

Faculté des sciences

Modelling of a collaborative bicycle insurance and description of the risk sharing mechanism based on the collaborative bicycle insurance Laka

Author: **Maxim MIZRAHI**
Supervisor: **Michel DENUIT**
Reader: **Cindy COURTOIS**
Academic year 2022–2023
Master [120] en sciences actuarielles

Acknowledgment

I would like to thank all the people who, in one way or another, have helped and supported me throughout my academic career.

I would also like to thank Michel Denuit for not only agreeing to be my supervisor for this master thesis, but also for the time and energy he devoted to it. His observations and comments were all very pertinent.

I also wish to mention Stéphanie Brixhe, who helped me conduct the interview that added so much value to my master thesis.

Finally, I want to thank Allianz for providing me with their rates, without which the data required for this master thesis would not have been available.

To conclude, I have a special dedication to my parents, who have supported me throughout my entire academic career.

Abstract

This master thesis undertakes an in-depth examination of collaborative bicycle insurance, focusing primarily on the pioneering model presented by Laka. From a methodological point of view, a rigorous path is mapped out, involving meticulous calculations and the construction of portfolios of insureds using uniform and binomial distributions. The master thesis is further enriched by empirical insights from an illuminating interview with GRACQ.

As the master thesis progresses, it becomes clear that collaborative insurance presents significant advantages and related challenges. Ethical considerations are omnipresent, juxtaposed with the art of judicious cost management. The crucial importance of well-founded databases is emphasised, while avenues for future research are outlined. In sum, this master thesis is a substantive scholarly contribution, shedding deep light on the potentially transformative role of collaborative bicycle insurance within the broader field of insurance paradigms.

Contents

1	Introduction	1
1.1	Peer-to-peer insurance & Mutual insurance	2
1.2	Laka	3
1.2.1	General information about Laka	3
1.2.2	Technical aspects of Laka	4
2	Theory	7
2.1	Methodology	7
2.2	Data Set	7
2.3	Distribution	8
2.4	Stop-Loss	10
2.5	Conditional Mean Risk Sharing	11
2.6	Pricing	19
2.7	Surplus	25
2.8	Variance of the CMRS	27
3	Preparation	30
3.1	Data Set	30
3.2	Portfolio of insureds	32
4	Case Studies	33
4.1	Uniform distribution	34
4.2	Binomial (0,5) distribution	41
5	Interview with GRACQ : Perceptions and thoughts on collabora- tive bike insurance	47
6	Conclusion	49
7	Annexe	52
7.1	Graphs	52
7.2	Tables	60

7.3	Interview with GRACQ	66
7.4	R code	70

Chapter 1

Introduction

Nowadays, conventional insurance is increasingly frowned upon. The first reason is that people may feel that they are paying more than they are receiving in benefits without necessarily understanding the hidden costs behind an insurance premium which can lead to confusion or misunderstandings. Since policyholders don't fully understand the complexity of insurance companies, they can also have a perception of unfairness because they are paying for coverage that they may never need, while others may receive payouts for claims. The second reason is that insurance companies are businesses that aim to make a profit, which can sometimes conflict with the interests of policyholders.

For the reasons mentioned above some policyholders are now considering collaborative insurance, however, this is still a relatively little-used option. A collaborative insurance is in fact a fund supplied by the insureds themselves allowing its insureds to use this fund in case a predefined event would occur. Obviously, this description is very simplistic and in reality there are many more parameters that come into play (an intermediary who manages the fund, a reinsurer, etc...).

This type of insurance is more and more appreciated by the general public because it guarantees an almost total transparency between the insured and the insurer, who are in a way the insured. This can help to solve the problem mentioned above because the costs of each premium are usually very detailed in collaborative insurances, which allows the insured to understand why they are paying and there will be less of a feeling of unfairness. In addition, it allows policyholders to be more cautious and try to have as few claims as possible because this would increase the cost of the fund and therefore increase the premiums for each policyholder. Moreover, P2P insurances are often less expensive than a classic insurance. However, this is not always the case, collaborative insurance models rely on a community of members to pool their resources and share risk. If the community is small or lacks

diversity, the risk may be concentrated, and premiums may be higher than expected.

Obviously, P2P insurance has other advantages and shortcomings than those mentioned above. Therefore the goal of this master thesis is to model a collaborative bicycle insurance and describe its risk sharing based on the P2P bike insurance Laka to try to understand how companies like Laka are viable and how they work. This master thesis is divided into four parts. The first deals with the general definition of P2P insurance and its comparison with mutual insurance. Then, the methodology will be explained in detail, using a simplified dataset. In the third part, the methodology will be applied to more realistic datasets, to see the impact on premiums and surpluses. Finally, an interview will be conducted with a group of cyclists, the GRACQ, to discuss collaborative bicycle insurance.

1.1 Peer-to-peer insurance & Mutual insurance

P2P insurance often entails the use of a platform to link people who are interested in insuring comparable risks. These people make financial contributions to a pool that are used to settle claims if one of the participants has a loss (Jake Frankenfield (2021) [2]).

Contrarily, the mutual insurance business model has been in use for many years. Owners of mutual insurance businesses are its policyholders, who also choose the board of directors that will lead the business. The company receives premiums from policyholders, which are then used to pay claims and administrative costs. All gains made by the business are distributed to policyholders as dividends or lower premiums (Julia Kagan (2021) [3]).

Although policyholders pool their premiums and share risks in both P2P and mutual insurance, there are some significant distinctions between the two models. The following are some of the main variations:

- **Ownership:** There is typically no established ownership structure in peer-to-peer insurance. The company that owns the platform that connects policyholders could be for-profit or charity. With mutual insurance, customers own the insurance provider.
- **Control:** In P2P insurance, the platform typically has the final say about the handling of claims and the administration of the insurance plan. Policyholders in mutual insurance choose the board of directors, which is responsible for making these choices.

- Risk Selection: P2P insurance systems often include eligibility requirements, and they reserve the right to refuse users depending on their risk profiles. All policyholders are accepted under mutual insurance, and premiums are frequently determined using a communal rating system that does not take individual risk into consideration.

The following two examples illustrate these two types of insurance. Firstly, there is Lemonade which is an example of a P2P insurance platform. An AI-powered platform is used by Lemonade, a tech-based insurance provider, to connect policyholders who want to pool their premiums and insure similar risks. The premiums that policyholders pay go into a pool that is used to settle claims when a member suffers an insured loss. For its services, Lemonade charges a flat rate, and any overage is donated to charity through its Giveback program. In the US, Germany, and the Netherlands, Lemonade provides insurance coverage for renters, homeowners, and pet owners (*Lemonade Insurance* [9]).

Then, to illustrate mutual insurance, State Farm turns out to be a good example as it provides a variety of insurance policies, such as life, house, and auto insurance. State Farm’s policyholders, who choose a board of directors to lead the business, own it. The company receives premiums from policyholders, which are then used to pay claims and administrative costs. All gains made by the business are distributed to policyholders as dividends or lower premiums. One of the biggest insurance providers in the country, State Farm has been around for almost 100 years (*State Farm Insurance* [10]).

1.2 Laka

1.2.1 General information about Laka

Since my master thesis is based on Laka, it is interesting to present this insurance. Therefore, Laka is a peer-to-peer (P2P) bicycle insurance with headquarters in London. It was established in 2017 and now has operations in Germany, the Netherlands, Belgium, France, and the United Kingdom (*Laka Collective Fietsverzekeringen* [4]).

Laka’s insurance strategy is the collective coverage theory, in which policyholders join a pool to share the cost of claims. Laka employs an independent claims assessment to calculate the reimbursement when a member submits a claim. Each collective member contributes a portion of the claim’s cost through a separate monthly payment after it has been distributed among all members. By doing this, Laka’s policyholders avoid paying for the costs associated with a traditional

insurance firm and instead pay the cost of claims that arise inside their group.

The fact that policyholders' monthly contributions are limited to a specific sum, known as the maximum premium, is one of the distinctive aspects of Laka's P2P model. It appears that Laka insures the high risk portion of their policy portfolio, which allows them to assure their policyholders that their premium will not exceed a certain amount.

Moreover, Laka gives its policyholders a number of advantages, including:

- No excess payments: Unlike most conventional insurance plans, this policy does not require policyholders to pay an excess fee or a deductible when filing a claim. The following example shows how an excess fee works : you have car insurance with a 500€ excess fee. If you get into an accident and the cost to repair your car is 2000€, you would be responsible for paying the first 500€, and the insurance company would cover the remaining 1500€.
- Global coverage: Laka's policy covers bicycle equipment and incidents that occur anywhere in the world.
- Digital claims: Via the Laka mobile app, policyholders can file claims and get payments within hours.
- Overall, by utilizing a P2P model that enables policyholders to split the expense of claims and pay a reasonable price for their coverage, Laka provides a distinctive and novel approach to insurance.

It offers guarantees for theft, damage, vandalism, travel abroad, on or in your car and a return home service (in case you are stuck in the middle of nowhere with a broken bike).

1.2.2 Technical aspects of Laka

When an estimate with Laka is carried out, Laka gives details of the maximum premium to be paid monthly. Indeed, the premium varies each month according to the number of claims the collective has suffered during the previous month but will never exceed the maximum given by Laka. The premium is calculated as follows (*Laka Collective Fietsverzekeringen* [4]):

- a share of the total number of claims reported by the collective
- a compensation of 25% for Nationale-Nederlanden (and for Laka)

- an insurance tax of 9,25%.

For example, if someone insured his 1900€ worth electric bike, the insured will never pay more than 10€ per month. This amount is made up of the following elements (*Laka Collective Fietsverzekeringen* [4]):

- A premium for the previous month's claims: 7€ (this is the maximum we can participate in the previous month's claim)
- Compensation for Laka and Nationale-Nederlanden (insurance): 1,75€ ($7 \times 25\%$)
- Insurance tax: 0,81€ ($((7 + 1,75) \times 9,25\%)$)

It thus appears that the maximum total monthly premium will be 9,56€, which represents approximately a maximum premium of 120€ per year. This premium covers theft, damage and breakdown.

In addition, Laka sends a summary at the end of each month that clearly indicates the premium for the month's claims and also breaks down what the premium consists of for that month. As it can be seen, the advantage of a platform like Laka is the transparency as well as its very attractive prices. In other words, the insured knows clearly how his premium is composed and where his money goes. Moreover, as Laka offers monthly premiums, the insured can stop insuring his bike at any time of the year. There is therefore a great flexibility with Laka, which is not found with traditional insurers who generally offer contracts on an annual basis. On the other hand, the disadvantage of this type of insurance is the uncertainty. Indeed, despite the assurance that the premium of the insured will not exceed a certain amount, the premium is still variable and depends on the other insureds, which could scare some risk-averse insureds who will rather turn to a classic insurance, where the premiums are fixed and annual.

As a comparison, the previous example of the 1900€ electric bicycle can be used. As it has been discussed previously, the maximum annual premium that Laka can ask is 120€. Nevertheless, this is a maximum premium, however, it is very unlikely that every month the customer has to pay the maximum premium of 10€. For Allianz, the annual premium for the same bike is 80€. In both cases, the premium covers theft, damage and breakdown. However, the guarantees that Laka offers are much more extensive. The main differences are explained underneath :

- Regarding bicycle theft, Allianz amortizes the price of the bicycle, which means that the bicycle loses value every year as it ages. This means that if

the bike is stolen, the insured will only be reimbursed for the purchase price of the amortized bike. Whereas Laka pays for replacement costs up to the insured amount stated in the policy.

- The following elements are not covered by Allianz but are covered by Laka: casing, cables, bell, battery and other removable elements.
- There is a deductible of 25€ per damage caused to the bike, which can go up to 50€ depending on the type of bike. This is the case for Allianz but not for Laka.
- It is important to highlight that for some types of bikes at Allianz, in addition to the amortization of the price of the bike, there is a deductible of 20% of the price of the bike in case of theft.
- In addition, there is a possibility of retraction at Laka : as discussed previously, the premiums at Laka are monthly, which means that the insured can decide to stop the contract at any time to insure his bike during the year. This is obviously impossible with Allianz because the premiums are annual.

The maximum premium offered by Laka is therefore higher than the one offered by Allianz. However, as it has been demonstrated earlier, the premium is actually much lower. In fact, 120€ is the maximum annual premium. Moreover, the guarantees are more extensive with Laka, which also provides a flexibility that cannot be found with Allianz.

Since a more detailed understanding of Laka has been provided, the methodology used will be explained in order to do an actuarial modelling of Laka and describe the risk-sharing mechanism. The methodology will be explained using a simplified dataset, but based on realistic assumptions. The same method will then be used again, albeit with much more realistic datasets.

Chapter 2

Theory

2.1 Methodology

Regarding the way I wanted to proceed with my thesis, I wanted to base my thesis on the article *allocation of benefits in mutual aids and survivor funds* (Denuit & Robert, 2022) which describes risk sharing and how to calculate premiums in a mutual aid fund. However, this has proved to be problematic. In fact, in the article mentioned previously, it is estimated that the people, who are participating financially in the mutual aid fund and have suffered a pre-defined event, share the totality of the fund. However, in my study, only the price of the bike is returned to the insured, which is called a compensation repair and the fund is not divided between the insureds who have suffered a theft. Otherwise, they would risk receiving less or more than the price of their bike. It is therefore impossible to apply this article to my master thesis.

As an alternative, I will opt for a pure peer-to-peer approach based on the course LACTU2010 (Michel Denuit (2022) [5]) where the amount to be paid by the participants at the end of the period is not known and then calculate the upper risk layer of each of the portfolios I have created and have this upper risk layer insured by a classical insurer.

2.2 Data Set

The following simplified dataset is used to present the methodology. It will be explained in the following section (Preparation) how these theft frequencies are calculated.

Price	Proportion	Theft frequency
1000€	25%	3%
2000€	50%	2,5%
3000€	25%	1,5%

The previous proportions for $n = 100, 1000$ and 10000 insureds will be applied in order to examine how the distribution of costs and risks behaves.

The loss of a person i is denoted by X_i . The loss for each category of insured is therefore defined as follows :

$$X_i = \begin{cases} 1000 & \text{with a probability of 3\%} \\ 0 & \text{with a probability of 97\%} \end{cases} \quad i = 1, \dots, 0.25n$$

$$X_i = \begin{cases} 2000 & \text{with a probability of 2,5\%} \\ 0 & \text{with a probability of 97,5\%} \end{cases} \quad i = 0.25(n+1), \dots, 0.75n$$

$$X_i = \begin{cases} 3000 & \text{with a probability of 1,5\%} \\ 0 & \text{with a probability of 98,5\%} \end{cases} \quad i = 0.75(n+1), \dots, n$$

The total loss of the pool with n insured is denoted as following :

$$S_n = \sum_{i=1}^n X_i$$

2.3 Distribution

The whole methodology has been done on R using the `distr` package available in R (Peter Ruckdeschel, Matthias Kohl, Thomas Stabla, Florian Camphausen (2022) [11]). The R code is available in the appendix.

To model the portfolio of insureds, a distribution composed of several binomial distributions is used, with one binomial distribution per category of insureds. To exemplify, for the bikes worth 1000€, the following binomial distribution is applied:

$$Bin(n * 25\%, 3\%)$$

n = Number of insureds

The three binomial random variables are convolved together, as there are three different price categories, to obtain the portfolio of insureds. To convolve these 3 binomial distributions, the `distr` package uses a convolution algorithm. In order to understand how this convolution algorithm works an example with two binomial distributions is provided. The convolution algorithm works by computing the convolution of two binomial distributions $F_1 = \text{Bin}(k_1, p_1)$ and $F_2 = \text{Bin}(k_2, p_2)$ to obtain their sum $F_1 + F_2$. The convolution of two probability distributions is defined as the distribution of the sum of two independent random variables, each distributed according to one of the original distributions.

To compute the convolution of F_1 and F_2 , the probability mass function (pmf) of the sum of two independent random variables X_1 and X_2 needs to be computed, where X_1 is distributed according to F_1 and X_2 is distributed according to F_2 . The pmf of the sum of two independent random variables is given by the convolution of their pmfs, which is defined as:

$$(F_1 + F_2)(n) = P(X_1 + X_2 = n) = \sum P(X_1 = i)P(X_2 = n - i)$$

where the sum is taken over all possible values of i such that $0 \leq i \leq k_1$ and $0 \leq n - i \leq k_2$.

In the case of binomial distributions, the pmf is given by the binomial formula:

$$P(X = i) = \binom{k}{i} * p^i * (1 - p)^{(k-i)}$$

where $\binom{k}{i}$ is the binomial coefficient, which counts the number of ways to choose i successes out of k trials.

Substituting the binomial formula into the convolution formula :

$$(F_1 + F_2)(n) = \sum \binom{k_1}{i} * p_1^i * (1 - p_1)^{(k_1-i)} * \binom{k_2}{n-i} * p_2^{(n-i)} * (1 - p_2)^{(k_2-n+i)}$$

where the sum is taken over all possible values of i such that $0 \leq i \leq k_1$ and $0 \leq n - i \leq k_2$.

This formula can be computed efficiently using the Fast Fourier Transform (FFT) algorithm, which is implemented in the R package `distr`. The resulting distribution $F_1 + F_2$ is also a binomial distribution with parameters $\text{Bin}(k_1 + k_2, p_1 + p_2)$ (Peter Ruckdeschel, Matthias Kohl (2012), [12]).

To add more than two binomial distributions together, the convolution operator can be used repeatedly.

After convolving the three binomial distributions, Figure 5 is obtained, it displays the distribution functions of the total cost according to the number of insureds.

For extreme values, it is impossible to calculate conditional mean risk sharing, as the probabilities are too close to zero (as can be seen on figure 5). The whole methodology part will be implemented on a range between two quantiles of levels $\frac{\epsilon}{2}$ and $1 - \frac{\epsilon}{2}$ with $\epsilon = 10^{-6}$.

2.4 Stop-Loss

In order to insure the high layer of the risk like the one covered by the Nationale-Nederlanden for Laka, the pure premiums will be calculated. Insuring the high layer of the risk helps to balance the risk pool, as it ensures that the cost of insurance is limited for all members.

Additionally, by insuring the high layer of the risk, it will avoid unpleasant surprises and assure the policyholders that their premium will never exceed a certain amount.

Then, the Stop-Loss pure premium is calculated as follows :

$$\mathbb{E}[(X - K)_+]$$

$$K = \text{Stop-Loss Strike}$$

The stop-loss premium is computed as follows for a portfolio of 100 policyholders and a strike K :

$$\mathbb{E}[(S_{100} - K)_+] = \sum_{k=1}^{200} (k1000 - K)_+ P(S_{100} = k1000)$$

And the exact same method is applied for different strikes and different policyholder portfolios. Figure 6 displays the pure premium versus the strike.

2.5 Conditional Mean Risk Sharing

Conditional mean risk sharing is a type of risk sharing arrangement where individuals or entities agree to share the risks they face in a certain situation. $h_i^{cm}(S)$ defines the participation of the insured i facing the total loss of the pool S . The participation of the insured i under the conditional mean risk sharing rule is :

$$h_i^{cm}(S) = \mathbb{E}[X_i|S]$$

This means that each party will bear a portion of the risk, but the total risk across all parties will be fully shared because :

$$\sum_{i=1}^n h_i^{cm}(S) = \sum_{i=1}^n \mathbb{E}[X_i|S] = S$$

Conditional mean risk sharing allocates the full risk and it ensures that each party bears a fair share of the risk based on their expected loss :

$$\mathbb{E}[h_i^{cm}(S)] = \mathbb{E}[\mathbb{E}[X_i|S]] = \mathbb{E}[X_i]$$

By agreeing to share the risk, each party is effectively pooling their resources, which allows them to collectively absorb the full impact of the risk.

Overall, the conditional mean risk sharing rule appears to be a reasonable requirement because it is Pareto-optimal within large pool, it ensures that participants have an interest in keeping the total loss of S_n as small as possible and it also satisfies the following four axioms :

- Axiom AF (Actuarial fairness) : The expected value of the allocation for each agent is equal to the expected value of the initial risk. $\mathbb{E}[h_i(X_n)] = \mathbb{E}[X_i]$ for all $i = 1, 2, \dots, n$
- Axiom RF (Risk fairness) : Each agent's allocation should remain within the limits of their maximum potential loss. $h_i(X_n) \leq \max[X_i]$ for all $i = 1, 2, \dots, n$
- Axiom RA (Risk anonymity) : The value that each agent receives from the allocation is determined by the total risk's realized value. $h_i(X_n) = h_i(S_n)$ for every $i \in \{1, 2, \dots, n\}$
- Axiom OA (Operational anonymity) : The allocation to a single agent remains unchanged even when the risks of two other agents merge. $h_i(X_n)$ is determined only by the pair (X_i, S_n) for each i and X_n .

Formal proof for previous axioms can be found in Jiao, Z., Liu, Y., & Wang [8].

The following example illustrate how the calculations are made. In this example, the insured number 1 has a bike costing 1000€, there are a thousand insured in the portfolio and a total cost (S) of 50000€. The possible loss of the insured 1 is :

$$X_1 = \begin{cases} 1000 & \text{with a probability of 3\%} \\ 0 & \text{with a probability of 97\%} \end{cases}$$

$$\begin{aligned} h_1^{cm}(50000) &= \mathbb{E}(X_1 | S = 50000) \\ &= 1000 * P(X_1 = 1000 | S = 50000) + 0 * P(X_1 = 0 | S = 50000) \end{aligned}$$

$$\begin{aligned} h_1^{cm}(50000) &= 1000 * P(X_1 = 1000 | S = 50000) \\ &= 1000 * \frac{P(X_1 = 1000 \cap S = 50000)}{P(S = 50000)} \end{aligned}$$

$$h_1^{cm}(50000) = 1000 * \frac{P(X_1 = 1000) * P(S - X_1 = 49000)}{P(S = 50000)}$$

The probability that the total loss of the portfolio without X_1 is equal to 49000€ and that the total portfolio is equal to 50000€ is equal to :

$$P(S - X_1 = 49000) = 3,406493\%$$

$$P(S = 50000) = 3,179133\%$$

It results that :

$$h_1^{cm}(50000) = 1000 * \frac{0,03 * 0,03406493}{0,03179133} = 32,14549$$

The calculation are carried out for all the costs of the number of insured (with a different S) and for each category of insured (with a different X). These same calculations are done for the 3 portfolios of insureds (100, 1000 and 10000 insured). For each category of insured, the outcomes are :

$$h_i^{cm}(s) = 1000 * \frac{0,03 * P(S - X_i = s - 1000)}{P(S = s)} \quad i = 1, \dots, 0.25n$$

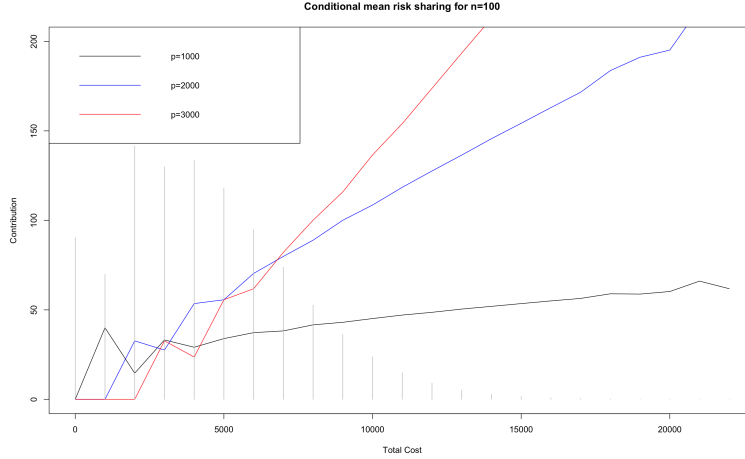


Figure 1: Conditional Mean Risk Sharing for $n = 100$

$$h_i^{cm}(s) = 2000 * \frac{0,025 * P(S - X_i = s - 2000)}{P(S = s)} \quad i = 0.25(n + 1), \dots, 0.75n$$

$$h_i^{cm}(s) = 3000 * \frac{0,015 * P(S - X_i = s - 3000)}{P(S = s)} \quad i = 0.75(n + 1), \dots, n$$

And figure 1 is plotted (the densities of the distribution are displayed on the graphs in grey, they are in per thousand).

