
Modélisation de la distribution de la sévérité en réassurance

*Mémoire de Master 2
en vue de l'obtention du diplôme d'actuariat*

Institut de Science Financière et d'Assurances

Année académique 2019-2020



Auteur
Tuteur entreprise
Tuteur académique

Biber Léo
Loljeeh Ramachandra
Loisel Stéphane



Département Analytics
QBE Re

Mémoire présenté le : 25 mai 2021

**pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA
et l'admission à l'Institut des Actuaires**

Par : Biber Léo

Titre : Modélisation de la distribution de la sévérité en réassurance

Confidentialité : NON (Durée : 1 an 2 ans)

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

*Membres présents du jury de Signature
l'Institut des Actuaires*

B. BALTESAR

B. POTENTIER

L. LAURENT

*Membres présents du jury de
l'ISFA*

S. LOISEL

E. MASIELLO

.....

Entreprise

Nom : QBE Re

Signature :

*Directeur de mémoire en entre-
prise*

Nom : Loljeeh Ramachandra

Signature :



Invité :

Nom :

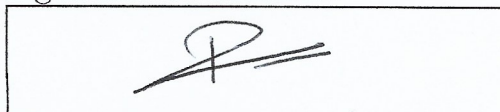
Signature :

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

Signature du responsable entreprise



Signature du candidat



Résumé

Mots-clés : Réassurance, Intervalles, Sévérité, Censure, Troncature, Tail, Body, Erlangs, Maximum de Vraisemblance, Mean-Excess, Algorithme EM

Il est d'usage en réassurance, de séparer en deux la distribution de la sévérité des sinistres : la partie "Tail" et la partie "Body". Pour la Tail, afin d'estimer une loi pouvant représenter ces sinistres, nous utilisons la théorie des valeurs extrêmes qui est une théorie asymptotique, et donc par définition, avant le seuil à partir duquel on peut appliquer la distribution, il faut avoir recours à d'autres méthodes. Les méthodes pour représenter le Body ne sont pas nombreuses et traditionnellement, on conserve une distribution empirique pour représenter les sinistres appartenant au Body. De plus, les outils d'estimation des paramètres de la loi de la Tail ne sont pas très nombreux et pas toujours très efficaces.

L'idée de ce mémoire est d'étudier d'autres lois possibles afin de modéliser la sévérité des sinistres, tout en conservant et en adaptant les outils actuels utilisés.

En réassurance, nous sommes confrontés au manque cruel de données, et finalement les prévisions des sinistres ne sont représentées que par des points fixes, c'est pourquoi, nous tâcherons de chercher à modéliser la distribution de la sévérité dans le cas où l'on posséderait des sinistres représentés par des intervalles et représentant possiblement une meilleure estimation de la réelle valeur des sinistres.

Abstract

Keywords : Reinsurance, Intervals, Severity, Censure, Truncation, Truncation, Tail, Body, Erlangs, Maximum Likelihood, Mean-Excess, EM Algorithm

In reinsurance, it is customary to separate the distribution of claims severity into two parts : the "Tail" and the "Body". For the Tail, in order to estimate a law that can represent these claims, we use the extreme value theory which is an asymptotic theory. So, before the threshold where the distribution can be separated, other methods must be used. There are not many methods to represent the Body and traditionally an empirical distribution is used to represent the claims belonging to the Body. Moreover, the tools for estimating the parameters of the Tail law are not very numerous and not always very efficient.

The idea of this thesis is to study other possible laws in order to model the severity of claims, while preserving and adapting the current tools used.

In reinsurance, we are confronted with a lack of data, and claims forecasts are only represented by fixed points, which is why we will try to model the distribution of severity in the event that we have claims represented by intervals and possibly representing a better estimate of the real value of the claims.

Remerciements

Ce mémoire n'aurait pas vu le jour sans l'aide de nombreuses personnes que j'aimerais modestement remercier en ces quelques lignes. Tout d'abord, un grand merci à mon tuteur professionnel, Indra Loljeeh pour son écoute, sa pédagogie, sa disponibilité ainsi que l'aide qu'il m'a apportée durant toute la durée de mon stage. Merci également à Alexandre Masquelein pour ses précieux conseils.

Je tiens à remercier Aurélien Dubois de m'avoir permis d'être accueilli chaleureusement au sein du département Analytics de QBE, où dès les premiers instants, j'ai eu la chance de faire partie intégralement de cette belle équipe.

Je tiens à remercier tout le corps enseignant de l'I.S.F.A. pour ces trois années d'encadrement, ainsi que mon tuteur académique Stéphane Loisel.

Ce mémoire constitue probablement la fin de mes années d'études et de ma vie étudiante, je tiens à remercier ma famille, mes parents et mon frère, de leur soutien toujours infaillible et inconditionnel ; sans qui tous mes projets n'auraient jamais vu le jour.

Table des matières

1	Introduction à la Réassurance	7
1.1	Présentation générale	7
1.2	Les contrats de Réassurance	7
1.2.1	Réassurance proportionnelle	8
1.2.2	Réassurance non-proportionnelle	9
1.3	Méthodes de tarification	11
1.3.1	Tarification à l'expérience	11
1.3.2	Tarification modèle collectif	12
1.3.3	Tarification à l'exposition	14
1.3.4	Autres méthodes de tarification	15
2	Modélisation classique de la sévérité en réassurance	16
2.1	Modélisation classique de la Tail	16
2.1.1	Generalized Pareto Distribution	16
2.1.2	QQ-plots	17
2.1.3	GQ plot	18
2.1.4	Mean-Excess plot	19
2.1.5	Estimateurs de ξ	20
2.2	Modélisation classique du Body	23
2.3	En pratique	24
2.4	But du mémoire	27
3	Modélisation de la Tail de la distribution de sévérité	28
3.1	Étude de nouvelles lois	28
3.1.1	Loi de Pareto	29
3.1.2	Loi Exponentielle	31
3.1.3	Loi de Weibull	34
3.1.4	Conclusion	37
3.2	Estimation pour des sinistres à valeur non-fixe	37
3.2.1	La fonction de vraisemblance	38
3.2.2	Cas avec des poids discrets	39
3.2.3	Cas avec des poids continus	40
3.2.4	Biais des estimateurs	43
3.2.5	Adaptation du Alpha-Plot et Lambda-Plot	47
3.2.6	Adaptation du Mean-Excess Plot	49
3.2.7	Comparaison des prix pour un contrat Excess of loss (XS)	50
3.2.8	Conclusion	51
4	Modélisation du Body de la distribution de sévérité	52
4.1	Principe et base de donnée	52
4.2	Les mélanges d'Erlangs	52
4.3	L'algorithme de maximisation de l'espérance (EM)	54
4.4	Ajustement avec différents nombres d'Erlangs dans le mélange	56
4.4.1	Temps d'exécution	58
4.4.2	Paramètres obtenus	58
4.4.3	Espérance et Variance	61
4.5	Tests de sensibilité	62
4.5.1	Test de sensibilité aux valeurs extrêmes	62
4.5.2	Test de sensibilité générale	64
4.6	Coût moyen	66
4.7	Critères d'information	68
4.8	Méthode de discrimination	69
4.9	Conclusion	70

5	Mise en application sur des données réelles	71
5.1	Base de données	71
5.2	Séparation de la Tail et du Body	73
5.2.1	Situation originelle avec des sinistres à valeur fixe	73
5.2.2	Nouvelle situation avec des sinistres à valeurs dans un intervalle de possibilités	74
5.3	Modélisation de la Tail	76
5.3.1	Situation originelle avec des sinistres à valeur fixe	76
5.3.2	Nouvelle situation avec des sinistres à valeurs dans un intervalle de possibilités	76
5.4	Modélisation du Body	77
5.4.1	Situation originelle avec des sinistres à valeur fixe	77
5.4.2	Nouvelle situation avec des sinistres à valeurs dans un intervalle de possibilités	78
5.5	Impact tarifaire	82
5.5.1	Modélisation globale retenue	82
5.5.2	Prix Excess of Loss (XS)	83
6	Conclusion	87
7	Bibliographie	89
8	Annexe	90

1 Introduction à la Réassurance

Tout d’abord, dans cette partie 1, nous allons faire une présentation générale de la réassurance afin de bien comprendre tous les enjeux de ce secteur de l’assurance. Puis en partie 2, nous verrons comment en théorie et en pratique la sévérité des sinistres est modélisée en réassurance. Nous verrons que cette modélisation présente de nombreux inconvénients, auxquels ce mémoire va tenter d’apporter des solutions alternatives. Ainsi, dans un premier temps, on étudiera comment représenter théoriquement la sévérité des sinistres en séparant cette sévérité en deux parties (Tail et Body de la distribution de sévérité des sinistres) afin de les étudier indépendamment. Puis dans un second temps, nous mettrons en application les solutions trouvées pour modéliser cette sévérité des sinistres, sur une base de données développée au sein de QBE dans le mémoire de MOURIER [2020]¹.

1.1 Présentation générale

Dans certains cas, les sociétés d’assurance ont comme les particuliers un besoin de s’assurer. Pour cela, elles doivent faire appel à un spécialiste : le réassureur. On a tendance à introduire la réassurance comme l’assureur des compagnies d’assurances, et les sociétés d’assurance faisant appel à un ou plusieurs réassureurs sont appelées « cédantes ».

La principale raison de faire appel à un réassureur est de se protéger contre la survenance de très grands sinistres. Mais pour divers autres arguments, ces cédantes vont également faire appel à un réassureur : pouvoir souscrire avec des limites plus grandes, s’adapter aux normes concernant les capitaux propres (Solvency II) ou encore équilibrer leur résultat. En échange d’une prime de réassurance, les compagnies d’assurance vont céder une partie de leur risque dont ils veulent s’alléger. Cette cession de la part des cédantes leur permet d’avoir plus de libertés et pouvoir par exemple tester l’entrée en gamme de nouveaux produits.

Les réassureurs ont donc un positionnement clé dans le secteur de l’assurance, en effet, ce sont eux qui supportent la majorité des grands risques. Historiquement, ces très grands risques sont principalement les catastrophes naturelles et les incendies, mais l’on voit de plus en plus de nouveaux risques encore peu étudiés : les attaques terroristes ou encore les pandémies. Par définition, ces événements étant inhabituels, nous avons peu de données. Une des problématiques principales de la réassurance est ce manque de données qui conduit les réassureurs à l’élaboration de modèles statistiques poussés afin de pallier ce manque.

QBE Re, entité de réassurance du groupe d’assurance QBE (Queensland Bankers Equity), est implantée à New-York, Londres, Dublin, Bruxelles, ainsi qu’aux Bermudes. L’équipe basée à Bruxelles est composée d’un effectif de 90 personnes, divisé en trois départements : Technique non-vie, Technique vie, et le département Analytics, dédié aux problématiques de Recherche et Développement. Les trois départements travaillent en étroite collaboration, afin d’élaborer une stratégie cohérente sur les différents marchés où QBE Re Brussels exerce son activité, à savoir l’Europe et l’Amérique. Ses principales branches sont la responsabilité Civile, le Marine ainsi que l’Incendie.

