Università Cattolica Del Sacro Cuore
Milano
Interfacoltà di Scienze Bancarie, Finanziarie e Assicurative - Economia
Corso di Laurea in Statistical And Actuarial Sciences

Optimal reinsurance treaties: assessment of capital requirement and profitability for a multi-line insurer

Tesi di Laurea di:
Alberto ZANOTTO
Matr. n. 4708222

Relatore: Chiar.mo Prof. Gian Paolo CLEMENTE
Correlatore: Chiar.mo Prof. Diego ZAPPA

Anno Accademico 2018/2019

Contents

Introduction

In the actuarial literature, optimal reinsurance problems have been heavily studied, but almost all the solutions rely on mathematical optimizations that are difficult to implement in practice. When the optimization is treated with numerical examples, the latter are usually too simple in terms of assumptions and reinsurance programs' complexity.

The aim of this thesis is to overcome this impracticality through the implementation of a solid optimization procedure based on the *frequency-severity* model, which is one of the mostly known methodologies in the insurance sector. The combination of a wide range of Quota Share and Excess of Loss contracts is generated and tested on a fictitious multi-line insurer, which is calibrated on the Italian insurance market.

The aggregation of different lines of business has been achieved through the use of *Vine Copulas*, a flexible instrument that extends the bivariate Archimedean Copulas to the multivariate context.

Unconventional applications of the *Panjer* algorithm have been introduced: from a backtesting technique to measure simulations' accuracy to an unbiased reinsurance pricing tool.

After the optimal reinsurance program has been determined, an additional question is raised and answered: *can the presence of reinstatements in an Excess of Loss improve reinsurance in terms of risk return trade-off?*

Chapter 1

Reinsurance

The term *Reinsurance* can be easily described as the insurance for insurers: the reinsurer agrees to assume a portion of insurance risks in charge of the insurer, where the latter pays premium to compensate the reinsurer.

Reinsurance plays a major role in the insurance business since it diminishes the impact of claims losses and stabilizes the underwriting result of the direct insurer. Reinsurance is a fundamental tool that allows to obtain a more secure and reliable insurance system.

Many drivers are involved in the demand of a reinsurance policy:

- Risk transfer: an insurance company may decide to split its portfolio with a reinsurance company in order to decrease the risk exposure.

- Loss experience stabilization: since the insurance result uncertainty depends on the volatility of the claims frequency and severity, reinsurance comes handy to mitigate adverse fluctuations that could destabilize the insurer placement in the market.

- Insurer's capital relief and capacity increase: reinsurance reduces the capital requirements of the direct insurer. Consequently, the insurer can access new opportunities like the underwriting of new policies. From an economical perspective, this driver is considered as one of the main reason to build an effective reinsurance plan. Also, the achievement of lower capital requirements for the company leads to a higher Solvency Ratio, an fundamental indicator under Solvency framework.

- Catastrophe protection: this point gained relevance in recent years for the increasing number of both natural and man-made catastrophes. Without reinsurance, many insurers may risk being in ruin every time a relevant catastrophe occurs. The reinsurer plays as safety net towards the insurance market in case of such extreme events. Note that the reinsurance company is able to receive this kind of risks, characterized by a huge exposure, thanks to an adequate amount of own capital and a very diversified portfolio.

The features and characteristics of a reinsurance coverage are the result of a reciprocal trust and dialogue between the two parts.

Many conflicts of interest emerge:

On one hand, both the insurer and the reinsurer calibrate their decisions by taking into account own needs and objectives.

On the other hand, both parties give huge emphasis to reasonable loadings and commissions involved in the agreement: a weak compensation can damage the reinsurer for assuming risks recklessly, while a too burdensome compensation depletes the underwriting result and the own funds of the direct insurer.

For these reasons, facing the reinsurance topic from only one point of view may lead to unpractical or unused solutions.

Afterall, a reinsurance contract take place only when both parties benefit from it.

1.1 Reinsurance Basics

1.1.1 Facultative vs Treaty

When we talk about traditional reinsurance, we should distinguish between facultative and treaty reinsurance.

A treaty reinsurance shall include all the risks of a specific class of the ceding company's business, and the reinsurer accepts the block of business within the terms of the reinsurance contract.

Instead, a facultative reinsurance enables the direct insurer to choose on an individual basis which risk to include in the cover, and consequently, the reinsurer has the right to accept or reject those risks one by one. This kind of cover is usually suitable for risks that aren't covered by treaties already in force.

There are two other types, given by a mix of the previously described ones. In the facultative-obligatory treaty the ceding insurer can choose which risks include in the agreement and the reinsurer can only accept or reject the selection made. Vice versa, in an obligatory-facultative treaty the reinsurer is the only one that can decide from the whole portfolio which risks will be considered.

In this thesis the focus will be on treaties to emphasize reinsurance effect on the whole underwriting result. The analysis of simulated ad hoc facultative reinsurance agreements, where just few risks are covered, may lead to a lack of concrete results.

1.1.2 Proportional vs Non-Proportional

Both facultative and treaty reinsurance can be further divided in 2 groups:

- In proportional reinsurance premium and losses are shared proportionally between ceding company and reinsurer. The latter usually pays a ceding commission to the direct insurer to reimburse for expenses linked to the underwriting of the policy. This kind of contract protects against both adverse frequency and adverse severity of claims.

- In non-proportional reinsurance the claims amount is transferred to the reinsurer in relation to the excess of a specified retention. Typically, the price of the policy is expressed as a percentage of the direct insurer's premiums. The protection granted by a non-proportional contract allows to smooth, or even delete, the worst case of claims losses during the coverage period.

A well detailed description of differences between proportional and non-proportional policies will be provided later on, since further specifications can be addressed once the main coverages are explained.

1.2 Proportional Reinsurance

1.2.1 Quota Share

In a quota share contract premiums and losses are shared according to a fixed ratio α between the two parties. The letter α corresponds to the so-called retention coefficient that represents the percentage of premiums and losses retained by the direct insurer.

Quota share contracts are by definition treaties, since all kind of risks in the portfolio are shared without any regards to the corresponding sum insured. It is important to specify the notations used to clarify any doubts:

$\widetilde{X}$ Original aggregate losses random variable

P Pure premiums with $P = E(\widetilde{X})$

λP Safety loading expressed as a percentage λ of the pure premium P

C Commissions expressed as a percentage c of the tariff premiums B

B Original tariff premiums with $B = (1+\lambda)P + cB$

αB Direct insurer's tariff premiums

$\alpha\widetilde{X}$ Direct insurer's losses random variable

$(1-\alpha)B$ Reinsurer's tariff premiums

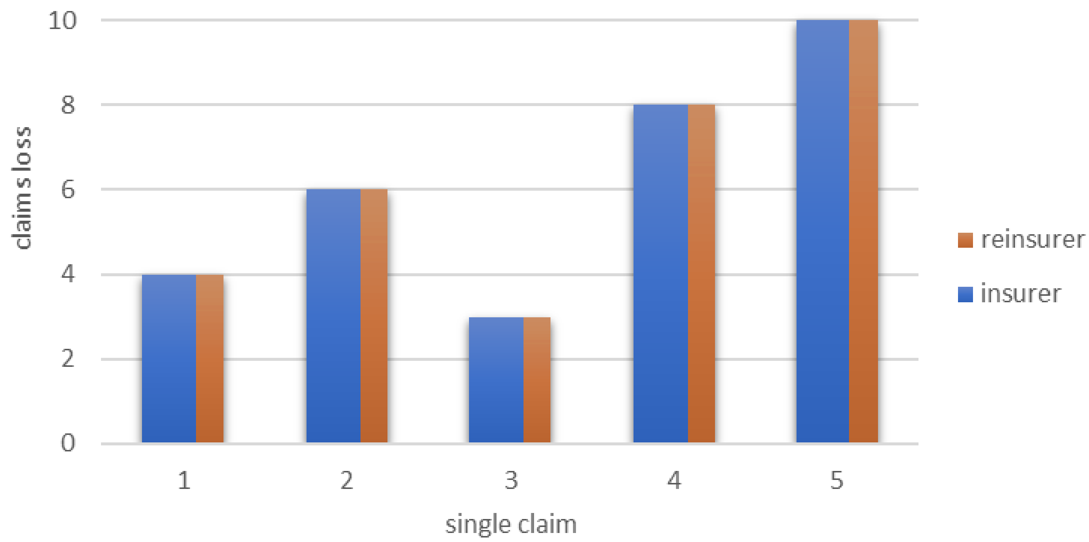
$(1-\alpha)\widetilde{X}$ Reinsurer's losses random variable

Thus, we can visualize the transfer of money between parties with the following scheme. Firstly, the premium payments:

$$Policyholders \xrightarrow{\text{pay B to}} DirectInsurer \xrightarrow{\text{pays } (1-\alpha)B \text{ to}} Reinsurer$$

Secondly, the claims payment in the opposite direction:

$$Policyholders \xleftarrow{\text{pays } \widetilde{X} \text{ to}} DirectInsurer \xleftarrow{\text{pays } (1-\alpha)\widetilde{X} + C^{re} \text{ to}} Reinsurer$$



As mentioned before, the reinsurer pays back to the direct insurer the ceding commission $C^{re} = c^{re}B$, with usually $c^{re} \leq c$.

In fact, in quota share the reinsurer can apply an additional safety loading to the direct insurer by acknowledging to him a commission rate c^{re} lower than c at the end of the treaty's coverage period.

The value assumed by c^{re} influences radically the nature of the contract since, in case of low values, the profitability from the direct insurer point of view is depleted.

It's important to mention that c^{re} can be either fixed or stochastic.

We will give a particular emphasis to this commission rate, since many quota share contracts feature nowadays the so-called sliding commissions, where the value depends on the loss ratio observed in the year.So, c^{re} is a random variable and, in this case, will be written as $\widetilde{c^{re}}$

In particular, $\widetilde{c^{re}}$ assumes high values in case of favourable loss ratio and vice versa.

Therefore, the reinsurer is more protected against two scenarios: the first, as said before, where we observe a large aggregate amount of losses compared to the earned premiums; the second were the direct insurer is transferring a portfolio of risks were an insufficient pricing is made.

It doesn't exist a unique way to build the sliding commissions, but in general they can be based upon:

- A step rule where the distribution function of the loss ratio is split into a fixed number of classes with each having a corresponding value of $\widetilde{c^{re}}$. So, given a realisation y of the loss ratio of the year, we will use the commission rate associated to the class where y falls within.

- A mathematical function that, given the loss ratio (or another reasonable index), returns the corresponding commissions. Therefore, the domain of $\widetilde{c^{re}}$ is not bounded to a discrete subset of values like the previous case, but it is a continuous interval. Usually, the mentioned interval has upper and lower bounds to avoid inappropriate commissions in case of extreme (favourable and unfavourable) events.

The adoption of stochastic commissions affects in a very interesting manner the variability of the insurance company's capital from year to year. Obviously, more protection for the reinsurer implies less protection for the direct insurer in case of unfavourable loss ratio.

Let assume two opposite scenarios under stochastic $\widetilde{c^{re}}$ framework:

1. During the year, small losses occurred, and the loss ratio assumes a very low value. Therefore, high amounts of commissions are paid back by the reinsurer to the insurer at the end of the year. The insurer has two good news at the same time: the underwriting result of the year has gone great, and the treaty cost is low.

2. Vice versa, huge losses in the year imply high loss ratio and low commissions paid back to the direct insurer. The latter has now two problems: the underwriting result has gone bad, and it is further worsened by high treaty's cost.

Thus, the presence of stochastic commissions amplifies the direct insurer's result of the year. From the cedant point of view, signing up this kind of coverage may be counter-intuitive when compared with a quota share with fixed c^{re}. Afterall, the insurer is relying on reinsurance to reduce risks, not to amplify them.

The reason why many quota share contracts with sliding commissions are signed up nowadays can be attributed to many drivers:

- As said before, a reinsurance contract is an agreement between two parties. So, the reinsurer may impose stochastic commissions as a safety measure against the possibility that the considered portfolio's riskiness has been underestimated by the ceding company.
- if both quota share with and without sliding commissions are offered by the same reinsurer,on average the second type is expected to be more expensive;
- Or we may reasonably think that both types are available on the market, but the quota share contract with low fixed commission is provided by a bad rated reinsurer. In this way, the reinsurance contract would be affected by a more probable credit risk.

There is a wide list of decisions that lead a company to prefer a contract over another one. We shall always remember that many trade-offs and market relations between parties are involved in a reinsurance agreement.

1.2.2 Surplus

In opposition to quota share, *surplus* contracts are by definition facultative, because the *retention line M*, fixed ex-ante, excludes all those risks characterized by a smaller *sum insured V*. Thus, the direct insurer obtains a more homogeneous portfolio of risks by retaining losses according a different retention coefficient α for each risk, such that:

$$\alpha = \begin{cases} 1 & \text{if } V < M \\ \frac{M}{V} & \text{if } V > M \end{cases} \tag{1.1}$$

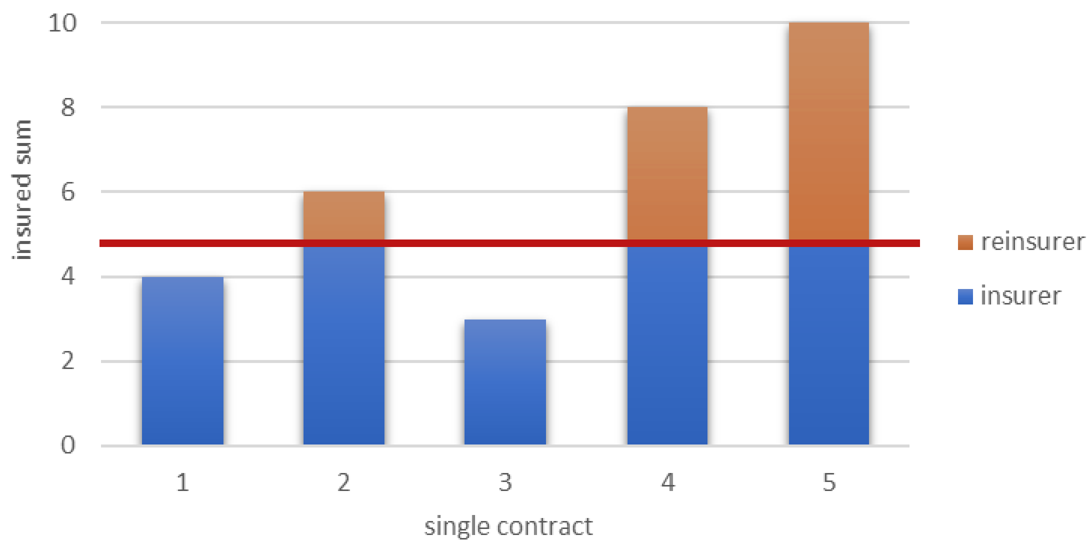
Let us denote the random variable *single claim cost* with $\widetilde{Z}$, where $0 < \widetilde{Z} \leq V$.

Under a surplus cover, $\alpha\widetilde{Z}$ is paid by the ceding company and $(1-\alpha)\widetilde{Z}$.

In case $V < M$, the whole single claim cost $\widetilde{Z}$ is in charge of the direct insurer, and, as mentioned before, the latter faces on his own only those risks characterized by a small sum insured.

Trivially, if case of loss $Z = V$, with $V > M$:

- The direct insurer pays $\alpha V = \frac{M}{V} V = M$
- The reinsurer pays $(1 - \alpha)V = (1 - \frac{M}{V})V = V - M$.



As mentioned by Antal [2], the reinsurance commissions are paid in the same manner as we discussed in the quota share contracts.

Regarding the presence of sliding commissions, we shall consider the reinsurer's loss ratio to calibrate the realizations of $\widetilde{c}^{re}$.

Note that, when we deal with quota share contracts, both the parties observe the same loss ratio because the treaty includes all risks in the considered business.

But with surplus, since we are dealing with a facultative reinsurance agreement, the direct insurer's loss ratio of the year differs from the reinsurer's one.

Surplus coverage is usually used in the property line of business, where the sum insured value is directly linked to the value of the asset insured. It may be hard to define the concept of sum insured in lines like General Third Party Liability (GTPL) or Motor Third Party Liability (MTPL), where physical damages to others may be involved.

1.3 Non-Proportional Reinsurance

1.3.1 Excess of Loss

In an excess of loss treaty *XL*, the ceding company transfers to the reinsurer a portion of each loss $\widetilde{Z}$ in excess of the *deductible* D. In practice a *limit* L is imposed on the reinsurer's payment for each claim. We are in front of the so-called L *xs* D (in words: L *in excess of* D).

Let us represent the mathematical notation:

The $i-th$ claim cost $\widetilde{Z}_i$ is split in two components:

$$\widetilde{Z}_i^{re} = (\widetilde{Z}_i - D)^+ - (\widetilde{Z}_i - (D+L))^+ = min(max(\widetilde{Z}_i - D, 0), L) \tag{1.2}$$

$$\widetilde{Z}_i^{in} = \widetilde{Z}_i - \widetilde{Z}_i^{re} \tag{1.3}$$

Where Z_i^{re} is the portion of claim paid by the reinsurer, and Z_i^{in} is the part paid by the direct insurer. As we can see, the reinsurer is in charge of the layer that goes from D to $D+L$. To be complete in the notation, the *Layer* operator is introduced:

$$Layer_{D,L}(\widetilde{Z}_i) \equiv \widetilde{Z}_i^{re} = min(max(\widetilde{Z}_i - D, 0), L) \tag{1.4}$$

The number of losses $\widetilde{N}$ that will occur during a year is a random variable, that we will suppose distributed like a $Poisson(E(\widetilde{N}))$. In this way, the *aggregate claim cost* $\widetilde{X}$ is defined as a combination of two random variables:

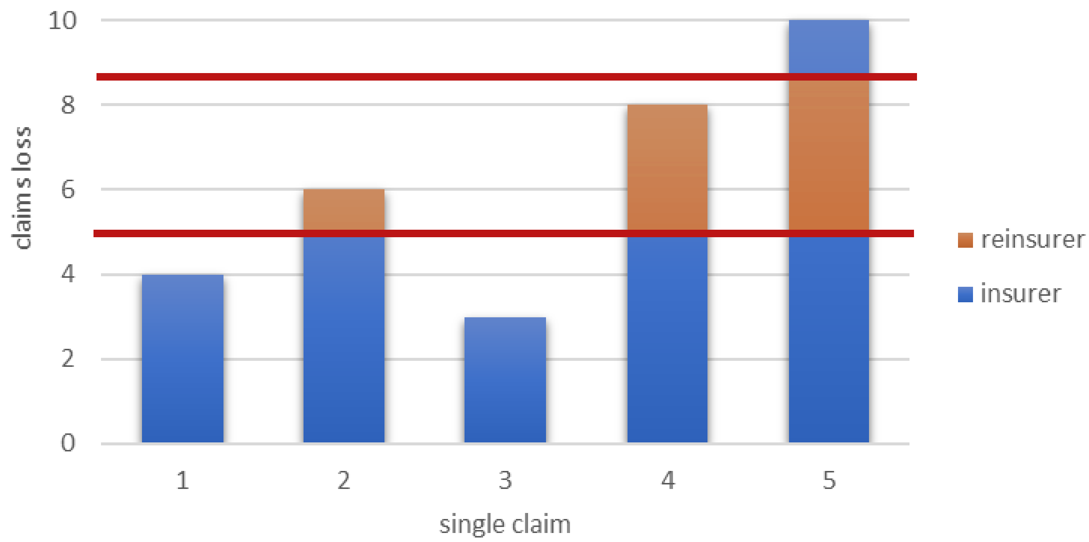
$$\widetilde{X} = \sum_{i=1}^{\widetilde{N}} \widetilde{Z}_i \tag{1.5}$$

Then, under an excess of loss treaty, $\widetilde{X}$ is split between the two parties as follows:

$$\widetilde{X}_{XL}^{re} = \sum_{i=1}^{N} \widetilde{Z}_i^{re} = \sum_{i=1}^{N} Layer_{D,L}(\widetilde{Z}_i) \tag{1.6}$$

$$\widetilde{X}^{in} = \sum_{i=1}^{N} \widetilde{Z}_i^{in} = \sum_{i=1}^{N} Layer_{0,D}(\widetilde{Z}_i) + Layer_{D+L,\infty}(\widetilde{Z}_i) \tag{1.7}$$

It's important to remind that a XL coverage protects the ceding company



only from individual claims which cost falls in the layer. In fact, in case a large number of small claims occurs, the direct insurer suffers huge losses anyways.

To be precise, the reinsurance agreement previously described is called *Working Excess of Loss* (WXL), since the layer is applied individually to each observed loss.

Instead, when we talk about the *Catastrophe Excess of Loss* (Cat XL), the deductible D and limit L are applied per event (e.g. earthquake, flood, hail). In fact, the occurrence of a catastrophic event affects a wide number of contracts at the same time.

Let us now compute the reinsurance pure premium P^{re} through the use of the so-called *stop loss transformation*.

First, we consider the expected value $E(\widetilde{Z}^{re})$. It is necessary to know the cumulative distribution function $F_{\widetilde{Z}}(z) = P(\widetilde{Z} \leq z)$ of the single claim $\widetilde{Z}$.

For the generic random variable $\widetilde{S}$, the stop-loss transformation is defined as:

$$E(\widetilde{S}) = \int_0^{\infty} 1 - F_{\widetilde{S}}(s) ds \tag{1.8}$$

In case of deductible D and infinite Limit L we have:

$$E(\widetilde{Z}^{re}) = E[(\widetilde{Z} - D)_+] = \int_D^\infty 1 - F_{\widetilde{Z}}(z)dz \tag{1.9}$$

Given the previous formula, we can compute trivially the result when L is finite:

$$\begin{aligned} E(\widetilde{Z}^{re}) = E[(\widetilde{Z} - D)_+] - E[(\widetilde{Z} - (D+L))_+] = \\ \int_D^\infty 1 - F_{\widetilde{Z}}(z)dz - \int_{D+L}^\infty 1 - F_{\widetilde{Z}}(z)dz = \int_D^{D+L} 1 - F_{\widetilde{Z}}(z)dz \end{aligned} \tag{1.10}$$

Given the expected value $E(\widetilde{N})$ of the claims frequency, the reinsurance pure premium P^{re} is obtained by:

$$P^{re} = E(\widetilde{X}^{re}) = E(\widetilde{N})E(\widetilde{Z}^{re}) \tag{1.11}$$

1.3.2 Stop Loss

Stop Loss is a non-proportional reinsurance contract where the reinsurer pays the portion $\widetilde{X}^{re}$ of the aggregate claim losses $\widetilde{X}$ that exceed the *aggregate deductible AD*, but only up to the *aggregate limit AL*. Trivially, a Stop Loss is an Excess of Loss treaty applied to aggregate annual losses.

Recalling (1.5), under a Stop Loss contract, we can split $\widetilde{X}$ into two components:

$$\widetilde{X}^{re}_{SL} = Layer_{AD,AL}(\widetilde{X}) = Layer_{AD,AL}(\sum_{i=1}^{\widetilde{N}} \widetilde{Z}_i) \tag{1.12}$$

$$\widetilde{X}^{in} = \widetilde{X} - \widetilde{X}^{re} \tag{1.13}$$

In this scenario, the direct insurer is exposed in a different manner with respect to the XL: Stop Loss treaty does not give any relevance to the behaviour of the individual claims loss and protects the ceding company against the overall unfavourable result on aggregate basis.

Obviously, every insurance company would opt for such a strong protection, if it wasn't for the related price. In fact, the safety loading imposed by the reinsurer is noteworthy, in order to protect himself from the high volatility of the random variable $\widetilde{X^{re}}$.

Note that a precise estimation of the so-called tail behaviour is one of the hardest actuarial tasks, and it has been studied in numerous researches.

Let us compute the reinsurance pure premium for a stop loss contract. We need to know $F_{\widetilde{X}}(x)$ to proceed.

Recalling the stop loss transformation of Equation (1.9), we have:

$$P^{re} = E(\widetilde{X}^{re}) = \int_{AD}^{AL} 1 - F_{\widetilde{X}}(x)dx \tag{1.14}$$

1.3.3 Excess of Loss with Aggregate Deductible and Aggregate Limit

We now consider a mixture of an Excess of Loss with layer *(L xs D)* and of a Stop Loss with aggregate layer *(AL xs AD)*. We can indicate the reinsurer's aggregate claim loss of this form of contract as:

$$\begin{aligned}\widetilde{X}^{re} = Layer_{AD,AL}(\widetilde{X}^{re}_{XL}) = Layer_{AD,AL}(\sum_{i=1}^{\widetilde{N}} Layer_{D,L}(\widetilde{Z}_i)) = \\ min(max(\sum_{i=1}^{\widetilde{N}}(Layer_{D,L}\widetilde{Z}_i))) - AD, 0), AL) = \\ min(max(\sum_{i=1}^{\widetilde{N}}(min((max(\widetilde{Z}_i - D, 0), L)) - AD, 0), AL)\end{aligned} \tag{1.15}$$

The resulting formula can be easily analysed step by step. Let us imagine that N claims occurred:

- The single claim loss Z_i is transferred to the reinsurer according to the Layer L xs D;
- Repeat step 1 for all the N occurred claims to obtain N values of Z_i^{re}, with $i = 1, 2, \ldots, N$;
- Obtain a provisional aggregate reinsurer loss by summing up all the Z_i^{re};
- Apply the Aggregate Layer AL xs AD to the provisional aggregate loss.

1.3.4 Excess of Loss with Reinstatements

Usually, the Aggregate Limit AL is set equal to a multiple of the individual claim Limit L. If $AL = (K+1)L$, we are in front of an Excess of Loss L xs D with K *reinstatements*.

At the start of the contract, the direct insurer pays the initial premium P^L_{AD} for the *Original Aggregate Layer L xs AD*. Afterwards, if a claim loss Z_i leads the amount $\sum Z^{re}_i$ to fall within the Aggregate Layer *(K+1)L xs AD*, the direct insurer pays an additional premium to restore the used up part of the layer.

The additional premium is called *reinstatement premium*, and it is proportional to both P^L_{AD} and to the used portion of the layer. The capacity of each reinstatement is equal to L, and the total premium of the $k-th$ reinstatement is expressed as a percentage of P^L_{AD}. We call it $P^L_{AD+kL} = c_k P^L_{AD}$, with $c_k \geq 0$. The $k-th$ reinstatement covers the Aggregate Layer $L\ xs\ (AD + (k-1)L)$.

If $c_k = 0$, we are in front of the so-called *free* reinstatements, where the initial premium P^L_{AD} already includes the price of the whole Aggregate Layer *AL xs AD*. Instead, with $c_k \geq 0$, we have the so-called *paid* reinstatements.

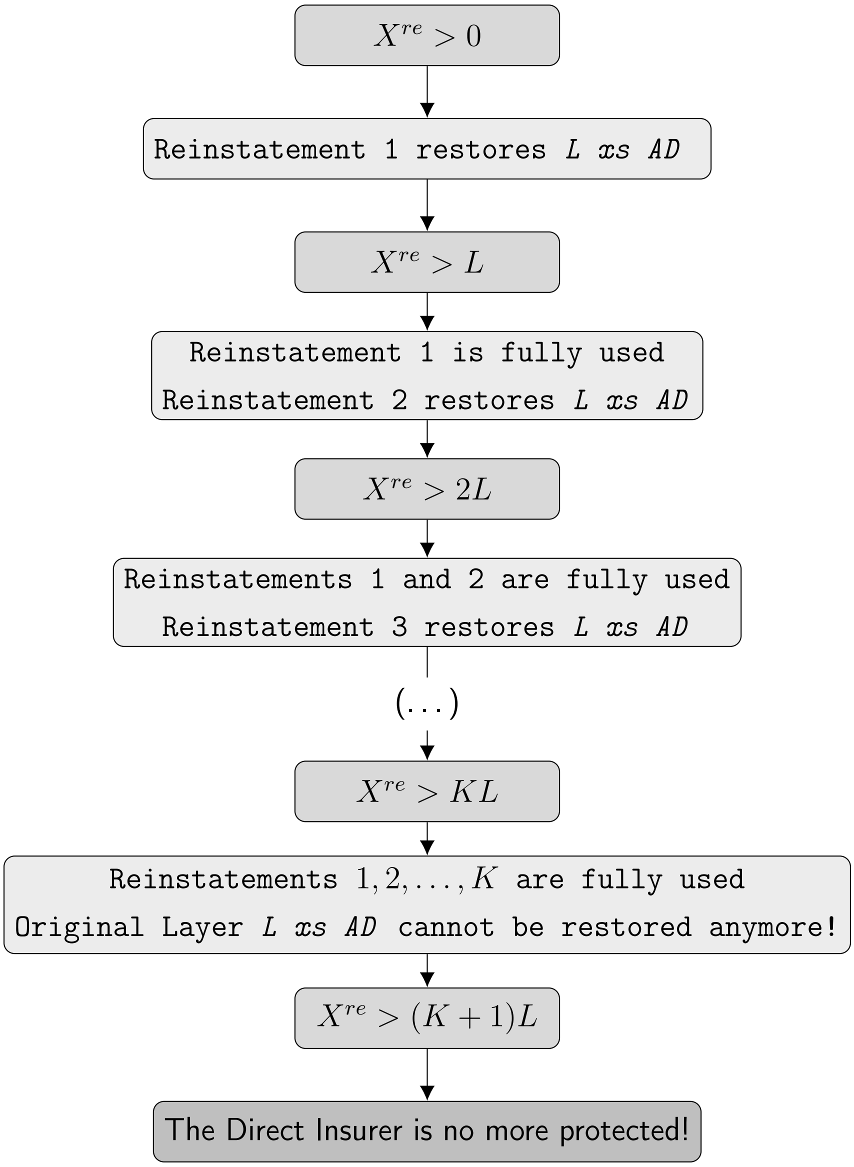
Usually in practice, $c_k = c\ \forall k = 1, 2, \ldots, K$ and assumes the value $0,5$ or 1. As c tends to 1, the initial premium P^L_{AD} and the stochastic premiums balance, and P^L_{AD} decreases.

Obviously, in case $c_k = 1$, we have:

$$P^L_{AD} = P^L_{AD+L} = P^L_{AD+2L} = \cdots = P^L_{AD+KL}. \tag{1.16}$$

With free reinstatements, the reinsurance premium is deterministic, while, with paid reinstatements, we deal with a random variable since the premium is paid pro rata. Trivially, the reinsurance loss and the stochastic premiums are highly correlated random variables.

To clarify how the reinstatements work, a graphical example is provided, in which a Excess of Loss with K reinstatements is bought and the loss in the year is big enough to consume the whole capacity of the treaty:



As mentioned before, the k-th reinstatement premium is proportional to the amount of layer to be reinstated, and it is given by:

$$c_k P_{AD}^{L} \frac{Layer_{AD+(k-1)L,L}(\widetilde{X}_{XL}^{re})}{L} \tag{1.17}$$

To simplify the notation, we will denote the r.v. amount covered by the $k-th$ reinstatement with:

$$\widetilde{r}_{k-1} = Layer_{AD+(k-1)L,L}(\widetilde{X}_{XL}^{re}) \tag{1.18}$$

The term $\frac{\widetilde{r}_{k-1}}{L}$ represents the proportion of layer reinstated. The only random variable involved in the formula (1.17) is $\widetilde{X}^{re}_{XL}$. So, it is fundamental to deeply understand the behaviour of the aggregate claims amount to price such a contract.

Since the total random premium is defined as the sum of the deterministic initial premium P^L_{AD} and of the stochastic reinstatements premiums $\sum_{k=1}^{K} c_k P^L_{AD} \frac{\widetilde{r}_{k-1}}{L}$, we can express it as:

$$\widetilde{P}^{AL}_{AD} = P^L_{AD}(1 + \frac{1}{L}\sum_{k=1}^{K} c_k \widetilde{r}_{k-1}) \tag{1.19}$$

By applying the Expected Value operator, we have:

$$E(\widetilde{P}^{AL}_{AD}) = P^L_{AD}(1 + \frac{1}{L}\sum_{k=1}^{K} c_k E(\widetilde{r}_{k-1})) \tag{1.20}$$

Now, our focus is to determine P^L_{AD}, since it is the only component we cannot estimate directly. In fact:

- Recalling equation (1.15), note that the term $E(\widetilde{P}^{AL}_{AD})$ is equivalent to the pure premium paid for an excess of loss $LxsD$ with aggregate layer $ALxsAD$. Therefore:

  $$E(\widetilde{P}^{AL}_{AD}) = E(Layer_{AD,AL}(\widetilde{X}^{re}_{XL})) \tag{1.21}$$

  Assume now that we are able to determine the cumulative distribution function $F_{\widetilde{X}^{re}_{XL}}(x)$, with $\widetilde{X}^{re}_{XL} = \sum_{i=1}^{\widetilde{N}} Layer_{D,L}(\widetilde{Z}_i)$. This point is achievable through the use of methodologies like *Panjer Algorithm*.

  Hence, using the stop loss transformation of equation (1.9), we can compute the following term:

  $$E(\widetilde{P}^{AL}_{AD}) = \int_{AD}^{AL} 1 - F_{\widetilde{X}^{re}_{XL}}(x)dx \tag{1.22}$$

- $E(\widetilde{r}_{k-1})$ computation is based on the same concepts of $E(\widetilde{P}^{AL}_{AD})$. In fact, recalling the previous formula, we have:

$$E(\widetilde{r}_{k-1}) = E(Layer_{AD+(k-1)L,L}(\widetilde{X}^{re}_{XL})) = \int_{AD+(k-1)L}^{L} 1 - F_{\widetilde{X}^{re}_{XL}}(x)dx \tag{1.23}$$

Therefore, it is possible to compute the initial premium P^L_{AD} by rewriting equation (1.20) as:

$$P^L_{AD} = \frac{E(\widetilde{P}^{AL}_{AD})}{(1 + \frac{1}{L}\sum_{k=1}^{K} c_k E(\widetilde{r}_{k-1}))} \tag{1.24}$$

In case $c_k = c \ \forall k = 1, \ldots, K$, the previous equation can be easily simplified:

$$P^L_{AD} = \frac{E(\widetilde{P}^{AL}_{AD})}{(1 + \frac{c}{L}\sum_{k=1}^{K} E(\widetilde{r}_{k-1}))} = \frac{E(\widetilde{P}^{(K+1)L}_{AD})}{(1 + \frac{c}{L}E(\widetilde{P}^{KL}_{AD}))} \tag{1.25}$$

Trivially, in case of *free reinstatements*, the deterministic premium P^{AL}_{AD} is equal to the numerator of the previous formula.

To better grasp how the reinstatements work in practice, a detailed numerical example is provided.

Suppose we are dealing with a XL coverage with layer *15 xs 10* with $AD = 20$ and 2 reinstatements. Hence, $AL = (1 + 2)15 = 45$.

Imagine we observed 8 individual claim costs Z_i, represented in the following table:

1. The 1^{st} claim Z_1 is 20, and through the application of the Layer *15 xs 10*, we have the corresponding claim amount $Z^{re}_1 = 10$. Thus, the reinsurer's aggregate claim amount X^{re} is still 0, since 10 is lower than the $AD = 20$.

2. $Z_2 = 5$, that is lower than $D = 10$. So, this claim loss is in charge of the direct insurer ($Z^{re}_2 = 0$,$Z^{in}_2 = 5$).

