

**Mémoire présenté le :
pour l'obtention du diplôme
de Statisticien Mention Actuariat
et l'admission à l'Institut des Actuares**

Par : Mossad ABDELSALAM

Titre du mémoire : Tarification des tranches peu travaillantes d'un traité de réassurance en excédent de sinistre pour la branche RC AUTO.

Confidentialité : NON OUI (Durée : 1 an 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus.

Membres présents du jury de la
filière :

Signature :

Entreprise :

Nom : *BPCE Assurances IARD*

Signature :

Directeur de mémoire en
entreprise

Membres présents du jury de
l'Institut des Actuares :

Signature :

Nom : *SACRÉ Emile*

Signature : *SACRÉ*

Invité :

Nom :

Signature :

**Autorisation de publication et de mise
en ligne sur un site de diffusion de
documents actuariels (après expiration
de l'éventuel délai de confidentialité)**

Signature du responsable
entreprise :

Signature du candidat :

Mossad

Résumé

L'objectif de ce mémoire est de faire évoluer le modèle de tarification du traité en excédent de sinistre pour la garantie responsabilité civile automobile de BPCE Assurances IARD en se concentrant sur les tranches peu travaillantes.

De par la complexité et la faible occurrence des sinistres extrêmes, les méthodes de tarification par expérience comme la méthode du Burning Cost, ou fréquence-sévérité ne permettent pas d'établir une évaluation correcte sur des tranches peu travaillantes.

Pour déterminer le coût de nos tranches de réassurance, le modèle historique de BPCE Assurances IARD repose sur l'approche probabiliste basée sur la méthode de fréquence-sévérité. Nous avons proposé un modèle à deux lois afin d'affiner notre modélisation en apportant une attention particulière aux sinistres extrêmes. Ce modèle repose sur la théorie des valeurs extrêmes et plus précisément sur la méthode du dépassement de seuil. Une attention particulière est portée sur la détermination du choix du seuil des extrêmes. Une fois ce seuil déterminé, nous avons calculé les charges cédées sur chacune de nos tranches de réassurances. Nous avons ensuite confronté les résultats obtenus par les modèles candidats que nous avons proposés dans cette étude, aux résultats obtenus à partir du modèle historique de BPCE Assurances IARD ainsi qu'à nos données réelles.

À l'issue de cette phase de comparaison, nous permettant ainsi de juger la pertinence et la sensibilité des modèles aux sinistres extrêmes, nous avons retenu le modèle à deux lois issues de la théorie des valeurs extrêmes modélisant les sinistres extrêmes avec une loi de Pareto Généralisée. Une tarification de chaque tranche de notre traité est alors déduite du modèle retenu. Ce modèle nous permet ainsi de mieux tarifer les tranches peu travaillantes par rapport aux autres modèles expérimentés, répondant ainsi à la problématique de notre étude.

Mots-clés : Réassurance, Non-vie, IARD, théorie des valeurs extrêmes, seuil des extrêmes, tarification, cotation, burning-cost, fréquence, sévérité, loi marché, Pareto généralisée.

Abstract

The aim of this dissertation is to develop the excess of loss treaty pricing model for BPCE Assurances IARD's motor third party liability cover, focusing on the less-worked sections.

Due to the complexity and low occurrence of extreme claims, experience-based pricing methods such as Burning Cost, or frequency-severity, do not allow us to establish a correct valuation on lightly worked layers.

To determine the cost of our reinsurance tranches, BPCE Assurances IARD's historical model relied on the probabilistic approach based on the frequency-severity method. We have proposed a two-law model to refine our modeling, with particular attention to extreme claims. This model is based on extreme value theory, and more specifically on the threshold crossing method. Particular attention is paid to determining the choice of extreme threshold. Once the threshold had been determined, we calculated the ceded expenses on each of our reinsurance tranches. We then compared the results obtained by the candidate models we proposed in this study with the results obtained from BPCE Assurances IARD's historical model, and with our actual data.

At the end of this comparison phase, enabling us to judge the relevance and sensitivity of the models to extreme claims, we chose the two-law model derived from extreme value theory, which models extreme claims with a Generalized Pareto distribution. The model is then used to derive a rate for each bracket in our treaty. This model enables us to better price the less hard-working slices than the other models tested, thus responding to the problematic of the dissertation.

Keywords: Reinsurance, Non-life, IARD, extreme value theory, threshold exceedances, pricing, rating, burning cost, frequency, severity, heavy-tailed distributions, generalized Pareto distribution.

Note de synthèse

La cotation des tranches peu travaillantes en réassurance est un enjeu complexe pour les assureurs. Le sujet de ce mémoire est d'analyser les différentes méthodologies nous permettant de correctement modéliser les sinistres extrêmes afin de faire évoluer la méthode de cotation du traité de réassurance en excédent de sinistre pour la branche responsabilité civile automobile de BPCE Assurances IARD.

Les traités de réassurance non proportionnels sont très flexibles dans leur définition et leur fonctionnement. Contrairement aux traités proportionnels, les traités non proportionnels divergent en termes de définition et de fonctionnement. En effet, la prime de réassurance pour les traités non proportionnels n'est pas fixée mécaniquement par une règle de partage préétablie. En d'autres termes, la répartition des risques entre le réassureur et l'assureur n'est plus aussi claire que dans les cas de réassurance proportionnelle. C'est pourquoi les traités non proportionnels permettent une tarification individualisée du contrat, indépendante de la tarification originale des risques ou des polices couvertes.

Pour certaines tranches d'un traité de réassurance non proportionnelle en Excédent de Sinistre, les cédantes sont protégées contre des risques et des sinistres extrêmes qui ont une occurrence très faible. Par conséquent, les actuaires de réassurance disposent de très peu de données pour modéliser ces risques, ce qui rend la tarification du contrat plus complexe.

Nous pouvons alors segmenter les sinistres en deux catégories, les sinistres de forte fréquence, mais à faible coût qui représentent la très grande majorité de notre portefeuille et les sinistres de faible fréquence mais, à très fort coût. Dans notre portefeuille des sinistres graves, on observe un sinistre minimum à 250 351 €, et un sinistre maximum à 20 120 105 €. La moyenne et l'écart-type de notre échantillon sont respectivement de 377 400 € et 1 196 357 €.

Note échantillon d'observations présente donc une très forte dispersion des valeurs avec un écart-type très élevé et des valeurs accusant de la présence de sinistres extrêmes, ce qui rend difficile l'utilisation d'une modélisation classique basée sur une distribution de probabilité normale.

Pour modéliser ces sinistres, différentes approches sont possibles, telles que la méthode du Burning Cost ou la méthode de fréquence-sévérité. Dans notre étude, nous allons, à partir de la méthode fréquence-sévérité, faire évoluer ce modèle afin de l'adapter aux sinistres extrêmes.

Le modèle de tarification fréquence-sévérité est une méthode de modélisation statistique courante qui permet d'estimer les primes de réassurance pour un portefeuille de risques donné. Cette méthode se base sur la séparation des sinistres en deux composantes distinctes : la fréquence des sinistres et leur sévérité.

La fréquence des sinistres est le nombre de sinistres survenus sur une période donnée, tandis que la sévérité représente le montant de chaque sinistre. Notre étude portera uniquement sur la modélisation des coûts, reprenant ainsi le modèle déjà existant de BPCE Assurances IARD pour la modélisation de la fréquence. Dans ce modèle, la loi de fréquence est calibrée sur une taille de portefeuille définie. En utilisant cette méthode, on peut estimer la prime pure en fonction de la fréquence et de la sévérité des sinistres. Pour cela, on utilise souvent des distributions statistiques pour modéliser la fréquence et la sévérité des sinistres, comme la loi de Poisson pour la fréquence et la loi Gamma pour la sévérité. On peut ensuite combiner ces distributions pour obtenir une distribution jointe qui permet de modéliser les deux composantes simultanément. Une fois que l'on a déterminé la distribution jointe, on peut estimer la prime pure par simulations de Monte Carlo.

Cependant, ces méthodes ne sont pas adaptées aux tranches peu travaillantes en raison de la faible occurrence des sinistres extrêmes. En effet, ces modèles classiques supposent souvent que les données suivent une distribution symétrique et relativement régulière. Dans le cas d'un échantillon avec des

sinistres extrêmes, la distribution peut être fortement asymétrique et avoir une queue de distribution très longue, ce qui peut affecter les estimations des paramètres de la distribution.

La méthode adoptée pour modéliser les sinistres extrêmes afin de tarifier les tranches peu travaillantes est celle utilisant l'approche par la théorie des valeurs extrêmes.

La théorie des valeurs extrêmes (TVE) est une branche des statistiques qui permet de modéliser le comportement de la queue de distribution d'une variable aléatoire, c'est-à-dire les valeurs extrêmes ou les événements rares. Cette méthode est basée sur l'hypothèse que les valeurs extrêmes suivent une distribution de probabilité des valeurs extrêmes.

Le principe de cette théorie est que les valeurs extrêmes d'une distribution suivent une loi de probabilité différente de celle des valeurs centrales. En effet, les valeurs extrêmes sont souvent plus éloignées de la moyenne et exercent une influence plus importante sur les risques encourus.

La théorie des valeurs extrêmes s'appuie sur deux résultats importants :

- Le théorème de Fisher-Tippett-Gnedenko, qui décrit les trois types de lois de probabilité qui peuvent caractériser les valeurs extrêmes d'une distribution.
- Le théorème de Pickands-Balkema-de Haan, qui permet d'estimer la distribution des valeurs extrêmes en se focalisant sur les excès au-delà d'un certain seuil.

Plus précisément, la théorie des valeurs extrêmes utilise des distributions de probabilité comme la distribution de Gumbel, la distribution de Fréchet ou la distribution de Weibull, qui ont des queues de distribution plus épaisses que les distributions normales ou gaussiennes. Ces distributions ont des propriétés mathématiques qui permettent de modéliser les événements extrêmes de manière plus précise que les méthodes statistiques classiques. Cette théorie est par conséquent adaptée pour modéliser les sinistres extrêmes en réassurance, car les sinistres les plus importants sont souvent rares, mais peuvent avoir un impact significatif sur la prime de réassurance.

Contexte :

La tarification technique des tranches XS consiste à déterminer la prime pure des tranches, c'est-à-dire la valeur actuelle probable des sinistres qui vont être payés au titre du traité.

La première étape de calcul de la prime pure consiste à redresser les données historiques (primes d'assurance, sinistres et seuils) afin de le rendre représentatif des conditions de souscription et économiques de l'année de couverture à évaluer. En effet, étant donné la longue durée de liquidation (> 10 ans pour les sinistres corporels), des sinistres déclarés peuvent ne pas être totalement réglés dans un laps de temps restreint. Ces sinistres subissent l'évolution normale des principaux coûts (hospitalisations, assistance tierce personne...). Pendant la période de liquidation, il peut y avoir plusieurs chocs (réglementaires, jurisprudentiels) entraînant de l'inflation. Le coût des sinistres individuels est alors revalorisé en utilisant un indice de revalorisation décrivant l'évolution de l'inflation. Cette revalorisation des données d'expérience est appelée mise en « As If ».

Modélisation

Pour étudier les sinistres extrêmes, nous avons utilisé la méthode du dépassement de seuil (Peak Over Threshold), qui est une approche basée sur la modélisation de la distribution des excès de seuils.

Cette méthode repose sur le postulat que la distribution des excès de seuils (c'est-à-dire les sinistres supérieurs à un certain seuil) suit une loi de Pareto généralisée, qui peut être définie comme suit :

Soit X une variable aléatoire tel que $\sigma > 0, \mu < x^F$ et $\gamma \in \mathbb{R}$ alors, $X \sim GPD(\mu, \sigma, \gamma)$, si la fonction de répartition de X est de la forme :

$$GPD_{\mu,\sigma,\gamma}(x) = \begin{cases} 1 - \left(1 + \gamma \left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0 \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{si } \gamma = 0 \end{cases} \quad \forall x > 0, \text{ tel que } 1 + \frac{\gamma x}{\sigma} > 0.$$

- μ est le paramètre de position
- σ est le paramètre d'échelle
- γ est le paramètre de forme, il détermine la queue de la distribution, c'est-à-dire la rapidité avec laquelle elle décroît. Si γ est grand, la queue est plus lourde, tandis que si γ est petit, la queue est plus légère.

Cette approche est adaptée pour modéliser les sinistres extrêmes, car elle permet de se concentrer sur la loi des excès de seuils définie tel que

Si on note $x^F = \sup\{x \in \mathbb{R} | F(x) < 1\}$, le point extrémal de F.

Alors $\forall u < x^F$, on appelle fonction de répartition des excès la fonction suivante :

$$F_u(x) = \mathbb{P}(X - u \leq x | X > u) = 1 - \frac{1 - F(x + u)}{1 - F(u)}, x \geq 0, \text{ tel que } 0 \leq x \leq x^F - u$$

La fonction de distribution des excès représente la probabilité que X dépasse le seuil μ d'au plus la quantité x . En fixant un seuil des extrêmes approprié, on peut obtenir des données plus homogènes et mieux adaptées à la modélisation de la loi de Pareto généralisée et donc des sinistres extrêmes.

Pour appliquer cette méthode, il faut d'abord estimer les paramètres γ et σ de la loi de Pareto généralisée. Cela peut se faire en utilisant différentes techniques statistiques, telles que la méthode du maximum de vraisemblance ou la méthode des moindres carrés.

Dans un premier temps, nous avons déterminé le domaine d'attraction des données, car cela nous permet de nous assurer que les données extrêmes suivent effectivement une loi de probabilité de Pareto généralisé et d'appliquer certains outils de détermination spécifique au domaine d'attraction de Fréchet. En traçant le graphique quantile-quantile généralisé :

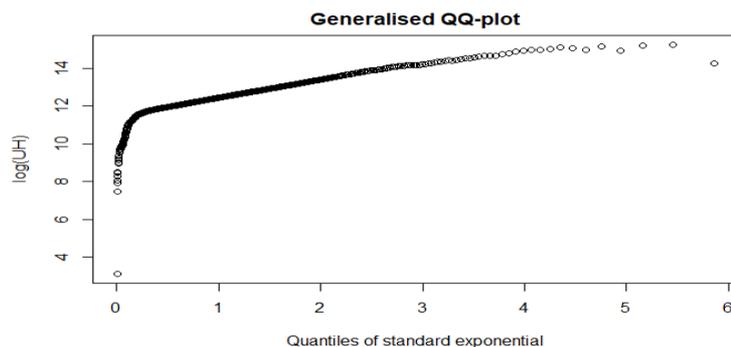


Figure 1 : Graphique Quantile-Quantile généralisé pour notre portefeuille

Nous remarquons que la courbe du quantile-plot généralisé est concave et croissante, et son allure est montante. On peut donc déduire que l'indice des valeurs extrêmes est positif et que les données appartiennent au domaine d'attraction de Fréchet. Le calcul du coefficient de stabilité de Hill vient alors nous confirmer le résultat précédent.

Maintenant que nous avons déterminé le domaine d'attraction de notre échantillon d'observation, il est important de porter une attention particulière au choix du seuil, car celle-ci affecte directement la précision et la fiabilité de l'estimation de la distribution de la Pareto Généralisée.

En effet, le seuil de déclenchement est le niveau minimum de perte au-delà duquel les sinistres sont considérés comme des sinistres extrêmes. Il doit être choisi avec soin, puisqu'un seuil trop élevé peut exclure des observations importantes de la modélisation, tandis qu'un seuil trop bas peut inclure des événements moins extrêmes, ce qui peut conduire à une estimation incorrecte de la distribution des valeurs extrêmes.

Si le seuil est trop bas, il est possible que la modélisation inclût des sinistres qui ne seront pas vraiment extrêmes, ce qui peut conduire à une surestimation de la fréquence et de la gravité des sinistres extrêmes. Cela peut entraîner des conséquences négatives sur la tarification et la gestion des risques, car les compagnies d'assurance et de réassurance peuvent prendre des décisions fondées sur des informations inexactes. Si le seuil est trop élevé, la modélisation peut exclure des sinistres réels qui sont considérés comme des sinistres extrêmes. Cela peut conduire à une sous-estimation de la fréquence et de la gravité des sinistres extrêmes, ce qui peut également entraîner des conséquences négatives sur la tarification et la gestion des risques.

Ainsi, pour déterminer le seuil des extrêmes, nous avons eu recours à plusieurs estimateurs :

L'estimateur de Hill ainsi que son graphique nous fournit une première estimation du choix du seuil :

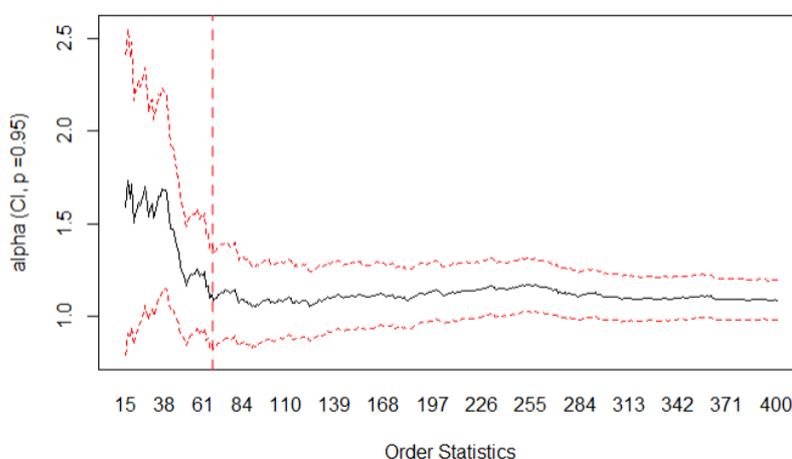


Figure 2 : Graphique de Hill pour notre portefeuille

En utilisant les critères de détermination du choix du seuil selon la méthode du graphique de Hill, on observe que la courbe des excès est approximativement linéaire et se stabilise à partir de la courbe verticale rouge, représentant un nombre d'excès égal à 67 et donc un seuil des extrêmes à 2 142 946€. Le 67^{ème} excès représente, notre 67^{ème} statistique d'ordre qui est le 67^{ème} plus grand sinistre de notre portefeuille.

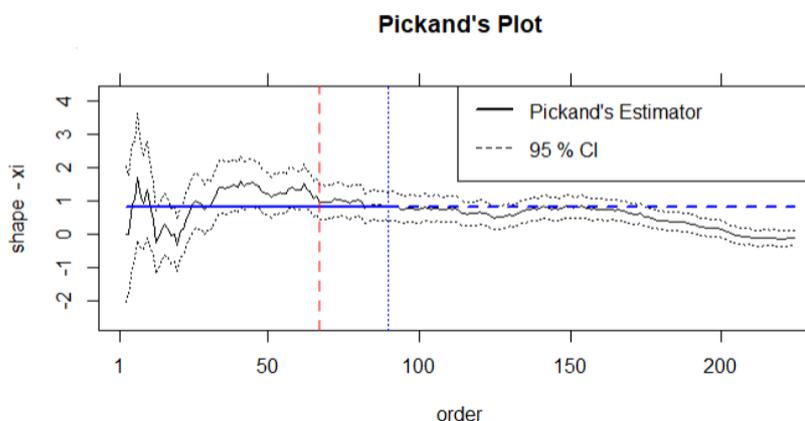


Figure 3 : Graphique de Pickands pour l'estimation du seuil des extrêmes de notre portefeuille.

L'utilisation d'autres outils permettant l'estimation du seuil des extrêmes comme le graphique de Pickands, la minimisation de l'erreur quadratique de l'estimateur de Hill ou encore le graphique de la moyenne des excès viennent alors corroborer notre première hypothèse émise par le graphique de Hill supposant que le seuil des extrêmes était atteint à un nombre d'excès égal à 67, soit 2 142 946 €.

Le seuil des extrêmes étant alors défini, nous pouvons alors estimer les paramètres de la loi de Pareto généralisée qui caractérise les sinistres au-delà du seuil. Pour ce faire, nous utilisons la méthode du maximum de vraisemblance et des moindres carrés. L'étude des graphiques quantile-quantile ainsi les tests statistiques comme le test de Kolmogorov-Smirnov ou d'Anderson-Darling sont alors utilisés pour déterminer quelle estimation choisir. Finalement, l'estimation par la méthode du maximum de vraisemblance sera retenue.

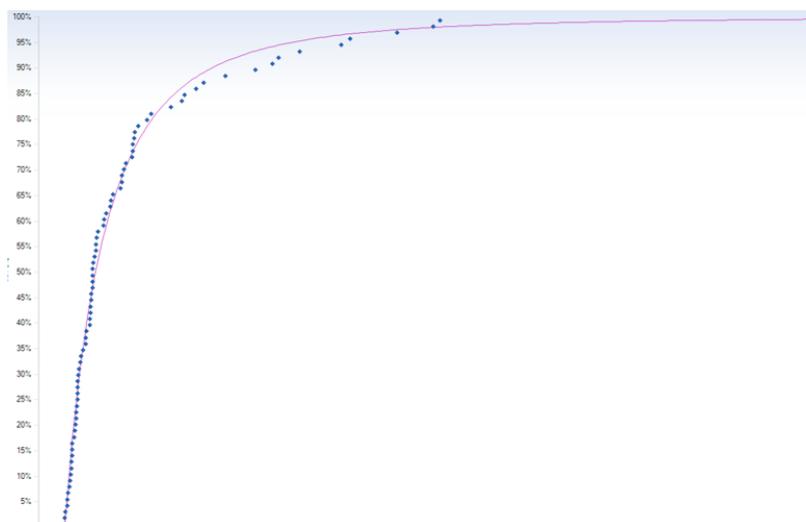


Figure 4 : Fonction de répartition estimée par le maximum de vraisemblance

Notre échantillon d'observation semble être en adéquation avec les paramètres de la loi retenue modélisant le coût des sinistres au-dessus du seuil.

Disposant de suffisamment de données pour modéliser la loi de sévérité en dessous du seuil des extrêmes, une approche classique de calibrage de loi est alors utilisée.

Les paramètres des lois de coûts et de fréquence en dessous et au-dessus du seuil étant déterminés, nous pouvons alors procéder à la modélisation de la charge annuelle en utilisant la méthode de Monte-Carlo. Ces étapes de modélisations sont réalisées sur les données BPCE Assurances IARD.

Pour les différentes modélisations mises en œuvre, une étape de comparaison entre les résultats obtenus par les modèles candidats et les données réelles de notre portefeuille est alors réalisée.

Conclusion

Parmi les différents modèles que l'on a proposés afin de modéliser notre sinistralité, seul le modèle à deux lois issues de la théorie des valeurs extrêmes a été retenu. Les différentes méthodes de validation que l'on a utilisées pour évaluer la performance de notre modèle ont été primordiales nous permettant ainsi d'assurer la qualité et la validité de nos résultats. Nous avons principalement utilisé des méthodes statistiques telles que la comparaison des ratios S/P et des charges cédées, l'ajustement de la distribution et les mesures de performance pour valider notre modèle.

Le modèle historique de BPCE Assurances IARD présentait certaines incertitudes quant à l'estimation de la charge des sinistres extrêmes. En effet, ce dernier a tendance à sous-estimer les coûts des

sinistres extrêmes, entraînant alors un impact direct sur le coût de la cession pour les tranches peu travaillantes.

Quelles que soient les tranches du traité, les modèles issus des lois marché présentaient des caractères moins prudeniels que le modèle historique ou celui issu de la théorie des valeurs extrêmes. Ces modèles ne sont donc pas retenus car ils sous-estiment la cession.

En appliquant le modèle généré par la théorie des valeurs extrêmes et en le comparant au modèle historique de BPCE Assurances IARD, il semblerait que la segmentation faite par la détermination du seuil des extrêmes ainsi que la modélisation au-delà de ce seuil par une loi de Pareto Généralisé résout les problèmes liés à la sous-tarification des tranches peu travaillantes, tout en conservant les bonnes propriétés du modèle historique pour tranches travaillantes et pour les sinistres en dessous du seuil des extrêmes. Ce modèle résout aussi les problèmes liés à l'instabilité du choix des lois de sévérité des sinistres extrêmes dans la modélisation historique. Ainsi, pour la modélisation actuelle et rétroactive effectuée en utilisant le modèle issu de la loi de Pareto Généralisée, le modèle choisi présentait un caractère plus prudentiel que le modèle historique qui avait tendance à sous-estimer le risque et à le tarifier de manière moins importante.

Les résultats de la modélisation sont alors exprimés :

		Moyenne	Ecart-type
Tranche inférieure au seuil des extrêmes	Fréquence	119.6	28.40
	Coût total	39 205 050	9 628 042
Tranche supérieure au seuil des extrêmes	Fréquence	16.4	7.17
	Coût total	57 366 029	32 245 603

Tableau 1 : Estimation de la fréquence et de la charge moyenne pour le modèle issu de la Pareto généralisée

À l'issue de cette phase de comparaison, le modèle issu de la théorie des valeurs extrêmes est alors validé et une tarification est alors déduite du modèle retenu pour les tranches de notre traité de réassurance.

Synthesis Note

The pricing of low-exposure layers in non-life reinsurance is a crucial issue for insurers. The subject of this thesis is to analyze the different methodologies that allow us to properly model extreme losses in order to evolve the pricing method of the excess of loss reinsurance treaty for the automobile liability branch of the BPCE Assurances IARD group.

Non-proportional reinsurance treaties are very flexible in their definition and operation. Unlike the most common proportional treaties such as quota share and excess of loss, the reinsurance premium for non-proportional treaties is not mechanically set by a pre-established sharing rule. In other words, the distribution of risk between the reinsurer and the insurer is no longer as clear-cut as in the case of proportional reinsurance. Therefore, non-proportional treaties allow for an individualized pricing of the contract, independent of the original pricing of the risks or policies covered.

For some layers of a non-proportional excess of loss reinsurance treaty, ceding companies are protected against extreme risks and losses that have not or have rarely occurred in the past. Consequently, reinsurance actuaries have very little data to model these risks, which makes pricing the contract more complex.

We can then segment claims into two categories: high-intensity but low-cost claims, which represent the vast majority of our portfolio, and low-intensity but high-cost claims. In our portfolio, we observe a minimum claim of 250 351 € and a maximum claim of 20 120 105 €. The mean and standard deviation of our sample are 377 400 € and 1196 357 €, respectively.

Our sample of observations presents a very high dispersion of values with a very high standard deviation and values showing the presence of extreme claims, which makes it difficult to use a classical modeling based on a normal probability distribution or a Poisson distribution.

To model these claims, different approaches are possible, such as the Burning Cost method or the frequency-severity method. In our study, we will use the frequency-severity method and evolve this model to adapt it to extreme claims.

The frequency-severity pricing model is a common statistical modeling method that estimates reinsurance premiums for a given portfolio of risks. This method is based on the separation of claims into two distinct components: the frequency of claims and their severity.

The frequency of claims is the number of claims that occur during a given period, while severity represents the average amount of each claim. Our study will focus solely on modeling costs, thus building on the existing internal model of BPCE Assurances IARD for modeling frequency. By using this method, we can estimate the pure premium based on the frequency and severity of claims, as well as the size of the portfolio. To do this, statistical distributions are often used to model the frequency and severity of claims, such as the Poisson distribution for frequency and the Gamma distribution for severity. These distributions can then be combined to obtain a joint distribution that allows for the simultaneous modeling of both components. Once the joint distribution has been determined, different methods can be used to estimate the pure premium, such as the Monte Carlo simulation method.

However, these methods are not suitable for thinly traded tranches due to the low occurrence of extreme losses. Indeed, these classical models often assume that the data follows a symmetric and relatively regular distribution. In the case of a sample with extreme losses, the distribution can be strongly asymmetric and have a very long tail, which can affect the estimates of the distribution parameters.

The method we have chosen to model extreme losses and rate thinly traded tranches is the approach using extreme value theory.

Extreme value theory (EVT) is a statistical method that allows modeling the behavior of the tail of a probability distribution, that is, the extreme values or rare events. This method assumes that extreme values follow a probability distribution of extreme values.

The principle of this theory is that the extreme values of a distribution follow a different probability law than the central values. Indeed, extreme values are often farther away from the mean and exert a greater influence on the risks incurred.

The theory of extreme values is based on two important results:

- The Fisher-Tippett-Gnedenko theorem, which describes the three types of probability distributions that can characterize the extreme values of a distribution.
- The Pickands-Balkema-de Haan theorem, which allows the estimation of the distribution of extreme values by focusing on the exceedances beyond a certain threshold.

Specifically, the theory of extreme values uses probability distributions such as the Gumbel distribution, the Fréchet distribution, or the Weibull distribution, which have heavier tails than normal or Gaussian distributions. These distributions have mathematical properties that allow for a more precise modeling of extreme events than classical statistical methods. Therefore, this theory is well-suited for modeling extreme losses in reinsurance, as the largest losses are often rare but can have a significant impact on the reinsurance premium.

Context:

The technical pricing of XS layers consists of determining the pure premium of the layers, which is the present value of the probable losses that will be paid under the treaty.

The first step in calculating the pure premium is to adjust the historical data (premiums, losses, and thresholds) to make it representative of the underwriting and economic conditions of the coverage year being evaluated. Due to the long settlement period (> 10 years for bodily injury claims), reported claims may not be fully settled within a limited period of time. These claims are subject to the normal evolution of major costs (hospitalizations, third-party assistance, etc.). During the settlement period, there may be several shocks (regulatory, jurisprudential) leading to a change in inflation. The cost of individual claims is then revalued using an index describing the inflation trend. This revaluation of experience data is called "AS-IF" adjustment.

Modeling

To study extreme losses, we used the Peak Over Threshold method, which is an approach based on modeling the distribution of threshold exceedances.

This method is based on the assumption that the distribution of threshold exceedances (i.e., losses above a certain threshold) follows a generalized Pareto distribution, which can be defined as follows:

Let X be a random variable such that $\sigma > 0, \mu < x^F$ et $\gamma \in \mathbb{R}$, then $X \sim GPD(\mu, \sigma, \gamma)$, if the cumulative distribution function of X is of the form:

$$G_{\mu, \sigma, \gamma}(x) = \begin{cases} 1 - \left(1 + \gamma \left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0 \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{si } \gamma = 0 \end{cases} \quad \forall x > 0, \text{ as } 1 + \frac{\gamma x}{\sigma} > 0.$$

Where:

- μ is the location parameter
- σ is the scale parameter
- γ is the shape parameter, which determines the tail of the distribution, i.e., how quickly it decreases. If gamma is large, the tail is heavier, while if gamma is small, the tail is lighter.

This approach is suitable for modeling extreme losses because it allows us to focus on the distribution of excesses above a certain threshold, defined as:

If we denote $x^F = \sup\{x \in \mathbb{R} | F(x) < 1\}$, the extreme point of F .

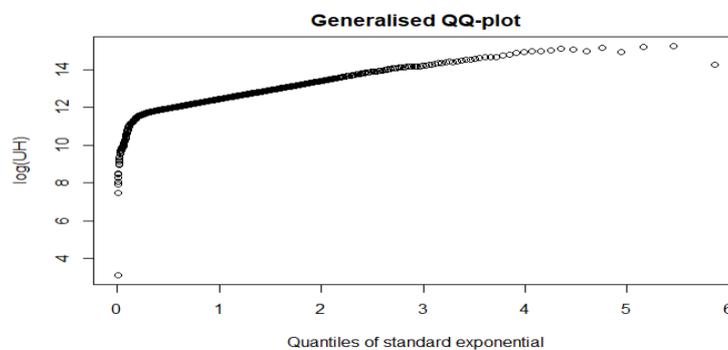
Then, for $\forall u < x^F$, we call the following function the excess distribution function:

$$F_u(x) = \mathbb{P}(X - u \leq x | X > u) = 1 - \frac{1 - F(x + u)}{1 - F(u)}, x \geq 0, \text{ tel que } 0 \leq x \leq x^F - u$$

The excess distribution function represents the probability that X exceeds the threshold μ by no more than the amount x . By setting an appropriate extreme threshold, we can obtain more homogeneous data that is better suited for modeling the generalized Pareto distribution and thus extreme losses.

To apply this method, we first need to estimate the parameters γ and σ of the generalized Pareto distribution. This can be done using various statistical techniques, such as maximum likelihood or least squares methods.

Initially, we determined the domain of attraction of the data, as this allowed us to ensure that the extreme data actually follows a generalized Pareto probability distribution and to apply specific determination tools to the Fréchet domain of attraction. By plotting the generalized quantile-quantile graph:

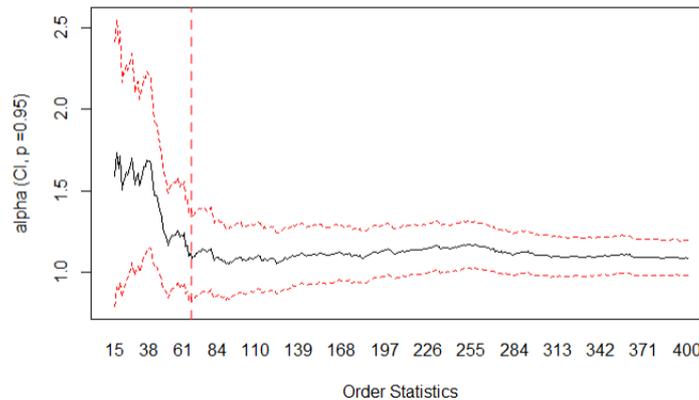


We notice that the generalized quantile-plot curve is convex and decreasing, and its shape is increasing. We can therefore deduce that the index of extreme values is positive and that the data belongs to the Fréchet domain of attraction. The calculation of the Hill stability coefficient confirms the previous result.

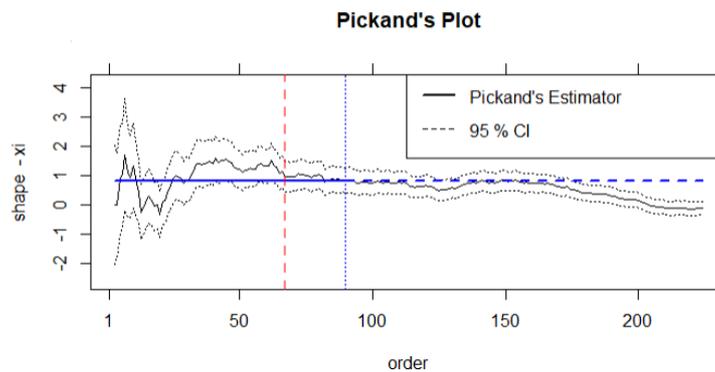
Now that we have determined the domain of attraction of our observation sample, it is important to pay particular attention to the choice of the threshold, as it directly affects the accuracy and reliability of the estimation of the Generalized Pareto Distribution.

Indeed, the triggering threshold is the minimum level of loss beyond which claims are considered as extreme claims. It must be chosen carefully, as a threshold that is too high may exclude important observations from the modeling, while a threshold that is too low may include less extreme events, which can lead to an incorrect estimation of the distribution of extreme values.

If the threshold is too low, it is possible that the modeling includes claims that are not really extreme, which can lead to an overestimation of the frequency and severity of extreme claims. This can have negative consequences on pricing and risk management, as insurance and reinsurance companies may make decisions based on inaccurate information. If the threshold is too high, the modeling may exclude real claims that are considered extreme claims. This can lead to an underestimation of the frequency and severity of extreme claims, which can also have negative consequences on pricing and risk management. Thus, to determine the extreme threshold, we used several estimators: the Hill estimator and its plot provide us with a first estimate of the threshold choice:

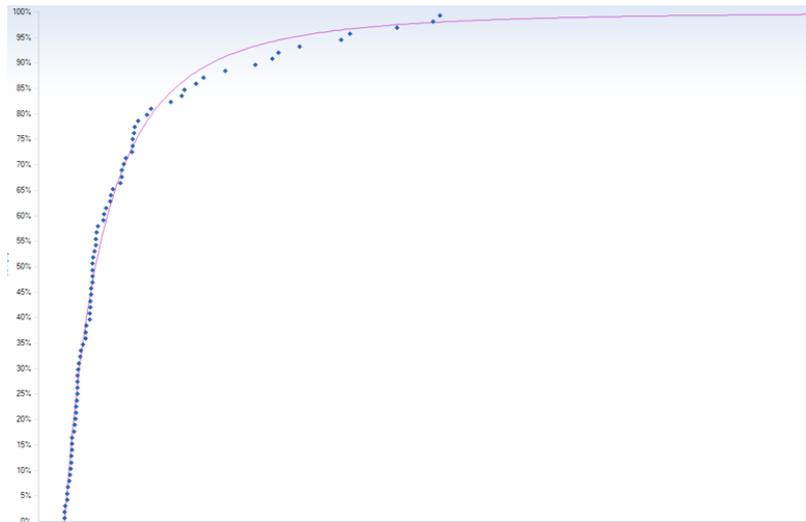


Using the threshold selection criteria according to the Hill plot method, we observe that the excess plot is approximately linear and stabilizes from the red vertical curve, representing a number of excesses equal to 67 and thus an extreme threshold of 2 142 946 €.



Using other methods for estimating the threshold of extremes such as the Pickands plot, the minimization of the mean squared error of the Hill estimator, or the mean excess plot, confirm our initial assumption made by the Hill plot that the threshold of extremes is reached at an excess number of 67, which corresponds to 2 142 946 €.

With the threshold of extremes defined, we can then estimate the parameters of the generalized Pareto distribution that characterizes the losses beyond the threshold. To do this, we use the maximum likelihood method and least squares. The study of the quantile-quantile plots and statistical tests such as the Kolmogorov-Smirnov or Anderson-Darling test are then used to determine which estimation to choose. Finally, the maximum likelihood estimation method will be chosen.



Estimated cumulative distribution function by maximum likelihood.

Our observation sample appears to be in line with the parameters of the chosen law modeling the cost of claims above the threshold. Having enough data to model the cost law below the extreme threshold, a classical approach to law calibration is then used. With the parameters of cost and frequency laws below and above the threshold determined, we can proceed with modeling the annual load using the Monte-Carlo method. These modeling steps are repeated for the historical model of BPCE Assurances IARD as well as for models created from market laws.

For the different proposed models, a comparison step between the results obtained by the candidate models and the actual data of our portfolio is then performed.

Conclusion

Of the various models we proposed to model our claims experience, only the two-law model based on extreme value theory was selected.

The various validation methods we used to assess the performance of our model were of prime importance in ensuring the quality and validity of our results. We mainly used statistical methods such as comparison of S/P ratios and ceded loads, distribution fitting and performance measures to validate our model.

The historical model of BPCE Assurances presented some uncertainties regarding the estimation of extreme claims costs. Indeed, it tends to underestimate the costs of extreme claims, thus having a direct impact on the cost of the cession for less active layers.

Regardless of the treaty layers, the models derived from market laws presented less cautious characteristics than the historical model or the one derived from extreme value theory. Therefore, these models are not retained.

By applying the model generated by extreme value theory and comparing it to the historical model of BPCE Assurances IARD, it appears that the segmentation made by determining the extreme threshold as well as modeling beyond this threshold with a Generalized Pareto distribution solves the problems related to the underpricing of less active layers, while retaining the good properties of the historical model for active layers and claims below the extreme threshold.

This model also solves the problems related to the instability of the choice of severity laws for extreme claims in the historical modeling.

		Mean	Standard deviation
Range below the threshold of extremes	Number of losses	119.6	28.40
	Total losses	39 205 050 €	9 628 042 €
Range above the threshold of extremes	Number of losses	16.4	7.17
	Total losses	57 366 029 €	32 245 603 €

Estimation of frequency and average load for the generalized Pareto model

Thus, for the current and retroactive modeling performed using the model derived from the Generalized Pareto distribution, the chosen model had a more cautious character than the historical model which tended to underestimate the risk and tariff it less. At the end of this comparison phase, the model derived from extreme value theory is validated and a pricing is deduced from the chosen model for the layers of our reinsurance treaty.

Remerciements

Je tiens à remercier Emilie SACRE, tutrice de ce mémoire, ainsi que Toan NGUYEN-HUU manager de l'équipe Réassurance de BPCE Assurances IARD, pour leur disponibilité, leur bienveillance, et les missions qui m'ont été confiées ainsi que leurs nombreux conseils tout au long de ma période d'alternance, qui m'ont permis de réaliser ce mémoire dans les meilleures conditions.

Merci également à toute l'équipe du département Actuariat de BPCE Assurances IARD pour son accueil chaleureux et son professionnalisme.

Pour finir, je remercie tout le corps professoral et administratif de l'ISUP pour la formation dispensée tout au long de mon master, et plus particulièrement Oliver LOPEZ, tuteur académique de ce mémoire et directeur de l'ISUP pour son encadrement et ses conseils.

Table des matières

Résumé	0
Abstract	1
Note de synthèse	2
Synthesis Note	8
Remerciements	14
Introduction	18
I Généralités sur la réassurance	19
1.1 Principe de la réassurance	19
1.2 Les types de réassurance	20
1.2.1 Les formes juridiques de la réassurance	20
1.2.2 La réassurance non-proportionnelle	21
1.2.3 La réassurance proportionnelle	24
1.3 Les clauses des traités	28
1.3.1 La clause de reconstitution de garantie	28
1.3.2 La clause de franchise Aggregate ou clause de franchise annuelle	29
2 Sinistralité	34
2.1 Type de branche	34
2.2 Notion de provisions IBNYR et IBNER	34
2.3 La méthode Chain Ladder	35
2.4 Notion de tranches travaillantes et non travaillantes	37
2.5 Indemnisation des victimes en Responsabilité Civile	38
2.5.1 Les postes de préjudices	41
2.5.2 Evaluation des sinistres	44
2.5.3 Les rentes	44
2.5.4 Indexation des rentes	45
2.5.5 Révision des rentes	47
2.5.6 Provisionnement	48
2.5.7 Indemnisation sous forme de capital :	49
2.6 Les rentes en réassurance	50
2.6.1 Clause de rentes en suivi additionnel	50
2.6.2 Clause de rachat	51

3	Cotation.....	54
3.1	Prime pure et commerciale	54
3.2	Tarification par la méthode du « Burning Cost » : approche et limites	55
3.3	Tarification par l'approche probabiliste : Méthode fréquence-coût sinistre	59
3.3.1	Présentation de la méthode.....	59
3.3.2	Impact du choix du seuil sur les lois.....	60
3.3.3	Modélisation de la fréquence.....	60
3.3.4	Modélisation de la sévérité	63
3.3.5	Estimation des paramètres de la loi.....	65
3.3.6	Ajustement des lois	68
3.3.7	Détermination de la charge annuelle par simulation de Monte-Carlo	72
4	Théorie des Valeurs Extrêmes.....	73
4.1	Introduction générale.....	73
4.1.1	Domaine d'attraction	74
4.2	Le théorème de Fisher-Tippett-Gnedenko (en anglais GEV pour Generalized Extreme Value) 75	
4.3	Modèle du dépassement de seuil (Peak Over Threshold).....	77
4.3.1	Distribution de Pareto généralisée	77
4.3.2	Loi des excès.....	77
4.3.3	Théorème de Pickands	78
4.4	Estimation des paramètres de la GPD	79
4.4.1	Estimation des paramètres de la GPD par la méthode du maximum de vraisemblance 79	
4.4.2	Estimation des paramètres de la GPD par la méthode des moments.....	80
4.4.3	Estimation des paramètres de la GPD par la méthode des moments pondérés	80
4.4.4	Estimation des paramètres de la GPD par la méthode des moindres carrés.....	81
4.4.5	Estimation des paramètres de la GPD par l'Estimateur de Hill.....	82
4.4.6	Estimation des paramètres de la GPD par l'Estimateur de Dekkers, Einmahl et De Hann 83	
4.4.7	Estimation des paramètres de la GPD par l'Estimateur de Pickands.....	84
5	Choix du seuil des extrêmes	84
5.1	Détermination du seuil des extrêmes par l'estimateur de Hill ou de Pickands.....	85
5.2	Détermination du seuil des extrêmes par la fonction des excès moyens	87
5.3	Détermination du seuil des extrêmes par le Gerstengarbe plot	88

6	Application de la méthode sur notre portefeuille	90
6.1	Etude de notre échantillon d'observation de base	91
6.2	Retraitement de nos données sinistres.....	92
6.2.1	Indexation	92
6.2.2	Redressement des primes.....	93
6.2.3	Mise en « AS-IF » de nos données.....	94
6.2.4	Projection à l'ultime.....	95
6.2.5	Traitement des postes capitalisés.....	96
6.3	Détermination du seuil des extrêmes	96
6.3.1	Détermination du domaine d'attraction des données.....	97
6.3.2	Choix du seuil des extrêmes.....	99
6.4	Modélisation de nos lois de sévérités.....	103
6.4.1	Modélisation au-dessus du seuil des extrêmes.....	103
6.4.2	Modélisation sous le seuil des extrêmes	107
6.5	Modélisation de la charge annuelle et application du traité de réassurance	108
6.6	Back-Testing et validation du modèle retenu	110
6.6.1	Comparaison de notre historique aux modèles à deux lois et historique de BPCE Assurances IARD.....	110
6.6.2	Construction d'un modèle issu d'une loi marché.....	112
6.6.3	Value at Risk (VaR)	113
6.6.4	Validation du modèle retenue.....	114
6.7	Cotation	115
6.8	Période de retour	115
6.9	Rétrospective sur la validité de notre modèle	117
	Conclusion	118
	Liste des figures.....	120
	Bibliographie.....	122

Introduction

La réassurance est un marché essentiel pour les compagnies d'assurance, qui transfèrent une partie de leurs risques à des réassureurs pour se protéger contre des pertes financières importantes en cas de sinistres. De plus, la nature complexe de la sinistralité découlant de cette activité rend difficile sa modélisation. Dans notre étude, nous allons nous intéresser à la garantie responsabilité civile, qui est à la fois la garantie minimale d'un contrat d'assurance automobile, mais aussi l'une des couvertures les plus importantes. Il est donc primordial de bien maîtriser la sinistralité concernant cette garantie. Cependant, elle est également sujette à une forte variation des sinistres, en particulier de sinistres graves et coûteux. En effet, nous pouvons segmenter cette sinistralité en deux parties :

- D'une part, nous avons des sinistres de faibles coûts mais, à très forte fréquence pour lesquelles nous avons assez de données historiques afin de correctement les modéliser.
- D'autre part, nous avons des sinistres aux coûts extrêmes mais à très faible fréquence pour lesquelles les données historiques sont très faibles.

Dans ce contexte, la modélisation des sinistres extrêmes est devenue un sujet de plus en plus important pour l'équipe de Réassurance de BPCE Assurances IARD. En effet, comprendre la fréquence et la sévérité de ces sinistres est essentiel afin de ne pas négliger l'impact de ces derniers dans les modélisations et de ne pas subir des répercussions importantes sur les résultats de la cession. Notre étude portera uniquement sur la modélisation des coûts, reprenant ainsi le modèle déjà existant de BPCE Assurances IARD pour la modélisation de la fréquence.

L'enjeu de ce mémoire est de construire un modèle global prenant en compte ces deux types de sinistralité afin d'estimer le coût d'un traité de réassurance en excédent de sinistre. Il est important de notifier que la faible volumétrie de nos données historiques concernant les sinistres extrêmes à une répercussion immédiate sur la tarification des tranches de notre traité de réassurance.

