

$$V_0 = 0 \tag{5.1.12}$$

it allows the computation of all V_m. (5.1.11) can be rewritten as

$$V_m = \frac{p_{x+m-1}+q_{x+m-1}}{p_{x+m-1}}(V_{m-1}+Pr_m)r - S_m - \frac{q_{x+m-1}}{p_{x+m-1}} D_m$$

so that

$$V_m - V_{m-1}r + S_m + \frac{q_{x+m-1}}{p_{x+m-1}}(D_m - V_{m-1}r) = Pr_m r\left(1+\frac{q_{x+m-1}}{p_{x+m-1}}\right)$$

with $Pr_m = vV_m - V_{m-1} + vS_m + q_{x+m-1}v(D_m - V_m - S_m)$, which we get directly from (5.1.11).

This equation has the following interpretation. Since the amount $V_m + S_m$ is needed at the end of the m-th year, the risked amount for the company is just

$$R_m = D_m - V_m - S_m, \tag{5.1.13}$$

which has to be paid if the insured dies. This happens with probability q_{x+m-1}. Thus

$$Pr_m^r = q_{x+m-1}v(D_m - V_m - S_m) \tag{5.1.14}$$

is its expected value. The difference

$$Pr_m - Pr_m^r = vV_m - V_{m-1} + vS_m \tag{5.1.15}$$

is saved.

Let us consider a number of examples.

Example 5.1 *Pure endowment insurance with level premiums* $D_i = 0$, $S_i = \delta_{i,n}$.

The present value of the insurance is $A^1_{x,n\rceil} = {}_n p_x v^n = D_{x+n}/D_x$. Thus the premium is $P = Pr_i = A^1_{x,n\rceil}/\ddot{a}_{x,n\rceil} = D_{x+n}/(N_x - N_{x+n})$ $i = 0,\dots,n-1$ and

$$V_m = {}_{n-m} p_{x+m} v^{n-m} - P \sum_{s=0}^{n-m-1} {}_s p_{x+m} v^s = A'_{x+m,n-m\rceil} - p\ddot{a}_{x+m,n-m\rceil} \tag{5.1.16}$$
$$= \frac{D_{n+n}}{D_{x+m}} - \frac{D_{x+n}}{N_x - N_{x+n}} \frac{N_{x+m} - N_{x+n}}{D_{x+m}}.$$

Example 5.2 *Term insurance for n years* In this case $S_i = 0$ and $D_i = 1$ for $I \leq i \leq n$. Its present value is

$$A^1_{x,n\rceil} = \frac{(M_x - M_{x+n})}{D_x}.$$

Thus the constant premium is given by

$$P = \frac{A^1_{x,n\rceil}}{\ddot{a}_{x,n\rceil}} = \frac{M_x - M_{x+n}}{N_x - N_{x+n}}$$

and

$$\begin{aligned}
V_m &= \sum_{s=0}^{n-m-1} {}_sp_{x+m}\, q_{x+m+1} v^{s+1} - P \sum_{s=0}^{n-m-1} {}_sp_{x+m} v^v \\
&= A^1_{x+m,n-m} - P\ddot{a}_{x+m,n-m\rceil} \\
&= \frac{M_{x+m} - M_{x+n}}{D_{x+m}} - \frac{M_x - M_{x+n}}{N_x - N_{x+n}} \frac{N_{x+m} - N_{x+n}}{D_{x+m}}.
\end{aligned}$$

As expected we have $V_0 = V_n = 0$.

Example 5.3 *Endowment insurance for n years.*

$$\begin{aligned}
D_i &= 1 \qquad i = 1, \dots, n, \qquad S_i = \delta_{i,n} \\
Pr_i = P &= \frac{(A^1_{x,n\rceil} + A^{\,1}_{x,n\rceil})}{\ddot{a}_{x,n\rceil}} = \frac{M_x - M_{x+n} + D_{x+n}}{N_x - N_{x+n}} \\
&= \frac{1}{\ddot{a}_{x,n\rceil}} - d \qquad d = (1 - v).
\end{aligned}$$

Then

$$\begin{aligned}
V_m &=_{n-m} p_{x+m} v^{n-m} + \sum_{s=0}^{n-m-1} {}_vp_{x+m}\, q_{x+m+v} v^{s+1} - P \sum_{s=0}^{n-m-1} {}_sp_{x+m} v^s \\
&= A_{x+m,n-m\rceil} - P\ddot{a}_{x+m,n-m\rceil} = 1 - d\ddot{a}_{x+m,n-m\rceil} - \frac{\ddot{a}_{x+m,n-m\rceil}}{\ddot{a}_{x,n\rceil}} + d\ddot{a}_{x+m,n-m} \\
&= 1 - \frac{\ddot{a}_{x+m,n-m\rceil}}{\ddot{a}_{x,n\rceil}}.
\end{aligned}$$

Finally, let us consider

Example 5.4 Pensions starting at $x + n$. In this case, $S_s = 1$ $n \leq s \leq \omega - x$, $D_s = 0$

$$P = \frac{{}_{n|}\ddot{a}_x}{\ddot{a}_{x,n\rceil}} = \frac{N_{x+n}}{N_x - N_{x+n}}.$$

Thus for $m \geq n$

$$V_m = \sum_{s=0}^{\omega-m-1} {}_vp_{x+m} v^{s+1} = \ddot{a}_{x+m} = \frac{N_{x+m}}{D_{x+m}}$$

and for $m < n$

$$V_m = {}_{n-m|}\,\ddot{a}_{x+m} - P\ddot{a}_{x+m,n-m} = \frac{N_{x+m}}{D_{x+m}} - \frac{N_{x+n}}{N_x - N_{x+n}} \frac{N_{x+m} - N_{x+n}}{N_{x+n}}.$$

In this case the risked amount is for $m \leq n$

$$Pr_m^r = q_{x+m-1} v(-V_m),$$

i.e. if the person dies before the beginning of his/her pension all the reserves remain with the company.

It is also possible to define and compute the reserves for continuous insurance models. We will consider this in the third section of this chapter.

Problem 5.1 Compute the reserves for a term fix insurance.

Problem 5.2 Compute the reserves for an m-year deferred term insurance.

Problem 5.3 What is the influence of the interest rate p on the reserves? Compute this for some simple examples, e.g. term insurance or whole insurance.

Summary: The reserves for an insurance contract are necessary for future liabilities. They can be computed prospectively and retrospectively. The recursion form is most useful in applications.

5.2 Reserves and Expenses

So far our considerations were restricted solely to net premiums. But insurance companies also have expenses at the beginning of a contract and during the contract period. In particular, the initial expenses, with commissions making up the lion's share, may easily exceed the first year's premium. Here are some examples:

Example 5.5 Whole life insurance $x = 40$ life table 60/62 male modified. Face value 10,000. At $p = 3\%$ (3.5%) level net premium 209.44 (195.94). So if the commission is 3.5% and other expenses 1.5% of the face value, the expenses of 500 exceed the premium by a factor of about 2.5.

Example 5.6 Endowment insurance $x = 40$ $n = 25$ life table as above, face value 10,000. At $p = 3\%$ (3.5%) the yearly level net premium is 311.19 (293.63).

In both cases, even a 10% fluctuation contribution would not help. In general the expenses during the first year by far exceed the expenses during the subsequent years. For these expenses the insurance companies have to be reimbursed, possibly when they occur. It is also clear that each contract should pay for the expenses it causes. Because of the high initial expenses the situation is thus rather complicated, because a company can hardly charge the initial costs and net premium as the gross premium. So the costs have to be spread out over the contract period. There are various means of allocating the expenses to each contract and spreading them out over the contract period. In particular, these methods influence the valuation of the reserves.

In the US the full preliminary term method of valuation has been used. It is based on the assumption that the first year's annual premium is used to cover the excess in expenses of this time. This part of the premium is then used as a 1-year term insurance whose terminal reserve covers the excess in cost. The remaining amount is then amortized over the rest of the period. The original contract then goes in effect with the second policy year with a one-year delay. The necessary premiums will then be based on this shortened period of payment.

Another method also used in the US is the Commissioners Reserve Valuation method. It is an extension of the preliminary term method. Its main objective is to avoid negative reserves. In Canada yet another method is used. In the remainder we will consider a method which was developed by the Swiss actuary Zillmer. This method, which is closely related to the FPT method, is widely used in Germany. In order to describe it we decompose the expenses for the insurance company into initial costs and running costs. The initial costs only arise in the first year. Thus

Initial Expenses Selling expense, commissions, risk classification, health examination.

Running Expenses *General expenses:* Advertising, marketing, actuarial and legal services, research investment management. These expenses should be allocated to each policy proportional to its face value.

Maintenance Collecting and accounting of premiums. Change of contracts and settlement. Correspondence with policy holder, claim investigation, legal expenses.

We can now write for the gross premium

$$\begin{aligned} \text{Gross premium} =& \text{Contrib. to initial Exp.} + \text{Net premium} \\ &+ \text{Flucutation term} + \text{Running Expenses} \end{aligned} \tag{5.2.17}$$

If the initial expenses are I then the contribution to the premium should be $I/\ddot{a}_{x,\overline{n}|}$ in the case of level premiums. The computation of the reserves will now be based on the Zillmer premiums Pr^Z, because we may assume that the running cost part is deducted immediately from the premiums to cover the corresponding expenses. Thus

$$Pr^Z_m = Pr_m + \frac{I}{\ddot{a}_{x,\overline{n}|}}. \tag{5.2.18}$$

Here Pr_m is the premium scheme for the insurance contract in question. Because the running costs are already accounted for, the reserves can now be computed as in the previous section, with Pr^Z instead of Pr. The reserves V^Z_m, however, should be initialized with

$$V_0 = -I. \tag{5.2.19}$$

Hence in principle the reserves can be computed as above. In particular, one still has the recursion relation

$$V_m^Z = \frac{(1+\frac{p}{100})}{p_{x+m-1}}(V_{m-1}^Z + Pr_m^Z) - S_m - \frac{q_{x+m-1}}{p_{x+m-1}}D_m. \tag{5.2.20}$$

Remark. It should be noted that in some cases the fluctuation term is built into the life table. Thus a separate consideration of the fluctuation term is not necessary.

Problem 5.4 Derive a formula for the Zillmerized reserves for a term insurance.

Problem 5.5 Repeat Problem 5.4 for an endowment insurance.

If one requires, as Zillmer did, that all V_m, $m \geq 1$ be non-negative, we have a means to determine the maximal allowable I. Since mortalities are increasing the condition

$$V_1^Z \geq 0 \tag{5.2.21}$$

will determine the maximal allowable initial expenses. If we have $V_1^Z = 0$, the premium in the first year will just cover the initial expenses and the risked amount.

Example 5.7 Endowment insurance: If $V_1^Z = 0$ we have $P^Z = I - q_x v$, since the premium in the first year just covers the initial expenses and the risked amount $q_x v$ (V1 5.1.15). In this case the Zillmerized premiums will be collected in the following years. But then there are no initial expenses so that

$$Pr_{x,n\rceil}^Z = Pr_{x+1,n-1\rceil}. \tag{5.2.22}$$

This allows one to compute the Zillmerized premiums. These depend clearly on the age of the insured, as (5.2.22) shows. In fact they grow rapidly with the age of the insured.

Problem 5.6 Determine the premiums and Zillmerized premiums for $x = 20$, 30, 40, 50 and $n = 10$, 20.

Cancellation of a contract

In general the insured is free to cancel an insurance contract. Since in general the insured has overpaid initially, he is entitled to get some of it back. The amount of overpayment for which no service of the insurance was rendered is just the reserves, in the case above the Zillmerized reserves. So in general the insured can expect about V_m^Z back if he cancels in the m-th year. Actually he will get a little less

$$\text{Return} = V_m - R_m \tag{5.2.23}$$

since cancellation of the contract changes the risk for the remaining policies. In real life this means that one should not cancel insurance contracts, because for the first 5 to 8 years hardly anything will be returned. The extent to which this happens depends largely on the initial expenses, i.e. commissions.

Summary: There are various ways to build the initial expenses, commissions into the calculations of premiums and reserves. The scheme by Zillmer is considered in more detail.

5.3 Reserves in the Continuous Description

We will now consider a general continuous description of an insurance policy for a person of age x. The functions D, S, $B : [0, n] \to [0, \infty)$ are called *term*, *endowment* and *premium rates* of the policy, if the following holds: $v^t = e^{-\delta t}$, $n = \omega - x$

1. If the insured dies at time t the amount $D(t)$ is paid to the beneficiary.
2. Until his/her death $S(t)$ is the endowment rate.
3. Premiums are paid with a rate $B(t)$.

The corresponding (continuous discounted) *loss function* L is then

$$L(t) = D(t)v^t + \int_0^t S(\tau)v^\tau d\tau - \int_0^t B(\tau)v^\tau d\tau. \tag{5.3.24}$$

L describes the discounted loss until t. L is clearly a random variable. The equivalence principle amounts to

$$E(L) = \int_0^{\omega-x} L(t)\mu(x+t)_t p_x dt = \int_0^{\omega-x} L(t)\mu(x+t)\exp\left(-\int_x^{x+t}\mu(s)ds\right) = 0. \tag{5.3.25}$$

The variance of the loss is then

$$\sigma_L^2 = E(L^2) - (E(L))^2. \tag{5.3.26}$$

Example 5.8 Whole life insurance $D(t) = 1$, $S(t) = 0$, $B(t) = P$

$$L(t) = v^t - P\frac{(1-v^t)}{\delta}, \qquad E(L) = 0 \text{ fixes } P$$
$$\sigma^2 = \sigma^2\left(v^t\left(1+\frac{P}{\delta}\right) - \frac{P}{\delta}\right) = \sigma^2\left(v^t\left(1+\frac{P}{\delta}\right)\right) = \left(1+\frac{P}{\delta}\right)\sigma^2(v^t)$$
$$\sigma^2(v^t) = {}^2\,\overline{A}_x - \overline{A}_x^2.$$

This latter expression is also the variance of the whole life insurance with an immediate single premium.

Problem 5.7 Determine the loss function L and variance for a term insurance.

Problem 5.8 Show that the loss function for a continuous endowment insurance satisfies $\sigma^2 = (1+P/\delta)^2({}^2\overline{A}_{x,n\rceil} - \overline{A}^2_{x,n})$.

The prospective reserves for a general insurance are now given by

$$V_t = \int_0^{n-t} D(t+s) v^s \mu(x+t+s)_s p_{x+t} ds \qquad (5.3.27)$$
$$+ \int_0^{n-t} \left(\int_0^s (S(\tau) - P(\tau)) d\tau \right) v^s \mu(x+t+s)_s p_{x+t} ds.$$

By partial integration this gives

$$V_t = \int_0^{n-t} [D(t+s)\mu(x+t+s) + S(t+s) - P(t+s)] v^s \, {}_s p_{t+x} ds. \qquad (5.3.28)$$

This in turn can be rewritten as

$$V_t = \int_t^n [D(s)\mu(x+s) + S(s) - P(s)]_{s-t} p_{x+t} \frac{v^s}{v^t} ds.$$

Since ${}_{s-t}p_{x+t} = {}_s p_x / {}_t p_x$ and $d/dt\, v^t \, {}_t p_x = -(\delta + \mu(x+t)) v^t \, {}_t p_x$ this shows

$$\frac{d}{dt}(V_t) = -D(t)\mu(x+t) + S(t) - P(t) + (\delta + \mu(x+t))V_t. \qquad (5.3.29)$$

This is the *Thiel differential* equation. Since it is linear the solution of (5.3.29) with the initial condition

$$V_0 = 0 \qquad (5.3.30)$$

is unique. It can be rewritten as

$$V_t' = P(t) + \delta V_t - S(t) - \mu(x+t)(D(t) - V_t). \qquad (5.3.31)$$

This means that the rate of change of the reserves is given by the rate of premiums plus the rate of interest paid on V_t minus the amount which has to be paid in the case of survival and the amount payable in the case of death. (5.3.31) obviously corresponds to the recursive definition of the reserves given in V2 (5.1.10). Above (5.3.24) we defined the random variable L, where

$$E(L(t)) = V_t.$$

Somewhat more generally let now

$$L_t(s) = D(s) + \int_t^s S(\tau) v^\tau d\tau - \int_t^s B(\tau) v^\tau d\tau \qquad (5.3.32)$$

be the random variable, which describes the prospective losses, because

$$E(L_t) = V_t. \qquad (5.3.33)$$

Example 5.9 Whole life insurance

$$L_t = v^{s-t} - P \frac{1 - v^{s-t}}{\delta},$$

s is the variable. This gives in example 5.9

$$\sigma^2(L_t) = \left(1 + \frac{P}{\delta}\right)^2 ({}^2\overline{A}_{x+t} - \overline{A}^2_{x+t}).$$

Example 5.10 Constant mortality μ and constant rate of interest δ

$$V_t = \int_0^\infty e^{-s\delta} \mu e^{-\mu s} ds - P \int_0^\infty s\mu e^{\delta s - \mu s} ds.$$

This expression does not depend on t, so $V_t = 0$.

Problem 5.9 Derive an expression for the loss function of an endowment insurance for n years. Give a formula for the variance.

Problem 5.10 Compute these expressions for $x = 40$, $n = 25$ for the Moivre law.

Example 5.11 Whole life insurance, $x = 40$ $\omega = 86$. Mortality according to Moivre $\mu(x) = \frac{1}{86-x}$ $\overline{A}_x = \int_0^{\omega - x} v^t\,{}_tp_x \mu(x+t) =$ payment by the insurance

$$\begin{aligned}\overline{A}_x &= \int_0^{\omega-x} e^{\int_x^{x+t} \delta + \frac{dy}{\omega - y}} \frac{dt}{\omega - x - t} = \int_0^{\omega - x} \exp\left(-\delta t - \ln \frac{\omega - x}{\omega - x - t}\right) \frac{dt}{\omega - x - t} \\ &= \frac{1}{\omega - x} \int_0^{\omega - x} e^{-\delta t} dt = (\omega - x)^{-1} \delta^{-1} (1 - e^{-\delta(\omega - x)}).\end{aligned}$$

Then the premium part is given by $P \cdot \overline{a}_x = P \cdot \int_0^{\omega - x} v^t\,{}_tp_x dt$ which one obtains by partial integration. Now we have

$${}_tp_x = \exp - \int_x^{x+t} \frac{dy}{\omega - y} = \frac{\omega - x - t}{\omega - x}.$$

Thus

$$\begin{aligned}\overline{a}_x &= \int_0^{\omega - x} {}_tp_x v^t dt = \int_0^{\omega - x} \left(1 - \frac{t}{\omega - x}\right) e^{-\delta t} dt = \frac{-1}{\delta} e^{-\delta t} \mid_0^{\omega - x} - \frac{1}{\omega - x} \int_0^{\omega - x} t e^{-\delta t} dt \\ &= \delta^{-1}(1 - e^{-\delta(\omega - x)}) - \frac{1}{(\omega - x)\delta^2} \int_0^{\delta(\omega - x)} u e^{-u} du = \delta^{-1}(1 - e^{-\delta(\omega - x)}) \\ &\quad + \frac{1}{\delta^2(\omega - x)} (u e^{-u} + e^{-u}) \mid_0^{\delta(\omega - x)} \\ &= \delta^{-1}(1 - e^{-\delta(\omega - x)}) + \delta^{-1} e^{-\delta(\omega - x)} + \delta^2 (\omega - x)^{-1} e^{-\delta(\omega - x)} - \delta^{-2} (\omega - x)^{-1} \\ &= \delta^{-1} - \delta^{-2} (\omega - x)^{-1} (1 - e^{-\delta(\omega - x)}).\end{aligned}$$

Since $V_0 = 0$ we get $\overline{A}_x = P\overline{a}_x = 0$ or

$$P = \frac{\delta^{-1} - \delta^{-2}(\omega - x)^{-1}(1 - e^{-\delta(\omega - x)})}{(\omega - x)^{-1}\delta^{-1}(1 - e^{-\delta(\omega - x)})} = \frac{\omega - x}{(1 - e^{-\delta(\omega - x)})} - \frac{1}{\delta}.$$

Then $V_t = A_{x+t} - P\overline{a}_{x+t}$.

6. Risk Theory

6.1 Risk

Before we begin with the analysis of risk, we have to consider how to define risk. In colloquial language risk is the chance for something bad happening. We described this by a random variable X, which describes the damage or loss. For insurance companies, however, risk is the chance of the occurrance of large claims, i.e. deviations from the expectation value. A convenient way to measure this is the standard deviation σ_X of X. This follows, for example, from the Tschebycheff inequality

$$\text{Prob}(|X - EX| \geq \varepsilon) \leq \varepsilon^{-p} E(|X - EX|^p) \quad p \geq 1, \tag{6.1.1}$$

which for $p = 2$ is

$$\text{Prob}(|X - EX| \geq \varepsilon) \leq \varepsilon^{-2} \sigma_X^2. \tag{6.1.2}$$

Another reason why the standard deviation is the sweetheart of all probabilists follows from the additivity of the variance for independent random variables

$$\sigma_Y^2 = \sum_{i=1}^{n} \sigma_{X_i}^2 \qquad \text{if } Y = \sum_{1}^{n} X_i \text{ and if the } X_i \text{ are independent.} \tag{6.1.3}$$

We used this law extensively in the first chapter. Insurance companies are not particularly interested in expressions like (6.1.1), but rather in $\text{Prob}(X \geq EX + K)$, i.e. the probability that the damage exceeds the net premium by the amount K, the reserves. If X is approximately normally distributed, we saw

$$\text{Prob}(X \geq EX + K) \approx 1 - \phi\left(\frac{K}{\sigma_X}\right), \tag{6.1.4}$$

where ϕ is the distribution function of the standard normal distribution. Thus it is important to know how good the normal approximation is for a large sum of independent random variables.

We will not pursue this problem here, but rather assume that the normal approximation holds. Generally, one would expect the normal approximation to hold if the portfolio is large and the individual claims are not too distinct from a normal distribution. This is also a considerable improvement on the Tschebycheff inequality, because for $\varepsilon = k \cdot \sigma_X$ we get

$$\text{Prob}(|X - EX| \geq k\sigma_X| = 2(1 - \phi(k)).$$

For $k = 2(3)$ this is 0.0456 (.0026), while the Tschebycheff inequality gives .25 and .111.

The risk for the insurance company cannot be computed simply from the claim. But one has to take into account all payments made. In order to see how to proceed let us consider a general life insurance of a person x with term, endowment and premium rates $(D,\ S,\ B)$. If this person dies in the year $x+\nu$ the present value of the loss for the insurance company is

$$L(\nu) = -\sum_{\mu=0}^{\nu-1} B_\mu v^\mu + \sum_{\mu=1}^{\nu-1} S_\mu v^{\mu+1} + D_\nu v^\nu. \tag{6.1.5}$$

This clearly defines a random variable. Since the probability of death in the year $x+\nu$ is ${}_{\nu-1}p_x q_{x+\nu-1}$ we get

$$E(L) = \sum_{\nu=1}^{\omega-x} L(\nu) q_{x+\nu-1\ \nu-1}p_x. \tag{6.1.6}$$

However, we have

$$E(L) = 0 \tag{6.1.7}$$

by the equivalence principle. Thus

$$\sigma^2(L) = \sum_{\nu=1}^{\omega-x} L(\nu)^2\ {}_{\nu-1}p_x q_{x+\nu-1}.$$

Insurances do not all begin in the same year. Thus we will have to compute the present value of the loss after t years. The loss of this policy discounted to the year $x+t$ is

$$_tL(v) = -\sum_{\mu=t}^{\nu-1} B_\mu v^{\mu-t} + \sum_{\mu=t}^{\nu-1} S_\mu v^{\mu-t} + D_v v^{\mu-t}. \tag{6.1.8}$$

Then, as in V3 we see

$$E({}_tL) = \sum_{\nu=t}^{\omega-x} {}_tL(\nu)\ {}_{\nu-1}p_x q_{x+\nu-1} = V_t, \tag{6.1.9}$$

the reserves in year t. Thus

$$\sigma^2({}_tL) = \sum_{\nu=t}^{\omega-x} L^2(\nu)\ {}_{\nu-1}p_x q_{x+\nu-1} - V_t^2 = E({}_tL^2) - V_t^2. \tag{6.1.10}$$

In many concrete cases with level premiums and term or endowment insurance the expressions can be simplified with the aid of commutation numbers.

Problem 6.1 Derive expressions for $\sigma^2({}_tL)$ in the case of

1. Term insurance with level (single) premium

2. Endowment insurance with level (single) premium.

Let us compute the risk for a term insurance with face value 1 for a particular year. Assume the reserves are V_0 and V_1 at the beginning and end of that year, respectively. If q is the mortality, the loss in the case of death is then $L = -(V_0 + B) + v$, where B is the premium. In the case of survival the loss is $L = -(V_0 + B) + vV_1$. Then

$$\begin{aligned} E(L) &= -q(V_0 + B) + qv - (1-q)(V_0 + B) + (1-q)vV_1 \\ &= -(V_0 + B) + qv + (1-q)vV_1 = 0 \end{aligned}$$

by the recursion formula for reserves. Thus

$$\begin{aligned} \sigma^2(L) = E(L^2) &= q(V_0 + B - v)^2 + (1-q)(V_0 + B - vV_1)^2 \\ &= pqv^2(1 - V_1)^2 \end{aligned} \tag{6.1.11}$$

when one uses $V_0 + B = qv + pvV_1$.

We now assume that the insurance company has a portfolio of N insurance policies. The loss function of the i-th policy will be denoted by L_i. If t_i is its age, the total loss L is given by

$$L = \sum \, {}_{t_i}L_i. \tag{6.1.12}$$

Even though one cannot always guarantee the independence of the L_i, $i = 1, \ldots, N$ in realistic portfolios, this assumption will hold to a high degree. Thus it is a good approximation and mathematically an extremely useful one. Thus we have

$$\sigma^2(L) = \sum_{i=1}^{N} \sigma^2({}_{t_i}L_i). \tag{6.1.13}$$

Since the ${}_{t_i}L_i$ only make sense in the discrete description for integer t_i, one will have to determine $\sigma({}_{t_i}L_i)$ for non-integer t_i, by linear interpolation. Then (6.1.12) defines the average risk for the portfolio for a particular year.

The discussion so far was based on the net premiums. But the gross premium also contain the administrative term and the fluctuation term Fl_i. Since the administrative terms are already used up, the total fluctuation term

$$Fl = \sum_{i=1}^{N} Fl_i$$

can be used by the insurance company to cover the risk. In addition, the company has reserves K and the company cannot satisfy its obligations if the losses L exceed $Fl + K$ or if the fluctuations in L exceed the reserves and $Fl + K$. If we assume L to be normally distributed, the probability for this is

$$1 - \phi\left(\frac{(\sum V_{t_i} + Fl + K)}{\sigma(L)}\right).$$

Since $V_{t_i} \geq 0$, this term can be estimated by

$$1 - \phi\left(\frac{(Fl + K)}{\sigma(L)}\right).$$

The consideration of the risk of a single policy or the consideration of the risk for a whole portfolio leads naturally to the question about the likehood of the ruin and time of ruin of an insurance company. Here ruin is understood to be the fact that claims exceed the premiums and reserves. Normally ruin does not imply that a company is bankrupt. The study of such questions is called collective risk theory. This area clearly occupies a central domain in insurance mathematics, and it is important also for the determination of the fluctuation terms. Despite this, we will not pursue this area here, and defer it to the second part, because it leads to a wide field of rather theoretical models about portfolios. In addition, collective risk theory is of lesser importance for life insurances, because these usually have large reserves and assets and the largest part of insurance policy involves surplus sharing, which is an additional safety measure for companies. But there is another safety device for insurances: they can insure themselves against extreme losses. This is called reinsurance.

Summary: If one wants to compute the risk of a life insurance contract, one has to take the reserves into account. Mathematically one does this by defining the loss variable L.

6.2 Reinsurance

Reinsurance is an insurance for insurance companies. It is thus a means of spreading the risk among several companies. Reinsurance can be applied for single risks, usually low-probability high-cost risks like oil tankers, planes, oil rigs, ... , or whole portfolios. Below we describe a general scheme for a reinsurance. Let $X_1, X_2, \ldots X_n$ describe the risks of a portfolio. Let N be the number of claims in a particular year. We order the claims according to their size

$$X_{N,1} \geq X_{N,2} \geq \ldots \geq X_{N,N},$$

i.e. $X_{N,1}$ is the largest claim and $X_{N,N}$ the smallest. In order statistics a similar procedure is used. The reinsurer will now cover a certain amount of the damage incurred. The part of $X_{N,i}$ covered by the reinsurer will be denoted by $f_i(X_{N,i})$. We assume that the f_i are measurable functions which satisfy

$$F_i(0) = 0, \; f_i(x) \leq x. \tag{6.2.14}$$

This means that in general the reinsurer will only cover part of the claim. Then

$$R = \Sigma f_i(X_{Ni}) \tag{6.2.15}$$

is the part of the total claims covered by the reinsurer. The $\{f_i\}$ define the reinsurance contract. Now a number of mathematical problems arise.

1. How should one choose the $\{f_i\}$ from the point of view of the company or the reinsurer, respectively?

2. How should the premiums be divided?

The functions f_i usually have the form

$$f_i(x) = c_i h(x) \text{ with } 0 \leq h(x) \leq x. \tag{6.2.16}$$

Example 6.1 1. $c_i = \alpha \in [0,1] \qquad h(x) = x$
This is called the *quota share reinsurance*. In this case the reinsurer covers the fraction α of the risk.

2. $c_i = 1$ $h_i(x) = \max(0, x - P)$
is an *excess of loss* reinsurance. Now the reinsurer covers the loss exceeding P.

3. $c_i = c_2 = \ldots = c_p = 1, \qquad c_j = 0$ otherwise, $h_i(x) = x$.
The reinsurer covers the p largest claims.

4. A combination of 2. and 3. is the excess of loss reinsurance with parameter p. In this case $c_1 = \ldots = c_p = 1, \qquad c_j = 0$ otherwise and $h_i(x) = \max(0, x - P)$. Here the reinsurer covers the p largest risks, but only those parts which exceed P.

5. Finally, the ECOMOR reinsurance covers the n largest risks, but only those parts which exceed the n-th largest claim. Thus the coverage R by the reinsurer is given by

$$R = \begin{cases} \sum\limits_{i=1}^{n} (X_{N-i+1} - X_{N-n+1}) \text{ if } n < N \\ \sum\limits^{N} X_i \text{ if } N \leq n \ . \end{cases}$$

Finally, we have the

6. **Global reinsurance**, which applies one of 1. or 2. to the whole portfolio or a part of the portfolio. This is also called stop loss reinsurance. In this case it is also possible to employ several reinsurances. Thus the first reinsurer covers all claims exceeding P_1 up to P_2, while the second reinsurer covers all claims exceeding P_2 up to $P_3, \ldots$.

The role of reinsurance for the insurer is to make the risk controllable. The following result shows that this is done best for the insurer by a global excess of loss reinsurance. This is plausible in as much excessive fluctuations in the global claim are covered by the reinsurer. To be precise, let X be the (global) risk with expectation value μ and let P be the ceiling for the loss. Then

$$E(X) = \mu = \int_0^\infty x dP_X = \left(\int_0^P x dP_X + P \int_P^\infty dp_x \right) + \int_P^\infty (x - P) x dP_X = \mu_1 + \mu_2 \tag{6.2.17}$$

describes the splitting up of the net premium into the part for the insurer μ_1 and for the reinsurer. Now assume we have another reinsurance policy, which is described by the function h and which leads to the same net premium

$$\mu_2 = \int_0^\infty h(x)dP_X \text{ and } \mu_1 = \int_0^\infty (x - h(x))\, dP_X. \tag{6.2.18}$$

Now the risk for the insured is

$$\begin{aligned}
&\int_0^\infty (x - h(x) - \mu_1)^2\, dP_X = \int_0^\infty ((x - h(x) - P) + (P - \mu_1))^2\, dP_X \\
&= \int_0^\infty \big((x - h(x) - P)^2 + 2(x - h(x) - P)(P - \mu_1)\big) + (P - \mu_1)^2 dP_X \\
&= \int_0^\infty (x - h(x) - P)^2 dP_X - (P - \mu_1)^2 \geq \int_0^P (x - h(x) - P)^2 dP_X - (P - \mu_1)^2 \\
&\geq \int_0^P (x - P)^2 dP_X - (P - \mu_1)^2.
\end{aligned}$$

If one chooses here $h(x) = \max(0, x - P)$, one has equality everywhere. Thus we see:

Theorem 6.1 Given the splitting of the net premiums $\mu = \mu_1 + \mu_2$ then the excess of loss reinsurance has the smallest risk for the insured.

Problem 6.2 This result is also true for several reinsurances. Prove this by induction.

While the distribution of the net premiums between insurer and reinsurer is not problematic, the previous result makes it plausible that the reinsurer will want a larger share of the fluctuation term, at least in the case of excess of loss reinsurances.

In the next result we will characterize the quota reinsurance. To do this let X be a global risk and let $h(X)$ be the part covered by the reinsurer. Then $f(X) = X - h(X)$ is the part covered by the insurer. Assume $\sigma_1^2 = \sigma_{h(X)}^2 < \sigma^2$, more precisely, assume $\sigma_h^2 = (1-t)^2\sigma^2 \qquad 0 < t < 1$.

Theorem 6.2 With $\sigma_1^2 = (1-t)^2\sigma^2$ fixed the quota reinsurance with parameter t gives a minimal risk for the original insurer.

Proof. We have $(1-t)^2\sigma^2 = \sigma_1^2 = \sigma^2(h(X))$

$$= \int_0^\infty (x - f(x) - E(X - f(X)))^2\, dP_X = \int_0^\infty (x - tx - S - E(x - tx - S))^2$$

where $S = f(X) - tX$. Then

$$\begin{aligned}
\sigma_1^2 =& (1-t)^2 \int_0^\infty (X - EX)^2 dP_X - 2(1-t) \int_0^\infty (X - EX)(S - ES) dP_X \\
&+ \int_0^\infty (S - ES)^2 dP_X = (1-t)^2\sigma^2 - 2(1-t) \int_0^\infty (X - EX)(S - ES) dP_X + \sigma_S^2.
\end{aligned}$$

Thus

$$\int_0^\infty (X - EX)(S - ES)dP_X = \frac{1}{2(1-t)}\sigma_S^2.$$

Then the risk for the primary insurer is $f(X) = X - h(X)$ and

$$\begin{aligned}\sigma_{f(X)}^2 &= \int_0^\infty (f - Ef)^2 dP_X = \int_0^\infty (tx + S(x) - E\,(tX + S(X))^2\, dP_X \\ &= t^2\sigma_X^2 + 2t\int_0^\infty (x - E(X))(S(X) - E(S(X)))dP_X + \sigma_S^2 \\ &= t^2\sigma_X^2 + \frac{t}{1-t}\sigma_S^2 + \sigma_S^2 = t^2\sigma_X^2 + \frac{\sigma_S^2}{1-t}.\end{aligned}$$

Thus we always have $\sigma_f^2 =$ risk of the primary insurer $\geq t^2\sigma_X^2$ and the minimum is attained for $S = 0$.

Because of Theorem 1 the excess of loss reinsurance is widely used for a global portfolio as well as for individual policies in order to control the risk. For life insurances it will be applied mainly for policies with a relatively large face value, while in non-life insurances it will be used for high-value low-probability cases, i.e. high-risk cases. Since the most sensible way of splitting up the contributions of the fluctuation term is based on the variances, we will finish this section by considering the variances. For a risk X we will denote by X^R and X^S the parts covered by the reinsurer and the insurer, respectively. Thus

$$X = X^S + X^R, \quad S = \text{self}, \quad R = \text{reinsurance}.$$

If P is the excess ceiling for the excess of loss reinsurance, we have for the distribution functions

$$F_{X^R}(x) = F_X(P + x). \tag{6.2.19}$$

Now let X describe a risk whose higher moments $E(X^\alpha)$ exist. Then

Lemma 6.1 For $\alpha > 0$ one has

$$E(X^\alpha) = \int_0^\infty x^\alpha dP_X = \int_0^\infty x^\alpha dF_X(x) = \alpha\int_0^\infty (1 - F_{(x)})x^{\alpha-1}\,dx.$$

Proof. This is a special case of Lemma 4.1. Now

$$\begin{aligned}E(X^\alpha) &= \lim_{c\to\infty} -\int_0^\infty x^\alpha d(1 - F_X(x)) = \lim_{c\to\infty} -\int_0^\infty x^\alpha d(1 - F_X(x)) \\ &= \lim_{c\to\infty} x^\alpha(1 - F_{(x)})|_0^c + \alpha\int_0^\infty x^{\alpha-1}(1 - F)dx \\ &= \alpha\int_0^\infty x^{\alpha-1}(1 - F)dx.\end{aligned}$$

The Lemma shows

$$E(X^R) = \int_0^\infty (1 - F_{X^R})dx = \int_P^\infty (1 - F_x(x))dx \tag{6.2.20}$$

$$= \int_P^\infty x dP_X - P(1 - F_X(P)) = E(X) - E(X^s), \tag{6.2.21}$$

$$E(X^{R2}) = 2\left[\int_P^\infty x(1 - F_X)dx - P\int_P^\infty (1 - F_X)dx\right]$$

and

$$E(X^S) = 2\int_0^P x(1 - F_X)dx. \tag{6.2.22}$$

We will now apply this to a portfolio of risks, each of which is reinsured with an excess of loss reinsurance with parameter P. The following discussion is actually more geared to non-life insurance than to life insurance, but we follow it nonetheless, because it indicates the first steps in collective risk theory. We denote by N the number of risks which require action by the insurer. Note that in this setting N is a random variable. Similarly, let N^R be the number of risks in the portfolio which exceed P. Assume P_X is the distribution of each individual risk. Let

$$p = P_X(X \geq P)$$

be the probability that the risk exceeds P. Then

$$\text{Prob}(N^R = n) = \sum_{i=n}^\infty \text{Prob}(N = i)\binom{i}{n}p^i(1-p)^{i-n}. \tag{6.2.23}$$

Now

$$E(N^R) = \sum n\,\text{Prob}(N^R = n) \tag{6.2.24}$$

$$= \sum_{i=o}^\infty \text{Prob}(N = i)\sum_{n=0}^i n\binom{i}{n}p^n(1-p)^{i-n} \tag{6.2.25}$$

$$= \sum_{i=0}^\infty \text{Prob}(N = i)ip = pE(N).$$

Similarly, one shows

$$\sigma_{N^R}^2 = p(1-p)E(N) + p^2\sigma_N^2. \tag{6.2.26}$$

Problem 6.3 Prove (6.2.26).

