University of Tartu
Faculty of Science and Technology
Institute of Mathematics and Statistics

NON-LIFE INSURANCE MATHEMATICS (MTMS.02.053)

Lecture notes

Lecturer: Meelis Käärik

Contents

1 Introduction. Definition of risk. Insurable risk

Almost every human activity is related to some risk(s). When planning a picnic, there is a risk that it will rain, when ordering a theater ticket, there is the risk that the performance is sold out, when driving a car, there is a risk of a traffic accident. There is endless amount of such examples, there are various risks that may influence us all the time. However, turns out that defining a risk itself is not that simple and obvious task. Although people mainly think in a similar area, there is no unique definition that includes all specifics of a risk. In the following we give few possible definitions:

- risk is a combination of threats;
- risk is a probability of something unpleasant to occur;
- risk is the uncertainty of loss;
- risk is the probability of loss;
- risk is uncertainty, tendency that the reality will differ from expectations;
- risk is a possibility of an unwanted negative outcome (which may be known).

Let us assume now that we have established a common understanding about the essence of a risk. The obvious question is how to deal with the risks. In general, there are four basic ways how individuals deal with risk:

1. Assumption, acceptance – a decision is taken that the level of risk is acceptable and no action is taken. For example, a cost benefit analysis of the possible alternatives could conclude that the most efficient solution is to take no action.
2. Elimination – all hazards to which one is exposed are removed. This is not always possible and can often have unpredictable side-effects. For example, pesticides can be used to eliminate the risk of crop failure, but they migh then pollute the environment.
3. Avoidance – behaviour is modified in order to avoid the undesirable exposure. For example, car is parked only in a secure garage to avoid the theft risk.
4. Transfer – risk is transferred to third party. This is the basis of insurance (and reinsurance) contracts.

Many risks involve economical factor and have financial consequences (i.e. measurable in monetary units). Such risks can also be divided into:

1. Speculative risk (dynamic risk) – either profit or loss is possible. Examples of speculative risks are betting, gambling, investing in stocks/bonds and real estate. Speculative risk is uninsurable as it violates the most fundamental concept of insurance – the insured should not gain from insurance.

2. Pure risk (absolute risk, static risk) – there is a chance of either loss or no loss, but no chance of gain; for example, either a building will burn down or it won't. Only pure risks are insurable.

Which criteria describe an insurable risk?

- the outcome must be financial, i.e., it involves a loss in value that can be measured in monetary units;
- the risk must be pure risk, i.e. the insured can not gain from it;
- fortuity, i.e. the events causing the loss must arise due to chance, the occurrence, timing, severity are not under the control of the insured (policyholder);
- the frequency and severity of the possible loss must be measurable;
- the probability of the occurrence must not be too high;
- the circumstances of the loss event must be clearly definable;
- the price of the transfer must be reasonable (in general).

Definition 1.1 (Insurance, I)**.** Insurance is a way of buying one off from the economical consequences of possible risks.

Definition 1.2 (Insurance, II)**.** Insurance is a way of redistributing the society's assets, which (in case when the party that suffered loss has an insurance policy) will help the suffered party, covering their loss on the credit of those policyholders who did not suffer the loss.

The insurance product is quite different from common "physical" products one can buy. It is possible (and even common) that no loss will occur during the insurance period, nevertheless, if a loss occurs, the insurer must cover the claim as specified by the contract. The occurrence of a loss event depends on many different factors. In non-life insurance one can make difference between risk factors and rating factors. Risk factors are those factors which are believed to influence directly the frequency or severity of a claim for a given risk exposure. Such factors are often difficult to obtain or measure reliably. Therefore, insurance companies gather information of related factors, which are easier to measure and manage. For example, traffic density in which a car is driven is clearly a significant risk factor, but it is very difficult to measure the traffic density accurately for all the routes of all policyholders. Some crude estimates that can be used here are the policyholder's address and also the purpose for which the vehicle is used.

In non-life insurance there is also a variety of different types of policy. This means that many different measures are required to describe risk exposure. For example, in motor insurance the usual measure of exposure is the vehicle year. However, for some types of risk (e.g. traffic accidents), the distance driven might be a better exposure indicator. On the other hand, distance driven does not reflcet properly the exposure to theft claims as they occur when the car is not being driven. Another question of importance for practical point of view is how accurately these measures can be obtained. Using previous example, the vehicle year is straightforward and requires no additional calculations. The distance driven, although theoretically reasonable, is rarely used in practice because of the difficulties in actual calculations.

The likelihood of a policyholder claiming against his or her insurance policy obviously depends on risk factors mentioned above. Another way of classifying the circumstances which make claim more or less likely is to identify the hazards to which a policyholder is exposed. In the following we give a broad classification of such hazards.

1. Legal hazards. These include changes in laws or imposition of new conditions under which insurers could become liable (especially under retrospective legislation).

2. Physical hazards. Physical or structural conditions which could increase the likelihood of loss. For example, faulty house wiring or outdated safety systems.

3. Moral hazards. Deviation from normal behaviour of the policyholder in order to gain financially from the insurance contract by taking certain actions within his or her control. In other words, the principle of fortuity is violated.

4. Personal hazards. Personal hazards are closely related to moral hazards. For example, people might be careless, or badly qualified, or just more accident prone than others and therefore impose an above average liability on the insurer.

To avoid or reduce these hazards, several rules and regulations are used. For example:

- the principle of deductible: risk is not transferred fully, the policyholder is still responsible for a small part of the risk;
- encouragement to increase security and safety levels to reduce physical risks, e.g. discounts of insurance premium if anti-theft devices are fitted;
- no cover for claims if they are caused while intoxicated by alcohol or under the influence of drugs, i.e. reducing personal risks;
- the conditions when and how much of the loss will be compensated need to be precicely fixed in the insurance contract, thus reducing possible moral risks.

In order to assure the policyholder's trust in insurers and to guarantee the solvency of insurers policyholders, the regulations are fixed by laws and strictly supervised by appropriate institutions:

- in European level – Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS);
- in Estonia – Financial Inspection.

It is also important to note that the current regulation regime Solvency I is being soon replaced by more dynamic and flexible regime Solvency II.

References

1. Booth P., Chadburn, R., Cooper, D., Haberman, S., James, D. (1999) *Modern Actuarial Theory and Practice.* Chapman & Hall / CRC.
2. Sandström, A. (2010) *Handbook of Solvency for Actuaries and Risk Managers: Theory and Practice.* Taylor & Francis.
3. Sundt, B. (1993) *An Introduction to Non-Life Insurance Mathematics.* VVW, Karlsruhe.

2 The cash-flow model of an insurance company

2.1 Transition equation

The financial operations of an insurer can be viewed in terms of a series of cash inflows and outflows. The inflow components are added to the reservoir of assets, while the reservoir is depleted by the outflow components.

Main inflow components for an insurance company are:

- premiums – the main income of an insurance company;
- reinsurance recoveries – an insurer may also transfer some risks or parts of risk further to a reinsurer, in this case the corresponding claims are also (partially) recovered by a reinsurer;
- income from investments – this includes interest payments, dividends, rental income, changes in value of assets;
- new capital issued and subscribed for;
- miscellaneous.

Main outflow components for an insurance company are:

- claim payments – the main outflow component;
- reinsurance premiums;
- expenses – includes commission paid, administration and operating expenses, we may also include taxes in this term;
- dividends paid to shareholders and bonuses paid to policyholders;
- miscellaneous.

It is obvious that the first two components in both lists are characterizing the insurance business while the other components are general and do not depend on company's business.

Let us now introduce some mathematical notation so we can write the whole cash-flow model as certain transition equation.

Notations for inflow (in period $[0, t]$) are following:

P_t – the premium income;

X_t^{Re} – recoveries from reinsurance;

I_t – the return from investments;

U_t^{new} – new capital issued and subscribed for;

and notations for outflow (in period $[0, t]$) are:

X_t – claims;

E_t – paid commission, administration and operating expenses;

P_t^{Re} – ceded (reinsured) reinsurance premium;

D_t – dividends paid to shareholders.

Let A_t denote the assets of the insurer at time t. Then, the flows and resulting assets can be expressed in the form of transition equation:

$$A_t = A_0 + P_t + X_t^{Re} + I_t + U_t^{new} - X_t - E_t - P_t^{Re} - D_t. \tag{2.1}$$

Model (2.1) is useful in various situations and depending on particular need, the terms and whole model can have different interpretations. Firstly, we can use it either as a discrete time model or a continuous time model. Since most of the reporting and revisions are required on annual basis, a discrete time model is justified to describe the yearly development. If one wants to monitor the cash-flow more precisely, a continuous time model is better suited.

Although we stated the model as referring to the whole operation of an insurer, the same principles can be applied to establish sub-models. For example one may want to concentrate on a narrower context and examine the development in certain subportfolio. One must be noted that application of the model to a subportfolio may present some problems of interpretation as to which assets are allocated to that subportfolio (but that does not change the fundamental principle).

The model can also be interpreted as either deterministic or stochastic. Many interesting and useful applications may be possible using deterministic model. However, modelling the uncertainty is one of the main challenges in risk theory and stochastic model clearly depicts the true essence of the situation.

From the insurance perspective it is also important to make distinction between 'paid' amounts or 'earned' and 'incurred' quantities. Depending on the quantities we use, the meaning of the model also changes.

Let us make the distinction

- P'_t – earned premium in period $[0, t]$;

- P_t – written premium in period $[0, t]$.

Written premium is the premium charged (or to be charged) for a policy or group of policies and it is (usually) fixed while signing the contract.

Earned premiums for an accounting year are those parts of premiums written in the year, or in previous years, which relate to risks borne in that accounting year. In so far as premiums written during the accounting year provide cover for risk in the next or subsequent accounting years, the part of the premium relating to those later periods is carried forward by establishing the unearned premium reserve (UPR).

This construction can be written as

$$P'_t = P_t - UPR_t + UPR_0,$$

where UPR_t is the unearned premium reserve at t (end of the accounting period) and UPR_0 is the initial unearned premium reserve at time 0 (beginning of the accounting period).

Similarly to premiums, we make distinction between

- X'_t – incurred claims in period $[0, t]$;
- X_t – paid claims in period $[0, t]$.

Incurred claims in the accounting year are defined as the total amount of claims arising from events which have occurred in the year (irrespective of when final settlement is made!). It should be noted that the actual settlement of some claims may be delayed considerably beyond the year in which the event giving rise to the claim occurs. This means that the *claims paid* will include amounts in respect of claims incurred in earlier accounting years, which should have been included in the reserve for reported but not settled claims (RBNS) brought forward from the previous accounting period.

In other words, the following relation between incurred and paid claims holds:

$$X'_t = X_t + RBNS_t - RBNS_0,$$

where $RBNS_t$ is the reserve for outstanding claims at t and $RBNS_0$ is the reserve for reported but not settled claims at time 0.

2.2 Risk reserve and solvency

Now, let us go back to model (2.1) and notice the following:

- X_t^{Re} ja P_t^{Re} are strongly related to process X_t and determined by reinsurance mechanics;

- P_t is determined by the claims process X_t and many other factors (market situation, marketing politics, global economical situation, legal regulations, human psychology, etc);
- investments, expenses, dividends are not related to the essence of insurance, thus they can be omitted from the model and dealt with separately.

Taking this remark into account we obtain a simplified surplus process U_t:

$$U_t = U_0 + P_t - X_t, \tag{2.2}$$

where U_0 is the initial surplus (at $t = 0$), P_t is written (or earned) premium during $[0, t]$ and X_t denotes paid (incurred) claims during $[0, t]$.

The model can be simplified further by considering the process P_t to be linear in time: $P_t = P \cdot t$.

The quantity U_t is also called *solvency margin* or *risk reserve*.

Definition 2.1 (Absolute solvency)**.** An insurer is said to be absolutely solvent if its liabilities do not exceed its assets, in other words $U_t \geq 0$

The absolute solvency is not the best criterion to use in practice, because it is too rough and it is too late to take any action if an insurer is already insolvent. Therefore, in current solvency I regime, there are two margins that specify when the regulatory authorities should take action: *required solvency margin* U_{min} and *minimum guarantee fund* U_{mgf}. In case the solvency margin of an insurance company falls below the required margin, i.e., $U_t < U_{min}$, then the supervision institutions may apply sanctions in order to save the investments of the policyholders and the shareholders. The solvency margin should never fall below a minimum guarantee fund, which is the absolute minimum amount of capital required.

Let us consider the expectations of different parties related to the solvency of the insurer. Policyholders' main interest is that the insurer stays solvent:

$$\mathbf{P}\{U_0 \cdot (1 + i) + P - X < U_{min}\} = \varepsilon,$$

where i is risk-free interest rate and $\varepsilon > 0$ is small.

Insurer's (or shareholders') on the other hand want to earn profit:

$$\mathbf{P}\{U_0 \cdot (1 + i) + P - X < U_0 \cdot (1 + i + j_{min})\} = \delta,$$

where j_{min} is the required return rate and $\delta > 0$ is small (but usually $\delta > \varepsilon$).

It can be seen that the solution to solvency equations above is the given by the following equalities.

$$U_0 = \frac{F^{-1}(1-\varepsilon) - F^{-1}(1-\delta) + U_{min}}{1+i+j_{min}}, \tag{2.3}$$

$$P = F^{-1}(1-\delta) + j_{min}U_0. \tag{2.4}$$

Formulas (2.3) and (2.4) guarantee to the shareholders capital return of at least j_{min} with probability $1-\delta$.

The obvious remaining problem is how to find $F(x)$. One possible simple approach is so-called Normal Power approximation:

$$F^{-1}(1-\varepsilon) \approx EX + z_{1-\varepsilon}\sqrt{VarX} + \frac{z_{1-\varepsilon}^2 - 1}{6} \cdot \frac{E(X-EX)^3}{VarX},$$

where $z_{1-\varepsilon}$ is the (1-ε)-quantile of standard normal distribution.

2.3 Managing insurance risk: risk pooling

Let us denote

- $P^{(1)}$ – individual insurance premium;
- n – number of contracts;
- P – whole premium for the period, $P = nP^{(1)}$;
- ω – a proportion factor, $U_0 = \omega P$;
- X – total claims.

Then we can write an even more simplified formula for the surplus process at $t = 1$:

$$U_1 = U_0 + P - X = \omega P + P - X = (\omega + 1)P - X,$$

Shareholders' perspective in this process can be measured in terms of the *return of capital*. The return of capital R is defined as

$$R = \frac{P-X}{U_0} = \frac{1}{\omega}\left(1 - \frac{X}{P}\right)$$

with expected value

$$ER = \frac{P-EX}{U_0} = \frac{1}{\omega}\left(1 - \frac{EX}{P}\right).$$

On the other hand, policyholders' perspective in this process is to minimize the *cost of insurance*. The (relative) cost of insurance L is defined as

$$L = \frac{P - X}{P} = 1 - \frac{X}{P}$$

with expected value

$$EL = \frac{P - EX}{P} = 1 - \frac{EX}{P}.$$

The characteristics R and L are connected through the proportion factor ω: $R = \frac{1}{\omega}L$ and $ER = \frac{1}{\omega}EL$

The probability of insolvency can be reduced by increasing the number of independent risks insured, n. Using the law of large numbers we get

$$\lim_{n\to\infty} \mathbf{P}\left\{\left|\frac{X}{n} - \frac{EX}{n}\right| < \varepsilon\right\} = 1.$$

Thus, we can characterize the *pooling of risk* principle, which underlies all insurance:

- the more contracts (the bigger n), the smaller can be the relative cost of insurance (as the actual claim amount is close to expected with high probability);
- the proportional factor can be decreased to increase the capital return R;
- the more contracts the less capital (relatively) is required to obtain sufficient capital return and acceptable cost of insurance.

Remark 2.1. It must be noted, however, that where there is heterogeneity amongst the risks, the law of large numbers may not be entirely valid!

Example 2.1. An insurer issues 10 000 identical policies with the following characteristics

- claim size (if it incurs) is 10 000 EUR;
- probability of incurrence is 0.05;
- individual insurance premium is 550 EUR.

The shareholders provide initial capital of 2.5 million Euros, expenses and investment income will be ignored. As the portfolio is large, we assume that the number of claims N is approximately normally distributed with

$$\mu = 10000 \cdot 0.05 = 500,$$
$$\sigma^2 = 10000 \cdot 0.05 \cdot 0.95 = 475.$$

Thus the incoming cash flow is $550 \cdot 10000 + 2500000 = 8000000$ EUR and the probability of insolvency ε can be calculated as

$$\varepsilon = \mathbf{P}\{X > 8000000\} = \mathbf{P}\{N > 800\} = 1 - \Phi(13.765) \approx 0.$$

The expected cost of insurance is $EL = 1 - \frac{EX}{P} = 1 - \frac{5000000}{5500000} \approx 0.09$ and the expected return of capital is $ER = \frac{EL}{\omega} = \frac{0.09}{0.45} = 0.2$.

Remark 2.2. Besides the general pooling principle, several risk mitigation techniques are used in order to manage the insurance risk:

- *diversification* involves accepting risks that are not similar in order to benefit from the lessened correlation of contingent events;
- *hedging* involves accepting risks with a strong negative correlation;
- *reinsurance* means transferring risks or parts of risks to a reinsurer.

References

1. Booth P., Chadburn, R., Cooper, D., Haberman, S., James, D. (1999) *Modern Actuarial Theory and Practice.* Chapman & Hall / CRC.
2. Daykin, C.D., Pentikäinen, T., Pesonen, M. (1994) *Practical Risk Theory for Actuaries.* Chapman & Hall, London.

3 Principles of the compensation calculation. Principles of the deductible

The calculation of compensation relies mainly on the following two characteristics:

- insurable value V – describes the value of the insurable object;
- sum insured S – usually determines the limit of the compensation.

The *insurable value* simply means the value of the insured object. When an insurance claim occurs, one typically calculates the insurable value immediately before the claim event.

In many classes of property insurance, a *sum insured* is specified. In most cases this is an upper limit for the compensation the insurance company will pay in the event of claim.

Both sum insured and insurable value can be given as a fixed sum or as a calculation principle (reinstatement value, replacement value, regular value, etc). If the insurable value and sum insured are both given as fixed sums, these sums regularly need to be equal. Otherwise we are talking about under-insurance (if $S < V$) or over-insurance (if $S > V$).

If a policyholder suffers a loss caused by an event covered by the insurance, then he receives compensation from the insurer. In the following we introduce the main principles used to calculate the compensation.

Let X be the actual loss and let $g(X)$ denote the calculated compensation. One should be noted that the calculated compensation is not yet the compensation that the policyholder actually receives, it is usually reduced by a deductible (this will be described later).

1. Pro-rata principle:
$$g(x) = \min\{1, \frac{S}{V}\} \cdot x,$$
takes into account the specifics of over- and under-insurance. The compensation is reduced proportionally by the ratio between the sum insured and the insurable value.

2. First risk principle:
$$g(x) = \min\{x, S\},$$
the loss is fully covered as long as it does not exceed the sum insured; if it exceeds the sum insured, then the sum insured is covered. The first risk principle is often used when it is difficult to determine an insurable value.

3. Full insurance principle:
$$g(x) = x,$$
is rarely used, the possible compensation is not limited.

In most cases the calculated compensation is reduced by a *deductible.*

There are several reasons for introducing deductibles:

a) loss prevention – to lower the probability of claim occurrence;

b) loss reduction – to lower the claim amount in case of a loss event;

c) avoidance of small claims (as administration cost are dominant when handling small claims);

d) premium reduction – the first three properties clearly simplify the insurer's risk management, in return the insurer can decrease the insurance premium.

Let $h[g(X)]$ denote the actual compensation paid to the claiming policyholder. There are 3 main principles of deductibles that can be applied.

A. Fixed amount deductible b:
$$h_1(g(x)) = \max\{0; g(x) - b\}.$$

B. Proportional deductible β:
$$h_2(g(x)) = (1 - \beta)g(x).$$

C. Franchise deductible d:
$$h_3(g(x)) = I_{\{g(x)\geq d\}}g(x).$$

Remark 3.1. Notice that

- principle A satisfies all requirements a)–d);
- principle B could not avoid the handling of small claims c);
- principle C does not satisfy b) and can work against it.

References

1. Booth P., Chadburn, R., Cooper, D., Haberman, S., James, D. (1999) *Modern Actuarial Theory and Practice.* Chapman & Hall / CRC.

2. Sundt, B. (1993) *An Introduction to Non-Life Insurance Mathematics.* VVW, Karlsruhe.

4 Premium principles

4.1 Desirable properties of premium principles. Classical premium principles

As discussed before, insurance can be considered as transfer of a risk X from policyholder to the insurer. The risk is of stochastic nature, so we can consider it as a random variable. By the construction, we also assume it is a non-negative random variable; risks taking negative values are not realistic in non-life insurance.

The insurance industry exists because people are willing to pay a price for being insured. A natural question related to the insurance process is how much should the insurer ask for transfer of a risk X? In other words, we need to specify some rules to determine a proper premium for a risk X. A premium principle is a rule P that to any risk X assigns a premium $P(X)$. The premium $P(X)$ is non-random and can be considered as a function of the distribution of X.

Before finding some suitable functions P, let us think what properties should such function have. A short list of natural properties is the following:

1) subadditivity: $\forall X, Y,\ P(X+Y) \leq P(X) + P(Y)$;

2) monotonicity: $\forall X, Y,\ P(X) \leq P(X+Y)$;

3) risk loading: $\forall X,\ P(X) \geq EX$;

4) premium is limited by maximal possible loss:
for $\forall X$ the inequality $\mathbf{P}\{X < P(X)\} < 1$ holds;

4') a weaker form of 4):
if $\forall X\ \exists m$ so that $\mathbf{P}\{X \leq m\} = 1$, then also $P(X) \leq m$.

Four commonly used classical premium principles are

(a) expected value principle $P_1(X) = (1+\alpha)EX,\ \alpha > 0$;

(b) standard deviation principle $P_2(X) = EX + \beta\sqrt{VarX},\ \beta > 0$;

(c) variance principle $P_3(X) = EX + \gamma VarX,\ \gamma > 0$;

(d) "combined variational principle" (compromise principle) $P_4(X) = EX + \beta_1\sqrt{VarX} + \gamma_1 VarX,\ \beta_1, \gamma_1 > 0$.

Turns out that the some of the properties 1)-4') are not well satisfied by these classical principles (prove it!).

One can also find premium as a solution for the following condition

$$\mathbf{P}\{X \geq P\} \leq \varepsilon,$$

where $\varepsilon > 0$ is the maximal tolerated ruin probability. In other words, we want to find a premium which ensures that the probability that this premium is sufficient to cover the claims is at least $1 - \varepsilon$. In that case we obtain

(e) the quantile principle: $P_5(X) = \min_P[\mathbf{P}\{X \leq P\} \geq 1 - \varepsilon]$.

The quantile premium calculation principle is often also called the *percentile principle*. No additional assumptions are set to risk X, but the amount of data must be sufficient to give usable estimates for the required quantiles.

4.2 Utility theory

There is an economic theory that explains how much insureds are willing to pay for transferring a possible loss. The theory postulates that a decision maker, generally without being aware of it, attaches a value $u(x)$ to his wealth x instead of just x, where $u(\cdot)$ is called his or her *utility function*. So, all decisions related to random losses are done by comparing the expected changes in utility. Although it is impossible to determine a person's utility function exactly, we can give some plausible properties of it. For example, more wealth would imply a higher utility level, so the utility function $u(\cdot)$ should be a non-decreasing function. It is also logical that "reasonable" decision makers are *risk averse*, which means that they prefer a fixed loss over a random loss with the same expected value. In conclusion, a utility function $u(x)$ is assumed to be increasing ($u'(x) > 0$) and concave, i.e. the relative value of money will decrease while x increases ($u''(x) < 0$).

