Mathematical Methods in Reinsurance

Dr. Peter Antal, Swiss Re

May 15, 2009

Chapter 1

Fundamentals

1.1 The history and current role of reinsurance

1.1.1 Historical background

Whereas the first known insurance cover (in connection with commercial seafaring) originated from the time before Christ, the oldest known treaty of a reinsurance nature was concluded in 1370 in Genoa. However, at that time coinsurance was the usual method of risk-sharing, i.e. insurers, having risks beyond their means to pay, insured these by sharing the risk with other insurers.

The increased number of risks arising from industrialization during the last century produced an ever greater need for reinsurance cover. The first professional reinsurance company, Cologne Re, was founded following a devastating fire in Hamburg in 1842. The loss from this event amounted to 18 million marks, whereas the local Hamburg Fire Fund only had 500,000 marks in reserve. This event assisted the final breakthrough of the need to share the risks of whole portfolios amongst several risk-carriers. By establishing professional reinsurance companies, the disadvantages of coinsurance - whereby a company could gain an insight into the business of another company and misuse this information to gain an unfair advantage - were eliminated. In addition, specialization allowed the development of new forms of reinsurance and worldwide multi-line activity allowed a better distribution of the risks. By providing better reinsurance protection, direct insurers were also able to offer their clients better conditions.

1.1.2 The current role of reinsurance

The question as to the role played by reinsurance has historically been answered with the following list of factors:

- smoothing out fluctuations
- risk transfer
- financing growth
- substitute for equity capital
- optimization of taxes
- liquidity provisions

While these are all valid points, they do not explain when reinsurance makes economic sense. For this, we need to look at reinsurance in a capital management context:

We can define an insurance company's *underwriting* risk by applying an appropriate risk measure ρ to the company's underwriting result R(the underwriting result is the random variable defined as the company's total premiums minus losses and expenses). We will deal with the question of risk measures later in detail, let's think of ρ at the moment as the expected 99%-shortfall, i.e. the average of the 1% worst outcomes when simulating the company's possible results. $\rho(R)$ can also be interpreted as the risk based capital (RBC) needed by the insurance company: holding this capital allows to meet the company's obligations versus the policyholders as long as the actual claims do not lead to a result which is worse than $\rho(R)$. By buying reinsurance, the capital need for the primary insurance company is reduced to $\rho(\widetilde{R})$, where $\widetilde{R}$ stands for the result after reinsurance. The capital relief is thus

$$K = \rho(R) - \rho(\widetilde{R}) > 0 \tag{1.1}$$

At the same time, the reinsurance company needs additional capital $\widetilde{K}$ for the newly assumed risks. Reinsurance thus creates economic value if $K > \widetilde{K}$.
Note that in practice the capital needed for an insurance company does not only depend on internal risk considerations but also on regulatory and rating agency constraints.

1.2 Contractual forms

1.2.1 General

A reinsurance treaty is a contractual agreement between a direct insurance company (sometimes also called primary insurance company) and a reinsurance company stipulating which share of (future) losses will be assumed by the reinsurance company (RI) and the premium which the direct insurance company (DI) is required to pay to the reinsurance company for this. The following points are characteristic:

- The duration of the treaty is fixed, in most cases for one year
- The treaty refers to either a well-defined DI's portfolio (e.g. all fire policies in Switzerland) or to a single risk (e.g. a production plant's fire policy). The first case is referred to as *obligatory* and the second case is referred to as *facultative* reinsurance.
- The premiums the DI receives from his clients (the insureds), are called *original premiums*. Losses which he is required to pay are correspondingly called *original losses*. The reinsurance treaty divides these losses into a loss deductible (borne by the DI) and a reinsurance loss (paid for by the RI). Thus:

$$\text{original loss} = \text{loss deductible} + \text{reinsurance loss}$$

Reinsurance treaties are usually divided into two categories: proportional and non-proportional. We will adhere to this division and will now explain the structure of the types of treaties which are in common use.

1.2.2 Proportional reinsurance treaties

Quota share reinsurance

In so-called *quota share reinsurance* original premiums and original losses are divided at a fixed ratio α between the direct insurer and the reinsurer. Thus:

Original loss: X_0	Original premium Π_0
RI loss: αX_0	RI premium $\alpha\Pi_0$
Loss deductible: $(1-\alpha)X_0$	Deductible premium $(1-\alpha)\Pi_0$

The original premiums are calculated in such a way that the direct insurer can pay for the losses as well as costs incurred (administration, agents etc.). The reinsurer participates in losses proportionally and in turn receives the corresponding proportion of premiums. In this way the reinsurer would, however, always have a better result than the direct insurer, since the reinsurer's costs are lower. To truly "share the fortunes" between direct insurer and reinsurer a share of the reinsurance premium is repaid to the direct insurer. This amount is referred to as reinsurance commission.

The transfer of money between insured, direct insurer and reinsurer can be represented as follows:

$$\text{Client} \;\; \underset{X_0}{\overset{\Pi_0}{\rightleftarrows}} \;\; \text{DI} \;\; \underset{\alpha X_0 + c}{\overset{\alpha\Pi_0}{\rightleftarrows}} \;\; \text{RI} \quad \text{(c stands for commission)}$$

Surplus reinsurance

Surplus reinsurance consists of a quota where the ratio α is not the same for all risks in the portfolio. Here a deductible M is first of all determined (in practice this is known as a retention line). Risks having a sum insured V smaller than M remain entirely with the DI (i.e. $\alpha = 0$). For other risks, premiums and losses are divided between RI and DI in the ratio $V - M : M$. The quantity which we have designated α in quota share is equal here to $\frac{V-M}{V}$. In the event of a total loss ($X_0 = V$) the DI pays M and the RI $V - M$. The commission rate is determined in the same way as for a quota share treaty.

1.2.3 Non-proportional reinsurance treaties

Excess of loss

In this form of reinsurance the RI takes on a share of each loss in excess of a previously agreed limit D , albeit only up to a limit C. The limit D is known as the deductible or sometimes as priority, C stands for the cover. The original loss X_0 is therefore divided here into a loss deductible X_{DI} and an RI loss X_{RI}, whereby

$$\begin{aligned} X_{RI} &= (X_0 - D)^+ - (X_0 - C - D)^+ \\ X_{DI} &= X_0 - X_{RI} \end{aligned}$$

The notation for this type of treaty is

C xs D

(in words: C excess D). For this cover the RI requires a premium which is calculated independently from the original premium. The area between D and D+C is often referred to as the layer belonging to the cover C xs D. Since we will often be needing the losses occurring in a layer as a function of the original loss, we will introduce the following notation:

$$L_{D,C}(X) := (X_0 - D)^+ - (X_0 - C - D)^+ \tag{1.2}$$

There are two types of excess of loss treaty:

1. WXL (working excess of loss) This is a per risk cover whereby the direct insurer retains a deductible of D in the case of every risk affected by a loss. This type of treaty protects the direct insurer from individual major losses.

2. Cat XL (catastrophe excess of loss) This is a per event cover common in property insurance, whereby the direct insurer retains a deductible D per event (e.g. earthquake, storm, hail). This type of treaty is used if many risks can be affected by a loss event at the same time.

Stop-loss reinsurance

Stop-Loss is an excess of loss on the DI's aggregate annual loss. In order to clarify this, we will consider a portfolio where N losses have occurred in a given year, which we denote by $X^1, \ldots, X^N$. The total reinsured losses of the year then look as follows in the case of an excess of loss or stop-loss cover (each having deductible D and cover C):

$$S_{XL} = \sum_{i=1}^{N} L_{D,C}(X^i) \tag{1.3}$$

$$S_{SL} = L_{D,C}\left(\sum_{i=1}^{N} X^i\right) \tag{1.4}$$

.

Excess of loss with AD and AL (aggregate deductible and aggregate limit)

This is a combination of excess of loss and stop-loss. A stop-loss AL xs AD on reinsured losses follows the excess of loss C xs D. The year's reinsured loss is thus

$$L_{AD,AL}(S_{XL}) = (S_{XL} - AD)^+ - (S_{XL} - AD - AL)^+ \tag{1.5}$$

The idea behind this structure is that the DI retains a large deductible for the "first" loss, (namely $D+AD$), and a small deductible (D) in the case of future losses. Often AL, the total annual liability, is given as a multiple of C, the liability per loss. A treaty is described as having k *reinstatements*, if $AL = (k+1)C$.

Largest claims reinsurance

In this treaty the RI takes on the year's largest (or the k highest) losses. It should, however, be noted here that this type of treaty is hardly ever used in current practice. For this reason we will not go into further detail.

Chapter 2

Some concepts of risk theory

2.1 Some distributions and their characteristics

In this lecture we are working with the so-called collective risk model, i.e. we consider a portfolio of risks which can lead to certain losses and do not differentiate which losses affect which risks (we will, however, be making an exception in exposure rating, see 4.1.2).

With N we indicate the number of (original) losses in the portfolio under examination and with $X_i \quad (1 \leq i \leq N)$ the loss amount from the i-nth loss. Further, we make the following assumptions:

1. N has a discrete distribution given by

   $\mathbb{P}[N = k] = p_k, \quad (k \in \mathbb{N}_0)$

2. The $X_i \quad (1 \leq i \leq N)$ are i.i.d with distribution F_X and independent of N.

The annual total loss is then

$$S = \sum_{i=1}^{N} X_i \tag{2.1}$$

First of all we will deal with the distributions used for modeling N and X.

2.1.1 Distributions for modeling the number of losses

To describe the number of losses in a portfolio, one of the following three distribution types is usually chosen:

1. *Poisson distribution with parameter* λ:

$$\begin{aligned} \lambda &> 0 \\ P[N=k] = e^{-\lambda}\lambda^k/k!\ , \qquad k &= 0,1,\dots \\ E[N] = Var[N] &= \lambda \end{aligned}$$

2. *Binomial distribution with parameters m, p*

$$\begin{aligned} m > 0,\ 0 \le p &\le 1 \\ P[N=k] = \binom{m}{k} p^k (1-p)^{m-k}, \qquad 0 \le k &\le m \\ E[N] = mp,\ Var[N] = mp(1-p) \end{aligned}$$

3. *Negative binomial distribution with parameters a, p (Polya(a,β))*

$$\begin{aligned} &a > 0,\ 0 \le p \le 1 \\ &P[N=k] = \binom{a+k-1}{k} p^a (1-p)^k, k = 0,1,\dots \\ &E[N] = a(1-p)/p,\ Var[N] = a(1-p)/p^2 \end{aligned}$$

Whereas binomial and Poisson distribution are also common and presumably well known outside actuarial mathematics, negative binomial distribution perhaps requires further explanation. The name originates from the relationship:

$$\binom{a+k-1}{k} p^a (1-p)^k = \binom{-a}{k} (p)^a (-1+p)^k$$

A second common parametric representation of this distribution is obtained from the substitution $\beta = 1/p - 1$ or $p = \frac{1}{1+\beta}$. This parametric representation is often termed Polya(a,β) distribution (in this case we have $E[N] = a \cdot \beta$). A third version (as used by Alois Gisler in his lectures) is using as parameters $\lambda = E[N]$ and a, i.e. the probabilities are given by

$$P_{a,\lambda}[N = k] = \binom{a+k-1}{k} * \left(\frac{a}{a+\lambda}\right)^a \left(\frac{\lambda}{a+\lambda}\right)^k$$

Exercise 1 *Prove the following properties of the negative binomial distribution:*

1. For $\lambda \in \Re^+$, $P_{a,\lambda}[N = k] \to e^{-\lambda}\lambda^k/k!$ as $a \to \infty$.
2. For $\lambda \in \Re^+$, $Q \sim \Gamma(a, a)$ (Gamma Distribution) and $N \sim$Poisson$(\lambda \cdot Q)$ we have: $P[N = k] = P_{a,\lambda}[N = k]$.

Here we would like to discuss a characteristic of the above distributions important for application to modeling the excess loss:

We consider those losses which exceed a given deductible D. The number of these losses is

$$N_D = \sum_{i=1}^{N} I_{\{X_i > D\}} \tag{2.2}$$

Let $\pi = P[X > D]$. Then:

Lemma 1 *The following holds for the distributions of N and N_D :*

Distribution of N	Distribution of N_D
Poisson(λ)	Poisson$(\lambda\pi)$
Binomial(m, p)	Binomial$(m, \pi p)$
Neg. Binomial(a, p)	Neg.Binomial$(a, \frac{p}{p+\pi(1-p)})$

Proof.

In the case of the Poisson distribution:

$$\begin{aligned} P[N_D = k] &= \sum_{n=0}^{\infty} P[N_D = k \mid N = n] e^{-\lambda}\lambda^n/n! \\ &= \sum_{n=k}^{\infty} \binom{n}{k} \pi^k (1-\pi)^{n-k} e^{-\lambda}\lambda^n/n! \\ &= e^{-\lambda}(\lambda\pi)^k/k! \sum_{m=0}^{\infty} \frac{1}{m!}(1-\pi)^m \lambda^m \\ &= e^{-\lambda\pi}(\lambda\pi)^k/k! \qquad \square \end{aligned}$$

The most elegant way of showing the Lemma for the other two distributions is by using the technique of generating functions. We will therefore be proving this in a following section. At this point we would like to introduce a term used very often in practice:

Definition 1 $f(D) := E[N_D] = E[N] \cdot P[X > D]$ *is called the excess frequency at deductible D.*

2.1.2 Distributions for modeling the loss amount

There are many probability distributions which can be used for modeling single loss amounts. Here, we intend to deal primarily with those which are significant in practice but are not necessarily dealt with in standard works.