As shown in figure 7, a strange behavior for small values of S appears. First, what stands out is that the CMRS for the 2000€ and 3000€ bikes are equal to 0 when $S = 1000$. This is quite normal because if $S=1000$, it means that it is necessarily a bike that cost 1000€ that has suffered a theft, it is not possible that a bike at 2000€ or 3000€ has suffered a theft. So :

$$h_{26-75}^{cm}(1000) = h_{76-100}^{cm}(1000) = h_{76-100}^{cm}(2000) = 0$$

The calculations for small values of S have to be checked to see if the error comes from the calculations:

$$\begin{aligned} h_{1-25}^{cm}(1000) &= 1000 * \frac{0,03 * P(S - X_{1-25} = 1000 - 1000)}{P(S = 1000)} \\ &= 1000 * \frac{0,03 * 0,09303918}{0,06977944} = 39,99997 \end{aligned}$$

$$\begin{aligned}
h_{1-25}^{cm}(2000) &= 1000 * \frac{0,03 * P(S - X_{1-25} = 2000 - 1000)}{P(S = 2000)} \\
&= 1000 * \frac{0,03 * 0,06906001}{0,1416002} = 14,63134
\end{aligned}$$

$$\begin{aligned}
h_{1-25}^{cm}(3000) &= 1000 * \frac{0,03 * P(S - X_{1-25} = 3000 - 1000)}{P(S = 3000)} \\
&= 1000 * \frac{0,03 * 0,1438436}{0,1299599} = 33,20492
\end{aligned}$$

$$\begin{aligned}
h_{26-75}^{cm}(2000) &= 2000 * \frac{0,025 * P(S - X_{26-75} = 2000 - 2000)}{P(S = 2000)} \\
&= 2000 * \frac{0,025 * 0,09256193}{0,1416002} = 32,68425
\end{aligned}$$

$$\begin{aligned}
h_{26-75}^{cm}(3000) &= 2000 * \frac{0,025 * P(S - X_{26-75} = 3000 - 2000)}{P(S = 3000)} \\
&= 2000 * \frac{0,025 * 0,0715685}{0,1299599} = 27,53485
\end{aligned}$$

$$\begin{aligned}
h_{26-75}^{cm}(4000) &= 2000 * \frac{0,025 * P(S - X_{26-75} = 4000 - 2000)}{P(S = 4000)} \\
&= 2000 * \frac{0,025 * 0,1428573}{0,1334973} = 53,50568
\end{aligned}$$

$$\begin{aligned}
h_{76-100}^{cm}(3000) &= 3000 * \frac{0,015 * P(S - X_{76-100} = 3000 - 3000)}{P(S = 3000)} \\
&= 3000 * \frac{0,015 * 0,0944568}{0,1299599} = 32,70668
\end{aligned}$$

$$\begin{aligned}
h_{76-100}^{cm}(4000) &= 3000 * \frac{0,015 * P(S - X_{76-100} = 4000 - 3000)}{P(S = 4000)} \\
&= 3000 * \frac{0,015 * 0,07011226}{0,1334973} = 23,63383
\end{aligned}$$

$$\begin{aligned}
h_{76-100}^{cm}(5000) &= 3000 * \frac{0,015 * P(S - X_{76-100} = 5000 - 3000)}{P(S = 5000)} \\
&= 3000 * \frac{0,015 * 0,1460353}{0,1181168} = 55,63637
\end{aligned}$$

From the calculations above, it can be seen that there are no errors in the calculations. Some values of the CMRS decrease as the cost of the portfolio S increases. These same calculations are computed for a portfolio of 1000 and 10000 policyholders which lead to figure 7. Since the extreme values are removed, for the 1000 insured portfolio the total cost starts at 11000 and ends at 86000 and for the 10000 portfolio, the total cost starts at 313000 and ends at 576000.

It is apparent from figure 7 that the larger the number of insureds, the less volatile the CMRS is. However, there are still small oscillations at the edges, the calculations need to be checked in order to see if there are any errors.

For $n = 1000$:

$$\begin{aligned}
h_{1-250}^{cm}(11000) &= 1000 * \frac{0,03 * P(S - X_{1-250} = 11000 - 1000)}{P(S = 11000)} \\
&= 1000 * \frac{0,03 * 0}{0,0000113065} = 0
\end{aligned}$$

In reality, the probability $P(S - X_{1-250} = 11000 - 1000)$ should not be equal to 0, as it is possible that 5 bicycles at 2000€ could be stolen, for example. The probability might be too small that it's considered to be zero by the R software. Indeed, if you look at $P(S - X_{1-250} = 12000 - 1000)$, you can already see that it's very low. Even if I ask R to give the result with 22 decimal places, which is the maximum on R, R returns 0.

$$\begin{aligned}
h_{1-250}^{cm}(12000) &= 1000 * \frac{0,03 * P(S - X_{1-250} = 12000 - 1000)}{P(S = 12000)} \\
&= 1000 * \frac{0,03 * 0,00001147671}{0,00002126966} = 16,18743
\end{aligned}$$

$$\begin{aligned}
h_{1-250}^{cm}(13000) &= 1000 * \frac{0,03 * P(S - X_{1-250} = 13000 - 1000)}{P(S = 13000)} \\
&= 1000 * \frac{0,03 * 0,00002157253}{0,0000413042} = 15,66851
\end{aligned}$$

$$\begin{aligned}
h_{1-250}^{cm}(14000) &= 1000 * \frac{0,03 * P(S - X_{1-250} = 14000 - 1000)}{P(S = 14000)} \\
&= 1000 * \frac{0,03 * 0,00004187425}{0,00007266322} = 17,28835
\end{aligned}$$

The calculations for the 2000€ and 3000€ bikes are not verified since there are no oscillations for small total cost on the graph on figure 7. However, there are still oscillations for high total cost for the 1000€ and the 2000€ bikes, the calculation for the participation will therefore be checked:

$$\begin{aligned}
h_{1-250}^{cm}(82000) &= 1000 * \frac{0,03 * P(S - X_{1-250} = 82000 - 1000)}{P(S = 82000)} \\
&= 1000 * \frac{0,03 * 0,00006588635}{0,00004919672} = 40,17729
\end{aligned}$$

$$\begin{aligned}
h_{1-250}^{cm}(83000) &= 1000 * \frac{0,03 * P(S - X_{1-250} = 83000 - 1000)}{P(S = 83000)} \\
&= 1000 * \frac{0,03 * 0,00004868053}{0,00003503529} = 41,68413
\end{aligned}$$

$$\begin{aligned}
h_{1-250}^{cm}(84000) &= 1000 * \frac{0,03 * P(S - X_{1-250} = 84000 - 1000)}{P(S = 84000)} \\
&= 1000 * \frac{0,03 * 0,00003466494}{0,00002567948} = 40,49725
\end{aligned}$$

$$\begin{aligned}
h_{251-750}^{cm}(84000) &= 2000 * \frac{0,03 * P(S - X_{251-750} = 84000 - 2000)}{P(S = 84000)} \\
&= 2000 * \frac{0,03 * 0,00003432439}{0,00002567948} = 93,86158
\end{aligned}$$

$$\begin{aligned}
h_{251-750}^{cm}(85000) &= 2000 * \frac{0,03 * P(S - X_{251-750} = 85000 - 2000)}{P(S = 85000)} \\
&= 2000 * \frac{0,03 * 0,00002514399}{0,00001870635} = 91,74531
\end{aligned}$$

$$\begin{aligned}
h_{251-750}^{cm}(86000) &= 2000 * \frac{0,03 * P(S - X_{251-750} = 86000 - 2000)}{P(S = 86000)} \\
&= 2000 * \frac{0,03 * 0,00001830586}{0,00001285742} = 97,78007
\end{aligned}$$

For $n = 10000$:

Figure 7 shows that there are some oscillations for small and large total cost but only for the 1000€ and 2000€ bikes, the calculations will be verified once again to see if there are any errors :

$$\begin{aligned}
h_{1-2500}^{cm}(314000) &= 1000 * \frac{0,03 * P(S - X_{1-2500} = 314000 - 1000)}{P(S = 314000)} \\
&= 1000 * \frac{0,03 * 9,092279 * 10^{-7}}{1,062923 * 10^{-6}} = 25,66210
\end{aligned}$$

$$\begin{aligned}
h_{1-2500}^{cm}(315000) &= 1000 * \frac{0,03 * P(S - X_{1-2500} = 315000 - 1000)}{P(S = 315000)} \\
&= 1000 * \frac{0,03 * 1,067677 * 10^{-6}}{1,246013 * 10^{-6}} = 25,70623
\end{aligned}$$

$$\begin{aligned}
h_{1-2500}^{cm}(316000) &= 1000 * \frac{0,03 * P(S - X_{1-2500} = 316000 - 1000)}{P(S = 316000)} \\
&= 1000 * \frac{0,03 * 1,251528 * 10^{-6}}{1,492996 * 10^{-6}} = 25,14800
\end{aligned}$$

$$\begin{aligned}
h_{2501-7500}^{cm}(315000) &= 2000 * \frac{0,03 * P(S - X_{2501-7500} = 315000 - 2000)}{P(S = 315000)} \\
&= 2000 * \frac{0,03 * 9,114871 * 10^{-7}}{0,00002567948} = 36,57615
\end{aligned}$$

$$\begin{aligned}
h_{2501-7500}^{cm}(316000) &= 2000 * \frac{0,03 * P(S - X_{2501-7500} = 316000 - 2000)}{P(S = 316000)} \\
&= 2000 * \frac{0,03 * 1,070309 * 10^{-6}}{0,00001870635} = 35,84435
\end{aligned}$$

$$\begin{aligned}
h_{2501-7500}^{cm}(317000) &= 2000 * \frac{0,03 * P(S - X_{2501-7500} = 317000 - 2000)}{P(S = 317000)} \\
&= 2000 * \frac{0,03 * 1,25459 * 10^{-6}}{0,00001285742} = 35,92284
\end{aligned}$$

$$\begin{aligned}
h_{1-2500}^{cm}(574000) &= 1000 * \frac{0,03 * P(S - X_{1-2500} = 574000 - 1000)}{P(S = 574000)} \\
&= 1000 * \frac{0,03 * 1,008647 * 10^{-6}}{8,892337 * 10^{-7}} = 34,02863
\end{aligned}$$

$$\begin{aligned}
h_{1-2500}^{cm}(575000) &= 1000 * \frac{0,03 * P(S - X_{1-2500} = 575000 - 1000)}{P(S = 575000)} \\
&= 1000 * \frac{0,03 * 8,855405 * 10^{-7}}{7,653734 * 10^{-7}} = 34,71014
\end{aligned}$$

$$\begin{aligned}
h_{1-2500}^{cm}(576000) &= 1000 * \frac{0,03 * P(S - X_{1-2500} = 576000 - 1000)}{P(S = 576000)} \\
&= 1000 * \frac{0,03 * 7,621911 * 10^{-7}}{6,712545 * 10^{-7}} = 34,06418
\end{aligned}$$

$$\begin{aligned}
h_{2501-7500}^{cm}(573000) &= 2000 * \frac{0,03 * P(S - X_{2501-7500} = 573000 - 2000)}{P(S = 573000)} \\
&= 2000 * \frac{0,03 * 1,322005 * 10^{-6}}{1,012821 * 10^{-6}} = 65,26347
\end{aligned}$$

$$\begin{aligned}
h_{2501-7500}^{cm}(574000) &= 2000 * \frac{0,03 * P(S - X_{2501-7500} = 574000 - 2000)}{P(S = 574000)} \\
&= 2000 * \frac{0,03 * 1,144177 * 10^{-6}}{8,892337 * 10^{-7}} = 64,33501
\end{aligned}$$

$$\begin{aligned}
h_{2501-7500}^{cm}(575000) &= 2000 * \frac{0,03 * P(S - X_{2501-7500} = 575000 - 2000)}{P(S = 575000)} \\
&= 2000 * \frac{0,03 * 1,005436 * 10^{-6}}{7,653734 * 10^{-7}} = 65,68272
\end{aligned}$$

The calculations above proves that there are no errors in the calculations. For extreme values, the function is therefore not monotonic, but this poses no problem when applying CMRS, since the values are located in areas of very low probability (as can be seen from the $P[S = s]$ probabilities shown in grey in figures 7) and the extremes quantiles have been removed. However, as the number of policyholders increases, the oscillations tend to be less significant.

2.6 Pricing

To price the collaborative insurance, two methods will be applied. In both methods, the high layer of the risk will be insured, as was previously mentioned. This means that a traditional insurer will assume responsibility for the costs in excess of this limit after a certain percentage in the cumulative distribution function. In the first method, a retention level will be set and then the premiums will be computed, and in the second method, the premiums will be set and the retention level will be computed. For the first method, the probability corresponding to the retention level will be set to 80% as an example, allowing the insurer to pay out only an average of one year out of every five years.

The results obtained in the previous section demonstrated that $h_i^{cm}(s)$ is not continuously increasing for the portfolio of 100 insureds. This portfolio therefore does not meet the conditions for applying the CMRS rule. As a result, regardless of the method used, the premiums cannot be calculated.

Therefore, for each of the datasets, the retention levels that correspond to these 80% ($\beta = 80\%$) will be determined :

$$\beta = F_{S_n}(w_n) \Leftrightarrow w_n = F_{S_n}^{-1}(\beta)$$

$$F_{S_{1000}}^{-1}(0, 8) = w_{1000} = 52000$$

$$F_{S_{10000}}^{-1}(0, 8) = w_{10000} = 463000$$

All the following calculations exploit the fact that :

$$S_n = \sum_{i=1}^n X_i = \sum_{i=1}^n \mathbb{E}[X_i|S_n] \quad (1)$$

The specific limit for each insured ($w_{i,n}$) is then determined using the formula below :

$$w_{i,n} = F_{h_i^{cm}(S_n)}^{-1}(F_{S_n}(w_n)) = \mathbb{E}[X_i|S_n = w_n]$$

$$S_n \geq w_n \Leftrightarrow h_i^{cm}(S_n) \geq w_{i,n} = F_{h_i^{cm}(S_n)}^{-1}(F_{S_n}(w_n))$$

Theorem 6 in Dhaene J, Denuit M, Goovaerts MJ, Kaas R, Vyncke D (2002) [6] demonstrates that the stop-loss premiums for the variable S can be expressed as the sum of stop-loss premiums associated with each individual comonotonic term in the sum S . Furthermore, upon examining the proof of Theorem 6 in Dhaene et al. [7], it can be observed that this decomposition applies not only to the stop-loss premiums but also to the random variables themselves. From equation 1, for any w_n such that $0 \leq F_{S_n}(w_n) \leq 1$, there exist $w_{1,n}, \dots, w_{n,n}$ such that $\sum_{i=1}^n w_{i,n} = w_n$ and the identities

$$(S_n - w_n)_+ = \sum_{i=1}^n (h_i^{cm}(S_n) - w_{i,n})_+$$

$$(w_n - S_n)_+ = \sum_{i=1}^n (w_{i,n} - h_i^{cm}(S_n))_+$$

both hold true. Formal proof for previous identities can be found in Dhaene and Linders [7, Chapter 7].

The respective contributions for each participant to the retained loss $\min(S_n, w_n)$ can then be computed :

$$\begin{aligned} \min(S_n, w_n) &= w_n - (w_n - S_n)_+ \\ &= \sum_{i=1}^n \min(h_i^{cm}, w_{i,n}) \end{aligned}$$

Thus, participant i contributes to the lower layer $(0, w_n)$ an amount $\min(h_i^{cm}(s), w_{i,n})$ where the individual retention levels $w_{i,n}$ is equal to $\mathbb{E}[X_i|S_n = w_n]$.

And the results for each dataset are :

$$w_{1-250,1000} = 32,73452 ; w_{251-750,1000} = 59,0674 ; w_{751-1000,1000} = 57,13051$$

$$w_{1-2500,10000} = 30,83667 ; w_{2501-7500,10000} = 52,7901 ; w_{7501-10000,10000} = 48,78312$$

The P2P insurance community, as previously mentioned, merely keeps the lower layer and purchases a stop-loss cover from a traditional insurer for the upper layer.

So, the insured must divide up the upper layer stop loss premium.

The stop loss premium will be distributed as follows, taking into account each party's individual contribution to the losses' top layer:

$$\pi_{i,n}^{SL} = (1 + \theta_{SL})\mathbb{E}[(h_i^{cm}(S_n) - w_{i,n})_+]$$

θ_{SL} represents the corresponding loading to obtain the premium charged by the insurer covering the upper layer. It can be assumed that $\theta_{SL} = 0,5$ for the numerical illustration.

And the results for each dataset are :

$$\pi_{1-250,1000}^{SL} = 0,4459995 ; \pi_{251-750,1000}^{SL} = 1,696142 ; \pi_{751-1000,1000}^{SL} = 2,625639$$

$$\pi_{1-2500,10000}^{SL} = 0,1550996 ; \pi_{2501-7500,10000}^{SL} = 0,541737 ; \pi_{7501-10000,10000}^{SL} = 0,7712756$$

To have the final premiums, the two premiums found are added: the P2P premium and the StopLoss premium. The P2P premium is simply equal to the specific limit for each insured :

$$\pi_{i,n}^{P2P} = w_{i,n} = h_i^{cm}(w_n) = E[X_i|S_n = w_n]$$

The following pure premium for each class of insured are obtained:

$$\pi_{i,n} = \pi_{i,n}^{P2P} + \pi_{i,n}^{SL}$$

$$\pi_{1-250,1000} = \pi_{1-250,1000}^{P2P} + \pi_{1-250,1000}^{SL} = 33,18052\text{€}$$

$$\pi_{251-750,1000} = \pi_{251-750,1000}^{P2P} + \pi_{251-750,1000}^{SL} = 60,76354\text{€}$$

$$\pi_{751-1000,1000} = \pi_{751-1000,1000}^{P2P} + \pi_{751-1000,1000}^{SL} = 59,75615\text{€}$$

$$\pi_{1-2500,10000} = \pi_{1-2500,10000}^{P2P} + \pi_{1-2500,10000}^{SL} = 30,99177\text{€}$$

$$\pi_{2501-7500,10000} = \pi_{2501-7500,10000}^{P2P} + \pi_{2501-7500,10000}^{SL} = 53,33184\text{€}$$

$$\pi_{7501-10000,10000} = \pi_{7501-10000,10000}^{P2P} + \pi_{7501-10000,10000}^{SL} = 49,55440\text{€}$$

It can be noted from the calculations above that premiums increase when the number of insured decreases because each member of a P2P insurance pool is exposed to a bigger proportion of the risk. As a result, each member's premium may increase if one of the members experiences a covered loss because the expense of that loss will be shared among fewer members.

To ensure that there is enough money in the pool to cover any losses and that each member is bearing their fair part of the risk, premiums are sometimes higher in smaller P2P insurance pools.

For the second method, Premiums will be arbitrary set as following :

$$\pi_{i,n} = 40\text{€} \quad i = 1, \dots, 0, 25n$$

$$\pi_{i,n} = 50\text{€} \quad i = 0, 25(n+1), \dots, 0, 75n$$

$$\pi_{i,n} = 70\text{€} \quad i = 0, 75(n+1), \dots, n$$

The retention level can be found w_n by solving the following equation :

$$\begin{aligned} \sum_{i=1}^n \pi_{i,n} &= w_n + \pi_{SL}(w_n) \\ &= w_n + (1 + \theta_{SL})\mathbb{E}[(S_n - w_n)_+] \end{aligned}$$

Thanks to this function, the retention level linked to the total premium is identified. figure 8 provides the plot of the following function:

$$f(w_n) = w_n + (1 + \theta_{SL})\mathbb{E}[(S_n - w_n)_+]$$

The total premiums for the portfolios of policyholders are as follow :

$$\begin{aligned} \sum_{i=1}^{1000} \pi_{i,1000} &= 52500\text{€} \\ \sum_{i=1}^{10000} \pi_{i,10000} &= 525000\text{€} \end{aligned}$$

A w_n that satisfy the following equation needs to be found :

$$52500 = w_{1000} + (1 + \theta_{SL})\mathbb{E}[(S_{1000} - w_{1000})_+]$$

$$525000 = w_{10000} + (1 + \theta_{SL})\mathbb{E}[(S_{10000} - w_{10000})_+]$$

In order to find w_n , all the premium for every w_n 10 by 10 are calculated. For 1000 insured, two solutions are found (because, as can be seen from figure 8, there are two possible w_n for a total premium of 52500€), the first is $w_{1000} = 26470$ and the second is $w_{1000} = 50360$. Therefore the probabilities of the retention levels can be computed:

$$P[S \leq 26470] = 2,630047\%$$

$$P[S \leq 50360] = 76,81598\%$$

As can be seen, the probability corresponding to the retention level for $w_{1000} = 26470$ is much too low and no insurer would accept to transfer such a high part of the risk, so only the second retention level $w_{1000} = 50360$ will be kept. In order to find a solution, w_{1000} needs to be round to 50000 because if a retention level is not rounded to the nearest thousand, the probability for the retention level will be equal to 0 since there are only 1000€, 2000€ and 3000€ bikes.