1.2 Les contrats de Réassurance

Qu’elle soit facultative (couvrant un ou plusieurs risques bien spécifiques) ou bien obligatoire (couvrant le portefeuille dans une sous-branche d’activité), il existe deux types de contrats en réassurance : la réassurance proportionnelle et la réassurance non-proportionnelle. Nous allons voir une présentation succincte de ces contrats de réassurance, pour plus de précisions, le lecteur pourra consulter les ouvrages de WALHIN [2012] et DELACROIX [2019]².

1. MOURIER [2020] - cf Bibliographie [1]

2. WALHIN [2012] et DELACROIX [2019] - cf Bibliographie [2] et [3]

1.2.1 Réassurance proportionnelle

Comme l'indique son appellation, la réassurance proportionnelle repose sur la proportionnalité des engagements entre cédante et réassureur. Pour un portefeuille donné, la cédante reçoit des primes de la part de ses assurés, et elle va en céder un pourcentage au réassureur en échange du même pourcentage de sa sinistralité. Dans cette famille, on retrouve deux principaux types de contrat : le Quote-part et l'Excédent de plein.

Le Quote-part (Quota-share)

Pour ce type de contrat, on fixe un taux de cession à l'avance. Notons τ ce taux de cession pour lequel la cédante s'engage donc à reverser au réassureur $\tau\%$ des primes de ses assurés. En contrepartie, le réassureur va devoir reverser à la cédante $\tau\%$ des sinistres qui vont survenir. Nous avons une homogénéité de la cédante pour ce type de contrat, c'est-à-dire que le taux de cession reste inchangé peu importe le risque et la somme assurée.

Comme de nombreux sinistres, petits et grands, vont se produire au cours d'une année, le traité Quote-part prévoit une clause indiquant la fréquence des relevés comptables. Ainsi, pour chaque période convenue, la cédante donnera son résultat technique pour cette période :

$$\text{Résultat technique} = \text{Primes} - \text{Sinistres} - \text{Frais} - \text{Courtage} - \text{Commission}$$

En fonction des sinistres survenus la cédante ou le réassureur versera à l'autre ce qu'il lui doit. Si le résultat est positif, la cédante versera $\tau\%$ du résultat, et s'il est négatif, c'est au réassureur de lui verser ce pourcentage.

Prenons un taux de cession de 70 % et des périodes de règlement de 3 mois sur une année, alors si les primes perçues par la cédante sont de 2 M par période et que les sinistres en cumulés valent pour chaque période respectivement : 3.5 M, 1.8 M, 0.8 M et 3 M ; alors le partage se fera de la manière suivante :

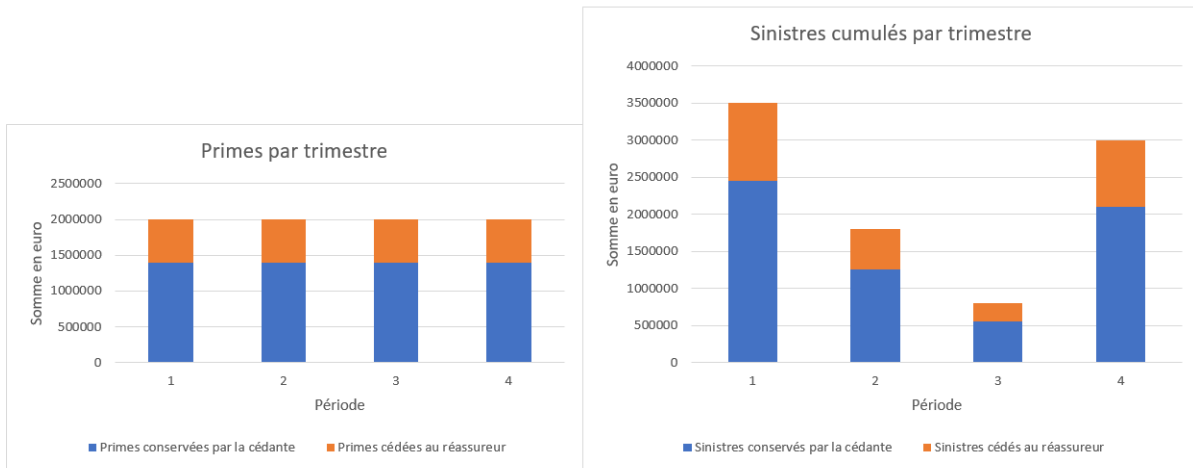


FIGURE 1 – Quote-part

L'Excédent de plein (Surplus)

Pour l'Excédent de plein, il est établi la somme maximale que la cédante ne veut pas dépasser pour chaque sinistre, cette somme est appelée « plein de rétention », on note :

- C : la capacité maximale de l'Excédent de plein
- P : la prime du risque
- R : le plein de rétention
- K : la valeur assurée
- τ : le taux de cession

On définit le taux de cession pour le risque i comme : $\tau_i = \min\left(\max\left(\frac{K_i - R}{K_i}; 0\right); \frac{C}{K_i}\right)$.

Pour chaque risque, si la valeur assurée est inférieure au plein de rétention, le réassureur n'aura pas à intervenir. Ainsi pour les petits risques qui constituent la majorité du portefeuille, la cédante conserve tous ces risques. Concernant les grands risques, plus le risque a une somme assurée grande, plus la cession est grande ; on retrouve ainsi la notion d'homogénéité. Ce type de traité permet à la cédante de garder la majorité de ses primes perçue, d'avoir une protection en cas de survenance d'un grand risque et de payer au maximum le plein de rétention pour ces risques. Illustrons cela, prenons un plein de rétention de 100 et 4 risques différents, pour lesquels nous avons respectivement :

- Une somme assurée de 200, 150, 100 et 50
- Des primes de 70, 50, 30 et 20
- Une sinistralité de 100, 100, 100 et 100

Alors le partage se fera de la manière suivante :

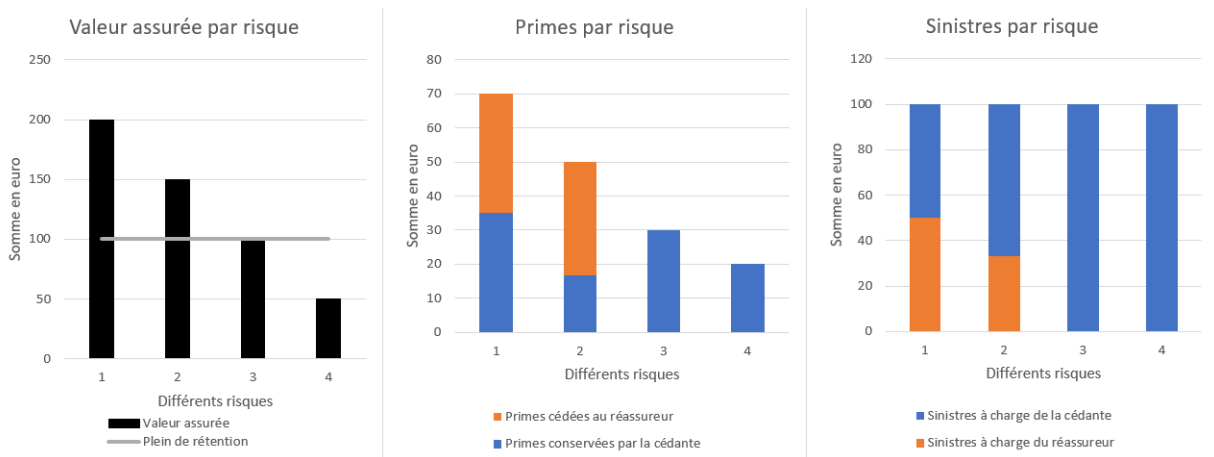


FIGURE 2 – Excédent de plein

Pour les risques 3 et 4, la valeur assurée étant en dessous du plein de rétention, le réassureur n'intervient donc pas. En revanche pour les risques 1 et 2, la valeur assurée dépassant le plein de rétention, on calcule le taux de cession et une partie des primes de la cédante est reversée au réassureur. En contrepartie, le réassureur prend en charge une partie des risques 1 et 2.

1.2.2 Réassurance non-proportionnelle

En réassurance non-proportionnelle, un seuil est fixé pour définir le montant à partir duquel le réassureur va intervenir et un autre seuil pour établir sa limite d'intervention. Si le montant du sinistre est en dessous du premier seuil, appelé priorité alors la cédante prend tout en charge. Mais si ce montant dépasse le deuxième seuil, appelé limite, alors le réassureur paiera au maximum la couverture, qui est la différence entre les deux

seuils. Les sorts de la cédante et du réassureur sont moins liés que dans le cas de la réassurance proportionnelle, en effet ici, une perte pour une cédante n'implique pas une perte pour le réassureur, ce qui n'est pas le cas pour un Quote-part ou un Surplus.

L'Excédent de sinistre (Excess of loss)

Un contrat Excédent de sinistre est communément noté C XS P avec C la couverture et P la priorité, et on le note XS ou XL (pour Excess of loss). Il s'agit des deux limites exposées précédemment. Ainsi, une sinistralité en dessous de la priorité n'engagera que la cédante, une sinistralité comprise entre la priorité et la limite (définie comme $L=C+P$) engagera la cédante à hauteur de P puis le réassureur à partir de P, et une sinistralité dépassant la limite engagera la cédante à hauteur de P plus tout ce qui dépasse la limite et engagera le réassureur à hauteur de C. Mathématiquement, en notant X la sinistralité, la charge du réassureur sera de : $\min(C; \max(0; X-P))$. Comme les engagements ne sont plus proportionnels, le réassureur conviendra d'une prime via la méthode de tarification de son choix. Illustrons cela avec un contrat 50 XS 100 avec 4 sinistralités différentes : 175, 100, 125 et 50.

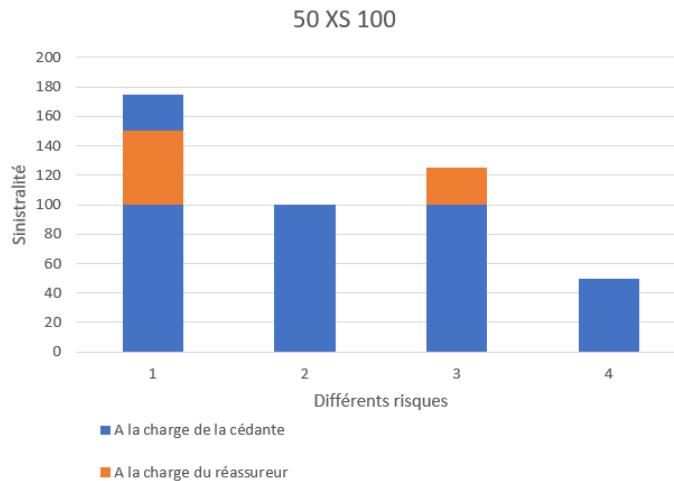


FIGURE 3 – Excédent de sinistre

En pratique, pour le risque 1, si la cédante pense que la sinistralité va être importante, et qu'elle veut bien se couvrir pour les grands risques, elle va mettre en place un deuxième Excédent de sinistre à partir du seuil 150 (à la limite du premier XS); ou bien elle peut souhaiter un XS avec une couverture plus grande.