Pareto distribution

This distribution is commonly used in reinsurance. The distribution function and the density are given by

$$F_X(x) = \begin{cases} 1 - (\frac{x}{x_0})^{-\alpha} & x > x_0 \\ 0 & \text{else} \end{cases} \tag{2.3}$$

$$f_X(x) = \begin{cases} \alpha x_0^{\alpha} x^{-\alpha-1} & x > x_0 \\ 0 & \text{else} \end{cases} \tag{2.4}$$

The parameters x_0 and α are both strictly positive. A minimum loss amount is determined by x_0. The parameter α defines the tail behavior of the distribution. If the moments of the Pareto distribution are investigated it can be seen that the n-th moment only exists for $n < \alpha$. The following formulas apply for the expected value and the variance:

$$E[X] = x_0 \frac{\alpha}{\alpha - 1} \qquad (\alpha > 1) \tag{2.5}$$

$$Var[X] = x_0^2 \frac{\alpha}{(\alpha-1)^2(\alpha-2)} \qquad (\alpha > 2) \tag{2.6}$$

There are various ways of parameterizing the Pareto distribution. The above is particularly advantageous as it has been in practice shown that a typical value α can be associated with a certain loss potential. The following rules of thumb have been established:

Loss potential	α
Earthquake/storm	≈ 1
Fire	≈ 2
Fire in industry	≈ 1.5
Motor liability	≈ 2.5
General liability	≈ 1.8
Occupational injury	≈ 2

It should, however, be noted that the effective loss amount distribution can only be approximated using the Pareto distribution in a certain interval $[x_0, M]$, although this is often sufficient for practical purposes.

A further noteworthy characteristic of the Pareto distribution is its behavior in forming certain conditional distributions: If X is Pareto(x_0, α)-distributed and $D \geq x_0$, then

$$P[X > x|X > D] = \frac{P[X > \max(x, D)]}{P[X > D]} = \begin{cases} (\frac{x}{D})^{-\alpha} & x > D \\ 1 & \text{else} \end{cases} \tag{2.7}$$

i.e. we once again obtain a Pareto distribution with the same "shape parameter" α.

In reinsurance one is often interested in the loss amount in a layer with deductible D and cover C. In this case all moments exist since the underlying random variable is bounded above (by the constant C). If X_0 is Pareto(x_0, α)-distributed and $X_{RV} = L_{D,C}(X_0)$, then with $R = D + C$:

$$E[X_{RV}] = \begin{cases} x_0 \ln(1+\frac{C}{D}) & \alpha = 1 \\ x_0(\frac{x_0}{D} - \frac{x_0}{R}) & \alpha = 2 \\ \frac{x_0}{\alpha-1}\left((\frac{x_0}{D})^{\alpha-1} - (\frac{x_0}{R})^{\alpha-1}\right) & \text{else} \end{cases} \quad (2.8)$$

$$Var[X_{RV}] = \begin{cases} 2x_0(C - D\ln(1+\frac{C}{D})) & \alpha = 1 \\ 2x_0^2(\ln(1+\frac{C}{D}) - (1-\frac{D}{R})) & \alpha = 2 \\ 2x_0\left[\frac{R}{1-\alpha}\left((\frac{x_0}{D})^{\alpha-1} - (\frac{x_0}{R})^{\alpha-1}\right) + \right. & \\ \left. \frac{x_0}{\alpha-1}\left((\frac{x_0}{D})^{\alpha-2} - (\frac{x_0}{R})^{\alpha-2}\right)\right] & \text{else} \end{cases} \quad (2.9)$$

The MBBEFD distribution

This is a class of two-parameter distributions introduced by Bernegger[2] which has proved useful particularly in modeling degrees of loss in fire insurance. The name MBBEFD stands for Maxwell-Boltzmann-Bose-Einstein-Fermi-Dirac, which is intended to suggest the similarity with the corresponding functions from statistical mechanics. The distribution has two parameters b and g ($b \geq 0$ and $g > 1$) and the support $[0, 1]$. The distribution function is given by

$$F(x) = \begin{cases} 0 & x \leq 0 \text{ or } g = 1 \text{ or } b = 0 \\ 1 & x \geq 1 \\ 1 - \frac{1}{1+(g-1)x} & x \in (0,1) \text{ and } g > 1 \text{ and } b = 1 \\ 1 - b^x & x \in (0,1) \text{ and } g > 1 \text{ and } bg = 1 \\ 1 - \frac{1-b}{(g-1)b^{1-x} + (1-gb)} & \text{else} \end{cases}$$

As already mentioned, this distribution is particularly suitable for describing degrees of loss, i.e. describing the loss amounts measured in terms of the maximum possible loss (MPL) or the sum insured. The degree of loss 1 thus corresponds to a total loss. A positive total loss probability $P[X = 1] = 1/g$ is characteristic of the MBBEFD distribution. There is also a variation of this distribution defined on the interval $[0, \infty)$. We will, however, not be dealing with this here. Interested readers should refer to [2]. The expected value of the MBBEFD distribution is given by

$$E[X] = \begin{cases} 1 & g = 1 \text{ or } b = 0 \\ \frac{\ln(g)}{g-1} & g > 1 \text{ and } b = 1 \\ \frac{b-1}{\ln(b)} & g > 1 \text{ and } bg = 1 \\ \frac{\ln(gb)(1-b)}{\ln(b)(1-gb)} & \text{else} \end{cases}$$

In practice the total loss probability $P[X = 1] = p_0$ and the expected value $E[X] = \mu$ are given and it is necessary to find the relevant distribution from the MBBEFD class. It can always be said that:

$$g = \frac{1}{p_0}$$

In defining b, once again various cases must be differentiated:

$$b(\mu, g) = \begin{cases} 0 & \text{if } \mu = 1 \\ 1/g & \text{if } \mu = \frac{1}{g^2 \ln g} \\ 1 & \text{if } \mu = \frac{\ln g}{g-1} \\ \infty & \text{if } \mu = \frac{1}{g} \end{cases}$$

In the general case of ($0 < b < \infty$ and $b \neq 1/g$ and $b \neq 1$) b must be determined numerically from the equation

$$\mu = \frac{\ln(gb)(1-b)}{\ln(b)(1-gb)} \tag{2.10}$$

The Benktander distribution

This involves a distribution introduced by Benktander [3], the asymptotic behavior of which lies between the exponential distribution and the Pareto distribution. It is used above all in liability business. The distribution function is given by

$$F_X(x) = \begin{cases} 1 - (\frac{x}{x_0})^{b-1} e^{-(a/b)(x^b - x_0^b)} & \text{if } x > x_0 \\ 0 & \text{if } x \leq x_0 \end{cases} \tag{2.11}$$

and for the parameters a, b, x_0 we have:

$$\begin{aligned} 0 &< a \\ 0 &< b \leq 1 \\ 0 &< x_0 \end{aligned}$$

In the case $b = 0$ we have a Pareto distribution with $\alpha = a + 1$, $b = 1$ gives an exponential distribution with parameter a shifted by x_0. Parameter b is scale invariant (currency exchange rate, inflation, etc.), but not, however, parameter a. If we let:

$$\begin{aligned} X' &= \gamma X \\ x_0' &= \gamma x_0 \end{aligned}$$

then X' once again has a distribution of the same type with parameters a', b', where

$$\begin{aligned} a' &= \frac{a}{\gamma^b} \\ b' &= b \end{aligned}$$

In motor liability business selecting $b \approx 0.5$ has proved useful. The choice of a depends on the unit and the currency in which the loss amount is measured due to the above scaling characteristic.

Generalized Pareto Distribution (GPD)

This distribution appears in connection with results from the extreme value theory. It involves a two-parameter distribution with the distribution function

$$G_{\xi,\sigma}(x) = \begin{cases} 1 - (1 + \xi x/\sigma)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - e^{-x/\sigma} & \text{if } \xi = 0 \end{cases}$$

whereby $\sigma > 0$ and $\xi \in \mathbb{R}$. The support is $[0, \infty)$ if $\xi > 0$ or $[0, -\sigma/\xi]$ if $\xi < 0$. In the case of $\xi = 0$ we have an exponential distribution. This distribution is often suitable for describing loss amounts in high-excess layers. In order to understand this more fully we need some results from the extreme value theory:

Definition 2 ***Definition 3*** *Let F be a distribution and F^n the distribution belonging to the successive maxima, i.e. if $(X_i)_{i=1,2,...}$ is a sequence of i.i.d random variables with distribution F , then F^n is the distribution of $M_n = \max(X_1, \ldots, X_n)$. We say that F is in the domain of attraction of an extreme value distribution, if there are successive numbers $a_n > 0$ and b_n, so that*

$$F^n(a_n x + b_n) \to H(x)$$

for a non degenerated *distribution H.*

A famous result from the extreme value theory (Fischer-Tippett, 1928) says that there are only three candidates for the distribution H, namely the Fréchet, the Gumbel and the Weibull distribution. The loss distributions of relevance for actuarial mathematics are either in the domain of attraction of the Fréchet distribution (e.g. if F = Pareto or loggamma) or that of the Gumbel distribution (F = normal, exponential, gamma, lognormal). The greatest problem in the applicability of these results in practice is the fact that if F is a distribution with finite support (in the real world the highest possible insured loss is surely finite), then the assumption of the statement is not fulfilled, since M_n then converges towards the right endpoint of the distribution and H is thus degenerated.

If the real loss distributions are nevertheless to be described by means of distribution functions with infinite support (there is little choice available), then the link with the general Pareto distribution can be reproduced using the following statement:

Theorem 1 *(Pickands-Balkema-deHaan, 1974)*

Let F be a loss distribution with infinite support and F_D the distribution of the excess loss above the deductible D, i.e.

$$F_D(x) = P[X - D \leq x \mid X > D]$$

Then there is a positive, measurable function $\sigma(D)$ and $\xi \in \mathbb{R}$, so that

$$\lim_{D\to\infty} \sup_{x\geq 0} \left|F_D(x) - G_{\xi,\sigma(D)}(x)\right| = 0$$

if and only if F lies in the domain of attraction of an extreme value distribution.

This statement gives us a theoretical basis for sensibly modeling the loss amount in high-excess layers by means of the generalized Pareto distribution.

Generalized Beta Distribution

This is a four parameter distribution that is rarely used in his general form, but it is very convenient for the common parametrization of a large class of distributions. In fact, many frequently used severity distributions turn out to be a special case of the generalized beta distribution. The distribution function is given by

$$F_{\alpha,\beta,\omega,\rho}(x) = \begin{cases} B(\alpha, \beta, \frac{x^\rho}{x^\rho+\omega}) & \text{if } x > 0 \\ 0 & \text{else} \end{cases}$$

where

$$B(\alpha, \beta, y) = \frac{\Gamma(\alpha+\beta) \int\limits_0^y u^{\beta-1}(u-1)^{\alpha-1} du}{\Gamma(\alpha)\Gamma(\beta)}$$

is the incomplete Beta function and the parameters $\alpha, \beta, \omega, \rho$ are all strictly positive.

3-Parameter sub-classes of the generalized beta distribution:

Inverse Burr	$\mathrm{IB}(\alpha, \omega, \rho) = \mathrm{GB}(1, \beta, \omega, \rho)$
Burr	$\mathrm{BR}(\alpha, \omega, \rho) = \mathrm{GB}(\alpha, 1, \omega, \rho)$
Transferred Gamma	$\mathrm{TG}(\alpha, \beta, \omega, \rho) = \mathrm{GB}(\alpha, \beta, 1, \rho)$
Generalized Pareto (3)	$\mathrm{GP}(\alpha, \beta, \omega) = \mathrm{GB}(\alpha, \beta, \omega, 1)$

Some 2-Parameter sub-classes of the generalized beta distribution

Beta	$\mathrm{B}(\alpha, \beta) = \mathrm{GB}(\alpha, \beta, 1, 1)$
Second Pareto	$\mathrm{GP}(\alpha, \omega) = \mathrm{GB}(\alpha, 1, \omega, 1)$
Weibull	$\mathrm{GP}(\alpha, \rho) = \mathrm{GB}(\alpha, 1, 1, \rho)$
Inverse Pareto	$\mathrm{IP}(\beta, \omega) = \mathrm{GB}(1, \beta, \omega, 1)$
Loglogistic	$\mathrm{LL}(\omega, \rho) = \mathrm{GB}(1, 1, \omega, \rho)$

In an empricial study investigating more than 100'000 large industrial losses, Aebischer [1]observed that excess loss distributions always showed two more features:

- Excess loss distributions are convex, i.e. have a monotonic density function
- The slope of the distribution function is finite and strictly greater than zero

Enforcing these two properties eliminates one parameter of the generalized beta distribution and boils down to the relationship

$$\beta = 1/\rho \tag{2.12}$$

2.2 Generating Functions

The generating functions discussed in this section are technical tools which often facilitate calculations using random variables. We will not be dealing here with existential questions but will assume that we are only considering random variables for which these functions are finite. This is in fact so in most cases. The Pareto distribution constitutes one exception, although it is mainly used only in finite layers where there is no problem with existence.

2.2.1 The moment generating function

For a random variable X we define the moment generating function as

$$M_X : \mathbb{R} \to \mathbb{R}^+, \quad M_X(s) = E[e^{sX}] \tag{2.13}$$

The following properties of M are important in practice:

- The distribution of X is uniquely determined by the moment generating function.
- $\frac{d^n}{ds^n} M\Big|_{s=0} = E[X^n]$
- If X_1, X_2 are independent random variables and $X = X_1 + X_2$, then: $M_X(s) = M_{X_1}(s) \cdot M_{X_2}(s)$

Example:

For a Poisson(λ) distributed random variable N the following applies:

$$M_N(s) = e^{\lambda(e^s-1)}. \tag{2.14}$$

From this it immediately follows that the sum of 2 independent Poisson variables with expected values λ_1 and λ_2 is again Poisson distributed with parameter $\lambda_1+\lambda_2$.

