DECREASING THE DEDUCTIBLE IN AN AUTOMOBILE INSURANCE POLICY

Rapport de groupe de travail Isfa

STEPHANE BONCHE
LUDOVIC BRAU
NEIL OLYMPIO

SUPERVISION: PIERRE THEROND

ISFA 2004 - 2005

Summary

Introduction

With a decrease of the numbers of accidents on the French roads, and the recent decision from the European Court of Justice not to discriminate between men and women when setting tariffs, the automobile insurance companies are often in the news.

The French legislation requires every car owner to insure his vehicle. To meet this requirement, insurers offer many possibilities through different kinds of policies such as policy with a deductible.

In automobile insurance, the role of deductibles is essential, not only for the insurance companies, but also for the policyholders.

We will focus on the impact of a decrease of a deductible on the policyholder's and insurer's behaviour. It seems indeed obvious that, for the insured driver, there is an arbitrage between reporting the claim and therefore paying a higher premium the year after, or not declaring the claim and paying a lower premium. This arbitrage takes place especially when the claim amount is a little higher than the amount of the deductible. This is called "hunger for bonus".

For the insurer, lowering the deductible means taking more risk and therefore means an increase of the premium asked to the insured driver.

1 Deductible

An automobile insurance policy is an agreement by which an insurer agrees to pay to a policyholder, a specific amount of money (claim payment or benefit), upon the occurrence of a particular loss defined in the policy. Per contra, the insured (or policyholder) has to pay a fixed amount of money (the premium). An automobile policy may cover different losses such as theft, accident, bodily injuries etc.
Deductibles are often attached to automobile insurance policies. Therefore we will see in this section the use of such a tool for both the insurance company and the policyholder.

1.1 *Definition*

Assuming we have an auto insurance contract with a deductible d. The deductible is the personal financial contribution a policyholder has to make in settlement of the damage. It is therefore the maximum amount of money a policyholder will pay before filing the claim to his insurance company. It never applies to bodily injury.

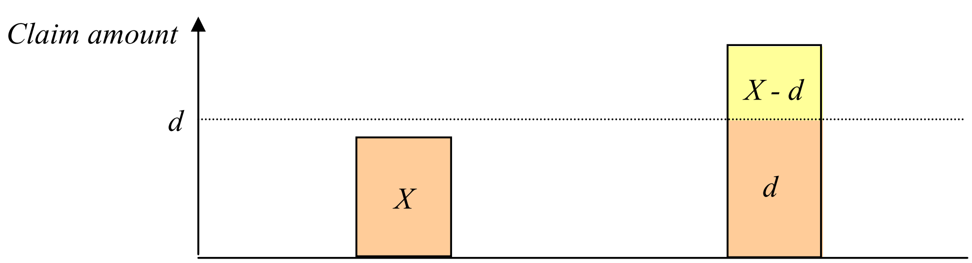
1.2 *How does it work?*

Let's consider a contract inclusive of a deductible of 200€. In case an accident occurs, there are two possibilities:

- The claim amount is lower or equal to 200€, then there will be no charge for the insurance company.
- The claim amount is greater than 200€: In that case, the insurer will have to pay for the loss in excess of the deductible (the difference between the actual claim and 200€).

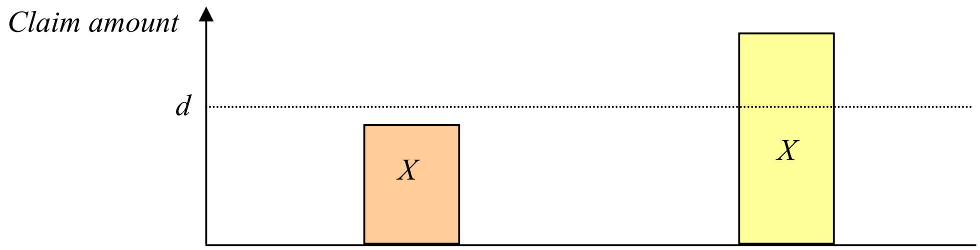
Example: If the loss is less than 200€, the insurer will not pay anything to the policyholder. On the other hand, if the loss is 500€, the insurer will pay 300€.

More formally, if X is a random variable representing the claims, and d is the deductible we have the following:



	$X \leq d$	$X > d$
Amount paid by the insurer	0	$X - d$
Amount paid by the insured	X	d
Total amount paid	X	X

There may also be another type of deductible, which works as stated below:



	$X \leq d$	$X > d$
Amount paid by the insurer	0	X
Amount paid by the insured	X	0
Total amount paid	X	X

This type of deductible is very dangerous for the insurer. Indeed, it's always the interest of the policyholder to have a claim under the deductible rather than below!

In this paper, we will focus on the first type of deductible, which is more frequent, especially in automobile insurance.

1.3 What is the purpose of a deductible?

There are at least two reasons for which an insurer decides to apply a deductible to a contract.

First of all, insurance companies include deductibles in insurance policies as a method of sharing risk with the businesses they insure. By setting a deductible, and therefore sharing the risk, the insurance companies expect their clients to be more responsible. It provides indeed an economic incentive for the policyholder to prevent losses. As a matter of fact, the more a person is involved in the cost of his claim, the more careful this person will be. This is the key principle in setting a deductible. As a consequence, it implies fewer claims to be proceeded by the insurer. In addition to that, processing claims has a cost for the insurer, and this is why the insurance companies will tend to reduce the number of claims with low amounts.

Moreover, attaching a deductible to an insurance contract is an alternative to increasing or decreasing the premium paid by the insured. We will explain in the following section the reasons for which the insurance companies would move the deductible instead of modifying the premiums.

On the other hand, setting a deductible may disappoint the insured since losses are not paid in full. It increases indeed the risk for which the insured remain responsible.

2 Decreasing the deductible

2.1 *Reasons for decreasing a deductible*

First of all, an insurance company may lower the deductible in order to attract more clients. We will explain it later.

Second, depending on the fluctuations of the market, the insurance companies may be tempted to modify the premium paid by the insured. The insurers may indeed increase the premiums when the market is hard, and lower the premiums when the market is softer.

The market has indeed three main components:

- Interest rates

The greatest part of the income for an insurance company comes from the financial markets. This is actually a specificity of the economic cycle of an insurance company: generally, in any other sector than insurance, the price of a good or service is paid at the time the good or service is delivered. Conversely, the insurance companies receive the price (e.g. the premium) before the product (e.g. the claim payment) is delivered. Thus it allows the insurance companies to invest the money received on financial markets to earn their income. That is the reason why the insurer growth is closely linked to the interest rates.

- Frequency of losses

Normally, only a small percentage of insured suffer losses. Those losses are paid out of the premiums collected. The entire pool compensate for the losses of a minority of policyholders. However, the more frequent a risk is, the more dangerous it is for the insurer. Therefore, if for any reason, the frequencies of claims tend to increase or decrease, the insurance companies should adapt the premiums to cover the losses they may undergo.

- Severity of losses

It is the amount of loss given that a loss has occurred. We have seen previously that the insurer will reimburse the difference between the amount of the loss and the deductible, provided there is a deductible.

Insurance companies have had some hard times recently because they kept the premiums at the same level in spite of a decrease in the number of accidents. However, we see in the table below that a decrease of the frequency does not imply a decrease of the premium. As a matter of fact, the severity of accidents should also be taken into account. This is exactly what happened for the French insurance companies. The cost of accidents had indeed increased even if their number has been reduced by a stricter legislation. This is why French insurer could not automatically lower the premiums asked. They had to wait to assess the impact of an increased individual amount of the claims. This is why the severity component is really important.

Therefore, given the same frequency of claims, increased severities of the claims represent a higher risk for the insurer. That is why the premium should be adapted to the fluctuations of the severity.

To simplify our explanation, we will now take into account the frequency and severity components of a loss.

Using mathematical tools, an insurance company can determine the average frequency of losses ($E[F]$) and the expected severity of losses ($E[S]$).

If α denote the average loss for the insurer, α is therefore given by: $E[F]*E[S] = \alpha$ (we consider here that the frequency of claims is independent from the severity of claims). Of course the insurer should ask for total premiums P such that $P > \alpha$.

In year N-1, let the situation of an insurance company be:

Average frequency	:	$E[F]$
Average severity	:	$E[S]$
Expected loss	:	$\alpha = E[F]*E[S]$
Premium	:	$P > \alpha$

In year N, let's now study the modifications that should be made to the premium asked, depending on the evolution of the frequency (F') and severity (S') components:

Average frequency	:	$E[F']$
Average severity	:	$E[S']$
Expected loss for year N	:	$\alpha' = E[F']*E[S']$
Premium	:	$P' > \alpha'$

		FREQUENCY	
		Increase : $F'>F$	Decrease : $F'<F$
SEVERITY	Increase : $S' > S$	$P' > P$	$\alpha > \alpha'$ implies $P' \leq P$
			$\alpha < \alpha'$ implies $P' > P$
	Decrease : $S' < S$	$\alpha > \alpha'$ implies $P' \leq P$	$P' \leq P$
		$\alpha < \alpha'$ implies $P' > P$	

When $\alpha' < \alpha$, it means that the risk is less important in the current year (*N*) than it was in the previous year. Thus, technically speaking, the insurer should lower the premium asked, in order to reflect the actual risk taken.

However, some insurers will still ask the same level of premium, in order not to decrease their income. In such a case, and in order not to loose clients, the insurers have an other option which is to lower the deductible.

As a matter of fact, if the insurer maintains the same premium whereas the premium could be decreased, the insured might want to choose another insurer. This is why, the insurance company in such a situation should make an effort and this effort might come from the level of the deductibles.

In that case, the policyholders will pay less if a loss occurs, and the insurer will still have the same amount of premium.

Let's now focus on the consequences of a decrease of deductible in an automobile insurance contract.

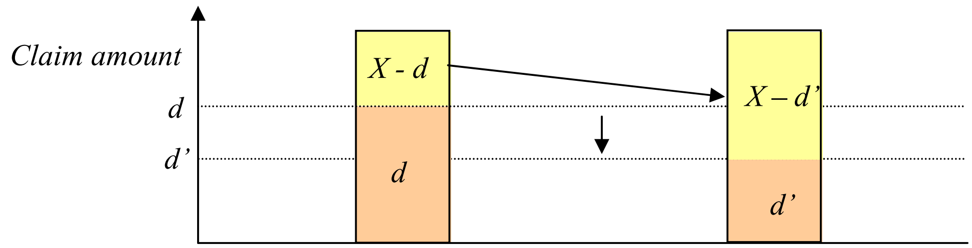
2.2 Consequences for the insurance company

Once an insurance company has lowered the premiums, this company will have to take more losses in charge. There are indeed two consequences for an insurer.

Let's consider a new deductible d' such that d' is lower than d (previous deductible).

- *Paying more for claims already filed*

If we compare the amount paid by the company for a loss amount of $X > d$ (it also implies $X > d'$), between year N-1 and year N, we'll notice that the insurer now has to pay more than he had to, a year earlier:



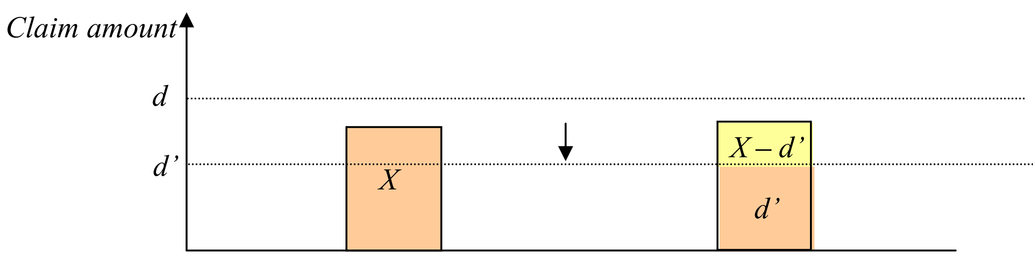
	Year N-1	Year N
Amount paid by the insurer	X-d	X-d'
Amount paid by the insured	d	d'
Total	X	X

Since $d' < d$, for a same amount of claim, the policyholder will pay less in year N than in year N-1. It implies that the insurer will pay more in year N than in year N-1.

- *Paying new claims*

Let's consider a claim X such that: $d' < X < d$

Such a claim was not paid by the insurance company in year N-1 since this claim was lower than the deductible. However, due to the new deductible, a part of this claim has now to be paid in year N, since the claim is higher than the new deductible:



	Year N-1	Year N
Amount paid by the insurer	0	X-d'
Amount paid by the insured	X	d'
Total	X	X

To sum up, lowering the deductible for a commercial purpose will imply more charges for the insurance company which then has no other choice than increasing the premiums. If the deductible is lowered in a soft market period (instead of decreasing the premiums), the increases amount of claims processed will be compensated by a decreased of the claims both in frequency and severity.

2.3 *Consequences for the policyholder*

We have seen previously that a policyholder would benefit from a decrease of the deductible provided that it is a soft market period. As a matter of fact, in a hard market period, the policyholder will pay a higher premium. Conversely, in a soft market period the insurance company has chosen not to increase the premium, therefore it is a real advantage for the policyholder.

3 Decreasing the deductible under a bonus-malus contract

3.1 *Bonus-malus contract*

Under the system discussed previously, the factor of personal history of the driver was not taken into account. In many countries, this factor plays an important role:

- Bonus-malus system influences the premium rate

Actually, Bonus is a reward which allows discounts for claim-free period. Conversely, Malus is a loading in the premium correlated to the number of accidents at fault for the policyholder. Both Malus and Bonus are expressed in a percentage of the premium asked to an insured with no history.

- Bonus-malus acts as a consideration for acceptance of the risk

In countries where the bonus-malus system is imposed by the authorities, insurance companies can compare the clients and their level of bonus since they have the same bonus-malus scale. Therefore, an insurance company will probably not accept to insure a driver with a high malus, since it will represent a higher risk. This solution is not always possible from a legal point of view. Thus the company has one solution which is to increase the premium asked.

The discount system offered acts as an incentive to the insured to drive with care. It then contributes indirectly to road safety.

Actually, a bonus-malus system corresponds to a merit-rating system. It penalises indeed policyholders at fault in accidents by surcharges, and rewards claim-free years by discounts. After each year, the price of the policy (the premium) moves up or down according to transition rules and to the number of claims at fault. With a bonus-malus system, the personal claim history plays an important role.

The regulatory environments of bonus-malus systems vary from one country to another and are extremely diversified. As an example, the rule in the UK is the "total freedom". It actually means that every insurer may design and use its own bonus-malus scale. Conversely, in countries like France and Switzerland, the bonus-malus systems are imposed by the authorities.

In countries where the authorities do not impose the bonus-malus system, it is difficult for an insurance company to accept a client coming from another insurer since the scales are different. Of course some insurance companies have set a conversion system to make it easier for them to assess quickly the level of driving of a potential client.

On the other hand, in countries where the tariff is imposed, an insurance company has no commercial pressure to match the premium to the risks by using any available information since the authorities often prohibit the use of some factors such as the gender, even though they may be significantly correlated to the losses. Actually, a recent order from the European Court of Justice confirmed that the bonus-malus system could still be used, and that no discrimination could be made between men and women.

3.2 The French Bonus-malus system

In France, the bonus-malus system concerns all insurance contracts for drivers of motor vehicles and motorcycles above 80 cm^3. Technically, the state of bonus or malus is represented by a coefficient expressed in percentage of the original premium. A coefficient below 100% indicates a bonus, whereas a coefficient above 100% represents a malus.

This coefficient is calculated by taking into account the claims history of the policyholder and by the use of transition rules.

- History of policyholder

Provided the insured driver is not responsible for any accident, he will get a bonus, and therefore his premium will be lowered. On the other hand, if the driver has had any accidents at fault, he will get a malus and will therefore be penalised since he will have to pay a higher premium for the next coverage period.

The coefficient is calculated two months before the maturity of the annual contract. As an example, for an automobile contract which maturity is 1 January every year, the calculation period would start November 1st and end October 31st the year after. During this period, the insurance company will assess the number of claims at fault and therefore calculate the coefficient.

CLAIM-FREE YEAR

The original coefficient is equal to 1 for a young driver. The initial premium is then multiplied by the coefficient of 0.95 (claim-free year). There is a limit to this coefficient which cannot be lower than 0.5 (50% of the initial premium). As shown in the following, it takes a good driver (no claim each year) 13 years to reach the coefficient of 0.5:

Number of claim-free years	Bonus
1	1.00 x 0.95 = 0.95
2	0.95 x 0.95 = 0.90
3	0.90 x 0.95 = 0.85
4	0.85 x 0.95 = 0.80
5	0.80 x 0.95 = 0.76
6	0.76 x 0.95 = 0.72
7	0.72 x 0.95 = 0.68
8	0.68 x 0.95 = 0.64
9	0.64 x 0.95 = 0.60
10	0.60 x 0.95 = 0.57
11	0.57 x 0.95 = 0.54
12	0.54 x 0.95 = 0.51
13	0.51 x 0.95=0.48 limited to 0.50
14+	0.50

Let us consider the case of a young driver with an initial premium of 250€. If this driver has had no accident at fault for 5 years in a row, his premium for his 6th year of contract will be: 250 x 0.76 = 190€.

CLAIMS AT FAULT

A claim at fault will directly imply a malus. Therefore the coefficient will be greater than 1. In case of a partial liability, the multiplying coefficient is equal to 1.125. On the other hand, if it is a total liability, the coefficient will be equal to 1.25.

There is an upper limit to the value of the malus: it cannot be greater than 3.50 which represents 250% of the initial premium.

Let us consider the case of a driver with initial premium 200€. Assuming this driver has got one accident with total liability the first year, 2 accidents the second year (one with total liability and one with partial liability), and finally, no accident the third year. Let's calculate the premium due for the fourth year of contract:

Year	Claims	Premium at maturity
1	1 with total liability	200 x 1.25 = 250€
2	1 with total liability 1 with partial liability	250 x 1.25 x 1.125 = 351.56€
3	0	351.56 x 0.95 = 333.98€

Even if a driver has got one total liability accident per year, he will not pay more than 200 x 3.50 = 700€. However, if such is the case, an insurance company is not likely to renew his policy. As a matter of fact, all policies are signed for a one year period. After that time, the insurer will assess the situation of the driver and propose a new premium adapted to the risk taken. Then, if an insurer finds out that a driver represents too much risk, he will not renew the policy.

- Transition rules

In addition to the personal claim history of the insured driver and the upper and lower limit of the coefficient (0.50 and 3.50), there are some transition rules in order not to penalise too much both "good" and "less good" drivers.

A policyholder with a bonus of 0.5 (the lower limit) for at least 3 years will not be penalised for the first claim at fault.

After 2 years with no claim at fault, the malus coefficient is brought back to 1.00. It enables a driver to come quickly to a reasonable premium if his behaviour has changed instead of having to wait for many years. Let's see the advantage of this rule:

Considering a driver who has had regularly accidents at fault, so that he has got the maximum malus possible (3.50). It would take 25 years to this driver to come back to a regular premium (coefficient of 1).

Number of claim-free years	Coefficient
1	3.50 x 0.95 = 3.33
2	3.33 x 0.95 = 3.16
3	3.16 x 0.95 = 3.00
4	3.00 x 0.95 = 2.85
5	2.85 x 0.95 = 2.71
.	.
.	.
.	.
24	1.02
25	1.02 x 0.95 = 0.97

With this rule, the driver does not have to wait longer than 2 years.