L'Excédent de perte (Stop-loss)

Il y a très peu de différences conceptuelles entre l'Excédent de sinistre et l'Excédent de perte. Ici, il faut prendre en compte la sinistralité totale, les sinistres sont agrégés sur l'année correspondante. Reprenons nos précédents risques, et un Excédent de perte de 250 XS 250. La sinistralité cumulée est donc : 175, 275, 400 et 450, et le partage des pertes sera fait de la manière suivante :

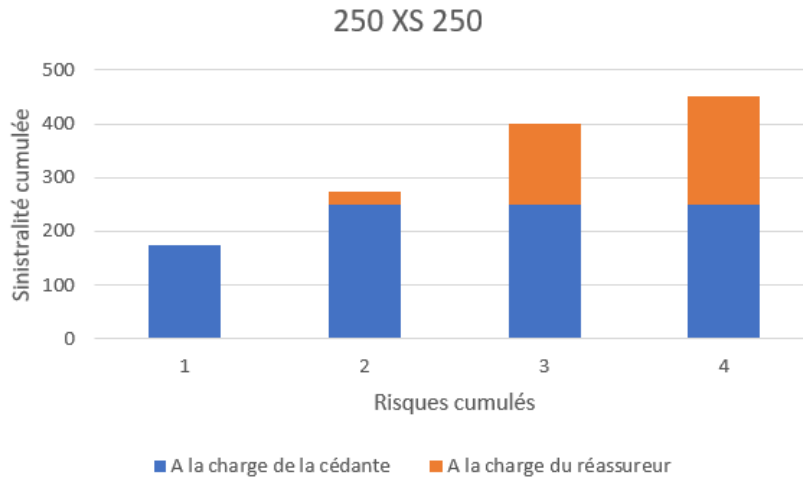


FIGURE 4 – Excédent de perte

De manière analogue à l'Excédent de sinistres, la cédante pourra demander au réassureur de faire co-exister plusieurs tranches.

1.3 Méthodes de tarification

Pour des contrats de réassurance proportionnelle, la tarification repose sur la proportionnalité des engagements entre cédante et réassureur. De son côté, l'assureur reversera au réassureur un pourcentage des primes qu'il a reçu de ses clients en échange du même pourcentage préalablement payé par le réassureur établi sur les pertes de l'assureur.

Cependant, la tarification des traités non-proportionnels n'est pas similaire à celle effectuée pour les traités proportionnels, l'existence d'une franchise et d'une couverture rend la tarification différente. Afin de tarifier ces différents contrats, il existe trois principales méthodes de tarification : tarification à l'expérience, le modèle collectif et la tarification à l'exposition.

1.3.1 Tarification à l'expérience

La tarification à l'expérience est la méthode la plus historique et intuitive. Elle repose sur l'historique des sinistres de la cédante. Ainsi comme toute méthode empirique, elle n'est robuste qu'à condition d'avoir assez de données ; et suppose forcément que l'évolution historique de la cédante soit représentative de l'évolution future des sinistres de la cédante. Pour cette méthode, il faudra :

- Posséder un portefeuille qui s'étale sur suffisamment d'années.
- Avoir un grand nombre de sinistres.
- Actualiser les sinistres et les primes avec un indice (inflation, salaires...).
- Regarder le "Burning Cost" moyen pour chaque tranche à tarifier (voir "Burning Cost" ci-dessous).
- Établir les charges du réassureur.
- Estimer la marge commerciale voulue.

Prenons un traité Excess of loss, de priorité P et de couverture C, et notons :

- P_i la prime de l'année $i \in \{1, \dots, n\}$ reçu par la cédante de ses clients
- C_{j,k_j} le k ème sinistre de l'année $j \in \{1, \dots, n\}$

Le Burning Cost (BC) à l'expérience est défini pour une tranche de réassurance comme le ratio entre la sinistralité dans la tranche et la prime.

$$BC_{\text{expérience}} = \frac{\text{Sinistralité}}{\text{Prime}} = \frac{\sum_{i=1}^n \sum_{k_i=1}^{k_n} \min(C, \max(C_{i,k_i} - P, 0))}{\sum_{i=0}^n P_i}$$

Le problème de cette méthode, en plus du fait qu'elle perde très fortement en robustesse lorsque le nombre de données décroît ; c'est qu'elle ne tient pas compte des modifications du portefeuille. De plus pour une tranche qui n'est pas touchée, le Burning Cost sera nul alors qu'on ne peut pas envisager un prix nul même si la priorité de la tranche se trouve être très élevée.

1.3.2 Tarification modèle collectif

Nous avons vu avec la méthode Burning Cost à l'expérience que deux problèmes majeurs étaient le nombre de sinistres dans un portefeuille ainsi que l'estimation d'un prix dans une tranche qui ne serait pas touchée empiriquement. Afin de corriger ces points, on peut envisager un modèle collectif qui se présente sous la forme suivante :

- Soit la suite $(X_i)_{i \geq 0}$ des variables aléatoires réelles représentant les montants individuels de sinistre par ordre de survenance.
- Soit N la variable aléatoire $\in \mathbb{N}$ représentant le nombre de sinistres.
- Soit S la charge totale des sinistres.

$$S = \sum_{i=0}^N X_i$$

Trois hypothèses doivent être faites pour rendre l'utilisation de ce modèle très pratique :

- Les variables $(X_i)_{i \geq 1}$ et N sont supposées indépendantes (H1)
- Les $(X_i)_{i \geq 1}$ sont indépendants et identiquement distribués (H2)
- $X_0 = 0$ (H3)

L'hypothèse 1, représente l'indépendance entre la fréquence et la sévérité des sinistres ; l'hypothèse 2 représente l'indépendance entre chaque sinistre (peu importe l'année de survenance, ils ne sont pas corrélés) ; et enfin l'hypothèse 3 représente le cas où il ne surviendrait aucun sinistre. Grâce à ces hypothèses, nous avons l'espérance de la charge totale équivalant au produit des espérances entre la sévérité et la fréquence des sinistres. Une modélisation séparée entre fréquence et sévérité peut être ainsi envisagée. Historiquement, en réassurance, il est utilisé un modèle Poisson pour la fréquence et Pareto pour la sévérité.

Afin de bien comprendre la structure de la chaîne de tarification en réassurance, voici un diagramme reprenant les principales étapes préalables à l'estimation de la distribution de la fréquence et de la sévérité des sinistres :

Les "inputs" sont :

- Les sinistres sur base des années fournies par la cédante.
- Les indices d'indexation (inflation, salaire, construction...).
- Les informations fournies par la cédante (selon la branche d'activités à tarifier).
- Le programme de réassurance à tarifier.

Quelques explications sur le prochain graphique :

- Retained amin : chaque année, le réassureur reçoit une liste de sinistres généralement supérieurs à environ 50 % d'un seuil établi par la cédante, "l'Amin" permet d'établir le seuil à partir duquel on possède toute l'information.
- Les coefficients IBNER : les coefficients de passage d'une année sur l'autre pour la fréquence des sinistres.
- Threshold : le seuil final retenu pour l'étude globale.
- Ultimates : la valeur des montants des sinistres après indexation, extrapolation et prise en compte du threshold.
- Payment pattern : les cadences de règlement pour chaque année.

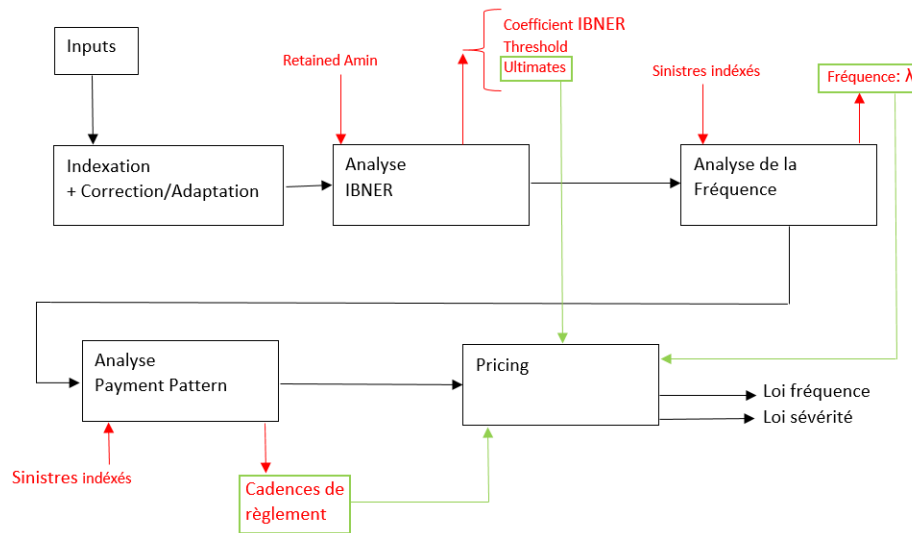


FIGURE 5 – Articulation de la chaîne de tarification

Les cinq étapes principales de cette chaîne de tarification sont :

1. L'indexation : les sinistres sont mis en "As-if", c'est-à-dire que l'on indexe les sinistres vers le futur avec un indice représentatif de cette évolution dans le temps (inflation, salaires...). Il s'agit de la valeur des sinistres étant survenus dans le passé, si jamais ils s'étaient produits aujourd'hui.
2. L'analyse IBNER : on étudie les coefficients de passage d'une année sur l'autre pour le développement des sinistres.
3. L'analyse de la fréquence : on étudie la fréquence des sinistres au-delà de certains seuils, choisis lors de l'étape 2.
4. L'analyse des "Payment Pattern" : on étudie à quelle fréquence les sinistres vont être réglés par les cédantes.
5. Le "pricing" : il s'agit de l'étape finale, durant laquelle à partir des lois définies précédemment, on détermine le prix de la couverture.

C'est sur les valeurs des "Ultimates" que ce mémoire se focalisera, ce qui correspond à l'étape 5 de cette chaîne de tarification. Ces valeurs ont été développées dans le mémoire de MOURIER [2020]³, lors de l'épate numéro 2, nous essayerons dans ce mémoire, d'en modéliser la distribution.