Maximum premium P that the insured with wealth x is willing to pay is the solution of the following equation:

$$u(x - P) = E[u(x - X)].$$

Similar line of argument applies from an insurer's viewpoint. Then the minimum acceptable premium by insurer with utility function $u(x)$ is found from:

$$u(x) = E[u(x + P - X)].$$

The problem is that, in general, the premium will depend on the insurer's surplus x, which makes it very difficult to apply in practice.

A known solution to last equation is given by function

$$u(x) = \frac{1}{a}(1 - e^{-ax}),\ a > 0.$$

Using this function we can calculate the premium:

$$\begin{aligned}
\frac{1}{a}(1-e^{-ax}) &= \frac{1}{a}E(1-e^{-a(x+P-X)});\\
e^{-ax} &= e^{-ax}e^{-aP}Ee^{aX};\\
e^{aP} &= Ee^{aX};\\
P &= \frac{1}{a}\ln Ee^{aX}.
\end{aligned}$$

Thus, we obtained a formula where the premium does not depend on surplus x anymore. The corresponding premium principle is called

(f) exponential principle: $P_6(X) = \frac{1}{a}\ln Ee^{aX}$.

The parameter $a > 0$ is called *risk aversion*, it represents the 'unwillingness' of the insurer to undertake the risk. It can be proved that exponential premium increases when a increases: the more risk averse one is, the larger premium one is willing to pay.

Turns out that the exponential principle satisfies the properties 1)-4') much better than the classical principles described above (prove it!). Unfortunately it has a quite complex form and it is quite difficult to apply this principle in practice.

Remark 4.1 (Allais paradox (1953))**.** Consider the following capital gains

$$X = \quad 1\,000\,000 \quad \text{with probability } 1$$

$$Y = \begin{cases} 5\,000\,000 & \text{with probability } 0.10\\ 1\,000\,000 & \text{with probability } 0.89\\ 0 & \text{with probability } 0.01\end{cases}$$

$$V = \begin{cases} 1\,000\,000 & \text{with probability } 0.11\\ 0 & \text{with probability } 0.89\end{cases}$$

$$W = \begin{cases} 5\,000\,000 & \text{with probability } 0.10\\ 0 & \text{with probability } 0.90\end{cases}$$

Turns out that having a choice between X and Y, many people choose X, but at the same time they prefer W over V. One can easily see that, assuming an initial wealth of 0, these choices contradict the expected utility hypothesis. It seems that the attraction of being completely safe is tronger than expected utility indicates, and induces people to make irrational decisions.

4.3 A note on terminology

The premiums are used in different contexts and with different meanings (depends on how many factors are taken into account). A short overview is given by the following list.

- *Pure premium*: the expected loss EX;
- *Risk premium*: pure premium EX plus *risk loading* D, where D is calculated depending on chosen premium principle;
- *Gross premium*: risk premium plus *expense loading*;

  $$GP = \frac{1}{1-k} \cdot [P(X) + K],$$

  where k denotes expenses proportional to premium (e.g. commission fees) and K denotes other (unproportional) expenses (e.g. expenses related to claim handling and adjustment);
- *Net premium*: by this we usually mean net of reinsurance (and agent's commissions or other related costs), i.e. either risk premium minus payments made for reinsurance or gross premium minus payments made for reinsurance, depending on situation.

We are mainly speaking of risk premiums unless specified otherwise.

References

1. Dickson, D.C.M. (2006) *Insurance Risk and Ruin.* Cambridge University Press.
2. Gray, R.J. & Pitts, S.M. (2012) *Risk Modelling in General Insurance. From Principles to Practice.* Cambridge University Press.
3. Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M. (2008) *Modern Actuarial Risk Theory: Using R.* Springer, Berlin/Heidelberg.
4. Straub, E. (1988) *Non-Life Insurance Mathematics.* Springer-Verlag / Association of Swiss Actuaries, Berlin/Heidelberg.
5. Sundt, B. (1993) *An Introduction to Non-Life Insurance Mathematics.* VVW, Karlsruhe.

5 Loss distributions

The aim of this chapter is to give a brief overview of distributions suitable to describe the claim *severity*. Claim severity refers to the monetary loss of an (individual) insurance claim. Obviously the final aim of an insurance company is to estimate the aggregate claim amount in order to find suitable premiums and reserves. The study of individual claims can be seen as the first step towards this objective.

5.1 Characteristics of loss distributions

Claim severity is usually modelled as a nonnegative continuous random variable.

The list of possible distributions is large:

- Exponential distribution
- Pareto distribution
- Weibull distribution
- Lognormal distribution
- Gamma distribution
- Burr distribution
- Loggamma distribution
- Generalized Pareto distribution
- Inverse Gaussian distribution
- ...

In the following we focus on the first five, i.e. exponential, Pareto, Weibull, lognormal and gamma distributions and give a review of their key characteristics.

5.1.1 Exponential distribution

Exponential distribution is a good reference distribution for examples, but usually too "optimistic" for real models. Exponential distribution is also often used to describe the inter-arrival times of loss events. For an exponentially distributed random variable X we write $X \sim Exp(\lambda)$, where $\lambda > 0$ is a parameter.

Main characteristics for the exponential distribution are:

- distribution function $F_X(x) = 1 - e^{-\lambda x},\ x \geq 0$,
- density function $f_X(x) = \lambda e^{-\lambda x},\ x \geq 0$,
- expectation $EX = \frac{1}{\lambda}$,
- variance $VarX = \frac{1}{\lambda^2}$.

5.1.2 Pareto distribution

The Pareto distribution is a power law probability distribution that is used in several models in economics and social sciences. The corresponding distribution family is a wide one, with several sub-families. We will first review the classical Pareto distribution and then focus on it's shifted version (so-called *American Pareto* distribution), which is most widely used in non-life insurance as a model for claim severity.

A. The classical Pareto, $X \sim Pa^*(\alpha, \beta),\ \alpha,\ \beta > 0$

Main characteristics for the classical Pareto distribution are:

- distribution function $F_X(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha},\ x \geq \beta$,
- density function $f_X(x) = \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}},\ x \geq \beta$,
- expectation $EX = \frac{\alpha\beta}{\alpha-1},\ \alpha > 1$,
- variance $VarX = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)},\ \alpha > 2$,
- moments $EX^n = \frac{\alpha\beta^n}{\alpha-n},\ \alpha > n$, but $EX^n = \infty$, $\alpha \leq n$.

B. American Pareto, $Y \sim Pa(\alpha, \beta), Y = X - \beta, X \sim Pa^*(\alpha, \beta),\ \alpha,\ \beta > 0$

The American Pareto distribution is obtained from classical Pareto distribution by shifting it to the origin.

Main characteristics for the American Pareto distribution are:

- distribution function $F_Y(y) = 1 - \left(\frac{\beta}{\beta+y}\right)^{\alpha},\ y \geq 0$,
- density function $f_Y(y) = \frac{\alpha\beta^{\alpha}}{(\beta+y)^{\alpha+1}},\ y \geq 0$,
- expectation $EY = \frac{\beta}{\alpha-1},\ \alpha > 1$,
- variance $VarY = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)},\ \alpha > 2$,
- moments $EY^n = \frac{\beta^n n!}{\prod_{i=1}^{n}(\alpha-i)},\ \alpha > n$, but $EY^n = \infty$, $\alpha \leq n$.

Parameter estimation using the method of moments:

$$\hat{\alpha} = \frac{2(m_2 - m_1^2)}{m_2 - 2m_1^2},$$

$$\hat{\beta} = \frac{m_1 m_2}{m_2 - 2m_1^2},$$

where $m_1 = \sum_{j=1}^{n} \frac{y_j}{n}$ and $m_2 = \sum_{j=1}^{n} \frac{y_j^2}{n}$.

5.1.3 Weibull distribution

The Weibull distribution is one of the main distributions in survival analysis and reliability analysis, it is also used in models for service and manufacturing times. Because of its shape it is also usable to model severities in non-life insurance. For a Weibull-distributed random variable X we write $X \sim W(\alpha, \lambda)$, $\alpha > 0, \lambda > 0$.

Main characteristics for the Weibull distribution are:

- distribution function $F_X(x) = 1 - e^{-\lambda x^\alpha}$,
- density function $f_X(x) = \alpha \lambda^\alpha x^{\alpha-1} e^{-\lambda x^\alpha}$,
- moments $EX^n = \lambda^n \Gamma(\frac{1+n}{\alpha})$.

We also note that if

$\alpha < 1$, then Weibull distribution is "between" exponential and Pareto;
$\alpha > 1$, then Weibull distribution has lighter tail than exponential;
$\alpha = 1$, then Weibull distribution reduces to exponential.

5.1.4 Lognormal distribution

The distribution function of lognormal distribution is found using the log transformation to reach normal distribution and standardization to reach standard normal distribution. Since the normal distribution is one of the most thoroughly studied distributions, the simple connection between lognormal and normal makes lognormal distribution also an appealing choice in different models. For a lognormally distributed random variable X we write $X \sim LnN(\mu, \sigma)$, $-\infty < \mu < \infty$, $\sigma > 0$.

Main characteristics are:

- its connection to normal distribution, if $Y \sim N(\mu, \sigma)$ and $X = e^Y$, then $X = LnN(\mu, \sigma)$,
- distribution function $F_X(x) = F_Y(\ln x) = \Phi(\frac{\ln x - \mu}{\sigma})$,

- density function $f_X(x) = \frac{1}{x} f_Y(\ln x)$,
- expectation $EX = e^{\mu + \frac{\sigma^2}{2}}$,
- variance $VarX = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$,
- moments $EX^n = e^{n\mu + \frac{1}{2}n^2\sigma^2}$.

Parameter estimation:

$$\hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} \ln x_j,$$

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{j=1}^{n}[\ln(x_j - \hat{\mu})]^2.$$

5.1.5 Gamma distribution

The gamma distribution can also be considered as a generalization of exponential distribution. Namely, if α is integer, then gamma distribution can be interpreted as a sum of α independent exponentially distributed random variables. Gamma distribution is widely used in different models for manufacturing processes and telecommunications, because of its shape it is also a suitable choice in risk and ruin theory. For a gamma-distributed random variable X we write $X \sim \Gamma(\alpha, \lambda)$, $\alpha > 0, \lambda > 0$.

Main characteristics are:

- distribution function $F_X(x) = \gamma(\alpha, x\lambda)$,
- density function $f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1}e^{-\lambda x}, \ x > 0$,
  where
  $$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx = (\alpha-1)\Gamma(\alpha-1) \quad \text{is gamma function,}$$
  $$\gamma(\alpha, t) = \frac{\int_0^t x^{\alpha-1}e^{-x}dx}{\Gamma(\alpha)} \quad \text{is incomplete gamma function,}$$
- expectation $EX = \frac{\alpha}{\lambda}$,
- variance $VarX = \frac{\alpha}{\lambda^2}$.

Parameter estimation using the method of moments:

$$\hat{\alpha} = \frac{m_1^2}{m_2 - m_1^2},$$

$$\hat{\lambda} = \frac{m_1}{m_2 - m_1^2}.$$

5.1.6 Mixture distributions. Exponential/Gamma example

Mixture distributions

a) enable us to include in models for claim amounts the variability amongst risks in a portfolio (that is, they allow us to model heterogeneity of risks);

b) provide a source of further heavy-tailed loss distributions;

c) shed further light on some distributions we have already met.

Let us consider the following model:

- size of each individual claim is exponentially distributed;
- parameter of this exponential distribution is a Gamma-distributed random variable, $\Lambda \sim \Gamma(\alpha, \beta)$;
- parameters α and β are known.

So, for each fixed λ the (conditional) individual claim size is

$$X|\Lambda = \lambda \sim Exp(\lambda)$$

and the density function of Λ is

$$f_\Lambda(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda},\ \lambda > 0.$$

Let us calculate the (marginal) distribution of X:

$$\begin{aligned}
f_X(x) &= \int_0^\infty f_{X,\Lambda}(x,\lambda)d\lambda \\
&= \int_0^\infty f_\Lambda(\lambda) f_{X|\Lambda}(x|\lambda)d\lambda \\
&= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}\cdot\lambda e^{-\lambda x}d\lambda \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty \lambda^\alpha e^{-(x+\beta)\lambda}d\lambda \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty \frac{f_\Gamma(\lambda)\Gamma(\alpha+1)}{(x+\beta)^{\alpha+1}}d\lambda \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+1)}{(x+\beta)^{\alpha+1}}\int_0^\infty f_\Gamma(\lambda)d\lambda \\
&= \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}},
\end{aligned}$$

where $\Gamma \sim \Gamma(\alpha+1, x+\beta)$.

In conclusion, loss size in this case is (American) Pareto-distributed, $X \sim Pa(\alpha,\beta)$, i.e. Pareto distribution is a mixture of exponential distributions with a Gamma mixing distribution.

5.2 Related functions in R statistical software

In R statistical software, there is a variety of functions available related to different probability distributions. Some most commonly needed for exponential distribution (available in package `stats`, rate = λ) are:

- `dexp(x, rate = 1, log = FALSE)` – density;
- `pexp(q, rate = 1, lower.tail = TRUE, log.p = FALSE)` – distribution function;
- `qexp(p, rate = 1, lower.tail = TRUE, log.p = FALSE)` – quantile function;
- `rexp(n, rate = 1)` – random number generation .

Similar functions are available for all distributions discussed:

- Pareto: `dpareto, ppareto, qpareto, rpareto` (package `actuar`);
- Weibull: `dweibull, pweibull, qweibull, rweibull` (package `stats`);
- Lognormal: `dlnorm, plnorm, qlnorm, rlnorm` (package `stats`);
- Gamma: `dgamma, pgamma, qgamma, rgamma` (package `stats`).

5.3 Model evaluation and selection

After a distribution has been estimated, we have to evaluate it to ascertain that the assumptions applied are acceptable and supported by the data. This should be done prior to using the model for prediction and pricing. Model evaluation can be done using graphical methods, as well as formal misspecification tests and diagnostic checks.

In graphical methods, one simply plots the proposed candidate distribution function against the empirical distribution function. If the proposed candidate distribution fits well, the plotted graphs should be close, if the proposed distributional assumption is incorrect, the plotted graphs will differ. Similarly, one can plot the probability density functions of candidate distributions against the histogram of observed data and compare their fit.

Formal misspecification tests can be conducted to compare the estimated model against a hypothesized model. When the key interest is the comparison of the distribution functions, we may use the Kolmogorov-Smirnov and/or Anderson-Darling test. The chi-square goodness-of-fit test is an alternative for testing distributional assumptions, by comparing the observed frequencies against the theoretical frequencies. The likelihood ratio test is applicable to test the validity of restrictions on a model, and can be used to decide if a model can be simplified.

We will study two less known methods which might be particularly useful in insurance related problems:

- method of empirical mean excess function (or mean residual life function);
- limited expected value comparison test.

5.3.1 Method of mean excess function

Let X be a continuous random variable and let $x > 0$.

Definition 5.1 (Mean excess function $e(x)$)**.**

$$e(x) = E(X - x|X \geq x) = \frac{\int_x^\infty (t-x) f_X(t)dt}{\mathbf{P}\{X \geq x\}}.$$

Assuming $\lim_{t\to\infty}(x-t)(1-F_X(t)) = 0$, we can also write

$$e(x) = \frac{\int_x^\infty (1-F_X(t))dt}{1-F_X(x)}.$$

Empirical mean excess function can be calculated as:

$$e_n(x) = \sum_{i=1}^{n} \frac{I_{\{x_i \geq x\}} x_i}{\#\{x_i \geq x\}} - x.$$

Mean excess function is used

- in various risk and extreme value models;
- for distribution fitting;
- tail estimation of heavy-tailed distributions;
- ...

Mean excess functions for the distributions discussed in this chapter have the following forms:

- $X \sim Pa(\alpha, \beta)$, $e(x) = \frac{\beta + x}{\alpha - 1}$;
- $X \sim Exp(\lambda)$, $e(x) = \frac{1}{\lambda}$;
- $X \sim \Gamma(\alpha, \lambda)$, $e(x) \approx const$, if x is large;
- $X \sim W(\alpha, \lambda)$, $e(x) \approx \frac{1}{\lambda x^{\alpha - 1}}$, if x is large;
- $X \sim LN(\mu, \sigma)$, $e(x) \approx \frac{const \cdot x}{\ln x}$, if x is large.

Graphs of these functions are shown in the next figure. As one can see, the graphs of different distributions are clearly distinguished. So one can simply calculate the empirical mean excess function and decide based on its behaviour. If this seems like a linear function with positive slope, then a Pareto distribution might be an appropriate model. If it is more like a constant for the larger x values, gamma might provide good fit. Something between these choices might suggest either lognormal or Weibull distribution.

5.3.2 Limited expected value comparison test

Another ad hoc test that actuaries sometimes use is closely related to mean excess functions. This test might be preferred in situations in which the data is *censored* (from above) and therefore follows certain truncated distribution. Also, in that case it is impossible to compute the empirical mean excess function for the tail part.

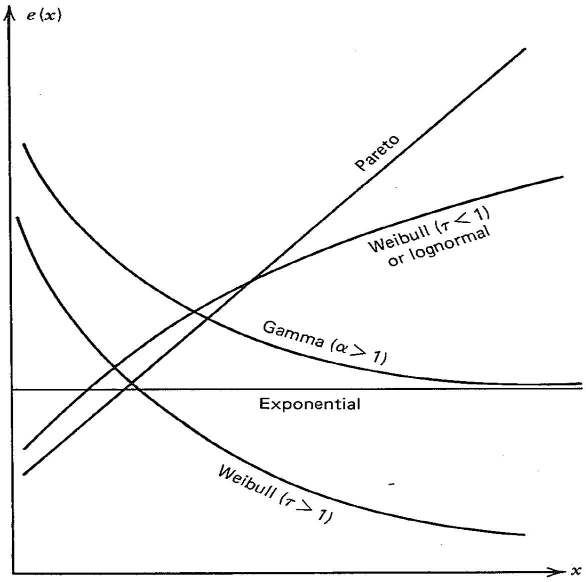


Figure 3.7. Mean residual life functions.

Definition 5.2. For any non-negative random variable X (or corresponding distribution $F(x)$) the limited expected value function $E[X;x]$ is defined by

$$E[X;x] = E(\min(X,x)) = \int_0^x ydF(y) + x(1-F(x)), x > 0.$$

An alternative formula might also prove useful in some situations:

$$E[X;x] = \int_0^x [1-F(y)]dy.$$

Following the same idea, the empirical limited expected value function can be calculated as

$$E_n[X;x] = \frac{\sum_{i=1}^n \min\{x_i, x\}}{n}.$$

For simplicity (and keeping in mind that the main aim is to apply the limited expected value function to a claim size distribution), we assume that

- X is continuous;
- X is nonnegative;

- $EX < \infty$.

It is easy to prove that the LEV-function $E[X;x]$ has the following general properties:

1. $E[X;x]$ is continuous, concave and nondecreasing function;
2. $EX = E[X;x] + e(x)(1 - F_X(x))$
3. $E[X;x] \to E(X)$, if $x \to \infty$;
4. $F(x) = 1 - (E[X;x])'$
   $\Rightarrow E[X;x]$ determines the distribution of X uniquely!
5. Limited expected value function of $aX + b$ is given by

$$E[aX + b; x] = aE\left[X; \frac{x-b}{a}\right] + b.$$

The goodness-of-fit of a proposed candidate distribution and the observed sample is measured by the following comparison test. Let us calculate the differences

$$d_i = \frac{E[X;x_i] - E_n[X;x_i]}{E[X;x_i]}, \quad i = 1, 2, \ldots, n - c,$$

and find the vector $\vec{D} = (d_1, \ldots, d_{n-c})$, where c is the number of censored observations. For simplicity we assume that the sample is ordered and the last c have not been observed. If $\vec{D}$ is close to null vector, it is reasonable to believe that the distribution corresponding to $E[X;x]$ fits given data. The main problem concerning this method is that there are no good criteria to decide when $\vec{D}$ is close enough to null vector (and when not).

5.4 Effects of coverage modifications to loss distributions

We use the following notations for quantities of interest (also certain subscripts may be used to specify the particular situation):

- X – actual loss size, this will be divided between insurer and policyholder, $X = Y + Z$;
- Y – insurer's part of loss;
- Z – policyholder's part of loss;
- $F_X(\cdot)$ – distribution function of X;

- $f_X(\cdot)$ – probability density function of X.

Let us also define a characteristic that shows the proportion of losses that given coverage limitation (e.g., deductible or an upper limit) allows us to eliminate.

Definition 5.3 (Loss elimination ratio (LER))**.**

$$LER = \frac{\text{amount of eliminated claims}}{\text{amount of total claims}} = \frac{\sum Z_i}{\sum X_i}.$$

The following examples show how the expected claim amounts change and what are loss elimination ratios using different coverage limitations.

Example 5.1. In case of fixed amount deductible b, we can express the actual loss size as a sum

$$X = Y_b + Z_b,$$

where

$$Y_b = \begin{cases} X - b, \text{ if } X > b, \\ 0, \text{ if } X \le b. \end{cases}$$

Then

$$\begin{aligned} E(Z_b) &= \int_0^b y f_X(y) dy + b[1 - F_X(b)] = E[X; b], \\ E(Y_b) &= E(X) - E[X; b] \end{aligned}$$

and

$$LER_b = \frac{E[X; b]}{E(X)}.$$

Example 5.2. If we apply fixed amount deductible and also take into account the inflation rate r, we get the following formulas:

$$\begin{aligned} X_r &= (1 + r)X, \\ Y_{b,r} &= \begin{cases} X_r - b, \text{ if } X_r > b, \\ 0, \text{ if } X_r \le b, \end{cases} \\ E(Y_{b,r}) &= (1 + r)\left(E(X) - E\left[X; \frac{b}{1 + r}\right]\right), \\ LER_{b,r} &= \frac{E\left[X; \frac{b}{1+r}\right]}{E(X)}. \end{aligned}$$

If the deductible is not adjusted by inflation, then

- it is not possible to calculate $E_n[X; \frac{b}{1+r}]$,
- $LER_{b,r}$ decreases if r increases.

Example 5.3. In case we apply an upper limit u to an individual claim amount, we get the following formulas:

$$Y_u = \begin{cases} X, \text{ if } X \leq u, \\ u, \text{ if } X > u, \end{cases}$$

$$E(Y_u) = \int_0^u x f_X(x)dx + u[1 - F_X(u)] = E[X; u],$$

$$LER_u = \frac{E(X) - E[X; u]}{E(X)} = 1 - \frac{E[X; u]}{E(X)}.$$

Example 5.4. In case we apply an upper limit u and also adjust by inflation rate r, then:

$$X_r = (1 + r)X,$$

$$E(Y_{u,r}) = (1 + r)E\left[X; \frac{u}{1 + r}\right],$$

$$LER_{u,r} = 1 - \frac{E\left[X; \frac{u}{1+r}\right]}{E(X)}.$$

If we also adjust the upper limit by inflation, i.e., $u^{'} = (1 + r)u$, the expectation is calculated as follows:

$$E(X_{u^{'},r}) = (1 + r)E\left[X; \frac{u^{'}}{1 + r}\right] = (1 + r)E[X; u].$$

Example 5.5. In case the coverage is modified by upper limit and deductible, then, without the effect of inflation we have:

$$E(Y_{b,u}) = E[X; u] - E[X; b],$$
$$= \int_b^u y dF(y) + u(1 - F(u)) - b(1 - F(b)) = \int_b^u (1 - F(y))dy,$$
$$LER_{b,u} = LER_b + LER_u = 1 - \frac{E[X; u] - E[X; b]}{E(X)}.$$

If we also take the inflation into account, these formulas change to:

$$E(Y_{b,u,r}) = (1 + r)\left(E\left[X; \frac{u}{1 + r}\right] - E\left[X; \frac{b}{1 + r}\right]\right),$$

$$LER_{b,u,r} = LER_{b,r} + LER_{u,r} = 1 - \frac{E\left[X; \frac{u}{1+r}\right] - E\left[X; \frac{b}{1+r}\right]}{E(X)}.$$

References

1. Goulet, V., et al. Package 'actuar': http://cran.r-project.org/web/packages/actuar/actuar.pdf
2. Gray, R.J. & Pitts, S.M. (2012) *Risk Modelling in General Insurance. From Principles to Practice.* Cambridge University Press.
3. Hogg, R. & Klugman, S. (1984) *Loss distributions.* Wiley, New York.
4. R Core Team. Documentation for package 'stats': http://stat.ethz.ch/R-manual/R-patched/library/stats/html/00Index.html
5. Tse, Y.-K. (2009) *Nonlife Actuarial Models.* Cambridge University Press.