3.3 *Hunger for bonus*

A bonus-malus system like the French one described previously has many advantages since it discriminates between those drivers who are rarely at fault and those who are often responsible for accidents.

From the insurer point of view, using a bonus-malus scale reflects an adjustment of the premium to the risk. It is compulsory in France, for every single driver to drive an insured vehicle, even if the driver himself is only an occasional driver.

This may then be a means for everyone to find an insurer, provided the premium asked does not exceed a malus of 3.5. (cf. previous section).

From the policyholder point of view, an insured will be rewarded provided he is a good driver. Moreover, with such a system, a driver is responsible only for his driving and does not have to pay for the other drivers.

However, we have seen that the bonus-malus system is based on the number of claims at fault reported in one year by the insured driver. Such a system does not take into account the amount of the claims reported, and allows the policyholder not to report all claims even if he is expected to.

As a consequence, some drivers will tend not to report all claims in order not to have a too high malus. This tendency is called the hunger for bonus.

The hunger for bonus will not occur for every loss. Actually, for a very expensive loss, the insured driver will probably rather have his loss reimbursed rather than a decreased premium the year after. Conversely, for a little claim amount, it might be more interesting for the insured driver not to report his claim and to have a lower premium the year after.

Of course each policyholder has got his own tolerance to risk and the threshold for bonus hunger varies from one driver to another. However, studies show that there is a fixed threshold, common to all rational insured drivers. For a loss below this threshold, it might be interesting for the insured driver not to declare the loss. And from that threshold, a rational policyholder will not retain a claim because the gain in premium is not worth the loss.

An insurance company should be very much aware to that bonus hunger strategy. As a matter of fact, it gives the insurer a wrong appreciation of the risk he is taking. Actually, if the driver has many accidents in a year, and if those accidents were not reported, the insurer will have a false estimation of the possible future losses related to this insured.

Let us illustrate the impact of the hunger for bonus:

Assuming we have a driver with initial premium 200€. This driver has had 3 accidents this year, with total liability, for a total cost of 195€. The driver has decided not to report his claims, let's study the impact of this decision if he has the same losses the year after and decides, this time to declare the losses.

In year *N*-1: No claims reported, so the premium for year *N* will be 200 x 0.95 = 190€.

In year *N*: The driver reports 3 claims for a total cost of 195€.

On the other hand, if the driver had reported his previous claims, his new premium for year *N* would have been: 200 x 1.25 x 1.25 x 1.25 = 390.625

We are indeed studying the behaviour of a policyholder who has the same frequency of accident in year *N*-1 and in year *N*. The only difference is that, in year N-1, the claim is not reported whereas the policyholder decides to report his claim the year after. If such is the case, the insurer does not have a proper assessment of the risk he is taking since he believes the policyholder is not likely to have no accidents a year and 3 the year after. Then the premium has not been adapted.

It is then obvious that, if all policyholders have the same behaviour, there is a high probability that the new premiums won't be enough to cover the losses since the insurance company will have had a false estimation of the risk taken.

3.4 Decreasing the deductible

In such a context, a decrease of the deductible could have an impact more important, than a deductible lowered without a bonus-malus system. As a matter of fact, let us consider the point of view of a driver.

Assuming we have a deductible of 300€:

Without bonus-malus system, it is useless not to report a claim:

	Claim reported	Claim not reported
Deductible	300€	
Claim	2 with total liability	
Total amount of claim	350€	
Premium year N+1	800€	800€
Payment by insurer	50€	0
Payment by insured	300€	350€

With a bonus-malus system,

	Claim reported	Claim not reported
Deductible	300€	
Claim	2 with total liability	
Total amount of claim	350€	
Premium year N+1	800 x 1.25 x 1.25 = 1250€	800€
Payment by insurer	50€	0
Payment by insured	300€	350€

Since there is not a huge difference between the amounts paid by the insurer when the claim is reported and when it is not, it is more interesting in that case to declare the claim.

Assuming we are in a soft market period and, rather than decreasing the premium, the insurance company decides to lower the deductible from 300€ to 200€.

Without Bonus-malus system, it is still useless not to report a claim:

	Claim reported	Claim not reported
Deductible	200€	
Claim	2 with total liability	
Total amount of claim	350€	
Premium year N+1	800€	800€
Payment by insurer	150€	0
Payment by insured	200€	350€

With a Bonus-malus system,

	Claim reported	Claim not reported
Deductible	200€	
Claim	2 with total liability	
Total amount of claim	350€	
Premium year N+1	800 x 1.25 x 1.25 = 1250€	800€
Payment by insurer	150€	0
Payment by insured	200€	350€

We can therefore understand that a policyholder would not be willing to report a claim due to the increase in premium for the following year. In addition to that, it might be more interesting to pay 350€ this year, than paying 1250 – 800 = 450€ as an increase of premium and 200€, that is, a total of 650€. Especially since this difference (650 – 350 =300€) could be invested in financial markets.

4 Aggregate loss models:

4.1 Introduction

Considering a simple situation with a random number N of events and individual payments amount (X_1, X_2,… X_N).

Hence, we can represent the aggregate losses S as:

$$S = X_1 + \ldots + X_N\,, \qquad S = \sum_{i=1}^{N} X_i \qquad\qquad N = 0, 1, 2, \ldots$$

An empty sum always return zero, hence $S = 0$ when $N = 0$.

In practice, N represent the total number of accident occurring during a year and X_i is the cost of the event i. Therefore, S represents the total losses for one year.

Hence, it's reasonable to assume that the number of event is independent of the amount of claim. More formally, the independence assumptions are :

1. Conditional on $N = n$, the random variables X_1, X_2,… X_N are iid random variables.
2. Conditional on $N = n$, the common distribution of the random variables X_1, X_2,… X_N does not depend on n.
3. The distribution of N does not depend in any way on the values of X_1, X_2, …

Obviously, modeling the distribution of N and the distribution of the X_is separately has some real advantages and gives a more accurate and flexible model:

- Only the expected number of claims changes as the number of insured policies changes.

- The effect of general economic inflation and additional claims inflation are reflected in the losses incurred by insured parties and the claims paid by insurances companies.

- Impacts on a specific layer are more easily studied. This is done by changing the specification of the severity distribution.

- The impact on claims frequencies of changing layer is better understood

- The understanding of the relative shapes of N and X_is is useful when modifying policy details.

Thus, S depends on both N and X_i. It's very useful to understand the relative shape when modifying policy details.

4.2 Properties of this model

We can compute the distribution function as below

$$\mathrm{F}_S(x) = \Pr(\mathrm{S} \le \mathrm{x}) = \sum_{n=1}^{\infty} \mathrm{P}(N = n)\, \mathrm{F}_{\mathrm{X}}^{*\mathrm{n}}(x)$$

Where $\mathrm{F}_{\mathrm{X}}(\mathrm{x}) = \Pr(X < \mathrm{x})$ and $\mathrm{F}_{\mathrm{X}}^{*\mathrm{n}}(x)$ is the "n-fold convolution" of the distribution function of *X*.

$$\mathrm{E}(S) = \sum_{i=1}^{\infty} \mathrm{E}(S|N = i) \Pr(N = i) = \sum_{i=1}^{\infty} i\, \mathrm{E}(X) \Pr(N = i) = \mathrm{E}(X)\, \mathrm{E}(N)$$

$$\mathrm{V}(S) = \mathrm{E}\left[\mathrm{V}(S|N)\right] + \mathrm{V}\left[\mathrm{E}(S|N)\right] = \mathrm{E}\left[N\, \mathrm{V}(X)\right] + \mathrm{V}\left[N\, \mathrm{E}(X)\right] = \mathrm{E}(N)\, \mathrm{V}(X) + \mathrm{E}(X)^2\, \mathrm{V}(N)$$

5 Excess of loss

We assume that the same deductible *d* applies to all claims.

Let N_1 be the number of accident for which $X < d$ and N_2 the number for which $X \ge d$. Thus, we have $N = N_1 + N_2$

Let $X_{1i} = X_i \mid X_i < d$
Let $X_{2i} = X_i \mid X_i \ge d$
Let $\alpha = \Pr(X_i \ge d)$

Hence, we can represent the aggregate losses S_2 assuming a deductible of *d* as:

$$S_2 = \sum_{i=1}^{N} \left(X_i - d\right)^+ = \sum_{i=1}^{N_2} X_{2i}$$

Because of the properties of the Poisson distribution demonstrated before, if we suppose that *N* has Poisson distribution with mean λ, then N_2 has Poisson distribution with mean α.λ . In the general case, we have:

$$\Pr(N_2 = k) = \sum_{i=k}^{\infty} \Pr[N_2 = k|N = i) = \sum_{i=k}^{\infty} \binom{k}{i} \alpha^k (1-\alpha)^{i-k} \Pr(N = i)$$

Independence issue.

Quite obviously, X_i and N_2 are not independent. Indeed, the distribution of X_i will have a direct impact on N_2. The interesting fact is that X_{2i} and N_2 are still independent variables. We have just "translated" the problem.

5.1 Properties

We can compute the distribution function as below

$$\mathrm{F}_{S_2}(x) = \Pr(\mathrm{S}_2 \le \mathrm{x}) = \sum_{n=1}^{\infty} \mathrm{P}(N_2 = n)\,\mathrm{F}_{\mathrm{X}_2}^{*\mathrm{n}}(x)$$

$$\mathrm{E}(S_2) = \mathrm{E}(X_2)\,\mathrm{E}(N_2) = \mathrm{E}((X-d)^+)\,\mathrm{E}(N) = [\mathrm{E}(X) - \mathrm{E}(X \wedge d)]\mathrm{E}(N) = \mathrm{E}(S) - \mathrm{E}(X \wedge d)\,\mathrm{E}(N)$$

$$\mathrm{V}(S_2) = \mathrm{E}(N_2)\,\mathrm{V}(X_2) + \mathrm{E}(X_2)^2\,\mathrm{V}(N_2) = \mathrm{E}(N)\,\mathrm{V}((X-d)^+) + \mathrm{E}((X-d)^+)^2\,\mathrm{V}(N)$$

We can easily see that $\mathrm{E}(S_2) < \mathrm{E}(S)$ if $d > 0$.

5.2 Increase of the deductible

Assume that the insurance company increases its deductible d to a new one d'.

Let N_1', N_2', X_{1i}', X_{2i}', S_2', P_2' and α' be the variable associated with the deductible d'.

Thus, we have:

$d' > d$	: the new deductible is higher.
$N_2' < N_2$	: there are fewer claim which are above d' than above d.
$P' < P$	: the insurance company takes less risk, thereby there is a price diminution for the policy.

Formally:

Let $\Delta d = d' - d$

$$P' = \mathrm{E}\left[S_2'\right] = \mathrm{E}\left[\sum_{i=1}^{N_2'} X_{2i}'\right] = \mathrm{E}\left[\sum_{i=1}^{N_2} \left(X_{2i} - \Delta d\right)^+\right]$$

X_{2i} and N_2 are independent as well as $(X_{2i} - \Delta d)^+$ and N_2.

Hence:

$$\begin{aligned} P' &= \mathrm{E}[N_2]\,\mathrm{E}\left[(X_{2i} - \Delta d)^+\right] \\ &= \mathrm{E}[N_2]\left(\mathrm{E}[X_{2i}] - \mathrm{E}[X_{2i} \wedge \Delta d]\right) \\ &= P - \mathrm{E}[N_2]\,\mathrm{E}[X_{2i} \wedge \Delta d] \\ &< P \end{aligned}$$

6 Decrease of the deductible

6.1 First step: Severity variable

There is here a lack of information. The insurer knows nothing about accident with a cost below the deductible. Truncated data present more of a challenge. There are two ways to proceed. One is to shift the data by subtracting the truncation point, the deductible, from each observation. The other is to accept the fact that there is no information about values below the truncation point but then attempt to fit a model for the original population. Therefore it would be easy with this model to compute the new price after a decrease of the deductible. Hence, we can use the method of moment by computing E [$X^k \mid X \geq d$] or the maximum likelihood :

We have:

$$\begin{aligned}
\mathrm{E}\left[(X-d)^{+}\right] &= \mathrm{E}\left[(X-d)^{+} \middle| (X-d)^{+} > 0\right] \times \Pr[X > d] + \mathrm{E}\left[(X-d)^{+} \middle| (X-d)^{+} = 0\right] \times \Pr[X \leq d] \\
&= \mathrm{E}\left[X - d \middle| X > d\right] \times \Pr[X > d] + 0 \\
&= \mathrm{E}\left[X \middle| X \geq d\right] \times \alpha - d\alpha
\end{aligned}$$

By the same way, letting k an integer,

$$\mathrm{E}\left[(X^k - d^k)^{+}\right] = \mathrm{E}\left[X^k \middle| X \geq d\right] \times \alpha - d^k \alpha$$

If we observe that

$$\mathrm{E}\left[(X^k - d^k)^{+}\right] = \mathrm{E}\left[X^k\right] - \mathrm{E}\left[X^k \wedge d^k\right]$$

We can easily calculate the k-th moment of $X \mid X \geq d$:

$$\mathrm{E}\left[X^k \middle| X \geq d\right] = \frac{\mathrm{E}\left[X^k\right] - \mathrm{E}\left[X^k \wedge d^k\right]}{\alpha} + d^k$$

For the maximum likelihood, we have:

$$L = \prod_{j=1}^{n} \frac{f(x_j)}{1 - F(d)} = \prod_{j=1}^{n} \frac{f(x_j)}{\alpha}$$

If the insurer wants to decrease the deductible from d to d’, we firstly have to shift the data by subtracting d – d’ from each observation.

1.1.1. For the exponential distribution:

k-th moment

$$\alpha\,\mathrm{E}\left[X^k \middle| X \geq d\right] = \mathrm{E}\left[X^k\right] - \mathrm{E}\left[X^k \wedge d^k\right] + d^k \alpha = \lambda^{-k} k!\,(1-\Gamma(k+1; d\lambda)) - d^k\,\mathrm{e}^{-d\lambda} + d^k\,\mathrm{e}^{-d\lambda}$$

$$\mathrm{E}\left[X^k \middle| X \geq d\right] \quad = \lambda^{-k} k!\,(1-\Gamma(k+1; d\lambda))\,\mathrm{e}^{d\lambda}$$

$$\mathrm{E}\left[X \; \middle| X \geq d\right] \quad = \lambda^{-1} + d$$

maximum likelihood

$$L = \prod_{j=1}^{N} \frac{\lambda\,\mathrm{e}^{-\lambda X_j}}{\mathrm{e}^{-\lambda d}} \Rightarrow \frac{\partial l}{\partial \lambda} = \frac{N}{\lambda} + Nd - \sum X_i$$

$$\hat{\lambda} = \frac{N}{\sum X_i - Nd}$$

We can observe that both methods give the same result.

1.1.2. For the pareto distribution:

k-th moment

$$\alpha\,\mathrm{E}\left[X^k \middle| X \geq d\right] = \mathrm{E}\left[X^k\right] - \mathrm{E}\left[X^k \wedge d^k\right] + d^k \alpha = \frac{\theta^k k!}{(\alpha-1)\ldots(\alpha-k)}(1-\beta(k+1, \alpha-k; \frac{d}{d+\theta})) + 0$$

$$\mathrm{E}\left[X^k \middle| X \geq d\right] \quad = \frac{(d+\theta)^\alpha \theta^{k-\alpha} k!}{(\alpha-1)\ldots(\alpha-k)}(1-\beta(k+1, \alpha-k; \frac{d}{d+\theta}))$$

$$\mathrm{E}\left[X \; \middle| X \geq d\right] \quad = \frac{(d+\theta)^\alpha \theta^{1-\alpha} k!}{(\alpha-1)}(1-\beta(2, \alpha-1; \frac{d}{d+\theta}))$$

maximum likelihood

$$L = \prod_{j=1}^{N} \frac{\dfrac{\alpha\theta^\alpha}{(Xj+\theta)^{\alpha+1}}}{\dfrac{\theta^\alpha}{(d+\theta)^\alpha}} = \prod_{j=1}^{N} \frac{\alpha(d+\theta)^\alpha}{(Xj+\theta)^{\alpha+1}} \Rightarrow \frac{\partial l}{\partial \lambda} = \frac{N}{\alpha} + N\ln(d+\theta) - \sum \ln(Xj+\theta)$$

$$\hat{\alpha} = \frac{N}{\sum_{i=1}^{N} \ln(Xj+\theta) - N\ln(d+\theta)}$$

1.1.3. For the Weibull distribution:

k-th moment

$$\alpha\,\mathrm{E}\left[X^k \middle| X \geq d\right] = \mathrm{E}\left[X^k\right] - \mathrm{E}\left[X^k \wedge d^k\right] + d^k\alpha = \theta^k\Gamma(1+k/\tau)\left[1-\Gamma(1+k/\tau);(d/\theta)^\tau\right] + 0$$

$$\mathrm{E}\left[X^k \middle| X \geq d\right] \quad = \theta^k\Gamma(1+k/\tau)\left[1-\Gamma(1+k/\tau;(d/\theta)^\tau)\right]\exp((d/\theta)^\tau)$$

$$\mathrm{E}\left[X \;\middle| X \geq d\right] \quad = \theta^k\Gamma(1+1/\tau)\left[1-\Gamma(1+1/\tau;(d/\theta)^\tau)\right]\exp((d/\theta)^\tau)$$

maximum likelihood

$$L = \frac{\tau^N \exp(N(x/\theta)^\tau)\prod_{i=1}^{N}(Xi^{\tau-1})\exp\left(-\theta^{-\tau}\sum Xi^\tau\right)}{\theta^{N\tau}} \Rightarrow \frac{dl}{d\theta} = \frac{\tau}{\theta^{\tau+1}}\sum_{i=1}^{N} Xi^\tau - \frac{N\tau}{\theta} - \frac{Nd^\tau}{\theta^{\tau+1}}$$

$$\hat{\theta} = \left(\frac{1}{N}\sum_{i=1}^{N} Xi^\tau - \frac{d^\tau}{\tau}\right)^{1/\tau}$$

1.1.4. For the Gamma distribution:

k-th moment

$$\alpha'\,\mathrm{E}\left[X^k \middle| X \geq d\right] = \mathrm{E}\left[X^k\right] - \mathrm{E}\left[X^k \wedge d^k\right] + d^k\alpha' = \tau^{-k}\,\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}(1-\Gamma(a+k;d\tau)) + 0$$

$$\mathrm{E}\left[X^k \middle| X \geq d\right] \quad = \tau^{-k}\,\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}\frac{(1-\Gamma(a+k;d\tau))}{(1-\Gamma(a;d\tau))}$$

$$\mathrm{E}\left[X \;\middle| X \geq d\right] \quad = \tau^{-1}\,\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}\frac{(1-\Gamma(a+1;d\tau))}{(1-\Gamma(a;d\tau))}$$

maximum likelihood

$$L = \frac{\prod_{i=1}^{N}(Xi^{\alpha-1})\exp\left(-\tau\sum Xi\right)}{\tau^{-N\alpha}\left(\Gamma(\alpha)\right)^N\left(1-\Gamma(\alpha;d\tau)\right)^N} \Rightarrow \frac{dl}{d\tau} = -\sum_{i=1}^{N} Xi + \frac{N\alpha}{\tau} - N\frac{d}{d\tau}\ln\left(1-\Gamma(\alpha;d\tau)\right)$$

$$\frac{d}{d\tau}\Gamma(\alpha;d\tau) = \frac{d}{d\tau}\frac{1}{\Gamma(\alpha)}\int_0^{d\tau} t^{\alpha-1}\,\mathrm{e}^{-t}\,dt = \frac{1}{\Gamma(\alpha)}\frac{d}{d\tau}\mathrm{F}(d\tau) = \frac{1}{\Gamma(\alpha)}d\left(d\tau\right)^{\alpha-1}\,\mathrm{e}^{-d\tau}$$

$$\Gamma(\alpha;d\tau) = 1 - \sum_{j=0}^{\alpha-1}\frac{x^j\,\mathrm{e}^{-x}}{j!} \quad \text{if } \alpha \text{ is an integer}$$

Hence, with $\alpha = 2$, we have :

$$\Gamma(2) = 1 \Rightarrow -N\frac{d}{d\tau}\ln\left(1-\Gamma(\alpha;d\tau)\right) = \frac{Nd^2\tau\,\mathrm{e}^{-d\tau}}{1+d\tau}$$

Then,

$$\frac{dl}{d\tau} = 0 \Rightarrow m \overset{\Delta}{=} \frac{1}{N}\sum_{i=1}^{N} Xi = \frac{2}{\tau} + \frac{d^2\tau}{1+d\tau}$$

$$\Leftrightarrow \tau^2\left(d^2 - dm\right) + \tau(2d - m) + 2 = 0$$

And

$$\Delta = -4d^2 + 4dm + m^2$$

$$\hat{\tau} = \frac{m - 2d \pm \sqrt{\Delta}}{2d\left(d - m\right)}$$

Δ is strictly positive because all the Xi are above the deductible and then $m > d$.
To compute τ, we choose the positive solution that makes l maximum.