3. $Z_3 = 40$, split among reinsurer and insurer with $Z^{re}_3 = 15$ and $Z^{in}_3 = 25$. Now, $\sum_{i=1}^{3} Z^{re}_i = 25 > 20 = AD$. Therefore, the reinsurer pays $X^{re} = 5$ and the direct insurer pays the additional premium to reinstate the layer, reducing the 1^{st} reinstatement capacity by 5.

4. $Z_4 = 25$ with $Z^{re}_4 = 15$ and $Z^{in}_4 = 10$. Note that, in comparison with the previous case, the reinsurer is always in charge of 15 for the presence of

Table 1.1: Reinstatement numerical example

Claim number i	1	2	3	4	5	6	7	8
Z_i	20	5	40	25	15	25	35	20
$X = \sum Z_i$	20	25	65	90	105	130	165	185
$Z_i^{re} = Layer_{10,15}(Z_i)$	10	0	15	15	5	15	15	10
$\sum Z_i^{re}$	10	10	25	40	45	60	75	85
$X^{re} = Layer_{20,45}(\sum Z_i^{re})$	0	0	5	20	25	40	**45**	45
$X^{in} = X - X^{re}$	20	25	60	70	80	90	120	140
Original layer	15	15	15	15	15	5	**0**	0
Reinstatement 1	15	15	10	**0**	0	0	0	0
Reinstatement 2	15	15	15	10	5	**0**	0	0

$L = 15$. The reinsurer pays 15 and the insurer reinstate the layer using both the 1^{st} and the 2^{nd} reinstatement. So, the 1^{st} capacity is completely used to pay 10, and the second one is reduced by the remaining 5.

5. $Z_5 = 15$, with $Z_5^{re} = 5$ and $Z_5^{in} = 10$. So, the layer is reinstated by reducing by 5 the 2^{nd} reinstatement.

6. $Z_6 = 25$, with $Z_6^{re} = 15$ and $Z_6^{in} = 10$. Therefore, also the 2^{nd} reinstatement is depleted, and the original layer is reduced by the remaining 10. From now on, the layer won't be reinstated anymore.

7. $Z_7 = 35$, with $Z_7^{re} = 15$ and $Z_7^{re} = 20$. But, since the original layer capacity is only 5, the reinsurer will pay 5, and the insurer the remaining 30. In this moment, $\sum_{i=1}^{7} Z_i^{re} = 75 > 65 = AD + AL$, so the reinsurer does not have any further obligation towards the insurer.

8. $Z_8 = 20$, with $Z_7^{re} = 10$ and $Z_7^{re} = 10$. But, as said before, the whole claim is in charge of the direct insurer now.

In conclusion, the reinsurer pays 45, the maximum possible according to the agreement. The direct insurer pays 140 for the claims (instead of 185 without the reinsurance) and an additional amount for the reinsurance premiums related to the original layer and the two reinstatements, expressed as

$$P_{20}^{15} + P_{35}^{15} + P_{50}^{15} = P_{20}^{15}(1 + c_1 + c_2). \qquad (1.26)$$

It's important to mention that the analysed numerical example shows an extreme case: It's unusual to observe the reinsurer in charge of the whole Aggregate Limit AL.

In fact, the Non-Proportional treaties protect the ceding company from extreme unfavourable losses. Typically, the reinsurer's profit is generated in those years where there aren't any obligations towards the ceding company, in particular when $\sum Z_i^{re} < AD$.

This typology of contract adds a significant amount of complexity into the reinsurance framework. But it's important to remark that the presence of aggregate deductibles and limits is common in Excess of Loss reinsurance context. For this reason, this topic will be deeply analysed and commented later in the thesis with adequate case studies.

Chapter 2

The Frequency-Severity Approach

To understand clearly reinsurance and its effects in terms of risk mitigation, an adequate model is required. The most known in literature is the Frequency-Severity approach. Recalling equation (1.5), the scope is to estimate the aggregate claim amount $\widetilde{X}$ of the direct insurer, which is defined as:

$$\widetilde{X} = \sum_{i=1}^{\widetilde{N}} \widetilde{Z}_i$$

In the standard frequency-severity approach, the actuary estimates distributions for the number of claims $\widetilde{N}$ (frequency) and the single claim cost $\widetilde{Z}$ (severity), because the main goal is to define the distribution of $\widetilde{X}$, both gross and net of reinsurance. The question may sound obvious: why not estimate $\widetilde{X}$ directly? The first issue is connected to the fact that excess of loss reinsurance works on $\widetilde{Z}$, and therefore the estimation of $\widetilde{Z}$ distribution shouldn't be avoided. The second and biggest problem is related to the estimation accuracy. Let us take few steps back.

To estimate properly a distribution, a large amount of observations is needed to achieve an accurate result. Since the random variable $\widetilde{X}$ is the *annual* aggregate amount, the actuary would have too little data to perform a proper fit. But, if we consider how many single claims are observed annually, it is a whole different situation: $\widetilde{Z}$ would be way easier to estimate in comparison.

But it is not ended: how could we estimate the frequency if it is annual too? In literature, the main distribution associated to $\widetilde{N}$ is the Poisson, which comes very handy for this purpose due to one of its properties: the sum of independent Poisson-distributed random variables is Poisson-distributed. In particular, if $\widetilde{N}_i \sim Poisson(\lambda_i)$ for $i = 1, \ldots, n$ are independent, then $\widetilde{N} = \sum_{i=1}^{n} \widetilde{N}_i \sim Poisson(\sum_{i=1}^{n} \lambda_i)$.

This property is more important than it might appear at the first glance: instead of studying $\widetilde{N}$, it is now possible to focus on the frequency component $\widetilde{N}_i$ of the single policyholder in the portfolio. In this way, we are even able to study more precisely the frequency component, and associate the proper volatility to this random variable. Now, more than ever, Generalized Linear Models (GLM) are used to perform frequency and severity fitting per contract, based on the characteristics of the single policyholder.

Usually, in the actuarial literature $\widetilde{Z}$ and $\widetilde{N}$ are assumed independent. This assumption makes simpler both the theory and the simulation procedure to obtain $\widetilde{X}$. Nonetheless, in case more lines of business (e.g. *A,B,C*) are modelled, an estimation of the underlying dependency structure between $\widetilde{X}_A, \widetilde{X}_B, \widetilde{X}_C$ is needed. In this case, each single $\widetilde{X_i}$ is computed separately, and in a second step they are aggregated considering the dependency. The **copula** aggregation is a very popular method to achieve this operation. With the introduction of Solvency II, copulas are a hot topic in insurance for both internal model and ORSA (own risk and solvency assessment). The most known copula is the Gaussian Copula, due to being the simplest and the most flexible one in a high-dimensional case.

In the next sections many distributions and dependencies are explained to give a good overview over all the main topics and tools underlying the frequency-severity approach used in the thesis.

2.1 Distributions

In this section the only distributions that will be presented are the ones needed to understand the reasoning in this thesis. Further discussions on other distributions are avoided in order to maintain a fluid argumentation.

Normal Distribution Known to be the most famous distribution in economics and finance, it is one of the most inappropriate choice to model insurance risks. Since the normal distribution has a huge relevance in statistics, it is often used as a benchmark to make comparison with other distributions. For this reason, it is worthy to give a quick summary on its properties.

The density function of a random variable $\widetilde{Y} \sim Normal(\mu, \sigma)$ is equal to

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

with

$$y \in (-\infty, \infty), \qquad \mu = E(\widetilde{Y}), \qquad \sigma = \sigma(\widetilde{Y})$$

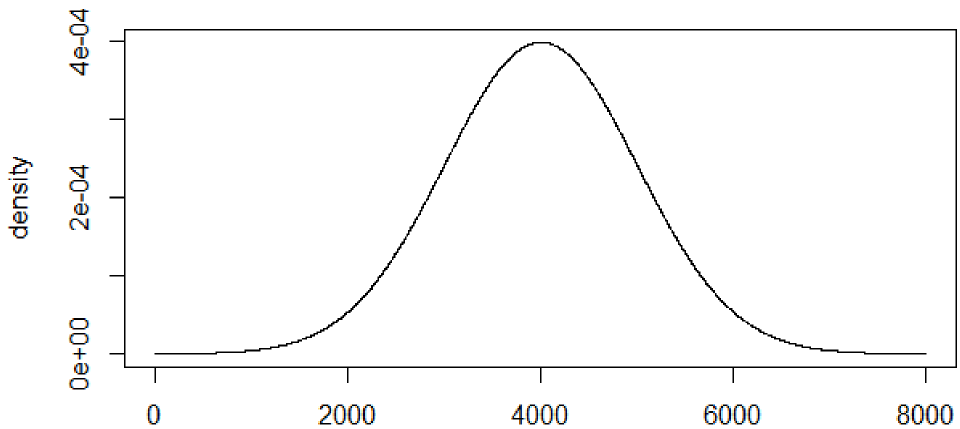


Figure 2.1: Normal(4000,1000) density function

The normal distribution is characterized by a perfect symmetry, and for the severity fitting, it is a serious problem. Let's investigate the main issues:

- the mathematical domain of $\widetilde{Z}$ is $\mathbb{R}^+$, and by using a Normal we would observe claims with also negative values, which don't make sense;
- in Non-Life insurance, the skewness index $\gamma(\widetilde{Z})$ is positive, while, a Normal, instead, is characterized by a perfect symmetry;
- insurance risks exhibit extreme tails, which are not present in a Normal.

The Normal is often mentioned for the possibility to determine its quantiles for every chosen level of confidence $(1 - \alpha)$ through the use of the standard Normal tables. We need first to define the standardized Normal $\widetilde{S}$ as

$$\widetilde{S} = \frac{\widetilde{Y} - \mu}{\sigma}$$

Then, it is possible to associate for the same level of confidence $(1-\alpha)$ the quantile $y_{1-\alpha}$ of $\widetilde{Y}$ and the quantile $s_{1-\alpha}$ of $\widetilde{S}$:

$$s_{1-\alpha} = \frac{y_{1-\alpha} - \mu}{\sigma}$$

Since $s_{1-\alpha}$ can be found on standard Normal tables, we isolate $y_{1-\alpha}$:

$$y_{1-\alpha} = s_{1-\alpha}\, \sigma + \mu$$

The most important level of confidence in insurance is 99.5% due to the fact that Solvency II evaluates at that level the capital requirement using the Value at Risk. The corresponding standard Normal quantile is $s_{0.995} = 2.58$.

This multiplier assumes a huge importance when comparing distributions: if we are dealing with a positive skewed distribution, we would expect to have a value greater than 2.58 in order to reach the quantile with level of confidence 99.5%.

LogNormal Distribution The LogNormal distribution is often used in actuarial context for both single claim cost $\widetilde{Z}$ and aggregate claim cost $\widetilde{X}$. It is obtainable through a transformation of the Normal distribution, such that:

if $\widetilde{Y} \sim Normal(\mu, \sigma)$ and $\widetilde{W} \sim LogNormal(\mu, \sigma)$, then $\widetilde{W} = exp(\widetilde{Y})$.

The LogNormal density function is defined as

$$f_{\widetilde{W}}(w) = \frac{1}{\sigma w \sqrt{2\pi}} e^{-\frac{1}{2}\frac{(ln(w)-\mu)^2}{\sigma^2}}$$

It is crucial important to know that, unlike the Normal, the parameters μ and σ are not respectively the mean and the standard deviation of the LogNormal. In fact, they are obtained through the following formula:

$$\sigma = \sqrt{ln(1 + CV_{\widetilde{W}}^2)}$$

$$\mu = ln(E(\widetilde{W})) - \frac{\sigma^2}{2}$$

where the so-called *coefficient of variation* $CV_{\widetilde{W}} = \frac{\sigma(\widetilde{W})}{E(\widetilde{W})}$.

This distribution is characterized by a positive skewness and a long right tail, which come handy to model the severity component $\widetilde{Z}$.

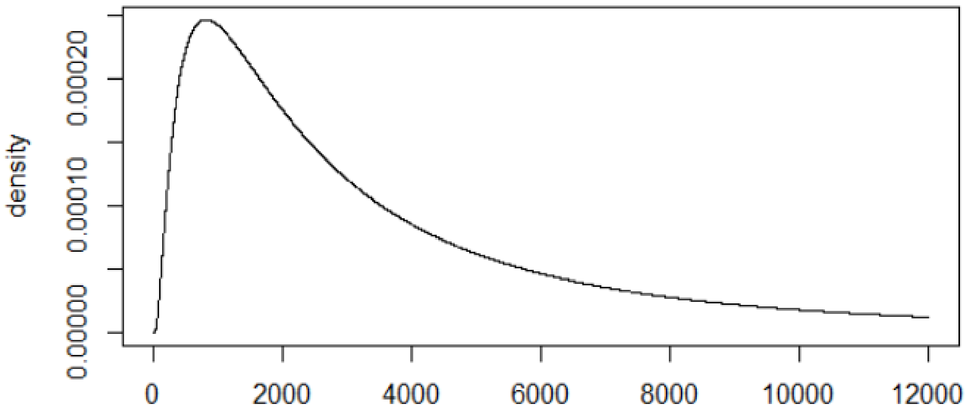


Figure 2.2: $LogNormal(8, 1.1)$ density function

Poisson Distribution As mentioned before, the Poisson distribution is used in the actuarial context to model the number of claims $\widetilde{N}$. It is a discrete probability distribution, where the random variable $\widetilde{N} \in \mathbb{N} = \{0, 1, 2, \dots\}$.

The probability mass function of the generic random variable $\widetilde{N} \sim Poisson(n)$ is defined as:

$$P(\widetilde{N} = k) = e^{n} \frac{n^k}{k!}$$

with $k \in \mathbb{N}$.

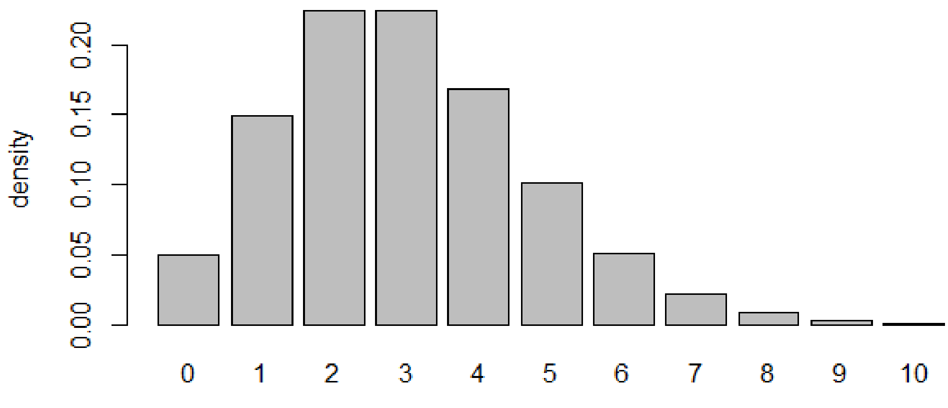


Figure 2.3: $Poisson(3)$ density function

Note that the Poisson distribution is characterized by only one parameter n, and therefore, all of its moments depend on it. The main results are listed below:

$$E(\widetilde{N}) = \sigma^2(\widetilde{N}) = n$$

$$\gamma(\widetilde{N}) = \frac{1}{\sqrt{n}}$$

It is interesting to see that, for high values of the parameter n, the distribution is almost symmetric. The fact that the mean and the variance are equal is a huge limitation, since in real insurance scenarios this assumption does not hold. In fact, its coefficient of variation is equal to:

$$CV(\widetilde{N}) = \frac{1}{\sqrt{n}}$$

which tends to zero too as the portfolio increases in size. One could think that this type of relation is understandable, since it is true that, through the contracts' diversification, the insurer is able to reduce the risks in relative terms. But, it is also true that a perfect diversification is not realistic.

To overcome this issue, one possible solution is the use of a Negative Binomial, which can be expressed as a $Poisson(n\widetilde{Q})$, where $\widetilde{Q}$ is distributed as a $Gamma(h, h)$. This feature is described more in detail in the next paragraphs.

Gamma Distribution The Gamma distribution is a continuous and positive-only distribution with two parameters (α, β). α is the shape parameter and β the rate parameter. A random variable $\widetilde{Q} \sim Gamma(\alpha, \beta)$ has its density function defined as:

$$f_{\widetilde{Q}}(q) = \frac{\beta^\alpha q^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta q}$$

with

$$q, \alpha, \beta > 0 \qquad \Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

In statistics the Gamma distribution is used to describe the time required for α events to occur in a $Poisson(\beta)$ process. In the actuarial context it can be used to model the single claim random variable $\widetilde{Z}$ for light positive skewed data, like non-catastrophe claims.

The main results are:

$$E(\widetilde{Q}) = \frac{\alpha}{\beta}$$

$$\sigma^2(\widetilde{Q}) = \frac{\alpha}{\beta^2}$$

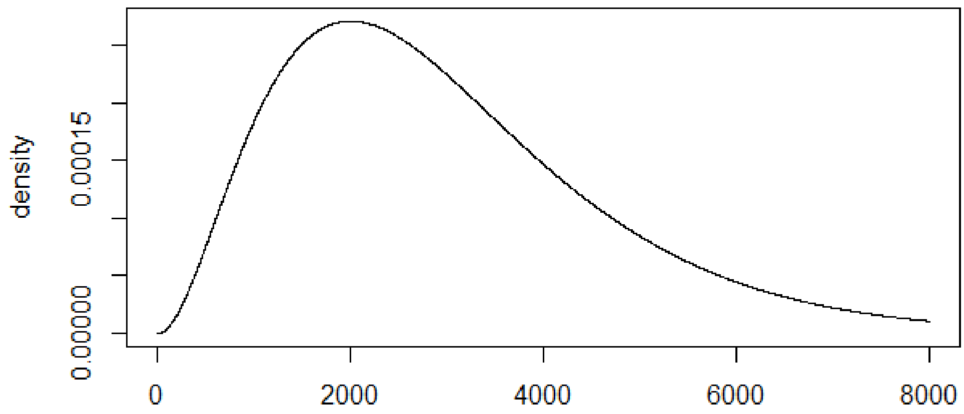


Figure 2.4: $Gamma(3, 0.001)$ density function

$$\gamma(\widetilde{Q}) = \frac{2}{\sqrt{\alpha}}.$$

Hence, the Gamma distribution is always positive skewed.

Negative Binomial Distribution The negative binomial is a discrete and positive only distribution used to model the claims number $\widetilde{N}$ random variable in the actuarial literature. A huge emphasis is given to this distribution since, compared to the Poisson, we are able to reject the assumption for which mean and variance of $\widetilde{N}$ are equal.

A negative binomial random variable $\widetilde{N} \sim NB(k, p)$ is characterized by the following probability mass function:

$$P(\widetilde{N} = n) = \binom{n+k-1}{n} p^k (1-p)^n. \tag{2.1}$$

Its main moments are

$$E(\widetilde{N}) = k\frac{(1-p)}{p}$$

$$\sigma^2(\widetilde{N}) = k\frac{1-p}{p^2}$$

$$\gamma(\widetilde{Q}) = \frac{2-p}{\sqrt{k(1-p)}}$$

Another interesting formulation for our purposes is the one that connects a particular Poisson to the Negative Binomial one: given $\widetilde{N} \sim Poisson(n \cdot \widetilde{Q})$

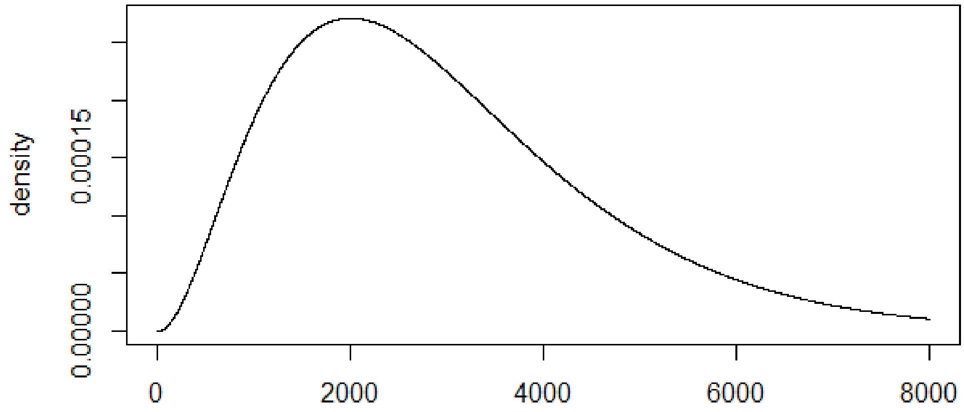


Figure 2.5: $NB(12, 0.004)$ density function

with $\widetilde{Q} \sim Gamma(\alpha, \beta)$, the claims count random variable can be expressed as

$$\widetilde{N} \sim NB(\alpha, \frac{\beta}{\beta + n}). \tag{2.2}$$

Since the presence of $\widetilde{Q}$ is justified in order to add variability to the Poisson without affecting the mean, its parameters are both equal to a number h such that $\widetilde{Q} \sim Gamma(h, h)$ has

$$E(\widetilde{Q}) = \frac{h}{h} = 1$$

$$\sigma^2(\widetilde{Q}) = \frac{h}{h^2} = \frac{1}{h}$$

$$\gamma(\widetilde{Q}) = \frac{2}{\sqrt{h}}.$$

Therefore, we are in front of

$$\widetilde{N} \sim NB(h, \frac{h}{h + n}) = NB(h, p). \tag{2.3}$$

with the following moments:

$$E(\widetilde{N}) = h\frac{(1 - p)}{p} = h\frac{1 - \frac{h}{h+n}}{\frac{h}{h+n}} = h\frac{n}{h} = n$$

$$\sigma^2(\widetilde{N}) = h\frac{1-p}{p^2} = \frac{n(h+n)}{h} = n + \frac{n^2}{h} = n + n^2\sigma^2_{\widetilde{Q}}$$

$$\gamma(\widetilde{Q}) = \frac{2-p}{\sqrt{h(1-p)}} = \frac{2n+h}{\sqrt{nh(h+n)}}.$$

This bridge between these two formulations enables the opportunity to perform useful insights and comparisons in terms of variability.

The Negative Binomial isn't the only solution in order to add variability to the pure Poisson case. For example, in case $\widetilde{Q} \sim Exponential$, it is possible to express $\widetilde{N} \sim Geometric$. These alternatives won't be explored in the thesis, but it's worth noticing the reader of other ways to model $\widetilde{N}$.

2.2 Dependence measures

The main dependence measures between couples of variables are the Pearson correlation, Kendall's Tau and Spearman correlation.

Linear (Pearson) correlation coefficient It is a measure of the strength of a linear association between two variables X and Y, such that:

$$\rho(X,Y) = \frac{COV(X,Y)}{\sqrt{\sigma^2(X)\sigma^2(Y)}} \in [-1,1]$$

where $\sigma^2(X)$ and $\sigma^2(Y)$ are the variances of X and Y, which must be finite, and it is not always satisfied when dealing with heavy tailed distributions.

The Pearson correlation coefficient is the very popular in statistics, but its linearity is a huge pitfall in the actuarial context: generally only extreme events present a strong dependence. Assuming a linear correlation measure would underestimate the correlation between the tails of the distributions, and it would overestimate the one between common events. Therefore, since capital requirement measures rely on the estimation of extreme quantiles, the passage from a linear to a non-linear dependence structure would increase in a significant way the estimated risk on the shoulders of the insurance company.

Kendall's Tau It is a rank correlation measure that measures the concordance between couples of elements (X_i, Y_i) and (X_j, Y_j) (with $i < j$) from a bivariate population. Given a sample n of (X, Y), Kendall 's tau estimator is:

$$\hat{\tau}_k = \binom{n}{2}^{-1} \sum_{i<j} sign((x_i - y_i)(x_j - y_j)) \in [-1, 1]$$

In case $\tau = 1$, there is a perfect agreement between X and Y, and the rankings are exactly the same. On the opposite, in case $\tau = -1$, the two random variables are ordered in the opposite way.

Spearman correlation coefficient It measures the strength and direction of association between two ranked variables. It assumes high values if the ranks of the two variable are similar, and it can be seen as the linear correlation coefficient between ranks. Considering an n number of observations, r_{xi} and r_{yi} the ranks of X and Y, $\bar{r}_x$ and $\bar{r}_y$ the mean rank, an estimator of Spearman's rho is given by:

$$\hat{\rho}_s = \frac{\sum_{i=1,...,n}(r_{xi} - \bar{r}_x)(r_{yi} - \bar{r}_y)}{\sqrt{\sum_{i=1,...,n}(r_{xi} - \bar{r}_x)^2 \sum_{i=1,...,n}(r_{yi} - \bar{r}_y)^2}}$$

2.3 Copula functions

The Copula functions are one of the most popular tools to aggregate different random variables with an underlying non-linear dependence.

A d-dimensional Copula $C : [0, 1]^d \to [0, 1]$ is a multivariate cumulative distribution function with uniform marginals ($U_i \in [0, 1]$):

$$C(u_1, ..., u_d) = \mathbb{P}(\tilde{U_1} \leq u_1, \tilde{U_2} \leq u_2, ..., \tilde{U_d} \leq u_n)$$

Sklar theorem (1959) shows that if F is a d-dimensional distribution function with marginals $F_1, ..., F_d$, there exists a Copula C such that:

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))$$

and so:

$$C(u_1, u_2, ..., u_d) = \mathbb{P}(\tilde{X}_1 \leq F_1^{-1}(u_1), \tilde{X}_2 \leq F_2^{-1}(u_2), ..., \tilde{X}_d \leq F_d^{-1}(u_d))$$

The joint density assumes the following form:

$$f(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \cdot \prod_{i=1}^{d} f_i(x_i)$$

and the conditional density is equal to

$$f(x_1|x_2, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \cdot \prod_{i=2}^{d} f_i(x_i)$$

The main Copula classes that will be considered are Archimedean and Elliptical.

2.3.1 Elliptic Copulas

The elliptical Copulas are based on elliptic distributions and share a link with the linear correlation coefficient. A huge advantage of these copulas is the possibility to specify different linear correlation index between the marginals involved in the copula aggregation.

Gaussian Copula It corresponds to a multivariate gaussian distribution with a matrix linear correlation matrix $R \in [-1, 1]^{dxd}$. It holds if the univariate marginals are gaussian and the relation between them is described by an unique Copula function such that:

$$C_R^{Normal}(u_1, ..., u_d) = \frac{1}{|R|^{\frac{1}{2}}} exp\Big\{-\frac{1}{2} S^T (R^{-1} - \mathbb{I}) S\Big\}$$

with $S_j = \Phi^{-1}(u_j)$, $\phi{-}1$ the inverse of the CDF of the standard Normal, $\mathbb{N}(0, 1)$, $\mathbb{I}$ the identity matrix of size d.

The reason of the popularity behind the Gaussian Copula in the insurance context is its only input R. Since the aggregation of risks in Solvency II is based upon correlation matrices, this copula provides the easiest implementation for an internal model. The other subtle reason is the fact that the use of this copula usually implies an underestimation of the risk, and with that, a reduction in the capital requirements of the company. In this way, companies are able to appear less risky and more capitalised to the market.

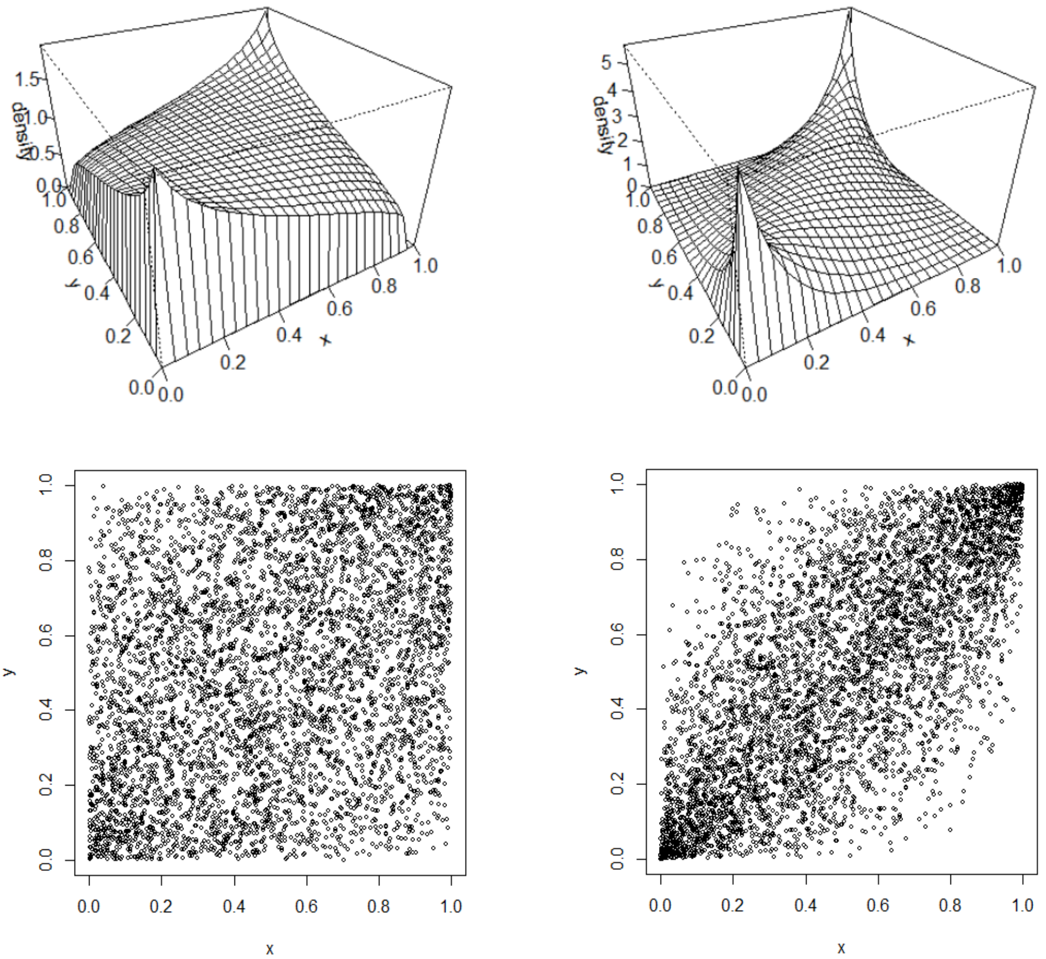


Figure 2.6: Bi-variate Gaussian copula with ρ=0.25 (left) and $\rho = 0.75$ (right)

Student's T It is the Copula associated with the multivariate Student-t distribution and has a matrix as a parameter too, its PDF is:

$$C_{R,v}^{Student}(u_1, ..., u_d) = \frac{1}{|R|^{\frac{1}{2}}} \cdot \frac{\Gamma(\frac{v+d}{2})}{\Gamma(\frac{v}{2})} \cdot \left(\frac{\Gamma(\frac{v}{2})}{\Gamma(\frac{v+1}{2})}\right)^d \cdot \frac{(1 + \frac{1}{v}S^T R^{-1} S)^{\frac{v+d}{2}}}{\prod_{i=1}^d (1 + \frac{S_i^2}{v})^{\frac{v+1}{2}}}$$

where $S_i = T_v^{-1}(u_i) T_v^{-1}$ is the inverse of the CDF of the univariate Student-t distribution with v degrees of freedom. The v degrees of freedom and the matrix parameter have a huge impact on the heaviness of the tails.

Recalling the moral hazard of companies, this copula has the same advantage of the Gaussian one, but depends also on v that shapes the estimated risk situation. The choice of v clearly has a significant impact, and it is easier for

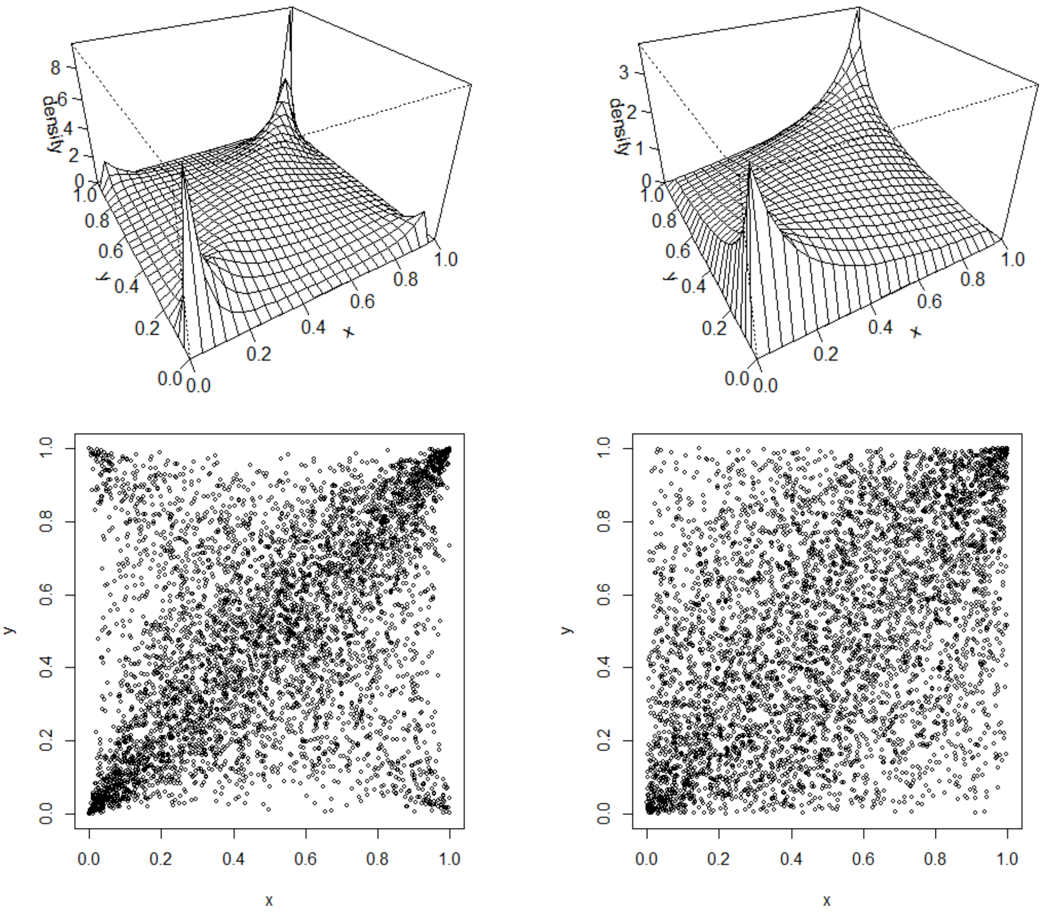


Figure 2.7: Bi-variate T-Student copula with ρ=0.5 and $v = 1$ (left) and $v = 8$ (right)

the insurer to prefer the Gaussian copula, which requires less inputs.

2.3.2 Archimedean Copula

Archimedean copulas are popular for being easy to build and for specifying interesting non-linear correlation structures. A huge pitfall is that they are used almost in bi-variate cases, since all the marginals involved in the Archimedean copula are dependent with the same strength. For this reason this type of copulas are called bi-variate copulas. But, this limitation has been solved through the implementation of the so-called *Vine Copula* aggregation. Let us investigate the Gumbel and Clayton copulas, which are very used in practice.