Après toutes ces étapes, nous avons notre distribution en loi de la sévérité ainsi que de la fréquence des sinistres. Afin d'obtenir un prix commercial, il ne reste plus qu'à prendre en compte les charges du réassureur, les cadences de règlement et toutes les marges commerciales à ajouter à nos primes pures déterminées préalablement.

1.3.3 Tarification à l'exposition

Nous allons présenter cette méthode de manière succincte, un lecteur intéressé pourra se référer aux mémoires réalisés au sein de QBE Re⁴, par LOLJEEH [2012] et GILLIER [2018] traitant de la tarification à l'exposition.

La méthode de tarification à l'exposition a été développée afin de :

- Combler l'éventuel manque de données, rendant la tarification à l'expérience inenvisageable.
- Utiliser les informations données par la cédante concernant le portefeuille.

Cette méthode est très simple à utiliser une fois les courbes d'exposition connues. Afin de construire une courbe d'exposition, il faut introduire le taux de destruction qui correspond au ratio entre le coût d'un sinistre et la somme assurée correspondante. En notant X le coût du sinistre et M la somme assurée, le taux de destruction par définition vérifie :

$$0 \leq \tau = \frac{X}{M} \leq 1$$

La courbe d'exposition est définie par une fonction, qui est le rapport entre le coût moyen des sinistres pour une tranche illimitée partant de P et une tranche illimitée partant de 0. Cette fonction, que l'on notera G , est exprimée en fonction de notre priorité (P) normalisée; ainsi :

$$G(u) = \frac{\mathbb{E}(N)\mathbb{E}(\min(X, P))}{\mathbb{E}(N)\mathbb{E}(X)} = \frac{\mathbb{E}(\min(\tau, u))}{\mathbb{E}(\min((\tau, 1))}$$

Avec N la variable représentant le nombre de sinistres, qui disparaît complètement de l'expression. Au numérateur et au dénominateur nous avons deux coûts moyens, c'est pourquoi G ne dépend donc pas de la fréquence des sinistres ni de la somme assurée M , et est donc uniquement lié à la distribution de sévérité. Si l'on note F_τ la fonction de répartition du taux de destruction, alors :

$$G(u) = \frac{\int_0^u (1 - F_\tau(t)) dt}{\int_0^1 (1 - F_\tau(t)) dt} = \frac{\int_0^u (1 - F_\tau(t)) dt}{\mathbb{E}(\tau)}$$

3. MOURIER [2020] - cf Bibliographie [1]

4. LOLJEEH [2012] et GILLIER [2018] - cf Bibliographie [4] et [5]

Cette fonction G est caractérisée par : sa stricte croissance, sa concavité sur son domaine de définition $([0; 1])$, $G(0)=0$ et $G(1)=1$.

Pour la tarification d'un traité Excess of loss C vs P (avec $L= C+P$ la limite), en prenant en compte une unique année du portefeuille de la cédante, on aura donc avec les notations suivantes :

- Soit M le nombre de bandes dont on dispose de la somme assurée
- Soit ASI_k la somme assurée pour la k -ème bande
- Soit P_k la prime pure de la k -ème bande
- $\frac{L}{ASI_k}$ le degré de destruction minimum dans la bande correspondante
- $\frac{P}{ASI_k}$ le degré de destruction maximum dans la bande correspondante
- P_G la prime pure globale

$$P_G = \sum_{k=1}^M \left[G_k \left(\frac{L}{ASI_k} \right) - G_k \left(\frac{P}{ASI_k} \right) \right] * P_k$$

À noter, qu'en pratique, on dispose de courbe d'exposition seulement pour certains risques (risques industriels, petites valeurs assurées...). Il n'est pas possible d'avoir de manière triviale une courbe d'exposition pour chaque valeur assurée. Même s'il ne s'agit pas d'une méthode purement empirique, elle est très sensible au nombre de sinistres constituant le portefeuille. De plus, la supposition de l'homogénéité au sein d'une bande donnée peut s'avérer une hypothèse très forte selon la taille des bandes fournies par les cédantes.

1.3.4 Autres méthodes de tarification

Comme vu précédemment, les méthodes exposées ont des avantages, mais également des inconvénients. La méthode à l'exposition calibrée est une combinaison entre la tarification à l'expérience et celle à l'exposition. La principale idée est de conserver la tarification à l'expérience lorsque l'on a assez d'observations dans une tranche de tarification, on appelle ces bandes les « tranches travaillantes », puis ensuite d'utiliser la tarification à l'exposition lorsque l'on manque de données. Dans chaque cas, il faudra regarder le ratio « BC à l'expérience » / « BC à l'exposition » et ajuster le modèle avec ce ratio.

Une autre méthode est une « méthode combinée », développée en interne au sein de la compagnie QBE Re, est une version améliorée de la méthode de tarification à l'exposition calibrée. Elle a été mise en place dans le but d'intégrer les informations données par la cédante. Le lecteur intéressé par cette méthode pourra se référer à l'ouvrage rédigé par l'équipe interne de QBE en 2006⁵.

5. DESMEDT et al. [2012] - cf Bibliographie [6]

2 Modélisation classique de la sévérité en réassurance

Grâce au modèle collectif, nous pouvons séparer l'analyse de la fréquence de celle de la sévérité. Dans un premier temps, nous allons présenter les différents outils nous permettant d'accéder à la distribution de la sévérité. Puis dans un second temps, nous regarderons les outils qu'en pratique les Analystes utilisent lors d'une tarification afin de choisir la meilleure distribution de sévérité en accord avec les observations fournies.

Tout d'abord, l'ensemble des sinistres peut être divisé en trois parties : les sinistres "attritionnels" (sinistre de faible intensité), les sinistres dits "larges", et les CAT (catastrophes naturelles). Les CAT font l'étude d'un traitement à part qui ne sera pas vu dans ce mémoire. Les sinistres "attritionnels" et "larges" n'ont a priori aucune raison d'avoir un comportement identique. De manière analogue, l'étude de la sévérité peut être divisée en deux parties :

- La Tail : il s'agit de la queue de la distribution, la partie pour laquelle nous n'avons pas assez de sinistres pour envisager une tarification sur base de l'empirique. Il faut donc trouver la meilleure distribution théorique possible qui colle avec nos quelques sinistres observés.
- Le Body : il s'agit du corps de la distribution, la partie dans laquelle la grosse majorité des sinistres se trouve, et pour laquelle une tarification sur base de l'empirique est envisageable.

2.1 Modélisation classique de la Tail

2.1.1 Generalized Pareto Distribution

La modélisation de la sévérité théorique pour la Tail est présentée à titre informatif et afin d'avoir une compréhension plus pointue des raisons qui ont motivé la proposition de modélisation de la sévérité dans ce mémoire. C'est pourquoi elle n'est pas entièrement détaillée dans ce mémoire, un lecteur intéressé par la modélisation de la sévérité via la théorie des valeurs extrêmes pourra se référer à l'article rédigé par ROBERT [2019]⁶.

La théorie des valeurs extrêmes nous suggère que la Tail d'une distribution de sévérité des sinistres peut être décrite par des distributions du type Generalized Pareto Distribution (GPD). Ces distributions remplissent les caractéristiques suivantes :

Soit $X \sim \text{GPD}(\mu, \sigma, \xi)$, avec $\mu \in \mathbb{R}$ le paramètre de position, $\sigma \in \mathbb{R}^{*+}$ le paramètre d'échelle et $\xi \in \mathbb{R}$ le paramètre de forme. La fonction de répartition de X est alors :

$$F_{(\mu, \sigma, \xi)} = \begin{cases} 1 - \left(1 - \frac{\xi(x - \mu)}{\sigma}\right)^{\frac{-1}{\xi}} & \text{pour } \xi \neq 0 \\ 1 - \exp\left(-\frac{x - \mu}{\sigma}\right) & \text{pour } \xi = 0 \end{cases} \quad \text{Où } \begin{cases} x \geq 0 & \text{si } \xi \geq 0 \\ \mu \leq x \leq \mu - \frac{\sigma}{\xi} & \text{si } \xi < 0 \end{cases}$$

6. ROBERT [2019] - cf Bibliographie [7]

Propriétés intéressantes

- Si $\xi = \mu = 0$ alors la distribution GPD est la distribution Exponentielle.
- Si $\xi > 0$ et $\mu = \frac{\sigma}{\xi}$ alors la distribution GPD est une Pareto de paramètre d'échelle $A = \frac{\sigma}{\xi}$ et de paramètre de forme $\alpha = \frac{1}{\xi}$.
- $Y \sim G_{\xi, \sigma, 0} \Leftrightarrow Y + \mu \sim G_{\xi, \sigma, \mu}$ où μ est le paramètre de position.
- Et donc $X - A \mid X > A \sim G_{\xi, \sigma, \mu} \Leftrightarrow X \mid X > A \sim G_{\xi, \sigma, \mu + A}$; très utile dans notre cas, puisque la GPD à trois paramètres va pouvoir être utilisée pour modéliser la distribution des montants des pertes supérieures à un seuil fixé.

ξ est aussi appelé l'indice de queue, c'est lui qui va nous donner des informations sur l'épaisseur de la queue de la distribution.

- $\xi > 0$ correspond à une queue épaisse; exemples : loi Cauchy, Pareto, Student.
- $\xi = 0$ correspond à une queue intermédiaire; exemples : loi Gamma, Weibull, Normale, Log-Normale.
- $\xi < 0$ correspond à une queue fine; exemples : loi Uniforme, Beta.

Sur base d'un historique de sinistres, plusieurs outils vont nous permettre de déterminer les trois paramètres μ, σ et ξ , les plus aptes à représenter nos données. Ils sont présentés dans les parties suivantes.

2.1.2 QQ-plots

Le diagramme quantile-quantile (QQ plot) permet d'estimer si la distribution théorique envisagée pour notre modèle décrit bien la distribution empirique observée. Il est ici utilisé pour déterminer la finesse de la queue de la distribution. En abscisse, nous allons avoir les quantiles empiriques de la distribution observée alors qu'en ordonnée, nous allons faire correspondre les quantiles de la loi théorique supposée. Si jamais les points sont bien alignés sur une droite alors la queue empirique se comporte comme la queue de la loi théorique testée. De plus, si la forme du graphique est concave alors la queue empirique est plus lourde que la queue théorique. Et inversement, si la forme est convexe alors la queue empirique est plus fine que la queue théorique.

Notons U la fonction quantile d'ordre $x : U(x) = F^{-1}(1 - \frac{1}{x}) = \inf\{y > 0 : F(y) \geq 1 - \frac{1}{x}\}$ avec $x > 1$.

Comme vu précédemment, le diagramme QQ-plot va nous permettre de deviner la finesse de la queue. Le but étant de se faire un avis sur le poids de la queue de distribution, le signe de ξ sera analysé lors de la prochaine étape.