6 Risk models

The aggregate loss of a portfolio of insurance policies is the sum of all losses incurred in the portfolio. There are two main approaches to model such loss: the *individual risk model* and the *collective risk model*.

6.1 Individual risk model

Individual risk model is mostly used in case of group insurance (with fixed group size), it is also more common in health and life insurance. In the individual risk model we consider a risk portfolio with n policies. We also denote

- X_k – claim amount corresponding to k-th policy (in case there occurs a loss);
- Y_k – risk outcome corresponding to k-th policy (0 or X_k);
- S – aggregate (total) claim amount.

The general assumptions for individual risk model are:

1) risks Y_k are independent;
2) number of risks n is fixed;
3) each risk Y_k can cause at most 1 claim;
4) claims X_k may come from different distributions.

Thus, the aggregate (or total) claim amount can be calculated as

$$S = Y_1 + \ldots + Y_n.$$

The calculation of expectation and variance for S is obvious due to construction:

$$ES = EY_1 + \ldots + EY_n$$

and

$$VarS = VarY_1 + \ldots + VarY_n.$$

On the other hand, the distribution of S is usually quite hard to find. Note that typically most policies have zero loss, so that Y_k is zero for these policies. In other words, Y follows a mixed distribution with probability mass at point zero.

Example 6.1. Consider the following portfolio from medical insurance

Coverage	Number insured	Expected cost per insured	Std. dev per insured
Single	786	76	42
Family	592	187	77

Then

$$ES = 786 \cdot 76 + 592 \cdot 187 = 170\ 440,$$
$$VarS = 786 \cdot 42^2 + 592 \cdot 77^2 = 4\ 896\ 472.$$

Using the normal approximation, we can find

$$F_S(x) \approx \Phi(\frac{x - 170440}{2212.8})$$

and, e.g.,

$$\mathbf{P}\{S > 175000\} = 1 - F_S(175000) = 1 - \Phi(2.06) = 0.0197.$$

6.2 Collective risk model

The aims of collective risk model are:

- to describe the distribution of total claim amount with some known distribution;
- to include only the policies that actually caused claims (in order to reduce the amount of work).

Thus, we define S as a random sum

$$S = \sum_{i=1}^{N} X_i,$$

where N is a random variable denoting the number of claims.

We also assume that

- the claim severities $X_1, X_2, \ldots$ do not depend on the number of claims N;
- for any fixed n the individual claims $X_1, \ldots, X_n$ are i.i.d. random variables.

We can see that if all the risks Y_i follow the same distribution, individual risk model can be considered as a special case of collective risk model with $\mathbf{P}\{N = n\} = 1$.

Let us now denote the distributions of interest:

- $F(x) := \mathbf{P}\{X_k \le x\}$ – distribution of individual claim X_k;
- $G(x) := \mathbf{P}\{S \le x\}$ – distribution of total claim amount S.

Then

$$\begin{aligned}
G(x) = \mathbf{P}\{S \le x\} &= \mathbf{P}\left\{\bigcup_{k=0}^{\infty}\{S \le x \text{ and } N = k\}\right\} \\
&= \sum_{k=0}^{\infty}\mathbf{P}\{S \le x \text{ and } N = k\} = \sum_{k=0}^{\infty}\mathbf{P}\{N = k\}\mathbf{P}\{S \le x | N = k\} \\
&= \sum_{k=0}^{\infty}\mathbf{P}\{N = k\}\mathbf{P}\{X_1 + \ldots + X_k \le x\} \\
&= \sum_{k=0}^{\infty}\mathbf{P}\{N = k\}F^{*k}(x),
\end{aligned}$$

where F^{*k} denotes the n-fold convolution of F.

Remark 6.1 (Convolution of distributions)**.** Let $X_i,\ i = 1, \ldots, k$ be independent random variables with distributions P_i, then the distribution of $X_1 + \ldots + X_k$ (say, P) is called the convolution of distributions P_i. Similar notion is used for distribution functions and probability density functions: the distribution function $F_{X_1+\ldots+X_k}$ is called the convolution of (independent) distribution functions F_{X_i} and denoted by

$$F_{X_1+\ldots+X_k} = F_{X_1} * F_{X_2} * \ldots * F_{X_k}.$$

If X_i-s have same distribution, the corresponding convolution is denoted by F^{*k}.

For two independent random variables X and Y:

- in general: $F_{X+Y}(s) = \int F_X(s-y)dF_Y(y)$;
- if X is continuous: $f_{X+Y}(s) = \int f_X(s-y)dF_Y(y)$;
- if both X and Y are continuous: $f_{X+Y}(s) = \int f_X(s-y)f_Y(y)dy$.

We also recall the notion of moment generating function of a random variable: for any random variable Z, its moment generating function is defined by

$$M_Z(t) = E(e^{tZ}).$$

The key property of moment generating function is that its n-th order derivative at zero gives n-th raw moment of corresponding random variable, i.e.

$$EZ^n = M_Z^{(n)}(0).$$

Let us now focus on the moment generating function of aggregate claim amount S. It can be shown that the moment generating function of aggregate claim amount M_S can be calculated using the moment generating function of individual claims M_X and the moment generating function of the claim number M_N by the following formula:

$$M_S(t) = M_N(\ln M_X(t)).$$

Furthermore, this relation allows us to calculate expectation and variance of aggregate claim amount by

$$\begin{aligned} ES &= EN \cdot EX, \\ VarS &= (EX)^2 \cdot VarN + EN \cdot VarX. \end{aligned}$$

We also mention that the construction used to define the aggregate claim amount in collective risk model is actually a special form of definition of a *compound distribution*.

Definition 6.1 (Compound distribution)**.** Let $S = \sum_{i=1}^{N} X_i$ be a (random) sum of random variables X_i, where X_i are i.i.d. and $N \perp X_i$. Then S has a compound distribution of N.

The distribution of N is called the *primary distribution* and the distribution of X (where X follows the same distribution as X_i; since X_i-s are i.i.d. we can simply use X for brevity) is called the *secondary distribution*.

There are three classical choices for claim number N:

- binomial distribution ($EN = np,\ VarN = np(1-p),\ EN > VarN$);
- Poisson distribution ($EN = VarN = \lambda$);
- negative binomial distribution ($EN = \frac{\alpha(1-p)}{p},\ VarN = \frac{\alpha(1-p)}{p^2},\ EN < VarN$).

Thus we can talk about compound binomial distribution, compound Poisson distribution and compound negative binomial distribution.

6.3 Direct estimation of total claim amount

Besides severity-frequency models it is also possible to estimate the aggregate claim amount directly.

There are few different options:

- normal approximation (in case 2 first moments can be estimated):

$$F(x) \approx \Phi\left(\frac{x-\mu}{\sigma}\right);$$

- normal Power approximation (in case we can also estimate skewness γ):

$$F(x) \approx \Phi\left(-\frac{3}{\gamma} + \sqrt{\frac{9}{\gamma^2} + 1 + \frac{6}{\gamma}\frac{x-\mu}{\sigma}}\right);$$

- translated gamma approximation:
  - set $S = k + Y$, where $Y = \Gamma(\alpha, \lambda)$ and k is some constant
  - set parameters α, λ ja k equal to those of S, i.e. solve

$$\begin{aligned}\mu &= k + \frac{\alpha}{\lambda},\\ \sigma^2 &= \frac{\alpha}{\lambda^2},\\ \gamma &= \frac{2}{\sqrt{\alpha}}.\end{aligned}$$

6.4 Conclusions

The aggregate claim amount S can be estimated

- directly; or
- through distributions of individual claim amount X and claim frequency N.

Modelling the aggregate claim amount through distributions of X and N has some distinct advantages:

- The expected number of claims changes as the number of insured policies changes. Growth in the volume of business needs to be taken into account.
- The effects of inflation are reflected in the individual losses and are difficult to take into account on aggregate amounts, especially when deductibles and policy limits do not depend on inflation.

- The impact of changing individual deductibles and policy limits is more easily studied.
- The impact on claims frequencies of changing deductibles is better understood.
- Data that are heterogeneous in terms of deductibles and limits can be combined to obtain the hypothetical aggregate claim amount distribution (useful when data from several years with different conditions is combined).
- It is easier to take into account the influence of reinsurance.
- The shape of the distribution of S depends on the shapes of both distributions of N and X. The understanding of the relative shapes is useful when modifying policy details.

6.5 Calculation of aggregate claim amount distribution in R

Function `aggregateDist` (in package `actuar`): returns a function to compute the distribution function of the aggregate claim amount distribution in any point.

Most important arguments of `aggregateDist`:

- `method` – method to be used:
  - `method="recursive"` – Panjer recursion;
  - `method="convolution"` – uses convolutions;
  - `method="normal"` – normal approximation;
  - `method="npower"` – Normal Power approximation;
  - `method="simulation"` – uses simulations from empirical distribution;
- `model.freq` – frequency distribution*;
- `model.sev` – severity (individual claim amount) distribution*.

  * Exact usage will depend on the choice of the method, see the documentation of package `actuar`.

References

1. Dickson, D.C.M. (2006) *Insurance Risk and Ruin.* Cambridge University Press.

2. Goulet, V., et al. Package 'actuar': `http://cran.r-project.org/web/packages/actuar/actuar.pdf`

3. Gray, R.J. & Pitts, S.M. (2012) *Risk Modelling in General Insurance. From Principles to Practice.* Cambridge University Press.

4. Klugman, S.A., Panjer, H.H., Willmot, G.E. (1998) *Loss Models: From Data to Decisions.* Wiley, New York.

5. Panjer, H.H., Willmot, G.E. (1992) *Insurance risk models.* Society of Actuaries.

6. Tse, Y.-K. (2009) *Nonlife Actuarial Models.* Cambridge University Press.

7 Claim number distribution

In this section we study some distributions suitable to describe the number (or frequency) of claims.

7.1 The (a, b)-class of counting distributions

Definition 7.1 (Counting distribution)**.** A counting distribution is a discrete distribution with only non-negative integers in its domain.

We typically use a counting distribution to model the number of occurrences of a certain event, for example "number of car accidents in a year".

Definition 7.2 (The (a, b)-class)**.** The (a, b)-class (or $(a, b, 0)$-class) is a two-parameter family of counting distributions that satisfy the following recursion:

$$p(k) = (a + \frac{b}{k})p(k-1),\ k = 1, 2, \ldots \tag{7.1}$$

for some $a,\ b \in \mathbb{R}$. The family is usually denoted by $\mathcal{R}$ and a particular counting distribution with parameters a and b is denoted by $\mathcal{R}(a, b)$.

Example 7.1 (Poisson distribution)**.** Let us have $N \sim Po(\lambda)$.

Then

$$p(k) = \mathbf{P}\{N = k\} = \frac{\lambda^k}{k!}e^{-\lambda},\ \ p(k-1) = \mathbf{P}\{N = k-1\} = \frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}.$$

In conclusion

$$p(k) = \frac{\lambda}{k}p(k-1)$$

, i.e., $a = 0$ and $b = \lambda$ or, equivalently,

$$N \sim \mathcal{R}(0, \lambda).$$

Thus, the region of parameters (a, b) covered by the Poisson distribution is

$$\{(a, b) : a = 0,\ b > 0\}.$$

Example 7.2 (Binomial distribution)**.** Let us have $N \sim Bin(n, p)$.

Then

$$p(k) = C_n^k p^k (1-p)^{n-k},\ \ p(k-1) = C_n^{k-1} p^{k-1} (1-p)^{n-k+1}$$

and

$$\frac{p(k)}{p(k-1)} = \frac{n+1-k}{k}\frac{p}{1-p} = -\frac{p}{1-p} + \frac{n+1}{k}\frac{p}{1-p}$$

or, equivalently,

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

NB! Check what happens if $k > n$!

Thus $a = -\frac{p}{1-p}$ and $b = -(n+1)a$ and the region of parameters (a, b) covered by binomial distribution is

$$\{(a, b) : a < 0,\ b = -ma,\ m = 2, 3, \ldots\}.$$

Example 7.3 (Negative binomial distribution)**.** Let us have $N \sim NBin(\alpha, p)$. Then

$$p(k) = C^k_{\alpha+k-1} p^\alpha (1-p)^k, \quad p(k-1) = C^{k-1}_{\alpha+k-2} p^\alpha (1-p)^{k-1}$$

and

$$\frac{p(k)}{p(k-1)} = \frac{(k+\alpha-1)(1-p)}{k} = (1-p) + \frac{(\alpha-1)(1-p)}{k}.$$

Thus $a = (1-p)$ and $b = (\alpha-1)a$ and the region of parameters (a, b) covered by negative binomial distribution is

$$\{(a, b) : 0 \leq a \leq 1,\ b > -a\}.$$

Theorem 7.1 (The (a, b)-class theorem)**.** *The class $\mathcal{R}$ contains the Poisson, the negative binomial, and the binomial distributions, and these are the only non-degenerate members.*

Idea of the proof:

1. Previous examples show that the mentioned distributions belong to $\mathcal{R}$
2. Consider the remaining 3 regions separately:
   (a) $a + b \leq 0$;
   (b) $a \geq 1,\ a + b > 0$;
   (c) $a < 0,\ b \neq -ma$ for any $m = 2, 3, \ldots$

7.2 Examples of collective risk models

Example 7.4 (Compound Poisson model)**.** Let us have $N \sim Po(\lambda)$. Then $EN = VarN = \lambda$ and $M_N(t) = \exp\{\lambda(e^t - 1)\}$, general formulas for $M_S(t)$, ES and $VarS$ simplify to

$$\begin{aligned} M_S(t) &= \exp\{\lambda(M_X(t) - 1)\}, \\ ES &= \lambda EX, \\ VarS &= \lambda VarX + \lambda(EX)^2 = \lambda EX^2. \end{aligned}$$

Also $E(S - ES)^3 = \lambda EX^3$ (prove it!), thus skewness is given by

$$\eta_3(S) = \frac{E(S-ES)^3}{\sqrt{(VarS)^3}} = \frac{\lambda EX^3}{\sqrt{(\lambda EX^2)^3}} \geq 0,$$

since $X \geq 0$.

Recall now that the sum of independent Poisson distributed random variable is also Poisson distributed. Such additivity is obviously very desirable property for a compound model to have as well. It is important when establishing the relations between distributions of aggregate claim amounts in different aggregation levels. Therefore the question whether the sum of independent compound Poisson random variables also has such property is naturally of great interest.

Theorem 7.2 (Sum of independent compound Poisson random variables)**.** *Let $S_1, S_2, \ldots, S_n$ be independent random variables such that S_i is compound Poisson distributed with parameters λ_i and $F_i(x)$). Then $S_1 + \ldots + S_n$ is compound Poisson distributed with parameters λ and $F(x)$, where*

$$\lambda = \sum_{i=1}^{n} \lambda_i \quad \text{and} \quad F(x) = \frac{1}{\lambda}\sum_{i=1}^{n} \lambda_i F_i(x).$$

Proof. Notice that $F(x)$ is a distribution function (it is weighted average of distribution functions with positive weights which sum to 1). The corresponding moment-generating function:

$$M(t) = \int_0^\infty e^{tx}\frac{1}{\lambda}\sum_{i=1}^{n}\lambda_i f_i(x)dx = \frac{1}{\lambda}\sum_{i=1}^{n}\lambda_i\int_0^\infty e^{tx} f_i(x)dx = \frac{1}{\lambda}\sum_{i=1}^{n}\lambda_i M_i(t),$$

where $f_i(x)$ and $M_i(x)$ are the probability density function and moment-generating function corresponding to $F_i(x)$.

Let $M_S(t)$ denote the moment-generating function for S.
Then (as $S_1, S_2, \ldots, S_n$ are independent):

$$M_S(t) = Ee^{tS} = Ee^{tS_1 + \ldots + tS_n} = \prod_{i=1}^{n} Ee^{tS_i}.$$

On the other hand:

$$Ee^{tS_i} = M_S(t) = \exp\{\lambda_i(M_i(t) - 1)\},$$

which implies

$$M_S(t) = \exp\left\{\sum_{i=1}^{n} \lambda_i(M_i(t) - 1)\right\} = \exp\{\lambda(M(t) - 1)\},$$

where

$$\lambda = \sum_{i=1}^{n} \lambda_i \quad \text{and} \quad M(t) = \frac{1}{\lambda}\sum_{i=1}^{n} \lambda_i M_i(t).$$

Since the moment-generating function determines the distribution uniquely, S is compound Poisson distributed with parameters S and $F(x)$. □

In conclusion, with the compound Poisson model the estimation of each individual risk in homogeneous classes gives immediately an estimate for the distribution of total claim amount as well. So it is clearly an appealing choice to model the aggregate claim amount.

Example 7.5 (Compound binomial model)**.** Let us have $N \sim Bin(n, p)$, then $EN = np$, $VarN = np(1 - p)$ and $M_N(t) = (p \cdot e^t + 1 - p)^n$.

Then

$$\begin{aligned} M_S(t) &= (p \cdot M_X(t) + 1 - p)^n, \\ ES &= npEX, \\ VarS &= npVarX + np(1-p)(EX)^2 = npEX^2 - np^2(EX)^2. \end{aligned}$$

The expression for the third central moment is

$$E(S - ES)^3 = npEX^3 - 3np^2EX^2EX + 2np^3(EX)^3$$

(prove it!) and skewness is calculated as

$$\eta_3(S) = \frac{npEX^3 - 3np^2EX^2EX + 2np^3(EX)^3}{\sqrt{(npEX^2 - np^2(EX)^2)^3}}.$$

NB! Skewness can be either positive or negative!

Example 7.6 (Compound negative binomial model)**.** Let us have $N \sim NBin(\alpha, p)$, then $EN = \frac{\alpha(1-p)}{p}$, $VarN = \frac{\alpha(1-p)}{p^2}$ and $M_N(t) = p^\alpha(1-(1-p)e^t)^{-\alpha}$.

Then

$$\begin{aligned} M_S(t) &= \frac{p^\alpha}{(1-(1-p)M_x(t))^\alpha}, \\ ES &= \frac{\alpha(1-p)}{p}EX, \\ VarS &= \frac{\alpha(1-p)}{p}VarX + \frac{\alpha(1-p)}{p^2}(EX)^2 = \frac{\alpha(1-p)}{p}EX^2 + \frac{\alpha(1-p)^2}{p^2}(EX)^2. \end{aligned}$$

Skewness is positive, but the exact formula is quite complex.

Lastly, we introduce yet another criterion for fitting the claim number distribution from the (a, b)-class of distributions.

We can rewrite the (a, b)-class condition as

$$\frac{p(k)}{p(k-1)} = a + \frac{b}{k},$$

which implies $k\frac{p(k)}{p(k-1)} = ka + b$.

In other words the quantity $k\frac{p(k)}{p(k-1)}$ is a linear function of k.

Moreover, the slope a clearly distinguishes the candidate distributions:

- for Poisson $a = 0$;
- for binomial $a < 0$;
- for negative binomial $a > 0$.

In practice, one can plot

$$k \cdot \frac{\hat{p}(k)}{\hat{p}(k-1)} = k \cdot \frac{\text{policies with } k \text{ claims}}{\text{policies with } k-1 \text{ claims}}$$

against k.

References

1. Gray, R.J. & Pitts, S.M. (2012) *Risk Modelling in General Insurance. From Principles to Practice.* Cambridge University Press.

2. Hesseläger, O. (1998) *Lecture Notes in Non-Life Insureance Mathematics.* HCO Tryk, Copenhagen.

3. Klugman, S.A., Panjer, H.H., Willmot, G.E. (1998) *Loss Models: From Data to Decisions.* Wiley, New York.

4. Tse, Y.-K. (2009) *Nonlife Actuarial Models.* Cambridge University Press.

8 Panjer recursion and discrete Fourier transform

8.1 Panjer recursion

Recall that we have obtained the following formula for calculating the distribution function for aggregate claim amount:

$$G(x) = \sum_{n=0}^{\infty} \mathbf{P}\{N = n\} F^{*n}(x).$$

The problem is that this approach is usually not feasible except for certain special choices for the claim frequency and claim severity distributions. In this section, we take a new approach and assume that all severities are positive and integer-valued. While this might seem unexpected, in practice this assumption actually holds: all amounts are measured as multiples of some monetary unit!

Now, the probability mass function corresponding to aggregate claim amount is

$$g(x) = \mathbf{P}\{S = x\} = \sum_{n=0}^{\infty} \mathbf{P}\{N = n\} f^{*n}(x), \ x = 0, 1, \dots,$$

where

$$f^{*n}(x) = \mathbf{P}\{X_1 + \dots + X_n = x\} = \sum_{y=1}^{x} f(y) f^{*(n-1)}(x - y)$$

and $f(x) = \mathbf{P}\{X_i = x\}$, $x = 1, 2, \dots$.

Then the next result allows us calculate all the probabilities $g(x)$ sequentially, starting from $g(0)$.

Theorem 8.1 (Panjer recursion)**.** *Let $S = \sum_{i=1}^{N} X_i$, where X_i-s are i.i.d. positive integer-valued random variables independent of N. Let the distribution of N belong to the (a, b)-class of counting distributions.*

Then

$$g(x) = \mathbf{P}\{S = x\} = \sum_{y=1}^{x} \left(a + \frac{by}{x} \right) f(y) g(x - y), \ x = 1, 2, 3 \dots$$

and $g(0) = \mathbf{P}\{N = 0\}$.

Proof. A. $g(0) = \mathbf{P}\{S = 0\} = \mathbf{P}\{N = 0\}$, since $X_i > 0$.