1.1.5. For the Lognormal distribution:

k-th moment

$$\alpha\,\mathrm{E}\left[X^k \middle| X \geq d\right] = \mathrm{E}\left[X^k\right] - \mathrm{E}\left[X^k \wedge d^k\right] + d^k\alpha = \exp(k\mu + k^2\sigma^2/2)\left[1 - \Phi\left(\frac{\ln d - \mu - k\sigma^2}{\sigma}\right)\right] + 0$$

$$\mathrm{E}\left[X^k \middle| X \geq d\right] = \exp(k\mu + k^2\sigma^2/2)\left[1 - \Phi\left(\frac{\ln d - \mu - k\sigma^2}{\sigma}\right)\right]\left(1 - \Phi\left(\frac{\ln d - \mu}{\sigma}\right)\right)^{-1}$$

$$\mathrm{E}\left[X \middle| X \geq d\right] = \exp(\mu + \sigma^2/2)\left[1 - \Phi\left(\frac{\ln d - \mu - \sigma^2}{\sigma}\right)\right]\left(1 - \Phi\left(\frac{\ln d - \mu}{\sigma}\right)\right)^{-1}$$

maximum likelihood

$$L = \prod_{i=1}^{N} \frac{1}{Xi\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(Xi) - \mu)^2}{2\sigma^2}\right](1 - \Phi(d))^{-1}$$

$$L = \frac{1}{\left(\prod_{i=1}^{N} Xi\right)\sigma^N (2\pi)^{N/2}\left(1 - \Phi\left(\frac{\ln d - \mu}{\sigma}\right)\right)^N} \exp\left[\sum_{i=1}^{N} -\frac{(\ln(Xi) - \mu)^2}{2\sigma^2}\right]$$

$$l = -\sum_{i=1}^{N} \ln(Xi) - N\ln(\sigma) - \frac{N}{2}\ln(2\pi) - N\ln\left(1 - \Phi\left(\frac{\ln d - \mu}{\sigma}\right)\right) - \sum_{i=1}^{N} \frac{(\ln(Xi) - \mu)^2}{2\sigma^2}$$

$$\frac{dl}{d\mu} = \sum_{i=1}^{N} \frac{\ln(Xi) - \mu}{\sigma^2} - \frac{N}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\ln d - \mu)^2}{2\sigma^2}\right) \frac{1}{1 - \Phi\left(\frac{\ln d - \mu}{\sigma}\right)} = 0$$

$$\frac{dl}{d\sigma} = -\frac{N}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{N} (\ln(Xi) - \mu)^2 - \frac{N}{\sqrt{2\pi}} \frac{\ln d - \mu}{\sigma^2} \exp\left(-\frac{(\ln d - \mu)^2}{2\sigma^2}\right) \frac{1}{1 - \Phi\left(\frac{\ln d - \mu}{\sigma}\right)} = 0$$

1.1.6. Test of fitness

In order to test our estimators and check their accuracy we have realised two type of simulation. First, we have used SAS to simulate 1000 realisation for each distribution above. Then we have used Excel to compute estimators end Chi-2 test. Parameters are such that realisations are approximately between 0 and 100 000. To test our estimators, we begin by only use simulation data above deductible of 70000, 50000 and 30000.

We have also indicated the error made between estimators and true parameters. The Chi-2 test is clearly not the better one for continuous distribution. But it is easy to compute in Excel and with appropriate choice for intervals using during the test, we have compute interesting result.

Table 6.1. Simulations - Estimators

	Exponential		Pareto		Weibull		Gamma	
True parameters	1 / l =	15000	a =	0.6	q =	30000	1 / t =	10000
			q =	10	t =	1.5	a =	2
Deductible 70 000	1 / l =	8642	a =	0.49	q =	53558	1 / t =	12526
	error =	42.39%	error =	18.33%	error =	78.53%	error =	25.26%
	info =	0.80%	info =	0.50%	info =	2.20%	info =	0.40%
	Chi-2 =	0.66	Chi-2 =	0.66	Chi-2 =	0.004	Chi-2 =	0.09
Deductible 50 000	1 / l =	13987	a =	0.42	q =	42095	1 / t =	9625
	error =	6.75%	error =	30.00%	error =	40.32%	error =	3.75%
	info =	3.20%	info =	0.50%	info =	11.20%	info =	33%
	Chi-2 =	0.8	Chi-2 =	0.7	Chi-2 =	0	Chi-2 =	0.2
Deductible 30 000	1 / l =	13830	a =	0.51	q =	35900	1 / t =	9520
	error =	7.80%	error =	15.00%	error =	19.67%	error =	4.80%
	info =	14.00%	info =	75.00%	info =	36.00%	info =	18%
	Chi-2 =	0.77	Chi-2 =	0.64	Chi-2 =	0	Chi-2 =	0.5
Deductible 0	1 / l =	15157	a =	0.59	q =	29600	1 / t =	9720
	error =	1.05%	error =	1.67%	error =	-1.33%	error =	2.80%
	info =	100.00%	info =	100%	info =	100%	info =	100%
	Chi-2 =	0.59	Chi-2 =	0.54	Chi-2 =	0.35	Chi-2 =	0.48

Because distributions are really different, work only on deductible criteria may not be relevant for make comparison between distributions. As an example, we only have 5 realisations above 50000 for the Pareto distribution, whereas the Weibull one has 110 realisations. To make our test comparison relevant, we have also tested our parameters for deductible which are adjusted for each distribution to have the same level of information. Information = 70% means that above the deductible of 5500 for the exponential, or above the deductible of 8.4 for the Pareto, we exactly have 300 realisations. In other word, we have the same portion of the distribution's tail for compute estimator.

Table 6.2. Simulations - Estimators

	Exponential		**Pareto**		**Weibull**		**Gamma**	
True parameters	1 / l =	15000	a =	0.6	q =	30000	1 / t =	10000
			q =	10	t =	1.5	a =	2
Information 30%	1 / l =	14627	a =	0.58	q =	37223	1 / t =	9657
	error =	2.49%	error =	3.33%	error =	24.08%	error =	3.43%
	d =	18500	d =	66.3	d =	33600	d =	23280
	Chi-2 =	0.56	Chi-2 =	0.55	Chi-2 =	0	Chi-2 =	0.8
Information 50%	1 / l =	15013	a =	0.58	q =	34466	1 / t =	9384
	error =	0.09%	error =	3.33%	error =	14.89%	error =	6.16%
	d =	10550	d =	22.3	d =	22700	d =	16600
	Chi-2 =	0.4	Chi-2 =	0.54	Chi-2 =	0.0003	Chi-2 =	0.64
Information 70%	1 / l =	15008	a =	0.59	q =	31960	1 / t =	9496
	error =	0.05%	error =	1.67%	error =	6.53%	error =	5.04%
	d =	5500	d =	8.4	d =	14900	d =	11030
	Chi-2 =	0.5	Chi-2 =	0.53	Chi-2 =	0.008	Chi-2 =	0.61
Information 100%	1 / l =	15157	a =	0.59	q =	29600	1 / t =	9720
	error =	1.05%	error =	1.67%	error =	1.33%	error =	2.80%
	d =	0	d =	0	d =	0	d =	0
	Chi-2 =	0.77	Chi-2 =	0.54	Chi-2 =	0.35	Chi-2 =	0.48

We notice that the estimator for the exponential distribution gives really good results with less information (4%) on the distribution. The Pareto and Gamma distribution's parameters estimator need at least 20 % of the realisation to be computed. The estimator for the Weibull distribution has to be computed with at least 60 % of all the realisations.

The second type of simulation was only traded in SAS. It's allow us to simulate a high number of data and test the accuracy of our estimator by using a Kolmogorov test which is more appropriate in such situation. The table are directly produced by our SAS Code:

Estimateurs Conditionnels de la loi exponentielle 1/l = 15000

Obs	Franchise	Information	Estimateur
1	0	1.00000	15029.33
2	1000	0.93544	15032.47
3	10000	0.51411	15039.75
4	30000	0.13596	15131.66
5	50000	0.03638	14885.07

Estimateurs Conditionnels de la loi de Pareto a=0.6 q=10

Obs	Franchise	Information	Estimateur
1	0	1.00000	0.60065
2	1000	0.06178	0.59696
3	10000	0.01560	0.58889
4	30000	0.00823	0.59362
5	50000	0.00612	0.59948

Estimateurs Conditionnels de la loi de Weibull q=30000 t=1.5

Obs	Franchise	Information	Estimateur
1	0	1.00000	30020.95
2	1000	0.99390	30062.28
3	10000	0.82393	31321.22
4	30000	0.36913	36337.01
5	50000	0.11737	42755.93

Estimateurs Conditionnels de la loi Gamma 1/t=10000 a=2

Obs	Franchise	Information	Estimateur
1	0	1.00000	9998.64
2	1000	0.99504	10001.36
3	10000	0.73631	9993.62
4	30000	0.19790	10033.77
5	50000	0.03974	10299.50

To have a graphical view, we have drawn for each distribution (Exponential, Pareto...), the distribution function of the simulated data, and the one for each level of information : 30 %... 100 % :

Chart 6.1. Exponential Repartition Function

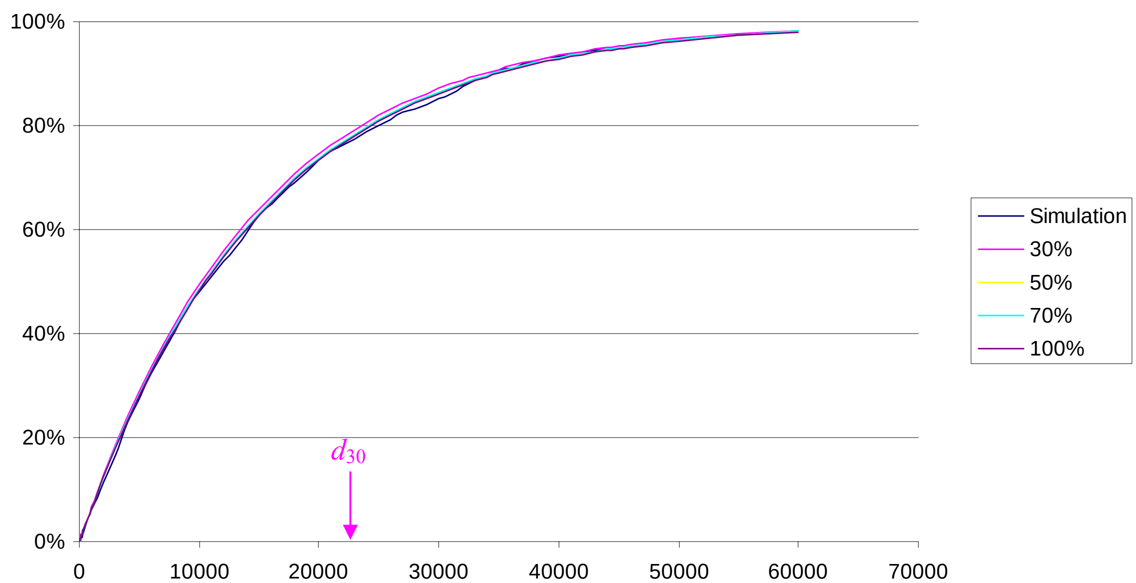


Chart 6.2. Pareto Repartition Function

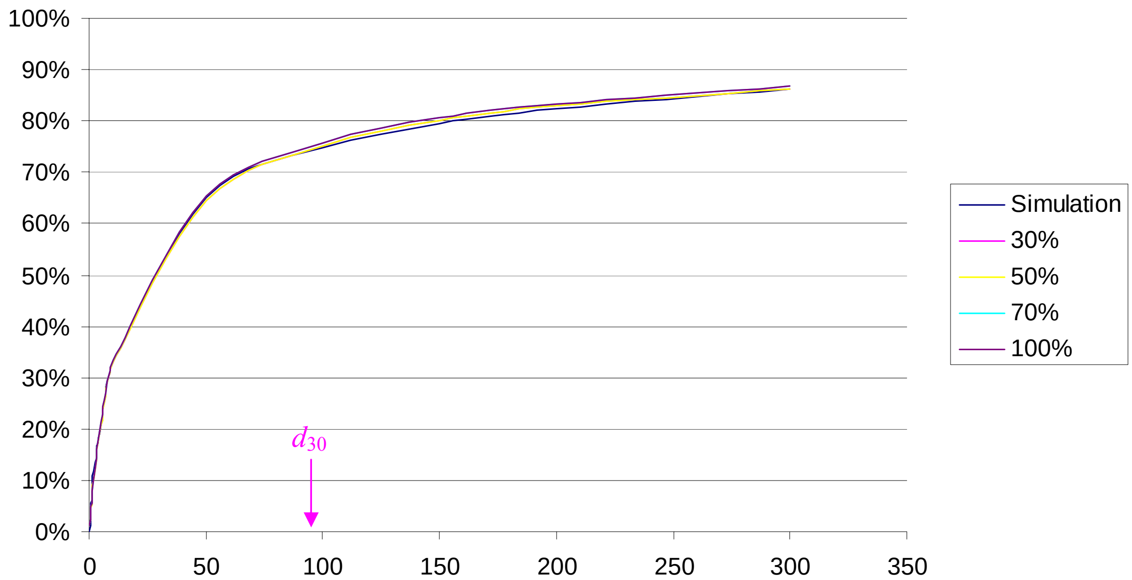


Chart 6.3. Weibull Repartition Function

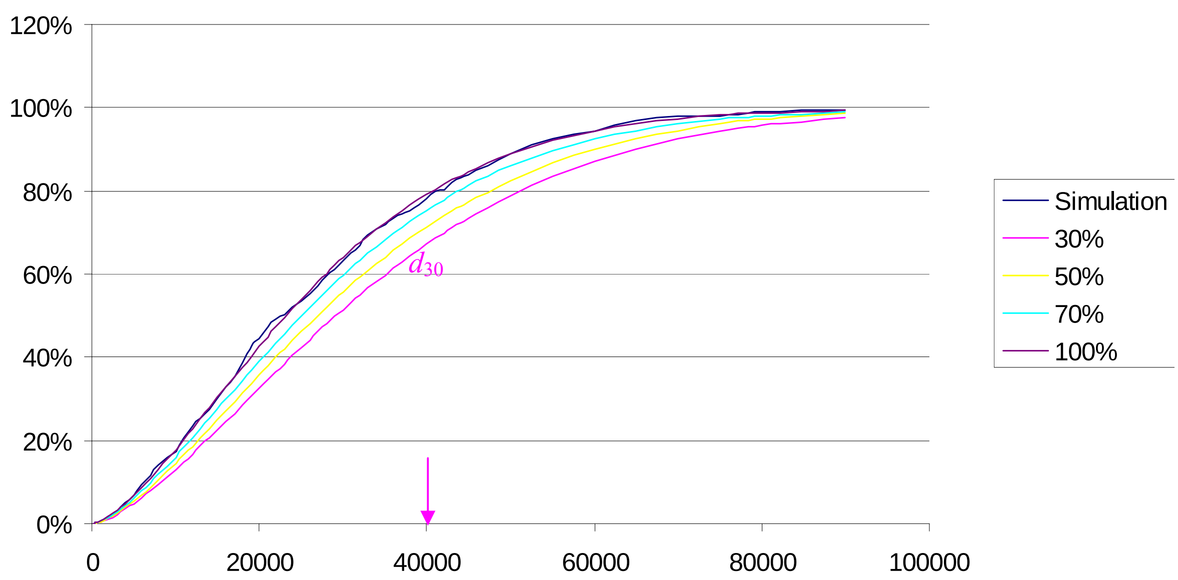
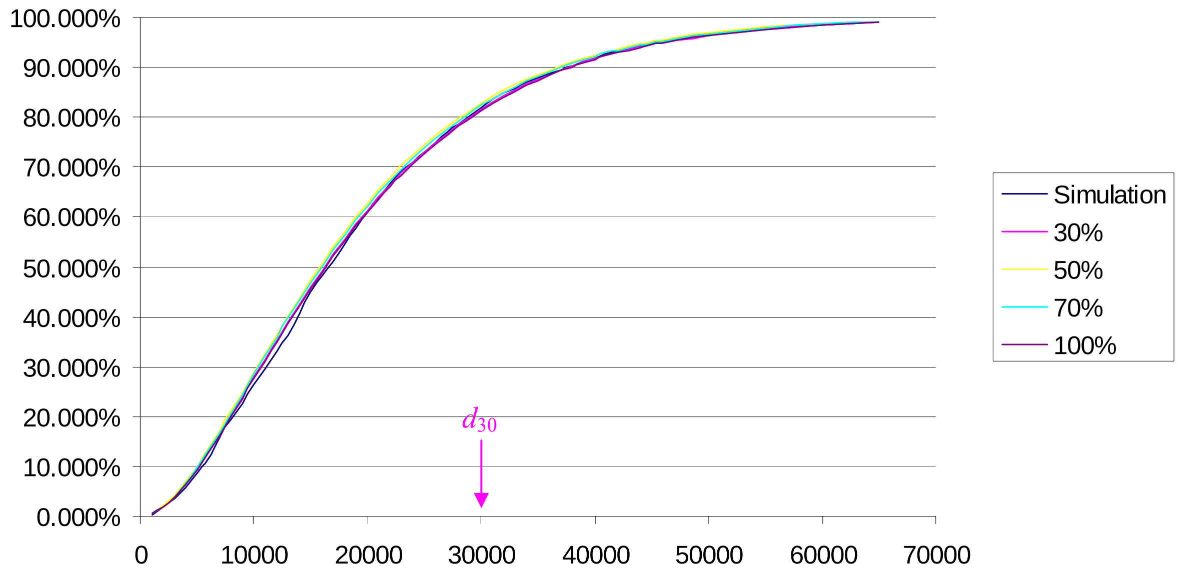


Chart 6.4. Gamma Repartition Function



It seems that our estimator gives really good result. In order to really see what's going on about accuracy of our estimators, we have drawn the density function for each distribution with the true parameters (yellow), and estimators based on only 30 % of information (in blue). We observe that estimators for the exponential and Pareto distributions under estimate the tail whereas the Weibull and Gamma distribution have a more prudential distribution with higher probabilities in the tail.