L'idée pour les lois théoriques à tester est de prendre une loi avec $\xi > 0$ pour cela nous prendrons la loi de Pareto; et une loi avec $\xi = 0$, pour cela nous prendrons la loi Exponentielle. De plus, nous testerons également la loi de Weibull ($\xi = 0$) en complément de la loi Exponentielle. Cela nous permettra d'identifier les cas où la distribution est plus lourde que la distribution exponentielle sans être de type Pareto.

Par définition du diagramme QQ-plot, on trace donc pour les différentes lois choisies, les points suivants :

Soit (X_1, \dots, X_n) , les n sinistres observés, triés par ordre croissant.

Exponentiel quantile plot : $\left(-\log\left(\frac{j}{n+1}\right), X_{n-j+1} \right)$ pour $j=1, \dots, n$

Pareto quantile plot : $\left(-\log\left(\frac{j}{n+1}\right), \log(X_{n-j+1}) \right)$ pour $j=1, \dots, n$

Weibull quantile plot : $\left(\log\left(-\log\left(\frac{j}{n+1}\right)\right), \log(X_{n-j+1}) \right)$ pour $j=1, \dots, n$

Pour illustrer cela, simulons une loi Normale et regardons les QQ-plot obtenus pour nos trois lois théoriques envisagées :

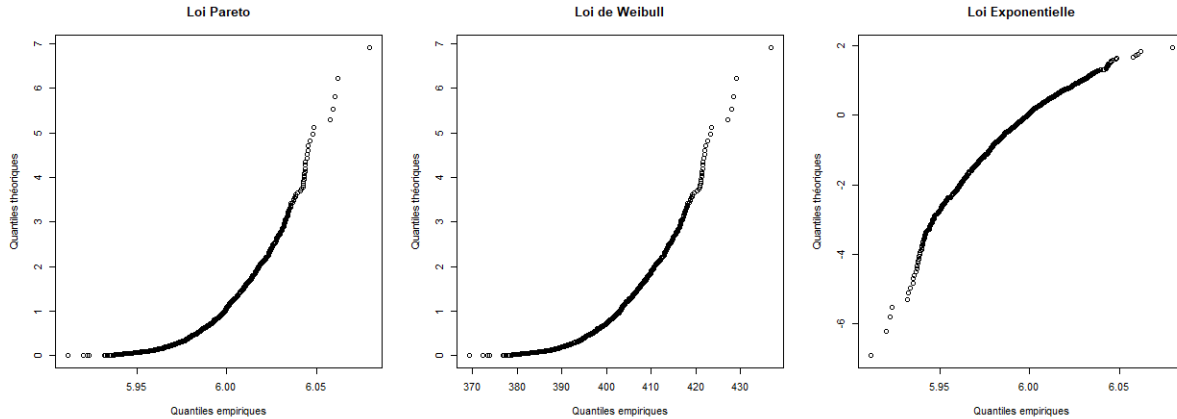


FIGURE 6 – QQ plot

Il apparaît que le QQ-plot de la loi Pareto est convexe, celui de la loi de Weibull est également convexe, mais que celui de la loi Exponentielle est concave. On peut donc conclure que la distribution empirique (ici une loi Normale) possède une queue plus fine que la loi de Pareto (c'est bien le cas car $\xi_{pareto} > 0$ alors que $\xi_{normale}=0$). Et que la queue de la distribution empirique est plus épaisse que celle de la loi Exponentielle, mais plus fine que la loi de Weibull.

2.1.3 GQ plot

Le Generalized Quantile Plot (GQ) est très utile pour déterminer le signe de ξ . Comme précédemment, on compare les quantiles empiriques avec les quantiles théoriques, mais cette fois-ci issus d'une GPD. La pente dans la partie extrême du graphique indique si ξ est positif, nul ou négatif. C'est très facile à interpréter. En reprenant les notations précédentes :

Generalized quantile plot : $\left(-\log\left(\frac{j+1}{n+1}\right), \log(UH_j) \right)$ pour $j=1, \dots, n-1$

$$\text{Où } UH_j = X_{n-j} \left(\frac{1}{j} \sum_{k=1}^j \log(X_{n-k+1}) - \log(X_{n-j}) \right) \text{ pour } j=1, \dots, n-1$$

La suite $(UH_j)_{j=1, \dots, n-1}$ est utilisée comme approximations des valeurs moyennes empiriques des excès. La pente de la dernière "ligne" du graphique estimera ξ . Ainsi, si la pente est positive alors $\xi > 0$, si elle est nulle alors $\xi = 0$, et si elle est négative alors $\xi < 0$. Illustrons cela avec trois lois : Pareto, Exponentielle et Uniforme.

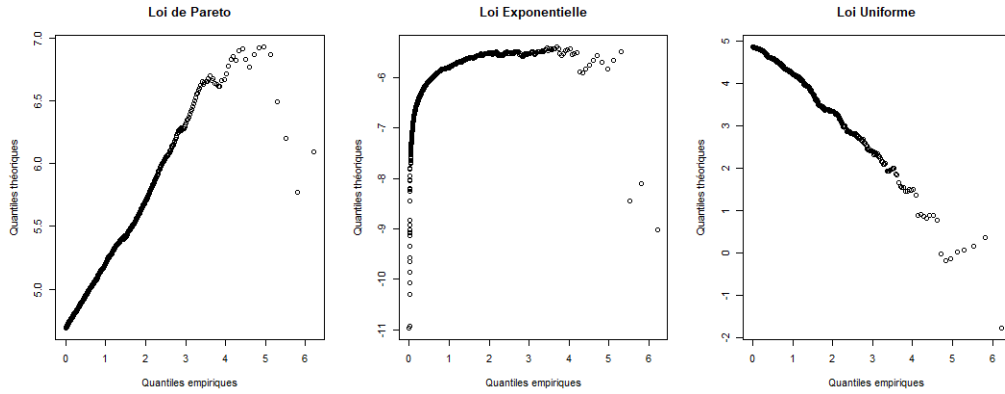


FIGURE 7 – Q-Q plot

Ce graphique illustre nos propos avec les trois lois : Pareto ($\xi > 0$), Exponentielle ($\xi = 0$) et Uniforme ($\xi < 0$) ; les pentes affichées correspondent bien avec le signe de l'indice de queue des lois.

2.1.4 Mean-Excess plot

Le "mean excess" (que l'on pourra noter : ME) est la fonction des excès moyens. La "mean excess fonction" que l'on notera e peut être définie pour $X > 0$ par :

$$e(x) = \mathbb{E}(X - x | X > x), \text{ pour } x \geq 0 \text{ et en supposant } \mathbb{E}(X) < \infty.$$

De plus, si $X \sim \text{GPD}_{(\sigma, \xi)}$ alors $X - x | X > x \sim \text{GPD}_{(\sigma + \xi x, \xi)}$ et l'on a :

$$\mathbb{E}(X - x | X > x) = \frac{\sigma + \xi x}{1 - \xi} = \frac{\sigma}{1 - \xi} + \frac{\xi}{1 - \xi} x$$

De cette expression affine, il est facile de déduire selon la pente de la "mean excess fonction", le signe de ξ .

Exemples

- Distribution de Pareto à partir de A , avec $\xi = \frac{1}{\alpha}$, on a pour $x > A$, $e(x) = \frac{x}{\alpha - 1}$; ainsi le ME augmente linéairement à partir du seuil A .
- Distribution Exponentielle à partir de A , avec $\xi = 0$, on a pour $x > A$, $e(x) = \frac{1}{\lambda}$; ainsi le ME est constant à partir du seuil A .

À noter que le ME devient ainsi très utile pour déterminer le seuil à partir duquel on a un comportement similaire à notre loi théorique envisagée pour modéliser la queue de la distribution empirique.

En reprenant les mêmes notations que précédemment, l'estimateur empirique du mean excess sera noté :

$$\hat{e}_n(x) = \frac{\sum_{i=1}^n X_i \mathbb{1}_{(X_i \geq x)}}{\sum_{i=1}^n \mathbb{1}_{(X_i \geq x)}} - x \text{ pour } x \geq 0$$

Bien qu'en pratique on prendra, pour $k=1, \dots, n-1$: $x = X_{n-k}$, soit $E_k = \hat{e}_n(X_{n-k}) = \frac{1}{k} \sum_{i=1}^k X_{n-i+1} - X_{n-k}$.
 Et afin d'avoir le Mean Excess Plot, on tracera les points $(X_{n-k}; E_k)$, pour $k=1, \dots, n-1$.

Par soucis de simplification et d'identification des potentielles lois, on comparera les lois par rapport à la loi Exponentielle (bien qu'en réalité, on puisse être plus léger qu'une loi Exponentielle; la racine d'Exponentielle par exemple; et avoir un ME également asymptotiquement constant) :

- À noter que E_k peut en réalité être interprété comme une estimation de la pente du diagramme quantile exponentiel à l'extrémité droite $\left(-\log\left(\frac{j}{n+1}\right), X_{n-j+1} \right)$ pour $j=1, \dots, n$.
 - Les distributions dont la queue est plus lourde que celle de la distribution exponentielle : le quantile exponentiel a une forme convexe pour les plus grandes observations \rightarrow Et le ME plot augmente.
 - Les distributions dont la queue est plus légère que celle de la distribution exponentielle : le quantile exponentiel a une forme concave pour les plus grandes observations \rightarrow Et le ME plot décroît.
 - Les distributions avec queue de type exponentielle : le quantile exponentiel est linéaire pour les plus grandes observations \rightarrow Et le ME est asymptotiquement constant.

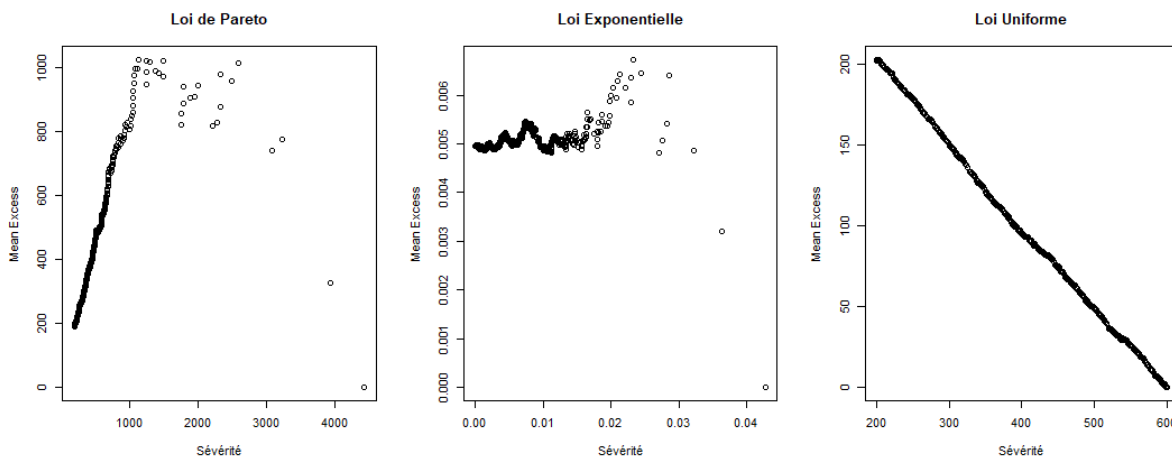


FIGURE 8 – ME plot

Ce graphique illustre nos propos avec les trois lois simulées : Pareto ($\xi > 0$), Exponentielle ($\xi = 0$) et Uniforme ($\xi < 0$) : les pentes affichées correspondent bien avec le poids de la queue de distribution de ces lois.