B. Use probability generating function $P_X(t) = Et^X = \sum_{i=0}^{\infty} p(i) t^i$, $p(i) = \mathbf{P}\{X = i\}$

It can be proved that $P_S(t) = P_N(P_X(t))$ and therefore also

$$P'_S(t) = P'_N(P_X(t))P'_X(t).$$

Let us calculate

$$\begin{aligned}
P'_N(t) &= \sum_{i=0}^{\infty} i \cdot t^{i-1} p(i) = \sum_{i=1}^{\infty} i \cdot t^{i-1} \left(a + \frac{b}{i}\right) p(i-1) \\
&= \sum_{i=0}^{\infty} (i+1) \cdot t^i \left(a + \frac{b}{i+1}\right) p(i) = \sum_{i=0}^{\infty} ((i+1)a + b) p(i) t^i \\
&= (a+b) P_N(t) + \sum_{i=0}^{\infty} ia \cdot p(i) t^i = (a+b) P_N(t) + at P'_N(t),
\end{aligned}$$

which leads us to

$$P'_N(t) = \frac{a+b}{1-at} P_N(t).$$

Now

$$P'_S(t) = \frac{a+b}{1 - aP_X(t)} P_N(P_X(t)) \cdot P'_X(t)$$

and

$$[1 - aP_X(t)] \cdot P'_S(t) = (a+b) \cdot P_S(t) \cdot P'_X(t). \tag{8.1}$$

Since, by definition of probability generating function, we have

$$P_S(t) = \sum_{x=0}^{\infty} g(x) t^x \quad \text{and} \quad P'_S(t) = \sum_{x=0}^{\infty} x \cdot g(x) t^{x-1}$$

and

$$P_X(t) = \sum_{y=1}^{\infty} f(y) t^y \quad \text{and} \quad P'_X(t) = \sum_{y=1}^{\infty} y \cdot f(y) t^{y-1},$$

we can rewrite (8.1) as

$$\sum_{x=0}^{\infty} x g(x) t^{x-1} - a \sum_{y=1}^{\infty} \sum_{x=0}^{\infty} x f(y) g(x) t^{x+y-1} = (a+b) \sum_{x=0}^{\infty} \sum_{y=1}^{\infty} y f(y) g(x) t^{x+y-1}$$

or, equivalently,

$$\sum_{x=1}^{\infty} x g(x) t^{x-1} - a \sum_{x=1}^{\infty} \sum_{y=1}^{x} (x-y) f(y) g(x-y) t^{x-1} = (a+b) \sum_{x=1}^{\infty} \sum_{y=1}^{x} y f(y) g(x-y) t^{x-1},$$

from where

$$xg(x) - a\sum_{y=1}^{x}(x-y)f(y)g(x-y) = (a+b)\sum_{y=1}^{x} yf(y)g(x-y).$$

In conclusion, we have

$$g(x) = \sum_{y=1}^{x}\left(a + \frac{by}{x}\right) f(y)g(x-y),\ x = 1, 2, 3\ldots.$$

□

Now, the probabilities can be calculated sequentially as follows:

- $g(1) = (a+b)f(1)g(0)$,
- $g(2) = (a+\frac{b}{2})f(1)g(1) + (a+b)f(2)g(0)$,
- $g(3) = (a+\frac{b}{3})f(1)g(2) + (a+\frac{2b}{3})f(2)g(1) + (a+b)f(3)g(0)$,
- $\ldots$

Finally, let us note the main advantages of Panjer recursion:

- it is easy to implement;
- it is "cheaper" than brute-force computation; and
- it allows "exact" evaluation of compound distributions.

On the other hand, it also has some known weaknesses:

- the method is still still quite resource-expensive, especially in case of large number of lattice points;
- precision of the result depends highly on how precise are the estimates for distributions of N and X.

8.2 The discrete Fourier transform method

Let us first recall two important definitions.

Definition 8.1 (Characteristic function)**.** The characteristic function φ of a random variable X is defined as

$$\varphi(t) = Ee^{itX},$$

where i is the imaginary unit.

Definition 8.2 (Discrete Fourier Transform (DFT))**.** Discrete Fourier transform for a sequence of M complex numbers c_0, c_1, $\dots c_{M-1}$ is defined by

$$\phi(k) = \sum_{t=0}^{M-1} c_t e^{-\frac{2\pi i}{M}tk}, \quad k = 0, \dots, M-1. \tag{8.2}$$

Using these definitions, for a discrete random variable X with probability mass function f and possible values 0,1,...,$M-1$ and choosing in Formula (8.2) $c_t = f(t)$, we can write

$$\phi(k) = \sum_{t=0}^{M-1} f(t)e^{-\frac{2\pi i}{M}tk} = Ee^{i\cdot(-\frac{2\pi k}{M})X} = \varphi\left(-\frac{2\pi k}{M}\right), \quad k = 0, \dots, M-1.$$

Similarly, one can move back from the characteristic function φ to the probability mass function f using the inverse transform:

$$f(t) = \frac{1}{M}\sum_{k=0}^{M-1} \phi(k)e^{\frac{2\pi i}{M}tk} = \frac{1}{M}\sum_{k=0}^{M-1} \varphi\left(-\frac{2\pi k}{M}\right)e^{\frac{2\pi i}{M}tk}, \quad t = 0, \dots, M-1.$$

To apply these ideas to the aggregate claim amount, let us recall the collective model setup:

$$S = \sum_{i=1}^{N} X_i,$$

where X_i-s are *iid* and $X_i \perp N$.

It can be proved that the characteristic function φ_S can be expressed as

$$\varphi_S(t) = P_N(\varphi_X(t)),$$

where P_N is the probability generating function of N. The proof is similar to the proof of moment generating function and probability generating function for S (use conditioning by N).

Now the following DFT algorithm can be used to estimate the aggregate claim amount:

1. Discretize the severity distribution and find the required values of the probability mass function:

$$f_X(0), f_X(1), \dots, f_X(M-1).$$

2. Using DFT, find the values for the characteristic function of severity:
$$\varphi_X\left(-\frac{2\pi\cdot 0}{M}\right), \varphi_X\left(-\frac{2\pi\cdot 1}{M}\right), \ldots, \varphi_X\left(-\frac{2\pi\cdot (M-1)}{M}\right).$$

3. Apply the relation $\varphi_S(t) = \varphi_N(\varphi_X(t))$ to find the values for compound characteristic function:
$$\varphi_S\left(-\frac{2\pi\cdot 0}{M}\right), \varphi_S\left(-\frac{2\pi\cdot 1}{M}\right), \ldots, \varphi_S\left(-\frac{2\pi\cdot (M-1)}{M}\right).$$

4. Apply the inverse DFT to these values to obtain the values for probability mass function g of S:
$$g(0), g(1), \ldots, g(M-1).$$

Straightforward application of DFT results in the same complexity as the Panjer recursion ($O(n^2)$ operations), a more efficient result is obtained using certain Fast Fourier transform (FFT) algorithms, where the complexity is reduced to $O(n \log n)$. The main idea of FFT algorithms relies on the factorization of properly chosen M (e.g., choose $M = 2^k$). In case of heavy-tailed distributions and moderate to big amount of data, the FFT approach outperforms Panjer recursion by far.

Example 8.1. Consider a compound Poisson model, i.e. $S = \sum_{i=1}^{N} X_i$, where $N \sim Po(\lambda)$.

Then the probability generating function P_N is given by
$$P_N(t) = \exp\{\lambda(t-1)\}$$
and the characteristic function of the compound distribution is
$$\varphi_S(t) = \exp\left\{\lambda\left(\varphi_X(t) - 1\right)\right\}.$$

Now, assume we have the following objects defined in R:

- `f` – an M-element vector consisting of the values of pmf of X at $0, \ldots, M-1$;
- `lambda` – the parameter of Poisson distribution of N.

Then the aggregate probability mass function g can be calculated by

```
g=Re(fft(exp(lambda*(fft(f)-1)),inverse=T)/M)
```

Note that R uses unscaled versions for both forward and inverse Fourier transforms (this is why we need to divide by M).

References

1. Gray, R.J. & Pitts, S.M. (2012) *Risk Modelling in General Insurance. From Principles to Practice.* Cambridge University Press.

2. Hesseläger, O. (1998) *Lecture Notes in Non-Life Insureance Mathematics.* HCO Tryk, Copenhagen.

3. Shevchenko P. V. (2010) Calculation of aggregate loss distributions. *The Journal of Operational Risk* 5(2), 3–40.

9 Introduction to classical ruin theory

Recall that while studying the cash-flow model we already encountered the surplus process of an insurance company in a quite simplified setting. In this section we study this process more thoroughly.

9.1 Setup

In the classical risk process, an insurer's surplus at fixed time $t > 0$ is determined by three quantities: the amount of surplus at time 0, the amount of premium income received up to time t, and the amount paid out in claims up to time t:

$$U(t) = u + ct - S(t),$$

where

- u is the insurer's initial surplus (at time 0);
- c is the insurer's (constant) rate of premium income per unit time (for simplicity it is assumed to be received continuously);
- $S(t)$ is the aggregate claims process.

The aggregate claims process is the only random component in the equality, and is modelled similarly to compound models studied so far:

$$S(t) = \sum_{i=1}^{N(t)} X_i,$$

where X_i are i.i.d. random variables representing the individual claim amounts and $\{N(t), t \geq 0\}$ is a counting process for the number of claims.

In the classical risk process model it is assumed that $\{N(t), t \geq 0\}$ is a Poisson process and the aggregate claims process $\{S(t), t \geq 0\}$ is thus a compound Poisson process.

Definition 9.1. A counting process $\{N(t), t \geq 0\}$ is called a *Poisson process*, if

(1) $N(0) = 0$;

(2) the increments of the process $N(s_i + t_i) - N(s_i)$ are independent for disjoint intervals $(s_i, s_i + t_i)$, $i = 1, 2, \ldots$;

(3) the distribution of the number of events in an interval of length t is a Poisson distributed random variable with parameter λt, i.e. for arbitrary $s, t \geq 0$

$$\mathbf{P}\{N(s+t) - N(s) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \; n = 0, 1, 2, \ldots.$$

The quantity λ is called the *intensity* of the underlying Poisson process.

Definition 9.2. A random process $\{S(t), t \geq 0\}$ is called *compound Poisson process*, if

$$S(t) = \sum_{i=1}^{N(t)} X_i, \quad t > 0,$$

where $\{N(t),\; t \geq 0\}$ is a Poisson process and $\{X_i,\; i = 1, 2, \ldots\}$ are i.i.d. random variables independent of $\{N(t), t \geq 0\}$.

It can be shown that in the Poisson process with intensity λ the interarrival times (the times between events) are exponentially distributed random variables with parameter λ.

In the context of a compound Poisson process representing an aggregate claims process, for any fixed time $t > 0$, the distribution of the time until the next claim is exponential with parameter λ.

Using the properties of the compound Poisson model, we can see that the average growth of the surplus in a time interval with length t is given by

$$\begin{aligned} E[U(s+t) - U(s)] &= E[u + c \cdot (s+t) - S(s+t) - u - c \cdot s + S(s)] \\ &= c \cdot t - E[S(s+t) - S(s)] = c \cdot t - \lambda t EX. \end{aligned}$$

So, in average, the growth of surplus is positive if

$$c > \lambda EX. \tag{9.1}$$

The condition (9.1) is called *net profit condition* and it can be shown that if this condition does not hold, then $\psi(u) = 1$ for all $u \geq 0$. Therefore, from now on we assume that the net profit condition holds, i.e. $c > \lambda EX$. It is often convenient to write $c = (1 + \theta)\lambda EX$, where θ is the premium *loading factor* or *safety loading*:

$$\theta = \frac{c}{\lambda EX} - 1.$$

So, with net profit condition, we have $\theta > 0$.

9.2 Definitions of ruin probability

The probability of ruin in infinite time, also known as the *ultimate ruin probability*, is defined as

$$\psi(u) = \mathbf{P}\{U(t) < 0 \text{ for some } t > 0\}.$$

So, $\psi(u)$ is the probability that insurer's surplus falls below zero at some time in the future, i.e. the insurer turns insolvent. Given model is a continuous time model. For discrete time model, the ultimate ruin probability is expressed as

$$\psi_r(u) = \mathbf{P}\{U(t) < 0 \text{ for some } t, t = a, 2a, 3a, \ldots\}.$$

If ruin occurs under the discrete time model, it must also occur under the continuous time model, the opposite is not true. However, as a becomes small, so that we are 'checking' the surplus level very frequently, then $\psi_r(u)$ should be a good approximation for $\psi(u)$.

One can also define *survival probability* $\phi(u)$, i.e. the probability of non-ruin as

$$\phi(u) = 1 - \psi(u).$$

The *finite time ruin probability* is defined by

$$\psi(u, t) = \mathbf{P}\{U(s) < 0 \text{ for some } s, 0 < s \leq t\},$$

i.e., the insurer becomes insolvent at some time in the interval $(0, t]$. Similarly, the discrete time ruin probability in finite time is defined by

$$\psi_r(u, t) = \mathbf{P}\{U(t) < 0 \text{ for some } s, s = a, 2a, 3a, \ldots, t\}.$$

9.3 The adjustment coefficient and Lundberg's inequality

In order to give an upper bound for the ruin probability, it is important to measure the risk of given surplus process. This is done using certain *adjustment coefficient* (or *Lundberg exponent*) R, which takes into account the aggregate claims and the premium income and is defined as the unique positive root of the following equality in r:

$$\lambda M_X(r) - \lambda - cr = 0, \tag{9.2}$$

where $M_X(r)$ is the moment generating function of claim severity X.

In general, the adjustment coefficient equation has one positive solution:

- $M_X(t)$ is strictly convex, since $M_X''(t) = E(X^2 e^{tX}) > 0$;

- $\lambda M_X(t) - \lambda - ct$ is decreasing at zero, if net profit condition (9.1) holds, since $M_X'(0) = EX < \frac{c}{\lambda}$;
- $M_X(t) \to \infty$ continuously (with few exceptions).

There are several equivalent forms to equation (9.2):

- $\lambda + cR = \lambda M_X(R)$,
- $1 + (1+\theta)R \cdot EX = M_X(R)$,
- $e^{Rc} = Ee^{RS}$,
- $c = \frac{1}{R} \ln M_S(R)$,

where S is the total claims amount in an interval of length 1 and $M_S(R)$ is its moment-generating function.

Theorem 9.1 (Lundberg's exponential bound for the ruin probability)**.** *For a compound Poisson risk process with an initial capital u, premium per unit of time c, and an adjustment coefficient R, satisfying* (9.2) *the following inequality for the ruin probability holds:*

$$\psi(u) \leq e^{-Ru}.$$

9.4 Top-down model for premium calculation

The ruin theory also allows us to shed some more light to the premium principles. Namely, we can use the ruin process model to find certain estimates for the risk loading coefficients of premium principles.

Let us start with the following question: how large should be the initial capital u and the premimum c in order to remain solvent at all times with a prescribed probability?

By the Lundberg inequality, the ruin probability $\psi(u)$ is bounded from above by e^{-Ru}, where R can be found as the root of $e^{Rc} = Ee^{RS}$. Now, if we set the upper bound equal to ε, then

$$R = \frac{|\ln \varepsilon|}{u}. \tag{9.3}$$

Hence, the premium c can be calculated for given initial capital u and allowed ruin probability ε using the following formula:

$$c = \frac{1}{R} \ln(Ee^{RS}), \tag{9.4}$$

where R is specified by (9.3), and we recognize the exponential principle with risk aversion equal to the adjustment coefficient R.

Let us assume now that the total premium (in time unit) S has finite (non-central) moments, $E[S^k] < \infty$, $k = 1, 2, 3...$ It is a known fact that the moment generating function defines the distribution of S uniquely. The logarithm of the moment generating function is called *the cumulant generating function* (further referred to as *cgf*) which is denoted as $\kappa_S(t)$ and is presented as

$$\kappa_S(t) = \ln M_S(t) = \ln Ee^{tS}.$$

Using the Taylor expansion on the last formula, we obtain

$$\kappa_S(t) = \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!},\ t = 1, 2, 3...,$$

where κ_n denotes the n-th *cumulant* of random variable S. Now the premium for exponential principle (9.4) can be simply rewritten as

$$c = \frac{1}{R}\kappa_S(R),$$

and applying the Taylor expansion will result in

$$\begin{aligned} c &= \frac{1}{R}\kappa_S(R) = \frac{1}{R}\sum_{n=1}^{\infty} \kappa_n \frac{R^n}{n!} = \frac{1}{R}\left(\kappa_1 R + \kappa_2 \frac{R^2}{2!} + ...\right) \\ &\approx \kappa_1 + \frac{1}{2}R\kappa_2 = ES + \frac{1}{2}R \cdot VarS, \end{aligned}$$

where the approximation is justified if the risk aversion R is small. Thus the variance principle can be considered as certain simplification of the exponential principle and in the top-down method framework the risk loading factor γ is determined by the Lundberg exponent, $\gamma = \frac{1}{2}R$. Also, the following properties hold true:

- if the insurer wants to rise the loading factor γ two times, then the tolerated ruin probability ε decreases to ε^2;
- if the insurer wants to double the loading factor γ and keep the same tolerated ruin probability, the initial capital u can be halved.

Let us now consider the aspect of dividends in this framework. Assume that operating capital suppliers want to receive the dividend proportional to the provided capital u. Let the yearly dividend rate be denoted by i. One must now take into account that the dividend part is not included into costs but must be added to them, therefore the premium must be also increased by iu. The risk process (for one year) can then be presented as

$$U_1 = u + (c + iu) - S - iu.$$

Thus, the question to be answered is: how big should the premium be to ensure the non-ruin probability $1 - \varepsilon$ and to cover the dividends paid to the shareholders? Consider the variance premium calculation principle with loading factor $\gamma = \frac{1}{2}R$, where R is calculated as in (9.3). Then the corresponding premium can be presented by

$$P(S) = ES + \frac{|\ln \varepsilon|}{2u} VarS + iu. \tag{9.5}$$

Such premium is more likely to be achieved if the initial capital is chosen so that it makes the premium to be as low as possible. After setting the first derivative of $P(S)$ with respect to u equal to zero, we get

$$P'(S) = -VarS\frac{|\ln \varepsilon|}{2u^2} + i = 0,$$

or, equivalently,

$$\frac{VarS|\ln \varepsilon|}{2u^2} = i.$$

The the initial capital u can be calculated from

$$u = \sqrt{VarS}\sqrt{\frac{|\ln \varepsilon|}{2i}}. \tag{9.6}$$

Thus, the higher is the required dividend rate i, the lower is the optimal size of initial capital u. To obtain the required premium, the derived initial capital u must be substituted into (9.5). After some simplifications it comes out that the premium is equal to

$$\begin{aligned} P(S) &= ES + \frac{\sqrt{i|\ln \varepsilon|}\sqrt{VarS}}{\sqrt{2}} + \frac{\sqrt{i|\ln \varepsilon|}\sqrt{VarS}}{\sqrt{2}} \\ &= ES + \sqrt{2i|\ln \varepsilon|}\sqrt{VarS}. \end{aligned}$$

As $\sqrt{VarS}$ is the standard deviation of claim amount S, the derived premium is the standard deviation premium with the loading factor $\beta = \sqrt{2i|\ln \varepsilon|}$.

References

1. Dickson, D.C.M. (2006) *Insurance Risk and Ruin.* Cambridge University Press.

2. Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M. (2008) *Modern Actuarial Risk Theory: Using R.* Springer, Berlin/Heidelberg.

3. Ross, S.M. (2009) *Introduction to Probability Models.* Elsevier/Academic Press.

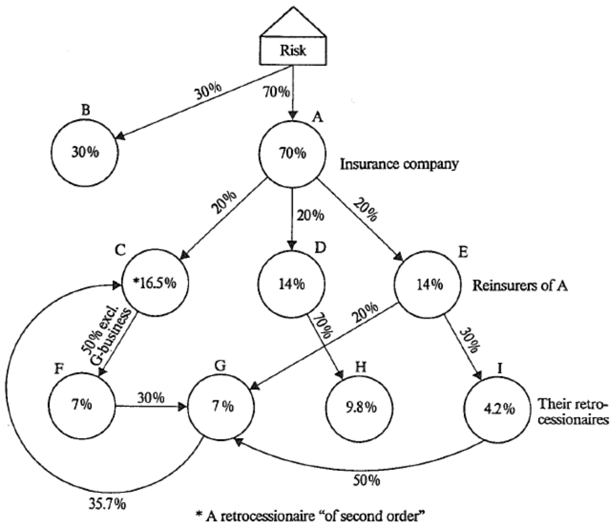
10 Reinsurance

Similarly to "regular" insurance, the risks and liabilities can grow too high for a single insurer to handle, thus there is a need to transfer portions of risk portfolios to other parties by some form of agreement in order to reduce the likelihood of having to pay a large obligation resulting from an insurance claim.

There are two general solutions

- reinsurance – "insurance for insurers";
- co-insurance – cooperation with other insurers.

In the reinsurance process, the direct (initial) insurer is called either *ceding company* or *cedant* or *cedent*, for the reinsurer we may also use the term reinsurance company. An example of the reinsurance process is shown in the following figure.



Some world's biggest names in reinsurance business are Swiss Re, Munich Re, Hannover Re, Lloyd's, Allianz.

10.1 Types of reinsurance

Reinsurance can be classified by the type of reinsurance contract:

- facultative reinsurance;
- compulsory (treaty) reinsurance.

By the compensation mechanics, reinsurance can be divided into the following categories:

- proportional reinsurance:
  - *quota share*;
  - *surplus*;
- non-proportional reinsurance:
  - *excess of loss (XL)*;
  - *stop loss.*

Before studying the reinsurance mechanics in different categories, we need to introduce some related notation. Let us denote

- S, X, P – gross total claim amount, severity, premium;
- $\tilde{S}, \tilde{X}, \tilde{P}$ – net total claim amount, severity, premium;
- $\hat{S}, \hat{X}, \hat{P}$ – reinsurer's total claim amount, severity, premium.

10.1.1 Proportional reinsurance

The main principle of proportional reinsurance is that the proportion of premiums equals to the proportion of claim amount.

Remark 10.1. In practice, this is not entirely true, since there are, in fact, certain commission fees that need to be taken into account separately. Commission fee is a fee that the reinsurer pays to ceding company for bringing the reinsurer a new client. Still, the commission fees do not change the general principle and are not of interest from our point of view, thus, we may safely ignore them.

So, we can illustrate the idea of proportional reinsurance by the following formula:

$$\frac{\hat{P}}{P} = \frac{\hat{X}}{X}.$$

Main advantage of proportional reinsurance is simple calculation of premiums and compensation, while the main drawback is that both parties have to handle all claims (not depending on claim size).

There are two types of proportional reinsurance. In *quota share* reinsurance, everything is determined by fixed retention proportion α, $0 < \alpha < 1$.

Then, all claims in given portfolio are divided using proportion α as follows:

$$\tilde{X} = \alpha X, \qquad \hat{X} = (1-\alpha)X,$$
$$\tilde{S} = \alpha S, \qquad \hat{S} = (1-\alpha)S.$$

Similar holds for premiums:

$$\tilde{P} = \alpha P, \qquad \hat{P} = (1-\alpha)P.$$

Another type of proportional reinsurance is the *surplus* reinsurance. Whereas quota share reinsurance scales down each risk in the same proportion, a surplus treaty is more flexible: the proportion of each risk reinsured may vary. To obtain this flexibility, we need to introduce more parameters than one single proportion factor. The ceded amount is defined by the following parameters:

- ceding company retains at most a certain amount m of each risk (*one line*);
- the exceeding part is reinsured but only up to a certain multiple l of the retention (*number of lines*);
- proportion of liabilities is determined individually depending on the risk size Q (consider, e.g., possible maximal loss (PML), sum insured or estimated maximum loss (EML));
- (automatic underwritten) capacity of the treaty is thus $L = (l+1)m$.
- larger risks $Q > L$ will require another reinsurance layer or facultative reinsurance.

Under such construction the initial insurer's (cedant's) part is calculated as:

$$\tilde{X} = \begin{cases} X, & Q \leq m; \\ \frac{m}{Q}X, & m < Q \leq L. \end{cases}$$

NB! The quantities Q and m are fixed when contract is underwritten and do not depend on the actual claim amount X. If $X = Q$ then $\tilde{X} = m$, i.e. initial insurer's maximal payout can not exceed m.

Example 10.1 (Surplus treaty)**.** Consider a 9-line surplus treaty ($l_1 = 9$) with one line $m_1 = 1\ 000\ 000$ EUR, i.e. the treaty capacity is $L_1 = (l_1 + 1)m_1 = 10\ 000\ 000$ EUR. Let the cedant also have another surplus treaty with $m_2 = 10\ 000\ 000$, $l_2 = 3$ and $L_2 = 40\ 000\ 000$. Consider a risk such that risk size (Q) is 21 million EUR.