Gumbel Copula Often useful for catastrophe risks, it is defined as:

$$C(u, v) = exp\{-[(-lnu)^{\theta} + (-lnv)^{\theta}]^{1/\theta}\}$$

where $\theta \in [1, \infty)$ and the Generator of the Copula is given by: $\phi_{\theta}(t) = (-lnt)^{\theta}$

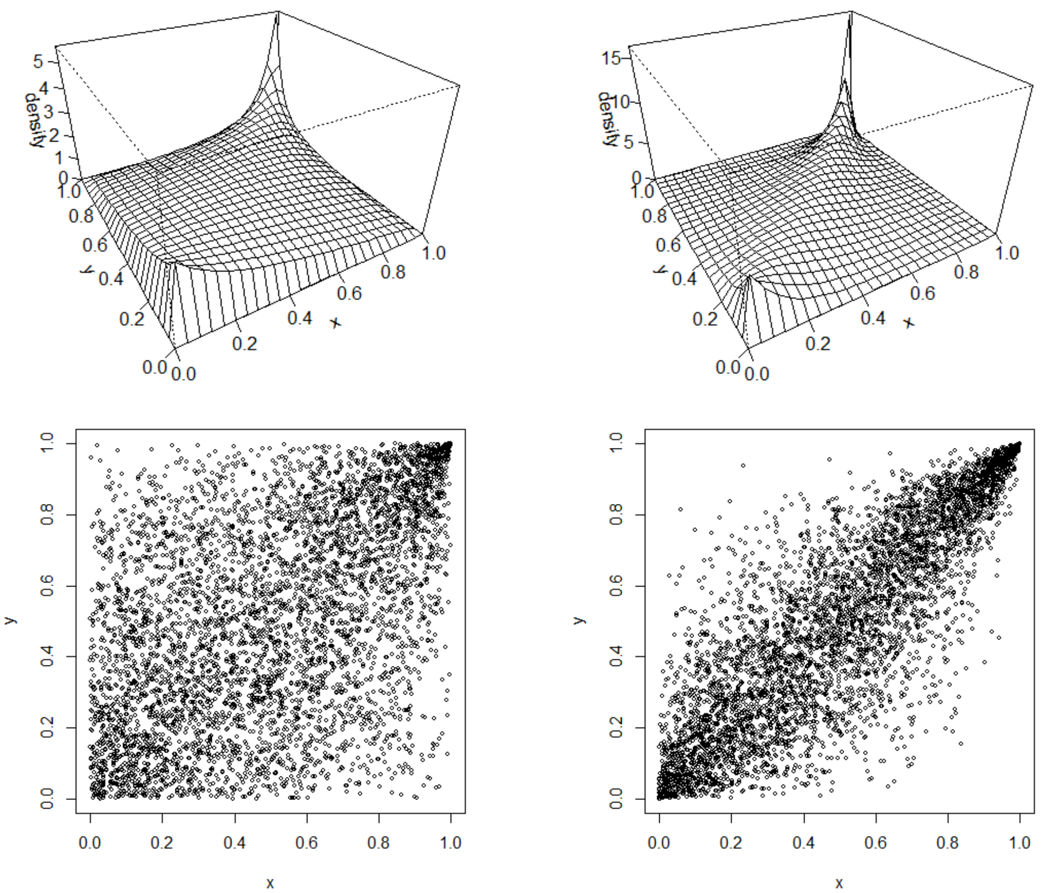


Figure 2.8: Bi-variate Gumbel copula with $\theta = 1.5$ (left) and $\theta = 3$ (right)

The shape of this copula underlies a particular correlation between both the left and right tails of the random variables: it is more probable to observe both the random variables performing well or bad at the same time. It makes sense if we think about it in the catastrophe context: there are years where there aren't notable natural catastrophes around the world, and so, insurance covers against this kind of risk will return a profit. Instead, in other years we observe many significant catastrophes and many covers will suffer a huge loss.

Clayton Copula The Clayton Copula (1978) CDF is defined as:

$$C(u, v) = max\{[u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0\}$$

where $\theta \in [-1, 0) \cup (0, \infty)$ and the generator of the Copula is given by: $\phi_\theta(t) = \frac{1}{\theta}(t^{-\theta} - 1)$

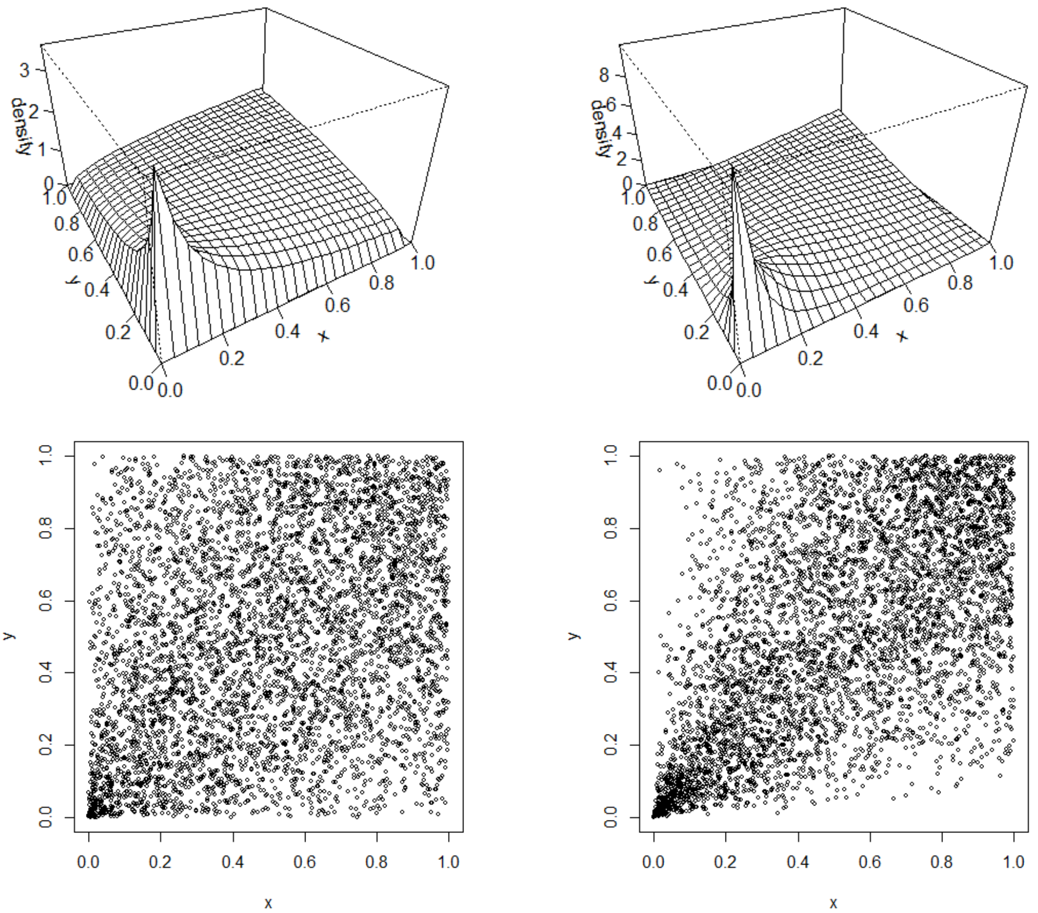


Figure 2.9: Bi-variate Clayton copula with $\theta = 0.5$ (left) and $\theta = 1.5$ (right)

This copula describes a stronger dependence between the left tails of the distributions. So, in case we are considering as random variables the profit of two lines of business, it is more probable to see both performing bad. This point is crucial: when we consider as random variables the claim amounts of two lines, the Clayton copula is not suitable since we would expect a correlation between the right tails. The solution is to use the so-called *Mirror Clayton*, obtained easily as the survival function of a Clayton copula:

$$C^{mirror}(u, v) = 1 - C(u, v)$$

Let's see how a Clayton Copula is generated in a simple bi-variate example:

1. two independent uniform realizations u_1, v_2 are sampled;
2. in the case of a Clayton copula, the correlated marginal is computed as $u_2 = [u_1^{-\theta}(v_2^{-\theta/1+\theta} - 1) + 1]^{-1/\theta}$;
3. the couple (u_1, u_2) is treated as the cumulative probabilities of the marginal random variables $(\widetilde{X}_1, \widetilde{X}_2)$. Therefore, we compute the corresponding quantiles through the inverse operations $F_{\widetilde{X}_1}^{-1}(u_1)$ and $F_{\widetilde{X}_2}^{-1}(u_2)$;
4. the aggregate result is given by the sum $F_{\widetilde{X}_1}^{-1}(u_1) + F_{\widetilde{X}_2}^{-1}(u_2)$.

Thanks to already built packages in many programming software (like *R*), it is possible to draw multivariate samples from the chosen copula, and so, the user is not obliged to know the formula specified at point 2. It is important to note that we need to know the CDF $(F_{\widetilde{X}_1}, F_{\widetilde{X}_2}, \dots)$ of our marginals to have a result expressed in monetary amounts (found in points 3 and 4). In case of a frequency-severity approach, $(F_{\widetilde{N}}, F_{\widetilde{Z}})$ are estimated, but there isn't a way to obtain the resulting exact distribution of the aggregate claim amount $F_{\widetilde{X}}$ for each line of business. The main approaches used to estimate it will be discussed in the next chapter.

2.3.3 Vine Copula

Vine copulas are able to build flexible dependency structures using bi-variate copulas. The copula aggregation is divided in blocks of bi-variate aggregations, in such a way that each block can be characterized with different choice of copula and parameters. Of course, as the dimension of the copula increases, the number of parameters involved increases too, which could lead to overfitting issues and long computational times. The order of aggregation is a very discussed topic in literature because the outcome strongly depends on it. There are several techniques to select the best order, like the Traveling Salesman Problem, the Maximum Spanning Tree and Bayesian approaches.

$$f(X_1,\dots,X_d)=\prod_{j=1}^{d-1}\prod_{i=1}^{d-j}c^*_{i,(i+j)}\cdot\prod_{k=1}^{d}f_k(X_k)$$

where

$$c^*_{i,(i+j)}=C_{i,(i+j)|(i+1),\dots,(i+j-1)}(F_i(X_i|X_{i+1},\dots,X_{i+j-1}),F_{i+j}(X_{i+j}|X_{i+1},\dots,X_{i+j-1})).$$

In a 3 dimensional case we have that one of the possible order of aggregation can be decomposed as follows:

$$\begin{aligned}f(x_1,x_2,x_3)=f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3)\cdot\\ C_{12}(F_1(x_1),F_2(x_2))\cdot C_{23}(F_2(x_2),F_3(x_3))\cdot\\ C_{13|2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2))\end{aligned}\quad(2.4)$$

where the first row contains the marginals, the second the unconditional pairs and the third the conditional pair.

The decomposition present in $c^*_{i,(i+j)}$ is not unique and it depends on the order of aggregation. Bedford and Cooke [3] introduced a graphical vine structure to visualize the order of pair copula. Vines organize the $d(d-1)/2$ bi-variate copulas of a $d-$dimensional Pair-Copula Construction (PCC) in $d-1$ trees.

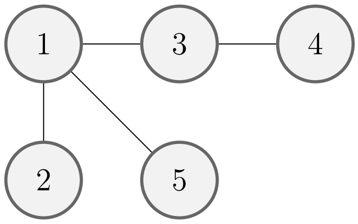
The most general vine is the Regular one, called $R-vine$.

$V=(T_1,\dots,T_{n-1})$ is an R-vine of n elements if:

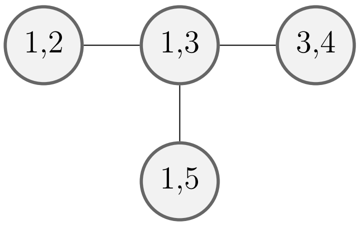
1. T_1 is a tree with nodes $N_1=\{1,\dots,n\}$ and a set of edges E_1;
2. for $i=2,\dots,n-1$, T_i is a tree with nodes $N_i=E_{i-1}$ and edge E_i;
3. for $i=2,\dots,n-1$ and $\{a,b\}\in E_i$ with $a=\{a_1,a_2\}$ and $b=\{b_1,b_2\}$ it must hold that the cardinality of $(a\cap b)=1$ (proximity condition).

To better understand how a Vine Copula can be structured, a 5-dimensional example is provided, showing one of the possible configurations. Starting from the tree T_1, the aggregation develops as follows:

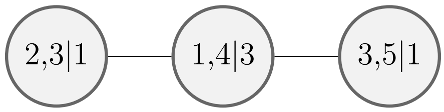
T_1



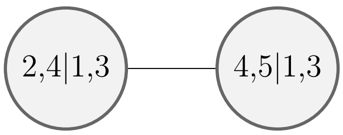
T_2



T_3



T_4

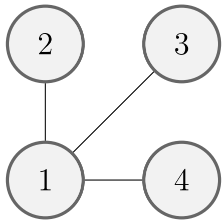


Two types of R-vines are described in literature:

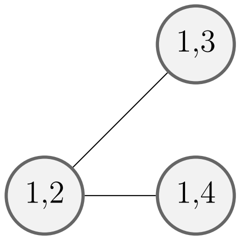
- Canonical (C-Vine);
- Drawable (D-Vine).

C-Vines The C-vines are characterized by a unique node that is connected to all the others. The structure is star-shaped and in a 4-dimensional example we have:

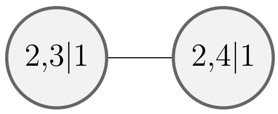
T_1



T_2



T_3



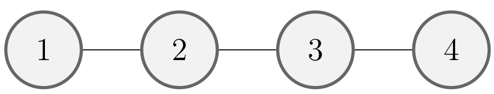
Note that, given the starting configuration of T_1, there isn't a unique way in which the following trees can be structured. In fact, in T_2, also the nodes $(1,3)$ or $(1,4)$ could have been the central ones. While, given T_2, T_3 is the only combination left. Therefore, in a general d-dimensional case, given T_1, there are $\frac{1}{2}(d-1)!$ number of possible C-vines combinations.

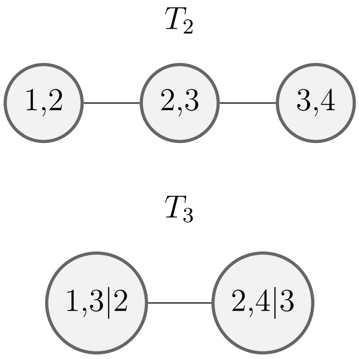
Note that the 3-dimensional case treated with equation (2.4) can be structured as a C-vine where x_2 is the central node connected to both x_1 and x_3. This scenario connects easily with the D-vines.

D-vines The D-vines are characterized by a linear structure. Unlike the C-vine, once T_1 is chosen, the following trees are uniquely defined. In fact, by setting the starting sequence of nodes, there is only one way to aggregate.

A 4-dimensional case is provided:

T_1





In the 3-dimensional case treated with equation (2.4) the starting tree is a path composed sequentially by x_1, x_2, x_3. Trivially, in a 3-dimensional case, there is no difference between a C-vine and D-vine.

Chapter 3

Methods to compute the Aggregate loss distribution

Now that distributions have been explained in the previous chapter, a brief review of methods to compute the aggregate loss distribution $\widetilde{X}$ is presented. Many alternatives are available, where each one has its pro and cons.

First of all, $\widetilde{X}$ will be estimated in the collective risk model framework, and, as mentioned in equation (1.5):

$$\widetilde{X} = \sum_{i=1}^{\widetilde{N}} \widetilde{Z}_i.$$

There are two important assumptions underlying this model, used both in practice and literature:

1. $\widetilde{Z}_i$ are independent and identical distributed;

2. $\widetilde{Z}$ and $\widetilde{N}$ are independent.

The first assumption tells us that all the claims are originated from very similar contracts, in terms of risk covered and contract's limitations. In practice the insurer portfolio, even when referring to a single line of business, is composed by a wide variety of contracts, due to different deductibles, limits and risk exposure involved. Usually, one of the solution is to apply clustering techniques to obtain smaller groups of contracts which are homogeneous. In this thesis the argumentation will follow the first assumption.

The second assumption is very important in the calculation and modelling procedure: by assuming independence between frequency and severity, we are able to easily compute $\widetilde{X}$'s simulations and distributions with the approaches in the following sections.

The mean and the variance of $\widetilde{X}$ can be calculated through exact formulas as follows:

$$E(\widetilde{X}) = E(\widetilde{N})E(\widetilde{Z})$$
$$\sigma^2(\widetilde{X}) = E(\widetilde{N})\sigma^2(\widetilde{Z}) + \sigma^2(\widetilde{N})E(\widetilde{Z})^2 \tag{3.1}$$

Note that no assumptions on $\widetilde{N}$ and $\widetilde{Z}$ distributions is made.

If $\widetilde{N} \sim Poisson(n)$, the variance and the skewness index are equal to:

$$\sigma^2(\widetilde{X}) = nE(\widetilde{Z}^2)$$
$$\gamma(\widetilde{X}) = \frac{nE(\widetilde{Z}^3)}{\sigma^3(\widetilde{X})} = \frac{E(\widetilde{Z}^3)}{n^{1/2}E(\widetilde{Z}^2)^{3/2}}. \tag{3.2}$$

In case we assume that $\widetilde{N} \sim Poisson(n \cdot \widetilde{Q})$ the variance and the skewness index can be calculated as:

$$\sigma^2(\widetilde{X}) = nE(\widetilde{Z}^2) + n^2E(\widetilde{Z})^2\sigma^2(\widetilde{Q})$$
$$\gamma(\widetilde{X}) = \frac{nE(\widetilde{Z}^3) + 3n^2E(\widetilde{Z})E(\widetilde{Z}^2)\sigma^2(\widetilde{Q}) + n^3E(\widetilde{Z}^3)\gamma(\widetilde{Q})\sigma^3(\widetilde{Q})}{\sigma^3(\widetilde{X})}. \tag{3.3}$$

By setting the variance $\sigma^2(\widetilde{Q})$ and the skewness $\gamma(\widetilde{Q})$ equal to zero, the results are referred to the case of a pure $Poisson(n)$.

The coefficient of variation $CV(\widetilde{X})$ is equal to:

$$CV(\widetilde{X}) = \frac{\sigma(\widetilde{X})}{E(\widetilde{X})} = \frac{\sqrt{nE(\widetilde{Z}^2) + n^2E(\widetilde{Z})^2\sigma^2(\widetilde{Q})}}{nE(\widetilde{Z})} \tag{3.4}$$

In these formulas lies a fundamental concept regarding the difference between assuming a pure Poisson or a Negative Binomial. The portfolio diversification (an high value of the expected number of claims n) produces asymptotically interesting results when we look to the relative indexes like $CV(\widetilde{X})$ and $\gamma(\widetilde{X})$: with $n \to \infty$

$Poisson(\cdot)$	n	$n\widetilde{Q}$
CV	0	$\sigma(\widetilde{Q})$
$\gamma(\widetilde{X})$	0	$\gamma(\widetilde{Q})$

Therefore, a big size insurance company could possibly diversify totally its risk under the pure Poisson assumption. While, in the other case, no matter the size of the portfolio, a non-diversified component will always remain. The resulting distribution of $\widetilde{X}$ is clearly more risky in case the frequency is distributed according to a Negative Binomial.

3.1 Distribution fitting based on moments

The most straightforward method is to estimate directly the distribution of $\widetilde{X}$ using the moments of $\widetilde{N}$ and $\widetilde{Z}$.

The simplest approach is to set possible distribution candidates for $\widetilde{X}$ and fitting the parameters for each in such way that the fitted moments are equal to the ones described in the previous formulas.

Another solution is to use approximation formulas like the Normal Power and Wilson-Hilferty. Those are based on transformation of Normal distribution in such a way that the resulting distribution assumes the desired characteristics.

Normal Power The Normal Power formula is the following:

$$F_{\widetilde{X}}(x) \approx Normal\left(-\frac{3}{\gamma(\widetilde{X})} + \sqrt{\frac{9}{\gamma^2(\widetilde{X})} + 1 + \frac{6}{\gamma(\widetilde{X}} \cdot \left(\frac{x - E(\widetilde{X})}{\sigma(\widetilde{X})}\right)}\right)$$

based upon the assumption that the normalized realization can be approximated by a standard normal y such that

$$\frac{x - E(\widetilde{X})}{\sigma(\widetilde{X})} \approx y + \frac{1}{6}\gamma_{\widetilde{X}}(y^2 - 1) \qquad if \quad x > E(\widetilde{X}).$$

This formula usually provides a good proxy in case $\gamma(\widetilde{X}) < 1$, therefore, it is not suitable to model heavy tail risks. The other problem is that Normal Power can be used only for values greater than the mean.

Wilson-Hilferty An alternative approximation, valid also for the left tail, is the Wilson-Hilferty that determines the cumulative density function as

$$F_{\widetilde{X}}(x) \approx Normal\left(\left(\frac{\gamma(\widetilde{X})}{6}-\frac{6}{\gamma(\widetilde{X})}\right)+3\left(\frac{2}{\gamma(\widetilde{X})}\right)^{2/3}\cdot\left(\frac{x-E(\widetilde{X})}{\sigma(\widetilde{X}}+\frac{2}{\gamma(\widetilde{X})}\right)^{1/3}\right).$$

Usually this approximation returns better results compared to the Normal Power. But, also in this case, as $\gamma(\widetilde{X})$ increases, the formula loses the ability to approximate adequately the distribution.

The version presented in this thesis are based only on the first three cumulants (mean, variance and skewness)

3.2 Monte Carlo simulations

Known to be the most used method in practice and literature, Monte Carlo simulations provide solid results with an easy implementation. The outcome of this methodology is a simulated distribution of $\widetilde{X}$, which approximates with increasing accuracy the exact one as the number of simulations increases. It is important to find an adequate balance between accuracy of the simulations and the algorithm's time demand. The steps to simulate $\widetilde{X}$ distribution are the following:

1. for the single simulation a value n is sampled from the distribution of $\widetilde{N}$;
2. n realisations are sampled from the distribution of $\widetilde{Z}$;
3. obtain $\widetilde{X} = \sum_{i=1}^{n} \widetilde{Z}_i$;
4. repeat the first three steps for the number of simulations desired;

A realistic estimation of $\widetilde{N}$ and $\widetilde{Z}$ is crucial in order to avoid biased results. If this step is not performed correctly even a high number of simulations will lead to misunderstand the real underlying risk, to which the company is exposed.

An important reminder when using Monte Carlo simulation is to compare the moments of simulated data with the exact one calculated through formulas (3.1). In this way the user is able to verify quickly the accuracy of the simulations. An particular emphasis should be given to the skewness index γ, which ,in case $\widetilde{Z}$ is very skewed, requires a huge number of simulations to converge to the exact one. An underestimation of γ usually means that the extreme events, the so-called *tail* of the distribution, are not represented properly by the simulations. A common shortcoming of simulation methodologies is that every run will return a different result, and for this reason it may be dangerous to blindly give full credibility to such results. Despite these problems, Monte Carlo simulations are able to model the most difficult insurance and reinsurance scenarios.

It is worthy to acknowledge that the model proposed in this section may be too much time consuming to deal with reinsurance treaties. In fact, every time a new excess of loss treaty is tested, the whole model should re-run to cut properly the single claims $\widetilde{Z}_i$.

An advanced extension of Monte Carlo simulations is described in detail later in the thesis. Building the proper simulation environment is the stepping stone to investigate reinsurance.

3.3 Panjer Algorithm

The Panjer algorithm is based on a recursive relationship between $\widetilde{N}$ and $\widetilde{Z}$, where both are discrete random variables. The algorithm requires the frequency to be modelled with a distribution that belongs to the *(a,b,0)* class of distributors, which satisfy the following recursive property:

$$P(\widetilde{N} = n) = (a + \frac{b}{n})P(\widetilde{N} = n - 1).$$

There are only three distributions in that class:

- *Binomial(n,p)*, with $a = \frac{-p}{1-p}$ and $b = \frac{p(n+1)}{1-p}$;
- *Poisson(λ)* with $a = 0$ and $b = \lambda$;
- *Negative Binomial(k,p)* with $a = (1 - p)$ and $b = (1 - p)(k - 1)$.

Since $\widetilde{Z}$ is usually modelled with a continuous distribution, its density function should be discretised in such a way that its discrete realisations are multiples of the so-called *span h*.

The Panjer algorithm defines the probability mass function of $\widetilde{X}$ as:

$$P(\widetilde{X} = x) = \frac{1}{1 - aP(\widetilde{Z} = 0)} \sum_{z=1}^{x} (a + \frac{b \cdot z}{x}) P(\widetilde{Z} = z) P(\widetilde{X} = x - z) \qquad (3.5)$$

with the starting condition

$$P(\widetilde{X} = 0) = P(\widetilde{N} = 0).$$

and, because of the discrete $\widetilde{Z}$, the domain of $x = j \cdot h$, where $j \in \mathbb{N}$

The dimension of h impacts on both the accuracy and the time required by the algorithm: a tiny h would produce a discrete $\widetilde{Z}$ with a small loss of information, but the algorithm will require more time to run. The strength of the algorithm is the ability to determine with accuracy the tail of the distribution of $\widetilde{X}$. Imagine using a span of 1000 euro to discretize $\widetilde{Z}$: for low values of Z it may appear a rough approximation, but for extreme claim events the loss of information is almost nothing.

But, the algorithm can presents a numerical problem: in case $P(\widetilde{N} = 0)$ is so small that the computer can only represent it with a zero, the whole calculation fails. This scenario, called *numerical underflow*, happens when dealing with frequency component characterized by a high mean. It has been found by Kaas et al. [18] that a $Poisson(n)$ with $n \sim 727$ generates underflow. In the case of $\widetilde{N} \sim Poisson(n)$, the solution is to perform the recursion by using $Poisson(\frac{n}{2^s})$ and applying convolutions s times on $\widetilde{X}$. Instead, in case of a negative binomial $NB(k, p)$, the shrinking should be applied only to the size parameter k. The use of convolutions increases the computational times.

Due to these reason, the Panjer algorithm's application is often preferred to study the risk transferred with a non-proportional treaty. In fact, this type of contracts affects only the tail of the distribution of $\widetilde{Z}$ and the expected number of large claims is usually very low (less than 727). Furthermore, the user can set a high value of h to cut computational times without losing much information.

Obviously the algorithm is hard to be applied in practice directly by the reinsurer since usually only the insurer has the whole info to model in detail the risk. In fact, without a proper estimation of $\widetilde{N}$ and $\widetilde{Z}$ the algorithm is not suitable.

Later on, it will be shown how the algorithm, compared to Monte Carlo simulations, is able to cut computational times from hours to seconds even when dealing with the risk of the direct insurer. Panjer algorithm provides a faster and more precise alternative in risk modelling under certain circumstances. The cons, with respect to Monte Carlo simulations, are the more difficult implementation and the less flexibility when dealing with the presence of non-proportional reinsurance with Aggregate Deductibles, Aggregate Limits and reinstatements.

Chapter 4

SCR in Solvency II and Risk Theory

Solvency II is a regulatory standard for European insurance and reinsurance companies that went live on 1 January 2016. Its aim is to establish capital requirements and risk management standards that properly describe the risk faced. The purpose of this chapter is to present briefly the main features of Solvency II that will be treated in the thesis. Describing the whole regulation is avoided in order to obtain a more focused discourse.

The framework is composed by three main areas, called *pillars*:

- Pillar 1 defines the quantitative capital requirements (the amount of financial resources needed by the (re)insurer to be considered solvent) and the market-consistent valuation of assets and liabilities;

- Pillar 2 describes the qualitative supervisory review process, where the regulators incentive the companies to perform a better internal control and risk management;

- Pillar 3 introduces public disclosures requirements, which oblige companies to publish annually a report on financial and solvency situation.

Regarding Pillar 1, Solvency II introduced two capital requirements indicators, called respectively Solvency Capital Requirement *(SCR)* and Minimum Capital Requirement *MCR*. The company's own capital has to be greater than these two measures. The MCR is characterized by the following floor and cap:

$$25\%SCR \leq MCR \leq 45\%SCR$$

In case the amount of resources fall below the SCR, a regulatory intervention would be triggered, while, in case of a breach of the MCR, the company would be declared insolvent and unable to continue its business.

The so-called Standard Formula *(SF)* of Solvency II for SCR calculation is based on the Value-at-Risk with level of confidence 99.5% *(VaR 99.5%)* with one-year time horizon. If an insurance company's own funds are equal to the SCR, the default occurs on average once every 200 years. From another perspective, every year on average 1 over 200 insurers bankrupts. Of course, in practice defaults are rarer to occur since companies' own capital is much higher then the SCR. And, as mentioned before, recovery plans must be applied in case of infringement of the SCR.

Other than the standard formula, companies have the opportunity to develop an internal VaR 99.5% calculation method, which must be approved and validated by the supervisory authorities. There are three possibilities:

- Undertaking Specific Parameters *(USP)*, where the company computes its own parameters and substitutes the one provided by the Standard Formula. The reason why the USP is made is because the parameters of the SF may not be well representative of the risk profile of the company. This procedure is not difficult to be approved since it doesn't modify the structure provided by the regulation;

- Internal Model *(IM)*, where the company builds from scratch a calculation procedure to determine its own capital requirements. The methodologies involved may be very technical, and the approval from the supervisor may take time due to the significant changes. The insurer is required to demonstrate that the internal model plays a fundamental role in the decision-making processes;

- Partial Internal Model, where the company uses an IM only for some risk modules, and uses SF or USP for the remaining part.

The need of a rigorous process of validation and approval by the supervisor is driven by the fact that every company's attempt is to reduce its own capital requirement via application of USP or IM. In fact, the calibration of the SF's formulas advantages small size insurers and penalizes large insurers. Hence, anti-selection occurs and the permission to apply the USP or IM is requested by only those who receive a benefit from them. In particular, small insurance companies are usually characterized by a low diversification benefit of the portfolio and by a high relative variability of the losses, and the SF's volatility parameters may underestimate their riskiness.

4.1 Valuation of Assets and Liabilities

According to Article 75 and 76 of Solvency II Directive, assets and liabilities shall be valued at fair value. The choice behind the use of the fair value approach is to limit subjectivity. In particular:

- assets should be valued at the amount for which they could be exchanged between knowledgeable willing parties in an arm's length transaction;
- liabilities should be valued at the amount for which they could be transferred, or settled, between knowledgeable willing parties in an arm's length transaction.

In the insurance context, the most important liabilities are technical provisions, which represent the current amount insurance and reinsurance undertakings would have to pay if they were to transfer their insurance and reinsurance obligations immediately to another insurance or reinsurance undertaking.

The technical provisions *(TP)* shall be calculated in a prudent and reliable manner as it follows:

$$TP = BE + RM \tag{4.1}$$

where, according to Article 77 of the directive:

- BE = Best Estimate, which shall be equal to expected present value of future cash flows, using the relevant risk free interest rate term structure to discount published monthly by EIOPA. The BE shall be calculated using credible information and realistic assumptions, taking account of all the cash in-flows and out-flows required to settle the insurance and reinsurance obligations. In formula, given the expected value of future cash flows $F_1, F_2, \ldots, F_T$ and the interest spot rates $r(0,1), r(0,2), \ldots, r(0,T)$, we have:

$$BE = \sum_{t=1}^{T} \frac{E(F_t)}{(1 + r(0,t))^t} \tag{4.2}$$

with $r(0,t) = 1 + i(0,t) + adj$, where i is the risk free interest rate and adj is an adjustment, usually applied to mitigate the effects of short term market volatility.

- RM = Risk Margin, which is a buffer needed to ensure that the value of technical provisions is equal to the amount that insurance and reinsurance undertakings would be expected to require in order to take over and meet the insurance and reinsurance obligations. RM is computed using the cost of capital approach, that consists in the determination of the cost of providing an amount of eligible own funds equal to the SCR necessary to support the insurance and reinsurance obligations over the lifetime thereof. The RM formula involves the present value of the projection of the SCR until liabilities' run-off, such that:

$$RM = CoC \cdot \sum_{t=0}^{T} \frac{SCR_t}{(1 + i(0,t+1))^{t+1}} \tag{4.3}$$

where CoC is the Cost of Capital rate, fixed at 6% by EIOPA, $i(0,t+1)$ is the risk free interest spot rate and the SCR is referred only to 4 risks: underwriting, default, operational and material market risk.

An useful key indicator in Solvency II is the so-called *Solvency Ratio (SR)*, given by:

$$SR = \frac{EOF}{SCR} \tag{4.4}$$

where EOF denotes the *Eligible Own Funds*, which are the amount of own capital used to cover the SCR. The EOF includes only the Own funds that satisfy a specific mechanism of tiers. The steps are represented briefly in Figure (4.1). To simplify the computations, usually in risk theory the numerator of equation (4.4) is set equal to the Excess of Assets over Liabilities.

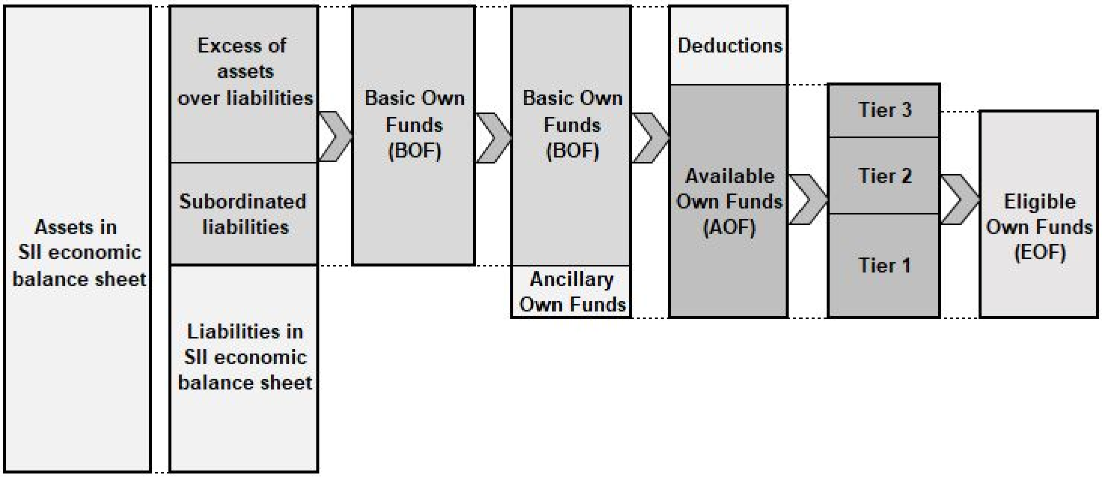


Figure 4.1

4.2 SCR Calculation

On the thesis' computational side, the SCR will be calculated according to three different approaches:

1. a SF approach using the risk structure and the parameters provided by the directive;
2. a Quasi-USP approach using the risk structure provided by the directive and ad-hoc parameters based on the characteristics of the insurer;
3. an IM approach, where the risk is adequately computed using risk theory methodologies calibrated on the characteristics of the insurer. This model refers partially to some Standard Formula's assumptions to keep some consistency with the previously described approaches.

Hence, the aim of the following sections is to describe all the elements needed to fully understand the SCR computation under the Solvency II framework.

4.2.1 SCR in Solvency II

Coming back to the definition of SCR, it takes into account all quantifiable risks. According to the first Pillar of Solvency II, the calculation of the SCR is divided into the modules of **Figure 4.2**, which are then aggregated using correlation matrices provided by the directive.