2.1.5 Estimateurs de ξ

Le but de cette partie est de présenter rapidement les différents estimateurs envisageables pour ξ , selon la distribution théorique intuitée à l'aide des outils présentés précédemment. Un lecteur intéressé par un meilleur détail de l'obtention de ces différents estimateurs pourra se référer à l'article rédigé par LADOU-CETTE [2011]⁷. Pour estimer ξ différentes approches sont ainsi possibles :

7. LADOU-CETTE [2011] - cf Bibliographie [8]

- Les méthodes POT (Pick Over Threshold), qui se préoccupent de la loi des excès en se basant sur les GPD. Pour cela, on fixe un seuil A à partir duquel on souhaite modéliser notre GPD. Dans ces méthodes, on y retrouve :

— **Le Maximum Likelihood Estimator (ML estimator)**

La log-vraisemblance pour Y_1, \dots, Y_{N_A} , issus d'une GPD $_{(\sigma, \xi = \frac{1}{\gamma})}$, avec N_A le nombre d'observations dépassant le seuil A, et n le nombre total d'observations, se note :

$$\text{Pour } \gamma \neq 0 \rightarrow \log(\mathcal{L}(\sigma, \gamma)) = -N_A \log(\sigma) - \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^{N_A} \log\left(1 + \gamma Y_i \sigma\right)$$

avec $1 + \frac{\gamma Y_i}{\sigma} > 0$ pour $i = 1, \dots, N_A$

$$\text{Pour } \gamma = 0 \rightarrow \log(\mathcal{L}(\sigma, 0)) = -N_A \log(\sigma) - \frac{1}{\sigma} \sum_{i=1}^{N_A} Y_i$$

Les estimateurs $\hat{\gamma}_{N_A}^{ML}$ et $\hat{\sigma}_{N_A}^{ML}$ se déduisent en dérivant la log-vraisemblance par rapport à chacun des paramètres et en résolvant $\frac{\partial \log(\mathcal{L}(\sigma, \gamma))}{\partial \gamma} = \frac{\partial \log(\mathcal{L}(\sigma, \gamma))}{\partial \sigma} = 0$.

Cet estimateur est asymptotiquement sans biais.

On pourrait penser que plus N_A est grand (en abaissant le seuil A par exemple), plus l'estimateur est robuste, sauf que la théorie des valeurs extrêmes étant asymptotique, plus l'on prend des données basses, plus l'on inclut des données qui n'appartiennent pas à la Tail, et l'on introduit donc un biais.

— **La méthode des moments (MOM)**

L'idée de ces estimateurs est de faire correspondre les moments théoriques avec les moments empiriques, ce qui revient au final à résoudre (pour $\gamma < \frac{1}{2}$) :

$$\hat{\gamma}_{N_A}^{MOM} = \frac{1}{2} \left(1 - \frac{\bar{Y}^2}{S_Y^2}\right) \text{ et } \hat{\sigma}_{N_A}^{MOM} = \frac{\bar{Y}}{2} \left(1 + \frac{\bar{Y}^2}{S_Y^2}\right)$$

avec $\bar{Y} = \frac{1}{N_A} \sum_{i=1}^{N_A} Y_i$ et $S_Y^2 = \frac{1}{N_A - 1} \sum_{i=1}^{N_A} (Y_i - \bar{Y})^2$

— **La méthode du percentile élémentaire (EPM estimator)**

Il s'agit d'une méthode par estimation de la médiane (aucune restriction sur γ) :

$$\hat{\gamma}_{N_A}^{EPM} = \text{median} \{\hat{\gamma}_{i,j} : i > j\} \text{ et } \hat{\sigma}_{N_A}^{EPM} = \text{median} \{\hat{\sigma}_{i,j} : i > j\}$$

$$\text{où } \hat{\sigma}_{i,j} \text{ et } \hat{\gamma}_{i,j} \text{ sont solutions de : } \begin{cases} 1 - (1 + \hat{\tau}_{i,j} Y_i)^{-\frac{1}{\hat{\gamma}_{i,j}}} = p_i \\ 1 - (1 + \hat{\tau}_{i,j} Y_j)^{-\frac{1}{\hat{\gamma}_{i,j}}} = p_j \end{cases}$$

avec $i \neq j$ et donc $Y_i \neq Y_j$; $\hat{\tau}_{i,j} = \frac{\hat{\gamma}_{i,j}}{\hat{\sigma}_{i,j}}$; et $p_i = \frac{i}{n+1}$.

En pratique l'estimateur EPM donne de très bons résultats pour $\gamma < 0$.

— **L'estimateur de Hill**

Il s'agit en réalité du ML estimateur pour une distribution de Pareto. Pour $k=1, \dots, n-1$:

$$\hat{H}_k = \frac{1}{k} \sum_{j=1}^k \log(X_{n-j+1}) - \log(X_{n-k}), \text{ avec } (X_1, \dots, X_n) \text{ les } n \text{ observations triées par ordre croissant.}$$

Le quantile plot pour une Pareto est linéaire, et la pente de cette partie linéaire estime l'indice de queue ξ . H_k est un estimateur de cette pente.

Par construction, partant d'une distribution de Pareto, cet estimateur n'est asymptotiquement pas biaisé, mais il le devient si l'on n'a pas exactement une loi de Pareto, avec par exemple une GPD différente.

- Les méthodes basées sur les quantiles de la distribution, qui sont très liées aux diagrammes QQ-plot introduits lors de la partie précédente. Soit (X_1, \dots, X_n) les n observations triées par ordre croissant :

— **L'estimateur de Pickands**

Introduit par Pickands en 1975 :

$$\hat{\gamma}_k^P = \frac{1}{\log(2)} \log \left(\frac{X_{n-\lfloor \frac{k}{4} \rfloor} - X_{n-\lfloor \frac{k}{2} \rfloor}}{X_{n-\lfloor \frac{k}{2} \rfloor} - X_{n-k}} \right), \text{ pour } k=1, \dots, n \text{ et } \lfloor x \rfloor \text{ la partie entière de } x$$

Cet estimateur est asymptotiquement normal.

Il est très simple mais irrégulier par rapport à la variable k .

En pratique, il n'est pas applicable pour des échantillons de petite et moyenne taille, car il est très volatile.

— **L'estimateur de Dekkers-Einmalh-de Hann (DEdH)**

$$\hat{\gamma}_k^{DEdH} = H_k + 1 - \frac{1}{2} \left(1 - \frac{H_k^2}{H_k^*} \right)^{-1}, \text{ pour } k=1, \dots, n-1$$

$$\text{avec } H_k^* = \frac{1}{k} \sum_{j=1}^k (\log(X_{n-j+1}) - \log(X_{n-k}))^2$$

et H_k défini dans l'expression de l'estimateur de Hill.

Il s'agit d'une adaptation de l'estimateur de Hill, pour avoir une cohérence pour tout $\gamma \in \mathbb{R}$.

Il est asymptotiquement normal.

— **L'estimateur de Hill généralisé**

$$\hat{\gamma}_k^H = \frac{1}{k} \sum_{j=1}^k \log(UH_j) - \log(UH_{k+1}), \text{ pour } k=1, \dots, n-2$$

Il s'agit d'un estimateur de la partie extrême (droite) du diagramme GQ, nous menant à l'estimation de γ .

C'est une adaptation directe de l'estimateur de Hill qui peut être comparé à l'estimateur DEpH.

Les deux estimateurs s'écartent de celui de Hill afin de généraliser et adapter au cas où $\gamma \in \mathbb{R}$.

— **L'estimateur Zipf**

$$\hat{\gamma}_k^Z = \frac{\frac{1}{k} \sum_{j=1}^k \left(\log \left(\frac{k+1}{j+1} \right) - \frac{1}{k} \sum_{i=1}^k \log \left(\frac{k+1}{i+1} \right) \right) \log(UH_j)}{\frac{1}{k} \sum_{j=1}^k \left(\log \left(\frac{k+1}{j+1} \right) \right)^2 - \left(\frac{1}{k} \sum_{j=1}^k \log \left(\frac{k+1}{j+1} \right) \right)^2}$$

Il s'agit également d'un estimateur de la partie extrême (droite) du diagramme GQ, nous menant à l'estimation de γ (aucune restriction sur son domaine de définition).

Une caractéristique intéressante est la fluidité des réalisations en fonction de k , ce qui atténue dans une certaine mesure le problème du choix de $k \in \mathbb{R}$.

- Les méthodes de régression :

- **Le modèle de régression exponentiel (ERM)**

Les rapports logarithmiques échelonnés des espacements sont définis par :

$$Z_j = j \log \left(\frac{X_{n-j+1} - X_{n-k}}{X_{n-j} - X_{n-k}} \right), \text{ pour } j=1, \dots, k$$

On peut faire l'approximation : $Z_j \approx \frac{\gamma}{1 - (\frac{j}{k+1})^\gamma} E_j$ avec E_1, \dots, E_{k-1} i.i.d. variables aléatoires

exponentielles standard. Cette approximation est un prolongement de l'estimateur ERM au cas où $\gamma \in \mathbb{R}$. Les Z_j sont approximativement répartis de manière exponentielle. Il s'agit de l'estimateur par maximum de vraisemblance pour γ (qui partage de nombreuses propriétés avec l'estimateur POT ML). Il est asymptotiquement sans biais uniquement si l'échantillon provient d'une GPD.

- **Le modèle de régression exponentiel raffiné (Refined ERM)**

Inspiré du modèle de régression exponentiel, on y introduit des paramètres de variation régulière de second ordre réduisant le biais de l'estimateur. Il étend au cas où $\gamma \in \mathbb{R}$.

2.2 Modélisation classique du Body

La Théorie des Valeurs Extrêmes est une théorie asymptotique, et donc par définition, avant le seuil où l'on peut l'appliquer, il faut avoir recours à d'autres méthodes. Les sinistres situés avant ce seuil, ne figurent donc pas dans la Tail de distribution et se retrouvent ainsi dans le Body de la distribution. La grande majorité des sinistres constituent ainsi cette partie. Contrairement à la Tail, nous disposons ici d'un nombre relativement important de sinistres. C'est pourquoi l'on décide de faire confiance à l'empirique. L'hypothèse sous-jacente est que l'on suppose que le passé va avoir tendance à se reproduire. Même si l'on fait confiance au passé, il ne faut pas oublier de prendre en compte le temps qui passe et les différents facteurs à ajouter à nos sinistres : ils sont mis en "As-if", c'est-à-dire que l'on indexe les sinistres vers le futur avec un indice représentatif de cette évolution dans le temps (inflation, salaires...). Il s'agit de la valeur des sinistres étant survenus dans le passé, si jamais ils s'étaient produits aujourd'hui.