Under the first treaty the cedant's proportion is

$$\frac{m_1}{Q} + \frac{Q - m_1(l_1+1)}{Q} = \frac{12}{21}$$

and the reinsurer's proportion is

$$1 - \frac{12}{21} = \frac{9}{21}.$$

Under the second treaty the cedant's proportion is

$$\frac{m_2}{Q} = \frac{10}{21}$$

and the reinsurer's proportion is

$$1 - \frac{10}{21} = \frac{11}{21}.$$

For a claim $X = 21$ million EUR the reinsurance part is

$$\hat{X} = \frac{9}{21} \cdot 21 + \frac{11}{21} \cdot 21 = 9 + 11 = 20 \text{ million EUR}$$

and cedant's part is 1 million EUR.

For a claim $X = 7$ milllion EUR the reinsurance part is

$$\hat{X} = \frac{9}{21} \cdot 7 + \frac{11}{21} \cdot 7 = \frac{20}{3} = 6\frac{2}{3} \text{ million EUR}$$

and cedant's part is $\frac{1}{3}$ million EUR.

10.1.2 Non-proportional reinsurance

The traditional forms of non-proportional reinsurance cover are known as *excess of loss* (XL) and *stop loss*. Both provide cover once claims exceed a certain level and usually have a limited insured amount.

In case of excess of loss reinsurance, the reinsurer will cover the part of claims which exceed a certain *excess point* or retention (also called *first risk*), up to agreed limit (called *second risk*).

Excess of loss reinsurance can be further classified:

- Working XL – applied to each claim;
- Aggregate XL – applied to aggregate claims;

- Cat XL – applied to losses occurred in a fixed period after a catastrophe.

Claims are divided by initial insurer and reinsurer as follows:

$$\tilde{X} = \min(X, M)$$

and

$$\hat{X} = \max(0, X - M),$$

where M is the retention of initial insurer.

Usually the XL-treaty covers certain layer A exceeding threshold M (*layer A xs M*). In such case

$$\hat{X} = \min(A, \max(0, X - M)) = \min(M + A, X) - \min(M, X).$$

In stop-loss contracts the aggregate covers are linked to the cedant's gross premium income during a 12 month period (P) through the loss ratio $\frac{S}{P}$: in case $\frac{S}{P}$ exceeds a fixed proportion, exceeding part is covered by reinsurer.

Usually the reinsurer's liability is also limited, i.e. the reinsurer will cover a layer defined by $m_1 < \frac{S}{P} \leq m_2$.

For example $m_1 = 1.05$, $m_2 = 1.3$ Then,

- in case $\frac{S}{P} \leq 1.05$ everything is covered by cedant, $\tilde{S} = S$;
- if $1.05 < \frac{S}{P} \leq 1.3$ then $\tilde{S} = 1.05 \cdot P$ and $\hat{S} = S - 1.05 \cdot P$;
- if $\frac{S}{P} > 1.3$ then $\tilde{S} = S - 0.25 \cdot P$ and $\hat{S} = 0.25 \cdot P$.

Example 10.2. Consider a portfolio with the following risks:

Risks	House	Block of flats	Industry
A	100 000		
B	200 000		
C	300 000		
D		2 000 000	
E			10 000 000
F			20 000 000
Premium rate	0.1%	0.2%	0.3%
Losses	B: 150 000		F: 2 550 000
Sum insured	600 000	2 000 000	30 000 000
Gross premium	600	4000	90 000
Loss burden	150 000		2 550 000

Consider also the following reinsurance contracts:

1. Quota of 30% (i.e. retention $\alpha = 70\%$);
2. Surplus 4 max after 100 000, 500 000, 2 500 000;
3. Fac. 7.5 million after 12.5 million;
4. Excess of loss (WXL) 4 million xs 1 million;
5. Stop loss 8 million after 2 million (Aggregate XL 8 million xs 2 million).

Then the corresponding reinsurance claims and premiums are shown in the following table:

	Premiums ($\hat{P}$)	Claims ($\hat{S}$)
1	28 380	810 000
2	55 800	1 350 000
3	22 500	956 250
4	?	1 550 000
5	?	700 000

10.2 The effect of reinsurance to claim distributions

Let us now study how the modifications to the cover by reinsurance affect the claim distributions.

In case of proportional reinsurance the proportion α fixed (either the same for all policies in case of quota share, or different for each risk depending on its size in case of surplus), so the claim severity distribution $\tilde{X}$ is found from:

$$F_{\tilde{X}}(x) = \mathbf{P}\{\tilde{X} \leq x\} = \mathbf{P}\{\alpha X \leq x\} = \mathbf{P}\{X \leq \frac{x}{\alpha}\} = F_X(\frac{x}{\alpha}).$$

The number of claims (claim frequency) stays the same for both parties.

In case of non-proportional (XL) reinsurance, the distribution of $\tilde{X} = \min(M, X)$ is given by

$$F_{\tilde{X}}(x) = \begin{cases} F_X(x), & x < M, \\ 1, & x \geq M. \end{cases}$$

Then, the expected individual and aggregate claim amounts for initial insurer are:

$$E\tilde{X} = E(\min(M, X)) = E[X; M],$$
$$E\tilde{S} = E\tilde{N}E\tilde{X} = EN \cdot E[X; M].$$

In practice, the reinsurer only needs to know the claims he is (partially) compensating, i.e. claims exceeding M. Therefore we introduce the following variables which are important from reinsurer's point of view:

- individual claim size: $\hat{X}_M := X - M|X > M$;
- number of claims: $\hat{N}_M := \sum_{i=1}^{N} I_{\{X_i > M\}}$.

The distribution function corresponding to $\hat{X}_M$ is

$$\begin{aligned} F_{\hat{X}_M}(x) :&= \mathbf{P}\{\hat{X}_M \le x\} = \mathbf{P}\{X - M \le x|X > M\} \\ &= \mathbf{P}\{M < X \le x + M|X > M\} = \frac{\int_M^{x+M} f_X(t)dt}{1 - F_X(M)} \\ &= \frac{F_X(x+M) - F_X(M)}{1 - F_X(M)}, \quad x \ge 0 \end{aligned}$$

and the probability density function of $\hat{X}_M$ is

$$f_{\hat{X}_M}(x) = \frac{f_X(x+M)}{1 - F_X(M)}, \quad x \ge 0.$$

By construction, the expectation of $\hat{X}_M$ is

$$E\hat{X}_M = E(X - M|X > M) = e(M) = \frac{EX - E[X;M]}{1 - F_X(M)}.$$

On the other hand, since $\tilde{X} + \hat{X} = X$,

$$E\hat{X} = EX - E\tilde{X} = EX\left(\frac{EX - E\tilde{X}}{EX}\right) = EX\left(1 - \frac{E[X;M]}{EX}\right).$$

Denote $c(x) := 1 - \frac{E[X;x]}{EX}$, then

$$E\hat{X}_M = \frac{c(M) \cdot EX}{1 - F_X(M)} = \frac{E\hat{X}}{1 - F_X(M)}.$$

Expected number of claims $\hat{N}_M$ can be calculated:

$$E\hat{N}_M = E(\sum_{i=1}^{N} I_{\{X_i > M\}}) = EN(1 - F_X(M))$$

In conclusion, the expected total claim amount $\hat{S}_M$ is given by

$$E\hat{S}_M = E\hat{X}_M E\hat{N}_M = \frac{E\hat{X}}{1 - F_X(M)} EN(1 - F_X(M)) = E\hat{X}EN = E\hat{S}.$$

References

1. Booth P., Chadburn, R., Cooper, D., Haberman, S., James, D. (1999) *Modern Actuarial Theory and Practice.* Chapman & Hall / CRC.

2. Gray, R.J. & Pitts, S.M. (2012) *Risk Modelling in General Insurance. From Principles to Practice.* Cambridge University Press.

3. Straub, E. (1988) *Non-Life Insurance Mathematics.* Springer-Verlag / Association of Swiss Actuaries, Berlin/Heidelberg.

11 Reserving

An insurance company's technical reserve is the amount set aside to meet company's principal insurance liabilities. Reserves are required in respect of business written, both earned and unearned.

Technical reserves serve serveral purposes:

- to enable the company to meet and administer its contractual obligations to policyholders;
- to provide management information;
- for tax purposes;
- to assist in sale and purchase negotiations;
- to advise on portfolio reinsurance;
- ...

The reserves held in respect of insurance related liabilities fall into three main categories:

- reserves in respect of unexpired (or unearned) exposure:
    1. unearned premium reserve (UPR);
    2. deferred acquisition costs (DAC);
    3. additional unexpired risk reserve (AURR);
- contingent reserves:
    1. catastrophe reserves;
    2. claims equalization reserves;
- reserves in respect of earned exposure:
    1. notified outstanding claims or reported but not settled claims (RBNS);
    2. incurred but not reported claims (IBNR);
    3. incurred but not enough reported (IBNER) on existing notified claims.

11.1 Unearned premium reserve (UPR)

The situation related to premium payments can be briefly described as follows:

- premiums are paid in advance;
- policies start at different times;
- insurance periods and accounting periods generally do not match;
- the insurer has contractual obligations to provide coverage beyond the accounting date.

Thus, the following two questions need to be answered:

1. How do estimate the sufficient size of the UPR?
2. When can we say that a premium is finally earned?

In general, the UPR in respect of an individual policy can be expressed using the following formula:

$$UPR(t_i) = P_0 \cdot F(t_i) \cdot (1-k),$$

where

- t_i is the accounting date;
- P_0 is the premium received from given policy;
- $F(t_i)$ is the proportion of the not exposed cover for this policy;
- k is the proportion of expenses (acquisition expenses are treated as earned).

A straightforward basic formula for the proportion factor is

$$F(\cdot) = \frac{\text{period of unexpired cover}}{\text{duration of original policy}}.$$

Such approach is also called the $\frac{1}{365}$-method. Proportions are calculated for each individual policy.

In practice it is unusual to calculate a UPR on an individual policy basis; generally the premiums received are aggregated within similar classes of business by month or quarter and the UPR is calculated on the assumption that the premiums were received, on average, half way through the period. So we can talk about

- monthly ('24ths' or $\frac{1}{24}$-) method;
- quarterly ('eights' or $\frac{1}{8}$-)method.

Then, the formula corresponding to $\frac{1}{24}$-method is

$$F(t) = \frac{2t-1}{24}, \; t = 1, 2, \ldots, 12.$$

Similar construction holds for quarterly calculations.

Example 11.1. Let the accounting date be 31.12.2012 and let us consider policies that started in March 2012. Then $t = 3$ and $F(3) = \frac{2\cdot3-1}{24} = \frac{5}{24}$, i.e. for each policy started in March 2012 a proportion $\frac{5}{24}$ from premiums is still unearned at the end of year 2012.

In case the risk exposure is not uniformly spread over the policy duration one can

- use a more complex proportion function $F(t)$;
- use additional reserves for periods with higher risk intensity.

11.2 Reserves in respect of earned exposure

The process underlying the appearance of a claim in the insurance company's books involves:

1) the occurrence of an insured event causing a loss to the policyholder;
2) the policyholder being aware of the loss and subsequently advising the insurer via a claim form;
3) the insurer processing the claim form and establishing a case reserve which might lead to a payment.

There are different options for establishing and estimation of the claim reserves:

- straightforward – claim-based approach (for reported claims);
- statistical methods for estimating required future provisions (especially important in IBNR calculation)
  - chain ladder method;
  - Bornhütter-Ferguson method;

– other stochastic approaches (GLM).

The claim development pattern is usually depicted in the following form (*run-off triangle*):

	Development period j				
Year of origin i	0	1	2	...	J
0	C_{00}	C_{01}	C_{02}	...	C_{0J}
1	C_{10}	C_{11}	...		
2	C_{20}	...			
...	...				
I	C_{I0}				

Here C_{ij} denotes the claims (numbers or amounts) for year of origin i ($i = 0, \ldots, I$, we can also use simply calendar years here) and development year j ($j = 0, \ldots, J$). Development period 0 means "current" or "running" period and the number of development periods J is defined by the nature of particular risks. The number of usable years of origin is defined by the history insurance company has. The usual assumption is that $I = J$; although, it is possible that the insurance company has more historical information about the particular reserve development ($I > J$), in this case we assume that all this information is included in the first row of the run-off triangle.

The development pattern is defined by the choice for the year of origin, which can be

- year of underwriting;
- year of accident (claim occurrence);
- year of reporting;

and the development year, which can be

- year of accident (occurrence);
- year of reporting;
- year of claim settlement (year of final payment).

Depending on the choice of years of origin and development, the run-off triangles can describe the development of different stages of different reserves. Obviously the development period has to correspond to a later process than the year of origin in order to get a meaningful development pattern.

For example, if the origin period is accident year and

- the development period is reporting year, the corresponding run-off triangle describes development in IBNR reserve;
- the development period is claim settlement year, the corresponding run-off triangle describes development in total claims reserve.

11.2.1 The chain ladder method

The most widely used reserving method is the *chain ladder* method. There are many variations, but they all have the same objective: to extract from the loss development triangle a pattern for the claim run-off that can be used to extrapolate the less mature years of account. The method is very simple and is based on the assumption that the development proportions remain the same (or at least are similar) in the future. If this assumption holds, the behaviour of future claims behaviour can be described by certain *development factors* or, equivalently, through some *proportion factors*.

More precisely, the chain ladder model assumes that the development between successive periods of development can be described as

$$C_{i,j+1} \approx f_j C_{ij}, \quad i = 0, \ldots, I, \; j = 0, \ldots, J-1,$$

where

- C_{ij} is the (cumulative) claim amount (or number) corresponding to year of origin i and development period j;
- $C_{i,j+1}$ is the (cumulative) claim amount (or number) corresponding to year of origin i and development period $j+1$;
- f_j is the development factor between development periods j and $j+1$.

Equivalently, the chain ladder model can be written as

$$C_{ij} \approx S_i R_j,$$

where

- C_{ij} is the (cumulative) claim amount (or number) corresponding to year of origin i and development period j;
- S_i is the ultimate claim amount (or number) for year of origin i;
- R_j is the proportion of the ultimate that has emerged by the end of the on jth development period.

The approximate equalities are used to stress out that there is no fixed relationship, we just propose a model for predictions (alternatively, one can specify a model with exact equalities and include error terms).

The estimation of development factors is done using all the available information between successive development periods:

$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{ij}}, \quad j = 0, \dots, J-1.$$

The cumulative development factors describe development pattern between given development stage j and the ultimate (or final) stage J. Thus, the estimates are found as

$$\hat{F}_j = \prod_{k=j}^{J-1} \hat{f}_k, \quad j = 0, \dots, J-1.$$

Also, the inverses of cumulative factors represent the proportion of total claims emerged by the end of given development period. Thus, the the estimates for proportions are simply found by

$$\hat{R}_j = \frac{1}{\hat{F}_j}, \quad j = 0, \dots, J-1.$$

As the basic chain ladder method does not take into account factors depending on particular calendar years (e.g. inflation) the following adjustment can be made to obtain a more flexible model:

$$C_{ij} \approx S_i R_j \lambda_{ij},$$

where λ_{ij} is the inflation coefficient.

In order to apply the influence of inflation to correct datum it is necessary to consider incremental data while correcting with inflation, i.e. in first step we apply the inflation to increments and in second step we get cumulative sums from inflation-adjusted increments.

The calculation of different reserves using chain ladder method is shown in the following examples. The key step is step 5 where the development factors and proportions are found.

Claims reserving example - paid claims

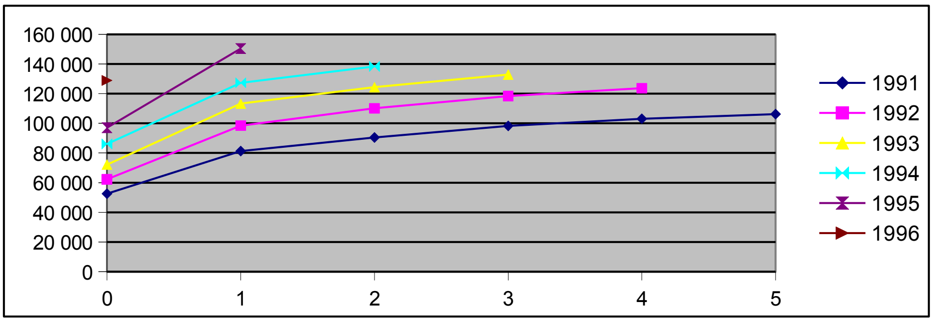
Step 1 – Incremental paid claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	52 546	28 729	9 186	7 816	4 885	3 102
1992	62 285	36 210	11 601	8 250	5 336	
1993	72 173	41 126	11 041	8 543		
1994	86 135	41 224	11 050			
1995	97 068	53 408				
1996	128 982					

Step 2 – Cumulative paid claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	52 546	81 275	90 461	98 277	103 162	106 264
1992	62 285	98 495	110 096	118 346	123 682	
1993	72 173	113 299	124 340	132 883		
1994	86 135	127 359	138 409			
1995	97 068	150 476				
1996	128 982					
Column sum	499 189	570 904	463 306	349 506	226 844	106 264
Column sum (excl. last value)	370 207	420 428	324 897	216 623	103 162	0

Step 3 – Graph



Step 4 – Individual development factors

	Development period				
Year of origin	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
1991	1,547	1,113	1,086	1,050	1,030
1992	1,581	1,118	1,075	1,045	
1993	1,570	1,097	1,069		
1994	1,479	1,087			
1995	1,550				
1996					

Step 5 – Development factors

	Development period				
	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
Development factor	1,542	1,102	1,076	1,047	1,030
	0 – ultimate	1 – ultimate	2 – ultimate	3 – ultimate	4 – ultimate
Cumulative factor	1,972	1,279	1,160	1,079	1,030
Inverse	0,507	0,782	0,862	0,927	0,971

Claims reserving example - paid claims

Step 6 – Individual development factors (step 4 extended)

Year of origin	Development period 0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
1991	1,547	1,113	1,086	1,050	1,030
1992	1,581	1,118	1,075	1,045	*1,030*
1993	1,570	1,097	1,069	*1,047*	*1,030*
1994	1,479	1,087	*1,076*	*1,047*	*1,030*
1995	1,550	*1,102*	*1,076*	*1,047*	*1,030*
1996	*1,542*	*1,102*	*1,076*	*1,047*	*1,030*

Step 7 – Cumulative paid claims (step 2 extended)

Year of origin	Development period 0	1	2	3	4	5
1991	52 546	81 275	90 461	98 277	103 162	106 264
1992	62 285	98 495	110 096	118 346	123 682	127 401
1993	72 173	113 299	124 340	132 883	139 153	143 337
1994	86 135	127 359	138 409	148 893	155 918	160 606
1995	97 068	150 476	165 823	178 383	186 799	192 416
1996	128 982	198 906	219 192	235 794	246 920	254 344

Step 8 – Estimated reserve

Year of origin	Reserve
1992	3 719
1993	10 454
1994	22 197
1995	41 940
1996	125 362
Total	**203 673**

Step 9 – Estimated ultimate claims at each development period

Year of origin	Development period 0	1	2	3	4	5
1991	103 617	103 928	104 969	106 009	106 264	106 264
1992	122 822	125 947	127 753	127 656	127 401	
1993	142 321	144 877	144 281	143 337		
1994	169 853	162 856	160 606			
1995	191 412	192 416				
1996	254 344					

Step 10 – Incremental payments

Year of origin	Development period 0	1	2	3	4	5
1991	52 546	28 729	9 186	7 816	4 885	3 102
1992	62 285	36 210	11 601	8 250	5 336	*3 719*
1993	72 173	41 126	11 041	8 543	*6 270*	*4 184*
1994	86 135	41 224	11 050	*10 484*	*7 025*	*4 688*
1995	97 068	53 408	*15 347*	*12 560*	*8 417*	*5 617*
1996	128 982	*69 924*	*20 286*	*16 602*	*11 126*	*7 425*

Claims reserving example - paid claims

Step 11 – Estimated future payments by calendar year

	Development period					
Year of origin	1997	1998	1999	2000	2001	Total
1991						
1992	3 719					3 719
1993	6 270	4 184				10 454
1994	10 484	7 025	4 688			22 197
1995	15 347	12 560	8 417	5 617		41 940
1996	69 924	20 286	16 602	11 126	7 425	125 362
Total	105 743	44 055	29 707	16 742	7 425	203 673

Step 12 – Back-fitted incremental claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	53 888	29 214	8 475	6 936	4 648	3 102
1992	64 607	35 025	10 161	8 316	5 573	
1993	72 689	39 406	11 432	9 356		
1994	81 446	44 154	12 809			
1995	97 577	52 899				
1996	128 982					

Step 13 – Residuals (actual – expected)

	Development period					
Year of origin	0	1	2	3	4	5
1991	-1 342	-485	711	880	237	0
1992	-2 322	1 185	1 440	-66	-237	
1993	-516	1 720	-391	-813		
1994	4 689	-2 930	-1 759			
1995	-509	509				
1996	0					

Example of IBNR

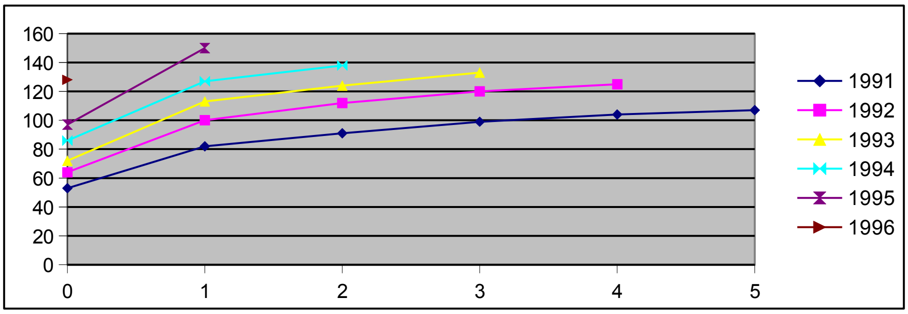
Step 1 – Incremental numbers of reported claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	53	29	9	8	5	3
1992	64	36	12	8	5	
1993	72	41	11	9		
1994	86	41	11			
1995	97	53				
1996	128					

Step 2 – Cumulative numbers of reported claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	53	82	91	99	104	107
1992	64	100	112	120	125	
1993	72	113	124	133		
1994	86	127	138			
1995	97	150				
1996	128					
Column sum	500	572	465	352	229	107
Column sum (excl. last value)	372	422	327	219	104	0

Step 3 Graph



Step 4 – Individual development factors

	Development period				
Year of origin	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
1991	1,547	1,110	1,088	1,051	1,029
1992	1,563	1,120	1,071	1,042	
1993	1,569	1,097	1,073		
1994	1,477	1,087			
1995	1,546				
1996					

Step 5 – Development factors

	Development period				
	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
Development factor	1,538	1,102	1,076	1,046	1,029
	0 – ultimate	1 – ultimate	2 – ultimate	3 – ultimate	4 – ultimate
Cumulative factor	1,962	1,276	1,158	1,076	1,029

Example of IBNR

Inverse	0,510	0,784	0,864	0,930	0,972

Step 6 – Individual development factors (step 4 extended)

	Development period				
Year of origin	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
1991	1,547	1,110	1,088	1,051	1,029
1992	1,563	1,120	1,071	1,042	*1,029*
1993	1,569	1,097	1,073	*1,046*	*1,029*
1994	1,477	1,087	*1,076*	*1,046*	*1,029*
1995	1,546	*1,102*	*1,076*	*1,046*	*1,029*
1996	*1,538*	*1,102*	*1,076*	*1,046*	*1,029*