For 10000 insured, the result is $w_{10000} = 524970$. Then, the probability linked to the retention level can be computed :

$$P[S \leq 524970] = 99,76624\%$$

It can be seen from the previous calculation that the probability corresponding to the retention level for $w_{10000} = 524970$ is really high, so maybe the premiums could be lowered even more to find a lower probability. As done previously, w_{10000} is rounded to the nearest thousand so $w_{10000} = 525000$

For the premiums mentioned above, the probability corresponding to the retention level is 76,8% approximately for a portfolio of 1000 insured and approximately 99,8% for a portfolio of 10000 insured. Since the premiums are already known, it is possible to calculate the stop-loss premiums from the retention levels. First the specific retention levels for each policyholder is calculated :

$$w_{i,1000} = \mathbb{E}[X_i | S_{1000} = w_{1000}] = 32,14549 \quad i = 1, \dots, 250$$

$$w_{i,1000} = \mathbb{E}[X_i | S_{1000} = w_{1000}] = 56,9309 \quad i = 251, \dots, 750$$

$$w_{i,1000} = \mathbb{E}[X_i | S_{1000} = w_{1000}] = 53,99256 \quad i = 751, \dots, 1000$$

$$w_{i,10000} = \mathbb{E}[X_i | S_{10000} = w_{10000}] = 32,71152 \quad i = 1, \dots, 2500$$

$$w_{i,10000} = \mathbb{E}[X_i | S_{10000} = w_{10000}] = 59,44358 \quad i = 2501, \dots, 7500$$

$$w_{i,10000} = \mathbb{E}[X_i | S_{10000} = w_{10000}] = 58,40978 \quad i = 7501, \dots, 10000$$

Now that the retention levels for each insured are known, the calculation for their specific stop loss premiums can be realised :

$$\pi_{i,1000}^{SL} = (1 + \theta_{SL}) \mathbb{E}[(h_i^{cm}(S_{1000}) - w_{i,1000})_+] = 0,6379628 \quad i = 1, \dots, 250$$

$$\pi_{i,1000}^{SL} = (1 + \theta_{SL}) \mathbb{E}[(h_i^{cm}(S_{1000}) - w_{i,1000})_+] = 2,392203 \quad i = 251, \dots, 750$$

$$\pi_{i,1000}^{SL} = (1 + \theta_{SL}) \mathbb{E}[(h_i^{cm}(S_{1000}) - w_{i,1000})_+] = 3,647634 \quad i = 751, \dots, 1000$$

$$\pi_{i,10000}^{SL} = (1 + \theta_{SL}) \mathbb{E}[(h_i^{cm}(S_{10000}) - w_{i,10000})_+] = 0,0009157543 \quad i = 1, \dots, 2500$$

$$\pi_{i,10000}^{SL} = (1 + \theta_{SL}) \mathbb{E}[(h_i^{cm}(S_{10000}) - w_{i,10000})_+] = 0,0033501555 \quad i = 2501, \dots, 7500$$

$$\pi_{i,10000}^{SL} = (1 + \theta_{SL}) \mathbb{E}[(h_i^{cm}(S_{10000}) - w_{i,10000})_+] = 0,0050337999 \quad i = 7501, \dots, 10000$$

Finally, the P2P premiums can be computed:

$$\pi_{i,1000}^{P2P} = \pi_{i,1000} - \pi_{i,1000}^{SL} = 39,36204 \quad i = 1, \dots, 250$$

$$\pi_{i,1000}^{SL} = \pi_{i,1000} - \pi_{i,1000}^{SL} = 47,60780 \quad i = 251, \dots, 750$$

$$\pi_{i,1000}^{SL} = \pi_{i,1000} - \pi_{i,1000}^{SL} = 66,35237 \quad i = 751, \dots, 1000$$

$$\pi_{i,10000}^{SL} = \pi_{i,10000} - \pi_{i,10000}^{SL} = 39,99908 \quad i = 1, \dots, 2500$$

$$\pi_{i,10000}^{SL} = \pi_{i,10000} - \pi_{i,10000}^{SL} = 49,99665 \quad i = 2501, \dots, 7500$$

$$\pi_{i,10000}^{SL} = \pi_{i,10000} - \pi_{i,10000}^{SL} = 69,99497 \quad i = 7501, \dots, 10000$$

To summarize, the results are displayed on tables 21 and 22. As it can be observed, w_i is not equal to π_{P2P} , this means that there are transfers between the groups.

2.7 Surplus

Now that the premiums have been calculated, each policyholder could receive a surplus in the case :

$$S_n \leq w_n$$

Indeed when claims are above the retention rate, policyholders receive a surplus, and therefore a reduction in their premium at the end of the year. Surplus is a very useful method because it reduces moral hazard for the following reasons:

- Sense of mutual responsibility : In collaborative insurance, policyholders see themselves as part of a community where they collectively share the risks. They know that their actions can have an impact on the group as a whole. As a result, they tend to act more responsibly to avoid increasing overall risks and harming other members.

- Incentive to minimise costs: Since the surplus accrues directly to members, policyholders are more motivated to minimise costs and claims. They are more likely to adopt preventive measures to reduce the risk of claims, as this will help them to maintain the higher surplus and therefore reduce future premiums.

The surplus in collaborative insurance thus creates a bond of responsibility and mutual commitment between policyholders, encouraging more responsible behaviour and reducing moral hazard.

However, as mentioned in the introduction, having a variable premium scares risk-averse policyholders, who are more likely to turn to traditional fixed-premium insurance.

The expected surplus for each category of insured is computed as following :

$$\mathbb{E}[(w_{i,n} - h_{i,n}^{cm}(S_n))_+]$$

Per example, for the insured 1, who has a 1000€ bike and $n = 1000$:

$$\mathbb{E}[(w_{1,1000} - h_{1,1000}^{cm}(S_{1000}))_+] = \sum_{k=1}^{2000} (w_{1,1000} - h_{1,1000}^{cm}(k * 1000))_+ P(S_{1000} = k * 1000)$$

Tables 1, 2, 3 and 4 displays the surplus expectation for each category and pricing method.

Group	1000€	2000€	3000€
π_i	33,18€	60,76€	59,76€
$P[\text{cash-back}_i \geq 0]$	80%	80%	80%
$\mathbb{E}[\text{cash-back}_i]$	3,032239	10,199707	13,883452

Table 1: table
Surplus for $n = 1000$

Group	1000€	2000€	3000€
π_i	40€	50€	70€
$P[\text{cash-back}_i \geq 0]$	76,8%	76,8%	76,8%
$\mathbb{E}[\text{cash-back}_i]$	2,571176	8,527245	11,426831

Table 2: table
Surplus for $n = 1000$

Group	1000€	2000€	3000€
π_i	30,99€	53,33€	49,55€
$P[\text{cash-back}_i \geq 0]$	80%	80%	80%
$\mathbb{E}[\text{cash-back}_i]$	0,9400878	3,1513332	4,2974117

Table 3: table
Surplus for $n = 10000$

Group	1000€	2000€	3000€
π_i	40€	50€	70€
$P[\text{cash-back}_i \geq 0]$	99,8%	99,8%	99,8%
$\mathbb{E}[\text{cash-back}_i]$	2,712150	9,445888	13,413241

Table 4: table
Surplus for $n = 10000$

As can be seen from tables 1 to 4, as the number of policyholders increases, the expected cash-back decreases. This can be explained by changes in risk distribution. Indeed, with more policyholders, the overall risk is distributed among a larger group of people. This means that the potential losses are spread over a greater number of individuals, reducing the individual contribution to the common pool. As a result, the excess funds available for cash-back are smaller for each policyholder.

Secondly, it can be noted that increasing the retention level leads to an increase in the expected cash-back. This phenomenon turns out to be normal, as increasing the level of retention also means increasing the likelihood of cashback. Furthermore, as the policyholder pays more when the retention level increases, this means that the policyholder will benefit from a higher cashback.

2.8 Variance of the CMRS

The variance of the conditional mean risk sharing is then calculated as a function of the insured population as follows :

$$Var[\mathbb{E}[X_i|S_n]] = \mathbb{E}[(\mathbb{E}[X_i|S_n])^2] - (\mathbb{E}[\mathbb{E}[X_i|S_n]])^2$$

n = number of insured

To calculate the variance, it is necessary to separate each group of policyholders according to the price of their bike. First, n is fixed which is the number of insureds in

the portfolio and for each group of insureds, the conditional expectations $\mathbb{E}[X_i|S_n]$ are calculated by varying S_n , the cost of the portfolio of n insureds. Secondly, the variance of the conditional expectations $Var[\mathbb{E}[X_i|S_n]]$ is calculated. Finally, the process is repeated by varying n by steps of 1000, from 1000 to 30000 which is illustrated in figure 9.

Figure 9 shows that the variance of the conditional mean risk sharing for all policyholder bike prices starts to settle at about 15000 policyholders and around 10000 policyholders for the standard deviation.

The next step is to calculate the evolution of the quantiles as a function of the number of policyholders which is depicted in figure 2. In the figure, the blue line represents the median (50% quantile) and the two ends of the curve correspond to the 25% quantile (bottom line) and the 75% quantile (top line). As can be seen from figure 2, the interquartile range (the gap between the 25% and 75% quantiles) narrows as the number of policyholders increases. The interquartile range stabilizes at around 10,000 policyholders. As the standard deviation and interquartile range stabilize at around 10000 policyholders, the minimum number of policyholders is set at 10000.

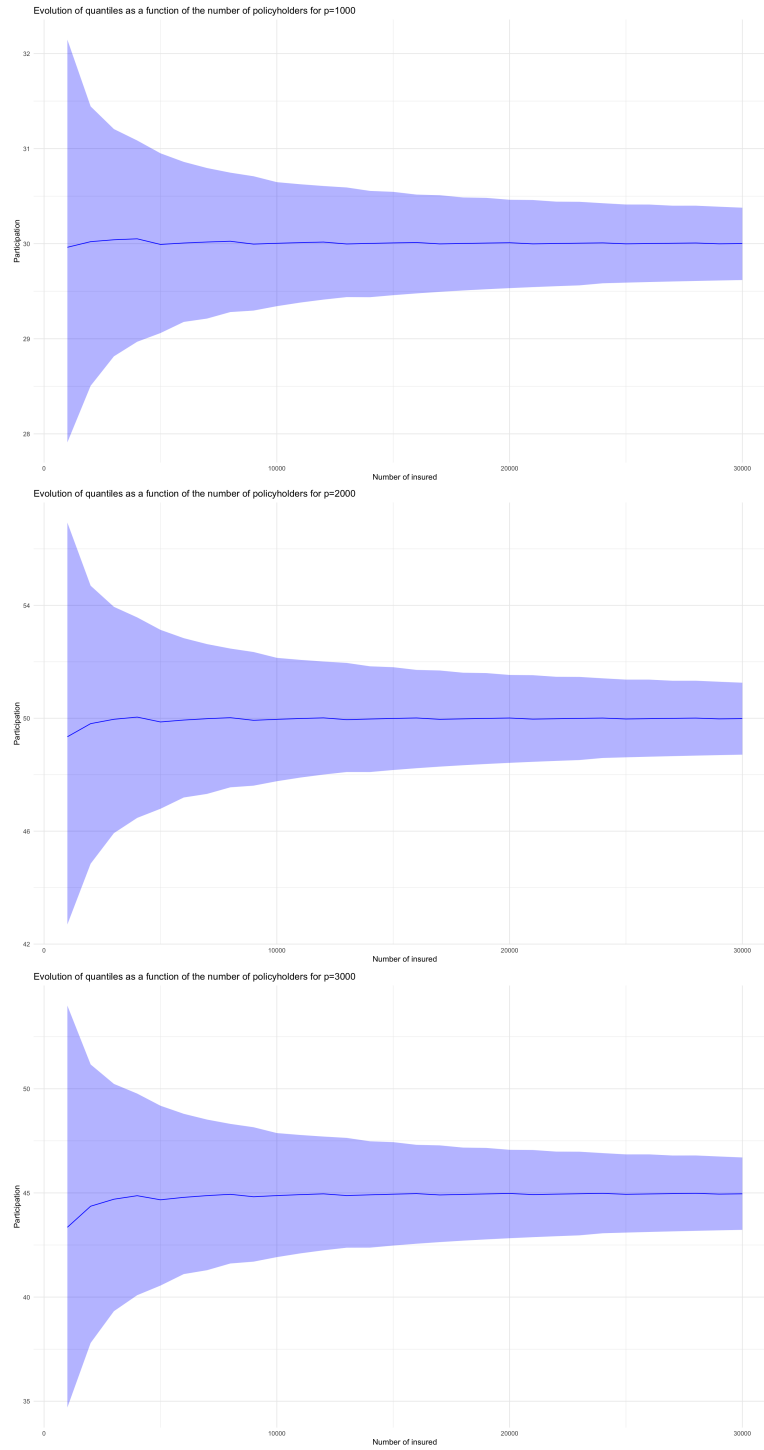


Figure 2: Evolution of quantiles as a function of the number of policyholders for $p=1000$ (upper panel), for $p=2000$ (middle panel) and for $p=3000$ (lower panel)

Chapter 3

Preparation

3.1 Data Set

To be able to do my master thesis, I tried to contact the Laka platform to have access to their database but unfortunately they refused. I then tried to collect public statistical data on bicycles in Belgium but unfortunately this data was not accurate enough to determine the frequency of theft according to the price of the bicycle.

I therefore had to find another solution. For this reason, as an alternative, I contacted classic insurers to ask for their rates to be able to make a segmentation of the market and thus obtain thefts and accidents frequencies. However, only Allianz and Axa replied. If I wanted to know the prices of other insurers, I could only make online quotes, but that would have taken too much time. In fact, every time I changed the slightest characteristic of the contract the price altered. Nevertheless, it turns out that Allianz offers rates for theft cover only and grouped rates for the accident and breakdown covers. Axa, on the other hand, has grouped rates for theft, accident and breakdown covers. This made it difficult for me to determine the frequency of claims on grouped rates. For this reason, I decided, with the agreement of my master thesis sponsor Mr Michel Denuit, to concentrate my master thesis on theft cover alone, as this is the only cover that has its own rate.

The rates for the theft coverage offered by Allianz vary according to the price of the bike, so I have to find theft frequencies according to the price of the bike. I therefore need to find the theft frequencies based on Allianz premiums.

First, I calculate the proportion of the commercial premium in relation to the price of the bike. For example, for a bike costing 450€, the commercial premium is 28,80€, which represents approximately 6,4% of the price of the bike. Afterwards,

I divided the proportions of the commercial premiums by 2 to obtain the theft frequencies on table 5. Indeed, according to my master thesis sponsor and myself, we thought that these theft frequencies were rather realistic.

Price	Commercial premium	Proportion of the commercial premium	Theft frequency
450€	28,80€	6,4%	3,2%
500€	30,60€	6,12%	3,06%
600€	35,40€	5,9%	2,95%
700€	38,40€	5,49%	2,74%
800€	44,40€	5,55%	2,78%
900€	53,40€	5,93%	2,97%
1000€	63€	6,3%	3,15%
1200€	66€	5,5%	2,75%
1400€	69€	4,93%	2,46%
1600€	75€	4,69%	2,34%
1800€	81€	4,5%	2,25%
2000€	87€	4,35%	2,18%
2500€	102€	4,08%	2,04%
3000€	117€	3,9%	1,95%
3500€	132€	3,77%	1,89%
4000€	144€	3,6%	1,8%
4500€	162€	3,6%	1,8%
5000€	180€	3,6%	1,8%

Table 5: Allianz premiums and theft frequencies

All commercial premiums are annual and the prices represent the upper bracket. For example, a bike worth 550€, will have an annual commercial premium of 35,40€.

To find the thefts frequencies from the Allianz commercial rates, the following assumptions has been applied:

- Tax on the insurance contract: 9,25%
- Commissions to intermediaries: 15%
- General expenses: 25%
- Claims handling costs: 20%
- Profit margin/remuneration of solvency capital: 25%

As it can be noted, if all these percentages are added up, the result is 94,25%. This justifies why dividing the theft frequencies by 2, since the assumptions mentioned above add almost 100% to the pure premium. The commercial premium is then divided by 2 to obtain the pure premiums (see table 23).

3.2 Portfolio of insureds

The next question is how to allocate the policyholders to each price range and how many policyholders to put in the portfolio ? As explained in the methodology section, the minimum number of insured in the portfolio is approximately 10000. So for the case studies, the first step consists of reducing the number of classes to 10, i.e classes ranging from 500€ to 5000€ in increments of 500€. Then 2 portfolios of 10000 insured are created:

- a portfolio of 10000 policyholders evenly distributed
- a portfolio of 10000 policyholders distributed using a binomial distribution (0.5) (figure 10)

Chapter 4

Case Studies

For the case studies, a dataset with 10000 policyholders is used, whose theft frequencies are presented in table 6 (for more details on the choice of dataset, please refer to the preparation section).

Dataset	
Price	Theft frequency
500€	3,06%
1000€	3,15%
1500€	2,4%
2000€	2,18%
2500€	2,04%
3000€	1,95%
3500€	1,89%
4000€	1,8%
4500€	1,8%
5000€	1,8%

Table 6: Theft frequencies

For the 1500€ price range, Allianz did not offer a rate. Thus the theft frequency has been calculated by using the theft frequencies of a 1400€ and 1600€ bike:

$$\frac{2,46\% + 2,34\%}{2} = 2,4\%$$

In terms of methodology, the exact same method as in the theory section has been applied. As explained in the preparation section, 2 different distributions for the policyholders will be used. The methodology will be applied to two different dataset:

a uniformly distributed portfolio of policyholders and a portfolio of policyholders distributed according to a 0,5 binomial distribution.

4.1 Uniform distribution

For the uniform distribution, the dataset is distributed as in table 7.

Dataset distribution	
Price	Number of insured
500€	1000
1000€	1000
1500€	1000
2000€	1000
2500€	1000
3000€	1000
3500€	1000
4000€	1000
4500€	1000
5000€	1000

Table 7: Dataset uniformly distributed

As in the theory section, each category is assigned a binomial distribution and then all the distribution are convolved together to obtain an overall distribution:

$$Bin(1000, 3.06\%) + Bin(1000, 3.15\%) + Bin(1000, 2.4\%) + Bin(1000, 2.18\%) + Bin(1000, 2.04\%) + Bin(1000, 1.95\%) + Bin(1000, 1.89\%) + Bin(1000, 1.8\%) + Bin(1000, 1.8\%) + Bin(1000, 1.8\%)$$

The result is the following distribution illustrated in figure 11.

To determine which specific retention level for each policyholder to implement for this dataset, the following function is plotted (for more information on this function, please refer to the theory section):

$$f(w_n) = w_n + (1 + \theta_{SL})\mathbb{E}[(S_n - w_n)_+]$$

It is again assumed that $\theta_{SL} = 0, 5$.

This leads to figure 3 which shows which retention level gives the lowest overall premium since it appears to be a minimum.

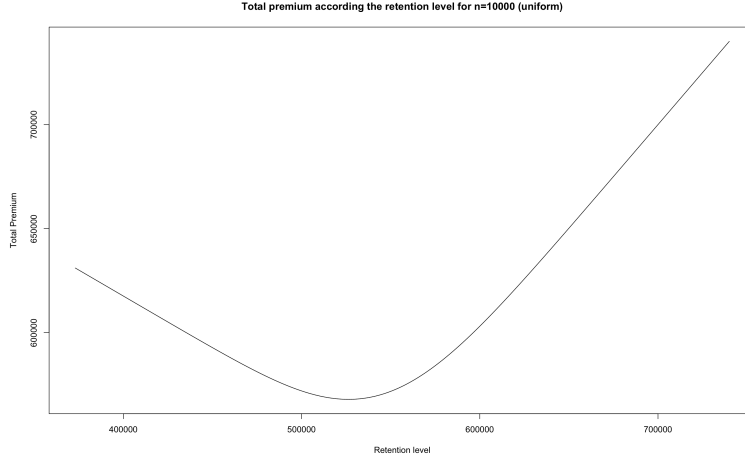


Figure 3: Total premium according the retention level for $n=10000$ (uniform)

The retention level that minimize the overall premium is $w_{10000} = 526500$. It is now possible to calculate the probability of this retention level:

$$P(S_{10000} \leq 526500) = 33,67332\%$$

As can be seen, the probability corresponding to the retention level is very low. This would mean that the insurer would pay out on average 2 years out of 3, which that it is highly unlikely that an insurer would accept to insure a risk with such a high probability of loss. However, the calculations for the retention level $w_{10000} = 526500$ will still be done, but also the calculations for retention level $w_{10000} = 566500$ and $w_{10000} = 615500$ will be done in parallel, because:

$$P(S_{10000} \leq 566500) = 70\%$$

$$P(S_{10000} \leq 615500) = 95\%$$

This allows to have a probability of 70% and 95%, which is a much more acceptable level for an insurer.

With the retention levels now available, the premiums can be easily calculated (in exactly the same way as in the theory section).

The tables 24, 25 and 26 therefore show the P2P and Stop Loss premiums for each retention level. The P2P premium and Stop Loss premium are then added together to obtain the pure premium (table 27). As can be seen, increasing the retention level from 526500 to 566500 has very little impact on the pure premium.

Please note that when talking about pure premium, the corresponding loading (θ_{SL}) is included to obtain the premium charged by the insurer covering the top layer.

First, it is necessary to determine the percentage that the platform intends to deduct. As explained in the introduction, Laka takes 25% of the premium to insure the upper part of the risk (Nationale-Nederlanden) and for its fees. It is therefore difficult to estimate Laka's costs on the basis of this 25%. However, on Laka's UK website, the fee is set at 20% of the premium. Therefore, $\theta_{P2P} = 20\%$. The platform revenue is calculated as follows, since the total premium includes the platform fee:

$$\frac{\sum_{i=1}^n \pi_i}{1.2} \times 0.2$$

Thus, the platform will take 20% of the pure premium and the following revenue for the platform are obtained:

$$\begin{aligned} w_{10000} &= 526500 : \frac{567929,7}{1,2} \times 0,2 = 94654,94\text{€} \\ w_{10000} &= 566500 : \frac{579034,6}{1,2} \times 0,2 = 96505,77\text{€} \\ w_{10000} &= 615500 : \frac{616910,5}{1,2} \times 0,2 = 102818,4\text{€} \end{aligned}$$

As mentioned above, here it is the revenue that is mentioned, not the profit. There are lots of costs involved in calculating profit, such as the cost of hosting the website, paying employees, bank charges, etc. . . All these costs are difficult to estimate, as they are highly variable and take many parameters into account. For this reason, only the revenue is calculated here. The platform have to see if, taking all these costs into account, it is still viable and the revenue sufficient. If the revenue is not sufficient to cover all costs, the percentage corresponding to the retention level and the percentage retained by the platform will have to be readjusted in order to become viable in the long term.

Taxes on insurance contacts of 9.25% must now be added, to give the commercial premiums shown in table 8.

Bike price	$w_{10000} = 526500$	$w_{10000} = 566500$	$w_{10000} = 615500$	Allianz premiums
500€	16,82€	16,87€	17,03€	30,60€
1000€	34,85€	35,05€	35,71€	63€
1500€	40,08€	40,43€	41,59€	
2000€	48,84€	49,42€	51,31€	87€
2500€	57,48€	58,33€	61,13€	102€
3000€	66,33€	67,50€	71,41€	117€
3500€	75,46€	77,00€	82,23€	132€
4000€	82,62€	84,55€	91,14€	144€
4500€	93,49€	95,94€	104,38€	162€
5000€	104,48€	107,51€	118,05€	180€

Table 8: Commercial premiums for uniformly distributed dataset

The premiums for the uniform distribution are much lower than those for Allianz. Therefore, in the event of a problem, the platform can increase its premiums and still remain cheaper than conventional insurance.

In order to find a retention level w_{10000} , the following amount is imposed on individual premiums:

$$\pi_i = (1 + \theta_{P2P})\mu_i$$

μ_i represents the pure premium, which means the price of the bike is multiplied by the theft rate to calculate μ_i and θ_{P2P} represents the compensation for the platform, which is 20% in this case. The premiums shown in Table 9 are then calculated for each bike price, giving a total premium of 654060.