Plus l'on aura de sinistres dans le "Body", plus l'estimation empirique sera robuste. Ici, pour représenter un "Body empirique", nous avons simulé une loi Exponentielle avec 10, 40 et 200 observations. Il est dur de reconnaître la forme de la fonction de répartition d'une loi Exponentielle lorsque l'on manque d'observations.

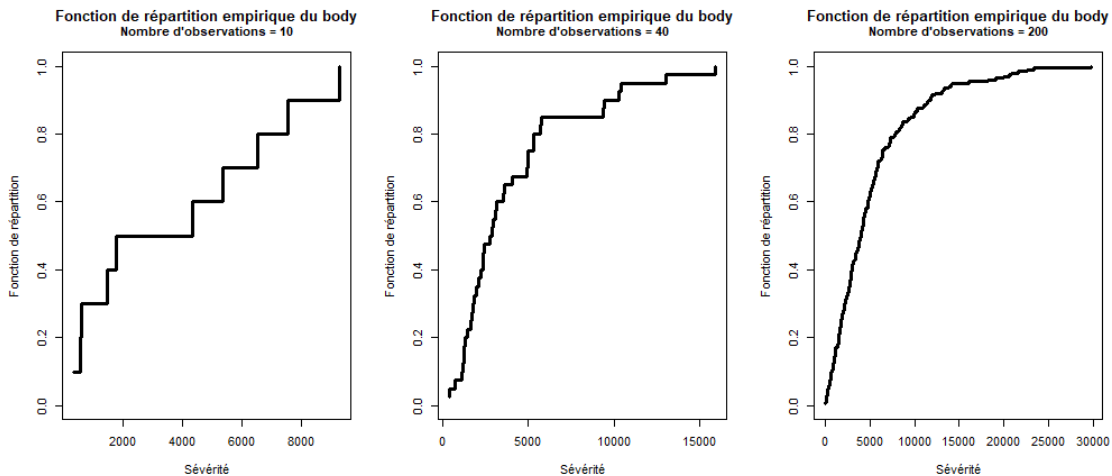


FIGURE 9 – Fonction de répartition empirique d'une loi Exponentielle

Moins l'on aura de sinistres, plus la distribution comportera de grands sauts. Prenons par exemple un traité Excess of Loss (C xs P) : mettons que l'on ait une distribution avec beaucoup de grands sauts avec notamment un sinistre valant $P+\epsilon$ avec ϵ infiniment petit. Dans cette tranche (C xs P) la sévérité moyenne est beaucoup moins importante que si jamais le sinistre en $P+\epsilon$ valait en réalité $P-\epsilon$; puisque ce sinistre vient baisser la sévérité moyenne des autres sinistres se trouvant dans cette tranche. La différence de valeur sur le sinistre est infime alors que la différence de prix pour la tranche (C xs P) peut s'avérer très grande. Si l'on augmente le nombre d'observations dans le body, ce problème est évidemment atténué, puisque les sauts seront de plus en plus petits.

2.3 En pratique

Pour la Tail

L'ajustement de la Tail avec les distributions GPD fait intervenir une famille de distributions à trois paramètres. À travers les différents estimateurs que nous avons mentionnés, il apparaît clairement qu'ils sont très sensibles au nombre de données, et perdent en robustesse lorsque nous n'avons pas assez d'observations. C'est pourquoi, en pratique cette famille de distribution n'est utilisée que pour établir des modèles marché (sur une branche d'activités avec l'ensemble des cédantes) ou bien pour des tarifications individuelles (une seule cédante) avec des portefeuilles qui possèdent au moins 200 sinistres. En réassurance une des principales problématiques est le manque de données, et très souvent nous n'avons pas ces 200 observations sur lesquelles s'appuyer. De plus, la Théorie des Valeurs Extrêmes est une théorie asymptotique, et par définition, plus on s'approche de la Tail moins on a de données. Ainsi, avec un jeu d'observations donné, plus on ajoute des sinistres bas dans la Tail, plus on a tendance à biaiser les estimateurs de la loi modélisant la Tail ; mais d'un autre côté, moins on a de données, plus les estimateurs sont volatils.

En pratique, on se limitera à principalement deux familles de distribution pour modéliser la Tail : les lois de Pareto ($\xi > 0$) et les lois Exponentielles ($\xi = 0$). Pour cela, nous prendrons deux outils : le ME plot et un graphique représentatif du paramètre de la loi de la Tail. Les deux outils sont utilisés ensemble afin de déterminer le seuil de séparation entre la Tail et le Body, ainsi le paramètre de la loi de la Tail.

Ainsi comme nous l'avons vu avec le ME plot, une pente croissante nous indique une loi de Pareto alors qu'une ligne constante nous indique une loi Exponentielle. Le seuil à partir duquel se dégage une tendance, nous indique le seuil à partir duquel on est potentiellement en présence de telle ou telle loi et où l'on peut séparer la Tail et le Body. En parallèle, on utilise un Alpha-Plot (pour la loi de Pareto) ou un Lambda-Plot (pour la loi Exponentielle), qui sont des graphiques sur lesquels on représente les valeurs du paramètre de la loi de la Tail obtenues par maximum de vraisemblance. Un graphique stable et constant nous indiquera alors que l'on modélise bien la bonne loi à partir du seuil où le graphique devient constant.

Pour l'estimation du paramètre, on estime par maximum de vraisemblance :

Loi de Pareto : Si $X \sim \text{Pareto}(A, \alpha)$ avec x_1, \dots, x_n ses n observations, alors la vraisemblance s'écrit :

$$* \quad \mathcal{L}(\alpha) = \prod_{i=0}^n \frac{\alpha A^\alpha}{x_i^{\alpha+1}} = \alpha^n A^{n\alpha} \frac{1}{\left(\prod_{i=0}^n x_i\right)^{\alpha+1}}$$

$$\Rightarrow \ell(\alpha) = n \ln(\alpha) + n\alpha \ln(A) - (\alpha + 1) \sum_{i=0}^n \ln(x_i)$$

$$\begin{aligned} \Rightarrow \frac{\partial \ell(\alpha)}{\partial \alpha} &= \frac{n}{\alpha} + n \ln(A) - \sum_{i=0}^n \ln(x_i) = \frac{n}{\alpha} - \sum_{i=0}^n \ln\left(\frac{x_i}{A}\right) \\ \Rightarrow \frac{\partial \ell(\alpha)}{\partial \alpha} = 0 &\Leftrightarrow \hat{\alpha} = \frac{n}{\sum_{i=0}^n \ln\left(\frac{x_i}{A}\right)} \end{aligned}$$

Vérifions le biais de l'estimateur :

$$\begin{aligned} * \quad X \sim \text{Pareto}(A, \alpha) &\Rightarrow \sum_{i=0}^n \ln\left(\frac{X}{A}\right) \sim \text{Gamma}(n, \alpha) \text{ et donc } \left(\sum_{i=0}^n \ln\left(\frac{X}{A}\right)\right)^{-1} \sim \text{Inv-Gamma}(n, \alpha^{-1}) \\ \mathbb{E}(\hat{\alpha}) &= n \mathbb{E}\left(\frac{1}{\sum_{i=0}^n \ln\left(\frac{X}{A}\right)}\right) = \frac{n}{n-1} \alpha \Rightarrow \boxed{\hat{\alpha} = \frac{n-1}{\sum_{i=0}^n \ln\left(\frac{x_i}{A}\right)}} \end{aligned}$$

Loi Exponentielle : Si $X \sim \text{Exponentiel}(\lambda)$ avec x_1, \dots, x_n ses n observations, alors la vraisemblance s'écrit :

$$\begin{aligned} * \quad \mathcal{L}(\lambda) &= \prod_{i=0}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=0}^n x_i} \\ \Rightarrow \ell(\lambda) &= n \ln(\lambda) - \lambda \sum_{i=0}^n x_i \Rightarrow \frac{\partial \ell(\lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=0}^n x_i \\ \Rightarrow \frac{\partial \ell(\lambda)}{\partial \lambda} = 0 &\Leftrightarrow \hat{\lambda} = \frac{n}{\sum_{i=0}^n x_i} \end{aligned}$$

Vérifions le biais de l'estimateur :

$$\begin{aligned} * \quad X \sim \text{Exponentiel}(\lambda) &\Rightarrow \sum_{i=0}^n X_i \sim \text{Gamma}(n, \lambda) \\ \mathbb{E}(\hat{\lambda}) &= n \mathbb{E}\left(\frac{1}{\sum_{i=0}^n X}\right) = \frac{n}{n-1} \lambda \Rightarrow \boxed{\hat{\lambda} = \frac{n-1}{\sum_{i=0}^n x_i}} \end{aligned}$$

En pratique, pour les tracés des graphiques on estimera le ME plot ainsi que le paramètre aux points : $x = X_{n-k}$, avec $k=1, \dots, n-1$ et (X_1, \dots, X_n) les n observations triées par ordre croissant.

$$\text{Pour la loi de Pareto : } \hat{\alpha}_k = \frac{k-1}{\sum_{j=1}^k \ln\left(\frac{X_{n-j+1}}{X_{n-k}}\right)} ; \text{ et pour la loi Exponentielle : } \hat{\lambda}_k = \frac{k-1}{\sum_{j=1}^k X_{n-j}}$$

Et l'on tracera les points : $(X_{n-k}, \hat{\alpha}_k)$ et $(X_{n-k}, \hat{\lambda}_k)$, pour $k=1, \dots, n-1$.