7. Pensions

7.1 Government pension schemes

The basic means to provide income in retirement are

1. government pensions
2. occupational pension schemes
3. pensions from insurances
4. savings, bonds
5. employment
6. family.

Of these, government pensions are the most important, not only by sheer volume. Nearly all countries – all governments in western Europe – have a state pension scheme to alleviate poverty in retirement. Since this basic state pension is provided by the state it is usually granted on the condition of citizenship or at least residency. The basic pension can be provided in two forms:

1. a uniform flat rate basic pension for all
2. an income-related pension.

The advantage of (1.) is its simplicity. It is the case that states will have to provide support for retired people in the form of social welfare anyway. Thus this model comes up again and again in debates on the reform of the pension system. In Germany (1.) has been proposed repeatedly by politicians engaged in this area. Its basic disadvantage is its unfairness. It is apparent to everyone that those who laboured hard during their working life should have a better position in retirement. Thus all current pension schemes in Europe are more or less income-related. There are two basics forms: either the pension is proportional to lifetime earnings or to income over the last k-years (final averaging). Here k can be 1, 5 or usually 10. In reality, the actual models practiced in most countries are a combination of (1.) and (2.). Thus the pension Pusually is of the form

$$P = B + aL(k).$$

However, computations become even more complicated because there may be a further term related to additional voluntary contributions made during one's working life. Also the number of children may be relevant, e.g. in France.

In Germany, for example, a pension is computed by the so-called *"Rentenformel" (pensions formula).* This formula depends on the time the pensioner was insured, the sum of all contributions made and the state of the pension system. Thus pensions are defined yearly by the so-called *"Rentenanpassungsgesetz" (law tying pensions to the national average wage).* In Britain the pension consists of the basic state pension (BSP) and the state earnings-related pension scheme (SERPS). However, as noted above, this usually provides only a very low support. The basic pension, for example, amounts only to one third of the average income of a working person. In France the pension amounts to about 50 % of a worker's average income during his/her ten best paid years if he/she has worked for more than 37 years. But an additional 10 % is paid for 3 or more children.

In Spain a worker can expect a pension which corresponds to his/her average income over the last 8 years if he/she has worked for more than 35 years.

Usually retirement is compulsory and begins at the age of 65, or in some cases at 60 – for women, for example. Early retirement or an incompleted working life usually leads to some reduction in the pension. The recent high unemployment rate has led to several schemes of early retirement with minimal reduction in the pension. In general, the government pension scheme is obligatory for all. Those people whose income exceeds a certain limit, may opt out (contract out), however, and obtain private insurance.

The funds for current pensioners are generally provided by those actively working, i.e. a certain amount of their income is deducted from their pay cheque. It is rather difficult, however, to determine the exact amount of the compulsory contribution which goes into the pension scheme, because it also covers unemployment benefit, accident insurance, insurance for dependents, etc. In some cases, employees may increase their future benefits by additional voluntary contributions. Pensions usually increase with time. In many cases they are tied to the retail index. In Germany there have recently been discussions – fall 1999 – to tie the increase to inflation, which is lower. All this shows that government pensions are a rather tricky business. They are more determined by politics and law than by mathematics. Moreover, it is apparent that all these laws have to be overhauled for demographic reasons and because of the current high rate of unemployment. The present system cannot be continued because the share of pensioners is rising dramatically. The increase in pensioners, higher benefits, an increase of eligible beneficiaries, their increased life expectancy and the high rate of unemployment have led to a steady growth of public expenditure for pensions which has been faster than the rise in available economic resources. Thus the whole system needs to be scrutinized. Politicians, however, have little inclination to do this because they have worked out particularly favorable pension systems for themselves and the necessary cutbacks are not easy to sell. Several trends in the pension system are obvious already:

1. Additional private pension provisions will become increasingly important and
2. other forms of provision for old age, like bonds, stock, ... will have to be considered too.
3. The pension systems will become more flexible.

4. The basic government pension will be reduced.

Pensions are usually paid as an annuity, for good reason. In some rare cases they may be paid out as a lump sum. In the case of death after retirement a certain fraction – determined by law – of the pension is paid to the widowed spouse and dependent children who have no income of their own.

In Germany an additional complication and demand for the government pension system arose through the reunification: people from Eastern Germany did not pay into this system, which now benefits a great number of them. For the other European countries, but also in particular for Germany, people seeking asylum may cause similar, but less severe problems for these governments.

Since government pensions are in general barely adequate, the occupational pension schemes form an important supplement. The advantage of such a scheme is to tie employees to a firm. In addition, it gives the employer the chance to honour good service. In Germany this scheme is voluntary and not as widespread ($\sim$ 5–10 %) as in Great Britain, where about 55 % of all working people are eligible for an occupational pension. In France this scheme is compulsory, where as in Spain it is mainly restricted to larger firms. The size and range of these schemes depend on the firm and, to a large extent, on the tax law. Particular provisions have to be made if an employee leaves the firm. Usually an employee is only entitled to the pension benefits if he or she has been in service with this firm for a certain time. The benefits be paid as a deferred annuity, due at retirement. But there may also be other transfer schemes. In France this is regulated by AGIRC (Association pour la gestion des institutions de retraites de cadres) or ARCO (Ass. de regiones de retraites complementaires). In addition to pensions, the occupational schemes also cover pensions related to disability or pensions for dependents. Such occupational pension schemes may be bought from insurance companies in the form of group insurance or individual insurances.

Currently, the complete government and employment pension scheme is being reworked in all countries because the demographic factor causes severe demand. Thus there will be considerable changes in the near future. As with life insurances, one can expect a greater variety and more flexible terms of provision for old age. In particular, it will be necessary to treat all forms of provision for old age, like insurances, bonds, savings, etc. on the same footing. Greater mobility within Europe will also require a harmonization of different schemes. Until now we are far from this, since harmonization encompasses legal conditions for such schemes, tax laws and insurance laws. No trend for such a development is discernible as of now. The governments are rather torn between finding new sources for taxes and encouraging people to save for old age.

For the occupational pension scheme one will need

1. A legally independent institution whose sole purpose is the administration and control of occupational pensions. The extent of ties to the employer is open and has to be settled.

2. Flexible pension schemes.

3. Flexible premiums.

4. Financial management of premiums, pensions and reserves as for insurances with a guaranteed minimum of benefits.

5. Tax exemption for premiums and tax reduction for pensions.

6. Freedom for investment of premiums. This, however, creates the problem of the distribution of gains and risk.

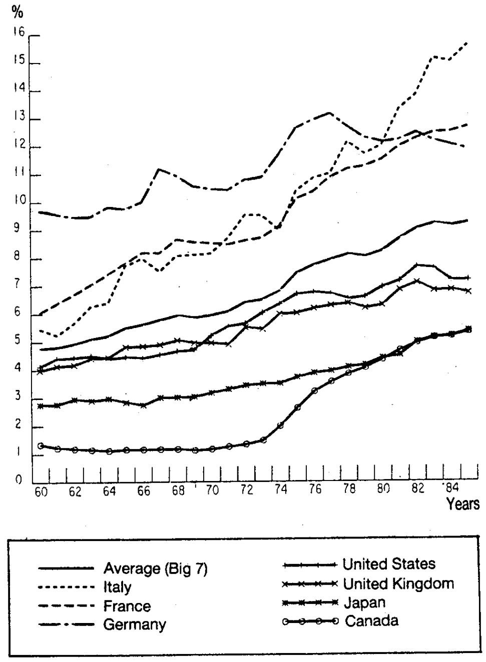
7. Global control of assets and investment.

All this should be on the same basis throughout Europe.

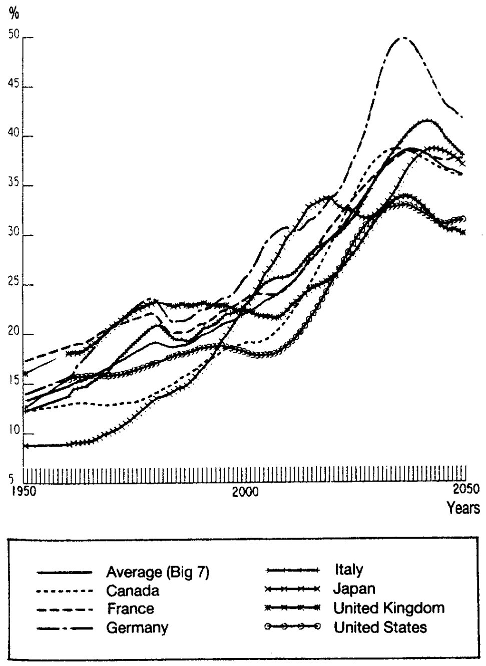
The pension system in Europe

Financing the system	Current premiums, government subsidies	Ge, GB, F, Es, I, NL, A, CU
Eligible persons	1. Whole population 2. Working people 3. Special groups	NL, GB (BSP), CH, E, Ge Ge, GB, I, F, Es, NL
Type of benefits	Old-age pension, disability pension, pension for dependents	GB, Ge, F, Es, I, NL, A, CH
Size of benefits	fixed depending on the size of earings and duration of employment	GB (BSP), CH, NL, Ge, GB (SERPS), I, F, Es
Age of retirement Men/Women	Ge (65/65), GB (65/60), F (60/60), Es (65/65)	
Premium as % of income	Ge ($\sim$ 20 UH), GB ($\sim$ 20 UH) F (16), E (28 H), U = unemployment benefit, H = health benefit	
Employer's contribution	Same amount as employees GB, Ge, F, Es about 4 times as much as employees.	
Percentage of employees in occ. pension scheme	I, Es $\sim$ 30 %, GB, Ge $\sim$ 50 %, F, NL, CH $\sim$ 100 %	

Summary. There are still major differences in the national pension systems within the EU. For demographic reasons and because of the high unemployment rate these systems have to be overhauled. Economic globalization favors harmonization of pension schemes and taxation. The pension schemes will become more flexible and individual planning will become more important. So far life insurances only play a supplementary role in the pension system. In the near future they will become more important.



Public pension expenditures as a percentage of GDP



Old-age dependency ratio for the seven major OECD countries 1950–2050

Literature

P. Booth, R. Chadburn, D. Cooper, S. Haberman, D. James: Modern Actuarial Theory and Practice. Chapman Hall 1999

K. Heubeck: Pensionsversicherung und Pensionsfonds. Zeitschift für die gesamte Versicherungswissenschaft 1998
OECD: Social Policy Studies # 5. Reforming Public Pensions.
N. Heinen, H. Storck: Kollektive Personenversicherung in Europa. Schriftenreihe Angew. Versicherungsmathematik Vol. 26

7.2 Funding Methods

In order to control the incident cost of pension schemes, funding methods are used. Most of these methods use some form of advance funding. This means contributions are made to meet future benefits before these benefits are due. The contributions are generally based on the equivalence principle, i.e. expected contributions equal expected benefits. This can be enacted as an individual principle or as a collective principle for a whole group. The methods used most commonly in the US or GB are defined in terms of a *standard fund*. Under normal conditions this fund suffices to meet all future obligations. The standard fund is replenished at a *standard contribution rate* (US normal cost). Because of varying economic and demographic factors, deviations from the standard fund (US accrued liability) are inevitable. Thus the standard contribution has to be supplemented by the distribution (amortization) of these gains or losses.

Pay-as-you-go

This form of funding means that today's benefits are paid for by those presently working. It is generally only used for the basic state pension. Its main disadvantage is its instability with regard to demographic and social changes, as we can witness at present. A large increase in retired persons and unemployment, together with a decrease in the number of working people, requires additional government funds via taxes. In Germany reunification has aggravated the problems. In France PAYG is used for the compulsory industry-wide pension system. In Italy it is also used to a large extent.

Most of the occupational pension schemes use the same form of advance funding, because it provides better security and enables the future cash flow to be planned better. Also the fact that every person pays for his/her own benefit is more natural and avoids intergenerational cross-subsidy.

For simplicity, we will only deal with net premiums and cost in the remainder. Since income and retail prices grow, all these evaluations have to be made on a dynamic basis. The extent of this may vary, however. We will then use the following notations:

$x(\overline{x})$	(average) age of entry
$n(\overline{n})$	(average) number of years until retirement
m	number of years in employment
$L_{x+m}(\Delta)$	size of benefits during period Δ taking into account inflation and growing salary
$a^a_{x+m,\overline{n-m}\rceil}$	present value of a unit annuity computed on the basis of the same dynamics as L
G	sum of all salaries and income
G_{x+m}	individual salary at age $x+m$
s	yearly trend in salaries
P_{x+m}	single yearly premium
N	normal contribution quotient as part of the salary.
S	standard fund (accrued liability)
V	value of the pension fund.

Projected Unit Method

The standard fund S is defined as the present value of all accrued benefits, taking into account future salary increases until retirement. The standard contribution rate is defined as the present value of the liabilities expected to be accrued in the following year. Thus

$$S = \sum L_{x+m}(m) \tag{7.2.1}$$
$$N = \frac{(\sum P_{x+m})}{G} = \frac{\sum L_{x+m}(1)}{G}.$$

The correction contribution will then be used to bring V to its standard value.

This method offers high security to its members. Its stability depends on the distribution of ages and salaries. If they remain stable, so will the fund. Thus it requires a steady stream of new entrants because it considers only the liabilities which are expected to arise in the following year. Thus it is particularly useful for large funds. It is now the standard method applied in the US and GB.

The Partly projected unit method uses only a restricted dynamic for a limited number of salaried years. This also applies to the *Current Unit Method* which uses the dynamics for only 1 year.

Entry Age Method (individual entry age) The standard fund S is defined as the present value of all liabilities of the present membership minus the present value of the standard contribution rate paid for all members throughout their future expected membership.

The standard contribution rate is the present value of all future liabilities expected to be accrued for an average new entrant. Thus we have

$$G_{\overline{x}} N = \frac{L_{\overline{x}}(n)}{a^a_{\overline{x},\overline{\overline{n}}\rceil}}. \tag{7.2.2}$$

The standard fund is then the prospective reserve

$$S = \sum \left(L_x(n) - NG_{x+m}\, a^a_{x+m,n-m\rceil} \right). \tag{7.2.3}$$

In the individual age entry method one has

$$NG = \left(\sum P_x \right) = \frac{\sum L_x(n)}{a^a_{x,n\rceil}}. \tag{7.2.4}$$

This method is rather similar to the aggregate method, which we will discuss now.

Aggregate Method This method has no standard fund nor a standard contribution rate. The latter is calculated as the present value of the total liabilities of the present membership minus the present value of assets. With contribution rate r we have

$$\sum rG_{x+m}\, a^a_{x+m,n-m\rceil} = \sum L_{x+m}(n) - V. \tag{7.2.5}$$

Attained Age Method In this method the future liabilities are spread evenly over the future. The standard contribution is thus the average contribution throughout the future membership. Thus

$$NG = \sum P_{x+m} = \frac{\sum L_{x+m}(n-m)}{a^a_{x+m,n-m\rceil}}. \tag{7.2.6}$$

S is defined as in the projected unit method and the correction contribution is defined as for that method.

Compared to the Entry Age Method, the contribution is defined with respect to the present age and not the entrant age. Thus as a rule the contribution tends to be higher, resulting in a higher surplus.

Defined Contribution In GB there are funds which define both the contribution and the final benefits.

The Use of Funding Methods in the EU

Country	Form of organization	Funding Method	Re-evaluation
B	Pension fund group insurance	proj. unit meth. aggregate method + indiv. premiums	annually
DK	Pension fund group insurance	individual premiums "	annually
F	Pension scheme with enforced membership	PAYG (several years)	
D	Pension reserves	Teilwertverfahren ≈ individual entry age method	annually
I	Pension funds	Partly PAYG	3–5 years
NL	Pension funds	Current unit, projected unit for funds without reinsurance Attained age otherwise	annually
E	Pension funds	Projected unit method, Entry age	1–3 years
GB	Pension funds	Mainly projected unit method	3 years

Literature

N. Heinen, H. Storck: Kollektive Personenversicherung in Europa. Schriftenreihe Angew. Versicherungsmathematik Vol. 26
P. Booth, R. Chadburn, D. Cooper, S. Haberman, D. James: Modern Actuarial Theory and Practice. Chapman Hall 1999

7.3 Mathematics for Pensions

In the previous sections we have mainly considered the social and economic boundary conditions for government or occupational pension schemes. In this section we will develop the mathematical aspects of insurance for pensions or pension funds. It therefore, of course, also applies to some extent to government pension schemes.

Altogether the mathematics is rather similar to that of life insurances, though slightly more complicated. For this reason we will only treat the new aspects in further detail. As in life insurance, these considerations always refer to a particular group of people which is homogeneous with respect to the properties studied, e.g. men; women or members of a particular occupation such as farmers, laborers, civil servants, etc. In life insurance a member of such a group can only be in one of two states: alive or dead. Now we have to consider several states. Here are some examples:

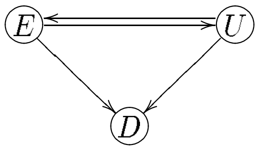
Example 1:
Employed
Unemployed
Dead

Example 2:
Employed
Unemployed
Disabled
Dead

Example 3:
Employed
Unemployed
Retired
Dead

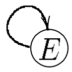
Example 4:
Employed
Unemployed
Retired
Dead
Cancellation of contract.

In setting up such models one has to consider the transition between these states. Possible transitions will be indicated by an arrow, so that these examples can be represented by a *directed graph.*
Example 1:



E = employed
U = unemployed
D = dead

The two top arrows indicate the transitions: An employed person losing his job → and an unemploymed person finding one. The other arrows indicate the death of an employed or unemployed person. In principle one should also consider the transitions employed → employed and unemployed → unemployed. Sometimes this is written as



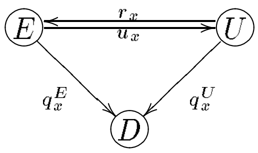
Problem 7.1 Write down the graph representations of the other examples.

Problem 7.2 Write down the graph representations of an example which takes into account several degrees of disability.

In the above model the state "dead" is called an *absorbing* state, i.e. no arrows lead away from it. The state "cancelled" will also be obsorbing if people are not able to return to their particular pension scheme.
In life insurance the important probabilistic quantities were q_x, the transition probabilities

$L \xrightarrow{q_x} D$ $\qquad L$ = Living

Here we have to take into account several such transition probabilities.
Example 1:



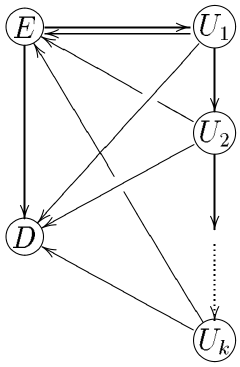
Thus we have
q_x^E = mortality of an x-year-old employee
r_x = the rate with which x-year-old unemployed people find a job
The other quantities have similar interpretations.

The term decrement means the decrease or termination of a given status. Here *decrement* will also be used as a transition probability. Thus a life table can be considered as a single decrement theory, while here we have to consider a *multiple* *decrement theory* or a theory of competing risks. The mathematical framework for such a theory is the theory of *Markov Processes*. Such processes are characterized by the fact that they have *no memory*, i.e. the transition rates depend only on the current state, but not on the previous history. There is a well-developed theory of Markov processes. However, here we will refrain from delving too much into it; we will only describe what we need. Markov processes can be described with continuous or with discrete time. The latter processes are called Markov chains. For us the time variable will then be age x of the insured person. Since the transition probabilities (q_x) depend on x, such processes are called inhomogeneous.

The situation is actually even more complex, since we have seen that the q_x also depend on time. In general, this is not compatible with a Markov model, which allows for only one time, namely the age. This can be remedied by considering the q_x to be only x-dependent for the time period under consideration. This, however, requires

1. a regular readjustment of the decrement tables – every 10 years or so
2. a proper adherence to the Markov idea would require *generation decrement tables*, i.e. decrement tables developed for a particular cohort. As we have seen with life tables, this is not feasible, thus all decrement tables are derived from *period data*.

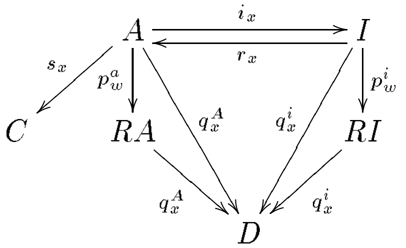
There is another complication on the road to a fully fledged Markov Theory. Quite often the transition probabilities depend on the time spent in a particular state. This is particularly appearent with "disability". Usually mortality for newly disabled people is quite high, because it is generally the consequence of an accident or a serious illness. But as the years go by, mortality approaches the mortality of actively working people. For working people mortality may also depend on periods of unemployment since quite often those people who suffer longer periods of unemployment have a questionable status of health. Such considerations lead to the theory of semi-Markov processes. We will not follow this line of thought because in practice these difficulties are usually ignored by using averaged transition rates. There is another device to take this into account. The next model graph solves this by taking into account several classes $U_1, \ldots, U_k$ of unemployed people. Thus U_k could mean having been unemployed for k-years.



For practical purposes, the main difficulty arises from the fact that the classes tend to get smaller and smaller as the states get more and more refined. This in turn increases the error in the decrements considerably, so that eventually one has to compromise and use averaged decrements. Thus such very refined models or models with continuous time are of little practical importance.

The model which we will study in more detail considers the following classes.

C People who have canceled their contract. If membership to the particular pension scheme is mandatory, this class may be empty. We will generally assume that C is an absorbing state, i.e. people may not return to their pension scheme once they have canceled their membership.
A actively working people.
RA retired people who took retirement whilst in employment. Often retirement is mandatory upon reaching a particular age.
I people of ill health, those who are unable to work because of a disability.
RI retired people of ill health.
D the dead.



The corresponding graph of this model is given by the following plan, with the age-dependent decrement rates. The interpretations are obvious. As pointed out above, the decrement table is determined as a period table. The transition probabilities i_x, s_x, ... are determined as for life insurance in the following steps:

1. determination of the raw data $\hat{i}$, $\hat{s}_x$, ...

2. smoothing

3. adjusting.

The raw data are usually determined by:

$$\hat{i}_x = \frac{\text{Number of people in the group of age } x \text{ who become disabled}}{\text{Number of people in the group of age } x} \tag{7.3.7}$$

with a similar definition for the other transition probabilities. The group can either be the number of persons within the portfolio of an insurance company, or, if available, several companies (see next section for an example). Smoothing the data is carried out by one of the methods described for life insurance. However, as yet there is no functional description for "i_x", "r_x", ...

In contrast to life insurance, an additional difficulty arises in the determination of the disability decrements "i_x", and reactivation probability "r_x" because the disability status is ill-defined. While the status "dead" is, apart from some dubious suicide cases, uniquely determined, disability requires a number of conditions for a valid insurance case. However, a number of subjective and objective factors determine the risk of disability. For one, it is not always clear what disability means: unable to work at all or unable to work in a particular profession. Here is a scale which describes disability:

1. Unable to work in one's profession.

2. Unable to work.
3. Unable to work at all.
4. Complete inability to work.
5. Needs support and help (various degrees).

Secondly, doctors may disagree in their judgement. Finally, the status, circumstances of life, social conditions and the expected pension determine the intensity to acquire the disability status. Political reasons also determine the risk of invalidity because, in times of high unemployment, the tendency to define someone as unable to work is higher. This can be expressed, e.g. via political decisions, particular laws or decisions in administration.

In order to determine these decrement rates we use the period approach. Thus let ℓ_x^a ($a =$ active, ill health, ...) be the number of people of age x and check to what extent they remain in that particular state. For the actively working state a, we have, for example

$$\ell_{x+1}^a = \ell_x(1 - q_x^a - i_x), \qquad x < w-1,\ w = \text{ retirement age.} \tag{7.3.8}$$

In general, one considers the decrement rates with respect to one attribute only. These are called net decrement rates. If we denote these by $\tilde{q}_x$, $\tilde{i}_x$, ... , we have under the assumption of independence

$$\ell_{x+1}^a = \ell_x(1 - \tilde{q}_x^a)(1 - \tilde{i}_x). \tag{7.3.9}$$

This differs from (7.3.8) by the cross term $\ell_x^a \tilde{q}_x \tilde{i}_x$, i.e. persons who die and become invalid within the same year. If one assumes that the transitions occur evenly distributed over the year, one finds that in 50 % of all cases i_x appears before q_x. Thus

$$\tilde{q}_x = \frac{q_x}{(1 - i_{x/2})} \approx q_x\left(1 + \frac{i_x}{2}\right).$$

Argue similarly for $\tilde{i}_x$. Thus

$$\tilde{q}_x = q_x\left(1 + \frac{i_x}{2}\right) \qquad \text{and } \tilde{i}_x = i_x\left(1 + \frac{q_x}{2}\right). \tag{7.3.10}$$

The probabilities $\tilde{q}_x$, $\tilde{i}_x$ are called *partial transition probabilities.*

In general one will have to ascertain what is being measured, partial transition probabilities or transition probabilities. If one measures q_x by

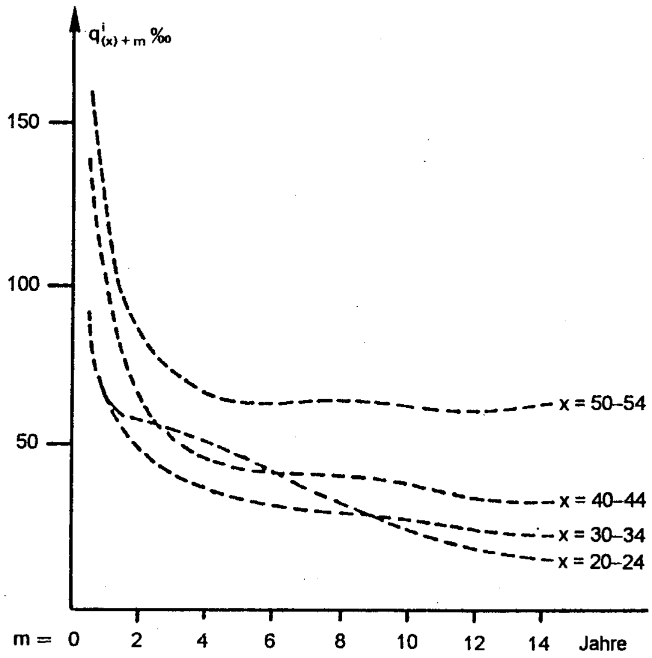
$$\hat{q}_x = \frac{(\text{number of persons of age } x \text{ who die before the } x+1 \text{ birthday})}{\ell_x}$$

one really measures $\tilde{q}_x$. Similarly, one will estimate $\hat{i}_x$.

These raw transition probabilities will then be smoothened and corrected and finally yield decrement tables which can be used to calculate premiums. So far, "formulae" for the q_x, i_x, ... have not been used, though it would be an interesting task to find suitable "transition laws".

Problems

1. Formulate the above model for h states – e.g. partial invalidity.
2. What type of differences do you expect for $(i_x)_{x=20}^{65}$ for miners and accountants? Consider in particular the values for $25 \leq x \leq 40$ and $45 \leq x \leq 60$.
3. Why would you expect the mortality of people of ill health to be higher initially but then tend to general mortality.
4. List a number of personal and economic reasons for a person to seek or avoid the status of ill health.
5. How would you modify (7.3.10) if you had several transitions i_x, $q_x, \ldots s_x$.



Mortality of disabled persons as a function of duration of disability m and begin x of disability.

Appendix 1: *Markov Chains*

We consider a time discrete Markov process $X(t)$, $t = 0, 1, \ldots$ with a finite state space $\mathcal{S} = \{1, \ldots, n\}$ i.e. a Markov chain. The Markov property then means

$$\begin{aligned} &P(X(t) = j \mid X(t_1) = j_1, X(t_2) = j_2, \ldots, X(t_k) = j_k) \\ &\quad = P(X(t) = j \mid X(t_k) = j_k) \quad \text{if } t_1 < t_2 < \ldots < t_k < t \end{aligned}$$

Here P stands for probability and $P(X \mid Y)$ for the conditional probability of X given Y. The Markov probability thus states that the transition to state j at t depends only on the state at the last observed time, but not on the times before. We can therefore define the transition probabilities

$$P_{ij}(t, s) = P(X(s) = j \mid X(t) = i).$$

These satisfy the Chapman Kolmogoroff equations

$$P_{ij}(t,s) = \sum_{k=1}^{n} P_{ik}(t,r)P_{k,j}(r,s) \qquad t \leq r \leq s.$$

This follows easily within the calculus of conditional probabilities because we sum over all intermediate states at time r. In matrix form, this last relation can be written for short as

$$P(t,s) = P(t,r)P(r,s).$$

Thus in particular

$$P(t,t+k) = P(t,t+1)P(t+1,t+2)\dots P(t+k-1,t+k).$$

This in turn shows that the $P(t) = P(t,t+1)$ determine the whole process

$$P(t,s) = P(t)\dots P(s-1) \qquad t < s.$$

Within the Markov framework a general insurance program is now easy to realize. In this case the claims of an x-year-old insured person are determined by $(c_j(t) \mid t = 1,2,\dots,w-x,\ j=1,\dots,n)$. This means that the insurance has to pay c_j at the end of year t if the insured person is in state j at t. The present value of this is

$$\sum_{t=0}^{w-x} P_{i,j}(x,x+t+1)c_j(t+1)v^{t+1}$$

if the person is in state i at x. For the premiums we have similarly

$$\sum_{t=0}^{w-x} P_{i,j}(x,x+t)Pr_j(t)v^t.$$

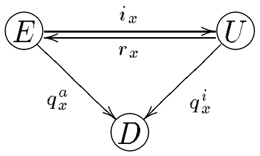
With these formulae it is now possible to develop the theory as for life insurance, by computing reserves, $\dots$. The projected reserves are then given by

$$\begin{aligned} {}_iV_m = & \sum_{t=0}^{w-x-m} P_{ij}(x+m,x+t+m+1)c_j(t+1)v^{t+1} \\ & - \sum_{t=0}^{w-x-m} P_{ij}(x+m,x+t+m)Pr_j(t)v^t \end{aligned}$$

if the insured person is in state i at age $x+m$. It should, however, be clear that even such an innocent-looking expression as

$$P_{i,j}(x,x+t+1)$$

involves summation over n^{t+1} terms. Thus simplifications via commutation numbers is only possible in exceptional cases. Even for the simple case $x < \omega =$ retirement age and $C = \phi$ the above system is still a 3 state system



with transition matrix $$P(x) = \begin{pmatrix} p_x^a & i_x & q_x^a \\ r_x & p_x^i & q_i^x \\ 0 & 0 & 1 \end{pmatrix}$$

In this case one can reduce the computation of $P(x, x+t)$ essentially to that of 2×2 matrices. However, the computation is still quite complicated. A more convenient method is the recursive definition of actuarial functions. Even if this Markov formulation of actuarial problems is quite straightforward, there remains the problem of determining all transition rates.

Appendix 2: Definitions of disability in various countries.

Germany: One distinguishes between inability to work in one's profession (Berufsunfähigkeit) and inability to work at all (Erwerbsunfähigkeit). The former is defined by a reduction of professional capabilities and abilities to less than 50%. In this case the disability pension is 2/3 of that for general disability.

France: Here one distinguishes between 3 levels of disability.

1. Invalides capables d'exercer une activité rémunerée.
2. Invalides absolument incapables d'exercer une profession quelconque.
3. Invalides qui, étant absolument incapables d'exercer une profession, sont, en outre, dans l'obligation d'avoir recours à l'assistance d'une tierce personne pour effectuer les actes ordinaires de la vie courante.

In the domain of the self-employed one also speaks of

Incapacité de travail permanente ou invalidité professionelle

and

Incapacité fonctionelle.

The latter describes general disability. Both forms are considered in various stages (10%, 20%, ... 100%) and both together determine the level of disability.

Great Britain: Here one also differentiates between two levels of disability

- Disability for the purpose of this policy means that the insured person is totally unable by reason of sickness or accident to follow his/her own occupation and is
    1. not following any other for profit or reward.
    2. any other suitable occupation.

USA: The following stricter definition is widely used

The insured person has become totally disabled by bodily injury or disease so that he/she is and will be thereby wholly prevented from performing any work, following any occupation or engaging in any business for remuneration or profit, and such disability has already continued uninterruptedly for a period of at least six months.

7.4 Premiums for Pensions

We will now compute the premiums for the above 3-state model, for which we neglect reactivation and cancellation. In reality, this of course is an overestimate of the premiums. However, a more complete analysis can be developed along the lines indicated above. As for life insurance we have

$$\ell^a_{x+1} = \ell^a_x(1 - q^a_x - i_x), \qquad \ell^a_{x_0} = 100,000 \quad x_0 \leq x \leq w. \tag{7.4.11}$$

Here x_0 is a suitable starting age, e.g. $x_0 = 20$, while w is the age of retirement, $w = 65$ in most cases. For the group of invalids we assume

$$\ell^i_{x_0} = 0 \qquad \ell^i_{x+1} = \ell^i_x(1 - q^i_x) + \ell^a_x i_x. \tag{7.4.12}$$

Problem 7.3 Formulate (7.4.11), (7.4.12) for several classes of disability.

Problem 7.4 Formulate (7.4.11) and (7.4.12) including reactivation.

Beyond retirement age $x \geq w$ the model simplifies to:

$$RI \xrightarrow{q^I_x} D \qquad \text{and} \qquad RA \xrightarrow{q^A_x} D$$

with

$$\ell^I_{x+1} = (1 - q^J_x)\ell^I_x, \qquad \ell^A_{x+1} = \ell^A_x(1 - q^A_x). \tag{7.4.13}$$

Annuity during employment

This describes an annuity of value 1 which is paid as long as the insured person is actively at work. Its present value is

$$a^a_x = \sum_{\nu=1}^{w-x} {}_\nu p^a_x v^\nu = \sum_{\nu-1}^{w-x} \frac{\ell^a_{x+\nu}}{\ell^a_x} v^\nu = \sum_{\nu-1}^{w-x} \frac{\ell^a_{x+\nu} v^{\nu+x}}{\ell^a_x v^x} \tag{7.4.14}$$

$$\ddot{a}^a_x = \sum_{\nu=0}^{w-x-1} {}_\nu p^a_x v^\nu.$$

If one defines

$$D^a_x = \ell^a_x v^x \quad \text{and}$$

$$N^a_x = \sum_{\nu=0}^{w-x-1} D^a_{x+\nu}$$

one gets

$$\ddot{a}^a_x = \frac{N^a_x}{D^a_x}. \tag{7.4.15}$$

Problem 7.5 What does $\ddot{a}^a_{x,n\rceil} = (N^a_x - N^a_{x+n})/D^a_x$ stand for?

The formulae (7.4.14) and (7.4.15) are also correct if reactivation is included. Then, however, the recursion relation for the ℓ^a_x (7.4.11) has to be modified (Problem 7.4).

Disability annuity

In case of disability insurance, reserves have to be put aside for disability annuities. For an annuity of face value 1 the present value is

$$\ddot{a}^i_x = \sum_{\nu=0}^{w-x-1} {}_\nu p^i_x v^\nu = \sum_{\nu=0}^{w-x-1} \frac{\ell^i_{x+\nu} v^{x+\nu}}{\ell^i_x v^x} = \sum_{\nu=0}^{w-x-1} \frac{D^i_{x+v}}{D^i_x} = \frac{N^i_x}{D^i_x}. \tag{7.4.16}$$

Then

$$\ddot{a}^i_{x,n\rceil} = \frac{(N^i_x - N^i_{x+n})}{D^i_x}$$

is the present value of a temporary annuity.

With this the disability insurance of an actively working person of age x can be computed as follows. The probability of this person becoming disabled in the n-th year is given by the probability to be working actively in year $n-1$ and then becoming disabled. Thus it is ${}_{n-1}p^a_x i_{x+n-1}$. Since invalidity payments have to be made immediately after becoming disabled, the corresponding present value is

$${}_{n-1}p^a_x i_{x+n-1} \ddot{a}_{x+n-1/2} ={}_{n-1} p^a_x i_{x+n-1} \frac{1}{2}(\ddot{a}_{x+n} + \ddot{a}_{x+n-1}).$$

Altogether, this gives

$$\ddot{a}^{ai}_x = \sum_{\nu=1}^{w-x} {}_{\nu-1}p^a_x i_{x+\nu-1} \ddot{a}_{x+\nu-1/2} v^{\nu-1/2}. \tag{7.4.17}$$

For the last two formulae we have assumed that the disability annuity begins about mid year.

Problem 7.6 With $D^{ai}_x = \sum \ell^a_x i_x v^{x+1/2} \ddot{a}^i_{x+1/2}$ and

$$N^{ai}_x = \sum_{\nu=0}^{w-x-1} D^{ai}_x \quad \text{show} \quad \ddot{a}^{ai}_x = \frac{N^{ai}_x}{D^{ai}_x}.$$

Problem 7.7 Formulate the problem of a disability insurance with inflation correction.

Problem 7.8 Quite often a pension insurance is deferred by a fixed amount of k years. Think of a disability insurance for employees. But this may also be the waiting period. Compute the present value for this case.

Problem 7.9 Disability insurance is generally only one part of the insurance package. In addition, one has to take into account pensions (annuities) in retirement. This can mean two things: retired while actively working and retired while being disabled. What changes do you have to make in (7.4.16) to cover the disability retirement annuity? What is the present value of actively retired persons?

As a final example, we will analyze widow's pensions. Similar considerations also hold for pensions for dependents.

Problem 7.10 Sometimes reactivation is considered only for the ill health group, i.e. one has $\ell^1_{x+1} = \ell^q_x(1 - i_x - q^a_x)$ and $\ell^i_{x+1} = \ell^i_x(1 - q^i_x - r_x)$. Derive the above formulae on this basis of these laws. What is the effect of this assumption?

Widows pension

This case really covers several contingencies which we write out in full

1. Person x married to a wife of age y dies in year $x + n$ as an actively working person or as an actively retired person.
2. Person x married to a wife of age y dies in year $x + n$ as a disabled person or as a retired disabled person.