Group	500€	1000€	1500€	2000€	2500€	3000€	3500€	4000€	4500€	5000€
π_i	18,36€	37,80€	43,20€	52,32€	61,20€	70,20€	79,38€	86,40€	97,20€	108€

Table 9: $\pi_i = (1 + \theta_{P2P})\mu_i$, uniformly distributed dataset

Therefore, the platform revenue can be calculated:

$$\frac{654060}{1,2} \times 0,2 = 109010€$$

Now that the premiums are available, the retention level can be calculated using the function plot in figure 3 and it can be seen that $w_{10000} = 653930$. However, as in the methodology section, if $w_{10000} = 653930$, then $P[S_{10000} = 653930] = 0$ since the total cost of claims is 500. Therefore it is necessary to round up w_{10000} so

$w_{10000} = 654000$. Thus, the stop loss premium can be calculated with the retention level and the P2P premium can be determined as follows:

$$\pi_{P2P} = \pi_i - \pi_{SL}$$

The P2P and stop-loss premiums shown in table 28 are obtained and the commercial premiums are calculated once the tax insurance of 9,25% has been added:

Bike price	Commercial premiums	Allianz commercial premiums
500€	20,06€	30,60€
1000€	41,30€	63€
1500€	47,20€	41,59€
2000€	57,16€	87€
2500€	66,86€	102€
3000€	76,69€	117€
3500€	86,72€	132€
4000€	94,39€	144€
4500€	106,19€	162€
5000€	117,99€	180€

Table 10: Commercial premiums, uniformly distributed dataset

Once again, the premiums are well below Allianz premiums, despite a very high retention probability ($P(S_{10000} \leq 654000) = 99,4\%$).

Table 11, 12, 13 and 14 show the surpluses (the expected cash-backs have been calculated in exactly the same way as in the methodology).

Bike price	$P[\text{cash-back}_i \geq 0]$	$\mathbb{E}[\text{cash-back}_i]$
500€	33,67%	0,040675
1000€	33,67%	0,165537
1500€	33,67%	0,282881
2000€	33,67%	0,452761
2500€	33,67%	0,655498
3000€	33,67%	0,892847
3500€	33,67%	1,165084
4000€	33,67%	1,433697
4500€	33,67%	1,793433
5000€	33,67%	2,188132

Table 11: Surplus, uniformly distributed dataset for $w_{10000} = 526500$

Bike price	$P[\text{cash-back}_i \geq 0]$	$\mathbb{E}[\text{cash-back}_i]$
500€	70%	0,126143
1000€	70%	0,517678
1500€	70%	0,892346
2000€	70%	1,440680
2500€	70%	2,104097
3000€	70%	2,891218
3500€	70%	3,806109
4000€	70%	4,725439
4500€	70%	5,963445
5000€	70%	7,340523

Table 12: Surplus, uniformly distributed dataset for $w_{10000} = 566500$

Bike price	$P[\text{cash-back}_i \geq 0]$	$\mathbb{E}[\text{cash-back}_i]$
500€	95%	0,285236
1000€	95%	1,180199
1500€	95%	2,051886
2000€	95%	3,341242
2500€	95%	4,922116
3000€	95%	6,822347
3500€	95%	9,059613
4000€	95%	11,347315
4500€	95%	14,445664
5000€	95%	17,937990

Table 13: Surplus, uniformly distributed dataset for $w_{10000} = 615500$

Bike price	$P[\text{cash-back}_i \geq 0]$	$\mathbb{E}[\text{cash-back}_i]$
500€	99,4%	0,420023
1000€	99,4%	1,746835
1500€	99,4%	3,053802
2000€	99,4%	5,000107
2500€	99,4%	7,406671
3000€	99,4%	10,323704
3500€	99,4%	13,785958
4000€	99,4%	17,365142
4500€	99,4%	22,231937
5000€	99,4%	27,763762

Table 14: Surplus, uniformly distributed dataset for $\pi_i = (1 + \theta_{P2P})\mu_i$

As can be seen from the tables above, the higher the retention level, the higher the expected cashback. Furthermore, a higher retention level implies a higher probability of obtaining a cashback will be. It is then up to the platform to see what it prefers, because as explained in the methodology section, cashback is a good way of avoiding moral hazard, with a low expected cashback, the effect on moral hazard may not be as good as expected. In addition, having a high expected cashback may play a role in the psychological aspect of the policyholder and may make it easier to build loyalty. The platform has to find what suits it best between higher premiums but higher expected cashback, or lower premiums but lower expected cashback.

4.2 Binomial (0,5) distribution

For the binomial distribution, the dataset is distributed as in table 15.

Dataset distribution	
Price	Number of insured
500€	98
1000€	439
1500€	1172
2000€	2051
2500€	2461
3000€	2051
3500€	1172
4000€	439
4500€	98
5000€	10

Table 15: Dataset with binomial distribution

If the total number of policyholders is added together, the result is 9991, not 10000, because to obtain integers approximations have to be made for each category of policyholder. The policyholders are distributed exactly as shown in figure 10.

As in the theory, each category is assigned a binomial distribution and then all the distribution are convolved together to obtain an overall distribution :

$$Bin(98, 3.06\%) + Bin(439, 3.15\%) + Bin(1172, 2.4\%) + Bin(2051, 2.18\%) + Bin(2461, 2.04\%) + Bin(2051, 1.95\%) + Bin(1172, 1.89\%) + Bin(439, 1.8\%) + Bin(439, 1.8\%) + Bin(10, 1.8\%)$$

The result is the following distribution illustrated in figure 12.

As for the uniform distribution, the following function is plotted:

$$f(w_n) = w_n + (1 + \theta_{SL})\mathbb{E}[(S_n - w_n)_+]$$

This gives figure 4 and a possible retention level that minimize the overall premium.

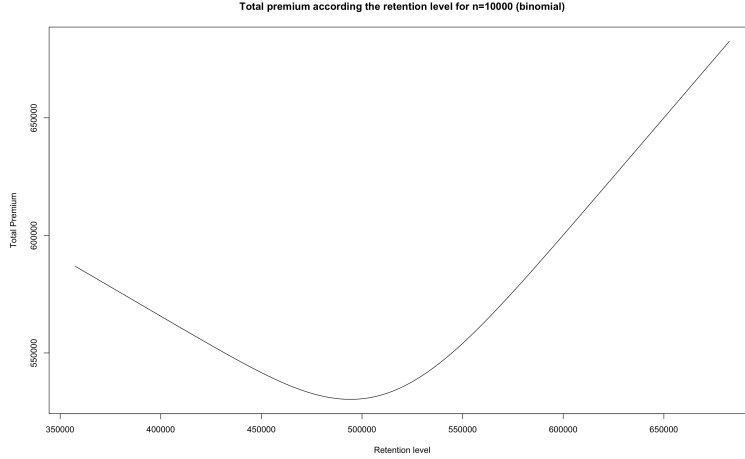


Figure 4: Total premium according the retention level for $n=10000$ (binomial)

As it can be seen in figure 4, the retention level that minimizes the overall premium is $w_{10000} = 494500$. It is now possible to calculate the probability of this retention level :

$$P(S_{10000} \leq 494500) = 33,82668\%$$

As for the uniform distributed dataset, the probability corresponding to the retention level is very low. The calculations for retention level $w_{10000} = 529500$ and $w_{10000} = 571500$ will also be done in parallel, because :

$$P(S_{10000} \leq 529500) = 70\%$$

$$P(S_{10000} \leq 571500) = 95\%$$

This makes it possible to try out different retention levels and observe the impact on premiums.

Now that the retention levels are available, the premiums can be easily computed (in exactly the same way as in the theory section). Therefore, P2P and stop-loss premiums are for each retention level shown in tables 29, 30 and 31.

The P2P premium and Stop Loss premium are then added together to obtain the pure premium (table 32). Once again, the pure premiums for $w_{10000} = 494500$ and $w_{10000} = 529500$ are very close. The same assumptions as for the uniformly distributed dataset for θ_{P2P} and for θ_{SL} will be used.

The following revenue for the platform are obtained:

$$w_{10000} = 526500 : \frac{530269,2}{1,2} \times 0,2 = 88378,20\text{€}$$

$$w_{10000} = 566500 : \frac{540130,4}{1,2} \times 0,2 = 90021,73\text{€}$$

$$w_{10000} = 615500 : \frac{572697,6}{1,2} \times 0,2 = 95449,50\text{€}$$

Taxes on insurance contacts of 9,25% must now be added, to give the commercial premiums shown in table 16.

Bike price	$w_{10000} = 494500$	$w_{10000} = 529500$	$w_{10000} = 571500$	Allianz premiums
500€	16,84€	16,90€	17,08€	30,60€
1000€	34,91€	35,16€	35,92€	63€
1500€	40,19€	40,61€	41,95€	
2000€	49,02€	49,70€	51,91€	87€
2500€	57,74€	58,74€	62,02€	102€
3000€	66,69€	68,08€	72,66€	117€
3500€	75,93€	77,76€	83,89€	132€
4000€	83,21€	85,49€	93,23€	144€
4500€	94,23€	97,13€	107,06€	162€
5000€	105,38€	108,97€	121,41€	180€

Table 16: Dataset with binomial distribution : commercial premiums

Once again, the premiums are much lower than those for Allianz. In the event of a problem, the platform can increase its premiums and still remain cheaper than conventional insurance. However, the premiums are slightly higher than those for the uniform distributed dataset. This is due to the fact that there are only 3760 bicycles priced strictly below 2500€ compared with 3770 for bicycles priced strictly above 2500€. There is therefore a slight over-representation of the more expensive bikes in the dataset, hence the slight increase in the price of the premium.

As with the uniform distribution, the individual premiums will be set in the following way :

$$\pi_i = (1 + \theta_{P2P})\mu_i$$

And the calculations leads to the same premiums as in table 9.

For the revenue, it is not the same as in the uniformly distributed dataset because the total premium is different. In this case the revenue is:

$$\frac{612494.2}{1,2} \times 0,2 = 102082,37\text{€}$$

Now that the premiums are available, the retention level can be calculated using the function plot in figure 4 and shows that $w_{10000} = 654060$. As for the uniform distributed dataset, $w_{10000} = 612430$ cannot be used because $P[S_{10000} = 612430] = 0$ since the total cost of claims is per 500. So $w_{10000} = 612500$. Therefore it is possible to calculate the stop loss premium with the retention level and to calculate the P2P premium like it has been done for uniform distributed dataset (table 33).

Then the tax insurance of 9,25% is added in order to calculate the commercial premiums and the same premiums as in table 10 are obtained. In this case, the retention probability is $P(S_{10000} \leq 612500) = 99,7\%$.

Regarding the surplus (once again, the expected cash-backs have been calculated in exactly the same way as in the methodology) in tables 17, 18, 19 and 20.

Bike price	$P[\text{cash-back}_i \geq 0]$	$\mathbb{E}[\text{cash-back}_i]$
500€	33,8%	0,046706
1000€	33,8%	0,189800
1500€	33,8%	0,323650
2000€	33,8%	0,517053
2500€	33,8%	0,747207
3000€	33,8%	1,015868
3500€	33,8%	1,323100
4000€	33,8%	1,624970
4500€	33,8%	2,028710
5000€	33,8%	2,470243

Table 17: Surplus, dataset with binomial distribution : $w_{10000} = 494500$

Bike price	$P[\text{cash-back}_i \geq 0]$	$\mathbb{E}[\text{cash-back}_i]$
500€	70%	0,146785
1000€	70%	0,602304
1500€	70%	1,037584
2000€	70%	1,674465
2500€	70%	2,444530
3000€	70%	3,357532
3500€	70%	4,417932
4000€	70%	5,482321
4500€	70%	6,915047
5000€	70%	8,507240

Table 18: Surplus, dataset with binomial distribution : $w_{10000} = 529500$

Bike price	$P[\text{cash-back}_i \geq 0]$	$\mathbb{E}[\text{cash-back}_i]$
500€	95%	0,330998
1000€	95%	1,371047
1500€	95%	2,385496
2000€	95%	3,887944
2500€	95%	5,732640
3000€	95%	7,952726
3500€	95%	10,569923
4000€	95%	13,250450
4500€	95%	16,882575
5000€	95%	20,981274

Table 19: Surplus, dataset with binomial distribution : $w_{10000} = 571500$

Bike price	$P[\text{cash-back}_i \geq 0]$	$\mathbb{E}[\text{cash-back}_i]$
500€	99,7%	0,526248
1000€	99,7%	2,196071
1500€	99,7%	3,851105
2000€	99,7%	6,325977
2500€	99,7%	9,401359
3000€	99,7%	13,146329
3500€	99,7%	17,613294
4000€	99,7%	22,260501
4500€	99,7%	28,592464
5000€	99,7%	35,824260

Table 20: Surplus, dataset with binomial distribution : $\pi_i = (1 + \theta_{P2P})\mu_i$

As expected and as it is for the uniformly distributed dataset, the higher the retention level is the higher the expected cash-back will be. However, cashback is slightly higher for the dataset distributed using a binomial distribution. Again, this can be explained by a slight over-representation of expensive bikes. This means that the less expensive bikes, which are more numerous in the uniformly distributed dataset, will have a greater impact on premiums, and so cash-back will be relatively lower.

As it can be observed, a slight over-representation of expensive bikes has had a slight impact on premiums, cash-backs and platform's incomes. The distribution of the dataset is therefore important, and it is very important to choose the right distribution in order to get the most accurate representation of reality and avoid unpleasant surprises.

Chapter 5

Interview with GRACQ : Perceptions and thoughts on collaborative bike insurance

During the research phase for this master thesis on modelling collaborative bicycle insurance with a focus on theft cover, I had the opportunity to conduct an informative interview with Ms Stéphanie Brixhe, a representative of the Groupe de Recherche et d'Action des Cyclistes Quotidiens (GRACQ) based in Brussels and Wallonia. The discussion provided valuable insights into the practical and ethical aspects of collaborative bike insurance, as well as an overview of the dynamics of cyclists' attitudes to theft protection.

GRACQ is a well-established association with over 18000 members across the Brussels and Walloon regions. Although GRACQ does not currently offer collaborative bike insurance, its commitment to promoting everyday cycling and road safety has led to a considered approach to the issue of bike insurance.

During the interview, Ms Brixhe shared information about the current insurance offered to GRACQ members, which focuses on legal protection in the event of cycling accidents. This protection gives members access to legal and medical advice in the event of an incident on the road. However, protection against bicycle theft is not included in this offer.

The idea of collaborative insurance, such as the one proposed by Laka, was discussed during the interview. Stéphanie Brixhe expressed a positive interest in this model, pointing out that GRACQ is in favour of collaborative approaches rather than traditional insurance companies.

However, concerns were also raised during the interview. Stéphanie Brixhe pointed out that the successful implementation of such insurance would require careful management and cost control to ensure its accessibility and effectiveness. She also touched on the complexity of legal proceedings in the event of serious accidents, illustrating the importance of legal protection for cyclists.

The issue of fairness was also addressed. GRACQ members were sensitive to the existing imbalance between the various stakeholders on the road. Although some members expressed concerns about the cost of bike insurance, the interview showed that the general perception of bike insurance was positive. The benefits of cover, particularly in the event of theft, were recognised, although issues of cost and accessibility were raised.

All in all, the interview with Ms Brixhe from GRACQ provided some interesting insights into how cyclists perceive bike insurance, and highlighted the importance of taking both economic and ethical aspects into account when modelling collaborative bike insurance. The thoughts shared during this interview enriched the understanding of the subject and helped to further guide the research into the design of collaborative bicycle insurance proposed in this master thesis.

Chapter 6

Conclusion

In this master's thesis, the concept of collaborative bicycle insurance has been extensively explored and thoroughly modeled the risk-sharing mechanism exemplified by Laka. The primary objective was to comprehensively understand this innovative approach to insurance, its advantages, disadvantages, and potential areas for improvement. As this study comes to an end, the significance of collaborative bike insurance has been reflected, its inherent strengths and weaknesses, and the critical role of accurate data in refining its outcomes.

Collaborative bicycle insurance, as exemplified by Laka's model, represents a departure from traditional insurance norms. It introduces a sense of communal responsibility and solidarity among cyclists, transcending the conventional transactional relationship between policyholders and insurers. The most prominent advantage of this model is the alignment of interests between policyholders and the insurance platform, fostering transparency and trust. Moreover, by pooling resources and distributing risks across a collective, collaborative insurance has the potential to offer more competitive premiums, thus making insurance coverage more accessible to a wider spectrum of cyclists.

However, as with any innovative concept, collaborative bicycle insurance is not immune to challenges. One key concern revolves around managing the collective and ensuring equitable cost distribution. Striking a balance between financial contributions from participants and ensuring affordability is a delicate undertaking. The effectiveness of such platforms hinges on active participation, necessitating sustained efforts to cultivate a sense of community and shared responsibility.

This exploration has illuminated several areas where collaborative bicycle insurance models could benefit from enhancement. Foremost among these is the improvement of the underlying data quality and granularity. Although the data sources

available have been effectively exploited, a more complete and realistic database would undoubtedly enable more accurate results to be obtained. Collaborative platforms could explore collaborations with governmental bodies, insurers, and other stakeholders to gain access to more reliable datasets. This, in turn, would enable more precise risk assessment, premium calculation, and loss projections, ultimately enhancing the overall efficacy of collaborative insurance models.

Another avenue for improvement lies in incorporating advanced predictive modeling techniques. Although the approach employed straightforward statistical distributions, the evolving realm of data science offers opportunities for more sophisticated algorithms. Machine learning and artificial intelligence could be leveraged to discern patterns, detect anomalies, and refine risk estimation, contributing to a more resilient and adaptable risk-sharing mechanism.

In conclusion, this master's thesis has delved into the realm of collaborative bicycle insurance, unraveling its nuances and shedding light on its mechanisms, benefits, and challenges. The central goal of this research - to model and elucidate the risk-sharing mechanism underpinning collaborative bike insurance exemplified by Laka - has been successfully achieved through meticulous research, thorough data analysis, and thoughtful deliberation.

Collaborative bicycle insurance, with its emphasis on community, shared responsibility, and affordability, holds the potential to disrupt the insurance sector. It presents an alternative to conventional insurance paradigms, founded on a sense of unity and collective support. While challenges exist, particularly in terms of cost management and participant engagement, they are not insurmountable.

Looking towards the future, the evolution of collaborative bike insurance is captivating. With technological advancements, richer datasets, and more sophisticated modeling techniques, these platforms could further reshape the insurance landscape. The insights and lessons derived from this research extend beyond bicycle insurance, influencing the design of collaborative insurance models in diverse domains.

In essence, the success of collaborative bicycle insurance hinges on the delicate balance between the collective's interests and the pragmatic aspects of risk management. It is a testament to the potency of innovation, community, and collaboration in shaping a more inclusive and responsive insurance ecosystem.

Bibliography

- [1] Michel Denuit & Christian Y.Robert (2022). *Allocation of benefits in mutual aids and survivor funds*.
- [2] Jake Frankenfield (2021). *Peer-to-Peer (P2P) Insurance*. <https://www.investopedia.com/terms/p/peertopeer-p2p-insurance.asp>. Accessed 10 April 2023.
- [3] Julia Kagan (2021). *Mutual Insurance Company: Definition and How They Invest*. <https://www.investopedia.com/terms/p/peertopeer-p2p-insurance.asp>. Accessed 10 April 2023.
- [4] *Laka Collective Fietsverzekeringen*. (n.d.) Laka Collective Fietsverzekeringen. <https://laka.co/be/>. Accessed 12 March 2023.
- [5] Michel Denuit (2022). *Modelling, measuring and managing property and casualty insurance risks*.
- [6] Dhaene J, Denuit M, Goovaerts MJ, Kaas R, Vyncke D (2002). *The concept of comonotonicity in actuarial science and finance: theory*. Insur Math Econ 31:3–33
- [7] Dhaene J, Linders D (2019). *Foundations of risk measurement*.
- [8] Jiao, Z., Liu, Y., & Wang, R. (2022). *An axiomatic theory for anonymized risk sharing*. arXiv preprint arXiv:2208.07533.
- [9] *Lemonade Insurance*. <https://www.lemonade.com/>. Accessed 10 April 2023.
- [10] *State Farm Insurance*. <https://www.statefarm.com/>. Accessed 10 April 2023.
- [11] Peter Ruckdeschel, Matthias Kohl, Thomas Stabla, Florian Camphausen (2022). *R-Package "distr"*. <https://distr.r-forge.r-project.org/distr.html>.
- [12] Peter Ruckdeschel, Matthias Kohl (2012). *General Purpose Convolution Algorithm in S_4 -Classes by means of FFT*.