2.2.2 The cumulant generating function

The cumulant generating function φ_X is the logarithm of the moment generating function:

$$\varphi_X : \mathbb{R} \to \mathbb{R}, \quad \varphi_{\mathbb{X}}(s) = \ln\left(M_X(s)\right) \tag{2.15}$$

The value

$$k_j = \frac{d^j}{ds^j}\varphi\Big|_{s=0} \qquad (j \geq 1) \tag{2.16}$$

is called the k-th cumulant of the distribution of X. Quite analogous to the moment generating function we have:

- The distribution of X is uniquely determined by the cumulant generating function.
- If X_1, X_2 are independent random variables and $X = X_1 + X_2$, then:

$$\varphi_X(s) = \varphi_{X_1}(s) + \varphi_{X_2}(s)$$

- The cumulants can be expressed in terms of the moments $\alpha_j = E[X^j]$, in particular the following applies:

$$k_1 = \alpha_1 \tag{2.17}$$

$$k_2 = \alpha_2 - \alpha_1^2 \quad (= Var[X]) \tag{2.18}$$

- For i=2,3 we have: $k_i = E[(X - E[X])^i]$.[1]

[1] Note that this characteristic does not apply to $i \geq 4$. Thus $E[(X - E[X])^4 = k_4 + 3k_2^2$ for example.

Example:

For a Poisson(λ) variable: $k_j = \lambda$ (independent of j!)

2.2.3 The probability generating function

For random variables N with values in $\mathbb{N}$ the probability generating function

$$W_N : \mathbb{R} \to \mathbb{R}^+, \quad W_N(s) = E[s^N] \tag{2.19}$$

is often considered.

We have as characteristics :

$$W_N(s) = \sum_{k=0}^{\infty} s^k P[N = k] \tag{2.20}$$

$$P[N = k] = \frac{d^k}{ds^k} W_N \Big|_{s=0} \cdot \frac{1}{k!} \tag{2.21}$$

Example: if $N \sim$ Poisson(λ) then: $W_N(s) = e^{\lambda(s-1)}$

2.3 The Aggregate Loss Process

This section deals with the aggregate loss process $S = \sum_{i=1}^{N} X_i$. First of all we will deduce some important properties of the distribution of S.

2.3.1 Distribution function, generating functions and moments

Proposition 1 *The distribution function of S is given by*

$$F_S(s) = \sum_{k=0}^{\infty} P\left[N = k\right] \; F_X^{*k}(s)$$

where F_X^{*k} is the k-th convolution power of F_X.[2]

Proof.

[2] The convolution of two distribution functions F_1 and F_2 is defined as $F_1 * F_2(x) = \int_{-\infty}^{\infty} F_1(x-y) dF_2(y)$

Follows directly from conditioning to N and from the fact that the distribution function of a sum of independent random variables is obtained by the convolution of the individual distribution functions.

Proposition 2

$$\begin{aligned} M_S(z) &= M_N(\varphi_X(z)) = W_N(M_X(z)) &(2.22)\\ \varphi_S(z) &= \varphi_N(\varphi_X(z)) &(2.23)\end{aligned}$$

Proof.

$$\begin{aligned} M_S(z) = E[e^{zS}] &= \sum_{n=0}^{\infty} E[e^{zS} \mid N = n] \, \mathbb{P}[N = n] \\ &= \sum_{n=0}^{\infty} E[e^{zX_1 + \cdots + zX_N}] \, \mathbb{P}[N = n] \\ &= \mathbb{E}[(M_X(z))^N] \\ &= \mathbb{E}[e^{N \ln M_X(z)}] = M_N(\varphi_X(z)) \end{aligned}$$

We will now consider two applications:

Application 1

Let $S_1, \ldots, S_k$ be independent compound Poisson-distributed random variables with Poisson parameters λ_i and loss severities $X_{i,j}$ $(1 \le i \le k,\ 1 \le j \le N_i)$. For the distribution of $X_{i,j}$ we assume that the moment generating functions M_{X_i} exist (since the $(X_{i,j})_{j=1\ldots N_i}$ are i.i.d. for fixed i, we will omit the index j). Furthermore let $S = \sum_{i=1}^{k} S_i$. According to (2.23) and (2.14):

$$\begin{aligned} \varphi_S(z) &= \sum_i \lambda_i (M_{X_i}(z) - 1) \\ &= \lambda \left(\sum_i \frac{\lambda_i}{\lambda} M_{X_i}(z) \right) - \lambda &(2.24)\end{aligned}$$

whereby we have let $\lambda = \sum_i \lambda_i$. However, the expression in brackets is the moment generating function of a random variable with the distribution function $F = \sum_i \frac{\lambda_i}{\lambda} F_{X_i}$ (it should be noted that this is not like the distribution of $\sum_i \frac{\lambda_i}{\lambda} X_i$). Thus the distribution of S is also compound Poisson with Parameter λ and single loss severity distribution F. We will see an important application of this result in the chapter on exposure rating .

Application 2

As a second application we prove Lemma 1 for the negative-binomial distribution. The proof of the binomial distribution is carried out in the same manner and is left to the reader as an exercise. We want to show that if $N \sim$ negative-binomial (a, p), then:

$$N_D \sim \text{negative-binomial } (a, \frac{p}{p + \pi(1-p)})$$

where $\pi = P[X > D]$ and $N_D = \sum_{i=1}^{N} I_{\{X_i > D\}}$.
First of all we note that the moment generating function of N is given by

$$M_N(z) = \left(\frac{p}{1-(1-p)e^z}\right)^a$$

Furthermore we have for the moment generating function of $Y = I_{\{X>D\}}$:

$$\begin{aligned} M_Y(z) &= E\left[\exp\{z \cdot I_{\{X > D\}}\}\right] \\ &= 1 - \pi + \pi e^z \end{aligned}$$

The moment generating function of $N_D = \sum_{i=1}^{N} Y_i$ is according to Proposition 2

$$\begin{aligned} M_{N_D}(z) &= M_N(\ln(M_Y(z)) \\ &= \left(\frac{p}{1-(1-p)(1-\pi+\pi e^z)}\right)^a \\ &= \left(\frac{\frac{p}{p+(1-p)\pi}}{1-(1-\frac{p}{p+(1-p)\pi}e^z)}\right)^a \end{aligned}$$

and this leads to the assertion.

The so-called Wald-Identities are very useful for calculating the expected value and variance of the aggregate loss distribution:

Proposition 3 *(Wald-Identities)*

$$E[S] = E[X] \cdot E[N] \tag{2.25}$$

$$Var[S] = Var[X] \cdot E[N] + (E[X])^2 \cdot Var[N] \tag{2.26}$$

$$= E[N] \cdot \left(E[X^2] + (Q-1)(E[X])^2\right) \tag{2.27}$$

where Q *stands for* $\frac{Var[N]}{E[N]}$.

Proof.

(2.25) results immediately from $E[S] = E[E[S/N]] = E[N \cdot E[X]]$. In order to obtain (2.26) we first of all note that

$$\begin{aligned} E[S^2/N] &= E\left[\sum_{i=1}^{N} X_i^2 + \sum_{i \neq j} X_i X_j\right] \\ &= N \cdot E[X^2] + (N^2 - N) \cdot (E[X])^2 \end{aligned}$$

From $Var[S] = E[E[S^2/N]] - (E[S])^2$ (2.26) follows. The variants (2.27) follow in the same way using $E[N^2] = E[N] \cdot Q + (E[N])^2$. □

Note: The class consisting of the Poisson, binomial and negative-binomial distributions is often called the Panjer class (named after the Canadian mathematician Harry Panjer). For this distribution the above-defined quantity Q is called the Panjer factor. The following then applies:

Poisson	$Q = 1$
Binomial	$Q < 1$
Neg. Binomial	$Q > 1$

In the case of N $\sim$ Poisson(λ) the simplified formula:

$$Var[S] = \lambda \cdot E[X^2]$$

follows due to $Q = 1$

In concluding this section we would like to discuss a method with which aggregate loss distributions with negative binomial distributed numbers can be traced back to aggregate loss distributions with Poisson distributed number of losses. This involves the so-called *Ammeter transformation* which we will define in the following theorem:

Theorem 2 *(Ammeter Transformation)*

We consider an aggregate loss process (N, X), *where* $N \sim$*Polya*(a, β). *The Ammeter transform of this process is an aggregate loss process* $(\tilde{N}, \widetilde{X})$, *where*

$$\tilde{N} \sim Poisson(a \cdot \beta) \quad (2.28)$$

$$M_{\widetilde{X}} = 1 - \frac{\ln(1 - \beta(M_X - 1))}{\beta} \quad (2.29)$$

Therefore :

$$\sum_{i=1}^{N} X_i \overset{d}{\sim} \sum_{i=1}^{\tilde{N}} \widetilde{X}_i$$

ie we can replace the Polya process with a Poisson process having the same expected value[3]*, if we transform the distribution of the loss amount according to (2.29).*

Proof.

We show that both processes have the same moment generating function. The moment generating function of $S = \sum_{i=1}^{N} X_i$ is

$$\begin{aligned} M_S(z) &= W_N(M_X(z)) \\ &= \frac{1}{(1-\beta(M_X(z)-1))^a} \end{aligned}$$

on the other hand

$$\begin{aligned} M_{\widetilde{S}}(z) &= \exp(a \cdot \beta(M_{\widetilde{X}} - 1)) \\ &= \exp\left\{a \cdot \beta \cdot \left(1 - \frac{\ln(1-\beta(M_X(z)-1))}{\beta}\right) - 1\right\} \\ &= (1-\beta(M_X(z)-1))^{-a} \end{aligned}$$

and from this follows the assertion.

2.3.2 The distribution of excess losses

In this section we aim to investigate the distribution of the excess-loss burden $\sum_{i=0}^{N}(X_i - D)^+$. The main result is as follows

Proposition 4 *Let $D > 0$ and N_D be the number of losses greater than D (see (2.2)). Then:*

$$\sum_{i=0}^{N}(X_i - D)^+ \quad \overset{d}{\sim} \quad \sum_{i=0}^{N_D}(Y_i - D)^+$$

where Y_i are i.i.d random variables with distribution $P[. \quad | X > D]$ and independent of N_D.

[3] Note that $E[N] = a \cdot \beta$

Proof.

P is used to denote the probability measure belonging to the model (N, X) and $\mathbb{P}$ the measure belonging to the model (Y, N_D). Then:

$$\begin{aligned}
P\left[\sum_{i=1}^{N}(X_i - D)^+ \leq s\right] &= \sum_{n=0}^{\infty} P\left[\sum_{i=1}^{N}(X_i - D)^+ \leq s \;\mid\; N_D = n\right] \cdot P[N_D = n] \\
&= \sum_{n=0}^{\infty} P\left[\sum_{i=1}^{n}(X_i - D)^+ \leq s \;\mid\; X_1, \ldots, X_n > D\right] \cdot P[N_D = n] \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left[\sum_{i=1}^{n}(Y_i - D)^+ \leq s\right] \cdot \mathbb{P}[N_D = n] \\
&= \mathbb{P}\left[\sum_{i=1}^{N_D}(Y_i - D)^+ \leq s\right]
\end{aligned}$$

□

Example: Poisson-Pareto model

Let the loss amount X Pareto $(x_{0,}\alpha)$ and N, the number of losses larger than x_0 be Poisson(λ) distributed. The total annual loss in layer (D,C) then has the same distribution as when the model (N_D, Y) is considered instead of the model (N, X) where

$$\begin{aligned}
Y &\sim \text{Pareto}(D, \alpha) \qquad (2.30) \\
N_D &\sim \text{Poisson}(\lambda_D)
\end{aligned}$$

with $\lambda_D = \lambda \cdot P[X > D] = \lambda \cdot (\frac{x_0}{D})^\alpha$.

2.3.3 The Panjer Algorithm

As we have seen, the distribution function of the aggregate loss is equivalent to

$$F_S(z) = \sum_{i=0}^{\infty} F_X^{*\,i}(z)\; P[N = i]$$

This expression cannot be exactly evaluated for most distributions so that it is necessary to rely on numerical methods. For discrete severity distributions the convolutions can be calculated by computer, though this is rather time-consuming despite present-day technology.

A simple recursive formula for the numerical calculation of F_S was developed by Panjer. This algorithm only functions for discrete severity distributions which means that continuous distributions must first be made discrete. We will deal with this point in greater detail later. We will first consider a mixed model (N, X) and assume the following:

1. The distribution of X is concentrated on the lattice $d\mathbb{N}$ (for some $d \in \mathbb{R}^+$) with $f_k = P[X = d \cdot k]$. The moment generating function M_X exists in an environment of 0.
2. The distribution of N is from the Panjer class, ie Poisson, binomial or negative binomial.

The assumption made in 1) does not cause problems in practice since one has to restrict oneself to distributions with finite support in the case of numerical models. As preparation we require the following result:

Lemma 2 *The distribution of N is from one of the three types above, if and only if there are constants $a, b \in \mathbb{R}$, so that the weights $p_k = P[N = k]$ fulfill the following relation:*

$$\begin{aligned} p_0 &\neq 0 \\ p_k &= (a + \frac{b}{k})\ p_{k-1} \end{aligned} \tag{2.31}$$

Proof.
See Sundt / Jewell (1981): "Further Results on Recursive Evaluations of Compound Distributions", ASTIN Bulletin
On the basis of our assumptions the distribution of $S = \sum_{i=1}^{N} X_i$ is also concentrated on the lattice $d\mathbb{N}$. This can be calculated recursively as follows:

Theorem 3 *(Panjer) The weights $g_k = P[S = d \cdot k]$ fulfill the following relation:*

$$\begin{aligned} g_0 &= \begin{cases} p_0 & \text{if } f_0 = 0 \\ W_N(f_0) & \text{if } f_0 > 0 \end{cases} \\ g_k &= \frac{1}{1 - af_0} \sum_{i=1}^{k} (a + i\frac{b}{k})\ f_i\ g_{k-i} \end{aligned} \tag{2.32}$$

Proof.