Step 7 – Cumulative reported claims (step 2 extended)

	Development period					
Year of origin	0	1	2	3	4	5
1991	53	82	91	99	104	107
1992	64	100	112	120	125	129
1993	72	113	124	133	139	143
1994	86	127	138	149	155	160
1995	97	150	165	178	186	191
1996	128	197	217	233	244	251

Step 8 – Estimated reserve

Year of origin	Number (not reported)	Average claim	Reserve
1992	4	989	3 956
1993	10	999	9 990
1994	22	1 003	22 066
1995	41	1 003	41 123
1996	123	1 008	123 984
Total	**200**		**201 119**

Step 9 – Estimated ultimate number of claims at each development period

	Development period					
Year of origin	0	1	2	3	4	5
1991	104	105	105	107	107	107
1992	126	128	130	129	129	
1993	141	144	144	143		
1994	169	162	160			
1995	190	191				
1996	251					

Step 10 – Incremental numbers of claims reporting

	Development period					
Year of origin	0	1	2	3	4	5
1991	53	29	9	8	5	3
1992	64	36	12	8	5	*4*
1993	72	41	11	9	*6*	*4*
1994	86	41	11	*11*	*7*	*4*
1995	97	53	*15*	*13*	*8*	*5*
1996	128	*69*	*20*	*17*	*11*	*7*

Example of IBNR

Step 11 – Estimated reporting of claims by calendar year

	Development period					
Year of origin	1997	1998	1999	2000	2001	Total
1991						
1992	4					4
1993	6	4				10
1994	11	7	4			22
1995	15	13	8	5		41
1996	69	20	17	11	7	123
Total	104	43	29	16	7	200

Step 12 – Back-fitted incremental claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	55	29	9	7	5	3
1992	66	35	10	8	5	
1993	73	39	11	9		
1994	81	44	13			
1995	98	52				
1996	128					

Step 13 – Residuals (actual – expected)

	Development period					
Year of origin	0	1	2	3	4	5
1991	-2	0	0	1	0	0
1992	-2	1	2	0	0	
1993	-1	2	0	0		
1994	5	-3	-2			
1995	-1	1				
1996	0					

Claims reserving example - inflation-adjusted paid claims

Step 1 – incremental paid claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	52 546	28 729	9 186	7 816	4 885	3 102
1992	62 285	36 210	11 601	8 250	5 336	
1993	72 173	41 126	11 041	8 543		
1994	86 135	41 224	11 050			
1995	97 068	53 408				
1996	128 982					

Inflation rates

Period	Inflation
1991-92	12,4%
1992-93	22,0%
1993-94	21,9%
1994-95	15,9%
1995-96	13,2%

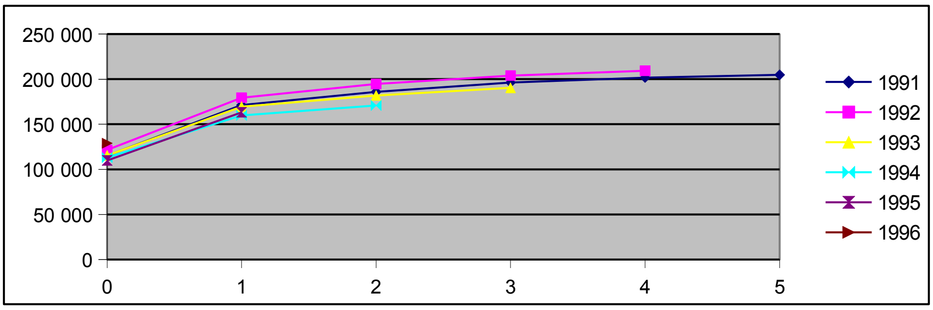
Step 1A – inflation-adjusted incremental paid claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	115 239	56 055	14 691	10 254	5 530	3 102
1992	121 528	57 911	15 220	9 339	5 336	
1993	115 427	53 957	12 498	8 543		
1994	113 008	46 666	11 050			
1995	109 881	53 408				
1996	128 982					

Step 2 – Cumulative paid claims

	Development period					
Year of origin	0	1	2	3	4	5
1991	115 239	171 294	185 985	196 240	201 770	204 872
1992	121 528	179 439	194 660	203 999	209 335	
1993	115 427	169 384	181 882	190 425		
1994	113 008	159 674	170 724			
1995	109 881	163 289				
1996	128 982					
Column sum	704 065	843 080	733 251	590 664	411 104	204 872
Column sum (excl. last value)	575 083	679 791	562 527	400 238	201 770	0

Step 3 – Graph



Claims reserving example - inflation-adjusted paid claims

Step 4 – Individual development factors

	Development period				
Year of origin	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
1991	1,486	1,086	1,055	1,028	1,015
1992	1,477	1,085	1,048	1,026	
1993	1,467	1,074	1,047		
1994	1,413	1,069			
1995	1,486				
1996					

Step 5 – Development factors

	Development period				
	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
Development factor	1,466	1,079	1,050	1,027	1,015
	0 – ultimate	1 – ultimate	2 – ultimate	3 – ultimate	4 – ultimate
Cumulative factor	1,732	1,181	1,095	1,043	1,015
Inverse	0,577	0,847	0,913	0,959	0,985

Step 6 – Individual development factors (step 4 extended)

	Development period				
Year of origin	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5
1991	1,486	1,086	1,055	1,028	1,015
1992	1,477	1,085	1,048	1,026	*1,015*
1993	1,467	1,074	1,047	*1,027*	*1,015*
1994	1,413	1,069	*1,050*	*1,027*	*1,015*
1995	1,486	*1,079*	*1,050*	*1,027*	*1,015*
1996	*1,466*	*1,079*	*1,050*	*1,027*	*1,015*

Step 7 – Cumulative paid claims in current money (step 2 extended)

	Development period					
Year of origin	0	1	2	3	4	5
1991	115 239	171 294	185 985	196 240	201 770	204 872
1992	121 528	179 439	194 660	203 999	209 335	212 553
1993	115 427	169 384	181 882	190 425	195 595	198 602
1994	113 008	159 674	170 724	179 263	184 130	186 960
1995	109 881	163 289	176 130	184 940	189 961	192 881
1996	128 982	189 089	203 960	214 161	219 975	223 357

Step 8 – Estimated reserve

	Reserve		
Year of origin	Current money (1996)	Nominal	Net present value
1992	3 218	3 605	3 400
1993	8 177	9 562	8 820
1994	16 237	19 646	17 795
1995	29 592	37 082	32 966
1996	94 375	115 415	103 845
Total	**151 600**	**185 310**	**166 826**

Claims reserving example - inflation-adjusted paid claims

Step 9 – Estimated ultimate claims at each development period

	Development period					
Year of origin	0	1	2	3	4	5
1991	199 559	202 337	203 673	204 666	204 872	204 872
1992	210 450	211 958	213 173	212 758	212 553	
1993	199 885	200 081	199 180	198 602		
1994	195 695	188 611	186 960			
1995	190 280	192 881				
1996	223 357					

Step 10 – Incremental payments (in 1996 money)

	Development period					
Year of origin	0	1	2	3	4	5
1991	115 239	56 055	14 691	10 254	5 530	3 102
1992	121 528	57 911	15 220	9 339	5 336	*3 218*
1993	115 427	53 957	12 498	8 543	*5 170*	*3 007*
1994	113 008	46 666	11 050	*8 539*	*4 867*	*2 831*
1995	109 881	53 408	*12 841*	*8 810*	*5 021*	*2 920*
1996	128 982	*60 107*	*14 870*	*10 202*	*5 814*	*3 382*

Step 10A – Incremental payments (with future inflation 12% per annum)

	Development period					
Year of origin	0	1	2	3	4	5
1991						
1992						*3 605*
1993					*5 790*	*3 772*
1994				*9 564*	*6 105*	*3 977*
1995			*14 382*	*11 051*	*7 054*	*4 595*
1996		*67 320*	*18 653*	*14 333*	*9 149*	*5 960*

Step 10B – Incremental payments (with future inflation 12%, discounted 6% per annum)

	Development period					
Year of origin	0	1	2	3	4	5
1991						
1992						*3 400*
1993					*5 462*	*3 357*
1994				*9 023*	*5 433*	*3 339*
1995			*13 568*	*9 835*	*5 923*	*3 640*
1996		*63 510*	*16 601*	*12 034*	*7 247*	*4 454*

Step 11 – Estimated future payments by calendar year (in 1996 money)

	Development period					
Year of origin	1997	1998	1999	2000	2001	Total
1991						
1992	3 218					3 218
1993	5 170	3 007				8 177
1994	8 539	4 867	2 831			16 237
1995	12 841	8 810	5 021	2 920		29 592
1996	60 107	14 870	10 202	5 814	3 382	94 375
Total	89 876	31 554	18 053	8 735	3 382	151 600

Claims reserving example - inflation-adjusted paid claims

Step 12 – Back-fitted incremental claims

Year of origin	Development period 0	1	2	3	4	5
1991	118 307	55 133	13 640	9 357	5 333	3 102
1992	122 743	57 200	14 151	9 708	5 533	
1993	114 687	53 446	13 222	9 071		
1994	107 964	50 313	12 447			
1995	111 383	51 906				
1996	128 982					

Step 13 – Residuals (actual – expected)

Year of origin	Development period 0	1	2	3	4	5
1991	-3 068	922	1 052	897	197	0
1992	-1 215	711	1 069	-369	-197	
1993	741	511	-724	-528		
1994	5 044	-3 647	-1 397			
1995	-1 502	1 502				
1996	0					

11.2.2 Loss ratio and Bornhütter-Ferguson method

Besides using only claim-based data for reserve estimation it is also possible to take into account the (relevant) premiums. A simplest calculation of reserves could be made by multiplying premiums by the expected loss ratio (obviously the true loss ratio is not yet known) to get an estimate for ultimate claim amount.

Example 11.2. Assume that the expected loss ratio is 0.8 and after development period j the loss ratio based on known claims is 0.71. Then one can estimate that the remaining IBNR reserve at the end of period j should be about $0.8 - 0.71 = 0.09$, i.e. 9% of premiums.

The Bornhütter-Ferguson method gives a way of combining the prior expectation of losses provided by simple loss ratio estimates with the actual rate of emergence of claims. The estimated ultimate claims for year of origin i at the end of developent period j (i.e. using the information available at the end of development period j) is given by

$$S_i = P_i \cdot IELR_i(1 - R_j) + C_{ij},$$

where

- P_i is the premium income received;
- $IELR_i$ is the initial expected loss ratio for year of origin i;
- R_j is the proportion factor at the end of the on jth development period;
- C_{ij} is the (cumulative) amount of claims for year of origin i and development year j.

11.2.3 Chain ladder as a generalized linear model

The methods considered so far are mainly based on averages and provide only point estimates. The other weak point is that there is no way to measure the goodnes of fit of proposed models. In the following we consider certain stochastic reserve estimation methods that also allow us to estimate the variability of predicted claims reserves.

Firstly, we note that several often used and traditional actuarial methods to complete a run-off triangle can be described by one generalized linear model (GLM). Let the random variables C_{ij}, $i = 0, 1, \ldots, I$ and $j = 0, 1, \ldots, J$ denote the random variables corresponding to the claim figure for year of origin i and year of development j, as previously. We consider the following multiplicative model

$$C_{ij} \approx \alpha_i \cdot \beta_j \cdot \gamma_k, \tag{11.1}$$

where α_i is the parameter that describes the effect of year of origin i, β_j is the parameter corresponding to development year j, and γ_k describes the effect of calendar year $k = i + j$.

The expected value of C_{ij} can be given as

$$EC_{ij} = \exp\{\ln \alpha_i + \ln \beta_j + \ln \gamma_k\},$$

or, equivalently,

$$\ln EC_{ij} = \ln \alpha_i + \ln \beta_j + \ln \gamma_k,$$

so there is a *logarithmic link* function (or *log-link*).

Turns out that this simple model allows generate quite a few reserving techniques, depending on the assuptions set on distribution of the C_{ij}.

If we restrict the model (11.1) to

$$C_{ij} \sim Po(\alpha_i \beta_j),\ \gamma_k \equiv 1, \tag{11.2}$$

where C_{ij} are independent for different $i, j = 1, \ldots, I;\ i+j \leq I$, then we obtain a model behaving as the chain ladder. Indeed, if the parameters $\alpha_i > 0$ and β_j are to be estimated by maximum likelihood, then model (11.2) is a multiplicative GLM with log-link. Also the (Poisson) distributional assumption allows to find estimates for other characteristics of the predicted claims reserves, such as variance.

It can be shown that the optimal parameters α_i and β_j produced by this GLM are equal to the estimates obtained by the chain ladder method.

11.2.4 Mack's stochastic model behind the chain ladder

In the previous subsection we saw that the chain ladder method can be considered as an algorithm to estimate the parameters of a simple GLM with two factors (year of origin and development year). In 1993 Thomas Mack describes 'the' stochastic model behind chain ladder as a different set of assumptions under which doing these calculations makes sense. As the proposed model is distribution-free, the maximum likelihood approach cannot be used, instead certain unbiased linear estimators for the mean squared errors are found. Let us recall that the *mean squared error* (MSE) of an estimate is the variance of the estimate plus the square of its bias. Therefore, if an estimate is unbiased, its MSE is equal to its variance, which makes MSE a convenient choice to model the variability of the predictions.

The Mack's model for chain ladder is described by three assumptions:

(A1) there exist proportion factors $f_j > 0$ such that for all $i \in \{0, 1, \ldots, I\}$ and $j \in \{0, 1, \ldots, J-1\}$ the conditional expectation for cumulative claims for the development period $j+1$ can be calculated by multiplying the cumulative claims for development period j by the proportion factor f_j:

$$E(C_{i,j+1}|C_{i1},\ldots,C_{ij}) = f_j C_{ij};$$

(A2) for different years of origin i and k the random variables $\{C_{i0},\ldots,C_{iJ}\}$ and $\{C_{k0},\ldots,C_{kJ}\}$ are independent;

(A3) there exist constants $\sigma_j^2 > 0$ such that for all $i = 0,1,\ldots,I$ and $j = 0,1,\ldots,J-1$ the following equality holds

$$Var(C_{i,j+1}|C_{i1},\ldots,C_{ij}) = \sigma_j^2 C_{ij}.$$

The assumption A1 establishes the existence of proportion factors and depicts the very essence of the chain ladder method.

The independence assumption A2 follows from the fact that the chain ladder algorithm does not take into account any dependencies between years of origin.

The variance assumption A3 follows from the fact that the proportion factor f_j is the C_{ij}-weighted mean of the individual development factors $\frac{C_{i,j+1}}{C_{ij}}$ and thus the variance $Var(\frac{C_{i,j+1}}{C_{ij}}|C_{i1},\ldots,C_{ij})$ should be inversely proportional to C_{ij}.

It can be shown that (Mack, 1993):

- the estimates of parameters f_j calculated by chain ladder method are unbiased;
- there exist unbiased estimators for the variance factors σ_j^2.

In conclusion, under assumptions A1–A3 there can be found formulas for mean squared errors of the claim reserve amounts that only use data from the chain ladder triangle. This allows to calculate the variability of the proposed predictions. The exact calculations and formulas are quite technical and are thus omitted. But it is important to remember that the constructed stochastic model

- formalizes and extends the idea of chain ladder method;
- gives exactly the same point estimates as the chain ladder method;
- gives an additional possibility to find certain interval estimates for the predicted claims reserves.

11.2.5 Chain ladder bootstrap

Bootstrapping is a simple and powerful method which enables the calculation of a number of different estimates of a random variable, using empirical data as an approximation of the true distribution. In other words, no distributional assumptions are made, the only assumption is that available data is representative enough for the underlying population.

In this section we briefly introduce the idea of bootstrap method and show how it can be applied to calculated errors associated with the predicted claims reserves.

Let us assume that the chain ladder development factors are found and the ultimate claims are estimated. We can now back-fit the claims in the run-off triangle based on the ultimate claims and development factors to obtain the run-off triangle (say, with components $\hat{C}_{ij}$) which follows exactly the proposed model:

	Development period j				
Year of origin i	0	1	2	...	J
0	$\hat{C}_{00}$	$\hat{C}_{01}$	$\hat{C}_{02}$	...	$\hat{C}_{0J}$
1	$\hat{C}_{10}$	$\hat{C}_{11}$	...		
2	$\hat{C}_{20}$	...			
...	...				
I	$\hat{C}_{I0}$				

Here $\hat{C}_{ij}$ is calculated from

$$\hat{C}_{ij} = \frac{\hat{S}_i}{\hat{f}_j \cdot \ldots \cdot \hat{f}_{J-1}},$$

where $\hat{S}_i$ and $\hat{f}_j$ denote estimated values for ultimate claims for year of origin i and proportion factor between development years j and $j+1$, respectively.

The cumulative back-fitted claims are found as differences between actual and expected values:

Year of origin i	Development period j				
	0	1	2	...	J
0	$C_{00} - \hat{C}_{00}$	$C_{01} - \hat{C}_{01}$	$C_{02} - \hat{C}_{02}$	...	$C_{0J} - \hat{C}_{0J}$
1	$C_{10} - \hat{C}_{10}$	$C_{11} - \hat{C}_{11}$	...		
2	$C_{20} - \hat{C}_{20}$	...			
...	...				
I	$C_{I0} - \hat{C}_{I0}$				

If we now move from cumulative differences to incremental differences to remove the cumulative effect, we obtain a run-off triangle which consists of the *residual errors.* The residual errors (or simply *residuals*) give us information how well the model fits to the known data and allows to make predictions about the future fit.

We now make assumption that the residual errors are random. This assumption should actually be tested to ensure there is no systematic component, but it is consistent with the chain ladder model. The significance of this assumption for the bootstrap method is that each error could equally well have arisen as the residual error from any other development period and year of origin.

Now we are ready to formulate the chain ladder bootstrap algorithm:

Step 1: fix the number of repetitions n.

Step 2: for each repetition $i = 1, \ldots, n$:

- 2.1) draw a random sample (allowing for replacement) from the set of residual errors (in other words, we produce an alternative set of equally likely outcomes, called the *pseudo data*);
- 2.2) create incremental pseudo claims by adding the pseudo residuals to actual claims;
- 2.3) cumulate the incremental pseudo claims;
- 2.4) apply the chain ladder method to cumulative claims as normally to obtain a reserve estimate.

Step 3: now we have n estimates for the reserves, which, by the assumptions provide a random sample of the distribution of the true reserve value. This sample can be used to calculate estimates for different characteristics of the distribution (e.g. variance).

In case the incremental claims are affected by inflation then the data should be adjusted by inflation before applying the bootstrap procedure.

References

1. Booth P., Chadburn, R., Cooper, D., Haberman, S., James, D. (1999) *Modern Actuarial Theory and Practice.* Chapman & Hall / CRC.

2. Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M. (2008) *Modern Actuarial Risk Theory: Using R.* Springer, Berlin/Heidelberg.

3. Mack, T. (1993) Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *Astin Bulletin*, **23**, 213–225.

12 Using individual history in premium calculation

12.1 Bayesian credibility theory

Let us consider the following motivating problem.

Example 12.1. An insurance company has a long-time customer with average claim amount per year $\bar{x}$ (e.g., $\bar{x} = 1600$). On the other hand, the estimated average claim for similar risks in the insurer's portfolio is μ (e.g. $\mu = 2500$). How much premium should the insurer ask from this customer? 1600? 2500? Or something in between?

To answer this question we need to estimate the *credibility* of the customer (or customer data), which can be mathematically described by the *credibility factor* z.

Let $z = 0.6$, then the (pure) premium is found as

$$P = 0.6 \cdot 1600 + 0.4 \cdot 2500 = 1960.$$

In general, the premium is found as a weighted average

$$P = z \cdot \bar{x} + (1 - z)\mu,$$

where

- $\bar{x}$ – estimated claim amount (or number) for some fixed risk;
- μ – reference value, expected claim amount (or number) obtained by analysis of similar risks (collateral information);
- z – credibility factor, $0 \leq z \leq 1$.

What properties should the credibility factor z have?

It should increase if

- there is more data about the particular risk;
- the precision of individual risks becomes greater.

It should decrease if

- there is more collateral information;
- the collateral risk becomes more relevant.

12.1.1 Poisson/gamma model

Suppose that the number of claims in a portfolio is Poisson distributed with parameter λ, where the value of λ is a realization of a gamma distributed random variable Λ ($\Lambda \sim \Gamma(\alpha, \beta)$). Suppose also that for some risk the claim numbers in past k years are known: $n_1, \ldots, n_k$.

The problem is to estimate based on this information (prior distribution and history) the (conditional!) expected claim number for the next year:

$$E(E(N|\Lambda)|n_1, \ldots, n_k) = E(\Lambda|n_1, \ldots, n_k).$$

Notice that this problem is closely related to the setup of Bayesian statistics: we need to estimate the *posterior distribution* of Λ based on the *prior distribution* of Λ and the observations $n_1, \ldots, n_k$.

By Bayes Theorem the distribution of $(\Lambda|n_1, \ldots, n_k)$ is $\Gamma(\alpha+\sum_{i=1}^k n_i, \beta+k)$ (check it!).

We are mostly interested in its expectation:

$$E(\Lambda|n_1, \ldots, n_k) = \frac{\alpha + \sum_{i=1}^k n_i}{\beta + k} = \frac{\beta}{\beta + k} \cdot \frac{\alpha}{\beta} + \frac{k}{\beta + k} \cdot \frac{\sum_{i=1}^k n_i}{k}.$$

Taking $z = \frac{k}{\beta+k}$, we can write

$$E(\Lambda|n_1, \ldots, n_k) = z\frac{\sum_{i=1}^k n_i}{k} + (1-z)\frac{\alpha}{\beta}.$$

Recall that we have initial estimates for the expectation based on two sources:

- $\mu = \frac{\alpha}{\beta}$ (based on $\Gamma(\alpha, \beta)$);
- $\bar{n} = \frac{\sum_{i=1}^k n_i}{k}$ (based on individual history).

Then the conditional expectation can be written as

$$E(\Lambda|n_1, \ldots, n_k) = z \cdot \bar{n} + (1-z)\mu,$$

i.e. it is a weighted average of the initial estimates.

It is easy to see that always $0 \leq z \leq 1$ and the more history we have the higher z.

12.1.2 Normal/normal model

Suppose that claim sizes X are normally distributed ($X \sim N(\theta, \sigma_1)$), where θ is fixed, but unknown (and σ_1 is assumed to be known). Let us also have claims information for some risk in past k years: $x_1, \ldots, x_k$.

Supposing now that the collateral information gives prior $\Theta \sim N(\mu, \sigma_2)$ distribution for θ, we have the following problem of Bayesian statistics: estimate the posterior distribution of Θ based on its prior distribution and observations $x_1, \ldots, x_k$.

By Bayes Theorem the posterior distribution of Θ is

$$N\left(\frac{\mu\sigma_1^2 + k\sigma_2^2\bar{x}}{\sigma_1^2 + k\sigma_2^2}, \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + k\sigma_2^2}\right),$$

where $\bar{x} = \frac{\sum_{i=1}^k x_i}{k}$.