Chart 6.5 Exponential Density

True parameter: $1/\lambda = 15000$
Estimator: $1/\lambda' = 14627$
Deductible = 18100

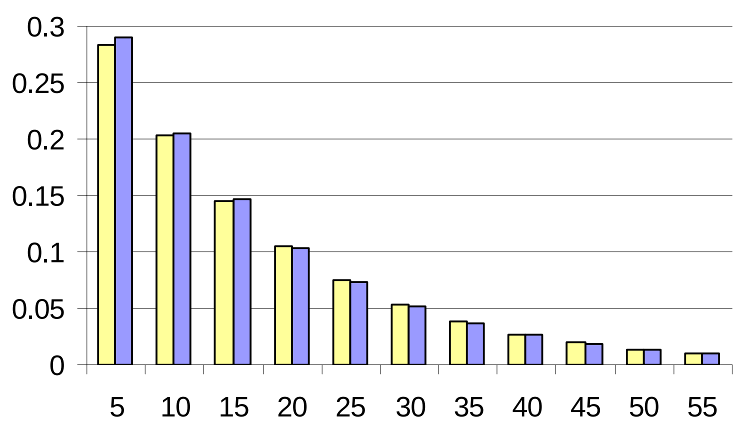


Chart 6.2. Pareto Density
True parameter: $\alpha = 0.6$, $\theta = 10$
Estimator: $\alpha = 0.58$, $\theta = 10$
Deductible = 65

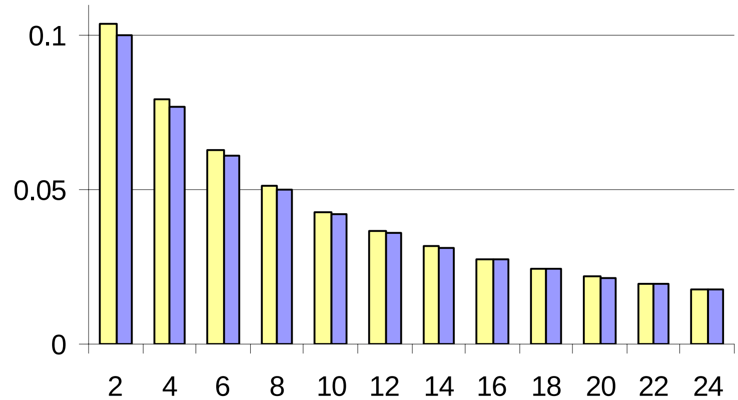


Chart 6.3. Weibull Density
True parameter: $\tau = 1.5$, $\theta = 30000$
Estimator: $\tau = 1.5$, $\theta = 37223$
Deductible = 34000

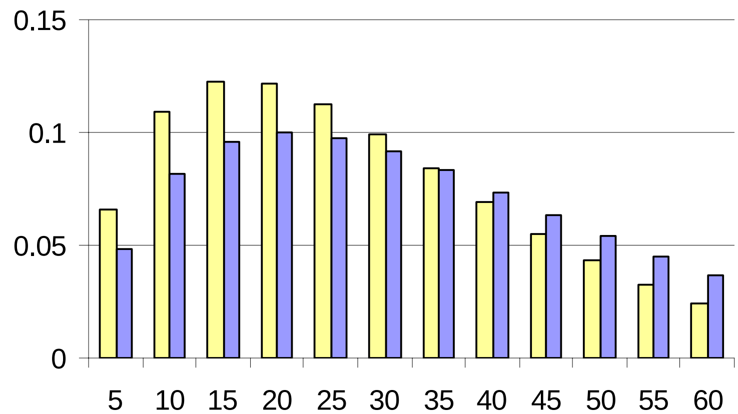
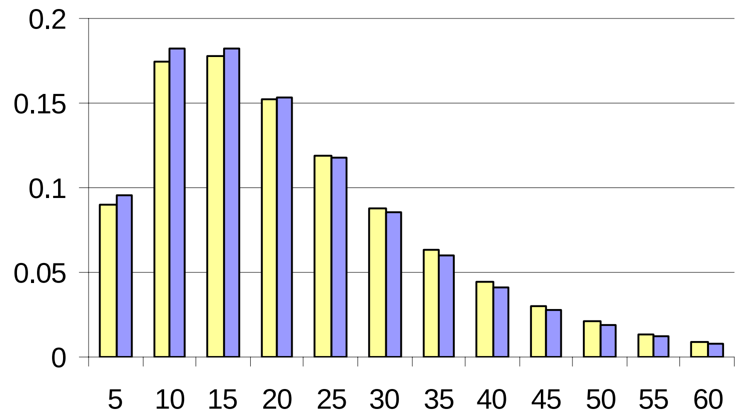


Chart 6.4. Gamma Density
True parameter: α=2 τ=1/10000
Estimator: α=2 τ=1/10000
Deductible = 34000



6.2 *Second step: Frequency variable*

An important component in analysing the effect of policy modifications pertains to the change in the frequency distribution of payments when the deductible is changed. We assume that the first step has been done and consequently, that α is determined. It's reasonable to assume that the imposition of coverage modifications does not affect the process that produces losses. Let's define the indicator random variable I_j by $I_j = 1$ if the jth loss results in a payment and $I_j = 0$ otherwise. Then I_j has a Bernoulli distribution with parameter α and its probability generating function is $P_{I_j}(z) = 1 + \alpha(z-1)$. It's clear that every I_j are mutually independent with N, and so N_2 has a compound distribution. Thus,

$$P_{N_2}(z) = P_N(1 + \alpha(z-1))$$

Hence:

N	Parameter for N_2
Poisson	$\lambda_2 = \alpha\,\lambda$
Binomial	$q_2 = \alpha\, q$
Negative binomial	$\beta_2 = \alpha\,\beta$ $r_2 = r$

The results can be further generalized to an increase or decrease in the deductible by setting α equals to :

$$\alpha = \frac{1 - F_X(d')}{1 - F_X(d)} \text{ where } d' \text{ is the new deductible}$$

It's very interesting to observe that only the new severity claim have to be computed. The frequency variable will just follow up in this situation. The approach yields very simple solutions. Nevertheless there are some important theoretical difficulties. There is no theoretical justification why the model for X is the same model has for X. Thus, there is no guidance for the choice of a model for X, one would prefer a more cautious model and others would select a more accurate model. Besides, the fact that N and N_2 have the same distribution is a direct consequence of the independence between frequency and severity.

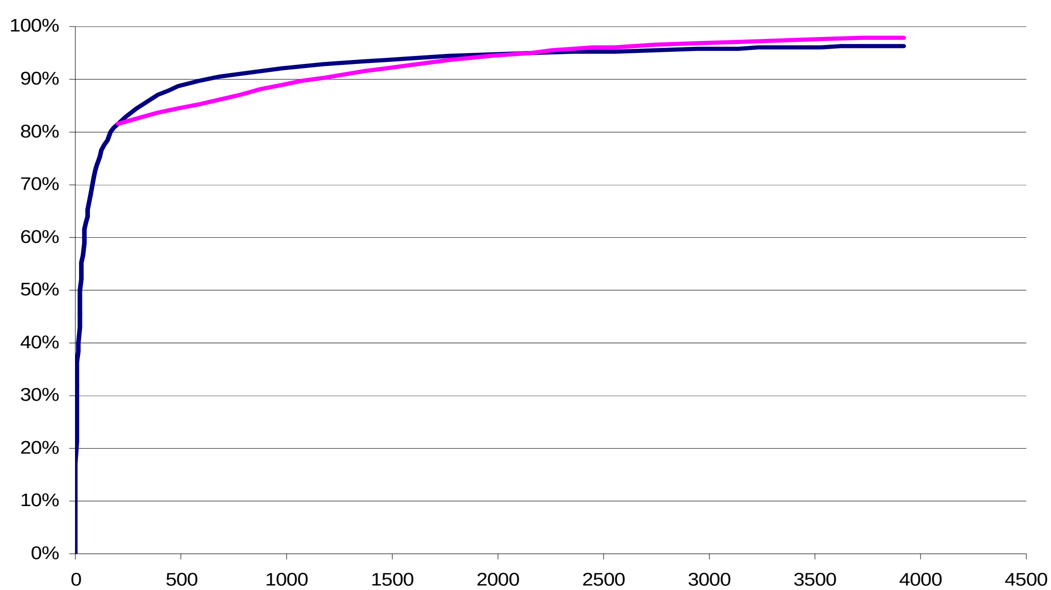
6.3 *Data analysis*

We have two set of car insurance data, both with a deductible of 200. Using our estimator, we have the following results:

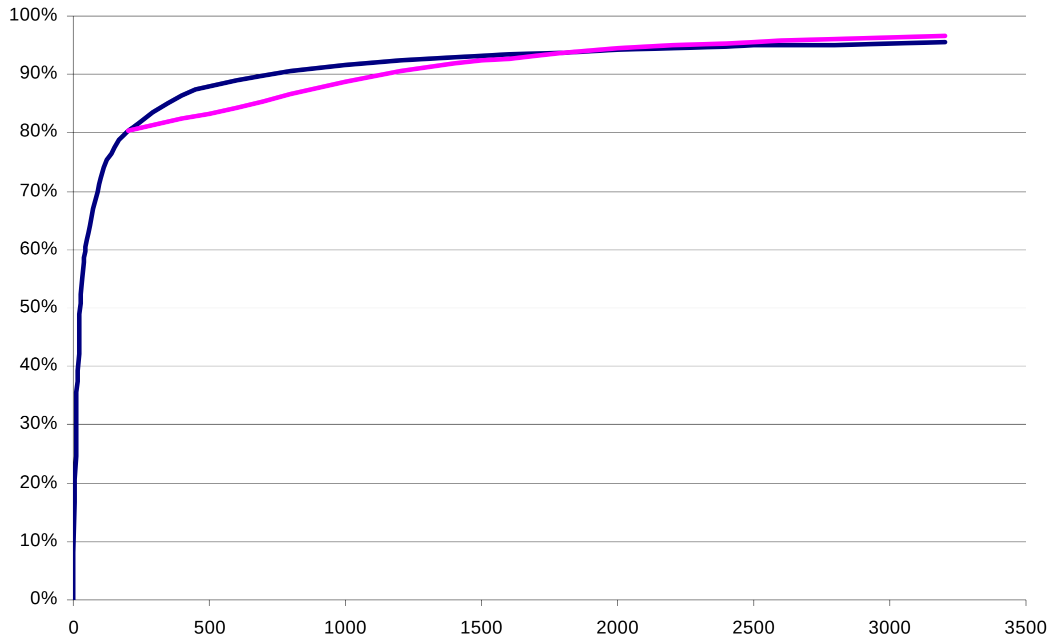
	Exponential		Pareto		Weibull		Gamma	
Contract 1	$1/\lambda =$	2219	$\alpha =$	0.53	$\theta =$	3358	$1/\tau =$	1195
			$\theta =$	10	$\tau =$	1.5	$\alpha =$	2
	p-val =	< 0.001	p-val =	< 0.001	p-val =	< 0.001	p-val =	< 0.001
Contract 2	$1/\lambda =$	1753	$\alpha =$	0.55	$\theta =$	2464	$1/\tau =$	959
			$\theta =$	10	$\tau =$	1.5	$\alpha =$	2
	p-val =	< 0.001	p-val =	< 0.001	p-val =	< 0.001	p-val =	< 0.001

P-value corresponds to the Kolmogorov test. All p-values indicate that none of those models correctly fit our data. We have chosen to use the Pareto distribution. Indeed, this distribution is commonly used in non-life insurance and has a prudential tail. First, we draw the observed density function and the Pareto one. The only real chart we can make is with conditional functions. Indeed, it is impossible to represent on the same chart the whole distribution of the adjusted Pareto and the portion of distribution of observed values without making assumption about where the truncated distribution begin. Indeed, we have no idea about the probability for an amount of claim to be above the deductible. In consequence, we have made the assumption that this probability is exactly the probability predicted by our model.

Density function Data 1



Density function Data 2



We don't have enough information about number of accident to adjust a frequency model. We assume that we expect 0.2 accidents per year above 200 for the contract number 1. We assume that the second portfolio is more dangerous with an expected number of 0.4.

Using this formula:

$$E[N] = \frac{E[N|X \geq 200]}{P[X \geq 200]} = \frac{E[N|X \geq 200]}{\left(\frac{10}{200+10}\right)^{\alpha}}$$

Total expected numbers of accident are:

Data 1: 0.2 / 0.184 = 1.08

Data 2: 0.4 / 0.195= 2.045

The expected amount of an accident for a deductible d between 0 and 200 given that an accident has occurring is:

$$E[X|X \geq d] = \int_d^\infty P[X \geq x|X \geq d]dx = \int_d^\infty \frac{P[X \geq x]}{P[X \geq d]}dx = \frac{d+10}{\alpha - 1}$$

The first moment does not exist with an $\alpha \leq 1$. In order to compute the expected mean, we assume a maximum loss amount of 100 000:

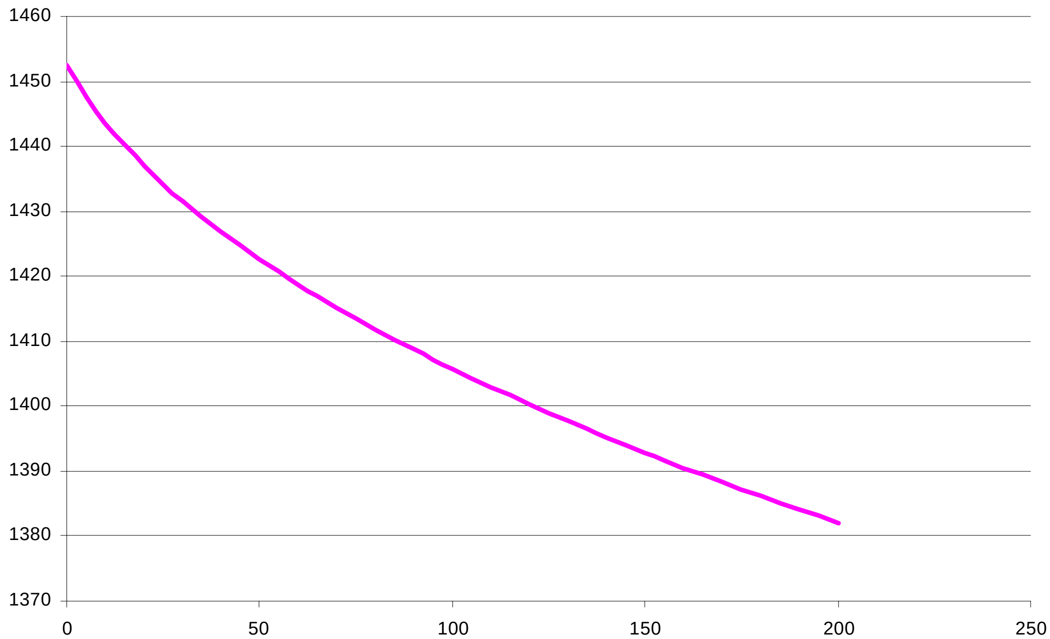
$$E[X|d \leq X \leq Max] = \int_d^{\max} P[X \geq x|d \leq X \leq Max]dx = \int_d^{\max} \frac{P[X \geq x]}{P[d \leq X \leq Max]}dx$$

$$E[X|d \le X \le Max] = \frac{10^{\alpha}}{\alpha-1}\left(\frac{1}{(d+10)^{\alpha-1}} - \frac{1}{(Max+10)^{\alpha-1}}\right)\left(\frac{10^{\alpha}}{(d+10)^{\alpha}}\left(1-\frac{10^{\alpha}}{(Max+10)^{\alpha}}\right)\right)^{-1}$$

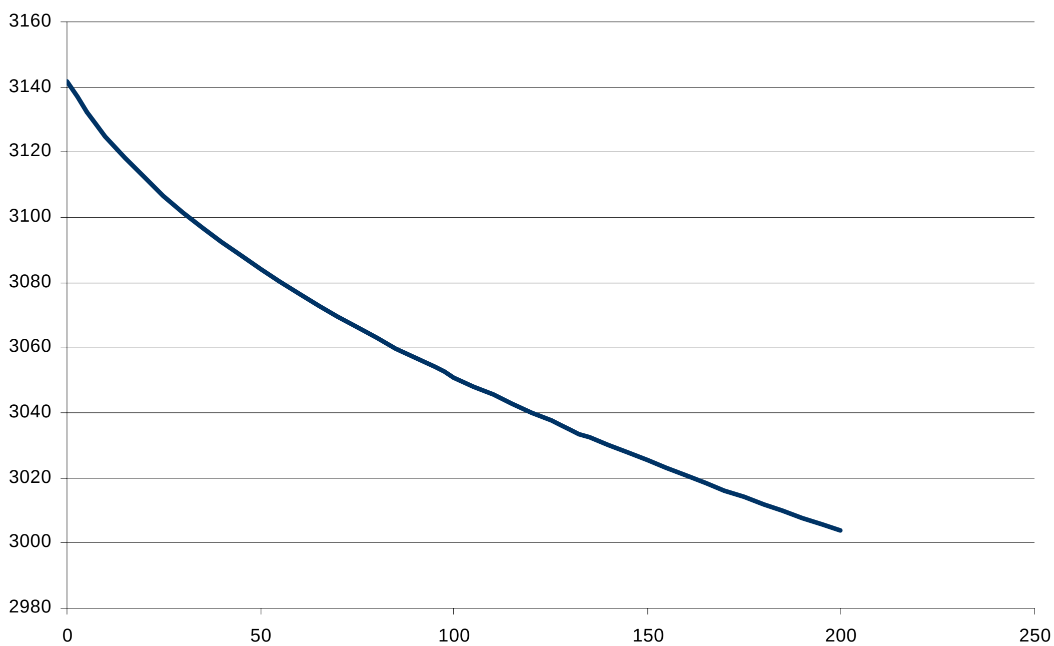
Thus, the premium with a deductible d is: $P_d = E[X|X \ge d] \times E[N] \times P[X \ge d]$

We are now able to represent the evolution of the premium against the decrease of the deductible for each contract:

Premium against deductible - Data 1



Premium against deductible - Data 2



Obviously our results are far from car reinsurance reality. Nevertheless, it allows interpretations. As expected, because of the increase of the number of accident, the premium increase when the deductible decrease. Data 2 has a lower α and then a most important tail, moreover, there is a higher probability of having accidents, and thus, it is obvious that the premium for Data 2 is higher. One may observe an acceleration of the increase of the premium. It is clear that with no deductible at all, one may expect a very high frequency.

This frequency effect is really difficult to predict. It is not reasonable to directly decrease the deductible to zero. One more prudential way may be to decrease the deductible step by step, and adjust the model year after year. Another possibility may be to estimate the premium for a decrease of 80 for example and applied the premium with a decrease of only 50.

Let's now focus on the Bonus Malus issue.

7 The Bonus Malus system

The insurer has to take in account the impact of the bonus malus system. Indeed, as we have seen in the first part, the hunger of bonus implies that the insurer does not have all information about the distribution of claims just above the deductible. However, it is clear that there exist an amount *K* from which the policyholder should declare a claim rather than hunting for bonus. It would be interesting to compute the premium estimator by using only the information that is above *K* (Which is not impacted by the hunger of bonus effect).