Proof here is essentially provided by the fact that the moment generating functions

$$\begin{aligned} M_S(z) &= \sum_k g_k \; e^{kdz} \\ M_X(z) &= \sum_k f_k \; e^{kdz} \end{aligned}$$

fulfill the following differential equation:

$$M_S' = aM_X M_S' + (a+b)M_X' M_S \tag{2.33}$$

Before proving (2.33) , we will show how then the Theorem follows. If we insert the definition of the moment generating functions into (2.33), we obtain by the comparison of coefficients

$$g_k \cdot k = \sum_{i,j:\ i+j=k} a\, f_i\, g_j\, k + b\, f_i\, g_j\ i$$

which directly provides (2.32). It thus remains to show (2.33):

By (2.31) we have

$$kp_k = a(k-1)p_{k-1} + (a+b)p_{k-1}$$

We multiply this equation by $\{M_X(z)\}^{k-1} M_X'(z)$ and sum over k. This provides

$$\begin{aligned} \sum_k kp_k \{M_X(z)\}^{k-1} M_X'(z) &= a \sum_k (k-1)p_{k-1}\{M_X(z)\}^{k-1} M_X'(z) \\ &\quad + (a+b) \sum_k p_{k-1}\{M_X(z)\}^{k-1} M_X'(z) \end{aligned}$$

With $M_S(z) = \sum_{k=0}^{\infty} p_k \{M_X(z)\}^k$ the statement follows.

2.4 Stop-loss transform and exposure curves

In this section we will be defining two very useful integral transformations and will be investigating their characteristics. In the following Z is always used to denote a positive random variable with finite expected value.

2.4.1 Stop-loss transformation

Definition 4 *The stop-loss transform of Z is the function*

$$slt_Z : \mathbb{R}^+ \to \mathbb{R}^+, \qquad slt_Z(u) = \int_u^\infty (1 - F_Z(s))\ ds \tag{2.34}$$

The name of this function derives from the fact that it provides precisely the expected loss in a stop-loss treaty with infinite cover as a function of the priority. The following applies:

Proposition 5

$$E[L_{u,\infty}(Z)] = slt_Z(u) \tag{2.35}$$

Proof.

In literature proof of this statement by means of partial integration is commonly found. The disadvantage of this method is that it only functions when the distribution of Z has a density - which is not always the case in practice. The following proof functions for any distributions:

$$\begin{aligned} slt_Z(u) &= \int_u^\infty P[Z > s]\ ds \\ &= \int_u^\infty \int_0^\infty I_{(s,\infty)}(z)\ dF_Z(z)\ ds \\ &= \int_0^\infty \int_u^\infty I_{(0,\ z)}(s)\ ds\ dF_Z(z) \\ &= \int_0^\infty (z-u)^+\ dF_Z(z) \\ &= E[(Z-u)^+] \qquad \square \end{aligned}$$

Proposition 6 *The stop-loss transform is a decreasing convex function with $slt_Z(0) = E[Z]$ and $slt_Z(\infty) = 0$.*

Proof.
The boundary values result directly from the definition. The convexity arises because the derivative $slt'_Z = F(z) - 1$ is increasing. $\square$
Example:

For $Z \sim$ Pareto (x_0, α) with $\alpha > 1$ the following applies: $slt_Z(u) = \frac{x_0^\alpha}{\alpha - 1} u^{1-\alpha}$.

Definition 5 *On the set of the distribution functions with finite expected values we define a partial ordering* $\prec$ *by*

$$F \prec G \Leftrightarrow slt_F(u) \leq slt_G(u) \quad \forall u \geq 0$$

This is called the stop-loss order. We say that the distribution G *is more dangerous than* F*.*

Proposition 7 *Let* F *and* G *be two distribution functions with their associated expected values* μ_F *and* μ_G*. Furthermore:*

1. $\mu_F \leq \mu_G$
2. $\exists\ \beta > 0$ *with* $F(x) \leq G(x)$ *for* $x < \beta$ *and* $F(x) \geq G(x)$ *for* $x \geq \beta$*.*

Then: $F \prec G$

Proof.

We have to show that for all $t \geq 0$

$$\int_t^{\infty} G(x) - F(x)\ dx \leq 0$$

For $t \geq \beta$ this is evident. For $t < \beta$ we have:

$$\int_t^{\infty} G(x) - F(x)\ dx \leq \int_{-\infty}^{\infty} G(x) - F(x)\ dx = \mu_F - \mu_G \leq 0 \qquad \square$$

Lemma 3 *The stop-loss order is preserved under mixture and convolution, i.e for two series of distributions* $(F_i)_{i\in I}$*,* $(G_i)_{i\in I}$ *with* $\forall\ i \in I$*,* $F_i \prec G_i$ *:*

$$\begin{aligned} \sum_{i\in I} p_i F_i &\prec \sum_{i\in I} p_i G_i \\ F_1 * F_2 &\prec G_1 * G_2 \end{aligned}$$

where $(p_i)_{i\in I}$ *is a set of weights with* $\sum_{i\in I} p_i = 1$*.*

Proof: Exercise

Corollary 1 *We consider two severity/frequency models* (N, X) *and* (N, Y) *with* $F_X \prec F_Y$*. Then the distribution of* $\sum_{i=1}^{N} X_i$ *is less dangerous than the distribution of* $\sum_{i=1}^{N} Y_i$*.*

Proof: Follows from the Lemma and the fact that $F_S = \sum_{i=1}^{\infty} F^{*i} \cdot P[N = i]$ *.*

Example: Upper bound for a stop-loss premium.

Let's consider a motor liability portfolio with the following underlying model:

Number of claims	$N \sim \text{Poisson}(\lambda)$
Single claim-size	$X \sim F_X$, unknown distribution
Aggregate Loss	$S = \sum_{i=1}^{N} X_i$

We further assume that the insured value (policy limit) is equal to m for all policies in the portfolio and that we know the average claim size $\mu = E[X]$. We want to derive an upper bound for the stop-loss risk premium $E[L_{D,\infty}(S)]$.

The idea is to find a claim-size (severity) distribution which is more dangerous than F_X. An obvious candidate is the random variable

$$Y = \begin{cases} 0 & \text{with probability } 1 - \frac{\mu}{m} \\ m & \text{with probability } \frac{\mu}{m} \end{cases}$$

Using Proposition 7 we immediately see that $F_X \prec F_Y$. Thus,

$$\begin{aligned} E[L_{D,C}(S)] &\leq E\left[L_{D,C}\left(\sum_{i=1}^{N} Y_i\right)\right] \\ &= E\left[L_{D,C}\left(N^* \cdot m\right)\right] \\ &= \sum_{k=0}^{\infty} (km - D)^+ \cdot e^{-\lambda^*} \frac{(\lambda^*)^k}{k!} \end{aligned}$$

where N^* is a Poisson random variable with parameter $\lambda^* = \lambda * P[Y > 0] = \lambda * \frac{\mu}{m}$.

Application in the discretisation of continuous distributions

Very often, eg when one wishes to use the Panjer algorithm, it is necessary to approximate a continuous distribution by means of a discrete distribution From an actuarial point of view one of the desirable characteristics is an approximation which is such that the expected value (ie the risk premium) is conserved. We would now like to present two methods where this characteristic is fulfilled. As we will see, in one case the approximation will be more dangerous whilst in the other case the original distribution will be more dangerous. Let F be a distribution function with mass q between a and b ($b > a$), ie $q = F(b) - F(a)$.

a) Method of mass concentration ($\rightarrow$ less dangerous distribution)

Concentrate the mass q in a point c ($a < c < b$), so that

$$qc = \int_a^b x \; dF(x)$$

b) Method of mass dispersion ($\rightarrow$ more dangerous distribution)

Distribute the mass q over the endpoints a and b so that

$$\begin{aligned} q &= q_1 + q_2 \\ q_1 a + q_2 b &= \int_a^b x \; dF(x) \end{aligned}$$

Lemma 4

$$G \prec F \prec H$$

Proof.

This statement is a direct consequence of Proposition 7 . Here, however, we want to prove a little more and show how the stop-loss transforms of G and H look like.

i) slt_G

For $a \le u \le c$ the following applies:

$$\begin{aligned} slt_G(u) &= \int_u^\infty (1 - G(s)) \; ds \\ &= (c - t)(1 - F(a)) + (b - c)(1 - F(b)) + \int_b^\infty (1 - F(s)) \; ds \end{aligned}$$

and consequently

$$slt'_G(u) = slt'_F(a) \quad \forall u \in [a, c]$$

and via analogous calculation the following is obtained

$$slt'_G(u) = slt'_F(b) \quad \forall u \in [c, b]$$

ii) slt_H

We want to show that the equation

$$slt'_H(u) = \frac{1}{b-a}(slt_F(b) - slt_F(a)) \quad \forall u \in [a,b] \tag{2.36}$$

holds. Let $u \in [a,b]$. The following then applies

$$\begin{aligned}
slt'_H(u)(b-a) &= -(1-H(a))(b-a) \\
&= \int_a^b (1-H(a))\ ds \\
&= -\int_a^b (s-a)\ dH(s) - (b-a)(1-H(b)) \\
&= -\int_a^b (s-a)\ dF(s) - (b-a)(1-F(b)) \\
&= -\int_a^b (1-F(s))\ ds \\
&= slt_F(b) - slt_F(a) \qquad \square
\end{aligned}$$

2.4.2 Exposure curves

Exposure curves, sometimes called deductible-credit curves, appear in rating excess of loss treaties. They are defined as follows:

Definition 6 *The exposure curve belonging to the random variable Z is the function*

$$\begin{aligned}
e_Z \quad &: \quad \mathbb{R}_0^+ \to [0,1], \qquad e_Z(u) = \frac{1}{E[Z]}\int_0^u (1-F_Z(s))\ ds \\
&= \quad \frac{1}{E[Z]}E[L_{0,u}(Z)]
\end{aligned} \tag{2.37}$$

If the distribution of Z has a finite support $[0,M]$ (eg MBBEFD), then the exposure curve belonging to the normalized variable $\frac{1}{M}Z$ defined on the interval $[0,1]$ is usually considered.

Proposition 8 *The exposure curve is an increasing concave function with $e_Z(0) = 0$ and $e_Z(\infty) = 1$ (or $e_Z(1) = 1$ in the normalized case).*

Proof.

The boundary values result directly from the definition. The concaveness derives from the fact that the derivative $e'_Z = 1 - F_Z$ is decreasing. $\square$

The exposure curve $e_Z(u)$ - as a function of the (possibly normalized) priority - shows which part of the total risk premium ($= E[Z]$) the DI retains when ceding the cover ∞ xs u, where the random variable Z describes the extent of loss.

Example (finite support):

An Excess of Loss reinsurance

$$C \text{ xs } D$$

is concluded on a DI's portfolio consisting of similar risks (eg detached houses). The reinsurer is aware of the following values:

- The original premium for the portfolio: Π_0.
- The DI's average loss ratio: 80%
- The sum insured V is assumed to be the same for all risks
- The exposure curve relevant to the class of risks under consideration[4]

  $e_Z : [0,1] \to [0,1]$

On the basis of the first two items of information, the following applies:

$$E[S] = 0.8 \cdot \Pi_0 = E[N] \cdot E[X]$$

The RI premium required for C xs D cover is thus equivalent to:

$$\begin{aligned} E[N] \cdot E[L_{D,C}(X)] &= \frac{0.8 \cdot \Pi_0}{E[X]} \int_D^{D+C} 1 - F_X(s) \; ds \\ &= 0.8 \cdot \Pi_0 \cdot \left(e_Z(\frac{C+D}{V}) - e_Z(\frac{D}{V}) \right) \end{aligned}$$

Note:

The MBBEFD distribution introduced in the section on loss-severity distributions is derived from the adaptation to existing empirically obtained exposure curves.

[4]In actual fact many reinsurers have such curves, developed as a result of many years of experience. Knowledge of the exposure curve is of course equivalent to knowledge of distribution. It is, however, preferable to present such information in the form of exposure curves rather than as distribution functions. The advantage of this is that the reinsurance premium can be calculated from the exposure curve without integration - a factor which was of course more important at the time these curves were introduced (before the information age!) than it is nowadays.

Chapter 3

Risk capital

3.1 Why loadings are required

3.1.1 Ruin-theoretical consideration

Let us imagine a company calculating a premium only accounting for expected losses and the costs incurred (administrative costs, brokerage, etc), in the hope that the available capital reserves would be sufficient to balance out the fluctuations between the expected and the actual results. Such a non-profit-making company would thus require for a treaty with a loss burden S, the premium

$$P = E[S] + K$$

where K is defined so that the sum of all K s from all treaties would provide the exact total costs of the reinsurance. (It is very difficult to attribute the exact costs to each treaty). Unfortunately any such reinsurer would with probability 1 be ruined after a finite period of time, regardless of how large his capital reserve might be. In order to show this we introduce the following quantities:

- S_i = the annual loss burden in the year i ($i = 1, 2, \ldots$)
- P_i = the premiums paid in year i
- K_i = the cost in year i

- $R_i = P_i - S_i - K_i =$ the result in year i
- $B_i = \sum_{k=1}^{i} R_k =$ the total balance following i years

Since we have already used P to represent the premiums, we will use $\mathbb{P}$ to indicate the probability measure. We now make the following assumptions:

1. $\sup_i E[S_{i+1} \mid \sigma(S_1, \ldots, S_i)] < \infty$
2. $\inf_i Var[S_{i+1} \mid \sigma(S_1, \ldots, S_i)] > 0$
3. $P_{i+1} = E[S_{i+1} \mid \sigma(S_1, \ldots, S_i)] + K_{i+1}$

It should be noted that we have not assumed independence, nor do the S_i all have the same distribution. The latter assumption would be completely unrealistic in practice as a company's portfolio structure and thus also the aggregate loss distribution generally varies from year to year.

Theorem 4 *If the annual losses $S_1, S_2, \ldots$ fulfill the above conditions, then the following applies for each $a > 0$:*

$$\mathbb{P}[T_a < \infty] = 1$$

where $T_a = \min\{k \geq 1, B_k \leq -a\}$.