Let us first compute the present value of an annuity for a widow of age y. As before, we define the evolution of the group of widows by

$$\ell^w_{y+t+1} = \ell^w_{y+t}(1 - q^w_{y+t} - h^w_{y+t}), \tag{7.4.18}$$

where q^w, h^w is the mortality and rate of remarriage of widows, respectively. Then

$${}_np^w_y = \frac{\ell^w_{y+n}}{\ell^w_y}$$

gives the probability of being a widow in n years hence. Hence

$$\ddot{a}^w_y = \sum_{\nu=0}^{\omega-y-1} {}_\nu p^w_y v^\nu = \sum_{\nu=0}^{\omega-y-1} \frac{\ell_{y+\nu} v^{\nu+y}}{\ell_y v^y}. \tag{7.4.19}$$

It should be noted that, in general, q^w_y is larger than q_y, so if the q_y from the general mortality table are used for the q^w_y, (7.4.19) overestimates the actual expenses.

Problem 7.11 In many cases insurance for disability demands a waiting period of k years. Formulate such a deferred disability insurance.

Problem 7.12 Sometimes a widow may be entitled to a part k of her pension if she remarries. Show that the present value for this is given by

$$k \sum_{\nu=0}^{\omega-y-1} {}_{\nu-1}p^w_y h^w_{y+\nu-1} v^{\nu+1/2} \ddot{a}^w_{y+\nu}$$

if remarriage is assumed to take place in mid year or is assumed to be spread evenly over the year.

Now it is easy to write down the present value of a widow's pension in case (a)

$$\ddot{a}^{aw}_{x/y} = \sum_{\nu=0}^{\omega-x-1} {}_xp^a_x q^a_{x+\nu} \ {}_{\nu+0.5}p_y \ddot{a}^w_{y+\nu+.5}. \tag{7.4.20}$$

In (7.4.20) we have assumed again that death is spread approximately evenly throughout the year. A similar formula will hold for disabled persons: just replace a by i. In order to cover also the case of an x-year-old actively working person married to a wife of age y, who becomes invalid before retirement, we get

$$\sum_{\nu=1}^{\omega-x} {}_{\nu-1}p_x i_{x+v-1} q^i_{x+\nu} \ {}_{\nu+1/2}P_y \ddot{a}_{y+\nu+1/2}. \tag{7.4.21}$$

The above present values have been determined as individual insurances. If one is considering a larger group of persons, one has take the average of the values. For this, one has to take into account the probability $m(x,y)$ of a person of age x being married to a woman of age y. Since this is too complicated in general, one writes

$$m(x,y) = \begin{cases} m(x) & \text{if } y = x - d(x) \\ 0 & \text{otherwise} \end{cases}$$

where $m(x)$ is the probability of an x-year old being married and where $d(x)$ is the average age difference between x and y. Since $d(x)$ grows from 2 at x about 25 to 6 at $x = 80$, because of remarriage, one can then easily compute the average present value of a widow's pension. One obtains

$$\ddot{a}^w_x = \sum_{\nu=0}^{\omega-x-1} {}_\nu p_x q_{x+\nu} \ {}_{\nu+1/2}p_{x+\nu-d(x)} \ddot{a}^w_{y(x)} v^{\nu+.5} \tag{7.4.22}$$

where $\ddot{a}^w_y$ is the present value of the expected widows pension of a $y(x) = x + d(x) +$ 1/2-year-old widow from that particular group.

Even though the mathematics for disability pensions is rather like mathematics in lifeinsurance, computations for premiums, reserves, ... are connected to considerably larger insecurities, because the decrement rates have larger variances and because external social and political conditions change. The long duration of contracts makes it particularly difficult for insurances with small portfolios or with a young portfolio.

It goes without saying that widower's pensions or pensions for dependents are computed using the same method. Even though these formulae look rather complicated, they are at most messy because of the many contingencies. From a practical point of view, they are easy to compute with a computer. The large number of contingencies, however, mean that many transition probabilities have to be determined. Thus the corresponding groups are small and the transition probabilities have relatively large errrors.

Once all transition rates are known, one can compute rates and reserves as for life insurances. The reserves are usually computed in their prospective form. For group insurances for larger firms the individual equivalence principle may be replaced by a collective equivalence principle. Thus one essentially takes the average of all policies. For basic state pensions, the PAYG method is usually used. The consequences of such a policy are visible to every one, when the number of pensioners is increasing, while the number of actively working people is decreasing.

Problem 7.13 Determine all actuarial functions – present value, level premiums, prospective reserves – for a long-term disability insurance for an x-year-old actively working person, when premiums are only paid during employment.

Problem 7.14 What type of changes arise if one considers two states of disability.

Problem 7.15 Sometimes increased disability insurance benefits are paid, if the disability was caused by an accident (at work). Formulate this.

Problem 7.16 Compute the reserves for retired active (disabled) persons.

Problem 7.17 Compute the reserves for a widow's pension if the widow gets $k \cdot 10^{-1}$ of the man's pension.

References

S. Haberman, E. Pitacco: Actuarial Models for Disability Insurance. Chapman Hall 1999. This book is rather theory-oriented and favors the continuous time approach. A rather similar approach is given by
H.G. Möller, H.J. Zwiesler: Mehrzustandsmodelle in der Berufsunfähigkeitsversicherung
K. Wolfsdorf: Versicherungsmathematik I. Teubner.

7.5 Permanent Health Insurance in GB

The model for disability insurance that we have considered above was oriented rather closely towards life insurance and considered mainly permanent disability. As we have seen, however, the reactivation rates can be considerable. Moreover, it is rather complicated to consider reactivation properly in a full model. Thus it is interesting to see other models for disability insurance. One such form is the "friendly society method", which has been practiced by "Friendly Societies", (mutual insurance groups), in Great Britain. In contrast to the above approach, they do not distinguish between disabled persons and those actively working. Instead, they consider the average duration during which an x-year-old member is disabled. These periods are usually measured in weeks. Thus they consider $Z_x^{u/w}$ = average number of weeks of disability expected for an x-year-old member between the u-th and $u+w$-th week. In general, only the combinations $(u \mid w) = 0/13, 13/13, 26/26, \ldots, 52/52$ and 104/everything are considered. Thus one writes

$$Z_x = Z_x^{0/13} + Z_x^{13/13} + Z_x^{26/26} + Z_x^{52/52} + Z^{104/\mathrm{all}}. \tag{7.5.23}$$

This also permits us to build in a waiting time. Thus, e.g.

$$Z_x^{13/\text{all}} = Z_x^{13/13} + Z_x^{26/26} + Z_x^{52/52} + Z^{104/\text{all}} \tag{7.5.24}$$

gives the total number of weeks of disability with a waiting period of 13 weeks.

As an example, consider the benefit of 1 per year, i.e. 1/52.18 per week for a member of insurance age x and actual age $x + k$

$$a_{x+k}(u, w) = 52.18^{-1} \frac{1}{D_{x+k}} \sum_{\nu=k}^{\omega-x-1} D_{x+\nu} Z_{x+k}^{u/w}$$

where (D_x) is the commutation number based on the life table for all (!) members, because no difference is made between disabled and active members. If membership premiums have to be paid also during ill health status, the net premium is

$$Pr_x(u, w) = \frac{a_{x_1}(u, w)}{a_{x,w-x\rceil}}.$$

If disabled persons do not have to pay remiums during the duration of invalidity, this expression has to be modified. If, for example, no premiums have to be paid in a period which begins m weeks after invalidity and lasts for n weeks, one gets

$$Pr_x(u, w; m, n) = \frac{Pr_x(u, w)}{(1 - Pr_x(m, n))}.$$

While this model is clearly simpler than the model discussed above, it has one flaw. This disadvantage arises if the ratio of active and invalid people in the portfolio differs from the assumed ratio which was used to derive the Z_x. Here obviously, the number of permanent invalids is particularly crucial.

In the US, another method, which is rather similar to the one above, is used. This approach is called the "continuance method" because it is based on the event and continuance of disability. Typical quantities are then terms like $p_{x,\pi}$, which give the probability that a person of age x becomes disabled within the next year and remains disabled for t periods of duration π. π is usually a week.

In both approaches, reactivation is thus one of the key features.

7.6 New decrement rates for occupational disability insurance in Germany

Since decrement rates also undergo large changes over the years, they have to be reevaluated regularly. The following notes refer to the computation of 1997 decrement tables of the German Actuary Association [N. Kolster, H. Loebus, W. Mörtelbauer: Neue Rechnungsgrundlagen für die Berufsunfähigkeitsversicherung DAV 1997]. The basis of these new decrement tables are data on occupational disability insurance in Germany for the years 1983 to 1993, notably 1987 to 1991, of the Munich reinsurance company? The volume of these data is about $8.2{\cdot}10^6$ person years for men and $2.0{\cdot}10^6$ person years for women. It was found that the i_x did not only depend on the type of occupation but also on the expected pension. The transition probabilities are determined in three steps, like the mortalities q_x for life insurance.

1. Determination of the raw data

2. Smoothing the raw data – Whittaker Henderson's method –

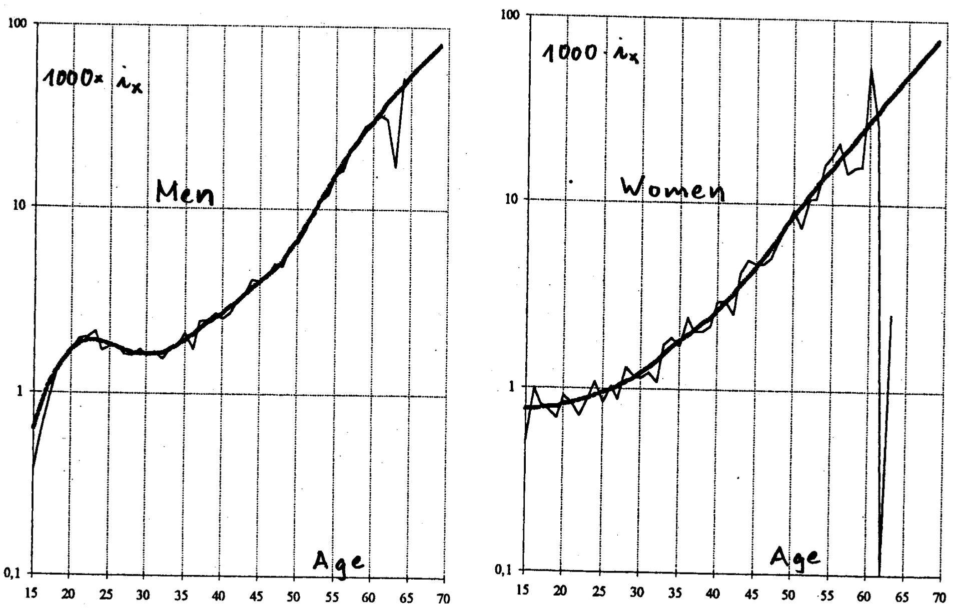
3. Extension of the data to the boundary by a modified Gompertz term. It was also found that i_x decreased for higher ages (men $x \geq 61$, women $x \geq 56$), contrary to expectations. The reason for this was probably early retirements. Thus corrections had to be made for this domain.

4. Adjustment of the data to include the influence of the fluctuation terms

$$\bar{i}_x = r_x(i_x + s_x^\alpha), \quad r_x = 1.1$$

   s_x^α = fluctuation correction to security level $\alpha = .001$. This adjustment makes the use of special fluctuation terms for the premiums superfluous. As for life insurance, neither the motivation for r_x nor for s_x^α is quite justified.

All in all, the derivation is rather similar to that in life insurance. The following graphs give the raw and smoothened disability decrement rates in a logarithmic scale



Disability decrement rates for men (x) and women

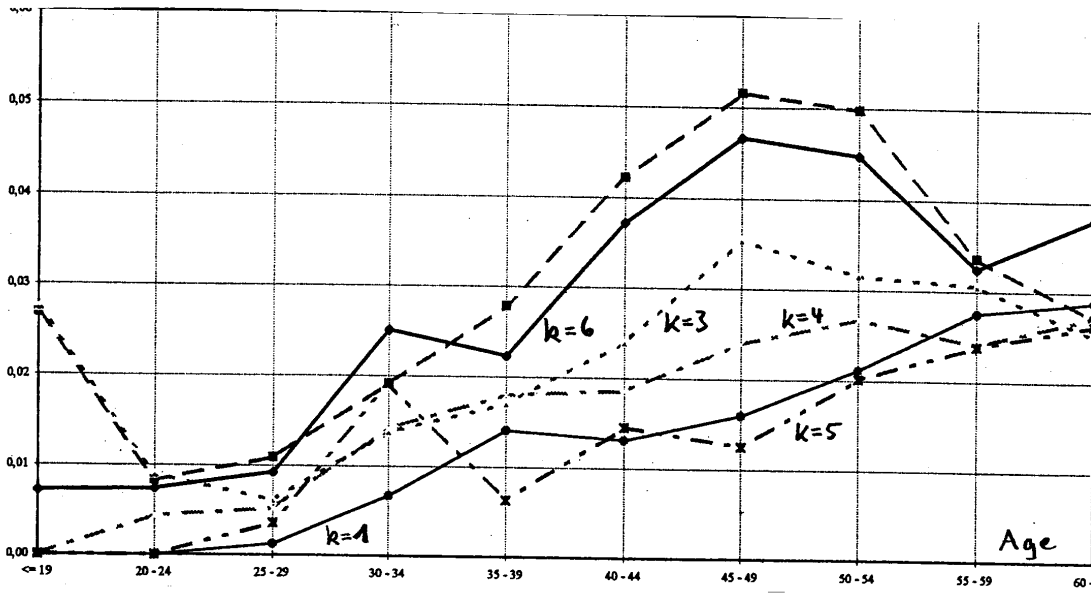
Age	L_x	$\hat{i}_x$	L_y	$\hat{i}_y$	Age	L_x	$\hat{i}_x$	L_y	$\hat{i}_y$
15	200	0,3836	373	0,5163	40	2424	2,5356	1740	2,9390
16	457	0,6383	835	0,9948	41	2260	2,7340	1569	2,9987
17	806	0,9020	1513	0,8184	42	2197	3,2330	1441	2,5418
18	1229	1,3029	2395	0,7625	43	2179	3,4948	1353	4,2580
19	1660	1,5419	3331	0,6901	44	2167	4,1125	1276	4,9877
20	2039	16897	4198	0,9169	45	2291	4,0799	1253	4,7086
21	2378	1,9874	4915	0,8294	46	2522	4,2758	1282	4,7185
22	2670	2,0173	5400	0,7138	47	2585	5,0542	1229	5,1174
23	2915	2,1608	5644	0,9012	48	2605	4,8576	1139	6,3817
24	3080	1,7089	5613	1,0816	49	2591	6,1564	1046	7,4862
25	3189	1,7970	5400	0,8434	50	2444	6,6506	914	9,3073
26	3236	1,7749	5087	1,0282	51	2262	7,7006	767	7,3770
27	3234	1,6304	4706	0,8715	52	2058	9,9519	631	10,5174
28	3157	1,6027	4291	1,2941	53	1798	11,1356	514	10,7012
29	2962	1,7174	3723	1,1510	54	1542	12,1749	403	16,4453
30	2797	1,5946	3268	1,1415	55	1326	15,5644	310	18,4578
31	2722	1,6745	2979	1,2126	56	1159	16,3313	259	21,4638
32	2667	1,5287	2813	1,0890	57	1020	21,8391	225	14,7640
33	2621	1,7292	2634	1,7306	58	929	24,3597	198	15,8084
34	2586	1,8220	2496	1,8798	59	710	28,5415	103	16,2122
35	2560	2,1111	2376	1,7032	60	508	30,5092	18	54,2668
36	2559	1,7196	2262	2,4343	61	419	32,6876	10	26,8265
37	2577	2,4692	2141	2,0297	62	341	31,0721	10	0,1000
38	2562	2,4874	2020	2,0441	63	262	17,3039	10	2,5647
39	2520	2,7253	1889	2,1793	64	13	51,7379	1	2,5647

If one analyzes the risk of invalidity with respect to profession and type of work, one finds considerable differences. Thus insurances will have to adjust these data with respect to the clients they represent. The following table gives the ratio of disability pensions with respect to the low-risk group of professional white-collar workers (data from New York Life Insurance 1968–73)

	Age group	20–29	30–39	40–49	50–59
Men	II	1,61	1,85	1,59	1,35
	III	2,63	2,90	2,24	1,77
	IV	(2,05)	3,41	2,99	2,37
Women	II	1,62	1,60	1,60	1,37
	III	2,56	2,39	2,75	2,05
	IV	(1,83)	(0,28)	1,90	1,18

Relative cost of disability during the first year in relation to that of group I

I Professional white collar
II Tradesmen, Foremen etc
III Skilled Craftsmen etc
IV Heavy Labores, Miners etc



Raw and smoothened mortality rates for disabled men in their k-th year of disability

Raw mortality rates of men age x depending on the year of invalidity

Alter x	q_x^1	q_x^2	q_x^3	q_x^4	q_x^5	q_x^6
≤ 19	0,007171	0,026767	0,027337	0,000000	0,000000	
20 – 24	0,007365	0,008269	0,008958	0,004333	0,000000	0,000000
25 – 29	0,009219	0,010892	0,006099	0,005214	0,003577	0,001265
30 – 34	0,025033	0,019109	0,013754	0,014207	0,019099	0,006676
35 – 39	0,022295	0,027835	0,016996	0,017987	0,006347	0,014077
40 – 44	0,037153	0,042218	0,023943	0,018602	0,014536	0,013070
45 – 49	0,046694	0,051623	0,035338	0,023911	0,012505	0,015990
50 – 54	0,044780	0,049785	0,031494	0,026736	0,020123	0,021144
55 – 59	0,032399	0,033587	0,030579	0,024020	0,023845	0,027497
60 – 64	0,037925	0,027094	0,025015	0,027225	0,026245	0,028799
≥ 65	0,000000	0,031549	0,019956	0,012048	0,004605	0,012698

Smoothened Rates

Alter x	q_x^1	q_x^2	q_x^3	q_x^4	q_x^5	q_x^6
$20-24$	$0,007685$	$0,006684$	$0,005547$	$0,004213$	$0,002750$	$0,001232$
$25-29$	$0,015936$	$0,013731$	$0,011418$	$0,009023$	$0,006508$	$0,003897$
$30-34$	$0,024286$	$0,020839$	$0,017342$	$0,013827$	$0,010253$	$0,006608$
$35-39$	$0,031988$	$0,027596$	$0,022965$	$0,018347$	$0,013798$	$0,009387$
$40-44$	$0,038429$	$0,033356$	$0,027782$	$0,022289$	$0,017152$	$0,012328$
$45-49$	$0,042101$	$0,036999$	$0,031086$	$0,025324$	$0,020251$	$0,015804$
$50-54$	$0,041800$	$0,037455$	$0,032112$	$0,027172$	$0,023169$	$0,019982$
$55-59$	$0,038314$	$0,034901$	$0,031221$	$0,027976$	$0,025801$	$0,024457$
$60-64$	$0,034787$	$0,031649$	$0,029465$	$0,028430$	$0,028224$	$0,028018$

Since the data for this analysis were not sufficiently large – 168 000 (20 000) person years for men (women), the observed fluctuations were large. Thus the age classes had to be grouped – ≤ 19, 20–24, 25–29, ... , 60–64, ≥ 65. From these data it is clear that the mortality of disabled people decreases dramatically with the duration of disability.

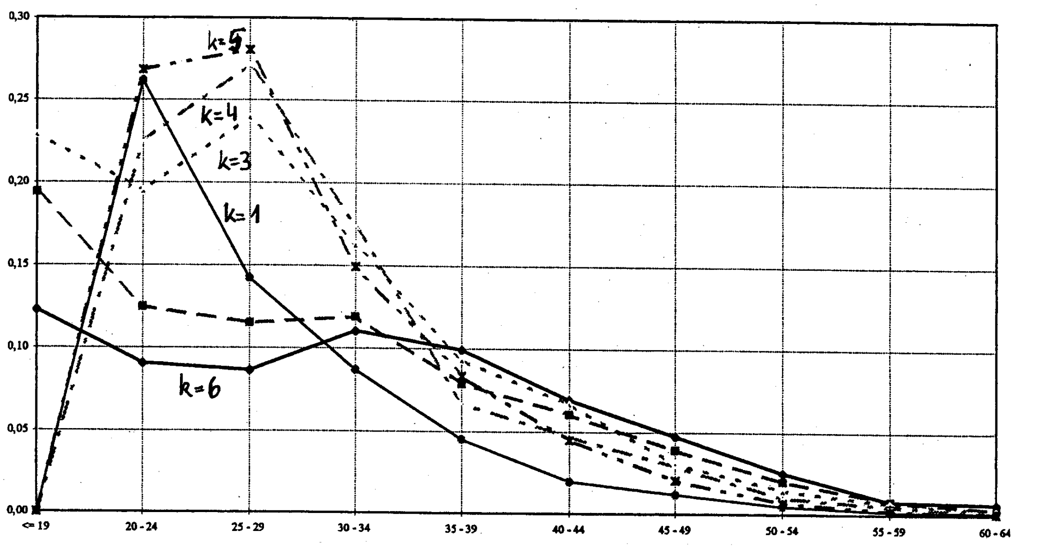
Reactivation intensities of disabled men raw data

Alter x	r_x^1	r_x^2	r_x^3	r_x^4	r_x^5	r_x^6
≤ 19	$0,122766$	$0,193814$	$0,228990$	$0,000000$	$0,000000$	
$20-24$	$0,090436$	$0,124645$	$0,194795$	$0,224915$	$0,268389$	$0,261802$
$25-29$	$0,086256$	$0,114846$	$0,239642$	$0,271602$	$0,280642$	$0,142093$
$30-34$	$0,110193$	$0,118689$	$0,162570$	$0,173723$	$0,149233$	$0,086752$
$35-39$	$0,099017$	$0,079016$	$0,092990$	$0,066908$	$0,083588$	$0,044929$
$40-44$	$0,069072$	$0,060478$	$0,066912$	$0,046968$	$0,044932$	$0,019687$
$45-49$	$0,047443$	$0,039055$	$0,030819$	$0,029101$	$0,020491$	$0,011944$
$50-54$	$0,025175$	$0,019952$	$0,014548$	$0,010205$	$0,006971$	$0,004726$
$55-59$	$0,007910$	$0,008013$	$0,006983$	$0,004852$	$0,002309$	$0,001639$
$60-64$	$0,006414$	$0,003191$	$0,002509$	$0,000884$	$0,000543$	$0,001252$
≥ 65	$0,041237$	$0,042141$	$0,045929$	$0,046990$	$0,036171$	$0,013956$

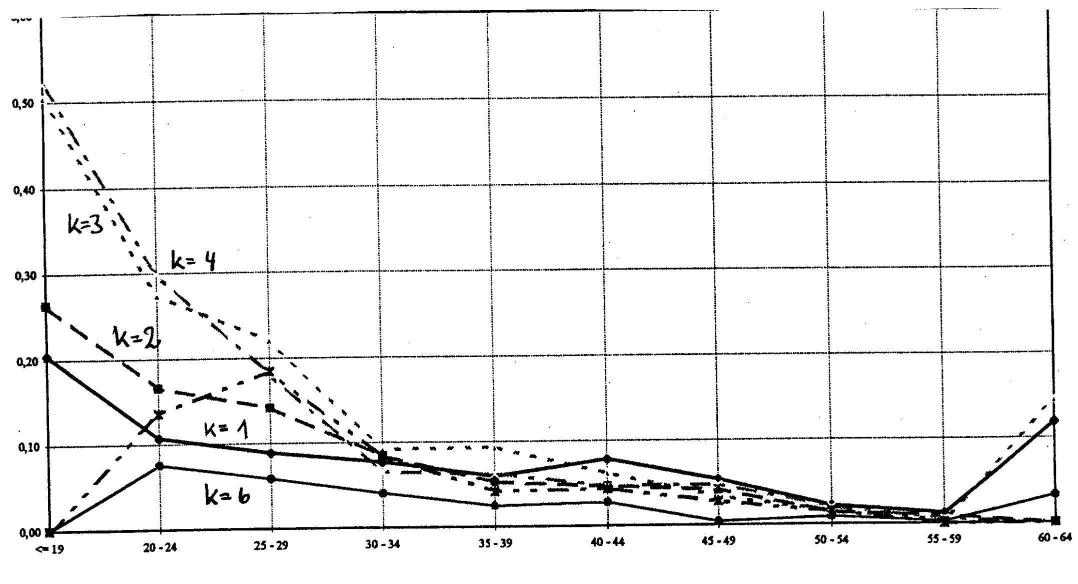
Smoothened data

Alter x	r_x^1	r_x^2	r_x^3	r_x^4	r_x^5	r_x^6
$20-24$	$0,089128$	$0,126030$	$0,223260$	$0,270415$	$0,305457$	$0,206696$
$25-29$	$0,093747$	$0,117763$	$0,210613$	$0,233105$	$0,240444$	$0,144046$
$30-34$	$0,101630$	$0,107002$	$0,164460$	$0,168048$	$0,164173$	$0,089081$
$35-39$	$0,094075$	$0,084112$	$0,108862$	$0,098093$	$0,097024$	$0,047503$
$40-44$	$0,071764$	$0,060630$	$0,065229$	$0,051716$	$0,049313$	$0,022001$
$45-49$	$0,047407$	$0,038794$	$0,032988$	$0,025999$	$0,020910$	$0,010656$
$50-54$	$0,025145$	$0,020455$	$0,014894$	$0,010970$	$0,006954$	$0,004605$
$55-59$	$0,009305$	$0,008524$	$0,006633$	$0,004347$	$0,001786$	$0,001737$
$60-64$	$0,004484$	$0,002287$	$0,001928$	$0,000386$	$0,000000$	$0,000000$

Like mortality, reactivation also depends on the duration of disability. This duration is also age-dependent because at a young age disability is largely due to accidents, while it depends more and more on diseases as x progresses.



Reactivation intensities for men as a function of age, K = years of disability



Reactivation intensities as a function of age, k = years of disability

For the actual decrement table the smoothened $r_x^{(i)}$ are decreased by a certain fraction – .32 – to cover the risk of fluctuation.

These data show that the model for disability presented in the previous section has to be modified in two respects

1. The reactivation has to be taken into account.

2. The duration of disability has to be considered too.

Thus we have for $\ell^i_{z,x}$ = number of disabled persons of age x who became disabled at the age of z

$$\ell^i_{z,x+1} = \ell^i_{z,x}(1 - q^i_{z,x})(1 - r_{z,x}).$$

With this and $p = 4\%$, one gets for the present value of disability pension. For the calculation the q^i_x and r_x have been adjusted for fluctuations.

ω	x	$\ddot{a}^{ai}_x$	$\ddot{a}^{ai}_y$
65	15	6,5036	6,8775
	20	7,4918	7,9533
	25	8,8421	9,3893
	30	10,2493	10,8500
	35	11,2704	11,8788
	40	11,4892	12,0589
	45	10,9520	11,4352
	50	9,6631	10,0104
	55	7,5436	7,7313
	60	4,4106	4,4612

ω	x	$\ddot{a}^{ai}_x$	$\ddot{a}^{ai}_y$
60	15	6,4659	6,8281
	20	7,3903	7,8274
	25	8,6117	9,1144
	30	9,8030	10,3334
	35	10,5150	11,0243
	40	10,3422	10,7854
	45	9,3030	9,6377
	50	7,3555	7,5479
	55	4,3531	4,4099

Problem: Why is $\ddot{a}^i_{20,45}$ so much smaller than $\ddot{a}^i_{40,25}$ and why do the $\ddot{a}^i_{x,\omega-x}$ decrease again with x?

Problem: Interpret the data above. Compare the i_x, q^i_x and r_x for men and women.

Literature

K. Allerdissen, T. Gebhardt, T. Schulz (1996): Angemessene Reservierung in der Lebensversicherung nach Änderung der biometrischen Rechnungsgrundlagen, Blätter der DGVM, Band XXII, Heft 3, pp. 563–580.

H.-R. Dienst: Zur aktuariellen Problematik der Invaliditätsversicherung, Schriftenreihe DGVM, Heft 27, Karlsruhe 1995.

M. Pannenberg (1997): Statistische Schwankungszuschläge für biometrische Rechnungsgrundlagen in der Lebensversicherung, Blätter der DGVM, Band XXIII, Heft 1.

8. Health Insurance

8.1 Mean Health Cost

Health insurance basically comes in three forms

1. Government health insurance
2. Occupational health insurance
3. Individual health insurance.

Of these, the government health insurance is by far the most important. In fact, in all countries of the EU, governments provide basic health care for most people. For demographic reasons, however, this system is facing serious difficulties. The main reasons for this are

1. Senior persons are considerably more costly than young ones. An 80-year-old costs the average 6 to 10 times as much as a 20-year-old.
2. Medical technology is becoming increasingly sophisticated and expensive and older people are much more prone to use it. Think, e.g. of methods to combat coronary diseases, cancer or diabetes and Alzheimer.

All these aspects will make it necessary to overhaul this system. Depending on the political spectrum governments vary between: “Everything that is feasible for everyone” or “Only that, which is necessary for everyone”. It is already clear that the problem we are facing will eventually lead to a strict curtailment of publicly subsidised medical service of the form “No bypass (chemotherapy) beyond the age of 75”. For this reason, the occupational health schemes and private health insurance will become increasingly important, even though this implies better service only for the rich. All this will have to be discussed in detail in various government ethics commussions. Here we shall largely concentrate on the individual health insurance, which at present either plays the role of an additional supplementary insurance or is largely restriced to the group of self-employed people who have opted out of the public system. The extent of this varies considerably between the countries.

As with life insurance premiums, health insurance is computed by the equivalence principle, i.e.

$$\text{Premiums} = \text{Expected Cost.} \tag{8.1.1}$$

Deviations from this are observed occasionally for children and seniorcitizens for political reasons. Beyond that, there are considerable differences between life and health insurance. The most important are

1. The costs can vary considerably from a simple visit to the doctor because of a cold to a heart transplant.
2. One can have multiple damage over the year.
3. The cost increase because of improvements in the medical service. Thus the increase usually exceeds the retail price index.
4. Health insurance is modularized. Coverage is offered for
   (a) visits to the doctor's
   (b) medicine and health treatment
   (c) hospital care including surgery
   (d) dental service
   (e) health cures
   (f) hospitalization insurance

   even though in the end this is sold as a package.

In the remainder we shall restrict ourselves to a particular class of risks. These risks depend on several factors like general health status, mode of living, environment, region, occupational class, gender and age. Of these, gender and age are the most important, and the risk is usually considered as a function of these quantities. The other factors play a minor role, though they may influence the whole portfolio. In addition, insurances may select "good" risks by demanding a health check-up. In the remainder, we will thus treat a fixed group of a particular gender and consider only age as a discriminating variable.

Consider a portfolio of L health insurance policies. In a particular year t under consideration, policy λ, $\lambda = 1, \ldots, L$ incurs n_λ insurance claims $X_1, \ldots, X_{n_\lambda}$. Here n_λ, the number of claims and X_i $i = 1, \ldots, n_\lambda$ are random variables. Most important for the insurance, however, is the total claim

$$Y_\lambda = \sum_{i=1}^{n_\lambda} X_i.$$

These still depend on age x of the insured person and on year t. Thus we write $Y(x,t)$. But for the moment we will neglect dependence on t and simply write $Y(x)$. The distribution of $Y(x)$ can be observed from the record of all x-year-olds in the portfolio, unfortunately only for the preceeding year. Since the individual policies are independent, we may assume that the $Y(x)$ are nearly normally distributed. Even if the individual $Y(x)$ are not normally distributed, this will hold approximately for the average, which is the only observable quantity anyway. The parameters can be estimated from the data $Y_1(x), \ldots, Y_\lambda(x)$ of all L_x policies of x-year-olds. An estimate for $E(Y(x)) = \mu_x$ is thus given by

$$E(Y(x))\hat{} = \hat{\mu}_x = \sum^{L_x} \frac{y_i(x)}{L_x}. \tag{8.1.2}$$

If group L_x under consideration is not fixed because policies terminate or new ones are acquired, one writes

$$L_x \approx \frac{1}{2}(L_{x\text{in}} + L_{x\text{fin}}) \tag{8.1.3}$$

where $L_{x\text{in}}$ ($L_{x\text{fin}}$) is the number of x-year-olds at the beginning (end) of the year. For the number $N(x)$ of claims we have as in (8.1.2)

$$E(N(x))\hat{} = \frac{\sum n_\lambda(x)}{L_x}. \tag{8.1.4}$$

Health problems are not always recognized to their full extent. Thus sometimes claims for a particular year are only claimed at a later date. Thus we write

$$Y(x) = Y_0(x) + Y_1(x) + \cdots + Y_k(x) \tag{8.1.5}$$

where $Y_i(x)$ denotes part of the risk which is claimed only in year i after the event. Usually insurances set a limit on k, i.e. on claims they are willing to honor. Then

$$e(x) = \frac{Y_0(x)}{(Y_0(x) + Y_1(x) + \cdots + Y_k(x))} \tag{8.1.6}$$

gives the fraction of all risks that are claimed immediately. Since $e(x)$ varies slowly, $\hat{\mu}_x$ can be estimated from the immediate claims $\hat{y}_0(x)$ in age group x

$$\hat{\mu}(x) = \frac{\hat{y}_0(x)}{e(x)}. \tag{8.1.7}$$

In general, age classes L_x are too small to yield sufficiently accurate estimates $\hat{\mu}_x$. For this reason, one groups several ages into one age class. Thus, one considers the classes 0–15, 16–20, 21–25, 26–30, … . Correspondingly, one only determines the $\hat{\mu}_{\overline{x}}$ for the central ages $\overline{x} = 8, 18, 23, \ldots$. Thus the function $(\hat{\mu}(x))_{x=20}^{\omega}$ or $(\hat{\mu}(\overline{x}))_{\overline{x}=K\cdot 5+3}$ contains most of the information to compute the premiums because $\hat{\mu}(x)$ is the expected yearly claim for an x-year-old. Since L_x may be rather small, the $\hat{\mu}(x)$ still show too much fluctuation. For this reason the $\hat{\mu}(x)$ are smoothed and adjusted. Then one obtains the lists $(\mu(x))$ and $\mu(\overline{x})$, respectively. In the remainder, we will ignore the age groups and work with the $(\mu(x))$ only.

We have noted above that, in general, the $\mu(x)$ can only be determined for the previous years. For this reason, one will have to devise suitable methods to project the $\mu(x)$ into the future. Thus assume the $\mu(x, t_1), \ldots \mu(x, t_k)$, $x = 21, \ldots, \omega$ are known for past times $t_1 < t_2 < \cdots < t_k < t$. In order to extrapolate these data to t we try to find a class of functions $f(x, t, a_1, \ldots, a_\ell)$ depending on the parameters $a_1, \ldots, a_\ell$, which accurately describes the data in the past. To fix the parameters, one considers

$$\sum_{i=1}^{k} (\mu(x, t_i) - f(x, t_i, a_1, \ldots, a_\ell))^2 = Q(a_1, \ldots, a_\ell) \tag{8.1.8}$$

and minimizes this expression with respect to the $a_1, \ldots, a_\ell$

$$Q(a_1, \ldots, a_\ell) \to \min .$$

If this minimum is attained at $(a_1^*, \ldots, a_\ell^*)$, the function $f(x, t, a_1^*, \ldots, a_\ell^*)$ is the best fit, and values $(f(x, t, a_1^*, \ldots, a_\ell^*))_{x=20}^{\omega}$ are the best guess for the $\mu(x)$ in year t. It should be noted that the a_i^* may vary with x. Above one may also choose a weighted sum in (8.1.8) which, e.g. gives the more recent data a higher weight. This process is called *nonlinear regression.* In its simplest form, *linear regression*, one chooses for the f linear function in t

$$f(x, t, a, b) = a(x)t + b(x). \tag{8.1.9}$$

In this case, there are formulae for $a^*(x)$ and $b^*(x)$. One may also choose f as a polynomial in t polynomial regression. The real problem, however, is the proper choice of a simple family of functions with few parameters which accurately describes the evolution of the $\mu(x, t)$. We have already seen this problem in the graduation of life tables. Function classes used most often in such problems are

1. linear functions $ax + b$
2. polynomials $a_0 + a_1 x + a_2 x^2 + \cdots + a_\ell t^\ell$
3. power functions ax^b
4. exponentials ae^{xb}

The choice of such a function depends very much on the experience of the mathematician. The above process may also be applied to transforms of the $\mu(x)$ like $\ell n \mu(x)$ or so.

The above approach computes the projection individually for each x, i.e. we have a pointwise extrapolation in t for each fixed x. This will, not in general, lead to a function $\mu(x, t)$ which depends smoothly on x. A more sophisticated approach will therefore consider fits in x and t simultaneously, i.e. instead of (8.1.8) we consider

$$\sum_{i=1}^{k} \sum_{x=20}^{\omega} (\mu(x, t_i) - f(x, t_i, a_1, \ldots, a_\ell))^2 L_x(t_i) = Q(a_1, \ldots, a_\ell). \tag{8.1.10}$$

Here, we have also introduced the weight factors $L_x(t_i)$ = size of age class x at t_i. If f depends smoothly on x and t, so will the extrapolated values $\mu(x, t) = f(x, t, a_1^*, \ldots, a_\ell^*)$. Here $a_1^*, \ldots, a_\ell^*$ are again the optimal parameters which minimize Q.

Here are three examples of such an approach

1. $f(x, t) = h_0(t) + h_1(t) g_1(x) + \cdots + h_m(t) g_m(x)$ with $h_i(t) = a_i t + b_i$ and $g_i(x)$ polynomials, exponentials or power functions.

   In view of the scarcity of data, one will have to try to minimize the number of parameters used. This can be achieved using the following example.

2. No change in the profile $f(x,t) = h(t)g(x)$. In this example, $h(t)$ describes the time change in the cost structure, while $g(x)$ describes the cost profile as a function of x. The idea behind this model is the assumption that the general profile will change very little. In this case, one might choose for $h(t)$ a quadratic polynomial.

3. The graduation model
Determine a class of functions $f(x, a_1, \ldots a_\ell)$ that accurately describes the data $(\mu(x,t))_{x=20}^{\omega}$. Consider the corresponding optimal values a_i^* as functions of t and apply regression to each of these functions. With $\tilde{a}_i(t)$ as regression expressions, we can then take

$$f(x, \tilde{a}_i(t), \ldots, a_\ell(t)) = \mu(x,t).$$

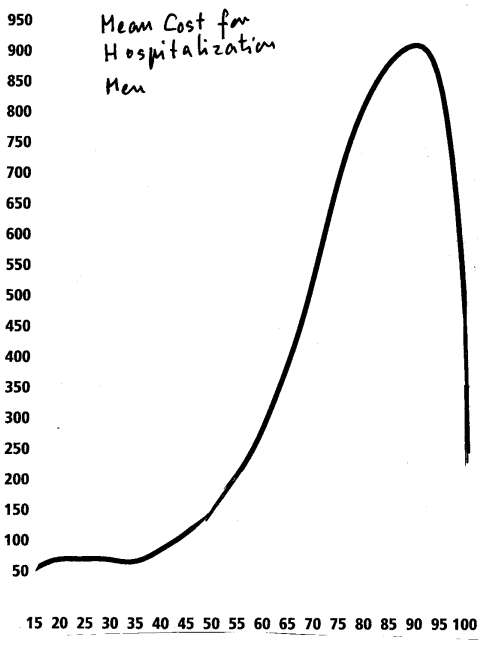
The philosophy behind this approach is that we can expect the $a_i^*(t)$ to vary slowly with t. This last approach has not yet been applied to real data.