En effet, lorsque, la cédante dispose de suffisamment de données sinistres franchissant la priorité de la tranche du traité, il est possible d'utiliser une seule distribution de probabilité pour estimer les coûts de réassurance associés. Cependant, lorsque la tranche de réassurance n'est que très peu touchée par des sinistres, il devient nécessaire d'estimer les coûts de manière plus précise en se concentrant sur la queue de la distribution des sinistres.

Le modèle de distribution de probabilité avec un calibrage classique des lois, qui est souvent utilisé pour estimer les coûts, ne permet pas de refléter correctement le risque associé aux tranches non-travaillantes. Nous explorons donc dans cette étude l'utilisation de modèles de sinistralité plus adaptés pour les sinistres extrêmes en assurance automobile, afin d'améliorer l'estimation des coûts de réassurance pour les tranches non-travaillantes.

Dans un premier temps nous aborderons les spécificités liées à la réassurance, aux types de sinistres ainsi que les principes de la tarification, ces points là nous permettent de définir notre cadre de travail ainsi que de mieux de comprendre les enjeux liés à notre étude. Dans un second temps, nous introduirons les différents outils mathématiques qui nous serviront à résoudre notre problématique et nous nous pencherons en particulier sur la théorie des valeurs extrêmes. Nous introduirons ensuite les différentes méthodes de détermination du choix du seuil des extrêmes ainsi que les méthodes d'estimation des lois des extrêmes. Pour finir, nous appliquerons la méthode au portefeuille de BPCE Assurances IARD, et nous procéderons à une étape de backtesting sur des données réelles afin de vérifier la robustesse et la pertinence des résultats du modèle proposé.

I Généralités sur la réassurance

I.1 Principe de la réassurance

L'objectif principal de la réassurance est de diminuer le risque pris par l'assureur.

Pour bien comprendre la réassurance, il faut tout d'abord rappeler certains principes de base de l'assurance puisque l'assurance est le fondement de la réassurance.

Pour une compagnie d'assurances, il est nécessaire de savoir quelle part de ses fonds propres elle risque de perdre au cours d'un exercice, et avec quelle probabilité. En effet, l'assurance est basée sur le calcul de probabilité de survenance d'un sinistre. À partir de cette probabilité, la compagnie d'assurance détermine la prime qu'elle demande à ses assurés pour couvrir leurs risques.

L'assurance est une opération par laquelle un assureur organise en mutualité une multitude d'assurés exposés à la réalisation de risques déterminés, et indemnise ceux d'entre eux qui subissent un sinistre grâce à la masse commune des primes collectés. Elle repose sur un fondement statistique : Plus le nombre d'expériences augmente, plus les écarts absolus augmentent et plus les écarts relatifs diminuent pour devenir quasiment négligeable pour un nombre très grand d'expérience, c'est la loi des grands nombres. L'assureur va donc sélectionner les risques qu'il prend en charge, se les répartir avec d'autres, et ajuster le montant de ses primes en conséquence. L'assureur doit par conséquent des risques homogènes, normaux, présentant approximativement les mêmes caractéristiques, ce qui n'est pas toujours le cas.

L'assurance est donc un mécanisme de mutualisation des risques, de sorte qu'ils se compensent entre eux.

Pour réduire sa probabilité de ruine, l'assureur pourrait décider d'inclure des chargements importants dans ses tarifs, mais cette option risquerait de le rendre moins attractives aux vues de ses concurrents. Il pourrait alors décider d'accroître ses fonds propres, mais ce processus nécessite une validation au plus haut niveau et peut être contraignant, ou alors de limiter ses souscriptions, mais encore une fois cela limitera son développement et lui offrira une clientèle restreinte. L'assureur préférera donc se réassurer, afin de pouvoir souscrire et se développer, tout en étant protégé contre des écarts de sinistralité.

Le principe de la réassurance est aussi simple que celui de l'assurance. De par sa nature, la réassurance est une transaction financière impliquant un transfert de risque (cession) d'une compagnie d'assurance (la cédante) vers une compagnie de réassurance (le réassureur) en échange d'une prime de réassurance.

Un contrat de réassurance existe sous plusieurs formes et couvre une période donnée. De manière générale, la plupart des contrats de réassurance ont une période de validité d'un an, et sont renouvelés chaque année.

Dans un contrat de réassurance, il existe une cédante et plusieurs réassureurs, cependant il peut aussi exister plusieurs cédantes et réassureurs, c'est le cas par exemple de la coassurance réassurée. Le réassureur qui a la plus grande part de la cession totale est appelé le réassureur apériteur. Le contrat de réassurance ne lie juridiquement que la cédante à son réassureur, par ce fait, la cédante devra rembourser ses assurés sinistrés même si le réassureur refuse de la payer (différente interprétation des clauses du contrat de réassurance) ou s'il ne peut tout simplement pas les régler (insolvabilité), l'assureur reste le seul responsable vis-à-vis de l'assuré.

La réassurance permet à une cédante d'obtenir certains avantages, notamment une réduction de son engagement sur des risques et une protection contre des pertes multiples ou importantes. La réassurance permet également à une cédante d'obtenir une capacité de souscription supérieure et donc de souscrire des polices portant sur des risques plus importants et plus nombreux, ce qui ne serait pas possible sans une augmentation de ses fonds propres.

Le principe de réassurance permet ainsi de maîtriser la probabilité de ruine d'un assureur, de compenser l'absence de fonds propres, de lisser les écarts statistiques sur un portefeuille, d'augmenter la possibilité de diversifier le portefeuille de risques des assureurs, d'apporter un soutien pour des risques hors norme et de répartir les risques.

La réassurance permet ainsi aux assureurs de réduire le risque émanant de leurs portefeuilles, ce qui leur offre de rester solvables et d'honorer leurs engagements en dépit d'événements extrêmes. Elle permet aussi de limiter la volatilité du résultat de la cédante, tout en réduisant le besoin en capitaux propres étant donné le transfert de risque. Dans le cadre de la réglementation Solvabilité II, le capital de solvabilité requis représente le capital cible nécessaire pour absorber des chocs exceptionnels. Les cédantes pourront faire appel à la réassurance pour atteindre un niveau acceptable de risque, qui correspond à une probabilité de ruine à un an inférieur à 0.5%.

Tout comme l'assureur direct, les réassureurs ont recours à la réassurance, appelée alors rétrocession, c'est l'opération par laquelle un réassureur cède à son tour, une partie des risques qu'il a réassurée à un rétrocessionnaire qui peut être une société d'assurance ou de réassurance.

Les différentes formes de réassurance suivent dans les grandes lignes ceux des assurances. Il s'exerce donc une première différence entre la réassurance vie ou réassurance de personnes et la réassurance non-vie, dite aussi de dommage.

La réassurance non-vie se divise ensuite entre : la réassurance de biens, la réassurance de responsabilité civile et la réassurance de personnes. Dans notre étude, nous traiterons exclusivement du cas de la Réassurance en responsabilité civile automobile.

1.2 Les types de réassurance

La réassurance peut avoir deux sens, selon si la compagnie cède ou accepte le risque. En réassurance passive, l'assureur cède à un réassureur, les risques qui excèdent sa capacité de souscription ou qu'il n'est pas en mesure de couvrir. A l'inverse, en réassurance active, le réassureur accepte les risques qu'une compagnie d'assurance souhaite lui céder, c'est l'acceptation en réassurance. De plus, on peut distinguer quatre formes de réassurance qui se distinguent par le caractère de la cession ou de l'acceptation.

1.2.1 Les formes juridiques de la réassurance

La réassurance se distingue en quatre modes : La réassurance facultative, la réassurance facultative-obligatoire ou obligatoire-facultative et la réassurance obligatoire.

Dans le cadre de la réassurance facultative, les opérations proposées par la cédante sont examinées risque par risque par le réassureur. Les deux parties sont libres de céder ou d'accepter un risque en totalité ou en partie. Historiquement, c'est la forme de réassurance la plus ancienne.

En réassurance facultative-obligatoire, l'assureur a la possibilité de céder ou non un risque, en totalité ou en partie, mais le réassureur a l'obligation d'accepter tout ce qui lui est cédé, selon des conditions

bien définies au préalable. Dans le cadre de la réassurance obligatoire-facultative, le raisonnement est antagoniste à la réassurance facultative-obligatoire, la cession peut être obligatoire et l'acceptation facultative.

En réassurance obligatoire, les opérations sont traitées au niveau d'une branche d'assurance du portefeuille de la cédante, la cédante et le réassureur établissent des obligations réciproques via le biais d'un traité dans lequel la société d'assurance s'engage, durant une période spécifique, à céder les risques d'une catégorie donnée, et la société de réassurance est dans l'obligation de les accepter. C'est le mode réassurance le plus couramment utilisé.

Dans notre étude, nous allons principalement nous concentrer sur ce dernier type de réassurance.

		Réassureur	
		Acceptation facultative	Acceptation obligatoire
Cédante	Cession facultative	Réassurance facultative	Réassurance facultative-obligatoire
	Cession obligatoire	Réassurance facultative obligatoire-	Réassurance obligatoire

Tableau I : Les quatre modes de réassurance

Les formes techniques de la réassurance

De plus, chaque mode de réassurance, peut prendre plusieurs formes, on en distingue deux : la réassurance proportionnelle et la réassurance non proportionnelle.

1.2.2 La réassurance non-proportionnelle

En réassurance non-proportionnelle, le transfert de risque de la cédante vers le réassureur ne se fait qu'au-delà d'un certain seuil de sinistre, ainsi leurs sorts sont beaucoup moins liés que dans le cas d'une réassurance proportionnelle, et on pourra retrouver des cas de traités non-proportionnels à l'issue desquels la quasi-totalité de la charge de sinistre reviendra à la cédante tandis que le réassureur n'aura qu'une faible part de récupérations à verser à la cédante.

Ainsi, le sort de la cédante et du réassureur est moins accolé, la cédante peut alors être en perte tandis que le réassureur peut ne pas l'être.

De plus, il n'existe aucun lien de proportionnalité entre la prime d'assurance versée par l'assuré à l'assureur et la prime de réassurance. La prime de réassurance est-elle déterminée lors du processus de tarification et de négociations avec les réassureurs en fonction du risque couvert.

En réassurance non proportionnelle, le réassureur prend à sa charge exclusivement :

- Soit tout sinistre au-delà d'un certain montant, c'est ce qu'on appelle un excédent de sinistre par risque et/ou par événement.
- Soit le montant annuel global des sinistres, au-delà d'un certain pourcentage ou d'un certain montant, dans ce cas-là, on parle d'excédent de perte annuelle, ou Stop-loss.

Les traités non proportionnels sont définis par une priorité ou déductible, correspondant à une priorité (que l'on notera d) et d'une portée ou limite (que l'on notera l). Le plafond est alors le résultat de la somme entre la priorité et la portée. Le réassureur prend alors à sa charge tout ou une partie du sinistre qui excède la priorité et dans la limite de la portée.

Un contrat non proportionnel se note *Portée XS Priorité*.

La priorité fonctionne comme une franchise, c'est le montant du sinistre qui reste à la charge de l'assureur. Ainsi, pour tous les sinistres inférieurs à la priorité, le réassureur ne paie rien.

La portée, quant à elle, correspond au montant maximum que le réassureur s'engage à couvrir. Pour tous les sinistres dépassant la somme de la priorité et de la portée, le réassureur ne prend seulement en charge la part du sinistre qui se trouve entre la priorité et le plafond. L'assureur est par conséquent dans l'obligation de prendre à sa charge la franchise (partie sous la priorité) ainsi que la partie du sinistre dépassant le plafond.

1.2.2.1 L'excédent de sinistre ou XS (Excess of Loss – XL) :

La réassurance en excédent de sinistre ne repose pas sur les mêmes principes que ceux des traités proportionnels. Contrairement à la réassurance proportionnelle qui est basée sur un partage des conditions originales entre l'assureur et le réassureur, en réassurance non proportionnelle le réassureur évalue son risque et le tarifie indépendamment des conditions originales.

L'assureur prend ainsi en charge la totalité des sinistres couverts par la branche visée dans le traité jusqu'à la priorité, quelles que soient la somme assurée et les sinistres qui dépassent ce seuil doivent être indemnisés par le réassureur dans la limite de la portée.

Il est fréquent que le besoin de couverture soit trop important pour faire l'objet d'un seul traité, la couverture est ainsi subdivisée en plusieurs tranches (ou layers), où chaque tranche donne lieu à un traité distinct, et dans le cas où le sinistre ou l'événement se produit, les réassureurs des différentes tranches paient leur part respective selon les différentes limites convenues dans les traités.

Dans un traité par excédent de sinistre, le contrat peut être établi deux manières différentes : par risques ou par événements.

Dans le cas d'un excédent de sinistre par risque, le réassureur s'engage à payer à chaque fois qu'un sinistre est supérieur à la priorité dans la limite de la portée pour chaque police donnée.

Dans le cas d'un excédent de sinistre par événement, le réassureur intervient seulement à chaque survenance d'événement (événement lié à une sécheresse, une tempête, ou un autre événement climatique...) et éventuellement sur plusieurs polices. L'événement est défini contractuellement dans le traité :

- Par sa nature : type de catastrophe : tempête, tremblement de terre, inondation, sécheresse... par exemple.
- Par la durée on définit la durée maximale pendant laquelle les dommages imputables à une même cause sont pris en compte dans l'événement. Si l'événement dépasse la durée maximale couverte, alors on ne prendra que la période comprise dans le traité le plus sinistré.
- Dans l'espace, la zone géographique où se produit l'événement couvert est délimitée.

Le montant de l'événement est alors constitué de la somme de tous les sinistres engendrés et dans la limite de la définition de l'événement. Dans ce cas, la priorité et la portée sont appliquées à l'événement et non pas individuellement à chaque sinistre.

Ces événements sont caractérisés par un nombre important de sinistres de faibles intensités pour lesquels le XS par risque n'aurait apporté que très peu ou pas de protection, car le montant individuel de chaque sinistre ne dépasse pas forcément la priorité, alors que le montant global de l'événement peut lui, largement dépasser le seuil de la priorité.

Le montant d'un sinistre ou d'un événement à la charge du réassureur se calcule donc, à partir de la priorité et de la portée, avec la formule suivante :

$$S_{réassureur} = \min(\max(S_{assureur} - d ; 0) ; l)$$

Avec : $S_{réassureur}$ montant du sinistre ou de l'événement à charge du réassureur

$S_{assureur}$: montant du sinistre ou de l'événement

d : la priorité l : la portée

	1	2	3	4	5	Total
Montant du sinistre	4 M€	0.5 M€	19 M€	3 M€	36 M€	62.5 M€
Montant à charge du réassureur	1 M€	0 M€	6 M€	0 M€	6 M€	13 M€
Montant à charge de l'assureur	3 M€	0.5 M€	13 M€	3 M€	30 M€	49.5 M€

Tableau 2 : l'exemple fictif d'un traité en excédent de sinistre par risque 6M XS 3M

Graphiquement, on peut illustrer cette répartition du risque en l'assureur et la cédante de la manière suivante :

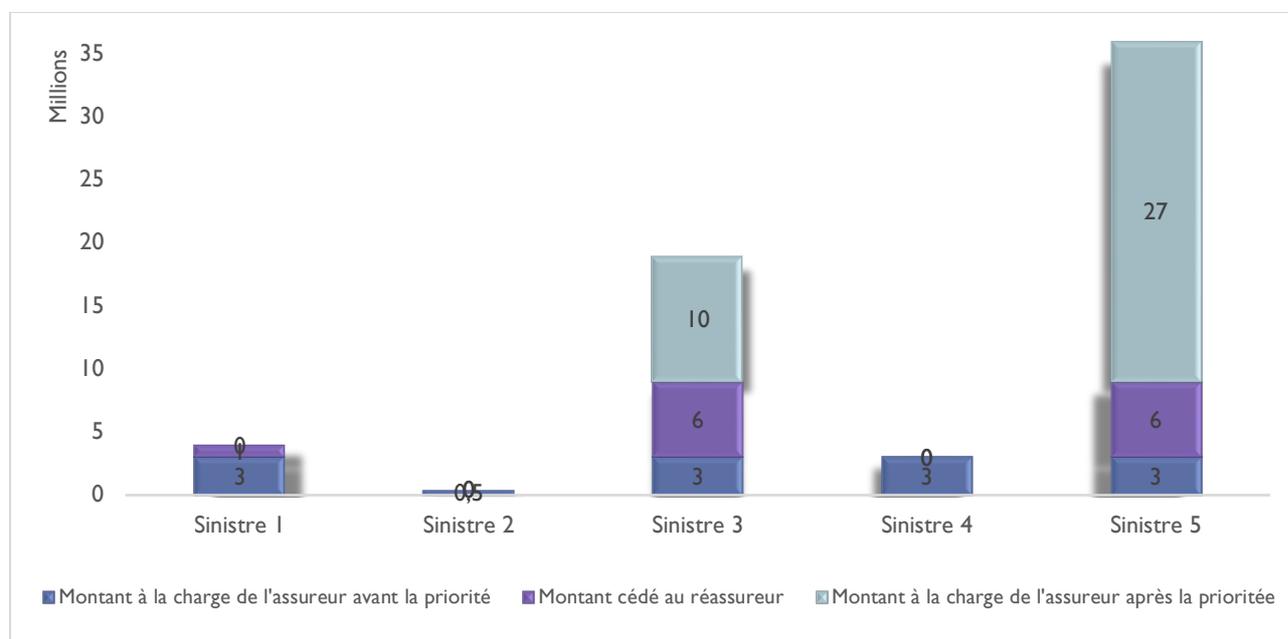


Figure 1 : Répartition de la charge des sinistres selon un traité en XS par risque

1.2.2.2 L'excédent de perte annuelle – Stop-Loss :

Le Stop-loss intervient lorsque la cédante cherche à se prémunir contre les mauvais résultats en s'intéressant directement à ses résultats plutôt qu'en se basant sur les sinistres qui ont eu lieu durant son année d'exercice.

L'excédent de perte annuelle protège l'assureur contre les mauvais résultats en appliquant la priorité et la portée directement au ratio de sinistres sur primes (que l'on notera dans la suite S/P), annuel de la cédante.

$$loss\ ratio\ S/P : \frac{\text{coûts de sinistres}}{\text{total des primes}}$$

La priorité et la portée sont alors exprimées en ratio S/P. La priorité est donc définie comme la sinistralité annuelle que l'assureur conserve à sa charge. Le réassureur s'engage donc à protéger la

cédante d'un montant maximum, le montant dépassant le seuil financier au-delà duquel l'assureur est obligatoirement en perte.

Ce traité offre une protection contre l'accumulation de sinistres sur une durée, et permet à la cédante, que le réassureur intervienne lorsque le ratio de perte excède la priorité et ce dans la limite de la portée.

Mathématiquement, la couverture du réassureur est définie par :

$$S_{réassureur} = \min\left(\max\left(0, \frac{S}{P} - d\right); l\right) \times P$$

Avec S/P , le rapport sinistres sur prime de l'année d'exercice
 P la somme des primes acquises par la cédante

Cas pratique : Prenons l'exemple d'un traité en excédent de perte 150% SL 100% :

Dans le cas où le ratio annuel de S/P est inférieur à 100%, le réassureur n'intervient pas.

Si le portefeuille de la cédante affiche un ratio S/P de 200%, alors le réassureur intervient et verse des récupérations à la cédante de manière à limiter le ratio de pertes à 100%, cette dernière touchera dans la limite de 150 %, 200% - 100 %, soit 100% des primes en récupérations.

Si le portefeuille de la cédante affiche un ratio S/P de 400%. Le réassureur intervient car le ratio S/P de la cédante dépasse la priorité qui est de 100%. La cédante touchera alors dans la limite de 150%, le montant du ratio S/P dépassant la priorité : 400% - 100% = 300%, donc elle touchera 150% des primes en récupérations.

Ce type de traité est généralement utilisé pour les branches qui possèdent une forte sinistralité mais qui est due à un grand nombre de petits sinistres ou encore dans le cas d'une branche en cours de développement très rapide.

Le risque de perte ne doit pas être supporté en totalité par le réassureur, la cédante doit elle-même assumer une partie de sa perte dans la branche protégée par le stop loss, c'est pourquoi les priorités sont très souvent supérieures à un loss ratio de 100% des primes. Dans certains cas, la couverture fournie par le réassureur envers la cédante peut être compromise ou limitée, c'est pourquoi certaines clauses faisant référence à des restrictions ou des limitations sont rédigées contractuellement.

1.2.3 La réassurance proportionnelle

En réassurance proportionnelle, l'assureur verse au réassureur une part de sa prime en échange de laquelle le réassureur prend en charge cette même part des sinistres de l'assureur. Les primes et les sinistres sont partagés proportionnellement entre l'assureur et le réassureur. Le sort de la cédante est alors très lié à celui de son réassureur, ce qui ne permet pas à la cédante d'agir impunément (remboursements excessifs des assurés par exemple) puisque son sort est directement lié à celui du réassureur. Du fait de la proportionnalité du partage des risques, ils ont des intérêts communs en ce qui concerne la performance du portefeuille couvert. Parmi les contrats de réassurance proportionnelle, on trouve les contrats de type Quote-Part, notés QP et les contrats de type excédent de pleins, notés XP.

1.2.3.1 La Quote-Part (QP)

La Quote-Part est la forme de réassurance la plus simple, pour laquelle un taux unique fixé est appliqué à tous les risques, la cédante partage ainsi une proportion équivalente de ses primes et des sinistres de son portefeuille au réassureur.

Avec un traité en quote-part, si on considère $\alpha \in [0; 1]$, le taux cédé au réassureur pour toutes les polices. La cédante cède alors une proportion α de ses primes et de ses sinistres au réassureur moyennant une commission qui correspond à la cession des frais de gestion de la cédante payée par le réassureur. Cette commission est généralement fixe mais peut être variable selon les résultats du ratio S/P.

Tout sinistre de charge S peut alors se décomposer en une partie αS à la charge du réassureur et une partie $(1 - \alpha)S$ à la charge de la cédante. Tout revient pour la cédante à déléguer une proportion α optimale de ses polices au réassureur.

Notons :

- n_s Le nombre de primes de la cédante.
- n_p Le nombre de sinistres de la cédante, alors :

	Risque total	Risque conservé	Risque cédé
Primes	$P = \sum_{i=1}^{n_p} P_i$	$\alpha P = \alpha \times \sum_{i=1}^{n_p} P_i$	$(1 - \alpha)P = (1 - \alpha) \times \sum_{i=1}^{n_p} P_i$
Sinistres	$S = \sum_{i=1}^{n_s} S_i$	$\alpha S = \alpha \times \sum_{i=1}^{n_s} S_i$	$(1 - \alpha)S = (1 - \alpha) \times \sum_{i=1}^{n_s} S_i$

Tableau 3 : Formules de la répartition des risques

Exemple : Prenons un portefeuille fictif d'un groupe d'assurance qui à recourt à de la réassurance avec un Quote-Part comprenant un taux de cession de 70 %.

	Garantie		
	Assuré	Cédé	Conservé
1	18.00	0.7*18 = 12.6	0.3*18=5.4
2	14.00	9.8	4.2
3	11.00	7.7	3.3
4	7.00	4.90	2.10
5	5.00	3.5	1.5
6	3.00	2.10	0.90
7	1.00	0.7	0.3

Tableau 4 : Portefeuille fictif d'un Quote-Part à 70%, limité à un plein de souscription de 13 M€, en M€

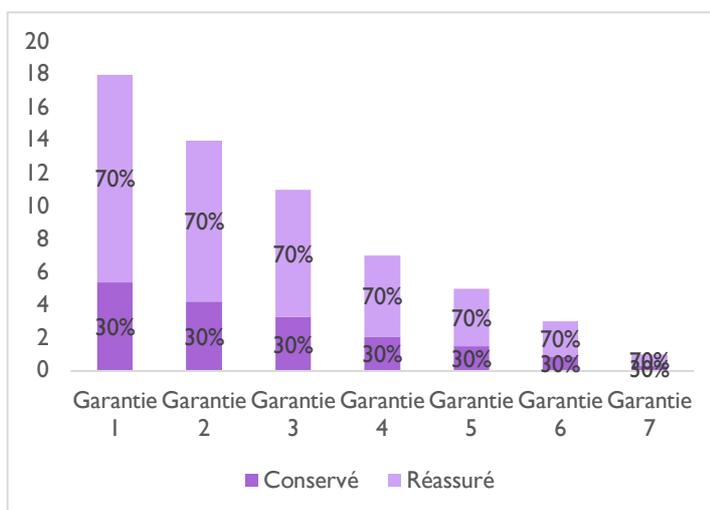


Figure 2 : Représentation d'un quote part avec un taux de cession à 70%

Dans cette forme de réassurance, l'assureur cède la même part sur les risques faibles que sur les risques importants, le profil de portefeuille conservé par le réassureur est semblable au portefeuille initial, seul le niveau des engagements est modifié.

Les traités en Quote-Part sont faciles à appliquer et à gérer à la fois pour le réassureur et l'assureur, cela permet aussi au réassureur de suivre le portefeuille de la cédante et permet une amélioration de la marge de solvabilité de la cédante. En revanche, après application de la réassurance, le portefeuille reste hétérogène, c'est-à-dire que tous les risques acceptés par la cédante font l'objet d'un partage proportionnel entre la cédante et le réassureur, et par conséquent il n'y a aucune distinction entre un risque de faible capital et un autre de plus haute intensité.

De plus, le Quote-Part n'offre pas de lissage des résultats dans le temps et offre un transfert considérable du montant des primes de la cédante vers le réassureur. Ainsi pour corriger cette non-homogénéité dans la réassurance, nous devons choisir une autre forme de réassurance, le traité en excédent de plein.

1.2.3.2 L'excédent de plein

Dans cette forme de réassurance, le taux de cession est calculé police par police sur la base de la somme assurée (noté SI) et non des sinistres, on parle alors d'excédent de plein par risque. L'intervention du réassureur se fait uniquement si la somme à assurer dépasse le plein de rétention (noté R), dans la limite du plein de souscription, formellement le taux de cession pour toutes les polices i se définit tel que :

$$\tau_i = \max(0, 1 - \frac{R}{SI_i})$$

L'excédent de plein peut aussi être vu par événement, dans ce cas le réassureur couvre un ensemble de sinistres individuels reliés par un même fait générateur et offre ainsi à l'assureur une protection contre l'accumulation de sinistres par événement. L'événement est défini contractuellement dans sa nature, dans l'espace et dans le temps.

Finalement, dans le cas de l'excédent de perte annuelle, l'assureur cherche à se prémunir contre les mauvais résultats puisque ce traité offre uniquement à l'assureur une protection contre l'accumulation de sinistres sur une durée. Le plein de rétention est donc défini comme la sinistralité annuelle que l'assureur conserve à sa charge.

Prenons l'exemple d'un traité en excédent de plein par sinistre avec un plein de rétention à $R = 6M\text{€}$, une somme à assurer $SI = 10M\text{€}$

	Capital			Prime			Charge du sinistre		
	Assuré	Cédé	Conservé	De base	Cédé	Conservé	Totale	Cédée	Conservée
1	18.00	10.00	8.00	0.75	0.42	0.33	12.00	6.67	5.33
2	7.00	1.00	6.00	0.35	0.05	0.30	10.00	1.43	8.57
3	3.00	0.00	3.00	0.15	0.00	0.15	2.00	0.00	2.00

Tableau 5 : Exemple d'un traité en excédent de plein par sinistre

$C_{cédé} = C_{assuré} - \text{Min}(C_{assuré}; 6M\text{€})$	$P_{cédée} = \frac{C_{cédé}}{C_{assuré}} \times P_{base}$	$S_{cédée} = \frac{C_{cédé}}{C_{assuré}} \times S_{totale}$
--	---	---

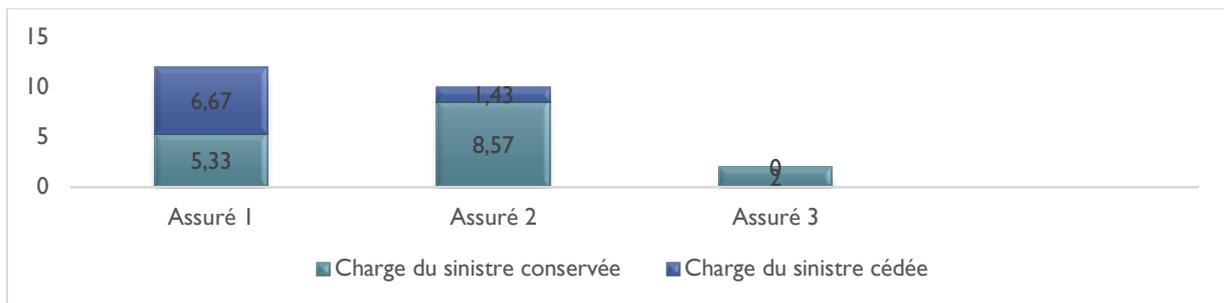


Figure 2 : Répartition de la charge cédée au réassureur pour un traité en excédent de sinistre par risque avec un plein de rétention à 6 M€, et une somme à assurer à 10 M€

L'objectif d'un tel traité de réassureur par rapport à un traité en Quote-Part, est surtout de limiter l'impact des fluctuations aléatoires sur le plein de rétention ainsi que sur la charge cédée au réassureur. Cela permet aussi une homogénéisation de la charge cédée au réassureur à la fois sur les sinistres à faible risque et sur les sinistres à risque plus élevé. Cela permet ainsi de niveler le risque en portefeuille, et permet à l'assureur de garder un engagement limité tout en ne cédant pas les petits risques pour lesquels il est relativement facile d'estimer le sinistre maximum possible. En revanche, l'assureur est tout de même très affecté en cas de sinistre très important.

Dans les deux cas, la participation du réassureur est proportionnelle aux primes et aux pertes de la cédante.

De plus, un tel traité permet une meilleure gestion des risques et des capacités. En effet, une quote-part serait moins adapté dans certaines situations que dans d'autres. Par exemple si on prend l'exemple d'une entreprise qui veut conserver un plein de rétention à 100 000€, et un taux de rétention à 99%, alors un traité en Quote-part nous garantira une capacité à 10M€. La question qui se poserait alors serait : pourrait-on se placer dans un tel traité, là où la cédante ne se basera pas sur un facteur d'exposition ?

1.2.3.3 Conclusion sur la réassurance proportionnelle :

Les traités en Quote-part permettent une meilleure gestion tant en termes de simplicité qu'en termes d'utilisations. Les entreprises qui font généralement appel à ce type de traités, sont typiquement des nouvelles sociétés qui ont des niveaux de fonds propres relativement limités et une appréciation difficile des résultats techniques des premières années, l'appui d'un réassureur leur permet ainsi d'augmenter leur marge de solvabilité.

Pour l'assureur, l'excédent de plein peut présenter de nombreux avantages : nivellement du portefeuille, conservation homogène, capacité, aliment conservé puisque beaucoup de petits et moyens risques sont conservés et la gestion se fait risque par risque. Mais pour le réassureur, l'excédent de plein présente un inconvénient majeur, en effet la différence de résultats entre la partie conservée et la partie réassurée, qui est essentiellement constituée par des sinistres à fort enjeu et dont on traite uniquement les pointes des risques, est déséquilibrée. S'il souscrit fréquemment des risques supérieurs à la capacité de son excédent de plein, l'assureur à tout intérêt à négocier avec ses réassureurs un deuxième excédent de plein. L'inconvénient majeur de ce type de traité est que le traitement est complexe et donc coûteux en l'absence de ressources informatiques.

1.3 Les clauses des traités

Dans chaque traité, il existe différentes clauses que l'on doit prendre en compte dans notre étude.

Ces clauses jouent un rôle essentiel pour définir les termes et conditions de réassurance. Elles servent à établir des règles de fonctionnement et permettent de garantir une compréhension mutuelle entre la cédante et le réassureur afin de définir les obligations et les droits de chaque partie.

1.3.1 La clause de reconstitution de garantie

Dans un traité non proportionnel, le réassureur couvre la cédante par le biais d'une portée d'un éventuel sinistre ou événement qui surviendrait, mais la cédante ne doit pas se trouver à découvert si jamais d'autres sinistres survenaient et absorbent toute la portée.

C'est pourquoi, une clause de reconstitution qui sert à renouveler la couverture de l'assureur après l'apparition d'un sinistre est mise en place, autrement dit cela sert à maintenir le même niveau de protection à la cédante. Ainsi dès lors que la portée est totalement ou partiellement utilisée par un sinistre, la clause de reconstitution s'applique.

Lorsque le réassureur accorde N reconstitutions de garantie, il s'engage à payer au maximum la garantie initiale ainsi que les N reconstitutions supplémentaire, soit N+1 fois la portée durant la période transcrite par le traité. Le nombre de fois peut être illimité.

La prime de réassurance versée par la cédante au réassureur ne constitue qu'une prime de base destinée à couvrir la portée. Au fur et à mesure de la survenance des sinistres, la cédante doit payer une surprime qui correspond à un renouvellement de la partie de la portée qui a été consommée, c'est la prime de reconstitution.

Par exemple : Supposons que la cédante dispose d'une assiette de prime de 2 000 000€, d'un taux de prime de réassurance de 1% des primes et d'une couverture en excédent de sinistre par risque de 2M XS 2M avec 1 reconstitution à 100% de la portée alors :

Alors, on a que le nombre de reconstitutions possible est de 1 (portée initiale) + 1 (une reconstitution) fois la portée du traité soit une garantie annuelle de 4 000 000€.

La prime de réassurance $\pi = 1\% \times 2\,000\,000\text{€} = 20\,000\text{€}$

Sinistres	Montant à charge de la cédante	Montant à charge du réassureur	Etat de la garantie annuelle	Reconstitution
3 M€	2 M€	1 M€	4 - 1 = 3 M€	$\pi \times 100\% \times \frac{1\,000\,000 \text{ (cession)}}{2\,000\,000 \text{ (portée)}}$
4 M€	2 M€	2 M€	1 M€	$\pi \times 100\% \times \left(\frac{1\,000\,000}{2\,000\,000}\right)$
4 M€	3 M€	1 M€	0 €	0 €

Figure 3 : Exemple d'application de reconstitutions des primes

Exemple 2 : On va maintenant supposer un encaissement de prime total à 10 00 000€, un taux de prime de réassurance à 2% du montant des primes, et une couverture en excédent de sinistre par risque 2 000 000 XS 1 000 000, avec 2 reconstitutions, une à 100% de la prime de réassurance et l'autre à 200%.

On en déduit que la garantie annuelle maximale de couverture de la cédante est de trois fois la portée du traité (une première portée initiale ainsi que deux autres reconstitutions), soit 6 000 000€.

La prime de réassurance $\pi = 2\% \times 10\,000\,000\text{€} = 200\,000\text{€}$

Sinistres	Montant à charge de la cédante	Montant à charge du réassureur	Etat de la garantie annuelle	Reconstitution
2 M€	1 M€	1 M€	5 M€	$\pi \times 100\% \times \frac{1\text{ M€}(\text{cession})}{2\text{ M€}(\text{portée})}$
3 M€	1 M€	2 M€	3 M€	$\pi \times 100\% \times \frac{1\text{ M€}(\text{cession allouée à la 1ère portée})}{2\text{ M€}}$ $+ \pi \times 200\% \times \frac{1\text{ M€}(\text{cession allouée à la 2ème portée})}{2\text{ M€}}$ (1ère reconstitution à 100%)
6 M€	4 M€	2 M€	1 M€	$\pi \times 200\% \times \frac{1\text{ M€}(\text{cession allouée à la 2ème portée})}{2\text{ M€}}$ (Les 2 reconstitutions ont été réalisées)

Figure 4 : Exemple d'application de reconstitutions

1.3.2 La clause de franchise Aggregate ou clause de franchise annuelle

La couverture de la cédante peut être limitée par un plafond annuel appelé Annual Aggregate Limit (AAL), ou Annual Aggregate Deductible (AAD).

Cette clause a pour but de restreindre l'engagement du réassureur et de diminuer le coût de la réassurance. En effet, sans l'application de cette clause, le réassureur s'engage à prendre à sa charge chaque sinistre ou événement dont le montant dépasse la priorité dans la limite de la portée et ceci de manière unitaire. Tandis qu'en cas d'application d'une AAL ou AAD, l'intervention du réassureur est soumise à une condition supplémentaire qui vient impacter sa participation :

- L'AAD fonctionne comme une franchise annuelle, appliqué à la somme des montants totale qui seraient normalement dus par la réassurance sans application de cette clause. La cédante garde à sa charge les premiers sinistres dont le montant est supérieur à la priorité du traité, jusqu'à ce que leur cumul dépasse le montant de la franchise annuelle. Une fois l'AAD consommée, le traité fonctionne de manière « normale ». Mathématiquement, si on considère N sinistres alors, la charge cédée au réassureur se traduit par :

$$\text{Cession} = \max(0, (\sum_{i=1}^N \min(\text{portée}, \max(0, \text{sinistre } i - \text{priorité}))) - \text{AAD})$$

En effet, cette formule ne peut être négative, car cela signifierait que l'AAD n'est que partiellement consommée et donc l'engagement du réassureur n'entre pas en vigueur. En revanche, lorsque le résultat est positif, l'AAD est complètement consommée et le résultat correspond au montant à charge du réassureur.

- L'AAL fonctionne comme une limite appliquée à la somme annuelle des montants dus par le réassureur, c'est-à-dire des montants sur lesquels ont déjà été appliquées la priorité et la portée prévue par le traité. Dans le cas d'une AAL, le nombre de reconstitutions est omis puisqu'il serait redondant avec cette clause, et l'AAL serait alors (N+1) x portée, avec N le nombre de reconstitutions. La charge cédée au réassureur se traduit donc par :

$$\text{Cession} = \min(\sum_{i=1}^N \min(\text{portée}, \max(0, \text{sinistre } i - \text{priorité})), \text{AAL})$$

Mathématiquement, dans le cas d'un traité avec AAL et AAD, si on considère N sinistres, avec r le montant des récupérations associé à chaque sinistre, la cédante touchera ainsi des récupérations se cumulant à un montant de :

$$Cession = \min(\max((\sum_{i=1}^N r_i) - AAD; 0); AAL)$$

Dans le cas où un traité comporte à la fois une clause de reconstitution et une clause de franchise annuelle, l'ordre de fonctionnement de ces deux clauses doit clairement être précisé, car :

- Si la reconstitution de garantie est appliquée en premier, on calcule toutes les reconstitutions au fur et à mesure des sinistres, puis à la fin de l'année on déduit l'AAD de la charge que le réassureur devra céder à la cédante.
- Si l'AAD est appliquée en premier, alors, on le déduit sur les premiers sinistres puis une fois que la capacité de l'AAD est totalement consommée, on ne peut céder au réassureur les sinistres qui dépassent la priorité dans la limite du nombre de reconstitutions de la portée.

Exemple de fonctionnement de la clause : Soit un traité en excédent de sinistre par risque : 3M XS 1M, avec une AAL à 4M sur la première tranche, et illimité XS 4M sur la seconde tranche avec une AAD à 1M.

Montant du sinistre	Montant à charge de la cédante avant AAL/AAD	Montant à charge du réassureur sur avant AAL	Consommation de l'AAL	Montant à charge du réassureur sur la tranche 1 après AAL	Montant à charge du réassureur sur la tranche 2 avant AAD	Consommation de l'AAD	Montant à charge du réassureur sur la tranche 2 après application de l'AAD	Montant à charge de la cédante après AAL et AAD
4,5 M€	1 M€	3 M€	3 M€ sur 4 M€	3 M€	0,5 M€	0,5 M€ sur 1 M€	0 €	1 M€
7 M€	1 M€	3 M€	4 M€ sur 4 M€	1 M€	3 M€	1 M€ sur 1 M€	2 M€	3 M€
9 M€	1 M€	3 M€	4 M€ sur 4 M€	0 €	5 M€	1 M€ sur 1 M€	5 M€	4 M€

Tableau 6 : Exemple sur le fonctionnement des aggregates

1.3.2.1 Clause de stabilité

La clause de stabilité a pour objectif de répartir le surcoût engendré par l'inflation entre la cédante et le réassureur. Elle permet de conserver le niveau économique de la charge des sinistres pour des années de survenance successives. Sans clause de stabilisation, dès lors qu'un sinistre atteint la priorité, uniquement le réassureur était impacté par l'aggravation des coûts due à l'inflation jusqu'à ce que le sinistre atteigne une valeur certaine et définitive. Ce type de clause est généralement présente dans les traités de réassurance non proportionnels qui couvrent une branche à long terme comme la responsabilité civile.

En effet, avec le temps, la charge du sinistre ainsi que les bornes d'un traité perdent en signification à cause de l'érosion monétaire et de l'inflation. Afin que l'on puisse garder la même application de la couverture du traité de survenance sur un sinistre dans le temps, il est nécessaire d'indexer les bornes du traité sur un indice économique correspondant à l'évolution des coûts des sinistres dans la branche. Par exemple, indice trimestriel Insee des salaires horaires des ouvriers relevant de la nomenclature des activités pour la santé humaine (id : I567402), pour la RC Automobiles par exemple.

L'indice de base est défini comme étant l'indice à une date donnée.

Les règlements et provisions sont alors stabilisés afin de ramener leurs montant comme s'ils avaient été effectués l'année de survenance du sinistre uniquement si le niveau de variation de l'indice de référence dépasse un certain seuil p (généralement 10%), par rapport à l'indice de base.

Les règlements stabilisés de chaque année ainsi que la provision restante stabilisée de chaque sinistre sont agrégés pour ainsi définir une nouvelle charge stabilisée du sinistre. On pourra ainsi déterminer deux nouvelles priorités et portées stabilisées.

Ce sont ces priorités et portées stabilisées qui seront alors appliquées à la charge actuelle du sinistre pour calculer la cession en réassurance. Chaque sinistre s'il doit être stabilisé dispose de bornes unique et spécifique qui lui sont propres. Si l'évolution de l'indice de référence de l'année à stabiliser dépasse le seuil contractuel, alors les règlements et la provision restante sont stabilisés à partir de la formule suivante :

$$\frac{\text{Montant de règlement ou provision de l'année comptable N} \times \text{indice de base}}{\text{indice de référence de l'année comptable N}} \mathbb{1}_{\frac{\text{indice}_{base}}{\text{indice}_i} > p}$$

Les limites du traité seront stabilisées, en multipliant chaque borne du traité par le coefficient de stabilisation qui est décrit par : $\frac{\text{charge totale du sinistre réelle}}{\text{charge totale du sinistre stabilisée}}$

Avec : Charge totale du sinistre réelle stabilisée = $\sum_{i=1}^n R'_i + P_{\text{stabilisée}}$

Où :

- $indice_{base}$ est l'indice de base
- $indice_i$ est l'indice de référence de l'année i
- R_i est le montant des règlements de l'année i
- R'_i est le montant des règlements stabilisé de l'année i tel que

$$R'_i = \begin{cases} \frac{R_i \times indice_{base}}{indice_i} & \text{si l'évolution de l'indice}_i > \text{seuil} \\ R_i & \text{sinon} \end{cases}$$

- n : le dernier règlement du sinistre
- P le montant de la provision restante du sinistre
- $P_{\text{stabilisée}}$ le montant stabilisé de la provision tel que :

$$P_{\text{stabilisée}} = \begin{cases} \frac{P \times indice_{base}}{indice_i} & \text{si l'évolution de l'indice}_i > \text{seuil} \\ P & \text{sinon} \end{cases}$$

Exemple : Soit un sinistre qui est survenu en 2016, la cédante est couverte par un traité en excédent de sinistre 100 000 XS 20 000 l'indice de base du traité est de 100, et le traité comprend une clause de stabilité avec un seuil de 10% de variation de l'indice de référence :

Année comptable	Règlement de l'année	Provision restante	Indice de base	Règlement stabilisés	Provision stabilisée
2016	10 000 €	-	100	10 000 €	
2017	30 000 €	-	103	30 000 €	
2018	6 000 €	-	105	6 000 €	
2019	13 000 €	-	107	13 000 €	
2020	10 000 €	-	111	$\frac{100(\text{indice de référence})}{111(\text{indice de base})} \times 10\,000$ = 9 009€	
2021	15 000 €	100 000€	114	13 158 €	87 719 €
Total	84 000 €	100 000		81 167€	87 719 €

Tableau 7 : Exemple de stabilisation

- La charge totale du sinistre est de : 184 000 €.
- La charge stabilisée totale du sinistre est donc de 168 886 €.
- Le coefficient de stabilisation vaut donc 1,08.

On peut donc déduire les nouvelles bornes du traité : 108 000 XS 21 600.

Le réassureur aura donc à sa charge : 108 000€. La cédante aura à sa charge : 21 600 + 54 400 = 76 000€.

1.3.2.2 Clause sur les rentes

Lors d'un préjudice corporel, la victime reçoit une indemnisation pour couvrir ses frais futurs et pour compenser ses pertes. Celle-ci peut se faire sous deux formes : soit en rente, soit en capital.

Lorsque l'indemnisation est sous la forme d'un capital, le règlement est unique, tandis que sous forme de rente, un versement régulier pendant une durée donnée et une somme d'argent prédéfinie, c'est l'arrérage, et il est versé à la victime. Le versement peut se faire de manière annuelle, dans ce cas-là on parle d'annuité, mais aussi mensuellement ou trimestriellement.

A la date de la transaction, nous devons calculer une somme d'argent qui prend en compte les intérêts financiers et l'espérance de vie de la victime qui sera nécessaire à la date d'attribution de la rente pour la payer jusqu'à l'extinction de celle-ci, c'est le capital constitutif de la rente (CCR).

Le capital constitutif de la rente (CCR), se calcul lors de la constitution de la rente grâce aux éléments suivants :

- L'arrérage de la rente initiale
- Un taux d'actualisation défini dans le traité
- L'âge de la victime
- Une loi de survie en invalidité définie dans le traité (par exemple TD 88-90 homologuée par l'arrêté du 27 avril 1993)

La valeur équivalente à un euro réévaluée à l'inflation est déterminée par la formule suivante :

$$K = \sum_i \frac{\prod_{j=1}^i (1 + inf_j)}{(1 + \tau)^i} \times iPx$$

Avec :

- inf_j le taux d'inflation à l'année j
- τ le taux d'actualisation prédéfini dans la clause
- iPx la probabilité de survie dans i années du crédirentier d'âge x

Une clause de rente a été introduite pour définir tous les aspects relatifs aux rentes dans le cadre d'un traité de réassurance.

Certains traités non proportionnels fonctionnent en rachat de rentes et d'autres en suivi.

- Dans le cas de d'un rachat de rente : Lorsqu'une rente est constituée, la méthode à retenir pour calculer l'indemnisation que doit prendre en compte le réassureur est celle du rachat de la rente par débit immédiat du réassureur. La rente n'est alors plus considérée dans la durée, mais à sa date d'attribution par le biais du CCR, qui est considéré comme montant unique correspondant à la totalité de la rente. Le réassureur se doit donc de payer la totalité de la part dépassant la priorité stabilisée.

En effet, le réassureur se voit dans l'obligation de participer à l'indemnisation de la rente par le versement de sa part du capital constitutif de rente. En revanche le réassureur ne sera donc plus engagé sur le suivi de la rente vis-à-vis de la cédante, et toute évolution notable sera à prendre en charge uniquement par la cédante. Toutefois, il existe des cas de révisions sur des durées limitées, durant laquelle le montant de la rente sur laquelle était engagé le réassureur peuvent se voir évoluer, comme lorsqu'il y a aggravation de l'état de santé et/ou de la situation du crédientier, dans ce cas un capital constitutif complémentaire est calculé sur lequel le réassureur sera engagé ou encore lors du décès du crédientier. En effet, en cas de décès de celui-ci, pendant une certaine période prédéfinie contractuellement suivant la date de constitution de la rente, la charge finale sur laquelle le réassureur sera engagé est ramenée au cumul des arrérages versés par la cédante.