Thus, the conditional expectation is

$$\begin{aligned} E(E(X|\Theta)|x_1, \ldots, x_k) &= E(\Theta|x_1, \ldots, x_k) \\ &= \frac{\mu\sigma_1^2 + k\sigma_2^2\bar{x}}{\sigma_1^2 + k\sigma_2^2} = \frac{\sigma_1^2}{\sigma_1^2 + k\sigma_2^2}\mu + \frac{k\sigma_2^2}{\sigma_1^2 + k\sigma_2^2}\bar{x}. \end{aligned}$$

Taking

$$z = \frac{k\sigma_2^2}{\sigma_1^2 + k\sigma_2^2} = \frac{k}{k + \frac{\sigma_1^2}{\sigma_2^2}},$$

the posterior expectation can be written as a weighted average of prior expectation μ and the mean value of observed claims $\bar{x}$:

$$E(\Theta|x_1, \ldots, x_k) = z \cdot \bar{x} + (1 - z)\mu,$$

One can see that again always $0 \leq z \leq 1$ and the more history we have the higher z.

12.2 Empirical Bayesian credibility theory

In Bayesian credibility models, the information came from two sources: a priori distribution of collateral data describing the between risk information and the individual history. In empirical Bayesian credibility theory (EBCT), the general idea stays the same, we simply drop the distributional assumption set to collateral data and study certain distribution-free models.

12.2.1 The Bühlmann credibility model

Let us consider a portfolio with n risks and let X_{ij}, $j = 1, \ldots, k$ denote the (aggregate) claim amounts (or numbers) that arise from risk i for k successive years.

Now, let us consider a similar model to Bayesian credibility model, where the distributional assumptions are replaced with certain structural assumptions:

- the distribution of each X_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, k$ depends on the value of an unknown parameter Θ_i, which is assumed to be an (unknown) random variable, but its value θ_i is fixed for given risk;
- given θ_i, the random variables X_{ij}, $j = 1, \ldots, k$ are i.i.d.;
- for different risks i and i' ($i \neq i'$), the pairs of variables (Θ_i, X_{ij}) and $(\Theta_{i'}, X_{i'j'})$, $j, j' = 1, \ldots, k$, are i.i.d. This assumption allows us to write simply Θ instad of Θ_i, since the distribution of Θ_i does not depend on i.
- the conditional mean and variance of X_{ij} are

$$E(X_{ij}|\Theta_i) = \mu(\Theta_i) = \mu(\Theta),$$
$$Var(X_{ij}|\Theta_i) = \sigma^2(\Theta_i) = \sigma^2(\Theta).$$

Now, the unconditional mean and variance for X_{ij} are calculated as

$$EX_{ij} = E(E(X_{ij}|\Theta)) = E\mu(\Theta),$$
$$VarX_{ij} = E(Var(X_{ij}|\Theta)) + Var(E(X_{ij}|\Theta)) = E\sigma^2(\Theta) + Var(\mu(\Theta)),$$

where the decomposition of variance is based on the *law of total variance*.

The variance components can be interpreted as follows:

- $E\sigma^2(\Theta)$ represents the expected value of the process variance, i.e. the *variance within risk*;
- $Var(\mu(\Theta))$ is the variance of the hypothetical means, i.e., the *variance between the risks*.

Now, assume that for past k years the claim amounts (or numbers) x_{ij} (the realizations of X_{ij}), $j = 1, \ldots, k$, for risk i are known. Then all the structural parameters, $E\mu(\Theta)$, $E\sigma^2(\Theta)$ and $Var(\mu(\Theta))$ can be estimated from the collateral data. The usual estimators are:

$$\widehat{E\mu(\Theta)} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n}\bar{x}_i = \frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k}x_{ij},$$

$$\widehat{E\sigma^2(\Theta)} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{k-1}\sum_{j=1}^{k}(x_{ij}-\bar{x}_i)^2$$

and

$$\widehat{Var(\mu(\Theta))} = \frac{1}{n-1}\sum_{i=1}^{n}(\bar{x}_i-\bar{x})^2 - \frac{1}{nk}\sum_{i=1}^{n}\frac{1}{k-1}\sum_{j=1}^{k}(x_{ij}-\bar{x}_i)^2.$$

Thus, we have again two sources of information: the individual history gives us mean value $\bar{x}_i = \frac{1}{k}\sum_{j=1}^{k} x_{ij}$ and the collateral information suggests overall mean $\bar{x}$.

Now, the formula for credibility premium for risk i in Bühlmann model can be written as

$$P_i = z\cdot\bar{x}_i + (1-z)\cdot\bar{x}$$

and the remaining question is how to calculate a proper credibility factor z.

Let us now recall the credibility factor for normal/normal model

$$z = \frac{k}{k+\frac{\sigma_1^2}{\sigma_2^2}},$$

where σ_1^2 and σ_2^2 can be interpreted as the within risk variance is the between risk variance, respectively. Turns out that similar property holds for Poisson/gamma model: the credibility factor is given by $z = \frac{k}{k+\beta}$ or, equivalently,

$$z = \frac{k}{k+\frac{\alpha/\beta}{\alpha/\beta^2}}$$

where we recognize that

- $\frac{\alpha}{\beta}$ is the expectation (and hence also the variance) of the Poisson distribution, so it describes the within risk variance;
- $\frac{\alpha}{\beta^2}$ is the variance of the gamma distribution, i.e. the between risk variance.

These expressions for credibility factor suggest the following formula for the credibility factor in empirical setup:

$$z = \frac{k}{k+\frac{\widehat{E\sigma^2(\Theta)}}{\widehat{Var(\mu(\Theta))}}}.$$

Notice that although $E\sigma^2(\Theta)$ and $Var(\mu(\Theta))$ are always positive, the proposed estimator for $Var(\mu(\Theta))$ can take negative values, as well. In that case, one can just take the corresponding estimate equal to zero, which implies the credibility factor z is also equal to zero (basically, this assumption means that there is no between risks variance, so the collateral information can be fully trusted). So, the value of credibility factor z will always satisfy $0 \leq z \leq 1$.

In conclusion, in the Bühlmann credibility model

- the credibility factor z is the same for all risks in the collective, it only has to be calculated once;
- the more information concerning individual risk (the bigger k) the bigger is z;
- large values of $E\sigma^2(\Theta)$ correspond to large variability from year to year within risks and implies smaller value of credibility factor z;
- large values of $Var(\mu(\Theta))$ mean large variability between risks and thus the data from other risks is not very informative or relevant, implying higher credibility factor (i.e. the data from an individual risk is more important).

12.2.2 The Bühlmann-Straub model

We saw that the Bühlmann model provides a convenient way to model a heterogeneous portfolio. At the same time, the Bühlmann model has a serious limitation that the the claim amounts (or numbers) for a risk are identically distributed for all years of exposure. In practice, this assumption is often violated. This issue is addressed in the Bühlmann-Straub model: the assumption of i.i.d. claims is relaxed by introducing an additional *exposure* or *risk volume* parameter.

In particular, let us consider again a portfolio with n risks and k successive years. Let Y_{ij} denote the aggregate claim amount for risk i in year j and let us introduce the claim amount per unit of risk volume (or per unit of exposure)

$$X_{ij} = \frac{Y_{ij}}{v_{ij}},$$

where v_{ij} represents the the risk volume. The risk volumes v_{ij} are assumed to be known and fixed for each risk i and year j.

Example 12.2. In practice, several different risk volumes are used, depending on the line of business and risk specifics:

- number of years at risk in motor insurance;
- total amount of wages in the collective health or collective accident insurance;
- sum insured in fire insurance;
- annual turnover in commercial liability insurance;
- annual premium (written or earned) by the ceding company in excess of loss reinsurance.

Similarly to Bühlmann model, a random variable X_{ij} is assumed to depend from a parameter Θ_i, which is unknown, but its value θ_i is fixed for each risk i .

The general assumptions of the Bühlmann-Straub model are:

- given θ_i, the random variables X_{ij}, $j = 1, \dots, k$, are independent;
- $E(X_{ij}|\Theta_i) =: \mu(\Theta_i)$ does not depend on j;
- $v_{ij} Var(X_{ij}|\Theta_i) =: \sigma^2(\Theta_i)$ does not depend on j;
- for different risks i and i' $(i \neq i')$, the pairs of variables (Θ_i, X_{ij}) and $(\Theta_{i'}, X_{i'j'})$, $j, j' = 1, \dots, k$, are i.i.d.;
- the risk parameters Θ_i, $i = 1, \dots, n$ are i.i.d.

The last assumption allows us to write simply Θ instad of Θ_i, since the distribution of Θ_i does not depend on i.

Then, the credibility premium for risk i can be found as

$$P_i = z_i \cdot \bar{x}_i + (1 - z_i) \cdot \widehat{E\mu(\Theta)},$$

where the meaning of $\bar{x}_i$ and the calculation of an estimate for $E\mu(\Theta)$ are slightly changed compared to Bühlmann model:

$$\bar{x}_i = \frac{\sum_{j=1}^k v_{ij} x_{ij}}{\sum_{j=1}^k v_{ij}} = \frac{\sum_{j=1}^k y_{ij}}{\sum_{j=1}^k v_{ij}}$$

and

$$\widehat{E\mu(\Theta)} = \bar{x} = \frac{\sum_{i=1}^n \sum_{j=1}^k v_{ij} x_{ij}}{\sum_{i=1}^n \sum_{j=1}^k v_{ij}} = \frac{\sum_{i=1}^n \sum_{j=1}^k y_{ij}}{\sum_{i=1}^n \sum_{j=1}^k v_{ij}},$$

where x_{ij} and y_{ij} are the realizations of X_{ij} and Y_{ij}, respectively.

Similarly, the estimates for variance components are now adjusted by risk volume as follows:

$$\widehat{E\sigma^2(\Theta)} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{k-1}\sum_{j=1}^{k} v_{ij}(x_{ij}-\bar{x}_i)^2$$

and

$$\widehat{Var(\mu(\Theta))} = \frac{1}{v}\left[\frac{1}{nk-1}\sum_{i=1}^{n}\sum_{j=1}^{k} v_{ij}(x_{ij}-\bar{x})^2 - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{k-1}\sum_{j=1}^{k} v_{ij}(x_{ij}-\bar{x}_i)^2\right],$$

where

$$v = \frac{1}{nk-1}\sum_{i=1}^{n}\sum_{j=1}^{k} v_{ij}\left(1-\frac{\sum_{j=1}^{k} v_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{k} v_{ij}}\right).$$

Now, the credibility factor is given by

$$z_i = \frac{\sum_{j=1}^{k} v_{ij}}{\sum_{j=1}^{k} v_{ij} + \frac{\widehat{E\sigma^2(\Theta)}}{\widehat{Var(\mu(\Theta))}}}.$$

By similar argumentation as before, we can see that $0 \leq z_i \leq 1$. Also, the general idea underlying this model is the same as in Bühlmann model. The main difference in the calculation of credibility factor is that the number of years k is changed to the aggregate risk volume, which is obviously a more informative characteristic. The calculation of estimates for structural parameters is also slightly changed, but their meaning is the same: $E\sigma^2(\Theta)$ describes the within risk variance and $Var(\mu(\Theta))$ is the between risks variance.

To sum up, we also mention the following:

- in case all the volumes v_{ij} are equal to 1, the Bühlmann-Straub model reverts to the Bühlmann model;
- z_i is increasing function of aggregate risk volume for risk i: higher risk volume implies higher credibility for that risk;
- although we formulated all the results assuming the variables Y_{ij} represent claim amounts, the same model can be applied to claim numbers, as well;
- the Bühlmann-Straub model is by far the most used credibility model in insurance practice.

12.3 Bonus-malus systems (No Claims Discount systems)

We consider the following situation

- insurer has sufficient history of its customers' (policyholders') claim behaviour;
- insurer is interested in keeping its "good" customers (who cause no or few claims);
- insurer is willing to offer certain discount to keep the good customers.

The main question we search answers for is how to construct a suitable discount system.

Example 12.3 (A three-level NCD system)**.** Let us consider the following simple system:

level	discount
0	0%
1	25%
2	40%

The rules how a policyholder can reach different discount levels will be specified later on.

In general, let us have n discount levels (categories) and let π_i denote the proportion of policyholders at i-th level ($\sum_i \pi_i = 1$). To give a formal description of an NCD system we need to recall some properties and results related to Markov chains.

Let us denote

$$\vec{\pi} = (\pi_0, \pi_1, \dots, \pi_n)$$

and

$$P = (p_{ij}) = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0n} \\ p_{10} & p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n0} & p_{n1} & \cdots & p_{nn} \end{pmatrix},$$

where p_{ij} are transition probabilities, i.e. probabilities that (based on the claim behaviour of the next year) a policyholder will move from discount level i to discount level j. Matrix P is called the *transition matrix.*

An important task for insurer is to estimate the vector $\vec{\pi}$. In case there exists a *steady state*, i.e. $\vec{\pi} = \vec{\pi} \cdot P$, we can find $\vec{\pi}$ by solving the corresponding system of equations.

Example 12.4. Let us now extend the NCD system described in previous example by the following rules:

- in case of a claim-free year the policyholder moves in the next year to the next higher level of discount (if possible);
- in case of claim(s) the policyholder moves in the next year to next lower level of discount (if possible).

Let the probability of a loss be 0.1. The corresponding transition matrix is:

$$P = \begin{pmatrix} 0.1 & 0.9 & 0 \\ 0.1 & 0 & 0.9 \\ 0 & 0.1 & 0.9 \end{pmatrix}$$

and $\vec{\pi} = \vec{\pi} \cdot P$ gives

$$\begin{cases} 0.1\pi_0 + 0.1\pi_1 = \pi_0 \\ 0.9\pi_0 + 0.1\pi_2 = \pi_1 \\ 0.9\pi_1 + 0.9\pi_2 = \pi_2 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases},$$

which has solution $\pi_0 = \frac{1}{91}$, $\pi_1 = \frac{9}{91}$ ja $\pi_2 = \frac{81}{91}$.

In practice the assumption that all customers have the same loss probability is obviously oversimplified. Still, the customers can be divided by their claim behaviour into "good" and "bad" categories and the loss probability can be estimated in those categories separately.

Let us extend the example even further Let the probability of "good" drivers having an accident be 0.1 and for "bad" drivers 0.2. Then the steady state distribution for "good" drivers is given in previous example and for "bad" drivers the proportions are (check it!): $\pi_0 = \frac{1}{21}$, $\pi_1 = \frac{4}{21}$ and $\pi_2 = \frac{16}{21}$.

Suppose that the full (individual) premium c. Then the average pure premium paid by a "good" driver is

$$\frac{1}{91}c + \frac{9}{91} \cdot 0.75c + \frac{81}{91} \cdot 0.6c = 0.619c$$

and the average pure premium paid by a "bad" driver is

$$\frac{1}{21}c + \frac{4}{21} \cdot 0.75c + \frac{16}{21} \cdot 0.6c = 0.648c.$$

Thus, in spite the fact that "bad" drivers are twice as likely to claim as "good" drivers, the premium they pay is only marginally higher (on average).

The following example shows how a discount system may affect the policyholders decision whether to make the claim. In other words, the probability that there incurs a loss is not the same as the probability that a clame is made!

Example 12.5. Consider the same 3-level discount system (0%, 25% ja 40%), where policyholders move between discount levels as specified in previous examples. Let the policyholders be divided into two groups: "good" drivers and "bad" drivers with accident probabilities 0.1 and 0.2, respectively. The probability of any driver having two or more accidents is so small it may be assumed to be zero. The cost of repair following an accident has lognormal distribution with parameters $\mu = 5$ and $\sigma = 2$ (in EUR). The annual premium for policyholders (without discount) is 500 EUR. A policyholder makes a claim following an accident only if the cost of the repair is greater than the win from the discounts of premium in the following three years in case no claim is made.

For each level of discount, calculate

(a) the cost of repair below which a policyholder will not claim;

(b) the probability that a policyholder will make a claim following an accident;

(c) the proportions of "good" and "bad" drivers assuming these proportions have reached a steady state.

(a) A policyholder will decide whether to make a claim following an accident comparing the sums the premiums payable in the next three years

(i) given that a claim is made for the cost of repair (and assuming no claims in the following years);

(ii) given that no claim is made for the cost of this repair (and assuming no claims in the following years).

A policyholder will make a claim only if the cost of repair is greater than difference between (i) and (ii), i.e.

- for discount level 0%: $(500 + 375 + 300) - (375 + 300 + 300) = 200$;
- for discount level 25%: $(500 + 375 + 300) - (300 + 300 + 300) = 275$;
- for discount level 40%: $(375 + 300 + 300) - (300 + 300 + 300) = 75$.

(b) Let X denote the loss size, then $\ln X \sim N(5, 2)$. The claim will only be made if it exceed certain threshold x, i.e.

$$\mathbf{P}\{X > x\} = \mathbf{P}\{\ln X > \ln x\} = 1 - \Phi\left(\frac{\ln x - 5}{2}\right).$$

The threshold x is different for each level of discount and found in (a).

Thus the claim probabilities in case of accident are:

- for discount level 0%: $1 - \Phi\left(\frac{\ln 200 - 5}{2}\right) = 1 - \Phi(0.149) = 0.441$;
- for discount level 25%: $1 - \Phi\left(\frac{\ln 275 - 5}{2}\right) = 1 - \Phi(0.308) = 0.379$;
- for discount level 40%: $1 - \Phi\left(\frac{\ln 75 - 5}{2}\right) = 1 - \Phi(-0.341) = 0.633$.

(c) Let us consider "good" and "bad" drivers separately. Since

$$\mathbf{P}\{\text{claim}\} = \mathbf{P}\{\text{claim}|\text{accident}\} \cdot \mathbf{P}\{\text{accident}\},$$

the transition matrix for "good" drivers is

$$P = \begin{pmatrix} 0.0441 & 0.9559 & 0 \\ 0.0379 & 0 & 0.9621 \\ 0 & 0.0633 & 0.9367 \end{pmatrix}.$$

At the steady state we obtain the following system of equations

$$\begin{cases} 0.0441\pi_0 + 0.0379\pi_1 = \pi_0 \\ 0.9559\pi_0 + 0.0633\pi_2 = \pi_1 \\ 0.9621\pi_1 + 0.9367\pi_2 = \pi_2 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases},$$

which has solution (check it!) $\pi_0 = 0.0024$, $\pi_1 = 0.0616$ ja $\pi_2 = 0.9360$.
The transition matrix for "bad" drivers is

$$P = \begin{pmatrix} 0.0882 & 0.9118 & 0 \\ 0.0758 & 0 & 0.9242 \\ 0 & 0.1266 & 0.8734 \end{pmatrix}$$

and the solution of steady state system gives

- $\pi_0 = 0.0099$,
- $\pi_1 = 0.1193$,
- $\pi_2 = 0.8708$.

References

1. Booth P., Chadburn, R., Cooper, D., Haberman, S., James, D. (1999) *Modern Actuarial Theory and Practice.* Chapman & Hall / CRC.
2. Bühlmann, H., Gisler, A. (2005) *A Course in Credibility Theory and its Applications.* Springer, Berlin/Heidelberg.

3. Gray, R.J. & Pitts, S.M. (2012) *Risk Modelling in General Insurance. From Principles to Practice.* Cambridge University Press.

4. Casualty Actuarial Society. (1998) *Foundations of Casualty Actuarial Science.* Casualty Actuarial Society.

5. Hossack, I.B., Pollard, J.H., Zehnwirth, B. (1995) *Introductory Statistics with Applications in General Insurance.* Cambridge University Press.

6. Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M. (2008) *Modern Actuarial Risk Theory: Using R.* Springer, Berlin/Heidelberg.

7. Tse, Y.-K. (2009) *Nonlife Actuarial Models.* Cambridge University Press.

13 The accounting framework

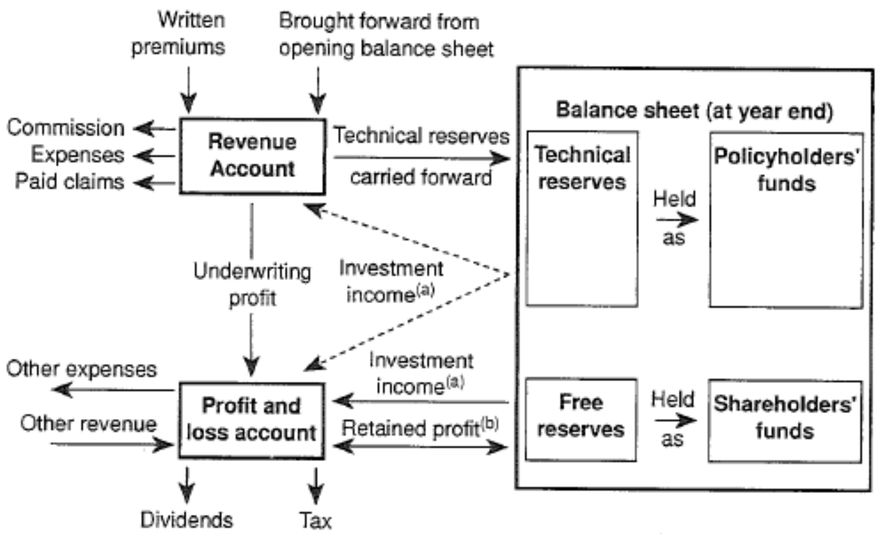
To understand the operation of a company it is important that the content of the accounts and the relationship between revenue accounts and balance sheets is clearly understood.

A brief description of key terms follows:

- the *revenue account* summarizes the cash flows incurred during the year due to the insurance business;
- the *profit and loss account* incorporates the insurance result from revenue account, as well as other cash flows (e.g. tax and dividend payments, investments, etc);
- the *balance sheet* provides a view of the total assets and the liabilities of the company at one moment in time.

The revenue and profit and loss accounts therefore provide the link between the opening (or *brought forward*) balance sheet and the closing (or *carried forward*) balance sheet.

The following figure demonstrates the connection between the various parts of the accounts and certain cash flows.



Notes:

a Investment income in respect of the technical reserves can be credited either to the revenue account or directly to the profit and loss account. If reserves are discounted explicitly, the investment return expected on the reserves is usually credited to the revenue account and the balance is credited to the profit and loss account. In this case, the balance in the revenue account labelled 'underwriting profit' is usually renamed 'insurance profit'.

b Retained profits increase the free reserves; losses reduce the free reserves.

Diagram from Taylor (1992).

13.1 The revenue account

The amount of premium included in the accounts is the earned premium. Recall that using the unearned premiums reserve (UPR), the total premium earned in an accounting year is found as:

UPR brought forward from the previous accounting year

+ the premium written in the accounting year

− the UPR carried forward to the next accounting year.

Similary, the insurer must account for the claims incurred during the year, not for the claims paid during the accounting year. Thus, using the claims reserves (RBNS and IBNR), the claims incurred are calculated as:

the claims paid

− claims reserves brought forward

+ claims reserves carried forward.

The items above, together with expenses associated with underwriting activities and commission, are combined in the revenue account to produce the *underwriting result* as follows:

earned premium

− incurred claims

− expenses and commission.

13.2 The profit and loss account

The profit and loss account is a continuation of the revenue account, reflecting those revenues and expenses not directly attributable to the underwriting activities.

A typical profit and loss account of a non-life insurance company includes the following items:

+ balance of profit and loss account brought forward

+ investment income

+ capital gains

− overhead expenses

+ exceptional items

\+ underwriting profit

= **overall result**

− tax liability

− dividend payments

= **balance of profit and loss account carried forward**

13.3 The balance sheet

The balance sheet gives a view of the financial position of the company. It is divided into two parts describing different aspects of the company's business:

- assets;
- liabilities and equity.

For a non-life insurance company, these parts are likely to include the following components.