Mean Health Cost per Head normalized to age 43
(M = medicine, D = dental care, DS = doctor visits)

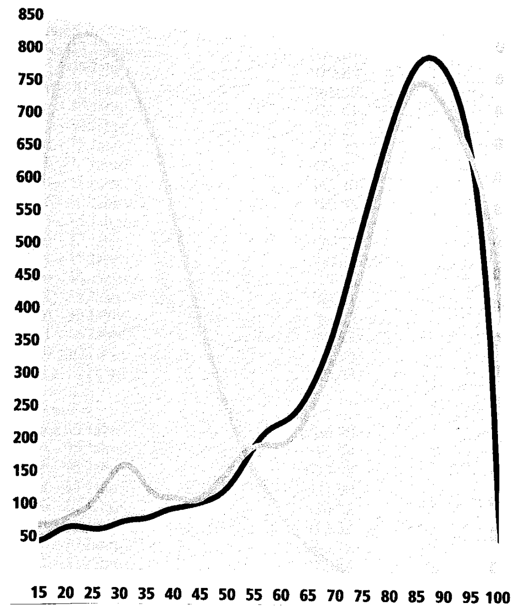
	women			men		
Age	M	D	DS	M	D	DS
≥ 95	471,35	12,70	231,38	588,77	26,48	369,51
91 – 95	512,08	22,08	259,93	771,80	40,89	422,13
86 – 90	496,19	39,21	231,09	688,40	60,07	355,77
81 – 85	487,86	56,92	207,63	666,83	74,32	330,61
76 – 80	412,67	77,82	179,72	613,13	95,15	285,30
71 – 75	355,29	92,64	155,04	526,05	107,71	246,19
66 – 70	306,56	105,65	143,15	425,69	121,47	210,44
61 – 65	257,05	116,41	137,04	318,75	122,05	180,72
56 – 60	221,95	112,75	133,90	240,51	118,26	164,39
51 – 55	181,54	108,29	128,78	185,25	110,02	142,14
46 – 50	132,29	101,82	112,32	135,78	103,97	118,47
41 – 45	100,00	100,00	100,00	100,00	100,00	100,00
36 – 40	89,20	100,17	102,69	85,58	96,51	88,23
31 – 35	80,91	94,86	114,49	66,24	98,39	78,64
26 – 30	69,95	77,75	102,65	59,71	85,06	70,67
21 – 25	66,55	61,98	88,61	55,53	58,77	62,27
16 – 20	64,51	60,13	62,82	73,39	63,01	66,58
16	46,78	64,32	45,94	67,85	63,92	79,61

	men		women	
Age	H	DH	H	DH
≥ 95	595,07	202,27	466,90	574,46
91–95	919,82	729,36	616,54	663,55
86–90	888,59	780,91	673,92	749,23
81–85	873,61	802,16	634,90	755,47
76–80	770,25	626,12	530,28	603,41
71–75	637,32	502,01	412,32	430,89
66–70	474,91	348,83	296,12	311,86
61–65	345,40	253,67	219,64	217,37
56–60	248,85	225,78	187,99	192,67
51–55	187,95	165,22	159,67	180,49
46–50	136,77	120,96	121,63	139,95
41–45	100,00	100,00	100,00	100,00
36–40	75,93	94,25	109,96	123,57
31–35	64,91	77,71	151,83	154,77
26–30	69,63	65,35	133,29	131,87
21–25	65,93	61,06	72,62	101,07
16–20	62,29	55,09	67,76	70,41
below 16	59,03	56,44	40,98	46,98

Mean cost per head in hospitalization H and days in hospital DH for men and women standardized to age 43 = 100



Days in hospital (normalized to 100 for age 40)



Problem: Describe the difference in mean health cost per head for medicine, hospitals and dental care. Explain the difference.

Problem: In most areas (hospitals, general medical services) the functions $\mu(x)$ have become steeper in the higher ages in recent years. Explain the reason for this.

Literature

G. Siegel: Gewichtete Ausgleichsverfahren für Kopfschadensreihen und Leistungsschätzungen in der PKV.
K. Wolfsdorf: Versicherungsmathematik I. Teubner Verlag.
K. Bohn: Die Mathematik der deutschen PKV. Deutsche Ges. für Vers. Mathematik, Heft II

8.2 Premiums in Health Insurance

In 8.1, we have seen how to determine and project the mean health cost per head; $(\mu(x,t))_{x=20}^{\omega}$. These values of course also represent the mean claim of an insured person against the insurance company. On the basis of the equivalence principle, this would also be the net premium. Now we have seen that in many cases the $\mu(x,t)$ increase more than tenfold from $x = 20$ to $x = 80$. On the other hand, the customer expects level premiums, apart from inflation correction. This is only possible if the younger people overpay while the older ones underpay. As with life insurance, this requires reserves and a study of the evolution of the population in the portfolio. In addition, interest has to be considered. Thus let ℓ_x be the number of persons of age x in the portfolio or group under consideration. As for life insurances, this number may be fictitious and normalized to $\ell_{x_0} = 10^5$ for $x_0 = 20$. The evolution of this group is then determined by

$$\ell_{x+1} = \ell_x(1 - q_x - w_x) \approx \ell_x(1 - q_x)(1 - w_x) \tag{8.2.11}$$

where q_x is the mortality of an x-year-old. w_x is the probability of an x-year-old canceling the contract. Cancellation can be the result of moving to another company or can be mandatory if the insured is no longer self-employed. The q_x can be taken from one of the life tables for life insurances. But sometimes health insurance companies use their own or special life tables. Health insurance companies usually only take risk premiums for ages $x \geq 20$ and build up reserves for ages $x \geq 20$.

Problem: For the rate of cancellation w_x one usually expects a function which decreases with x. Why?
Here are some examples:

1. $w_x = .05$
2. $w_{20} = 10^{-2} \cdot k$, $w_{80} = 0$ linear in between $10 \geq k \geq 4$
3. $w_{20} = k \cdot 10^{-2}$ $w_{80} = 0$ w a quadratic polynomial with $w'_{80} = 0$

With ℓ_x defined by (8.2.11)

$$_\nu p_x = \frac{\ell_{x+\nu}}{\ell_x} \tag{8.2.12}$$

gives the probability of an x-year-old being insured with the same company after ν years. Then

$$\ddot{a}_x = \sum_{\nu=0}^{\omega-x-1} {}_v p_x v^\nu \tag{8.2.13}$$

gives the present value of an annuity 1 due. As you might have guessed, this represents the premium side of health insurance in the case of a level premium. The present value of the costs is then represented by

$$L_x = \sum_{\nu=0}^{\omega-x} {}_v p_x \mu(x) v^{\nu+1} \tag{8.2.14}$$

or

$$L_x = \sum_{\nu=0}^{\omega-x} {}_v p_x \mu(x) v^{\nu+.5}. \tag{8.2.15}$$

Quite often this is estimated by

$$L_x = \sum_{\nu=0}^{\omega-x} {}_v p_x \mu(x) v^\nu = \sum_{\nu=0}^{\omega-x} \frac{\ell_{x+\nu} v^{x+\nu}}{\ell_x v^x} \mu(x) = \sum_{\nu=0}^{\omega-x} \frac{D_{x+\nu}}{D_x} = \frac{U_x}{D_x} \tag{8.2.16}$$

where the definition of the commutation numbers D_x and U_x can be read off from (8.2.16). Formula (8.2.15) can then be obtained by the factor $v^{1/2}$ in (8.2.16). With (8.2.15) or (8.2.16) the net level premium is then

$$Pr_x = \frac{L_x}{\ddot{a}_x}. \tag{8.2.17}$$

Formula (8.2.17) gives the net premium. But we still have to consider additional costs. These are the

1. Fluctuation term: As already discussed, the fluctuation can be computed according to the expectation value principle, according to the standard deviation principle or it can be proportional to the variance. But there are also other terms. In Germany the fluctuation term is usually built into the function mean cost per head in terms of correction factors. In the remainder we will therefore consider the fluctuation term as part of Pr_x, otherwise replace $Pr_x + Fl_x$ by Pr_x.

   In addition, we have to consider the administrative costs

2. Administrative costs
   α_1: Cost of acquisition, medical examinations (immediate cost)
   α_2: Marketing (running expenses)
   β: Cost for collecting and administration
   ρ: Cost for disbursement and claims regulation
   All these cost are considered proportional to the yearly or monthly premium. Finally, there might be cost independent of the premium γ.

Thus if B_x is the yearly premium, we can write

$$\begin{array}{rlll} B_x = & Pr_x + & \alpha_1 \dfrac{B_x}{\ddot{a}} + & (\alpha_2 + \beta + \rho) B_x + \gamma \\ & \text{Premium} & \text{immediate} & \text{running cost} \\ & + \text{Fluct.t.} & \text{cost} & \end{array} \tag{8.2.18}$$

We write this as

$$B_x = Pr_x + \Delta B_x + \gamma \tag{8.2.19}$$

or

$$B_x = \frac{Pr_x + \gamma}{(1 - \Delta)}. \tag{8.2.20}$$

The monthly premium is usually $1/12 \cdot B_x$. Since the majority of all policies go with level premiums, reserves have to be considered. To do this we proceed as for life insurances and consider a general set-up of premiums $(B_x)_{\nu=0}^{\omega-x}$. Then the present value of claims is

$$\mathcal{C}(0) = \sum_{\nu=0}^{\omega-x} {}_v p_x \mu_{x+\nu} v^\nu \tag{8.2.21}$$

if we assume payment at the beginning of the year. The present value of the premiums correspond to

$$Pr = \sum_{\nu=0}^{\omega-x} {}_\nu p_x B_\nu v^\nu. \tag{8.2.22}$$

Hence the prospective reserves after m years are

$$V_m = \sum_{\nu=0}^{\omega-x-m} {}_v p_{x+m} (\mu_{x+m+\nu} - B_{\nu+m}) v^\nu. \tag{8.2.23}$$

We could also define the reserves retrospectively. But as for life insurance, which is actually slightly more general, we have the equality of prospective and retrospective reserves. Of course, we also have the well-known recursion relation in this case

$$\left(1 + \frac{p}{100}\right)(V_{m-1} + B_{m-1}) = p_{x+m-1} V_m + \mu_{x+m-1}\left(1 + \frac{p}{100}\right). \tag{8.2.24}$$

The factor $(1 + p/100)$ with μ_{x+m-1} arises because we assumed the health cost to arise at the beginning of the year.

Problem 8.1 Verify (8.2.24).

From (8.2.24) we obtain

$$p_{x+m-1} V_m v = V_{m-1} + B_{m-1} - \mu_{x+m-1}$$

or

$$B_{m-1} = \mu_{x+m-1} + (p_{x+m-1} V_m v - V_{m-1}). \tag{8.2.25}$$

This formula gives the decomposition of the premium into the risked part and the part which is saved. The gross premium, which encompasses the fluctuation term, the initial cost and the running cost for administration can now be computed as for life insurances. So we will not follow this up.

In life insurance the insured person has a right to a substantial part of the reserves upon cancellation. This is generally not so with health insurance. But when changing the contract the reserves are usually taken into account.

Problem 8.2 Prove $B_{m-1} = \mu_{x+m-1} + (V_m v - V_{m-1}) - (q_{x+m-1} + \omega_{x+m-1}) V_m$.

There are a number of particularities for health insurance that have to be taken into account.

Excess of loss plan: With such a policy the insured will only be reimbursed $\mu_x - K$, if this is positive. Otherwise he will get nothing. Here K is fixed in advance.

Problem 8.3 Derive the most important actuarial function for such a policy.

Partial Benefits: There are policies which cover only a certain percentage, $k\%$, of the damage.

Problem 8.4 Give reasons why the net premium might differ from $k/100\cdot$ premium in this case.

Supplementary premium, surcharge: In general, the insurance demands a health check before issuing a policy. Nonetheless, it will also insure people with a known health status. Such persons, however, will have to pay an additional risk surcharge.

Problem 8.5 Describe the steps which are necessary to determine such a surcharge.

Problem 8.6 You wish to develop a health insurance policy for nonsmokers. How do you proceed?

In recent years, costs in the health sector have increased more than the retail index. This makes it necessary to correct the premiums. This can, of course, not be done at each moment when such changes have occurred. Insurances also like to have stability of premiums because a change in premiums causes a lot of administrative activity. What is needed, therefore, are criteria to when such changes in premiums are necessary. Apart from marketing points of view, the only criteria can be the cost. Thus let $\mathcal{C}_{ac}(t)$ denote the actual claims for a particular year t, while $\mathcal{C}(t) = \sum \ell_x \mu_x(x,t)$ gives the expected claims. If then

$$\frac{\mathcal{C}_{ac}(t)}{\mathcal{C}(t)} > 1 + a \tag{8.2.26}$$

with a fixed in advance, new premiums have to be calculated. Usually $a \approx .1$. There are a number of problems with this approach

1. $\mathcal{C}_{ac}(t)$ has to be estimated from the claims that have been made in that year.

2. If the fluctuation terms have been built into the series of $\{\mu(x,t)\}$, as is custumarily the case in Germany, corrections have to be made for this. In general, it is not always easy to see to what extent high claims $\mathcal{C}_{ac}(t)$ are fluctuation-induced.

3. Finally, the calculations in (8.2.26) have to be made in advance, i.e. $\mathcal{C}_{ac}(t)$ has to be estimated for the next year. For this one usually extrapolates $\mathcal{C}_{ac}(t-l-1), \ldots, \mathcal{C}_{ac}(t-1)$ to $\mathcal{C}_{ac}(t)$.

9. Nonlife insurance

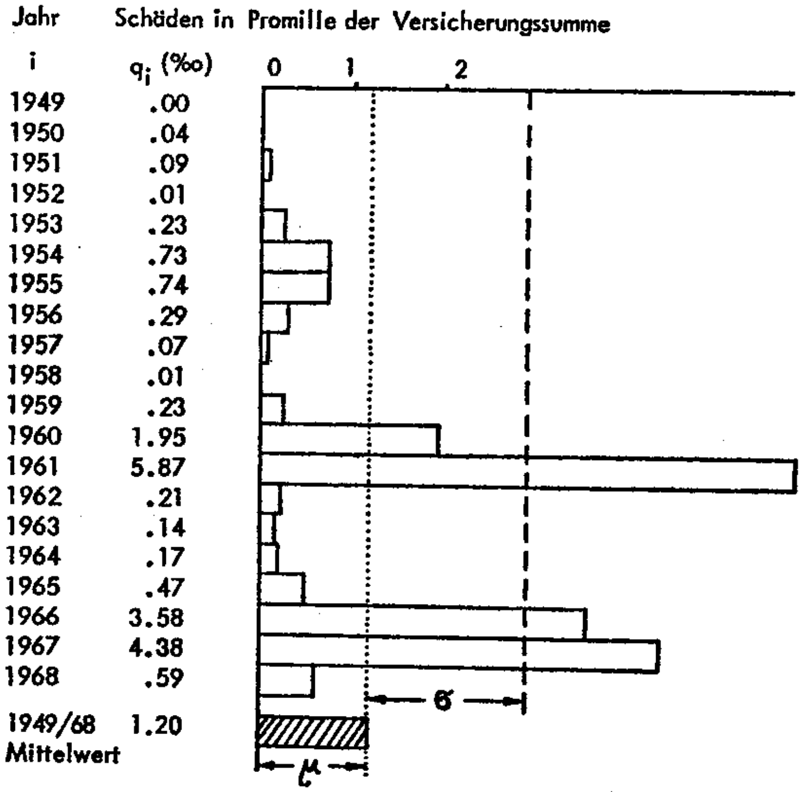
9.1 Fire insurance

Fire insurance and automobile insurance are in some sense the simplest form of insurance. For this reason we treated fire insurance examples first. What makes fire insurance so simple?

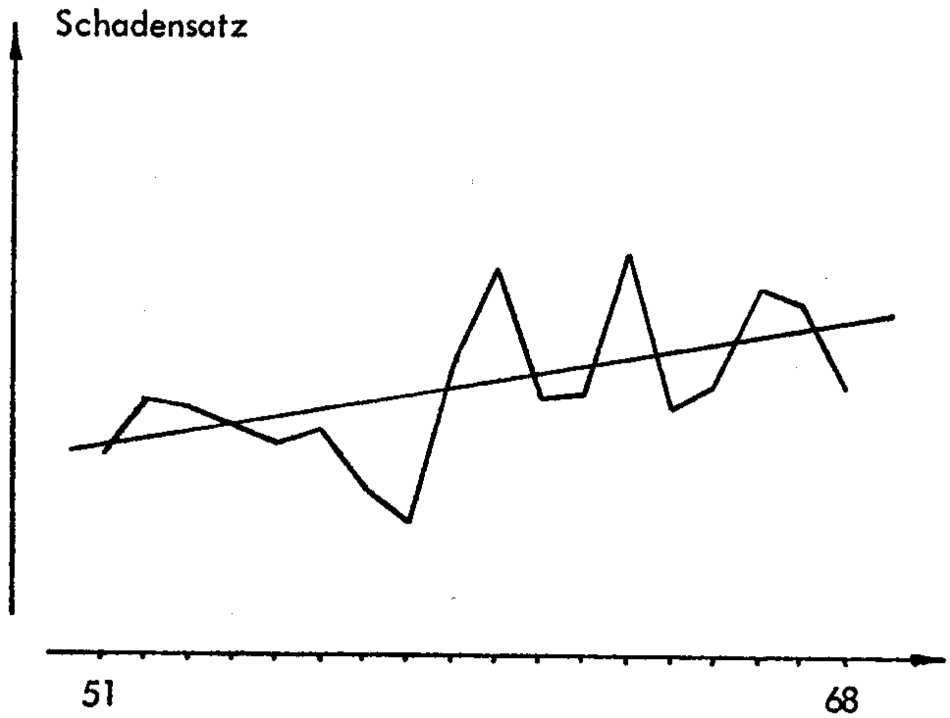
1. The insurance is made on a year-to-year basis, so there are no savings effects to consider. The reserves only have the function to compensate the fluctuations.

2. The variations in time, due to increasing cost, are small. They are usually of the order of the rate of inflation. Thus the rise in prices can be estimated easily from past records.

Variation in claims for a portfolio in fire insurance (steel mills) μ = mean value (Flach, Strauss)



Variation of claims in industrial fire insurance (Flach, Strauss)



What makes it complicated?

1. The risk depends on many factors. Important factors for family homes are

   (a) Type and build of house
   (b) Material used
   (c) Location
   (d) Fire prevention measures
   (e) Price of house

   For industrial fire insurance, factors are

   (a) Type and build of building
   (b) Use of building
   (c) Type of industry
   (d) Location and addess
   (e) Fire prevention measures
   (f) Value of objects

A Fire insurance statistics for steel mills (Flach, Strauss)

Claim cost in 10^3 DM	Number of claims	Mean Claim cost in 10^3 DM
0	305	0,23
0,5	259	0,74
1	374	1,47
2	264	2,47
3	199	3,49
4	163	4,51
5	125	5,54
6	178	7,01
8	139	9,04
10	98	11,0
12	112	13,5
15	111	17,6
20	141	25,0
30	147	39,0
50	152	71,1
100	93	136
200	64	310
500	20	705
1000	13	2346
5000	4	8317

Beneckert and Jung [Statistical Models of Claim distributions in Fire Insurance; Astin Bulletin **8** 1973, p. 1–25] found that the lognormal distribution gives a good description for claims in fire insurance for single brick houses.

For this reason, the fire insurance forms risk classes that are expected to have similar risk characteristics. Usually, one has some hundred or so classes. Thus the classes will in the case of fire insurance have about 10^3 policies. These classes are barely large enough.

2. A more serious feature in industrial fire insurance is the occurrence of very large claims – large fires in large establishments. As a consequence, the sum of all claims shows large fluctuations (see the graphs above).

3. The values of the various objects within each class may vary considerably. E.g. if you consider the class of saw mills, furniture factories, the value may vary from 10^5 to 10^8.

3 is easiest to overcome. In this case, one simply chooses the net premium proportional to the insured value. To make things more precise, we shall restrict ourselves now to a single class, e.g. the class of single detached family homes in suburban areas. If V is the face value of the insurance, the random variable X, which describes the risk in this class, is then

$$X = \frac{S}{V} \tag{9.1.1}$$

where S is the actual damage.

Then let V_i denote the sum of the face values of all policies in that class in year i and let D_i be the corresponding total damage i.e. the sum of all valid claims. Then

$$X_i = \frac{D_i}{V_i} \tag{9.1.2}$$

is the risk (rate) in year i for that class. We will assume that D_i and X_i are corrected for the general increase in the cost of living, so that the following assumptions are plausible for the years $1, \ldots, n$ of observation

1. The $X_1, \ldots, X_n$ are independent random variables
2. $E(X_i) = \mu$ is independent of i
3. $\sigma^2_{X_i} = \frac{\sigma^2}{V_i}$ with σ^2 independent of i.

The assumption that $\sigma^2_{X_i} = \sigma^2/V_i$ arises if one assumes D_i to be made up of V_i identical independent submits with the same risk characteristics, i.e. $D_i = X_{i_1} + \cdots + X_{i_K}$ with $i_K = V_i$. Then $\sigma^2_{D_i} = V_i\sigma^2$ to be proportional to $E(D) = \mu V_i$. This is indeed the case.

The problem is now to estimate ν and σ^2, because we can expect that each X_i is nearly normally distributed. For this estimate we need the following Lemma

Lemma 9.1 Let $X_1, \ldots, X_n$ be independent random variables with $\mu = E(X_i)$ but variances $\sigma_1^2, \ldots, \sigma_n^2$. Then

$$T = \sum_{i=1}^{n} c_i X_i, \qquad c_i = \left(\sum_{j=1}^{n} \frac{\sigma_i^2}{\sigma_j^2} \right)^{-1} \tag{9.1.3}$$

is an unbiased estimator for μ, with minimal variance among all estimates for μ, which are linear in the X_i. Its variance is $(\sum \sigma_i^{-2})^{-1}$.

Proof. Let $T' = \sum \alpha_i X_i$ be such an estimator. Since it is unbiased, we have $\sum \alpha_i = 1$. The X_i are independent. Thus we have

$$\sigma^2_{T'} = \sum \alpha_i^2 \sigma_i^2.$$

This has to be minimized by a proper choice of the α_i. To do this, we write for arbitrary ℓ

$$\sum \alpha_i^2 \sigma_i^2 = \sum \left(\alpha_i \sigma_i - \frac{\ell}{\sigma_i} \right)^2 + 2\ell - \ell^2 \sum_{i=1}^{n} \sigma_i^{-2}.$$

The sum on the right-hand side is clearly minimal if $\alpha_i = \frac{\ell}{\sigma_i^2}$. Since $\sum \alpha_i = 1$, we get $\ell = (\sum \sigma_i^{-2})^{-1}$.

Problem 9.1 Derive this result using the well-known method with Lagrange multipliers.

With this Lemma we can show

Theorem 9.1 1.
$$\overline{X} = \sum \frac{V_i}{V} X_i, \qquad V = \sum V_i$$
has minimal variance among all linear estimators for μ.
$$\sigma^2_{\overline{X}} = \frac{\sigma^2}{V} \tag{9.1.4}$$

2. $\overline{S}^2 = \frac{1}{n-1} \sum V_i (X_i - \overline{X})^2$ is an unbiased estimator for σ^2.

3. If all X_i are normally distributed
$$T = \sqrt{V} \frac{(\overline{X} - \mu)}{\overline{S}} \tag{9.1.5}$$
is t_{n-1} distributed.

Proof.

1. follows directly from the Lemma.

2. By replacing X_i, $\overline{X}$ by $X_i - \mu$, $\overline{X} - \mu$ if necessary, we may assume $\mu = 0$.
$$\begin{aligned} E(\overline{S}^2) &= \frac{1}{n-1} E\left(\sum_{i=1}^{n} V_i X_i^2 - 2 V_i X_i \overline{X} + V_i \overline{X}^2 \right) \\ &= \frac{1}{n-1} \left(\sum_{i=1}^{n} V_i \frac{\sigma^2}{V_i} \right) - \frac{2}{n-1} V E(\overline{X}^2) + \frac{V}{n-1} E(\overline{X}^2) \\ &= \frac{n}{n-1} \sigma^2 - \frac{V}{n-1} E(\overline{X}^2) = \sigma^2. \end{aligned}$$

3. The fact that $\overline{X}$ and $\overline{S}$ are independent and that $\overline{S}^2$ is χ^2_{n-1} distributed follows with the Helmert transformation. Since this is dealt with in most statistics texts in connection with the t-distribution, we will omit the proof.

The theorem now tells us how to estimate μ, which is essentially the net premium per face value. It also tells us how to estimate the fluctuations.

Above, we assumed that all X_i are nearly normally distributed with the same μ. Unfortunately, the actual observed values show large fluctuations in the domain of industrial fire insurance. These fluctuations are due to spurious large risks. These may be so large that they exceed the trend and make it impossible to recognize μ and thus estimate the net premium. For this reason, Flach, Strauss proposed to split X into a *base risk* or basic claim and an *excess risk* or peak claim, by artificially reducing the variation. They define
$$Y = \min(X, \mu + a\overline{S}) \tag{9.1.6}$$

Here a has to be chosen such that the Y_i are nearly normally distributed. One finds, by using a χ^2-test, that $a = 1$ is a suitable parameter for *all* classes, even though σ may vary considerably. Meanwhile insurances even split off a so called super excess risk. Such risks exceed $\mu + k\sigma$ with $k \approx 40$.

A fire is a rare event, thus, on the basis of a Bernoulli model or a Poisson model, one would expect

$$\sigma \approx \sqrt{p(1-p)} \approx \sqrt{\mu}.$$

For fire insurance this relation has not been found. Flach and Strauss found

$$\sigma \approx 1.0426 \cdot \mu^{1{,}002} \tag{9.1.7}$$

with

$$0.3021\mu^{1{,}02} \geq \sigma \geq 3.5973 \cdot \mu^{1.02}$$

a 95% confidence interval. For the base risks they found

$$\sigma \approx c \cdot \mu^{0.9335}. \tag{9.1.8}$$

Premiums in fire insurance

In order to compute the premium in fire insurance, let us return to our initial model and neglect excess risks for the moment. On the basis of the standard deviation principle,

$$Pr_i = E(D_i) + \lambda_i \sigma_{D_i}.$$

If one chooses $\lambda_i = \lambda/\sqrt{V_i}$ with λ independent of i, one gets

$$\frac{Pr_i}{V_i} = \mu + \lambda\sigma. \tag{9.1.9}$$

Since $\overline{X}$ and $\overline{S}$ are unbiased estimates for μ and σ, respectively, the empirical premium rate should be given by

$$\frac{Pr}{V} = \overline{X} + \lambda\overline{S}. \tag{9.1.10}$$

In order to determine λi, one uses a fixed probability α, which is the chance that the random variable $\overline{X}$ exceeds $\overline{x} + \lambda\overline{s}$, where $\overline{x}$, $\overline{s}$ are the actual measured values. One obtains

$$\lambda = \frac{t_{n-1}(\alpha)}{\sqrt{V}}. \tag{9.1.11}$$

It has been found that

$$\frac{\text{total damage}}{\text{base damage}} \approx \text{const.} = \gamma. \tag{9.1.12}$$

Of course $\gamma > 1$. Usually the commissions and cost of administration are taken proportional to the base premium rate, with a proportionality factor of β, $\beta \approx$

$3 - -7\%$. With this and the transition to the base premium, the estimated gross premium rate becomes

$$Pr = \frac{\gamma}{1-\beta}\left(\overline{Y} + \frac{t_{n-1}(\alpha)}{\sqrt{V}}\overline{S}_y\right). \tag{9.1.13}$$

Two remarks are in order with this formula.

1. This formula still has to be augmented by a trend factor that takes into account rising costs.
2. In the derivation of (9.1.13) we assumed solvency probability α for *each* class. In general, however, a fire insuance will have many classes. So the solvency condition should actually be for a portfolio which encompasses several (many) classes.

Above we had used the standard derivation principle for the fluctuation term. Even if (9.1.13) gives a nice and simple formula, we will now see that the variance principle might be better. Let us now consider this latter problem in more detail, because it also applies to other domains of nonlife insurance. Above we had seen

$$\begin{aligned}\text{Premium} =&\text{Expectation value of the risk } X + \text{Fluctuation term}\\ &\text{Administrative Term} + \text{Earnings}\end{aligned}$$

or formally

$$Pr = E(X) + Fl(X) + \lambda E(X) + \epsilon E(X). \tag{9.1.14}$$

It was necessary to introduce the fluctuation term to reduce the risk due to fluctuations for the insurance. Since the premium already contains the net premium $E(X)$, the risk for the insurance arises from the deviation of the actual claims. The risk is thus closely connected with the variance of the portfolio during each period. A part of the premium income, usually the part arising from the fluctuation term, can be invested and saved. These savings and gains from this investment form the reserves. In times when the fluctuations lead to a large demand in claims part of the reserves can and will be used to satisfy these claims. In times of low claims, the reserves will be stocked again. Thus even in non life insurance reserves play an important role. To put it concisely, we see that an insurance company works because it spreads the risk (fluctuations) over the whole portfolio. The role of the reserves is thus to spread the risk in time over many years over the portfolio. One of the difficulties in this context is then to distinguish fluctuations from actual trends.

Assume K are the reserves plus the capital of the insurance company set aside for insurance operations. Let r_1 be the interest rate of very secure fixed interest bonds and let r_2 be the interest rate the insurance market is expecting for these insurance operations. Since these are more risky we have $r_2 > r_1$ and at least $(r_2 - r_1)K$ has to be financed from the fluctuation terms. Now assume the company has the portfolios $Z_1, \ldots, Z_\ell$. In our case each Z_i corresponds to a particular risk class in fire insurance. Now we have to determine the contribution of each risk class Z_i to the total portfolio so that the overall bankrupcy probability is less than α, with α given in advance. In chapter I we had considered this problem for $\ell = 1$ assuming a normal distribution of Z. Let us now assume that all Z_i are normally distributed with $E(Z_i) = \mu_i$ and $\sigma^2_{Z_i} = \sigma_i^2$. Then with

$$Z = \sum Z_i$$

we have

$$\mu = \sum \mu_i = E(Z), \qquad \sigma^2 = \sigma_Z^2 = \sum \sigma_i^2. \tag{9.1.15}$$

Let us write now Pr_i for the premium part of the i-th subportfolio. Since the administrative part and the earnings do not contribute to our risk considerations, let us just write

$$Pr_i = \mu_i + Fl_i \quad \text{and} \quad Fl = \sum Fl_i.$$

The bankrupcy condition is then in the simplest case

$$\phi\left(\frac{K}{\sigma}\right) = 1 - \alpha, \quad \left(\sum Fl_i\right) = (r_2 - r_1)K.$$

In order to determine Fl_i one has now several options

1. Fl_i is chosen such that the bankruptcy probabilities are the same in each subportfolio.
2. $Fl_i = \lambda E(Z_i)$ with $\lambda = (r_2 - r_1)\ K/E(Z)$
3. $Fl_i = \lambda\sigma_i$ with $\lambda = (r_2 - r_1)\ K/\sum \sigma_i$.
4. $Fl_i = \lambda\sigma_i^2$ with $\lambda = (r_2 - r_1)\ K/\sum_{j=1}^n \sigma_j^2$.

We will now discuss these various rules of allocating the Fl_i to the total fluctuation term. The basic argument will be of a game theoretic nature in which subportfolios are united respectively split up into smaller units. In game theory this runs under the name of "coalitions". You are supposed to discuss these various cases in the form of problems with some simple examples.

Problem 9.2 Assume Z_1, Z_2 given with $E(Z_1) = \mu$, $E(Z_2) = k\mu$ and assume Z_2 consists of k independent subportfolios with the same distribution as Z_1. Then $\sigma_{Z_1} = \sigma$ and $\sigma_{Z_2} = \sqrt{k}\sigma$. Now compute the contribution of Z_1, Z_2 and each of the subportfolios of Z_2 to the overall fluctuation term Fl. Is this fair?

The considerations of problem 2 also show that the contribution of each subportfolio should be independent of the fact that it is part of another subportfolio or split up into smaller units. Thus we need a form of additivity. The expectation value certainly has this property. This makes case (2.), i.e. the expectation value principle a good candidate. However, consider

Problem 9.3 $E(Z_1) = E(Z_2) = \mu$ but $\sigma_{Z_2}^2 = k\sigma_{Z_1}^2 = k\sigma^2$ with k large, i.e. Z_2 is rather insecure because of its large variance. Now compute the contribution of each subportfolio to Fl and the risks arising from each part.

Since the standard deviation principle, case (3.) is *not* additive, this does not qualify either. So the only reasonable way, which remains, is the variance principle. This is not only additive in the above sense, but in game theoretic language also defines an (Nash) equilibrium in cooperative game theory. This means the variance or contribution to Fl cannot be improved by forming other, possibly larger subportfolios. On the basis of the variance principle (9.1.13) is now

$$Pr_i = \frac{\gamma}{(1-\beta)}\left[\overline{X}_i + \lambda \frac{\overline{S}_i^2}{\sum \overline{S}_j^2}(r_2 - r_1)K\right] \tag{9.1.16}$$

where λ is chosen such that

$$\phi\left(\frac{K}{(\sum \overline{S}_j^2)^{1/2}}\right) = 1 - \alpha.$$

Volume dependence in nonlife insurance

In many domains of nonlife insurance, not only in fire insurance, the risks have different volumes, i.e. the face value of the insurance policy may vary quite a bit. Let us thus assume that we have a fixed risk group in nonlife insurance and a portfolio of independent risks $Y_1, \ldots, Y_\ell$ with face values $v_1, \ldots, v_\ell$. Quite often it is assumed that there is a form of standard risk, which we assume to be Y_1 so that

$$E(Y_i) = E(Y_1) \cdot \frac{v_i}{v_1}, \qquad \sigma_{Y_1}^2 \cdot \frac{v_i}{v_1}. \tag{9.1.17}$$

Let $V = \sum v_i$ and $Y = \sum Y_i$. Then

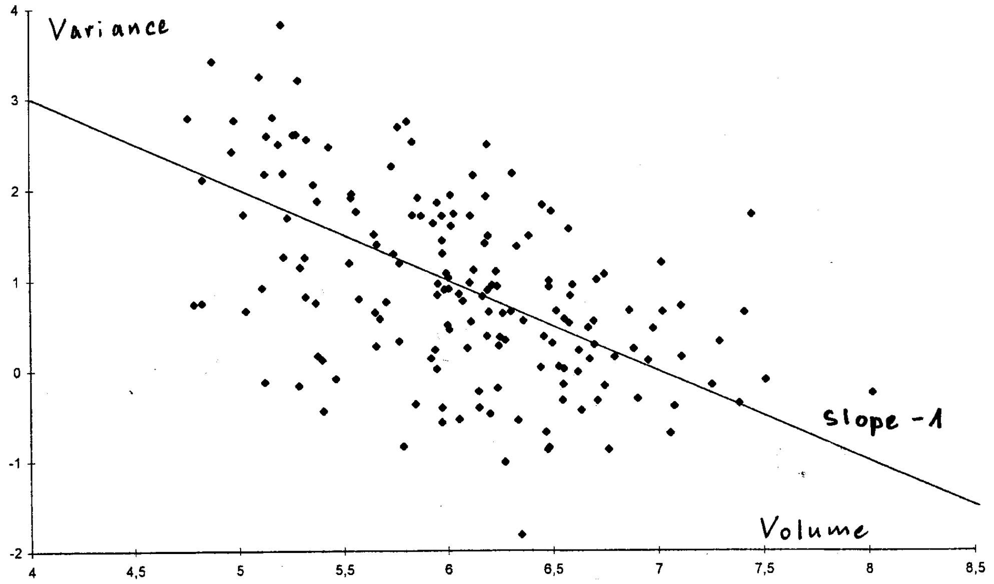
$$E(Y) = \frac{E(Y_1)}{v_1}V \quad \text{and} \quad \sigma_Y^2 = \sigma_{Y_1}^2 \cdot \frac{V}{v_1}. \tag{9.1.18}$$

In this case it is quite common to introduce the claim ratio $X_i = Y_i/v_i$ as a new quantity. With this one gets with $Y = X/v$

$$E(X_i) = \mu = E(X),\ \sigma_{X_i}^2 = \frac{\sigma^2}{v_i} \text{ and } \sigma_X^2 = \frac{\sigma^2}{V} \tag{9.1.19}$$

where $\sigma^2 = \sigma_{X_1}^2$.

Of course it has to be verified by statistical testing that the variance condition in (9.1.17) holds. For fire insurance this is indeed the case



For automobile insurance this will also hold, because there the v_i are nearly constant.

Literature:

D.Flach, J. Strauss: Analyse der deutschen Feuer-Industrie Statistik. Bl. der deutsch. Geschichte f"ur Vers., Vol. 9, p. 407–416, 1970.

9.2 Automobile Insurance

Within the EU about every other person possesses a car. Thus car and automobile liability insurance represents a large and important segment of the insurance industry. As usual for non life insurance the risks are classified according to their risk characteristics. For this one needs to know the various factors that influence the claimsize:

1. Technical variables:
   (a) Size and build of vehicle,
   (b) Type of car (sports car, van, ...),
   (c) Power of vehicle,
   (d) Amount of safety features (ABS system, airbags, re-enforced passenger compartment),
   (e) Age of vehicle,
   (f) Technical status of tires and parts,
   (g) Regular care or not.
2. Characteristics of driver:
   (a) Age of driver,

   (b) Health status of driver (coronary problems, reaction time),
   (c) Sex of driver,
   (d) Marital status, number of children,
   (e) Social status,
   (f) occupation,
   (g) Nationality,
   (h) Use by other people,
   (i) Psychic stability of driver.