- Dans le cas d'un suivi de rentes, le réassureur participe au fur et à mesure aux versements des arrérages de rentes jusqu'à l'arrêt de la rente et/ou du décès de la victime. Le suivi peut se faire de manière proportionnelle, le partage des arrérages selon la proportion (CCR – priorité) / portée, la priorité est alors calculée via le biais de la clause de stabilité juste avant le versement de la rente. Mais le suivi, peut aussi se faire de manière additionnelle, la rente est alors considérée dans la durée et nous ne différencions pas le paiement des arrérages des autres paiements relatifs au sinistre. Le réassureur commence donc à payer les arrérages dès lors que leur paiement cumulé dépasse la priorité stabilisée et ceux jusqu'à atteindre la portée stabilisée.

Il existe aussi d'autres clauses, comme celle de la participation aux bénéfices et de « No claim Bonus » par exemple, mais celles-ci ne sont pas incluses dans les traités XS de BPCE Assurances IARD.

1.3.2.3 Présentation du contexte de renouvellement et du traité étudié pour BPCE Assurances IARD

Les traités de réassurance ont tous une période d'effet, cela peut varier en fonction des réassureurs, mais pour la plupart, ils se renouvellent chaque année. En France, la période de renouvellement la plus importante est celle s'étalant d'octobre à fin décembre. C'est durant cette période que les traités sont revus.

La cédante reçoit alors des propositions de prix et de placement par des réassureurs ou des courtiers et décide de se placer avec l'un plutôt que l'autre en fonction de la protection prévue par le traité ainsi que de son coût de revient. Il est alors important pour une cédante de pouvoir évaluer le coût d'une protection afin de mettre en place des négociations qui permettront par la suite de réduire le coût du traité. Ainsi la tarification avant acceptation est un des indicateurs importants nécessaires à la décision de souscription d'un traité.

Les traités de réassurance non proportionnels de BPCE Assurances IARD sont classés principalement par type de risque : RC AUTO, Force de la Nature...

Dans notre étude, nous allons nous intéresser uniquement à la responsabilité civile auto.

Le traité que nous choisissons d'étudier dans l'objectif d'une possible mise en application pour BPCE Assurances IARD est un traité en excédent de sinistre par risque qui comporte les tranches suivantes :

- Tranche 1 : 3 M€ XS 6 M€.
- Tranche 2 : 3 M€ XS 9 M€.
- Tranche 3 : illimitée XS 12 M€.

Le traité comporte une clause de rachat à 1,50%, c'est-à-dire que tous les sinistres en rente seront recalculés avec un taux d'actualisation de 1,50%, la table de mortalité utilisée étant la table

réglementaire TD 88-90. Une clause de stabilisation est aussi ajoutée au traité, et l'indice de stabilisation utilisé est l'indice Insee des salaires horaires ouvriers relevant de la nomenclature des activités pour la santé humaine (id : 15677402), le seuil de déclenchement de cette clause est une variation de plus de 10% de l'indice de référence par rapport à l'indice de base.

2 Sinistralité

Le traité que nous souhaitons étudier pour une possible mise en application chez BPCE Assurances IARD, est un traité en excédent de sinistre pour la branche RC Auto.

Dans cette partie nous allons ainsi définir les différents types de branche auquel le portefeuille d'un assureur peut être confronté ainsi que différentes notions spécifiques au type de branche dont l'actuaire devra évaluer lors l'estimation de provisions. Nous verrons ensuite le cas des indemnisations des victimes en responsabilité civile.

2.1 Type de branche

On peut distinguer deux notions de branches d'assurance différentes selon la durée qui sépare la date de survenance d'un sinistre, la date de déclaration de ce dernier et la date de son règlement.

Les branches courtes, qui sont les branches d'assurances pour lesquelles la durée entre la date de déclaration d'un sinistre, l'estimation du montant à indemniser et le règlement effectif est dite courte, de manière générale cela n'excède jamais une durée de plus d'un an. Les branches courtes sont liées à la propriété, notamment avec des couvertures couvrant les risques d'incendies ou de vol par exemple.

Les branches longues, qui sont les branches d'assurances pour lesquelles la durée entre la date de déclaration d'un sinistre, l'estimation du montant à indemniser et le règlement effectif est dite longue, la durée peut généralement aller au-delà de quelques années. La responsabilité civile est une branche d'assurance réputée être longue dans le sens où l'assureur est encore amené à payer des indemnités relatives aux sinistres qui sont passés plusieurs années auparavant. Cela peut s'expliquer spécifiquement par le temps d'attente entre la survenance du sinistre et la décision de la cour de justice définitive et par le fait que beaucoup de sinistres ne sont déclarés que quelques années après leur année de survenance. L'actuaire doit d'autant faire plus attention au provisionnement du sinistre que l'estimation du coût à la date de déclaration peut être très différente de l'indemnisation effective qui sera payée quelques années plus tard. Des provisions sont donc constituées et suivies tout le long du traitement du sinistre afin de permettre une gestion adéquate de celui-ci.

2.2 Notion de provisions IBNYR et IBNER

Dans la partie précédente, nous avons vu que les paiements de sinistres peuvent s'étaler dans le temps pour des raisons judiciaires ou intrinsèques. Ainsi, à une date d'observation donnée il reste des provisions dont il est nécessaire d'estimer.

En effet, lorsqu'un sinistre est déclaré à une compagnie d'Assurances, cette dernière a tout intérêt à appréhender au plus juste les conséquences qui en résulteraient. La compagnie constitue ainsi un dossier sinistre, dans lequel elle consigne tous les éléments utiles à décrire le sinistre que son assuré vient de subir. Une composante importante de cette description est l'évaluation pécuniaire du sinistre. Pour les sinistres à développement long, la première évaluation est rendue difficile puisque le montant dû lui sera payé le plus souvent en plusieurs fois et sur une longue période. Pour faciliter l'analyse du

montant de sinistre, l'assureur sépare donc la part payée (connue) et la part restant à payer. Avec le temps, l'évaluation s'affine. Dans la réalité, l'évaluation initiale n'est pratiquement jamais exactement égale à la somme de tous les paiements. Ainsi, l'évaluation totale du sinistre est en suspens et évolue au cours du temps. En effet, les paiements de sinistres peuvent s'étaler dans le temps pour des raisons judiciaires (détermination des responsabilités, des indemnités) ou pour des raisons intrinsèques (consolidation de sinistres corporels graves par exemple). A une date d'observation donnée (par exemple lors de la tarification d'un traité), il reste des inconnues dont l'actuaire devra évaluer l'impact :

- Les sinistres qui sont survenus mais qui sont non encore déclarés à la date d'évaluation. Ces sinistres imposent la constitution de provisions dites Incurred but Not Yet Reported (IBNYR).
- Les sinistres qui sont survenus, et qui ont été déclarés à la date d'évaluation mais insuffisamment provisionnés. Ces sinistres imposent la constitution de provisions appelées Incurred But Not Enough Reserved (IBNER), ils correspondent à la différence entre le montant ultime des sinistres rattachés et observables dans la période étudiée et le montant courant de leur évaluation (payés + suspens).

Mais en réalité, cela peut différer, par exemple :

Pour la branche incendie, on peut noter des développements différents selon les pays ou les marchés. Aussi, pour les sinistres importants, y compris les catastrophes naturelles, les paiements peuvent s'étaler sur des périodes plus longues qu'une seule année. De même pour les branches longues, il existe de fortes différences entre les marchés. Par exemple pour la RC automobile, l'existence ou non de barèmes conditionne très fortement les développements observés. Il existe également au sein d'un même marché des comportements de développement très différents selon les cédantes, selon qu'elles privilégient ou non les règlements amiables ou contentieux.

Différentes natures d'indemnisation peuvent, bien entendu, se trouver présentes dans un même sinistre : les indemnités matérielles qui compensent la destruction d'un bien, les indemnités pour préjudice corporel touchant l'intégrité du corps humain et les indemnités pour pertes financières à la suite d'un sinistre matériel (pertes d'exploitation), d'un mauvais conseil (responsabilité civile professionnelle).

2.3 La méthode Chain Ladder

La méthode de Chain Ladder est souvent utilisée dans le cadre de la modélisation des coûts des sinistres en réassurance pour estimer les réserves de sinistres. Les réserves de sinistres correspondent à la somme des montants attendus pour les sinistres qui ont été signalés mais pas encore réglés, ainsi que pour les sinistres qui n'ont pas encore été signalés.

La méthode de Chain Ladder est utilisée dans ce contexte car elle permet d'estimer les paiements futurs de sinistres à partir des paiements passés, en supposant que les ratios de paiements cumulatifs sont constants dans le temps. Cette hypothèse est raisonnable pour les sinistres ordinaires, mais peut-être moins appropriée pour les sinistres extrêmes.

Cependant, la méthode de Chain Ladder reste un outil utile pour estimer les réserves de sinistres en réassurance, car elle est simple à utiliser et fournit des résultats rapides et fiables dans de nombreuses situations. Dans notre étude, elle sera également utilisée en combinaison avec d'autres méthodes plus sophistiquées pour améliorer la précision des estimations.

La méthode de Chain Ladder est basée sur des techniques de régression qui permettent de projeter les paiements futurs de sinistres à partir des paiements passés. Cela signifie que :

- Les sinistres sont considérés comme des variables aléatoires indépendantes identiquement distribuées selon une loi inconnue.
- Les paiements cumulatifs jusqu'à une année donnée suivent une loi de probabilité conditionnelle qui dépend de l'année d'évaluation.
- Les ratios de paiements cumulatifs, qui représentent les proportions de paiements cumulatifs entre deux années consécutives, sont constants au fil du temps.

En supposant que ces hypothèses sont vérifiées, la méthode de Chain Ladder peut être décrite en quatre étapes :

- Calcul des paiements cumulatifs pour chaque année :

Soit $P(i, j)$ le paiement cumulatif pour la j -ième année, où i est le nombre d'années depuis la première année de l'évaluation. On calcule les paiements cumulatifs pour chaque année en utilisant la relation :

$$P(i, j) = P(i, j - 1) + S(i, j)$$

Où $S(i, j)$ est le montant de paiement pour la j -ième année et $P(i, j - 1)$ est le paiement cumulatif jusqu'à la $(j-1)$ -ième année.

- Calcul des ratios de paiements cumulatifs :

On calcule les ratios de paiements cumulatifs comme suit : $R(j) = P(i, j) / P(i, j - 1)$

Où $R(j)$ est le ratio de paiements cumulatifs pour la j -ième année.

- Projection des paiements futurs en utilisant les ratios de paiements cumulatifs :

On projette les paiements futurs en utilisant les ratios de paiements cumulatifs. Pour cela, on calcule les paiements futurs pour chaque année $k > j$ comme suit : $P(i, k) = R(j) \times P(i, k - 1)$

Où $P(i, k - 1)$ est le paiement cumulatif jusqu'à l'année $k-1$.

- Estimation des réserves de sinistres en additionnant les projections de paiements futurs :

On estime les réserves de sinistres en additionnant les projections de paiements futurs pour chaque année $k > j$: $R(i) = \sum_{k=j+1}^n P(i, k)$

Où $R(i)$ est l'estimation de la réserve pour l'année i et n est la dernière année de l'évaluation.

Ainsi, dans notre étude, nous allons estimer la charge ultime des sinistres individuels en calculant les facteurs de développement via la méthode de Chain-Ladder.

Pour se faire, nous allons construire deux triangles :

- Dans le 1er triangle : la colonne j (année de développement j) contient la charge indexée des sinistres connus de l'année $j-1$ ou avant. On note cette charge $C'_{i,j}$ tel que :

$$C'_{i,j} = \sum_s C_{i,j,s} \times I_{s,j} , \quad \text{où } I_{s,j} = \begin{cases} 1 & \text{si } s \text{ connu en } j - 1 \\ 0 & \text{sinon} \end{cases}$$

Et $C_{i,j,s}$ est la charge indexée du sinistre s survenue en i vue en j

- Dans le 2ème triangle : la colonne j (année de développement j) contient la charge indexée des sinistres connus de l'année j ou avant. On note cette charge $C_{i,j} = \sum_s C_{i,j,s}$

Les coefficients de passage d'une année de développement j à une année de développement $j+1$ ou facteurs de développement f_j sont déterminés comme suit : $f_j = \frac{\sum_{i=1}^{n-j} C'_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}$.

Les facteurs de développement cumulés F_k sont définis comme suit : $F_k = \prod_{j=k}^J f_j$, pour $k \in \llbracket 1:J \rrbracket$.

On construit alors les triangles de liquidation tel que :

Année d'origine	Délai de règlement								
	0	1	...	j	...	$n-i$...	$n-1$	n
0	$C_{(0,0)}$	$C_{(0,1)}$...	$C_{(0,j)}$	$C_{(0,n-1)}$	$C_{(0,n)}$
1	$C_{(1,0)}$	$C_{(1,1)}$	
...		
i	$C_{(i,n-i)}$			
...				
$n-j$	$C_{(n-j,j)}$					
...						
$n-1$	$C_{(n-1,0)}$	$C_{(n-1,1)}$							
n	$C_{(n,0)}$								

Il suffit alors de valider les hypothèses de Chain-Ladder pour ainsi en déduire la charge ultime $C_{i,j,s}^u$ du sinistre s survenu en i est calculée en appliquant à la charge indexée le facteur de développement cumulé correspondant :

$$C_{i,j,s}^u = C_{i,j,s} \times F_{n-i+1}, \text{ où } F_k \text{ est le facteur de développement cumulé de l'année } k = n-i+1$$

En somme, la méthode de Chain Ladder utilise des techniques de régression pour projeter les paiements futurs de sinistres à partir des paiements passés, en supposant que les ratios de paiements cumulatifs sont constants au fil du temps.

2.4 Notion de tranches travaillantes et non travaillantes.

Précédemment, nous avons vu que les programmes de réassurance peuvent être sectionnés en plusieurs sous tranches, cette segmentation est principalement expliquée par l'appétence au risque des réassureurs. En effet certains réassureurs préfèrent couvrir des tranches basses fréquemment touchées parce que leurs priorités sont peu élevées (généralement pour un traité XS Auto par exemple), c'est ce qu'on appellera une tranche travaillante. A l'inverse d'autres réassureurs préfèrent couvrir des tranches hautes, avec une sinistralité moins élevée, c'est ce qu'on appellera une tranche non travaillante. La notion de tranche non travaillante reste tout de même très subjective, même si plusieurs facteurs nous permettent de suspecter ce caractère, comme :

- L'absence ou la faible sinistralité dans le passé
- Une priorité élevée par rapport à la moyenne du coût des sinistres
- Un Rate On Line (ROL) faible. Le ROL est le rapport entre la prime de réassurance et laouverture, ainsi si on observe un ROL faible, cela indique que la tranche a une faible chance d'être complètement touchée par des sinistres, autrement dit le tarif proposé est bas au vu de l'engagement associé à la tranche.

Notre étude se focalisera principalement sur la cotation des tranches non travaillantes, tranches où l'on retrouve un faible historique, avec une faible volumétrie du nombre de sinistres avec des coûts extrêmes.

2.5 Indemnisation des victimes en Responsabilité Civile

Dans le cas d'un sinistre avec préjudice corporel d'un tiers, le dénouement du dossier peut se faire de deux manières :

- Par transaction, faisant suite à un accord amiable avec la partie adverse
- Par voie judiciaire, faisant suite de la décision d'un juge qui déterminera les indemnités à verser.

En droit commun, deux types de victimes sont définis :

- Les victimes directes du dommage, c'est-à-dire les victimes blessées qui doivent obtenir réparation de leur préjudice et les proches des victimes décédées qui peuvent subir un préjudice moral mais aussi un préjudice économique.
- Les victimes indirectes, qui elles, correspondent aux proches de la victime blessée. L'indemnisation doit répondre au principe de réparation intégrale sous lequel la personne qui a subi un préjudice a droit à la réparation de celui-ci, en ce sens qu'elle doit être replacée dans une situation aussi proche que possible de celle qui aurait été la sienne si le fait dommageable ne s'était pas produit. Les victimes doivent donc être indemnisées le plus « justement » possible sans pertes, ni profits.

La loi du 5 juillet 1985 dite aussi loi Badinter est un texte fondamental en assurance car il prévoit le droit à l'indemnisation de tous les dommages engendrés par un accident de la circulation de manière rapide et convenable. Elle encourage les règlements amiables entre les parties et défend particulièrement les personnes vulnérables (personnes âgées, mineurs, handicapés). Son objectif principal est le droit à l'indemnisation plutôt que la recherche de la responsabilité, ce qui conduit à indemniser une victime (piéton, cycliste) même en cas d'absence de responsabilité du conducteur du véhicule impliqué. L'assureur de la victime dispose d'un délai de huit mois pour offrir un cadre d'indemnisation à la victime ou à défaut à proposer une avance si son état n'est pas stabilisé. Si le juge estime que l'offre est insuffisante, l'assureur peut se voir condamner à verser une pénalité de l'ordre de 15% du montant de l'indemnité. Cette loi offre aussi la possibilité à la victime en cas d'aggravation de son état de demander réparation de son préjudice dans un délai allant jusqu'à dix ans après la survenance du sinistre. Par cette loi, un suivi des indemnités versées est publié périodiquement. Si une rente est allouée, cette loi offre la possibilité au rentier de saisir le juge pour la conversion de la rente en capital sur la base d'une table de conversion réglementaire.

Les indemnités de réparation du préjudice peuvent prendre deux formes soit par le biais d'un versement d'un capital sur la base d'un barème de capitalisation ou autre fondement délibéré par les parties ou un juge ou par le biais d'une rente au bénéfice de la victime, ou de ses ayants droit, jusqu'à un terme prévu. A noter que le choix définitif du mode de versement relève de la décision des juges.

Une notion fondamentale qui mesure la gravité du préjudice est le taux d'Atteinte à l'Intégrité Physique et Psychique (AIPP) dont la définition est suite la suivante : « *Atteinte à l'Intégrité Physique et Psychique : Réduction définitive du potentiel physique et/ou psychique médicalement constatable ou médicalement explicable à laquelle s'ajoutent les douleurs et les répercussions psychiques que le médecin sait normalement liées à la séquelle ainsi que les conséquences dans la vie de tous les jours habituellement et objectivement liées à cette séquelle* ».

Le taux d'AIPP permet au médecin conseil d'évaluer le degré de gravité d'un dommage corporel sur la base d'une échelle allant de 0% à 100%. Elle est définie par la Commission Européenne (suite aux travaux de Trèves de 2000) comme :

- La réduction définitive du potentiel physique, psychosensoriel ou intellectuel résultant de l'atteinte à l'intégrité anatomo physiologique.

- Médicalement constatable donc appréciable par un examen clinique approprié complété par l'étude des examens complémentaires produits.
- L'apparition de phénomènes douloureux et les répercussions psychologiques normalement liées à l'atteinte séquellaire décrite ainsi que les conséquences habituellement et objectivement liées à cette atteinte dans la vie de tous les jours.

Par exemple, la perte d'un œil pourra être évaluée à un niveau de l'ordre de 20% d'AIPP, tandis que les séquelles d'un traumatisme crânien peuvent atteindre un taux de 100% d'AIPP.

Le calcul du taux d'AIPP ne peut intervenir qu'au moment de la consolidation de la victime, défini comme étant le moment où les différentes lésions se sont fixées en ayant pris un caractère permanent et qu'un traitement n'est plus nécessaire sauf pour éventuellement éviter une aggravation.

De manière générale, les deux modalités de versement des indemnités sont possibles, mais le recours à l'indemnisation en capital est la plus répandue, au regard de nombreux avantages pour les victimes :

- Le capital est transmissible aux proches au moment de la succession.
- L'indemnisation est perçue en une seule fois, et elle est gérée et placée par la victime ou ses ayants droits.
- Elle n'est pas soumise à imposition, seuls les intérêts financiers le sont.

Le versement de cette indemnité, suppose néanmoins une bonne gestion de ce capital pour pouvoir répondre aux besoins de la victime tout au long de sa vie et influe fortement la décision des juges qui doivent prémunir les victimes d'eux même mais aussi de leur entourage afin d'éviter d'éventuels abus. De ce point de vue, on peut noter que les indemnisations sous forme de rentes peuvent être plus adaptées à certaines victimes, notamment au regard des risques viager et économique portés par ce type d'engagement, mais n'ouvre aucun droit aux héritiers au moment du décès de la victime.

Du point de vue de la compagnie d'assurance qui doit verser cette indemnité, les deux formes d'indemnisations possèdent leurs avantages, en effet :

- L'attribution d'une rente permet un meilleur suivi de la victime et une indemnisation adéquate avec la survie de celle-ci, mais implique de prendre en charge toutes les revalorisations légales des rentes et les évolutions des taux.
- Le versement d'un capital à lui, un effet libératoire par la clôture comptable du dossier, mais en cas de décès de la victime, aucune récupération n'est possible.

En synthèse nous avons :

		Rentes	Capital
Pour la victime	Avantages	<ul style="list-style-type: none"> - Rentes indexées sur l'inflation. - Suivi du besoin de la victime sur la durée de vie. 	<ul style="list-style-type: none"> - Disponibilité du capital. - Liberté d'utilisation, gestion et placement. - Non soumis à l'impôt sur le revenu. - Transmissible lors de la succession.
	Inconvénients	<ul style="list-style-type: none"> ☞ Pas transmissibles lors de la succession. 	<ul style="list-style-type: none"> ☞ Négociations à mener sur le barème à retenir (taux et table de mortalité). ☞ Gestion prudente indispensable pour être en mesure de financer les besoins sur une période longue.
Pour l'assureur	Avantages	<ul style="list-style-type: none"> - Suivi de la victime tout au long de sa vie. - Résultat positif en cas de décès anticipé de la victime ou du crédirentier. 	<ul style="list-style-type: none"> - Pas d'aléa dû au risque de longévité de la victime. - Dossier de la victime clos (hors réouverture due à l'aggravation de l'état de santé de la victime). - Levier de négociation possible lors de discussions avec la victime.
	Inconvénients	<ul style="list-style-type: none"> ☞ Gestion opérationnelle lourde et suivi de la victime. ☞ Processus comptable complexe. ☞ Indexation des rentes à la charge de l'assureur. 	<ul style="list-style-type: none"> ☞ Négociations à mener sur le barème de capitalisation. ☞ Pas de bénéfice en cas de décès anticipé de la victime.

Tableau 8 : Synthèses des bénéfices pour la victime

2.5.1 Les postes de préjudices

L'indemnisation du dommage corporel recouvre de nombreux postes de préjudices qui sont divers et variés classifiés par une nomenclature, celle de Dintilhac.

Cette nomenclature constitue la référence Marché d'indemnisation des préjudices corporels et comporte une liste de postes de préjudices corporels qui concernent les victimes directes et indirectes. Cette liste n'est pas exhaustive mais permet de créer un cadre d'indemnisation. De plus, tous les préjudices ne sont pas forcément indemnifiables, cela dépend principalement de la gravité de l'accident corporel et des séquelles causées aux victimes. En revanche, son champ d'application est bien plus large que les victimes d'accident de la circulation puisqu'elle est applicable à toutes les victimes de dommage corporel, que la cause soit un accident de la route, médical ou toutes autres causes.

Pour procéder à l'évaluation du préjudice corporel, il est nécessaire de se baser sur une expertise médicale pour décrire la nature et la portée de chaque préjudice subi par une victime, qui à la suite de cela sera transformé en une indemnisation monétaire par un juge. C'est là tout l'apport de la nomenclature Dintilhac qui permet de faire le lien entre les points de vue médicaux, légaux et économiques nécessaires à la conversion monétaire de ces préjudices.

La nomenclature Dintilhac distingue principalement les préjudices temporaires et permanents selon que l'on se place avant ou après la date à partir de laquelle l'état de santé de la victime n'est plus susceptible d'amélioration par un traitement médical.

2.5.1.1 Les préjudices extra patrimoniaux

Les préjudices extra patrimoniaux temporaires

- Le Déficit Fonctionnel Temporaire (DFT) : Ce poste vient indemniser l'aspect non-économique de l'incapacité temporaire c'est-à-dire la gêne dans les actes de la vie courante rencontrée par la victime jusqu'à sa consolidation.
- Les souffrances endurées : Il s'agit des souffrances tant physiques que morales, endurées jusqu'à la consolidation. A noter qu'après consolidation, les souffrances endurées deviennent permanentes et elles sont intégrées au déficit fonctionnel permanent. De plus, l'appréciation de la sévérité de ces dernières est graduelle et échelonnée sur 7 niveaux.
- Le préjudice esthétique temporaire : La victime peut subir, pendant l'hospitalisation, une altération de son apparence physique, même temporaire, justifiant une indemnisation.

Les préjudices extra patrimoniaux permanents

- Le déficit Fonctionnel Permanent (DFP) : Ce poste est défini dans le rapport Dintilhac comme la réduction définitive du potentiel physique, psychosensoriel ou intellectuel appréciée par des examens médicaux approfondis à laquelle s'ajoutent les répercussions psychologiques. Le poste recouvre la perte d'intégrité mais également la dimension morale liée aux troubles subis dans les conditions d'existence.
- Le préjudice esthétique permanent : Ce préjudice renvoie au préjudice esthétique temporaire excepté que les altérations subies sont définitives et irréversibles.
- Le préjudice d'agrément : Le préjudice d'agrément vise exclusivement à réparer l'impossibilité pour la victime de pratiquer régulièrement une activité sportive ou de loisirs.
- Le préjudice sexuel : Ce préjudice recouvre 3 aspects distincts pouvant être altérés partiellement ou totalement, séparément ou cumulativement :

- L'atteinte morphologique des organes sexuels
 - La perte de libido
 - L'altération de la fertilité.
- Le préjudice d'établissement : Le préjudice d'établissement correspond au désespoir ne pouvoir réaliser ses projets personnels de vie comme fonder une famille, élever des enfants...
 - Les préjudices permanents exceptionnels : Pour l'indemnisation des victimes indirectes ou directes en cas de décès, on note également :
 - Les préjudices patrimoniaux
 - Les frais d'obsèques
 - Les frais divers
 - Les pertes de revenus des proches
 - Les préjudices extra patrimoniaux
 - Le préjudice d'affection : Il correspond au préjudice moral subi par les proches, parents ou non, mais pouvant justifier d'un lien affectif réel avec la victime.
 - Le préjudice d'accompagnement : Ce préjudice a pour objectif d'indemniser les troubles dans l'existence des proches liés à l'accompagnement de la victime jusqu'à la fin de sa vie.

2.5.1.2 Les préjudices patrimoniaux

Les préjudices patrimoniaux temporaires

- Les Dépenses de Santé Actuelles (DSA) : Les dépenses de santé recouvrent les frais médicaux et pharmaceutiques, non seulement les frais restés à charge de la victime, mais aussi ceux payés par les tiers (sécurité sociale, mutuelle d'assurance...), les frais d'hospitalisation et tous les frais paramédicaux (infirmiers, kinésithérapeutes...). Ils correspondent à tous les frais engagés pour soigner la victime peu importe l'organisme tiers payeur, dans ce sens ils peuvent être justifiés par les factures des différents intervenants.
- La Perte de Gains Professionnels Actuels (PGPA) : L'indemnisation est équivalente au coût économique du dommage pour la victime c'est-à-dire à l'ensemble de ses salaires nets sur la période d'arrêt, si l'employeur n'a maintenu aucun salaire. Le terme salaire inclut les primes et les indemnités qui font partie de la rémunération mais pas les frais non engagés pendant son arrêt (transport, nourriture...) Dans le cas où des indemnités journalières auraient été versées, les charges salariales (CSG et CRDS) correspondantes doivent être incluses dans l'indemnisation.
- Le préjudice scolaire, universitaire ou de formation : L'aspect patrimonial de ce préjudice résulte de la correspondance que l'on peut faire entre la scolarité d'un jeune et l'activité professionnelle d'un adulte. L'indemnisation de ce préjudice varie selon la perte ou non d'une année scolaire suite à l'accident et du niveau d'étude de la victime.
- Les frais divers : Il s'agit des frais de confort liés à l'hospitalisation comme la location d'une chambre individuelle ou de la télévision. Il convient d'accorder à la victime le confort dont elle aurait bénéficié si l'accident ne s'était pas produit. Il s'agit également des frais de transport et d'hébergement des proches pour visiter la victime puisque le moral du blessé peut agir sur l'évolution de son état de santé.

Les préjudices patrimoniaux permanents

- Les Dépenses de Santé Futures (DSF) : Ce poste est le pendant des dépenses de santé actuelles pour la période post consolidation.
- La Perte de Gains Professionnels Futures (PGPF) : De même, ce poste est l'équivalent de la perte de gains professionnels actuels pour la période post consolidation. La période de consolidation de la victime nécessitant généralement plusieurs années, la problématique de versement d'indemnités journalières disparaît. Le raisonnement indemnitaire porte donc sur les salaires nets de la victime.
- L'Incidence Professionnelle (IP) : Ce poste recouvre la dévalorisation sur le marché du travail qu'une victime peut avoir du fait de la dégradation de son état de santé (capacité physique et/ou cognitive). Il prend également en compte, pour quelqu'un contraint d'arrêter son activité, la perte d'épanouissement personnel, c'est-à-dire la perte de sociabilité (ne plus côtoyer ses collègues au quotidien) et la perte de gratification au travail (ne plus être valorisé par l'exercice de sa fonction).
- Les Frais d'Aménagement du Logement (FLA) : Il s'agit des dépenses consécutives à la réduction d'autonomie portant sur l'aménagement de l'habitat comme par exemple, les mises aux normes pour circuler en fauteuil roulant à l'intérieur du domicile.
- Les Frais de Véhicule Adapté (FVA) : La valorisation de ce poste ne correspond pas à la valeur totale du véhicule adapté mais à la différence de prix entre le prix de ce véhicule et celui dont se serait satisfait la victime en temps normal. Plus généralement, les préjudices d'aménagement du véhicule ou du logement doivent respecter le principe d'absence d'opportunités induit par le principe de la réparation intégrale.
- L'Assistance Tierce Personne (ATP) ou le Placement (PLAC) : La tierce personne est définie comme étant la personne qui apporte de l'aide à la victime incapable d'accomplir seule les actes essentiels de la vie courante comme :
 - L'autonomie locomotive : se laver, se coucher, se déplacer...
 - L'alimentation : manger, boire.

La valorisation de l'assistance tierce personne correspond donc à la rémunération des personnes venant apporter l'aide requise. Le recours à la tierce personne induit que la victime demeure vivre à son domicile ou chez un proche. Sinon elle peut être placée dans un établissement spécialisé, on parle alors de placement

	Préjudices patrimoniaux	Préjudices extra patrimoniaux
Temporaires (Avant consolidation)	Dépenses de Santé Actuelles Frais Divers Pertes de gains professionnels actuelles	Déficits Fonctionnels Temporaires Souffrances endurées Préjudices Esthétiques Temporaires
Permanents (Après consolidation)	Dépenses de Santé Futures Frais de logement adapté Frais de véhicule adapté Assistance Tierce Personne Pertes de gains professionnels Futures Incidence Professionnelle Préjudice Scolaire, universitaire ou d'études	Déficits Fonctionnels Permanents Préjudice d'Agrément Préjudices Esthétiques Permanents Préjudices Sexuels Préjudices d'Etablissement Préjudices Permanents Exceptionnels
Evolutif (hors consolidation)		Préjudices pour pathologies évolutives PEV

Tableau 9 : Synthèse de la nomenclature Dintilhac des postes de préjudices pour des victimes directes

	Préjudices patrimoniaux	Préjudices extra patrimoniaux
Cas de décès de la victime direct	Frais d'Obsèques Frais de Revenus des Proches Frais Divers des Proches	Préjudices d'Accompagnement Préjudices d'Affection
Cas de survie de la victime directe	Frais de Revenus des Proches Frais Divers des Proches	Préjudices d'Affection Préjudices Extra patrimoniaux Exceptionnels

Tableau 10 : Synthèse de la nomenclature Dintilhac des postes de préjudices pour des victimes indirectes

2.5.2 Evaluation des sinistres

2.5.3 Les rentes

Dans la partie en amont, nous avons vu que l'indemnisation de la victime peut se faire via le biais d'une rente. Une rente est le versement périodique pendant une durée donnée d'une somme d'argent prédéfinie, que l'on appelle un arrérage. La rente peut être viagère, c'est-à-dire versée jusqu'au décès de la victime, temporaire, annuelle ou fractionnée sur l'année. Le versement des arrérages peut être payable au début de la période, on dira que les arrérages sont versés à terme échoir. Il arrive également que le juge décide de faire courir une rente à partir d'une date antérieure à la date de jugement, appelée date d'effet. Les arrérages dus entre la date d'effet et de jugement sont appelés des arrérages échus, à ce titre l'indemnisation se fait de manière simple :

- On multiplie l'arrérage décidé par le juge par le nombre d'échéances déjà écoulées depuis la date d'effet décidée par le juge.
- On ne prend pas en compte le taux d'intérêt financiers ou d'espérance de vie dans les calculs.
- On ne prend pas en compte l'indexation de manière rétrospective.

Exemple : Si un juge décide d'attribuer en 2021 une rente dont l'annuité est de 1 000 € et dont la date d'effet est en 2011, alors les arrérages échus à verser en 2008 sont égaux à $1\ 000 \times 10$ soit 10 000€.

L'entrée en service de la rente peut se faire de manière immédiate ou différée.

La somme d'argent nécessaire à la date d'attribution de la rente pour payer la rente jusqu'à l'extinction de celle-ci, en tenant compte de l'espérance de vie des rentiers et des intérêts financiers correspond au Capital Constitutif de la Rente (CCR).

D'un point de vue actuariel, le CCR est la capitalisation des arrérages futurs à la date d'attribution de la rente en fonction des probabilités de survie fournis par la table de mortalité ; les flux futurs à considérer dans le cadre de la capitalisation sont la combinaison entre l'arrérage et la probabilité de survie.

Ainsi, pour permettre le calcul du CCR d'une rente, nous avons besoin de connaître :

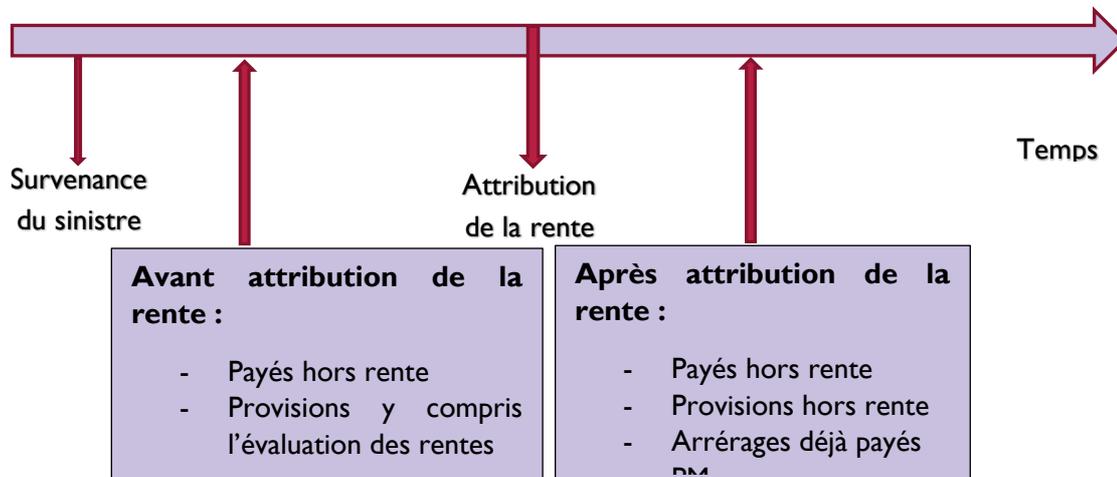
- Les caractéristiques de rente (à terme échu, échoir, et si elle se fait de manière annuelle, trimestrielle ou mensuelle).
- L'arrérage.
- L'âge de la victime.
- Le taux de capitalisation
- La table de mortalité.

En assurance non-vie, les rentes RC peuvent toucher deux types de provisions :

- La provision pour Sinistres A payer lorsque la rente est potentielle, c'est-à-dire tant que la décision (amiable ou judiciaire) n'est pas arrêtée. Cette provision correspond à la valeur estimative des dépenses en principal et en frais, tant internes qu'externes, nécessaires au règlement de tous les sinistres survenus non payés (sinistres déjà survenus, estimation dossier par dossier & IBNR (Incurred But Not Reported, les sinistres pas encore déclarés)
- La provision Mathématique (PM) lorsque la rente est servie. C'est la provision constituée par l'assureur dans son bilan pour servir les arrérages restant à payer d'une rente en tenant compte des intérêts financiers et de l'espérance de vie du crédirentier.

D'un point de vue actuariel, la provision mathématique d'une rente est la capitalisation des arrérages futurs en fonction de l'espérance de vie escomptée à une date postérieure à la date d'attribution. Le calcul de la PM est identique à celui du CCR, hormis que contrairement au calcul du CCR, le taux de capitalisation et la table de mortalité à utiliser sont encadrés légalement. De plus, la PM doit être recalculée à chaque arrêté de comptable et prend en compte la longévité observée du crédirentier ainsi que les changements de taux d'intérêts. Les articles R331-6, R331-7, R331-10, R331-12 du code des assurances sont relatifs à la provision mathématique en assurance non-vie. Ils stipulent les conditions de calcul de cette provision :

- Table de mortalité : TD88-90.
- Taux d'actualisation : inférieur à 60% du taux moyen d'emprunt de l'Etat français sur base semestrielle, sans dépasser 3,5%.
- Age des rentiers : la date de naissance du rentier est reportée au 31 décembre le plus voisin. Le montant de la provision par rente est le produit du montant annuel des arrérages à verser par l'euro de rente correspondant aux caractéristiques du dossier (âge, terme, différé...).



2.5.4 Indexation des rentes

Pour conserver le pouvoir d'achat d'une rente, la rente peut être revue annuellement, on parle ainsi de rente indexée. Toute rente est soumise à indexation de son arrérage, cependant le traitement en assurance diffère selon la branche d'assurance concernée par la rente.

On peut donc distinguer deux cas :

En automobile :

La validité de l'indexation des rentes allouées en réparation du préjudice causé par un véhicule a été reconnue en Jurisprudence par deux décisions de la chambre mixte de la Cour de cassation, qui stipule que les rentes allouées aux victimes d'accident de la circulation soit par transaction soit par décision judiciaire sont majorées de plein droit.

L'indexation doit s'effectuer selon le taux de majoration annuelle des rentes accident du travail communiqué par le Ministère de l'Economie (Article L434-17 du code de la Sécurité Sociale).

De cette loi résulte aussi la création d'un Fonds de revalorisation des rentes financé par une fraction de la prime d'assurance responsabilité civile automobile, dont l'objectif est le remboursement annuel de l'indexation des rentes. Historiquement le Fonds a été créé par la Caisse Centrale de Réassurance, mais l'article 82 de la loi du 1^{er} août 2003 de sécurité financière a transféré au Fonds de Garantie des Assurances Obligatoires (FGAO) la prise en charge de l'indexation des rentes. Le problème étant que le FGAO est très rapidement devenu déficitaire et à partir de 2008, le FGAO n'a jamais produit de compte équilibré, jusqu'à atteindre en 2012, un déficit de plus de 97 millions d'euros, les fonds propres du FGAO sont alors négatifs et sont estimés à -443 millions d'euros, tandis qu'à cette date 9000 rentes sont concernées et les provisions mathématiques sont estimées à plus de 700 millions d'euros.

Depuis 2013, il est alors décidé le désengagement du FGAO, les majorations futures des rentes sont à la charge des compagnies d'assurances qui doivent provisionner cet engagement par la prise en compte d'un taux d'inflation défini.

En non automobile :

Il n'existe pas de cadre légal concernant l'indexation mais dans la plupart des circonstances, les rentes non automobiles sont indexées par référence au cadre fixé pour les rentes automobiles. L'indice de revalorisation est alors celui prévu à l'article du code la Sécurité Sociale (Article L434-17). Contrairement aux rentes automobiles, les majorations d'annuités des rentes non automobiles, conséquences de l'indexation ne sont pas prises en charge par le FGAO, et dans ce cas, l'indexation des rentes non automobiles reste à la charge des assureurs.

En pratique, une rente est indexée selon les variations d'un indice de référence déterminé à l'attribution de la rente. A chaque échéance, en fonction des variations de cet indice, on calcule un pourcentage d'indexation à appliquer à l'arrérage indexé précédent afin de déterminer le nouveau arrérage indexé. L'indexation de la rente sera alors égale à la différence entre l'arrérage indexé et l'arrérage initial. Certains assureurs prennent en compte l'indexation des rentes en jouant sur le taux technique égal ou proche de 0%, ce qui revient à considérer que les intérêts financiers compensent l'indexation des rentes.

La Valeur Actuelle Probable (VAP) d'un euro de rente revalorisée à un taux de fixe de i' % est calculé de la manière suivante :

$$\frac{l_{x+1}}{l_x} \times \frac{1}{(1+i)^1} + \frac{l_{x+2}}{l_x} \times \frac{(1+i')}{(1+i)^2} + \dots + \frac{l_w}{l_x} \times \frac{(1+i')^{w-x-1}}{(1+i)^{w-x}}$$

Avec :

- x : âge du bénéficiaire
- l_x : nombre de survivant à l'âge x dans la table de mortalité
- i : taux d'intérêt technique. Correspond au taux de rémunération du capital
- w : âge limite

Exemple : Soit un arrérage de 1 000€ soumis à l'indexation selon un indice prédéfinis.

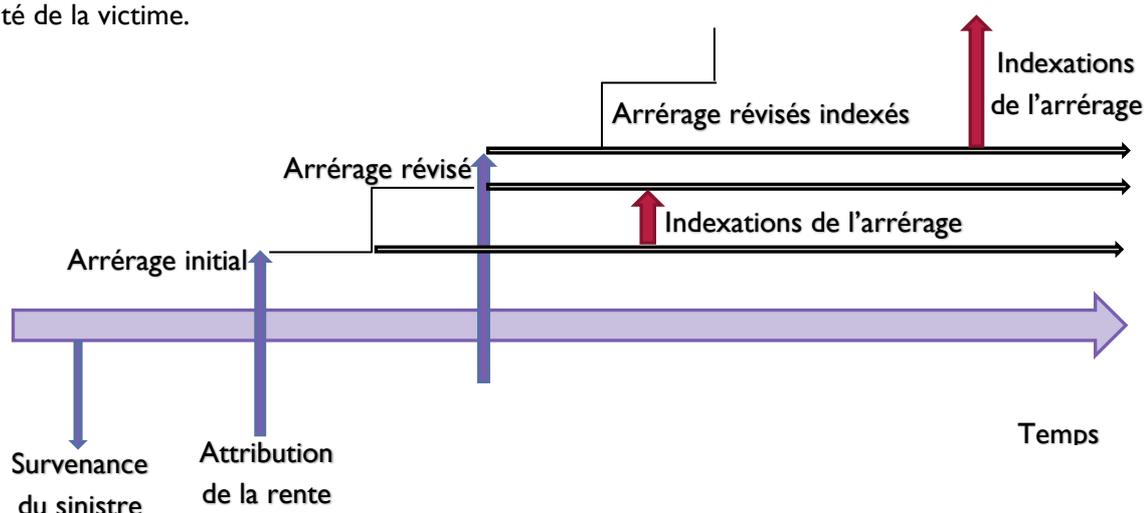
Année	2019	2020	2021	2022
Indice	100	104	109	113
Pourcentage d'indexation		$\frac{104}{100} - 1 = 4\%$	4,8%	3,7%
Arrérage indexé	1 000€	$1\ 000 \times (1 + 4\%) = 1\ 040\text{€}$	1089€	1129€
Indexation de la rente	0€	$1\ 040 - 1\ 000 = 40\text{€}$	89€	129€

Tableau 11 : Exemple d'arrérage soumis à l'indexation

2.5.5 Révision des rentes

Une rente étant versée durant la vie entière, il est courant que la victime ou sa famille demande une révision de la rente pour tenir compte de la modification de l'état de santé de la victime ou de la modification de l'état situationnel de la victime. Un complément de rentes peut ainsi être décidé par un juge ou convenu par transaction, notamment en cas d'évolution défavorable de l'état de santé de la victime.

A noter qu'en France, les rentes ne peuvent pas être révisées à la baisse en cas d'amélioration de l'état de santé de la victime.



Mathématiquement la valeur actuelle probable d'un 1€ de rente (revalorisée à un taux fixe de i' %) révisée après n années est traduite par la formule suivante :

$$\left[\left(\frac{1}{(1+i')} \times \left(\frac{l_{x+1}}{l_x} \times \frac{1}{(1+i'')^1} + \frac{l_{x+2}}{l_x} \times \frac{1}{(1+i'')^2} + \dots + \frac{l_{x+n}}{l_x} \times \frac{1}{(1+i'')^n} \right) \right] + \left[\left(\frac{1+r}{(1+i')} \times \left(\frac{l_{x+n+1}}{l_{x+n}} \times \frac{1}{(1+i'')^{n+1}} + \frac{l_{x+n+2}}{l_{x+n}} \times \frac{1}{(1+i'')^{n+2}} + \dots + \frac{l_w}{l_{x+n}} \times \frac{1}{(1+i'')^{w-x-n}} \right) \right] \right]$$

Avec $i'' = \frac{(1+i)}{(1+i')} - 1$ et r le taux de révision après n années.

Les formules de calcul de l'euro de rente sont les mêmes que l'indemnisation soit versée sous forme de rente ou de capital, ce qui diffère est les porteurs de risques viagers et financiers. Dans le cas

d'une indemnisation sous forme de rente les risques sont pris en charge par les compagnies d'assurances. Tandis que dans le cas d'un versement sous forme de capital, la victime subit les conséquences d'un potentiel risque que les produits financiers, générés par le placement de ce capital ne couvrent pas les futures inflations. De par ailleurs, la victime est soumise à un risque de longévité, en effet sa durée de vie résiduelle peut être supérieure à celle induite par la table de mortalité utilisée. On peut aussi observer une différence dans la cadence de règlement et dans les paramètres d'évaluation.

2.5.6 Provisionnement

Pour le provisionnement des rentes en norme comptable française, les compagnies d'assurances se réfèrent au Code des Assurances qui stipule que le calcul des provisions techniques de prestations d'invalidité se font à partir :

- D'une table de mortalité TD 88-90 (table de décès, réactualisée sur base du recensement fait en 1988/90).
- D'un taux d'actualisation qui ne peut excéder 60% du Taux Moyen des Emprunts de l'Etat français calculé sur base semestrielle (TME), sans pouvoir dépasser au maximum 3,5%.

A noter que les assureurs sont autorisés à répartir sur cinq ans au plus l'actualisation de leurs PM de rentes.

Dans le cas d'évaluation de rentes potentielles, les moyens de provisionnement ne sont pas réglementés et l'assureur évalue le provisionnement des rentes potentielles comme bon lui semble. Toutefois, dans la plupart des cas, les assureurs estiment l'arrérage potentiel et l'évaluation de la rente est alors égale à la capitalisation de cet arrérage.