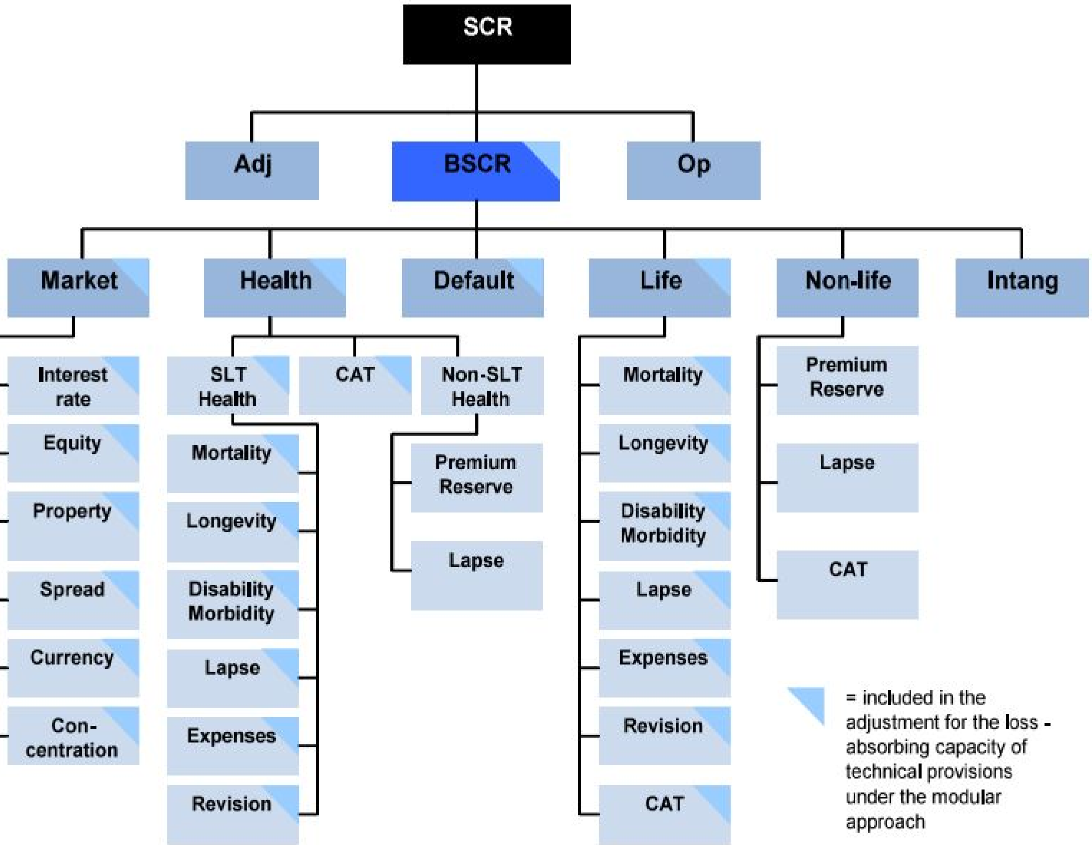


Figure 4.2

The SCR is determined as:

$$SCR = BSCR + Adj + SCR_{op} \tag{4.5}$$

where

- $BSCR$ = Basic Solvency Capital Requirement given by the combination of the six risk modules (market, health, default, life, non-life and intangible assets) present in the level below;
- SCR_{op} = Solvency Capital Requirement for operational risk;
- Adj = Adjustment for the risk absorbing effect of technical provisions and deferred taxes.

In this step of the calculation, no diversification benefit is present by summing the three components. Instead, for the BSCR calculation, the directive provides correlation matrices in the standard formula and the insurer is able to diversify the individual capital requirements belonging to the six risk modules that compose the BSCR. The formula is the following:

$$BSCR = \sqrt{\sum_{ij} \rho_{ij} \cdot SCR_i \cdot SCR_j} + SCR_{intangible} \tag{4.6}$$

where

- i and j are respectively the index of rows and columns of the correlation matrix presented in **Figure (4.3)**;
- ρ_{ij} is the correlation between the risk i and j;
- SCR_i, SCR_j are the capital requirements of the risk i and j;
- $SCR_{intangible}$ is the capital requirement for intangible asset risk.

j i	Market	Default	Life	Health	Non-life
Market	1				
Default	0.25	1			
Life	0.25	0.25	1		
Health	0.25	0.25	0.25	1	
Non-life	0.25	0.5	0	0	1

Figure 4.3

In this thesis the main module considered is the *Non-life*, in particular the premium risk sub-module. Hence, the following paragraphs will be focused mainly on the topics of interest contained in this module.

SCR Non-Life Underwriting Risk

The Non-life underwriting risk represents the risk coming from non-life insurance contracts, belonging to both already existing business and new ones expected to be written in the next 12 months.

As shown in Figure 4.2, the solvency capital requirement for Non-life underwriting risk,denoted with SCR_{NL}, is composed by 3 SCR:

- non-life premium and reserve NL_{pr};
- non-life lapse NL_{lapse}, given by possible incorrect assumptions regarding lapse and renewal options of non-life insurance contracts;
- non-life catastrophe NL_{CAT}, due to the possible loss arising from insufficient pricing and reserving assumptions related to extreme events.

These three sub-modules' SCR are aggregated as following:

$$SCR_{NL} = \sqrt{\sum_{ij} CorrNL_{ij} \cdot NL_i \cdot NL_j} \tag{4.7}$$

using the correlation matrix depicted in Figure 4.4.

CorrNL	NL_{pr}	NL_{lapse}	NL_{CAT}
NL_{pr}	1		
NL_{lapse}	0	1	
NL_{CAT}	0.25	0	1

Figure 4.4

The sub-module that assumes a particular interest in the thesis is the non-life premium risk module.

Standard Formula Non-Life Premium and Reserve Risk The calculation of the SCR_{pr} for this risk is based on the concept of standard deviation, which is defined as the multiplication between a volume V and a volatility factor σ. The volatility factor should be seen as a coefficient of variation, and not as a standard deviation itself. This reasoning is implemented in the Standard Formula as

$$SCR_{pr} = 3 \cdot \sigma_{nl} \cdot V_{nl}. \tag{4.8}$$

The choice to use a multiplier equal to 3 is connected to the definition of Value-at-Risk: $3 \cdot \sigma(X)$ approximates the distance between the quantile at level 99.5 and the mean of the distribution. As mentioned before, 2.58 is used in case of Normal distributed random variables. Therefore, the Standard Formula assumes that premium and reserve risk is positively skewed.

Both V_{nl} and σ_{nl} are computed as the aggregation of premium and reserve volumes and volatility factors of 12 segments listed in Figure (4.5). The structure is the following:

$$V_{nl} = \sum_{s=1}^{12} V_s = \sum_{s=1}^{12} (V_{prem} + V_{res})$$

$$\sigma_{nl} = \frac{1}{V_{nl}} \cdot \sqrt{\sum_{s,t} CorrS_{s,t} \cdot \sigma_s \cdot V_s \cdot \sigma_t \cdot V_t} \tag{4.9}$$

CorrS	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1: Motor vehicle liability	1											
2: Other motor	0,5	1										
3: MAT	0,5	0,25	1									
4: Fire	0,25	0,25	0,25	1								
5: 3rd party liability	0,5	0,25	0,25	0,25	1							
6: Credit	0,25	0,25	0,25	0,25	0,5	1						
7: Legal exp.	0,5	0,5	0,25	0,25	0,5	0,5	1					
8: Assistance	0,25	0,5	0,5	0,5	0,25	0,25	0,25	1				
9: Miscellaneous.	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1			
10:Np reins. (casualty)	0,25	0,25	0,25	0,25	0,5	0,5	0,5	0,25	0,25	1		
11:Np reins. (MAT)	0,25	0,25	0,5	0,5	0,25	0,25	0,25	0,25	0,5	0,25	1	
12:Np reins. (property)	0,25	0,25	0,25	0,5	0,25	0,25	0,25	0,5	0,25	0,25	0,25	1

Figure 4.5

The inputs required for the calculation of the volumes V_s of each segment s are:

- PCO_s is the Best estimate for claims outstanding net of recoverables from reinsurance contracts and special purpose vehicles;
- P_s is the estimate of premiums to be earned by the undertaking during the following 12 months;
- $P_{last,s}$ is the earned premiums by the undertaking during the last 12 months;
- $FP_{existing,s}$ is the expected present value of premiums to be earned by the undertaking after the next 12 months for already existing contracts;
- $FP_{future,s}$ is the expected present value of premiums to be earned by the undertaking for contracts that will be recognized in the following 12 months. Only the premiums earned after 12 months from the initial recognition are considered.

The volume V_s is equal to the sum between the premium volume $V_{prem,s}$ and the reserve volume $V_{res,s}$, considering the geographical diversification DIV_s as

$$V_s = (V_{prem,s} + V_{res,s}) \cdot (0.75 + 0.25 \cdot DIV_s)$$

where

$$V_{prem,s} = max(P_s; P_{last,s}) + FP_{existing,s} + FP_{future,s}$$

$$V_{res,s} = PCO_s$$

$$DIV_s = \frac{\sum_j (V_{prem,j,s} + V_{res,j,s})^2}{\sum_j (V_{prem,s} + V_{res,s})^2}$$

where j denotes the geographical segments and $(V_{prem,j,s}; V_{res,j,s})$ the volumes related only to the obligations situated in the geographical segment j.

DIV_s is set to 1 for segments $6, 10, 11, 12$. In case all the risks belong to the same area, DIV_s is trivially equal to 1 and V_s is equal to the simple sum between $V_{prem,s}$ and $V_{res,s}$.

The volatility factor σ_s of the individual segment s is defined through the aggregation of premium $\sigma_{prem,s}$ and reserve $\sigma_{res,s}$:

$$\sigma_s = \frac{\sqrt{(\sigma_{prem,s}V_{prem,s})^2 + \sigma_{prem,s}\sigma_{res,s}V_{prem,s}V_{res,s} + (\sigma_{res,s}V_{res,s})^2}}{V_{prem,s} + V_{res,s}}.$$

This formula can be seen as a risk aggregation with an underlying correlation between premium and reserve risk equal to 0.5.

In Table (4.1) the values of Standard Formula's σ_s are represented for each segment s that compose the non-life underwriting risk. The NP_{lob} factors are applied to σ_{prem} of segments 1,4 and 5 to consider the risk-mitigating effect of per risk excess of loss reinsurance in force for those segments. $NP_{lob} = 80\%$ in case the aforementioned reinsurance is present, 100% otherwise.

Note that a fixed 80% is a simplification that doesn't take in account the real amount of risk mitigation that the insurer benefits from the reinsurance contract. In fact, under particular distributional assumptions of the severity and frequency component, the actual reduction is very limited and the NP factor overestimates the risk reduction. For example, choosing a Negative Binomial for the frequency component strongly limits the mitigation benefit of reinsurance.

An insurance undertaking through the USP can propose to the supervisor its own $\sigma_{prem,s}^{USP}$ and $\sigma_{res,s}^{USP}$ for each segment calibrated ad-hoc on its portfolio. According to how many years of historical experience its business has, the resulting volatility factor used calculation is computed as a weighted average between the Standard Formula's one and the proposed one. Other than that, the structure provided by the Standard Formula shall remain the same when using USP.

Instead, in case proportional reinsurance like Quota Share is present, the volume V_s decreases to $\alpha \cdot V_s$, where α is the Quota Share's retention coefficient of the undertaking. In this way, proportional reinsurance aim is to influence the volumes, and the Non-proportional's aim is to reduce the volatility factors. These considerations are very practical and realistic due to the structure of such contracts.

Table 4.1: Non-life premium and risk sub-module's segments and σ

Segment	$\sigma_{premium}$	$\sigma_{reserve}$
1. Motor vehicle liability insurance and proportional reinsurance	$10\% \cdot NP_{lob}$	9%
2. Other motor insurance and proportional reinsurance	8%	8%
3. MAT insurance and proportional reinsurance	15%	11%
4. Fire insurance and proportional reinsurance	$8\% \cdot NP_{lob}$	10%
5. 3rd-party liability insurance insurance and proportional reinsurance	$14\% \cdot NP_{lob}$	11%
6. Credit insurance and proportional reinsurance	12%	19%
7. Legal expenses insurance and proportional reinsurance	7%	12%
8. Assistance insurance and proportional reinsurance	9%	20%
9. Miscellaneous insurance and proportional reinsurance	13%	20%
10. Np reinsurance (casualty)	17%	20%
11. Np reinsurance (MAT)	17%	20%
12. Np reins (prop)	17%	20%

4.2.2 SCR in Risk Theory

In this section the considerations are referred mainly to the premium risk with reference to both actuarial literature and Standard Formula's structure.

In risk theory's literature the capital requirement calculation relies on the estimate of the distribution's quantiles, without the approximation $3\sigma V$ provided by the Standard Formula.

The SCR is computed following the definition Value-at-Risk at level of confidence 99.5%:

$$SCR = quantile_{0.995}(\widetilde{X}) - E(\widetilde{X}) = quantile_{0.995}(\widetilde{X}) - P \tag{4.10}$$

where P is the pure premium.

Under this notation, the formula is comparable with equation (4.8) of Standard Formula. Many authors proved that the multiplier 3 overestimates the risk for big size insurers, that have a more diversified portfolio, and underestimates for small size insurers.

But, according to risk theory, equation (4.10) doesn't represent properly the capital requirement for the generic risk $\widetilde{X}$. In fact, the insurer, before depleting its own capital, uses to cover the losses both the pure premium P and the safety loading $\lambda \cdot P$. Hence, the SCR assumes another definition, denoted as follows with SCR^{RT}:

$$SCR^{RT} = quantile_{99.5\%}(\widetilde{X}) - P(1+\lambda). \tag{4.11}$$

Trivially, $SCR^{RT} > SCR$ in case of positive loadings ($\lambda > 0$).

The SCR^{RT} formulation gives a better emphasis to reinsurance scenarios: when buying reinsurance, the direct insurer reduces the quantile, the expected loss and the safety loading. Of course, the insurer agrees to buy the cover only if the capital requirement decreases with it, that it is almost always true in practice, except for reinsurance contracts with excessive loading $\lambda_{re} \cdot P_{re}$. To this purpose, SCR^{RT} is able to capture a possible incoherence, while equation (4.10) isn't and it will always be lower when net of reinsurance. Other than that, using the SCR formula according to Solvency II to diagnose the risk-profit situation of the company is a suitable choice if also a profit indicator, like Return on Equity (ROE), is taken in consideration.

The ROE between time $t-1$ and t computed in $t-1$ is defined as:

$$ROE(t-1,t) = \frac{E(return_t)}{equity_t} = \frac{E(\widetilde{U}_t - U_{t-1})}{U_{t-1}} \tag{4.12}$$

where U is the own funds of the insurer. Of course, the equity is not equal to the own funds, even if they have a similar definition. The ROE considered here will be considered as an indicator in the Solvency II framework.

The own funds $\widetilde{U}$ at the end of time t are a random variable that can be defined as:

$$\widetilde{U}_t = U_{t-1} + B_t - \widetilde{X}_t - E_t \tag{4.13}$$

where

$$B_t = P_t(1+\lambda) + c \cdot B_t$$

and it is assumed that the expenses are deterministic and equal to the expense loading, such that $E_t = c \cdot B_t$.

Note that the formula of $\widetilde{U}_t$ is a simplification that doesn't consider taxes, dividends, return of investment and inflation.

Therefore, the equation (4.12) can be rewritten as

$$ROE(t-1,t) = \frac{E(B_t - \widetilde{X}_t - E_t)}{U_{t-1}} = \frac{E(P_t(1+\lambda) - \widetilde{X}_t)}{U_{t-1}} = \frac{\lambda \cdot P}{U_{t-1}} \tag{4.14}$$

The capital requirement can be computed also using the quantiles of $\widetilde{U}$, since $\widetilde{U}$ can be considered as a reversed and shifted $\widetilde{X}$, which is the only random variable. In fact, by setting $U_{t-1} = 0$, the capital requirement of equation (4.10) is obtained as the following value-at-risk:

$$SCR = -quantile_{0.5\%}(\widetilde{U}_t) + \lambda \cdot P \tag{4.15}$$

since

$$\widetilde{U}_t = P_t(1+\lambda) - \widetilde{X}_t$$

and

$$SCR = quantile_{99.5\%}(\widetilde{X}_t) - P_t = -quantile_{0.5\%}(-\widetilde{X}_t) - P_t =$$
$$= -quantile_{0.5\%}(-\widetilde{X}_t) - P_t - \lambda P_t + \lambda P_t = -quantile_{0.5\%}(P_t(1+\lambda) - \widetilde{X}_t) + \lambda P_t.$$

The quantile at level 0.5% of $\widetilde{U}_t$ is the yearly loss (or negative profit) in own capital observed on average 1 year every 200 years. It is noticeable in equation (4.15) that, according to Solvency II's logic, the expected profit $\lambda \cdot P$ is not used to cover potential losses. In fact, in this formulation, the SCR is still calculated as the difference between the quantile and the mean of $\widetilde{X}$, since $E(\widetilde{U}_t = \lambda P)$ and $quantile_{0.5\%}(\widetilde{U}_t)$ measures the distance between the quantile and zero.

It is easily possible to obtain a risk theory SCR^{RT} of equation (4.11) as

$$SCR = -quantile_{0.5\%}(\widetilde{U}_t). \tag{4.16}$$

In this case, in presence of a loss $\widetilde{X}_t > P_t(1+\lambda)$, before decreasing the own funds, the loading $\lambda \cdot P$ is depleted.

Chapter 5

Pricing and optimization principles

The definition of Optimal Reinsurance is a difficult task since it depends on the choice of metrics, level of confidence and which constraints shall be satisfied at company level. The most straightforward solution is adopting in the calculation risk measures and level of confidence provided by Solvency II. Note that the optimization treated here is from the point of view of the direct insurer, and not of the reinsurer.

The main difficulty when dealing with optimal reinsurance is the fact that, in order to evaluate the efficiency of a particular reinsurance structure, the user shall be able to compute both the mitigation effects and the overall price of reinsurance. Usually, the reinsurance premiums are not available in advance for many contracts and the solution is their estimation through the use of pricing principles or pricing curves. On the other side, the risk mitigation of extreme events is based on the chosen loss model, which is just a simplification of the real world.

Through the concept of efficient frontier, a set of optimal contracts can be determined. Usually the frontier is defined using a profit and a risk indicator. A further step can be taken to pass from a 2-dimensional indicator (set of optimal reinsurance structures) to a 1-dimensional one, and obtain subsequently a single optimal treaty that maximize/minimize such indicator.

5.1 Efficient Frontier

The starting point to define an efficient frontier is the concept of Pareto efficiency related to reinsurance programs:

- if A and B cost the same premium, but A reduces more risk than B, then A $\succ$ B
- if A and B mitigate the same amount of risk, but A costs less than B, then A $\succ$ B

In the case where A doesn't dominate B, the two reinsurance structures are not comparable: if A mitigates more risk but requires a higher premium than B, then it is not possible to affirm which structure is objectively better. The set of programs that are not strictly dominated by any other program is the *efficient frontier*.

Once understood this notion, a measure to quantify the amount of risk reduction is needed. VaR provides an useful indicator under the level of confidence chosen by the company, since it can be easily linked to Solvency II capital requirement. The Tail VaR is a valid alternative to consider more deeply the underlying risk after the chosen quantile.

The other ingredient to build the efficient frontier is the reinsurance premium, which can be extrapolated from pricing curves or pricing principles. In this thesis the approach is based on the computations through pricing principles, like *Proportional Hazard* and *Standard Deviation principle*, that will be discussed adequately in the next sections. It should be always kept in mind that the uncertainty connected to reinsurance pricing influences the frontier.

Several measures can be compared in a 2 dimensional plot to find an efficient frontier, and each combination leads to different results. For example, Parodi [23] proposes the trade-off between reinsurance premium and VaR, which is a straight-forward approach. A good alternative to the use of the premium is the profit ceded to the reinsurer, where we neglect the expected mitigated loss. Why?

- A treaty characterized by a high premium and a tiny safety loading has the advantage of transferring a considerable amount of risk without losing too much profit. Choosing the ceded profit rather than the ceded premium is a good way to keep more focus on key performance indexes. The Quota Share reinsurance is an example of such contract;

- Excess of Loss contracts are usually characterized by relative small pure premiums and high safety loadings since they deal with more volatile and infrequent risks. Despite the small premium, such contract can return significant benefit in terms of reduction of VaR, and therefore, considering only the ceded profit as comparison metric with other contracts may be a more suitable choice.

The same reasoning can be applied to the choice of a risk measure: the VaR itself can be decomposed in expected and unexpected loss. If two treaties cost the same and provide the same VaR, it doesn't mean that both return the same unexpected loss (or SCR under a level of confidence of 99.5%). Therefore, since the selection of the SCR is a very important factor, neglecting the expected loss also in the risk measure can lead to an interesting alternative optimal frontier.

5.1.1 Total Cost of Risk

The main problem connected to the efficient frontier is the fact that, since all the programs that compose it are not comparable, an optimal reinsurance doesn't objectively exist. Therefore, the introduction of other metrics is needed to rank this set of contracts and determine the best option.

An interesting measure is the *Total Cost of Risk* (denoted with TC), which, from the point of view of an insurer that buys reinsurance, is defined as:

$$TC = E(\widetilde{X}^{in}) + CoC \cdot (VaR_{99,5\%}(\widetilde{X}^{in}) - E(\widetilde{X}^{in})) + P^{re} \tag{5.1}$$

where

- $E(\widetilde{X}^{in})$ is the expected loss of the direct insurer net of reinsurance;
- CoC is the *cost of capital* rate;
- $(VaR_{99,5\%}(\widetilde{X}^{in}) - E(\widetilde{X}^{in}))$ is the SCR under Solvency II framework;
- P^{re} is the reinsurance premium.

The goal of this optimization is to minimize the total cost through reinsurance. An approach is to apply such minimization to the programs that lie on the frontier. An insurer without reinsurance is characterized by higher expected losses and SCR, but zero reinsurance premiums. In literature a reinsurance program is efficient if:

$$P^{re} < \Delta(E(\widetilde{X}^{in})) + CoC \cdot \Delta(SCR)$$

where $\Delta(\cdot)$ indicates the reduction from gross to net of reinsurance. If the inequality holds, the reinsurance is able to create value for the direct insurer according to the Total Cost measure. Note that it's possible that such reinsurance contract may not exist at all, since in some scenario the reinsurance can only increase TC. This possible result doesn't mean that the reinsurance is useless, but that the TC may not be an adequate metric to evaluate the effectiveness of the treaty.

The definition of TC is not written on stone, and one can decide to modify the $VaR_{99,5\%}$ with another risk measure and another confidence level.

It is interesting to note that the profit of the insurer is somehow considered in the formula: if the reinsurer is applying positive safety loading P^{re} decreases the insurer's profit. In fact, the minimization of the total cost can be written as:

$$\begin{aligned} TC - E(\widetilde{X}^{gross}) = E(\widetilde{X}^{in}) - E(\widetilde{X}^{gross}) + CoC \cdot SCR + P^{re} = \\ E(\widetilde{X}^{re}) + CoC \cdot SCR + P^{re} = E(\widetilde{X}^{re}) + CoC \cdot SCR + E(\widetilde{X}^{re}) \cdot (1 + \lambda_{re}) = \\ \lambda_{re} \cdot E(\widetilde{X}^{re}) + CoC \cdot SCR \end{aligned} \quad (5.2)$$

where the term $\lambda_{re} \cdot E(\widetilde{X}^{re})$ is the profit ceded to the reinsurer. The minimization of the total cost can be expressed as the maximization of the Economic Value Added (EVA) as follows:

$$
\begin{aligned}
EVA = Profit - CoC \cdot SCR = \\
&= \lambda_{total} \cdot E(\widetilde{X}^{gross}) - \lambda_{re} \cdot E(\widetilde{X}^{re}) - CoC \cdot SCR = \\
&= (1 + \lambda_{total}) \cdot E(\widetilde{X}^{gross}) - TC
\end{aligned}
$$

The limitation provided by the total cost approach is that the variability of the loss around the mean is neglected. In fact, the SCR alone can only describe the extreme events, without giving enough emphasis to the volatility of the results. Also, with this approach the SCR becomes a variable itself that depends on the optimization, while in practice it is chosen in advance according to the risk appetite of the company.

For these reasons, an alternative approach is proposed in the next section.

5.1.2 Constrained Reinsurance Optimization

The aim of the proposed approach is the definition of an optimization criteria that takes into account and connects the most crucial performance indexes and the most realistic assumptions altogether. First of all, a clear distinction should be made between the indicators that we wish to optimize and indicators that must be fixed a priori:

- as mentioned before, the SCR is usually a target of the company that must be fulfilled through reinsurance. Therefore considering it as a parameter to be optimized may be inappropriate from the practical point of view;
- Companies fix their profit target with the use of many indicators: the ROE is one of them. The idea is to fix a certain level of profit under which the reinsurance shouldn't go. In this way the optimization will deal only with the reinsurance structures that satisfy the constraint, and subsequently the opportune risk-profit trade-off will be selected;
- the volatility is a good example of indicator to be minimized in order to obtain a good reinsurance structure. Of course, since the standard deviation is proportional to the volume of the business, a relative index like the Coefficient of Variation CoV is preferred;

- the presence of Non-proportional reinsurance may be driven by the existence in Solvency II of the NP reduction factor, which decreases by 20% the volatility factor for $MTPL$, $GTPL$ and $Fire$ lobs. In fact, even if the application of an Excess of Loss for one of these lobs may be inefficient, an insurer under Standard Formula would likely prefer to gain the benefits of the NP factor;

- when dealing with Quota Share contracts, the insurer may decide to retain at least a percentage s of its business, and avoid considering the contracts that transfer more than $s\%$;

- a delicate point is the criteria for splitting the layer of a Non-proportional reinsurance in order to diversify the risk transfer among more parties. In practice this component is mainly driven by the market relation between the insurer and the reinsurer. An estimate of the additional costs due to the split of the layer may be incorporated inside the optimization to diagnose the benefits and the drawbacks of such decision.

5.2 Pricing Reinsurance

What does it make a reinsurance contract worth buying? The biggest driver is its usefulness and how it can help the insurance company to achieve the adequate risk appetite. The other obvious driver is the price: even if the contract is very suitable to reach the management goals in terms of risk, a too high reinsurance loading would push down the expected profit of the company. In the results presented in this thesis, a huge variety of reinsurance contracts will be compared and, for this purpose, a well-functioning pricing principle is needed.

This topic should always be treated with white gloves, since there is often a significant gap between academic theory and practice. Several components are taken into account to build the price of a reinsurance treaty, and the main ones are:

- the expected losses $E(\widetilde{X}^{re})$;

- the uncertainty of $\widetilde{X}^{re}$;

- charges for reinsurer's expenses and profits;

- costs for the brokerage.

A premium principle π is a function that takes as argument a random variable $\widetilde{X}$ and returns the loaded premium. In general, a loaded premium can be expressed as:

$$\pi(\widetilde{X}^{re}) = E(\widetilde{X}^{re})(1 + \lambda) = P(1 + \lambda) \tag{5.3}$$

where P^{re} is the reinsurance pure premium and λP^{re} the safety loading.

Expressing the loaded premium as a percentage λ of the pure premium is a good way to communicate clearly the impact of the safety loading component. Under this framework, the difference between the various premium principles is the methodology used to compute λ.

In both the insurance and reinsurance context, there are several reasons behind the presence of the safety loading:

- the safety loading represents the expected profit of the contract, which is needed to accomplish company's growth goals.

- in the expected utility framework the pure premium is sufficient only for a risk neutral insurer/reinsurer. Since both insured and insurer/reinsurer are risk averse, the safety loading is required to perform the transfer of risk;

- according to the ruin theory, if no safety loading is present, the ruin will occur with certainty;

Since a strong economic component influences the price, it isn't a easy task for the direct insurer to forecast with accuracy how much a reinsurance plan will cost. Obviously, if only few treaties are slightly modified from one year to another, the plan's premiums are easier to predict. Usually, only reinsurers and reinsurance brokers have at disposal consistent data to build fitting price curves. This type of information has a significant relevance in the decision-making process of an insurance company, because knowing in advance the reinsurance price allows the entity to compute realistic risk and profit indicators, that describe in depth the economic situation. This type of information is the key to perform a good reinsurance optimization.

A crucial information used in practice to price reinsurance contracts is the portfolio's structure of the reinsurer: many pricing procedures are based on the allocation of capital in the various line of business. To fulfil management goals, the reinsurer applied a suitable premium, that depends on how the new contract diversifies in the current portfolio.

It is important to note that issues in pricing reinsurance are mainly regarding non-proportional reinsurance since the uncertainty of the risk is high, and the loading made by the reinsurer may vary according to the risk appetite and market conditions. For example, catastrophe reinsurance's price rises in a significant way after the occurrence of a relevant natural catastrophe. For simplicity, this type of price uncertainty won't be considered in the calculations.

In absence of this data, other analytical approaches can be applied in the attempt to approximate the complicate price making process. Every method will have a common uncertainty for the user: the loading component λ depends on a parameter, which can be highly subjective in few approaches. Therefore, even if the calculation rule is set, the choice of the parameter drives the result.

The following pricing principles can be used for both insurance and reinsurance contracts. In practice, these principles are not realistic for insurance premiums, since the market competition on prices is very relevant in the pricing process.

Usually in literature, when non-proportional reinsurance premiums are computed, the expenses are not considered in the calculations. Therefore, from now on, the non-proportional reinsurance tariff premium B will be expressed as

$$B = \pi(\widetilde{X})$$

5.2.1 Expected Value Principle

Known to be the most straight-forward premium principle, it is given by

$$\pi(\widetilde{X}^{re}) = P^{re}(1 + \lambda).$$

The choice of λ is very critical because no information about the variability of $\widetilde{X}$ is considered. Thus, the safety loading λP assumes only a mere economic meaning. Given the fact that in the thesis the same premium principle will be applied to many contracts, if λ is fixed and equal for every tested case, the safety loading would be insufficient for contracts covering only the most extreme scenarios. In fact, excess of loss contracts with a very high deductible D are characterized by a low expected value and a high volatility of claims. The total lack of flexibility of this pricing principle towards the reinsurance parameter's changes is a huge downfall when applied to non-proportional reinsurance.

This type of approach is more reasonable in the traditional insurance context, where the pricing is more based on competitive market strategies.

As mentioned before, the safety loading as percentage of the pure premium gives a clear idea of its size. For this reason, when numerical results will be presented in the next chapters under other premium principles, the safety loading will be expressed in the form λP.

5.2.2 Standard Deviation Principle

According to this principle, the reinsurance premium $\pi(\widetilde{Z}_i^{re})$ for the $i-th$ single contract is proportional to the standard deviation of the risk linked to the contract, such as:

$$\pi(\widetilde{Z}_i^{re}) = E(\widetilde{Z}_i^{re}) + \beta\sigma(\widetilde{Z}_i^{re}) \tag{5.4}$$

where β is a parameter chosen by the user. Note that, the aggregate premium paid for the reinsurance treaty is given by:

$$\pi(\widetilde{X}^{re}) = \sum_{i=1}^{N} \pi(\widetilde{Z}_i^{re}) \tag{5.5}$$

where N is the number of policies in the cedant's portfolio.

Under the assumption that $\widetilde{Z}_i^{re}$ are independent and identical distributed, we have:

$$\pi(\widetilde{X}^{re}) = \sum_{i=1}^{N} \pi(\widetilde{Z}^{re}) = N\pi(\widetilde{Z}^{re})$$

In literature, this approach is applied at portfolio level, where, instead of considering the single claim $\widetilde{Z^{re}}$ to price, the aggregate claim amount $\widetilde{X}^{re}$ is used:

$$\pi(\widetilde{X}^{re}) = E(\widetilde{X}^{re}) + \beta\sigma(\widetilde{X}^{re}) \tag{5.6}$$

Obviously, the parameter β assumes a lower value in comparison to the formula 5.4.

When the standard deviation principle is applied using $\widetilde{Z}^{re}$, we are not considering any diversification benefit of the portfolio. But, is it necessarily a bad thing? We have to remember that, when risks are transferred from the direct insurer to the reinsurer, the latter's portfolio may be already well diversified by other contracts signed with other companies. For this reason, it may not be coherent to consider the diversification of the standalone portfolio when pricing the treaty.

In both cases, the standard deviation principle has some drawbacks. The most important for our application is that the standard deviation measures the volatility of both left and right tail of the claims distribution. This fact may create incoherence in the reinsurance context because of particular distribution found in non-proportional treaties. In fact, for high deductibles D, both $\widetilde{Z}^{re}$ and $\widetilde{X}^{re}$ are at the same time continuous (over $\mathbb{R}^+$) and discrete (in 0). Moreover, in case a limit L is present, the distribution of $\widetilde{X}^{re}$ becomes "spiky", due to $\widetilde{Z}^{re}$ being discrete also in L. The standard deviation may not be the best indicator to describe this peculiar behaviour.

This premium principle can create some difficulties in case of reinstatements: the risk is composed also by the stochasticity of the reinstatement premiums, which depend on the loaded initial premium.

Therefore, the loaded premium $\pi(\widetilde{P}^L_{AD})$ is itself a random variable. Recalling the notation of equation (1.20), it is determined by solving the following equation ([26]):

$$E(\widetilde{P}^L_{AD}(1+\widetilde{R}) - \widetilde{X}^{re}) = \beta \cdot \sigma(\widetilde{P}^L_{AD}(1+\widetilde{R}) - \widetilde{X}^{re}) \tag{5.7}$$

which can be rewritten as

$$(\widetilde{P}^L_{AD})^2 \cdot \left[E(1+\widetilde{R})^2 - \beta^2\sigma^2(1+\widetilde{R})\right] +$$
$$2\widetilde{P}^L_{AD} \cdot \left[\beta^2 Cov(1+\widetilde{R}, \widetilde{X}^{re}) - E(1+\widetilde{R})E(\widetilde{X}^{re})\right] + \left[E(\widetilde{X}^{re})^2 - \beta^2\sigma^2(\widetilde{X}^{re})\right] = 0$$

where

- $\widetilde{R} = \frac{1}{L}\sum_{k=1}^{K}(c_k E(\widetilde{r}_{k-1}))$ is the reinstatement cost as percentage of the initial premium $\widetilde{P}_{AD}^{L}$;

- the terms in parenthesis can be seen respectively as a, b and c of second order equation, which can be solved with respect to $\widetilde{P}_{AD}^{L}$.

The main problem is that, under the standard deviation principle, the system may not have any solution. This fact may become a significant limitation when this type of contract is studied. But, if the solution exists, this pricing principle provides interesting results. In fact, the standard deviation principle is subadditive:

$$\sigma(X_1 + X_2) \leq \sigma(X_1) + \sigma(X_2). \tag{5.8}$$

which comes handy when splitting layers into smaller sub-layers.

Imagine we are in front of two scenarios:

1. the scope is to price two consecutive layers for a non proportional reinsurance;

2. the scope is to price the combined layers altogether;

Due to the inequality, using the standard deviation premium principle and the same parameter β, the sum of the loaded premiums of case 1 is greater than the loaded premium of case 2. From an economic point of view, it means that it is more convenient to choose only one reinsurer instead of two to transfer a risk. At the same time, when two or more reinsurers are involved, the counterparty default risk is reduced. Solvency II framework takes into account this diversification and reduces the capital requirement when the reinsured risk is split among more entities.

5.2.3 Variance Principle

Similar to the previous case, the variance premium principle is defined as:

$$\pi(\widetilde{Z}_i^{re}) = E(\widetilde{Z}_i^{re}) + \beta\sigma^2(\widetilde{Z}_i^{re}). \tag{5.9}$$

Due to the variance being in a larger scale than the standard deviation, the parameter β is often very small. For this reason, the variance premium principle is very unstable to price reinsurance contracts, which are characterized by risks with a significant variance.

The variance holds the opposite relation with respect to the standard deviation:

$$\sigma^2(\widetilde{X}_1 + \widetilde{X}_2) \geq \sigma^2(\widetilde{X}_1) + \sigma^2(\widetilde{X}_2). \tag{5.10}$$

Hence, it means that if a layer is split in infinite parts, the safety loading tends to zero. This property is usually not desired, especially when pricing consecutive layers.