To determine this amount *K* we begin by expose a first simple approach of the bonus malus effect. In a second part, we will introduce a more elaborate one and use it for our sets of data.

7.1 *First approach*

First Model:

Let us consider a one period time, from year 0 to year 1.

Assumption 1 : we consider claim amounts above the deductible
Assumption 2 : when there is an accident, it is always with a total liability

P_0 : Initial premium
S_0 : Amount of claim in year 0
P_1 : Premium in year 1 if the claim is reported
P_1' : Premium in year 1 if the claim is not reported
D : Deductible

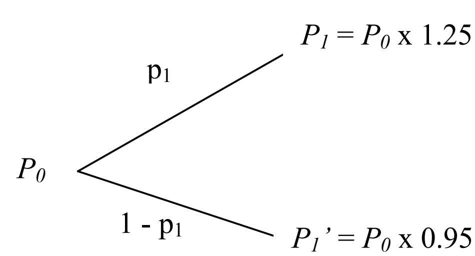
We consider a claim will be reported if: $S_0 + P_1' > D + P_1$ (1)

$$\Leftrightarrow S_0 > D + P_1 - P_1'$$

We have: $P_1 = P_0 \times 1.25$ and $P_1' = P_0 \times 0.95$

Thus, (1) $\Leftrightarrow S_0 > D + P_0 \times (1.25\text{-}0.95)$

Eventually, a claim is reported if $S_0 > \alpha_0 = D + 0.3 \times P_0$
From year 0 to year 1 a claim above α_0 (threshold) is reported. We then have the following situation:



Let's describe p_1 and $1 - p_1$:

The premium in year 1 is P_1 if:

- There is a claim
  and
- The claim amount is above α_0

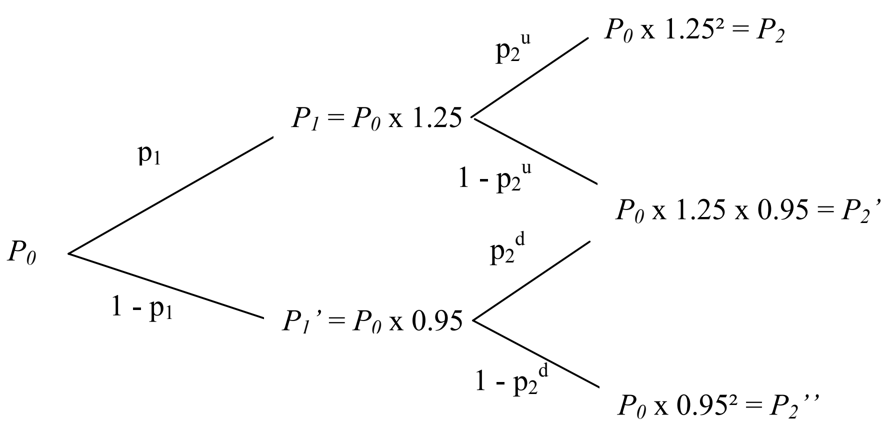
Thus p_1 = Pr[there is a claim | Previous premium is P_0] x Pr[$S_0 > \alpha_0$]

The premium in year 1 is $P1'$ if :

- There is no claim
  or
- There is a claim **and** the claim amount is below α_0

Thus $1 - p_1$ =
Pr[there is no claim | Previous premium is P_0] + Pr[there is a claim| Previous premium is P_0] x Pr[$S_0 < \alpha_0$]

Now let's consider an agent from period 1 to period 2 :



Let's now describe the probabilities:

- p_2^u = Pr[there is a claim | Previous premium is P_1] x Pr[$S_1 > \alpha_1^u$]

- $1 - p_2^u$ = Pr[there is no claim | Previous premium is P_1]
  + Pr[there is a claim| Previous premium is P_1] x Pr[$S_1 < \alpha_1^u$]

- p_2^d = Pr[there is a claim | Previous premium is P_1'] x Pr[$S_1 > \alpha_1^d$]

- $1 - p_2^d$ = Pr[there is no claim | Previous premium is S_1']
  + Pr[there is a claim| Previous premium is S_1'] x Pr[$S_1 < \alpha_1^d$]

With $\alpha_1^u = D + 0.3 \times P_1$ and $\alpha_1^d = D + 0.3 \times P_1'$

Let's now say we have three types of drivers: good, neutral and bad.

Assuming the insurer has no history of claims, we have to compute the probabilities since they will not be given by the transition matrix in that case.
Let's say we have the following transition matrix:

	Bad	Neutral	Good
Bad	0.7	0.3	0
Neutral	0.5	0	0.5
Good	0	0.3	0.7

Let us now set the probabilities of having an accident depending on the state of the driver:

Pr[Claim | Bad driver] = 0.7
Pr[Claim | Neutral driver] = 0.5
Pr[Claim | Good driver] = 0.3

Finally, in order to know how good a driver is depending on the amount of premium he has paid the previous year, let's assume the following:

Current period	Premium of previous period	Status of driver
1	P_0	Neutral
2	P_1	Bad
	P_1'	Good
3	P_2	Bad
	P_2'	Neutral
	P_2''	Good

Example:

- If we consider that the insurer has the actual distribution for the frequency and severity of claims. It is therefore necessary to calculate the different probabilities since they are not given by the transition matrix in that example.

- p_2^u = Pr[there is a claim | Previous premium is P_1] x Pr[$S_1 > \alpha_1^u$]

- If the previous premium is P_1, we see in the table above that the driver is considered a bad driver, then :

- $p_2^u = \dfrac{\text{Pr[there is a claim and Previous premium is P1]}}{\text{Pr[Previous premium is P1]}}$ x Pr[$S_1 > \alpha_1^u$]

And, with the tables above, we are able to compute the different tresholds.

However, in this model, the policyholder does not take into account the impact of his decision on the future years. Thus (1) is not a sufficient condition to report a claim.
We should therefore consider a model which includes the impact of the decision on the rest of the tree.

Second Model

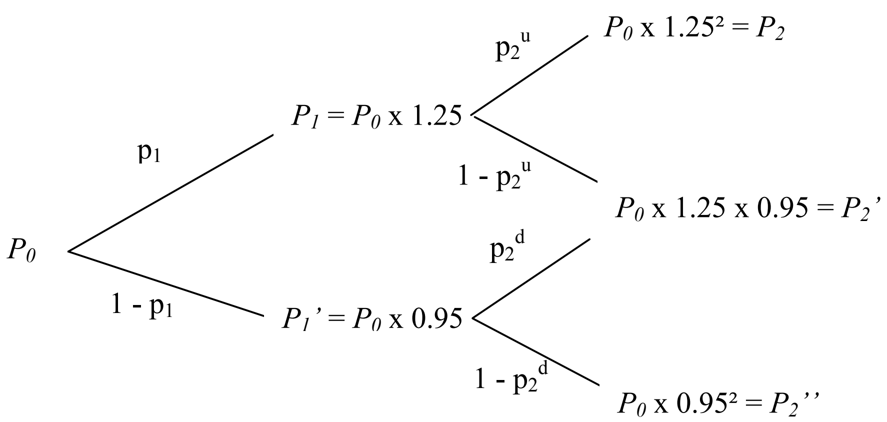
At time 0 let's rewrite condition (1) : We consider a claim will be reported if

$$S_0 + E_1 > D + E_2$$

with E_1 : Expected future premiums when claim not reported
E_2 : Expected future premiums when claim reported

$$(1) \Leftrightarrow S_0 > D + E_2 - E_1 = \alpha_0$$

Let's calculate E_1 and E_2 for a 2 year-period.



$$E_1 = \frac{P1'}{1+i} + \frac{1}{(1+i)^2} \times [P_2' \times p_2^d + (1 - p_2^d) \times P_2'']$$

$$E_2 = \frac{P1}{1+i} + \frac{1}{(1+i)^2} \times [P_2 \times p_2^u + (1 - p_2^u) \times P_2']$$

Here, let's say the insurer has got the history of past claims and that is transition matrix is not an assumption but is given by the history of policyholders he has in his portfolio. In that case, p_2^u and p_2^d are given by the transition matrix. Then, α_0 is totally known. In time 1, we will consider the same assumptions and obtain α_1. With that model, we can therefore compute all thresholds for all periods. We shall explain it in the next section.

7.2 *Formal approach*

We consider a bonus-malus system with n states. Let $\boldsymbol{T}$ be the matrix transition between states. $\boldsymbol{T} = [p_{ij}]$, where p_{ij} = Pr(be in state j in year t+1 / was in state i in year t). We assume a reasonable situation in which each state is recurrent. It means that a policyholder in state i could be, after several years in state j, with a positive probability, and this for all states i and j. Furthermore, we classify states: the better the driver is, the smaller the index of his state is.

Let $\boldsymbol{\beta} = [\beta_{\mathrm{i}}]_{\mathrm{i}=1,\ldots n}$ be the price's coefficients vector and P be the premium. A policyholder in state i will pay a premium of $\beta_{\mathrm{i}}P$. For example, coefficients may be greater than one if the policyholder is a bad conductor (i close to n) and less than one if he is good (i close to 1). We assume that one or more claims declared make the policyholder move towards the state just above, and that no claim declared make him move towards the state just below. It is clear that if the policyholder is in state 1 (or n) and has to move below (or above) this state, he simply won't move. Moreover, we assume that the transition matrix stays the same year after year. This is a strong hypothesis because we don't take into account the past historic of policyholders.

Now we consider a policyholder in state 3 for example. He is confronted to an accident with a cost of S. He has two possibilities. First, declare this accident to his insurance company. In this case, he will only pay a deductible of d. But he will move to the state 4 for the next year. Second, he does not declare this accident and pays. As a consequence, he will move to the "better" state 2. The policyholder will declare a claim if, and only if:

$$d + \text{cost of being in state 4 rather than being in state 2} < S$$

Without loosing generality we assume that β_1 is 1. And therefore the better driver will pay a premium of P. We will now speaking in terms of cost of being in a state i rather than in state 1. In each state the driver may make economy if he moves toward a better state: he will pay less. So being in a state i have a cost in comparison to the state 1. Let C_i this cost. The policyholder will declare a claim if:

$$d + C_4 < S + C_2$$

Now, we have to determine $\boldsymbol{C} = [C_i]$:

Let $\boldsymbol{P_t} = [P_{\mathrm{i}}]_{\mathrm{i}=1,\ldots n}$ be the vector of premiums P_i paid at time t for a policyholder in state i.

$$\boldsymbol{P_t} = \boldsymbol{\beta}'P$$

$$\mathrm{E}[\boldsymbol{P_{t+1}}] = \boldsymbol{T}.\boldsymbol{\beta}'P, \qquad \mathrm{E}[\boldsymbol{P_{t+2}}] = \boldsymbol{T}^2.\boldsymbol{\beta}'P, \qquad \ldots$$

Let r be the actuarial rate (annual).

$\boldsymbol{I} = (0, 0, \dots, 1, \dots, 0)$: size n with a 1 in position i.

$\boldsymbol{Id} = (1, \dots, 1)$: size n with 1 everywhere.

➔ The cost of being in state i rather than being in state 1 is the sum of the present values of the difference between future premiums paid in state i and future premiums paid in state 1. (Expected values).

The cost of being in state i rather than being in state 1 is:

$$C_i = \boldsymbol{I}\,.\,\boldsymbol{C}$$

$$\boldsymbol{C} = (\beta' - \boldsymbol{Id}')\,.\,P + \boldsymbol{T}.\,(\beta' - \boldsymbol{Id}')\,.\,P.\,(1+r)^{-1} + \boldsymbol{T}^2.\,(\beta' - \boldsymbol{Id}')\,.\,P.\,(1+r)^{-2} + \dots$$

To compute an example with Matlab, we assume that we have only tree state with the following transition matrix:

	Good	Neutral	Bad
Good	0.7	0.3	0
Neutral	0.5	0	0.5
Bad	0	0.1	0.9

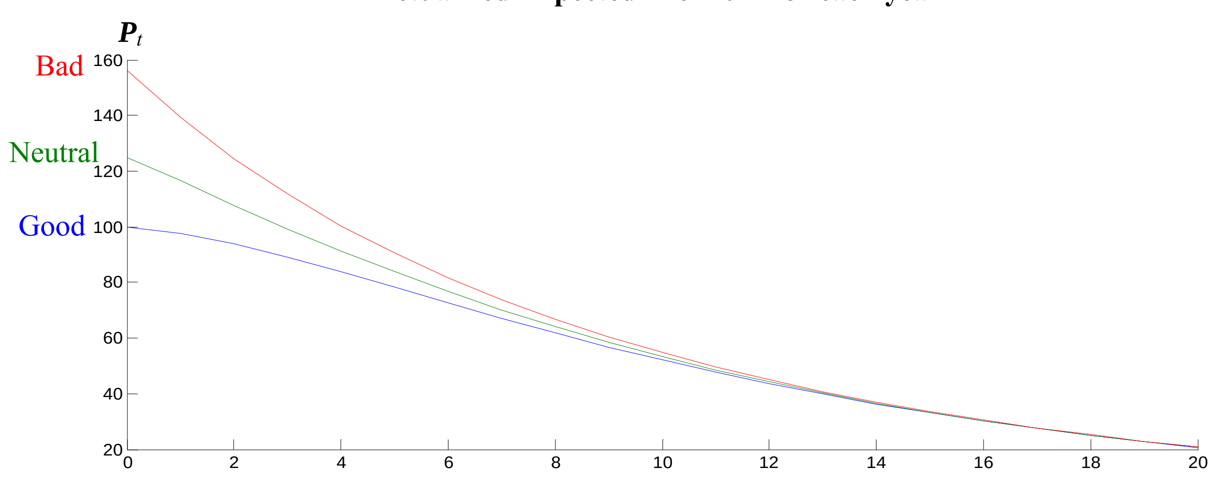
In that model, a neutral driver may become a good or a bad one with equal probabilities. A good driver has a high chance to stay good, and a bad one has a high probability to stay bad.

Moreover, we suppose:

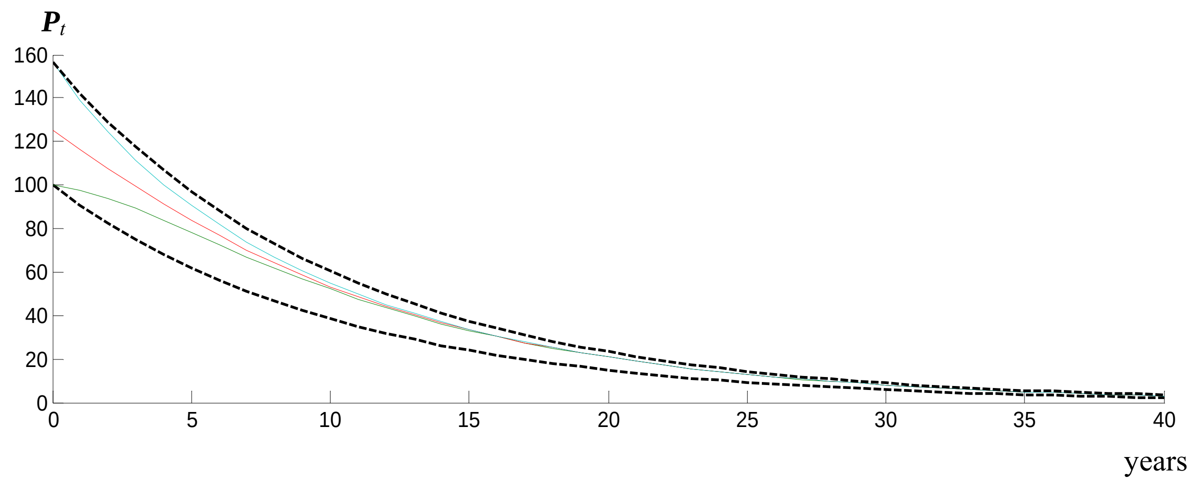
$P = 100$, $\beta = [1, 1.25, 1.25^2]$, $r = 0.1$, Deductible = 1000

We have drawn the actualized value of the expected premium for each year: $\boldsymbol{P_t}\,(1+r)^{-t}$. Beginning in year zero, a bad driver pays $1.23^2\,P$ and Good one P. One year after, the bad driver expects to pay $0.1*1.23\,P + 0.9*1.23^2\,P$ and so on. Obviously, whatever the year, a Good driver expects to pay less than a neutral one and the Neutral driver expects less than the Bad one.

Actualized Expected Premium for each year



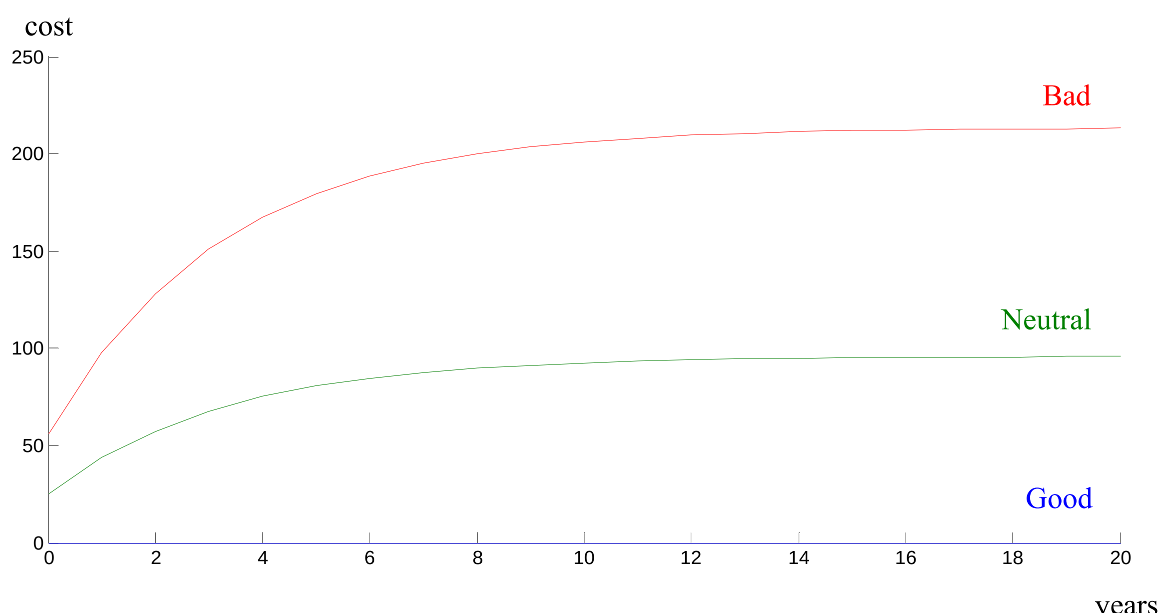
Due to actualization, all expected premiums reach zero after 20 years. For each year, a good driver expects to pay less than a neutral one, and so on. One may observe that there is no difference of being a good or a bad driver for expected premium in 15 years. This is of course a direct consequence of the assumption of lack of memory in the transition matrix. In a more realistic model, if we take into account all the past of drivers, all expected premium will also reach zero due to actualization. We have represented the evolution of expected premiums as if a driver stays always in a good state or in a bad state. In consequence, whatever the choice of modelisation for the transition matrix, we are sure that those expected premiums are always between these two curves. This chart shows that expected premium will reach zero after 40 years.