Le montant de la provision par rente est le produit du montant annuel des arrérages à verser par l'euro de rente correspondant aux caractéristiques du dossier de la victime.

Calcul d'un euro de rente :

A la date comptable du 31/12/N, le calcul de l'euro de rente est donné par la formule ci-dessous en considérant :

- x l'âge retenu du rentier.
- $N_{naissance}$ L'année de naissance du rentier
- $a = \begin{cases} 1 & \text{si la naissance du rentier a eu lieu dans le premier semestre} \\ 0 & \text{sinon} \end{cases}$
- w l'âge maximum de la table TD88-90
- l_x Nombre de survivant d'âge x issu de la table TD88-90
- i Taux technique tel que $i = \min(3,5\%; 60\% \text{ du TME publiée au } 01/01/N + 1)$

Si on pose :

- $v = \frac{1}{1+i}$
- $n^P x = \frac{l_{x+n}}{l_x}$, qui correspond alors à la probabilité pour un individu d'âge x d'atteindre l'âge $x + n$.

Alors $n^E x = (1+i)^{-n} n^P x = V^n \cdot n^P x$ correspond au facteur d'escompte viager, qui est l'engagement de verser 1€ dans n années à un individu d'âge x .

Notons alors le nombre de commutations $D_x = V^x l_x$ et $N_x = \sum_{i=x}^w D_i$ alors :

$${}_n^E x = \frac{v^{x+n} l_{x+n}}{v^x l_x} = \frac{D_{x+n}}{D_x}$$

Rente viagère annuelle temporaire différée : Soit $m|n^a x$ l'euro de rente correspondant à l'engagement de payer annuellement 1€ à terme échu pendant n années à partir de m années à un individu d'âge x alors :

$$m|n^a x = \sum_{j=m+1}^{m+n} j E_x = \frac{N_{x+m+1} - N_{x+m+n+1}}{D_x}$$

Rente annuelle viagère immédiate : Soit a_x l'euro de rente correspondant à l'engagement de payer annuellement 1€ jusqu'au décès de la victime, alors :

$$a_x = \sum_{j=1}^{w-x} j E_x = \frac{N_{x+1}}{D_x}$$

Dans le cas où la rente n'est plus annuelle mais mensuelle, trimestrielle ou semestrielle (sans différé, par la règle de l'escompte viager, ${}_{m+p}E_x = {}_mE_x \cdot {}_pE_{x+m}$.

2.5.7 Indemnisation sous forme de capital :

Lorsque l'indemnisation se fait sous forme de capital, la valorisation de l'euro de rente est réalisée via le biais de barèmes de référence qui sont utilisés par les différentes cours d'appels, on retrouve notamment, le barème de la gazette du palais depuis 2004, ainsi que le barème de capitalisation pour l'indemnisation des victimes (BCIV) depuis 2002. Ces barèmes sont mis à jour perpétuellement par suite de la publication de nouvelles tables de mortalité.

En synthèse nous avons donc :

Modalité de paiement	Référence légale	Utilisation du barème	Tables et taux à utiliser
Capital	Pas de barème officiel mais recommandations des	<ul style="list-style-type: none"> - Indemnisation de la victime - Provisionnement de rente potentielle 	<ul style="list-style-type: none"> - BCIV - Gazette du Palais
Rentes	Article A.331-10 du Code des Assurances	<ul style="list-style-type: none"> - Provisionnement uniquement 	<ul style="list-style-type: none"> - TD88-90, taux à 60% du TME avec un maximum à 3.5% - Prise en compte de l'inflation à 2.25% pour les accidents post 2013

Tableau 12 : Tableau synthèse de l'indemnisation sous forme de rente ou capital

2.6 Les rentes en réassurance

Comme vu précédemment, les aspects relatifs aux rentes sont complexes et variés, ce qui entraîne que les clauses de rentes figurant dans les traités de réassurance en excédent de sinistre le sont devenues également. Cette partie a pour vocation d'une part de présenter les différentes clauses de rentes et d'autre part de passer en revue les points essentiels que doivent aborder et définir les clauses de rentes dans un traité de réassurance.

L'objectif principal d'une clause de rente est de définir tous les aspects relatifs aux rentes dans le cadre d'un traité de réassurance. Elle se doit de spécifier comment dans le calcul de la perte nette définitive sont pris en compte : les rentes, l'indexation des rentes, les révisions d'arrérages et décès. Mais aussi comment le CCR est calculé et la manière dont le provisionnement des rentes doit être effectué.

Plusieurs solutions s'offrent ainsi à l'assureur pour se couvrir contre le risque d'indexation des rentes.

2.6.1 Clause de rentes en suivi additionnel

Le principe de cette clause est simple, on considère la rente dans la durée et on ne différencie pas : le paiement des arrérages des autres paiements ou les provisions mathématiques des autres provisions relatives au sinistre. Une clause de rente en suivi additionnel a pour objectif principal d'effectuer un suivi de la rente de sa date d'attribution jusqu'à sa clôture ou ce que la limite de garantie du réassureur ait été atteinte.

Les arrérages indexés sont pris en charge par le traité comme n'importe quel autre paiement rattaché au sinistre et sont intégrés dans le calcul de la perte nette définitive. Une limite de la prise en charge de l'indexation des rentes exprimée en fonction du Capital Constitutif de Rente est prédéfinie par le traité. Le réassureur intervient dès lors que la perte nette définitive atteint la priorité, dans la limite de couverture du traité. La prise en charge de l'indexation peut se faire de deux manières :

Soit, l'arrérage et son indexation sont pris en charge dans le traité comme un paiement relatif au sinistre et on considère alors les arrérages indexés dans le calcul de la perte nette définitive

Soit, l'arrérage est pris en charge de manière classique mais l'indexation des rentes est partagée de manière proportionnelle entre le réassureur et la cédante. On doit alors différencier les arrérages fixes (non indexés) dans le calcul la perte nette définitive du partage de l'indexation de la rente de manière proportionnelle qui se fait selon la manière suivante : A l'attribution de la rente, on calcule le CCR de la rente (hors indexation) et on y ajoute les autres frais hors rentes (payés et suspens), nous donnant ainsi le montant total estimé du sinistre. On applique les bornes du traité de réassurance et on calcule la proportion du montant total estimé du sinistre théoriquement à charge du traité si tel est le cas. Puis ce pourcentage est ensuite utilisé à chaque échéance pour partager l'indexation de la rente entre le réassureur et la cédante.

Dans le cadre d'un suivi additionnel d'une rente, on ne peut pas procéder à la réouverture d'une rente car elle est suivie jusqu'à sa limite de couverture par le traité. Le calcul du CCR servant à définir la limite de garantie doit être correctement explicité dans le traité. A savoir que le traité doit spécifier : les caractéristiques de la rente (à terme échu/échoir, mensuelle/trimestrielles/annuelles ...), l'arrérage fixe, la table de mortalité et le taux de capitalisation.

Cette clause présente de nombreux avantages à la fois pour le réassureur et la cédante.

Du point de vue de la cédante, elle offre une couverture plus large car elle prend en compte la révision des arrérages et l'indexation de la rente et ne subit pas seules les éventuelles mortalités tardives des victimes dans la limite du traité.

Du point de vue du réassureur, il peut fractionner ses paiements, lui permettant ainsi de ne pas déboursier un capital qui peut parfois être conséquent.

En revanche, cette méthode comprend aussi des inconvénients notables à savoir que la cédante prend un risque de crédit important en cas de défaillance du réassureur et engendre une gestion des dossiers qui peut être longue.

C'est pourquoi il s'agit de la solution de réassurance qui protège le mieux la cédante contre l'inflation. Pour le réassureur, cette couverture engendre un surcroît d'aléa étant donné la protection contre le risque de longévité de la victime mais également le risque d'inflation sur des durées qui peuvent être longues.

D'autres solutions de réassurance sont mises en place permettant ainsi un équilibre dans la protection fournis par la réassurance.

Exemple d'application de la clause de rente en suivi additionnel des rentes :

Rentes	Priorité	Plafond	Table de mortalité
Suivi additionnel	1M €	2 x CCR	TD 88-90

Tableau 13 : Caractéristiques de la couverture de réassurance du traité

Simulation d'un sinistre de référence RCC :

En milliers d'euros.	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2025
Sinistre de référence survenance 2020	2020	2108	2205	2311	2421	2543	2515	2485	2446	2418	2388	2356	2322	2291	2261	2227
Provision de l'indexation (moy.simulation)	911	978	1051	1130	1215	1307	1272	1238	1204	1170	1137	1103	1070	1037	1005	973
Autre préjudices (réglés en t=0)	294	314	339	366	391	421	428	432	428	433	436	438	437	438	441	439
Priorité stabilisée	1000	1005	1010	1016	1022	1029	1035	1041	1046	1052	1057	1063	1068	1073	1078	1083
Rentes versées	0	0	0	0	0	54	108	163	219	276	334	393	453	514	575	638
Charge cédée	1020	1103	1195	1295	1399	1567	1588	1595	1590	1584	1579	1574	1569	1563	1558	1554
Charge conservée	1000	1005	1010	1016	1022	976	927	890	857	834	809	783	754	727	703	673

Tableau 14 : Exemple suivi additionnel des rentes

2.6.2 Clause de rachat

L'objectif de cette clause est de simplifier la gestion de la rente entre la cédante et le réassureur et de réduire la durée du traité. Elle permet de fixer à l'avance le mode de calcul d'un capital qui sera versé par le réassureur à la cédante en cas de paiement par rente à la victime.

Son principe de fonctionnement est simple, on considère la rente non pas dans la durée mais à sa date d'attribution par le biais du CCR tel qu'il est défini dans la clause de rentes, et on considère le CCR comme étant l'unique montant correspondant à la totalité de la rente.

Le rachat de la rente peut se faire de manière immédiate dès l'attribution de la rente ou différée.

Par rapport au rachat immédiat, le rachat différé permet à l'assureur de bénéficier d'une période de couverture "réelle" du risque de base d'inflation. Par rapport au suivi additionnel, les relations entre

assureur et réassureur sont limitées dans le temps à un maximum de 20 ans, ce qui permet de limiter le risque de contrepartie.

Dans le cas d'un rachat différé, la clause est en réalité une combinaison entre une clause de suivi additionnel et en rachat immédiat car :

- De l'attribution à la fin de la période de différé, les rentes sont suivies de façon additionnelle
- A la fin de la période différée, les rentes sont rachetées en calculant la capitalisation des arrérages futurs.

Pour le calcul du CCR, il est nécessaire de connaître l'âge de la victime, les caractéristiques de la rente, l'arrérage, le taux de capitalisation et la table de mortalité à utiliser. Comme à l'attribution de la rente on ne peut pas calculer un CCR sur un arrérage indexé puisqu'on ne connaît pas les variations futures de l'indice de référence choisi, on utilise alors un arrérage fixe mais avec un taux de capitalisation fixe, qui fait augmenter le montant de capital que la cédante doit recevoir au moment du sinistre.

Pour l'indexation, il y a deux possibilités :

- Soit l'indexation de la rente n'est pas prise en compte par le traité, et reste à la charge de la cédante et tous les CCR de rente sont déterminés en utilisant un taux de capitalisation sans indexation contractuelle.
- Soit l'indexation de la rente est prise en compte par le traité, dans ce cas, on utilise un taux de capitalisation avec indexation contractuelle.

En cas de révisions des arrérages ou de décès anticipé de la victime, il est tenant de rouvrir la rente, en théorie, une clause de rentes en rachat ne permet pas la réouverture d'une rente dans le cas de la révision des arrérages ou du décès de la victime ou du crédirentier. L'objectif principal de d'une clause de rentes en rachat immédiat est une commutation de la rente ce qui n'est pas compatible avec une réouverture en cas de révision des arrérage ou de décès. Pour les clauses de rentes en rachat différé, l'objectif étant d'envisager le cas des décès anticipés ou le cas des révisions des arrérages durant la période de différé, il ne serait alors pas cohérent d'envisager une telle chose.

Cette protection comporte de nombreux avantages :

Du point de vue de la cédante, cela permet de limiter dans le temps le risque de crédit, de réduire les frais de gestions du traité et de profiter du bénéfice d'un éventuel décès anticipé de la victime ou du crédirentier.

Du point de vue du réassureur, cela permet de raccourcir la durée d'un traité car les sinistres sont pour eux théoriquement clos et ne sont pas impactés par des cas de mortalités tardives.

Malheureusement, il y a toujours un risque de crédit mais dans une moindre mesure que dans le cas d'une clause de rente en suivi additionnel, le réassureur ne bénéficie pas des décès anticipés et la cédante est la seule partie impactée par le risque de mortalités tardives.

Exemple d'application de la clause de rachat des rentes :

Rentes	Priorité	Taux d'escompte	Table de mortalité
Rachat	1M€	1,50%	TD 88-90

Tableau 15 : Caractéristiques de la couverture de réassurance du traité en exemple

Simulation de la clause avec un sinistre de référence RCC :

En milliers d'euros.	Année de survenance T=0	Année de consolidation (T=5ans)	
	Calcul brut	Calcul brut	Calcul réassurance
Capital constitutif de rente hors indexation (CCR référence vue à fin 2020 $= 911*(1+6)^2$)	911	1 307	1 201
Coût de l'indexation (moy. simulation)	294	421	0
Autres préjudices (réglés en t=0) (Coût moyen de référence - CCR de référence= $(1536*(1+6)^2)-911$)	815	815	815
Total	2 020	2 543	2 016

Tableau 16 : Exemple de simulation de la clause de rente

En milliers d'euros.	Avant clause de stabilité	Après clause de stabilité (Priorité stabilisée à 1019K€)
Coût de la cession (A l'inventaire de la date de consolidation (5ans))	1 016	988
Coût à la charge de l'assureur (A l'inventaire de la date de consolidation (5ans))	1 526	1 555
Total	2 543	2 543

Tableau 17 : Exemple de calcul de cession avec clauses

En conclusion : Pour la couverture du risque de l'indexation des rentes, plusieurs solutions sont possibles tel que le rachat immédiat ou différé des rentes dès leurs attributions ou le suivi additionnel des rentes.

Toutes ces solutions ont leurs avantages et inconvénients, le choix de l'assureur dépendra de son appétence aux risques, ses fonds propres, son historique de sinistralité et sa maturité sur les risques RC Automobile.

3 Cotation

A présent que nous avons vu toutes les spécificités liées à un traité de réassurance qui peuvent impacter notre cotation. Nous allons nous intéresser à l'objectif de ce mémoire, qui est de présenter une tarification d'un traité en excédent de sinistre en RC auto, en répondant à certaines problématiques notamment liées aux sinistres extrêmes, à cet effet, nous présenterons uniquement les méthodes de tarification qui peuvent être utilisées dans ce cadre.

Pour effectuer une cotation, l'actuaire se sert des statistiques historiques, principalement d'un historique conséquent de sa sinistralité (fréquence et sévérité des sinistres), mais aussi de l'historique d'encaissement, et d'une mesure d'exposition du portefeuille en nombre ou en somme assurée. Toutes ces informations permettent à l'actuaire de modéliser la sinistralité de la cédante et d'estimer une prime pure et une prime commerciale.

La tarification consiste à effectuer une première modélisation avant la souscription d'un contrat. En pratique, un exercice important pour les actuaires consiste à se poser la question si, pour les générations antérieures d'affaires, la tarification ne souffre pas de biais systématiques (sur/sous tarification). Différentes approches sont possibles, en relation avec les données et pratiques de provisionnement pour une segmentation spécifique.

3.1 Prime pure et commerciale

Les composantes classiques de la tarification technique sont la prime pure, les chargements de sécurité et de profitabilité, ces deux composantes formant la prime de risque.

La prime pure sera complétée par des chargements économiques (par exemple, des frais internes de gestion, de courtage) et des ajustements pour la prise en compte des primes d'ajustement (par exemple, prise en compte des primes de reconstitutions payantes) pour former la prime commerciale.

La prime pure est la valeur actuelle probable des sinistres qui vont être payés au titre du traité. Le montant de cette prime pure serait suffisant à la garantie d'un résultat (Prime - Sinistres) nul, si les sinistres n'étaient pas aléatoires. Autrement dit, il s'agit du meilleur estimateur « Best Estimate » des coûts futurs engendrés par les sinistres au traité. En branche courte, l'impact de l'escompte peut être négligé.

Le taux commercial effectif auquel la cédante est soumise repose sur un taux technique, lui-même basé sur un taux pur qui est ensuite augmenté par différentes charges telles que :

- Le coefficient de sécurité C_S , permet de couvrir la volatilité autour du taux pur liée aux écarts entre la sinistralité future observée et prédite. Cette volatilité est souvent la conséquence d'un faible nombre de données pour estimer les paramètres du modèle. Ajouté au taux pur, le coefficient de sécurité permet d'obtenir le taux technique.
- Les frais de gestion du traité f_g permettent de couvrir les frais engendrés par l'acceptation d'une affaire, comme l'enregistrement comptable des flux ou la gestion des sinistres.
- Les frais de courtage f_c , permettent de rémunérer l'intermédiaire de réassurance qui est intervenu pour placer l'affaire.
- Les frais financiers f_r diminuent le taux et représentent les intérêts générés par le placement des provisions de primes et sinistres du portefeuille réassuré.

Le taux commercial est alors exprimé par la formule suivante :

$$\text{Taux commercial} = \underbrace{\text{Taux pur} \times (1 + C_s)}_{\text{Taux technique}} \times (1 + f_g) \times \frac{1}{1 - f_c} \times \frac{1}{1 + f_r}$$

Dans la suite de notre étude, nous allons uniquement nous concentrer sur l'estimation du taux pur de réassurance, la prime pure associée reflète l'espérance des coûts des sinistres.

3.2 Tarification par la méthode du « Burning Cost » : approche et limites

La première méthode de tarification que nous avons décidé d'introduire est la méthode dite du Burning Cost. Cette cotation repose sur l'expérience du portefeuille de la cédante en prenant en compte la sinistralité passée. On entend par là, le rapport entre les sinistres survenus à une époque pendant une période d'exposition, et les expositions de l'époque. Cette méthode de tarification est très populaire en réassurance et quelles que soient les approches adoptées, elle constitue une référence incontournable de perception du comportement statistique du traité non proportionnel.

En décomposant cette méthode de tarification en plusieurs étapes, cela donne :

- Première étape, mise en AS-IF : On redresse l'échantillon de statistiques historiques afin de le rendre représentatif des conditions de souscription et économiques de l'année de couverture à évaluer.
- Seconde étape : Pour chaque année d'expérience, on calcule les charge annuelles AS-IF du contrat à tarifier qui doivent tenir compte des spécificités du traité : clauses applicables dont les limites annuelles.
Puis on rapporte ces charges aux assiettes de primes AS-IF, et on obtient un taux annuel estimé AS-IF.
- Troisième étape : On calcule le taux moyen Burning cost sur les n années d'expérience disponible lorsque la série des taux est stable, ou de projeter un taux probable lorsque l'on décèle une tendance.

Si on considère un traité en excédent de sinistre par risque, sans AAL et AAD, avec un nombre de reconstitutions gratuites et illimitées sur une tranche : c XS d Avec :

- d la priorité et c la portée
- n : le nombre d'années historiques disponible pour procéder à la tarification
- $j \in \llbracket 1 : n \rrbracket$ l'année historique
- $(N_j)_{1 \leq j \leq n}$: le nombre de sinistres survenus l'année j
- $(X_{i,j})_{1 \leq i, j \leq n}$; le montant AS-IF du sinistre i survenu durant l'année j
- $(S_j)_{1 \leq j \leq n}$: la charge totale supportée par la réassurance pour l'année j
- $(EPI_j)_{1 \leq j \leq n}$: l'assiette de primes AS-IF de l'année j

Mathématiquement si pour chacune des années j on exprime le montant à la charge du réassureur par la formule suivante : $(S_j)_{1 \leq j \leq n} = \sum_{i=1}^{N_j} \min(\max(X_{i,j} - d; 0) : c)$

Le taux de burning cost (BC_{pur}), est alors la moyenne des montants à charge pondérée par l'assiette de primes AS-IF de l'année associée, mathématiquement cela est traduit par :

$$BC_{pur} = \frac{1}{n} \times \sum_{j=1}^n \frac{S_j}{EPI_j}$$

On obtient par la suite la prime de réassurance en multipliant le taux de burning cost pur par l'assiette de primes estimée pour l'année à tarifier. Cette méthode par conséquent est le ratio entre les pertes enregistrées par le réassureur et les encaissements de la cédante sur plusieurs années d'expérience.

Cas pratique et exposition des limites de la méthode :

Soit un traité en excédent de sinistre couvrant la RC Corporelle AUTO selon le programme suivant :

- T1 : 3M XS 3M
- T2 : 6 XS 6M
- T3 : illimité XS 12M
- Reconstitutions illimitées et gratuites

Pour le renouvellement de notre traité 2021, nous souhaitons proposer une tarification de nos tranches, on dispose des données d'expérience revalorisées sur un horizon d'expérience de 10 ans allant de 2010 à 2020.

Année	EPI	Sinistres	Charge cédée en TI	Charge cédée en TI cumulée	Taux BC annuel
2010	100.234 M€	3.4 M€ ; 1.2 M€ ; 4.6 M€ ; 6 M€ ;	0.4 M€ ; 0 M€ ; 1.6 M€ ; 3 M€ ;	5 M€	5 M€/100.234 M€ = 5,0 %
2011	112.345 M€	0.250 M€ ; 1.8 M€ ; 4.5 M€ ; 5.5 M€ ;	0 M€ ; 0 M€ ; 1.5 M€ ; 2.5 M€ ;	4 M€	3,6 %
2012	118 M€	3.5 M€ ; 4 M€ ;	0.5 M€ ; 1 M€ ;	1.5 M€	1,3 %
2013	121.3 M€	0.45 M€ ; 7.6 M€ ;	0 ; 3 M€ ;	3 M€	2,5 %
2014	124 M€	3.6 M€ ; 4.5 M€ ;	0.6 M€ ; 1.5 M€ ;	2.1 M€	1,7 %
2015	125 M€	1.8 M€ ; 5.7 M€ ;	0 M€ ; 2.7 M€ ;	2.7 M€	2,1 %
2016	127.45 M€	4.2 M€ ; 5.2 M€ ;	1.2 M€ ; 2.2 M€ ;	3.4 M€	2,7 %
2017	129.54 M€				0,0 %
2018	130.23 M€	4.3 M€ ; 5.3 M€ ; 12 M€ ;	1.3 M€ ; 2.3 M€ ; 3 M€ ;	6.6 M€	5,1 %
2019	132.4 M€	1.5 M€ ; 4.2 M€ ; 5.25 M€ ;	0 M€ ; 1.2 M€ ; 2.25 M€ ;	3.45 M€	2,6 %
2020	137.3 M€	3.4 M€ ; 3.1 M€ ; 5.2 M€ ;	0.4 M€ ; 0.1 M€ ; 2.2 M€ ;	2.7 M€	2,0 %

Tableau 18 : Exemple d'application de la méthode Burning cost

En 2021, l'assiette de primes (EPI) estimée est de 143 450 000€. On remarque que notre tranche T1 est travaillante, en effet chaque année, au moins un sinistre touche notre traité pour cette tranche, sauf en 2017, ou aucun sinistre n'a été reporté.

Chaque année on a calculé le taux de burning cost annuel. On calcule ensuite le taux pur de burning cost, qui est la moyenne pondérée des taux de burning cost annuel, nous donnant un taux pur de 2,59%.

La prime de réassurance associée à la tranche T1 est donc : $2,59\% \times 143\,450\,000$ (EPI estimés de l'année tarifé) = 3 709 220 €. Si on souhaite tarifer notre T2, on remarque que très peu de sinistres franchise la limite de notre seconde tranche, en effet :

Année	EPI	Sinistres	Charge cédée en T2	Charge cédée en T2 cumulée	Taux BC annuel
2010	100.234 M€	3.4 M€ ; 1.2 M€ ; 4.6 M€ ; 6 M€ ;			0,0 %
2011	112.345 M€	0.250 M€ ; 1.8 M€ ; 4.5 M€ ; 5.5 M€ ;			0,0 %
2012	118 M€	3.5 M€ ; 4 M€ ;			0,0 %
2013	121.3 M€	0.45 M€ ; 7.6 M€ ;	1.6 M€	1.6 M€	1,3%
2014	124 M€	3.6 M€ ; 4.5 M€ ;			0,0%
2015	125 M€	1.8 M€ ; 5.7 M€ ;			0,0%
2016	127.45 M€	4.2 M€ ; 5.2 M€ ;			0,0%
2017	129.54 M€				0,0%
2018	130.23 M€	4.3 M€ ; 5.3 M€ ; 12 M€ ;	6 M€	6 M€	4,6%
2019	132.4 M€	1.5 M€ ; 4.2 M€ ; 5.25 M€ ;			0,0%
2020	137.3 M€	3.4 M€ ; 3.1 M€ ; 5.2 M€ ;			0,0%

Tableau 19 : Limite de l'application de la méthode du Burning Cost

La T2 n'est que très peu travaillante, sur 10 années d'historiques, seulement 3 sinistres on franchit cette tranche. Le taux de burning cost associé à cette tranche est de 0,60% et si on procédait à la

tarification de cette tranche par cette méthode, on obtiendrait une prime pure de 863 914, alors que le montant maximum atteint de cette tranche est de 12M€, soit 6M€ à la charge du réassureur.

On comprend clairement la limitation d'une telle technique de tarification qui permet difficilement de prévoir toutes les réalisations probables futures. Ce constat est d'autant plus accentué lorsque la tranche n'est pas travaillante, ce qui est clairement le cas de notre T3 qui n'a jamais été atteinte.

En effet, la méthode du burning cost fournit des taux nuls, ce qui engendre un coût de couverture nul, or il existe tout de même une exposition sur cette tranche, où des sinistres peuvent largement dépasser les 12M€. Cette méthode atteint donc sa limite et ne nous permet pas d'estimer la prime pure d'une tranche non travaillante.

Pour estimer la prime pure d'une tranche non-travaillante, l'actuaire dispose de modèles probabilistes. Il peut d'une part modéliser la fréquence, et d'autre part la sévérité des sinistres. Ce mémoire se focalise uniquement sur la modélisation de la sévérité des traités de réassurance en excédent de sinistre.

On peut alors conclure que :

La tarification Burning cost est une méthode de tarification fiable lorsque l'on dispose d'une forte expérience passée.

Cependant, cette méthode comporte des limites, notamment le fait qu'elle suppose que l'historique passée se répétera dans l'avenir, ce qui est acceptable seulement si les politiques de souscription restent stables. De plus, elle ne tient pas compte de l'exposition de chaque tranche (profil de portefeuille), et ne permet pas de coter les tranches non-travaillantes pour lesquelles très peu ou pas du tout de sinistres traversent la priorité. En conséquence, cette méthode est peu adaptée aux expositions de très faible fréquence mais de forte sévérité, et à la cotation des options alternatives de clauses annuelles.

3.3 Tarification par l'approche probabiliste : Méthode fréquence-coût sinistre

Comme nous venons de le voir, la tarification par la méthode du Burning Cost nous a permis d'obtenir une évaluation de la prime pure de réassurance, mais cette méthode est limitée à des tranches travaillantes.

Pour remédier à cela, le modèle interne de BPCE Assurances utilise un outil qui permet d'estimer un taux pur de réassurance par une approche probabiliste. L'enjeu du mémoire est d'améliorer et d'optimiser l'outil de tarification utilisé afin de le rendre plus performant et plus stable.

En réassurance, un portefeuille est caractérisé par deux composantes :

- Le nombre de sinistres par année qui correspond à la fréquence, que l'on notera N .
- Le coût individuel de chaque sinistre qui correspond à la sévérité et que l'on notera X_i (avec i le $i^{\text{ème}}$ sinistre).

La fréquence et la sévérité sont deux variables aléatoires que l'on considère comme indépendantes. Le montant de tous les sinistres X_i sont considérés comme indépendants et identiquement distribués. Ils suivent tous la loi du montant de sinistre notée X . De plus, les deux variables aléatoires, N et X_i sont toutes les deux définies en dessous et au-dessus d'un certain seuil X_0 . Ce seuil qui sera déterminé plus tard, représente un fort enjeu dans la modélisation de notre prime pure. En effet, actuellement le choix de ce seuil est contraint par les seuils de communication revalorisés par année d'observation ainsi que par les valeurs de la franchise contractuelle la plus basse du traité de réassurance. Un des enjeux du mémoire sera donc de redéfinir le choix du seuil afin de permettre une meilleure modélisation des tranches peu ou non travaillantes.

En effet, toute la cotation dépend du choix de ce seuil. En choisissant un seuil, on tronque la distribution des sinistres à droite en excluant les observations au-delà du seuil, et à gauche en gardant seulement les observations qui sont supérieures ou égales au seuil. Cela permet de ne considérer que les observations les plus extrêmes et de mieux ajuster une distribution de probabilité pour modéliser ces dernières. Choisir un seuil d'analyse (ou de modélisation) trop bas peut revenir à se concentrer sur l'analyse de la survenance des "petits" sinistres et ne pas expliquer suffisamment la survenance des plus grands sinistres qui seront à charge du programme XS et peuvent provenir d'un processus aléatoire différent. Choisir un seuil trop haut ne permettrait pas de correctement modéliser les tranches peu travaillantes dû au manque de sinistres sur cette partie-là.

Le modèle fréquence-sévérité consiste à spécifier les distributions de probabilité pour les variables N et X afin de modéliser la distribution de la charge annuelle par simulation de Monte-Carlo.

3.3.1 Présentation de la méthode

Le modèle probabiliste fréquence-coût est non seulement utilisé car l'hypothèse d'indépendance entre le nombre de sinistres et leur montant est respectée, mais aussi car il permet de prendre en compte l'historique et donc d'inclure une part totale de crédibilité dans la tarification tout en ayant une approche paramétrique du risque, d'où la possibilité de calculer la Value At Risk (VaR), et d'effectuer des tests de sensibilité sur les paramètres de distributions.

La méthode retenue pour la modélisation de la charge annuelle des graves consiste à modéliser séparément la sévérité et le nombre des sinistres, puis en déduire la charge annuelle par simulation de Monte-Carlo à partir des lois de la sévérité et du nombre de sinistres obtenues.

Cette approche consiste donc à modéliser les lois de fréquence et de sévérité en ne tenant compte que du comportement des sinistres en dessous et au-dessus d'un certain seuil.

3.3.2 Impact du choix du seuil sur les lois

Le choix du seuil implique que le choix des données historiques est limité et seuls les montants dépassant un seuil sont disponibles. Ainsi durant tout le processus de modélisation de la fréquence et de la sévérité, il faut prendre en compte l'impact du seuil de modélisation, et c'est pourquoi les distributions de probabilités sont normalisées pour prendre en considération la troncature.

On va considérer un seuil de modélisation noté s , et un second noté u , qui permettra d'illustrer la troncature à droite. Si on note X la variable aléatoire qui représente la sévérité modélisée et F sa fonction de répartition. Alors la fonction de répartition associée aux données tronquées à gauche et à droite aux seuils s et u ne vaut plus F mais :

$$\forall x \in \mathbb{R}, F_{s,u} = P(X \leq x | s \leq X \leq u)$$

Cette fonction de répartition peut s'exprimer à l'aide de la fonction de répartition non tronquée F tel que :

$$F_{s,u} = \begin{cases} 0 & \text{si } x \leq s \\ \frac{F(x) - F(s)}{F(u) - F(s)} & \text{si } s \leq x \leq u \\ 1 & \text{sinon} \end{cases}$$

Démonstration :

$$F_{s,u} = P(X \leq x | s \leq X \leq u)$$

Par la formule de Bayes on obtient que : $F_{s,u} = \frac{P(s \leq X \leq \min(x,u))}{P(s \leq X \leq u)}$

En passant à la fonction de répartition non tronquée F on a que : $F_{s,u} = \frac{F(\min(x,u)) - F(s)}{F(u) - F(s)}$

D'où :

$$F_{s,u} = \begin{cases} 0 & \text{si } x \leq s \\ \frac{F(x) - F(s)}{F(u) - F(s)} & \text{si } s \leq x \leq u \\ 1 & \text{sinon} \end{cases}$$

3.3.3 Modélisation de la fréquence

Maintenant que nous avons introduit le choix du seuil de modélisation, la première étape est de modéliser la fréquence au-dessus et en dessous de ce seuil.

La fréquence ne faisant pas l'objet de notre étude, l'approche utilisée par le groupe BPCE est conservée et nous présenterons uniquement la méthode utilisée.

Pour déterminer la loi de fréquence, trois lois sont les plus fréquemment utilisées :

La loi de poisson :

$$\left\{ \begin{array}{l} X(\Omega) = \mathbb{N} \\ P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \\ E(X) = \frac{1}{n} \sum x_i = \lambda \text{ et } V(X) = \lambda \end{array} \right. \quad \text{Avec } \lambda \in]0; +\infty[$$

Elle est la plus répandue et la plus simple à estimer. En effet, pour obtenir l'estimateur de la loi de Poisson tronquée, il suffit de calculer une simple moyenne du nombre de sinistres répertoriés au-dessus du seuil au cours des dernières années d'historique, ou de faire une régression pour tenir compte d'une tendance. On peut estimer le paramètre λ comme suit : $\hat{\lambda} = \bar{X}$

Concrètement, le critère le plus commun pour choisir une loi de Poisson est l'égalité de la moyenne et de la variance. Le processus de Poisson (homogène) est particulièrement utilisé en réassurance pour modéliser les événements rares, qui sont d'espérance et de variance proches. Si ce n'est pas le cas, d'autres lois de comptage peuvent être envisagées : la loi binomiale et la loi binomiale négative.

La loi binomiale : Soit n , le nombre d'expériences réalisées et p la probabilité de succès.

$$\left\{ \begin{array}{l} X(\Omega) = 1, 2, \dots, n \\ P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \\ E(X) = np \text{ et } V(X) = np(1-p) \end{array} \right.$$

L'utilisation de la loi binomiale dans l'assurance en cas de décès ou d'invalidité comme loi de fréquence correspond à un contexte de population assurée très homogène en âge et en classe de risque de mortalité. Ainsi, si l'on prend l'hypothèse d'un taux de mortalité moyen annuel q_x sur l'ensemble du portefeuille, la fréquence s'écrit :

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

Avec n le nombre d'assurés dans le portefeuille et X_i une suite de variables indépendantes et identiquement distribuées de Bernoulli, de paramètre q_x . Alors :

$$N \sim B(n, q_x) \text{ et } E(N) > \text{Var}(N)$$

La loi binomiale négative :

La loi binomiale négative est interprétée comme une « loi de Poisson mélange », c'est-à-dire une loi de Poisson dont le paramètre λ suit une loi Gamma. Soit $p \in]0, 1]$ et $q = 1-p$.

$$X \sim BN(r, p) \text{ avec } r: \text{Nb succès et } p: \text{Proba de succès}$$

$$\left\{ \begin{array}{l} X(\Omega) = \mathbb{N}^* \\ P(X = k) = \binom{k+n-1}{k} p^n q^k \\ E(X) = \frac{1}{n} \sum x_i = \bar{X} = \frac{r(1-p)}{p} \text{ et } V(X) = \frac{1}{n} \sum (x_i - \bar{X})^2 = \frac{r(1-p)}{p^2} \end{array} \right.$$

La loi binomiale négative se caractérise notamment par $E(N) < \text{Var}(N)$.

Alors on peut estimer les paramètres p et r comme suit : $\hat{p} = \frac{E(X)}{V(X)}$ et $\hat{r} = \frac{\hat{p}}{1-\hat{p}} \cdot E(X)$

Pour évaluer la fréquence de sinistre, dans un premier temps, nous projetons à l'ultime le nombre de sinistres graves après indexation via la méthode de Chain-Ladder.

3.3.3.1 Pour les sinistres graves au-dessus du seuil déterminé :

Le choix de la loi se fait en plusieurs étapes :

- **1ère étape** : Nous considérons comme nombre ultime, le nombre de sinistres dont la charge indexée est supérieure au seuil. Ici, nous ne projetons pas à l'ultime car le nombre de sinistres n'est pas suffisant pour une telle projection.
- **2ème étape** : Nous redressons de l'exposition le nombre ultime de sinistres par survenance en appliquant la formule suivante :

$$Nb \text{ redressé}_i = \frac{Nb_i \times Année \text{ Assurance}_{N+1}}{Année \text{ Assurance}_i}, \text{ où } i \text{ est la survenance}$$

- **3ème étape** : Nous excluons les survenances dont le redressement est trop fort
- **4ème étape** : Choix de la loi entre une Poisson et une Binomiale Négative avec l'aide de l'indice de Fisher
 - o Si $Indice \text{ de Fisher} = \frac{variance}{moyenne} \leq 1,05$, alors la loi de fréquence est une loi de Poisson
 - o Sinon la loi de fréquence est une binomiale négative

Un effet horizon est aussi ajouté, en effet il est possible de restreindre la période d'observation pour tenir compte des nouvelles tendances.

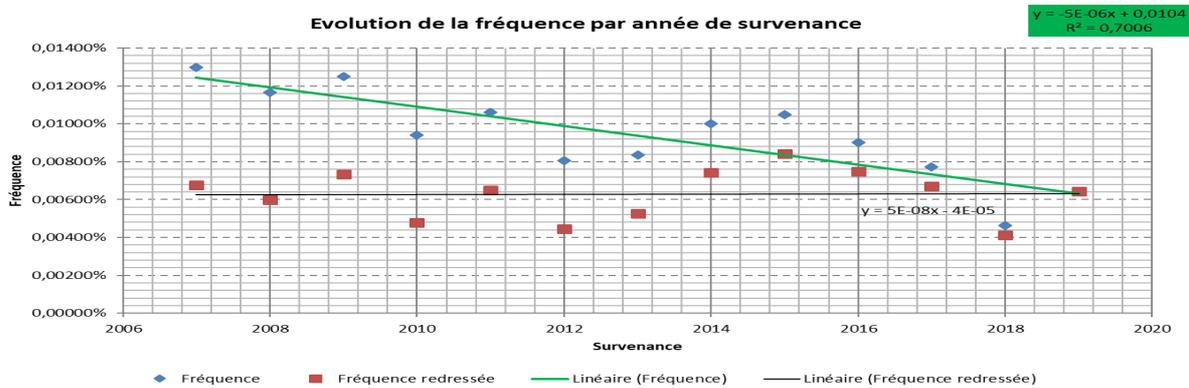
3.3.3.2 Pour les sinistres graves inférieurs au seuil

- **1ère étape** : Nous projetons à l'ultime et par survenance le nombre de sinistres graves dont la charge est inférieure au seuil déterminé via la méthode de Chain-Ladder.
- **2ème étape** : Nous retranchons par survenance le nombre de sinistres (au dossier/dossier) au-delà du seuil déterminé.
- **3ème étape** : Nous redressons de l'exposition le nombre de sinistres ultime ainsi obtenu

$$Nb \text{ redressé}_i = \frac{Nb_i \times Année \text{ Assurance}_{N+1}}{Année \text{ Assurance}_i}, \text{ où } i \text{ est la survenance}$$

- **4ème étape** : Nous choisissons l'horizon qu'il convient de considérer en fonction de l'évolution de la fréquence (Nombre ultime / AA)

- **5ème étape** : nous corrigeons le redressement de l'exposition par un redressement de l'allure de la fréquence (cf. graphe ci-dessous) :
 - o Redressement de l'allure de la fréquence avec l'aide de la droite de régression linéaire (fréquence x survenance) et par rapport au dernier point de la régression
 - o Passage au nombre par multiplication de la fréquence ainsi redressée par survenance par le nombre d'années d'assurance N+1



- **6ème étape** : Choix de la loi à l'issue de toutes ces étapes entre une Poisson et une Binomiale Négative avec l'aide de l'indice de Fisher
 - o Si $Indice\ de\ Fisher = \frac{variance}{moyenne} \leq 1,05$, alors loi de Poisson
Sinon loi binomiale négative

3.3.4 Modélisation de la sévérité

Avant toute chose, et comme le nom de l'approche probabiliste le laisse entendre, un choix concernant les lois de distribution continue qui vont être étudiées par la suite est à effectuer. Pour définir ces dernières, une bonne estimation de la « queue » de distribution est un critère important. Autrement dit, il va être recherché une ou des lois continues qui vont s'adapter convenablement aux différentes tranches définies par le choix du seuil mais aussi qui vont permettre d'extrapoler pour les événements rares. Usuellement, en réassurance, les lois utilisées lors de ce type d'étude sont :

- La loi de Weibull ;
- La loi Gamma ;
- La loi Log-Normale ;
- La loi Exponentielle ;
- La loi de Pareto ;
- La loi Log Logistic.

Les lois log-normale et de Pareto offrent le plus de sécurité au réassureur car elles ont les queues de distributions les moins lourdes parmi ces lois. Les lois Log-Normale et Weibull, sont souvent retenues dans pour modéliser les sinistres extrêmes pour la branche RC Auto, sachant aussi que la loi de Pareto est certainement l'une des lois les plus utilisées en réassurance pour modéliser la sévérité.

Loi de Weibull : On dit qu'une variable aléatoire X suit une distribution de Weibull de paramètres $a > 0$ et $b > 0$, si la densité de la variable aléatoire est de la forme :

$$f(x) = ab^{-a}x^{a-1}e^{-\left(\frac{x}{b}\right)^a} \quad 1_{x \geq 0}$$

La fonction de répartition associée s'exprime $\forall x \geq 0$ par : $F(x) = 1 - e^{-\left(\frac{x}{b}\right)^a}$

L'espérance et la variance d'un Weibull sont données par : $\mathbb{E}(X) = \frac{b}{a} \Gamma\left(\frac{1}{a}\right)$ et $\mathbb{V}(X) = \frac{b^2}{a} \left(2\Gamma\left(\frac{2}{a}\right) - \frac{1}{a} \Gamma^2\left(\frac{1}{a}\right)\right)$

Loi Gamma : On dit qu'une variable aléatoire X suit une distribution de Gamma de paramètres $k > 0$ et $\theta > 0$, si la densité de la variable aléatoire est de la forme :

$$f(x) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(0,k)} 1_{x \geq 0} \text{ où, } \Gamma(x, k) = \int_x^\infty t^{k-1} e^{-t} dt$$

La fonction de répartition associée s'exprime $\forall x > 0$ par : $F(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left[\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right]$ où, $\operatorname{erf}(x)$ est la fonction d'erreur de Gauss tel que $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$

L'espérance et la variance d'un Gamma sont données par : $\mathbb{E}(X) = k\theta$ et $\mathbb{V}(X) = k\theta^2$

Loi Log-Normal : On dit qu'une variable aléatoire X suit une distribution de Log-Normal de paramètres $\sigma > 0$, $\mu \in \mathbb{R}$, si la densité de la variable aléatoire est de la forme :

$$f(x) = \frac{1}{\sqrt{2\pi\sigma x}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2} 1_{x > 0}$$

La fonction de répartition associée s'exprime $\forall x \geq 0$ par : $F(x) = 1 - \frac{1}{\Gamma(0,k)} T\left(\frac{x}{\theta}, k\right)$ où, $\Gamma(x, k) = \int_x^\infty t^{k-1} e^{-t} dt$

L'espérance et la variance d'une loi Log-Normal sont données par : $\mathbb{E}(X) = e^{\mu + \frac{\sigma^2}{2}}$ et $\mathbb{V}(X) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$

Loi Exponentielle : On dit qu'une variable aléatoire X suit une distribution de loi Exponentielle, de paramètre $\lambda > 0$, si la densité de la variable aléatoire est de la forme :

$$f(x) = \lambda e^{-\lambda x} 1_{x \geq 0}$$

La fonction de répartition associée s'exprime $\forall x \geq 0$ par : $F(x) = 1 - e^{-\lambda x}$

L'espérance et la variance d'une Exponentielle sont données par : $\mathbb{E}(X) = \frac{1}{\lambda}$ et $\mathbb{V}(X) = \frac{1}{\lambda^2}$

Loi de Pareto : On dit qu'une variable aléatoire X suit une distribution de Pareto de paramètres $\alpha > 0$, x_{min} , si la densité de la variable aléatoire est de la forme :

$$f(x) = \alpha (x_{min})^\alpha x^{-\alpha-1} 1_{x \geq x_{min}}$$

La fonction de répartition associée s'exprime $\forall x \geq x_{min}$ par : $F(x) = 1 - (x_{min})^\alpha x^{-\alpha}$

L'espérance et la variance d'une Pareto sont données par :

- Si $\alpha \leq 1$ alors l'espérance n'existe pas.
- Si $\alpha \geq 2$ alors : $\mathbb{E}(X) = \frac{\alpha x_{min}}{\alpha - 1}$ et si $\alpha > 2$ $\mathbb{V}(X) = \frac{\alpha}{(\alpha - 2)(\alpha - 1)^2} (x_{min})^2$

Le paramètre α d'une loi de Pareto est un paramètre de forme, cela implique que plus α est petit, plus la probabilité d'avoir des montants de charge de sinistres élevés est grande. De manière générale, le

marché de la réassurance s'accorde sur des valeurs de références pour ce paramètre de forme selon les différentes catégories de garanties, dans le cas de la RC Auto, $\alpha \sim 2,5$.

Loi Log Logistic : On dit qu'une variable aléatoire X suit une distribution log Logistic de paramètres $\alpha > 0, \beta > 0$, si la densité de la variable aléatoire est de la forme :

$$f(x) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{[1 + (x/\alpha)^\beta]^2} 1_{x \geq 0}$$

La fonction de répartition associée s'exprime $\forall x \geq 0$ par : $F(x) = \frac{1}{1+(x/\alpha)^\beta}$

L'espérance et la variance d'une Log Logistic sont données par :

- Si $\beta > 1$ alors $E(X) = \frac{\alpha(\pi/\beta)}{\sin(\pi/\beta)}$
- Si $\beta > 2$ alors $V(X) = \alpha^2(2\beta/\sin(2\beta) - \beta^2)$
- Sinon l'espérance et la variance ne sont pas définies.

Le choix des lois de probabilités pour ajustement étant arrêté, l'étape suivante consiste à déterminer les paramètres des lois retenues. Pour ce faire, plusieurs méthodes existent.

3.3.5 Estimation des paramètres de la loi

La première étape afin d'estimer les paramètres de notre loi consiste à recueillir des données sur les sinistres passés pour une période donnée, en se concentrant sur les sinistres graves. Ces données sont utilisées pour estimer les paramètres de la distribution de la gravité des sinistres, tels que la moyenne et l'écart-type. On peut utiliser différents types de distributions, tels que la distribution de Pareto, la distribution log-normale ou la distribution exponentielle.

Une fois que les paramètres de la distribution ont été estimés, il est possible de calculer les probabilités de réalisation des sinistres extrêmes, en utilisant les formules statistiques appropriées. Ensuite, le coût attendu des sinistres extrêmes peut être estimé en multipliant la probabilité de réalisation du sinistre par le coût moyen prévu par exemple.

Il est important de noter que cette méthode repose sur l'hypothèse que les sinistres suivent une distribution statistique spécifique, ce qui peut ne pas être le cas dans la réalité. De plus, cette méthode peut nécessiter une grande quantité de données pour être efficace, ce qui peut rendre sa mise en œuvre difficile dans certains contextes.