Chapter 7

Annexe

7.1 Graphs

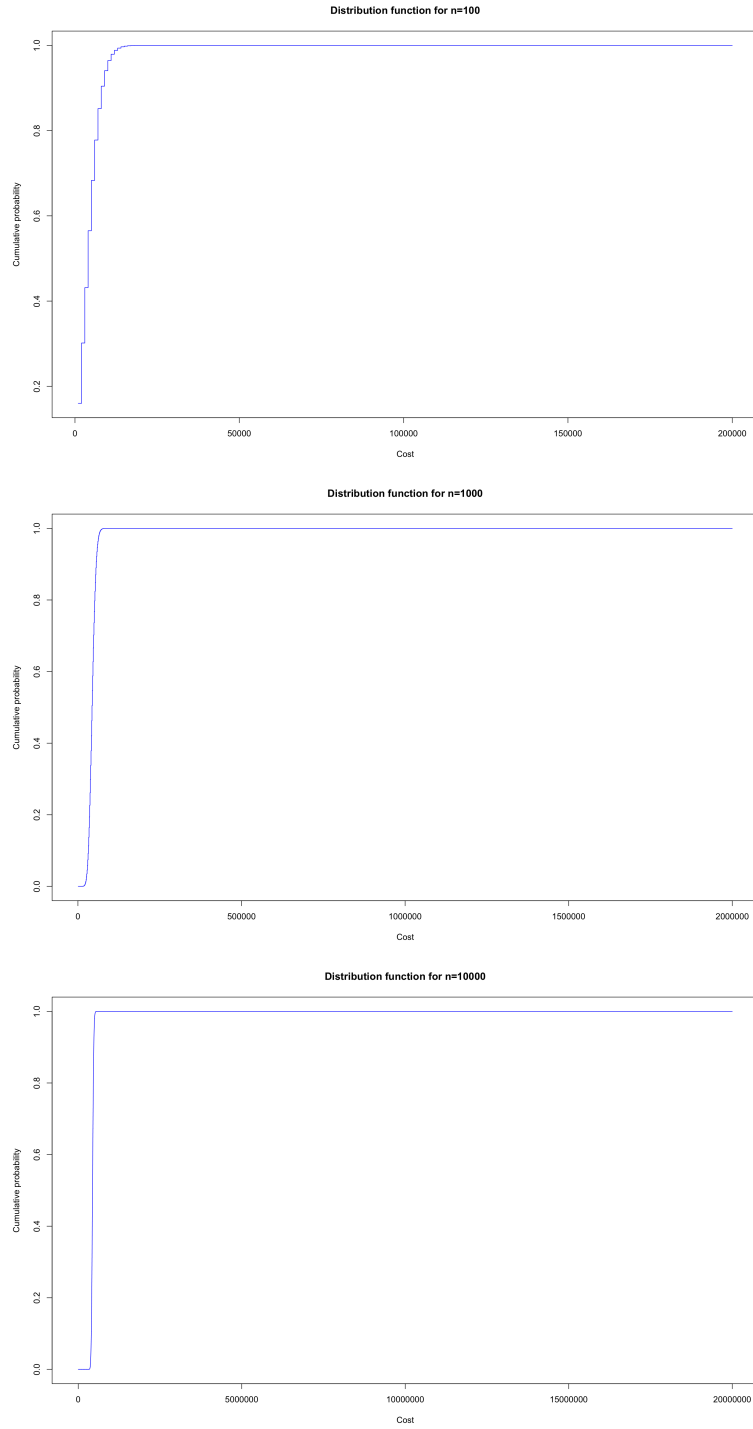


Figure 5: Distribution function for total claims for $n = 100$ (upper panel), $n = 1000$ (middle panel) and $n = 10000$ (lower panel)

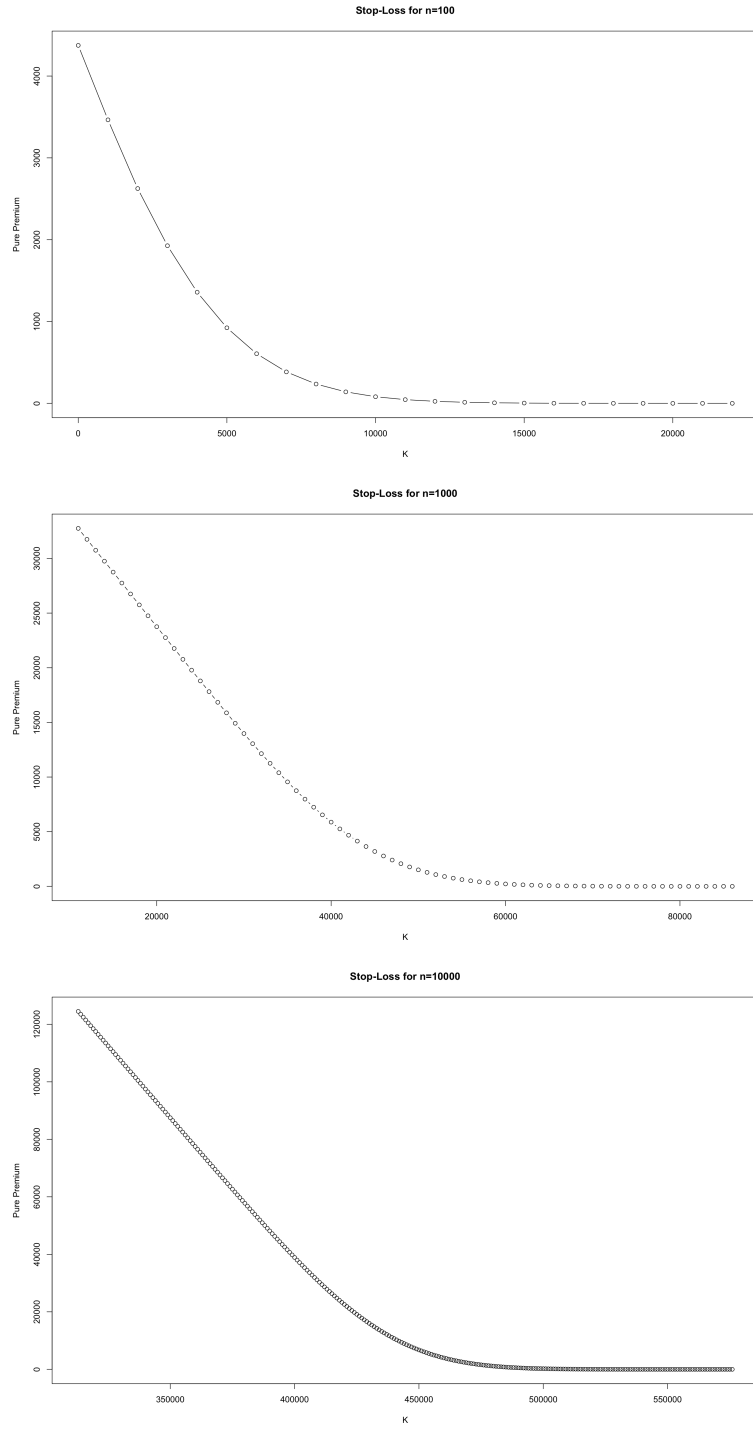


Figure 6: Stop-Loss for $n = 100$ (upper panel), $n = 1000$ (middle panel) and $n = 10000$ (lower panel)

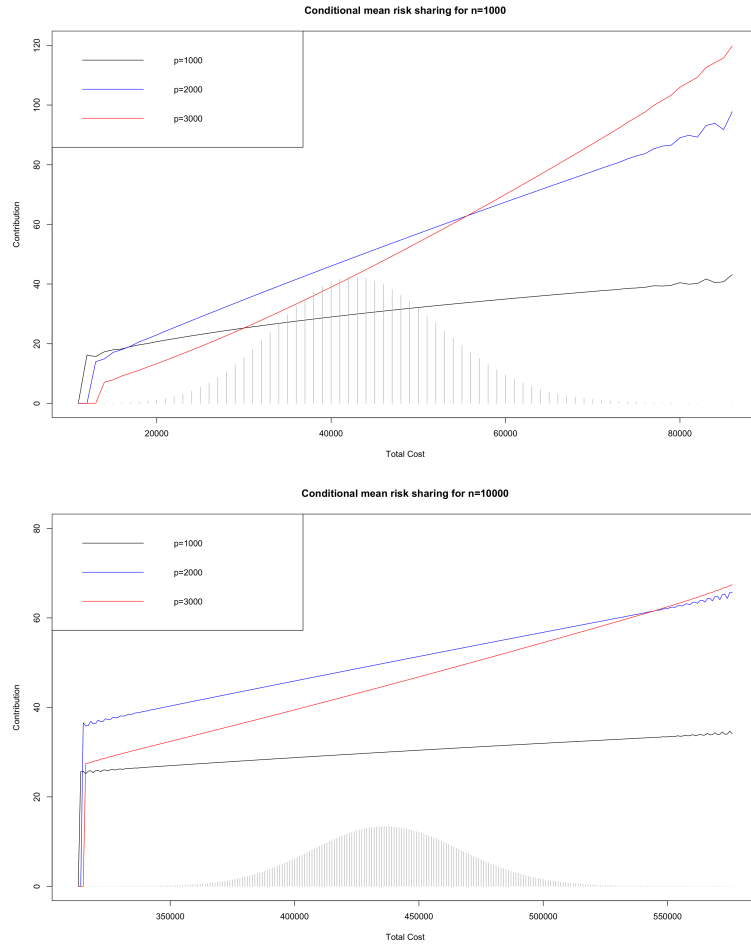


Figure 7: Conditional Mean Risk Sharing for $n = 1000$ (upper panel) and $n = 10000$ (lower panel)

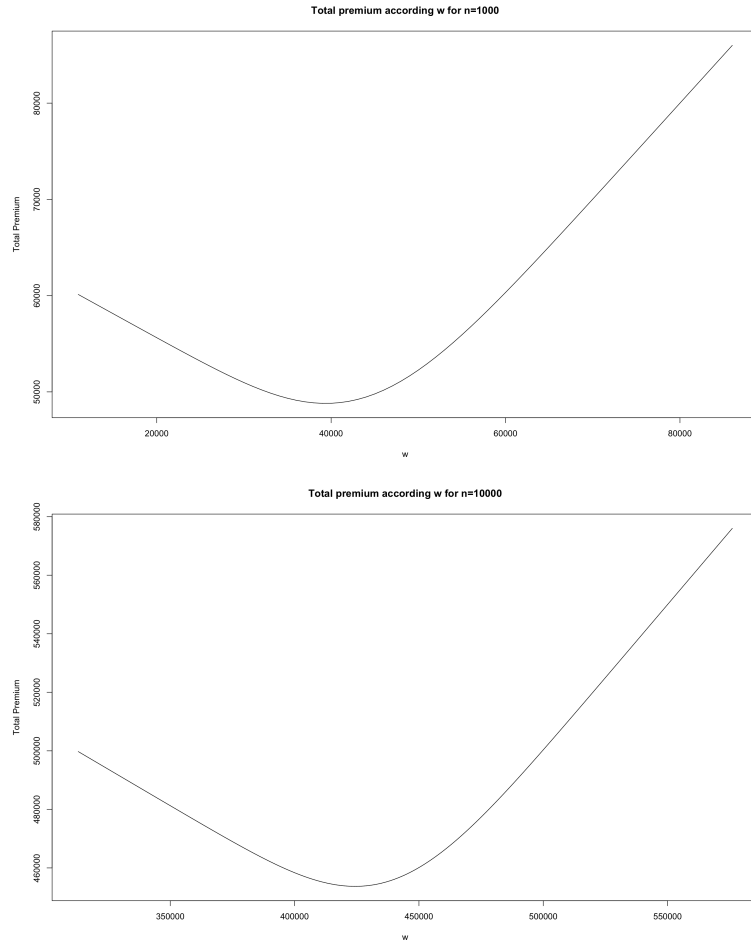


Figure 8: Total premium according to w for $n = 1000$ (upper panel) and $n = 10000$ (lower panel)

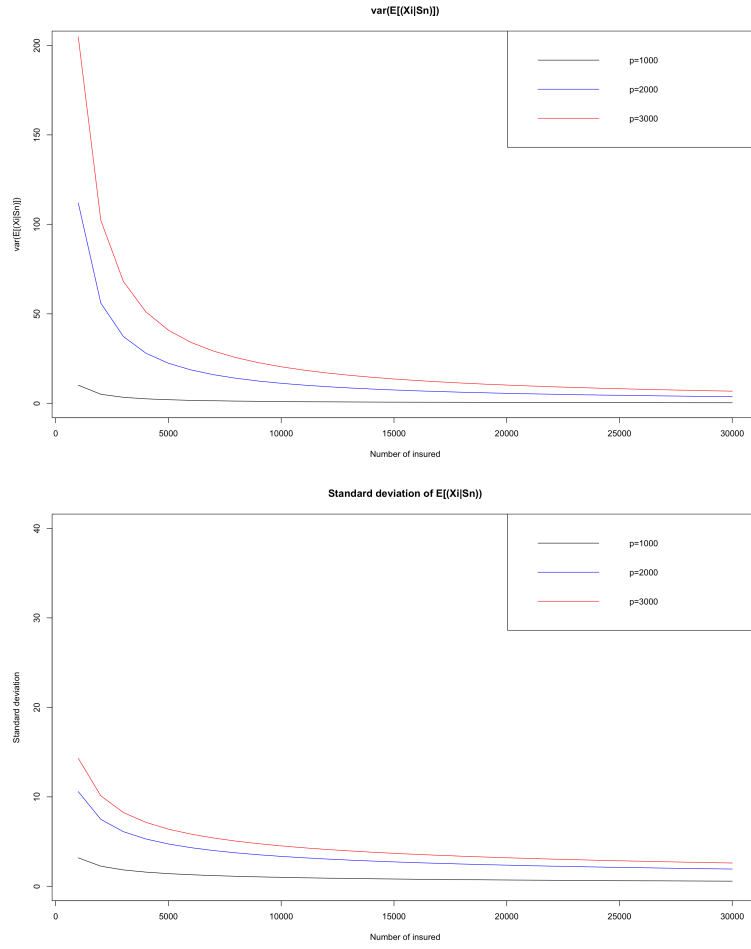


Figure 9: Variance of the participation according to the number of insured (upper panel) and standard deviation of the participation according to the number of insured (lower panel)

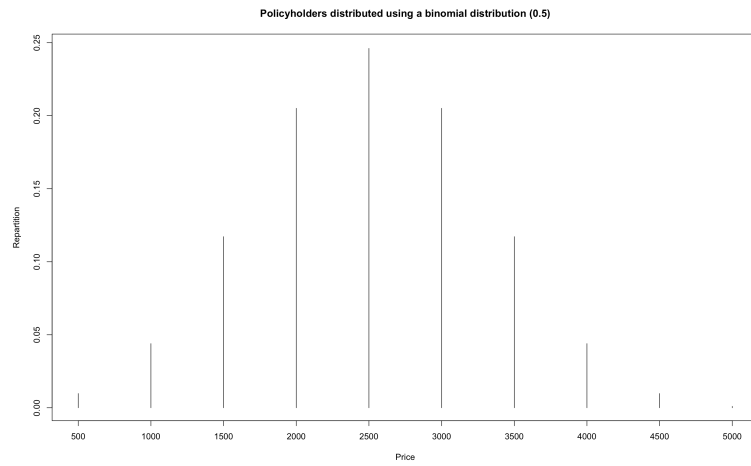


Figure 10: Policyholders distributed using a binomial distribution (0,5)

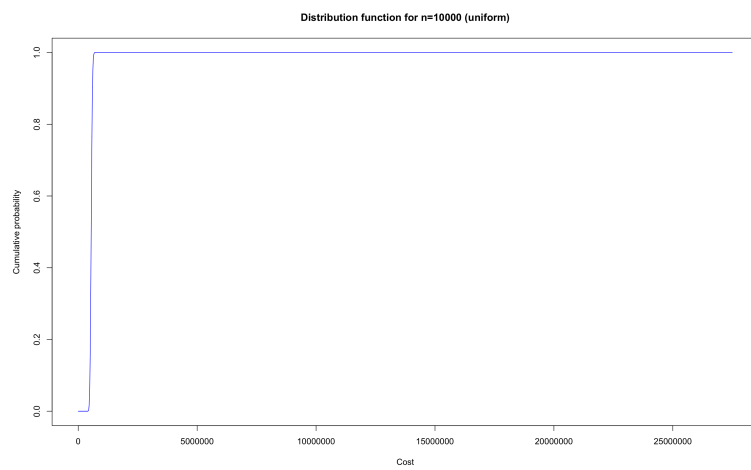


Figure 11: Distribution function for n=10000 (uniform)

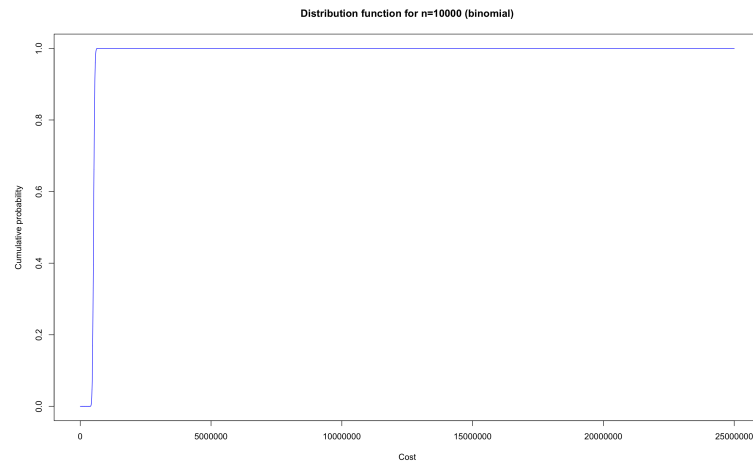


Figure 12: Distribution function for $n=10000$ (binomial)

7.2 Tables

Group	1000€	2000€	3000€
π_i	40	50	70
w_i	32,14549	56,9309	53,99256
π_{P2P}	39,36204	47,60780	66,35237
π_{SL}	0,6379628	2,392203	3,647634

Table 21: Premiums for $n = 1000$

Group	1000€	2000€	3000€
π_i	40	50	70
w_i	32,71152	59,44358	58,40978
π_{P2P}	39,99908	49,99665	69,99497
π_{SL}	0,0009157543	0,0033501555	0,0050337999

Table 22: Premiums for $n = 10000$

Price	Pure premium
450€	14,40€
500€	15,30€
600€	17,70€
700€	19,20€
800€	22,20€
900€	26,70€
1000€	31,50€
1200€	33€
1400€	34,50€
1600€	37,50€
1800€	40,50€
2000€	43,50€
2500€	51€
3000€	58,50€
3500€	66€
4000€	72€
4500€	81€
5000€	90€

Table 23: Pure premiums

Bike price	π_{P2P}	π_{SL}
500€	15,22845	0,1683435
1000€	31,20223	0,6949435
1500€	35,47935	1,205264
2000€	42,74754	1,957775
2500€	49,73756	2,876822
3000€	56,74127	3,977234
3500€	63,80311	5,267785
4000€	69,0466	6,580409
4500€	77,22329	8,3549
5000€	85,29007	10,34672

Table 24: Dataset uniformly distributed : $w_{10000} = 526500$

Bike price	π_{P2P}	π_{SL}
500€	15,39402	0,048183
1000€	31,88400	0,200503
1500€	36,65856	0,350656
2000€	44,65776	0,574327
2500€	52,53674	0,850954
3000€	60,60031	1,186235
3500€	68,89989	1,584150
4000€	75,39506	1,995341
4500€	85,26043	2,554207
5000€	95,21427	3,189023

Table 25: Dataset uniformly distributed : $w_{10000} = 566500$

Bike price	π_{P2P}	π_{SL}
500€	15,58188	0,005037
1000€	32,66608	0,021163
1500€	38,02694	0,037397
2000€	46,89995	0,061882
2500€	55,86031	0,092629
3000€	65,23529	0,130461
3500€	75,09216	0,176000
4000€	83,19788	0,223932
4500€	95,25241	0,289583
5000€	107,69429	0,365204

Table 26: Dataset uniformly distributed : $w_{10000} = 615500$

Bike price	$w_{10000} = 526500$	$w_{10000} = 566500$	$w_{10000} = 615500$
500€	15,40€	15,44€	15,59€
1000€	31,90€	32,08€	32,69€
1500€	36,68€	37,01€	38,06€
2000€	44,71€	45,23€	46,96€
2500€	52,61€	53,39€	55,95€
3000€	60,72€	61,79€	65,37€
3500€	69,07€	70,48€	75,27€
4000€	75,63€	77,39€	83,42€
4500€	85,58€	87,81€	95,54€
5000€	95,64€	98,40€	108,06€

Table 27: Pure premiums for uniformly distributed dataset

Bike price	π_{P2P}	π_{SL}
500€	18,35954	0,000463
1000€	37,79804	0,001960
1500€	43,19651	0,003491
2000€	52,31418	0,005822
2500€	61,19122	0,008783
3000€	70,18753	0,012469
3500€	79,36305	0,016950
4000€	86,37829	0,021707
4500€	97,17169	0,028306
5000€	107,96404	0,035960

Table 28: $\pi_i = (1 + \theta_{P2P})\mu_i$, uniformly distributed dataset

Bike price	π_{P2P}	π_{SL}
500€	15,21736	0,194027
1000€	31,15546	0,801517
1500€	35,39664	1,390522
2000€	42,61061	2,259671
2500€	49,53268	3,321886
3000€	56,45295	4,594374
3500€	63,41475	6,087511
4000€	68,55351	7,607150
4500€	76,58756	9,661666
5000€	84,49110	11,968623

Table 29: Dataset with binomial distribution : $w_{10000} = 494500$

Bike price	π_{P2P}	π_{SL}
500€	15,41008	0,055065
1000€	31,94923	0,229612
1500€€	36,76940	0,402269
2000€	44,83444	0,660043
2500€	52,79138	0,979729
3000€	60,94544	1,368141
3500€	69,34774	1,830277
4000€	75,94272	2,309381
4500€	85,94092	2,961141
5000€	96,03841	3,703165

Table 30: Dataset with binomial distribution : $w_{10000} = 529500$

Bike price	π_{P2P}	π_{SL}
500€	15,62708	0,005879
1000€	32,85452	0,024793
1500€	38,35619	0,043958
2000€	47,43931	0,072964
2500€	56,65959	0,109590
3000€	66,34950	0,154848
3500€	76,58019	0,209599
4000€	85,07204	0,267610
4500€	97,65110	0,347168
5000€	110,68842	0,439223

Table 31: Dataset with binomial distribution : $w_{10000} = 571500$

Bike price	$w_{10000} = 494500$	$w_{10000} = 529500$	$w_{10000} = 571500$
500€	15,41€	15,47€	15,63€
1000€	31,96€	32,18€	32,88€
1500€	36,79€	37,17€	38,40€
2000€	44,87€	45,49€	47,51€
2500€	52,85€	53,77€	56,77€
3000€	61,05€	62,31€	66,50€
3500€	69,50€	71,18€	76,79€
4000€	76,16€	78,25€	85,34€
4500€	86,25€	88,90€	98,00€
5000€	96,46€	99,74€	111,13€

Table 32: Dataset with binomial distribution : pure premiums

Bike price	π_{P2P}	π_{SL}
500€	18,35971	0,000292
1000€	37,79875	0,001246
1500€	43,19777	0,002231
2000€	52,31628	0,003724
2500€	61,19434	0,005659
3000€	70,19192	0,008085
3500€	79,36894	0,011061
4000€	86,38572	0,014275
4500€	97,18126	0,018739
5000€	107,97605	0,023952

Table 33: Dataset with binomial distribution : $\pi_i = (1 + \theta_{P2P})\mu_i$

7.3 Interview with GRACQ

Maxim Mizrahi: Oui, bonjour Maxim Mizrahi, je sais pas si vous vous rappelez de moi je vous avais appelé par rapport à mon mémoire, il y a une semaine à-peu-près, et vous m'avez dit de vous rappeler le trois ou le quatre pour juste vous poser quelques questions par rapport au GRACQ. Je ne sais pas si vous rappelez. Je veux savoir si c'est possible de vous poser juste quelques questions. Je ne sais pas si vous rappelez le sujet de mon mémoire.

Stéphanie Brixhe: Je ne me souviens plus précisément du sujet.

Maxim Mizrahi: C'est juste pour un mémoire, c'est pas du tout dans la réalité, mais je suis en train de mettre en place une assurance collaborative pour le vol de vélo, et juste vous poser quelques questions par rapport à ça. Si ça ne vous dérange pas. Est-ce que je peux commencer ? Si ça ou va ?

Stéphanie Brixhe: Oui oui

Maxim Mizrahi: La première question, c'est déjà: est-ce que, est-ce que les cyclistes, au GRACQ souscrivent une assurance vélo? Et si oui, c'est pour quelle garantie exactement ?

Stéphanie Brixhe: Une assurance en protection juridique. Les membres du GRACQ sont couverts par une assurance en protection juridique, mais pas par une protection vélo. La protection juridique c'est en cas d'accident où par exemple, vous seriez blessé. Vous avez, avec la protection juridique, la possibilité de prendre un avocat ou de prendre un médecin expert. Elle est déjà intervenue pour des cas de lié au code de la route. Par exemple, au moment du confinement, la région avait déclaré qu'elle mettait des signaux B22, B23 sur la rue de la loi, et il y a une soixantaine de cycliste qui ont franchi le feu rouge et la protection juridique a fait sauter un ou deux PV sur des membre du GRACQ enfaite.

Maxim Mizrahi: Ok très bien. Ensuite, la prochaine question, est-ce que vous savez, ce qu'est une assurance collaborative ou pas du tout ?

Stéphanie Brixhe: C'est une assurance mutualiste enfaite on va dire, c'est ça ?.