5.2.4 PH transform

The Proportional-Hazard transform is an interesting pricing principle proposed by Wang [29]. It recalls the stop-loss transformation of equation (1.8) and introduces an exogenous index $\rho \geq 1$, such that:

$$\pi(\widetilde{X}^{re}) = \int_0^\infty (1 - F_{\widetilde{X}_{re}}(x))^{1/\rho} \, dx. \tag{5.11}$$

Trivially, with $\rho = 1$ the stop-loss transformation is obtained.

The result of this formula can be seen as a risk-adjusted premium based on the exact shape and heaviness of the distribution tail. Knowing the cumulative distribution function $F_{\widetilde{X}_{re}}(x))$ is crucial, and in practice it can be estimated through the use of the methodologies explained in the previous chapter. The user could use simulation to estimate this element, or, way better, use the Panjer algorithm. As discussed before, the Panjer algorithm is particularly used in reinsurance scenarios due to its properties. All we need is the distribution of $\widetilde{N}$ and $\widetilde{Z}$ to obtain the discrete $F_{\widetilde{X}_{re}}(x))$. The use of simulations might be a good choice when considering risk structures that are not easy to implement through Panjer algorithm. A good example is the pricing of an Umbrella cover, where the lobs covered are correlated according to a copula structure.

In case there is a strong uncertainty around the estimated distributions of frequency and severity, due to the lack of historical data, a good idea would be to use a higher value of ρ to compensate.

The strength of this pricing principle is that, in case contractual limitation of the Excess of Loss are present (limit, deductible, aggregate limit, aggregate deductible and reinstatements), it is able to consider properly the underlying risk once the cumulative distribution function is estimated. Unlike the standard deviation and variance principles, the PH transform overcomes these difficulties and returns concrete risk-based results.

In case of reinstatements, by recalling equation (1.24),since $\widetilde{r}$ is comonotonic, Mata [20] proposes to use the PH-transform as follows:

$$\pi(X_{AD}^{L}) = \frac{\pi(\widetilde{X}_{AD}^{AL})}{(1 + \frac{1}{L}\sum_{k=1}^{K} c_k \pi(\widetilde{r}_{k-1}))} \tag{5.12}$$

Like the standard deviation principle, the PH-transform is subadditive.

Chapter 6

A reinsurance simulation model

In this chapter we will set up everything we need to perform the comparison between different reinsurance treaties. To do so, a frequency-severity model is implemented, with some additional features in order to overcome difficulties that appear when dealing with the reinsurance world. The whole topic will be explained also with computational programming tips to give a full overview to the reader, that is important to create a bridge between theory and calculations. The reference programming tool is *R*.

6.1 Building the simulation's environment

The choice of not using real data for the analysis has been made, and it is important to specify the underlying reasons: a well-functioning reinsurance tool must be comprehensive of the whole business model of the insurance company to achieve reasonable results. Therefore, in case a frequency-severity approach is chosen, it is crucial to estimate properly the risk distributions (of $\widetilde{N}$ and $\widetilde{Z}$) and the dependency structure. Curve fitting is a very delicate task, and, since it is very easy to model wrongly the risk, the reinsurance model would probably lead to unreliable results. The main focus of this thesis is the investigation of reinsurance applied in different contexts, and, the use of real data would shift the whole attention from the main topic to issues regarding fitting of parameters and distributions.

For these reasons it is better to skip the fitting procedure, and propose scenarios where we assume that the risk has already been modelled correctly.

The case study considers a fictitious insurer that operates in only three lines of business:

- Motor Own Damage (*MOD*), where the severity distribution is characterized by a small mean and volatility;

- General Third Party Liability (*GTPL*), known to be a *"long tail"* business (where both the expected value and standard deviation are significantly high);

- Motor Third Party Liability (*MTPL*), that can be seen in the middle between the previous two lines of business.

In practise, due to market demand, the biggest share of insurance portfolios is dedicated to MTPL business, followed by MOD, and ultimately by GTPL. This fact is taken into account in the choice of the respective frequency parameters.

Since only premium risk will be considered in this thesis, it is important to mention that the results presented are not fully representative of the insurer's situation. Obviously this choice was made to simplify the topic and to avoid to over complicate the approach. In case, the implementation of additional risks in this model is still feasible within the Solvency II framework.

Let's check out the general direction taken for the calculation, the problems that arise and the solutions applied.

All the procedures are made to determine the distribution of the aggregate claim amount $\widetilde{X}$ of the company in the year. Since in our scenario this fictional company has insurance risks originated by three different lines of business, $\widetilde{X}$ is given by:

$$\widetilde{X} = \widetilde{X}_{mtpl} + \widetilde{X}_{gtpl} + \widetilde{X}_{mod}$$

Therefore, it is needed to estimate the $\widetilde{X}$ for each LoB and associate a proper dependence structure through copula. Different types of copula will be treated to remark how much this choice is delicate. As explained in the previous chapter, in order to effectively use a copula aggregation, the distributions of the marginals $\widetilde{X}_{mtpl}, \widetilde{X}_{gtpl}, \widetilde{X}_{mod}$ must be estimated. A good way to estimate them is to use simulations, such that:

1. for the single simulation a value n is sampled from the distribution of $\widetilde{N}_{LoB}$;

2. n realisations are sampled from the distribution of $\widetilde{Z}_{LoB}$;

3. obtain $\widetilde{X}_{LoB} = \sum_{i=1}^{n} \widetilde{Z}_{i,LoB}$;

4. repeat the first three steps for the number of simulations desired;

5. repeat the steps above for each line of business.

The procedure seems really feasible explained in this way, but we are not considering the reinsurance: if a Excess of Loss treaty is applied, the single claim cost realisations $Z_{i,LoB}$ of step 2 are affected by the generic *Layer* operator. Then, on step 3 we would have $\widetilde{X}_{LoB} = \sum_{i=1}^{n} Layer(\widetilde{Z}_{i,LoB})$. Imagine how long it would be to repeat this simulation procedure for each possible configuration of an Excess of Loss treaty for each LoB! The obvious solution to this issue is to keep saved the same simulations and apply the desired Layer operators when needed. But here arises a problem from the computational point of view: storing every single claim $Z_{i,LoB}$ for each simulation takes too much computer memory. For those who didn't know, if the stored data is too much heavy, even simple computations will become slow.

One of the main goals in the architecture of this model is an adequate data structure. This point is crucial: since this model is based on simulations, if our attempt is to study the different effects of many reinsurance covers, we may change the parameters involved in the contracts many times. The solution is to generate less simulations and store information in a smarter way, instead of doing more simulations due to a lacking data structure.

Since the two main contracts applied in this thesis are the Quota Share and variants of the Excess of Loss, a smart solution would be to consider that:

- Quota share can be applied equivalently to $\widetilde{X}_{LoB}$, instead of $\widetilde{Z}_{i,LoB}$. Therefore, for this contract we don't need to store any single claim;

- Excess of Loss affects only those claims $Z_{LoB} > D$, where D is the deductible.

Hence, if we assume for example that the set of Excess of Loss treaties used in the thesis have at least $D > threshold$, we would need to store for individually only those claims greater than the threshold, and the remaining small claims as a single aggregate value. In this way, we may pass from storing thousands of values for each simulation, to a small number like 10.

For those interested in the practical point of view, this storing process can be achieved easily through the use of *lists*. Lists are objects which contain elements of different types like numbers, strings, vectors and another list inside it. We are interested especially in using a list containing lists.

The setup is the following:

$X_{j,LoB}$ is the aggregate claim amount value provided by the *j-th* simulation;

$Z_{i,j,LoB}$ is the $i-th$ single claim amount value provided by the *j-th* simulation;

$j = 1, 2, \ldots, J$ where J is the total number of simulations;

$i = 1, 2, \ldots, I$ where I is the total number of single claims generated in the simulation and depends on the sampling from the frequency distribution

Each LoB's simulations will be stored in a separated list. The *j-th* element of each of these lists is a list itself, containing info on $X_{j,LoB}$. It is composed by two components:

1. a vector containing all the single claims $Z_{i,j,LoB} > threshold$
2. a number representing the sum of of all the other claims $Z_{i,j,LoB} \leq threshold$

In this way all the small single claims are compressed in only one number, and we are able to spend way less computer memory. $X_{j,LoB}$ is obtained by summing the two components. Behind this structure lies a wonderful feature: if an Excess of Loss treaty is present, the corresponding Layer operator is applied to all the numbers contained in the first component, and then by summing the two components we obtain $X_{j,LoB}$.

With this data structure we have the chance to retrieve the simulated distribution of $\widetilde{X}_{LoB}$ for every Deductible $D > threshold$.

The general approach used in practise by actuaries is very similar to the one described here, but with few simplifications. The single claim cost $\widetilde{Z}$ is studied in two different components:

- *large* claims, which are those claims with value higher than a chosen threshold. The approach is to describe them with a very skewed distribution and simulate them one by one with a frequency-severity approach;

- *attritional* claims, which are the remaining small amounts. They are often simulated directly using an aggregate distribution $\widetilde{X}_{attritional}$. Hence, it happens that the distribution of attritional claims is fitted on just few yearly observations, and therefore, the fit may be lacking of accuracy. In general, attritional claims are studied as a residual part.

The main issues that may appear in this kind of framework are linked to the possible underestimation of risk coming from attritional claims. In fact, since these kind of claims are usually not affected by any non-proportional reinsurance treaty, it can happen that an unexpected huge number of small claims occurs in the year, and the reinsurance plan becomes almost useless. Therefore, one possible risk in this "*attritional-large*" framework is to overestimate the protection provided by non-proportional reinsurance over the covered portfolio.

6.2 Model parameterization

The parameters involved in the calculation are based on the Italian insurance market. The information are provided by ANIA's statistical appendix and publications.

To compute the safety loading coefficient λ and the expense coefficient c, a weighted average of the last 5 years has been performed in the computation of the Combined Ratio CR and Expense Ratio ER, such that:

$$CR = \frac{\sum_{i=1}^{5} B_{t-i} \cdot CR_{t-i}}{\sum_{i=1}^{5} B_{t-i}}$$

$$ER = \frac{\sum_{i=1}^{5} B_{t-i} \cdot ER_{t-i}}{\sum_{i=1}^{5} B_{t-i}}$$

where t is the current year, of which the CR and ER has not been observed yet. The CR in the calculation is not comprehensive of the run-off since the focus is set on the premium risk, and not the reserve risk.

The ER is defined as $\frac{E}{B}$, where E are the expenses, assumed deterministic. By definition ER is equal to the expense coefficient c, since by model assumption we have $c \cdot B = E$.

The Loss Ratio LR is equal to the difference $CR - ER$ or equivalently by $\frac{X}{B}$. In this case, the safety loading coefficient λ that it's present in the model is referred to the pure premium $P = B \cdot \frac{1-c}{1+\lambda}$, and therefore not equal to $(1 - CR)$, which is proportional to B. The following equation should be solved:

$$\lambda \cdot P = (1 - CR) \cdot B$$

Then

$$\lambda = (1 - CR) \cdot \frac{B}{P} = \frac{1 - CR}{LR}.$$

For the estimation of $\sigma(\widetilde{Q})$, the standard deviation of the Loss Ratios of the last 15 years has been considered. Assuming that the Italian insurance market is enough diversified, and recalling equation 3.4 and its asymptotic properties, we have that:

$$\lim_{n\to\infty} \sigma(\widetilde{LR}) = \lim_{n\to\infty} \sigma(\frac{\widetilde{X}}{B}) = \lim_{n\to\infty} \sigma(\frac{\widetilde{X}}{P}) \cdot \frac{P}{B} = \lim_{n\to\infty} \sigma(\widetilde{Q}) \cdot \frac{P}{B}$$

and, therefore,

$$\sigma(\widetilde{Q}) = \sigma(\widetilde{LR}) \cdot \frac{B}{P} = \sigma(\widetilde{LR}) \cdot \frac{1+\lambda}{1-c}$$

In terms of tariff premiums B, this insurer represents the $10 - th$ biggest Italian insurer for each lob. Therefore, given the distribution of premiums around the market, we are in front of a medium-sized insurer that operates in three lobs only.

As mentioned before, the *MTPL* lob is the most present in the portfolio, but at the same time, due to high competitiveness in the market, the safety loading coefficient λ assumes only $1,2\%$. Even if MTPL composes the $63,2\%$ of the tariff premiums B, the resulting profit is very limited with respect to the other lobs.

Table 6.1: Parameters for each line of business

	MTPL	*GTPL*	*MOD*
$E(\widetilde{Z})$	4.500	6.000	1.500
$CoV(\widetilde{Z})$	6	10	2
Policy limit	10.000.000	10.000.000	1.000.000
$E(\widetilde{N})$	50.000	10.000	30.000
$\sigma(\widetilde{Q})$	6,83%	12,37%	11,27%
λ	1,2%	6,7%	13,8%
E(Profit)	2.700.000	4.020.000	6.210.000
c	21.2%	32.3%	30.4%
B	288.959.391	94.564.254	73.577.586
% B	63,2%	16,1%	20,7%

The $GTPL$ is characterized by the higher $E(\widetilde{Z})$ and variability, and at the same time, it's the least diversified due to low number of claims.

From a risk perspective, the λ_{MOD} is too high considering that the coefficient of variation CoV of the single claim $\widetilde{Z}$ is the lowest observed.

The overall safety loading coefficient $\lambda_{total} = 3,92\%$, where

$$\lambda_{total} = \frac{\lambda_{MTPL}P_{MTPL} + \lambda_{GTPL}P_{GTPL} + \lambda_{MOD}P_{MOD}}{P_{MTPL} + P_{GTPL} + P_{MOD}}.$$

It is interesting to observe how risk theory literature and market's supply and demand clash each other, leading to a situation where the profit of MOD often supports the underwriting result of the other lobs inadequately priced.

The policy limit has been introduced for each lob to avoid encountering unreasonable results within the simulations. In fact, in practice, the single claim is capped to the limit specified in the contract, which changes across the market. The limit has been fixed in such a way that it is representative of the market and, at the same time, doesn't cut off the whole behaviour of the tail. Building a reinsurance treaty will be easier under a single claim cost $\widetilde{Z}$ with a limited domain.

Now we have everything we need to start the Monte Carlo simulations for each lob as explained in the previous paragraph, in order to keep track of large claims. The densities of $\widetilde{X}$ distributions for the three lobs are presented in Figure (6.1).

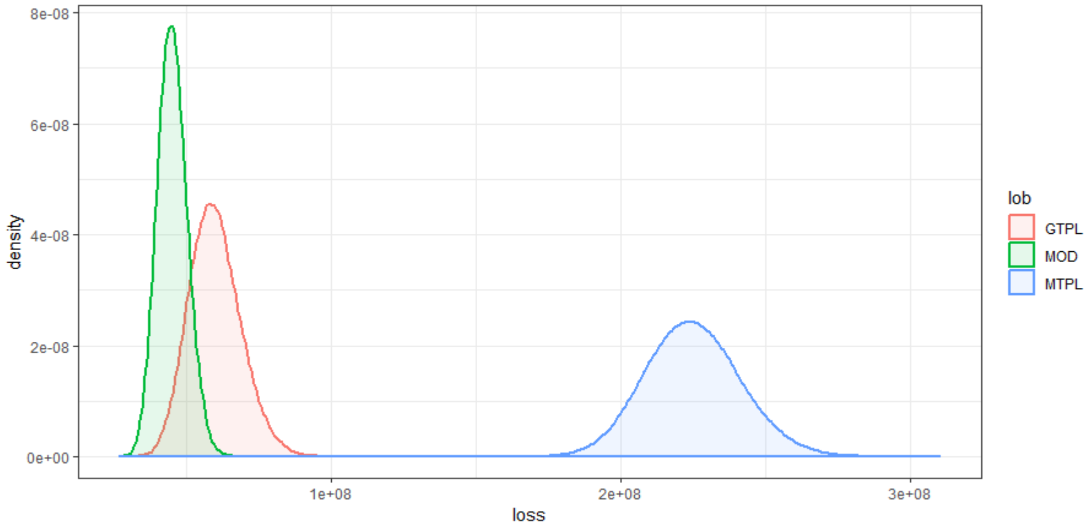


Figure 6.1

As expected, the *MTPL* is far from the other two lobs due to the dominance in terms of dimension. One should not be tricked graphically to think that *MTPL* is more risky also in relative terms. The main characteristics of the three distributions are presented:

	$MTPL$	$GTPL$	MOD
$E(\widetilde{X})$	224.853.164	59.755.974	44.991.474
$\sigma(\widetilde{X})$	16.366.551	8.880.263	5.119.400
$CoV(\widetilde{X})$	7,28%	14,86%	11,38%
$\gamma(\widetilde{X})$	0,158	0,373	0,229
$P(\widetilde{X} < P(1+\lambda))$	57,89%	70,34%	88,51%
$VaR99,5\%$	44.229.873	25.781.117	14.232.375
$TVaR99\%$	46.154.191	26.935.695	14.813.923

Table 6.2: simulation's summary

The most interesting results presented in the table are:

- CoV is lower for $MTPL$ due to a lower $\sigma(\widetilde{Q})$. Therefore, in relative terms, the $MTPL$ is less variable with respect to the other lobs;

- the probability of observing a profit in the lob is too low in the $MTPL$ case and too high inMOD case. Instead, $GTPL$ is priced coherently to the underlying risk;

- the Value-at-Risk for $MTPL$ is the highest due to the higher standard deviation; even if near in mean, $GTPL$ requires a significant greater capital requirement than MOD.

What's left now is the aggregation of these simulations in such a way that the underlying correlation is present.

6.3 Copula Aggregation

Coming back to our framework, the set $(X_{j,Mtpl}, X_{j,Gtpl}, X_{j,Mod})$ is uncorrelated. The copula aggregation is what is needed to create correlated sets of aggregate claims amount. But here a huge problem arises again: we would lose information in the attempt of applying a copula. Let's investigate what it is happening.

As explained in the previous chapters, the output of J copula simulations in this scenario would be a matrix with J rows and 3 columns. Each column represents a line of business and each row a triad of correlated simulations. Each cell of this matrix contains a number between 0 and 1, that can be seen as a cumulative probability.

Table 6.3: Copula simulation example

Simulation j	MTPL	GTPL	MOD
1	$e_{1,MTPL}$	$e_{1,GTPL}$	$e_{1,MOD}$
2	$e_{2,MTPL}$	$e_{2,GTPL}$	$e_{2,MOD}$
3	$e_{3,MTPL}$	$e_{3,GTPL}$	$e_{3,MOD}$
...	...	...	...
J	$e_{J,MTPL}$	$e_{J,GTPL}$	$e_{J,MOD}$

Table 6.4: Inverse CDF

Simulation j	MTPL	GTPL	MOD
1	$F^{-1}_{MTPL}(e_{1,MTPL})$	$F^{-1}_{GTPL}(e_{1,GTPL})$	$F^{-1}_{MOD}(e_{1,MOD})$
2	$F^{-1}_{MTPL}(e_{2,MTPL})$	$F^{-1}_{GTPL}(e_{2,GTPL})$	$F^{-1}_{MOD}(e_{2,MOD})$
3	$F^{-1}_{MTPL}(e_{3,MTPL})$	$F^{-1}_{GTPL}(e_{3,GTPL})$	$F^{-1}_{MOD}(e_{3,MOD})$
...	...	...	...
J	$F^{-1}_{MTPL}(e_{J,MTPL})$	$F^{-1}_{GTPL}(e_{J,GTPL})$	$F^{-1}_{MOD}(e_{J,MOD})$

The next step is to apply to the generic element $e_{j,LoB}$ the inverse of the cumulative distribution function $F^{-1}_{LoB}(e_{j,LoB})$ of the respective LoB to obtain the associated quantile.

Now, this table is representing values in the scale of measure of $\widetilde{X}$. Note that those who are interested in determining the capital requirements gross of reinsurance would be fine with this situation, because at this step they would have everything they need to proceed. The problems arise in the case reinsurance is treated: we had just lost simulation data about the structure of the aggregate claim amount. In particular, we are not able to retrieve in any way information about the large losses that lie behind each $F^{-1}_{LoB}(e_{j,LoB})$.

Solution 1: One straightforward solution to the retrieve the lost information is the following:

1. build the simulated cumulative distributions function for each lob with the desired combination of Excess of Loss treaties;
2. simulate from the copula as in table (6.3);
3. use the distributions obtained at step 1 to apply the inverse operation like in table (6.4).

This solution has too many flaws that must be discussed.

- First, this approach is not flexible in terms of computing times, since the procedure would be repeated from step 1 for every possible combination of non-proportional reinsurance. Not the best option if we want to study a large number of contracts.

- Second, non-linear correlation is assumed between gross of reinsurance $\widetilde{X}$, and the application of Non-proportional reinsurance affects such correlation. For example, assume that the correlation is described by a Gumbel Copula, which emphasises dependency in both left and right tail of the marginals. If the direct insurer decides to apply non-proportional reinsurance to each lob of its business, then the correlation between the claims net of reinsurance would be reduced, since the right tails are now partially protected. This fact is crucial for step 2 of the previous procedure: even if the dependency structure gross of reinsurance is assumed to be known, it is not possible to determine the new dependency after the application of non proportional reinsurance. Supposing that the correlation doesn't change, the actuary would overestimate the risk, or on another hand, would underestimate the benefits of the reinsurance plan.

Solution 2: Another approach is proposed since more suitable. The trick is to use at the fullest the simulation information stored in that particular list discussed before. Remember that, since this is a simulation approach, a bias will always be present and the results obtained should be treated as approximations of the scenario proposed. As the number of simulations increases, both the accuracy of this proxy and the time required by the machine increase.

The method involves a particular sorting of the elements contained in the lists of every lob of the company. As mentioned before, by summing the vector and the attritional value contained in each element, a simulated value of X is obtained. The concept is the following: reorder the simulated values X in all the three lists in such a way that the desired dependency structure is generated between the lists.

How could we perform such permutation?

The simulated cumulative distribution function of $\widetilde{X}_{MTPL}, \widetilde{X}_{GTPL}, \widetilde{X}_{MOD}$ is determined through the sum previously described. Then, we can associate each quantile to the simulated output of the Copula, which is composed by correlated $Uniform(0,1)$. Obviously, the copula results are expressed with more decimal numbers than the simulated CDF. Therefore, each value simulated by the copula is rounded until there are the same decimal numbers for both copula and CDF. For the sake of accuracy, it is also important to avoid replications inside each column of the copula (Table 6.3), which happens after the rounding up. It can also happen to obtain few zeros, which are definitively a problem since $F^{-1}_{LoB}(0)$ is not associated to any element of our lists.

A fast way to avoid the rounding up problems is the following:

- first thing, it is required that the length J of each list and of copula simulations is the same;
- the values of the copula are ranked by column. In this way each column contains $rank(e_{j,Lob})$ with support $1, 2, \ldots, J$;
- each simulated X is ranked too for each list. The mathematical support of the rank is the same of the previous step due to equal length J;
- each cell of the copula is substituted with the element (vector of large claims + attritional aggregate) which have the same rank.

At the end of this approach the user will obtain 3 reordered lists that are dependent according to the chosen dependency structure. The question is: which dependency structure?

The choice of using 3 lines of business is not casual: often in literature, when a copula aggregation is made, only 2 marginals are involved. The reason lies in the fact that, when more than 2 marginals are aggregated, Archimedean copulas can be used only under the assumption that the dependence has the same strength between all the marginals. Therefore, it is way easier to study a bi-variate case.

Keeping in mind that, according to Solvency II delegated acts, the correlation between the lobs is the one shown in Table 6.5, applying Archimedean copulas would be a difficult task. At this point *Vine copulas* play their role: we can extend any bi-variate copula to the 3-dimensional case.

6.3.1 Gaussian Copula

The Gaussian Copula is very easy to build in this tri-variate framework since we can simulate directly from it correlated uniforms that satisfy the correlation matrix shown before. This feature is one of the main reasons of its popularity.

But this copula isn't suitable to model tail dependency, which is the most common type of correlation present in financial and insurance risks. Due to its

Table 6.5: Correlation matrix

	MTPL	GTPL	MOD
MTPL	1	0, 5	0, 5
GTPL	0, 5	1	0, 25
MOD	0, 5	0, 25	1

simplicity, it was used back in 2007-2009 to model the credit risk in the banking sector. Many experts attribute part of the faults of the subprime mortgage crisis to this risks' underestimation. Obviously, there were many factors that contributed to that unfortunate event, but it is important to give adequate emphasis to this basic concept: underestimation of risks leads to distorted results, which will be used to make biased decisions at company's level.

Despite a possible underestimation of risk, the Gaussian copula is a good benchmark to compare the results obtained with more complicated Vine copula aggregations.

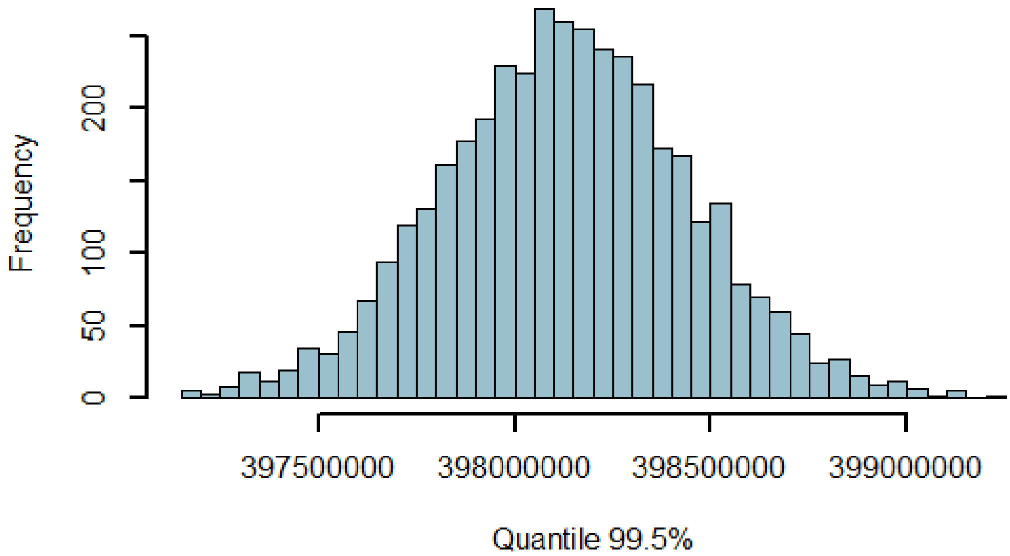
Adopting the correlation matrix of Solvency II, the empirical correlation between the uniform marginal of the copula are:

	MTPL	GTPL	MOD
MTPL	1	0, 483	0, 483
GTPL	0, 483	1	0, 239
MOD	0, 483	0, 239	1

It's important to note that the fact that the resulting correlation matrix is different from the one given as input is not a simulation error: in the definition of a Gaussian copula the Solvency II matrix is transformed through the so-called Cholesky decomposition. Therefore, the resulting correlations are not biased as it may seem.

A question that is interesting to analyze in the context of copula aggregations is the volatility of the results. Often, both in practice and literature, no diagnostic is performed to evaluate the variability of copula simulations. How much does the quantile move from one simulation to the other? In these paragraphs an answer will be provided and, consequently, the most appropriate copula simulation will be chosen to proceed in the comparison of the results.

The quantile at level $99,5\%$ of $\widetilde{X} = \widetilde{X}_{MTPL} + \widetilde{X}_{GTPL} + \widetilde{X}_{MOD}$ is plotted for each Gaussian copula simulation:



The results are presented in the following table:

min	5%	mean	95%	max
397.232.735	397.612.293	398.151.002	398.648.319	398.993.345

The percentage difference from the mean with level of confidence $(5\%; 95\%)$ is equal to $(-0,135\%; +0,125\%)$ and in absolute value $(-538.709; +497.316)$. The numbers confirm that the Gaussian copula simulations return stable results in relative terms, but it is always better to select a "central" simulation to obtain an unbiased capital requirement.

For stability purposes, the median simulation is selected and, recalling formula (4.10):

$$SCR = \widetilde{X}_{99,5\%} - E(\widetilde{X}) = 398.003.981 - 329.600.612 = 68.403.369$$

Note that the absolute interval of simulation results shown before, when compared to the $VaR_{99,5\%}$, is equal in percentage to the interval $(-0,788\%; +0,727\%)$.

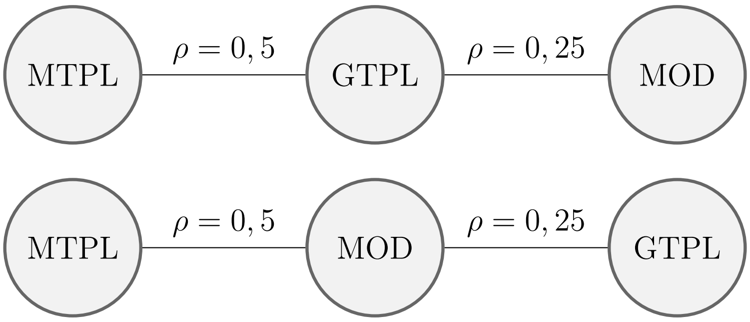
6.3.2 Vine Copula

Since the Vine Copula is composed by bi-variate copula, the order of aggregation is the main topic to be discussed. First of all, it is assumed that the lobs are correlated according to mirror Clayton copulas with a dependence measure comparable with the one provided by Solvency II. More precisely, the goal is to make the mirror Clayton copulas comparable with the Gaussian copula used in the previous paragraph. Under Gaussian copula, the Kendall τ is equal to $33,33\%$ when Spearman $\rho = 50\%$, and equal to $16,09\%$ when $\rho = 25\%$. The following relation connects the parameter θ of a Clayton Copula with the corresponding τ:

$$\theta = \frac{2 \cdot \tau}{1 - \tau}.$$

Hence, we obtain respectively $\theta = 1$ and $\theta = 0.3834$.

Since we are in a 3-dimensional scenario, the use of a C-Vine or D-Vine copula won't make any difference. All the possible starting structures of the Vine can be represented as:



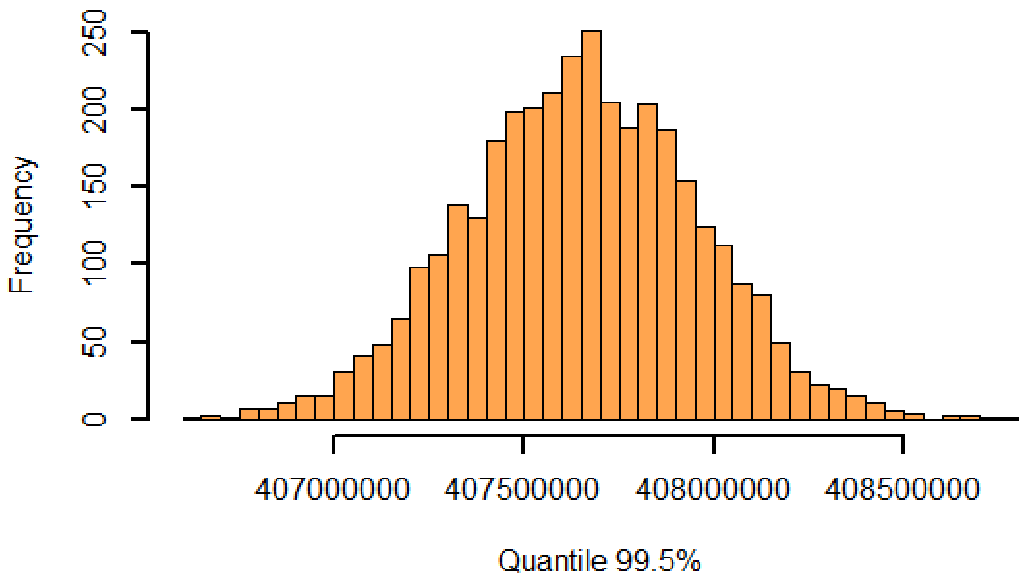
where the first and the second are equivalent in terms of underlying correlation. Therefore, the choice is restricted to only two aggregation structures.

Let's test empirically, according to each aggregation structure, how strong the correlations are between the vine copula's marginals.

1. If, as shown in the first structure, the first step is composed by the aggregation of the couples (*MTPL*,*GTPL*) and (*GTPL*,*MOD*), the resulting correlation matrix is:

   which is very similar to the Gaussian one, except for the correlation between *MTPL* and *MOD*. Since *MOD* is short tailed, the overestimation is very low.

	MTPL	GTPL	MOD
MTPL	1	0, 478	0, 523
GTPL	0, 478	1	0, 237
MOD	0, 523	0, 237	1



min	5%	mean	95%	max
406.718.483	407.165.576	407.649.081	408.125.925	408.517.815

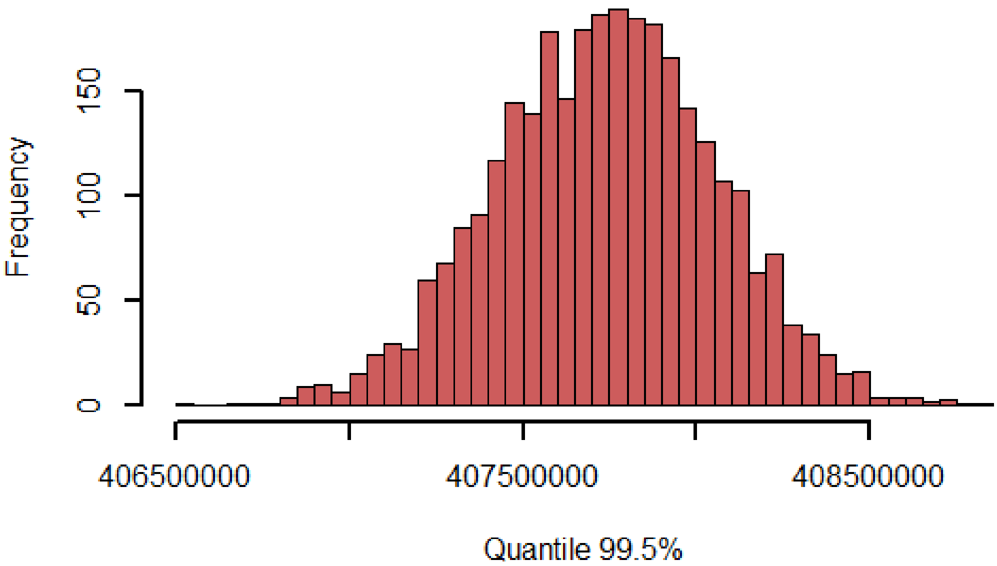
The results are presented in the following table:

The median simulation returns $SCR = 407.565.442 - 329.600.612 = 77.964.830$

2. Similarly, if the second structure is applied: the correlations, except for

	MTPL	GTPL	MOD
MTPL	1	0, 523	0, 478
GTPL	0, 523	1	0, 237
MOD	0, 478	0, 237	1

simulation error, are the same, but with *GTPL* and *MOD* switched. In this case the correlation between *MTPL* and *GTPL* is higher, which is not recommended due to risk overestimation.