Assets

- Investments
  - Equity
  - Fixed interest
  - Property
- Current assets
  - Debtors
  - Cash

Total assets

Liabilities and equity

- **Liabilities**
  - Technical reserves
    * RBNS carried forward
    * IBNR carried forward
    * UPR carried forward

    * DAC carried forward
    * URR carried forward
  - Other contingency reserves carried forward
  - Creditors
  - Long-term liabilities (loans, etc)

- **Equity**
  - Share capital
  - Retained earnings

Total liabilities + Equity

13.4 Key analytical statistics

Accounts can only provide summary information. They are particularly deficient in indicating the inherent uncertainty in both the amount and timing of asset receipts and liability outgo. This is a fundamental issue in insurance and to improve the presentation of information (e.g. for a supervisory perspective), certain financial characteristics of different insurers can be compared, and the performance of individual companies can be tracked through time. This includes characteristics related to premium growth, changes in the asset portfolio or in the solvency, and more:

- premium growth statistics, e.g.

$$\frac{\text{premium written in year } t}{\text{premium written in year } t-1}$$

  or

$$\frac{\text{premium written in year } t - \text{premium written in year } t-1}{\text{premium written in year } t-1};$$

- net income growth

$$\frac{\text{net income in year } t - \text{net income in year } t-1}{\text{net income in year } t-1};$$

- reinsurance related ratios

$$\frac{\text{net written premium}}{\text{gross written premium}};$$

$$\frac{\text{net paid claims}}{\text{gross paid claims}};$$

$$\frac{\text{net incurred claims}}{\text{gross incurred claims}};$$

$$\frac{\text{reinsurance debtors}}{\text{reinsurance recoveries due on paid claims}};$$

- the loss ratio
$$\frac{\text{incurred claims}}{\text{earned premiums}};$$
- the expense ratio
$$\frac{\text{expenses and commission}}{\text{written premiums}};$$
- the operating ratio
$$\text{loss ratio} + \text{expense ratio};$$
- the investment ratio
$$\frac{\text{investment returns}}{\text{written premiums}};$$
- solvency ratio
$$\frac{\text{free assets}}{\text{net written premiums}}.$$

We also mention some general ratios that are widely used in accounting framework (and are not limited to insurance business only):

- return on assets (ROA)
$$\frac{\text{net income}}{\text{total assets}}.$$
- return on equity (ROE)
$$\frac{\text{net income}}{\text{shareholder's equity}};$$
- current ratio
$$\frac{\text{current assets}}{\text{current liabilities}}.$$

13.5 A note on terminology. Premiums

Throughout this section (and the whole course) the term "premium" is used in different contexts having slightly different meaning. In the following we briefly recall the used terms for premiums in different situations.

1. **(Under)written premium:**
    - the premium charged (or to be charged) for a policy or group of policies;
    - fixed (usually) while signing the contract;
    - related date – starting date of the contract.
2. **Earned premium**
    - the premium recognized as revenue for the period in question;
    - requires implicit or explicit statement of the corresponding block of time.
3. **Accounted premium**
    - takes into account sent out invoices;
    - describes the premium volume of a portfolio without specifying the period when the contracts are signed;
    - related date – date when the invoice is issued.
4. **Received premium**
    - actual payments received on insurance company's bank account;
    - related to cash-flow;
    - related date – payment date.

In case of reinsurance, we may also categorize premiums as follows.

- **Direct premiums** – premiums arising from policies covering the customer's risks, other than risks from the customer's insurance/reinsurance or retrocession policies when the customer is an insurer (or reinsurer).
- **Reinsurance premiums** – premiums arising from policies covering risks from direct insurance policies;
- **Retrocession premiums** – premiums arising from policies covering risks from reinsurance policies.

References

1. Booth P., Chadburn, R., Cooper, D., Haberman, S., James, D. (1999) *Modern Actuarial Theory and Practice.* Chapman & Hall / CRC.

14 Solvency II

14.1 Background. Goals. Requirements

In general terms, *solvency* is the state of having more assets than liabilities. To be *solvent* means that one is in the state of solvency.

Currently (as of 2013), the regulations in insurance market are specified in the Solvency I directive, which was introduced already in 1973. The economy and risk management systems are changed a lot during last years, thus the Solvency I regime is outdated in many ways (it is static, volume-based, has so-called accounting view and does not take risk behaviour and economic factors into account). Therefore, in order to regulate the insurance market more effectively and to enhance the customer protection, a new dynamic regulation Solvency II is proposed and developed. Solvency II has so-called economic view, takes into account the current risk management strategies and allows much more dynamic capital allocation. Also, since the requirements specified in Solvency I proved to be not sufficient in many cases, several EU member states have already reformed their insurance regulations. Thus, the Solvency II regime also serves as unification of the solvency regulatory system. The Solvency II regime is expected to go live in January 1, 2014.

The specifications for Solvency II regime are obtained through a series of quantitative impact studies (QIS). Each following study refines the results of the previous study and provides more detailed and specific documentation required to adapt the Solvency II regime. In total, five quantitative impact studies were conducted, shortly referred to as QIS1 – QIS5.

The Solvency II framework is divided into three main areas (called pillars, similarly to the known banking supervisory system, Basel II)

- Pillar I (financial pillar): quantitative requirements;
- Pillar II (governance pillar): qualitative requirements – the supervisory review process;
- Pillar III (market conduct): statutory and market reporting.

The first pillar includes the calculation of the capital requirements (either according to standard formula or internal model, more details in next subsections). It also includes rules on provisioning and eligible capital.

The second pillar focuses on the supervisors and their review process (e.g., a company's internal control and risk management and the approval of using internal models in Pillar I).

The third pillar includes reporting to both the supervisor and the market. Reporting to market is required to promote market discipline and transparency.

The International Association of Insurance Supervisors (IAIS) gives 15 requirements for a solvency regime in relation to regulatory capital requirements:

1. A total balance sheet approach should be used in the assessment of solvency to recognize the interdependence between assets, liabilities, regulatory capital requirements and capital resources and to ensure that risks are appropriately recognized.

2. Regulatory capital requirements should be established at a level such that the amount of capital that an insurer is required to hold should be sufficient to ensure that, in adversity, an insurer's obligations to policyholders will continue to be met as they fall due.

3. The solvency regime should include a range of solvency control levels which trigger different degrees of intervention by the supervisor with and appropriate degree of urgency.

4. The solvency regime should ensure coherence between the solvency control levels established and the associated corrective action that may be at the disposal of the insurer and/or the supervisor. Corrective action may include options to reduce the risks being taken by the insurer as well as to raise more capital.

5. The regulatory capital requirements in a solvency regime should establish a solvency control level which defines the level above which the supervisor would not require action to increase the capital resources held or reduce the risks undertaken by the insurer. This is referred to as the *solvency capital requirement* (SCR) or *prescribed capital requiremen*t (PCR).

6. The SCR should be defined such that assets will exceed technical provisions and other liabilities with a specified level of safety over a defined time horizon.

7. The regulatory capital requirements in a solvency regime should establish a solvency control level which defines the supervisory intervention point at which the supervisor would invoke its strongest actions, if further capital is not made available. This is referred to as the *minimum capital requirement* (MCR).

8. The solvency regime should establish a minimum bound on the MCR below which no insurer is regarded to be viable to operate effectively.

9. The solvency regime should be open and transparent as to the regulatory capital requirements that apply. It should be explicit about the objectives of the regulatory capital requirements and the bases on which they are determined.
10. In determining regulatory capital requirements, the solvency regime should allow a set of standardized and, if appropriate, other approved more tailored approaches such as the use of (partial and full) internal models.
11. The solvency regime should be explicit as to where risks are addressed, whether solely in technical provisions, solely in regulatory capital requirements or if split between the two, the extent to which the risks are addressed in each. The regime should be explicit as to how risks and their aggregation are reflected in regulatory capital requirements.
12. The supervisor should set out appropriate target criteria for the calculation of regulatory capital requirements, which should underlie the calibration of a standardized approach.
13. Where the supervisory regime allows the use of approved more tailored approaches such as internal models for the purpose of determining regulatory capital requirements, the target criteria should also be used by those approaches for that purpose to ensure broad consistency among all insurers within the regime.
14. The solvency regime should be designed so that any variations to the regulatory capital requirement imposed by the supervisor are made within a transparent framework, are proportionate according to the target criteria and are only expected to be required in limited circumstances.
15. The solvency regime should be supported by appropriate public disclosure and additional confidential reporting to the supervisor.

To model the solvency capital requirement (SCR), the International Actuarial Association (IAA) has proposed five main *risk categories*:

1. insurance risk (or underwriting risk);
2. credit risk;
3. market risk;
4. operational risk;
5. liquidity risk.

The *underwriting risk* refers to the risk related to the businesses that will be written during the following year. In general, it is considered net of reinsurance, as the reinsurance risk will be dealt with in the default credit risk category. The underwriting process risk will thus be highly correlated with the credit risk.

Credit risk is the risk of not receiving promised payments (e.g., in company's investment portfolio, or reinsurance payments).

Market risk is introduced into an insurer's operations through variations in financial markets that cause changes in asset values, products, and portfolio valuation. Market risks relate to the volatility of the market values of assets and liabilities due to future changes in asset prices.

Operational risk is the risk of loss resulting from inadequate or failed internal processes, people, systems, or from external events. The most common risks that one may classify under operational risk are the failure in control and management, the failure in IT processes, human errors, fraud, jurisdictional and legal risks.

Liquidity risk is the risk that an insurer, although solvent, has insufficient liquid assets to meet his obligations (e.g. claim payments) when they fall due.

In the following we mostly concentrate on the problems related to the underwriting risk, keeping in mind that that the whole picture is more complex and includes other risk categories.

Generic risks related to underwriting that apply to most lines of business are:

- underwriting process risk – risk from exposure to financial losses related to the selection and approval of risks to be insured;
- pricing risk – risk that the prices charged by the company for insurance contracts will be ultimately inadequate;
- product design risk – risk that the company faces risk exposure under its insurance contracts that was unanticipated in the design and pricing of the insurance contract;
- claims risk (for each peril) – risk that many more claims occur than expected, or that some claims that occur are much larger than expected, resulting in unexpected losses;
- economic environment risk – risk that social conditions will change in a manner that has an adverse effect on the company;

- net retention risk – risk that higher retention of insurance loss exposures results in losses due to catastrophic or concentrated claims experience;
- policyholder behaviour risk – risk that the insurance company's policyholders will act in ways that are unanticipated and have an adverse effect on the company;
- reserving risk – risk that the provisions held in the insurer's financial statements for its policyholder obligations will prove to be inadequate.

In constructing solvency capital requirements (SCR), the following fundamental issues need to be discussed:

- valuation of assets and liabilities – the economic total balance sheet approach, i.e. assets and liabilities should be valued in a market-consistent way and their interactions should be a part of the solvency assessment;
- risk margins for uncertainty in assets and liabilities;
- risk measures for the volatility in assets and liabilities;
- modelling (risk categories, mitigation, diversification, etc).

14.2 Capital requirements in Solvency I

Before going to actual calculation in Solvency II, let us recall how the required capital is found under Solvency I regime.

Recall that in solvency I regime there are two important characteristics that specify when the regulatory authorities should take action:

- required (adequate) solvency margin;
- minimum guarantee fund.

In case the solvency margin of an insurance company falls below the required margin, supervisory organizations intervene and require certain actions to restore the adequate solvency margin. The solvency margin should never fall below a minimum guarantee fund, which is the absolute minimum amount of capital required.

The current required solvency margin for non-life insurance companies in EU is defined as the maximum of claim based and premium based indexes:

$$U_{min} = \max(P^*, S^*).$$

The claims and premium indices are defined as:

$$P^* = \begin{cases} 0.18 \cdot P \cdot R_{Re}, & for \quad P \leq 50 \cdot 10^6, \\ [0.18 \cdot 50 \cdot 10^6 + 0.16 \cdot (P - 50 \cdot 10^6)] \cdot R_{Re}, & for \quad P > 50 \cdot 10^6 \end{cases}$$

and

$$S^* = \begin{cases} 0.26 \cdot S \cdot R_{Re}, & for \quad S \leq 35 \cdot 10^6, \\ [0.26 \cdot 35 \cdot 10^6 + 0.23 \cdot (S - 35 \cdot 10^6)] \cdot R_{Re}, & for \quad S > 35 \cdot 10^6, \end{cases}$$

where P is the annual premium written, S is the average annual amount of claims and R_{Re} is the reinsurance ratio.

In other words, the calculation of required solvency margin is simple and straightforward, it is calculated as maximum of:

- 18% of premium written up to €50m plus 16% of premiums above €50m (adjusted by reinsurance);
- 26% of claims up to €35m plus 23% of claims above €35m (adjusted by reinsurance).

The minimum guarantee fund (MGF) is defined as

$$U_{mgf} = \max(v, \frac{1}{3}U_{min}),$$

where the margin v is between €2m and €3m, depending on the risk class.

Remark 14.1. All the limits in these calculations (€50m, €35m, €2m and €3m) are revised from time to time and are a subject to change.

14.3 Solvency II standard formula for non-life insurance

The calculation of capital requirements in Solvency II framework can be done in two ways:

- *standard formula* – a ready to use system with predetermined correlations between risks classes, volatilities, etc; this approach is especially convenient in the transition phase and also more suitable for smaller and medium companies;
- *internal model* – each insurance company can also develop its own model which will reflect better its individual risk profile.

14.3.1 Calculation of the solvency capital requirement SCR

Although with several predetermined relations and margins, the Solvency II standard formula framework is still very modular and dynamic (especially compared to Solvency I). The general idea of finding SCR consists of finding estimates for SCR in all modules at lower level of aggregation and then aggregating the results in order to get a higher level estimate. At the very top level, an insurance company is divided into parts depending on the company's profile (non-life, life, health, etc) and the aggregation rules are specified. In the next level, each part of the company is also divided into classes or modules of risks, then the solvency requirements are found for these classes and aggregated in order to obtain a requirement corresponding to higher level. Similar approach is applied in all levels of aggregation.

The SCR should be determined as

$$SCR = BSCR + CR_{OR} - Adj,$$

where $BSCR$ is the basic SCR, CR_{OR} is the capital charge for operational risk, and Adj is an adjustment for the risk-absorbing effect of future profit sharing and deferred taxes.

$BSCR$ is SCR before any adjustments, combining capital requirements for five major risk categories:

- CR_{NL} – the non-life underwriting risk module;
- CR_{LR} – the life underwriting risk module;
- CR_{HR} – the health underwriting risk module;
- CR_{CR} – the credit risk module;
- CR_{IAR} – the intangible assets risk module (intangible asset is an asset, other than a financial asset, that lacks physical substance).

$BSCR$ is found as

$$BSCR = CR_{IAR} + \sqrt{\sum_{i,j} \rho_{ij} \cdot CR_i \cdot CR_j},$$

where the i and j range over the values given in the list of categories above (NL, LR, HR, CR).

The dependence structure between these risks is assumed to be fixed and known, the values for ρ_{ij} are given in the standard formula documentation.

Remark 14.2. In many cases linear correlation is not an appropriate choice to describe the dependency (consider, e.g., skewed distributions, tail dependency). In such cases, the correlation parameters should be chosen in such a way as to achieve the best approximation of the *value at risk* (VaR) at 0.995 for the aggregated capital requirement[1]. In other words, there is a 1 in 200 chance that the insurer's surplus falls below the margin, so the capital level is intended to be such that the insurer could withstand a 1 in 200 years shock with sufficient assets remaining.

Let us now focus on the non-life insurance module and corresponding capital requirement CR_{NL}.

In the standard formula, there are two risk (sub)modules under the non-life underwriting risk module

- premium and reserve risk submodule;
- catastrophe risk (CAT risk) submodule.

Let the corresponding capital requirements be denoted as $CR_{NL,RP}$ for premium and reserve risk and $CR_{NL,CAT}$.

It is also assumed that the dependence between these risks $\rho_{RP,CAT}$ is 0.25. Conceptually, premium and reserve risks and the catastrophe risk should be independent, but a positive correlation factor may be appropriate in some situations, e.g., if underlying distributions are skewed and truncated. Then

Now, the capital requirement for the combined premium and reserve risks is determined as

$$CR_{NL,RP} = \rho_{0.995}(\sigma_{RP}) \cdot V_{RP},$$

where V_{RP} is a volume measure (or, more precisely, a combination of volume measures from reserve and premium volumes), σ_{RP} is combined standard deviation, and ρ is a charge function, defined by

$$\rho_{1-\alpha}(\sigma) = \frac{\exp\left(\Phi^{-1}(1-\alpha) \cdot \sqrt{\ln(1+\sigma^2)}\right)}{\sqrt{1+\sigma^2}} - 1.$$

In other words, the charge function ρ (with $\alpha = 0.005$) is chosen such that, assuming a lognormal distribution of the underlying risk, a capital charge

[1] Value at risk (VaR) for random variable X (with continuous and strictly increasing distribution function F) at given confidence level $\alpha \in (0,1)$ is defined as

$$VaR(\alpha) = F^{-1}(\alpha).$$

consistent with the value at risk at 0.995 confidence is produced. Using the lognormal assumption, the charge function is approximately $\rho_{0.995}(\sigma) \approx 3\sigma$.

The volume measures and standard deviations are calculated for each individual line of business (LoB) and then aggregated. For each LoB k let us denote

- $V_{R,k}$ – the volume measure for the reserve risk,
- $V_{P,k}$ – the volume measure for the premium risk,
- $V_{RP,k}$ – the volume measure for the premium and reserve risk,
- $\sigma_{R,k}$ – the volatility measure (standard deviation) for the reserve risk,
- $\sigma_{P,k}$ – the volatility measure (standard deviation) for the premium risk,
- $\sigma_{RP,k}$ – the volatility measure (standard deviation) for the premium and reserve risk.

Then, $V_{R,k}$ is defined as the best estimate for claims outstanding in LoB k. The volume measure for premium risk $V_{R,k}$ is defined as

$$V_{P,k} = \max\{P_k(1), P_k'(1), P_k(0)\} + M_k,$$

where $P_k(1)$ is the estimate of the net written premium in the forthcoming year, $P_k'(1)$ is the estimate of the net earned premium in the forthcoming year, $P_k(0)$ is the net written premium during the previous year, and M_k is the expected present value of net claims and expense payments that relate to claims incurred after the following year and are covered by existing contracts (multiyear contracts).

The volume measure for the premium and reserve risk within each LoB k is simply calculated as the sum of components' volume measures:

$$V_{RP,k} = V_{R,k} + V_{P,k}$$

and the volatility for the reserve and premium risks for each LoB k is calculated by

$$\sigma_{RP,k}^2 = \frac{(V_{R,k} \cdot \sigma_{R,k})^2 + (V_{P,k} \cdot \sigma_{P,k})^2 + 2\alpha V_{R,k} \cdot \sigma_{R,k} \cdot V_{P,k} \cdot \sigma_{P,k}}{V_{RP,k}^2},$$

where α is a dependence factor (in QIS4, α is set to 0.5).

All the required values for standard deviations $\sigma_{R,k}$ and $\sigma_{P,k}$ are specified in Solvency II documentation (for 12 general LoBs).

Now, the overall volume measure is just a sum of volume measures over all LoBs

$$V_{RP} = \sum_k V_{RP,k}$$

and the overall volatility measure is defined by

$$\sigma_{RP}^2 = \frac{1}{V_{RP}^2} \sum_{j,k} \rho_{j,k} \cdot V_j \cdot V_k \cdot \sigma_j \cdot \sigma_k,$$

where $\rho_{j,k}$ denotes the correlation between LoBs j and k and the needed values for all combinations of j and k are again specified in Solvency II documentation.

Lastly, without going into details, we also mention that there are two approaches to calculation of required capital in CAT-risk submodule:

- standardized scenario approach – detailed and comprehensive model;
- alternative (factor-based) approach – simplified model, used either in development phase, or in cases that are very clearly specified.

14.3.2 Calculation of the minimum capital requirement MCR

The MCR should ensure a minimum level below the amount of financial resources of an insurer should not fall. In case the resources fall below MCR, the most serious supervisory intervention is triggered.

The MCR is calculated using the following two-step algorithm:

1. Calculate MCR_L by

   $$MCR_L = MCR_{NL,NLt} + MCR_{NL,Lt} + MCR_{L,Lt} + MCR_{L,NLt},$$

   where

   - $MCR_{NL,NLt}$ denotes the MCR for non-life activities practiced on non-life technical basis,
   - $MCR_{NL,NLt}$ denotes the MCR for non-life activities practiced on life technical basis,
   - $MCR_{L,Lt}$ denotes the MCR for life activities practiced on life technical basis, and
   - $MCR_{L,NLt}$ denotes the MCR for supplementary obligations for life activities practiced on non-life technical basis.

All the components are calculated using the corresponding technical provisions and certain weight parameters provided by Solvency II documentation.

2. Squeeze the MCR_L into a corridor if needed: MCR_L should fit between $MCR_{Floor} = 0.25 \cdot SCR$ and $MCR_{Cap} = 0.45 \cdot SCR$.

Also, MCR must always be higher than an absolute floor $AMCR$:

$$MCR = \max\{MCR_L, AMCR\},$$

where the values for $AMCR$ depend on company's profile and are specified in the Solvency II documentation.

References

1. Sandström, A. (2010) *Handbook of Solvency for Actuaries and Risk Managers: Theory and Practice.* Taylor & Francis.

List of references

1. Booth P., Chadburn, R., Cooper, D., Haberman, S., James, D. (1999) *Modern Actuarial Theory and Practice.* Chapman & Hall / CRC.

2. Bühlmann, H., Gisler, A. (2005) *A Course in Credibility Theory and its Applications.* Springer, Berlin/Heidelberg.

3. Casualty Actuarial Society. (1998) *Foundations of Casualty Actuarial Science.* Casualty Actuarial Society.

4. Daykin, C.D., Pentikäinen, T., Pesonen, M. (1994) *Practical Risk Theory for Actuaries.* Chapman & Hall, London.

5. Dickson, D.C.M. (2006) *Insurance Risk and Ruin.* Cambridge University Press.

6. Goulet, V., et al. Package 'actuar': http://cran.r-project.org/web/packages/actuar/actuar.pdf

7. Gray, R.J. & Pitts, S.M. (2012) *Risk Modelling in General Insurance. From Principles to Practice.* Cambridge University Press.

8. Hesseläger, O. (1998) *Lecture Notes in Non-Life Insureance Mathematics.* HCO Tryk, Copenhagen.

9. Hogg, R. & Klugman, S. (1984) *Loss distributions.* Wiley, New York.

10. Hossack, I.B., Pollard, J.H., Zehnwirth, B. (1995) *Introductory Statistics with Applications in General Insurance.* Cambridge University Press.

11. Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M. (2008) *Modern Actuarial Risk Theory: Using R.* Springer, Berlin/Heidelberg.

12. Klugman, S.A., Panjer, H.H., Willmot, G.E. (1998) *Loss Models: From Data to Decisions.* Wiley, New York.

13. Mack, T. (1993) Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *Astin Bulletin*, **23**, 213–225.

14. Panjer, H.H., Willmot, G.E. (1992) *Insurance risk models.* Society of Actuaries.

15. R Core Team. Documentation for package 'stats': http://stat.ethz.ch/R-manual/R-patched/library/stats/html/00Index.html

16. Ross, S.M. (2009) *Introduction to Probability Models.* Elsevier/Academic Press.

17. Sandström, A. (2010) *Handbook of Solvency for Actuaries and Risk Managers: Theory and Practice.* Taylor & Francis.

18. Shevchenko P. V. (2010) Calculation of aggregate loss distributions. *The Journal of Operational Risk* 5(2), 3–40.

19. Straub, E. (1988) *Non-Life Insurance Mathematics.* Springer-Verlag / Association of Swiss Actuaries, Berlin/Heidelberg.

20. Sundt, B. (1993) *An Introduction to Non-Life Insurance Mathematics.* VVW, Karlsruhe.

21. Tse, Y.-K. (2009) *Nonlife Actuarial Models.* Cambridge University Press.