Maxim Mizrahi: Pas tout à fait. On va mettre en place un fond pour les assurés, où on va mettre toutes leurs primes dedans et puis, si jamais, en fonction des garanties auxquels ils souscrivent, si quelque chose arrive, ils peuvent utiliser ce fond et ça permet d'avoir beaucoup moins de coûts qu'une assurance classique, parce qu'il y a juste un coup de gestion du fond et un coup de gestion des sinistres, et puis ça s'arrête là. Il n'y a pas toute une assurance derrière, vu que les gens s'assurent entre eux. Et je me demandais si, maintenant que vous savez ce qu'est une assurance collaborative, vous savez peut-être quels sont les avantages d'une assurance collaborative par rapport à une assurance traditionnelle ?

Stéphanie Brixhe: Disons que nous, on n'est pas vraiment lié à une société d'assurance donc nous on travaille avec le bureau des balles et la dèsses, mais donc c'est un peu difficile de m'avancer pour tout le GRACQ, mais je suis certaine que

on est plus pour des choses collaboratives que des sociétés privées. Par contre, nous on ne pourrait pas mettre un système d'assurances collaboratives en place, parce qu'on a même réfléchi de le faire. Car il arrive que des personnes qui ne soient pas membres du GRACQ, je pense notamment à un monsieur à Liège, qui a eu un accident et il n'était pas membre du GRACQ, et cette personne était demandeuse d'emplois et avait très peu de moyens. Il a subi un accident vraiment très grave, avec quelqu'un qui a fait exprès de faire cet accident. On s'est demandé : est-ce que le GRACQ allait soutenir et aller prendre un avocat pour lui ? Et on a discuté de voir si on ne pouvait pas créer un fond. Et c'est très compliqué, surtout dans des cas d'accidents par exemple, ça peut chiffrer énormément les prestations d'un avocat dans un accident, à partir du moment où l'accident est grave. Parce que les compagnies de la partie adverse vont tout le temps essayé de rabaisser les montants et l'avocat va devoir te démontrer, avec tes médecins experts, etc. . . Ça prend énormément de temps et Parfois, ce sont des procès qui prennent dix ans.

Maxim Mizrahi: Ma prochaine question, est-ce que vous seriez favorable à une mise en place d'une telle assurance pour les moments du GRACQ ? Mais ce ne serait évidemment pas vous qui mettez tout en place. Ça existe déjà, je ne sais pas si vous connaissez Laka, c'est une assurance collaborative qui est déjà mise en place, et je sais qu'en France, ils ont un partenariat avec Décathlon, où les gens qui achètent un vélo chez Décathlon ont, je pense, un mois offert chez Laka, quelque chose comme ça, et c'est exactement le même principe que ce que je fais. Laka est donc une assurance qui met en place un fonds pour les qui assure les assurés entre eux en mettant en place un fond, et puis ils prennent un petit pourcentage sur leur prime pour pouvoir faire tourner la plateforme, parce qu'ils ont besoin d'employés etc. . . Je me demandais donc si vous seriez favorable à la mise en place d'une telle assurance au GRACQ. Évidemment, comme je vous ai dit, ce ne sera pas à vous de mettre en place le fond, c'est juste que vous souscrivez à une assurance qui fonctionne de cette manière.

Stéphanie Brixhe: En tous les cas, on ne serait pas contre, a Priori, mais on devrait s'assurer du bon fonctionnement de ce système et du coût. Mais a priori forcément, on ne sera pas contre ce genre de pratiques.

Maxim Mizrahi: Ok, je pense que c'est à-peu-près tout. Une petite dernière question un peu générale : quelle est votre perception par rapport aux assurances vélo en Belgique ? Vous voyez ça plutôt comme quelque chose de mal où il y a la volonté de faire du bénéfice ou plutôt comme quelque chose de bien ?

Stéphanie Brixhe: Je pense que, franchement, c'est super utile. Après le coût est conséquent pour les personnes qui font du vélo et malheureusement, les personnes qui ont moins de moyens sont souvent le moins bien assurées et ça c'est quand même assez problématique. C'est pour ça qu'en fait pour les personnes qui ont peu de moyens, le GRACQ peut vraiment être intéressant rien qu'avec cette protection

juridique. Ça ne concerne pas le vol de vélo, on est d'accord, mais au moins, c'est déjà le minimum. Comme je vous le disais ce monsieur à Liège c'est horrible ce qui lui est arrivé et en plus, il a dû se défendre tout seul. Enfin, voilà, on peut imaginer qu'en fait, au plus, vous êtes pauvre au moins vous savez vous défendre. Je trouve ça, je trouve ça profondément horrible. Mais je veux dire, je ne suis pas forcément contre les assurances vélo même si elles coûtent chères, forcément je ne peux être que pour parce que voilà, il y a une bénévoles qui vient de temps en temps ici. Elle s'est fait voler son vélo. Elle avait une assurance vol et elle s'est faite rembourser un vélo. Donc je ne vois pas comment je pourrais être contre. Je trouve qu'il y a vraiment des facilités à être assuré comme dans tout.

Maxim Mizrahi: Oui, évidemment, après ça dépend de la perception du risque de chaque personne. Il y en a qui juge que ce n'est pas nécessaire et qui voient un peu les assureurs comme des grands méchants. Et c'est pour ça que je vous pose la question.

Stéphanie Brixhe: J'ai fait une étude sur la protection juridique, donc j'ai appelé des membres qui ont eu des accidents. Eux ont rien contre notre protection juridique, mais bien contre les assurances des parties adverses et donc eux sont vraiment des grands méchant ça c'est clair.

Maxim Mizrahi: C'est ça, un peu le souci, c'est que les vélos ne s'assurent pas forcément avec une protection juridique, mais les autres usagers de la route ont ce type d'assurance et c'est ça qui pose un peu problème et qui crée un déséquilibre.

Stéphanie Brixhe: Non ce n'est même pas le fait que la partie adverse ai une protection juridique. La seule chose c'est que vous en tant que cycliste, même quand vous avez la protection juridique, par exemple quelqu'un me racontait qu'il a eu un accident, il a eu trois vertèbres cassées. Donc il a eu une lourde hospitalisation, 3 mois d'hospitalisation, l'assurance de la partie adverse est venu sur son lit en lui disant : « monsieur, bonjour, mais vous allez mieux ». Le gars répondu oui. Donc on lui a remis un décompte de ce qu'il aurait perdu comme mobilité, etc. . . . Que cette personne jugeât insuffisant. Après 3 mois d'hôpital, oui il allait mieux que la semaine suivant son accident quoi. Donc forcément, ce ne sont pas des entreprises caritatives et donc forcément, s'ils peuvent donner moins, ils vont donner moins et c'est la même chose dans tout ce qui est assurance. Sans contre expert c'est compliqué, c'est grâce au contre expert que vous arrivez à relever les montants que les assurances vous octroient. En incendie c'est la même chose, il y a toujours des contres expertises possibles mais sans protection juridique vous ne pouvez pas faire appel à un contre expert. Enfin vous pouvez le faire mais ça coute 1800 euros.

Maxim Mizrahi: Et vous savez, vous n'êtes pas obligé de me dire, mais à-peu-près à combien est la prime pour une protection juridique pour les vélos ?

Stéphanie Brixhe: Mais nous on bénéficie d'un tarif enfaite, parce qu'on assure les personnes en grand nombre. On est plus ou moins à 18 euros par an.

Maxim Mizrahi: Effectivement ce n'est pas très cher pour, pour une assurance juridique.

Stéphanie Brixhe: Non c'est ce que je vous dis, pour être concurrentiel à ça, c'est très compliqué. Si vous imaginez que vous avez un accident, dans lequel vous avez des années de frais d'avocat etc... C'est très cher.

Maxim Mizrahi: Juste pour avoir une idée, vous vous avez à-peu-près combien de cyclistes dans votre groupe? Et la deuxième question, c'est combien de sinistres vous avez par année environ ?

Stéphanie Brixhe: Dans les sinistres, par exemple il y en a qui ouvrent sinistre et puis pour finir, il y a rien qui se passe au niveau de la compagnie donc ça veut un peu rien dire. Vous avez des sinistres qui sont ouverts, la personne a eu peur, a des bleus, va à l'hôpital et puis enfaite il n'y a pas de suite qui est donnée parce que la majorité des accidents ne sont pas des accidents très graves. Moi je dirais qu'il doit y avoir un accident grave par an.

Maxim Mizrahi: Et vous avez à-peu-près combien de cyclistes dans le groupement ?

Stéphanie Brixhe: Dans notre groupement on a 18000 membres, mais pas 18000 assurés. On a 3000 membres cotisant.

Maxim Mizrahi: Et ça c'est pour toute la Belgique ou uniquement pour Bruxelles ?

Stéphanie Brixhe: Pour Bruxelles et la Wallonie.

Maxim Mizrahi: Je pense que c'est à-peu-près tout pour les questions. C'était très gentil de votre part d'avoir pris le temps de me répondre. Ça va beaucoup m'aider pour mon mémoire.

Stéphanie Brixhe: Eh bien bonne chance à vous alors.

Maxim Mizrahi: Merci beaucoup et je vous souhaite une très bonne journée. Au revoir.

7.4 R code

```
options(encoding = 'UTF-8')
install.packages("distr")
install.packages("distrEx")
install.packages("ggplot2")

rm(list=ls())

library("distrEx")
library("distr")
library("ggplot2")

#n=100
n=100 #Number of insured
lmax=n*0.25*1000+n*0.5*2000+n*0.25*3000 #Maximum cost

k<-seq(1000,lmax,by=1000) #All cost values

bin1 <- 1000*Binom(n*0.25,0.03) #First category
bin2 <- 2000*Binom(n*0.5,0.025) #Second category
bin3 <- 3000*Binom(n*0.25,0.015) #Third category

bin <- bin1 + bin2 + bin3 #Portfolio distribution

plot(k,p(bin)(k),type="s",main="Distribution▯function▯for▯n=100",
      xlab="Cost",ylab="Cumulative▯probability", col = "blue")

#Removing extrem values
eps <- 10−6/2
low <- q(bin)(eps)
up <- q(bin)(1−eps)

k <- seq(low,up,1000)

plot(k,d(bin)(k),type="l",main="Density▯function▯for▯n=100",
      xlab="Cost",ylab="Probability")

#Stop-Loss function
f <- function(x, k) {
```

```

    ifelse(x-k > 0, x-k, 0)
  }

stoploss <- c()

for (i in 1:(up/1000-low/1000+1)) {
  stoploss[i] <- sum(f(k,(i-1)*1000)*d(bin)(k))
}

plot(k, stoploss, type="b", main="Stop-Loss for n=100",
      xlab="K", ylab="Pure Premium")

#Participation

bin1m <- 1000*Binom(floor((n-1)*0.25), 0.03)
binm1 <- bin1m + bin2 + bin3

x1 <- c()

for (i in 1:((up-low)/1000+1)) {
  x1[i] <- (1000*d(binm1)((i-1)*1000-1000)*0.03)/d(bin)((i-1)*1000)
}

bin2m <- 2000*Binom(floor((n-1)*0.5), 0.025)
binm2 <- bin1 + bin2m + bin3

x2 <- c()

for (i in 1:((up-low)/1000+1)) {
  x2[i] <- (2000*d(binm2)((i-1)*1000-2000)*0.025)/d(bin)((i-1)*1000)
}

bin3m <- 3000*Binom(floor((n-1)*0.25), 0.015)
binm3 <- bin1m + bin2 + bin3m

x3 <- c()

for (i in 1:((up-low)/1000+1)) {
  x3[i] <- (3000*d(binm3)((i-1)*1000-3000)*0.015)/d(bin)((i-1)*1000)
}

```

```

plot(k,x1,type="l",main="Conditional_mean_risk_sharing_for_n=100",
      xlab="Total_Cost",ylab="Contribution", col = "black",ylim=c(0,200))
lines(k, x2, col = "blue", lty = 1)
lines(k,x3, col = "red", lty = 1)
lines(k,d(bin)(k)*1000,type="h",col="grey")
legend("topleft", legend = c("p=1000", "p=2000", "p=3000"),
      col = c("black", "blue", "red"), lty=1)

#n=1000
n=1000
lmax=n*0.25*1000+n*0.5*2000+n*0.25*3000

k<-seq(1000,lmax,by=1000)

bin1 <- 1000*Binom(n*0.25,0.03)
bin2 <- 2000*Binom(n*0.5,0.025)
bin3 <- 3000*Binom(n*0.25,0.015)

bin <- bin1 + bin2 + bin3
plot(k,p(bin)(k),type="s",main="Distribution_function_for_n=1000",
      xlab="Cost",ylab="Cumulative_probability", col="blue")

#Removing extrem values
eps <- 10^(-6)/2
low <- q(bin)(eps)
up <- q(bin)(1-eps)

k <- seq(low,up,1000)

plot(k,d(bin)(k),type="l",main="Density_function_for_n=1000",
      xlab="Cost",ylab="Probability")

stoploss <- c()

for (i in 1:(up/1000-low/1000+1)) {
  stoploss[i] <- sum(f(k,(i-1)*1000+low)*d(bin)(k))
}

plot(k,stoploss,type="b",main="Stop-Loss_for_n=1000",

```

```

xlab="K",ylab="Pure□Premium")

bin1m <- 1000*Binom(floor((n-1)*0.25),0.03)
binm1 <- bin1m + bin2 + bin3

x1 <- c()

for (i in 1:((up-low)/1000+1)) {
  x1[i] <- (1000*d(binm1)((i-1)*1000+low-1000)*0.03)
  /d(bin)((i-1)*1000+low)
}

bin2m <- 2000*Binom(floor((n-1)*0.5),0.025)
binm2 <- bin1 + bin2m + bin3

x2 <- c()

for (i in 1:((up-low)/1000+1)) {
  x2[i] <- (2000*d(binm2)((i-1)*1000+low-2000)*0.025)
  /d(bin)((i-1)*1000+low)
}

bin3m <- 3000*Binom(floor((n-1)*0.25),0.015)
binm3 <- bin1 + bin2 + bin3m

x3 <- c()

for (i in 1:((up-low)/1000+1)) {
  x3[i] <- (3000*d(binm3)((i-1)*1000+low-3000)*0.015)
  /d(bin)((i-1)*1000+low)
}

plot(k,x1,type="l",main="Conditional□mean□risk□sharing□for□n=1000",
      xlab="Total□Cost",ylab="Contribution", col = "black",ylim=c(0,120))
lines(k, x2, col = "blue", lty = 1)
lines(k,x3, col = "red", lty = 1)
lines(k,d(bin)(k)*1000,type="h",col="grey")
legend("topleft", legend = c("p=1000", "p=2000", "p=3000"),
col = c("black", "blue", "red"), lty=1)

```

```

w1000 <- q(bin)(0.8)

w11000 <- (1000*d(binm1)(w1000-1000)*0.03)/d(bin)(w1000)
w21000 <- (2000*d(binm2)(w1000-2000)*0.025)/d(bin)(w1000)
w31000 <- (3000*d(binm3)(w1000-3000)*0.015)/d(bin)(w1000)

SL11000 <- 1.5*sum(f(x1,w11000)*d(bin)(k))
SL21000 <- 1.5*sum(f(x2,w21000)*d(bin)(k))
SL31000 <- 1.5*sum(f(x3,w31000)*d(bin)(k))

purepremium1000 <- c(w11000+SL11000,w21000+SL21000,w31000+SL31000)

CB11000 <- sum(f(w11000,x1)*d(bin)(k))
CB21000 <- sum(f(w21000,x2)*d(bin)(k))
CB31000 <- sum(f(w31000,x3)*d(bin)(k))

#Second method

w <- seq(low,up,10)

stoploss2 <- c()

for (i in 1:(up/10-low/10+1)) {
  stoploss2[i] <- sum(f(w,(i-1)*10+low)*d(bin)(w))
}

Totalprem1000 <- w + 1.5*stoploss2

plot(w,Totalprem1000, type="l",
main="Total premium according to w for n=1000", ylab="Total Premium")

tab1000 <- data.frame(w,Totalprem1000)

pi1000 <- 40*n*0.25 + 50*n*0.5 + 70*n*0.25

w1000 <- tab1000[which(abs(Totalprem1000-pi1000)==min(
abs(Totalprem1000-pi1000))),1]
#Selecting the w so that it correspond to the pure premium
w1000 <- round(w1000,-3)
#Can only be in thousand otherwise the probability will be 0

```

```

p(bin)(w1000)

w11000 <- (1000*d(binm1)(w1000-1000)*0.03)/d(bin)(w1000)
w21000 <- (2000*d(binm2)(w1000-2000)*0.025)/d(bin)(w1000)
w31000 <- (3000*d(binm3)(w1000-3000)*0.015)/d(bin)(w1000)

SL11000 <- 1.5*sum(f(x1,w11000)*d(bin)(k))
SL21000 <- 1.5*sum(f(x2,w21000)*d(bin)(k))
SL31000 <- 1.5*sum(f(x3,w31000)*d(bin)(k))

P2P_1 <- 40 - SL11000
P2P_2 <- 50 - SL21000
P2P_3 <- 70 - SL31000

CB11000 <- sum(f(w11000,x1)*d(bin)(k))
CB21000 <- sum(f(w21000,x2)*d(bin)(k))
CB31000 <- sum(f(w31000,x3)*d(bin)(k))

#n=10000
n=10000
lmax=n*0.25*1000+n*0.5*2000+n*0.25*3000

k<-seq(1000,lmax,by=1000)

bin1 <- 1000*Binom(n*0.25,0.03)
bin2 <- 2000*Binom(n*0.5,0.025)
bin3 <- 3000*Binom(n*0.25,0.015)

bin <- bin1 + bin2 + bin3
plot(k,p(bin)(k),type="s",main="Distribution function for n=10000",
      xlab="Cost",ylab="Cumulative probability", col = "blue")

#Removing extrem values
eps <- 10^(-6)/2
low <- q(bin)(eps)
up <- q(bin)(1-eps)

k <- seq(low,up,1000)

```

```

plot(k,d(bin)(k),type="l",main="Density function for n=10000",
      xlab="Cost",ylab="Probability")

stoploss <- c()

for (i in 1:(up/1000-low/1000+1)) {
  stoploss[i] <- sum(f(k,(i-1)*1000+low)*d(bin)(k))
}

plot(k,stoploss,type="b",main="Stop-Loss for n=10000",
      xlab="K",ylab="Pure Premium")

bin1m <- 1000*Binom(floor((n-1)*0.25),0.03)
binm1 <- bin1m + bin2 + bin3

x1 <- c()

for (i in 1:((up-low)/1000+1)) {
  x1[i] <- (1000*d(binm1)((i-1)*1000+low-1000)*0.03)
    /d(bin)((i-1)*1000+low)
}

bin2m <- 2000*Binom(floor((n-1)*0.5),0.025)
binm2 <- bin1 + bin2m + bin3

x2 <- c()

for (i in 1:((up-low)/1000+1)) {
  x2[i] <- (2000*d(binm2)((i-1)*1000+low-2000)*0.025)
    /d(bin)((i-1)*1000+low)
}

bin3m <- 3000*Binom(floor((n-1)*0.25),0.015)
binm3 <- bin1 + bin2 + bin3m

x3 <- c()

for (i in 1:((up-low)/1000+1)) {
  x3[i] <- (3000*d(binm3)((i-1)*1000+low-3000)*0.015)
    /d(bin)((i-1)*1000+low)
}

```

```

}

plot(k,x1,type="l",main="Conditional mean risk sharing for n=10000",
      xlab="Total Cost",ylab="Contribution", col = "black",ylim=c(0,80))
lines(k, x2, col = "blue", lty = 1)
lines(k,x3, col = "red", lty = 1)
lines(k,d(bin)(k)*1000,type="h",col="grey")
legend("topleft", legend = c("p=1000", "p=2000", "p=3000"),
      col = c("black", "blue", "red"), lty=1)

w10000 <- q(bin)(0.8)

w110000 <- (1000*d(binm1)(w10000-1000)*0.03)/d(bin)(w10000)
w210000 <- (2000*d(binm2)(w10000-2000)*0.025)/d(bin)(w10000)
w310000 <- (3000*d(binm3)(w10000-3000)*0.015)/d(bin)(w10000)

SL110000 <- 1.5*sum(f(x1,w110000)*d(bin)(k))
SL210000 <- 1.5*sum(f(x2,w210000)*d(bin)(k))
SL310000 <- 1.5*sum(f(x3,w310000)*d(bin)(k))

purepremium10000 <- c(w110000+SL110000,
w210000+SL210000,w310000+SL310000)

CB110000 <- sum(f(w110000,x1)*d(bin)(k))
CB210000 <- sum(f(w210000,x2)*d(bin)(k))
CB310000 <- sum(f(w310000,x3)*d(bin)(k))

#Second method

w <- seq(low,up,10)

stoploss2 <- c()

for (i in 1:(up/10-low/10+1)) {
  stoploss2[i] <- sum(f(w,(i-1)*10+low)*d(bin)(w))
}

Totalprem10000 <- w + 1.5*stoploss2

plot(w,Totalprem10000, type="l",

```

```

main="Total_premium_according_w_for_n=10000", ylab="Total_Premium")

tab10000 <- data.frame(w, Totalprem10000)

pi10000 <- 40*n*0.25 + 50*n*0.5 + 70*n*0.25

w10000 <- tab10000[which(abs(Totalprem10000-pi10000)==min(
  abs(Totalprem10000-pi10000))),1]
#Selecting the w so that it correspond to the pure premium
w10000 <- round(w10000,-3)
#Can only be in thousand otherwise the probability will be 0

p(bin)(w10000)

w110000 <- (1000*d(binm1)(w10000-1000)*0.03)/d(bin)(w10000)
w210000 <- (2000*d(binm2)(w10000-2000)*0.025)/d(bin)(w10000)
w310000 <- (3000*d(binm3)(w10000-3000)*0.015)/d(bin)(w10000)

SL110000 <- 1.5*sum(f(x1,w110000)*d(bin)(k))
SL210000 <- 1.5*sum(f(x2,w210000)*d(bin)(k))
SL310000 <- 1.5*sum(f(x3,w310000)*d(bin)(k))

P2P_1 <- 40 - SL110000
P2P_2 <- 50 - SL210000
P2P_3 <- 70 - SL310000

CB110000 <- sum(f(w110000,x1)*d(bin)(k))
CB210000 <- sum(f(w210000,x2)*d(bin)(k))
CB310000 <- sum(f(w310000,x3)*d(bin)(k))

#Variance E[(Xi/Sn)] en fonction de n
#x1
var1 <- c()
for (n in seq(1000,30000,by=1000)) {
  bin1 <- 1000*Binom(n*0.25,0.03)
  bin2 <- 2000*Binom(n*0.5,0.025)
  bin3 <- 3000*Binom(n*0.25,0.015)

  bin <- bin1 + bin2 + bin3
  #Removing extreme values

```

```

low <- q(bin)(eps)
up <- q(bin)(1-eps)
k <- seq(low, up, 1000)
bin1m <- 1000*Binom(round((n-1)*0.25, 0), 0.03)
binm1 <- bin1m + bin2 + bin3
esp1 <- c()
for (i in 1:((up-low)/1000+1)) {
  esp1[i] <- (1000*d(binm1)((i-1)*1000+low-1000)*0.03)
  /d(bin)((i-1)*1000+low)
}
var1[(n/1000)] <- sum(esp1^2 * d(bin)(k)) - (sum(esp1*d(bin)(k)))^2
}