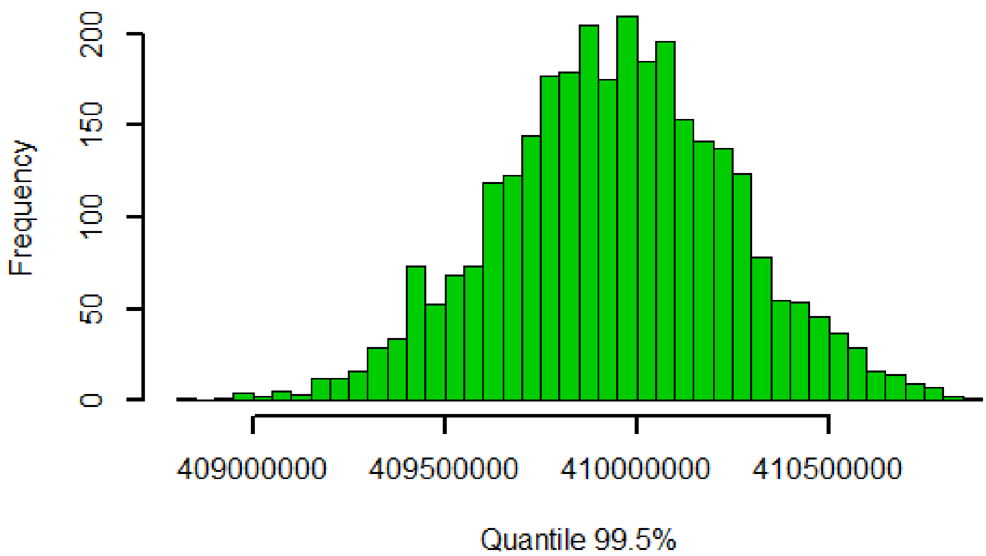
min	5%	mean	95%	max
409.029.100	409.463.085	409.934.218	410.395.067	410.784.235

with $SCR = 409.932.204 - 329.600.612 = 80.331.592$

3. The last case, instead, shows how a wrong aggregation order can create an unwanted correlation:

	MTPL	GTPL	MOD
MTPL	1	$0,478$	$0,478$
GTPL	$0,478$	1	$0,444$
MOD	$0,478$	$0,444$	1

In fact, the correlation of 25% between *GTPL* and *MOD* is overestimated in a significant manner. For this reason this aggregation structure is not considered.



min	5%	mean	95%	max
406.797.927	407.223.835	407.733.197	408.236.392	408.721.328

with $SCR = 407.732.206 - 329.600.612 = 78.131.595$

In conclusion, the first of the three orders is adopted for the vine copula aggregation since it seems to be the more comparable with the Gaussian copula case. In Figure (6.2) the correlation of the chosen vine copula is shown from the graphical point of view.

6.3.3 Clayton Copula

For comparison sake, a mirror Clayton copula will be simulated with underlying correlation $\rho = 50\%$ between all the lobs. Therefore, by recalling the previous results, the parameter θ of the tri-variate mirror Clayton is equal to 1.

The resulting correlation matrix is the following:

where the value $0,478$ is the same encountered in the previous tables.

As confirmed by the results, this type of mirror Clayton is characterized by higher quantiles and capital requirement:

with $SCR = 409.974.772 - 329.600.612 = 80.374.160$

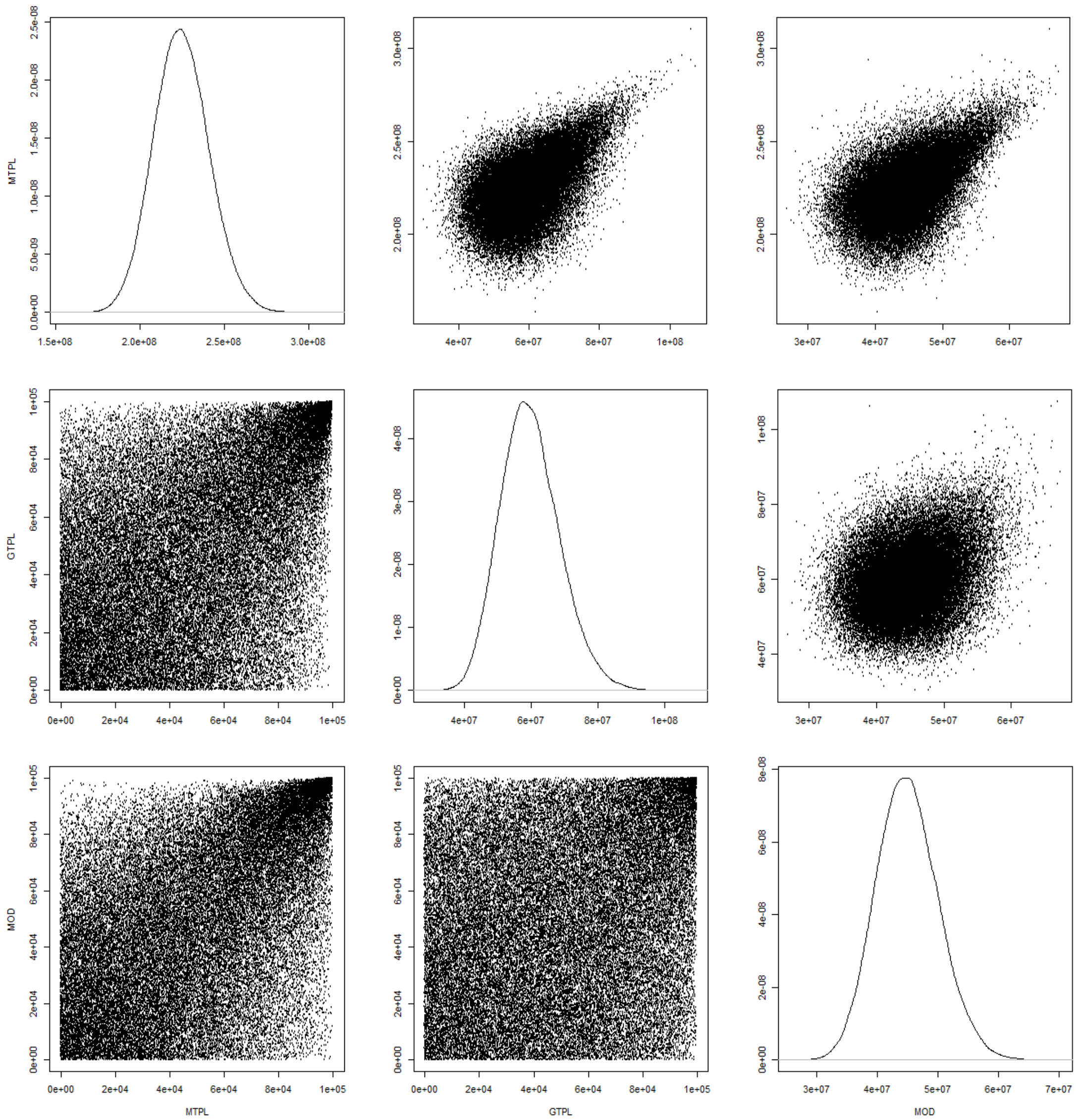
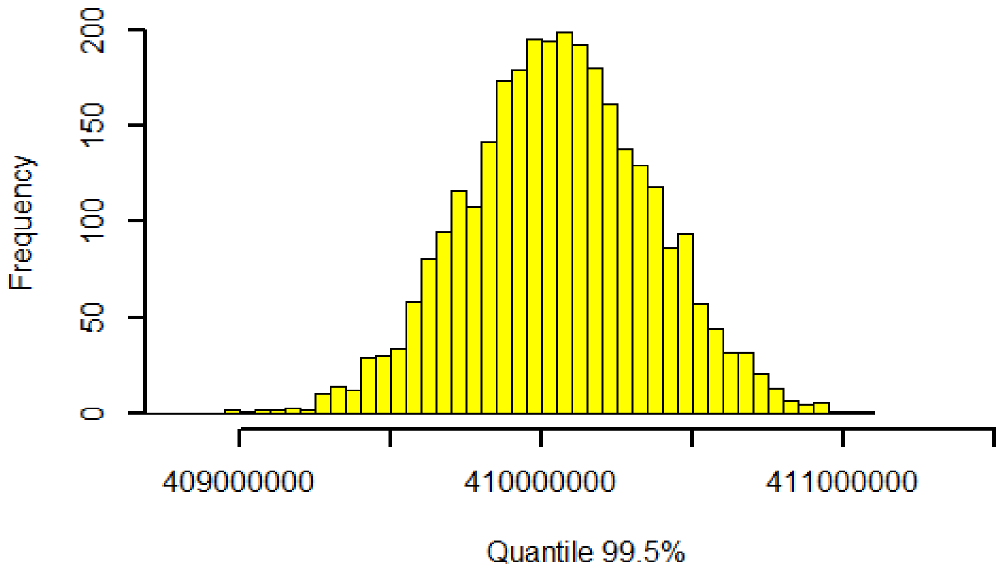


Figure 6.2: Results under Vine copula using mirror Clayton copulas. Histograms are represented on the diagonal, scatterplots on the upper triangle and rank-scatterplot on the lower triangle

	MTPL	GTPL	MOD
MTPL	1	0,478	0,478
GTPL	0,478	1	0,478
MOD	0,478	0,478	1



min	5%	mean	95%	max
409.175.924	409.514.831	410.035.277	410.529.311	410.825.955

Note that the VaR is very near to the vine copula's second scenario, where the correlation between *MTPL* and *GTPL* is strong enough ($= 0,523$) to counterbalance the lower correlation ($0,237$) between *GTPL* and *MOD*.

6.4 Gross SCR

In this section the capital requirement obtained with the different copula assumptions will be compared with the ones obtained under Solvency II approaches. First of all, the SCR in case of $\rho = 0$ and $\rho = 1$ are provided, to give a rough idea about a floor (neglecting negative correlations) and a cap :

$\rho = 0$	$\rho = 1$
51.808.138	84.229.034

Recalling formula (4.8), the distance between the quantile $99,5\%$ and the mean can be approximated as 3 times the standard deviation of $\widetilde{X}$. This concept is now tested by comparing the actual distance between quantile and mean with $3\sigma_S V$, where $\sigma_S = \frac{\sigma(\widetilde{X})}{V}$ is derived from the simulated distribution and V is set equal to the Best Estimate:

Copula	simulation	$3\sigma_S V$	σ_S	k
Gaussian	68.403.369	75.080.469	7,58%	2,73
Vine	77.964.830	76.005.350	7,68%	3,08
Full Clayton	80.374.160	77.091.111	7,79%	3,12

The table above shows how the approximation provided by Solvency II fits well only the Vine copula case. In fact, the multiplier k of the formula $k\sigma_S V$, which returns a SCR equal to the simulated one, is more near to 3 with the Vine. Instead, the SCR in the Gaussian copula case is overestimated by such formula, since multiplier $k = 2,73$ underlines that the distribution in the Gaussian case is more symmetric, and $k = 3$ would overestimate the skewness.

Another interesting comparison is the application of the Standard Formula using both the volatility factors (SF approach) provided by the regulation and the ones of the distributions (semi-USP approach). The underlying formula is $3\sigma V$, where σ is computed with formula (4.9), and, by assuming that only the Premium risk is present, the SCR is equal to:

USP	SF
75.106.699	88.735.619

Given that the Volumes V and volatility factors σ are:

lob	V	σ_{USP}	σ_{SF}
MTPL	225.000.000	$7,28\%$	10%
GTPL	60.000.000	$14,86\%$	14%
MOD	45.000.000	$11,38\%$	8%

It is interesting to note two things about the results:

- the SCR under semi-USP approach is very similar to the one obtained with $3\sigma V$ in the Gaussian copula case;

- the SCR under Standard Formula is the higher observed till now, since σ_{SF} for *MTPL* is significantly higher than the observed one. The SCR is even higher than the fully correlated case.

6.5 Panjer Backtesting

Here we are, all the lines of business have been simulated and, through the vine copula, we achieved the desired correlation. Now it's time to reinsure! Before setting the goals to reach through the reinsurance optimization, it is important to build all the tools needed. What's left to do is a stable pricing function that it is able to return to the user a reliable result. In the previous chapter two pricing principles were particularly useful for our purpouses:

- the *Standard Deviation* pricing principle;
- the *Proportional Hazard* pricing principle:

They both hinder particular properties and drawbacks, but they both share subadditivity, which comes very handy when dealing with reinsurance layering.

One can choose to build a pricing function in two ways: the rough and the accurate one. The first one is achieved simply by using the simulations to derive the result, while, with the second approach, the actuary can rely on other techniques, like Panjer algorithm, to avoid incorporating too much simulation noise in the calculations. For example, with regard to the standard deviation principle, one can use the standard deviation of the simulated loss transferred to the reinsurer to extract the corresponding loaded premium. But here comes the problem, are we sure that the simulations describe properly such long tailed risk, even when computing the price of extreme layers? Because it is fundamental to remember that an Excess of Loss deals with extreme events, and both risk assessment and pricing procedures should have a rigorous study behind.

As explained in the previous chapters, the Panjer algorithm is able to compute the aggregate claim amount $\widetilde{X}$, given the discrete distribution of the single claim $\widetilde{Z}$ and of the frequency $\widetilde{N}$ belonging to $(a, b, 0)$ family. Given a good discretization of $\widetilde{Z}$, the outcome is very interesting to access with accuracy the shape of very skewed distributions. In the reinsurance literature, the Panjer algorithm is literally praised since simulations usually fail where such technique succeeds: an extremely skewed $\widetilde{Z}$ and a low mean $\widetilde{N}$ may require millions of simulations to adequately describe the resulting $\widetilde{X}$. In our scenario, all the lobs are light-medium tailed, and we shouldn't suffer any significant loss in accuracy when using simulations.

We can use the Panjer algorithm to backtest the precision of the simulations and determine if the "rough" approach would be a good approximation.

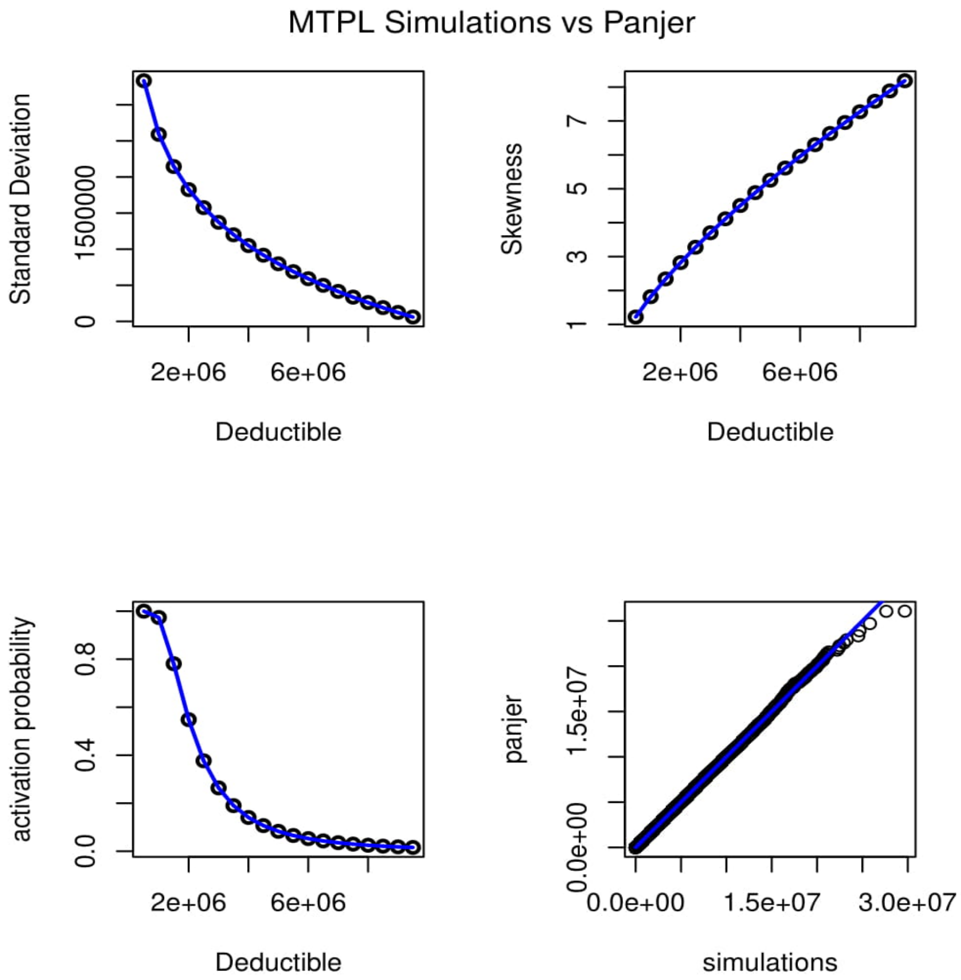


Figure 6.3: Simulation results (black dots) and Panjer output (blue line)

In Figure (6.3) the standard deviation, skewness, activation probability of the layer are compared between the simulations and the Panjer output for the MTPL line of business. All those characteristics refer to $\widetilde{X}_{re}$, where the XoL

Layer is always structured until the policy limit PL (assumed to be equal to 10.000.000 for MTPL), such that:

$$\widetilde{X}_{re} = \sum_{i=1}^{\widetilde{N}} Layer_{D,PL-D}(\widetilde{Z}_i) \tag{6.1}$$

The results obtained through simulations are extremely coincident with the one extracted from the Panjer algorithm, also in terms of skewness, which is a very delicate statistics to retrieve. Therefore, when layer operators are applied on these simulations, the results are solid and stable.

The 4^{th} graph in Figure (6.3) is a QQ-plot, with $D = 1.000.000$, and it shows how the simulations have been able to properly describe also the tail behaviour of the ceded risk. As long the points lie on the blue line, we are sure that the simulation process is adequately precise. We can clearly say that 100.000 simulations are enough to achieve a precise estimation of the risk. Of course, given a higher PL, this stability may not hold and a higher number of simulation might be required.

As in the previous case, Figure (6.4) shows the comparison between the GTPL simulations and the Panjer algorithm. The results are again very satisfying, also because this lob is the most volatile among the three proposed in the thesis. Comparing the skewness of GTPL with the MTPL one, it is possible to note that the latter is higher when we deal with Layer with high deductible. It may seem counter-intuitive since GTPL is characterized by more skewed and volatile $\widetilde{Z}$. But, actually, we should take into account how much extreme would it be to observe a huge ceded claim for both lines. In fact, GTPL is on a higher scale of measure and extreme layers for MTPL might not be as much extreme for GTPL.

Of course, since light tailed, also the MOD is described with precision through simulations. For the same reason as before, note in Figure (6.5), how the skewness of $\widetilde{X}_{re}$ explodes when the deductible increases. This is due to the fact that it is almost impossible to observe a large claim in this lob. Also the 3^{rd} graph confirms that the activation probability of the layer is very low compared to the MTPL and GTPL case. Therefore, due to low activation probability and low risk mitigation, it is clear that an Excess of Loss reinsurance would be almost useless when applied to this lob. Therefore, the considerations on

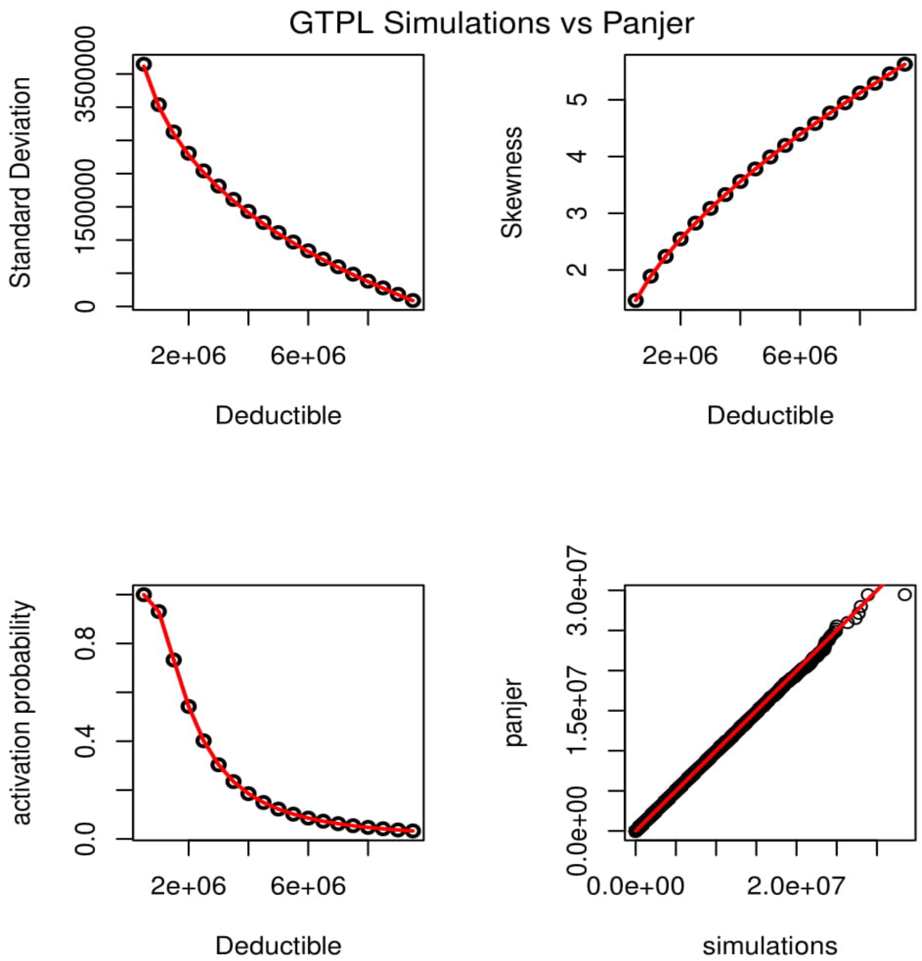


Figure 6.4: Simulation results (black dots) and Panjer output (red line)

non-proportional reinsurance will be limited to the other 2 lobs from now on. The QQ-plot of Figure (6.5) is displayed with $D = 100.000$ due to its lower scale.

Now that the accuracy of the simulations has been backtested, we can proceed studying the premium principles. Since the Panjer algorithm returns more precise and smoothed estimation of the distribution, it will be preferred for the computation of the safety loadings from now on. Of course, as shown before, premium principle calculations based on the simulations would be a good proxy in this scenario.

Following the reasoning of Equation (6.1) the safety loadings, expressed as

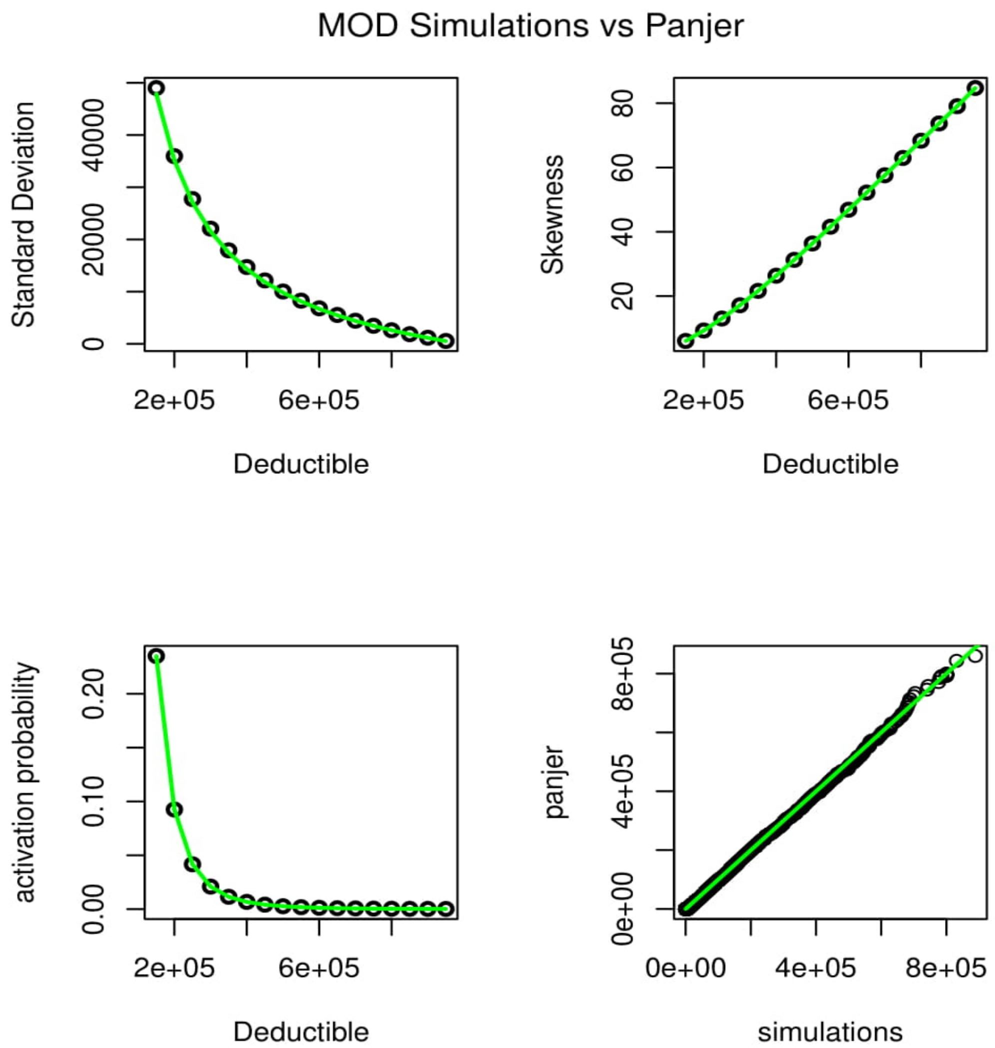


Figure 6.5: Simulation results (black dots) and Panjer output (green line)

percentage of the pure premium $P = E(\widetilde{X}_{re})$, are shown in Figure (6.6).

The graph on the left shows that the safety loading for MTPL increases almost linearly with the increase of the deductible. Instead, the GTPL is characterized by a slightly concave behaviour. It is also interesting to see that the loading for MTPL is higher in relative terms to the GTPL one when applying layers with a high attachment point. As described before, this fact is due to MTPL being less heavy tailed, and therefore the ceded losses for such layers are less likely to appear.

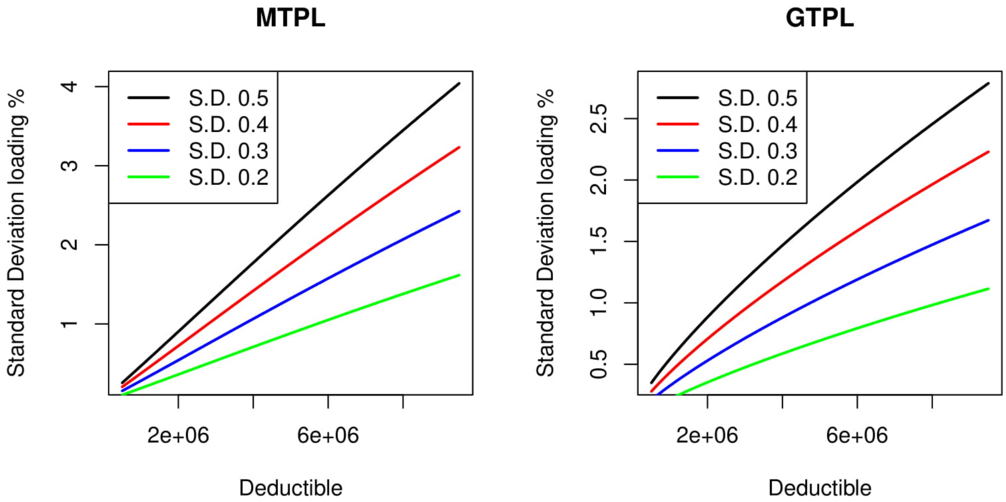


Figure 6.6: Reinsurance loading with Standard Deviation principle for different β

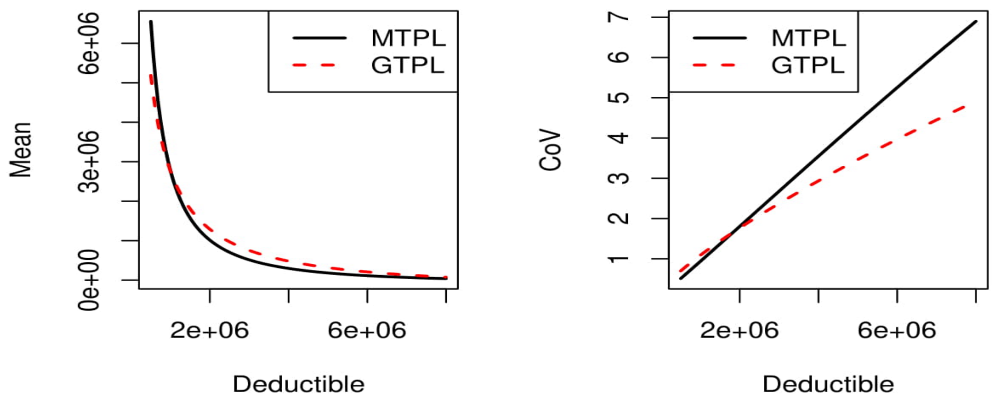


Figure 6.7: Mean and Coefficient of Variation of an Excess of Loss contract

As shown on the left graph in Figure (6.7), the mean of MTPL decreases faster then the GTPL one. Remember that the MTPL is characterized in our scenario by a way higher frequency component and represents the 63, 2% of the portfolio against the 20, 7% of GTPL. Therefore, MTPL starts the line from a higher point.

On the right we can see that in fact the Coefficient of Variation of MTPL increases more steeply than GTPL. The safety loading under standard deviation premium principle is equal to $\beta \cdot CoV$ when expressed as percentage of the mean.

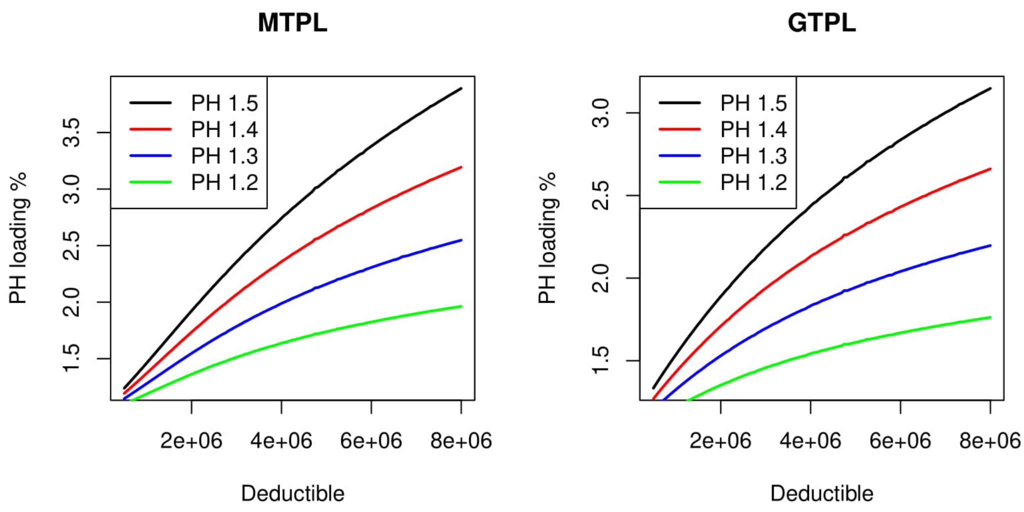


Figure 6.8: Reinsurance loading with Proportional Hazard for different ρ

In Figure (6.8) the different loadings according to Proportional Hazard principle are shown. In this case, the functions for both the lobs are concave. The PH pricing principle relies on the cumulative distribution function of the underlying risk, which is the exact output of Panjer algorithm. For this reason practitioners usually connect this premium principle with such algorithm. The choice between using the PH or Standard deviation principle to price reinsurance may depend on how they price the most extreme layers and how the subadditivity holds when a layer is split.

In Figure (6.9) the two pricing principles are compared on both the lobs using different parameters. It's interesting to note that empirically the two methods return the same price for layers with low deductible when

$$1 + \beta \approx \rho.$$

At the same time, when we deal with layers with high attachment point, the more concavity that characterizes the PH principle tends to create a significant gap between lines of the same color. This means that the user might choose a weighted mean of the two methods to give the desired relevance to layers with high D. The calibration of a reinsurance pricing curve is an interesting topic

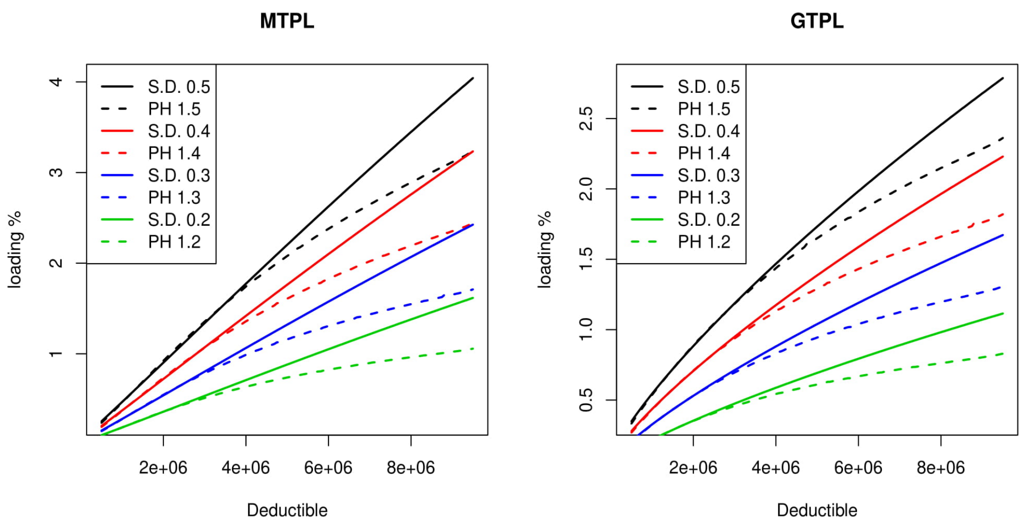


Figure 6.9: Reinsurance loading for both Standard Deviation and Proportional Hazard

worth to be further analyzed.

An important question at this point is: how can we determine which parameter and pricing principle could be a good choice to insert into the reinsurance optimization?

To have a vague idea of a realistic parameter, we can actually do compute in reverse which parameters the direct insurer could have used to price its own lines of business according to each principle. In a second step we can assume that the reinsurer might price its risk according to a similar risk aversion and economic competitiveness in the market.

In order to do so with Panjer, the distribution of gross aggregate claims $\widetilde{X}$ of the insurer should be computed. Note that we might face underflow when dealing with MTPL, since its frequency has $E(\widetilde{N}) = 50.000$, and therefore, $P(\widetilde{N} = 0)$ is approximated to 0 by the machine. Hence we will need to divide the size parameter of the Negative Binomial by 2^n and then convolve the process 2^n times. In this scenario $n = 1$ is enough to solve the underflow. The other two lobs are small enough to not return any underflow problem.

When computing the parameter β_{lob} of Standard Deviation principle from the point of view of the direct insurer, we have:

$$P_{lob} \cdot \lambda_{lob} = \beta_{lob} \cdot \sigma(\widetilde{X}_{lob}) \Rightarrow \beta_{lob} = \lambda_{lob} \frac{P_{lob}}{\sigma(\widetilde{X}_{lob})}$$

and, for each lob the outcome is:

	MTPL	GTPL	MOD
P	225.000.000	60.000.000	45.000.000
$\sigma(\widetilde{X})$	16.366.551	8.880.263	5.119.400
λ	1,2%	6,7%	13,8%
β	16,5%	45,3%	121,3%

Note that β is extremely high for MOD and very low for MTPL due to respectively low and high competitiveness in the insurance market. Recalling Table (6.2), GTPL have a reasonable probability of observing a profit in the year ($= 70,34\%$) and the safety loading applied by the direct insurer seems coherent with the underlying risk. Therefore, β_{GTPL} may be a good benchmark to calibrate the pricing principle.