Pour déterminer les paramètres de notre loi de sévérité et ainsi déterminer le montant de nos sinistres, nous allons, en utilisant les observations historiques, estimer les paramètres de ce modèle. Mathématiquement, on cherche à estimer un vecteur de paramètres $\theta = (\theta_1, \theta_2, \dots, \theta_p)$, d'une fonction de densité de probabilité noté f_θ , à partir d'un échantillon de variables aléatoires $X = (X_1, X_2, \dots, X_n)$ avec

$x = (x_1, x_2, \dots, x_n)$, notre échantillon d'observation de taille n .

Ainsi, pour répondre à notre problématique, plusieurs méthodes statistiques pour estimer les paramètres du modèle existent, dans notre étude, nous utiliserons :

- La méthode des moments ;
- La méthode du maximum de vraisemblance.

3.3.5.1 Méthode des moments

La méthode des moments est une technique statistique utilisée pour estimer les paramètres d'une distribution de probabilité.

Le processus de la méthode des moments consiste à évaluer les moments théoriques de la distribution de probabilité à leurs équivalents empiriques calculés à partir des données disponibles. Les moments théoriques sont les moments de la distribution de probabilité et sont calculés en fonction des paramètres de cette dernière.

Dans le cas de la loi de coût des sinistres extrêmes, les paramètres à estimer sont généralement la moyenne, la variance et le coefficient d'asymétrie. Pour appliquer la méthode des moments, on calcule d'abord les moments empiriques à partir des données de coût des sinistres extrêmes disponibles. Ensuite, on égale ces moments empiriques aux moments théoriques de la distribution de probabilité.

En résolvant ce système d'équations, on peut déterminer les valeurs des paramètres de la loi de coût des sinistres extrêmes. Pour la plupart des lois, l'utilisation de la moyenne et de la variance est suffisante. Cependant, pour certaines lois la moyenne suffit (par exemple, la loi de Poisson), ou, au contraire, le coefficient d'asymétrie est aussi nécessaire (loi de Weibull par exemple).

Une fois les paramètres estimés, il est possible d'utiliser la loi de coût des sinistres extrêmes pour modéliser les coûts futurs de sinistres extrêmes. Cependant, il convient de noter que la méthode des moments est une technique statistique parmi d'autres et qu'elle peut présenter des limites en fonction des données disponibles et des hypothèses de modélisation utilisées.

Soit X une variable aléatoire. On définit le moment d'ordre k de la distribution de X , tel que $\forall k \in \mathbb{N}^* : m_k = \mathbb{E}(X^k)$.

La méthode des moments repose sur l'égalité entre les p premiers moments théoriques et les p premiers moments empiriques pour obtenir l'estimateur des moments du vecteur de paramètres θ ainsi, on a que les estimateurs empiriques des moments que l'on note \hat{m}_k se définissent par :

$$\hat{m}_k = \frac{1}{n} \sum_{i=1}^n x_i^k$$

Ainsi pour obtenir les estimateurs des paramètres par la méthode des moments, il convient de résoudre l'équation vectorielle $(m_p)_{1 \leq k \leq p} = (\hat{m}_p)_{1 \leq k \leq p}$.

3.3.5.2 Méthode du maximum de vraisemblance

Soit $X = (X_1, X_2, \dots, X_n)$, un vecteur aléatoire tel que les variables aléatoires $(X_i)_{1 \leq i \leq n}$ sont supposées indépendantes et identiquement distribuées de fonction de densité f_θ . On note (x_1, x_2, \dots, x_n) l'échantillon d'observations de la réalisation des variables $(X_i)_{1 \leq i \leq n}$.

Alors la vraisemblance associée à l'échantillon s'exprime comme étant :

$$\mathcal{L}(x_1, x_2, \dots, x_n | \theta) = \prod_{i=1}^n f_\theta(x_i)$$

L'estimateur du maximum de vraisemblance de θ , est la valeur qui maximise la fonction de vraisemblance $\mathcal{L}(x_1, x_2, \dots, x_n | \theta)$ par rapport à θ et que l'on notera : $\hat{\theta}_n$.

$$\hat{\theta}_n = \operatorname{argmax}_\theta (\mathcal{L}(x_1, x_2, \dots, x_n | \theta)).$$

En introduisant la fonction logarithme on peut définir la fonction de log-vraisemblance qui nous permettra de simplifier nos calculs de maximisation :

$$l(\theta) = \sum_{i=1}^n \ln (f_{\theta}(x_i))$$

L'estimateur du maximum de vraisemblance est alors défini comme étant :

$$\widehat{\theta}_n = \operatorname{argmax}_{\theta}(l(\theta))$$

Par conséquent on définit les fonctions de score $\forall J \in \llbracket 1 : p \rrbracket$ par :

$$S_j(\theta) = \frac{\partial}{\partial \theta_j} \ln (l(\theta))$$

Tel que maximiser la fonction de log-vraisemblance revient à résoudre les p équations suivantes :

$$\forall J \in \llbracket 1 : p \rrbracket, S_j(\theta) = 0$$

La méthode du maximum de vraisemblance est une méthode statistique couramment utilisée pour estimer les paramètres d'une distribution de probabilité. Cette méthode consiste à trouver les valeurs des paramètres de la distribution qui maximisent la vraisemblance de l'échantillon observé.

La vraisemblance est la probabilité d'observer l'échantillon pour une valeur donnée des paramètres de la distribution. Par conséquent, la méthode du maximum de vraisemblance consiste à estimer les valeurs des paramètres de la distribution qui maximisent la probabilité d'observer l'échantillon. Cette probabilité est également appelée la fonction de vraisemblance.

Pour trouver les paramètres qui maximisent la fonction de vraisemblance, la méthode du maximum de vraisemblance utilise une technique d'optimisation, comme la méthode de Newton-Raphson ou la méthode de gradient, qui permettent de trouver les valeurs des paramètres pour lesquelles la fonction de vraisemblance est maximale.

Une fois que les paramètres optimaux ont été trouvés, la distribution de probabilité peut être utilisée pour estimer les coûts de sinistres, les primes d'assurance ou d'autres indicateurs de risque.

Il est important de noter que la méthode du maximum de vraisemblance est basée sur plusieurs hypothèses, notamment l'indépendance des observations, l'homogénéité des paramètres de la distribution et la représentativité de l'échantillon observé. Ces hypothèses doivent être examinées attentivement avant d'appliquer cette méthode pour estimer les paramètres d'une distribution de probabilité.

Cette méthode présente l'avantage d'être rigoureuse pour toutes les lois, et permet d'obtenir des écarts-types approximatifs pour les estimateurs des paramètres.

Ainsi, nous avons vu que le choix de la méthode utilisée repose essentiellement sur les propriétés recherchées pour l'estimateur des paramètres du modèle. Dans notre étude, ce sont ces deux méthodes qui sont utilisées. Ainsi lorsque la méthode du maximum de vraisemblance atteint ses limites, nous avons recours à la méthode des moments afin de déterminer les paramètres de nos lois. Il existe toutefois d'autres méthodes comme la méthode des quantiles ou bayésienne pour répondre à notre problématique.

3.3.6 Ajustement des lois

Dès lors que nous avons déterminé les différents paramètres des lois théoriques qui pourraient convenir à notre échantillon, il convient de déterminer laquelle parmi ces lois reflète le mieux nos observations. Différentes méthodes sont alors mises à notre disposition pour répondre à notre problématique :

- une méthode graphique avec les graphiques Quantile-Quantile qui permettent d'évaluer graphiquement la qualité de la distribution de probabilité.
- des méthodes statiques qui sont des mesures quantitatives qui permettent de comparer la qualité de l'ajustement des différentes lois candidates aux données observées.

3.3.6.1 Les graphiques Quantile-Quantile (QQ-plot)

Le graphique quantile-quantile (QQ-plot) est un outil de visualisation utile pour comparer une distribution théorique à une distribution empirique. Il permet de vérifier si les données observées sont compatibles avec une distribution donnée, en comparant les quantiles empiriques aux quantiles théoriques.

Soit $X = (X_1, X_2, \dots, X_n)$, un vecteur aléatoire tel que les variables aléatoires $(X_i)_{1 \leq i \leq n}$ sont supposées indépendantes et identiquement distribuées suivant la même loi qu'une certaine variable aléatoire X . On note alors la fonction de répartition de X par :

$$F_x(t) = \mathbb{P}(X \leq t)$$

On définit l'inverse généralisé $F^-(t)$ par :

$$F^-(t) = \inf\{x \in \mathbb{R}, F(x) \geq t\}, 0 \leq t \leq 1$$

La statistique d'ordre associé à $x = (x_1, x_2, \dots, x_n)$ est obtenu en ordonnant dans l'ordre croissant l'échantillon d'observation, que l'on note : $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$.

Le graphique quantile-quantile de la variable aléatoire X , est alors obtenu en traçant le graphique suivant :

$$\left\{ \left(\left(F^- \frac{i}{n+1}; x_{(i)} \right), \quad i \in \llbracket 1:n \rrbracket \right) \right\}$$

Graphiquement, le QQ-plot est créé en traçant les quantiles théoriques sur l'axe horizontal et les quantiles empiriques correspondants sur l'axe vertical. Si les données suivent la distribution théorique, les points sur le graphique doivent tomber sur une droite diagonale. Si les points sont décalés de cette diagonale, cela indique que les données ne suivent pas exactement la distribution théorique.

Si les données ont été générées par la même loi de X alors le graphique quantile-quantile doit avoir une apparence linéaire. Quand les points sont alignés sur une droite, les quantiles empiriques sont égaux aux quantiles du modèle, l'ajustement par le modèle est correct. On peut donc accepter l'hypothèse que les échantillons sont issus d'une même loi statistique.

S'il existe des points anormaux dans les données, ceux-ci peuvent être identifiés à l'aide du QQ-plot. Les points anormaux sont les points qui sont éloignés de la droite.

3.3.6.2 Les tests statistiques

Soit $X = (X_1, X_2, \dots, X_n)$, un vecteur aléatoire tel que les variables aléatoires $(X_i)_{1 \leq i \leq n}$ sont supposées indépendantes et identiquement distribuées avec une distribution de probabilité inconnue P .

Nous cherchons à tester l'hypothèse que P est égale à une certaine distribution particulière, pour cela nous allons utiliser les tests d'adéquations de Kolmogorov-Smirnov et d'Anderson-Darling. Pour accepter ou rejeter ces tests, nous utiliserons le critère de la p-value qui correspond à la probabilité pour que la statistique de test dépasse, sous l'hypothèse H_0 , la valeur seuil. Plus la p-value est proche de 0, plus l'hypothèse H_0 sera rejetée. Dans notre étude, nous utiliserons un seuil de 5%, nous accepterons donc l'hypothèse H_0 quand la p-value sera au-delà de 5%.

- Pour chaque loi candidate, nous pouvons comparer la distribution empirique de nos données avec la distribution théorique de chaque loi. Les tests les plus couramment utilisés sont les tests de Kolmogorov-Smirnov, le test de Chi-2 et celui d'Anderson-Darling. Ainsi en comparant les résultats des tests pour chaque loi candidate il convient de sélectionner celle qui présente le plus petit écart entre la distribution empirique de nos données et celle de la distribution théorique de la loi.

3.3.6.2.1 Test de Kolmogorov-Smirnov

Le test de Kolmogorov-Smirnov est un test non paramétrique utilisé pour déterminer si un échantillon suit une distribution particulière. Le test compare la distribution cumulative observée de l'échantillon avec la distribution cumulative théorique de la loi de probabilité choisie. Si la différence entre les deux distributions est significative, nous pouvons rejeter l'hypothèse nulle selon laquelle l'échantillon suit cette loi de probabilité.

La statistique de test D_n est calculée comme la plus grande différence absolue entre les deux distributions. Plus la valeur de D_n est élevée, plus il est probable que la loi de probabilité choisie ne convient pas bien aux données. Le test de Kolmogorov-Smirnov est largement utilisé car il est applicable à une large gamme de distributions de probabilité et ne nécessite pas de spécifier les paramètres de la distribution à l'avance.

Mathématiquement, soit X une variable aléatoire de fonction de répartition F_x .

Soit $F_0 : \mathbb{R} \rightarrow [0; 1]$ une fonction de répartition supposée continue, on souhaite tester l'hypothèse $H_0 : F_x = F_0$, contre l'hypothèse $H_1 : F_x \neq F_0$.

Le test de Kolmogorov-Smirnov repose sur les propriétés asymptotiques de la fonction de répartition empirique $F_n(t)$ définie par :

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n 1_{X_i \leq t}$$

Et par conséquent,

$$D_n = \sup_{t \in \mathbb{R}} |F_n(t) - F_x(t)|$$

Cette quantité représente bien la plus grande différence, en valeur absolue entre la fonction de répartition empirique et la fonction de répartition de la loi testée.

3.3.6.2.2 Test d'Anderson Darling

Le test d'Anderson-Darling est un autre test statistique utilisé pour déterminer si un échantillon suit une distribution particulière. Ce test est similaire au test de Kolmogorov-Smirnov, mais il accorde plus de poids aux queues de la distribution. Cela signifie que le test est plus sensible aux écarts de distribution dans les queues, plutôt qu'au centre de la distribution. C'est pourquoi ce test est plus pertinent que le test de Kolmogorov-Smirnov dans le contexte de l'étude des sinistres aux valeurs extrêmes. La statistique de test A_n^2 est calculée en utilisant la distance entre les deux distributions et leurs poids relatifs.

Elle est définie par :

$$A_n^2 = -n + \sum_{i=1}^n \frac{(2i-1)}{n} \ln \left(n F(X_{(i)}) + \ln (1 - (F(X_{(n+1-i)})) \right)$$

Où F est la fonction de répartition de X.

D'autres tests d'adéquations peuvent être utilisés :

- Le test de Shapiro-Wilk est utilisé pour déterminer si un échantillon suit une distribution normale. Le test compare les données de l'échantillon à la distribution normale attendue et produit une p-value. Si la p-value est inférieure à un niveau de signification spécifié (5% ici), alors nous pouvons rejeter l'hypothèse nulle selon laquelle l'échantillon suit une distribution normale. Cela suggère que la loi de probabilité choisie pour notre modèle ne doit pas être une distribution normale.
- Le test du χ^2 qui est un test statistique utilisé pour évaluer l'adéquation entre les données observées et la distribution de probabilité théorique. Le test compare le nombre d'observations attendues dans chaque catégorie de la distribution avec le nombre d'observations réelles dans chaque catégorie. La statistique de test χ^2 est calculée en utilisant la différence entre les observations réelles et les observations attendues pour chaque catégorie de la distribution.

3.3.6.3 Critères Statistiques

Les critères statistiques permettent de quantifier la qualité de l'ajustement. Ces méthodes statistiques sont des outils qui permettent de contribuer à la décision pour l'actuaire qui modélise les sinistres. Dans notre étude nous utiliserons : le critère d'information d'Akaike, Bayésien et celui de l'erreur quadratique moyenne.

3.3.6.3.1 Le critère d'information d'Akaike (AIC)

Le critère d'information d'Akaike (en anglais Akaike Information Criterion) Critère d'information d'Akaike (AIC) est un outil couramment utilisé pour sélectionner le meilleur modèle statistique pour un ensemble de données. Le critère d'Akaike prend en compte à la fois la qualité d'ajustement du modèle et sa complexité, en pénalisant les modèles plus complexes. Le calcul de l'AIC pour un modèle donné est donné par la formule suivante :

$$AIC = 2k - 2\log(L(\theta))$$

où k est le nombre de paramètres dans le modèle, et $L(\theta)$ est la fonction de vraisemblance maximale.

L'AIC est une mesure relative de la qualité de l'ajustement du modèle, ce qui signifie que les modèles avec les valeurs d'AIC les plus faibles sont considérés comme les meilleurs modèles car cela correspond au modèle ayant la vraisemblance la plus élevée.

3.3.6.3.2 Le critère d'information Bayésien (BIC)

Le critère d'information bayésien (en anglais Bayesian Information Criterion) est un outil similaire à l'AIC qui est également utilisé pour sélectionner le meilleur modèle statistique pour un ensemble de données. Comme l'AIC, le BIC prend en compte la qualité d'ajustement du modèle et sa complexité. Cependant, contrairement à l'AIC, le BIC pénalise davantage les modèles plus complexes.

Le calcul du BIC pour un modèle donné est donné par la formule suivante :

$BIC = k \log(n) - 2\ln(L(\theta))$ où k est le nombre de paramètres dans le modèle, n la taille de l'échantillon et $L(\theta)$ la fonction de vraisemblance maximale.

Le terme de $k \log(n)$, représente une pénalité de la sur-paramétrisation du modèle. Si on a $\log(n) > 2$ alors le critère BIC pénalise plus la sur-paramétrisation du modèle que le critère AIC. Le critère BIC favorise donc les modèles avec le moins de paramètres.

Comme pour l'AIC, les modèles avec les valeurs de BIC les plus faibles sont considérés comme les meilleurs modèles.

3.3.6.3.3 L'erreur quadratique moyenne

L'erreur quadratique moyenne est une mesure de l'exactitude de la prédiction pour un modèle donné. L'EQM est calculée en prenant la somme des carrés des écarts entre les valeurs observées et prédites, puis en divisant par le nombre total d'observations. Une EQM faible indique une bonne précision de la prédiction. L'EQM est donnée par la formule suivante :

$$EQM = (1/n) \sum (y - \hat{y})^2$$

où y est la valeur observée, \hat{y} est la valeur prédite, et n est le nombre total d'observations. L'EQM est souvent utilisée pour comparer les performances de différents modèles prédictifs et pour sélectionner le meilleur modèle.

En résumé, pour déterminer quelle loi de probabilité correspond le mieux à un échantillon de données, il est courant d'utiliser une combinaison de ces critères statistiques pour évaluer la qualité d'ajustement du modèle.

3.3.7 Détermination de la charge annuelle par simulation de Monte-Carlo

Une fois l'ajustement de nos lois effectué, il convient d'utiliser une méthode non paramétrique pour déterminer la charge annuelle de sinistre. Dans notre étude, nous avons donc convenu de modéliser séparément le nombre de sinistres obtenu, de la sévérité, puis nous allons en déduire une charge annuelle. Dans les parties précédentes, nous avons déterminé les lois de fréquence et de sévérité. Notons F la distribution de fréquence, et S la distribution de sévérité. Nous cherchons maintenant à déterminer la charge annuelle par simulation de Monte-Carlo.

Afin de répondre à cette problématique et prendre en compte l'incertitude inhérente à la modélisation des sinistres, nous utiliserons la méthode de Monte Carlo qui est une technique numérique de simulation stochastique. Le principe de base de la méthode de Monte Carlo est de générer des échantillons aléatoires d'une distribution de probabilité et de les utiliser pour estimer la moyenne, la variance, le maximum, le minimum, etc. Ces quantités sont ensuite utilisées pour estimer les caractéristiques de la distribution de probabilité.

La méthode de Monte-Carlo est fondée sur les principes de la loi des grands nombres et du théorème central limite. Si on considère (X_1, X_2, \dots, X_N) la suite i.i.d des coûts des sinistres avec une espérance $E[X_i] = \mu$ et une variance $Var[X_i] = \sigma^2$. Alors selon la loi des grands nombres, : $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N X_i = \mu$

Autrement dit, l'estimation de la charge moyenne des sinistres sur un grand nombre de simulations se rapproche de l'espérance de distribution des coûts des sinistres.

Et selon le théorème central limite, la somme (ou moyenne) des X_i suit approximativement une distribution normale avec une moyenne μ et une variance $\frac{\sigma^2}{N}$. Cela signifie que la distribution de la charge annuelle des sinistres sur un grand nombre de simulations tend vers une distribution normale, et l'erreur standard associée à l'estimation de la charge annuelle des sinistres diminue avec l'augmentation du nombre de simulations. Sur ces principes nous allons :

- Générer une variable aléatoire N à partir de la distribution de fréquence F . Cette quantité représente ainsi le nombre de sinistres.
- Pour chaque $i = 1, 2, \dots, N$, nous allons générer une variable aléatoire X_i à partir de la distribution de sévérité S . Cette quantité représente le montant du i -ème sinistre.
- La somme des montants totaux des sinistres pour cette première simulation A_1 résulte alors de la somme des coûts des sinistres. On a alors $A_1 = \sum_{i=1}^N X_i$ qui représente la charge annuelle pour notre première simulation.

Nous pouvons déterminer la moyenne, la variance, l'écart-type, le maximum, le minimum, de cette simulation. Puis en répétant ces étapes pour un grand nombre de simulations, et finalement la moyenne des charges annuelles de toutes les simulations est calculée, fournissant une estimation de la charge annuelle des sinistres.

Conformément à la loi des grands nombres, cette estimation tend vers la véritable charge annuelle des sinistres lorsque le nombre de simulation tend vers l'infini. De plus, selon le théorème de centrale limite l'erreur standard de l'estimation diminue comme $\frac{1}{\sqrt{N}}$, nous indiquant que la précision de l'estimation augmente avec le nombre de simulations.

Cette méthode nous offre un outil flexible pour l'évaluation des risques en utilisant l'approche probabiliste fréquence-sévérité. Elle nous permet de générer une distribution réaliste des montants totaux des sinistres, tout en tenant en compte de l'aléa et des corrélations dans les fréquences et sévérités des sinistres.

4 Théorie des Valeurs Extrêmes

4.1 Introduction générale

Dans les parties précédentes, nous avons vu que tarifier un traité de réassurance en excédent de sinistre par l'approche probabiliste consiste à calculer un coût moyen à charge du réassureur entre la priorité et le plafond de nos tranches de réassurance. L'une des difficultés que nous avons souligné est le fait que dans certains cas, la tranche de réassurance n'est pas travaillante. Par conséquent nous devons être en mesure d'estimer un coût moyen à charge dans des grandeurs qui appartiennent à la queue de distribution des sinistres pour lesquelles nous n'avons que très peu voire pas d'informations. L'étude de la queue de distribution des sinistres est donc une étape cruciale dans notre étude afin d'appréhender le risque lié aux sinistres extrêmes et la théorie des valeurs des extrêmes nous offre un cadre solide pour notre étude. Cette théorie se concentre sur les valeurs les plus élevées d'une distribution de données et utilise des modèles mathématiques pour estimer les probabilités d'événements extrêmes.

Dans le cadre de notre étude, la théorie des valeurs extrêmes est utilisée pour modéliser les sinistres les plus rares et les plus coûteux. Cette méthode est particulièrement utile pour estimer les niveaux de rétention nécessaires pour couvrir les sinistres extrêmes, ainsi que pour déterminer les primes de réassurance appropriées. Cette méthode permet également de mieux comprendre la fréquence et la gravité des événements extrêmes, ce qui est essentiel pour la gestion des risques et la planification financière à long terme.

Cette théorie est particulièrement adaptée à la modélisation de nos sinistres extrêmes car elle repose sur des distributions spécifiques qui sont conçues pour modéliser les queues supérieures des distributions de données. Ces distributions sont souvent utilisées pour modéliser les événements rares et coûteux qui se produisent dans de nombreux domaines, y compris l'assurance.

Un des principes de cette théorie est le fait que les valeurs extrêmes d'une distribution suivent une loi de probabilité différente de celle des valeurs centrales. En effet, les valeurs extrêmes sont souvent plus éloignées de la moyenne et exercent une influence plus importante sur les risques encourus.

La théorie des valeurs extrêmes s'appuie sur deux fondements importants :

- Le théorème de Fisher-Tippett-Gnedenko, qui décrit les trois types de lois de probabilité qui peuvent caractériser les valeurs extrêmes d'une distribution.
- Le théorème de Pickands-Balkema-de Haan, qui permet d'estimer la distribution des valeurs extrêmes en se focalisant sur les excès au-delà d'un certain seuil.

De plus, la théorie des valeurs extrêmes utilise des distributions de probabilité comme la distribution de Gumbel, la distribution de Fréchet ou la distribution de Weibull, qui ont des queues de distribution plus épaisses que les distributions normales ou gaussiennes. Ces distributions ont des propriétés mathématiques qui permettent de modéliser les événements extrêmes de manière plus précise que les méthodes statistiques classiques.

L'objectif de cette partie est de présenter les outils spécifiques à cette théorie qui nous permettront ainsi de modéliser les valeurs extrêmes des sinistres en réassurance.

4.1.1 Domaine d'attraction

Avant d'énoncer les différents théorèmes que nous allons utiliser dans notre étude, il convient de définir quelques notions liées au domaine d'attraction. Le domaine d'attraction d'une distribution de probabilité est l'ensemble des distributions de probabilité des sommes partielles normalisées d'une suite de variables aléatoires indépendantes et identiquement distribuées qui convergent en loi vers la distribution considérée. En d'autres termes, si une suite de variables aléatoires suit une loi de probabilité appartenant au domaine d'attraction d'une certaine distribution, alors la somme de ces variables normalisées converge en loi vers cette distribution.

Mathématiquement, soit X une variable aléatoire qui admet F comme fonction de répartition et Y une variable aléatoire qui admet H comme fonction de répartition non-dégénérée (i.e. une fonction de répartition qui n'est pas associée à une variable constante presque-sûrement).

On dit que F appartient au domaine d'attraction de H (i.e. $F \in D(H)$), s'il existe deux suites $(a_n) > 0$, et (b_n) , telles qu'en tout point de continuité x de H :

$$F^n(xa_n + b_n) \xrightarrow{\infty} H(x) \Leftrightarrow \frac{X_{(n)} - b_n}{a_n} \xrightarrow{\mathcal{L}} Y$$

Les suites $(a_n) > 0$, et (b_n) dépendent de F mais pour n fixé, sont des simples constantes de normalisation qui doivent satisfaire $\forall x \in \mathbb{R} : \mathbb{P}\left(\frac{X_{(n)} - b_n}{a_n} \leq x\right) \xrightarrow{\mathcal{L}} H(x)$.

4.1.1.1 Domaine d'attraction de la distribution de Fréchet

On dit que la distribution de F appartient au domaine d'attraction de la loi de Fréchet ϕ_α ($\alpha > 0$), si et seulement si : $\bar{F}(x) = x^{-\alpha}L(x)$

Avec L une fonction à variation lente (i.e. L mesurable et positive sur $]0; \infty[$, tel que $\forall x > 0$, $\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1$).

4.1.1.2 Domaine d'attraction de la distribution de Weibull

On dit que la distribution de F appartient au domaine d'attraction de la loi de Weibull ψ_α , si et seulement si : $\bar{F}(x^F - x^{-1}) = x^{-\alpha}L(x)$ et $x^F < \infty$

4.1.1.3 Domaine d'attraction de la distribution de Gumbel

On dit que la distribution de F appartient au domaine d'attraction de la loi de Gumbel ν , si et seulement si : $\lim_{x \rightarrow x^F} \frac{\bar{F}(x+tg(x))}{\bar{F}(x)} = \exp(-t)$, $t \in \mathbb{R}$

4.2 Le théorème de Fisher-Tippett-Gnedenko (en anglais GEV pour Generalized Extreme Value)

Soit $X = (X_1, X_2, \dots, X_n)$, un vecteur aléatoire tel que les variables aléatoires $(X_i)_{1 \leq i \leq n}$ sont supposées indépendantes et identiquement distribuées et de même fonction de répartition F .

On note $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$, l'échantillon précédent, ordonné de manière croissante.

On cherche la fonction de répartition du maximum de X , si on considère $M_n = \max(X_1, X_2, \dots, X_n)$, alors on a que :

$$\begin{aligned} \mathbb{P}(M_n \leq x) &= \mathbb{P}(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) \\ &= \mathbb{P}(X_1 \leq x) \mathbb{P}(X_2 \leq x) \dots \mathbb{P}(X_n \leq x) \text{ car i. i. d} \\ &= [F(x)]^n \end{aligned}$$

Notons maintenant, $x^F = \sup\{x | F(x) < 1\}$, le point extrémal de F , alors le support de F peut être borné si $x^F < \infty$ (lois uniforme, beta...) ou infini si $x^F = \infty$ (lois exponentielle, Pareto, normale ...).

On a alors que : $M_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} x^F$ et que la distribution asymptotique de M_n est dégénérée. F étant inconnue, cela n'apporte pas plus d'informations sur les queues de distributions.

Le théorème de Fisher-Tippett-Gnedenko stipule alors :

Soit F une fonction de répartition et G une fonction de répartition non-dégénérée, s'il existe deux suites $(a_n) > 0$, et (b_n) de réels et $\gamma \in \mathbb{R}$ un paramètre de forme telle que si $F \in D(G)$ alors la fonction de répartition G est du même type que la fonction de répartition de la loi GEV que l'on note H_γ tel que :

$$\mathbb{P}\left(\frac{X_{(n)} - b_n}{a_n} \leq x\right) \xrightarrow[n \rightarrow \infty]{} H_\gamma(x)$$

$$H_\gamma = \begin{cases} \exp\left(-\left(1 + \gamma x^{-\frac{1}{\gamma}}\right)\right) & \text{si } \gamma \neq 0 \\ \exp(-e^{-x}) & \text{si } \gamma = 0 \end{cases} \quad \forall x > -\frac{1}{\gamma}$$

Selon le signe de γ on définit 3 types de domaines d'attraction : domaine d'attraction de Fréchet de Weibull et de Gumbel.

Weibull $\gamma < 0$	Gumbel $\gamma = 0$	Fréchet $\gamma > 0$
- Uniforme	- Normale	- Pareto
- Beta	- Exponentielle	- Log-gamma
	- Log-normale	- Student
	- Gamma	- Cauchy
	- Weibull	

Tableau 20 : Exemple de domaine d'attractions pour des lois usuelles

Si $\gamma < 0$: Ce sont les distributions à queues fines et c'est le domaine d'attraction de Weibull, $x^F < \infty$ et la fonction de survie converge vers 0 à une vitesse polynomiale

Si $\gamma = 0$: Ce sont les queues de distributions à queues légères et c'est le domaine d'attraction de Gumbel. $x^F = \infty$ ou $x^F < \infty$ et la fonction de survie converge vers 0 à une vitesse exponentielle.

Si $\gamma > 0$: Ce sont les queues de distributions à queues lourdes et c'est le domaine d'attraction de Fréchet, $x^F = \infty$ et la fonction de survie converge vers 0 à une vitesse polynomiale.

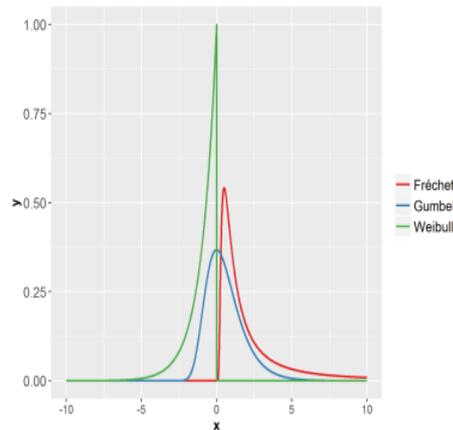


Figure 5 : Comparaison des densités des trois lois

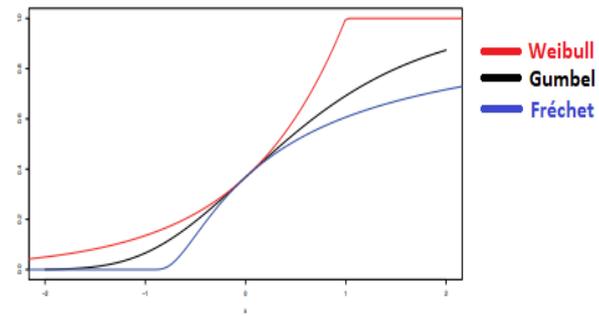


Figure 6 : Fonction de répartition de la loi EVD dans les trois domaines d'attractions

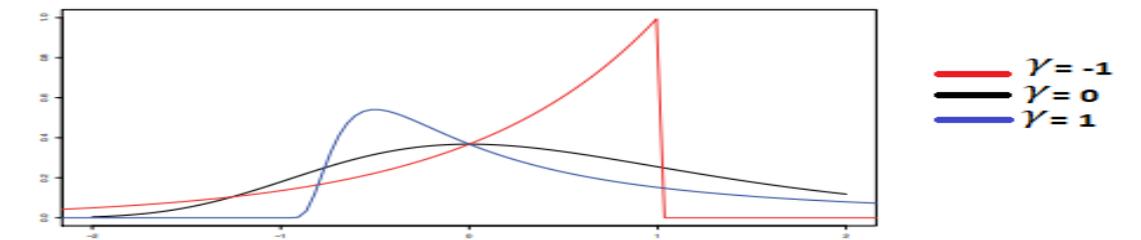


Figure 7 : Densité de la distribution EVD

Cette méthode repose uniquement sur l'étude du maximum de l'échantillon. Dans notre étude, nous avons décidé de nous intéresser à un domaine moins restrictif que uniquement l'étude du maximum de l'échantillon. Pour cela, nous allons nous aborder une autre approche de la théorie des valeurs extrêmes.

La modélisation des sinistres extrêmes en réassurance à l'aide de la distribution de GEV présente certaines limites. En effet, elle est basée sur l'hypothèse que les données suivent une loi de Gumbel, de Fréchet ou de Weibull, qui sont des lois de valeurs extrêmes. Cette hypothèse peut ne pas être vérifiée dans certains cas où les données ne suivent pas ces lois. La GEV suppose que les queues de distribution sont symétriques, ce qui peut ne pas être le cas dans certaines situations. De plus, cette méthode est souvent utilisée avec des données sur une période relativement courte, ce qui peut limiter la précision des estimations des paramètres.

La méthode du dépassement de seuil (excess loss method) est quant à elle utilisée pour contourner ces limites en se concentrant sur les excès au-delà d'un certain seuil plutôt que sur l'ensemble de la distribution. Cette méthode permet de modéliser la queue de la distribution de manière plus précise, ce qui permet de se concentrer sur les sinistres les plus graves et de mieux modéliser leur comportement. Cette méthode est particulièrement utile lorsque la distribution des sinistres ne suit pas une distribution continue comme la GEV, ou lorsque la période d'observation est relativement courte. Elle est donc particulièrement utile pour modéliser les événements extrêmes qui ont une faible probabilité de se produire, mais qui peuvent causer des pertes importantes.

4.3 Modèle du dépassement de seuil (Peak Over Threshold)

Le modèle POT (Peak Over Threshold) est une méthode alternative pour modéliser les valeurs extrêmes. Il est utilisé lorsque la modélisation par la GEV n'est pas adaptée pour représenter la distribution des valeurs extrêmes.

Le principe du modèle POT est de considérer uniquement les valeurs supérieures à un seuil préalablement fixé. On peut ainsi considérer que ces valeurs sont issues d'une distribution de probabilité généralisée Pareto (GPD). La GPD est une distribution à queue lourde, qui est appropriée pour modéliser les valeurs extrêmes. La méthode consiste à sélectionner les observations supérieures à un seuil μ , à estimer les paramètres de la distribution GPD à partir de ces observations, et à extrapoler la distribution au-delà du seuil. Puis nous verrons comment procéder à l'estimation des paramètres de la GPD.

4.3.1 Distribution de Pareto généralisée

On dit qu'une variable aléatoire X de paramètre $\sigma > 0, \mu < x^F$ et $\gamma \in \mathbb{R}$ suit une loi de Pareto généralisée que l'on note : $X \sim GPD(\mu, \sigma, \gamma)$, si la fonction de répartition de X est de la forme :

$$G_{\mu, \sigma, \gamma}(x) = \begin{cases} 1 - \left(1 + \gamma \left(\frac{x - \mu}{\sigma}\right)\right)^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0 \\ 1 - \exp\left(-\frac{x - \mu}{\sigma}\right) & \text{si } \gamma = 0 \end{cases}$$

$\forall x > 0$, tel que $1 + \frac{\gamma x}{\sigma} > 0$.

- μ est le paramètre de position
- σ est le paramètre d'échelle
- γ est le paramètre de forme, il détermine la queue de la distribution, c'est-à-dire la rapidité avec laquelle elle décroît. Si γ est grand, la queue est plus lourde, tandis que si γ est petit, la queue est plus légère.

La distribution de Pareto standard est un cas particulier de la distribution de Pareto généralisée pour lequel $\mu = 0$ et $\sigma = 1$

On a aussi les correspondances suivantes :

- Si $\gamma < 0$, alors $GPD(\mu, \sigma, \gamma) \rightarrow$ loi Béta
- Si $\gamma = 0$, alors $GPD(\mu, \sigma, \gamma) \rightarrow$ loi Exponentielle
- Si $\gamma > 0$, alors $GPD(\mu, \sigma, \gamma) \rightarrow$ loi de Pareto

4.3.2 Loi des excès

Soit X une variable aléatoire de fonction de répartition F .

Comme précédemment, notons $x^F = \sup\{x \in \mathbb{R} | F(x) < 1\}$, le point extrême de F .

Alors $\forall u < x^F$, on appelle fonction de répartition des excès la fonction suivante :

$$F_u(x) = \mathbb{P}(X - u \leq x | X > u) = 1 - \frac{1 - F(x + u)}{1 - F(u)}, x \geq 0, \text{ tel que } 0 \leq x \leq x^F - u$$

La fonction de distribution des excès représente la probabilité que X dépasse le seuil μ d'au plus la quantité x .

Dès lors, nous allons introduire le théorème de Pickands qui est un théorème fondamental dans la théorie des valeurs extrême lorsque l'on veut étudier une loi au-delà d'un certain seuil.

4.3.3 Théorème de Pickands

Nous avons, $F \in D(H_\gamma)$ si et seulement si :

$$\lim_{\mu \rightarrow x^F} \sup |F_u(x) - G_{\sigma(u),\gamma}(x)| = 0$$

Avec σ une fonction strictement positive et $G_{\sigma(u),\gamma}(x) \sim GPD(0, \sigma(u), \gamma)$.

On en déduit donc que, pour un μ assez grand, la loi des excès est approchée par une loi de Pareto Généralisée et F_μ est la fonction de répartition des excès au-delà du seuil de μ .

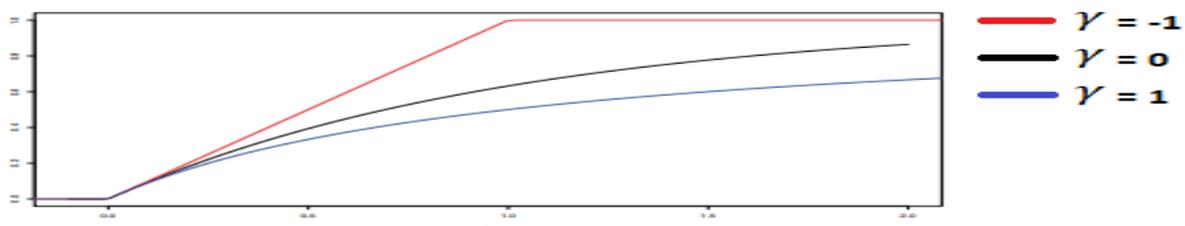


Figure 8 : Fonction de répartition de la GPD

En traçant la fonction de répartition de la GPD pour plusieurs valeurs de γ avec $\sigma = 1$ fixe, on voit clairement que plus γ est élevée, plus les valeurs extrêmes sont importantes.

Démontrons ce résultat : D'après l'approximation de la loi de maxima F^n par une $GEV(\mu, \sigma, \gamma)$, on a que :

$$F^n(x) \approx H_{\mu,\sigma,\gamma}(x) = \exp\left(-\left(1 + \gamma\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\gamma}}\right)$$

On a alors que : $n \ln(F(x)) \approx -\left(1 + \gamma\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\gamma}}$

Or $\lim_{x \rightarrow \infty} \ln(F(x)) = -(1 - F(x))$, nous avons donc que :

$$1 - F(x) \approx \frac{1}{n} \left(1 + \gamma\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\gamma}}$$

Or, d'après la loi des excès nous avons que : $F_\mu(x) = 1 - \frac{1-F(x+\mu)}{1-F(\mu)}$

Ainsi, en injectant la dernière équation on obtient que :

$$\begin{aligned} F_u(x) &= 1 - \frac{1 - F(x + u)}{1 - F(u)} = 1 - \left(1 + \gamma\left(\frac{x}{\sigma + \gamma(u - \mu)}\right)\right)^{-\frac{1}{\gamma}} = 1 - \left(1 + \gamma\frac{x}{\sigma(u)}\right)^{-\frac{1}{\gamma}} \\ &= G_{\sigma(u),\gamma}(x), \text{ avec } \sigma(u) = \sigma + \gamma(u - \mu) \end{aligned}$$

Le théorème de Pickands établit donc une relation entre la distribution des valeurs extrêmes d'une variable aléatoire et la distribution de ses excès au-dessus d'un seuil élevé. Ainsi, pour une séquence de seuils croissants, les excès au-dessus de chaque seuil convergent vers une distribution de valeurs extrêmes généralisée. Cette distribution est décrite par les mêmes paramètres que la GEV utilisée pour modéliser les valeurs extrêmes, mais avec une échelle et une localisation différente.

En d'autres termes, le théorème de Pickands permet de lier la distribution des excès au-dessus d'un seuil élevé à la distribution des valeurs extrêmes, et fournit une approximation pour la queue d'une distribution en se concentrant sur les valeurs qui dépassent un seuil élevé, ce qui permet de mieux comprendre le comportement de ces variables aléatoires rares mais importantes dans le contexte de notre étude. Il nous reste maintenant à déterminer les paramètres de la GPD.

4.4 Estimation des paramètres de la GPD

Précédemment, nous avons vu que au-dessus d'un certain seuil, la distribution des valeurs extrêmes était approchée par une loi de Paréto Généralisée. Ainsi, dans cette partie, nous allons nous intéresser à l'estimation des paramètres σ et γ de la loi de Pareto Généralisée dans le cas de l'approche du modèle du dépassement de seuil (POT). Ces paramètres doivent être estimés à partir de l'échantillon dépassant le seuil u , correspondant à la u_n valeur de notre échantillon d'observation (X_1, X_2, \dots, X_n) .

On note alors : $\forall i \in \llbracket 1 : u_n \rrbracket Y_i = X_{(n-u_n+i)} - X_{(n-u_n)}$ l'échantillon des observations au-delà du seuil. Les $(Y_i)_{i \in \llbracket 1 : k_n \rrbracket}$ sont supposées i.i.d et de même loi $GPD(\mu, \sigma, \gamma)$.

4.4.1 Estimation des paramètres de la GPD par la méthode du maximum de vraisemblance

La méthode du maximum de vraisemblance est une méthode couramment utilisée pour estimer les paramètres de la GPD dans le modèle POT. La méthode consiste à maximiser la fonction de vraisemblance, qui mesure la probabilité de l'ensemble des données observées en se concentrant sur les valeurs au-dessus d'un certain seuil étant donné les paramètres de la distribution.

La première étape de la méthode est de déterminer un seuil u approprié. Pour cela, on peut utiliser différentes méthodes, que l'on explicitera dans la partie suivante. Supposons que le seuil choisi soit u .

La fonction de vraisemblance pour les paramètres γ et σ est donnée par :

$$L(\gamma, \sigma | y_1, \dots, y_n) = (u_n/\sigma) [1 + \gamma(y_i - u)/\sigma]^{-1/\gamma - 1}$$

Le log de la fonction de vraisemblance est donné par :

$$\log L(\gamma, \sigma | y_1, \dots, y_n) = -u_n \ln \sigma - (1/\gamma) \sum_{i=1}^{u_n} \ln [1 + \gamma(y_i/\sigma)]$$

Lorsque $1 + \frac{\gamma}{\sigma} y_i > 0, \forall i \in \llbracket 1 : u_n \rrbracket$ sinon $L_{MV}(\gamma, \sigma) = -\infty$.

La méthode du maximum de vraisemblance consiste à trouver les valeurs des paramètres γ et σ qui maximisent la fonction de vraisemblance. Cela revient à résoudre l'équation suivante

- $d/d\gamma [\log L(\gamma, \sigma | y_1, \dots, y_n)] = (u_n/\gamma) - \sum_{i=1}^{u_n} (y_i/\sigma) [1 + \gamma(y_i/\sigma)]^{-1} = 0$
- $d/d\sigma [\log L(\gamma, \sigma | y_1, \dots, y_n)] = -u_n/\sigma + (1/\gamma) \sum_{i=1}^{u_n} [(y_i/\sigma) [1 + \gamma(y_i/\sigma)]^{-1-1/\sigma}] = 0$

Ces équations ne peuvent être résolues de manière explicite car il n'y a pas de solution analytique pour ces équations, il faut donc utiliser une méthode numérique telle que la méthode de Newton-

Raphson pour trouver les valeurs de γ et σ qui maximisent la fonction de vraisemblance. De plus, les recherches de Smith (1984), ont démontrés que les estimateurs du maximum de vraisemblance sont asymptotiquement gaussiens et efficaces pour $\gamma > -\frac{1}{2}$.

4.4.2 Estimation des paramètres de la GPD par la méthode des moments

La méthode des moments consiste à évaluer les moments empiriques de l'échantillon aux moments théoriques de la distribution GPD.

Soit X une variable aléatoire tel que, $X \sim GPD(u, \gamma, \sigma)$, alors les moments d'ordre un et deux existent si $\gamma < \frac{1}{2}$, et on a sous ces conditions :

$$\mathbb{E}(X) = \frac{\sigma}{1-\gamma} \text{ et } Var(X) = \frac{\sigma^2}{(1-\gamma)^2(1-2\gamma)}$$

On a ainsi que : $\gamma = \frac{1}{2} \left(1 - \frac{\mathbb{E}(X)^2}{Var(X)} \right)$ et $\sigma = \frac{\mathbb{E}(X)}{2} \left(1 + \frac{\mathbb{E}(X)^2}{Var(X)} \right)$

En remplaçant $\mathbb{E}(X)$ et $Var(X)$, par leurs estimateurs empiriques, ie :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } S_{n-1}^2(X) = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

On obtient ainsi les estimateurs des moments de γ et σ :

$$\widehat{\gamma}_n^M = \frac{1}{2} \left(1 - \frac{\bar{X}^2}{S_{n-1}^2(X)} \right) \text{ et } \widehat{\sigma}_n^M = \frac{\bar{X}}{2} \left(1 + \frac{\bar{X}^2}{S_{n-1}^2(X)} \right)$$

La méthode d'estimations des paramètres de la GPD par la méthode des moments repose sur un domaine de validité très restreint étant donné qu'ils sont asymptotiquement gaussiens que si $\gamma < \frac{1}{4}$.

4.4.3 Estimation des paramètres de la GPD par la méthode des moments pondérés

La méthode des moments pondérés est une version plus robuste de la méthode des moments traditionnelle lorsque les données contiennent des valeurs aberrantes ou des biais importants. Cette méthode vient remédier aux problèmes d'échantillons d'observations de petites tailles.

Soit X une variable aléatoire intégrable de fonction de répartition F . On appelle, le moment pondéré d'ordre $r \in \mathbb{N}$ et $s \in \mathbb{N}$ de X :

$$WM_X(r, s) = \mathbb{E} \left(X F^r(X) (1 - F(X))^s \right).$$

Pour $r = 0$, le théorème suivant nous donne la valeur du moment pondéré d'une loi $GPD_{\gamma, \sigma}$:

Soit X une variable aléatoire de loi de Pareto Généralisé de paramètres γ et σ , si $\gamma < 1$, on a que pour tout $s \in \mathbb{N}$:

$$WM_X(r, s) = \frac{\sigma}{(s+1)(s+1-\gamma)}$$

On peut ainsi écrire γ et σ en fonction de deux moments pondérés d'ordres s différents :

$$\gamma = \frac{4\mu_1 - \mu_0}{2\mu_1 - \mu_0} \text{ et } \sigma = \frac{2\mu_0\mu_1}{\mu_0 - 2\mu_1}$$

Cette méthode d'estimations des paramètres de la GPD reste toutefois très limitée au cas où l'indice des valeurs extrêmes $\gamma < 1$. En particulier, lorsque γ est très grand, la méthode peut conduire à des estimations biaisées et peu fiables des paramètres de la GPD.