Proof.
First of all we note that $(B_i)_{i \in \mathbb{N}}$ is a Martingale since on the basis of assumption 3:

$$E[B_{i+1} - B_i \mid \sigma(B_1, \ldots, B_i)] = 0$$

A Martingale with bounded increments (assumption 1) can, however, only behave in 2 ways, depending on whether the variance process

$$V_n = \sum_{i=1}^{n} E[R_i^2 \mid \sigma(R_1, \ldots, R_{i-1})]$$

for $n \to \infty$ remains finite or not. It converges towards a finite limit on the set $\{V_\infty < \infty\}$, whilst unlimited oscillation is to be observed on $\{V_\infty = \infty\}$, ie $\limsup_{i \to \infty} B_i = \infty$ and $\liminf_{i \to \infty} B_i = -\infty$. In our case, due to assumption 2, $V_\infty = \infty$ $\mathbb{P}$-a.s. applies and the balance B_i $\mathbb{P}$-a.s. thus achieves every negative level.

3.1.2 Economic consideration

Private insurers need a certain amount of equity capital to compensate for business fluctuations. However, they only obtain this capital in the long-term if they aim at an additional return to the return on risk-free investments, which flows back to the investors in the form of dividends and increased equity. It is the task of management to set an appropriate target for returns. Once this is done is must be defined how much an individual treaty should contribute towards this aggregate profit. The desired global returns must therefore be calculated in profit margins for the individual treaties. In keeping with this, the premium for a RI treaty must comprise the following 3 components:

$$P = P_{risk} + K + M$$

where

$$\begin{aligned} P_{risk} &= \text{the risk premium} = \text{expected loss} \\ K &= \text{the supplement for RI's costs} \\ M &= \text{the profit margin} \end{aligned}$$

The global target return on equity capital is given as

$$\sum_{\text{all treaties i}} M_i$$

In principle complete freedom exists in determining the individual M_i . We will deal with the loading principles used in practice in section 4.3.

3.2 The RAC concept

We would now like to deal in greater detail with the capital required by a reinsurance company and the associated mathematical concepts. This essentially involves the following two problems:

1. Determining the risk capital required

2. Dividing the aggregate capital amongst individual business units with regard to a uniform risk-adjusted performance measurement

The third problem in this respect is the above-mentioned calculation of a global target return on margins for individual treaties ($\rightarrow$ profit margin). We will be dealing with this point in greater detail later.

3.2.1 Determining Risk-Adjusted Capital (RAC)

Quantile and Shortfall as Risk Measures

In order to determine the risk adjusted capital, we first need to define risk. In our context risk is related to the potential negative deviation from the expected outcome of some future events. Let X be the total aggregate loss of the (re-) insurance company in a given period. We consider the following two risk measures associated with X :

1. The *translated p-quantile* of X is defined by

$$Q_p(X) = F_X^{-1}(p) - E[X]$$

   where $F_X^{-1}(p) = \inf\{x : F(x) \geq p\}$ is the generalized inverse of the distribution function.

2. The *p-Shortfall* of X is defined by

$$SF_p(X) = \frac{1}{1-p} \int_p^1 F_x^{-1}(y)\ dy - E(X)$$

Note that in the case where F_X is continuous, the shortfall is equal to the conditional expected value

$$E[X \mid X > F_X^{-1}(p)] - E(X)$$

Both risk measures are widely used in practice, but the shortfall has the advantage that it is a coherent risk measure whereas the quantile is not.

Definition 7 *A risk measure $\rho : \mathcal{L}^1 \to \mathbb{R}$ is called coherent, if the following holds:*

1. *Translation Invariance: For $\alpha \in \mathbb{R}$, $\rho(X + \alpha) = \rho(X)$*

2. *Homogeneity: For $\lambda \in \mathbb{R}^+$, $\rho(\lambda X) = \lambda \cdot \rho(X)$*

3. *Subadditivity: $\rho(X + Y) \leq \rho(X) + \rho(Y)$*

4. *Additivity for comonotonic risks: For $f, g : \mathbb{R}^+ \to \mathbb{R}^+$, increasing,*

$$\rho(f(X) + g(X)) = \rho(f(X)) + \rho(g(X))$$

We would like to have translation invariance in order to distinguish between the contribution to the expected value and the contribution to the volatility. Homogeneity is desired to make the risk measure independent of the unit chosen for risk quantification (especially the currency). Finally, subadditivity enables to quantify the diversification benefit. The reason while the quantile is not coherent lies in the fact that 3. is violated, i.e. subadditivity does not hold.

3.2.2 The Swiss Solvency Test

The Principles of the Swiss Solvency Test are the following:

- All assets and liabilities are valued market consistently

- Risks considered are market, credit and insurance risks

- Risk-bearing capital is defined as the difference of the market consistent value of assets less the market consistent value of liabilities, plus the market value margin

- Target capital is defined as the sum of the Expected Shortfall of change of risk-bearing capital within one year at the 99% confidence level plus the market value margin

- The market value margin is approximated by the cost of the present value of future required regulatory capital for the run-off of the portfolio of assets and liabilities

- Under the SST, an insurers capital adequacy is defined if its target capital is less than its risk bearing capital

- The scope of the SST is legal entity and group / conglomerate level domiciled in Switzerland

- Scenarios defined by the regulator as well as company specific scenarios have to be evaluated and, if relevant, aggregated within the target capital calculation

- All relevant probabilistic states have to be modeled probabilistically

- Partial and full internal models can and should be used. If the SST standard model is not applicable, then a partial or full internal model has to be used

- The internal model has to be integrated into the core processes within the company SST Report to supervisor such that a knowledgeable 3rd party can understand the results

- Public disclosure of methodology of internal model such that a knowledgeable 3rd party can get a reasonably good impression on methodology and design decisions

- Senior Management is responsible for the adherence to principles

The starting point for the Swiss Solvency Test (SST) is thus the economic balance sheet of the company. The available capital is defined as

$$CA(t) = Assets(t) - Liabilities(t) \tag{3.1}$$

whereby both assets and liabilities are valued market consistently. The market consistent valuation of liabilities means that one takes the market value -if exists - or the value of the replicating portfolio of traded financial instruments plus the cost of capital for the remaining "‘basis risk”. The replicating portfolio is a portfolio of financial instruments which are traded in a deep, liquid market, with cash flow characteristics matching either the expected cash flows of the policy obligations or, more generally, matching the cash flows of the policy obligations under a number of financial market scenarios. The replicating portfolio has to match the company specific cash flows, depending on the company specific expenses, claims experience etc. The cost of capital margin is defined as the cost for future regulatory capital which has to be set up for the liabilities. The cost of capital was set for 2008 as 6% over risk-free.

The Solvency Capital Requirement (SCR) captures the risk that the economic balance sheet of the company at $t = 1$ differs from the economic balance sheet at $t = 0$.

SCR $= \rho(CA(1)/(1+r) - CA(0)) + MVM$

(3.2)

where ρ stands for the 99% shortfall risk measure and r is the one year risk free rate.

The market value margin is the cost of future regulatory capital which has to be set up for the liabilities, i.e

MVM $= \sum_{t \geq 1} SCR(t)$

(3.3)

The capital SCR(t) , $t \geq 1$ has to be set up for the run-off risk (i.e. the risk that the actual claims will be higher than reserved ultimates at t=0) and the credit and market risk for the replicating portfolio.

3.2.3 Division of capital and performance measurement

Now that we have determined a company's total risk capital we wish to allocate it to the individual sub-units commensurate to their risk. Sub-units can be of an organisational nature (such as profit centers) or products (eg all non-proportional business). This division aims at judging and comparing the results of the individual sub-units. The riskier the business of a sub-unit is, the more capital should be allocated to it. We will then be in a position to make a direct comparison of the results of the individual areas by measuring the return against the capital allocated. This is referred to as return on risk-adjusted capital or in short RORAC.

It is important to recognise that the risk-adjusted capital defined in this way is an imaginary concept and not something which can be physically attributed to the various business segments. This RAC is (as we have seen in the last chapter) generally much lower than the risk capital which the unit would require if it were an independent company having the same risk tendencies.

It is now a question of defining a key for the risk-commensurate allocation of capital. For this we use the following simplified model:

- The entire company U is divided into n separate sub-units $U_1, \ldots, U_n$
- The results of the sub-units are $R_1, \ldots, R_n$, and those of the whole company $R = \sum R_i$
- $R_i = V_i \cdot Y_i$, where V_i is a measure of volume for the unit i.

For a given risk model (distribution of $(Y_1, \ldots, Y_n)$), the risk of the total portfolio depends only on the volume vector $\underline{V} = (V_1, \ldots V_n)$, i.e. $\rho = \rho(\underline{V})$. The risk of the unit i on a standalone basis is $\rho_i = \rho(0, \ldots, 0, V_i, 0, \ldots, 0)$ and the diversification benefit is defined as

$$\Delta\rho_{div}(\underline{V}) = \sum \rho_i - \rho(\underline{V}) \tag{3.4}$$

For quantifying the contribution of a unit to the total risk, we can use different principles:

1. The marginal principle

$$\Delta_i \rho = \frac{\rho(\underline{V} + \Delta V_i) - \rho(\underline{V})}{\Delta V_i} * V_i \tag{3.5}$$

Since in general $\sum \Delta_i \rho \neq \rho(\underline{V})$, it is better to replace $\Delta_i \rho$ with the normalized contribution

$$\tilde{\Delta_i} \rho = \rho(\underline{V}) * \frac{\Delta_i \rho}{\sum_j \Delta_j \rho} \tag{3.6}$$

2. The 'with and without' principle

$$\Delta_i \rho = \rho(\underline{V}) - \rho(V_1, V_2, \ldots, V_{i-1}, 0, V_{i+1}, \ldots, V_n) \tag{3.7}$$

This is a special case of the marginal principle with $\Delta V_i = -V_i$.

3. The Euler principle

$$\Delta_i \rho = \frac{\partial \rho(\underline{V})}{\partial V_i} * V_i \tag{3.8}$$

Note that according to the Euler Theorem for homogeneous functions we have for all $\lambda > 0$:

$$\sum \frac{\partial \rho(\lambda \underline{V})}{\partial (\lambda V_i)} * V_i = \rho(\underline{V}) \tag{3.9}$$

Especially, for $\lambda = 1$ we obtain

$$\sum \Delta_i \rho = \rho(\underline{V})$$

Theorem 5 *The Euler principle is the only allocation principle which satisfies the axioms of a coherent allocation defined by Denault.*

- *Full allocation:* $\sum K_i = K$, *whereby* K_i *is the capital allocated to unit* i *and* K *is the total capital.*

- *No undercut: the capital allocated to any subportfolio is smaller or equal to the capital determined for the subportfolio on a standalone basis, using the same risk measure.*
- *Symmetry: two units contributing the same risk get the same capital.*
- *Additivity over risk measures:* $\rho = \rho_1 + \rho_2, \Rightarrow K_i(\rho) = K_i(\rho_1) + K_i(\rho_2), \forall i$
- *Riskless allocation:* $\rho_i = 0 \Rightarrow K_i = 0$

Remark: The *covariance principle* is an Euler allocation with $\rho = \sigma$(standard deviation).

In this case we namely have

$$\Delta_i \rho = \frac{\partial \sigma}{\partial V_i} * V_i \tag{3.10}$$

Furthermore,

$$\frac{\partial \sigma}{\partial V_i} * V_i = \frac{V_i}{2} \frac{\partial \sigma^2}{\partial V_i} \frac{1}{\sigma} \tag{3.11}$$

Using

$$\sigma^2 = \sum_{k,j} Cov[R_k, R_j]$$

we get

$$\frac{V_i}{2} \frac{\partial \sigma^2}{\partial V_i} = Cov[R_i, R]$$

and using (3.11) we get

$$\Delta_i \rho = \frac{Cov[R, R_i]}{\sigma(R)} \tag{3.12}$$

ie. the capital must be allocated in proportion to the covariances of each sub-unit with the total business.

3.3 Appendix to Chapter 3: Some properties of Quantiles

Here we ant to discuss some properties of quantiles. We will especially prove that quantiles are not subadditive, but they are additive for comonotonic risks.

Definition 8 *Two random variables X and Y on a probability space $(\Omega, \mathcal{F}, P)$ are* comonotonic, *if there exists a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ and $(\widetilde{X}, \widetilde{Y})$ random variables on $\widetilde{\Omega}$ such that*

1. $(\widetilde{X}, \widetilde{Y}) \stackrel{d}{=} (X, Y)$
2. $\widetilde{X}(\omega_1) \leq \widetilde{X}(\omega_2)$ *implies* $\widetilde{Y}(\omega_1) \leq \widetilde{Y}(\omega_2) \quad \forall \omega_1, \omega_2 \in \widetilde{\Omega}$.

Proposition 9 *If X and Y have continuous marginals, it holds that*

X *and* Y *comonotonic* $\iff$ $Y = g(X)$ *a.s., with* $g = F_Y^{-1} \circ F_X$ *increasing*

Lemma 5 *Let X and Y denote two random variables such that there exists an increasing function g, such that Y=$g(X)$. In addition suppose that g is continuous. Let $x_p := F_X^{-1}(p)$ and $y_p := F_Y^{-1}(p)$ denote the quantile functions of X and Y. Then*

$$y_p = g(x_p), \tag{3.13}$$

that means, that the quantiles transform the same way like the random variables.

Proof.

(i) g increasing $\implies \quad \{\omega \in \Omega : X(\omega) \leq x_p\} \subseteq \{\omega \in \Omega : g(X(\omega)) \leq g(x_p)\}$,

$$\implies \quad P[g(X) \leq g(x_p)] \geq P[X \leq x_p] \geq p. \tag{3.14}$$

(ii) Since g is continuous we know that

$\forall \varepsilon > 0 \quad \exists \delta > 0$ such that $g(z) > g(x_p) - \varepsilon$, whenever $z > x_p - \delta$

thus $g(z) \leq g(x_p) - \varepsilon \quad \implies \quad z \leq x_p - \delta$, therefore $\{\omega \in \Omega : g(X(\omega)) \leq g(x_p) - \varepsilon\} \subseteq \{\omega \in \Omega : X(\omega) \leq x_p - \delta\}$,

hence

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 : \quad P[g(X) \leq g(x_p) - \varepsilon] \leq P[X \leq x_p - \delta] < p \tag{3.15}$$

$\Longrightarrow y_p = \inf\{z : P[g(X) \leq z] \geq p\} = g(x_p)$.