```

#x2

```

var2 <- c()
for (n in seq(1000, 30000, by=1000)) {
  bin1 <- 1000*Binom(n*0.25, 0.03)
  bin2 <- 2000*Binom(n*0.5, 0.025)
  bin3 <- 3000*Binom(n*0.25, 0.015)

  bin <- bin1 + bin2 + bin3
  #Removing extreme values
  low <- q(bin)(eps)
  up <- q(bin)(1-eps)
  k <- seq(low, up, 1000)
  bin2m <- 2000*Binom(round((n-1)*0.5, 0), 0.025)
  binm2 <- bin1 + bin2m + bin3
  esp1 <- c()
  for (i in 1:((up-low)/1000+1)) {
    esp1[i] <- (2000*d(binm2)((i-1)*1000+low-2000)*0.025)
    /d(bin)((i-1)*1000+low)
  }
  var2[(n/1000)] <- sum(esp1^2 * d(bin)(k)) - (sum(esp1*d(bin)(k)))^2
}

```

#x3

```

var3 <- c()
for (n in seq(1000, 30000, by=1000)) {
  bin1 <- 1000*Binom(n*0.25, 0.03)
  bin2 <- 2000*Binom(n*0.5, 0.025)

```

```

bin3 <- 3000*Binom(n*0.25,0.015)

bin <- bin1 + bin2 + bin3
#Removing extreme values
low <- q(bin)(eps)
up <- q(bin)(1-eps)
k <- seq(low,up,1000)
bin3m <- 3000*Binom(round((n-1)*0.25,0),0.015)
binm3 <- bin1 + bin2 + bin3m
esp1 <- c()
for (i in 1:((up-low)/1000+1)) {
  esp1[i] <- (3000*d(binm3)((i-1)*1000+low-3000)*0.015)
  /d(bin)((i-1)*1000+low)
}
var3[(n/1000)] <- sum(esp1^2 * d(bin)(k)) - (sum(esp1*d(bin)(k)))^2
}

k <- seq(1000,30000,by=1000)

plot(k, var1, type="l", main="var(E[(Xi|Sn)])",
      xlab="Number of insured", ylab="var(E[(Xi|Sn)])",
      col = "black", ylim=c(0,200))
lines(k, var2, col = "blue", lty = 1)
lines(k, var3, col = "red", lty = 1)
legend("topright", legend = c("p=1000", "p=2000", "p=3000"),
col = c("black", "blue", "red"), lty=1)

plot(k, sqrt(var1), type="l", main="Standard deviation of E[(Xi|Sn)]",
      xlab="Number of insured", ylab="Standard deviation",
      col = "black", ylim=c(0,40))
lines(k, sqrt(var2), col = "blue", lty = 1)
lines(k, sqrt(var3), col = "red", lty = 1)
legend("topright", legend = c("p=1000", "p=2000", "p=3000"),
col = c("black", "blue", "red"), lty=1)

#Quantiles plots
p25_1 <- c()
p50_1 <- c()
p75_1 <- c()
p25_2 <- c()

```

```

p50_2 <- c()
p75_2 <- c()
p25_3 <- c()
p50_3 <- c()
p75_3 <- c()
for (n in seq(1000, 30000, by=1000)) {
  bin1 <- 1000*Binom(n*0.25, 0.03)
  bin2 <- 2000*Binom(n*0.5, 0.025)
  bin3 <- 3000*Binom(n*0.25, 0.015)
  bin <- bin1 + bin2 + bin3
  q25 <- q(bin)(0.25)
  q50 <- q(bin)(0.5)
  q75 <- q(bin)(0.75)
  bin1m <- 1000*Binom(floor((n-1)*0.25), 0.03)
  binm1 <- bin1m + bin2 + bin3
  p25_1 <- c(p25_1, (1000*d(binm1)(q25-1000)*0.03)/d(bin)(q25))
  p50_1 <- c(p50_1, (1000*d(binm1)(q50-1000)*0.03)/d(bin)(q50))
  p75_1 <- c(p75_1, (1000*d(binm1)(q75-1000)*0.03)/d(bin)(q75))
  bin2m <- 2000*Binom(floor((n-1)*0.5), 0.025)
  binm2 <- bin1 + bin2m + bin3
  p25_2 <- c(p25_2, (2000*d(binm2)(q25-2000)*0.025)/d(bin)(q25))
  p50_2 <- c(p50_2, (2000*d(binm2)(q50-2000)*0.025)/d(bin)(q50))
  p75_2 <- c(p75_2, (2000*d(binm2)(q75-2000)*0.025)/d(bin)(q75))
  bin3m <- 3000*Binom(floor((n-1)*0.25), 0.015)
  binm3 <- bin1 + bin2 + bin3m
  p25_3 <- c(p25_3, (3000*d(binm3)(q25-3000)*0.015)/d(bin)(q25))
  p50_3 <- c(p50_3, (3000*d(binm3)(q50-3000)*0.015)/d(bin)(q50))
  p75_3 <- c(p75_3, (3000*d(binm3)(q75-3000)*0.015)/d(bin)(q75))
}

n <- seq(1000, 30000, by = 1000)

quantiles_data_1 <- data.frame(n, p25_1, p50_1, p75_1)

ggplot(quantiles_data_1, aes(x = n)) +
  geom_ribbon(aes(ymin = p25_1, ymax = p75_1), fill = "blue",
    alpha = 0.3) +
  geom_line(aes(y = p50_1), color = "blue") +
  labs(title = "Evolution of quantiles as a function of
the number of policyholders for p=1000",

```

```

      x = "Number_of_insured",
      y = "Participation") +
theme_minimal()

quantiles_data_2 <- data.frame(n, p25_2, p50_2, p75_2)

ggplot(quantiles_data_2, aes(x = n)) +
  geom_ribbon(aes(ymin = p25_2, ymax = p75_2), fill = "blue",
    alpha = 0.3) +
  geom_line(aes(y = p50_2), color = "blue") +
  labs(title = "Evolution_of_quantiles_as_a_function_of
the_number_of_policyholders_for_p=2000",
    x = "Number_of_insured",
    y = "Participation") +
  theme_minimal()

quantiles_data_3 <- data.frame(n, p25_3, p50_3, p75_3)

ggplot(quantiles_data_3, aes(x = n)) +
  geom_ribbon(aes(ymin = p25_3, ymax = p75_3), fill = "blue",
    alpha = 0.3) +
  geom_line(aes(y = p50_3), color = "blue") +
  labs(title = "Evolution_of_quantiles_as_a_function_of
the_number_of_policyholders_for_p=3000",
    x = "Number_of_insured",
    y = "Participation") +
  theme_minimal()

#Illustration

#Uniform distribution
#10000 insured
#Regular bikes

n=10000 #Number of insured
lmax=n/10*sum(seq(500,5000,500)) #Maximum cost

k<-seq(500,lmax,by=500) #All cost values

```

```

bin1 <- 500*Binom(n/10,0.0306)
bin2 <- 1000*Binom(n/10,0.0315)
bin3 <- 1500*Binom(n/10,0.024)
bin4 <- 2000*Binom(n/10,0.0218)
bin5 <- 2500*Binom(n/10,0.0204)
bin6 <- 3000*Binom(n/10,0.0195)
bin7 <- 3500*Binom(n/10,0.0189)
bin8 <- 4000*Binom(n/10,0.018)
bin9 <- 4500*Binom(n/10,0.018)
bin10 <- 5000*Binom(n/10,0.018)

bin <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10 #Portfolio distribution

plot(k,p(bin)(k),type="s",
main="Distribution of function for n=10000 (uniform)",
xlab="Cost",ylab="Cumulative probability", col = "blue")

#Removing extrem values
eps <- 10−6/2
low <- q(bin)(eps)
up <- q(bin)(1−eps)

k <- seq(low,up,500)

#Premium calculation

bin1m <- 500*Binom(floor((n−1)/10),0.0306)
binm1 <- bin1m + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10

bin2m <- 1000*Binom(floor((n−1)/10),0.0315)
binm2 <- bin1 + bin2m + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10

bin3m <- 1500*Binom(floor((n−1)/10),0.024)
binm3 <- bin1 + bin2 + bin3m + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10

```

```

bin4m <- 2000*Binom(floor((n-1)/10),0.0218)
binm4 <- bin1 + bin2 + bin3 + bin4m + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10

bin5m <- 2500*Binom(floor((n-1)/10),0.0204)
binm5 <- bin1 + bin2 + bin3 + bin4 + bin5m + bin6
+ bin7 + bin8 + bin9 + bin10

bin6m <- 3000*Binom(floor((n-1)/10),0.0195)
binm6 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6m
+ bin7 + bin8 + bin9 + bin10

bin7m <- 3500*Binom(floor((n-1)/10),0.0189)
binm7 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7m + bin8 + bin9 + bin10

bin8m <- 4000*Binom(floor((n-1)/10),0.018)
binm8 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8m + bin9 + bin10

bin9m <- 4500*Binom(floor((n-1)/10),0.018)
binm9 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9m + bin10

bin10m <- 5000*Binom(floor((n-1)/10),0.018)
binm10 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10m

x1 <- c()

for (i in 1:((up-low)/500+1)) {
  x1[i] <- (500*d(binm1)((i-1)*500+low-500)*0.0306)
  /d(bin)((i-1)*500+low)
}

x2 <- c()

for (i in 1:((up-low)/500+1)) {
  x2[i] <- (1000*d(binm2)((i-1)*500+low-1000)*0.0315)
  /d(bin)((i-1)*500+low)
}

```

```

}

x3 <- c()

for (i in 1:((up-low)/500+1)) {
  x3[i] <- (1500*d(binm3)((i-1)*500+low-1500)*0.024)
  /d(bin)((i-1)*500+low)
}

x4 <- c()

for (i in 1:((up-low)/500+1)) {
  x4[i] <- (2000*d(binm4)((i-1)*500+low-2000)*0.0218)
  /d(bin)((i-1)*500+low)
}

x5 <- c()

for (i in 1:((up-low)/500+1)) {
  x5[i] <- (2500*d(binm5)((i-1)*500+low-2500)*0.0204)
  /d(bin)((i-1)*500+low)
}

x6 <- c()

for (i in 1:((up-low)/500+1)) {
  x6[i] <- (3000*d(binm6)((i-1)*500+low-3000)*0.0195)
  /d(bin)((i-1)*500+low)
}

x7 <- c()

for (i in 1:((up-low)/500+1)) {
  x7[i] <- (3500*d(binm7)((i-1)*500+low-3500)*0.0189)
  /d(bin)((i-1)*500+low)
}

x8 <- c()

for (i in 1:((up-low)/500+1)) {

```

```

x8[i] <- (4000*d(binm8)((i-1)*500+low-4000)*0.018)
/d(bin)((i-1)*500+low)
}

x9 <- c()

for (i in 1:((up-low)/500+1)) {
  x9[i] <- (4500*d(binm9)((i-1)*500+low-4500)*0.018)
/d(bin)((i-1)*500+low)
}

x10 <- c()

for (i in 1:((up-low)/500+1)) {
  x10[i] <- (5000*d(binm10)((i-1)*500+low-5000)*0.018)
/d(bin)((i-1)*500+low)
}

#Premium calculation
w <- seq(low, up, 10)

stoploss2 <- c()

for (i in 1:(up/10-low/10+1)) {
  stoploss2[i] <- sum(f(w, (i-1)*10+low)*d(bin)(w))
}

Totalprem10000 <- w + 1.5*stoploss2

plot(w, Totalprem10000,
type="l", main="Total premium according to the
retention level for n=10000 (uniform)",
ylab="Total Premium", xlab="Retention level")

tab10000 <- data.frame(w, Totalprem10000)

w10000_1 <- tab10000[Totalprem10000==min(Totalprem10000),1]
#Selecting the w that minimize the premium value

w10000_2 <- q(bin)(0.7)

```

```

w10000_3 <- q(bin)(0.95)

purepremium10000_4 <- 1.2*c(500*0.0306,1000*0.0315 , 1500*0.024 ,
                           2000*0.0218 , 2500*0.0204 ,
                           3000*0.0195 , 3500*0.0189 ,
                           4000*0.018 , 4500*0.018 , 5000*0.018)

pi10000 <- 1000*sum(purepremium10000_4) #Total premium

w10000_4 <- tab10000[which(abs(Totalprem10000-
pi10000)==min(abs(Totalprem10000-pi10000))),1]
#Selecting the w so that the premium is equal to the pure premium

w10000_4 <- round(w10000_4,-3)

w110000_1 <- (500*d(binm1)(w10000_1-500)*0.0306)/d(bin)(w10000_1)
w210000_1 <- (1000*d(binm2)(w10000_1-1000)*0.0315)/d(bin)(w10000_1)
w310000_1 <- (1500*d(binm3)(w10000_1-1500)*0.024)/d(bin)(w10000_1)
w410000_1 <- (2000*d(binm4)(w10000_1-2000)*0.0218)/d(bin)(w10000_1)
w510000_1 <- (2500*d(binm5)(w10000_1-2500)*0.0204)/d(bin)(w10000_1)
w610000_1 <- (3000*d(binm6)(w10000_1-3000)*0.0195)/d(bin)(w10000_1)
w710000_1 <- (3500*d(binm7)(w10000_1-3500)*0.0189)/d(bin)(w10000_1)
w810000_1 <- (4000*d(binm8)(w10000_1-4000)*0.018)/d(bin)(w10000_1)
w910000_1 <- (4500*d(binm9)(w10000_1-4500)*0.018)/d(bin)(w10000_1)
w1010000_1 <- (5000*d(binm10)(w10000_1-5000)*0.018)/d(bin)(w10000_1)

SL110000_1 <- 1.5*sum(f(x1,w110000_1)*d(bin)(k))
SL210000_1 <- 1.5*sum(f(x2,w210000_1)*d(bin)(k))
SL310000_1 <- 1.5*sum(f(x3,w310000_1)*d(bin)(k))
SL410000_1 <- 1.5*sum(f(x4,w410000_1)*d(bin)(k))
SL510000_1 <- 1.5*sum(f(x5,w510000_1)*d(bin)(k))
SL610000_1 <- 1.5*sum(f(x6,w610000_1)*d(bin)(k))
SL710000_1 <- 1.5*sum(f(x7,w710000_1)*d(bin)(k))
SL810000_1 <- 1.5*sum(f(x8,w810000_1)*d(bin)(k))
SL910000_1 <- 1.5*sum(f(x9,w910000_1)*d(bin)(k))
SL1010000_1 <- 1.5*sum(f(x10,w1010000_1)*d(bin)(k))

purepremium10000_1 <- c(w110000_1+SL110000_1,w210000_1+SL210000_1,
                        w310000_1+SL310000_1,w410000_1+SL410000_1,

```

```

w510000_1+SL510000_1,w610000_1+SL610000_1,
w710000_1+SL710000_1,w810000_1+SL810000_1,
w910000_1+SL910000_1,w1010000_1+SL1010000_1)

```

#Expected cash-back

```

CB110000_1 <- sum(f(w110000_1,x1)*d(bin)(k))
CB210000_1 <- sum(f(w210000_1,x2)*d(bin)(k))
CB310000_1 <- sum(f(w310000_1,x3)*d(bin)(k))
CB410000_1 <- sum(f(w410000_1,x4)*d(bin)(k))
CB510000_1 <- sum(f(w510000_1,x5)*d(bin)(k))
CB610000_1 <- sum(f(w610000_1,x6)*d(bin)(k))
CB710000_1 <- sum(f(w710000_1,x7)*d(bin)(k))
CB810000_1 <- sum(f(w810000_1,x8)*d(bin)(k))
CB910000_1 <- sum(f(w910000_1,x9)*d(bin)(k))
CB1010000_1 <- sum(f(w1010000_1,x10)*d(bin)(k))

```

```

w110000_2 <- (500*d(binm1)(w10000_2-500)*0.0306)/d(bin)(w10000_2)
w210000_2 <- (1000*d(binm2)(w10000_2-1000)*0.0315)/d(bin)(w10000_2)
w310000_2 <- (1500*d(binm3)(w10000_2-1500)*0.024)/d(bin)(w10000_2)
w410000_2 <- (2000*d(binm4)(w10000_2-2000)*0.0218)/d(bin)(w10000_2)
w510000_2 <- (2500*d(binm5)(w10000_2-2500)*0.0204)/d(bin)(w10000_2)
w610000_2 <- (3000*d(binm6)(w10000_2-3000)*0.0195)/d(bin)(w10000_2)
w710000_2 <- (3500*d(binm7)(w10000_2-3500)*0.0189)/d(bin)(w10000_2)
w810000_2 <- (4000*d(binm8)(w10000_2-4000)*0.018)/d(bin)(w10000_2)
w910000_2 <- (4500*d(binm9)(w10000_2-4500)*0.018)/d(bin)(w10000_2)
w1010000_2 <- (5000*d(binm10)(w10000_2-5000)*0.018)/d(bin)(w10000_2)

```

```

SL110000_2 <- 1.5*sum(f(x1,w110000_2)*d(bin)(k))
SL210000_2 <- 1.5*sum(f(x2,w210000_2)*d(bin)(k))
SL310000_2 <- 1.5*sum(f(x3,w310000_2)*d(bin)(k))
SL410000_2 <- 1.5*sum(f(x4,w410000_2)*d(bin)(k))
SL510000_2 <- 1.5*sum(f(x5,w510000_2)*d(bin)(k))
SL610000_2 <- 1.5*sum(f(x6,w610000_2)*d(bin)(k))
SL710000_2 <- 1.5*sum(f(x7,w710000_2)*d(bin)(k))
SL810000_2 <- 1.5*sum(f(x8,w810000_2)*d(bin)(k))
SL910000_2 <- 1.5*sum(f(x9,w910000_2)*d(bin)(k))
SL1010000_2 <- 1.5*sum(f(x10,w1010000_2)*d(bin)(k))

```

```

purepremium10000_2 <- c(w110000_2+SL110000_2,w210000_2+SL210000_2,

```

```

w310000_2+SL310000_2,w410000_2+SL410000_2,
w510000_2+SL510000_2,w610000_2+SL610000_2,
w710000_2+SL710000_2,w810000_2+SL810000_2,
w910000_2+SL910000_2,w1010000_2+SL1010000_2)

```

#Expected cash-back

```

CB110000_2 <- sum(f(w110000_2,x1)*d(bin)(k))
CB210000_2 <- sum(f(w210000_2,x2)*d(bin)(k))
CB310000_2 <- sum(f(w310000_2,x3)*d(bin)(k))
CB410000_2 <- sum(f(w410000_2,x4)*d(bin)(k))
CB510000_2 <- sum(f(w510000_2,x5)*d(bin)(k))
CB610000_2 <- sum(f(w610000_2,x6)*d(bin)(k))
CB710000_2 <- sum(f(w710000_2,x7)*d(bin)(k))
CB810000_2 <- sum(f(w810000_2,x8)*d(bin)(k))
CB910000_2 <- sum(f(w910000_2,x9)*d(bin)(k))
CB1010000_2 <- sum(f(w1010000_2,x10)*d(bin)(k))

w110000_3 <- (500*d(binm1)(w10000_3-500)*0.0306)/d(bin)(w10000_3)
w210000_3 <- (1000*d(binm2)(w10000_3-1000)*0.0315)/d(bin)(w10000_3)
w310000_3 <- (1500*d(binm3)(w10000_3-1500)*0.024)/d(bin)(w10000_3)
w410000_3 <- (2000*d(binm4)(w10000_3-2000)*0.0218)/d(bin)(w10000_3)
w510000_3 <- (2500*d(binm5)(w10000_3-2500)*0.0204)/d(bin)(w10000_3)
w610000_3 <- (3000*d(binm6)(w10000_3-3000)*0.0195)/d(bin)(w10000_3)
w710000_3 <- (3500*d(binm7)(w10000_3-3500)*0.0189)/d(bin)(w10000_3)
w810000_3 <- (4000*d(binm8)(w10000_3-4000)*0.018)/d(bin)(w10000_3)
w910000_3 <- (4500*d(binm9)(w10000_3-4500)*0.018)/d(bin)(w10000_3)
w1010000_3 <- (5000*d(binm10)(w10000_3-5000)*0.018)/d(bin)(w10000_3)

SL110000_3 <- 1.5*sum(f(x1,w110000_3)*d(bin)(k))
SL210000_3 <- 1.5*sum(f(x2,w210000_3)*d(bin)(k))
SL310000_3 <- 1.5*sum(f(x3,w310000_3)*d(bin)(k))
SL410000_3 <- 1.5*sum(f(x4,w410000_3)*d(bin)(k))
SL510000_3 <- 1.5*sum(f(x5,w510000_3)*d(bin)(k))
SL610000_3 <- 1.5*sum(f(x6,w610000_3)*d(bin)(k))
SL710000_3 <- 1.5*sum(f(x7,w710000_3)*d(bin)(k))
SL810000_3 <- 1.5*sum(f(x8,w810000_3)*d(bin)(k))
SL910000_3 <- 1.5*sum(f(x9,w910000_3)*d(bin)(k))
SL1010000_3 <- 1.5*sum(f(x10,w1010000_3)*d(bin)(k))

```

```

purepremium10000_3 <- c(w110000_3+SL110000_3,w210000_3+SL210000_3,
                        w310000_3+SL310000_3,w410000_3+SL410000_3,
                        w510000_3+SL510000_3,w610000_3+SL610000_3,
                        w710000_3+SL710000_3,w810000_3+SL810000_3,
                        w910000_3+SL910000_3,w1010000_3+SL1010000_3)

```

#Expected cash-back

```

CB110000_3 <- sum(f(w110000_3,x1)*d(bin)(k))
CB210000_3 <- sum(f(w210000_3,x2)*d(bin)(k))
CB310000_3 <- sum(f(w310000_3,x3)*d(bin)(k))
CB410000_3 <- sum(f(w410000_3,x4)*d(bin)(k))
CB510000_3 <- sum(f(w510000_3,x5)*d(bin)(k))
CB610000_3 <- sum(f(w610000_3,x6)*d(bin)(k))
CB710000_3 <- sum(f(w710000_3,x7)*d(bin)(k))
CB810000_3 <- sum(f(w810000_3,x8)*d(bin)(k))
CB910000_3 <- sum(f(w910000_3,x9)*d(bin)(k))
CB1010000_3 <- sum(f(w1010000_3,x10)*d(bin)(k))

```

```

w110000_4 <- (500*d(binm1)(w10000_4-500)*0.0306)/d(bin)(w10000_4)
w210000_4 <- (1000*d(binm2)(w10000_4-1000)*0.0315)/d(bin)(w10000_4)
w310000_4 <- (1500*d(binm3)(w10000_4-1500)*0.024)/d(bin)(w10000_4)
w410000_4 <- (2000*d(binm4)(w10000_4-2000)*0.0218)/d(bin)(w10000_4)
w510000_4 <- (2500*d(binm5)(w10000_4-2500)*0.0204)/d(bin)(w10000_4)
w610000_4 <- (3000*d(binm6)(w10000_4-3000)*0.0195)/d(bin)(w10000_4)
w710000_4 <- (3500*d(binm7)(w10000_4-3500)*0.0189)/d(bin)(w10000_4)
w810000_4 <- (4000*d(binm8)(w10000_4-4000)*0.018)/d(bin)(w10000_4)
w910000_4 <- (4500*d(binm9)(w10000_4-4500)*0.018)/d(bin)(w10000_4)
w1010000_4 <- (5000*d(binm10)(w10000_4-5000)*0.018)/d(bin)(w10000_4)