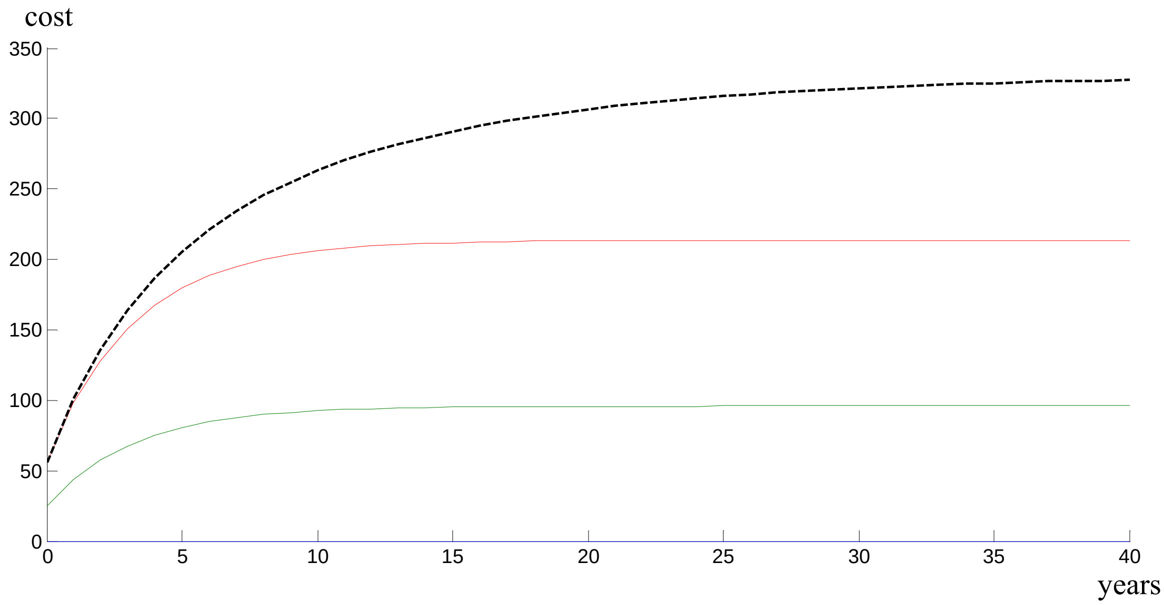
If we return to our model, we are interesting by compute $\boldsymbol{C} = [C_i]$. We have shown that $\boldsymbol{C}$ is an infinite sum. Due to previous charts, we expect convergence after 15 years. We have drawn the evolution of $\boldsymbol{C}$ against the number of years which were taken into account for the sum:

year 0 ➔ $(\beta' - \mathbf{Id}')\,.\,P$
year 1 ➔ $(\beta' - \mathbf{Id}')\,.\,P + \mathbf{T}.\,(\beta' - \mathbf{Id}')\,.\,P.\,(1+r)^{-1}$
year 2 ➔ ….

C



Of course for a good driver, this cost is zero. We can see that cost for bad and neutral driver converge after 15 years. If a bad driver planes to quit his insurance policy in the next year, being bad driver than rather a good one cost him 56. If he planes to keep his insurance as long as possible, it will cost 213. We have a strong assumption on the transition matrix. Like before, it is possible to determine a maximum for this cost whatever the model for transition is. We just draw the cost of stay in a bad state all the time.



Of course, it takes longer for having convergence; moreover, we are able to compute this exact maximum:

If we assume that a policyholder in the worst state has to pay a premium of βP, and the one in the better state pays a premium of P, then the cost of stay for all years in the worst state rather than the better one is difference of all actualized premiums:

$$C_{\max} = \beta P \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} - P \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} = \frac{P}{r}(\beta - 1)$$

In our example, C_{max} = 15625.5.

Now we are able to determine ***K*** for each state. We define:

C_G : cost of being in state G = Good driver rather than in G (obviously this cost is zero)
C_N : cost of being in state N = Neutral driver rather than in state G
C_B : cost of being in state B = Bad driver rather than in state G

C_{up}, C_{down} : cost of the state just above or below the actual state. This cost depends of the actual state:

if a driver is Bad, $C_{up} = C_N$, $C_{down} = C_B$
if a driver is Neutral, $C_{up} = C_G$, $C_{down} = C_B$
if a driver is Good, $C_{up} = C_G$, $C_{down} = C_B$

Using this notation, the policyholder report a claim S if $S + C_{down} \geq d + C_{up}$. For example, for a bad driver, he will report a claim if the amount of the claim plus the cost of staying in the bad state is more than pay the deductible d and the cost of being in state Neutral.

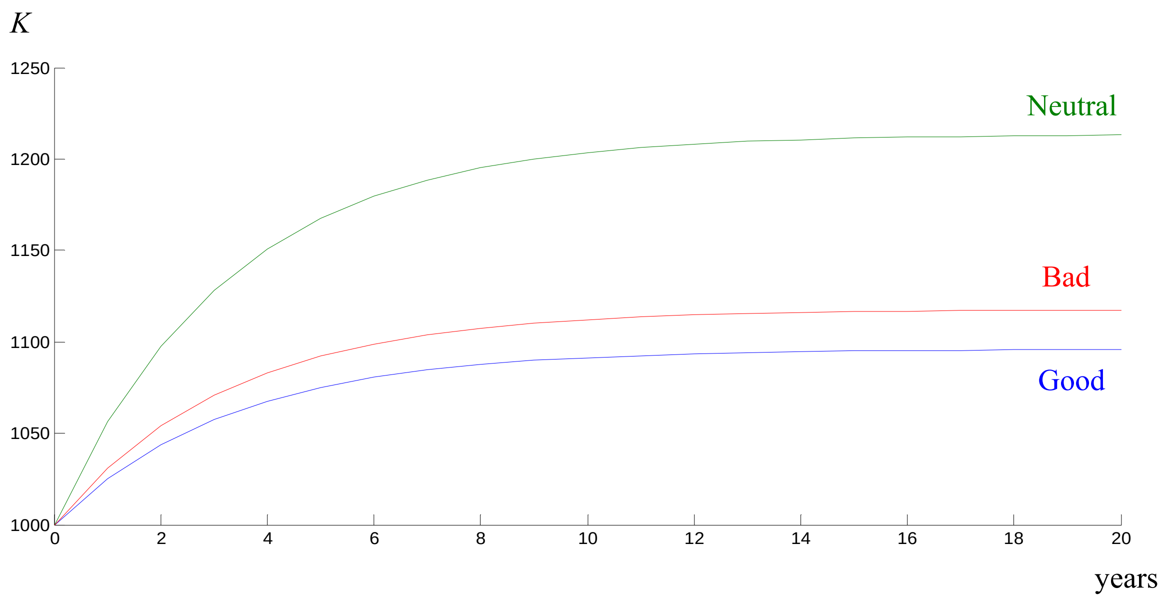
Thus,

$$\boldsymbol{K} = d\,\boldsymbol{Id} + \boldsymbol{C_{up}}' - \boldsymbol{C_{down}}'$$

Where $\boldsymbol{C_{up}} = [C_G, C_G, C_N]$ and $\boldsymbol{C_{down}} = [C_N, C_B, C_B]$

If we return to our example, we are now able to draw the value of $\boldsymbol{K}$ against the number of years the policyholder planes to keep his insurance.

Evolution of *K*



As expected, all thresholds are stable with more than 15 years of projection. It may be surprising to see that the highest threshold is for the Neutral State. In our example, the deal for the Neutral driver is quite important. On one hand, if he reports a claim, he will pass to the Bad state, and has a high probability to stay in this state, and on the other hand, if he does not report a claim, he will pass to the Good state and also will have a high chance to stay in this position. Using 10 years of projection, we have the following results:

$K_{\text{Good}} = 1089$
$K_{\text{Neutral}} = 1200$
$K_{\text{Bad}} = 110.$

If the driver will not take insurance the next year, it is obvious that the bonus malus has no effect for him, and he will report all claims that are above the deductible. Now, in our example, we are sure that all claims above 1200 are reported.

7.3 *Data analysis*

We will now use our previous bonus malus model to analyze data 1 and 2. If we assume the previous transition matrix to be true, a deductible of 200, an initial premium of 1382 for data 1 and 3003 for data 2 and factors β = [0.95, 1, 1.25], we obtain:

Data 1:	Data 2:
K_{Good} = 792	K_{Good} = 1487
K_{Neutral} = 1734	K_{Neutral} = 3535
K_{Bad} = 1142	K_{Bad} = 2247

Of course these values are far from the reality. Premiums are really high; thus, a malus has a huge impact. We have decided to decrease the effect of bonus malus by changing factors: β = [0.98, 1, 1.02], we obtain:

Data 1:	Data 2:
K_{Good} = 290	K_{Good} = 400
K_{Neutral} = 400	K_{Neutral} = 620
K_{Bad} = 300	K_{Bad} = 430

We now only take into account values above 400 for data 1 and 620 for data2 to compute estimator for the Pareto distribution:

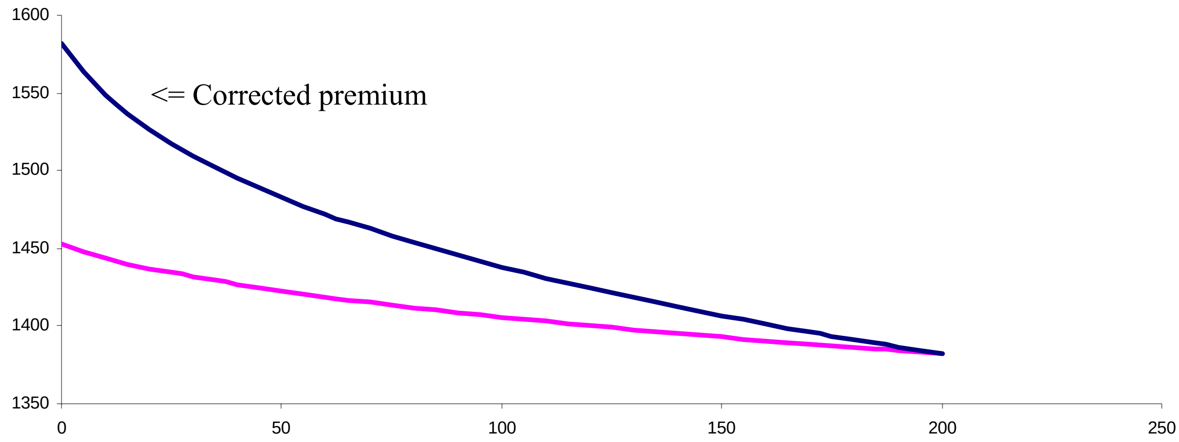
$\alpha_1 = 0.75$
$\alpha_2 = 0.93$

Those estimators are higher than in our first model. Indeed, due to the bonus malus system, we do not have all little claims in our data, and therefore, we underestimate the probability of having a small amount which means a smaller α for the Pareto distribution. On the other hand, we have underestimated the expected number of accident. As a result, there are two different effects:

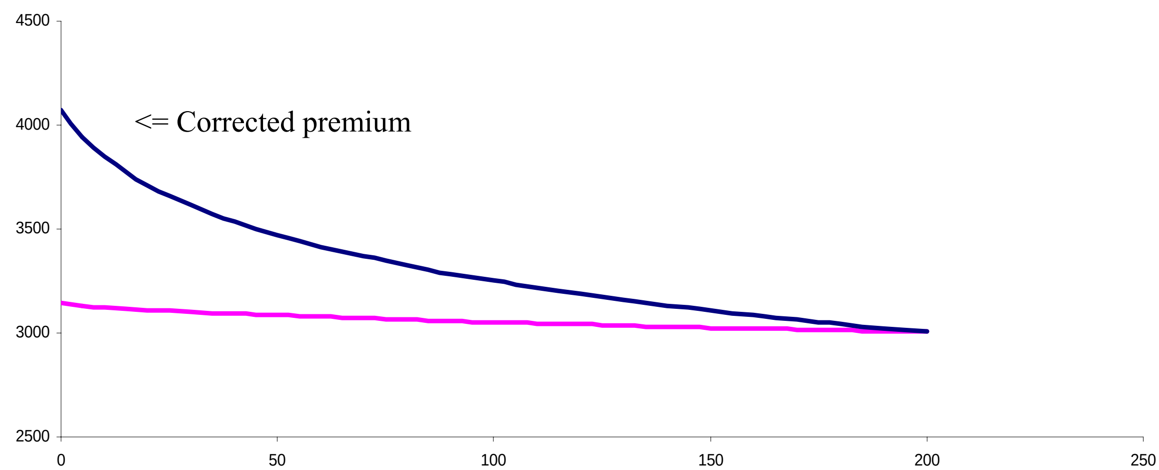
- decrease the expected severity
- increase the expected frequency

We do not have enough information to measure this impact on frequency. The insurer has to determine the expected number of accidents which are above K for each policy and use it to determine the premium. A prudential model may be compute by only take into account the frequency impact, and not the severity one. The insurer will then obtain a new amount of premium above the first one for each deductible. This is because we do not take into account the severity effect of the bonus malus. The new premium curve has to be adjusted in order to obtain the actual premium with the actual deductible.

Premium against deductible - Data 1



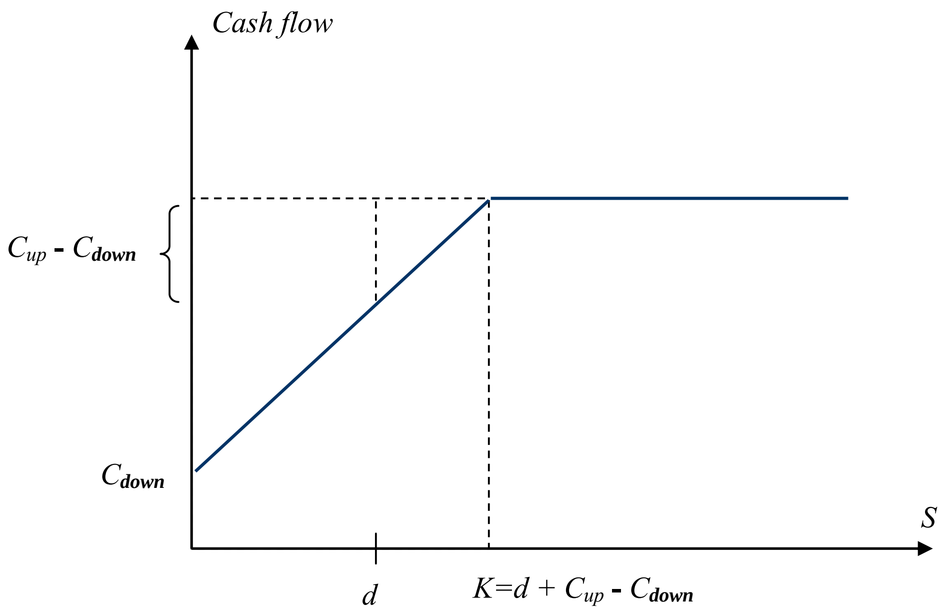
Premium against deductible - Data 2



We observe that, in both situations, the bonus malus system has an important impact on the premium. The insurer may lower the deductible very carefully. Indeed, any mistake may have huge consequences if the portfolio contains lots of policyholders.

7.4 Hunger of bonus, a financial option?

Using our previous notations, if an accident occurs, the policyholder will pay S + C if the amount of the accident is below K and $d + C_{up}$ if $S \geq K$. We obtain the following chart:



This situation seems to be the same as if the policyholder had sold a bond (C_{down}) and a put option with an exercise price of K. It would be interesting to develop a financial approach to this problem and try to find solution with option theory.

7.5 Appropriate Bonus Malus factors

If we assume that the insurer will pay for only one claim during the year, it is possible to find the correct factors β to applied. Assuming that premiums are paid at the beginning of the year, and claims at the end of the year, we have the following cash flows:

Accident: $S + C_{down}$ if $S \leq K$
$d + C_{up}$ if $S \geq K$

No accident: C_{down}

Premium: βP

We define:

Transition matrix:

	Good	Neutral	Bad
Good	P_{11}	P_{12}	0
Neutral	P_{21}	0	P_{23}
Bad	0	P_{32}	P_{33}

$q = \mathrm{E}[S \,/\, S \leq K]$
$\alpha = \Pr[S \geq K]$
P is the good driver premium

By equaling the premium and expected cash flows, we have:

$$1 \qquad P = C_G \; P_{11} + P_{12}\,[(\beta + C_G)\,(1\text{-}\alpha) + (d + C_N)\;\alpha]$$
$$2 \qquad P = C_G \; P_{21} + P_{23}\,[(\beta + C_G)\,(1\text{-}\alpha) + (d + C_B)\;\alpha]$$
$$3 \qquad P = C_G \; P_{32} + P_{33}\,[(\beta + C_N)\,(1\text{-}\alpha) + (d + C_B)\;\alpha]$$

1 and 2 give:

$$C_G \; P_{11} + P_{12}\,[(\beta + C_G)\,(1\text{-}\alpha) + (d + C_N)\;\alpha] = C_G \; P_{21} + P_{23}\,[(\beta + C_G)\,(1\text{-}\alpha) + (d + C_B)\;\alpha]$$

$\Rightarrow$ $(P_{12} - P_{23})\,(\beta\,(1\text{-}\alpha) + d\alpha\,) + P_{12}\; C_B\alpha = P_{23}\; C_B\;\alpha$

1 and 3 give:

$$C_G \; P_{11} + P_{12}\,[(\beta + C_G)\,(1\text{-}\alpha) + (d + C_N)\;\alpha] = C_G \; P_{32} + P_{33}\,[(\beta + C_N)\,(1\text{-}\alpha) + (d + C_B)\;\alpha]$$

$\Rightarrow$ $(P_{12} - P_{33})\,(\beta\,(1\text{-}\alpha) + d\alpha\,) + P_{12}\; C_B\;\alpha = C_G \; P_{32} + P_{33}\; C_N\,(1\text{-}\alpha) + P_{33}\; C_M\;\alpha$

Let $\Gamma = \beta\,(1\text{-}\alpha) + d\alpha$, this two equations yield:

$$C_B = \frac{\Gamma}{\dfrac{1-\alpha P_{33}}{P_{12}\left(P_{33}-P_{23}\right)} - \alpha}$$

Given a transition matrix, the insurer is able to compute appropriate factors by solving those equations.

Conclusion

We have seen that insurance companies have, for many years, decided to set deductibles to most of the policies. Thanks to the deductibles, insurers may at the same time provide an economic incentive to prevent losses, and have a risk-adjustment tool depending on the markets.

However, in spite of there being many advantages, deductibles should be carefully watched. As a matter of fact, when lowering a deductible in a hard market period, the insurer has to increase the premium asked to cover the risk and this might cause the policyholder to transfer to another insurance company.

In addition to that, if a deductible is decreased within a Bonus-malus system environment, there is a danger for the insurer which is the hunger for bonus. We have seen that there is an amount of loss under which no claim will be reported by a policyholder because it would increase too much his premium for the following year.

As a consequence, this makes it difficult for insurers to assess the real severity and distribution of the risks taken. Thus the insurance companies will try to determine the optimal threshold; that is the amount above which a driver will necessarily declare the claim.

Eventually, some authors suggest using a severity component in the Bonus-malus system, in addition to the frequency component already used. This would lead to an optimal Bonus-malus system for both the insurer and the policyholder.