In the following theorem we will show that in a case of two comonotonic random variables the sum of their quantiles is equal to the quantile of their sum.

Theorem 6 *Let X and Y denote two comonotonic random variables. Then we have*

$$F^{-1}_{X+Y}(p) = F^{-1}_X(p) + F^{-1}_Y(p)\;)$$

Proof.

Because of comonotonicity there exist continuous increasing functions $u, v : \mathbb{R} \to \mathbb{R}$ and a random variable Z such that

$$(X, Y) \stackrel{d}{=} (u(Z), v(Z))$$

and thus

$$X + Y \stackrel{d}{=} (u + v)(Z)$$

Using the previous Lemma and the fact that $u + v$ is a continuous increasing function we obtain

$$F^{-1}_{X+Y}(p) = (u + v)(F^{-1}_Z(p)) = u(F^{-1}_z(p)) + v(F^{-1}_z(p)) = F^{-1}_X(p) + F^{-1}_Y(p)\;)$$

We now show that quantiles are not subadditive by giving a counterexample:

Lemma 6 *If X and Y are independent random variables,* $X, Y \sim Pareto\left(1, \frac{1}{2}\right)$, *then*

$$P(X + Y \leq 2z) < P(X \leq z) \quad \forall z > 1. \tag{3.16}$$

Proof.

$$z > 1 \quad \Rightarrow \quad 2z > z + 1 \quad \Rightarrow \quad 2z - 1 > z \quad \Rightarrow \quad \sqrt{2z-1} > \sqrt{z}$$

$$\Rightarrow \quad \sqrt{2z-1} > \frac{z}{\sqrt{z}} \quad \Rightarrow \quad \frac{\sqrt{2z-1}}{z} < \frac{1}{\sqrt{z}} \quad \Rightarrow \quad \frac{2\sqrt{2z-1}}{2z} > \frac{1}{\sqrt{z}}$$

Hence

$$P(X + Y \leq 2z) = 1 - \frac{2\sqrt{2z-1}}{2z} \quad < \quad 1 - \frac{1}{\sqrt{z}} = P(X \leq z)$$

Let X, Y be independent $Pareto\left(1, \frac{1}{2}\right)$ distributed random variables and $0 < p < 1$.Then:

$$F_{X+Y}^{-1}(p) > F_X^{-1}(p) + F_Y^{-1}(p)$$

Proof.

Let $p \in (0,1)$.Then $z = F_X^{-1}(p) > 1$.Form the previous Lemma we then have

$$P\left[X + Y \leq 2z\right] < P\left[X \leq z\right] = p$$

Thus,

$$F_{X+Y}^{-1}(p) > 2z = F_X^{-1}(p) + F_Y^{-1}(p)$$

Chapter 4

Rating non-proportional reinsurance treaties

In this chapter we will investigate the methods for determining the premium necessary for a reinsurance treaty. Before going into further detail, we need as a first step to understand the components of the premium, i.e. all the costs incurred by a reinsurer. For this, we look at the economic value generated by an insurance contract: Every insurance contract generates a number of cash flows, some of them positive, others negative. In order to evaluate the economic value of the resulting cas-flow stream, we need first to define at which rate we discount these cash flows. The principle we use here is a strict separation of underwriting and investment activities: we assume that the investment department grants to the (re-)insurance operation a risk free return (for an appropriate asset management fee). Thus we discount all cas-flows at the risk free rate to get the present value (PV). Let's now look at all the cash-flows which incur for a reinsurance contract:

1. Premiums
2. Commissions
3. Claims
4. Acquisition Expenses

5. Runoff Expenses
6. Capital Costs
7. Taxes
8. Asset Management Fees

We assume that all items are present values, i.e. already discounted at the risk free rate. On the internal expense side we have differentiated between the acquisition expenses which are associated with new business written (marketing, underwriting, etc) vs. the runoff expenses which incur after the contract has been written (claims management, reserving,etc). The capital costs represent the target return on the capital allocated to this contract. We call (1)-(2)-(3)-(4)-(5)-(6)-(7)-(8) the economic profit generated by this contract. This profit is thus after deduction of all expenses and after tax. Note that the tax item consists of two components, namely the tax on the underwriting profit and the tax on the risk free return on the allocated capital. The latter item is sometimes also called "‘double taxation cost", reflecting that (re-)insurance companies are in a competetitive disadvantage vs. an investment fund because the investment income available to their shareholders is after tax (and shareholders will typically pay again tax on the dividends or capital gains).

Let's now consider a direct insurer's portfolio and take $X_1, \ldots, X_N$ to represent the original losses that the portfolio will produce during its treaty period, extending into the future. We now look at the most general case of a non-proportional reinsurance contract, namely an excess of loss treaty C xs D with an aggregate deductible AD and an aggregate limit AL. The reinsured loss from such a treaty is

$$S_{RV} = L_{AD,AL}(S_{XL})$$

where

$$S_{XL} = \sum_{i=1}^{N} L_{D,C}(X^i)$$

In the case $AD = 0$, $AL = \infty$ we have a simple excess of loss contract, and in the case $D = 0$, $C = \infty$ a stop loss AL xs AD. The rating procedure, i.e. the calculation of the required premium, can be divided into the following steps:

1. Determining the distributions of loss severity and number of losses
2. Determining the risk premium with respect to AD, AL and possibly other treaty conditions
3. Determining the payout pattern for losses and premiums and discounting them at the risk free rate.
4. Determining the loadings for items (4)-(8)so that the contract produces the required economic profit.

We will now discuss the individual steps in greater detail:

4.1 Determining the distributions of loss severity and numbers of losses

4.1.1 Determining loss severity distribution from previous loss experience

In this procedure the reinsurer does not use his own statistics but instead bases his calculations exclusively on the loss experience of the portfolio covered. Normally the direct insurer provides him with information regarding previous original losses from a certain observation period (usually 5-10 years). This does not, however, comprise all losses, but only those which exceed a certain threshold s_0 . Usually s_0 is fixed as half of the deductible. According to Proposition 4 we can determine the conditional distribution given $X > s_0$ instead of the original severity distribution. In order to be able to determine this loss distribution from observations, we must first of all make appropriate adjustments to the statistical material. Here, we must pay particular attention to the following factors:

1. Economic changes (inflation, wage levels)
2. Changes in the portfolio (growth, decline, other portfolio mixes)
3. Technical changes

4. Changes in legislation

5. Changes in insurance conditions

In the case of all of the above factors, it has to be considered whether they influence the number or the amount of losses (or both), and to what extent. Whereas the influence of changes in legislation and technical innovations can only be roughly estimated - if indeed at all, the influence of inflation or of changes in the size of a portfolio can be quantified quite accurately. For this reason we intend here to deal only with these two factors. Henceforth we will use the following notation:

1. - k : Number of years of observation
   - n_i : Number of losses (reported by DI) $\geq s_0$ in year i $(1 \leq i \leq k)$
   - $x_{i1}, \ldots, x_{in_i}$: reported losses $\geq s_0$ in year i

The effect of inflation (or deflation) is such that all losses in a given year i become more expensive by a factor q_i (or cheaper by a factor $q < 1$), ie we have to replace the observed original losses x_{ij} with $q_i x_{ij}$. Attention should also be paid to the fact that, following adjustment for inflation, the losses in year i are only known above the threshold $s_{o,i} = q_i s_0$ which means that all loss information is available only from the threshold $s_o^* = \max_i s_{0,i}$.
In order to now consider the change in the size of the portfolio, we first of all have to introduce a suitable measure of volume. Usual measures of volume are for example:

Branch	Measure of volume
Fire	Original premiums or sum insured
Motor liability	Original premiums or number of vehicles
Aviation loss	Fleet value of airline
Aviation liability	Passenger kilometers flown

We use V_i to represent the volume of the i-th observation year and V_T as the (expected) volume of the year to be rated (if the premiums are taken as a measure of volume, they must of course be adjusted to compensate for inflation and possible tariff adjustments). The effect of the change in volume depends on the type of cover.

In the case of per risk cover (eg fire WXL) it can be assumed that the number of losses will grow proportionally to the increase in volume, whereas the average loss amount will remain the same. In the case of per event covers (eg windstorm CatXL) it is not the number of events which increases but the number of risks which are affected by an event. Similarly, it is not the frequency of losses which changes but the severity of the losses. Due to the above considerations we must therefore apply the following corrections:

$$\begin{array}{lll} \text{CatXL:} & n_i \to n_i, & x_{ij} \to q_i \frac{V_T}{V_i} x_{ij} \\ \text{WXL:} & n_i \to \frac{V_T}{V_i} n_i, & x_{ij} \to q_i x_{ij} \end{array} \tag{4.1}$$

Exkursus: Effect of inflation on excess losses

We would now like to examine this point in greater detail as it is so important in practice. If an original loss increases in cost by a factor $a > 1$, then the following applies for losses in excess of priority D

$$\begin{aligned} E[L_{D,\infty}(aX)] &= \int_D^\infty (1 - F_{aX}(s))\, ds \\ &= \int_D^\infty (1 - F_X(\frac{s}{a}))\, ds \\ &= a \int_{D/a}^\infty (1 - F_X(s))\, ds \\ &= a \cdot E[L_{D/a,\infty}(X)] \end{aligned}$$

and subsequently for a finite layer

$$E[L_{D,C}(aX)] = a \cdot E[L_{D/a,C/a}(X)] \tag{4.2}$$

Although it is conceivable that there may be cases where the layer with priority D/a and cover C/a has a smaller loss expectation than the original layer (eg exponentially distributed losses with expected value $< D$), for all relevant cases in practice the opposite is the case, ie the loss in the layer grows by a factor $> a$ (ie disproportionately).

As an illustration of this we consider a Pareto (x_0, α) distributed original loss X. In accordance with the formula (2.8) the following then applies:

$$E[L_{D/a,C/a}(X)] = a^{\alpha-1} \cdot E[L_{D,C}(X)]$$

ie the loss in the layer increases by factor $a^\alpha > a$. *(end of exkursus)*

In branches having a longer settlement period (eg liability), only a rough estimate of the actual final loss burden is available in the case of most losses. For this reason the reinsurer needs to implement an IBNR-procedure on an individual loss basis in order to get from the previously observed data n_i and x_{ij} (and adjusted - in accordance with the above) to the actual numbers of losses (IBNR effect) resp. loss amounts (IBNER effect). The loss amount must then be discounted since the reinsurer is then able to invest premiums obtained until payment for losses is required.

We have now got as far as adjusting the data and can determine the distribution we are seeking. The first idea is to look at the empirical loss distribution, ie the distribution with the distribution function.

$$F_{emp}(x) = \frac{\#(X_{ij},\ X_{ij} \leq x)}{\sum_i n_i}$$

However, this distribution function only describes the distribution of the actual loss burden in a certain domain, in particular $F_{emp}(x) = 1$ if x is greater than the greatest of all losses observed in the past. This function is thus only of use in rating if a sufficient number of losses are available in the layer. Rating by means of the empirical distribution function is often termed "burning cost rating".

If the loss experience in the layer is not representative (eg if the greatest observed excess loss is smaller than the cover C, or - even more extreme - if so far no loss at all has been observed in the layer), then an additional model is needed for rating. In this case an attempt is made to adapt an analytical loss distribution to the observed data. The question arising here is how many of the observed losses should be taken into account for adapting the chosen distribution. Admittedly, we know all the losses $\geq s_0$, however, we only wish to rate the layer with deductible $D > s_0$. On the one hand it is desirable to consider as many losses as possible in making the adjustment to provide a reliable statistical basis; on the other hand the distribution of the small losses is often very different from the distribution of the higher losses and consequently distorts the adjustment.

In cases where a Pareto distribution is used for adjustment (which is very often the case in practice since it fits very well in the area of major losses) it is for example possible to proceed as follows in choosing an appropriate threshold, above which the loss experience is to be considered:

If the losses above a threshold s_1 are really Pareto distributed, then they are also so from every other threshold greater than $s_1$on account of (2.7). This can be used as follows to determine s_1 and the Pareto parameter α : If the parameter α is calculated for values of x_0 in the interval $[s_{0,}D]$ with a maximal likelihood estimate, ie with the estimator

$$\hat{\alpha} = \frac{\sum_i n_i(x_0)}{\sum_{i,j} \ln(X_{ij}/x_0)} \tag{4.3}$$

where $n_i(x_0)$ denotes the number of losses of the year i greater than x_0 (following correction), then $\alpha(x_0)$ should remain constant beyond a certain threshold s_1. Then we chose $x_0 = s_1$ and $\alpha = \alpha(s_1)$.

In practice this estimate is often vague given the small number of observations (resulting in a very large variance for the ML estimator). In cases where we have available a priori information about the possible range of the parameter, we therefore better use a Bayesian estimator. For the Pareto distribution this is very often the case, as we have outlined on page 11.