D'autres méthodes dérivées de la méthode des moments existent :

- La méthode de la médiane de la GPD : qui consiste à estimer le seuil à partir de la médiane des données de dépassement de seuil, puis à estimer les paramètres de la GPD en utilisant la méthode des moments avec ce seuil.
- La méthode de l'inverse de la fréquence moyenne : qui consiste à estimer le seuil à partir de l'inverse de la fréquence moyenne des données de dépassement de seuil, puis à estimer les paramètres de la GPD en utilisant la méthode des moments avec ce seuil.

Ces méthodes restent néanmoins limitées dans leurs cadres d'utilisation.

4.4.4 Estimation des paramètres de la GPD par la méthode des moindres carrés

La méthode des moindres carrés (SSQ) consiste à minimiser la somme des carrés des écarts entre les valeurs observées et les valeurs théoriques de la distribution GPD. Mathématiquement on peut exprimer cela comme :

Soit X_1, X_2, \dots, X_n un échantillon de variables aléatoires i.i.d suivant une distribution GPD avec des paramètres inconnus σ et γ . Soit k le nombre d'excès choisi et $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ les n valeurs ordonnées de l'échantillon. On choisit une valeur seuil u telle que $X_{(n-k)} < u < X_{(n)}$. On définit $Y_i = X_i - u$ pour $i = n - k + 1, n - k + 2, \dots, n$. On utilise ensuite la méthode des moments pour estimer σ et μ , respectivement :

$$\begin{cases} \hat{\sigma} = \frac{1}{k} \sum_{i=n-k+1}^n Y_i \\ \hat{\gamma} = u + \frac{\sigma}{k} \sum_{i=n-k+1}^n \left(1 - \left(\frac{u - X_i}{\hat{\sigma}} \right)_+ \right) \end{cases}$$

où $(x)_+ = \max(x, 0)$ est la fonction partie positive.

Ensuite, on peut calculer l'estimateur des moindres carrés $\widehat{\gamma}^{SSQ}$ en résolvant l'équation suivante :

$$\sum_{i=n-k+1}^n \left(1 - \left(\frac{u - X_i}{\hat{\sigma}} \right)_+ \right) - k \left(1 - \frac{\widehat{\gamma}^{SSQ} u}{\hat{\sigma}} \right) = 0$$

Où $\widehat{\gamma}^{SSQ}$ est le paramètre de forme de la distribution GPD. Cette équation peut être résolue numériquement à l'aide d'une méthode d'optimisation.

Enfin, on peut estimer la probabilité d'excès de seuil $u + c$ pour une constante $c > 0$ en utilisant la formule :

$$\mathbb{P}(X > u + c) = \frac{\widehat{\gamma}^{SSQ} c}{1 - \frac{\widehat{\gamma}^{SSQ} u}{\hat{\sigma}}}$$

Avec $u, \hat{\sigma}$ et $\widehat{\gamma}^{SSQ}$, sont les paramètres estimés.

4.4.5 Estimation des paramètres de la GPD par l'Estimateur de Hill

L'estimateur de Hill est basé sur le rang des valeurs de la queue de distribution. Il consiste à ajuster une droite sur un graphique des logarithmes des rangs de la queue de distribution par rapport aux logarithmes des valeurs correspondantes. L'indice des valeurs extrêmes est alors estimé comme la pente de cette droite.

Plus précisément, l'estimateur de Hill consiste à calculer la pente de la droite qui minimise l'erreur quadratique moyenne des logarithmes des rangs des valeurs qui sont supérieures à un seuil donné. Le seuil est choisi de manière à ne pas inclure de valeurs extrêmes évidentes dans le calcul de l'estimateur.

Soient X_1, X_2, \dots, X_n des variables aléatoires indépendantes et de même fonction de répartition F appartenant au domaine d'attraction de Fréchet avec un indice des valeurs extrêmes $\gamma > 0$. Soient $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ la statistique d'ordre obtenue à partir de l'échantillon d'observation (X_1, X_2, \dots, X_n) , alors l'estimateur de Hill de γ est défini par :

$$\widehat{\gamma^H}(k_n) = \frac{1}{k_n} \sum_{i=1}^n \ln(X_{(n-i+1)}) - \ln(X_{(n-k_n)})$$

où $1 \leq k_n \leq n$, est une valeur à choisir par l'utilisateur.

Au vu de l'expression de l'estimateur de Hill, le choix de k_n est très important car :

- Si k_n est trop petit alors nous aurons très peu d'observations pour calculer l'estimateur de Hill et donc une variance importante
- Si k_n est trop grand, nous aurons une mauvaise approximation par une loi de Pareto et par conséquent un biais important de l'estimateur de Hill.

Par exemple, traçons le Hill plot pour une loi de Fréchet de paramètre $\gamma = \frac{2}{3}$ et de taille $n = 500$, et étudions le comportement de l'estimateur de Hill en fonction de k_n :



Figure 5 : Hill plot pour la loi de Fréchet

On observe que le meilleur choix de k_n se situe dans la zone dite stable (entre 250 et 800) car les oscillations de l'estimateur de Hill sont très grandes lorsque k_n est entre 0 et 250, puis il devient stable entre 250 et 800 pour enfin pour $k_n > 800$, l'estimateur de Hill subit des déviations qui sont représentatives du biais.

En revanche si k_n est choisi de manière à avoir $k_n \xrightarrow{\infty} \infty$ et $\frac{k_n}{n} \xrightarrow{\infty} 0$, on peut montrer grâce au TCL que l'estimateur de Hill est asymptotiquement gaussien et qu'on a :

$$\sqrt{k_n} \left(\frac{\gamma_n^H(k_n) - \gamma}{\gamma} \right) \xrightarrow{\mathcal{L}} N(0,1)$$

L'une des principales limites de cet estimateur est qu'il s'applique uniquement à données d'échantillons qui sont dans le domaine d'attraction de Fréchet, néanmoins l'estimateur de Hill reste très largement utilisé dans la pratique de la modélisation de valeurs extrêmes en réassurance.

4.4.6 Estimation des paramètres de la GPD par l'Estimateur de Dekkers, Einmahl et De Hann

L'estimateur de Dekkers, Einmahl et de Hann (DEdH) est un estimateur non biaisé pour le paramètre de la GPD. Il est considéré comme un meilleur choix que l'estimateur de Hill dans les cas où le domaine d'attraction de la GPD n'est pas bien défini ou lorsque la queue de la distribution est très lourde.

L'estimateur DEdH est basé sur l'ajustement d'une droite à une transformation logarithmique des excès de seuil, avec des poids attribués en fonction de la distance entre les observations et le seuil. Il peut être obtenu en résolvant numériquement une équation qui dépend de l'estimateur de Hill et des quantiles de la distribution des excès de seuil. Il utilise une approche de maximisation de la vraisemblance pour estimer l'indice.

La méthode consiste à diviser les données en r blocs de taille m et à estimer les paramètres de la GPD pour chaque bloc en utilisant la méthode des moments pondérés. Ensuite, les k estimations des paramètres sont utilisées pour estimer l'indice des valeurs extrêmes en maximisant la fonction de vraisemblance.

L'estimateur est donné par la formule suivante :

$$\gamma^{\widehat{DEdH}}(k_n) = M_{k_n}^{(1)} + 1 - \frac{1}{2} \left(1 - \frac{(M_{k_n}^{(1)})^2}{M_{k_n}^{(2)}} \right)^{-1}$$

Où $M_{k_n}^{(r)} = \frac{1}{k_n} \sum_i^{k_n} (\ln(X_{(n-i+1)}) - \ln(X_{(n-k_n)}))^r$ et $M_{k_n}^{(1)}$ correspond à l'estimateur de Hill.

L'estimateur de Dekkers, Einmahl et de Hann est aussi asymptotiquement gaussien, et on a :

$$\sqrt{k_n} \left(\frac{\gamma_n^{\widehat{DEdH}}(k_n) - \gamma}{\sigma(\gamma)} \right) \xrightarrow{\mathcal{L}} N(0,1)$$

L'estimateur DEdH est donc une généralisation de l'estimateur de Hill valable pour tous les domaines d'attractions, cependant, il peut être computationnellement coûteux et nécessite une grande quantité de données pour être précis. En outre, il peut être sensible aux choix des paramètres et des seuils, ce qui peut affecter la stabilité et la précision de l'estimation.

4.4.7 Estimation des paramètres de la GPD par l'Estimateur de Pickands

Tout comme l'estimateur de Dekkers, Einmahl et de Hann, l'estimateur de Pickands est valable pour les trois domaines d'attraction.

L'estimateur de Pickands consiste à ajuster une droite sur un plot des quantiles de second ordre en fonction de la valeur de la fraction de données excédentaires au-dessus du seuil. La pente de cette droite est utilisée comme estimateur du paramètre de forme de la distribution GPD et on le définit à partir d'un certain nombre d'excès k_n par :

$$\widehat{\gamma^P}(k_n) = \frac{1}{\ln(2)} \ln \left(\frac{X_{(n-k_n)} - X_{(n-2k_n)}}{X_{(n-2k_n)} - X_{(n-4k_n)}} \right)$$

Comme les autres estimateurs, on peut montrer qu'il est asymptotiquement gaussien :

$$\sqrt{k_n} \left(\frac{\widehat{\gamma^P}(k_n) - \gamma}{\sigma(\gamma)} \right) \xrightarrow{\mathcal{L}} N(0,1)$$

Cet estimateur est considéré comme plus efficace que l'estimateur de Hill dans les cas où le domaine d'attraction est bien défini, en revanche il est également biaisé pour les petites valeurs de n .

5 Choix du seuil des extrêmes

Dans les parties précédentes, nous avons vu que l'estimation de la distribution de la Pareto Généralisée dépend très fortement du choix du seuil car cela affecte directement la précision et la fiabilité des résultats obtenus

En effet, le seuil de déclenchement est le niveau minimum de perte au-delà duquel les sinistres sont considérés comme des sinistres extrêmes. Il doit être choisi avec soin, car un seuil trop élevé peut exclure des observations importantes de la modélisation, tandis qu'un seuil trop bas peut inclure des événements moins extrêmes, ce qui peut conduire à une estimation incorrecte de la distribution des valeurs extrêmes.

Si le seuil est trop bas, il est possible que la modélisation inclût des sinistres qui ne sont pas vraiment extrêmes, ce qui peut conduire à une surestimation de la fréquence et de la gravité des sinistres extrêmes. Cela peut entraîner des conséquences négatives sur la tarification et la gestion des risques, car les compagnies d'assurances et de réassurance peuvent prendre des décisions fondées sur des informations inexactes.

Si le seuil est trop élevé, la modélisation peut exclure des sinistres réels qui sont considérés comme des sinistres extrêmes. Cela peut conduire à une sous-estimation de la fréquence et de la gravité des sinistres extrêmes, ce qui peut également entraîner des conséquences négatives sur la tarification et la gestion des risques.

D'autre part, si le seuil est choisi de manière inappropriée, cela peut entraîner des changements importants dans les estimations des paramètres de la distribution lorsque de nouvelles données sont ajoutées.

En conclusion, le choix du seuil de déclenchement est crucial pour obtenir des résultats précis et fiables dans le contexte de notre modélisation car s'il est choisi, cela peut conduire à des ajustements de la distribution qui ne correspondent pas bien aux données. Nous allons voir différentes méthodes analytiques et graphiques afin de sélectionner un seuil approprié à notre étude.

5.1 Détermination du seuil des extrêmes par l'estimateur de Hill ou de Pickands

Pour ces deux méthodes d'estimation du seuil des extrêmes on va considérer (X_1, X_2, \dots, X_n) notre échantillon de variables i.i.d et $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ notre statistique d'ordre, ordonnée de manière décroissante.

Le Hill plot est un graphique qui représente le logarithme des excès triés par ordre décroissant en fonction de leur rang. Cette méthode repose sur le principe que la GPD conserve sa loi par conditionnement, en effet on a la propriété suivante :

$$\mathcal{L}(X) \sim GPD(\gamma, \sigma), \text{ alors } \mathcal{L}(X|X > u) \sim GPD(\gamma, \sigma').$$

Et donc, lorsque la GPD est tronquée à gauche, elle conserve son paramètre de forme.

Ainsi, au-delà d'un certain seuil, l'estimation du paramètre de forme est plus ou moins du même ordre de grandeur pour des points de troncature différents, néanmoins il est important de trouver le nombre d'excès le plus faible afin de trouver un meilleur estimateur du seuil des extrêmes.

On va donc tracer l'estimateur de Hill qui a été défini au chapitre (9.4.4) en fonction du nombre d'excès k_n par :

$\forall \gamma > 0$

$$\left\{ k_n, \widehat{\gamma^H(k_n)} = \frac{1}{k_n} \sum_{i=1}^n \ln(X_{(n-i+1)}) - \ln(X_{(n-k_n)}) \right\}$$

Le seuil est déterminé visuellement en identifiant la courbe des excès qui est approximativement linéaire.

Voici un exemple d'un Hill plot d'une loi de Pareto :

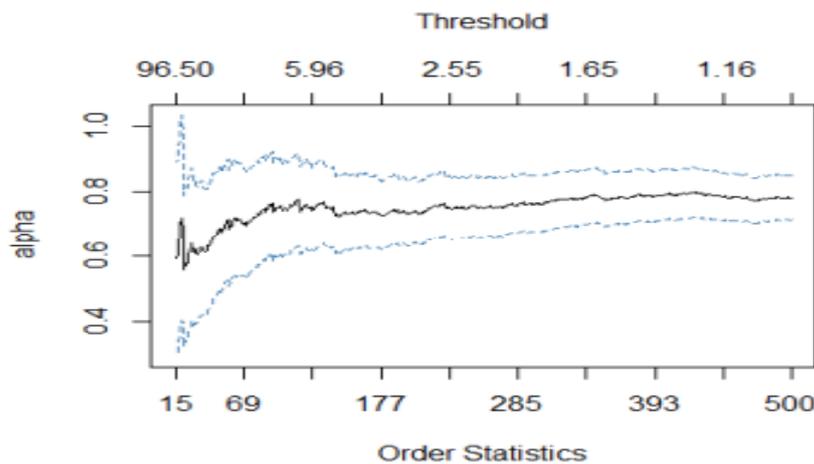


Figure 6 : Hill plot pour la loi de Pareto

Dans cet exemple, la stabilité se fait entre la statistique d'ordre 120 et 177.

Cette méthode graphique pour déterminer le seuil des extrêmes est très populaire mais reste néanmoins très restrictive dû au fait qu'elle est valable uniquement pour des données appartenant au domaine de Fréchet.

De manière analogue, on peut tracer le Pickands plot en fonction du nombre d'excès k_n :

$$\left\{ k_n, \quad \gamma^P(\widehat{k_n}) = \frac{1}{\ln(2)} \ln \left(\frac{X_{(n-k_n)} - X_{(n-2k_n)}}{X_{(n-2k_n)} - X_{(n-4k_n)}} \right) \right\}$$

Cette méthode présente l'avantage d'être valable dans les trois domaines d'attractions.

Comparaison de l'estimateur de Hill contre l'estimateur de Pickands pour des simulations de lois de *Pareto*(2 ; 5) et *GPD*(1.5; 1) :

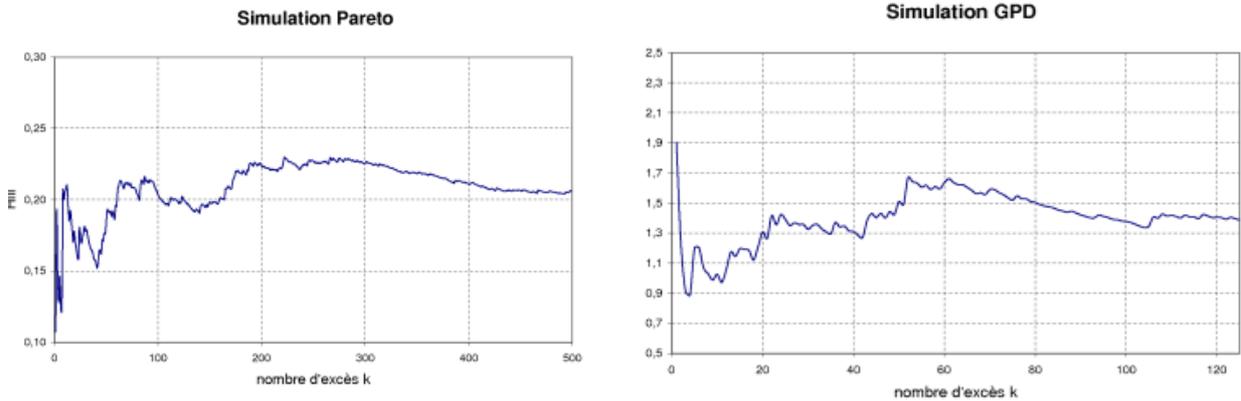


Figure 7: Estimateur de Hill pour des lois de Pareto et GPD

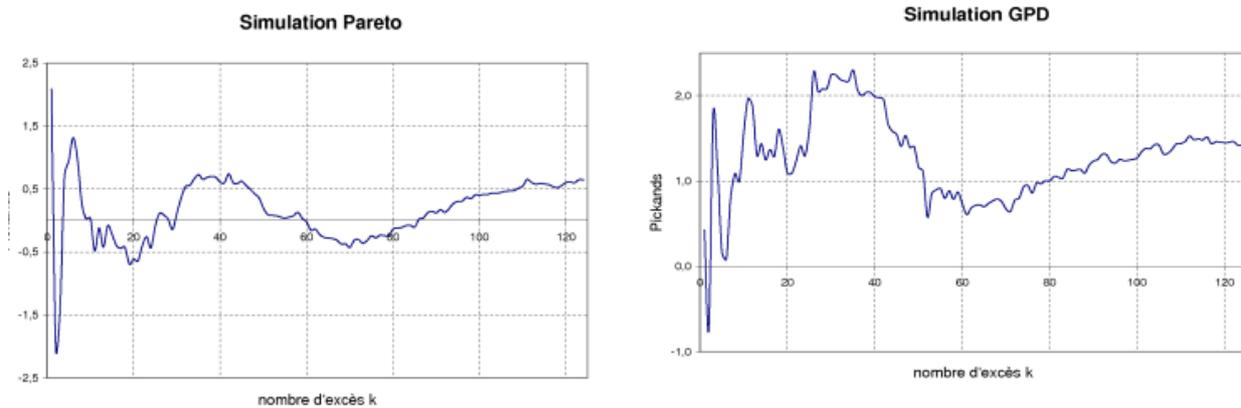


Figure 8 : Estimateur de Pickands pour des lois de Pareto et GPD

L'estimateur de Hill donne des valeurs proches des paramètres de la loi tandis que l'estimateur de Pickands présente de moins bons résultats que l'estimateur de Hill lorsque ce dernier est valable.

En effet l'estimateur de Hill fait intervenir la moyenne des logarithmes des observations, provoquant ainsi des résultats plus lissés et moins sensibles aux variations d'une observation à l'autre ce qui peut expliquer ce résultat.

5.2 Détermination du seuil des extrêmes par la fonction des excès moyens

La méthode consiste à calculer la moyenne des excès X_i au-dessus d'un seuil u pour différentes valeurs de u . On définit la fonction des excès moyens $e(u)$ comme étant la moyenne des excès au-dessus du seuil u :

$$e(u) = \mathbb{E}(X - u | X > u), \text{ avec } 0 \leq u \leq x_F \text{ avec } x_F \text{ le point terminal de } F.$$

Pour $X \sim GPD(\gamma, \sigma)$, la fonction des excès moyens s'écrit comme étant :

$$e(u) = \frac{\sigma}{1 - \gamma} + \frac{\gamma}{1 - \gamma} u$$

Cette fonction est donc linéaire en u ainsi en faisant varier le seuil u on peut utiliser le comportement de la fonction des excès moyens car si l'approche de la GPD est valide pour un premier seuil u_0 alors elle l'est pour $u > u_0$ par la propriété de stabilité de la GPD. En traçant $e(u)$ en fonction u il est donc assez simple de déterminer le seuil des extrêmes.

En pratique la fonction des excès moyens $e(u)$ est souvent approchée sur la base de l'approximation empirique $\widehat{e(u)}$ suivante :

$$\widehat{e(u)} = \frac{1}{\#\{X_i > u \mid i = 1 \dots n\}} \sum_{i=1}^n (X_i - u) 1_{X_i > u}$$

Dans le domaine d'attraction de Weibull, la fonction est donc décroissante, dans celui de Fréchet elle est croissante et dans celui de Gumbel, elle est constante, en traçant 3 lois appartenant à ces trois domaines d'attraction respectifs on a bien que :

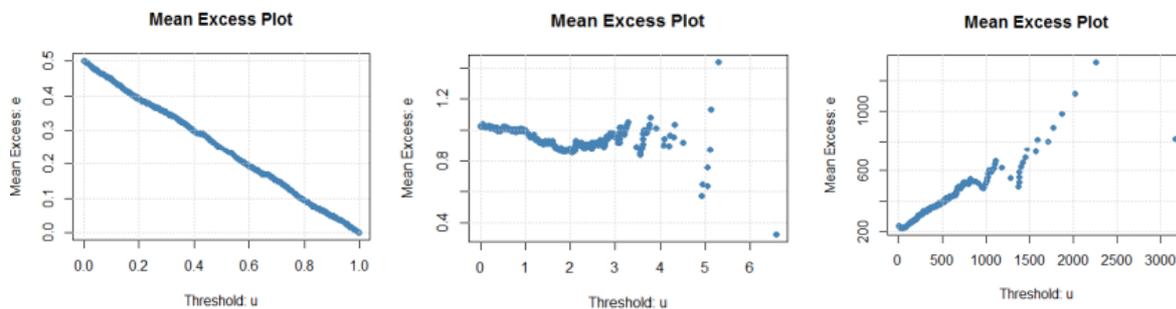


Figure 9 : Graphique de la fonction des excès moyens pour des loi respectives : uniforme, exponentielle et de Pareto

Si la distribution des excès suit bien une loi GPD, alors la fonction des excès moyens doit se comporter linéairement au-dessus d'un certain seuil.

Le seuil de stabilité est déterminé graphiquement en traçant la fonction des excès moyens en échelle logarithmique en abscisse et en ordonnée. On observe alors une zone linéaire où la fonction des excès moyens est approximativement droite. Le seuil de stabilité est alors choisi comme étant la plus petite valeur de γ à partir de laquelle cette zone linéaire est observée. Mais ce graphe peut aussi nous donner une indication sur l'appartenance du domaine d'attraction de notre échantillon.

5.3 Détermination du seuil des extrêmes par le Gerstengarbe plot

Cette méthode de détermination du seuil repose sur la détermination du point de départ de la région des extrêmes. L'idée étant que lors de l'entrée en zone des observations extrêmes, on s'attend à une variation du comportement de la série décrit ci-dessous :

Dans cette méthode, on va considérer x_1, x_2, \dots, x_n notre échantillon d'observation et $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ notre échantillon d'observation ordonnée de manière croissante et on va considérer la valeur suivante :

$$\Delta_i = x_{(i)} - x_{(i-1)}, i = 2 \dots n$$

En utilisant les travaux de la version séquentielle du test de Mann-Kendall on va chercher à identifier un changement dans la série et on va ainsi tracer :

$$\left\{ U_i = \frac{U_i^* - \frac{i(i-1)}{4}}{\sqrt{\frac{i(i-1)(i+5)}{72}}}, \quad i = 1, 2, \dots, n-1 \right\}$$

Avec $U_i^* = \sum_{k=2}^i n_k$ ou n_k est le nombre de valeur $\Delta_2, \dots, \Delta_k$ inférieur à Δ_k .

En répétant le même procédé mais cette fois-ci, en considérant notre échantillon $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ notre échantillon d'observation ordonnée de manière décroissante, on peut tracer une deuxième série \bar{U}_i . En traçant ces deux séries, on devrait observer un point d'intersection, déterminant ainsi le seuil des extrêmes.

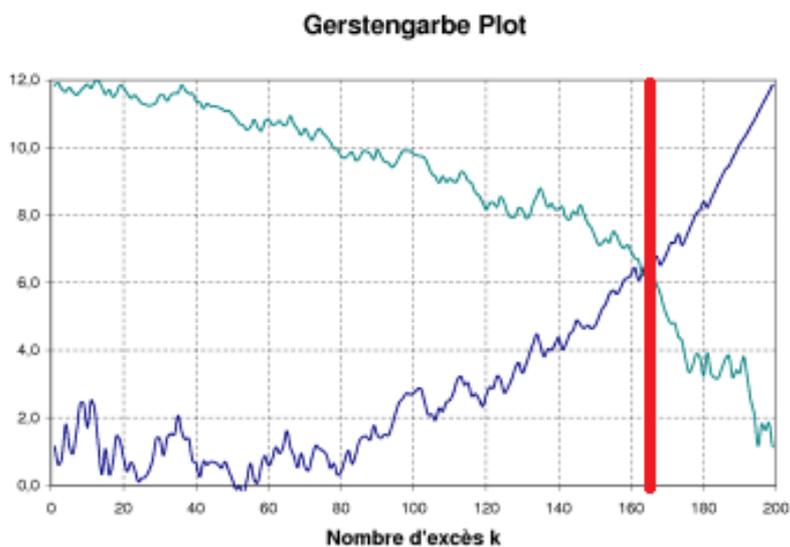


Figure 10 : Graphique de Gerstengarbe pour une loi de Pareto

En simulant $X \sim \text{Pareto}(10; 50)$ et en appliquant la méthode du Gerstengarbe Plot, on s'aperçoit que le point d'intersection des courbes se fait environ à la 165 statistique d'ordre, ce qui correspond à un seuil des extrêmes à 20,7. En traçant la fonction de répartition empirique on a :

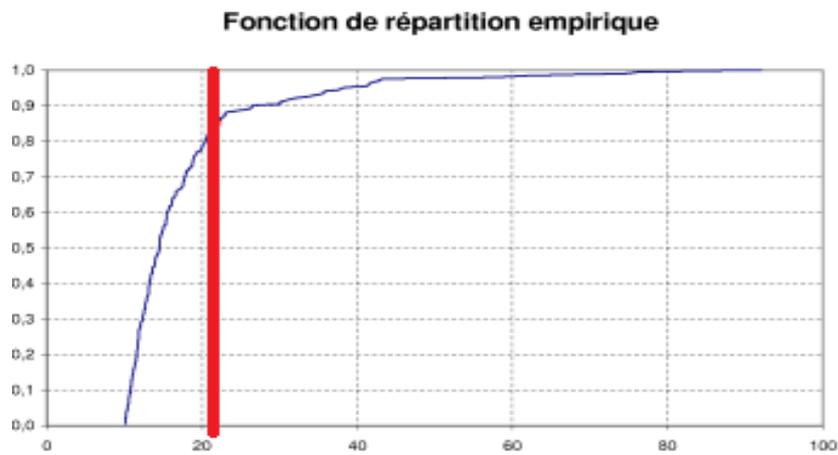


Figure 11 : Fonction de répartition de l'échantillon

On observe que ce seuil correspond identiquement à un point de cassure de la fonction de répartition ce qui nous indique le seuil d'entrée dans la zone des valeurs extrêmes.

Maintenant que nous avons présenté les différentes méthodes et outils nous permettant de construire notre modélisation, nous allons l'appliquer à notre portefeuille afin de répondre à notre problématique initiale.

6 Application de la méthode sur notre portefeuille

Dans les parties précédentes, nous avons déterminé les différentes méthodes qui nous permettent de répondre à notre enjeu initial, qui est l'amélioration de la cotation des tranches peu travaillantes de notre traité de réassurance en excédent de sinistre pour la branche RC AUTO. Nous rappelons que le traité que nous choisissons d'expérimenter dans l'objectif d'une possible mise en application pour BPCE Assurances IARD est un traité en excédent de sinistre par risque qui comporte les tranches suivantes :

- Tranche 1 : 3 M€ XS 6 M€.
- Tranche 2 : 3 M€ XS 9 M€.
- Tranche 3 : illimité XS 12 M€.

Dans notre étude, la méthode retenue pour la modélisation de la charge annuelle des graves consiste à modéliser séparément la sévérité du nombre de sinistres, puis en déduire la charge annuelle par simulation de Monte-Carlo à partir des lois de sévérités et de fréquences obtenues en dessous et au-dessous du seuil des extrêmes.

Dans ce mémoire la fréquence ne faisant pas l'objet d'une étude spécifique, la méthodologie de BPCE Assurances IARD sera conservée. Nous allons donc nous focaliser sur la modélisation de la sévérité. L'objectif de ce mémoire est donc d'améliorer la modélisation de cette sévérité et d'estimer ce coût moyen pour une tranche de réassurance donnée.

Comme nous l'avons vu dans l'approche probabiliste, ce coût moyen pour une tranche de réassurance s'exprime à l'aide de la fonction de survie de la variable aléatoire qui modélise les sinistres, ainsi nous allons estimer la distribution qui engendre nos données et qui permet de rendre compte du risque même dans les tranches hautes et peu travaillantes. Cette estimation est rendue difficile lorsque la tranche est peu travaillante car il faut être capable d'estimer un coût moyen dans des grandeurs qui appartiennent à la queue de distribution des sinistres. Parmi les méthodes présentées dans les parties en amont, nous avons opté pour la méthode basée sur la Théorie des valeurs extrêmes et sur le fondement du dépassement de seuil (Peak Over Threshold). La théorie des valeurs extrêmes et en particulier la méthode du Peak Over the Threshold nous offre un cadre solide pour cette étude. Cette méthode nous permettra ainsi d'alourdir la queue de distribution de nos données en la complétant par une loi de Pareto Généralisée.

Ainsi à l'aide de la méthode du Peak Over the Threshold, nous allons proposer une modélisation à deux lois qui permet d'accorder une attention particulière aux sinistres extrêmes sans pour autant uniquement focaliser notre modélisation sur ces sinistres.

Pour ce faire, commençons par une analyse des données, afin de comprendre la distribution des sinistres et l'impact des seuils de détection sur la modélisation des valeurs extrêmes. Ensuite, nous déterminerons le seuil des extrêmes à l'aide des différents outils à notre disposition et nous utiliserons la méthode du dépassement de seuil pour estimer les paramètres de la loi de Pareto généralisée et ainsi modéliser la distribution des sinistres extrêmes afin de déterminer le coût de chaque tranche de notre traité de réassurance. Cette modélisation de notre charge de réassurance aboutira par la suite à une analyse de l'impact des différents modèles sur le coût de cession.

Un point d'attention est porté sur l'étude du seuil, représentant un point crucial de notre étude. En effet, le seuil est utilisé pour déterminer quelles observations sont considérées comme des valeurs extrêmes et seront utilisées pour estimer les paramètres de la loi de Pareto généralisée. Le choix du seuil est donc essentiel pour garantir une estimation fiable des paramètres et une modélisation précise de la distribution des valeurs extrêmes. Cependant, le choix du seuil n'est pas trivial et peut varier

selon les données et le contexte. Un seuil trop bas peut entraîner une surestimation de la fréquence des sinistres extrêmes et une sous-estimation de leur gravité, tandis qu'un seuil trop élevé peut limiter le nombre d'observations extrêmes et compromettre la précision de l'estimation des paramètres. Cette étape aboutira sur une analyse de la modélisation effectuée sur notre portefeuille.

6.1 Etude de notre échantillon d'observation de base

A partir des données fournies par BPCE Assurances IARD, nous allons segmenter notre échantillon d'observations des sinistres graves reflétant la sinistralité de la compagnie. Pour ce faire nous définissons deux types de sinistres parmi les graves en fonction de leur sévérité :

- Pour commencer, la sinistralité grave est définie chez BPCE Assurances IARD pour tous sinistres dont le montant de la charge a dépassé le seuil des 250 k€. Parmi ces sinistres qualifiés de graves, on va alors distinguer deux catégories :
 - Les sinistres « sinistres 250K » qui correspondent aux sinistres dont le montant de la charge est compris entre 250 k€ et le seuil que l'on déterminera dans la suite de notre étude.
 - Les « sinistres extrêmes » qui correspondent aux sinistres dont le montant de la charge est supérieur au seuil que l'on déterminera dans la suite de notre étude.

Sur notre échantillon initial de 1 700 observations sur un horizon d'observations 2005 – 2022, on effectue un premier filtre qui nous permet de conserver uniquement les sinistres supérieurs à 250 000 € et qui sont pertinents pour notre modélisation, à savoir les sinistres graves. En utilisant un seuil d'attrition approprié, on peut s'assurer que les sinistres sélectionnés pour la modélisation sont tous des sinistres graves ou extrêmes. En excluant ainsi les sinistres dont le montant est inférieur à 250 k€, notre échantillon est alors réduit à 1 159 observations.

Ainsi sur notre nouvel échantillon d'observation, on observe ainsi un sinistre minimum à 250 351 €, et un sinistre maximum à 20 120 105€. La moyenne et l'écart-type de notre échantillon sont respectivement de 377 400 € et 1 196 357 €.

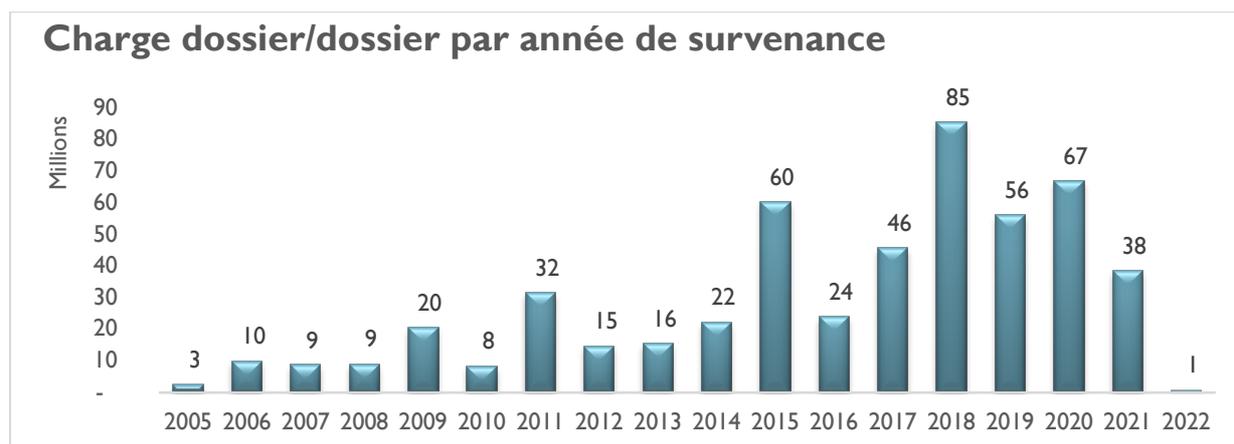


Figure 12 : Répartition de la charge des sinistres de notre portefeuille par année de survenance

On remarque que la charge dossier/dossier n'est pas homogène en fonction des années de survenances. En effet la charge dossier/dossier n'est pas représentative des conditions de souscriptions économiques, de l'inflation ou de l'évolution de la taille du portefeuille. Une étape clé est donc de traiter nos données.

6.2 Retraitement de nos données sinistres

Une étape fondamentale dans le cadre de notre étude consiste à redresser un échantillon de statistiques historiques afin de le rendre représentatif des conditions de souscriptions économiques de l'année de couverture à évaluer mais aussi de prendre en compte le redressement des rentes à 1,5%, c'est-à-dire de prendre en compte l'inflation. Ce redressement est appelé la mise en « AS-IF ».

Ce redressement doit prendre en compte les changements éventuels de politique de souscription, de sélection des risques, de composition du portefeuille, les changements éventuels de politique de provisionnement des sinistres (provisions pour sinistres à payer) pour les branches à développement long. Les statistiques redressées sont les sinistres historiques, les assiettes de primes historiques et les sommes assurées.

6.2.1 Indexation

L'indexation consiste à mettre en as-if N+1 (N étant l'année de la modélisation et N+1 l'année de survenance à modéliser) la charge de sinistres (règlements et suspens). Ceci afin de rendre les sinistres homogènes dans la modélisation et de se mettre en situation de l'année N+1.

Étant donné la longue durée de liquidation (> 10 ans pour les sinistres corporels), des sinistres déclarés peuvent ne pas être totalement réglés dans un laps de temps restreint. Ces sinistres subissent l'évolution normale des principaux coûts (hospitalisations, assistance tierce personne...). Pendant la période de liquidation, il peut y avoir plusieurs chocs (réglementaires, jurisprudentiels) entraînant une modification de l'inflation.

Le coût des sinistres individuels est revalorisé en utilisant un indice de revalorisation décrivant l'inflation économique sur les capitaux assurés (ainsi que d'autres éléments comme la sur-inflation dans le cas des branches couvrant les dommages corporels).

Il est à noter que le choix de l'indice permettant le redressement de la charge (règlements et suspens) de l'inflation est primordial dans l'analyse de provisionnement. Une hausse de l'inflation engendre un sous provisionnement et une baisse de l'inflation entraîne une surestimation des provisions.

Deux indices sont utilisés :

- Indice FFA (reconstitué avec l'indice des salaires de base par activité - Activités pour la santé humaine) pour les dommages corporels (RC Corporelle, RCVP Corporelle, GC et GAV).
- Indice FFB pour les dommages matériels (RCVP Matérielle et RC Propriétaire ou Usager).

Une fois le calcul des indices FFA, et FFB fait, nous devons effectuer une projection de ces indices.

Le choix de la méthode de projection dépend des données disponibles et des hypothèses sur les tendances futures. Les méthodes courantes comprennent l'utilisation de tendances historiques, l'analyse de tendances économiques, ou des modèles de prévision statistiques. Nous avons opté pour une projection dépendante de l'historique en fonction d'une tendance.

Dans notre étude, on choisira une tendance historique 2011-2021, ce choix a été effectué car il est plus représentatif de l'évolution de l'inflation et des coûts des réclamations que celui de la période complète de sinistralité.

Par exemple, pour l'indice FFA, sur les 10 dernières années, l'évolution de cet indice est d'environ 2,5%, on se servira donc de cette variation moyenne afin de projeter dans les années futures l'indice FFA.

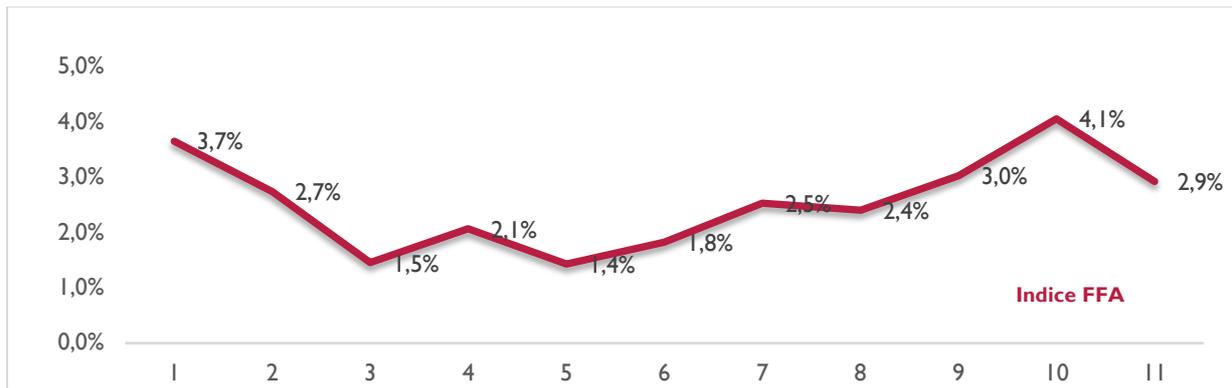


Tableau 21 : Evolution de l'indice FFA sur les 10 dernières années

Calcul de la charge indexée :

Pour un sinistre s de survenance i , la charge indexée vue en j se calcule de la façon suivante :

$$\text{Règlements décumulés indexés : } R_{i,j,s}^* = R_{i,j,s} \times \frac{\text{Indice}_{N+j}}{\text{Indice}_{i+j-1}}, \text{ pour } j = 1, \dots, N - i + 1$$

où $R_{i,j,s}$ sont les règlements décumulés du sinistre s survenu en i pour l'année de développement j

$$\text{Suspens indexés : } S_{i,j,s}^* = S_{i,j,s} \times \frac{\text{Indice}_{N+j}}{\text{Indice}_{i+j-1}}, \text{ pour } j = 1, \dots, N - i + 1$$

où $S_{i,j,s}$ sont les règlements décumulés du sinistre s survenu en i pour l'année de développement j

$$\text{Charge indexée : } C_{i,j,s} = \left(\sum_{k=0}^j R_{i,k,s}^* \right) + S_{i,j,s}^*, \text{ pour } j = 1, \dots, N - i + 1.$$

Dans l'exemple ci-dessous, un sinistre survenu en 2014, avec des règlements effectués 2 ans et 3 ans après la survenance, est indexé comme s'il survenait en 2020 :

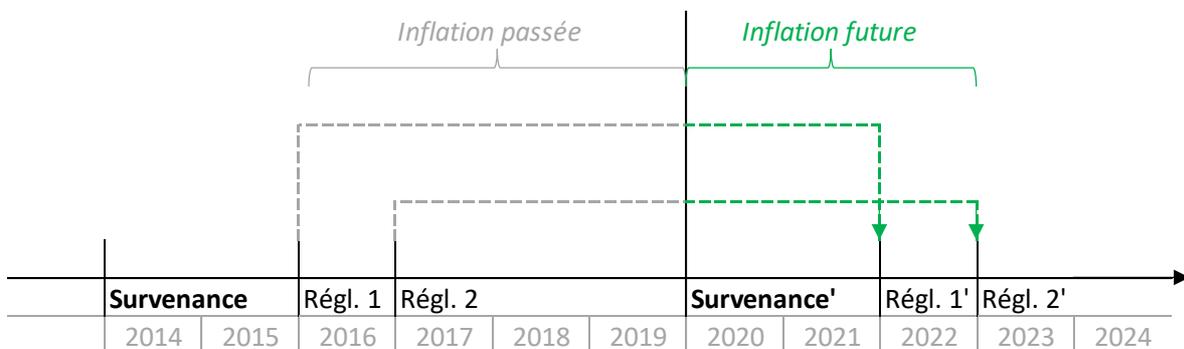


Figure 13 : Exemple d'indexation de la charge d'un sinistre

6.2.2 Redressement des primes

À cette dernière étape, s'ajoute le redressement des primes, car l'historique des primes doit prendre en compte les modifications tarifaires de base de l'assureur. Le redressement des primes permet de recalculer les assiettes de primes qui auraient été émises (et acquises) sur les années d'expérience si les conditions économiques et tarifaires avaient été celles de l'année de couverture. Cette étape permet d'obtenir une taille de portefeuille équivalente à l'année de modélisation.

Pour chaque année d'expérience passée k , nous disposons du nombre de polices souscrites n_{b_k} , et de la prime moyenne appliquée par police par l'assureur $P_{police,k}$, et des montants de primes historiques P_k . Les primes sont redressées de la manière suivante :

$$P_{AS-IF,k} = P_{police, \text{année de cotation}} \times n_{b_k} = P_{\text{année de cotation}} \times \frac{n_{b_k}}{n_{b_{\text{année de cotation}}}}$$

6.2.3 Mise en « AS-IF » de nos données

En appliquant la méthode précédemment décrite à notre échantillon d'observations, on obtient une nouvelle variable représentant la charge des sinistres après indexation.

Sur nos données, on observe maintenant, une charge minimum de sinistre indexée à 352 711 € et un maximum à 20 878 124 €. La moyenne et l'écart-type de notre échantillon sont respectivement de 511 442 € et 1 159 129 €.

En analysant notre échantillon, on remarque que le sinistre maximum n'a pas grandement évolué contrairement à certains sinistres. On sait que plus le sinistre est ancien, plus le niveau d'inflation entre l'année de survenance et l'année actuelle a des risques d'être important. Ce sinistre s'est déclaré récemment, ce qui peut expliquer ce faible taux de variation entre la charge brute et la charge indexée.

Pour illustrer cet écart, nous avons dressé un tableau reflétant de l'évolution par année de survenance des sinistres après indexation.

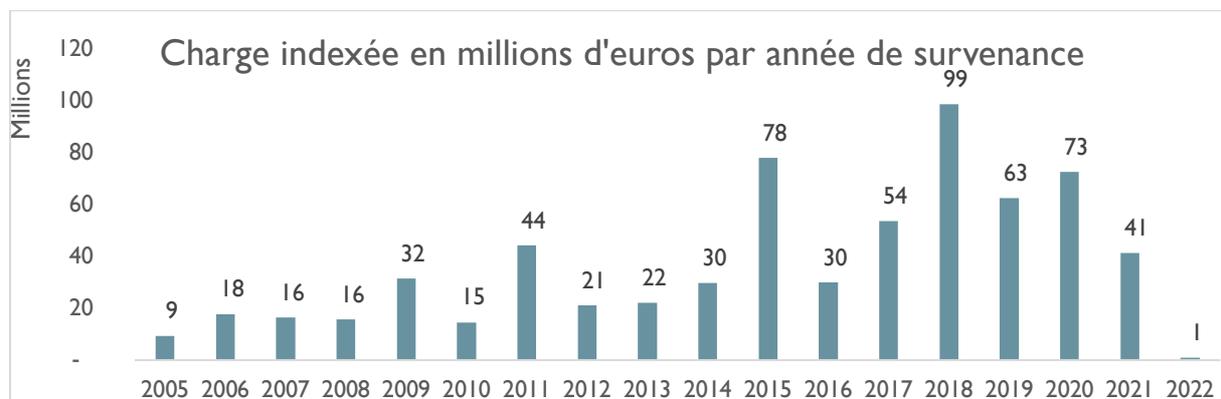


Figure 14 : Mise en AS-IF de nos données portefeuille par année de survenance

Nous remarquons ainsi, que plus les années de survenances des sinistres sont anciennes, plus l'évolution de la charge indexée est élevée pour enfin se rapprocher de l'évolution moyenne des indices utilisés pour la mise en as-fi de nos sinistres.

Cette indexation a aussi un impact sur notre type de sinistralité. En effet cette évolution des coûts de sinistralité fait basculer des sinistres qui ont été classifiés comme inférieur à notre seuil attritionnel (<250K€) à des sinistres maintenant classifiés comme étant graves. Cela a aussi un impact sur des sinistres qui sont catégorisés comme étant des sinistres étant inférieurs au seuil des extrêmes qui passeront alors ce seuil, pour ainsi devenir des sinistres extrêmes.

6.2.4 Projection à l'ultime

En s'appuyant sur la méthode de Chain-Ladder décrite en 2.3, nous allons estimer la charge ultime des sinistres individuels en calculant les facteurs de développements.