Appendix

Appendix 1. Frequency distribution

We will introduce a special class of counting distribution: the (a,b,0) class. Counting distributions are quite essential in an insurance context. They describe the number of events; that is losses for the insured or claims made to the insurance company. A separate vision of the number of claims and the size of claims is more useful to have a better understanding of a variety of issues surrounding insurance than having only information about total losses. We will focus on the distribution from the (a,b,0) class : The Poisson, Negative Binomial and Binomial distributions. This class is very easy to use in practice and gives most of the time very good results in modelling situations. Using only one or two parameters, the dimensionality of the information is greatly reduced when using parametric distribution. This is a very important point when searching for robust models to smooth empirical data. In a second part, we will have a look at an interesting property of this class in use.

1 The (a,b,0) Class

Let N be a discrete positive random variable with probability function pk. N is a member of the (a,b,0) class of distributions if there exists constants a and b such that :

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k} \qquad k = 1, 2, 3, ...$$

The boundary condition, the value of p_0, can be obtained since the probabilities must add up to 1. It can be shown that the only counting distributions which satisfy this recursive formula are the Poisson, Binomial, Negative Binomial and Geometric distributions. Consequently we begin by having a look at the basic material for all these distributions.

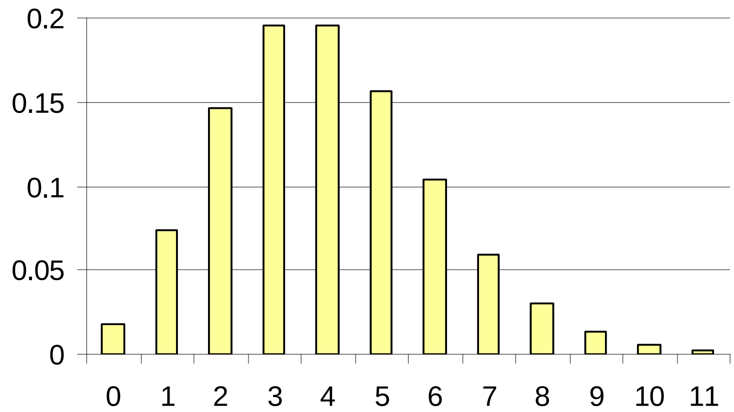
2 The Poisson Distribution - λ

2.1 Properties

$p_0 = e^{-\lambda}$, $a = 0$, $b = \lambda$, Distribution : $p_k = \dfrac{e^{-\lambda}\lambda^k}{k!}$

$E[N] = \lambda$, $Var[N] = \lambda$, Probability generating function : $P(z) = e^{\lambda(z-1)}$

Chart 4.1. Density λ = 4



It can be seen that for the Poisson distribution, the variance is equal to the mean. A first approach of smoothing empirical data consist in computing mean and variance and testing a Poisson model if they are relatively close. The Poisson distribution has some interesting properties that are very useful in modelling such as :

Theorem 1.1

Let (N_i), $i = 1$ to n have independent Poisson distribution with parameter λ_i.

Then $N = N_1 + \ldots + N_n$ is a Poisson variable with parameter $\lambda = \lambda_1 + \ldots + \lambda_n$.

Proof :

The probability generating function of the sum of independent random variables is the individual probability generating functions. The sum of Poisson random variables gives :

$$P_N(z) = \prod_{j=1}^{n} P_{N_j}(z) = \prod_{j=1}^{n} \exp\left[\lambda_j (z-1)\right]$$

$$= \exp\left[\sum_{j=1}^{n} \lambda_j (z-1)\right]$$

$$= e^{\lambda(z-1)}$$

The probability generating function is unique and therefore N follows a Poisson distribution with parameter λ.

Theorem 1.2

Let N be a Poisson variable with mean λ. Suppose that each event of N can be classified into one of m types with probabilities p1,…, pm independent of all other events and let Ni be the number of type i.

Then (Ni), i = 1 to m are mutually independent Poisson random variables with means λ.pi .

Proof :

The joint probability function of ($N_1 = n_1$, … , $N_m = n_m$) is given by

$$\Pr\left(N_1 = n_1,\ldots,N_m = n_m\right) = \Pr\left(N_1 = n_1,\ldots,N_m = n_m \middle| N = n\right) \times \Pr(N = n)$$

$$= \frac{n!}{n_1!n_2!\ldots n_m!} p_1^{n_1}\ldots p_m^{n_m} \frac{e^{-\lambda}\lambda^n}{n!}$$

$$= \prod_{j=1}^{m} e^{-\lambda p_j} \frac{\left(\lambda p_j\right)^{n_j}}{n_j!}$$

where $n = n_1 + n_2 +\ldots+ n_m$.

Furthermore, the marginal probability function of N_j is given by :

$$\Pr\left(N_j = n_j\right) = \sum_{n=n_j}^{\infty} \Pr\left(N_j = n_j \middle| N = n\right) \times \Pr(N = n)$$

$$= \sum_{n=n_j}^{\infty} \binom{n}{n_j} p_j^{n_j} (1 - p_j)^{n-n_j} \frac{e^{-\lambda}\lambda^n}{n!}$$

$$= e^{-\lambda} \frac{(\lambda p_j)^{n_j}}{n_j!} \sum_{n=n_j}^{\infty} \frac{\left[\lambda(1 - p_j)\right]^{n-n_j}}{(n - n_j)!}$$

$$= e^{-\lambda} \frac{(\lambda p_j)^{n_j}}{n_j!} e^{\lambda(1-p_j)}$$

$$= e^{-\lambda p_j} \frac{(\lambda p_j)^{n_j}}{n_j!}$$

Hence the joint probability function is the product of the marginal probability functions, establishing mutual independence.

This last proposition is very interesting. For example, we can easily differentiate claims which are above or below a limit like a deductible. In addition, it’s useful to remove or add a part of an insurance coverage.

2.2 Estimation of λ

For the Poisson distribution, the maximum likelihood gives the same result as the method of moments estimators. Let nk be the number of times in which a frequency of k events occurred. Thus, the likelihood and loglikelihood are :

$$L = \prod_{k=0}^{\infty} p_k^{n_k} \qquad\qquad l = \sum_{k=0}^{\infty} n_k \log p_k$$

With

$$\log p_k = -\lambda + k \log \lambda - \log k! \qquad \text{and} \qquad n = n_0 + n_1 + \ldots$$

we have

$$l = -\lambda n + \sum_{k=0}^{\infty} k\, n_k \log\lambda - \sum_{k=0}^{\infty} n_k \log k! \qquad \text{and} \qquad \frac{dl}{d\lambda} = -n + \sum_{k=0}^{\infty} k\, n_k \frac{1}{\lambda}$$

By setting the derivative of the loglikelihood to zero, the estimator of λ becomes :

$$\hat{\lambda} = \frac{\sum_{k=0}^{\infty} k\, n_k}{n} = \frac{\sum_{i=1}^{n} N_i}{n}$$

Consequently, the mean estimator and variance estimator are :

$$\mathrm{E}\left(\hat{\lambda}\right) = \mathrm{E}(N) = \lambda \qquad\qquad \mathrm{Var}\left(\hat{\lambda}\right) = \frac{\mathrm{Var}(N)}{n} = \frac{\lambda}{n}$$

Fisher's information :

$$I(\lambda) = -n\,\mathrm{E}\left[\frac{d^2}{d\lambda^2} \log p_N\right] = -n\,\mathrm{E}\left[\frac{d^2}{d\lambda^2}(-\lambda + N\log\lambda - \log N!)\right] = n\,\mathrm{E}\left(\frac{N}{\lambda^2}\right) \qquad I(\lambda)^{-1} = \lambda / n$$

Thus, the maximum likelihood estimator is asymptotically normally distributed with mean λ and variance λ /n . Hence the (1-α)% confidence interval for the true value of the parameter is :

$$\left[\hat{\lambda} - \Phi(\alpha/2)\sqrt{\hat{\lambda}/n}\ ; \lambda + \Phi(\alpha/2)\sqrt{\hat{\lambda}/n}\right]$$

We can test formally the distribution by using a classic chi-square test statistic :

$$Q = \sum_{k=0}^{\infty} \frac{\left(n_k - n\Pr\left(N = k, \hat{\theta}\right)\right)^2}{n \Pr\left(N = k, \hat{\theta}\right)}$$

3 The Negative Binomial distribution – r,β

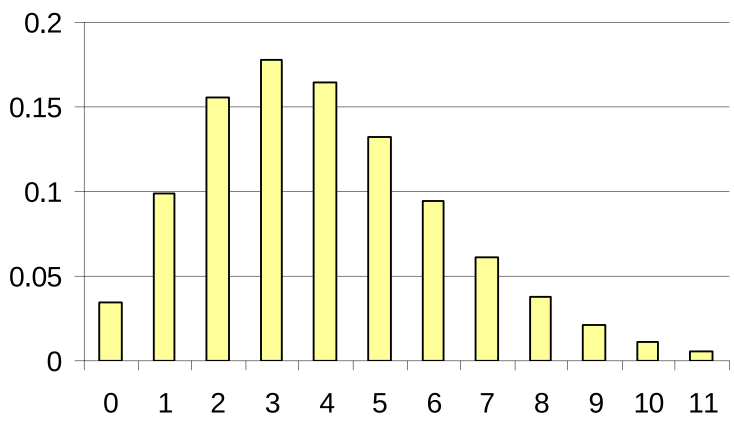
3.1 Properties

$p_0 = (1+\beta)^{-r}$, $\quad a = \dfrac{\beta}{1+\beta}$, $\quad b = (r-1)\dfrac{\beta}{1+\beta}$, $\quad$ Distribution :

$$p_k = \binom{k+r-1}{k}\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1+\beta}\right)^k$$

$E[N] = r\beta$, $\quad Var[N] = r\beta(1+\beta)$, $\quad$ Probability generating function : $P(z) = [1 - \beta(z-1)]^{-r}$

Chart 4.2. Density r = 10, β = 0.4



The negative binomial distribution has two parameters, consequently, it has more flexibility than the Poisson. Moreover, for a particular situation, if observed variance is larger than observed mean, then the negative binomial might be a better choice than the Poisson distribution.

The geometric distribution is a special case of the negative binomial distribution when r = 1. When r < 1, the negative binomial's tail decays more slowly than the geometric one, the opposite situation is observable when r > 1.

It can be seen that for the Negative Binomial distribution, the variance exceeds the mean. This information suggests that for a particular set of data, if the observed variance is larger than the observed mean, the Negative Binomial might be a better candidate than the Poisson distribution to fit the distribution of the set of data.

One way to generate a negative binomial is as a mixture of Poisson distributions :

Let Λ be a gamma distribution (α,θ) and f its probability function.

$$p_k = E[\Pr(N = k|\Lambda)] = \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!} f(\lambda)\, d\lambda$$

$$\mathrm{f}(\lambda) = \frac{\lambda^{\alpha-1} e^{-\frac{\lambda}{\theta}}}{\theta^{\alpha}\Gamma(\alpha)} \quad \Rightarrow \quad p_k = \frac{1}{k!}\frac{1}{\theta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} e^{-\lambda(1+\frac{1}{\theta})} \lambda^{k+\alpha-1}\,\mathrm{d}\,\lambda$$

$$\text{Let } \lambda' = \lambda(1+\frac{1}{\theta}) \quad \Rightarrow \quad p_k = \frac{1}{k!}\frac{1}{\theta^{\alpha}\Gamma(\alpha)}\frac{\theta^{k+\alpha}}{(1+\theta)^{k+\alpha}}\Gamma(k+\alpha)$$

$$\text{Thus} \qquad p_k = \binom{k+r-1}{k}\left(\frac{1}{1+\beta}\right)^{r}\left(\frac{\beta}{1+\beta}\right)^{k}, \quad r = \alpha,\ \beta = \theta$$

3.2 Estimation of r, β

The method of moments gives :

$$r\beta = \frac{\sum_{k=0}^{\infty} k\, n_k}{n} = \frac{\sum_{i=1}^{n} N_i}{n}$$

$$r\beta\,(1+\beta) = \frac{\sum_{k=0}^{\infty} k^2\, n_k}{n} - \left(\frac{\sum_{k=0}^{\infty} k\, n_k}{n}\right)^2$$

The likelihood and loglikelihood are :

$$L = \prod_{k=0}^{\infty} p_k^{n_k} \qquad\qquad l = \sum_{k=0}^{\infty} n_k \log p_k$$

Thus,

$$l = \sum_{k=0}^{\infty} n_k \left[\log\binom{r+k-1}{k} - r\log(1+\beta) + k\log\beta - k\log(1+\beta)\right]$$

Hence,

$$\frac{\partial l}{\partial \beta} = \sum_{k=0}^{\infty} n_k \left(\frac{k}{\beta} - \frac{r+k}{1+\beta}\right)$$

$$\frac{\partial l}{\partial r} = -\sum_{k=0}^{\infty} n_k \log(1+\beta) + \sum_{k=0}^{\infty} n_k \frac{\partial}{\partial r}\log\frac{(r+k-1)...r}{k!}$$

$$= -n\log(1+\beta) + \sum_{k=0}^{\infty} n_k \frac{\partial}{\partial r}\sum_{m=0}^{k-1}\log(r+m)$$

Setting these equations to zero yields :

$$\hat{r}\,\hat{\beta} = \frac{\sum_{k=0}^{\infty} k\, n_k}{n} = \hat{\mu}$$

$$n\log(1+\beta) = \sum_{k=0}^{\infty} n_k \left(\sum_{m=0}^{k-1}\frac{1}{\hat{r}+m}\right)$$

Because $\hat{\mu}$ must be the simple mean μ, the maximisation problem is reduced to one dimension.

To test if the negative binomial is a good fit, we can compose the chi-square test. In order to determine if the negative binomial distribution is a significantly better fit than the Poisson, we can use the likelihood ratio-test as below :

H_0 : Poisson loglikelihood = l_0

H_1 : Negative binomial loglikelihood = l_1

The test statistic with a chi-square distribution and one degree of freedom is :

$$Q = 2\,(l_1 - l_0)$$

The degree of freedom is the number of restriction from the negative binomial model to the Poisson model.

4 The Binomial Distribution - q , m

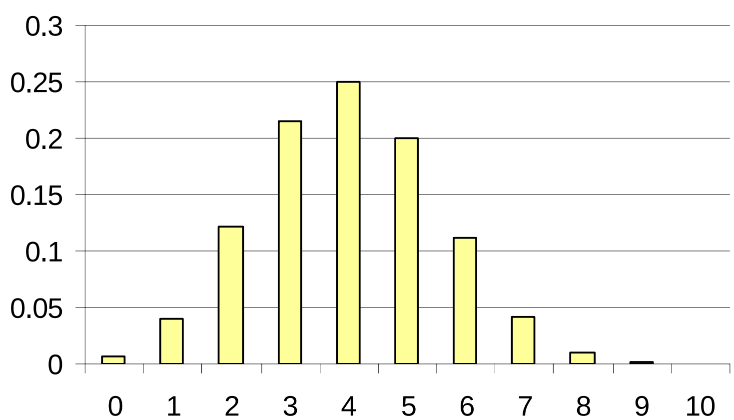
4.1 Properties

$$p_0 = (1-q)^m, \qquad a = -q/(1-q), \qquad b = (m+1)q/(1-q),$$

$$\text{Distribution : } p_k = \binom{m}{k} q^k (1-q)^{m-k},\ k = 0,.., m$$

$E[N] = mq$, Var$[N] = mq(1-q)$, Probability generating function : $P(z)=[1- q(z-1)]^m$

Chart 4.3. Density m = 10, r = 0.4



The range of values with positive probabilities has finite length. For example, this is useful in modelling the number of individuals injured in an automobile accident.

It can be seen that for the Binomial distribution, the mean exceeds the variance. Hence, for a particular set of data, if the observed mean is larger than the observed variance, the Binomial model might be a good candidate to fit the distribution.

4.2 Estimation of m, q

Firstly, we assume that m is known. The only one parameter q needs to be estimated. Thus, q represents the probability of some event and is estimated with the method of moments:

$$\hat{q} = \frac{\text{Number of observed events}}{\text{Maximum number of possible events}}$$

The loglikelihood is

$$l = \sum_{k=0}^{m} n_k \log p_k$$

$$l = \sum_{k=0}^{m} n_k \left[\log\binom{m}{k} + k\log q + (m-k)\log(1-q) \right]$$

$$\frac{\partial l}{\partial q} = \frac{1}{q}\sum_{k=0}^{m} kn_k - \frac{1}{1-q}\sum_{k=0}^{m}(m-k)n_k$$

Setting this equal to zero yields

$$\hat{q} = \frac{1}{m}\frac{\sum_{k=0}^{m} kn_k}{\sum_{k=0}^{m} n_k}$$

Hence, the maximum likelihood gives the same result as the method of moments estimators when m is fixed.

When *m* is unknown, the maximum likelihood estimator of *q* is $\hat{q} = \frac{1}{\hat{m}}\frac{\sum_{k=0}^{m} kn_k}{\sum_{k=0}^{m} n_k}$

To find this estimator, we starting by fixing m to the largest observation, then compute q and calculate the loglikelihood at these values. After, we increase m by one and so on until a maximum is found.

5 The (a,b,0) Class in use

The values of a, b and p_0 for the distribution detailed above are summarized in Table 1.1 :

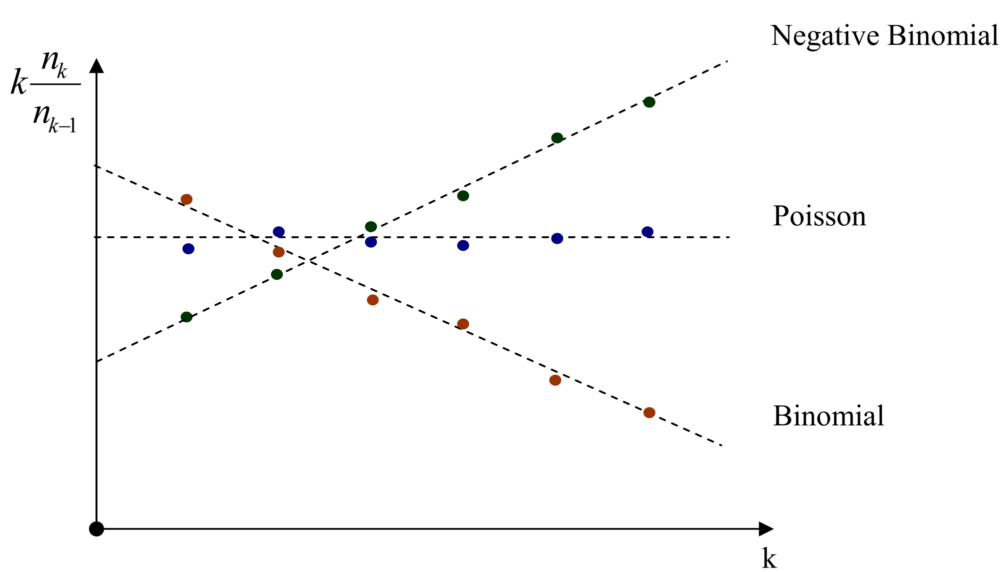
Table 1.1 Members of the (*a*, *b*, 0) class

Distribution	a	b	p_0
Poisson	0	λ	$e^{-\lambda}$
Binomial	$-\dfrac{q}{1-q}$	$(m+1)\dfrac{q}{1-q}$	$(1-q)^m$
Negative binomial	$\dfrac{\beta}{1+\beta}$	$(r-1)\dfrac{\beta}{1+\beta}$	$(1+\beta)^{-r}$
Geometric	$\dfrac{\beta}{1+\beta}$	0	$(1+\beta)^{-1}$

It is interesting to see that a graphical analysis will indicate which of the distributions should be selected for fitting. Indeed, the recursive formula can be rewritten as :

$$k\frac{p_k}{p_{k-1}} = ak + b \qquad k = 1, 2, 3, \ldots$$

Let n_k be the number of events by time k. One can plot $k\dfrac{n_k}{n_{k-1}}$ against k. If one of these models is to be selected, the observed values should form a straight line and the value of the slope indicate the right model : 0 for the Poisson, negative for the Binomial and positive for the Negative binomial.