Example:

We investigate a fire portfolio of a primary insurer and want to fit a Pareto(x_0, α)-distribution to l observed losses. As a rule of thumb we have:

$$1.5 \leq \alpha \leq 2$$

We assume that α itself a gamma distributed random variable with expected value $E[\alpha] = 1.8$ and standard deviation $\sigma(\alpha) = 0.3$. The gamma distribution has the density

$$f(x) = \frac{c^\gamma}{\Gamma(\gamma)} x^{\gamma-1} e^{-cx} I_{\{x>0\}}$$

and we have

$$E[\alpha] = \frac{\gamma}{c}; \qquad Var[\alpha] = \frac{\gamma}{c^2}$$

leading in our case to the values $c = 20$, $\gamma = 36$. The important result to note is that the conditional distribution of

$$\alpha \;/\; X_1, \ldots, X_l$$

is again a gamma distribution with parameters

$$\begin{aligned}
\gamma' &= \gamma + l \\
c' &= c + T, \text{ with } T = \sum_{j=1}^{l} \log(x_j / x_0)
\end{aligned}$$

The Bayes-estimator is thus given by

$$\widehat{\alpha} = E\left[\alpha\right] = \frac{\gamma'}{c'} = \frac{\gamma + l}{c + T}$$

4.1.2 Determining loss severity distribution according to the exposure method

With this method the reinsurer draws on the composition of the portfolio and the direct insurer's premiums. It only works for the case of property insurance (fire insurance) where the direct insurer provides a so-called risk profile, whereby he grades his portfolio into different classes according to the sums insured and lists the original premium for each class. Here is an example of a typical risk profile for an industrial fire portfolio:

Sum insured	Average SI	Number of risks	Premium rate	Loss ratio
1-5 Mio	2.8 Mio	56440	1.92 ‰	70%
5-10 Mio	7.0 Mio	6700	1.81 ‰	70%
10-30 Mio	17 Mio	3520	1.60 ‰	70%
30-100 Mio	54 Mio	860	1.15 ‰	70%
100- 500 Mio	220 Mio	250	1.00 ‰	70%

The reinsurer now determines an appropriate loss distribution for every class of risk. This is carried out with the help of market statistics or using his own statistics from proportional reinsurances where he knows the extent of all individual losses. In accordance with the above risk profile we now know for each class i:

- the sum insured V_i (identical for all risks in the class[1])
- the number of risks m_i
- the normalized loss severity distribution, ie. the distribution of X_i/V_i
- the original premium P_i
- the direct insurer's loss ratio k_i[2]

The advantage of the normalized loss distribution compared to the absolute loss distribution is that it is not affected by inflation and currency fluctuations. In addition, it can be assumed that this distribution is constant for each risk class. In order to model the aggregate loss of class i we use the mixed model (N_i, Ξ_i), whereby Ξ_i denotes the normalized loss-severity variable of the class i . The aggregate loss of the period covered by the treaty in class i is then

$$S_i = \sum_{j=1}^{N_i} \Xi_i^j \cdot V_i$$

The portfolio's aggregate loss is of course

$$S = \sum S_i$$

We determine the expected value $\lambda_i = E[N_i]$ using the relation

$$E[S_i] = k_i \cdot P_i = \lambda_i \cdot V_i \cdot E[\Xi_i]$$

thus,

$$\lambda_i = \frac{k_i P_i}{V_i E[\Xi_i]}$$

In the case where all N_i are Poisson distributed the aggregate loss of the "individual model" has - due to (2.24) - the same distribution as the aggregate loss in a collective

[1] This approximation can be improved further by subdividing each class i into sub-classes, assuming a certain distribution of sums insured per class.

[2] so that $E[S_i] = k_i P_i$

model with a Poisson (λ) distributed number of loss variables N , where $\lambda = \sum \lambda_i$, and a loss-severity variable Y with the distribution

$$F_Y = \sum_i \frac{\lambda_i}{\lambda} F_i$$

From this it ensues that the following also applies for the distribution of aggregate loss in the layer C xs D :

$$\sum_i \sum_{j=1}^{N_i} L_{D,C}(X_i^j) \overset{d}{\sim} \sum_{j=1}^{N} L_{D,C}(Y^j)$$

and instead of the individual model we can consider the above collective model.

Where the number of losses is best described by means of a Polya distribution, it is possible to move on to a Poisson model and a transformed loss severity distribution for each band using an Ammeter transformation. However, the Ammeter transformation $N \to \tilde{N}$, $X \to \widetilde{X}$ only guarantees that $S_i \overset{d}{\sim} \widetilde{S}_i$, whereas the corresponding aggregate loss distributions in the layer are no longer the same, ie

$$\sum_{j=1}^{N_i} L_{D,C}(X_i^j)$$

does not have the same distribution as

$$\sum_{j=1}^{\tilde{N}_i} L_{D,C}(\widetilde{X}_i^j)$$

For this reason it is not necessary to transform the loss-severity distribution itself, but rather the distribution of $L_{D,C}(X)$. The disadvantage of this method is that it is only possible to rate a single layer. (Practical cases usually involve an entire reinsurance program consisting of several layers).

Exposure rating of natural hazard covers

The procedure in the case of *natural hazard covers* is somewhat different. A physical model is first of all used to determine the loss distribution whereby the natural hazard under consideration is described using a family of random variables $(M_1, \ldots, M_l)$. For an earthquake, one variable (the magnitude) is sufficient ; other natural hazards such as wind storm require several variables. For the sake of simplicity we will limit ourselves here to one variable M. It is now a question of establishing the distribution of M, whereby it is of course enough to investigate values of M which can lead to physical losses. This distribution is usually determined using statistical data relating to past events. It has to be noted, however, that the magnitudes of historical events are not independent of each other (this is obvious for earthquakes), so the distribution to be determined is the conditional distribution for the next period, given the observed history.

The next step involves creating a relation between M and the degree of loss of the class of risk under consideration. This requires a somewhat different risk profile than in the case of fire insurance. The risk-prone areas are divided into zones (→ CRESTA zones) and the sum of the insured values is registered for each zone. This is necessary because the intensity of an event with a certain magnitude depends on the zone (in the case of earthquakes it is for example the distance from the epicenter which is significant). In addition, a further division into various risk classes per zone is required in order to model the various degrees of vulnerability to loss. This depends on many factors such as the construction quality of buildings, the stability of the ground, symmetry etc. For every zone i and class of risk j a function $q_{i,j}$ is determined, so that $\Xi_{i,j} = q_{i,j}(M)$. The aggregate loss for the portfolio is now

$$S_{nat} = \sum_{k=1}^{N_{nat}} \sum_{i,j} T_{i,j} \cdot q_{i,j}(M^k)$$

where N_{nat} represents the number of the natural hazard events considered for the treaty period and $T_{i,j}$ the sum of all insured values in the corresponding zone and risk class.

4.1.3 Determining the distribution of the number of losses

In practice either the Poisson or the Polya distribution is almost always used in calculations. The reasons for this have already been explained several times in the preceding sections. In practice it is therefore usually a question of deciding which of the two distributions is most appropriate to the observed loss data. It surely makes sense first of all to use the Panjer factor (see page 22). More demanding statistical methods such as the Chi square test can of course also be used. The problem with such tests is that often in reinsurance only a very small data base is available.

It is nevertheless possible to rely on certain rules of thumb; it is known for example that in the case where waiting times between individual loss events are not independent (eg earthquakes), the Polya distribution is often more appropriate than the Poisson distribution.

4.2 Determining the risk premium

We have seen in the last section how the distributions of loss severities and numbers of losses can be determined. If a fixed premium is asked for a treaty this can be calculated as follows:

Using the Panjer algorithm the distribution function $F_{S_{XL}}$ of

$$S_{XL} = \sum_{i=1}^{N} L_{D,C}(X^i)$$

is first determined. The desired premium is the expected value of S_{RI} , ie. according to (2.35)

$$E[S_{RI}] = \int_{AD}^{AL} 1 - F_{S_{XL}}(s)\ ds$$

In practice, however, there is not usually a fixed premium but a loss-dependent premium function $P(X^1, X^2, \dots X^N)$. The risk premium then corresponds to a function $P_{risk}(\cdot)$, which fulfills the relation.

$$E[P_{risk}(X^1, X^2, \dots X^N)] = E[S_{RI}] \tag{4.4}$$

We would now like to discuss the most commonly used types of premium functions:

Reinstatement premiums

For treaties having this kind of premium the aggregate limit is given as a multiple of the cover C , ie $AL = (k+1)C$. The constant k represents the number of reinstatements. Here the cedent is provided with the cover C for a basic premium P_{bas}. It is further conceivable that the payment Z made by a reinsurer "eats away" part of the cover C and that in the case of an additional loss event the cedent is left with only the cover $(C - Z)^+$. He than has to "top up" the cover to the level C by paying the appropriate additional premium.

We imagine that k pots each with the contents C are available for reinstatement, whereby the price of each pot is fixed as a multiple β_i $(1 \leq i \leq k)$ of the basic premium. We now consider the following example as an illustration:

A treaty has the cover 20 xs 10 and 2 reinstatements with $\beta_1 = 1$, $\beta_2 = 0.5$. For the sake of simplicity we make $AD = 0$ in this example. Let us now assume that the following losses occur in the given sequence:

$$\begin{aligned} X^1 &= 15 \\ X^2 &= 27 \\ X^3 &= 38 \\ X^4 &= 22 \end{aligned}$$

The RI pays the amount 5 for the first loss. Without reinstatement the cedent would only be left with the cover 15 xs 10. He can now increase the cover on the original amount of 20 by buying the required amount (ie 5) at the price of the first pot .

$$P_{bas} \cdot \beta_1 \cdot \frac{L_{D,C}(X^1)}{C} = P_{bas} \cdot \frac{1}{4}$$

He now has the full cover of 20, the pot 1 still contains 15, pot 2 remains untouched and thus still contains 20.

In the case of the next loss of 27 the RI pays 17 and the cedent can top up his cover by buying 15 from pot 1 and 2 from pot 2. The additional premium to be paid is thus

$$P_{bas} \cdot \frac{3}{4} + P_{bas} \cdot \frac{1}{2} \cdot \frac{2}{20}$$

The RI pays 20 of the next loss (38). The cover can, however, only be topped up to 18 (for the price $P_{bas} \cdot \frac{1}{2} \cdot \frac{18}{20}$) and all pots are now empty. The RI pays 12 of the fourth loss and there are no more additional premiums to be paid. However, the cedent is only left with cover 6, ie for a possible further loss only the cover 6 xs 10 is available.

We will formulate the situation for the general type of case. The RI loss is provided by the formula (1.5) and the premium function has the following form in view of the above considerations

$$P_{risk} = P_{bas} \cdot \left(1 + \sum_{i=0}^{k-1} \frac{\beta_{i+1}}{C} \cdot L_{AD+iC\ ,C}\ (S_{XL}) \right) \tag{4.5}$$

The basic premium can now be calculated from (4.4) and we obtain:

$$P_{bas} = E[S_{RI}] \cdot \left(1 + \sum_{i=0}^{k-1} \frac{\beta_{i+1}}{C} \cdot E[L_{AD+iC\ ,C}\ (S_{XL})] \right)^{-1} \tag{4.6}$$

To calculate the basic premium the distribution of the aggregate annual loss S_{XL} is needed. This is usually carried out in practice using the Panjer algorithm. In many special cases the formula (4.6) becomes much simpler, particularly if all β_i are equal. Let $w_{k,\beta}$ be the basic premium for k reinstatements with $\beta_i = \beta\ \forall i$ and $AD = 0$. Then with (4.6) we obtain:

$$\begin{aligned}
w_{\infty,0} &= E[S_{XL}] \qquad (\infty \text{ no. of free reinstatements}) \\
w_{\infty,\beta} &= \frac{w_{\infty,0}}{1 + \beta/C \cdot w_{\infty,0}} \qquad (\infty \text{ no of reinstatements at rate } \beta) \\
w_{k,0} &= E[\max\{S_{XL}, (k+1)C\}] = w_{\infty,0} - slt_{S_{XL}}((k+1)C) \\
w_{k,\beta} &= \frac{w_{k,0}}{1 + \beta/C \cdot E[L_{0,\ kC}(S_{XL})]} = \frac{w_{k,0}}{1 + \beta/C \cdot w_{k-1,0}}
\end{aligned}$$

Exercise: Modification of the above formula for the case $AD \neq 0$.

Slide with fixed loading

This type of loss-dependent premium can in principle be applied to all RI treaties. The premium function only depends on the reinsured aggregate annual loss S_{RI} :

$$P_{risk} = P_{risk}(S_{RI}) = \begin{cases} m & \text{if } S_{RI} \le m - l \\ S_{RI} + l & \text{if } m - l \le S_{RI} < M - l \\ M & \text{if } S_{RI} \ge M - l \end{cases}$$

The premium is thus equal to the loss S_{RI} plus a fixed loading l, whereby a minimal premium m and a maximum premium M are fixed. Written somewhat differently:

$$\begin{aligned} P_{risk}(S_{RI}) &= m + (S_{RI} - m + l)^+ - (S_{RI} - M + l)^+ \\ E[P_{risk}(S_{RI})] &= m + slt_{S_{RI}}(m - l) - slt_{S_{RI}}(M - l) \end{aligned} \tag{4.7}$$

Usually the minimum premium m and the loading l are first of all defined and M is then calculated so that (4.4) is fulfilled.

Slide with progressive loading

This is a variant of the type of premium described above, whereby the loading is not fixed but rather proportional to the loss:

$$P_{risk} = P_{risk}(S_{RI}) = \begin{cases} m & \text{if } S_{RI} \cdot a \leq m \\ S_{RV} \cdot a & \text{if } m \leq S_{RI} \cdot a < M \\ M & \text{if } S_{RI} \cdot a \geq M \end{cases}$$

or written differently

$$\begin{aligned} P_{risk}(S_{RI}) &= m + a \cdot (S_{RI} - m/a)^+ - a \cdot (S_{RI} - M/a)^+ \\ E[P_{risk}(S_{RI})] &= m + a \cdot slt_{S_{RI}}(m/a) - a \cdot slt_{S_{RI}}(M/a) \end{aligned} \tag{4.8}$$

4.3 The Loading for Profit and Capital Costs

4.3.1 General considerations

We have seen, both from an economic and a risk-theoretical point of view that loading is necessary in addition to the expected losses and expenses. We now consider a portfolio with the associated RAC and a return target on this capital. With this the planned total profit G is given and the margins (loading) for the individual treaties are to be determined in such a way that

$$G = \sum_{\text{all treaties i}} M_i$$

We do not intend to define the precise nature of the RI treaty here; the following applies in principle to all types of treaty. In this chapter we will represent the total reinsured loss of the treaty i with S_i and the total loss of the reinsured portfolio $(= \sum_i S_i)$ with S, ie we omit the index "RI". It should, however, be remembered that it is a question not of original loss but of RI loss.