3. Use of vehicle:
   (a) Average amount of mileage per year,
   (b) Commercial use or use within the family only.
4. Geographical variables:
   Location of main use of vehicle (urban or rural areas).
5. Other factors:
   (a) Age of insurance policy,
   (b) Type of coverage.

Even though most of these variables are selfexpanatory, a few remarks are in order. While the influence of I1, I2, I5 and I6 on the liability are immediately obvious, I4 lowers the claim size and thus the liability of the second party. On the other hand it has been largely the improvements in I4 as legal standards that the claims per vehicle decreased in recent years, despite an increase in traffic density. Even though many of the individual factors like sex, age, ... influence the probability to have an accident, it is quite often not accepted politically to use these variables, because these factors may be discriminatory. This applies in particular to sex age and nationality, even though it is known for example that foreigners contribut up to 30% more to accidents. The bonus-malus system is criticised, because its basic principle goes against the idea of insurance, which means sharing of risk. On the other hand it might be seen as an expression for the experience of the driver.

Two things are apparent from this list. There are too many variables. If one has n variables and each of these variables takes on 2 (3, 4) values, one is automatically led to $2^n(3^n, 4^n)$ classes. Secondly these variables are correlated and of different importance. So the question arises, which variables are important and how do we find them. As a first step let us look at a rough breakdown of the costs.

Components of Claim Costs (J. Lemaire)

Repair of material damage	33 %
Legal costs	4.5 %
Bodily injury	60 %
other factors	2.5 %

Even though this explains the domainance of "bodily injury" and the necessity for safety measures like safety belts, air bags, re-enforced passenger compartments, head rests, ... , it does not give us any clue about the important variables. What we need is a reduction of the 20 or so parameters to two or three variables, which capture the essentials of the claims in automobile insurance. These key variables should be largely independent. To formulate this mathematically let $X_1, \ldots, X_{20}$ be the variables above. We are looking for new variables

$$Z_i = f_i(X_1, \ldots, X_{20}) \qquad i = 1, 2, 3$$

so that the claims X are functions of the Z_i

$$X = g(Z_1, Z_2, Z_3)$$

and so that the Z_i are independent or at least uncorrelated. This problem of selecting the important independent variables arises in many statistical applications. But there does not exist a general method as of now how to proceed in this case. Only in a very special case does *factor analysis* give a solution. Factor analysis has been invented by psychologists (Spearman) in connection with IQ tests, in order to determine the principal components of intelligence. Even though we will not apply this methods to automobile insurance date, we will briefly present an outline. This method depends on the assumption:

> The data have to be distributed according to a multivariate normal distribution.

In our case we would plot the actual observed claim variables as functions of $X_1, \ldots,$ X_{20}. The data are thus realizations of a normal distribution in $\mathbb{R}^{20}$ and if we plot the observed valued $x_1, \ldots, x_n$, $x_i \in \mathbb{R}^{20}$ we get a cloud of points in 20-dimensional space. The density of this cloud is approximately constant on ellipsoids – this is a consequence of the assumption of multivariate normality. Now rotate the coordinate system so that the principal axes of these ellipsoids agrees with the coordinate axes. The new coordinates $z_1, \ldots, z_{20}$ are related to the old ones by a linear transformation

$$z_i = \sum a_{ij} x_j.$$

Since the z_i agree with the coordinate axes, the random variables

$$Z_i = \sum a_{ij} x_j$$

are uncorrelated and thus even independent. In most practical problems it turns out that these ellipsoids are rather flat in certain direction and extended in very few directions, i.e. they look more like a higher dimensional ellipsoidal pancake. It is now obvious that the flat coordinates contribute little to the variation, which is essentially due to the extended ones. So, if there are essentially only the extended coordinates Z_1, Z_2, Z_3, those with the largest semiaxes, we have found the most relevant independent variables. In addition, we will have to assume that the claimsize depends essentially on the variables Z_1, Z_2, Z_3. So far, such an analysis has not been carried out for automobile insurance, even though it would certainly help in locating the essential factors. On the other hand it should be clear that the claim data will be far from a multivariate normal distribution. This latter fact is largely due to the predominance of the factor bodily injury, which may lead to very large claim sums

Problem 9.4 Explain this latter statement in more detail. Compare the distribution of claims resulting from repair and material damage and from bodily injury.

As a result of all this each nation has chosen its own system of variables for premium rating. Increased traffic and commerce within Europe, however, makes it desirable to have a uniform rating system within Europe. Below we give a list of classification within various European countries.

France: A point system based on

1. Rating of vehicle (15 groups).
2. Geographical area.
3. Occupation of policy holder.
4. Duration since obtaining a drivers licence, age and sex.

The rating is thus based on a linear combination of these factors.

Great Britain:

1. Rating of vehicle.
2. Age of policy holder.
3. Geographical area.
4. no claim discount, bonus malus system.

Germany:

1. Power of engine and type of vehicle.
2. Geographical area.
3. Bonus malus system of no claim discount.
4. Special ratings for civil servants and farmers.

In the remainder we will outline the German system in more detail. As noted above it is based on the discretised variables:

1. Power (size of car) (Index $i = 1, \ldots, I$)

2. Geographical Area (Index $j = 1, \dots, J$)
3. Bonus, malus system (Index $\ell = 1, \dots, L$)

These variables are introduced separately for all kinds of vehicle classes like: trucks, motorcycles, passenger vehicles; taxis, ... Thus we shall stick to one of these classes, say private car to be specific, because this is the largest group anyway. Then let

$$\begin{aligned} N_{i,k,\ell} = &\text{number of policies in power class } i \\ &\text{area group } k \text{ and bonus class } \ell. \end{aligned}$$

Similarly let $M_{i,k,\ell}$ be the number of claims and $S_{ik\ell}$ the total claim sum per year for that group. Of course we assume that the claim sums are trend corrected. In Germany these data are collected from all automobile insurances by the HUK Verband and the statistics is available to all members of this association. Then

$$\hat{C}_{iK\ell} = \frac{S_{iK\ell}}{N_{iK\ell}} \tag{9.2.20}$$

is the average claim sum per policy. $\hat{C}_{ik\ell}$ is also an estimate for the netpremium in this class. Since the bonus-malus rating is introduced separately, we also need

$$\hat{C}_\ell = \frac{(\sum_{i,k} C_{ik\ell} \cdot N_{ik\ell})}{\sum_{i,k} N_{ik\ell}} = \frac{\sum_{i,k} S_{ik\ell}}{\sum_{i,k} N_{ik\ell}}. \tag{9.2.21}$$

For the computation of the trend and relative claim costs we also need

$$\mathcal{C} = \frac{\sum_{ik\ell} S_{ik\ell}}{\sum_{i,k\ell} N_{ik\ell}}. \tag{9.2.22}$$

The relative claims are then $\hat{C}_{ik\ell}/C = \hat{R}_{ik\ell}$. The net premium estimates $\hat{C}_{ik\ell}$ still show too much variation. So we need a method respectively a model too smoothen these data. Simon and Bailey (1960) have proposed a multiplicative model to achieve this. For this they write

$$R_{ik\ell} = X_i Y_k Z_\ell. \tag{9.2.23}$$

The variables X, Y, Z are determined by minimizing

$$\sum_{i,k,\ell} N_{ik\ell} \frac{(\hat{R}_{ik\ell} - X_i Y_k Z_\ell)^2}{(X_i Y_k Z_\ell)} = Q(X, Y, Z) \to \min. \tag{9.2.24}$$

This is done iteratively by using the initial values

$$X_i^{(0)} = \sum_{K,\ell} \frac{N_{ik\ell} \cdot \hat{R}_{ik\ell}}{\sum_{k\ell} N_{ik\ell}} \tag{9.2.25}$$

$$Y_k^{(0)} = \sum_{i,\ell} \frac{N_{ik\ell} \cdot \hat{R}_{ik\ell}}{\sum_{i,\ell} N_{ik\ell}}$$

$$Z_\ell^{(0)} = \sum_{ik} \frac{N_{ik\ell} \cdot \hat{R}_{ik\ell}}{\sum_{ik} N_{ik\ell}}$$

and the iteration

$$\begin{aligned} Q(X, Y^{(m)}, Z^{(m)}) &\to \min \qquad \text{to determine } X^{(m+1)} \\ Q(X^{(m+1)}, Y, Z^{(m)}) &\to \min \qquad \text{to determine } Y^{(m+1)} \\ Q(X^{(m+1)}, Y^{(m+1)}, Z) &\to \min \qquad \text{to determine } Z^{(m+1)} \end{aligned} \tag{9.2.26}$$

The net premium C_{ijk} is then

$$C_{ik\ell} = X_i Y_k Z_\ell \cdot \mathcal{C}. \tag{9.2.27}$$

Actually this is first averaged over ℓ, the bonus malus variable, which gives the so called 100% net premium. From this $C_{ik\ell}$ is computed by using a fixed set of bonus-malus factors.

In their 1979 paper Chang and Fairley propose instead an additive model for smoothing the data. Thus they write

$$R_{i,k,\ell} = \mu + X_i + Y_k + Z_\ell. \tag{9.2.28}$$

This model leads to problems related to analysis of variance. A more general approach to rate making in automobile insurance is presented in T. Mach 1997.

Problem 9.5 Write down explicitly the solution of the minimization problem (9.2.26)

$$Q(X, Y^{(m)}, Z^{(m)}) \to \min$$

in terms of R_{ikj}, $Y^{(m)}$, $Z^{(m)}$.

Literature:

R.A. Bailey, L.J. Simon: Two studies in automobile insurance rate making. Astin Bull. 1960

L. Chang, B. Fairley: Pricing Automobile Insurance under multivariate classification of risks. The Journal of risk and insurance 1979, p. 75–98.

J. Lemaire: Automobile Insurance. Kluwer Nijhoff Publ. 1985.

A.E. Maxwell: Multivariate Analysis in Behavioural Research. Chapman Hall 1977.

T. Mack: Schadensversicherungsmathematik. Schriftenreihe Angewandte Versicherungsmathematik Nr. 28

9.3 Claims, claim number and size of claims

As we have seen with fire insurance and automobile insurance, the standard actuarial method of studying the claims is to classify insurance policies according to their characteristics. This decomposes the portfolio of an insurance into distinct classes. In each such class the risks have a similar behaviour, i.e. they have the same risk characteristics. In order to describe the risk behaviour within each class, the actuary will use suitable models. These models are developed and tested usually on past data. The method of updating the models or parameters in these models runs under the name of credibility and will be discussed later. Modelling the data is a method to compress, understand, and explain the data and extend the conclusions into the future.

Once one has fixed a particular risk class, the important information is given by the number of claims per period and the claim sizes. Both terms are generally random variables. In the past, it has been customary to use certain standard distributions for these variables. These distributions depend on one or more parameters and are usually flexible enough to be applicable to a wide range of phenomena. But it may well be that in certain domains of non-life insurance, other distributions are more suitable. The choice and application of such models distributions depends to a great extent on the ability of the actuary. A good actuary will thus always find the right distributions, find efficient means to test the parameters and extend the results well into the future. But it should be clear that a probabilistic model is always a compromise between simplicity, computability and accuracy.

The number of claims

Let us begin with a bit of notation. For a random variable X, $P(X \in A)$ will denote the probability that X takes on values in $A \subset \mathbb{R}$. If $A = [a, b]$, we will also write $P(a \leq X \leq b)$ and interpret $P(a < X \leq b), \ldots$ similarly. Thus $P(X < a) = F(a)$ is the disribution function. The distribution of X will be denoted by P_X and if this distribution has a density, we write f_X for this density. Similarly, we will write $E(X)$ and σ_X^2 for the expectation value and variance of X, respectively.

Here we want to describe random variables which only have the values $0, 1, 2, \ldots$. Let X be such a random variable and let

$$p_k = P([X = k]) \qquad k = 0, 1, \ldots \tag{9.3.29}$$

be the probability that X has the value k. Then

$$m_l = \sum_{k=0}^{\infty} \quad p_k k^l \tag{9.3.30}$$

is the l-th moment. Clearly, $m_1 = E(X)$ and $m_2 - m_1^2 = \sigma_X^2$. A standard method to compute the moments uses the *generating function* M

$$M(z) = \sum_0^{\infty} p_k z^k. \tag{9.3.31}$$

M is called the generating function because

$$\frac{d^l M}{d_z^l}(1) = \sum_0^{\infty} p_k k(k-1)\ldots(k-l+1). \tag{9.3.32}$$

In particular, one has $m_1 = M'(1)$, $m_2 - m_1 = M''(1), \ldots$. Besides the generating function, the characteristic function

$$E(\exp(itX)) = \sum_{k=0}^{\infty} e^{itk} p_k = C_X(t) \tag{9.3.33}$$

is also used widely. $\mathcal{C}_X$ is thus the Fourier transform of the distribution of X. One of its key properties is

Problem 9.6 Let X and Y be independent random variables. Then

$$C_{X+Y}(t) = \mathcal{C}_X(t) \cdot \mathcal{C}_Y(t). \tag{9.3.34}$$

Remark. This property is also valid for arbitrarily distributed X, Y.
The real valued counterpart of the characteristic function $\mathcal{C}$ is the **moment generating function**.

$$L_x(t) = E(\exp tX) = \sum_{n=0}^{\infty} p_k e^{tk}. \tag{9.3.35}$$

This function may not always exist. If it does, we see

$$\frac{d^k}{dt} L_X(0) = m_k. \tag{9.3.36}$$

Problem 9.7 Prove (9.3.36).

Problem 9.8 Show $L_{X+Y}(t) = L_X(t) \cdot L_Y(t)$ if X and Y are independent.

L is thus the Laplace transform. The important properties of L and C are that they characterize the distributions uniquely. We will now discuss particular claim number distributions. As noted above, it is extremely useful for an actuary to know many distributions and to use them.

Binomial distribution, $\mathbf{B(n,p)}$

When a trial with the outcomes "0" and "1" is repeated n times, one finds k-times the occurrence of "1" with the probability

$$b(n,k,p) = \binom{n}{k} p^k (1-p)^{n-k} \quad 0 < p < 1, \quad 0 \le k \le n. \tag{9.3.37}$$

Here p is the probability of finding "1" in a single trial. This binomial law $B(n,p)$ is one of the best known laws in probability theory. Thus it is treated in all textbooks. One has

$$E(X) = np \qquad \sigma_X^2 = np(1-p) \tag{9.3.38}$$

and

$$M(z) = (1 - p + pz)^n. \tag{9.3.39}$$

The Poisson distribution $\mathbf{P}(\lambda)$

For $n \to \infty$ and $np \to \lambda > 0$ the binomial distribution converges weakly to the Poisson distribution. Its range is $0, 1, 2, \dots$ and

$$p_k = e^{-\lambda}\frac{\lambda^k}{k!} \qquad k = 0, 1, \dots . \tag{9.3.40}$$

Since $np \to \lambda$ for $n \to \infty$, one also says that the Poisson distribution describes **rare** events which occur with rate λ. One has

$$E(X) = \lambda, \quad \sigma_X^2 = \lambda \tag{9.3.41}$$

and

$$M(z) = \exp(\lambda(z-1)). \tag{9.3.42}$$

There is another method to derive the Poisson distribution which makes it very dear to many probabilists. We make the following assumption

1. The number of claims occurring in any two disjoint intervals in time are independent.
2. The number of claims occurring in $[t, t+\triangle]$ is for small $\triangle$ essentially proportional to $\triangle$, more precisely number of claims in this time interval is $\lambda\triangle + \sigma(\triangle)$. σ is the Landau symbol.
3. The chance that two or more claims occur is $\sigma(\triangle)$.

Then the number of claims in $[s, s+t]$ is k with probability

$$p_k(t) = e^{-\lambda t}(\lambda t)\frac{k}{k!} \tag{9.3.43}$$

Proof. Set $s = 0$ and let $p_k(t)$ be the probability that k claims occurred in $[0, t]$. Then

$$\begin{aligned} p_k(t+\triangle) = p_k(t)(1 - \triangle\lambda - \sigma(\triangle)) + p_{k-1}(t)(\triangle\lambda + \sigma(\triangle)) \\ + p_{k-2}(t)\sigma(\triangle) + \dots . \end{aligned}$$

Now divide by $\triangle$ and collect all $\sigma(\triangle)$ terms. One gets

$$\frac{p_k(t+\triangle) - p_k^{(t)}}{\triangle} = -\lambda p_k^{(t)} + \lambda p_{k-1}(t) + \sigma(\triangle)/\triangle[-p_k(t) + p_{k-1}(t) + p_{k-2}^{(t)} \dots]$$

for $\triangle \to 0$ we have: $\sigma(\triangle)/\triangle \to 0$ and the left-hand side converges towards p_k'. Thus

$$p_k'(t) = \lambda(p_{k-1} - p_k)(t). \tag{9.3.44}$$

With $P_0(0) = 1$ and $P_{-1} \equiv 0$ this equation has the unique solution given by (9.3.39). This can be shown by induction on k.

Remark. The assumptions show that the Poisson process, which we have studied above, is a pure birth process. It also shows that the *waiting time* for the first claim (event) to occur is exponentially distributed with parameter λ.

Remark. In practice, the assumptions above will not always be satisfied. The occurrence of one event, e.g. may trigger the next. In fire insurance, for example, a fire may spread to the next house. Fog may cause the collision of several cars. This leads to Poisson processes with contagion. Moreover, external factors may influence the occurrence of claims, i.e. the process is not homogeneous and λ may vary with t - inhomogeneous Poisson, processes. e.g. in a dry summer, fires may occur more readily.

Remark. The "success" and wide application of the Poisson distribution rests essentially with its mathematical simplicity and the following that result.

Theorem 9.2 Let $X_1, X_2, \ldots, X_k$ be independent Poisson distributed random variables with parameters $\lambda_1, \ldots, \lambda_k$. Then $Y = X_1 + \ldots + X_k$ is Poisson distributed with parameter $\lambda = \lambda_1 + \ldots + \lambda_k$.

Problem 9.9 Prove this theorem directly by induction on k, which essentially reduces the problem to $k = 2$.

Problem 9.10 Prove the theorem by means of the moment generating function or characteristic function.

Remark. The disadvantage of the Poisson distribution in application lies with the fact that the P_k decay too rapidly for large k. It depends only on 1 parameter. Thus a fit of data may be difficult.

The maximum of the Poisson distribution lies close to λ, the expectation value. This follows easily from the recursion relation

$$p_{k+1} = \frac{\lambda}{k+1} p_k \qquad p_0 = e^{-\lambda}. \tag{9.3.45}$$

Problem 9.11 Use (9.3.45) to show that the Poisson distribution has its maximum near λ.

Nonetheless it is sometimes useful to have a good approximative formula for the p_k, particularly for large λ. This is e.g. the Anscombe approximation

$$p_k = \phi\left[\frac{3}{2}\left(k + \frac{5}{8}\right)^{2/3} \lambda^{-1/6} - \frac{3}{2}\sqrt{\lambda} + \frac{1}{24\sqrt{\lambda}}\right] \tag{9.3.46}$$

Here $\emptyset$ is the distribution function of the standard normal distribution.

Negative Binomial distribution $\mathbf{NB(p, r)}$

The negative Binomial distribution has many applications in practical problems, because it depends on two parameters and has a longer tail, i.e. it decreases more slowly. It is used for example to model the days in hospitals in health insurance or to model the number of purchases of particular goods. It can be obtained from a continuous superposition of Poisson variables with a Gamma density. For the $NB(p, r)$, $0 < p < 1$, $r > 0$ we have

$$p_k = \binom{k+r-1}{k} p^r (1-p)^k \quad k = 0, 1, 2, \ldots . \tag{9.3.47}$$

For real r the symbol $\binom{k+r-1}{K}$ has to be expressed with the aid of the Γ-function, $\Gamma(x+1) = x!$. But we can also write for any $r \geq 0$ $\binom{r}{k} - r(r-1)\ldots(r-k+1)/k!$. For integer r p_k is the probability for k additional trials to achieve r successes in an independently repeated Bernoulli experiment $B(1, p)$ with parameter p. For an NB distributed random variable one has

$$EX = \frac{r(1-p)}{p} \quad \sigma_X^2 = \frac{r(1-p)}{p^2} \tag{9.3.48}$$

and

$$M(z) = \left(\frac{p}{(1-(1-p)z)}\right)^r . \tag{9.3.49}$$

For $r = 1$ one obtains the geometric distribution.

Logarithmic distribution

The logarithmic distribution has a long tail. It has only one parameter p, $0 < p < 1$. Its use is thus rather restricted. The density is given by

$$p_k = \frac{p^k}{k} \frac{1}{(-\ln(1-p))}. \tag{9.3.50}$$

One finds

$$EX = \frac{p}{((1-p)\tau)}, \; \sigma_X^2 = \frac{p(\tau - p)}{((1-p)^2\tau^2)} \qquad \tau = -\ln(1-p). \tag{9.3.51}$$

Generalized Poisson distribution $\mathbf{GP}(\lambda, \theta), \theta > \mathbf{0}, \mathbf{0} \leq \lambda < \mathbf{1}$

The defects of the Poisson distribution are partly remedied with the generalized Poisson distribution, which is defined by

$$p_K = \frac{1}{k!}(\theta + k\lambda)^{k-1} \exp(-\theta - k\lambda) \quad k = 0, 1, \ldots \tag{9.3.52}$$

with

$$E(X) = \frac{\theta}{1-\lambda} \quad \sigma_X^2 = \frac{\theta}{(1-\lambda)^3}. \tag{9.3.53}$$

As a final example we will consider the mixed Poisson distribution. At the same time this example shows us how to obtain larger classes of claim number distributions, by mixing them.

Mixed Poisson distribution $\mathbf{G_\pi P}(\lambda)$

Let $\pi(\lambda)$ be an arbitrary distribution of the parameter λ on $(0, \infty)$. Then define

$$p_k = \int_0^\infty \frac{\lambda^k}{k!} e^{-\lambda} d\pi(\lambda) \qquad k = 0, 1, 2, \dots . \tag{9.3.54}$$

This distribution is thus a continuous or discrete superposition of Poisson variables. We have noted above already that the negative binomial distribution is of this type with π a Gamma distribution. Superpositions of 2 Poisson variables have been used previously in automobile claim numbers. In this case one finds

$$E(X) = \mu = \int_0^\infty \lambda d\pi(\lambda), \quad \sigma_X^2 = \mu + \int_0^\infty (\lambda - \mu)^2 d\pi(\lambda) \tag{9.3.55}$$

and

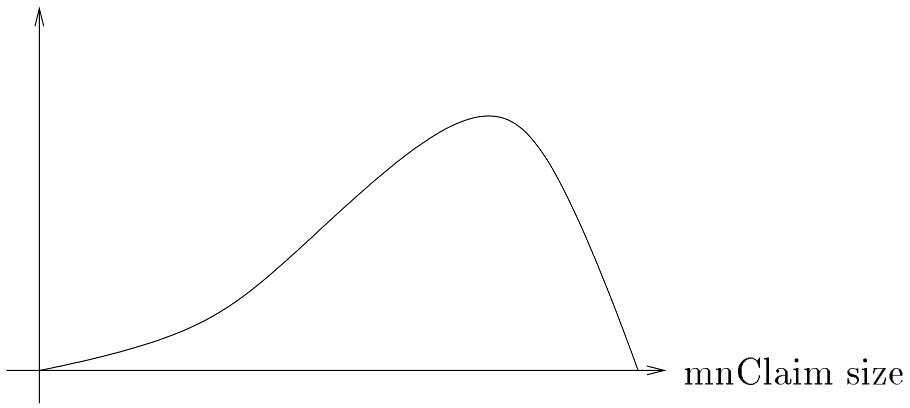
$$M(z) = \int_0^\infty \exp(-\lambda(z-1)) d\pi(\lambda).$$

Problem 9.12 Prove (9.3.55).

Particular mixed Poisson distributions are the Neyman Type A or the Delaporte distributions. Details about those and other lesser known distributions can be found in the monograph. Distributions in Statistics: Discrete distributions by N.L. Johnson and S. Kotz Wiley 1969.

The size of claims

Besides the number of claims, one wants to knwo something about their sizes too. The size distribtution generally has the form



Problem 9.13 What type of claim size distributions do you expect. Give examples and arguments.

So we will have to find distributions, which can approximate such a shape and are easy enough to handle mathematically. It is custumary in this area to use contiunous distributions only.

The uniform distribution $\mathbf{U(a, b)}$

The density of the uniform distribution is given by

$$f(x) = \frac{1}{b-a}, \quad a \le x \le b. \tag{9.3.56}$$

Its mean value is $1/2(a+b)$ and its variance $(b^2 - a^2)/12$. The uniform distribution model is the simplest model and used only for small claims or quotas. The model is not very flexible either.

Problem 9.14 Determine the moment generating function in this case.

One of the most useful distribution for claim sizes but also for the description of life times, times of latency, ... is the Gamma distribution.

The Gamma distribution $\Gamma(\alpha, \beta)$

$\alpha, \beta > 0$. Its density is given by

$$f(x) = \frac{\alpha\beta}{\Gamma(\beta)} x^{\beta-1} \exp(-\alpha x) \quad x > 0 \tag{9.3.57}$$

where $\Gamma(\beta) = \int_0^\infty x^{\beta-1} e^{-x} dx$ is the Gamma function, which gives the distribtuion its name. Its expectation value and variance are

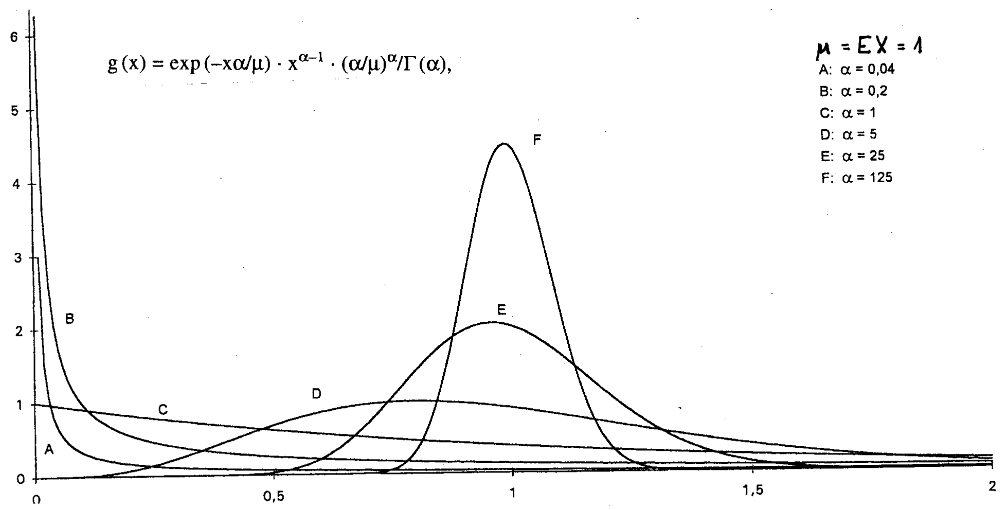
$$E(X) = \frac{\beta}{\alpha}, \qquad \sigma_X^2 = \frac{\beta}{\alpha^2}. \tag{9.3.58}$$

The characteristic function

$$C(t) = \left(\frac{\alpha}{\alpha - it}\right)^\beta. \tag{9.3.59}$$

Problem 9.15 Prove (9.3.59) and derive from this that the distribution of the sum of two independent $\Gamma(\alpha, \beta_1)$ and $\Gamma(\alpha, \beta_2)$ distributed random variables is $\Gamma d(\alpha, \beta_1 + \beta_2)$ distributed. Generalize this to k summands.

Here are some plots of the density of the Γ-distribution with $\alpha = \beta$.



The Gamma distribution is widely used, mainly to describe small claims. The property described in Problem 9.15 makes it particularly useful to describe sums of claims. For $\beta = 1$ one gets the exponential distribution, which we had already encountered as the waiting time distribution for the Poisson process.

In fire insurance we have the situation that the risks X_i $i = 1, \dots, n$ satisfy $E(X_i) = \mu v_i$ and $\sigma^2_{X_i} = \sigma^2 v_i$ with v_i the volume (face value) of the risk X_i. From the shape of the density it is plausible that the Gamma distribution might be a suitable model for X_i. The corresponding parameters are then

$$\frac{\beta_i}{\alpha_i} = \mu v_i \quad \text{and} \quad v_i\sigma^2 = \frac{\beta_i}{\alpha_i^2}.$$

This gives $v_i\sigma^2 = \mu v_i/\alpha_i$ or $\alpha_1 = \mu/\sigma^2 = \alpha$ and $\beta_i = \mu^2 v_i/\sigma^2$. Thus the sum $X_1 + \cdots + X_n = S$ has a Gamma distribution with parameter $\alpha = \mu/\sigma^2$ and $\beta = \mu^2/\sigma^2 V$ where $V = \sum v_i$. Even though the individual distributions are mainly concentrated near 0, the sum distribution will look much like a normal distribution, if n is not too small.

These considerations also show that we may add the risks of different portefolio classes along the same lines, as long as σ^2/μ is approximately constant. Hence σ^2/μ should be used in fire insurance to classify portfolios.

Problem 9.16 What is the waiting time distribution for k events with a Poisson process?

It is a particular Gamma distribution, also known as Erlang distribution. Erlang used it to describe incoming calls in telephone exchanges. A related distribution is the

Weibull distribution $\mathbf{W}(\alpha, \beta)$ $\quad \alpha, \beta > \mathbf{0}$

It density is given by

$$f(x) = \alpha\beta x^{\beta-1}\exp(-\alpha x^\beta) \qquad x > 0. \tag{9.3.60}$$

This distribution was introduced by the Swedish physicist W. Weibull when studying forces needed to break certain materials. It arises as a lifetime distribution if the force (intensity) of mortality grows like x^β, with x-the age. Besides describing lifetimes $W(\alpha, \beta)$ is also used to model reliability. One finds

$$EX = \alpha^{-1/\beta}\Gamma\left(1 + \frac{1}{\beta}\right), \quad \sigma_X^2 = \alpha^{-2/\beta}\left\{\Gamma\left(1 + \frac{1}{\beta}\right) - \Gamma\left(1 + \frac{1}{\beta}\right)^2\right\}. \tag{9.3.61}$$

As a claimsize distribution one uses $W(\alpha, \beta)$ with $\beta \geq 1$ for small and $\beta < 1$ for large claims.

Problem 9.17 Show that if X is exponentially distributed with parameter α, $X^{1/\beta}$ is $W(\alpha, \beta)$ distributed.

Lognormal distribution $\mathbf{LN}(\mu, \sigma)$ $\quad \sigma > 0$

Even though the normal distribution played such a dominant role previously, it is hardly used in claim size distribution. Mainly because of its symmetry. A variant of the normal distribution the Lognormal distribution, however, has applications in areas like distribution of the size of fragments, income distribution, fire insurance, automobile insurance and reinsurance. A random variable X is lognormal distributed, if it has nonnegative values and if $\ln X$ is normally distributed. Thus there is also a type of multiplicative central limit theorem. Its density is given by

$$f(x) = (2\pi\sigma^2)^{-1/2} x^{-1} \exp\left(-\frac{1}{2\sigma^2}(\ln x - \mu)^2\right) \quad x > 0. \tag{9.3.62}$$

Moreover

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{x}\right), \quad \sigma_X^2 = ((\exp\sigma^2) - 1)\exp(2\mu + \sigma^2). \tag{9.3.63}$$

Pareto distribution $\mathbf{P(a, b)}$

The Pareto distribution is named after the italian mathematician and economist W. Pareto, who found it when he studied the income distribution within an economy. It is widely used in fire insurance for wooden houses and for large claims, because it has a long tail. Its density is given by

$$f(x) = \frac{b}{a}\left(\frac{a}{a+x}\right)^{b+1} \quad \text{for } x > 0. \tag{9.3.64}$$

Since it has such a long tail, its moments m_n will only exist for $n < b$

$$m_n = \frac{a^n n!}{\prod_i^n (b-i)}. \tag{9.3.65}$$

A distribution with very nice properties, which unfortunately has not been widely used in insurance is the inverse Gauss distribution.

Inverse Gauss distribution $\mathbf{IG}(\lambda, \mu)$ $\quad \lambda, \mu > 0$

$$f(x) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left(-\frac{\lambda}{2\mu^2 x}(x - \mu)\right) \quad x > 0. \tag{9.3.66}$$

For this distribution all moments exist and

$$EX = \mu \quad \sigma_X^2 = \frac{\mu_3}{\lambda}. \tag{9.3.67}$$

A nice feature of the IG distribution is the fact that the sum of independent IG distributed random variables is again IG distributed. More precisely we have

Lemma 9.2 1. Let $X \sim IG(\lambda, \mu)$ then $cX \sim IG(c\lambda, c\mu)$.

2. $\theta = \lambda/\mu$ and $X \sim IG(\lambda, \mu)$ then $\lambda X/\mu^2 \sim IG(\theta^2, \theta)$
3. Let $X_i \sim IG(\xi\mu_i^2, \mu_i)\, i = 1, \dots, n,\ \xi > 0$ be independent then $X_1 + \cdots + X_n \sim IG(\xi\mu^2, \mu)\ \mu = \sum \mu_i$.

The proof of (1.) is trivial, while (2.) is a consequence of (1.). (3.) can be reduced to $n = 2$ by induction. This result has to be shown by means of the moment generating function.

As a final example we will consider the Beta distribution. It is concentrated in $[0, 1]$ and thus particularly useful in describing claim ratios. In addition it is the conjugate distribution of the Binomial distribution. Hence it is used in Bayesion Statistics.

Beta distribution

Its density is given by

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \quad a, b > 0 \quad 0 < x < 1 \tag{9.3.68}$$

while

$$E(X) = \frac{a}{a+b}, \quad \sigma_X^2 = \frac{ab}{(a+b+1)(a+b)^2}. \tag{9.3.69}$$

As you might have guessed this list of distributions is by far not complete, but I think these are the most important ones. If you should work is nonlife insurance, it might be necessary to study some more. Other, possibly important distributions may be found in the book by Johnson and Kotz "Distributions in Statistics I, II", Wiley 1970. Here you can also find important aspects about parameter estimators and additional properties. It is particularly important to know the estimators, because in practical problems, one will have to determine these parameters.

Problem 9.18 Panjer has discussed claim number distributions, where the weights p_k satisfy n rekursion relation of the form

$$p_k = \frac{(a + \frac{b}{k})}{p_{k-1}}.$$

Which of the above claim number distributions satisfy such relations?

Problem 9.19 The third normalized central moment $E((X - EX)^3) \cdot \sigma^{-3}$ of a random variable is called the *skewness*. Compute the skewness for the Poisson and normal distribution.

The skewness is an important parameter of a random variable to measure the deviation from the normal law.

Problem 9.20 What does the sign of the skewness of a random variable tell you? Sketch the density of a distribution with positive (negative) skewness.

Problem 9.21 Show that

$$f(x) = \frac{(\frac{\alpha}{\mu})^\alpha}{\Gamma(\alpha)} \cdot x^{\alpha-1} \exp\left(-\frac{x\alpha}{\mu}\right)$$

gives another parametrization of the Γ-distribution with μ the expectation value and μ^2/α as variance as parameters.

9.4 Transformations of distributions

The simplest method to generate new distributions from old ones into apply monotone differentiable transformations g.

Lemma 9.3 Let Y be a random variable with range $A \subset \mathbb{R}$ and continuous density of its distribution f_Y. Let g be differentiable and strictly monotone on A. Then the random variable $X = g(Y)$ has a distribution with density f_X

$$f_X(x) = \frac{f_Y(g^{-1}x)}{g'(g^{-1}(x))} \quad x \in g(A). \tag{9.4.70}$$

Problem 9.22 Prove the Lemma by considering the distribution functions $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$ and the fact that $F_X' = f_X$.

The simplest type of transformations are

1. Translation: $g(x) = x + c$
2. Scaling transformation: $g(x) = a \cdot x$.
   The combination of translation and scaling transformation gives the class of
3. affine linear transformations $g(x) = ax + b$

Problem 9.23 Show that the general normal distribution $N(\mu, \sigma)$ can be obtained from the standard normal distribution $N(0, 1)$ by an affine lineare transformation.

Problem 9.24 Let Y be as in the Lemma and let y be an affine linear transformation. Determine EX and σ_X^2 interms of $\mu = EY$ and $\sigma^2 = \sigma_Y^2$.

Problem 9.25 In a class of distributions $\mathcal{C}$ depending on the parameters $\alpha, \beta, \ldots$. α is called a *scaling parameter*, if the action of $g(x) = ax$ leaves $\mathcal{C}$ invariant and acts on the parameter α only, $\alpha \to a\alpha$. In which of the claim size distributions in the previous section can you find scaling parameters?

Problem 9.26 Assume Y has a continuous distribution f_Y. Determine the distribution of

1.) $X = \ln Y$ if $A = \mathbb{R}_+$
2.) $X = e^Y$ if $A = \mathbb{R}$
3.) $X = Y^{1/c}$ if $A = \mathbb{R}_+$ and $c > 0$
4.) In what cases of the distributions in the previous section have such distributions be used.

Besides these transformations there are other forms, which are interesting in connection with insurance problems.

Benefit Limitation

Let X be the claim random variable and assume the benefits are limited by L. Then the benefit random variable $Y = \min(X, L)$ has the density $f(x)$ for $x \leq L$ and $\int_L^\infty f(x)dx$ as a point mass in L.

Deductible M

Only the claims in excess of M are honored. The benefits are $Y = \max(X - M, 0)$. In this case the density of the benefits Y is $f(x + M)$ for $x > 0$ but there is a point mass $\int_0^M f(x)dx$ in 0.

Premium refund

There is a refund π if the claims do not exceed M otherwise only the claims exceeding M are honored. In this case the benefits are

$$Y = \begin{cases} \pi & X \leq M \\ X - M & X \geq M \end{cases}$$

This given the density $f(x + M)$ in $x \geq 0$ but a point mass $\int^M f(x)dx$ in π.

Another method to modify distributions is *truncation*. In its simplest form one chooses a subset $A \subset \mathbb{R}$. Then the random variable X_A restricted to X has the distribution

$$P_{X_A}(B) = \frac{P_X(A \cap B)}{P_X(A)}. \tag{9.4.71}$$

This method applies to continuous as well as to discret distributions, e.g.

$$f(x) = 2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}$$

is the density of a normal random variable truncated to $\mathbb{R}_+$.

Problem 9.27 Determine $\ln Y$ of an exponentially distributed random variable.