It is possible to compute β also on portfolio level (β_{tot}), but keep in mind that MOD probably doesn't need any Excess of Loss treaty. Also, MOD's safety loading is extremely high compared to the underlying risk, and therefore, it will shift up the loading criteria. Since the direct insurer would probably avoid to transfer MOD to the reinsurer, such β_{tot} can be interpreted as the relative target profitability of the reinsurer, which is in line with the insurer's one. The obvious consequence is that the price of Excess of Loss treaties for MTPL and GTPL is raised to take in consideration the absence of MOD's profitability into the reinsurance portfolio.

Therefore, using the selected vine copula simulation of the previous section, we can aggregate the lobs and derive the standard deviation through simulation or Panjer. The result is the following:

$$\beta_{tot} = \lambda_{tot} \frac{P}{\sigma(\widetilde{X})} = 3,92\% \frac{10^6(225 + 60 + 45)}{25.335.116} = 51,06\%$$

The same can be applied to Proportional Hazard principle, where the goal is to find the ρ that satisfies the following equation:

$$P_{lob} \cdot \lambda_{lob} = \int_0^\infty (1 - F_{\widetilde{X}_{lob}}(x))^{1/\rho} \, dx$$

Through a solver algorithm the following parameters have been determined for each lob and in total after vine copula aggregation:

	MTPL	GTPL	MOD	Total
ρ	1,19	1,54	2,83	1,60

As before, GTPL is in the middle between the other two lobs. For MTPL the rule $1 + \beta \approx \rho$ holds quite well when comparing the two pricing principles. Note that the rule holds empirically for Deductible D in an approximate range of $(500.000; 4.000.000)$, but in this case we are analyzing trivially the case $D = 0$, where the whole portfolio of the insurer is hypothetically ceded.

Since an Excess of Loss applied to MOD would cover against extremely rare losses, both the parameter β and ρ assume huge values. For this reason it might be a good idea to avoid considering XoL contracts on such lob. The actuary can decide to include such contract in the optimization algorithm, but it would be almost useless and it would cost in computing times. As a general rule, leave out everything unnecessary is the best choice to manage only feasible combinations of reinsurance that the actuary may consider worth to compare.

Seen the previous results, the choice of the parameters for each method are the following:

	MTPL	GTPL
β	0,2	0,45
ρ	1,2	1,45

We can now start to list some feasible assumptions for the reinsurance optimization and build on them an algorithm that returns all the possible reinsurance combinations.

Chapter 7

Optimal Reinsurance

In the previous sections, all the needed elements have been explained to understand the chosen parametrization and successfully build on it the simulation environment, together with the subsequent aggregation between lines of business. Now that all the pieces come together, we have in our hands the distribution of the aggregate claim amount $\widetilde{X}$ gross of reinsurance. It's time to determine the optimal reinsurance.

To perform the reinsurance optimization on the simulations, a huge number of possible reinsurance treaties has to be tested and compared. One can decide to optimize all the results with regard to the single line of business or the whole business. The path treated in this thesis is the second one, since dealing with a standalone lob might take to results that are not optimal on an aggregate basis, which is the thing that matters the most. A method to reduce the dimension of all the possible programs must be defined to focus the attention on the most efficient.

As a first step, it is important to compute some profitability and risk indexes gross of reinsurance, to set up some constraints for the optimization. Assuming that the own capital U of the company is equal to the 20% of the tariff premiums B of the year, we have:

$$U = 20\% \cdot 457.101.231 = 91.420.246$$

and therefore, if the results obtained through Vine Clayton aggregation are considered, the Solvency Ratio $SR$is:

$$SR = \frac{U}{SCR} = \frac{91.420.246}{77.964.830} = 117,26\%.$$

Given the parametrization of the lobs on the italian market, the expected profit in the year is equal to

$$E(Profit) = \sum_{lob} \lambda_{lob} \cdot P_{lob} = \lambda_{tot} \cdot \sum_{lob} P_{lob} = 3,92\% \cdot 330.000.000 = 12.936.000$$

and, subsequently, the expected Return on Equity ROE is

$$E(ROE) = \frac{E(Profit)}{U} = \frac{12.936.000}{91.420.246} = 14,15\%$$

.

From the graphical point of view is possible to observe the different quantiles of ROE distribution, showing a high volatility of results and significant probability to perform either greatly and poorly in the year.

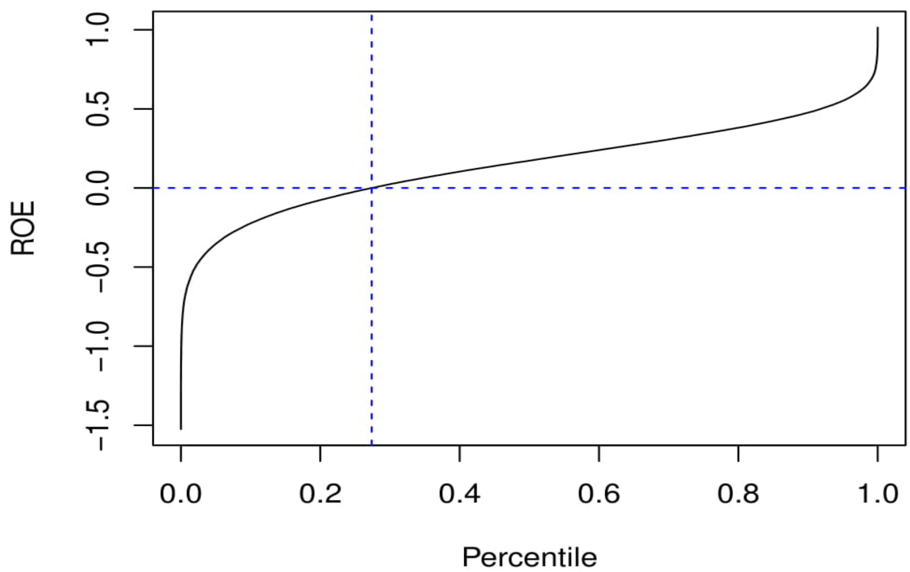


Figure 7.1: the horizontal dotted line represents $ROE = 0$ and the vertical one the corresponding percentile

Since the SCR is a risk measure that quantifies extreme risk, it is worth to consider the Coefficient of Variation CoV to give also emphasis to the volatility around the mean. Recalling the previous results:

$$CoV = \frac{\sigma(\widetilde{X})}{E(\widetilde{X})} = 7,69\%.$$

Let us also take into account the most important quantiles of the Combined Ratio CR. Under the hypothesis that expenses E are deterministic and equal to the expense loading $c \cdot B$:

$$CR = \frac{\widetilde{X} + E}{B} = \frac{\widetilde{X}}{B} + c = \widetilde{LR} + c$$

Note that Figures (7.1) and (7.2) show the same graph but mirrored and on another scale of measure.

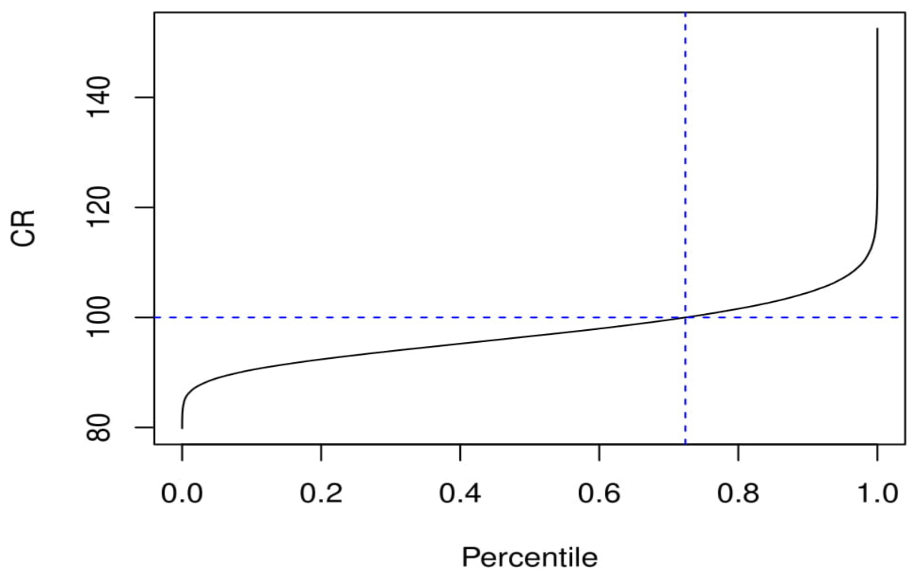


Figure 7.2: the horizontal dotted line represents $CR = 100\%$ and the vertical one the corresponding percentile

Now some assumptions are made on feasible goals that the insurance company may want to achieve through reinsurance:

- the target Solvency Ratio under Vine Copula aggregation is between 150% and 170%;
- the expected ROE should not fall below the 10%;
- applying an Excess of Loss reinsurance to both MTPL and GTPL is required in order to benefit from the NP factor provided by Solvency II;
- Quota Share and Excess of Loss contracts will be mixed in order to define the population of treaties. The Excess of Loss is applied first, and then the Quota Share is applied with a cost proportional to the tariff premiums net of the cost of the Excess of Loss;
- for the selection of all the possible Quota Share, a step of 5% is applied to discretize such contract and deal with realistic percentages, that are likely to be negotiated between insurer and reinsurer;
- for the selection of all possible Excess of Loss, we won't consider subsequent layers and the presence of possible reinstatements, aggregate limits and aggregate deductibles. The choice has been made in order to reduce the computational times, since otherwise there would be millions and millions of combinations of treaties. In a second step the layer's split and use of reinstatement is analyzed;
- the deductible D is defined in the range $(500.000; 2.000.000)$ with a step of 250.000. The reason behind is that, given the previous assumption, there is only one layer per lob, and we don't want the Layer to overfit the simulations by protecting extreme events only at a low price. Therefore, the deductible's domain is such to determine a good attachment point for the treaty;
- the limit L have been discretized with a 2 millions step in such a way to study how far should the layer go;

The reason to set up such assumptions is mainly to reduce execution times to a manageable amount. In order to have a wider view on the results, the simulated treaties will have at least a Solvency Ratio equal to 145% and a $ROE = 8,5\%$ at least.

Since the attachment point of the Excess of Loss is low, there isn't a significant difference in pricing with standard deviation principle or proportional hazard when $\beta \approx 1 + \rho$, as shown in Figure (6.9). The standard deviation principle will be used from now on to price Excess of Loss treaties. Instead, when pricing Quota Share, we will assume that the reinsurer requires an additional compensation by setting the reinsurance commission loading $c_{lob}^{re} = 95\% c_{lob}$ for each lob.

Let us investigate different optimizations: the main difference between them is the choice of which variable to use on the x and y axes of the plot. What we aim for is to find reinsurance programs with interesting trade-off between profit and risk under the aforementioned constraints.

7.1 Total Cost Optimization

The aim of this optimization is to minimize the Total Cost TC of the ceding insurer, that recalling 5.2 is equal to minimizing the following term:

$$\lambda_{re} \cdot E(\widetilde{X}^{re}) + CoC \cdot SCR$$

that are respectively the ceded profit to the reinsurer and the cost of capital of holding that amount of SCR. From now on, we will talk about *"Total Cost"* when referring to this term.

The cost of capital rate CoC is assumed to be equal to 6%, like the one present in Solvency II. Hence, gross of reinsurance we have:

$$0 + 6\% \cdot 77.964.830 = 4.677.890$$

Parodi [23] proposes to use as axes in the plot the reinsurance premium and the Value at Risk. As shown in Figure (7.3), the reinsurance premium and the Value at Risk are dependent in an almost linear way. The main reason behind such result is given by the fact that the reinsurance pricing under standard deviation principle is a risk based approach, and therefore, the price is proportional to the ceded risk. This linearity is not desired in the optimization context since there isn't a clear convex efficient frontier on which we can extract the best programs.

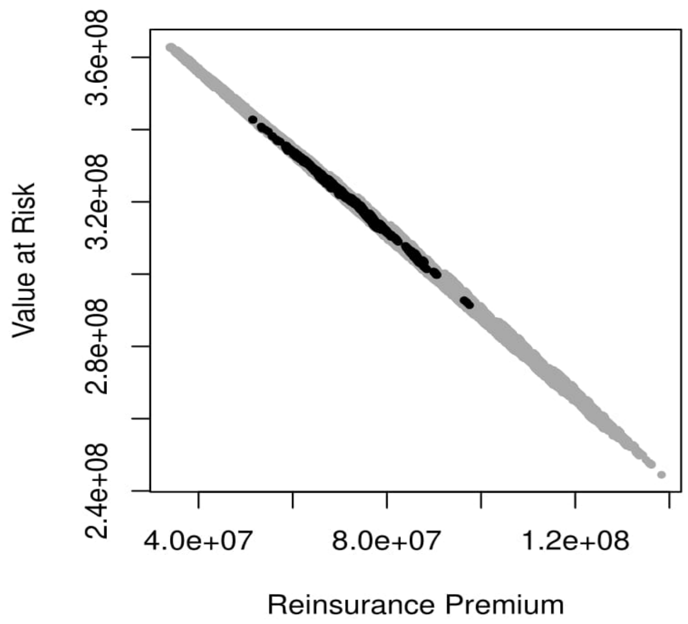


Figure 7.3: the gray points represent all the possible reinsurance programs and the black points the programs that satisfy the constraints

Be careful that this result implies that a frontier approach is not adequate for these two axes. The main problem of the chosen metrics is the fact that the reinsurance premium doesn't catch how much profit the direct insurer is ceding to the reinsurer, and the Value at Risk doesn't quantify the underlying SCR.

Therefore, we can try to minimize the Total Cost and see which contract is suggested:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.500.000	6.000.000	75%	100%	750.000	6.000.000	100%

Note that the application of excess of loss treaties has been avoided to MOD since it is almost useless and slows down computations. The following plot shows a closer look to the previous one, where only the black dots are considered. The efficient frontier that minimize both premium and Value at Risk is signaled in red, and the program that minimize the total cost is the blue point.

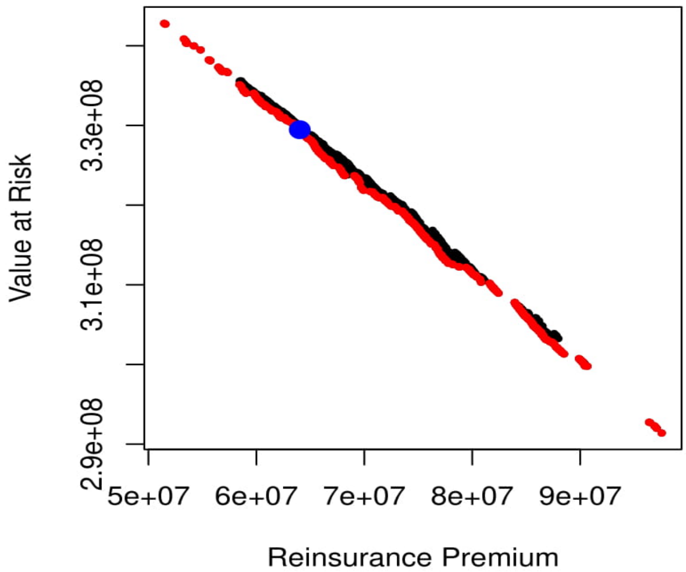


Figure 7.4: The red dots represent the efficient frontier and the blue dot the program that minimize the Total Cost

Due to the thin shape of the programs, the concept of frontier may appear not so coherent. But through it we can reduce the amount of programs by roughly 5 times, which is not negligible.

The main indicators of the direct insurer after the application of the program are:

CoV	Solvency Ratio	E(ROE)	TC
7,46%	150,59%	10,68%	6.817.467

With respect to the gross of reinsurance case, the *CoV* has been reduced. Note that the Total Cost is not improved passing from gross to net of reinsurance. Why? The main reason lies in the choice of using a Negative Binomial for the frequency component of each lob: due to its asymptotic behaviours, the excess of loss treaties' effectiveness is limited because the risk of observing a huge number of small losses is still present.

The trade-off between ceded profit and SCR reduction is such that any reinsurance will lead to an increase of the Total Cost under the imposed constraints. It is not a coincidence that the Solvency Ratio constraint has been met in its lower bound (= 150%). In fact, given the aforementioned trade-off, TC is minimized under constraint when the Solvency Ratio is at its lowest, in our case 150%.

But the result is even more complex: it is easily noticeable that the TC formulation can advantage Quota Share contracts applied to MTPL, since the ceded profit would be low and the SCR would be minimized by a significant amount. But, at the same time, such contracts are characterized by a high premium due to the predominant volume of MTPL over the total. Because of the choice of the x and y axes, these Quota Share are excluded when the efficient frontier is considered. Therefore, the optimal contract under these conditions is the result of complex trade-offs and it doesn't represent an obvious minimization of the Total Cost.

Another interesting alternative is to use the SCR (or equivalently the Solvency Ratio) instead of the Value at Risk. In this way the frontier would minimize the trade-off a more useful indicator.

As shown in Figure (7.5), we are in front of a non-linear frontier by changing the Value at Risk with the Solvency Ratio, which doesn't include the mean of retained losses in the numerator. The application of the frontier reduces the number of programs by 40 times.

The chosen program is defined as:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
2.000.000	6.000.000	75%	100%	500.000	8.000.000	100%

With the following characteristics:

CoV	Solvency Ratio	E(ROE)	TC
7,39%	152,89%	10,47%	6.945.862

With respect to the one seen before, this program reduces more the CoV and increase by a small amount the Solvency Ratio at cost of a small amount

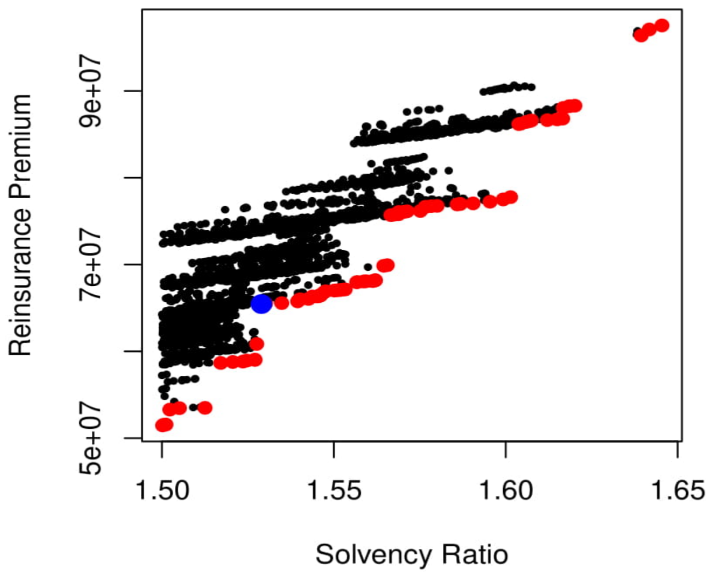


Figure 7.5: The red dots represent the efficient frontier and the blue dot the program that minimize the Total Cost

of expected ROE. It's hard to tell which one is better, but we can observe that this program covers more the GTPL, which is the most risky lob in the portfolio. Of course, since the frontier is composed by just few programs, it is not surprising to see a higher TC.

But why does this program in particular achieve the lowest TC over the frontier? The reason is clearer from a graphical point of view: the optimal program is the furthest from the lower right corner of the figure. In fact, if we aim for a convex efficient frontier, such program would be one of the first to be cut out of the selection. Hence, the Total Cost optimization and concepts of efficient frontier clash each other when dealing with Solvency ratio and reinsurance premiums as metrics.

7.2 Convex Frontier Optimization

As seen in Figure (7.5), the efficient frontier can be fuzzy and composed by non-continuous linear segments. This is given by the fact that:

1. many constraints on the feasible programs have been set in both terms of

structure and targets. Passing from an unconstrained to a constrained structure reduces significantly the dimension of the selection, and it may create gaps;

2. only Quota Share, Excess of Loss and a combination of the two have been simulated. Note that, even given these limited dimensions, the computation of the combinations is time consuming;

3. to build the program simulation algorithm, both Excess of Loss and Quota Share have been discretized. The idea behind is that, using a discretization with minimum loss of information, the shape would be way more completed and convex;

4. only 3 lines of business are considered. Increasing the number of lobs would provide a smoother frontier because the possible gaps would be filled.

Regarding the first point, the insurance company decides carefully its own goals in terms of profitability and riskiness. Therefore, studying only the programs that may be used in the practical context is the main goal. On the second point, Quota Share and Excess of Loss are the most used reinsurance contracts used in practice, and the results can provide useful insights. The third point is justified by the standard practice: it is unusual to observe a Quota Share where the parameter α is not a multiple of 5%.

But, does this frontier's non-convexity help in the selection of a optimal program? The answer is yes, since in many cases we can build a convex frontier and reduce once more the number of programs considered. To give an idea, a convex frontier has been built in Figure (7.6) on the basis of Figure (7.5).

In this way, the convex frontier is a subset of the non-convex one, which is a subset itself of the whole set of constrained programs.

The reason why the convexity is a desired property is the following: from a graphical point of view, when we move from a program on the convex frontier to another one, the best trade-off is achieved by moving on the next program that lie on convex frontier. Following this reasoning, it is possible to cancel out from the selection all the programs that don't lie on the convex frontier.

The optimal programs that lie on the convex frontier of Figure (7.6), ordered by Solvency ratio, are described in the following tables.

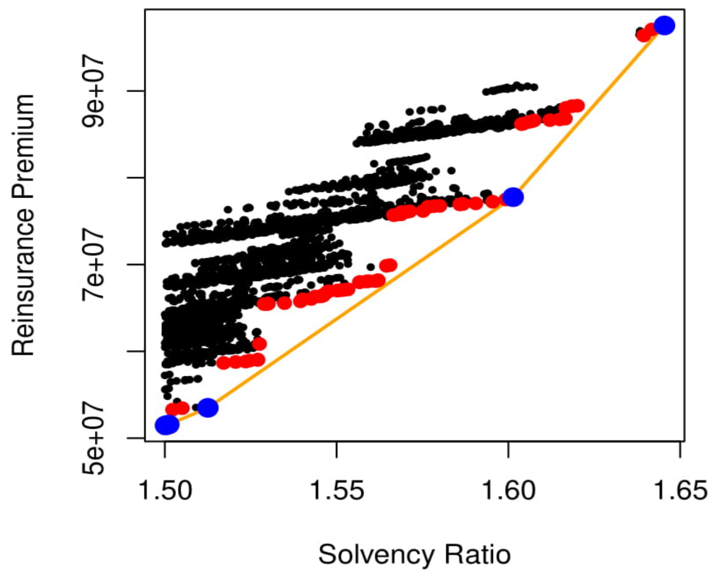


Figure 7.6: The orange line connects the programs that lie on the convex frontier (blue), leaving out the previous non-convex efficient frontier (red).

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
750.000	6.000.000	85%	100%	500.000	9.500.000	90%
750.000	9.250.000	85%	100%	500.000	9.500.000	90%
500.000	4.000.000	85%	100%	500.000	8.000.000	90%
1.250.000	6.000.000	70%	100%	500.000	8.000.000	100%
2.000.000	4.000.000	60%	100%	750.000	4.000.000	100%

CoV	Solvency Ratio	E(ROE)
7,23%	150,02%	10,06%
7,23%	150,11%	10,02%
7,23%	151,25%	10,01%
7,40%	160,14%	10,01%
7,66%	164,54%	10,01%

We can extrapolate interesting information from these results:

- the first two programs are almost identical, except for L_{MTPL}, and are almost overlapped from the graphical point of view. If their profit and risk indexes are compared, it is possible to note that there isn't a significant improvement of the Solvency ratio when covering the last extreme segment of MTPL;
- compared to the 2^{nd} program, the 3^{rd} one is able to achieve a +1% of Solvency ratio in exchange of a really small amount of ROE, and the meanwhile perform the same CoV;
- the last 3 programs show with constant profitability a trade-off between CoV and Solvency ratio;
- the ROE constraint has been met in its lower bound, and there aren't optimal contracts with a expected ROE that distances significantly the 10%.

Usually the frontier is composed by 10 or less programs. One step further is to reduce such dimension by applying the tangent line to the frontier. The slope of the tangent line can be derived by connecting the first and the last point of the convex frontier. The aim is to derive which programs are characterized by interesting trade-offs, and not to blindly select the tangent one. Care and judgement are always needed in every step of the optimization, and the fully automation of the process is not suggested, since the mere mathematical optimization may be misleading in certain situations.

In this scenario the tangent contract is the 4^{th} one, which presents an interesting trade-off between CoV and Solvency ratio. As described before, the 3^{rd} program almost dominates the first two, and their presence in the convex frontier might indicate that the reinsurance premium is not a coherent choice for the y-axis, since it doesn't clearly explain how much profit the company is ceding to the reinsurer.

Therefore, the optimizations presented in the next sections are based on the following three metrics:

- ROE as profit measure;

- Solvency Ratio as risk measure for extreme quantiles;
- CoV as risk measure for general volatility of the process.

For each optimization two out of three metrics will be chosen for the graphical representation. The resulting efficient frontier will be composed by programs that satisfy the constraint of the remaining metric.

7.2.1 ROE and SCR optimization under volatility constraint

With this optimization the Solvency Ratio and the ROE will be used on the x-axis and y-axis respectively. As usual, the programs will be filtered by the constraints $SR \geq 150\%$ and $CoV_{net} \leq CoV_{gross}$ and $ROE \geq 10\%$. The advantage of this optimization is that you don't need to be an actuary to comprehend immediately the results, since the axes are assigned to very clear metrics. In comparison, the CoV might require an additional effort to be fully understood.

On the left of Figure (7.7), the subset of programs is displayed in black. On the right graph a close up of the selection is represented, where the efficient frontier is red and the 3 programs that lie on the convex frontier are in blue.

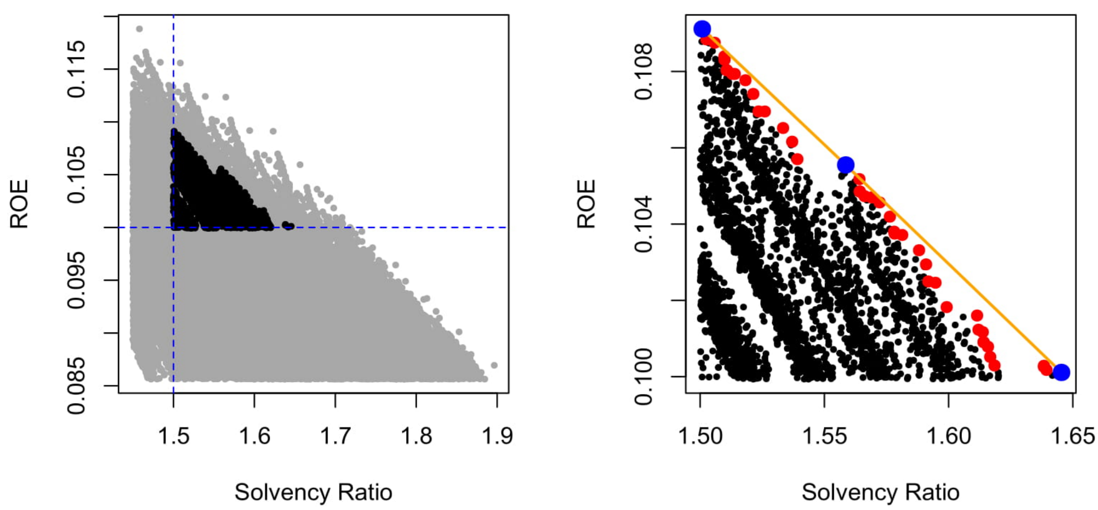


Figure 7.7

The 3 optimal programs in order of Solvency ratio are:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
2.000.000	6.000.000	70%	100%	1.500.000	4.000.000	100%
2.000.000	4.000.000	65%	100%	1.250.000	4.000.000	100%
2.000.000	4.000.000	60%	100%	750.000	4.000.000	100%

With the following characteristics:

CoV	Solvency Ratio	E(ROE)
7,65%	150,08%	10,91%
7,67%	155,86%	10,55%
7,66%	164,54%	10,01%

The results are very interesting and show some peculiarities:

- in each program the CoV is really near to the one gross of reinsurance, which is equal to $7,69\%$. Therefore the constraint has been satisfied in its upper bound, and the resulting programs maximize the trade-off between Solvency ratio and ROE without any significant reduction of CoV;

- all the programs rely heavily on the Quota Share of MTPL to adjust the Solvency ratio. The reason behind is that ceding such lob creates more equilibrium and a better risk allocation in the portfolio;

- In conjunction with the Quota Share, the Excess of Loss' layer of GTPL shifts towards more frequent risks to provide more coverage;

- the convex frontier is almost linear, even if the efficient frontier isn't;

- there is no explicit optimum with Solvency ratio = 160% because of the huge gap between the second and the third program;

Due to the low use of Non-proportional reinsurance and the wide use of proportional reinsurance to drive the result, this optimization share some similarities with the idea of optimal capital allocation. In fact, due to the portfolio structure of the insurer, this optimization attempts to reduce the MTPL volume over the total, and establish a better equilibrium between lobs.

7.2.2 Risk optimization under profit constraint

With this approach both the risk measures (Solvency Ratio and CoV) are represented on the axes, while the ROE constraint is used as a filter on the resulting graph. The aim is to optimize the trade-off between the general volatility and the tail heaviness under the desired profit constraint.

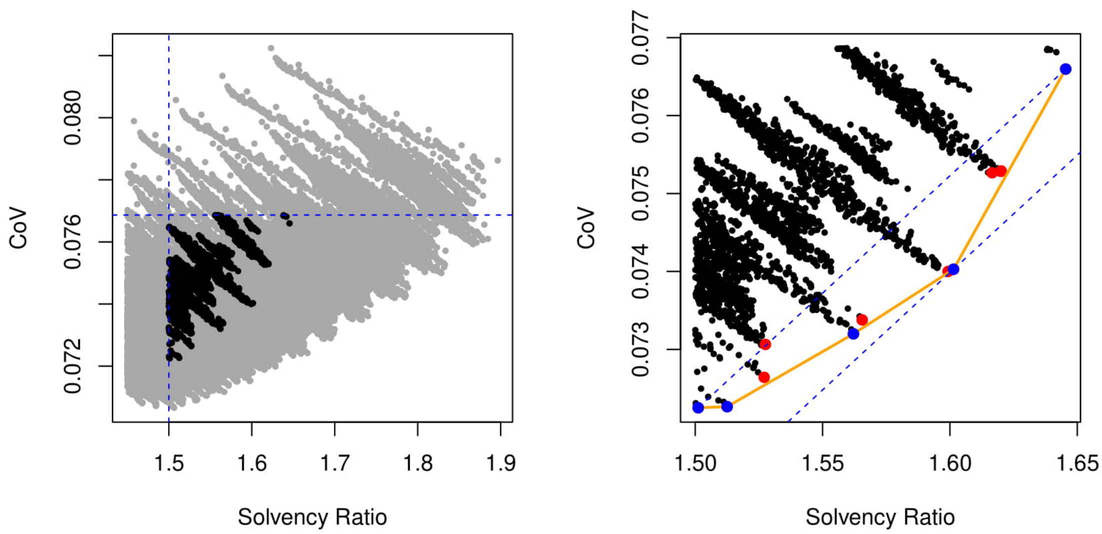


Figure 7.8

Looking at the left plot of Figure (7.8), the shape of the programs is interesting since it is composed by multiple oblique lines. Such phenomenon is given by the fact that there is a trade-off between SCR and CoV and, to reduce the computational times, all the programs returned by the simulation algorithm are subject to the constraint $ROE \geq 8,5\%$. Therefore, also the grey points in the plot are subject to a profit constraint and show this behavior. The subset under $ROE \geq 10\%$ assumes a spikier shape, which can be observed more precisely on the right of Figure (7.8).

As before, the efficient frontier is determined in red and, subsequently, the convex frontier is built. Note that, due to the particular shape of the subset, the efficient frontier is composed by few programs, and the use of the convexity reduces the number of selected programs by a small amount. The way the plot is structured makes the selection of optimal programs really clear and stable.

The optimal programs, ordered by Solvency ratio, are:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
750.000	9.250.000	85%	100%	500.000	9.500.000	90%
500.000	4.000.000	85%	100%	500.000	8.000.000	90%
750.000	9.250.000	75%	100%	500.000	9.500.000	100%
1.250.000	6.000.000	70%	100%	500.000	8.000.000	100%
2.000.000	4.000.000	60%	100%	750.000	4.000.000	100%

With the following characteristics:

CoV	Solvency Ratio	E(ROE)
7,23%	150,11%	10,02%
7,23%	151,25%	10,01%
7,31%	156,21%	10,02%
7,40%	160,14%	10,01%
7,66%	164,54%	10,01%

It is worthy to note that:

- the profit constraint has been meet in its lower bound, with ROE just slightly higher than 10%. In fact, the optimal trade-off between the two risk measures comes at the cost of reducing the expected profit;
- the 1^{st} program is almost dominated by the 2^{nd}. The $CoV = 7,22\%$ can be seen as the minimum value achievable under the Solvency ratio and the ROE constraints. If compared to the gross case ($CoV = 7,69\%$), the maximum reduction in relative terms is roughly equal to $6,1\%$, which is way far from the 20% assumed by Solvency's Standard Formula NP factor;
- the 5^{th} program appeared also in the previous optimization, since it meets the upper bound of CoV constraint and it is characterized by the highest Solvency ratio among the subset of programs.

The 4^{th} program is the tangent one. The reason behind this result is justified by several drivers. The program relies on a solid protection with an Excess of

Loss on GTPL and doesn't use any Quota Share on such lob to avoid sharing profit. At the same time it combines on MTPL a 70% Quota Share with an Excess of Loss that doesn't cover extreme losses after 7.250.000, due to their low occurrence. Hence, we see a balanced trade-off between the single reinsurance contracts that compose such program in order to reach the desired goals.

It is possible to note in Figure (7.8) that there is another program (in red), located really close to the 4^{th} one of the convex frontier. Such program is really similar, but with $D_{MTPL} = 1.500.000$ and $L_{GTPL} = 9.500.000$. Therefore, pushing the GTPL limit to the cap limit and reducing MTPL's protection is slightly inefficient in terms of SCR-CoV trade-off, but it can be considered as a valid alternative.