The theory of course suggests breaking down the RAC to the level of individual acceptances using the principle of covariance, according to which the loading should be calculated using the formula:

$$M_i = \frac{G}{Var[S]} \cdot Cov[S_i, S] \tag{4.9}$$

Later we will deal with an example where this principle can also be implemented successfully in practice. Unfortunately this is often not the case since the covariance of a single treaty with the whole portfolio cannot be determined. For this reason the correlation is often disregarded in practice and attention is only given to the fluctuation of the individual treaty. This provides the following principles:

1. The variance principle

$$M_i = \frac{G}{\sum_j Var[S_j]} \cdot Var[S_i] \tag{4.10}$$

In the case of uncorrelated risks this of course corresponds to the covariance principle .

2. The standard deviation principle

$$M_i = \frac{G}{\sum_j \sigma(S_j)} \cdot \sigma(S_i) \tag{4.11}$$

3. The root rate on-line principle

   This principle is particularly common for excess of loss in property but can in principle also be applied to all RI covers with finite maximum liability H. In the case of excess of loss, H is precisely the cover which we previously called with C. The so-called risk rate on-line of treaty i is defined by

$$r_i = \frac{E[S_i]}{H_i} \tag{4.12}$$

   and the loading concept now reads

$$M_i = \frac{G}{\sum_j H_j \sqrt{r_j}} \cdot H_i \cdot \sqrt{r_i} \tag{4.13}$$

   Note: If only total losses are possible, ie

$$S_i = N_i \cdot H_i$$

   and the number of losses N_i is Poisson distributed, then:

$$\begin{aligned} \sigma(S_i) &= H_i \sqrt{E[N_i]} \\ &= H_i \cdot \sqrt{r_i} \end{aligned}$$

   In this way the amount $H_i \cdot \sqrt{r_i}$ is also a measure of fluctuation corresponding to the standard deviation given the above conditions.

For all of the above principles the problem arises in practice that at the time of rating the reinsurer does not know his future portfolio and thus has to estimate the relevant parameters on the basis of the current or planned portfolio. Depending on the principle applied, he thus determines a coefficient α, so that

$$M_i = \alpha_1 \cdot Var[S_i] \qquad \text{or} \tag{4.14}$$

$$M_i = \alpha_2 \cdot \sigma(S_i) \qquad \text{or} \tag{4.15}$$

$$M_i = \alpha_3 \cdot H_i \cdot \sqrt{r_i} \tag{4.16}$$

These principles, whilst convenient in practice, result in a few problems from a theoretical point of view. One of these is that in dividing the cover into part coverages they are generally *non-additive*. In order to illustrate this we consider an excess of loss C xs D. If we divide this layer into two sublayers C_1 xs D_1 and C_2 xs D_2, so that

$$D_1 + C_1 = D_2$$
$$C_1 + C_2 = C$$

then $S_1 = \sum_{j=1}^{N} L_{D_1\ ,C_1}(X_j)$ and $S_2 = \sum_{j=1}^{N} L_{D_2\ ,C_2}(X_j)$ are of course positively correlated and thus

$$Var[S] > Var[S_1] + Var[S_2]$$

Using the variance principle for the margins this means:

$$M > M_1 + M_2$$

In this way a treaty would become increasingly cheaper on being divided up into part treaties. Similarly this problem occurs with the other two principles. The covariance principle is by contrast additive and thus in this way also satisfactory. However, no simplification is possible for the formulae (4.14 - 4.16) which makes implementation even more difficult.

In order to alleviate the problem of lacking additivity, the so-called infinitesimal ROL principle was introduced. The idea is to divide the cover H into infinitesimal part coverages dh and then to integrate the loading for these part coverages. Due to

$$r = \frac{1}{dh} \int_{d}^{d+dh} 1 - F_S(s)\ ds$$

then for an infinitesimal cover dh attaching at the deductible d

$$r = 1 - F_S(d)$$

applies and according to (4.16) the appropriate loading is equal to

$$m = \alpha \cdot dh\sqrt{1 - F_S(d)}$$

For the total loading this gives

$$M = \alpha \int_d^{d+H} \sqrt{1 - F_S(x)}\ dx \tag{4.17}$$

and this loading principle is clearly additive.

4.3.2 The loading on-line concept for Cat-XL treaties (according to Bernegger, 1994)

We will now deal with one example where the covariance principle can be successfully implemented in practical terms. The example involves margins for Cat-XL treaties. As a reminder: Cat-XL treaties are excess of loss treaties where the basic original loss comprises the total of all losses in a direct insurer's portfolio arising from one natural disaster (wind storm, earthquake, flood, hail).

We now consider a certain natural hazard potential (eg earthquakes in Japan) and assume that the appropriate RAC - and thus the global target for returns - has already been determined for the reinsurer's entire NP portfolio relating to this potential. By NP portfolio we mean the total of all Cat-XL treaties:

$$C_1 \text{ xs } D_1, \ldots, C_l \text{ xs } D_l$$

These treaties relate to the portfolios of various direct insurance companies, whereby the risks are all located in the same geographical area and exposed to the relevant natural hazard. Individual layers may of course overlap. We now want to determine the covariance of an individual Cat-XL layer using the reinsurer's whole portfolio. For this purpose we need to establish a relation between the loss in the layer and the aggregate loss of the portfolio (we assume the distribution of the latter to be known since major reinsurers have good natural hazard models at their disposal and can determine the loss distribution of their portfolios quite accurately).

First of all we intend to create an approximation for the relation between the loss in a layer and the aggregate loss of the portfolio. For this we first standardize all values (covers, deductibles and losses) with the portfolio's whole cover ($= \sum C_i$) and assume that the individual covers C_i are small compared to the whole coverage. We represent the standardized covers as δ_i and in accordance with our assumption $\delta_i \ll 1$. For the other variables we retain the old notation. For small covers the probability $P[L_{D_i,\delta_i}(X_i) > 0]$ can be approximated using the quotients of the ROL and the expected number of losses, since for ROL r_i the following applies:

$$r_i = E[N] \cdot \frac{1}{\delta_i} \int_D^{D+\delta_i} 1 - F_{X_i}(s)\; ds$$

and due to

$$\lim_{\delta_i \to 0} \frac{1}{\delta_i} \int_D^{D+\delta_i} 1 - F_{X_i}(s)\; ds = 1 - F_{X_i}(D) = P[L_{D_i,\delta_i}(X_i) > 0]$$

we have on account of $\delta_i \ll 1$, the approximation

$$P[L_{D_i,\delta_i}(X_i) > 0] \approx \frac{r_i}{\lambda} \tag{4.18}$$

where $\lambda := E[N]$. It should be noted that X_i represents the original loss (arising from an event) to the risks covered below the layer i . The individual X_i are naturally correlated and we would now like to provide a simple model of this correlation. We use Z to represent the aggregate loss for the portfolio arising from one event, ie

$$Z = \sum_i L_{D_i,\delta_i}(X_i) := \sum_i Y_i$$

We now make the following assumption: There exists a decreasing function $\xi : \mathbb{R}^+ \to \mathbb{R}^+$, so that :

$$Y_i = (Z - \xi(r_i))^+ - (Z - \xi(r_i) - \delta_i)^+ \tag{4.19}$$

applies.

This then means that a loss occurs in layer i at the very moment when the total loss for the portfolio exceeds $\xi(r_i)$. This assumption further implies that if a layer is affected by a loss event then this applies to every other layer with a higher rate on line. The function ξ can be identified using the approximation (4.18), since

$$1 - F_Z(\xi(r)) = P[Z > \xi(r)] = P[Y > 0] = \frac{r}{\lambda}$$

and thus

$$\xi(r) = F_Z^{-1}(1 - r/\lambda)$$

We are now in a position to calculate the covariance of the aggregate annual loss with the aggregate annual loss in the layer. For the sake of simplicity we will assume that the number of events is Poisson distributed. In this case the covariance we are seeking is equal to

$$\lambda \cdot E[Z \cdot Y_i]$$

We now have:

$$E[Z \cdot Y] = \int_{\xi}^{\xi+\delta} z(z-\xi)\ dF_Z(z)\ +\ \int_{\xi+\delta}^{\infty} z \cdot \delta\ dF_Z(z) \tag{4.20}$$

For the sake of simpler notation we have left out the index i and have written ξ instead of $\xi(r)$. Developing the above formula (4.20) with respect to δ now provides:

$$E[Z \cdot Y] = \delta \left(\xi(1 - F_Z(\xi)) + \int_{\xi}^{\infty} (1 - F_Z(z))\ dz \right) + O(\delta^2) \tag{4.21}$$

We can now write the formula for loading: If M is the given total margin for the portfolio then the margin

$$M \cdot \frac{\lambda\ E[Z \cdot Y]}{\lambda\ E[Z^2]} = M \cdot \frac{E[Z \cdot Y]}{E[Z^2]}$$

must be requested for the layer under consideration. For the loading on line for a layer with ROL r with (4.21) we obtain after disregarding the terms $O(\delta^2)$:

$$\begin{aligned} LOL(r) &= \frac{M}{E[Z^2]} \left(\xi(r)(1 - F_Z(\xi(r))) + slt_Z(\xi(r))\right) \\ &= \frac{M}{E[Z^2]} \left(F_Z^{-1}(1 - r/\lambda) \cdot r/\lambda + slt_Z(F_Z^{-1}(1 - r/\lambda))\right) \end{aligned} \tag{4.22}$$

We would now like to summarize the considerations just made in a theorem:

Theorem 7 *(Bernegger, 1994)*

Let C_1 xs $D_1, \ldots, C_l$ xs D_l represent the amount of all of a reinsurer's Cat-XL treaties relating to a given natural hazard potential. The number of the corresponding events is Poisson(λ) distributed. For a covariance between the loss in a layer i with $C_i \ll \sum_j C_j$ and the aggregate reinsured loss Z the following applies

$$Cov[S_i, Z] \approx \lambda \cdot C_i \cdot \left(F_Z^{-1}(1 - r_i/\lambda) \cdot r_i/\lambda + slt_Z(F_Z^{-1}(1 - r_i/\lambda))\right)$$

where $r_i = \frac{E[S_i]}{C_i}$ represents the associated ROL. For calculating a global target return on margins for the individual treaties, this provides the formula (4.22) for the loading on line.

4.3.3 The dependency of loading on the reinsurer's share

So far we have been neglecting an important factor in our considerations: a reinsurer seldom underwrites a treaty alone but instead generally only receives a share $a \leq 1$ of the treaty. This means that every reinsurer involved in the treaty has a proportional amount of responsibility for loss payments and receives a corresponding share of the total premium. The total premium is negotiated between the direct insurer and the reinsurer with the largest share (the so-called leading reinsurer).

If a_i then represents the share of the reinsurer in consideration as part of the treaty i and S_i signifies the treaty's reinsured loss, then the total loss for the reinsurer's portfolio is equal to

$$S = \sum_{\text{all treaties i}} a_i S_i$$

As a result of this, using the variance or covariance principle the premium does not depend on the share in a linear manner and therefore the treaty's total premium in particular depends on the share of the leading reinsurer. In the case of the variance principle we need to modify the loading formula (4.14) as follows:

$$M_i = \alpha \cdot \ a_i^2 \cdot Var[S_i] \tag{4.23}$$

here M_i is the loading which the leading reinsurer with share a_i needs for himself and the total premium for the treaty is thus equal to

$$\frac{M_i}{a_i} = \alpha \cdot \ a_i \cdot Var[S_i]$$

ie the cost of the treaty depends on the share of the leading reinsurer.

In practice the situation often arises where the leading reinsurer's rating is already available and the other reinsurers then have to decide which share of the treaty they wish to underwrite. The reinsurer's underwriters then have to solve the following task:

Given: P , the premium for 100% of the treaty

α, the loading factor for the variance loading

K, the reinsurer's costs

Sought: a, the optimal share

Furthermore, we assume that the loading principle is based on the variance principle. This provides the following condition

$$
\begin{aligned}
a \cdot E[S] + \alpha \cdot a^2 \cdot Var[S] + K &\leq a \cdot P \\
a^2(\alpha \cdot Var[S]) + a(E[S] - P) + K &\leq 0 \qquad (4.24)
\end{aligned}
$$

It should be noted that the costs do not depend on the underwritten share.

Discussion:

- If there is no real solution to the inequality (4.24), this means that participation in business is not acceptable for this reinsurer.
- If for there are two real solutions $a_1 \leq a_2$ which make the left hand side of (4.24) equal to zero:

  $$a_{1,2} = \frac{P - E[S] \pm \sqrt{(P - E[S])^2 - 4K \cdot \alpha \cdot Var[S]}}{2\alpha \cdot Var[S]}$$

  which are then automatically positive due to the above formula (except if $P < E[S]$, but then the reinsurer should keep well away from this treaty anyway), then every share $a_1 \leq a \leq a_2$ is acceptable for the reinsurer. The optimal share is a_2, since at that point the expected profit is greatest.
- The factor α and the costs K vary for different insurers and consequently so do the optimal shares.

Bibliography

[1] Aebischer, C.: *Industry Specific Convex High xs Distributions*, Preprint, 2006

[2] Bernegger, S. : *The Swiss Re Exposure Curves and the MBBEFD Distribution Class*, ASTIN Bulletin Vol 27, No.1, 1997, pp. 99-111

[3] G. Benktander/C.O. Segerdahl: *On the Analytical Representation of Claim Distributions with Special Reference to Excess of Loss Reinsurance*, Transactions of the 16th International Congress of Actuaries, Brussels 1960, Vol I, page 631

[4] Chevre, S. : *Modeling Dependent Risks,* Diploma Thesis in Mathematics, ETH Zurich, 1999

[5] Föllmer, H. : *Wahrscheinlichkeitstheorie I,* Lecture Notes ETHZ, 1988

[6] Gisler, A. : *Nicht-Leben Versicherungsmathematik*, Lecture Notes ETHZ, 1999

[7] Keller,P. (Federal Office of Private Insurance): *The Swiss Solvency Test*, Presentation for "Insurance Bureau of Canada" in Toronto, 28.03.2007