En effet, pour calculer les IBNER et prendre en compte les évolutions potentielles de sinistres, nous allons projeter à l'ultime la charge des sinistres individuels en calculant les facteurs de développements isopérimètres. Ces triangles permettent de modéliser uniquement l'effet de vieillissement pour des sinistres en j et $j+1$. Dans cette étape il est important d'exclure certains link ratios atypiques afin de ne pas avoir d'influence excessive car certains link ratios peuvent exercer une influence excessive sur les résultats du modèle. Cela peut se produire si un seul link ratio a une valeur extrêmement élevée ou basse, qui est très différente des autres link ratios. Cela peut biaiser les résultats du modèle et conduire à une mauvaise estimation des coûts futurs. L'exclusion de certains link ratios peut être effectuées car ils peuvent ne pas être représentatifs de la tendance globale des données, en raison d'une fluctuation aléatoire des sinistres. Dans ce cas, il peut être justifié de les exclure pour obtenir une estimation plus représentative des coûts futurs.

Après application de cette méthode, nous validons les hypothèses du Chain-Ladder via la méthode des coefficients de détermination R^2 des vecteurs de la charge des sinistres indexés :

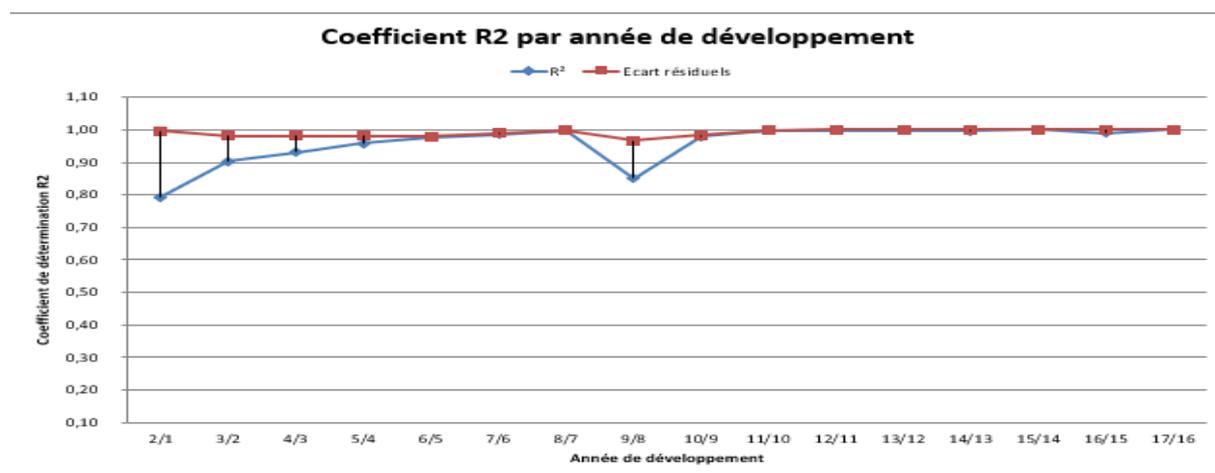


Figure 15: Coefficient R^2 par année de développement

Le coefficient de détermination R^2 prenant en compte les exclusions des link ratios est proche de 1 et valide par conséquent les hypothèses de Chain-Ladder. Cela indique une bonne adéquation du modèle de régression linéaire aux données. Ainsi la charge ultime de chaque sinistre est calculée en appliquant à la charge indexée le facteur de développement cumulé correspondant et le redressement de la taille du portefeuille, nous obtenons :

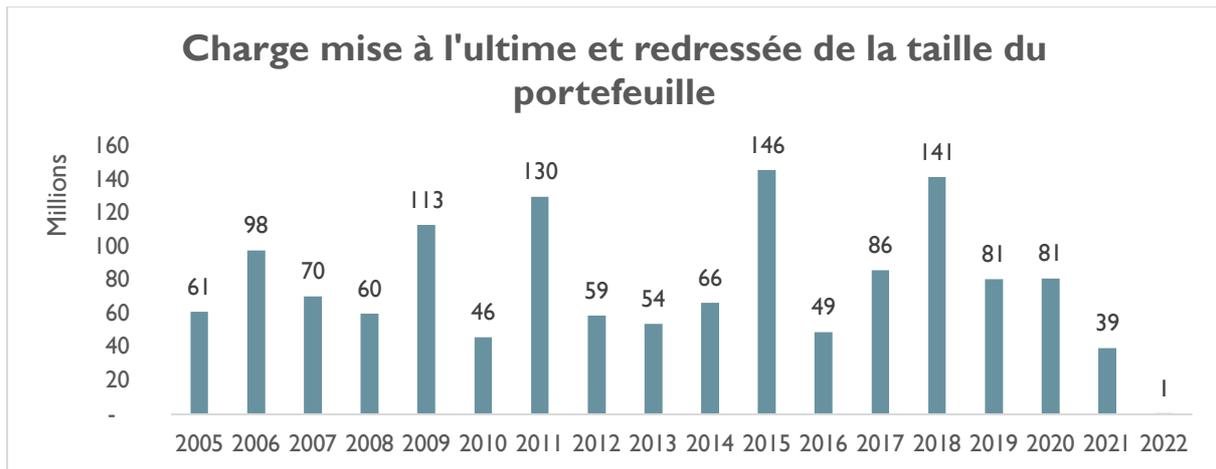


Figure 16 : Mise à l'ultime de la charge as-if, redressée de l'exposition par année de survenance

Nous observons maintenant que par année de survenance, le coût des sinistres est relativement homogène, hormis pour certaines années qui semblent plus sinistrées que d'autres.

6.2.5 Traitement des postes capitalisés

Pour les sinistres de responsabilité civile corporelle automobile, nous retraits les rentes potentielles en mettant à jour les paramètres d'actualisation :

- Postes avec règlement prévisible sous forme de rentes : mise à jour du capital constitutif avec la table TD 88/90 et un taux d'actualisation $i\%$ correspondant au taux d'actualisation convenu dans le traité de réassurance.
- Postes avec règlement en capital : actualisation avec la Gazette du Palais 2018

Notre étude ne se concentrant pas sur cette partie, nous utiliserons un outil interne à BPCI Assurances IARD nous permettant de calculer les coefficients d'ajustement pour mettre à jour le montant des sinistres avec la charge mise à jour des postes capitalisés.

Ainsi pour tous les sinistres disposant d'une rente, ce coefficient est calculé comme suit :

$$\text{Ajustement}_s = \frac{\text{Charge des postes capitalisés (valeur à } i\%)_s - \text{Charge des postes capitalisés (valeur initial)}_s}{\text{Charge de sinistre (valeur initial)}_s}$$

On peut ainsi déterminer la charge ultime actualisée par la formule suivante : $C_{i,j,s}^a = C_{i,j,s}^u \times (1 + \text{Ajustement}_s)$.

6.3 Détermination du seuil des extrêmes

La première étape que nous avons effectué nous a permis de redresser notre échantillon historique afin de le rendre représentatif des conditions de souscriptions économiques, de l'inflation, de l'évolution de la taille du portefeuille mais aussi d'effectuer la projection à l'ultime de nos sinistres et le traitement des postes capitalisés. Maintenant que nous disposons d'un échantillon retraité, nous pouvons réaliser la première étape de la construction de notre modèle : la détermination du seuil des extrêmes.

Dès lors que le seuil des extrêmes est déterminé, nous pourrions segmenter notre sinistralité et modéliser nos lois de sévérité en dessous et au-dessus de ce seuil. Cette partie est donc dédiée à la détermination et l'optimisation de ce seuil des extrêmes, afin de disposer d'un seuil des extrêmes idéal, nous permettant ainsi une modélisation adéquate de nos lois de sévérités.

6.3.1 Détermination du domaine d'attraction des données

Une étape préliminaire est cependant importante, elle correspond à la détermination du domaine d'attraction de nos données. Dans le cadre de notre étude, cette étape permet de savoir si la queue de la distribution des données est épaisse ou fine. Cela a un impact direct sur le choix du seuil, qui est une étape cruciale dans la modélisation de la loi de valeurs extrêmes. En effet, si la queue de la distribution est épaisse, l'indice des valeurs extrêmes est strictement positif, ce qui signifie que les pertes extrêmes sont plus fréquentes que ce qui est prévu par une distribution normale. Dans ce cas, il est important de choisir un seuil élevé pour capturer ces pertes extrêmes et éviter des sous-estimations dans la modélisation. En revanche, si la queue de la distribution est fine, l'indice est négatif et il est important de choisir un seuil plus bas pour éviter de fausses détections de pertes extrêmes. Par conséquent, la détermination du domaine d'attraction des données permet de mieux choisir le seuil pour une modélisation plus précise et adaptée à la réalité des pertes extrêmes en réassurance.

De plus, certains outils graphiques pour déterminer le seuil, sont valides que pour certaines valeurs de l'indice de la loi de valeurs extrêmes. En effet, si les données appartiennent au domaine d'attraction de Fréchet, les outils graphiques tels que le Hill plot peuvent être utilisés pour estimer le seuil des valeurs extrêmes, mais ils ne sont pas valides pour les données qui n'appartiennent pas à ce domaine d'attraction. Par conséquent, il est important de connaître le domaine d'attraction des données avant de choisir la méthode pour déterminer le seuil.

Pour déterminer le domaine d'attraction, il existe différentes méthodes. Une méthode couramment utilisée est la méthode par le quantile plot généralisé

En comparant les courbes obtenues avec les courbes de référence pour les distributions de Fréchet, Gumbel et Weibull, on peut déterminer le domaine d'attraction des observations :

- Si la courbe pour la distribution de Fréchet est la plus proche de la diagonale, alors les observations appartiennent au domaine d'attraction de Fréchet.
- Si la courbe pour la distribution de Gumbel est la plus proche de la diagonale, alors les observations appartiennent au domaine d'attraction de Gumbel.
- Si la courbe pour la distribution de Weibull est la plus proche de la diagonale, alors les observations appartiennent au domaine d'attraction de Weibull.

Sinon, cela suggère que les données n'appartiennent pas à ce domaine d'attraction. De manière générale on a:

- Si la courbe est convexe et décroissante : les données appartiennent probablement au domaine d'attraction de Fréchet.
- Si la courbe est concave et croissante : les données appartiennent probablement au domaine d'attraction de Gumbel.
- Si la courbe est droite : les données appartiennent probablement au domaine d'attraction de Weibull.

Pour le domaine d'attraction de Fréchet, on s'attend à observer une droite croissante dans le quantile plot généralisé. En effet, la loi de Fréchet modélise les queues de distribution épaisses, donc les valeurs extrêmes sont plus fréquentes que pour une distribution avec des queues fines. Ainsi, plus la valeur de la variable aléatoire est grande, plus la probabilité que cette valeur soit extrême est élevée.

En revanche, pour le domaine d'attraction de Weibull, on s'attend à observer une droite décroissante dans le quantile plot généralisé. La loi de Weibull modélise les queues de distribution fines, donc les valeurs extrêmes sont plus rares que pour une distribution avec des queues épaisses. Ainsi, plus la valeur de la variable aléatoire est grande, moins la probabilité que cette valeur soit extrême est élevée.

Voici un exemple de quantile plot généralisé pour des lois appartenant respectivement au domaine d'attraction de Fréchet (en bleu), Gumbel (en noir) et Weibull (en rouge) :

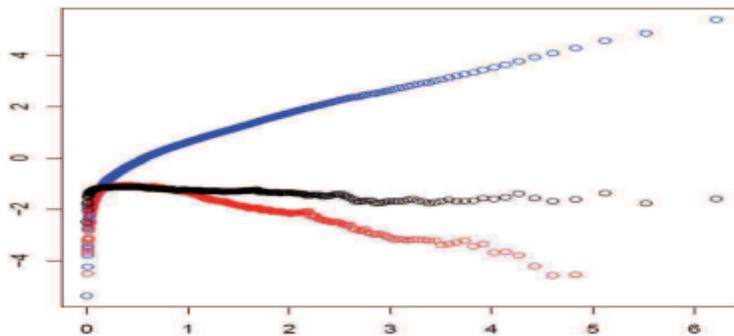


Figure 17 : Graphique du quantile-plot généralisé pour des lois appartenant aux domaines d'attractions de Fréchet, Gumbel et Weibull

Voici l'allure du Quantile-plot généralisé sur nos données :

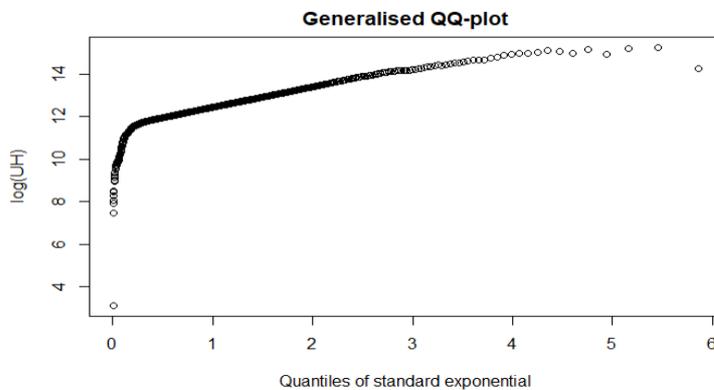


Figure 18 : Graphique Quantile plot généralisé sur notre échantillon

On voit clairement que la courbe du quantile-plot généralisé est convexe et décroissante, et son allure est croissante. On peut donc considérer que les données appartiennent au domaine d'attraction de Fréchet et que par conséquent, d'après la partie 4.1.1, on peut déduire que l'indice des valeurs extrêmes est positif.

Une seconde méthode consiste à estimer le coefficient de stabilité de Hill, qui permet d'estimer la queue de distribution de la distribution des valeurs extrêmes.

Le coefficient de stabilité de Hill est défini comme suit :

$$\hat{k}_n = \frac{1}{d} \sum_{i=1}^d \ln \left(\frac{X_{n,i:n} - X_{n,d:n}}{X_{n,d:n} - X_{n,1:n}} \right)$$

où $X_{n,1:n}, \dots, X_{n,n:n}$ sont les n valeurs triées par ordre décroissant et $d < n$ est le nombre de valeurs considérées pour estimer \hat{k}_n .

Une fois que le coefficient de stabilité de Hill a été estimé, il est possible d'estimer le domaine d'attraction de la distribution des valeurs extrêmes. En général, on considère que la distribution des valeurs extrêmes est de type GPD si $\hat{k}_n > 0$, de type Fréchet si $\hat{k}_n = 0$, et de type Weibull si $\hat{k}_n < 0$.

En appliquant la méthode à notre échantillon d'observations, celle-ci vient corroborer notre hypothèse selon laquelle nos données appartiennent au domaine d'attraction de Fréchet.

6.3.2 Choix du seuil des extrêmes

Dans cette partie, nous allons traiter du choix du seuil, pour cela nous allons faire référence à la partie 4.3 qui nous suggère que pour un seuil u , suffisamment grand et correctement choisi, la loi de Pareto Généralisée est une bonne approximation pour nos données.

6.3.2.1 Graphique de Hill

La partie précédente, nous a permis de déterminer le domaine d'attraction de notre échantillon d'observation : nos données sont supposées appartenir au domaine d'attraction de Fréchet. Par conséquent, on peut tracer le graphique de Hill afin de déterminer une première approximation du seuil des extrêmes.

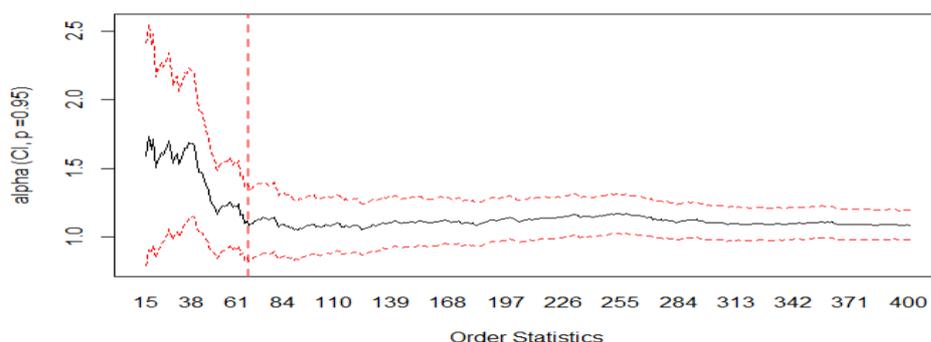


Figure 19 : Graphique de Hill sur notre portefeuille

En utilisant les critères de détermination du choix du seuil selon la méthode du graphique de Hill, on observe que la courbe des excès est approximativement linéaire et se stabilise à partir de la courbe verticale rouge, représentant un nombre d'excès égal à 67 et donc un seuil des extrêmes à 2 142 946€.

6.3.2.2 Graphique de Pickands

En traçant le graphique de Pickands sur nos données d'observations selon la méthode décrite en 5.1 Détermination du seuil des extrêmes par l'estimateur de Hill ou de Pickands, on obtient le graphique suivant :

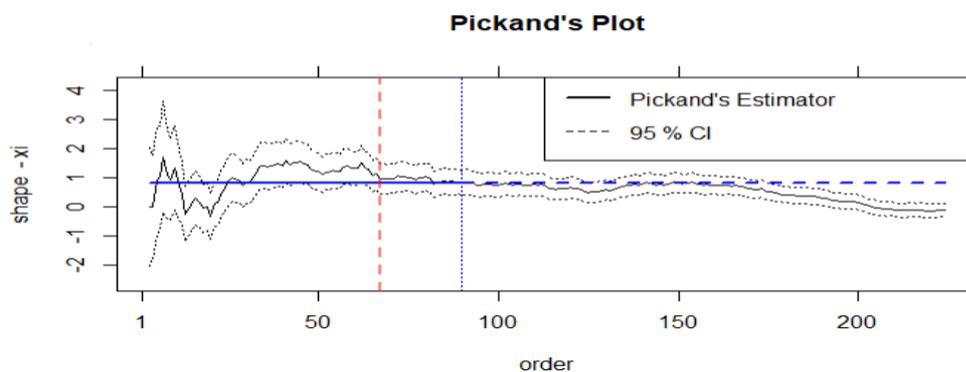


Figure 20 : Graphique de Pickands pour notre portefeuille

La courbe des excès est approximativement linéaire à partir de la ligne verticale rouge, représentant un nombre d'excès de 67, soit un seuil des extrêmes estimé à 2 142 946€.

Le graphique de Pickands vient corroborer l'hypothèse émise par le graphique de Hill.

6.3.2.3 Optimisation de l'estimateur de Hill par la méthode de minimisation de l'erreur quadratique asymptotique

Notre échantillon d'observations étant relativement faible, et les estimateurs de l'indice des valeurs extrêmes sont asymptotiques. Par conséquent, choisir un nombre optimal de statistiques d'ordre extrêmes k_{opt} n'est pas facile à manipuler. En effet, lorsque k est petit, la variance de l'estimateur est grande et l'utilisation de k introduit un grand biais, or l'équilibrage de la variance et du biais est important dans les applications de la théorie des valeurs extrêmes car cela réduit l'erreur quadratique moyenne.

On peut donc appliquer une méthode d'optimisation du choix du seuil basée sur la minimisation de l'erreur quadratique moyenne. En appliquant cette méthode à l'estimateur de Hill, on cherche à sélectionner un nombre d'excès k_{opt} qui minimise son erreur quadratique moyenne :

Soit $\gamma^{\widehat{H}}(\widehat{k}_n)$ l'estimateur de Hill de l'indice des valeurs extrêmes γ défini en 5.1 :

$$\gamma^{\widehat{H}}(\widehat{k}_n) = \frac{1}{\widehat{k}_n} \sum_{i=1}^n \ln(X_{(n-i+1)}) - \ln(X_{(n-\widehat{k}_n)})$$

Où $X_{(n-\widehat{k}_n)}$ est une estimation du seuil des extrêmes et \widehat{k}_n un entier compris entre 1 et n . En appliquant cette méthode, on cherche un seuil $u_{opt} = X_{(n-\widehat{k}_{opt})}$, avec \widehat{k}_{opt} une estimation de la valeur k_{opt} de k_n qui minimise l'erreur quadratique moyenne asymptotique de l'estimateur de Hill défini par :

$$AMSE(\gamma^{\widehat{H}}(\widehat{k}_n)) = \gamma^2 \left(\frac{1}{\widehat{k}_n} + \frac{\lambda^2}{(1-\rho)^2} \left(\frac{n}{\widehat{k}_n} \right)^{2\rho} \right)$$

Cette fonction est minimisée pour $k_{opt} = \left\lceil \left(\frac{(1-\rho)^2 n^{-2\rho}}{-2\rho\lambda^2} \right)^{\frac{1}{(1-2\rho)}} \right\rceil$ où p et λ sont des paramètres du second ordre contrôlant la vitesse de convergence de l'estimateur de Hill.

En appliquant cette méthode à notre échantillon d'observation on obtient :

```
$second.order.par
[1] 1.0092845 -0.7409946

$k0
[1] 76

$threshold
[1] 2119419

$tail.index
[1] 1.107346
```

Figure 21 : Minimisation sur R de l'erreur quadratique moyenne asymptotique de l'estimateur de Hill

Cette méthode d'optimisation du seuil, nous fournit donc une valeur du seuil des extrêmes estimé à 2 119 419 €.

6.3.2.4 Graphique de Gerstengarbe :

En utilisant la méthode du graphique de Gerstengarbe, nous cherchons à identifier un changement lors de l'entrée en zone des observations extrêmes, on s'attend à une variation du comportement de la série décrite en 5.3 *Détermination du seuil des extrêmes par le Gerstengarbe plot.*

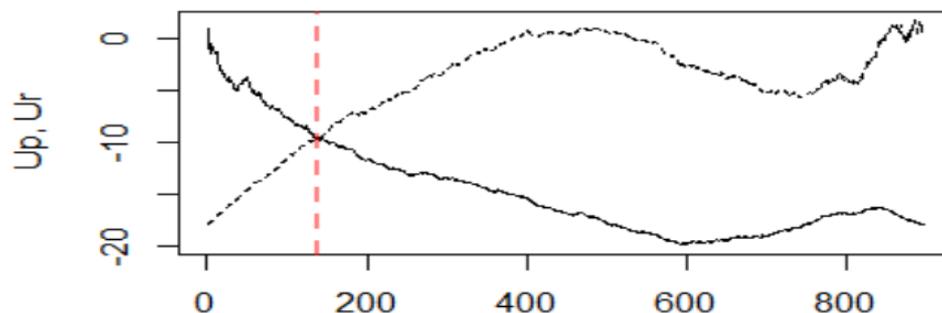


Figure 22 : Graphique de Gerstengarbe sur notre échantillon de données

La courbe en rouge, représente l'intersection des deux séries se faisant pour un nombre d'excès estimé à 167.

Nous pouvons conclure que la méthode du graphique de Gerstengarbe, nous indique un choix du seuil des extrêmes estimé à 946 426 €. Cette valeur du seuil des extrêmes reste toutefois très éloignée des valeurs trouvées par les deux premières méthodes utilisées.

6.3.2.5 Graphique des excès moyens

Nous allons maintenant tracer le graphique des excès moyens selon la méthode décrite en 5.2 *Détermination du seuil des extrêmes par la fonction des excès moyens :*

Sur ce même graphique, on trace aussi les nombres d'excès déterminés par les autres outils graphiques :

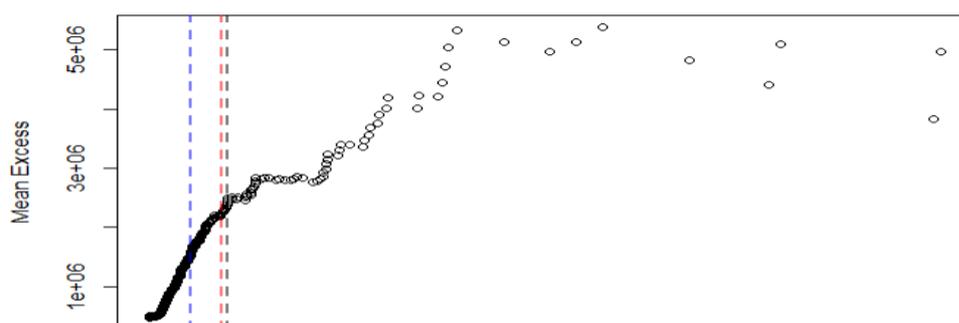


Figure 23 : Graphique des excès moyens sur notre portefeuille

La courbe verticale bleue représente le seuil déterminé par la méthode du graphique de Gerstengarbe, en rouge par la méthode de minimisation de l'erreur quadratique asymptotique moyenne de l'estimateur de Hill et en noire par la méthode Hill et Pickands. La courbe des excès moyens est plus ou moins linéaire entre la courbe verticale noire et rouge. Cette méthode tend à confirmer le choix du seuil choisi par la méthode du graphique de Hill ou de minimisation de l'erreur quadratique moyenne.

6.3.2.6 Conclusion

Méthode d'estimation	Graphique de Hill	Graphique de Pickands	MSE	Graphique de Gerstengarbe	Graphique des excès moyens
Seuil des extrêmes	2 142 946 €	2 142 946 €	2 119 419 €	946 426 €	2 142 946 €

Figure 24 : Récapitulatif des seuils des extrêmes candidats

Ainsi nous avons donc plusieurs valeurs de seuils des extrêmes candidates. Le seuil déterminé par la méthode de Gerstengarbe étant trop éloigné de celui trouvé par les autres méthodes nous laisse penser que cette valeur de seuil est un mauvais candidat. En revanche un seul et unique seuil est déterminé à plusieurs reprises favorisant ainsi son choix, c'est celui représentant un nombre d'excès à 67 soit 2 142 946 €. De plus, les méthodes du graphique des excès moyens et de minimisation de l'erreur quadratique asymptotique de l'estimateur de Hill tendent à montrer que les seuils trouvés par ces méthodes sont relativement proches de ce dernier.

Pour la suite, nous considérons ainsi un seuil des extrêmes à **2 150 000 €**.

En traçant la fonction de répartition empirique avec le seuil des extrêmes nous avons :

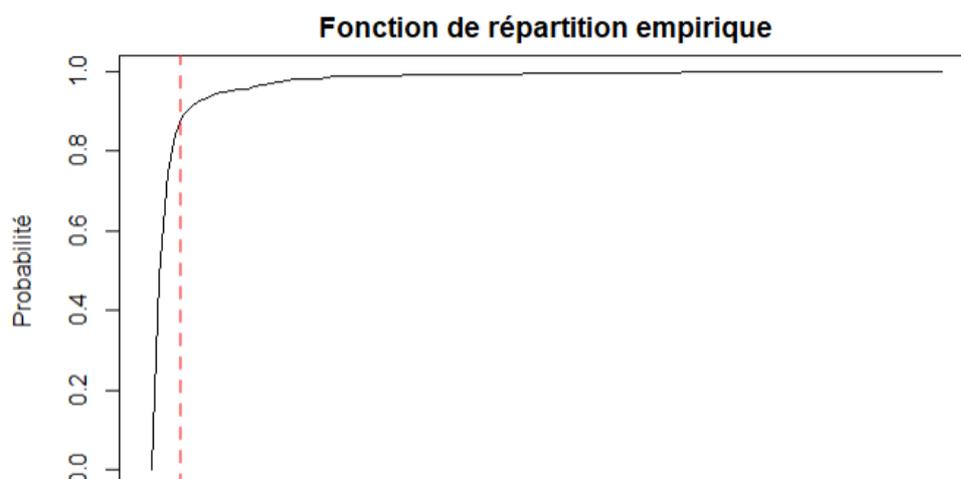


Figure 25 : Fonction de répartition empirique de nos données.

On observe aussi, que ce seuil correspond identiquement à un point de cassure de la fonction de répartition ce qui nous indique le seuil d'entrée dans la zone des valeurs extrêmes, ce qui vient fortifier notre choix du seuil des extrêmes.

Maintenant que nous avons déterminé le seuil des extrêmes, nous pouvons catégoriser nos sinistres graves de la manière suivante :

- Sinistres 250K : sinistre dont le montant est compris entre 250 000 € et 2 150 000 €.
- Sinistre extrême : sinistre dont le montant est supérieur à 2 150 000 €.

Sur nos données après indexation des sinistres : on observe que les sinistres 250K représentent 53% du nombre des sinistres de notre échantillon d'observations pour seulement 10% de la charge totale tandis que les sinistres extrêmes représentent 7% du nombre de nos sinistres pour 56% de la charge totale de nos données. En classifiant les sinistres ainsi, on s'aperçoit ainsi de la faible volumétrie des données dans la partie représentant les sinistres extrêmes mais représentant pourtant une partie très importante de la charge totale.

6.4 Modélisation de nos lois de sévérités

Le choix du seuil des extrêmes déterminé, nous allons maintenant déterminer les lois de sévérités en dessous et au-dessous du seuil des extrêmes.

6.4.1 Modélisation au-dessus du seuil des extrêmes

Commençons par modéliser notre sévérité au-dessus du seuil des extrêmes. Le modèle à deux lois que nous souhaitons élaborer est basé sur la théorie des valeurs extrêmes et plus précisément sur méthode du dépassement de seuil émanant du théorème de Pickands. Ce dernier nous permet d'approcher la distribution de notre sinistralité pour les sinistres dépassant la charge des 2 150 000€ par une distribution de Pareto Généralisée.

Dans la suite, nous utiliserons un logiciel interne à BPCE Assurances IARD afin de déterminer les différents paramètres de la loi de Pareto généralisée tout en prenant en compte les différents aléas liés à la troncature (ou censure) de nos données.

Nous utiliserons ainsi notre logiciel interne permettant de réaliser le calibrage de notre loi de sévérité ainsi que les différents tests statistiques et critères numériques.

Nous souhaitons modéliser nos sinistres supérieurs à 2 150 000 € par une distribution de Pareto généralisée par conséquent, il y a deux paramètres à estimer : le paramètre d'échelle σ et le paramètre de forme γ . Dans la partie précédente, nous avons pu déterminer que notre échantillon de données appartenait au domaine d'attraction de Fréchet, par conséquent et par définition, le paramètre de forme γ est positif.

En utilisant la méthode du maximum de vraisemblance et des moindres carrés, on obtient les résultats suivants :

	Maximum de vraisemblance	Moindres carrés
Seuil minimum	2 150 000 €	2 150 000 €
Seuil maximum	100 000 000 €	100 000 000 €
σ	1 318 347	1 320 146
γ	0.47	0.46
Moyenne théorique	4 702 927 €	4 702 927 €
Ecart-type théorique	5 154 307 €	6 428 212 €

La méthode du maximum de vraisemblance nous fournit une moyenne et un écart-type théorique, plus proche de nos données que celle des moindres carrés.

Traçons les graphiques quantile-quantile de nos lois de Pareto Généralisée estimés par les deux méthodes :

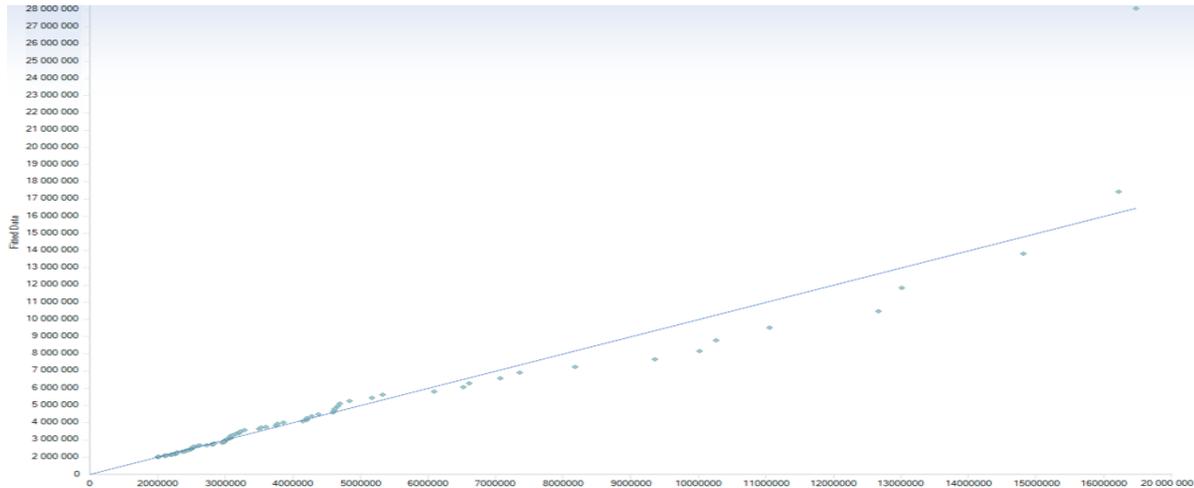


Figure 26 : Graphique Quantile-Quantile pour la GPD estimé par la méthode du Maximum de vraisemblance

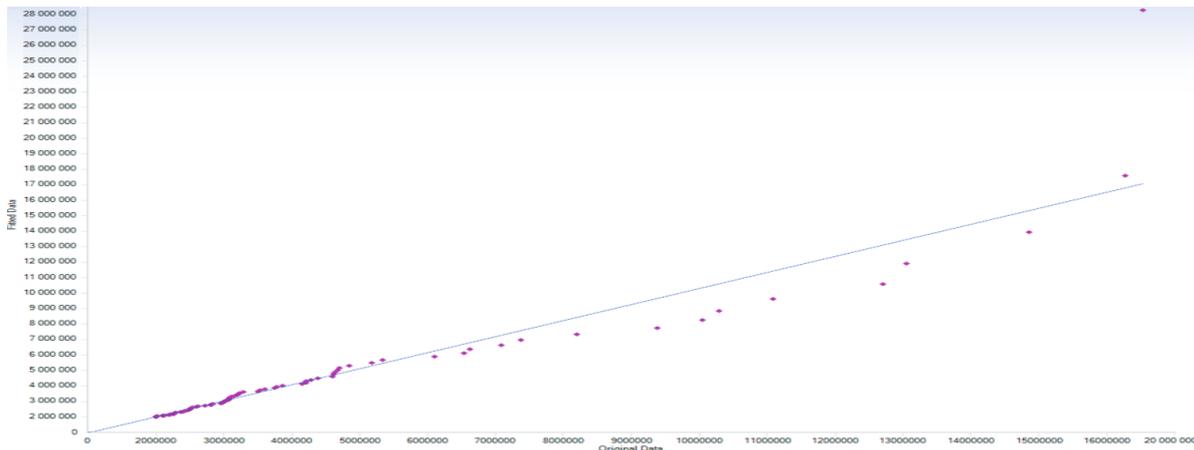


Figure 27: Graphique Quantile-Quantile pour la GPD estimé par la méthode des moindres carrés

Les QQ-plot sont très similaires, en revanche on remarque pour la GPD estimé par la méthode du maximum de vraisemblance, les points les extrêmes sont légèrement plus rapprochés de la droite linéaire.

En effectuant les tests statistiques on obtient alors :

Tests statistiques	Kolmogorov-Smirnov p-valeur = 0.955	Anderson-Darling p-valeur = 0.955	SSQ
GPD MLE	0.053	0.147	0.025
GPD moindres carrés	0.057	0.319	0.026

Les résultats des tests de Kolmogorov-Smirnov et Anderson-Darling nous indiquent que tous deux rejettent l'hypothèse nulle des tests car la p-valeur est inférieure au seuil de significativité. Pour l'estimation des paramètres de la GPD selon la méthode du maximum de vraisemblance, les trois tests donnent des valeurs inférieures à celles engendrées par l'estimation des paramètres selon la méthode des moindres carrés, cela signifie que la méthode du maximum de vraisemblance a une meilleure adéquation aux données que la deuxième loi.

Pour les raisons énumérées précédemment, nous utiliserons les paramètres estimés par la méthode du maximum de vraisemblance dans la suite de notre étude.

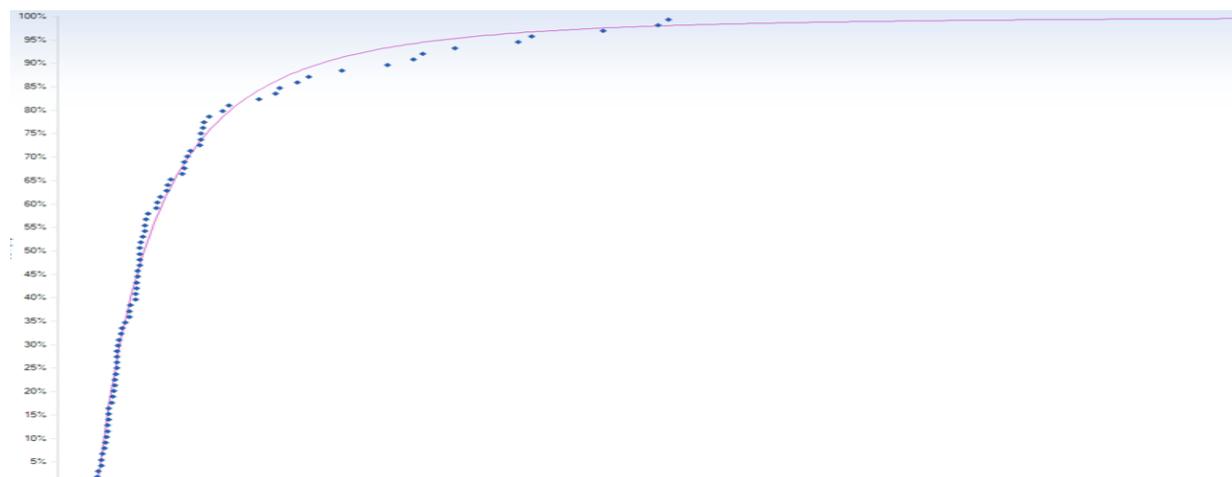


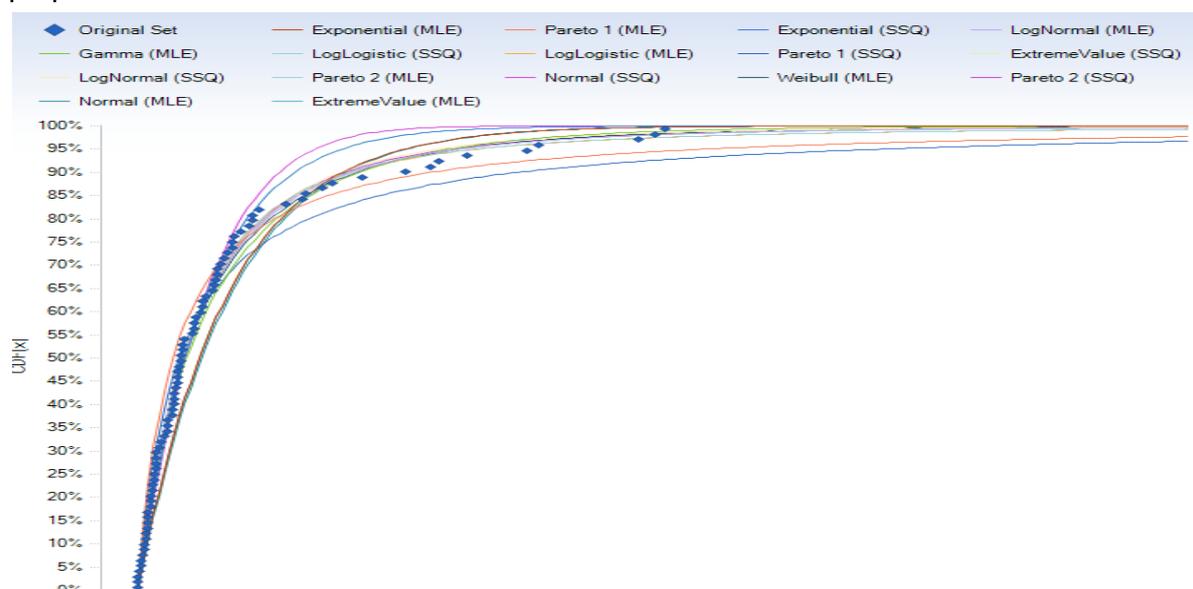
Figure 28 : Fonction de répartition estimée par le maximum de vraisemblance

Notre échantillon d'observation semble être en adéquation avec la loi modélisant la sévérité des sinistres extrêmes.

6.4.1.1 Validation du choix de notre loi de Pareto Généralisée au-dessus du seuil des extrêmes.

Afin de valider notre choix de modéliser la charge de nos sinistres par une loi de Pareto Généralisée, nous avons souhaité expérimenter d'autres lois usuelles pour modéliser la charge de nos sinistres extrêmes. Au-delà du seuil des extrêmes nous décidons de tester les lois candidates suivantes : Weibull, Log-Normal, Pareto, Gamma, Log-Logistic Exponentielle, Normal en comparaison de la loi de Pareto Généralisée estimée.

À l'issue de cette phase de comparaison, la loi de Pareto généralisée estimée par la méthode du maximum de vraisemblance a la moyenne théorique et l'écart-type théorique qui se rapproche le plus de nos données réelles, et les tests statistiques de Kolmogorov-Smirnov et d'Anderson Darling ainsi que les différents critères numériques utilisés sont plus concluants pour la loi de Pareto Généralisée que pour les autres lois candidates.



En les comparant individuellement avec la loi de Pareto généralisée estimée par la méthode du maximum de vraisemblance, on observe que cette dernière adhère mieux à notre jeu de données que les autres lois candidates.

6.4.1.2 Sensibilité du seuil des extrêmes sur l'estimation de notre GPD

Le choix du seuil des extrêmes est crucial dans l'approche du dépassement de seuil pour la modélisation des sinistres extrêmes, car il conditionne la distribution des données utilisées pour l'estimation des paramètres de la loi de GPD. Plus précisément, si le seuil est trop élevé, alors peu de données seront disponibles pour l'estimation des paramètres et la variance de l'estimation sera élevée. À l'inverse, si le seuil est trop bas, la distribution des données risque de ne plus être assez proche de la loi de GPD, ce qui entraînera des biais importants dans les estimations des paramètres.

Il est donc important de trouver le seuil optimal qui permet d'obtenir des estimations précises et non biaisées des paramètres de la loi de GPD. Pour cela, nous avons utilisé divers outils graphiques afin de résoudre cette problématique, le choix du seuil n'étant pas similaire pour toutes les méthodes et d'une personne à l'autre le choix du seuil des extrêmes peut être différent. Dans cette partie, nous allons nous intéresser à la sensibilité du choix du seuil sur l'estimation des paramètres de notre loi de Pareto Généralisé.

Ainsi pour tous les seuils candidats exprimés en partie 6.3.2 *Choix du seuil des extrêmes*, ainsi que pour certains seuils paliers, nous avons procédé à l'estimation des paramètres de la loi de Pareto Généralisée selon la méthode du maximum de vraisemblance. Les résultats sont les suivants :

Seuil des extrêmes candidat	$\hat{\sigma}$	$\hat{\gamma}$
946 426 €	489 741	1.6
1 500 000 €	793 243	1.1
1 800 000 €	864 338	0.86
2 150 000 €	1 318 347	0,47
2 400 000 €	1 298 124	0.51
3 000 000 €	1 648 357	0.53

Tableau 22 : Estimation des paramètres de la GPD selon différents choix du seuil

Dans le cadre de notre étude, nous avons opté pour un choix du seuil des extrêmes à 2 150 000€. On remarque que le paramètre d'indice de queue $\hat{\gamma}$ de notre loi de Pareto généralisée se stabilise à partir de ce seuil.

Lorsque le paramètre d'indice de queue de la loi GPD se stabilise pour un certain seuil, cela signifie que les événements extrêmes au-dessus de ce seuil sont suffisamment rares et similaires en termes de sévérité pour permettre une estimation fiable de l'indice de queue de la GPD. En d'autres termes, cela indique que le seuil choisi permet de capturer suffisamment d'événements extrêmes pour que l'estimation des paramètres de la GPD soit stable et fiable. Cela indique également que la loi GPD est bien adaptée à la modélisation des événements extrêmes de l'échantillon étudié.

Il est important de noter que le choix du seuil des extrêmes dépend également du contexte et de l'objectif de l'analyse. Par exemple, si l'on s'intéresse principalement aux sinistres les plus rares et les plus coûteux, il peut être justifié de choisir un seuil plus élevé, tandis que si l'on s'intéresse à l'ensemble de la distribution des sinistres, un seuil plus bas peut être plus approprié. Dans le contexte de notre étude, il est donc envisagé de considérer le plus petit seuil tel que le paramètre d'indice de queue de la loi de Pareto Généralisée se stabilise afin d'avoir suffisamment de données au-dessus du seuil.

6.4.2 Modélisation sous le seuil des extrêmes

De manière analogue à la partie précédente, nous allons déterminer quelle loi modélise le mieux notre sévérité en dessous du seuil des extrêmes.

Pour ce faire, nous allons expérimenter les lois suivantes : Weibull, Log-Normal, Pareto, Gamma, Log-Logistic Exponentielle, Normal, Pareto Généralisé, comme lois candidates à notre modèle sous le seuil des extrêmes. Ces lois seront estimées selon la méthode du maximum de vraisemblance et la méthode des moindres carrés.

En comparant les différents graphiques Quantile-Quantile et les graphes des fonctions de répartition par rapport à notre échantillon de base : Les lois de Pareto généralisé, Weibull et Log-Normal, estimés par la méthode des moindres carrés semblent être concurrentes pour la modélisation des sinistres en dessous du seuil des extrêmes.

En revanche les résultats des tests statistiques de Kolmogorov-Smirnov, d'Anderson-Darling et l'erreur quadratique moyenne ainsi que la comparaison des valeurs de la moyenne et de l'écart-type théorique par rapport à notre échantillon de base tendent finalement à opter pour une modélisation des sinistres en dessous du seuil par une loi suivant une distribution Log-Normale de paramètres $\mu = 281724$ et $s = 362433$.

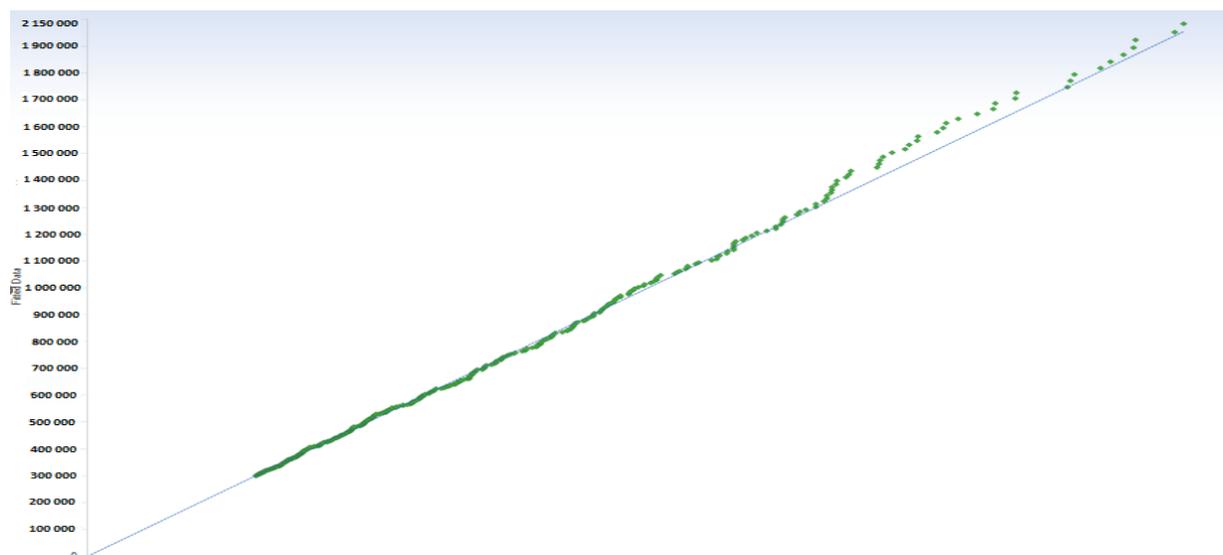


Figure 29 : Graphique Quantile-Quantile de la loi Log-Normal

Sur le graphique Quantile-Quantile ainsi que sur le graphique de la fonction de répartition, la loi Log-Normal semble être en adéquation avec nos données. Nous choisirons ainsi cette loi pour modéliser notre sinistralité en dessous du seuil des extrêmes.