7.2.3 Cov-Profit Optimization under SCR constraint

Imagine the insurer has a precise SCR goal to reach through reinsurance. Which programs are able to achieve at least the profitability constraint, while at the same time reducing the volatility of the result? This type of optimization gives an answer to this question by using ROE and CoV as axes and SCR as filter in the graph. In this section different SCR targets will be analyzed and the respective optimum will be determined. In Figure (7.9) the imposed

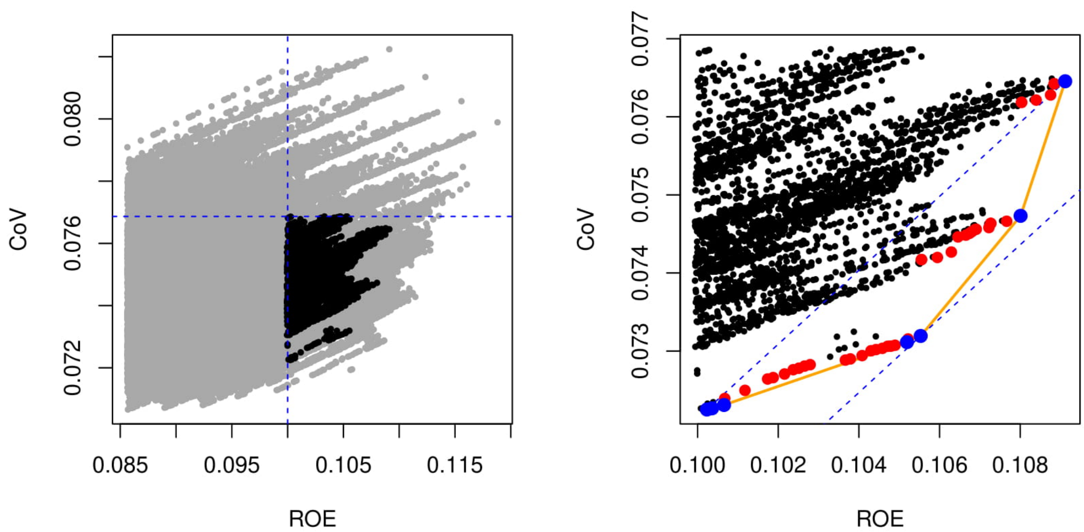


Figure 7.9: Optimization with Solvency ratio ≥ 150%

constraint on Solvency ratio is as usual equal to 150%. On the right, efficient and convex frontiers have been found, and, to avoid overwhelming the reader with too many programs, only the tangent one is provided:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.000.000	8.000.000	80%	100%	500.000	8.000.000	100%

With the following characteristics:

CoV	Solvency Ratio	E(ROE)
7,32%	150,01%	10,55%

The most close program to the tangent one shares its structure, except for having $L_{MTPL} = 6.000.000$ and $L_{GTPL} = 9.500.000$. Hence, the only difference is choosing which lob to cover more and which less with Excess of Loss. As expected, all the programs that lie on the convex frontier have a Solvency Ratio $\approx 150\%$, and the constraint is met in its lower bound.

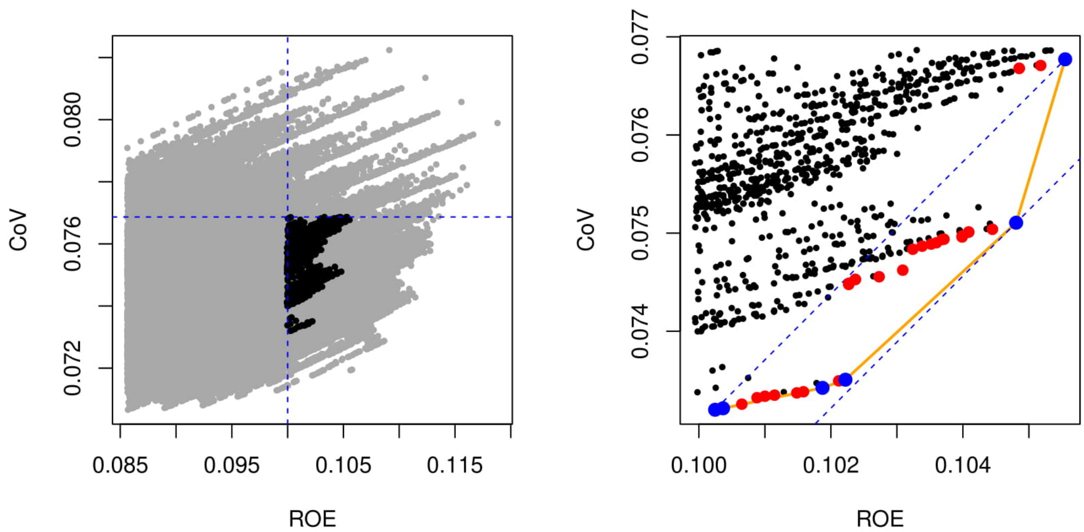


Figure 7.10: Optimization with Solvency ratio $\geq$ 155%

Changing the Solvency ratio constraint from 150% to 155%, the bold points of Figure (7.10) are reduced in number, and a new frontier is determined.

Looking at the right graph, the tangent program almost shares its first place with another program in terms of tangency. Therefore, both programs are shown below:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.000.000	8.000.000	75%	100%	500.000	8.000.000	100%
2.000.000	4.000.000	70%	100%	500.000	4.000.000	100%

CoV	Solvency Ratio	E(ROE)
7,35%	155,02%	10,22%
7,51%	155,44%	10,48%

To reach the $SCR \geq 155\%$ constraint, the two programs choose two different paths:

1. the first protects more the extreme events relying on more Excess of Loss;
2. the second suggests to buy more Quota Share on MTPL and spend less on Non-proportional treaties for both MTPL and GTPL.

It is difficult to assess which one is better, but we can note that the first is identical to the tangent program with $SR = 150\%$, except for $\alpha_{MTPL} = 75\%$ instead of 70%. Supposed that the tail heaviness might have been underestimated or it is uncertain due to lack of robust statistics, the best choice is to prefer the first program over the second to protect the insurer by this kind of risk. The danger of choosing the second contract is to blindly follow a higher profit without issuing any question about the reliability of distributional assumptions.

This is a good example of why the automation of the whole optimization process is not a good idea: critical thought and judgement should always be present when deciding whether the tangent program might be the best for the company.

Shifting now the constraint to $SR \geq 160\%$, the number of remaining programs is reduced significantly. As depicted in Figure (7.11), the convex frontier is now linear and composed by only two programs:

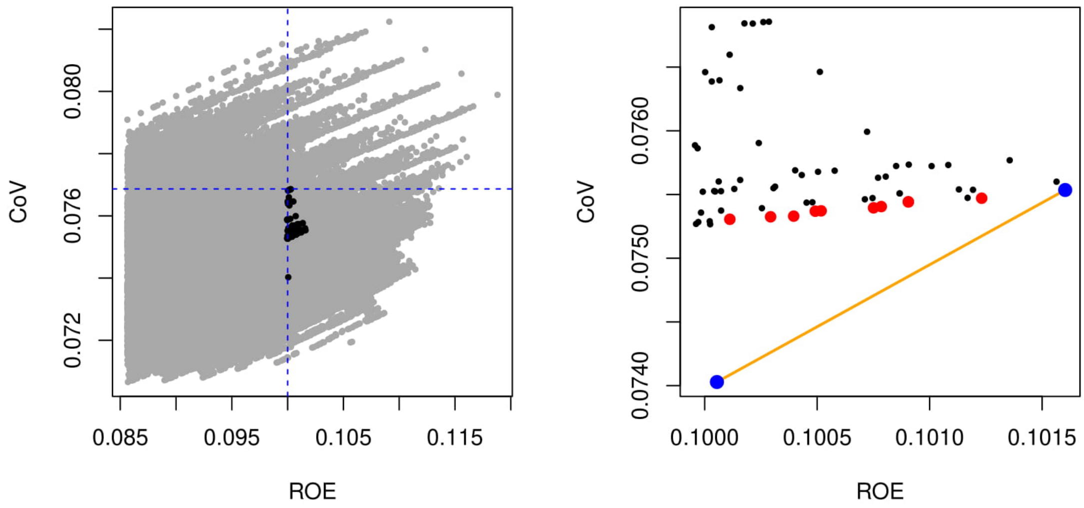


Figure 7.11: Optimization with Solvency ratio ≥ 160%

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.250.000	6.000.000	70%	100%	500.000	8.000.000	100%
2.000.000	4.000.000	65%	100%	500.000	4.000.000	100%

CoV	Solvency Ratio	E(ROE)
7,40%	160,14%	10,01%
7,55%	161,15%	10,16%

The results are interesting because it is the 3^{rd} time we meet the 1^{st} program on the tangent of convex frontiers. In fact, this program is characterized by an interesting trade-off between CoV and SCR, while satisfying ROE constraint in its lower bound. On the other hand, the 2^{nd} program recalls the one observed with Solvency Ratio ≥ 155%.

Since there are too few programs to compare given the current conditions, a good idea is to study what happens when the ROE constraint is less binding. The question is the following: is the 1^{st} program the tangent one when limitations change?

For this purpose, the case with $ROE \geq 9\%$ is proposed:

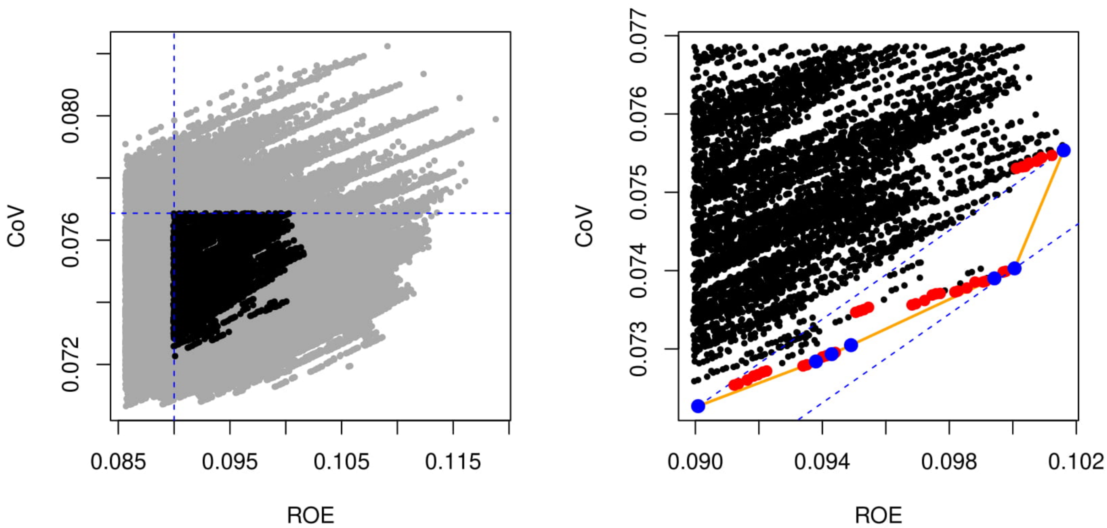


Figure 7.12: Optimization with Solvency ratio $\geq$ 160%, but with profit constraint $ROE \geq 9\%$ instead of 10%

The plot on the left of Figure (7.12) is the same of Figure (7.11), but with the vertical blue line shifted to the left. Now that way more programs are present, a proper convex frontier is determined and confirms that the 1^{st} program is the tangent one. It almost shares the tangency with its slightly alternative structure that we met before:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.250.000	6.000.000	70%	100%	500.000	9.500.000	100%

CoV	Solvency Ratio	E(ROE)
7,39%	160,35%	9,94%

Recalling the fact that the tails of the distribution might be underestimated in the fitting procedure, this program provides an useful alternative.

It is better to not go further with Solvency Ratio constraints, since $SCR \geq 165\%$ would include zero programs under $ROE \geq 10\%$.

7.3 Can reinstatements improve reinsurance?

Summarizing all the results obtained, the most recurring program that appeared in the reinsurance optimizations is:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.250.000	6.000.000	70%	100%	500.000	8.000.000	100%

CoV	Solvency Ratio	E(ROE)
7,40%	160,14%	10,01%

All the Excess of Loss treaties have an unlimited Aggregate Limit AL, which is not always true in practice. Usually, AL is a multiple of the limit L, and sometimes reinstatements are present. Assuming that the reinsurer is not willing to sell treaties with unlimited AL, we can re-structure the program with:

$$AL = 2 \cdot L$$

for both MTPL and GTPL. But, since GTPL is characterized by a long tail, $AL_{GTPL} = 2 \cdot 8 = 16$ is not high enough to provide a risk mitigation similar to unlimited AL case. In fact, the resulting program would have Solvency ratio equal to $158,04\%$.

To have something comparable, the intuition is to increase L_{GTPL} in order to have an aggregate limit that comprehends almost all the risk. From now on the following alternative treaty will be considered:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.250.000	6.000.000	70%	100%	500.000	9.500.000	100%

which have been proved to be on the convex frontier when the profitability constraint is $ROE \geq 9\%$ instead of 10%. The characteristics of the program under the 2 different AL assumptions are:

AL	CoV	Solvency Ratio	E(ROE)
∞	7,39%	160,35%	9,94%
$2L$	7,40%	159,73%	9,97%

We can see that the trade-off between Solvency ratio and expected ROE is not very convenient when passing from the unlimited scenario to the limited one. But, at least the 2^{nd} program is easier to find in the reinsurance market.

Now that we have layers with a finite AL, it is possible to study if the application of reinstatements on such layers can provide some beneficial effects to the direct insurer. As explained in the previous chapters, when dealing with paid reinstatements, the reinsurance premium is divided in two components:

- a deterministic premium, which is paid at the start of the contract;
- a stochastic premium, which is paid in case a loss hits the layer and the remaining AL decreases.

The advantage of such reinsurance clause is that, in case a small amount of losses occurs in the year, the premium paid is lower with respect to the case with free reinstatements (when there isn't any stochastic component in the premium). The stochastic premium is proportional to the percentage of layer to be reinstated and to a coefficient c_k, which in practice is equal to 50% or 100%.

On the other side, due to the fact that now also the premium is a random variable, the SCR and the CoV increase.

An algorithm is built to test all the possible splits of the aforementioned layers for both MTPL and GTPL, and for each of the resulting layers all the possible combinations of reinstatement (with $c_k = 0\%$, 50% and 100%) will be analyzed. In addition, also different scenario with $AL = L$ or $2L$ are inserted inside in case of no reinstatements. In this way, starting from one program we can extrapolate a massive amount of alternatives, and check where they lie in the various optimizations. We'll call this set of programs as *layered programs*.

In the following paragraph we will repeat the optimizations by comparing the previous efficient and convex frontier with the new alternatives to see if any new program is able to perform better.

Solvency Ratio and ROE optimization On the left of Figure (7.13) the optimization is represented for the overall programs as already seen in the previous sections. On the right the efficient and convex frontiers are compared to the set of layered programs depicted in green. As usual the constraints applied are Solvency ratio $\geq 150\%$, $ROE \geq 10\%$ and $CoV_{net} \leq CoV_{gross}$.

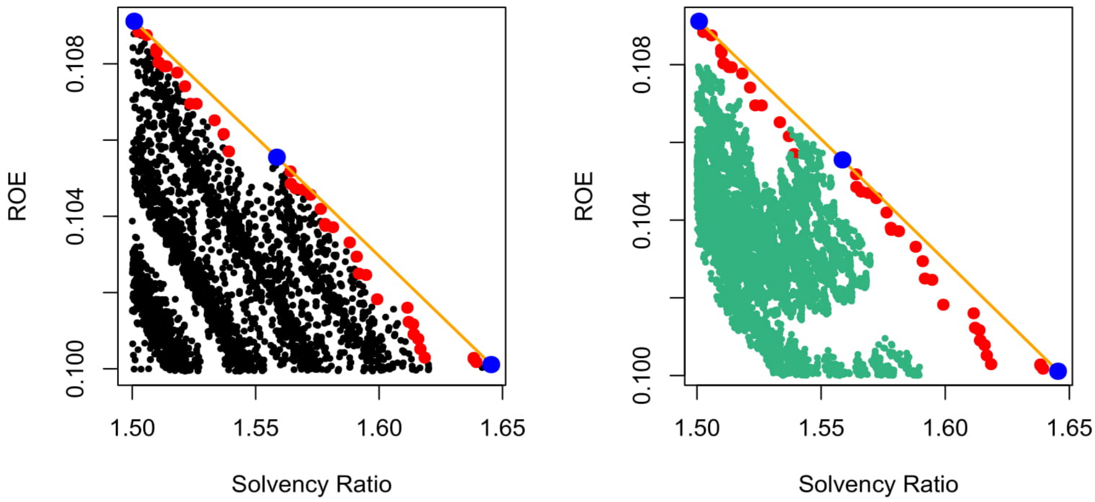


Figure 7.13

On the right plot we can see that no layered program lies on the convex frontier, and therefore, under this choice of axes, the layering procedure returns only sub-optimal programs. It means that using Excess of Loss contracts with $AL = \infty$ is the best choice in terms of ROE and Solvency ratio trade-off. The result of course depends on the property of the standard deviation pricing principle, which is characterized by a significant sub-additivity. In fact, due to this property two consequent layers cost more than an unique layer. Modifying the pricing assumptions might generate a layered program that is more efficient than the current programs on the convex frontier.

As seen before, the convex frontier is composed by programs that meet the CoV constraint in its upper bound. By setting a more binding constraint, the convex frontier would shift to the left and we might be able to observe that some layered programs are able to perform better than the tangent programs.

Solvency Ratio and CoV optimization As before, the optimization made in the previous section is shown on the left of Figure (7.14), while on the right the layered programs in green are displayed in comparison with the frontiers.

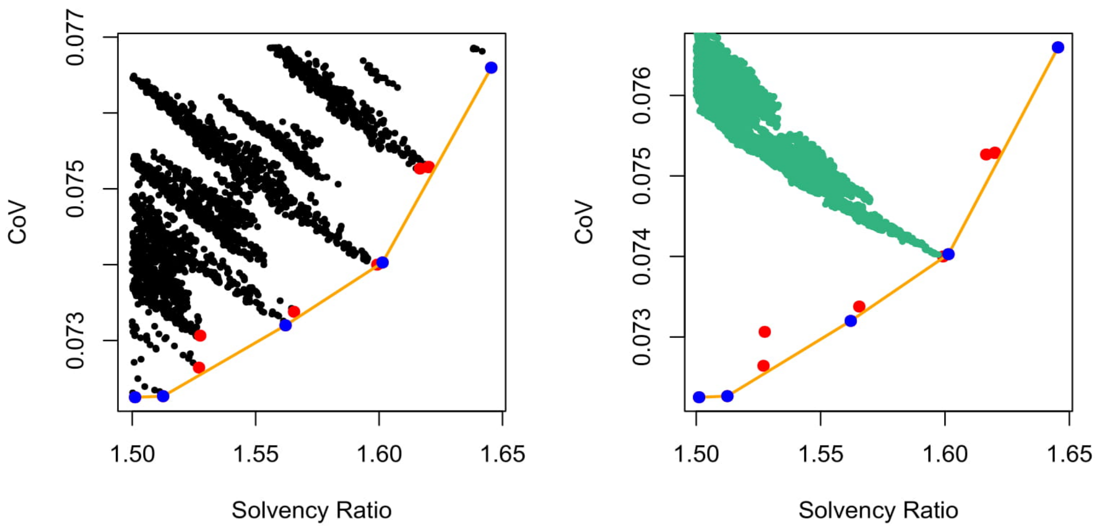


Figure 7.14

Since all the layered programs are derived from the same starting program, they are expressed as an extension of it under this choice of axes. In fact, the aim of the layering is to increase the profitability in exchange of increasing the risk due to lower protection. From the graphical point of view it is clear that every layered program is worsening both Solvency ratio and CoV. But, recalling that all the programs on the convex frontier meet the ROE constraint in its lower bound, the comparison between them and the layered programs is unfair since the latter are characterized by equal or higher profitability by definition. It is possible to increase the desired profit to investigate where the layered programs are located.

In Figure (7.15) the constraint $ROE \geq 10,5\%$ has been set, and a new convex frontier has been defined. Few programs (green) are now located under the tangent line, which is a sign that the layering procedure can improve reinsurance. Of course the *improvement* is measured in terms of trade-off between increased profit and decreased protection.

On the right plot, the light-green dot represents the best layered program in terms of distance from the convex frontier.

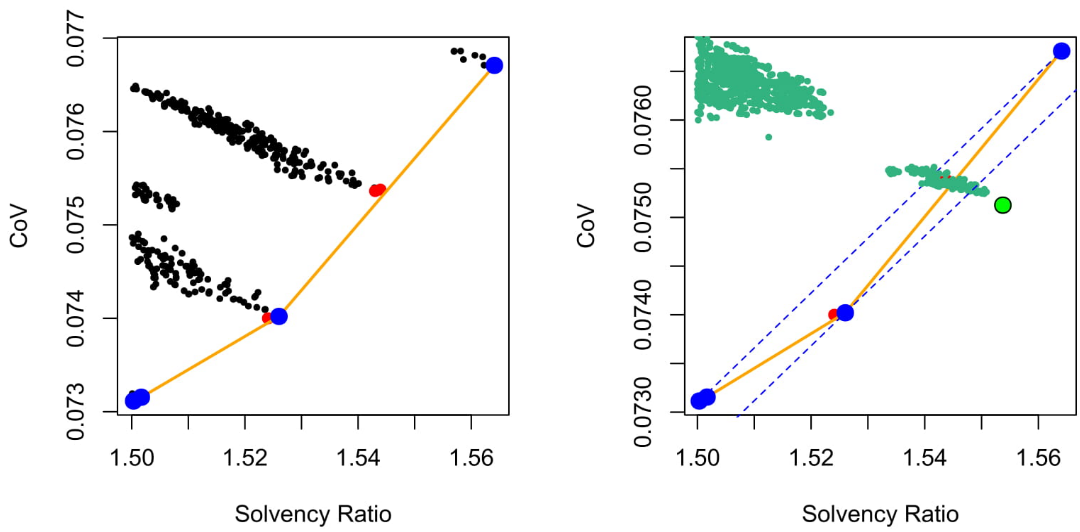


Figure 7.15

To have a context for the results, the tangent program (blue) is shown:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.250.000	4.000.000	75%	100%	500.000	6.000.000	100%

CoV	Solvency Ratio	E(ROE)
7,40%	152,60%	10,50%

While the best layered program achieves the following results:

CoV	Solvency Ratio	E(ROE)
7,51%	155,37%	10,51%

Note that both the tangent and the best layered program meet the *ROE* constraint in its lower bound, and making the profit constraint even more binding would exclude both of them. The difference between the two is given by a trade-off between CoV and Solvency ratio.

Recalling that the best layered program is based on:

D_{MTPL}	L_{MTPL}	α_{MTPL}	α_{MOD}	D_{GTPL}	L_{GTPL}	α_{GTPL}
1.250.000	6.000.000	70%	100%	500.000	9.500.000	100%

it is defined by the following additional characteristics (L and AL are in millions):

MTPL						GTPL					
1^{st} Layer			2^{nd} Layer			1^{st} Layer			2^{nd} Layer		
L	AL	c	L	AL	c	L	AL	c	L	AL	c
6	6	0%	-	-	-	9,5	19	100%	-	-	-

The best layered program doesn't rely on the split of the layer for both MTPL and GTPL due to the subadditivity of the pricing principle. Instead, it decreases AL_{MTPL} from 12 to 6 millions and buys 1 paid reinstatement to $GTPL$, leaving untouched AL_{GTPL}. The choice behind the reduction of AL_{MTPL} is that the probability of observing a reinsurance loss $\widetilde{X}^{re}_{MTPL} > 6.000.000$ is roughly 5% and the price of buying $AL_{MTPL} = 12.000.000$ is not particularly convenient. On the other side, $AL_{GTPL} = 19.0000.000$ is needed since the lob is long tail.

The presence of the reinstatement needs few numbers to be justified:

- if $c = 0$ the reinstatement is free, the reinsurance premium is fully deterministic and equal to 6.776.850;
- if $c = 100\%$ the reinstatement is paid and the initial deterministic premium is equal to 4.174.960. Due to the stochastic component, the insurer can pay at maximum 8.349.920 (2 times the initial premium since $c = 100\%$) in case $\widetilde{X}^{re}_{GTPL} \geq 9.500.000$;
- the paid reinstatement is more convenient than the free one only when $\widetilde{X}^{re}_{GTPL} \leq 9.500.000 \cdot \frac{4.174.960}{6.776.850} = 5.920.400$, where the corresponding probability is equal roughly to $67,90\%$.

ROE and CoV optimization Now we will deal with the last optimization, where the target Solvency ratio must be set in advance to extrapolate the results. First of all, in Figure (7.16) the constraint Solvency ratio $\geq 150\%$ is set.

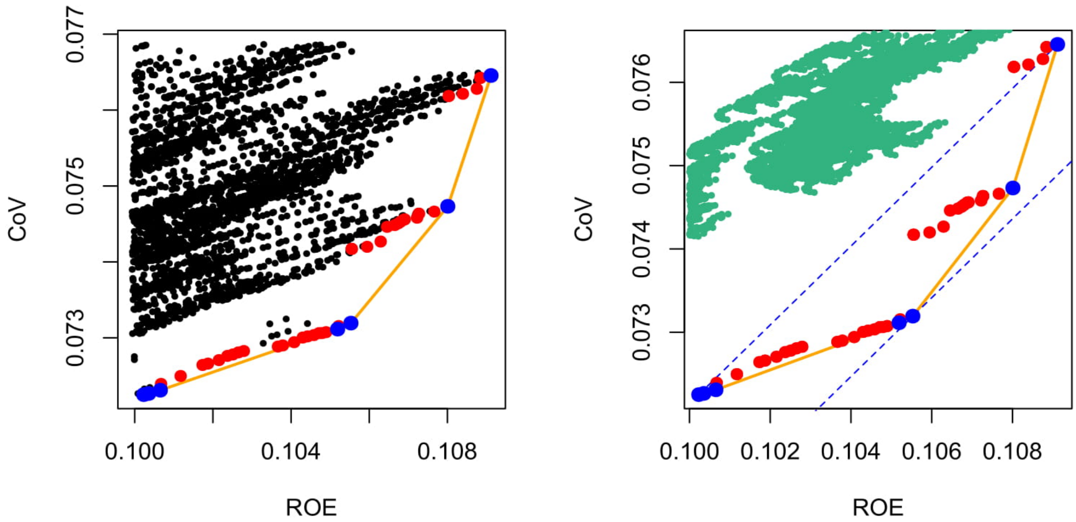


Figure 7.16

On the right plot we can see that all the layered programs are not even slightly near to the efficient frontier. The reason behind is that the starting optimal program have Solvency ratio = 160%, and the layering procedure doesn't reduce it to the lower bound 150%. Therefore, this target is inappropriate to compare the new set of programs with the old ones.

It is interesting to note that, if the chosen starting program achieved a lower Solvency ratio, through the layering procedure we would have observed some layered programs comparable under the this target. Of course, choosing a initial program on the convex frontier is suggested in order to obtain interesting layered outcomes.

It is important to underline that in this section we are only testing if layering can improve reinsurance, and not defining the best layered program among all the possible initial programs, since we are currently considering only a single initial program.

In Figure (7.17) the Solvency ratio target is set to 155%.

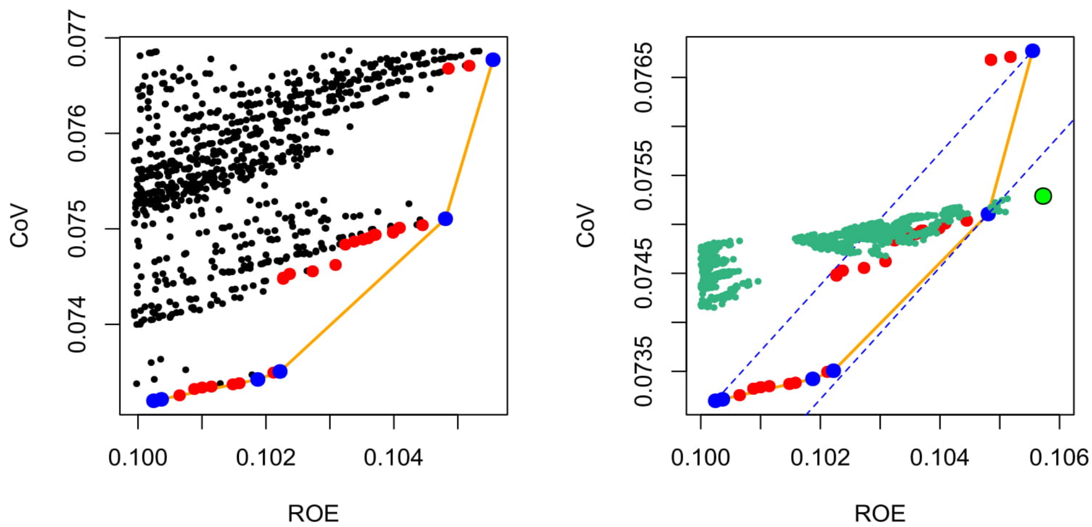


Figure 7.17

On the right plot, the light green dot represents the best layered program in terms of tangency. It is exactly the program found in the Solvency ratio and CoV optimization under $ROE \geq 10,5\%$. It is a good new that under two different optimizations the results are the same.

What if the insurer is willing to rely on splitting the same reinsurance layer among two different reinsurer? The following program is characterized by the split of MTPL layer, while still being under the tangency line:

MTPL						GTPL					
1^{st} Layer			2^{nd} Layer			1^{st} Layer			2^{nd} Layer		
L	AL	c	L	AL	c	L	AL	c	L	AL	c
2	4	0%	4	8	100%	9,5	19	100%	-	-	-

CoV	Solvency Ratio	E(ROE)
7,53%	155,08%	10,51%

Compared to the best layered program, it performs really similar to it in terms of CoV and expected ROE, but achieves a Solvency ratio of $155,08\%$

instead of 155, 37%. Therefore, it is almost dominated by the best layered program. To remain a good reinsurance scheme in terms of tangency, the GTPL layer should not be split, since the current reinstatement strategy is solid. Instead, the MTPL new split features an interesting reasoning:

- the first layer relies on the free reinstatements since, having a low attachment point, is will be hit by claims with high probability, and feature a stochastic premium wouldn't be convenient;
- instead, the second layer features one paid reinstatement since it covers only the most extreme claims, and the probability of full deterioration of such layer is low.

Since the starting program has a Solvency ratio = 159, 73% and the layering procedure can only decrease it, no layered program is present when dealing with the threshold 160% Solvency ratio.

Conclusions

In this thesis an optimal reinsurance simulation model has been implemented and explained in detail. The calibration of the lines of business has been based on the Italian insurance market to attribute more pragmatism to the results.

The reinsurance optimization has been quantified on profit and risk measures to investigate at best the insurer's economic situation. Through the comparison of different optimization criteria, one reinsurance scheme in particular has emerged from the wide set of simulated programs. The dimension of the latter has been drastically reduced by using the concepts of convex frontier and tangent program. The more constraints are set, the more the convex frontier is able to delete possible program candidates.

Exploring the benefits of the reinstatements' presence in the optimal program has highlighted that such clause, under adequate reinsurance design and pricing assumptions, is able to improve the risk return trade-off of the insurer.

Since insurance companies already have a reinsurance program, the proposed methodology can be extended to improve the current scheme without any drastic change from year to year: the set of possible reinsurance programs can be defined by applying small shocks to the parameters of the current reinsurance contracts. Analyzing the resulting programs under the adequate optimization procedures can suggest which modifications are more relevant to restructure the program. Ironically, this alternative version of the algorithm requires significantly lower computational time with respect to the one introduced in this thesis, and it can be easily implemented to fit the insurance companies' needs.

Acknowledgements

I would like to express my gratitude towards my thesis supervisor, Professor Gian Paolo Clemente, for the incisive guidance in the process of this thesis. His ability to understand and encourage creativity is one of the main reasons for which this thesis has come to life.

I am grateful to my family for the never-ending support over the years. In particular to my mother Marina for introducing me to the actuarial world, where I always feel stimulated. And I thank my brother Andrea for sharing with me so many interests and insightful conversations.

I would also like to acknowledge the help that both old and new friends gave to me during this short but intense path. I truly appreciate it, thank you.

Bibliography

[1] ANIA (2019). L'assicurazione Italiana 2018-2019.

[2] Antal, P. (2009). Mathematical Methods in Reinsurance. Swiss Reinsurance Company

[3] Bedford, T., Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial Intelligence 32, 245–268.

[4] Brechmann, E.C., Schepsmeier, U. (2013). Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software, Volume 52, Issue 3

[5] Buhlmann, H. (1996). "Mathematical Methods in Risk Theory", second ed. Springer-Verlag, Berlin.

[6] Campana, A.: Reinstatement premiums for an excess of loss reinsurance contract. Giornale dell'Istituto Italiano degli Attuari LXIX (2006) 55-68

[7] Canestraro D., Dacorogna M., 2010, The influence of risk measures and tail dependencies on capital allocation, SCOR Paper #7, available on http://www.scor.com/en/sgrc/scor-publications/scor-papers.html

[8] Clark, David R. 1996. Basics of Reinsurance Pricing. CAS Study Notes 41–43: 1–52.

[9] Clemente, G. P., The Effect of Non-Proportional Reinsurance: A Revision of Solvency II Standard Formula, RISKS, 2018; 6 (2): 1-13. Available at http://hdl.handle.net/10807/119209

[10] Clemente, G. P., Savelli, N., Zappa, D., The Impact of Reinsurance Strategies on Capital Requirements for Premium Risk in Insurance, RISKS, 2015; (Marzo): 164-182. Available at http://hdl.handle.net/10807/67288

[11] Daykin, Chris D., Pentikeainen, Teivo; Pesonen, Martti (1994). "Practical Risk Theory for Actuaries", Chapman and Hall, London.

[12] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D. (2002) The concept of comonotonicity in actuarial science and fi nance: theory. Insurance: Mathematics and Economics 31(1), 3-33.

[13] Dutang, C., Goulet, V., and Pigeon, M. (2008). actuar: An R package for actuarial science. Journal of Statistical Software, 25, 7, 1–37. Available at http://www.jstatsoft.org/v25/i07

[14] EIOPA. 2017. Consultation Paper on EIOPA's First Set of Advice to the European Commission on Specific Items in the Solvency II Delegated Regulation. Technical Report. Frankfurt.

[15] EIOPA. 2018. Consultation Paper on EIOPA's Second Set of Advice to the European Commission on Specific Items in the Solvency II Delegated Regulation. Technical Report. Frankfurt.

[16] European Commission. 2015. Commission Delegated Regulation (EU) 2015/35 supplementing Directive 2009/138/EC of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), 10 of October 2014. Official Journal of the EU 58: 1–797

[17] Guerra, M. and Centeno, M.L. (2012) Are quantile risk measures suitable for risk-transfer decisions? Insurance: Mathematics and Economics, 50(3), 446–461.

[18] Kaas, R., GoovaertsJan, M., Denuit, D. (2008). Collective risk models. In: Modern Actuarial Risk Theory. Springer, Berlin, Heidelberg

[19] Krämer, N., Schepsmeier, U. (2011). Introduction to vine copulas. NIPS Workshop, Granada.

[20] Mata, A.J. (2000). Pricing excess of loss reinsurance with reinstatements. ASTIN BULLETIN, Vol. 30, No. 2, 2000, pp. 349-368.

[21] Nelsen, R.B. (2006). An Introduction to Copulas. 2nd edition, Springer, New York

[22] Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions. ASTIN Bulletin (International Actuarial Association), 12, 1, 22–26.

[23] Parodi, P. (2014). Pricing in General Insurance. A Chapman & Hall book. Taylor & Francis

[24] Savelli, N.; Clemente, G.P. Modelling Aggregate Non-Life Underwriting Risk: Standard Formula vs Internal Model. Giornale dell'Istituto Italiano degli Attuari 2009, LXXII, 301–338.

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

[26] Sundt, B. (1991) On excess of loss reinsurance with reinstatements. Bulletin of the Swiss Association of Actuaries, Heft 1, 51-65

[27] Swiss Federal Office of Private Insurance (2006) Technical document on the Swiss Solvency Test. FINMA

[28] Walhin, J.F., Paris, J.: Excess of loss reinsurance with reinstatements: premium calculation and ruin probability of the cedent. Proceedings of the third IME Congress, London (1999)

[29] Wang, S. (1995) Insurance Pricing And Increased Limits Ratemaking By Proportional Hazards Transforms. Insurance: Mathematics and Economics,

[30] Young V.R. (2004) Premium Principles. In Encyclopedia of Actuarial Science, vol. 3, ed. J. Teugels and B. Sundt. John Wiley & Sons.