It should be clear that the above list of distributions and transformations is by no means exhaustive. More can be found in the above cited books by Johnson and Kotz or publications in actuarial journals.

9.5 Distributions of Claim sums

Assume we are given the claim number distribution N and the distribution of individual claims $X_1, \dots, X_N$. Then the total claim sum is

$$S = \sum_{i=1}^{N} X_i. \tag{9.5.72}$$

One calls a sum like (9.5.72) a *random sum*, because the range of the summation is random. In order to compute the distribution of S, the *sum of claims* the following assumptions are natural and usually made

$$\begin{aligned} N, X_1, \dots, X_N &\text{ are independent random variables} \\ \text{the } X_1, \dots, X_n &\text{ are identically distributed} \end{aligned} \tag{9.5.73}$$

If X and Y are independent random variables with densities f_X and F_Y then $X+Y$ has the density

$$g(x) = (f_X * f_Y)(x) = \int f_X(y) f_Y(x-y) dy. \tag{9.5.74}$$

This g is called the convolution of f_X and f_Y. This shows that the distribution of $X_1 + \cdots + X_n$ is given by

$$f * f * \cdots * f = f^{*n} \tag{9.5.75}$$

the n-fold convolution product. Here we have of course assumed that each X_i has the density f with its distribution. In f^{*n} one will have to interpret f^{*0} as δ_0, the point mass 1 in 0. With this and $p_k =$ Prob $(N = k)$ we see that the distribution of S from (9.5.72) is given by

$$\text{distr. } S = p_0 \cdot \delta_0 + \sum_{K=1}^{\infty} f^{*k}(x) = P_S. \tag{9.5.76}$$

The first summand reflects the fact that with probability p_0 no claim arises, while the claim density of the rest is given by this infinite convolution sum. Thus mathematical convenience and computability requires to use only densities f for which f^{*n} is computable or a least well approximable. In practical problems it will also be necessary to truncate the infinite sum and estimate the corresponding error.

Let us first determine some elementary important parameters

$$E(S) = \sum_{K=0}^{\infty} \text{Prob } (N = k) k\mu \quad \text{where } \mu = EX, \sigma^2 = \sigma_X^2$$

or

$$E(S) = E(N)E(X). \tag{9.5.77}$$

This then gives

$$\begin{aligned} E(S^2) &= \sum_{K=1}^{\infty} E(S^2 \mid N = k) \text{ Prob } (N = k) = \sum_{K=1}^{\infty} p_K [k\sigma^2 + k^2\mu^2] \\ &= E(N)\sigma^2 + \mu^2 E(N^2) = E(N)\sigma^2 + \mu^2\sigma_N^2 + \mu^2 E(N)^2. \end{aligned}$$

Thus

$$\sigma_S^2 = E(N)\sigma_X^2 + E(X)^2\sigma_N^2. \tag{9.5.78}$$

Lemma 9.4 Let X, Y be independent random variables with continuous (discrete) densities f_X, $f_Y(p_{X,k}, p_{Y,k})$ then their characteristic functions $\mathcal{C}_X, \ldots$ respectively moment generating functions L satisfy

$$\mathcal{C}_{X+Y}(t) = \mathcal{C}_X(t)\mathcal{C}_Y(t), \ L_{X+Y}(t) = L_X(t)L_y(t). \tag{9.5.79}$$

Proof. We have $\mathcal{C}_X(t) = E(e^{itX})$ and $L_X(t) = E(e^{tX})$. Now the result follows from a general property of independent random variables.

$$E(f(X)g(Y)) = E(f(X))E(g(Y)) \tag{9.5.80}$$

and the fact that

$$e^{it(X+Y)} = e^{itX} \cdot e^{itY}.$$

The proof for L is similar, if L_X and L_Y exist.

Problem 9.28 Prove (9.5.80) by approximating f and g by functions of the type

$$\sum_{i=1}^{n} a_i \chi_{A_i} \quad \chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

and thus reduce it to $f = \chi_A$ and $g = \chi_B$.
Now let S be a claim sum as above. Then the characterisitc function $\mathcal{C}_S$ of S is

$$\mathcal{C}(S) = E(e^{its}) = \sum_{k=0}^{\infty} p_k E(e^{it(X_1+\cdots+X_k)}) = \sum_{k=0}^{\infty} p_k \mathcal{C}_X^k(t).$$

Thus

$$\mathcal{C}(S) = M_N(\mathcal{C}_X(t)) \tag{9.5.81}$$

where $M_N(x) = \sum p_k X^k$ is the generating function of N.

Problem 9.29 Derive a similar result for the moment generating function L, if it exists.

Example 9.1 Assume N is $P(\mu)$ distributed-Poisson with parameter μ and assume X is exponentially distributed with density $\lambda e^{-\lambda x}$ on $[0,\infty)$. Then $X_1+\cdots+X_k$ is Gamma distributed with density $\lambda^k\ x^{k-1}/(k-1)!\ e^{-\lambda x}$. Thus the total claim distribution S has a density

$$e^{-\mu}\cdot\delta_0+\sum_{k=1}^{\infty}e^{-\mu}\frac{\mu^k}{k!}\lambda^k\frac{x^{k-1}}{(k-1)!}e^{-\lambda x}.$$

Now the generating function M of the Poisson distribution is

$$M(x)=\sum_{0}^{\infty}e^{-\mu}\frac{\mu^n}{n!}x^n=e^{-\mu+\mu x}$$

while the moment generating function of the exponential distribution is for $|t|<\lambda$

$$L(t)=\int_0^\infty\lambda e^{tx}e^{-\lambda x}dx=\frac{\lambda}{\lambda-t}.$$

Thus $L_S(t)=\exp(-\mu(1-\lambda/\lambda-t))$.

Example 9.2 Assume that N is geometrically distributed $p_k=(1-p)p^k$, $0<p<1$ and assume X is exponentially distributed as in Example 1. Then the total claim distribution has the density

$$(1-p)\delta_0+\sum_{k=1}^{\infty}(1-p)p^k\lambda^k\frac{X^{k-1}}{(k-1)!}e^{-\lambda x}=(1-p)\delta_0+p(1-p)\lambda e^{-\lambda x}\sum_{K=0}^{\infty}(p\lambda X)\frac{k}{k!}$$
$$=(1-p)\delta_0+p(1-p)\lambda e^{-\lambda x+p\lambda x}=(1-p)\delta_0+p(1-p)\lambda e^{-\lambda(1-p)x}.$$

This is a mixture of a discrete distribution in 0 and an exponential distribution with parameter $(1-p)\lambda$.

We had defined the moment generating function L_X of X by $L_X(t)=E(e^{tX})$, if this expression exists. Now let

$$\varphi_X(t)=\ln(L_X(t)) \tag{9.5.82}$$

be the logarithm generating function.

Problem 9.30 Show for independent random variables X, Y

$$\varphi_{X+Y}=\varphi_X+\varphi_Y.$$

Let N be a claim number and X be a claim size distribution with Prob $(N=k)=p_k$. Then the claim sum distribution S satisfies

$$L_S(t)=E(e^{tS})=\sum p_kL_X(t)^K=\sum p_ke^{k\cdot\varphi_X(t)}=L_N(\varphi_X(t))$$

or

$$\varphi_S(t)=\varphi_N(\varphi_x(t)). \tag{9.5.83}$$

Problem 9.31 Express $E(X)$ and σ_X^2 in terms of derivatives of φ and derive the expression for $E(S)$ and σ_S^2.

Even though the Poisson distribution is generally liked in insurance mathematics, because of its simplicity, it is generally not a very realistic model for a claim number distribution. On the other hand the negative Binomial distribution has much better properties. For this reason the following result of Ammeter is important.

Theorem 9.3 *(Ammeter Transformation)* Let $0 < p < 1$ and let r be fixed. Assume N is $NB(p,r)$ distributed. Then

$$\sum_{k=0}^{\infty} \binom{r+k-1}{k} (1-p)^r p^k f^{*k} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} g^{*k}, \quad \lambda = r|\ln p| \tag{9.5.84}$$

where $g = (-\ln(1-p))^{-1} \sum_{n=0}^{\infty} \frac{p^n}{n} f^{*n}$

In the remainder we will use the notation $P(\lambda, X)$ or $NB(p,r,X)$ for the claim sum distribution if N is $P(\lambda)$ or $NB(p,r)$ distributed and where X gives the distribution of single claims. We may also replace X by the symbol for its distribution.

The following result gives us the claim sum distribution of a whole portfolio, if each subclass has a Poisson sum distribution.

Theorem 9.4 Let $S_i = P(\lambda_i, X_i)$ be independent Poisson claim sum distributions. Then $S = S_1 + \ldots + S_n$ has a $P(\lambda, X)$ distribution where $\lambda = \lambda_1 + \cdots + \lambda_n$ and $X = \lambda_1 X_1 + \ldots \lambda_n X_n$.

Proof. The generating function of $P(\lambda)$ is $m(z) = \sum_0^\infty e^{-\lambda}\lambda^k/k! z^k = e^{\lambda(z-1)}$. Thus if ψ_i is the characteristic function of X_i, the characteristic function $\mathcal{C}_i$ of S_i and $\mathcal{C}$ of S is

$$\begin{aligned} \mathcal{C}_i(t) &= \exp(\lambda_i(\psi_i - 1)) \\ \mathcal{C}(t) &= \exp[\lambda_1(\psi_1 - 1) + \cdots + \lambda_n(\psi_n - 1)] = \exp[\lambda(\psi - 1)] \end{aligned}$$

because $\psi(t) = 1/\lambda \sum \lambda_i \psi_i$ is the characteristic function of X.

With similar methods we can prove a form of distributive law for claim size distribution.

Theorem 9.5 Let $N_1, \ldots, N_K$ be independent claim number distributions and let X be a fixed claim size distribution. Assume S_i is $N_i(X)$ distributed. Let N be the sum distribution of $N_1 + \cdots + N_k$. Then $S = S_1 + \cdots + S_k$ has the distribution $N(X)$.

Proof. Let m_i be the generating function of N_i and ψ the characteristic function of X. Then the characteristic function of $N_i(X)$ is $m_i(\psi(t)0$. Thus the characteristic function of $S_1 + \cdots + S_k$ is $m_1(\psi(t)) \cdot \cdots \cdot m_k(\psi(t)) = (m_1 \cdot \cdots \cdot m_k)(\psi(t))$ and the result follows since $m_1(z) \ldots m_k(z) = m(z)$ is the claim number distribution of $N = N_1 + \cdots + N_k$.

Problem 9.32 Let N_1 and N_2 be independent claim number distributions with generating functions m_1 and m_2. Then $N_1 + N_2$ has the generating function $m_1 \cdot m_2$. (Hint.: Use $m_i(z) = E(z^{N_i})$.

Even though this theory of claim sum distributions looks rather elegant, two basic problems remain.

1. Explicit computation of the claim sum distributions. This can be done rarely in closed form. Thus one has to resort to numerical methods or approximations. Numerically the claim sums can either be computed directly or by fast Fourier transform.

2. Estimating the parameters and checking the goodness of fit.

Both these problems have been treated extensively in insurance mathematics literature. However, since these are rather specialized topics we will go into it any further.

9.6 The individual model of risk theory

In the study of claims arising within a portfolio or several portfolios one distinguishes the so called *individual model* and *collective model*. In the individual model one keeps track which individual policy causes a claim. While this approach seem to be very appealing, it causes extensive numerical and approximation problems. In the collective model which we have looked at in section 9.4, one just considers the sum of all claims of a portfolio without regard to the origin of these claims. It is clear that collective models can be used to approximate individual ones. We will now study the individual model.

Let n be the number of policies and let $Y_i, i = 1, \ldots, n$ be the loss of policy i during the period under consideration. Its distribution function and density will be denoted by F_i and f_i respectively. Finally let q_i be the probability that a claim arises from the i-th policy. Then let

$$I_i = \chi_{(0,\infty)}(Y_i) = \begin{cases} 1 & \text{if} \quad Y_i > 0 \\ 0 & \text{if} \quad Y_i = 0. \end{cases} \tag{9.6.85}$$

Then I_i is $B(1, q_i)$ distributed. Finally let X_i $i = 1, \ldots, n$ be the positive losses of the i-th policy, if these occur. Denote by H_i the distribution function of X_i. Then

Problem 9.33 Show $H_i(x) = P(X_i \leq x) = (F_i(x) - (1 - q_i))/q_i$.

Moreover

$$Y_i = I_i \cdot X_i. \tag{9.6.86}$$

It follows from (9.5.75) and (9.5.76) that

$$E(Y_i) = q_i E(X_i), \quad \sigma^2_{Y_i} = q_i E(X_i^2) - q_i^2 E(X_i)^2. \tag{9.6.87}$$

As usual we now make the assumption that all $Y_1, \ldots, Y_n$ are independent. Then also $I_1, \ldots, I_n$ are independent, while $X_1, \ldots, X_n$ are conditionally independent if $I_i = 1$, $i = 1, \ldots, n$. This shows for

$$S = \sum_1^n Y_i \tag{9.6.88}$$

$$E(S) = \sum \mu_i q_i,\ \sigma_S^2 = \sum q_i E(X_i^2) - \sum q_i^2 \mu_i \quad \mu_i = EX_i. \tag{9.6.89}$$

The distribution of S is then the finite convolution product of

$$f_S = ((1-q_1)\delta_0 + q_1 f_1) * \cdots * ((1-q_n)\delta_0 + q_n f_n). \tag{9.6.90}$$

Since this is too complicated in most cases one approximates the distribution of S in the following fashion. One mixes the variables $X_1, \ldots, X_n$ by defining the distribution function

$$\overline{F} = \frac{1}{n} \sum F_i. \tag{9.6.91}$$

This is the distribution of a random variable $\overline{Y}$. Now replace $Y_1, \ldots, Y_n$ by independent copies of $\overline{Y}$. Thus one sets $Y_1 = \overline{Y}_1$, $Y_2 = \overline{Y}_2, \ldots$ where $\overline{Y}_i$ has the same distribution as $\overline{Y}$ and of course all $\overline{Y}_i$, are assumed to be independent. This makes sense, because the individual claim distributions are hardly observable anyway. Equivalently we can define the distribution function H of a random variable $\overline{X}$ by

$$H = \frac{1}{nq} \sum q_i H_i \quad \text{with } q = \frac{1}{n} \sum q_i. \tag{9.6.92}$$

Then

$$E(\overline{X}) = \int x dH(x) = \frac{1}{nq} \sum q_i \int x dH_i(x) = \frac{1}{nq} \sum q_i E(X_i).$$

In order to compare the individual model with the averaged distribution $\overline{Y}$ respectively $\overline{X}$ consider now the following collective model:

1. Let $X_1, \ldots, X_n$ be a set of independent identically distributed random variables with distribution function H.

2. Let N be a claim number distribution, independent of the X_i with $E(W) = nq$.

Now consider

$$S_{co} = \sum_0^N X_i \qquad co = \text{collective model}. \tag{9.6.93}$$

The condition on $E(N)$ is necessary, because we want

$$E(S_{co}) = E(N) \cdot E(\overline{X}) = \frac{E(N)}{nq} \sum q_i E(X_i) = E(S_{ind}) \tag{9.6.94}$$

$$S_{ind} = \text{ sum variable according to the individual model}$$

In order to compare both models, let us look at the variances.

$$\begin{aligned}
\sigma_{co}^2 &= E(N)\sigma_{\overline{X}}^2 + E(\overline{X})^2\sigma_N^2 \\
&= E(N)(E(\overline{X}^2) - (E\overline{X})^2) + E(\overline{X})^2\sigma_N^2 \\
&= E(N)(E(\overline{X}^2) + (\sigma_N^2 - E(N))E(\overline{X})^2 \\
&= E(N)\frac{1}{nq}\sum E(X_i^2)q_i + (\sigma_N^2 - E(N))\left(\frac{1}{nq}\sum q_i E(X_i)\right)^2 \\
&= \sum q_i E(X_i) + \frac{(\sigma_N^2 - E(N))}{E(N)^2}\left(\sum q_i\mu_i\right)^2 \\
&= \sigma_{S_{ind}}^2 + \sum q_i^2\mu_i^2 + \frac{(\sigma_N^2 - E(N))}{E(N)^2}\left(\sum q_i\mu_i\right)^2.
\end{aligned}$$

The idea now is to look for claim number distributions for which

$$\sigma_{S_{co}}^2 - \sigma_{S_{in}}^2$$

is minimal or at least rather small, i.e.

$$\sigma_{co}^2 - \sigma_{in}^2 = \sum E(Y_i)^2 + (\sigma_N^2 - nq)\cdot\frac{\sum(EY_i)^2}{(nq)^2} \to \text{ small.} \tag{9.6.95}$$

A good guess for N would be the binomial distribution $B(n, q)$. In this case the difference is

$$\sum E(Y_i)^2 \cdot \left(1 - \frac{1}{n}\right).$$

The Poisson distribution gives $\sum E(Y_i)^2$, because the second summand vanishes.

The above discussion was based essentially on the first two moments of S_{co} and S_{ind}. Now we will use the characteristic function in order to derive an approximation for the individual model. We had seen above that the distribution of Y_j is given by $\delta_0(1-q_j) + q_j f_j$. Thus its characteristic function $\mathcal{C}_i$ is given by

$$\mathcal{C}_i(t) = E(e^{it}X_j) = (1-q_j) + q_j\psi_j$$

where

$$\psi_j(t) = \int_0^\infty e^{itx} f_j(x)dx$$

is the characteristic function of X_j. The independence of $Y_1, \ldots, Y_n$ now shows for the characteristic function $\mathcal{C}$ of S_{ind}

$$\mathcal{C} = \prod^n (1 + q_j(\psi_j - 1)) \tag{9.6.96}$$

or

$$\ln\mathcal{C} = \sum_{j=1}^n \ln(1 + q_j(\psi_j - 1)) \approx \sum q_j(\psi_j - 1) = q(\psi - 1)$$

with $q = \sum q_j$ and $\psi = q^{-1}\sum q_j\psi_j$. Hence

$$\mathcal{C} \approx e^{q(\psi-1)}. \tag{9.6.97}$$

This expression is clearly the characteristic function of a distribution $P(q,X)$ where ψ is the characteristic function of X with

$$X = \frac{1}{q}\sum q_i X_j. \tag{9.6.98}$$

Thus we have approximated S_{ind} by a Poisson sum variable. $S_P = P(q,X)$.

Problem 9.34 Prove $E(S_{ind}) = \sum q_i E(X_i) = E(S_P)$ but

$$\sigma^2_{S_{ind}} \geq \sigma^2_{S_P}.$$

The above approximation is naturally called the PSV approximation, i.e. Poisson sum variable approximation.

The error in the above approximation arose essentially from the approximation of the logarithm. We have

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}.$$

Thus the error in $\ln \mathcal{C}$ is approximately given by

$$\sum q_i^2$$

because we used only the first order approximation. This can be made precise. To do so, define for two distributions (measures) P, Q on $\mathbb{R}$

$$d(P,Q) = \sup_A |P(A) - Q(A)|.$$

Then one can show the d is a metric (distance) and that the distance of S_{ind} and S_P can be estimated by $\sum q_i^2$.

Problem 9.35 Show that d defines a metric on all probability distributions.

The estimate for the error shows that the Poisson sum approximation is good for small-and medium-sized portfolios. The above PSV approximation can be improved. Such improvements have been derived by

P.S. Kornya Transact. Soc. of Act. **35**, 1983, pp. 823–858 and

S.M. Ross Prob. in Engineering and Inf. Sci. **1**, 1987, pp. 163–174.

We will, however, not pursue this any further.

In the remainder, we will concentrate mainly on the collective model. For this reason, we will briefly summarize the different features.

Collective Model

The portfolio is treated as a black box for which the claim number distribution N and claim size distribution X is defined based on experience from previous years. In as much as these distributions are parametrized these are estimated and extended. Large claims may be treated separately. While the aproximation is generally unbiased, i.e. it gives the correct mean, its variance is larger than that of the individual model.

Individual Model

Each individual policy in the portfolio is taken into account. Thus the actual composition of the portfolio is considered. For estimation, however, past experience has to be used. Then claim number variables and claim size variables are assigned to each policy.

There is another approximation – the Edgeworth approximation for the total claim sum, which may be quite useful, independent of the use of the individual or collective model. This approximation is based on two observations, namely

1. The larger the portfolio, the more the total claim sum will approximate a normal distribution.

2. For the total claim distribution, higher moments can be determined or estimated. These moments in turn determine the deviation from the normal distribution.

Usually this is measured in terms of the standardized moments

$$\begin{aligned} E((X-EX)^3)\cdot\sigma^{-3} &\qquad \text{the skewness or} \\ E((X-EX)^4)\cdot\sigma^{-4}-3 &\qquad \text{kurtosis.} \end{aligned}$$

But, as we will see, higher moments may also be considered. The skewness and kurtosis are good measures for the deviation from the normal distribution. (Why?)

Thus let S be a random variable with density f, distribution function F, mean value μ and standard deviation σ. If the standardized variable

$$Z=\frac{(S-\mu)}{\sigma}$$

does not differ much from a standard normal variable with moment-generating function

$$\begin{aligned} L(x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-1/2t^2}e^{-tx}dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-1/2[t^2+2tx+x^2]}e^{1/2x^2}dt \\ &= \exp\frac{1}{2}x^2 \end{aligned} \tag{9.6.99}$$

we can write

$$L_Z(t)=e^{t^2/2}\sum_{i=0}^{\infty}a_it^i. \tag{9.6.100}$$

Let ϕ be the density of the standard normal distribution. Then one can show by induction

$$t^ie^{t^2/2}=\int_{-\infty}^{\infty}e^{tu}(-1)^i\phi^{(i)}(u)du \qquad i=0,1,\ldots. \tag{9.6.101}$$

Thus

$$
\begin{aligned}
L_Z(t) &= \sum_{i=0}^{\infty} a_i \int_{-\infty}^{\infty} e^{tu}(-1)^i \phi^{(i)}(u)du \\
&= \int_{-\infty}^{\infty} e^{tu} \left(\sum_{i=1}^{\infty} a_i(-1)^i \phi^{(i)}(u) \right) du
\end{aligned}
$$

and the density of Z is given by

$$f_Z(t) = \sum_{i=0}^{\infty} a_i(-1)^i \phi^{(i)}(t). \tag{9.6.102}$$

In general, one will replace the infinite sum by a finite sum. Then

$$\tilde{f}_Z(t) = \sum_{i=0}^{n} a_i(-1)^i \phi^{(i)}(t) \tag{9.6.103}$$

defines the Edgeworth approximation of order n.

Problem 9.36 Prove (9.6.101).

Problem 9.37 Derive the expression for f and F from (9.6.102).

Problem 9.38 Compute the a_i $i = 0, 1$ for an Edgeworth approximation of order 1 in terms of the first 3 moments of S.

Problem 9.39 Show that $\mu_k = E(Z^k)$ gives $a_0 = 1$, $a_1 = a_2 = 0$, $a_3 = \mu_3/6$ and $a_4 = (\mu_4 - 3)/24$, $a_5 = 0$, $a_6 = \mu_3^2/72$.

Problem 9.40 Show that $P(Z \leq z) = \sum_{i=0}^{n}(-1)a_i\Phi^{(i)}(z)$ for the Edgeworth approximation of order n. Here Φ is the distribution function of the standard normal distribution.

The Edgeworth approximation can be improved in various ways. One such method is the normal power approximation which is described in the book by R. Beard et al.

The normal power approximation

The Edgeworth expansion to order 3 is

$$P_3(Z \leq z) = G(z) = \Phi(z) + \frac{\gamma}{6}\Phi^3(z) \qquad \text{with } \gamma = E(Z^3).$$

But

$$
\begin{aligned}
\Phi^3(z) &= \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt \right)^{(3)} = \frac{1}{\sqrt{2\pi}} \left(e^{-z^2/2} \right)'' \\
&= \frac{1}{\sqrt{2\pi}} (-ze^{-z^2/2})' = (z^2 - 1)\Phi'(z).
\end{aligned}
$$

Now consider the Taylor expansion of the Edgeworth expansion to first order

$$\begin{aligned} H(z+\Delta) = \mathrm{Edge}(z+\Delta) &= \mathrm{Edge}(z) + \mathrm{Edge}'(z)\Delta \\ &\approx \Phi(z) - \frac{\gamma}{6}\Phi'(z)\cdot(z^2-1) + \Phi'(z)\Delta. \end{aligned}$$

Choose $\Delta = \gamma/6(z^2-1)$. Then

$$H\left(z + \frac{\gamma}{6}(z^2-1)\right) \approx \Phi(z)$$

and if we set $y = z + \gamma/6(z^2-1)$ we get

$$z = \left(\frac{q}{\gamma^2} + 1 + \frac{6}{\gamma}y\right)^{1/2} - \frac{3}{\gamma}.$$

This implies for the distribution function F_S of the total claim distribution S

$$F_S(s) \approx \Phi\left(\left(\frac{q}{\gamma^2} + 1 + \frac{6}{\gamma}\cdot\frac{S-\mu}{\sigma}\right)^{1/2} - \frac{3}{\gamma}\right) \qquad \text{where} \quad \begin{array}{rcl} \mu &=& E(S) \\ \sigma^2 &=& \sigma_S^2 \end{array} \tag{9.6.104}$$

This is the normal power approximation for S. For the compound Poisson process we had seen:

$$E(S) = \lambda\mu_1, \quad \sigma_S^2 = \lambda E(X^2) = \lambda\mu_2 \quad \text{and} \quad E((S-ES)^3) = \lambda E(X^3) = \lambda\mu_3$$

where X is the single claim variable. Thus in this case

$$F_S(s) = \Phi\left(\left(\frac{q\lambda\mu_2^3}{\mu_3} + 1 + \frac{6\mu_2}{\mu_3}(s-\lambda\mu_1)\right)^{1/2} - \frac{3\lambda^{1/2}\mu_2^{3/2}}{\mu_3}\right). \tag{9.6.105}$$

Literature

R.E. Beard, T. Pentikäinen, E. Pesonen: Risk Theory. The stochastic basis of insurance. Chapman Hall 1984.
K.J. Schröter: Verfahren zur Approximation der Gesamtschadensverteilung. Karlsruher Reihe. Beiträge zur Versicherungswirtschaft.

9.7 Panjer's Method

We consider again the collective model with N the claim number distribution and X the individual claim distribution. Let again M_N be the generating function of N

$$M_N(x) = \sum_{k=0}^{\infty} P(N=k)x^k = \sum_{k=0}^{\infty} p_k x^k \tag{9.7.106}$$

and let L_X be the moment generating function of X

$$L_X(t) = E(e^{tX}). \tag{9.7.107}$$

Moreover let

$$\varphi_X(t) = \ln L_X(t). \tag{9.7.108}$$

For the (collective) claim sum variable S

$$S = \sum_1^N X_i \tag{9.7.109}$$

we have then

$$L_S(t) = E(e^{tS}) = \sum p_k L_X(t)^k = \sum p_k e^{k\varphi_X(t)}. \tag{9.7.110}$$

Many claim number distributions satisfy a recursion relation of the following type

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}. \tag{9.7.111}$$

With this we get

$$\begin{aligned} L_S'(t) = & \sum_{k=1}^{\infty} \left(a + \frac{b}{k}\right) p_{k-1} k \varphi_X'(t) e^{(k-1)\varphi_X(t)} e^{\varphi_X(t)} = \\ & a \sum_{k=1}^{\infty} (k-1) p_{k-1} \varphi_X'(t) e^{(k-1)\varphi_X(t)} e^{\varphi_X(t)} + (a+b)\varphi_X'(t) L_X(t) \cdot \\ & \sum p_{k-1} e^{(k-1)\varphi_X(t)}. \end{aligned}$$

This gives

$$L_S'(t) = a L_S'(t) L_X(t) + (a+b) L_X'(t) L_S(t). \tag{9.7.112}$$

Now one can either try to solve (9.7.112) directly, numerically or use a recursion method. Since the latter method is more commonly used in insurance mathematics, we will only treat this method here.

To apply this method, one discretizes X first, i.e. one assumes that claims are only determined to the next Euro (cent) or other currency unit. Thus we may assume that X, the claim size, has a discrete distribution. Assume

$$q_k = P_X(\{k\}) = P(X = k)$$

is the probability that the claim is k Euro. Then

$$L_X(t) = \sum q_k e^{tk} \qquad \text{with } q_0 = 0.$$

If each claim is given in integer Euro, then so is the sum, i.e. S also has a discrete distribution. Thus we write

$$L_S = \sum s_k e^{tk}$$

where s_k is the probability that the total claim sum is exactly k Euro. Insert this in (9.7.112)

$$\sum_{k=0}^{\infty} ks_k e^{tk} = a\sum_{j=0}^{\infty} js_j \left(\sum_{\ell=0}^{\infty} q_\ell e^{t\ell}\right) e^{tj} + (a+b)\sum_{j=0}^{\infty} s_j e^{tj} \sum_{i=1}^{\infty} iq_i e^{it}$$

and compare coefficients

$$ks_k = a\sum_{j=0}^{k-1} js_j q_{k-j} + (a+b)\sum_{j=0}^{k-1}(k-j)s_j q_{k-j}. \tag{9.7.113}$$

This can be rewritten as

$$ks_k = \sum_{j=0}^{k-1} s_j q_{k-j}[ak + b(k-j)]$$

or

$$s_k = \sum_{j=0}^{k-1} s_j q_{k-j}\left[a + b\frac{(k-j)}{k}\right] \tag{9.7.114}$$

or with $\ell = k - j$

$$= \sum_{\ell=1}^{k} s_{k-\ell} q_\ell \left[a + b\frac{\ell}{k}\right].$$

With an increasing refinement of this discretization one can then approximate the original continuous problem again. From (9.7.114) one obtains in this way for the density $f_S = g$ of S

$$g(x) = P(N=1)\cdot f_X(x) + \int_0^x \left(a + b\frac{u}{x}\right) f_X(u) g(x-u)du. \tag{9.7.115}$$

This is a Volterra integral equation of the second kind. There are methods known in which g can be approximately computed from this. This formula allows us to compute the s_k recursively.

Example 9.3 Assume N is Poisson $P(\lambda)$ distributed. Then $a = 0$ and $b = \lambda$ so that

$$ks_k = \lambda\sum_{j=0}^{k-1} s_k (k-j) q_{k-j}.$$

Remark. The method to analyze distributions by means of their generating functions, has a long tradition in the theory of Markov processes, particularly birth and death processes.

Problem 9.41 Let X be a random variable with a continuous distribution on $[0,\infty)$. Describe ways to discretize the distribution of X.

One can also use (9.7.111) to derive recursion relations for the higher moments of $S = \sum_{i=1}^{N} X_i$. For this, let again $L = L_{S-c}$ and L_X be the moment-generating functions of $S - c$, and X, respectively. Let M_X be the generating function of X. Then we have

$$L_{S-c}(t) = E(e^{(S-c)t}) = e^{-ct} L_S(t)$$

and

$$L_S(t) = M_N(L_S(t)).$$

From this we get by differentiation

$$L'(t) = -cL(t) + e^{-ct} M_N'(L_X(t)) L_X'(t). \tag{9.7.116}$$

Moreover, $M_N(x) = \sum_{i=0}^{\infty} p_i x^i$ where $P(N = i) = p_i$. Thus from (9.7.111) we get

$$\begin{aligned} M_N'(x) &= \sum_{i=0}^{\infty} p_i i x^{i-1} = \sum_{i=1}^{\infty} (ai + b) p_{i-1} x^{i-1} \\ &= (a + b) M_N(x) + \sum_{i=1}^{\infty} a(i - 1) x^{(i-1)} p_{i-1} \\ &= (a + b) M_N(x) + ax M_N'(x). \end{aligned}$$

This can be rewritten as

$$M_N' = \frac{(a + b)}{(1 - ax)} M_N. \tag{9.7.117}$$

Insert this in (9.7.116). Then

$$(1 - aL_X) L' = [(a + b) L_X' + acL_X - c] L. \tag{9.7.118}$$

Differentiate this expression k times and use the product rule for differentiation

$$(f \cdot g)^{(n)} = \sum \binom{n}{k} f^{(k)} g^{(n-k)}$$

$$(1 - aL_x) L^{(k+1)} - a \sum_{i=1}^{k} \binom{k}{i} L_X^{(i)} L^{(k-i+1)}$$

$$= \sum_{i=0}^{k} \binom{k}{i} [(a + b) L_X^{(i+1)} + acL_X^{(i)}] L^{(k-1)} - cL^{(k)}.$$

Now set $t = 0$ and use $L^{(k)}(0) = E((S - c)^k)$. Then we obtain finally for $k = 0, 1, \ldots$

$$\begin{aligned} (1 - a) E((S - c)^{k+1}) = \sum_{j=0}^{k} \binom{k}{j} \left[\left(\frac{k + 1}{j + 1} a + b \right) m_{j+1} + acm_j \right] \cdot \\ E((S - c)^{k-j}) - cE((S - c)^k) \end{aligned} \tag{9.7.119}$$

where $m_i = E(X^i)$. For $c = 0$ and $\mu_k = E(S^k)$ we get

$$(1 - a) \mu_{k+1} = \sum_{j=0}^{K} \binom{k}{j} \left(\frac{k + 1}{j + 1} a + b \right) m_{j+1} \mu_{k-j}.$$

Problem 9.42 Use (9.7.119) to compute a recursion relation for the central moments $c = E(S)$ of S.

Panjer's method has been extended by Sundt [Astin Bull. **22** 1992, pp. 61–80] to higher order recursions. Other recursive methods are discussed in the book "Verfahren zur Approximation der Gesamtschadensverteilung" by K.J. Schröter, Karlsruher Reihe.

One should not forget, however, that the computation of the distribution of S is usually the simpler part. More important are suitable estimates for the actual claim sum distribution.

10. Risk in Insurance

10.1 Solvency

Above, we defined risk as an adverse event, but we found it necessary also to restrict ourselves to risks which could be evaluated in monetary terms. We saw then that the basis of the insurance business is the transfer of risk from the insured to the insurer. For this, the insured has to pay a premium. This risk transfer is sensible because losses in a collective tend to even out. Because of the central limit theorem, the premium for a risk X will have the form

$$Pr(X) = E(X) + \text{Fluct.}(X) + \text{Admin. Term} + \text{Profit.} \tag{10.1.1}$$

Since there is not much to say about the administrative term and the profit (both are usually taken to be linear functions of $E(X)$ or $Pr(X)$), we will now concentrate on the first two terms.

$E(X)$ and $Fl(X)$ are usually estimated on the basis of past experience. Thus insurers will generally only insure risks which can be estimated and calculated properly. Thus the risk for the insurer arises if the actual claims exceed the computed (expected) claims. This can happen for several reasons.

1. *Fluctuations*
   In particular, nonlife insurance large claims may cause large fluctuations, as we have seen with fire insurance.

2. Error in estimation of parameters from previous years. This will mainly effect estimates of $E(S)$ and possibly σ_S^2.

3. Change in conditions and structure of the risk or technological conditions may change, for example, for political reasons. Example: After the collapse of the Eastern bloc and more intense East-West traffic, the number of thefts of luxury cars rose dramatically, resulting in a considerable loss for the insurance companies.

Fluctuations are generally controlled via a fluctuation term in the premium and the reserves. Thus, in times of low claims, reserves are built up to be used to finance large claims. The role of reserves is thus to spread risk over time. Thus it is only natural for the reserves to counter fluctuations, i.e. the contingency reserves or fluctuation reserves are somehow proportional to the standard deviation of the insured risks

$$\text{Fl. Reserves} \approx k \cdot \sigma_S \qquad S = \text{total sum of risks} \tag{10.1.2}$$

or

$$\text{Fl. Reserves} \approx k \cdot V \cdot \sigma_Y \qquad \begin{aligned} V &= \text{total volume of all policies} \\ Y &= \text{claim quotient.} \end{aligned} \tag{10.1.3}$$

k will generally vary between 3 and 6, in order to keep the risk of technical ruin sufficiently low. In most countries, this factor k is fixed by law. In Germany, e.g. $k = 4.5$ (6) for insurance against hailstorm (insurance of credits), while σ is determined from claims over a period of 10 to 20 years. For high risks like floods, earthquakes or some industrial fire insurances a large k is needed, while fire insurance for detached homes might only need $k = 3$. If the reserves are determined by law, the law must also regulate how the reserves are brought to their standard value (10.1.3), once low or large claims have occurred. In addition, it has to be settled on what basis σ is computed – on the whole portfolio or merely parts on of it. The fluctuation reserves are also called claims equalization reserves because their main function is to smoothen the claim process over the years.

With the aid of the Normal power approximation the somewhat naive forms of the fluctuation terms (10.1.2) can be made more precise. We had seen that the Normal power approximation gives for the distribution function F_X of a risk X

$$F_X\left(x + \frac{\gamma}{6}(x^2 - 1)\right) \approx \Phi(x) \qquad \gamma = \frac{E((X - EX)^3)}{\sigma_X^3}$$

where Φ is the distribution function of the standard normal distribution.