```

```

SL110000_4 <- 1.5*sum(f(x1,w110000_4)*d(bin)(k))
SL210000_4 <- 1.5*sum(f(x2,w210000_4)*d(bin)(k))
SL310000_4 <- 1.5*sum(f(x3,w310000_4)*d(bin)(k))
SL410000_4 <- 1.5*sum(f(x4,w410000_4)*d(bin)(k))
SL510000_4 <- 1.5*sum(f(x5,w510000_4)*d(bin)(k))
SL610000_4 <- 1.5*sum(f(x6,w610000_4)*d(bin)(k))
SL710000_4 <- 1.5*sum(f(x7,w710000_4)*d(bin)(k))
SL810000_4 <- 1.5*sum(f(x8,w810000_4)*d(bin)(k))
SL910000_4 <- 1.5*sum(f(x9,w910000_4)*d(bin)(k))
SL1010000_4 <- 1.5*sum(f(x10,w1010000_4)*d(bin)(k))

```

```

P2P110000_4 <- purepremium10000_4[1] - SL110000_4
P2P210000_4 <- purepremium10000_4[2] - SL210000_4
P2P310000_4 <- purepremium10000_4[3] - SL310000_4
P2P410000_4 <- purepremium10000_4[4] - SL410000_4
P2P510000_4 <- purepremium10000_4[5] - SL510000_4
P2P610000_4 <- purepremium10000_4[6] - SL610000_4
P2P710000_4 <- purepremium10000_4[7] - SL710000_4
P2P810000_4 <- purepremium10000_4[8] - SL810000_4
P2P910000_4 <- purepremium10000_4[9] - SL910000_4
P2P1010000_4 <- purepremium10000_4[10] - SL1010000_4

```

#Expected cash-back

```

CB110000_4 <- sum(f(w110000_4,x1)*d(bin)(k))
CB210000_4 <- sum(f(w210000_4,x2)*d(bin)(k))
CB310000_4 <- sum(f(w310000_4,x3)*d(bin)(k))
CB410000_4 <- sum(f(w410000_4,x4)*d(bin)(k))
CB510000_4 <- sum(f(w510000_4,x5)*d(bin)(k))
CB610000_4 <- sum(f(w610000_4,x6)*d(bin)(k))
CB710000_4 <- sum(f(w710000_4,x7)*d(bin)(k))
CB810000_4 <- sum(f(w810000_4,x8)*d(bin)(k))
CB910000_4 <- sum(f(w910000_4,x9)*d(bin)(k))
CB1010000_4 <- sum(f(w1010000_4,x10)*d(bin)(k))

```

#Binomial (0.5) distribution

#10000 insured

#Regular bikes

n=10000 #Number of insured

*lmax=sum(seq(500,5000,500)*round(dbinom(1:10,10,0.5)*n)) #Maximum cost*

k<-seq(500,lmax,by=500) #All cost values

```

bin1 <- 500*Binom(round(dbinom(1,10,0.5)*n),0.0306)
bin2 <- 1000*Binom(round(dbinom(2,10,0.5)*n),0.0315)
bin3 <- 1500*Binom(round(dbinom(3,10,0.5)*n),0.024)
bin4 <- 2000*Binom(round(dbinom(4,10,0.5)*n),0.0218)
bin5 <- 2500*Binom(round(dbinom(5,10,0.5)*n),0.0204)

```

```

bin6 <- 3000*Binom(round(dbinom(6,10,0.5)*n),0.0195)
bin7 <- 3500*Binom(round(dbinom(7,10,0.5)*n),0.0189)
bin8 <- 4000*Binom(round(dbinom(8,10,0.5)*n),0.018)
bin9 <- 4500*Binom(round(dbinom(9,10,0.5)*n),0.018)
bin10 <- 5000*Binom(round(dbinom(10,10,0.5)*n),0.018)

bin <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10 #Portfolio distribution

plot(k,p(bin)(k),type="s",main="Distribution function
#####for n=10000(binomial)",xlab="Cost",
      ylab="Cumulative probability", col = "blue")

#Removing extrem values
eps <- 10^(-6)/2
low <- q(bin)(eps)
up <- q(bin)(1-eps)

k <- seq(low,up,500)

#Premium calculation

bin1m <- 500*Binom(round(dbinom(1,10,0.5)*n)-1,0.0306)
binm1 <- bin1m + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10

bin2m <- 1000*Binom(round(dbinom(2,10,0.5)*n)-1,0.0315)
binm2 <- bin1 + bin2m + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10

bin3m <- 1500*Binom(round(dbinom(3,10,0.5)*n)-1,0.024)
binm3 <- bin1 + bin2 + bin3m + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10

bin4m <- 2000*Binom(round(dbinom(4,10,0.5)*n)-1,0.0218)
binm4 <- bin1 + bin2 + bin3 + bin4m + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10

bin5m <- 2500*Binom(round(dbinom(5,10,0.5)*n)-1,0.0204)
binm5 <- bin1 + bin2 + bin3 + bin4 + bin5m + bin6

```

```

+ bin7 + bin8 + bin9 + bin10

bin6m <- 3000*Binom(round(dbinom(6,10,0.5)*n)-1,0.0195)
binm6 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6m
+ bin7 + bin8 + bin9 + bin10

bin7m <- 3500*Binom(round(dbinom(7,10,0.5)*n)-1,0.0189)
binm7 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7m + bin8 + bin9 + bin10

bin8m <- 4000*Binom(round(dbinom(8,10,0.5)*n)-1,0.018)
binm8 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8m + bin9 + bin10

bin9m <- 4500*Binom(round(dbinom(9,10,0.5)*n)-1,0.018)
binm9 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9m + bin10

bin10m <- 5000*Binom(round(dbinom(10,10,0.5)*n)-1,0.018)
binm10 <- bin1 + bin2 + bin3 + bin4 + bin5 + bin6
+ bin7 + bin8 + bin9 + bin10m

x1 <- c()

for (i in 1:((up-low)/500+1)) {
  x1[i] <- (500*d(binm1)((i-1)*500+low-500)*0.0306)
  /d(bin)((i-1)*500+low)
}

x2 <- c()

for (i in 1:((up-low)/500+1)) {
  x2[i] <- (1000*d(binm2)((i-1)*500+low-1000)*0.0315)
  /d(bin)((i-1)*500+low)
}

x3 <- c()

for (i in 1:((up-low)/500+1)) {
  x3[i] <- (1500*d(binm3)((i-1)*500+low-1500)*0.024)

```

```

    /d(bin)((i-1)*500+low)
  }

x4 <- c()

for (i in 1:((up-low)/500+1)) {
  x4[i] <- (2000*d(binm4)((i-1)*500+low-2000)*0.0218)
    /d(bin)((i-1)*500+low)
}

x5 <- c()

for (i in 1:((up-low)/500+1)) {
  x5[i] <- (2500*d(binm5)((i-1)*500+low-2500)*0.0204)
    /d(bin)((i-1)*500+low)
}

x6 <- c()

for (i in 1:((up-low)/500+1)) {
  x6[i] <- (3000*d(binm6)((i-1)*500+low-3000)*0.0195)
    /d(bin)((i-1)*500+low)
}

x7 <- c()

for (i in 1:((up-low)/500+1)) {
  x7[i] <- (3500*d(binm7)((i-1)*500+low-3500)*0.0189)
    /d(bin)((i-1)*500+low)
}

x8 <- c()

for (i in 1:((up-low)/500+1)) {
  x8[i] <- (4000*d(binm8)((i-1)*500+low-4000)*0.018)
    /d(bin)((i-1)*500+low)
}

x9 <- c()

```

```

for (i in 1:((up-low)/500+1)) {
  x9[i] <- (4500*d(binm9)((i-1)*500+low-4500)*0.018)
  /d(bin)((i-1)*500+low)
}

x10 <- c()

for (i in 1:((up-low)/500+1)) {
  x10[i] <- (5000*d(binm10)((i-1)*500+low-5000)*0.018)
  /d(bin)((i-1)*500+low)
}

#Premium calculation
w <- seq(low, up, 10)

stoploss2 <- c()

for (i in 1:(up/10-low/10+1)) {
  stoploss2[i] <- sum(f(w, (i-1)*10+low)*d(bin)(w))
}

Totalprem10000 <- w + 1.5*stoploss2

plot(w, Totalprem10000, type="l", main="Total_premium_according
the_retention_level_for_n=10000_(binomial)", ylab="Total_Premium",
xlab="Retention_level")

tab10000 <- data.frame(w, Totalprem10000)

w10000_1 <- tab10000[Totalprem10000==min(Totalprem10000), 1]
#Selecting the w that minimize the premium value

w10000_2 <- q(bin)(0.7)

w10000_3 <- q(bin)(0.95)

purepremium10000_4 <- 1.2*c(500*0.0306, 1000*0.0315, 1500*0.024,
2000*0.0218, 2500*0.0204, 3000*0.0195,
3500*0.0189, 4000*0.018, 4500*0.018,
5000*0.018)

```

```

pi10000 <- sum(1.2*c(98*500*0.0306,439*1000*0.0315 , 1172*1500*0.024 ,
2051*2000*0.0218 , 2461*2500*0.0204 , 2051*3000*0.0195 ,
1172*3500*0.0189 , 439*4000*0.018 , 98*4500*0.018 ,
10*5000*0.018)) #Total premium

w10000_4 <- tab10000[which(abs(Totalprem10000-pi10000)==min(
abs(Totalprem10000-pi10000))),1]
#Selecting the w so that the premium is equal to the pure premium

w10000_4 <- 612500

w110000_1 <- (500*d(binm1)(w10000_1-500)*0.0306)/d(bin)(w10000_1)
w210000_1 <- (1000*d(binm2)(w10000_1-1000)*0.0315)/d(bin)(w10000_1)
w310000_1 <- (1500*d(binm3)(w10000_1-1500)*0.024)/d(bin)(w10000_1)
w410000_1 <- (2000*d(binm4)(w10000_1-2000)*0.0218)/d(bin)(w10000_1)
w510000_1 <- (2500*d(binm5)(w10000_1-2500)*0.0204)/d(bin)(w10000_1)
w610000_1 <- (3000*d(binm6)(w10000_1-3000)*0.0195)/d(bin)(w10000_1)
w710000_1 <- (3500*d(binm7)(w10000_1-3500)*0.0189)/d(bin)(w10000_1)
w810000_1 <- (4000*d(binm8)(w10000_1-4000)*0.018)/d(bin)(w10000_1)
w910000_1 <- (4500*d(binm9)(w10000_1-4500)*0.018)/d(bin)(w10000_1)
w1010000_1 <- (5000*d(binm10)(w10000_1-5000)*0.018)/d(bin)(w10000_1)

SL110000_1 <- 1.5*sum(f(x1,w110000_1)*d(bin)(k))
SL210000_1 <- 1.5*sum(f(x2,w210000_1)*d(bin)(k))
SL310000_1 <- 1.5*sum(f(x3,w310000_1)*d(bin)(k))
SL410000_1 <- 1.5*sum(f(x4,w410000_1)*d(bin)(k))
SL510000_1 <- 1.5*sum(f(x5,w510000_1)*d(bin)(k))
SL610000_1 <- 1.5*sum(f(x6,w610000_1)*d(bin)(k))
SL710000_1 <- 1.5*sum(f(x7,w710000_1)*d(bin)(k))
SL810000_1 <- 1.5*sum(f(x8,w810000_1)*d(bin)(k))
SL910000_1 <- 1.5*sum(f(x9,w910000_1)*d(bin)(k))
SL1010000_1 <- 1.5*sum(f(x10,w1010000_1)*d(bin)(k))

purepremium10000_1 <- c(w110000_1+SL110000_1,w210000_1+SL210000_1,
w310000_1+SL310000_1,w410000_1+SL410000_1,
w510000_1+SL510000_1,w610000_1+SL610000_1,
w710000_1+SL710000_1,w810000_1+SL810000_1,
w910000_1+SL910000_1,w1010000_1+SL1010000_1)

```

#Expected cash-back

```
CB110000_1 <- sum(f(w110000_1,x1)*d(bin)(k))
CB210000_1 <- sum(f(w210000_1,x2)*d(bin)(k))
CB310000_1 <- sum(f(w310000_1,x3)*d(bin)(k))
CB410000_1 <- sum(f(w410000_1,x4)*d(bin)(k))
CB510000_1 <- sum(f(w510000_1,x5)*d(bin)(k))
CB610000_1 <- sum(f(w610000_1,x6)*d(bin)(k))
CB710000_1 <- sum(f(w710000_1,x7)*d(bin)(k))
CB810000_1 <- sum(f(w810000_1,x8)*d(bin)(k))
CB910000_1 <- sum(f(w910000_1,x9)*d(bin)(k))
CB1010000_1 <- sum(f(w1010000_1,x10)*d(bin)(k))
```

```
w110000_2 <- (500*d(binm1)(w10000_2-500)*0.0306)/d(bin)(w10000_2)
w210000_2 <- (1000*d(binm2)(w10000_2-1000)*0.0315)/d(bin)(w10000_2)
w310000_2 <- (1500*d(binm3)(w10000_2-1500)*0.024)/d(bin)(w10000_2)
w410000_2 <- (2000*d(binm4)(w10000_2-2000)*0.0218)/d(bin)(w10000_2)
w510000_2 <- (2500*d(binm5)(w10000_2-2500)*0.0204)/d(bin)(w10000_2)
w610000_2 <- (3000*d(binm6)(w10000_2-3000)*0.0195)/d(bin)(w10000_2)
w710000_2 <- (3500*d(binm7)(w10000_2-3500)*0.0189)/d(bin)(w10000_2)
w810000_2 <- (4000*d(binm8)(w10000_2-4000)*0.018)/d(bin)(w10000_2)
w910000_2 <- (4500*d(binm9)(w10000_2-4500)*0.018)/d(bin)(w10000_2)
w1010000_2 <- (5000*d(binm10)(w10000_2-5000)*0.018)/d(bin)(w10000_2)
```

```
SL110000_2 <- 1.5*sum(f(x1,w110000_2)*d(bin)(k))
SL210000_2 <- 1.5*sum(f(x2,w210000_2)*d(bin)(k))
SL310000_2 <- 1.5*sum(f(x3,w310000_2)*d(bin)(k))
SL410000_2 <- 1.5*sum(f(x4,w410000_2)*d(bin)(k))
SL510000_2 <- 1.5*sum(f(x5,w510000_2)*d(bin)(k))
SL610000_2 <- 1.5*sum(f(x6,w610000_2)*d(bin)(k))
SL710000_2 <- 1.5*sum(f(x7,w710000_2)*d(bin)(k))
SL810000_2 <- 1.5*sum(f(x8,w810000_2)*d(bin)(k))
SL910000_2 <- 1.5*sum(f(x9,w910000_2)*d(bin)(k))
SL1010000_2 <- 1.5*sum(f(x10,w1010000_2)*d(bin)(k))
```

```
purepremium10000_2 <- c(w110000_2+SL110000_2,w210000_2+SL210000_2,
                        w310000_2+SL310000_2,w410000_2+SL410000_2,
                        w510000_2+SL510000_2,w610000_2+SL610000_2,
                        w710000_2+SL710000_2,w810000_2+SL810000_2,
```

w910000_2+SL910000_2,w1010000_2+SL1010000_2)

#Expected cash-back

```
CB110000_2 <- sum(f(w110000_2,x1)*d(bin)(k))
CB210000_2 <- sum(f(w210000_2,x2)*d(bin)(k))
CB310000_2 <- sum(f(w310000_2,x3)*d(bin)(k))
CB410000_2 <- sum(f(w410000_2,x4)*d(bin)(k))
CB510000_2 <- sum(f(w510000_2,x5)*d(bin)(k))
CB610000_2 <- sum(f(w610000_2,x6)*d(bin)(k))
CB710000_2 <- sum(f(w710000_2,x7)*d(bin)(k))
CB810000_2 <- sum(f(w810000_2,x8)*d(bin)(k))
CB910000_2 <- sum(f(w910000_2,x9)*d(bin)(k))
CB1010000_2 <- sum(f(w1010000_2,x10)*d(bin)(k))

w110000_3 <- (500*d(binm1)(w10000_3-500)*0.0306)/d(bin)(w10000_3)
w210000_3 <- (1000*d(binm2)(w10000_3-1000)*0.0315)/d(bin)(w10000_3)
w310000_3 <- (1500*d(binm3)(w10000_3-1500)*0.024)/d(bin)(w10000_3)
w410000_3 <- (2000*d(binm4)(w10000_3-2000)*0.0218)/d(bin)(w10000_3)
w510000_3 <- (2500*d(binm5)(w10000_3-2500)*0.0204)/d(bin)(w10000_3)
w610000_3 <- (3000*d(binm6)(w10000_3-3000)*0.0195)/d(bin)(w10000_3)
w710000_3 <- (3500*d(binm7)(w10000_3-3500)*0.0189)/d(bin)(w10000_3)
w810000_3 <- (4000*d(binm8)(w10000_3-4000)*0.018)/d(bin)(w10000_3)
w910000_3 <- (4500*d(binm9)(w10000_3-4500)*0.018)/d(bin)(w10000_3)
w1010000_3 <- (5000*d(binm10)(w10000_3-5000)*0.018)/d(bin)(w10000_3)

SL110000_3 <- 1.5*sum(f(x1,w110000_3)*d(bin)(k))
SL210000_3 <- 1.5*sum(f(x2,w210000_3)*d(bin)(k))
SL310000_3 <- 1.5*sum(f(x3,w310000_3)*d(bin)(k))
SL410000_3 <- 1.5*sum(f(x4,w410000_3)*d(bin)(k))
SL510000_3 <- 1.5*sum(f(x5,w510000_3)*d(bin)(k))
SL610000_3 <- 1.5*sum(f(x6,w610000_3)*d(bin)(k))
SL710000_3 <- 1.5*sum(f(x7,w710000_3)*d(bin)(k))
SL810000_3 <- 1.5*sum(f(x8,w810000_3)*d(bin)(k))
SL910000_3 <- 1.5*sum(f(x9,w910000_3)*d(bin)(k))
SL1010000_3 <- 1.5*sum(f(x10,w1010000_3)*d(bin)(k))

purepremium10000_3 <- c(w110000_3+SL110000_3,w210000_3+SL210000_3,
                        w310000_3+SL310000_3,w410000_3+SL410000_3,
                        w510000_3+SL510000_3,w610000_3+SL610000_3,
```

w710000_3+SL710000_3,w810000_3+SL810000_3,
w910000_3+SL910000_3,w1010000_3+SL1010000_3)

#Expected cash-back

```
CB110000_3 <- sum(f(w110000_3,x1)*d(bin)(k))
CB210000_3 <- sum(f(w210000_3,x2)*d(bin)(k))
CB310000_3 <- sum(f(w310000_3,x3)*d(bin)(k))
CB410000_3 <- sum(f(w410000_3,x4)*d(bin)(k))
CB510000_3 <- sum(f(w510000_3,x5)*d(bin)(k))
CB610000_3 <- sum(f(w610000_3,x6)*d(bin)(k))
CB710000_3 <- sum(f(w710000_3,x7)*d(bin)(k))
CB810000_3 <- sum(f(w810000_3,x8)*d(bin)(k))
CB910000_3 <- sum(f(w910000_3,x9)*d(bin)(k))
CB1010000_3 <- sum(f(w1010000_3,x10)*d(bin)(k))
```

```
w110000_4 <- (500*d(binm1)(w10000_4-500)*0.0306)/d(bin)(w10000_4)
w210000_4 <- (1000*d(binm2)(w10000_4-1000)*0.0315)/d(bin)(w10000_4)
w310000_4 <- (1500*d(binm3)(w10000_4-1500)*0.024)/d(bin)(w10000_4)
w410000_4 <- (2000*d(binm4)(w10000_4-2000)*0.0218)/d(bin)(w10000_4)
w510000_4 <- (2500*d(binm5)(w10000_4-2500)*0.0204)/d(bin)(w10000_4)
w610000_4 <- (3000*d(binm6)(w10000_4-3000)*0.0195)/d(bin)(w10000_4)
w710000_4 <- (3500*d(binm7)(w10000_4-3500)*0.0189)/d(bin)(w10000_4)
w810000_4 <- (4000*d(binm8)(w10000_4-4000)*0.018)/d(bin)(w10000_4)
w910000_4 <- (4500*d(binm9)(w10000_4-4500)*0.018)/d(bin)(w10000_4)
w1010000_4 <- (5000*d(binm10)(w10000_4-5000)*0.018)/d(bin)(w10000_4)
```

```
SL110000_4 <- 1.5*sum(f(x1,w110000_4)*d(bin)(k))
SL210000_4 <- 1.5*sum(f(x2,w210000_4)*d(bin)(k))
SL310000_4 <- 1.5*sum(f(x3,w310000_4)*d(bin)(k))
SL410000_4 <- 1.5*sum(f(x4,w410000_4)*d(bin)(k))
SL510000_4 <- 1.5*sum(f(x5,w510000_4)*d(bin)(k))
SL610000_4 <- 1.5*sum(f(x6,w610000_4)*d(bin)(k))
SL710000_4 <- 1.5*sum(f(x7,w710000_4)*d(bin)(k))
SL810000_4 <- 1.5*sum(f(x8,w810000_4)*d(bin)(k))
SL910000_4 <- 1.5*sum(f(x9,w910000_4)*d(bin)(k))
SL1010000_4 <- 1.5*sum(f(x10,w1010000_4)*d(bin)(k))
```

```

P2P110000_4 <- purepremium10000_4[1] - SL110000_4
P2P210000_4 <- purepremium10000_4[2] - SL210000_4
P2P310000_4 <- purepremium10000_4[3] - SL310000_4
P2P410000_4 <- purepremium10000_4[4] - SL410000_4
P2P510000_4 <- purepremium10000_4[5] - SL510000_4
P2P610000_4 <- purepremium10000_4[6] - SL610000_4
P2P710000_4 <- purepremium10000_4[7] - SL710000_4
P2P810000_4 <- purepremium10000_4[8] - SL810000_4
P2P910000_4 <- purepremium10000_4[9] - SL910000_4
P2P1010000_4 <- purepremium10000_4[10] - SL1010000_4

```

#Expected cash-back

```

CB110000_4 <- sum(f(w110000_4,x1)*d(bin)(k))
CB210000_4 <- sum(f(w210000_4,x2)*d(bin)(k))
CB310000_4 <- sum(f(w310000_4,x3)*d(bin)(k))
CB410000_4 <- sum(f(w410000_4,x4)*d(bin)(k))
CB510000_4 <- sum(f(w510000_4,x5)*d(bin)(k))
CB610000_4 <- sum(f(w610000_4,x6)*d(bin)(k))
CB710000_4 <- sum(f(w710000_4,x7)*d(bin)(k))
CB810000_4 <- sum(f(w810000_4,x8)*d(bin)(k))
CB910000_4 <- sum(f(w910000_4,x9)*d(bin)(k))
CB1010000_4 <- sum(f(w1010000_4,x10)*d(bin)(k))

```