Appendix 2. Severity Distributions

Using notations:

$\alpha > 0$, $x > 0$, $a > 0$, $b > 0$

$$\Gamma(\alpha;x) = \frac{1}{\Gamma(\alpha)}\int_0^x t^{\alpha-1}\,\mathrm{e}^{-t}\,dt$$

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\,\mathrm{e}^{-t}\,dt$$

$$\beta(a,b;x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^x t^{a-1}(1-t)^{b-1}\,dt \text{ with } 0 < x < 1$$

1 The Normal Distribution - μ, σ

1.1 Properties

A random variable X is called normal with parameter μ and σ if its distribution function is given by :

$$\Pr[X \le x] = \Phi\left(\frac{x-\mu}{\sigma}\right) \text{ where } \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-y^2/2)dy,\ x \in \mathrm{IR}$$

$$E[X] = \mu \qquad Var[X] = \sigma$$

Theorem : Let X be N (μ_1, σ_1) let Y be N (μ_2, σ_2), two independent random variables,

We then have :

$$X + Y \text{ is } N(\mu_1+\mu_2,\ \sqrt{\sigma_1^{\,2}+\sigma_2^{\,2}})$$

1.2 Estimation of Parameters

Assuming we have N observations : $X_1,\ldots,X_N$

$$l = \frac{1}{\sigma^N}\frac{1}{\sqrt{2\pi}^{\,N}}\exp\left[-\frac{1}{2}\sum_{i=1}^{N}\left(\frac{Xi-\mu}{\sigma}\right)^2\right]$$

$$L = -N\ln\sigma - \frac{1}{2}\sum_{i=1}^{N}\left(\frac{Xi-\mu}{\sigma}\right)^2$$

$$\frac{dL}{d\mu} = -\sigma \sum_{i=1}^{N} \left(\frac{Xi - \mu}{\sigma} \right) \qquad \text{and} \qquad \frac{dL}{d\mu} = 0 \Rightarrow \qquad \hat{\mu} = \frac{1}{\mu} \sum_{i=1}^{N} Xi$$

$$\frac{dL}{d\sigma} = -\frac{N}{\sigma} + \sum_{i=1}^{N} \frac{(Xi - \mu)^2}{\sigma^3} \qquad \text{and} \qquad \frac{dL}{d\sigma} = 0 \Rightarrow \qquad \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{N} (Xi - \hat{\mu})^2$$

2 The Exponential Distribution - λ

2.1 Properties

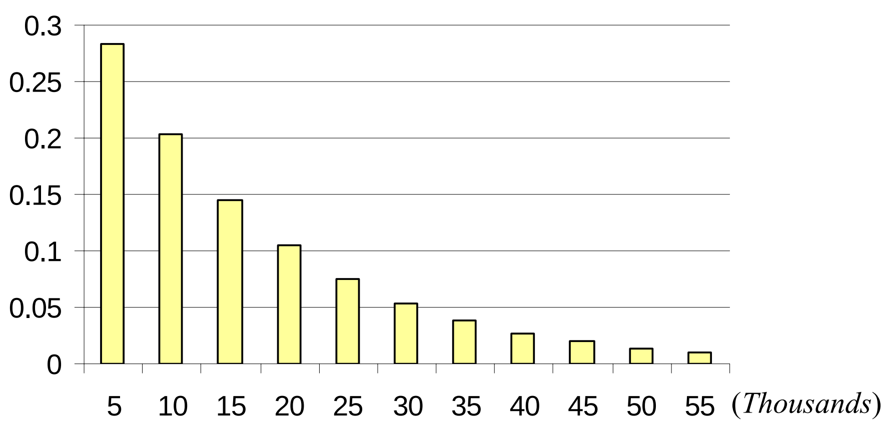
A continuous random variable X is called exponential with parameter $\lambda > 0$ if its density function is given by :

$$f(t) = \lambda \exp(-\lambda t) \text{ for } t \geq 0 \text{, and} \qquad F_X(t) = 1 - \exp(-\lambda t)$$

$$E[X] = 1/\lambda \qquad Var[X] = 1/\lambda^2 \qquad E[X^k] = \theta^k k! \text{, if k is an integer}$$

$$E\left[(X \wedge x)^k\right] = \lambda^{-k} k! \Gamma(k+1; x\lambda) + x^k e^{-x\lambda} \text{ if } k > -1 \text{ is an integer}$$

Chart 5.1 Density, 1/λ = 15000



The times between two realisations of a Poisson distribution follows an exponential distribution.

Example : The interarrival time between two customers at a post office is a random variable that might be exponential.

Let's keep in mind that the exponential distribution represents claim amounts that are relatively not dangerous for an insurance company since the survival function decreases exponentially (thin distribution tail).

2.2 Estimation of Parameters

Assuming we have N observations : $X_1,\ldots,X_N$

$$L = \lambda^n \exp(-\lambda \sum_{i=1}^{N} Xi) \qquad l = N \ln \lambda - \lambda \sum_{i=1}^{N} Xi \qquad \frac{dl}{d\lambda} = \frac{N}{\lambda} - \sum_{i=1}^{N} Xi$$

We then set the derivative of the loglikelihood to zero to obtain : $\hat{\lambda} = \dfrac{N}{\sum_{i=1}^{N} Xi}$

3 The Pareto Distribution – α,θ

3.1 Properties

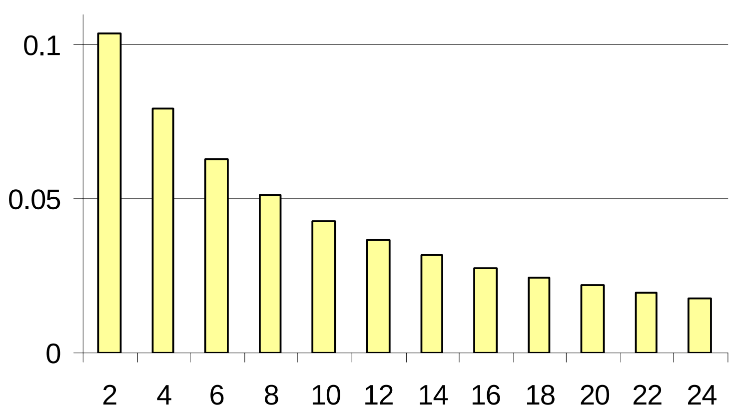
$$F_X(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha} \qquad f(x) = \frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}}$$

$$E[X] = \frac{\theta}{\alpha-1} \qquad E[X^k] = \frac{\theta^k k!}{(\alpha-1)\ldots(\alpha-k)} \quad \text{for } k \text{ integer}$$

$$E\left[(X \wedge x)^k\right] = \frac{\theta^k \Gamma(k+1)\Gamma(\alpha-k)}{\Gamma(\alpha)} \beta[k+1, \alpha-k; x/(x+\theta)] + x^k \left(\frac{\theta}{x+\theta}\right)^{\alpha}$$

Note that only α is the true parameter. The value of θ must be set in advance.

Chart 5.2. Density, $\alpha = 0.6$, $\theta = 10$



1.2. Estimation of Parameters

Assuming we have N observations : $X_1,\ldots,X_N$

$$L = \frac{(\alpha\theta^{\alpha})^N}{\prod_{i=1}^{N}(Xi+\theta)^{\alpha+1}} \qquad l = \ln(\alpha^N\theta^{N\alpha}) - \sum_{i=1}^{N}\ln((Xi+\theta)^{\alpha+1})$$

$$\frac{dl}{d\alpha} = \frac{N}{\alpha} + N\ln(\theta) - \sum_{i=1}^{N}\ln(Xi+\theta)$$

We then set $\frac{dl}{d\alpha}=0$, thus we have : $\hat{\alpha} = \frac{N}{\sum_{i=1}^{N}\ln(Xi+\theta) - N\ln(\theta)}$

4 The Weibull - τ, θ

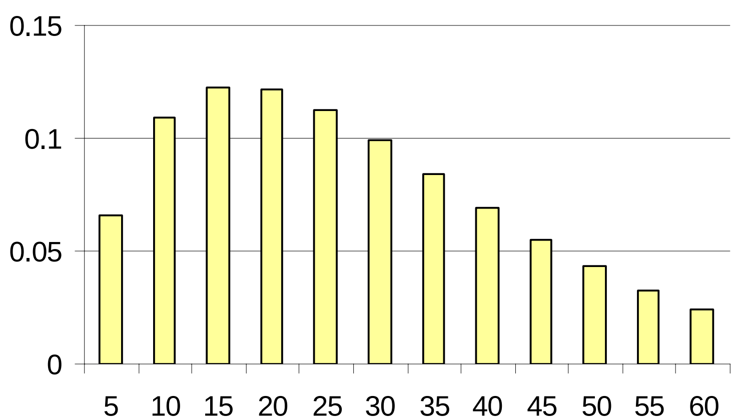
4.1 Properties

$$F(x) = 1 - e^{-(x/\theta)^{\tau}} \qquad f(x) = \frac{\tau(x/\theta)^{\tau}\, e^{-(x/\theta)^{\tau}}}{x}$$

$$E[X] = \theta\,\Gamma(1+1/\tau) \qquad E[X^k] = \theta^k\,\Gamma(1+k/\tau)$$

$$E\left[(X\wedge x)^k\right] = \theta^k\Gamma(1+k/\tau)\Gamma(1+k/\tau;(x/\theta)^{\tau}) + x^k\exp(-(x/\theta)^{\tau})$$

Chart 5.3. Density τ=1.5 , θ=30000



We can note that if X has a Weibull distribution with parameters τ and θ, and if $Z = X^{\tau}$, then Z has an exponential distribution with mean θ^{τ}.

4.2 *Estimation of Parameters*

Assuming we have N observations : $X_1,\ldots,X_N$, and τ is given

$$L = \frac{\tau^N \prod_{i=1}^{N} (Xi^{\tau-1}) \exp\left(-\theta^{-\tau} \sum Xi^{\tau}\right)}{\theta^{N\tau}}$$

$$l = N\ln(\tau) + (\tau-1)\sum_{i=1}^{N} \ln(Xi) - \theta^{-\tau}\sum Xi^{\tau} - N\tau\ln(\theta)$$

$$\frac{dl}{d\theta} = \frac{\tau}{\theta^{\tau+1}}\sum_{i=1}^{N} Xi^{\tau} - \frac{N\tau}{\theta}$$

We then set $\frac{dl}{d\theta} = 0$, thus we have : $\hat{\theta} = \left(\frac{1}{N}\sum_{i=1}^{N} Xi^{\tau}\right)^{1/\tau}$

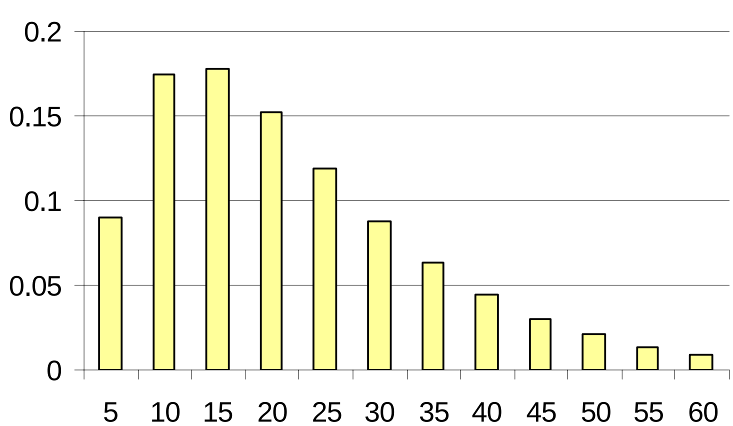
5 The Gamma distribution – α, τ

$$F(x) = \Gamma(\alpha; x\tau) \qquad f_X(x) = \begin{cases} \dfrac{x^{\alpha-1}\tau^{\alpha}\exp(-x\tau)}{\Gamma(\alpha)}, & \mathit{if} \quad x \geq 0, \\ 0, & \mathit{otherwise}. \end{cases}$$

$$\mathrm{E}[X] = \frac{1}{\tau}\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \qquad \mathrm{E}[X^k] = \tau^{-k}\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$$

$$\mathrm{E}\left[(X \wedge x)^k\right] = \tau^{-k}\alpha(\alpha+1)\ldots(\alpha+k-1)\Gamma(\alpha+k; x\tau) + x^k\left[1-\Gamma(\alpha; x\tau)\right] \text{ if k is an integer}$$

Charte 5.3. Density α=2 τ=1/10000



This distribution is called standard gamma distribution when $\tau = 1$, then

$$f_X(x) = \frac{x^{\alpha-1}\exp(-x)}{\Gamma(\alpha)}, \qquad x \in \Re^+.$$

5.1 Properties

Remark 1

The Khi-Square distribution with n degrees of freedom is a special case of the gamma distribution when $\tau = 0{,}5$ and $\alpha = \dfrac{n}{2}$.

The Exponential distribution $\varepsilon(\tau)$ is a special case of the gamma distribution when $\alpha = 1$.

Remark 2

We obtain the Erlang distribution if $\alpha \in \mathrm{N}^*$ and we have:

$$F_X(x) = 1 - \sum_{j=0}^{\alpha-1}\exp(-x\tau)\frac{(x\tau)^j}{j!}, \qquad x \ge 0$$

Its Laplace Transform is $(1+\dfrac{t}{\tau})^{-\alpha}$

Its moment generating function is $(1-\dfrac{t}{\tau})^{-\alpha} \qquad si \quad t < \tau$

Remark 3

Let (X_i), $i = 1$ to n have independent Gamma distribution such $X_i \sim Gamma(\alpha_i, \tau)$

Then $X = \sum_{i=1}^{n} X_i$ is a Gamma variable such $X \sim Gamma(\sum \alpha_i, \tau)$

5.2 Estimation of Parameters

Assuming we have N observations : $X_1,\ldots,X_N$, and α is given

$$L = \frac{\prod_{i=1}^{N}(Xi^{\alpha-1})\exp\left(-\tau\sum Xi\right)}{\tau^{-N\alpha}\left(\Gamma(\alpha)\right)^N} \qquad l = (\alpha-1)\sum_{i=1}^{N}\ln(Xi) - \tau\sum Xi - N\alpha\ln(\tau^{-1}) - N\ln\left(\Gamma(\alpha)\right)$$

$$\frac{dl}{d\tau} = -\sum_{i=1}^{N} Xi + \frac{N\alpha}{\tau}$$

We then set $\frac{dl}{d\tau} = 0$, thus we have : $\hat{\tau} = \frac{N\alpha}{\sum_{i=1}^{N} Xi}$

6 The Log-Normal distribution -μ,σ

Let Y be a Normal distribution $Nor(\mu, \sigma^2)$ and $X = \exp Y$, then

$$F_X(x) = \begin{cases} \Phi(\frac{\ln(x) - \mu}{\sigma}), & if \quad x > 0, \\ 0, & otherwise. \end{cases}$$

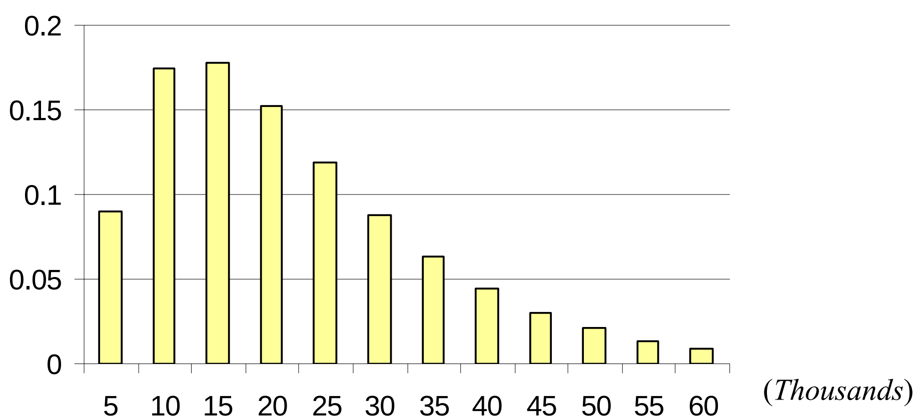
and its density is

$$f_X(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{\ln(x) - \mu}{\sigma})^2), & if \quad x > 0, \\ 0, & otherwise. \end{cases}$$

$$\mathrm{E}[X^k] = \exp(k\mu + k^2\sigma^2/2)$$

$$\mathrm{E}\left[(X \wedge x)^k\right] = \exp(k\mu + k^2\sigma^2/2)\,\Phi\left(\frac{\ln x - \mu - k\sigma^2}{\sigma}\right) + x^k\left[1 - F(x)\right]$$

Chart 5.4. Density μ = 0.4, σ = 01.2



6.1 Properties

Remark 1

We obtain the moment of the log-normal distribution with the moment generating function of the normal distribution, indeed we have:

$$E[X] = E[\exp Y] = M_Y(1) = \exp(\mu + \frac{\sigma^2}{2})$$
$$E[X^2] = M_Y(2) = \exp(2\mu + 2\sigma^2)$$

Thus, $\quad V[X] = \exp(2\mu)\exp(\sigma^2)(\exp(\sigma^2) - 1)$

6.2 Estimation of Parameters

Assuming we have N observations : $X_1,\ldots,X_N$, and α is given

$$L = \prod_{i=1}^{N} \frac{1}{Xi\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(Xi)-\mu)^2}{2\sigma^2}\right] = \frac{1}{\left(\prod_{i=1}^{N} Xi\right)\sigma^N (2\pi)^{N/2}} \exp\left[\sum_{i=1}^{N} -\frac{(\ln(Xi)-\mu)^2}{2\sigma^2}\right]$$

$$l = -\sum_{i=1}^{N} \ln(Xi) - N\ln(\sigma) - \frac{N}{2}\ln(2\pi) - \sum_{i=1}^{N} \frac{(\ln(Xi)-\mu)^2}{2\sigma^2}$$

$$\frac{dl}{d\mu} = \sum_{i=1}^{N} \frac{2(\ln(Xi)-\mu)}{2\sigma^2} = 0$$
$$\frac{dl}{d\sigma} = -\frac{N}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{N} (\ln(Xi)-\mu)^2 = 0$$

Thus we have :

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} \ln(Xi)$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N} (\ln(Xi) - \hat{\mu})^2$$

Bibliography

CHARPENTIER A., DENUIT M. [2004] *Mathématiques de l'assurance non-vie.* Economica.

JEAN LEMAIRE [1985] *Automobile Insurance, Actuarial models.* Economica.

PARTRAT CH., BESSON J.L. [2004] *Assurance non-vie. Modélisation, simulation*. Economica.

STUART A. KLUGMAN, HARRY H. PANJER, GORDON E. WILLMOT JEAN LEMAIRE [2004] *Loss Models, From data to decisions.* Second Edition Wiley.

Fédération Francaise des Sociétés d'Assurances www.ffsa.fr

Esurance www.esurance.com