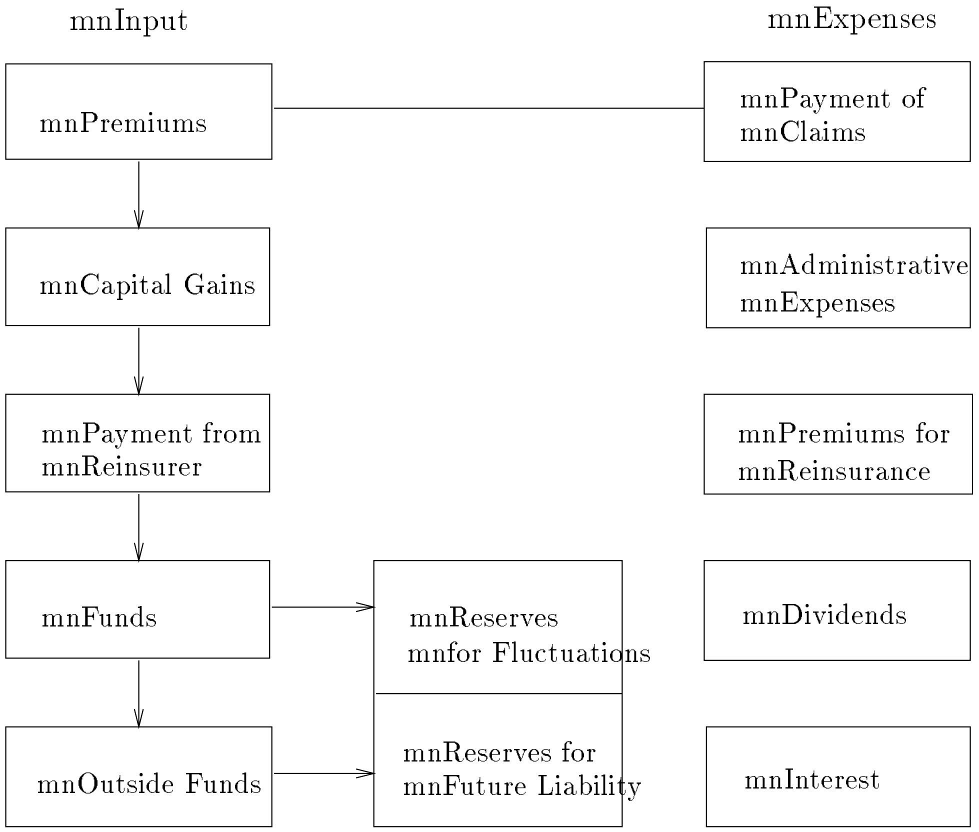
For a given ϵ let now z_ϵ be such that $\phi(z_\epsilon) = 1 - \epsilon$. Then

$$x_\epsilon = \mu + z_\epsilon \sigma + \frac{1}{6}(z_\epsilon^2 - 1)\sigma\gamma \tag{10.1.4}$$

is approximately the ϵ-critical value for X, i.e. the probability that X exceeds x_ϵ is less than ϵ. So, in a sense this formula gives a first order correction to the fluctuation term, which is due to the skewness. In applications X will be the claimsum S of a subportfolio for one year. For such cases this NP approximation is used for the Finish and Norwegian solvency control.

The risks in (2.) and (3.) are generally handled by reinsurance, i.e. the insurers themselves have an insurance against such risks. We have considered this before in but we will take up this problem again.

The flow of capital within an insurance can be represented as follows:



Of course, reserves also have to be put aside for other future obligations. In nonlife insurance, this happens, e.g. for an insurance extending over several years, like a warranty. Thus reserves fall into at least two groups: those for future liabilities and those to control fluctuations. Even though reserves are funds for an insurance, they differ from normal funds because they may be due in short notice. This should be reflected, of course, in the investment policies of a company. It is obvious that funds tied for a long term yield more interest than short-term funds. Similarly, funds invested in risky enterprises will yield more dividends. Finally, a company may borrow funds on the market, usually at a higher interest. Thus it is also the obligation of the actuary to decide with others on the proper management of finances.

10.2 The Ruin Problem

In insurance business, risk is transferred from the insured to the insurer. Even though the risk for the insurer is generally much smaller than for the insured, it is nonetheless not zero. It could, for example, occur that all claims in the portfolio become active. In this, admittedly extremely rare case, the claims will exceed the premium income and the reserves. More generally, we will speak of *ruin* for the insurance if reserves and premiums do not suffice to satisfy all claims. In general, this will of course not mean that the insurer is bankrupt because money may be obtained on the market or other insurers may come to the rescue. The chance of ruin is thus part of insurance business. In general, as insurer will operate in such a way that the ruin probability is kept below a certain limit. This limit naturally depends on the risk the insurer is willing to face and the market situation. An important role to honour liabilities is played by the reserves. This we have already seen with life insurances. But also for nonlife insurance, the role of reserves should be apparent.

The ruin problem for a casino

Before we begin our discussion let us consider a very simple model of the ruin problem, a simple roulette game. We assume that the casino has initial reserves and that the player plays either red or black with a winning probability of p – for a casino $p = 18/37$. The stake is always 1. If the player wins he gets 2. Thus in this case the casino looses 1 and if the paper looses the casino guins 1. Let X_i be the loss of the casino in i-th game. Then the X_i, $i = 1, 2, \ldots$ are independent and identically distributed with

$$P((X_i = 1) = p \qquad \text{and} \qquad P(X_i = -1) = 1 - p.$$

Thus the reserves of the casino after n games are

$$K_n = K - \sum_1^n X_i.$$

We are interested in the ruin probability $\psi(K)$ for the casino i.e.

$$P\left(\bigcup_{n=1}^{\infty}[K_n \leq 0]\right).$$

We have clearly $\psi(0) = 1$ and if we look at the first game we can derive a recursion relation

$$\begin{aligned}\psi(K) &= P([X_1 = 1]) \cdot \psi(K-1) + P([X_1 = -1])\psi(K+1) \\ &= p\psi(K-1) + (1-p)\psi(K+1).\end{aligned}$$

This is a second order difference equation, which can be solved by the Ansatz

$$\psi(K) = \lambda^K$$

which gives $\lambda = 1$ or $p/1-p$. Together with the constraint $\psi(0) = 1$ this gives

$$\psi(K) = A + (1-A)\left(\frac{p}{1-p}\right)^K$$

where A is an as yet undetermined constant. It is now plausible that $\pi(K)$ tends to 0 if the reserves are very large. Indeed later we will prove the Cramér Lundberg inequality, which gives $\psi(K) \to 0$ for $K \to \infty$. Thus $A = 0$ and

$$\psi(K) = \left(\frac{p}{1-p}\right)^K = e^{-r_p K} \qquad \text{with } r_p = -\ln\frac{p}{1-p}.$$

The following considerations will be more of a theoretical nature. Thus we will neglect all administratve costs, trends in expenses, inflation, and interest. Let I be the time interval we consider. If $I = [0, T]$ ($I = [0, \infty)$), we speak of a continuous time problem with finite (infinite) time horizon. If $I = \{0, 1, 2, \ldots, T\}$ ($I = \mathbb{N}$), we speak of a discrete time problem. The latter case arises if insurance operations are considered on a year-to-year basis.

Let Pr_t (S_t) denote the total premium income (claims cost) in $[0, t)$, $t \in I$. If K is the initial reserve we are interested in, the ruin probability of the insurance, i.e.

$$\psi(Pr, K) = P\left(\sup_{t \in I} S_t - Pr_t > K\right). \tag{10.2.5}$$

In this section we will study this problem for a special continuous time-infinite time horizon model. We will assume that the claim arrival process is Poisson and that the premiums arrive at a constant rate A. The claim process is therefore a compound Poisson process, where the individual claims X are distributed according to P_X. Assume $F(F^{*n})$ is the distribution function of X (the sum of n independent copies of X) then $P(S_t \leq x) = \sum_{K=1}^{\infty} e^{-\lambda t}\ (\lambda t)^k/k!\ F^{*k}(x)$. Let t be the time of the first claim and x the amount. Then ruin will not occur at all if it does not occur at the first jump

$$x \leq K + At$$

and if it does not occur later either

$$X_s = x \leq K + As - x \qquad \text{for } s > t.$$

The probability for this is

$$\lambda e^{-\lambda t} dt dP_X(x) \cdot \varphi(K - x + At)$$

where $\varphi(K) = 1 - \psi(K)$ and where dP_X is the distribution of claim sizes. Integrating over all claim sizes and times for the first jump, we get

$$\varphi(K) = \int_0^\infty \lambda \left(\int_0^{K+At} e^{-\lambda t} \varphi(K - x + At) dP_X(x)\right) dt. \tag{10.2.6}$$

With the substitution $u = K + At$, this gives

$$\varphi(K) = \frac{\lambda}{A} \int_K^\infty e^{-\lambda/A(u-K)} \left(\int_0^u \varphi(u - x) dP_X\right) du. \tag{10.2.7}$$

In order to eliminate at least one integration, we differentiate with respect to K

$$\varphi'(K) = \frac{-\lambda}{A} \int_0^K \varphi(K - x) dP_X(x) + \frac{\lambda}{A}\varphi(K). \tag{10.2.8}$$

The solution of this equation satisfies

$$\varphi(y) - \varphi(0) = \frac{\lambda}{A} \int_0^y \varphi(y - x)(1 - F(x)) dx \tag{10.2.9}$$

where F is the distribution function of X. One can easily see this by the differentiation of (10.2.9) and partial integration. Now we will show shortly that

$$\varphi(0) = 1 - \frac{\lambda}{A}\int_0^\infty (1 - F(x))dx = 1 - \frac{\lambda}{A}E(X). \tag{10.2.10}$$

Thus $\psi(0, A) = \lambda\mu/A$, $\mu = E(X)$, independent of the details of the claim distribution. The integral equation

$$\varphi(y) - 1 + \frac{\lambda\mu}{A} = \frac{\lambda}{A}\int_0^y \varphi(y - x)(1 - F(x))dx \tag{10.2.11}$$

cannot be solved in closed form in general. However, if X is assumed to be exponentially distributed with parameter $\mu = E(X)$, we get

$$\varphi(y) - 1 + \frac{\lambda\mu}{Y} = \frac{\lambda}{A}\int_0^y \varphi(y - x)e^{-x/\mu}dx. \tag{10.2.12}$$

The solution is then

$$\varphi(y) = 1 - \frac{\lambda\mu}{A}\exp\left(-\frac{y}{\mu}\left(1 - \frac{\lambda\mu}{A}\right)\right) \qquad \text{if } \lambda\mu < A. \tag{10.2.13}$$

Problem 10.1 Show that this function is the unique solution of (10.2.12).

This shows that ruin will happen eventually if the premium does not exceed the average claim rate, $\lambda\mu$. Summarizing, we see that the ruin probability with initial capital K is

$$\psi(K, A) = \frac{1}{1-\delta}\exp\left[-\lambda\frac{\delta}{1+\delta}\frac{K}{\lambda\mu}\right] \qquad \text{where } A = (1+\delta)\lambda\mu. \tag{10.2.14}$$

δ thus represents the fluctuation term factor according to the expectation value principle.

Problem 10.2 Let $I_1 = [0, T]$, $I_2 = [0, \infty)$, $I_3 = \{0, 1, \dots, n\}$ and $I_4 = \{0, 1, \dots\}$ and denote the corresponding ruin probabilities by $\psi_i(Pr, K)$, $i = 1, 2, 3, 4$. Then $\psi_3 \le \psi_4 \le \psi_2$ and $\psi_1 \le \psi_2$.

Theorem 10.1 Let $= [0, T)$, $T > 0$ and let $G_t(x) = P(S_t \le x)$. Then

$$1 - \psi(0, A) = \varphi(0) = \frac{1}{AT}\int_0^{AT} G_T(x)dx.$$

Proof. It follows from the Takacs' book: Combinatorial methods of Stochastic Processes. Wiley 1967, Th. 1, p. 37, that

$$P(S_t \le At,\ 0 \le t \le T \mid S_T = x) = 1 - \frac{x}{AT}.$$

Then

$$\begin{aligned}1-\psi(0,A)=\varphi(0)&=\int_0^{AT} P(S_t\le At,\ 0\le t\le T\mid S_T=x)dP_{S_t}(x)\\&=\int_0^{AT}\left(1-\frac{x}{AT}\right)dG_T(x)=\int_0^{AT}dG_T(x)-\frac{1}{AT}\int_0^{AT}x\,dG_T(x).\end{aligned}$$

Now use partial integration on the second integral. This gives

$$\varphi(0)=G_T(AT)-\frac{1}{AT}\left[-AT(1-G_T(AT))+\int_0^{AT}(1-G_T(x))dx\right].$$

From this the result follows easily.

Theorem 10.2 With μ and δ as above, $\lim_{T\to\infty}\varphi_T(0)=\delta/(1+\delta)$.

Proof. It follows from Theorem 10.1 that

$$\begin{aligned}\varphi_T(0)=1-\frac{1}{AT}&\int_0^{AT}(1-G_T(x))dx\\&=1-\frac{1}{AT}\int_0^{\infty}(1-G_T(x))dx+\frac{1}{AT}\int_{AT}^{\infty}(1-G_T(x))dx.\end{aligned}$$

Since for a general distribution function F_X which is concentrated on $(0,\infty)$

$$\int_0^\infty(1-F_X(x))dx=E(X)$$

and since the second term is $\lambda\mu T/AT$, it remains to show that the second integral converges to 0. With Tschebyschew's inequality, however, one sees

$$1-G_T(x)\le P(|S_T-\lambda T\mu\mid\ge x-\lambda T\mu)\le\frac{\lambda TE(X^2)}{(x-\lambda\mu T)^2}.$$

Use this to estimate the second integral. Now

$$\frac{1}{AT}\int_{AT}^{\infty}\frac{\lambda TE(X^2)}{x-\lambda\mu T)^2}dx=\frac{\lambda E(X^2)}{AT-\lambda\mu T}$$

shows the desired result.

Theorem 10.2 shows that an insurance which only charges the net premium will be ruined eventually.

In general, it is quite difficult to compute ruin probabilities if the claim number distribution is not Poisson and the claim size distribution is not of exponential type. However, at least for the case of a Poisson type claim number distribution one has the famous inequality of Cramér.

Theorem 10.3 *(Cramér)* Let $I=[0,\infty)$ and let the claim process be compound Poisson with paramter λ and single claim distribution function $F=F_X$. Moreover, let the premium rate be constant. $A > E(S_1)$ and let K be the initial reserves. Then the ruin probability ψ can be estimated by

$$\psi(K, A) \leq e^{-rK} \qquad \text{(Cramér-Lundberg inequality)} \tag{10.2.15}$$

where r is the solution of

$$\frac{A}{\lambda} = \int_0^\infty e^{ry}(1 - F(y))dy. \tag{10.2.16}$$

Proof. We note that the distribution function of the claims process S_t is given by

$$G_t(x) = \sum_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}F^{*n}.$$

Since the premium rate A is constant, we will not explicitly refer to it, and will denote by $\psi_n(K)$ the probability that ruin occurs with initial reserve K before the $(n+1)st$ claim. Then

$$\psi(K, A) = \lim \psi_n(K)$$

and it suffices to show

$$\psi_n(K) \leq e^{-rK}.$$

This will be done by induction on n. Since $n = 0$ gives

$$\psi_0(K) = \begin{cases} 1 & \text{for} \quad K < 0 \\ 0 & \text{for} \quad K \geq 0 \end{cases} \quad \text{and thus } \psi_0(K) \leq e^{-rK} \text{ it suffices}$$

to show the validity of the inequality for $n + 1$ under the assumption that it holds for n. Now

$$\psi_{n+1}(K) = \int_0^\infty \lambda e^{-\lambda t} \int_0^\infty \psi_n(K + At - x)dF(x)dt.$$

Here the inner integrals gives the probability that the first n claims amount to a loss of x. The first factor gives the probability that the n-th claim occurs t time units later. By induction assumption this can be estimated by

$$\leq \int_0^\infty \lambda e^{-\lambda t}\int_0^\infty e^{-r(K+At-x)}dF(x) = e^{-rK}\int_0^\infty e^{rx}\left(\int_0^\infty \lambda e^{-\lambda t - rAt}dt\right)dF(x)$$
$$e^{rK}\frac{\lambda}{\lambda + rA}\int_0^\infty e^{rx}dF(x) = e^{-rK}\frac{\lambda}{\lambda + rA}\left(1 + \frac{Ar}{\lambda}\right) = e^{-rK}.$$

Thus it remains to show

$$\int_0^\infty e^{rx}dF(x) = \left(1 + \frac{Ar}{\lambda}\right). \tag{10.2.17}$$

We will do this but connect this proof with some further properties of r, the *solvency, adjustment* or *Lundberg coefficient.*

Consider a random variable X such that: $EX > 0$, $P([x < 0]) > 0$. In our situation, $-X$ will be a risk process at a particular time. Thus $P([X < 0])$ essentially means that an insurance may have losses through its policies. Now consider

$$f(t) = E(e^{-tX}) \qquad t \in \mathbb{R}.$$

Then

$$f(0) = E(1) = 1;\ f'(t) = -E(Xe^{-tX}) \quad \text{so} \quad f'(0) = -EX < 0$$

while

$$f''(t) = E(X^2 e^{-tX}) > 0.$$

Thus f is strictly convex and for $t \to \infty$ we get $f(t) \to \infty$. Thus there is a second strictly positive solution r of

$$f(t) = 1. \tag{10.2.18}$$

This solution is called the *Lundberg coefficient*. It need not exist because $E(e^{tX})$ need not exist. But we will only consider the case in which the moment-generating function exists for small t. This is the case if the claim size is bounded.

Problem 10.3 Show $E(e^{-tX}) \to \infty$ for $t \to \infty$ if $P([X < 0]) > 0$.

Problem 10.4 Show for $g(t) = \ln(E(e^{-tX}))$, $g'(0) = EX$, $g''(0) = \sigma_X^2$.

We can use this to get an estimate the Lundberg coefficient r. To do so expand g (Problem 10.4) up to see second order around 0

$$g(t) \approx tEX + \frac{t^2}{2}\sigma_X^2.$$

Then $g(t) = 0$ gives

$$r \approx \frac{2EX}{\sigma_X^2} \tag{10.2.19}$$

as an approximation for r. This shows that distributions with a large variance are more dangerous as regards the ruin. As an example consider a normal distribution $N(\mu, \sigma)$

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-tX} e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-t\mu} \int_{-\infty}^{\infty} e^{-x^2/2\sigma^2 - tX} dx \\ &= e^{-t\mu} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[\frac{1}{2\sigma^2}(X^2 + 2t\sigma^2 x + t^2\sigma^4 - t^2\sigma^4)\right] dx \\ &= e^{-t\mu} e^{t^2\sigma^2/2} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}(x + t\sigma^2)\right] dx = e^{-t\mu} e^{t^2} \frac{\sigma^2}{2}. \end{aligned}$$

If we solve $f(t) = 1$ in this case, we get $e^{t\mu} = e^{t^2}\ \sigma^2/2$ or $t\mu = t^2\ \sigma^2/2$. But this just gives (10.2.19).

Let us now consider the risk in insurance with $X = A - S$, where A is the premium and S the claims in one period. Then

$$1 = E(e^{-(A-S)t})$$

or

$$e^{At} = E(e^{tS}) = \sum e^{-\lambda}\frac{\lambda^k}{k!}L_Z(t)^k = e^{-\lambda+\lambda L_Z(t)}.$$

Thus

$$\frac{A}{\lambda}t = -1 + L_Z(t) = -1 + \int_0^\infty e^{tX}dF(x). \tag{10.2.20}$$

This is the desired equation (10.2.17). Partial integration gives

$$\frac{A}{\lambda}t = t\int_0^\infty (1 - F_Z(x))e^{tx}dx. \tag{10.2.21}$$

Division by t makes it another useful equation to determine the Lundberg coefficient.

Example 10.1 For exponentially distributed claims we saw in (10.2.13) that the ruin probability was given by

$$\psi(y) = \frac{\lambda\mu}{A}\exp - \left(\frac{\frac{(1-\lambda\mu}{A}}{\mu}y\right).$$

Let us compute the Lundberg coefficient for that case. For this we use (10.2.20)

$$\frac{A}{\lambda}t = -1 + \mu\int_0^\infty e^{tx-\mu x}dx = -1 - \frac{\mu}{(t-\mu)} \qquad \text{for } t < \mu.$$

This gives

$$\frac{A}{\lambda}(\mu - t) = 1 \qquad \text{or} \qquad r = \mu - \frac{\lambda}{A}.$$

Theorem 10.3 can be refined somewhat. With the same assumptions as in the theorem, one has the Cramér Lundberg inequality

$$\psi(K, A) \approx \frac{(A/\lambda - E(Z))}{[r\int_0^\infty e^{rx}x(1 - F_Z(x))dx]}. \tag{10.2.22}$$

Here Z is again the single claim distribution.

Problem 10.5 Compare this formula with the exact result for exponentially distributed Z.

Problem 10.6 Improve the estimate (10.2.19) by also taking into account the 3rd and 4th moment.

Problem 10.7 Along the lines of the proof of Theorem 10.3, show: Let Y_1, Y_2, ... be a sequence of identically distributed independent random variables with distribution function F. Let r be the Lundberg coefficient for F. Then for $n \geq 0$

$$\psi(u) = P\left(\left[\sup_n \sum_{i=1}^{n} Y_i > u\right]\right) \leq e^{-ru}.$$

The above estimates are somewhat unrealistic. Meanwhile, a number of results have been derived for the ruin probability under more realistic assumptions, e.g. P. Embrechts, H. Schmidli: Ruin estimations for a general insurance model. Adv. Appl. Prob. **26**, 1994.

All in all, these results show the importance of reserves and adequate fluctuation terms. It is also apparent that one might use the Lundberg coefficient, or rather its inverse, as a measure for the risk.

Problem 10.8 How would interest on the reserves influence the ruin probability?

Problem 10.9 Can one extend these results on the ruin probability also to claim distributions for which the claim number distribution is negative binomial? (Hint: Ammeter Transformation)

Literature

H. Bühlmann: Mathematical Models in Risk Theory. Springer 1970.
C. Hipp, R. Michel: Risikotheorie. Verlag Versicherungswissenschaft 1990.
W.R. Heilmann: Grundbegriffe der Risikotheorie. Verlag Versicherungswissenschaft 1987.
E. Straub: Non-Life Insurance Mathematics. Springer Verlag 1988.

11. Rate Making

Even though this chapter primarily adresses non-life insurance, some aspects of it are also relevant to life insurance. The following considerations are an abstraction of the situation which we have already encountered. In rate making, the actuary is confronted with the following problems.

1. Distribution of claims

   (a) The actuary has to determine the distribution of claims and their trends. In order to describe the risks as well as accurately as possible he has to determine the key variables and form risk classes to differentiate the premiums. The premium should reflect the particular risk as closely as possible. Differentiate as much as necessary and as little as possible. Determine adequate models for forecasting and description of claims.

   Introduce benefits which induce the customer to avoid or minimize claims (Bonus-malus, deductibles, premium refunds).

   (b) Mathematical methods.
   Choose parametrized models to describe claims, e.g. the models discussed above or generalized linear models. Statistical methods to determine parameters and their evolution – maximum likelihood, χ^2-tests of fit, time series. To determine the key variables use analysis of variance, factor analysis multiple regression or cluster analysis. In practice, more than two or three key variables cannot be treated because then the classes become too small. Independence of these parameters suggests a multiplicative model (Simon Bailey). But one can also use an additive model. Particular methods, e.g. distributions with longtails, are needed to describe large claims.

2. Fixing the fluctuation term, fluctuation reserves.
   Determine the variance and skewness of the distribution. How does the variance depend on the volume and expectation value. Fix fluctuation reserves and decide on reinsurance plans.

   (a) Mathematical methods.
   Estimation, goodness of fit tests for variance, skewness and the Lundberg coefficient. Expected utility or losses.

3. Controlling products, developing new ones.
   The performance of new products has to be followed up and new ones have to be designed. All this in view of competition. This forces the insurer to differentiate and modularize.
   The customer (insured) expects constant premiums based on transparent principles and reimbursement under straightforward conditions.

(a) Methods.
Study of market, demand, law and tax conditions. Analyse products of competitors.

Even though multiple regression or analysis of variance plays an important role in problems depending on several factors, we will not go into this here because these techniques are usually treated in statistics courses. The same applies to the use of generalized linear models.

11.1 Principles for rate making

We have seen above that a premium for risk X usually has the form

$$\begin{aligned}\text{Premium} =&\text{Expect. value} + \text{Fluctuation term} + \text{Administr. term} \\ &+ \text{Const.} + \text{Profit.}\end{aligned} \tag{11.1.1}$$

The profit and administrative term are usually taken proportional to the expectation value or the premium. For the fluctuation term we had

$$\text{Fl. term} = \begin{cases} \lambda, \quad E(X) & \text{expectation value principle} \\ \lambda_2 \quad \sigma_X & \text{standard deviation principle} \\ \lambda_3 \quad \sigma_X^2 & \text{variance principle.} \end{cases} \tag{11.1.2}$$

These are, in fact, the standard principles which are mostly used. Of these, the expectation value principle is used because of its simplicity. Only the variance principle allows some theoretical justification. Nonetheless, there remains the problem of finding suitable methods to determine the first two terms in (11.1.1). One obvious and plausible modification of the variance principle is the *semivariance principle*

$$\text{Fl. term} = \lambda_4 E((X - EX)_+^2), \tag{11.1.3}$$

i.e. only the claims that exceed EX are counted. In the remainder, we will call any real valued function H of X a *premium principle*. In that case, $H(X)$ would be the premium for risk X. One way to find and judge possible premium principles, is to list desirable properties

1. $H(X) \geq E(X)$.
   As we have seen, this is a must for any premium principle.

2. $H(X) \leq F_X^{-1}(1)$.
   $F_X^{-1}(1)$ gives the maximal damage one can expect with risk X. Then it is only sensible that $H(X)$ should not exceed this because then nobody would buy this insurance.

3. $X = c$, then $H(X) = c$.
   If X is constant and does not fluctuate, no fluctuation term is needed. Of all principles above, only the semivariance principle satisfies this condition.

4. $H(X+c) = c + H(X)$ Translation invariance.
   Even though this is plausible, the expectation value principle violates translation invariance.

5. $H(cX) = cH(X)$ $c > 0$ (positive) Homogeneity.

6. X, Y independent then $H(X+Y) \leq H(X) + H(Y)$ Subadditivity.
   In the case of equality, one speaks of additivity.

7. Convexity: $H(\lambda X + (1-\lambda)Y) \leq \lambda H(X) + (1-\lambda)H(Y)$ $0 \leq \lambda \leq 1$.

Finally, one wants the premium principles to be robust because the premiums are usually computed on the basis of past experience. So one does not want any sensitive dependence on outliers etc.

Problem 11.1 Check to what extent the premium principles above satisfy the "desirable" properties 1, ... , 7.

The variance principle has one undesirable property. We have

$$nH\left(\frac{X}{n}\right) = n\left[E\left(\frac{X}{n}\right) + \lambda_3\sigma^2_{X/n}\right] = E(X) + n\lambda_3\frac{1}{n^2}\sigma^2_X = E(X) + \lambda_{3/n}\sigma^2_X$$

i.e. by splitting up the risk one could in principle reduce the premium.
Actuaries have devised a multitude of premium principles. These are discussed in: M. Goovaerts, F. DeVylder, J. Haezendonck: Insurance Premiums. North Holland 1984.
However, since most have only a theoretical importance, we will restrict ourselves to only two such methods.

Expected utility principle

Assume the insurer has a utility function u. Then one can expect

$$u(0) = 0,\ u' > 0. \tag{11.1.4}$$

By replacing u with the equivalent utility $u/u'(0)$, we can also achieve $u'(0) = 1$. The expected utility principle now demands to choose $H(X)$ such that

$$E(u(H(X) - X)) = u(0) = 0. \tag{11.1.5}$$

Example 11.1 Exponential principle: $u(x) = 1/a(1 - e^{-ax})$ $a > 0$

$$H(X) = a^{-1}\ln E(e^{aX}). \tag{11.1.6}$$

Problem 11.2 Which properties are satisfied by the exponential principle?

Problem 11.3 Apply the exponential principle to the normal and exponential distribution.

This approach has two drawbacks. The first is that the use of utility functions along with the expected utility principle is problematic. The second is that even mildly complicated utility functions will lead to unwieldy expressions. In order to somewhat counter the second reproach, assume u is analytic. Then

$$E(u(H-X)) = \sum_0^\infty \frac{u^{(k)}(0)}{k!} E((H-X)^k) = f(H) \qquad H = H(X).$$

Since H is determined by $f(H)=0$ and since H will be close to EX, we can apply Newton's method with $H = E(X)$ and get to first order

$$H(X) \approx E(X) - \sum_0^\infty \frac{u^{(k)}(0)}{k!}(-1)^k \frac{\mu_k}{\sum_0^\infty \frac{u^{(K+1)}(0)}{k!}\mu_k(-1)^k} \tag{11.1.7}$$

where $\mu_k = E((X-EX)^k)$ is the k-th central moment. To third order, i.e. $u^{(IV)}(0) = u^{(IV)}(0)\cdots \approx 0$ this becomes

$$H(X) \approx E(X) - \frac{\left(\frac{u''(0)}{2}\sigma_X^2 - \frac{u'''(0)}{6}\mu_3\right)}{\left(1 + \frac{u'''(0)}{2}\sigma_X^2\right)}. \tag{11.1.8}$$

This makes sense only if $u''(0) < 0$, i.e. if the deciding person is risk averse. In that case, $u'''(0) = 0$, i.e. second-order approximation, we are back to our variance principle with risk aversion coefficient $-u''(0)/2$.

Problem 11.4 Expand the exponential principle to 3rd order and determine the corresponding premium.

The Loss Function Principle

Let again $H = H(X)$ be the premium. If a value x for X is realized, the insurance has a loss which we describe by $L = L(H(X), x)$. Thus the *loss* is actually a random variable, namely $L = L(H(X), X)$. The loss function principle now means minimizing the expected losses, i.e. determine $H = H(X)$ such that

$$E(L) = E(L(H,X)) \to \min. \tag{11.1.9}$$

Example 11.2 $L(H,x) = (x-H)^2$ then $H(X) = E(X)$.

Problem 11.5 Show that $L(H,x) = (e^{ax} - e^{aH})$ leads to the exponential principle.

In order to weigh large deviations more strongly, let

$$L(H,x) = e^{cx}(x-H)^2 \qquad \text{with } c > 0. \tag{11.1.10}$$

Then

$$H(X) = \frac{E(Xe^{cX})}{E(e^{cX})}. \tag{11.1.11}$$

This latter form is called the *Esscher Principle.*

Problem 11.6 Prove that (11.1.11) gives the correct solution to (11.1.10).

Problem 11.7 Apply the Esscher principle to normally distributed X.

Problem 11.8 Do the same for an exponentially distributed X.

Problem 11.9 Let u be a utility function with $u(0) = 0$. What premium principle does one get with the loss function

$$L(H, x) = \int_x^H u(t - x)dt?$$

Since most of these nonstandard principles are more of a theoretical nature we will not discuss them any further.

11.2 Credibility Theory

Credibility theory is a method to improve estimates, used in rate making, on the basis of actual claim experience. It may also be described as an adaptive estimation of parameters or risk theory. It is basically a Bayesian technique. Here we will only consider the simplest situation of estimating the expectation value.

For this we assume that the claim distribution has a density – the discrete case is quite similar – which depends on a parameter θ. Since θ itself is not known exactly, this lack of knowledge manifests itself through the fact that θ is distributed according to some probability distribution $U = U(\theta)$. This function U thus represents our knowledge of the situation before a new set of measurements or observations is made. Therefore U is also called the a *priori distribution*. If U is a very broad distribution with a large variance, our knowledge is rather vague. If it is narrow, our knowledge is more precise. But $u(\theta)$ can also be interpreted that we are dealing with a collective with different risk behaviour described by θ. This means that the actual portfolio is inhomogeneous, where θ measures this inhomogeneity. We are thus dealing with the situation where this inhomogeneity has not been resolved so that we are forced to find a suitable premium for the whole portfolio. Now a new observation is made and $x = (x_1, \ldots, x_n)$ is found. The new, a posteriori distribution can then be calculated by means of the Theorem of Bayes. Before starting this, we first need the definition of a conditional probability. Let A, B be measurable subsets of the probability space Ω with $P(B) \neq 0$. Then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

is called the *conditional probability of* A under the condition of B.

Theorem 11.1 *(Bayes)* Let A_i $i = 1, 2, \ldots$ be a disjoint decomposition of the probability space Ω and let B be an event with $P(B) > 0$. Then

$$P(A_j \mid B) = \frac{P(A_j)P(B \mid A_j)}{\sum_i P(A_i)P(B \mid A_i)}. \tag{11.2.12}$$

This theorem is shown in most elementary books on probability. Important for us, however, is the interpretation

$$A_j \to P(A_j)$$

represents the a priori probability, B the observation and

$$A_j \to P(A_j \mid B)$$

the a *posteriori probability*. If the probability space Ω is discrete and if the A_i are its minimal parts, the atoms, formula (11.2.12) gives us indeed the change from the a priori distribution to the a posteriori distribution. If we are dealing with $\Omega = \mathbb{R}$ or $\mathbb{R}_+$ and continuous densities, the corresponding formulae can be obtained by a limiting argument.

Before we write down these formulae, let us first familiarize ourselves with compound distributions. We will do this only for the continuous case. One gets the discrete case by replacing integrals by sums and vice versa. Thus let $f(x, \theta)$ be a probability density for $x \in \mathbb{R}$, depending on the parameter θ. Assume θ is distributed with density $u(\theta)$. Then

$$\begin{aligned} P(A) &= \int \left(\int \chi_A(x) f(x, \theta) dx \right) u(\theta) d\theta \\ &= \int P(A \mid \theta) u(\theta) d\theta \qquad \chi_A(x) = \begin{cases} 1 & \text{for} \quad x \in A \\ 0 & \text{for} \quad x \notin A \end{cases} \end{aligned} \tag{11.2.13}$$

defines a probability measure. We had encountered this situation before.

Example 11.3 Consider a Poisson distribution whose parameter λ is itself exponentially distributed. This case was discussed before. Then

$$P(\{k\}) = \int \frac{\lambda^k}{k!} e^{-\lambda} \mu e^{-\mu\lambda} d\lambda.$$

Remark. In (11.2.13), x could be a vector and θ a vector index if the integrals are multiple integrals.

Using (11.2.13) one can easily show for a compound distribution

$$\mu = E(X) = \int \left(\int x f(x, \theta) dx \right) u(\theta) d\theta = \int \mu_X(\theta) u(\theta) d\theta$$

and

$$\begin{aligned} \sigma_X^2 &= \int \left(\int (x-\mu)^2 f(x,\theta) dx \right) = \int \left(\int (x - \mu(\theta)) + (\mu(\theta) - \mu))^2 f(x, \theta) dx \right) u(\theta) d\theta \\ &= \int \int \left((x - \mu(\theta))^2 f(x, \theta) dx \right) u(\theta) d\theta + \int \left(\int (\mu(\theta) - \mu)^2 f(x, \theta) dx \right) u(\theta) d\theta \end{aligned}$$

because the cross term vanishes (Why?). Thus we get

$$\sigma_X^2 = \int \sigma_X^2(\theta) u(\theta) d\theta + \int (\mu(\theta) - \mu)^2 u(\theta) d\theta = E(\sigma_{X(\theta)}^2) + \sigma_U^2.$$

Let us now move to the general situation. Let $f(x,\theta)$ be the density for a risk X and let θ be the (vector) parameter which describes this risk. Assume θ is distributed according to the a priori distribution u. Now we perform n independent measurements of X with the result $(x_1,\dots,x_n)=x$. Then the a posteriori density $\tilde{u}(\theta \mid x_1,\dots,x_n)$ is given by

$$\tilde{u}(\theta \mid x_1,\dots,x_n) = f(x_1,\theta)\dots f(x_n,\theta)u(\theta)N \tag{11.2.14}$$

where N is a normalization constant. Here we assume, of course, that $X_1(\theta),\dots,X_n(\theta)$ with θ fixed are independent. In order to relate (11.2.14) to (11.2.12), note that B corresponds to the measurement of $(x_1,\dots,x_n)$, $u(\theta)$ to $P(A_j)$ and $\tilde{u}(\theta \mid x_1,\dots,x_n)$ to $P(A_j \mid B)$. Thus $L(x,\theta)=f(x_1,\theta)\dots f(x_n,\theta)$ gives us the likelihood for θ given $(x_1,\dots,x_n)$. (11.2.14) can be rewritten as

$$\tilde{u}(\theta \mid x) = \frac{L(x,\theta)u(\theta)}{\int L(x,\theta)u(\theta)d\theta} \tag{11.2.15}$$

which is even closer to (11.2.12). Expectation values which involve $\tilde{u}$ will then be denoted by $E(\cdots \mid x_1,\dots,x_n)$. Then

$$\mu_{\text{cred}} = E(\mu_X(\theta) \mid x_1,\dots,x_n) = E(E(X(\theta))) \tag{11.2.16}$$

is the desired credibility estimator for μ because it minimizes the risk

$$E((\mu(\theta)-\hat{\mu})^2 \mid x_1,\dots,x_n) \to \min \tag{11.2.17}$$

subject to the measurement of $x_1,\dots,x_n$. For the proof that $\hat{\mu}=\mu_{\text{cred}}$ is the desired minimum, one needs Hilbert space techniques, so we will not dwell any further on it, but rather consider an example.

Example 11.4 Assume the claim number distribution is compound Poisson with a priori distribution of Gamma type, with density

$$u(\lambda) = \frac{c^\gamma}{\Gamma(\gamma)}e^{-c\lambda}\lambda^{\gamma-1}.$$

The expectation value of U is γ/c and its variance γ/c^2. Now assume n measurements were with the results $n_1,\dots,n_n$. Then

$$\tilde{u}(\lambda) = \left(e^{-\lambda}\frac{\lambda^{n_1}}{n_1!}\right)\dots\left(e^{-\lambda}\frac{\lambda^{n_n}}{n_n!}\right)\frac{c^\gamma}{\Gamma(\gamma)}e^{-c\lambda}\lambda^{\gamma-1}N \tag{11.2.18}$$

and the expectation value is the expectation value of another Gamma distribution

$$\begin{aligned}\mu_{\text{cred}} =& E(\lambda \mid n_1,\dots,n_n) = \frac{\left(\int_0^\infty \lambda e^{-n\lambda}\lambda^{n_1+\dots n_K+\gamma-1}d\lambda\right)}{\int_0^\infty e^{-n\lambda}\lambda^{n_1+\dots+n_K+\gamma-1}d\lambda} \qquad (11.2.19)\\ =&\frac{(N+\gamma)}{(c+n)} = \frac{n}{n+c}\left(\frac{1}{n}\sum n_i\right)+\frac{c}{n+c}\frac{\gamma}{c} \qquad \text{where } n=\sum n_i.\end{aligned}$$

This example is, of course, particularly simple because the Gamma distribution is a *conjugate distribution* for the Poisson distribution. Conjugate distributions play an important role in this theory because computation then becomes particularly simple. They are defined as follows: Let $\mathcal{D}$ be a class of 'a priori' distributions in x. The class of parameter distributions $\mathcal{D}_c$ is called conjugate to $\mathcal{D}$ if the following holds: $u \in \mathcal{D}_c$ then $\tilde{u} \in \mathcal{D}_c$ also

$$\mathcal{D}_c \ni u \xrightarrow[\text{conditioning}]{\sim} \tilde{u} \in \mathcal{D}_c \qquad \tilde{u}(x) = \tilde{u}(x \mid x_1, \ldots, x_n)$$

a priori distribution a posteriori distribution

Problem 11.10 Show that the Beta distribution class is conjugate to the Binomial distributions.

In the case of parametrized conjugate distributions the conditioning map, the transition from u to $\tilde{u}$, becomes a map in parameter space.