6.5 Modélisation de la charge annuelle et application du traité de réassurance

Dans les étapes précédentes, nous avons déterminé les différentes lois de sévérités en dessous et au-dessus du seuil des extrêmes. De plus, nous avons à l'aide de la méthodologie retenue par BPCE Assurances IARD déterminé les lois de fréquences sur ces différentes segmentations.

Maintenant que nous avons à notre disposition toutes les lois de sévérités et de fréquences en dessous et au-dessus du seuil des extrêmes, nous pouvons créer notre modèle de loi et ainsi déterminer notre charge annuelle. Le modèle à deux lois que nous avons établi se décrit de la manière suivante :

- En dessous du seuil des extrêmes, la loi de sévérité suit une distribution Log-Normale X^{inf} avec une loi de fréquence P^{inf} définie sous ce seuil.
- Au-dessus du seuil des extrêmes, la loi de sévérité suit une distribution de Pareto généralisé X^{sup} avec une loi de fréquence P^{sup} définie elle aussi sous ce seuil.

Ainsi pour chaque segmentation on calcule la charge annuelle correspondant à :

$\{Y^i = \sum_{k=1}^N X_k^i, i = inf \text{ ou } sup\}$ et la fonction de répartition de la charge annuelle sur chaque segmentation est donnée par : $F_{Y^i}(x) = P(Y^i \leq x) = \sum_{n=0}^{\infty} P(N = n) \times P(Y^i \leq x | N = n)$.

La charge annuelle est donc modélisée par la variable $Y = Y^{inf} + Y^{sup}$

Pour le calcul de la distribution de la charge annuelle, nous allons utiliser la méthode de simulations de Monte-Carlo décrite en 3.3.7 Détermination de la charge annuelle par simulation de Monte-Carlo.

Pour cela nous utiliserons un logiciel interne à BPCE Assurances IARD.

Ainsi pour chaque segmentation ($i = inf \text{ ou } sup$) de notre modèle (inférieur ou supérieur au seuil des extrêmes), nous allons via l'outil :

- Générer un nombre N de sinistres correspondant à la réalisation de la loi de fréquence retenue sur la segmentation
- Générer N valeurs de $X_{k=1,2...N}^i$
- Calculer $Y^i = \sum_{k=1}^N X_k^i$

Pour les sinistres qui sont modélisés au-dessus du seuil des extrêmes, nous pouvons appliquer notre traité de réassurance ainsi :

- Pour chaque coût des sinistres $k = 1,2 \dots N$ on applique alors notre traité de réassurance en excédent de sinistre.
- On calcule alors la charge annuelle par tranches cédées au réassureur Y_{TR}^{sup} avec $R = 1,2,3$ nos trois tranches de réassurances.

On réitère alors ces étapes $B = 40\ 000$ fois afin d'obtenir B valeurs de Y^i et Y_{TR}^{sup} constituant ainsi leurs distributions.

Pour chaque itération de ce procédé, nous simulons simultanément les sinistres inférieurs et supérieurs au seuil des extrêmes nous donnant ainsi les valeurs Y^{inf}, Y^{sup} de la charge annuelle de chaque segmentation. On obtient alors notre charge annuelle $Y = Y^{inf} + Y^{sup}$ et en déduit alors la distribution de notre charge annuelle et de notre programme de réassurance, nous donnant par exemple le coût moyen de notre charge annuelle ou de notre programme d'assurance.

Pour les sinistres au-dessus du seuil des extrêmes, chaque itération nous génère aussi le montant des charges annuel relative à notre cession au réassureur selon la tranche T1, T2 et T3 définis par notre

traité, respectivement $Y_{T^1}^{sup}$, $Y_{T^2}^{sup}$ et $Y_{T^3}^{sup}$. A noter que l'outil utilisé prend en compte les différentes clauses de notre traité de réassurance, notamment les clauses d'AAL et AAD. De plus, il est important de tenir compte de la clause de stabilité, ainsi pour maintenir la valeur réelle des limites et des primes tout au long de la période de couverture, les portées du traité doivent être ajustées pour refléter l'impact de l'inflation. Pour ce faire, nous calculons un facteur de stabilité issu des cadences décumulée et du taux d'inflation. Nous pouvons ainsi recalculer les portées et priorités stabilisées pour les tranches de notre traité de réassurance, afin d'appliquer correctement notre cession par sinistre.

On obtient ainsi :

		Moyenne	Ecart-type
Tranche inférieure au seuil des extrêmes	Fréquence	119.6	28.40
	Coût total	39 205 050 €	9 628 042 €
Tranche supérieure au seuil des extrêmes	Fréquence	16.4	7.17
	Coût total	57 366 029 €	32 245 603 €

Tableau 23 : Estimation de la fréquence et de la charge moyenne pour le modèle issu de la Pareto généralisée

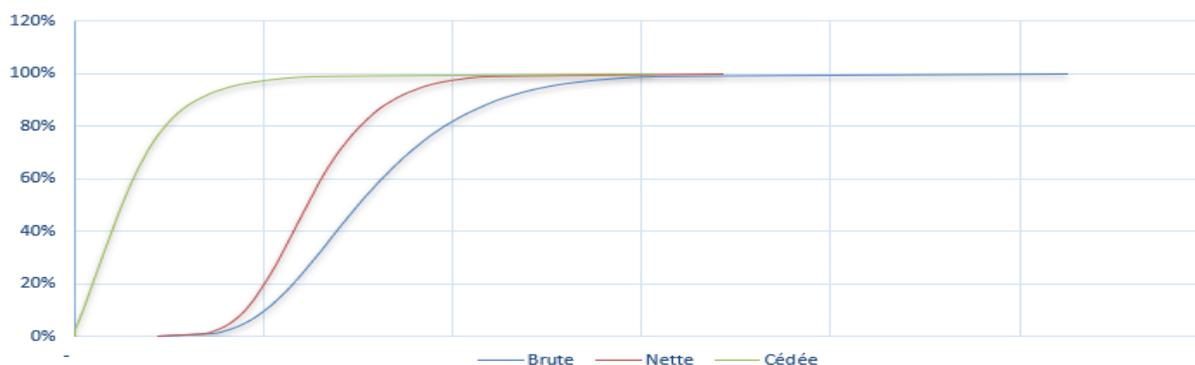


Figure 30 : Distribution de la charge brute, nette et cédée RC AUTO

L'outil nous permet ainsi de connaître les résultats de notre programme de réassurance par tranche en estimant un coût moyen :

	T1	T2	T3	Total
Moyenne	9 821 376 €	2 838 826 €	2 443 265 €	15 103 468 €
Ecart-type	11 231 134 €	3 182 617 €	2 993 902 €	17 407 654 €

Tableau 24 : Estimation de la cession en fonction des tranches de notre traité pour le modèle issu de la Pareto généralisée.

6.6 Back-Testing et validation du modèle retenu

Afin de conclure de l'efficacité de notre nouveau modèle issu de la théorie des valeurs extrêmes, nous avons effectué une étape de back-testing. En répétant le même procédé que décrit précédemment nous avons modélisé le coût de notre cession sur des années de survenances N et N-I pour le modèle à deux lois issues de la théorie des valeurs extrêmes mais aussi pour le modèle historique de BPCE Assurances IARD. Le but de cette partie est de comparer ces deux modèles.

Comparaison de notre historique aux modèles à deux lois et historique de BPCE Assurances IARD.

Cette étape va nous permettre de tester la qualité de notre modélisation. En effet, ayant connaissance des sinistres survenus sur les années N et N-I, nous allons comparer les résultats du modèle à deux lois issues de la théorie des valeurs extrêmes ainsi que le modèle historique à ceux réellement survenus.

Pour ce faire, dès lors que le processus de modélisation est répété pour ces modèles, nous calculons pour chaque tranche de notre portefeuille, la cession ainsi que les ratios $S(\text{cédée simulé})/P$ et nous allons les comparer les uns aux autres. On obtient les tableaux suivants :

Pour la charge dossier/dossier réelle :

		T1	T2	T3	Total
Charge cédée	N-I	9 759 417 €	2 730 981 €	1 969 182 €	14 459 580 €
	N	11 249 480 €	3 161 754 €	2 862 389 €	17 273 624 €

Tableau 25: Comparaison de la cession par tranches pour la charge dossier/dossier en N-I et N

Pour le modèle historique de BPCE Assurances IARD :

		T1	T2	T3	Total
Charge cédée	N-I	9 702 885 €	2 668 707 €	1 471 753 €	13 843 346 €
	N	11 193 201 €	3 023 233 €	1 427 503 €	15 643 939 €

Tableau 26: Comparaison de la cession par tranches pour la charge issue de la modélisation historique de BPCE Assurances IARD en N-I et N

Pour la tranche T1 : La charge cédée dossier/dossier en N-I est de 9 759 417 € et en N, elle est de 11 249 480 €.

- La tendance d'évolution de la charge dossier/dossier est respectée.
- En N-I, le modèle est très proche de la charge cédée historique.
- En N, le modèle est très proche de la charge cédée historique.

Pour la tranche T2 : La charge cédée dossier/dossier en N-I est de 2 730 981 € et en N, elle est de 3 161 754 €.

- La tendance d'évolution de la charge dossier/dossier est respectée.
- En N-I, le modèle sous-estime légèrement le coût de la cession.
- En N, le modèle sous-estime légèrement le coût de la cession.

Pour la tranche T3 : La charge cédée dossier/dossier en N-I est de 1 969 182 € et en N, elle est de 2 862 389 €.

- La tendance d'évolution de la charge dossier/dossier n'est pas respectée.
- En N-I, le modèle sous-estime le coût de la cession.
- En N, le modèle sous-estime le coût de la cession.

Pour le modèle issu de la loi de Pareto Généralisée :

		T1	T2	T3	Total
Charge cédée	N-I	9 821 376 €	2 838 826 €	2 443 265 €	15 103 468 €
	N	11 231 134 €	3 182 617 €	2 993 902 €	17 407 654 €

Tableau 27 : Comparaison de la charge cédée pour la modélisation issue de la loi de GPD en N-I et N

Pour la tranche T1 : La charge cédée dossier/dossier en N-I est de 9 759 417 € et en N, elle est de 11 249 480 €.

- La tendance d'évolution de la charge dossier/dossier est respectée.
- En N-I, le modèle est très proche de la charge cédée historique.
- En N, le modèle est très proche de la charge cédée historique.

Pour la tranche T2 : La charge cédée dossier/dossier en N-I est de 2 730 981 € et en N, elle est de 3 161 754 €.

- La tendance d'évolution de la charge dossier/dossier est respectée.
- En N-I, le modèle surestime légèrement le coût de la cession.
- En N, le modèle est très proche de la charge cédée historique.

Pour la tranche T3 : La charge cédée dossier/dossier en N-I est de 1 969 182 € et en N, elle est de 2 862 389 €.

- La tendance d'évolution de la charge dossier/dossier est respectée.
- En N-I, le modèle surestime très légèrement le coût de la cession.
- En N, le modèle surestime légèrement le coût de la cession.

Au global, on s'aperçoit que le modèle historique de BPCE Assurances IARD a tendance à bien estimer les coûts de la cession pour la tranche la plus basse (T1). En revanche plus la tranche est élevée (T2 puis T3) et donc moins travaillante, plus on observe une forte irrégularité et instabilité quant à l'estimation des coûts de cessions.

Pour le modèle issu de la loi de Pareto généralisée, on observe que sur la tranche basse le modèle estime très bien les pertes et a tendance à être plus prudent sur les tranches hautes en comparaison de notre historique.

On peut conclure que pour la tranche 1 les résultats sont assez proches car ils concernent une tranche travaillante, en revanche pour les tranches 2 et 3 qui sont peu ou très peu travaillante, le modèle historique de BPCE Assurances IARD présente un risque de sous-tarifcation. Le modèle issu de la loi de Pareto généralisé semble fournir des résultats plus cohérents car étant mieux adaptés à la modélisation des queues de distributions.

Il semblerait ainsi que le modèle historique soit instable de par sa capacité à sous-estimer la tranche peu travaillante d'une année à l'autre ou au global, tandis que le modèle avec GPD semble avoir une certaine tendance et une régularité quant à son estimation du risque sur ces tranches et au global.

Par mesure de confidentialité, nous ne pouvons pas faire apparaître les résultats des ratios S/P. En revanche, il en résulte que la comparaison des différents ratios S/P prédits par le modèle à deux lois issu de la théorie des valeurs extrêmes présente des résultats cohérents avec les ratios S/P réels, permettant ainsi de nous assurer que le modèle que l'on a construit est capable de capturer les variations et les tendances entre les sinistres et les primes, ce qui vient conforter le choix de notre modèle issu de la théorie des valeurs extrêmes pour modéliser notre cession.

6.6.1 Construction d'un modèle issu d'une loi marché

Dans cette partie, nous avons essayé d'optimiser le modèle à deux lois issues de la théorie des valeurs extrêmes afin de le rendre plus performant.

Nous avons utilisé une loi marché afin de modéliser les sinistres extrêmes. Dans cette partie nous allons ainsi construire deux modèles intermédiaires afin de les comparer à notre modèle issu de la loi GPD.

Les données fournies nous donnent l'approche à adopter pour les sinistres dépassant notre seuil des extrêmes, un second seuil ainsi que la loi marché au-delà de ce second seuil. Le premier modèle que nous expérimentons est basé sur notre modèle issu de la loi de Pareto généralisée :

Les sinistres en dessous du seuil des extrêmes sont modélisés de la même manière que précédemment, reprenant ainsi les mêmes paramètres de loi de fréquence et de coût que celles du modèle GPD.

Les sinistres au-dessus du seuil, sont eux aussi estimés de la même manière que précédemment. En revanche, lorsque nous procédons à la modélisation de la charge annuelle par la méthode de Monte-Carlo, nous allons apporter quelques modifications au modèle initial.

Pour les sinistres en dessous, il n'y a aucuns changements dans le procédé :

- On génère un nombre N de sinistres correspondant à la réalisation de la loi de fréquence retenue sur cette segmentation
- On génère N valeurs de $X_{k=1,2...N}^{inf}$
- Calculer $Y^{inf} = \sum_{k=1}^N X_k^{inf}$

En revanche pour les sinistres au-dessus du seuil des extrêmes :

- On génère un nombre N de sinistres correspondant à la réalisation de la loi de fréquence retenue sur cette segmentation
- Un deuxième filtre est alors appliqué pour chaque réalisation :
 - o Si le montant de cette dernière est supérieur au seuil u_2 (avec u_2 supérieur à notre seuil des extrêmes) alors le coût du sinistre est modélisé par la loi marché.
 - o Sinon il n'est pas modifié, et le sinistre conserve sa charge modélisée par la GPD.
- On a donc généré N valeurs de $X_{k=1,2...N}^{sup}$
- Calculer $Y^{sup} = \sum_{k=1}^N X_k^{sup}$
- Pour chaque coût des sinistres $k = 1,2 \dots N$ on applique alors notre traité de réassurance en excédent de sinistre.
- On calcule alors la charge annuelle par tranches cédées au réassureur Y_{TR}^{sup} avec $R = 1,2,3$ nos deux tranches de réassurances.

On réitère alors ces étapes $B = 40\,000$ fois afin d'obtenir B valeurs de Y^i et Y_{TR}^{sup} constituant ainsi leurs distributions, et comme précédemment on en déduit alors la distribution de notre charge annuelle et de notre programme de réassurance.

Nous créons donc deux modèles issus de cette loi marché, le premier étant le modèle qui modélise la sévérité au-dessus du seuil des extrêmes par une loi de Pareto Généralisée et le second par le modèle historique de BPCE Assurances IARD.

Pour ces deux modèles mixtes, la charge moyenne annuelle modélisée au-dessus du seuil des extrêmes est significativement basse par rapport à celle modélisée par les modèles GPD et historique (écart relatif de -12% et -16%).

Concernant les programmes de réassurance de ces deux lois :

		T1	T2	T3	Total
Charge cédée	N-1	9 702 885 €	2 562 319 €	945 653 €	13 210 858 €
	N	9 582 059 €	2 485 456 €	1 708 379 €	13 775 895 €

Tableau 28 : Comparaison de la charge cédée pour la modélisation mixte entre la loi marché et le modèle historique BPCE Assurances IARD en N-1 et N

		T1	T2	T3	Total
Charge cédée	N-1	9 734 461 €	2 642 880 €	1 249 497 €	13 626 839 €
	N	9 682 012 €	3 010 634 €	1 767 493 €	14 460 141 €

Tableau 32 : Comparaison de la charge cédée pour la modélisation issue de la loi marché et Pareto Généralisée en N-1 et N

Pour les modèles mixtes résultant de la combinaison de notre modèle GPD ou historique à BPCE Assurances IARD avec celui d'une loi marché, on observe que les charges cédées sur les deux tranches élevées (T2 et T3) sont très largement en dessous de la charge cédée dossier/dossier pour les années N-1 et N. Les deux modèles issus de la loi marché ont tendance à sous-estimer le coût de la cession sur ces deux tranches et ainsi qu'au global. De plus les résultats des comparaisons des ratios S/P des modes issus de la loi marché ne semblent pas être cohérents avec les résultats des ratios S/P historiques. Il semblerait donc que ces modèles ne soient pas bien adaptés à la modélisation de nos valeurs extrêmes.

6.6.2 Value at Risk (VaR)

Le calcul de la VaR consiste à déterminer la perte maximale que l'entreprise pourrait subir sur une période donnée avec une certaine probabilité. En d'autres termes, cela permet d'évaluer le montant de fonds propres nécessaires pour faire face aux pertes potentielles sur la période donnée.

Elle est déterminée en choisissant un niveau de confiance, généralement 95% ou 99%. Ensuite, on détermine la distribution des pertes possibles sur une période donnée, en utilisant des données historiques ou des modèles de simulation. Le quantile d'ordre $p \in]0,1[$ associé à la variable aléatoire X de fonction de répartition F_x est défini par : $x_p = F^{-1}(p) = \sup \{x \in \mathbb{R}, F(x) \geq p\}$

On peut alors estimer la VaR pour les modèles que l'on a choisi de tester à partir du seuil des extrêmes :

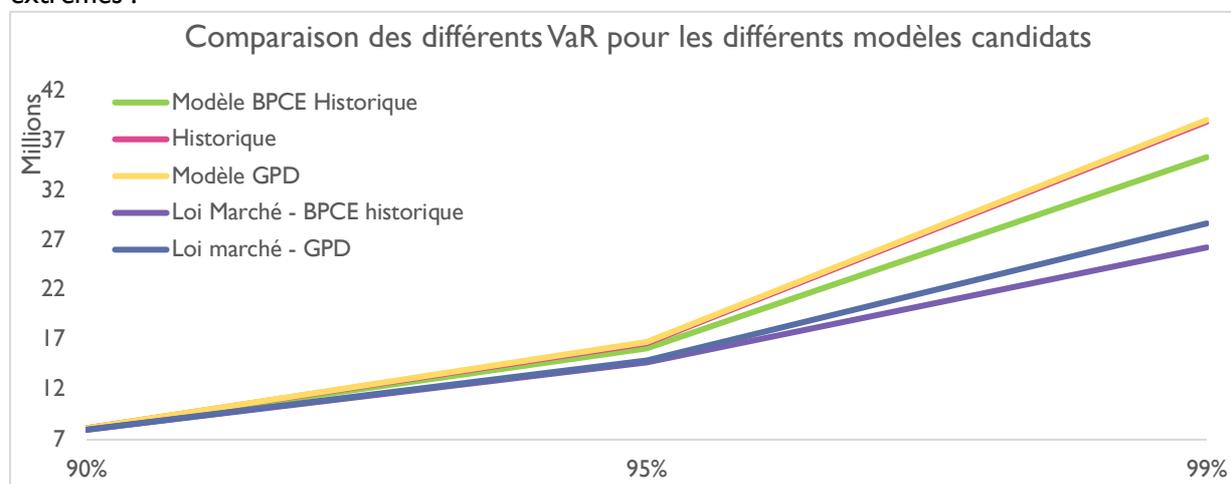


Figure 31 : VaR pour différentes p-values pour les modèles candidats

Ces quantiles permettent d'estimer la taille maximale de la perte probable et d'évaluer le montant de capital nécessaire pour couvrir cette perte. Ils sont également utilisés pour déterminer le niveau de rétention et le montant de la prime à payer pour couvrir cette rétention.

On remarque ainsi que les modèles issus de la loi marché sous-estiment tous les deux la cession.

Le modèle historique de BPCE Assurances IARD a tendance à bien estimer le risque pour tous les quantiles hormis celui à 99 % où il sous-estime légèrement le risque ne reproduisant pas une queue de distribution suffisamment épaisse. En revanche, le modèle issu de la Pareto généralisée a tendance à légèrement surestimer le risque sur les quantiles 90 % et 95 % et à bien estimer le risque sur le quantile à 90%.

6.6.3 Validation du modèle retenue

Au vu des différents résultats présentés dans les parties précédentes sur les charges cédées, les ratios S/P et de la VaR, Il est préférable d'opter pour le modèle à deux lois issues de la théorie des valeurs extrêmes que nous avons construit.

En effet ce modèle a tendance à légèrement surestimer le risque sans l'être de manière exagérée au risque de trop tarifier notre traité. Ce caractère lui vaut ainsi d'être plus prudent que les autres modèles proposés.

De plus, lors de l'ajustement de la loi de Paréto Généralisée à nos données, on s'aperçoit que la distribution prédite par le modèle correspond étroitement à la distribution de nos sinistres. Cela suggère que le modèle capture correctement les caractéristiques des événements extrêmes. Ce caractère a notamment été validé lors de la validation des tests de Kolmogorov-Smirnov et d'Anderson Darling mais aussi lors de la mesure des performances utilisés telles que le calcul du coefficient de détermination R^2 , du coefficient de corrélation et de l'erreur quadratique moyenne. Cela peut être justifié par le fait que la loi de Pareto généralisée est une distribution couramment utilisée pour modéliser les valeurs extrêmes, car elle possède des propriétés mathématiques appropriées pour modéliser les queues de distributions lourdes.

Ces différents résultats viennent valider le choix du modèle à deux lois issues de la théorie des valeurs extrêmes pour modéliser notre sinistralité.

En revanche, il aurait été pertinent d'effectuer différentes comparaisons de ces résultats afin d'identifier des périodes où les modèles présentent des écarts significatifs par rapport aux données historiques nous permettant ainsi d'identifier les zones où le modèle pourrait nécessiter des ajustements ou une réévaluation. De plus, des tests de sensibilités peuvent compléter notre analyse. En effet, si notre modèle est capable de reproduire les variations des différents résultats de cessions ou ratios de S/P en réponse à des changements de conditions du marché tel que l'inflation, des sinistres ou d'autres facteurs cela nous pourrait nous permettre de renforcer la crédibilité de notre modèle.

Ces étapes peuvent être vues dans la continuité de la validation de ce modèle au sein de BPCE Assurances IARD.

6.7 Cotation

Le modèle issu de la loi de Pareto généralisée étant validé, nous pouvons alors effectuer la cotation de notre traité pour toutes les tranches.

Avec la méthode de fréquence-coût et l'utilisation de la méthode de Monte-Carlo, nous avons pu déterminer la distribution de la charge annuelle et ainsi déduire le coût moyen cédé par tranches :

		T1	T2	T3	Total
Charge cédée	N-I	9 821 376 €	2 838 826 €	2 443 265 €	15 103 468 €
	N	11 231 134 €	3 182 617 €	2 993 902 €	17 407 654 €

Tableau 33 : Coût de la cession pour les tranches de notre traité

Cette charge annuelle moyenne pour chaque tranche représente la prime pure de réassurance.

Ensuite, la cédante peut appliquer les charges d'exploitation, les frais généraux et les marges bénéficiaires pour calculer la prime commerciale.

Les frais de charges d'exploitation, les frais généraux et les marges bénéficiaires sont généralement déterminés par la compagnie en fonction de différents critères tels que les coûts opérationnels, les exigences de rentabilité, les coûts de gestion, les dépenses de marketing et les risques liés aux activités de l'entreprise. Ces coûts sont généralement pris en compte dans le calcul de la prime commerciale qui est facturée aux clients de la compagnie d'assurance ou de réassurance.

Les frais de charges d'exploitation comprennent les coûts directs et indirects liés à la gestion des contrats d'assurance, tels que les salaires des employés, les frais de bureau, les dépenses informatiques, etc. Les frais généraux, quant à eux, représentent les coûts non liés directement à la production de contrats d'assurance, tels que les dépenses de marketing et de publicité, les frais juridiques et administratifs, etc. Les marges sont déterminées par les compagnies d'assurances ou de réassurances en fonction de leurs objectifs de rentabilité, des coûts de capital et des risques associés à leurs activités.

En pratique, les conditions différenciées pour la prime de réassurance sont déterminées par les négociations entre la cédante et la compagnie de réassurance, qui se basent sur l'évaluation des risques spécifiques de chaque contrat de réassurance.

Il convient de noter que les coûts associés aux frais de charges d'exploitation, aux frais généraux et aux marges bénéficiaires peuvent varier considérablement d'une entreprise à l'autre et dépendent des stratégies commerciales et des objectifs financiers de chaque entreprise.

Dans notre cas, ces frais sont déduits en fonction des primes et de nos modélisations historiques afin d'en déduire une prime commerciale.

6.8 Période de retour

Dans la partie précédente, nous avons déterminé, pour chaque tranche du traité que nous voulons expérimenter une prime pure. Dans cette partie nous allons voir les différents outils nous permettant de quantifier le niveau de risque associé à chacune des tranches et donc de juger de la qualité de notre couverture.

L'estimation de la période de retour nous permet de quantifier la probabilité de survenance d'un sinistre supérieur à un certain seuil de déclenchement. Elle correspond à la durée moyenne entre deux

événements de même ampleur. Plus précisément, la période de retour représente le temps moyen au-delà duquel une perte est susceptible de se produire.

En connaissant la période de retour pour chaque tranche, l'assureur peut ajuster le niveau de rétention de la tranche en fonction de sa tolérance au risque.

On appelle u_m le niveau de retour associé à une période de retour n .

Soit T_n la période de retour d'un sinistre de niveau u_n , ayant une probabilité d'être dépassé $p_n = \frac{1}{T_n}$.

Dans le cas où $(X|X > u)$ est modélisée par une loi de Pareto généralisée de paramètre σ et γ on a :

$$\mathbb{P}(X > u_n) = \frac{1}{n}$$

$$\mathbb{P}(X > u) \left(1 + \left(\frac{\gamma(u_n - u)}{\sigma} \right)^{\frac{1}{\gamma}} \right)^{-\frac{1}{\gamma}} = \frac{1}{n}$$

D'où : $z_n = u + \frac{\sigma}{\gamma} + ((nP(X > u))^\gamma - 1)$

Si la période de retour est très faible, cela signifie que la probabilité de dépasser le montant de rétention est élevée, et l'assureur peut décider d'augmenter la couverture de réassurance pour se protéger contre ce risque.

Pour les trois tranches de notre traité, on calcule les différentes périodes de retour :

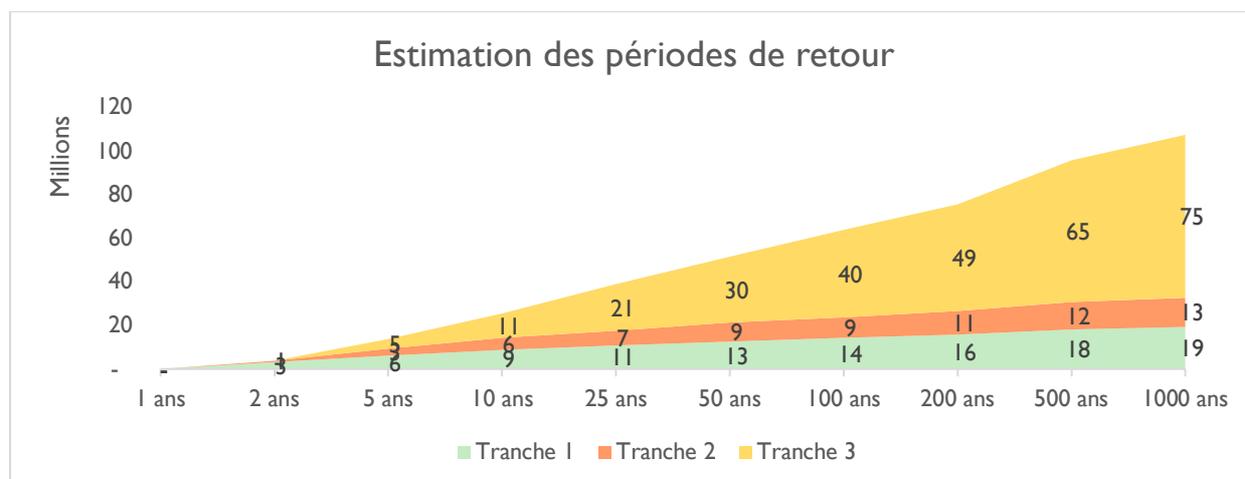


Figure 32 : Périodes de retour pour nos différentes tranches

Pour une période de retour de 5 ans, la perte maximale estimée est en moyenne de 4.6 M€. Cela signifie que l'on s'attend à une perte d'environ 3.5 M€ tous les 5 ans en moyenne.

Dans les données utilisées pour notre modélisation, un sinistre à 20 M€ était quantifié. Ce montant de sinistre est estimé à une période de retour de 100 ans en moyenne.

6.9 Rétrospective sur la validité de notre modèle

En ce qui concerne le choix du modèle pour modéliser la sinistralité, nous avons opté pour la méthode du dépassement de seuil et la loi de Pareto généralisée pour modéliser notre sévérité pour les sinistres extrêmes. Cette approche nous a permis de traiter spécifiquement les sinistres extrêmes, ce qui est pertinent dans le contexte de notre étude. Cependant, il est essentiel de souligner certaines limites liées à nos choix.

Tout d'abord, le choix de la Pareto généralisée pour modéliser les coûts de cession au-dessus du seuil des extrêmes repose sur des hypothèses simplificatrices sur la distribution des sinistres. Nous avons supposé que cette loi est adéquate pour capturer la queue de distribution des sinistres extrêmes, mais il est possible que d'autres lois ou modèles pourraient mieux correspondre à nos données. C'est pourquoi une étape de validation de l'utilisation de cette loi et de comparaison avec plusieurs modèles pour modéliser nos sinistres extrêmes a été effectuée. En revanche, l'utilisation de modèles plus variés pour les comparer à nos modèles candidats, nous aurait permis d'accroître la pertinence de notre validité. En utilisant cette méthode, nous avons supposé que les sinistres extrêmes suivent une distribution stable et constante au fil du temps. Cela peut ne pas tenir compte des différentes caractéristiques des sinistres ou d'autres facteurs qui pourraient influencer les sinistres extrêmes.

De plus, parmi les différents modèles que l'on a proposés afin de modéliser notre sinistralité, seul le modèle à deux lois issues de la théorie des valeurs extrêmes a été retenu. Les différentes méthodes de validation que l'on a utilisées pour évaluer la performance de notre modèle ont été primordiales nous permettant ainsi d'assurer la qualité et la validité de nos résultats. Cependant, en rétrospective, il est important de noter certaines limites et incertitudes associées à notre approche de validation.

Nous avons principalement utilisé des méthodes statistiques telles que la comparaison des ratios S/P et des charges cédées, l'ajustement de la distribution et les mesures de performance pour valider notre modèle. Bien que ces méthodes aient fourni des informations précieuses sur la cohérence entre les prédictions du modèle et les données historiques, elles ne tiennent peut-être pas compte de toutes les nuances et variations possibles dans notre portefeuille de sinistralité. A titre d'exemple, des tests de sensibilités sur des changements de nos conditions de marché tels que l'inflation, des sinistres ou d'autres facteurs auraient pu compléter notre analyse et nous permettre de renforcer la crédibilité de notre modèle.

En conclusion, bien que nous ayons pris des mesures pour valider notre modèle et choisi une méthode adaptée pour modéliser notre sinistralité, il est important de reconnaître les limites et les incertitudes associées à notre choix. Une approche explorant des modèles alternatifs de ceux utilisés dans notre étude à titre de comparaison pourrait compléter notre étude par exemple.

Cependant, en comparant les différents résultats obtenus des estimations des périodes de retour, des ratios S/P et celles des primes pures exprimées pour chacune de nos tranches on peut conclure que la couverture que nous avons décidé d'expérimenter est suffisamment efficace pour couvrir les pertes attendues pour des périodes de retour allant jusqu'à 25 ans. Au-delà de cette période, les pertes potentielles peuvent dépasser la prime pure et nécessiter une réévaluation de la couverture proposée.

On peut donc estimer que la couverture que nous souhaitons expérimenter semble bien nous convenir.

Conclusion

L'assurance automobile représente une part importante du marché de l'assurance en France, la responsabilité civile est la garantie minimale d'un contrat d'assurance automobile, il est donc primordial de bien maîtriser la sinistralité concernant cette garantie. L'objet de notre étude consistait à faire évoluer le modèle interne de tarification du traité en excédent de sinistre de BPCE Assurances IARD pour la garantie responsabilité civile automobile, en se focalisant sur la modélisation de la loi de sévérité des sinistres extrêmes.

La modélisation des sinistres extrêmes en réassurance est une tâche complexe qui nécessite une compréhension approfondie de l'analyse de données, des modèles statistiques et des méthodes de simulation. En effet la garantie responsabilité civile étant une branche à développement longue, pour être en mesure d'exploiter notre base de sinistres, nous avons dû prendre en compte les différents aspects relatifs à l'inflation et à notre traité de réassurance comme la prise en compte des IBNR afin de travailler sur des charges mises à l'ultime, des postes capitalisés ou encore de la clause de stabilité.

Le modèle historique de tarification de ce traité mis en place par BPCE Assurances IARD prévoyait une charge de sinistre modélisé en deçà de ce que l'on constate à posteriori sur la charge dossier/dossier, ainsi qu'une sous-estimation de la charge cédée sur la tranche non travaillante de notre traité de réassurance. Le modèle de BPCE Assurances IARD était donc adapté pour estimer la sévérité d'une tranche de réassurance travaillante, mais l'inconvénient se situe pour les tranches non-travaillantes du traité de réassurance faisant intervenir la queue de distribution. Il était donc difficile à la fois d'estimer la charge pour la réassurance dans ces tranches ainsi que d'avoir un modèle stable dans le temps.

Pour répondre à notre problématique, nous avons utilisé la méthode du dépassement de seuil issue de la théorie des valeurs extrêmes. Ce modèle consiste à définir un seuil au-delà duquel les sinistres sont considérés comme extrêmes, puis à modéliser la distribution de ces sinistres en utilisant une loi de Pareto Généralisée.

Le seuil est déterminé en sélectionnant un niveau de perte qui correspond aux sinistres les plus élevés et les plus rares. Il est déterminé à l'aide de différents outils issus de la théorie des valeurs extrêmes comme, la méthode de l'estimateur de Hill, de Gerstengarbe ou de minimisation de l'erreur quadratique moyenne asymptotique. Au-delà du seuil, les sinistres sont alors modélisés par une loi suivant une distribution de Pareto Généralisée.

Toutefois, il est important de porter une attention particulière au choix du seuil des extrêmes car pour un même jeu de données, plusieurs valeurs de celui-ci peuvent être choisies d'un actuaire à l'autre. Certaines méthodes ont récemment été développées pour pallier cette problématique comme la méthode de Thompson qui développe une procédure automatisée pour l'estimation du seuil et la quantification de l'incertitude. La méthode crée une grille uniformément espacée de valeurs de seuil possibles (entre la médiane et le quantile empirique à 98 %). Pour chaque seuil potentiel, la GPD est ajustée (en utilisant une estimation de maximum de vraisemblance) et les différences entre les paramètres d'échelle modifiés pour les seuils voisins sont calculées. Ils supposent l'asymptotique normalité de ces différences d'échelle et utilisent le fait qu'elles sont centrées autour de zéro si la GPD est le modèle correct. Ils combinent ces différences et les traitent comme un échantillon de variables aléatoires normales. Une procédure de sélection progressive (en augmentant l'incrément de seuil par incrément) est ensuite appliquée jusqu'à ce que le test du Khi-2 sous l'hypothèse de la normalité et de la moyenne zéro ne soit pas rejeté. À ce stade, le seuil est considéré comme étant cohérent avec une GPD. Une approche de bootstrap simple est utilisée pour quantifier l'incertitude du seuil sur les conclusions finales.

Un choix de seuil inapproprié peut entraîner une estimation biaisée des paramètres de la GPD, car si le seuil est trop bas, cela peut inclure des observations qui ne sont pas vraiment extrêmes et qui ne sont donc pas représentatives de la queue de la distribution. À l'inverse, si le seuil est trop élevé, cela peut exclure des observations importantes et réduire la précision de l'estimation. Il est donc important de trouver le bon équilibre dans le choix du seuil, en s'appuyant sur plusieurs méthodes de sélection du seuil.

En appliquant le modèle généré par la théorie des valeurs extrêmes et en le comparant au modèle historique de BPCE Assurances IARD, il semblerait que la segmentation faite par la détermination du seuil des extrêmes ainsi que la modélisation au-delà de ce seuil par une loi de Pareto Généralisé résout les problèmes liés à la sous-tarification des tranches peu travaillantes, tout en conservant les bonnes propriétés du modèle historique pour tranches travaillantes et pour les sinistres en dessous du seuil des extrêmes. Ce modèle résout aussi les problèmes liés à l'instabilité du choix des lois de sévérité des sinistres extrêmes dans la modélisation historique. Ainsi pour la modélisation actuelle et rétroactive effectuée en utilisant le modèle issu de la loi de Pareto Généralisée, le modèle choisi présentait un caractère plus prudentiel que le modèle historique qui avait tendance à sous-estimer le risque et à le tarifier de manière moins importante. Une tarification des tranches de notre traité de réassurance est alors estimée.

La modélisation par la théorie des valeurs extrêmes présente toutefois certaines limites.

Pour la modélisation de la fréquence et de la sévérité, la méthode du dépassement de seuil suppose que les événements extrêmes ont une distribution stationnaire dans le temps, ce qui peut être limitatif si les pratiques de souscription ou les événements du marché changent.

Cette méthode présente aussi des limites liées aux données tronquées et au choix du seuil. Les données tronquées peuvent biaiser les résultats si elles ne sont pas correctement traitées. Et comme nous l'avons vu, le choix du seuil peut également avoir un impact important sur les estimations des paramètres et sur les résultats finaux de la modélisation. Il est donc important d'être conscient de ces limites et d'utiliser des méthodes appropriées pour traiter les données tronquées et pour choisir le seuil. Malgré ces limites, la modélisation par la théorie des valeurs extrêmes reste une approche très fiable pour modéliser les sinistres extrêmes en réassurance, en particulier lorsque d'autres méthodes de tarification ne sont pas adaptées.

Une méthode de Machine Learning pourrait être envisagée pour répondre à la problématique de la tarification des tranches peu travaillante, l'apprentissage profond (deep learning).

En effet, cette méthode permet de modéliser des données complexes et de détecter des modèles non linéaires. Il est possible d'utiliser des réseaux de neurones pour prédire la fréquence et la sévérité des sinistres en fonction de différents facteurs tels que les caractéristiques des polices d'assurance, les données économiques et démographiques, et d'autres variables pertinentes.

L'avantage du deep learning est qu'il peut être utilisé pour prédire les sinistres sur des seuils très différents, pas seulement les seuils extrêmes. Cela permettrait une tarification plus précise et plus complète de l'exposition au risque, qui prendrait en compte des événements moins extrêmes mais potentiellement plus fréquents. Cependant, il convient de noter que le deep learning nécessite des données de hautes qualités et en quantité suffisante pour permettre un apprentissage optimal des modèles. Il est donc important de disposer d'un ensemble de données solide et diversifié pour obtenir des résultats fiables.

Liste des figures

Figure 1 : Répartition de la charge des sinistres selon un traité en XS par risque	23
Figure 2 : Répartition de la charge cédée au réassureur pour un traité en excédent de sinistre par risque avec un plein de rétention à 6 M€, et une somme à assurer à 10 M€.....	27
Figure 3 : Exemple d'application de reconstitutions des primes	28
Figure 4 : Exemple d'application de reconstitutions.....	29
Figure 5: Densité de la distribution EVD	76
Figure 6 : Fonction de répartition de la loi EVD dans les trois domaines d'attractions	Erreur ! Signet non défini.
Figure 7 : Comparaison des densités des trois lois	Erreur ! Signet non défini.
Figure 8 : Fonction de répartition de la GPD	78
Figure 9 : Hill plot pour la loi de Fréchet	82
Figure 10 : Hill plot pour la loi de Pareto.....	85
Figure 11: Estimateur de Hill pour des lois de Pareto et GPD.....	86
Figure 12 : Estimateur de Pickands pour des lois de Pareto et GPD.....	86
Figure 13 : Graphique de la fonction des excès moyens pour des loi respectives : uniforme, exponentielle et de Pareto.....	87
Figure 14 : Graphique de Gerstengarbe pour une loi de Pareto	88
Figure 15 : Figure 15 : Fonction de répartition de l'échantillon.....	89
Figure 16 : Répartition de la charge des sinistres de notre portefeuille par année de survenance	91
Figure 17 : Exemple d'indexation de la charge d'un sinistre.....	93
Figure 18 : Mise en AS-IF de nos données portefeuille par année de survenance	94
Figure 19: Coefficient R ² par année de développement.....	95
Figure 20 : Mise à l'ultime de la charge as-if, redressée de l'exposition par année de survenance	96
Figure 21 : Graphique du quantile-plot généralisé pour des lois appartenant aux domaines d'attractions de Fréchet, Gumbel et Weibull.....	98
Figure 22 : Graphique Quantile plot généralisé sur notre échantillon.....	98
Figure 23 : Graphique de Hill sur notre portefeuille.....	99
Figure 24 : Graphique de Pickands pour notre portefeuille.....	99
Figure 25 : Minimisation sur R de l'erreur quadratique moyenne asymptotique de l'estimateur de Hill	100
Figure 26 : Graphique de Gerstengarbe sur notre échantillon de données	101
Figure 27 : Graphique des excès moyens sur notre portefeuille	101
Figure 28 : Récapitulatif des seuils des extrêmes candidats	102
Figure 29 : Fonction de répartition empirique de nos données	102
Figure 30 : Graphique Quantile-Quantile pour la GPD estimé par la méthode du Maximum de vraisemblance	104
Figure 31: Graphique Quantile-Quantile pour la GPD estimé par la méthode des moindres carrés	104
Figure 32 : Fonction de répartition estimé par le maximum de vraisemblance.....	105
Figure 33 : Graphique Quantile-Quantile de la loi Log-Normal.....	107
Figure 34 : Distribution de la charge brute, nette et cédée RC AUTO	109
Figure 35 : VaR pour différentes p-values pour les modèles candidats	113
Figure 36 : Périodes de retour pour nos différentes tranches	116

Tableau 1 : Les quatre modes de réassurance.....	21
Tableau 2 : l'exemple fictif d'un traité en excédent de sinistre par risque 6M XS 3M.....	23
Tableau 3 : Formules de la répartition des risques	25
Tableau 4 : Portefeuille fictif d'un Quote-Part à 70%, limité à un plein de souscription de 13 M€, en M€ Figure 2 : Représentation d'un quote part avec un taux de cession à 70%.....	25
Tableau 5 : Exemple d'un traité en excédent de plein par sinistre	26
Tableau 6 : Exemple sur le fonctionnement des agregates.....	30
Tableau 7 : Exemple de stabilisation	31
Tableau 8 : Synthèses des bénéfices pour la victime.....	40
Tableau 9 : Synthèse de la nomenclature Dintilhac des postes de préjudices pour des victimes directes	43
Tableau 10 : Synthèse de la nomenclature Dintilhac des postes de préjudices pour des victimes indirectes	44
Tableau 11 : Exemple d'arrérage soumis à l'indexation.....	47
Tableau 12 : Tableau synthèse de l'indemnisation sous forme de rente ou capital	49
Tableau 13 : Caractéristiques de la couverture de réassurance du traité	51
Tableau 14 : Exemple suivi additionnel des rentes	51
Tableau 15 : Caractéristiques de la couverture de réassurance du traité en exemple.....	52
Tableau 16 : Exemple de simulation de la clause de rente	53
Tableau 17 : Exemple de calcul de cession avec clauses	53
Tableau 18 : Exemple d'application de la méthode Burning cost.....	56
Tableau 19 : Limite de l'application de la méthode du Burning Cost.....	57
Tableau 20 : Exemple de domaine d'attractions pour des lois usuelles	75
Tableau 21 : Evolution de l'indice FFA sur les 10 dernières années.....	93
Tableau 22 : Estimation des paramètres de la GPD selon différents choix du seuil.....	106
Tableau 23 : Estimation de la fréquence et de la charge moyenne pour le modèle issu de la Pareto généralisée.....	109
Tableau 24 : Estimation de la cession en fonction des tranches de notre traité pour le modèle issu de la Pareto généralisée.....	109
Tableau 25: Comparaison de la cession par tranches pour la charge dossier/dossier en N-I et N..	110
Tableau 26: Comparaison de la cession par tranches pour la charge issue de la modélisation historique de BPCE Assurances IARD en N-I et N	110
Tableau 27 : Comparaison de la charge cédée pour la modélisation issue de la loi de GPD en N-I et N.....	111
Tableau 28 : Comparaison de la charge cédée pour la modélisation issue de la loi marché et historique BPCE Assurances IARD en N-I et N.....	113

Bibliographie

- [1] – Anderson C. - O'Hagan A. - Tancredi A. - (2006), 'Accounting for threshold uncertainty in extreme value estimation', *Extremes*, Vol. 9, 87-106
- [2] – Aminat T. - (2014) - 'Modélisation par exposition des sinistres extrêmes pour la branche Marine Cargo'.
- [3] - PARTRAT C. et BESSON J-L - (2005) 'Assurance Non-Vie Modélisation', Economia.
- [4] – Mathieu P. (2012) - 'Analyse des solutions actuarielles en tarification des traités de réassurance non-proportionnels Non-Vie'. Mémoire d'Actuariat.
- [5] – Antoine P. (2021), 'Modélisation de la sévérité des traités en excédent de sinistre, approche par la théorie des valeurs extrêmes', Mémoire d'Actuariat.
- [6] – Duc Hien V. – (2016), 'Modèle de provisionnement des sinistres graves et son allocation économique aux différentes succursales d'AXA Corporate Solutions', Mémoire d'actuariat.
- [7] – Jacques B. – Christian P. – (2003) 'La réassurance : Approche technique', ED-Economia.
- [8] – Payre B. – (2012), 'Charge ultime nette de réassurance en RC corporelle : 2 modèles stochastiques pour les flottes automobiles.'
- [10] – Anis B. – (2010), 'Statistiques des valeurs extrêmes dans le cas de lois discrètes', Cours ESSEC.