We saw above that the credibility estimator for the mean value is a convex combination of the apriori mean value γ/c and the mean of the new measurements. This is in general not always true. Nevertheless, one uses the many computable examples, where it holds, as a motivation to define the *linear credibility estimator* by

$$\mu_{\ell,c} = b + a\frac{1}{n}\sum X_i = b + a\overline{X} \tag{11.2.20}$$

where a and b are chosen such that

$$E_\theta([E(\mu_X(\theta) \mid x_1, \ldots, x_n]) - b + a\overline{X}(\theta)]^2) \to \min \tag{11.2.21}$$

is minimal. Here E_θ denotes the expectation value with respect to the a priori distribution. One can show that

$$\begin{aligned} E_\theta([\mu_X(\theta) - (b + a\overline{X}(\theta))]^2) =& E_\theta([\mu_X(\theta) - E(\mu_X(\theta) \mid x_1, \ldots, x_n)]^2) \\ &+ E_\theta([E(\mu_X(\theta) \mid x_1, \ldots, x_n) - a - b\overline{X}(\theta)]^2) \end{aligned} \tag{11.2.22}$$

so that it suffices to minimize the left-hand side of (11.2.22) if one looks for the optimal linear credibility estimator. The solution for this problem is

$$a = \frac{\sigma_\theta^2(\mu_X(\theta))}{1/nE_\theta(\sigma_X^2(\theta)) + \sigma_\theta^2(\mu_X(\theta))} \qquad b = (1-a)\mu_X, \tag{11.2.23}$$

i.e. the linear credibility estimator is a convex combination of the old (a priori) mean value μ_X and the empirical mean $\overline{X}$. (11.2.23) also shows that the influence of the new data is larger if n is large or if the a priori information is rather vague.

In actual application the a priori distribution is not known and has to be estimated first. In Bayesian statistics this is the domain in which beliefs and subjective probability enter. In applications one can try to obtain information on the a priori distribution by exerience rating or from the knowledge of similar or related insurances. Experience rating is use a great deal in industrial insurance, like power plants, ships or planes. In group life insurance or motor car insurance information from similar insurances may be used.

Although this is but a first sketch of credibility, we have to leave it at that. More details of actual applications in various cases can be found in the literature or the following books.

Credibility theory has an interesting application to the bonus-malus system in automobile insurance. Based on data from 1961 F. Bichsel [Mitt. der Vereinig. schweizerischer Versicherungsmathematiker 1964, 119–130]

Number of policies with k-claims in 1961

k	number of policies	Poisson model	mixed * Poisson model
0	103 704	102 629	103757
1	14 075	15 922	13 934
2	1 766	1 235	1 871
3	255	64	251
4	45	3	34
5	6	–	5
6	2	–	1
	119 853	119 853	119 853

* The mixed Poisson model is a compound distribution with Gamma density.

developed the following model. Even though the χ^2-test does not quite accept the compound Poisson distribution, he used it nonetheless and wrote

$$P(N(t) = k \mid \theta) = \frac{1}{k!} e^{-\theta t} (\theta t)^k.$$

For θ we assume a Gamma distribution with density

$$f(\theta) = \frac{\alpha\beta}{\Gamma(\beta)} \theta^{\beta-1} \exp(-\alpha\theta)$$

with $E(\theta) = \beta/\alpha$ and $\sigma_\theta^2 = \frac{\beta}{\alpha^2}$. With this density

$$E(N(t)) = t\frac{\beta}{\alpha} \qquad \sigma_N^2 = t\frac{\beta}{\alpha}\left(1 + \frac{t}{\alpha}\right).$$

This can also be computed with the aid of general results on expectation and variance for compound distributions. The a posteriori density after k-claims is then

$$f(\theta \mid k) = e^{\theta t} \frac{(\theta t)^k}{k!} f(\theta).$$

This is again a Gamma density. The a posteriori expectation value, which is the exact credibility net premium, is then

$$e(k, t) = \frac{\beta + k}{\alpha + t}.$$

From this Bichsel computed the premium levels

Theoretical premium levels according to Bichsel

$t\backslash k$	0	1	2	3	4
1	87 %	173 %	260 %	346 %	433 %
2	76 %	153 %	229 %	306 %	382 %
3	68 %	137 %	205 %	273 %	341 %
4	62 %	123 %	185 %	247 %	309 %
5	56 %	113 %	169 %	226 %	282 %
6	52 %	104 %	155 %	207 %	259 %

Literature:

H. Bühlmann: Mathematical methods in risk theory. Springer Verlag 1970.
M.J. Goovaerts, R. Kaas, A.E. Van Herrwaarden, T. Banwelinckx: Effective actuarial methods. North Holland 1990.
B. Sundt: An introduction to non-life insurance mathematics. VVW Karlsruhe 1993.

11.3 Sharing of Risks

Insurance is the method of transferring risk from the insured to the insurance company. Even if the risk of an insurance company can be controlled much better, insurance companies also need an insurance for large or extreme risks. Think for example of an airline crash, a fire in a large factory or the stranding of an oiltanker. The damage in these cases can easily sum up to several hundred million Euro.

In this section we will therefore discuss the sharing of risk between the insured and the insurance on one hand and the sharing of risk between insurance and reinsurance on the other hand. The emphasis will be on the second part though.

Sharing of risk between client and insurance

There are basically two forms of risk sharing practiced at this level.

Proportional risk sharing: The insurance covers only the part cX, $0 < c < 1$ of the risk X. Mathematically this causes no problem. This method is practiced in health insurance. But also underinsurance against theft or breakage is a form of proportional risk sharing.

Non-proportional risk sharing

The *excess of loss insurance* (XL) is the most common form in this domain. In this case the insurance covers only

$$Z = \max(X - a, 0) \quad \text{and } Y = \begin{cases} X & X \leq a \\ a & X > a \end{cases} \tag{11.3.24}$$

remains with the insured. a is called the *deductible*, excess point or retention.
The reason for these forms of risk sharing are from the point of view of the

Client: Reduction of the premium
Reduction of the moral risk
Reduction of paper work with small claims.

Insurance: No administration of small claims. Induction of the client to risk avoiding measures, control of the moral risk.
The excess of loss insurance is practiced a great deal in auto motor insurance. There are also forms of insurance with an upper limit u. In this case the client retains

$$X_c = \begin{cases} X & \text{if} \quad X < a \\ a & \text{if} \quad a < X < u \\ X - u & \text{if} \quad X \geq a + u \end{cases} . \tag{11.3.25}$$

The risk for an insurance

In its work the insurance faces a number of risks. These are

Fluctuations: The fluctuations are the usual risk in insurance. They become smaller as the portfolio grows. They are handled on one hand by the fluctuation part of the premium and the fluctuation reserves. Extreme fluctuations are handled by catastrophe reserves or are controlled by reinsurance.

Structural changes: In this case there is a change in the parameters of the risk process. This can happen through changes in external political or economic conditions, new laws, new treatments in health insurance etc. The risk in this case is proportional to the portfolio. Sufficient reserves allow the insurance to cope with such eventualities. But there have been cases where the insurance ended up with considerable losses. An example are the thefts of luxury car after the reunification in Germany or the AIDS epidemic. For this also reinsurance may come to the aid.

Stochastic errors in predictions: In some cases premiums are determined on insufficient data or flawed data. In this case also reinsurance allows the insurance to control the extent of the risk.
The advantages of reinsurance for an insurance are thus

1. Reduction of the risk
2. A possible and permissable increase in the volume
3. A reinsurance reduces the necessary reserves and the need to obtain funds on the market
4. The premiums (fluctuation part) can be reduced.

The advantages for the reinsurance are

1. Share of work with the insurer, no administrative work and risk checking for the reinsurer
2. Diversification.

Forms of reinsurance

We have already discussed several forms of reinsurance. Nevertheless we will list the basic forms again

Proportional sharing of risk

Quota share reinsurance: The reinsurer covers certain percentage of the risk. Quite often this form is also practiced as a *coinsurance*, where several insurances each cover a certain part of a large risk

Surplus reinsurance:

The insurer fixes a number c and then the amount

$$Y = \begin{cases} X & X \leq c \\ \alpha X & X \geq c \end{cases} \qquad 0 < \alpha < 1 \tag{11.3.26}$$

is covered by the insurer.

Another form of this reinsurance is defined as follows. On fixes a standard volume v_0 (face value). Then the insurer covers the part $\min(1, v_0, v)$ of the risk, i.e.

$$Y = \min\left(1, \frac{v_0}{v}\right) X$$

e.g. if in a life insurance $v_0 = 1$ Mio. Euro, then the insurer will only cover 30 % of a 3 mi life insurance.

Non proportional sharing of risk

The traditional forms in this domain are the *excess of loss* and *stop loss* reinsurance. The first was already discussed above. The excess of loss reinsurance with *limit* L is defined as

$$Y = \begin{cases} X & \text{if} \quad X \leq a \\ a & \text{if} \quad a \leq X \leq L \\ X - L & \text{if} \quad X \geq L \end{cases} . \tag{11.3.27}$$

Usually the insurer will get a further reinsurer for the part that exceeds L.

This insurance may be applied to single policies in a portfolio, to a subportfolio or to the sum of all claims of a year. In the latter case one speaks of a stop loss reinsurance.

Other forms of reinsurance where already covered in section 7.2. So we will not repeat any of it here.

In general a reinsurance contract for a portfolio (X_i) can be considered as a family of maps $\mathfrak{R} = (R_n)$ from $\mathbb{R}_+^n$ to $\mathbb{R}_+ - n$ is the number of claims so that the following holds: If the actual claimnumber is $n = N$, $R_n(X-1, \ldots, X_n)$ is the part of the claim that remains with the insurer. For simplicity we write now $\mathfrak{R} = R_N(X_1, \ldots, X_N)$, where as usual N is the claimnumber variable.

Example 11.5 1. $R_n(X-1, \ldots, X_n) = \sum Y_i$ with $\varphi_i = \begin{cases} x_i & X_i \leq a \\ a & X_i \geq a \end{cases}$.
Individual XL insurance

2. $R_n(X_1, \ldots, X_n) = c \sum X_i 0 < c < 1$ quota reinsurance.

3. $R_n(X_1, \ldots, X_n) = X_{(1)} + \cdots + X_{(n-k)}$
Where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ (order statistics). The reinsurer covers the k largest claim.

Definition 11.1 A reinsurance contract is *collective* for a portfolio $(X_i)_{i=1}^n$ if there is a function r so that

$$R_n(X_1, \ldots, X_n) = r(X_1 + \cdots + X_n).$$

The following result shows that in general collective contracts are better than individual contracts.

Theorem 11.2 Let $\mathfrak{R} = R_N$ be a given reinsurance contract. Then the set of all reinsurance contracts with the same expectation value as $\mathfrak{R}$ contains a collective contract $\tilde{\mathfrak{R}}$ with

$$\sigma^2(\tilde{\mathfrak{R}}_N(X_1, \ldots, X_n) \leq \sigma^2(R(X_1, \ldots, X_N)) \tag{11.3.28}$$

$$\sigma^2\left(\sum^n X_i - \tilde{R}_N\right) \leq \sigma^2\left(\sum X_i - R(X_1, \ldots, X_N)\right).$$

The proof is based on conditional expectations with respect to $S = \sum^N X_i$ the total claim sum and Jensens inequality. The result means that collectice reinsurance contracts lead to smaller risks for the insurer and reinsurer, if risk is interpreted as the variance, provided the net premiums coincide

$$E(R_N) = E(\tilde{R}_N).$$

The optimality of the stop loss contract is characterized by

Theorem 11.3 Among all collective reinsurance contracts with

$$E(S - R(S)) = c > 0 \qquad X = \sum^N X_i$$

the stop loss contract R_0 is optimal in the sense

$$\sigma^2(R_0(S)) \leq \sigma^2(R(S)).$$

We have seen the proof of this already in 7.2. This result is plausible, because the stop loss contract eliminates all large risks, and exactly these contribute most to the variance.

Similarly we had seen that the quota share contract minimizes the variance of the claimpart of the reinsurer.

This shows that the stop loss contract and the quota share contract are singled out by optimality criteria. The individual XL contracts are also optimal in a sense, because they reduce the risk for large individual claims and thus the ruin probability.

These considerations are rather one sided though, because a reinsurance is a redistribution of risk between insurer and reinsurer. Thus both points of view should be taken into account simultaneously. Looking at it this way risk exchange and redistribution of risk can be considered as a nonzero sum game, where one looks for Pareto optimal solutions. We will not pursue this any further here.

Premiums for the Stop Loss contract

Let $S = \sum^N X_i$ and let P be the deductible. Then the net premium for the reinsurer is

$$\mu_2 = E((S-P)_+) = \int_P^\infty (s-P)dG(s) = E(S) - P - \int_0^P (s-P)dG(s)$$

where G is the distribution function of S. Partial integration now gives

$$\mu_2 = E(S) - \int_0^P (1-G(s))ds. \tag{11.3.29}$$

Thus for a computation of μ one needs the distribution of S. The variance for the reinsurer is

$$\sigma^2((S-P)_+) = \int_P^\infty (s-P)^2 dG(s) - \mu_2^2. \tag{11.3.30}$$

For the insurer we get for the net premium

$$\mu_1 = \int_0^\infty \min(S,P)dG(s) = \int_0^P SdG(s) + P\cdot(1-G(P)). \tag{11.3.31}$$

Thus one considers the function

$$e(P) = \frac{(\int_0^P SdG(s) + P(1-G(P)))}{E(S)}. \tag{11.3.32}$$

Problem 11.11 Prove the following properties of e

1. $e(0) = 0, \quad e(P) \to 1$ for $P \to \infty$
2. $e'(P) = 1 - G(P)/E(S)$
3. $e''(P) = -G'(P)/E(S) \leq 0.$

Problem 11.12 Compute e for the case of Poisson distributed claim numbers and exponentially distributed claim sizes.

Problem 11.13 Compute the function e for a log-normal distributed variable X as a function of the deductible P.

Problem 11.14 Compute the splitting up of the net premium between insurer, reinsurer and a second reinsurer, when X is exponentially distributed, and the first reinsurer has an XL contract with deductible P and limit L. The second reinsurer has an XL contract with deductible L.

Problem 11.15 Compute the distribution for an individual XL contract with deductible P, when the claim number is Poisson and the claim size exponentially distributed.

Two problems remain now:

1. How can we choose an optimal deductible or an optimal quota and

2. How do we split up the premium between insurer and reinsurer.

The first problem depends to a large extent or the size of the insurance. The choice of an optimal P depends on the reserves, the ability and cost to muster additional reserves versus the cost of reinsurance. Clearly there is no simple mathematical recipe for choosing P.

Similar considerations apply to 2. However, the so called "Burning Cost Method", which was originally used for fire insurance is used a lot. With this method the net premium of the reinsurer, without fluctuation term, as part of the full premium is given by

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{N_i} (X_{ij} - P)}{\sum_{i=1}^{n} P_i}.$$

Here n is the number of years, P the deductible N_i the number of claims above P, X_{ij} the size of the j-th claim exceeding P in year i and P_i the total premium income in year i.

With this method one assumes therefore a constant premium ratio for the reinsurer, which is estimated from experience of the last n years.

Since this is not always satisfactory variants of this methods are used too.

Literature:

Kremer: Lecture notes insurance mathematics, Univ. of Hamburg.

T. Mack: Schadensversicherungsmathematik, Schriftenreihe Ang. Vers. Math. Vol. 28, 1997.

11.4 IBNR – Reserves in Non-life Insurance

In life insurance the main function of reserves is to cover future liabilities. In non-life insurance part of the reserves, the fluctuation reserves, are used to spread the risk of fluctuations in time. But there are also reserves needed for future liabilities, even if the contracts usually run on a year to year basis. In accident insurance and car liability insurance, claims are usually reported immediately, but it may take years before the full extent of damage is known. In product liability it may even take years, before damage is discovered, but as in fire insurance the settlement process may take time. For all these types of contingencies actuaries have developed a kind of shorthand notation.

Types of claims

IBNR	Incurred but not reported
IBNER	Incurred and reported but not enough reserved
IBNS	Incurred but not settled
RBNS	Reported but not settled

Leaving aside catastrophe reserves for excessive claims, we shall now concentrate on IBNR reserves. In contrast to single claim reserves long term reserves for a whole portfolio are determined by statistical methods. In doing so, one has to take into account inflation and claim inflation, which is a type of superimposed inflation resulting from changes in law or repair. For simplicity we will assume that our data are trend corrected.

One writes now S_{ik} for claims which arose in year i and were settled in year $(k-1)$ after they occured. Let us assume that it takes n years to settle all claims then we get the following triangle of settlement, where n also denotes the present year.

$$\begin{array}{ccccc}
S_{11} & S_{12} & \dots & S_{1,n-1} & S_{1,n} \\
S_{21} & S_{22} & \dots & S_{2,n-1} & \\
\vdots & & \cdot^{\cdot^{\cdot}} & & \\
S_{n-1,1} & S_{n-1,2} & & & \\
S_{n,1} & & & &
\end{array}$$

The right diagonal thus denotes all claims, which have to be settled this year. In principle one could write down separate triangles for reported and settled claims where the latter is clearly more accurate. We will not make this distinction here.

A similar triangle can be setup for the number of claims N_{ik} and the volumes (face values).

Let now

$$C_{ik} = S_{i1} + \cdots + S_{ik} \tag{11.4.33}$$

be the accumulated claims from year i up to settlement year k. In the remainder we will denote the estimators for all quantities by and overhead $\wedge$. Thus

$$\hat{R}_i = \hat{S}_{i,n+2-i} + \cdots + \hat{S}_{i,n} \tag{11.4.34}$$

is the estimated reserve for year n need for claims from year i. The idea is now to extend experience from previous years into the future. The basic method to do this is the so called *chain ladder method.* In this method one writes

$$C_{i,k+1} = f_K C_{ik} \tag{11.4.35}$$

and assumes that the f_k are independent of i, the year, when the claims arose. One thus assumes that the structure of the claims and the way they are settled does not change too much. In addition it should be noted that the f_k will be close to 1 for large k, because most claims are settled early. On the other hand one should also consider that larger claims usually take time to settle. Implicitly we assume

CL 1: $E(C_{i,k+1} \mid C_{i,1}, \ldots, C_{ik}) = C_{i,k} f_k$

CL 2: The claims from different years are independent.

With (11.4.35) one gets

$$C_{i,n} = C_{i,n+1-i} f_{n+1-i} f_{n+2-i} \cdots f_{n-1} \tag{11.4.36}$$

and

$$R_i = C_{i,n} - C_{i,n+1-i} = C_{i,n+1-i}(f_{n+1-i} \cdots f_{n-1} - 1) \tag{11.4.37}$$

or

$$\hat{R}_i = C_{i,n+1-i}(\hat{f}_{n+1-i} \hat{f}_{n+2-i} \cdots \hat{f}_{n-1} - 1). \tag{11.4.38}$$

If one assumes independence of the S_{ik}, which is a somewhat doubtful assumption, one has

$$\hat{f}_k = \frac{\sum_{i=1}^{n-K} C_{i,k+1}}{\sum_{i=1}^{n-K} C_{i,k}} \qquad k = 1, 2, \ldots, n-1. \tag{11.4.39}$$

The chain ladder method has a number of disadvantages and advantages.

Advantages: The method is easy to apply. In particular models, e.g. Poisson distributed claim numbers, the method agrees with the maximum likelihood estimators.

The $\hat{f}_k$ are unbiased and uncorrelated.

Disadvantages: The pattern of occurrence and settlement of claims should not change. The method may be unstable, e.g. $R_i = 0$ if $C_{i,n+1-i} = 0$.

f_{n-1}, is least reliable, because it is based on one datum only. This also applies to $C_{n,1}$.

In order to alleviate this disadvantage, one may replace unreasonable $C_{n,1}$ by

$$v_n \cdot \frac{\sum_1^n C_{i,1}}{\sum_1^n v_i} = \tilde{C}_{n,1} \tag{11.4.40}$$

where v_i is the volume (premium) = sum of face values of year i. In this case other implausible $C_{i,1}$ may be replaced by $C_{i,n+1-i}/\hat{f}_1 \ldots \hat{f}_{n-i}$. This then leads to the estimate

$$\hat{r}_1 = \frac{(\sum_{i=1}^n C_{i,n+1-i})}{\sum(\hat{f}_1 \ldots \hat{f}_{n-i} v_i)}$$

for the first year claim ratio. With $\hat{r}_1$ the estimate for $C_{n,1}$ becomes $\hat{r}_1 v_n$.

Similarly one may replace other unreasonable $C_{i,2,\ldots}$ by such more robust values. This method is known under the name of *Cape Cod technique*. Since the ratios f_{n-1}, f_{n-2}, f_{n-3}, f_{n-4}, ... depend only on few data points, one may also smoothen the $\hat{f}_{n-k}$ by an Ansatz of the form

$$\hat{f}_k = 1 + ae^{-bk} \qquad \text{for } k \geq n-4.$$

Remark. If one computes the reserves by (11.4.36), (11.4.37) one should actually take into account the interest and compute the present value.

Example 11.6

Year of settlement of occur						Akk. Claims					Reserv.	Pr.
	1	2	3	4	5	1	2	3	4	5	Reserv.	Pr.
1	50	60	30	20	10	50	110	140	160	170		200
2	60	60	25	20		60	120	145	165		10.3	220
3	80	70	40			80	150	190			40.2	280
4	70	60				70	130				66,9	250
5	50					50					74,6	300

$$\hat{f}_4 = \frac{1,7}{1,6} = 1.0625 \quad \hat{f}_3 = \frac{325}{285} = 1.14-35 \quad \hat{f}_2 = \frac{475}{380} = 1.25 \quad \hat{f}_1 = \frac{510}{310} = 1.645.$$

Since the first value $S_{5,1} = 310/1250 \cdot 300 = 74,4$ and obtain a reserve $\hat{R}_5 = 111$.

If inflation for the last years is known and can be predicted reasonably into the future, it is not necessary to correct the data for trend. One can then work directly with the original claim triangle and adjust the reserves. Since this does not give any new insights, we will not follow this up.

Another multiplicative method is the

Payment Ratio Method

With this method the $\hat{f}_k$ are estimated by

$$\hat{f}_K = \frac{1}{n-k} \sum_{1}^{n-k} \frac{C_{i,k+1}}{C_{i,k}}. \tag{11.4.41}$$

Example 11.7 The same data as above give

$$\hat{f}_4 = \frac{1,7}{1,6} \quad \hat{f}_3 = 1,1404 \quad \hat{f}_2 = 1,249 \quad \hat{f}_1 = 1.983.$$

Problem 11.16 Apply the chain ladder technique to

	1	2	3	4	5	Premium
1	20	30	30	20	20	100
2	40	30	50	0		110
3	30	30	30			120
4	50	40				130
5	10					140

Problem 11.17 Apply the payment ratio method to this example.

Problem 11.18 Apply the Cape Cod method to the example above.

Problem 11.19 Apply the Cape Cod method to the system of Problem 11.16.

As a final example we consider the method of independent claim ratios.

Independent Claim Ratios

In this case one assumes that the claim ratios

$$s_k = \frac{S_{in}}{v_i} \tag{11.4.42}$$

are independent of i. In this case the best estimate for s_K is

$$\hat{S}_k = \frac{(\sum_{i=1}^{n+1-k} S_{in})}{\sum_{i=1}^{n+1-k} v_i} \tag{11.4.43}$$

so that

$$\hat{R}_i = (\hat{S}_{n+2-i} + \cdots + \hat{S}_n)v_i. \tag{11.4.44}$$

Problem 11.20 Apply this method to both examples above.

All this shows that there are a number of plausible estimates for the IBNR reserves. In particular circumstances these estimates can be improved. For example one might use also data going back more than n years, or use results from similar portfolios. Finally credibility theory might be used to improve the estimates. Altogether it is up to the actuary to come up with reasonable estimates for the reserves.

12. Asset-Liability-Management

12.1 Investment

With an insurance contract the insured entrust their premiums to the insurance to be used in times of need. The largest part of these premiums represent therefore a liability and have to be used so that these liabilities can be met. Apart from reinsurance the main tool for this are the various forms of reserves. Insurance companys usually invest their money to get higher returns. In savings accounts we will usually get only about 3 to 4%, while insurances have managed to net between 7% and 8%. An investment can be characterized by severed parameters

1. Volume
2. Returns
3. Duration, Availability, liquidity
4. risk
5. transaction costs, marketability
6. legal constraints, taxation

A number of general points can be made regarding theses parameters

1. Risk and return are usually inversely correlated, i.e. high risk investments usually give the highest returns, but possibly also losses. In fact the numerous investment scandals in recent years are simple evidence that even seasoned bankers may loose large sums.

2. Long form commitments yield higher returns. But they carry the risk of inflation and change of economic parameters.

3. Insurances will have to plan their investments according to the structure of their liabilities. Thus life insurances or pension insurances with long term commitments will prefer long term relatively secure investments with little cash for immediate needs. Non-life-insurances on the other hand, have to be able to obtain cash on a short term notice.

4. In particular with foreign investments the transaction cost (exchange costs) play an important role. Low marketability usually results in higher transaction costs. This may hold for property.

5. Various insurance laws restrict the investment politics of an insurance and usually enforce a form of diversification for higher security.

Forms of investment

Cash deposit Immediately avialable, low return, very secure, no risk apart from inflation.

Government bonds Committed for a fixed number of years. Very secure, danger of inflation. The price P, dividend D, duration n (years) and interest rate p for one currency unit are related by

$$P = D \cdot a_{n\rceil} + v^n \qquad \begin{array}{l} \text{where } v = (1 + p/100)^{-1} \\ a_{n\rceil} = v + v^2 + \cdots + v^n \end{array} \tag{12.1.1}$$

In this formula v^n is the present value of one Euro, due in n years, while the first summand describes an annuity D for n years.

Because of the risk of inflation some governments have also issued index linked bonds. With government bonds from foreign countries one has to take into account also transaction costs and inflation in the particular country.

Equities

Equity investment is the major method for insurance companies for their operations. Since equities can be traded, equity investments can be available immediately, but possibly with losses. Since firms tend to grow with the economy, equities are generally secure from inflation. The greatest risk with equities arises in care of insolvency of the firm or overvalued equities at times of purchase. So this type of investment requires a careful analysis of the financial prospects of the firm in question. Currently equities in electronic computer and biotech industries tend to be overvalued, i.e. the actual rate does not correspond to the proper value of the firm. The result can be short term fluctuations in the stock index.

Property investment Property, i.e. land or buildings are generally difficult to trade and they have to be managed in order to get the proper return in form of rent. Berides the rent the increase in value is the major aspect of returns. For property investment legal side conditions and taxes play an important role. Generally property is difficult to value but for long tences investment this is considered rath or secure.

Derivatives (futures, options, swaps and currency forwards) These are financial instruments to control exposure to a particular investment market. They are therefore mainly used as a protection against the volatility of investment (risk hedging). Thus options and futures may by tied to the Dow Jones index. This offers protection if the own investment portfolio correlates to this index. Options can also be used to prepare for an investment (long call options). Currency forwards are used for hedging foreign exchange risks. These are essentially due to changes in the exchange rate.

Literature

P. Booth, R. Chadburn, D. Cooper, S. Haberman, D. James: Modern actuarial theory and practice. Chapman Hall 1999

The mathematics of derivatives is a highly interesting al beit complex topic of insurance mathematics. The article by H. Föllmer "'Zur Versicherungsmathematik von Derivaten"'. Der Aktuar 3 1997 158-161 gives an overview of this.

Problem 12.1 An equity provides you with dividends of, which grow at a constant rate g. Then the value V_0 is given by

$$V_0 = d_1 \sum_{j=1}^{\infty} \frac{(1+g)^{j-1}}{(1+i)^j} \qquad i = \frac{p}{100}.$$

Problem 12.2 Simplify this exprersion.

12.2 Asset Liability Management

The premium income and reserves of an insurance are invested. These investments and cash reserves are used to meet the liabilities of the insuran. This requires that investments and liabilities match some how. More precisely asset-liability management aims at maximizing the returns subject to meeting of all liabilities. This requires

1. *Study of the liabilities*
   (a) Model calculation of the given portfolio subject to various conditions for premiums, interest rate, taxation and changes is parameters
   (b) Estimation of the influence of changes in the portfolio
   (c) study of various other scenarios and their influence on gains and losses
2. *Study of the asset side*
   (a) Forming of asset classes
   (b) Analysis of returns and performance in recent years
   (c) Study of various scenarios – best, worst case and most probable case analysis.
3. *Study of dependence of liabilities and assets*
   In most cases, when investment and liabilities are matched, there are dependences, e.g.. Fixed liabilities in life insurances should be matched by fixed money assets of comparable volume and duration.

   For pensions insurances, where the greatest risk lies with inflation, index linked bonds.

   Health insurance liabilities with its rise in cost might be matched by equrities.

In non-life insurance liabilities are usually short term with considerable fluctuations and increase in price. In addition reserves are needed for IBNR claims. This requires a certain amount of cash, a proportion of investment available on a short term basis (a year or so) and long term assets. Develop scenarios and models for such dependence

4. *Optimization of various strategies*

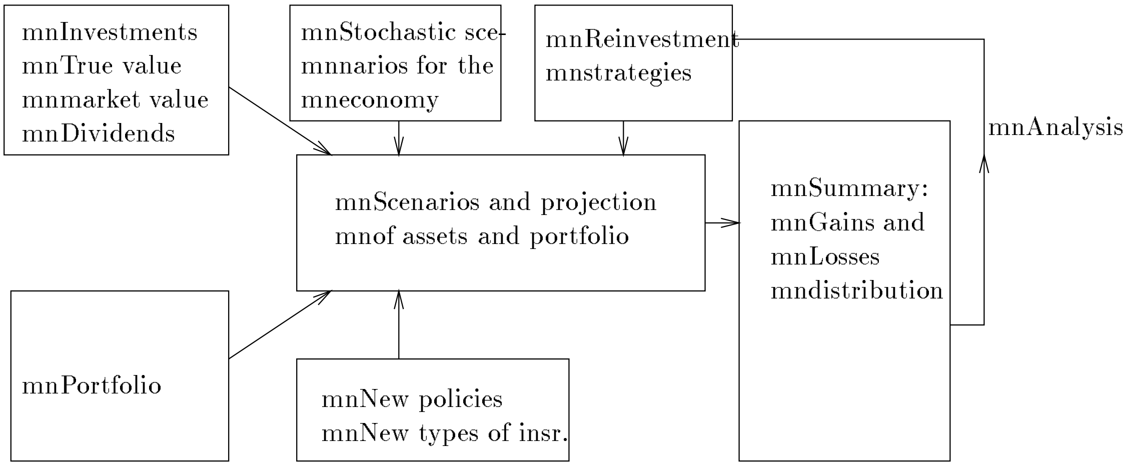
   (a) Model calculation of gains and losses for the portfolio subject to various scenarios for interest rates, economic conditions and probabilities.

   (b) Choose an optimal strategy for investment.

   (c) Choose an optimal strategy for changing and modyfiing portfolio.

   (d) Reinsurance considerations.

These interdependences can be described by the following flow chart.



Below we give an example of such a scenario.

One of the most important methods to reduce the investment risk is *diversification*. Sure enough, the returns are not as dramatic as with some biotech or electronic and computer stocks, but they are a lot less risky. Even if there is no law or rule that relates riskiness and returns, it is rather obvious that risky investments will only performed if they correspond to a higher yield. Here is a simple example that shows how diversification reduces the risk, if this is interpreted as the variance. Let $E(Z_i)$ be the expected return from the i-th investment for 1 currency unit and let V_i be the volume invested there. Let $V = \sum V_i$ be the total volume and $x_i = V_i/V$. Then the investment of the whole portfolio Z yields

$$E(Z) = V \cdot \sum x_i E(Z_i).$$

While the variance is given by

$$\sigma_Z^2 = V^2 \cdot \sum x_i^2 \sigma_{Z_i}^2 + \sum_{i<j} x_i x_j \sigma_{ij}$$

where σ_{ij} is the covariance between Z_i and Z_j. One knows that one has always

$$|\sigma_{ij}| \leq \sigma_i \sigma_j$$

which is a consequence of the Cauchy Schwartz inequality. The interesting case, however, arises if a large part of these investments is uncorrelated, i.e. $\sigma_{ij} = 0$. In this case, or even better if $\sigma_{ij} < 0$ the overall variance is smaller.

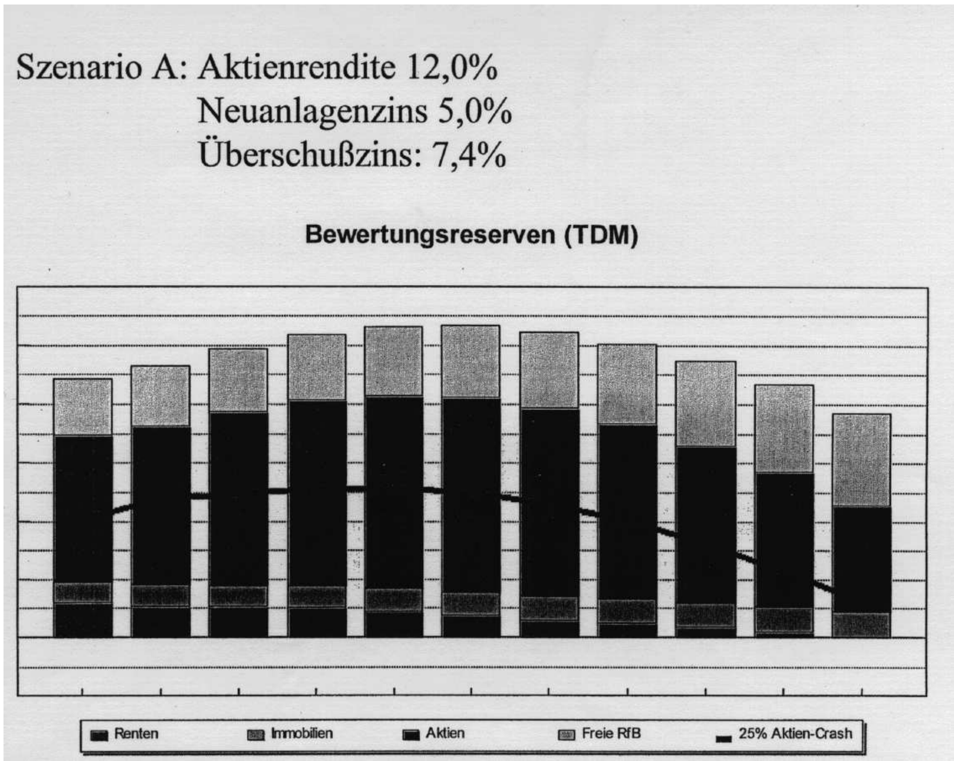
Life insurance The investment politics of a life insurance company will depend to a large degree on the mix of its business. The non-profit business with fixed money terms will require long term fixed interest securities or long term bends. An increasing number of contracts is now tied to funds or equities. So, this fixes the investment too, to some degree. Both these types of contracts need few free reserves. Most difficult to evaluate are contracts which leave the insured options upon maturity of the contract. In addition investment has to be made with an eye upon the incoming stream of premiums and new contracts. Altogether life insurance companies will be relatively conservative in their asset allocation, because their customers expect this. So this also requires a high degree of diversification.

Non-life insurances The liabilities are generally short term with larger fluctuations. So, within the EU inflation is not much of an issue. Cash reserves are important. So are short term fixed interest securities. Equities give uncertain long term returns and are rather volatile, so non-life insurances tend to invest less in equities than life-insurances.

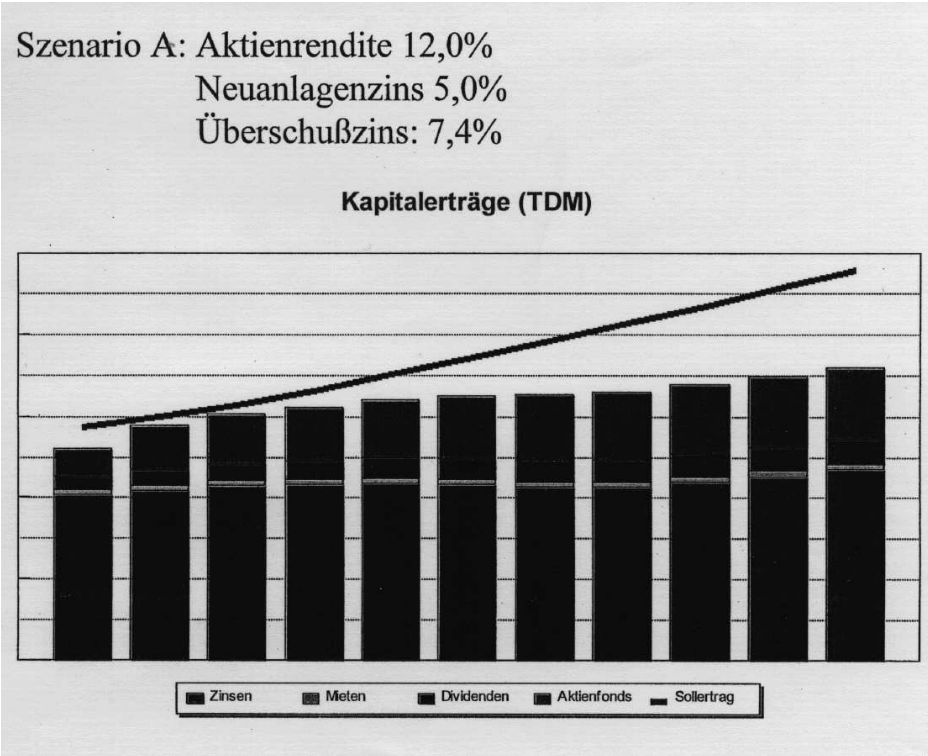
All these assessments that are performed to describe the economical side of an insurance are finally summarized in terms of the *embedded value* of the company. This is the value of the company for the stockholders and it is made up of the present value of all future gains plus the capital of the company. The embedded value can be used to follow the success of the company and may be used to compare the efficiency of several companies, though this is problematic. In order to compute the present value one has to compute all policy classes (several hundred) with actual (not modified) parameters including realistic interest rates and on this basis evaluate all policies (usually about 10^5 to 10^6). In a nutshell the embedded value gives the stockholdess an idea what their company is worth.

Asset-Liability model

Scenario: yield of stocks 12 %, yield of new investments 5 % interest or surplus.



Gains (1 000 DM) Reserves (in 1 000 DM)



Reserves (in 1 000 DM)