Illustrons ces deux cas :

Simulons des observations issues d'une loi de Pareto en rajoutant un bruit important au début de la distribution, là où la sévérité est faible; le ME plot et Alpha-plot nous donnent :

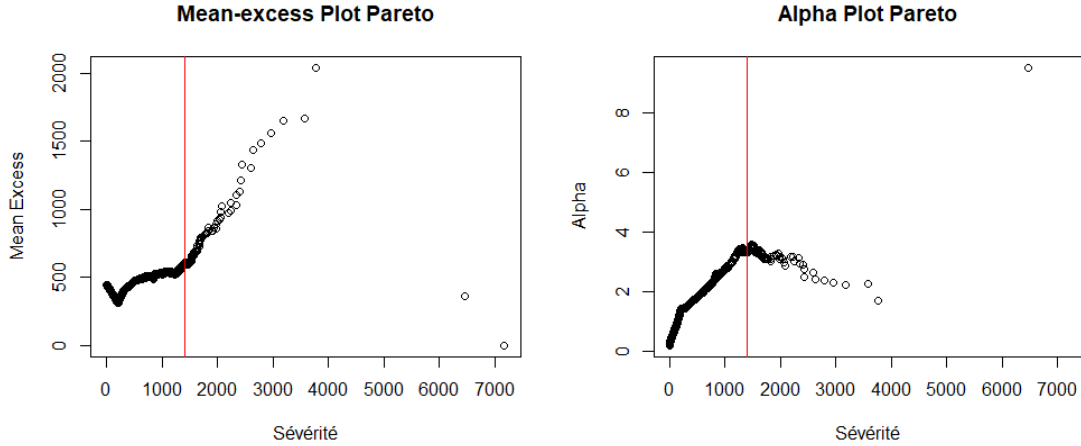


FIGURE 10 – ME plot et Alpha plot

À partir du seuil tracé en rouge on reconnaît bien un comportement du type Pareto, avec un ME plot linéairement croissant et un Alpha-plot approximativement constant.

De même, simulons des observations issues d'une loi Exponentielle en rajoutant un bruit important au début de la distribution, là où la sévérité est faible ; le ME plot et Alpha-plot nous donnent :

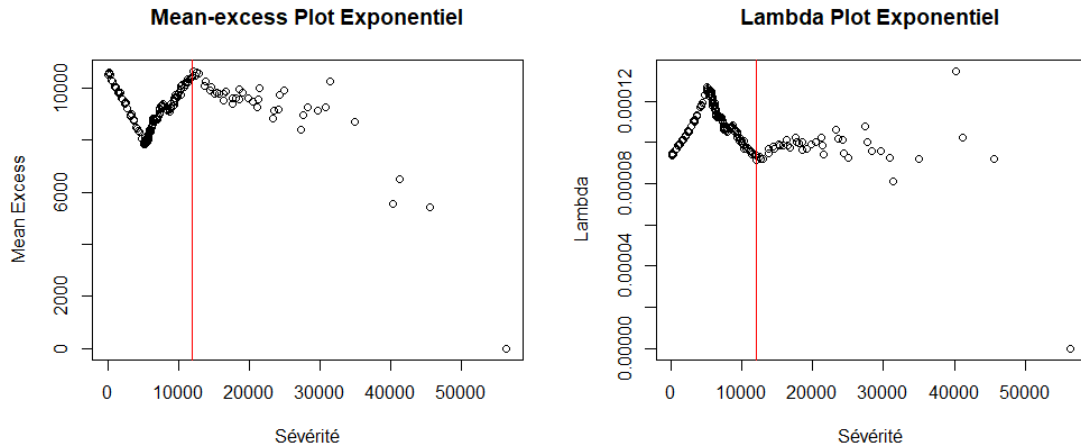


FIGURE 11 – ME plot et Lambda plot

Ici également, à partir du seuil tracé en rouge on reconnaît bien un comportement du type Exponentiel, avec un ME plot constant et un Lambda-plot approximativement constant.

Pour le Body

Pour le Body de la distribution, en pratique, on garde bien la distribution empirique comme expliqué dans la partie précédente. Bien évidemment, ce jugement reste subjectif et ainsi cela peut varier d'une personne à une autre. Mais dans la majorité des cas, on estime que le Body est constitué de suffisamment de sinistres et l'on opte pour une modélisation empirique.

2.4 But du mémoire

Comme nous avons pu le voir, même si la modélisation de la sévérité faite en pratique est une bonne adaptation de la modélisation théorique, elle présente de nombreux inconvénients.

Pour le Body de la distribution, le fait d’avoir une distribution empirique induit des prix segmentés et très sensibles au nombre de sinistres ainsi qu’à leur position dans la distribution, effet accentué selon le type de contrat que l’on souhaite tarifer (voir section 2.2). C’est pourquoi il serait intéressant d’avoir une méthode nous permettant d’avoir une distribution lissée pour modéliser le Body. Cette méthode, par soucis d’objectivité se doit d’être la même pour chaque Analyste.

Pour modéliser la Tail de la distribution, on envisage finalement peu de familles de lois potentielles. Et ces distributions envisagées sont à support illimité, il serait désirable d’étudier d’autres distributions, à support non-illimité (loi censurée ou tronquée par exemple) et qui pourraient très bien modéliser la Tail de la distribution.

De plus, pour l’ensemble de la distribution, nous considérons uniquement les sinistres comme des points fixes, en tronquant ainsi certaines informations fournies par les cédantes. Il serait intéressant d’intégrer le reste des informations fournies afin de prendre en compte la volatilité autour de ces sinistres. En prenant comme hypothèse que l’on réussisse à avoir un intervalle de possibilité pour la valeur de chaque sinistre, on pourrait alors adapter notre modélisation de la sévérité. En effet, un autre problème est le seuil choisi pour séparer la Tail du Body : mettons que l’on ait 10 sinistres dans la Tail de la distribution et 5 sinistres juste avant le seuil choisi, décaler le seuil très légèrement va faire intégrer ces 5 autres sinistres à la Tail, faisant ainsi baisser lourdement la sévérité de la loi ajustée. Considérer un intervalle de possibilités pour chaque sinistre, permettrait d’éviter cette problématique de seuil, et la modélisation sera beaucoup moins sensible et affectée par le choix du seuil. Une estimation, basée sur l’ensemble des informations des sinistres, permettant d’avoir un intervalle de possibilités pour chaque sinistre semble plus refléter la réalité. En considérant qu’un sinistre n’a pas une valeur fixe, comment pouvons-nous modéliser la distribution globale de la sévérité de ces sinistres ?

Afin de concentrer toute notre attention sur la sévérité des sinistres, la modélisation de la fréquence des sinistres n’est pas abordée dans ce mémoire. Comme expliqué, un des objectifs de ce mémoire est de trouver une distribution adaptée pour des sinistres pouvant prendre différentes valeurs dans un intervalle, on cherche donc à trouver cette distribution en sachant que l’on possède cette forme de sinistres. L’estimation de ces sinistres est réalisée au sein de QBE par le mémoire de MOURIER [2020]⁸.

8. MOURIER [2020] - cf Bibliographie [1]

3 Modélisation de la Tail de la distribution de sévérité

Cette partie porte sur l'étude de nouvelles lois potentielles, non étudiées au sein de la compagnie QBE afin de modéliser la distribution de la sévérité des sinistres. Dans un premier temps, sur base de sinistres fixes, nous envisagerons l'étude de lois censurées à droite, illimitées ou tronquées à droite, ainsi que leur mise en application dans un cadre de tarification. Puis, dans un second temps, nous considérerons des sinistres non-fixes, prenant leurs valeurs dans un intervalle de possibilités ; et nous étudierons différentes lois à ajuster sur base de ces sinistres non-fixes.

3.1 Étude de nouvelles lois

De façon assez systématique, afin de modéliser la queue de la distribution de la sévérité des sinistres, il est utilisé des distributions de Pareto et Exponentielle. La distribution Exponentielle est utilisée comme alternative à la distribution de Pareto, possédant une queue plus légère. Ces deux distributions ont été généralisées par les Generalized Pareto Distribution. On a introduit les distributions GPD avec comme principaux arguments :

- Il s'agit d'une généralisation des lois de Pareto et Exponentielle.
- Elles permettent d'apporter des lois à support limité.
- Elles ont du sens par rapport à la Théorie des Valeurs Extrêmes.

Il serait cependant intéressant d'étudier d'autres distributions, plus simples à ajuster. L'idée est de considérer d'autres familles de distributions, on reprendrait les lois Exponentielles et de Pareto, auxquelles on rajouterait la loi de Weibull afin d'identifier les cas où la queue de distribution est plus lourde que celle de la loi Exponentielle, mais sans être de type Pareto. L'étude de leur comportement sera faite dans différents cas :

- Illimité (comme ce qui est fait actuellement, en considérant en plus la loi de Weibull)
- Tronqué à droite (on aurait a priori droit de penser qu'un sinistre ne peut théoriquement pas être supérieur à 15M si la valeur maximale observée sur l'ensemble des sinistres empiriques sur les 20 dernières années vaut 8M)
- Censuré à droite (en reprenant la même idée que la troncature, mais en "coupant" la fonction de survie et non en la "lissant")

L'objectif est d'étudier les mêmes outils que l'on utilise actuellement avec la loi de Pareto illimitée et la loi Exponentielle illimitée, c'est-à-dire le Mean-Excess Plot et le graphique du paramètre calculé par Maximum de Vraisemblance. On écrira le Mean-Excess plot dans les trois situations envisagées (illimitée, tronquée à droite et censurée à droite), avec les notations suivantes :

- X une variable aléatoire à valeurs dans \mathbb{R}^+
- $u \in \mathbb{R}^+$
- $b \in \mathbb{R}^+$, la valeur de la troncature droite
- F la fonction de répartition de X
- C la valeur de la censure droite
- F^T la fonction de répartition tronquée à droite de la variable aléatoire X , c'est-à-dire $F^T(x) = \begin{cases} \frac{F(x)}{F(b)} & \text{si } x \leq b \\ 1 & \text{si } x \geq b \end{cases}$
- F^C la fonction de répartition censurée à droite de la variable aléatoire X , c'est-à-dire $F^C(x) = \begin{cases} F(x) & \text{si } x \leq C \\ 1 & \text{si } x \geq C \end{cases}$

$$\begin{aligned}
* \quad \mathbb{E}_u^I &= \mathbb{E}(X^I - u | X^I > u) = \frac{1}{1 - F(u)} \int_u^\infty (1 - F(x)) dx \\
* \quad \mathbb{E}_u^T &= \mathbb{E}(X^T - u | X^T > u) = \frac{1}{1 - F^T(u)} \int_u^\infty (1 - F^T(x)) dx \\
* \quad \mathbb{E}_u^C &= \mathbb{E}(X^C - u | X^C > u) = \frac{1}{1 - F^C(u)} \int_u^\infty (1 - F^C(x)) dx
\end{aligned}$$

Pour l'estimation par Maximum de Vraisemblance, on supposera que la variable aléatoire X possède n observations (x_1, \dots, x_n) , f sa densité, S sa fonction de survie, et que Θ représente l'ensemble des paramètres de la loi. La fonction de vraisemblance dans les trois situations envisagées, s'exprimera comme :

$$\text{Illimitée : } \mathcal{L}(\Theta) = \prod_{i=1}^n f_{\Theta}(x_i)$$

$$\text{Tronquée à droite : } \mathcal{L}(\Theta) = \prod_{i=1}^n \frac{f_{\Theta}(x_i)}{F_{\Theta}(b)} \text{ avec } b \text{ la troncature droite}$$

$$\text{Censurée à droite : } \mathcal{L}(\Theta) = \prod_{i=1}^n f_{\Theta}(t_i)^{d_i} S_{\Theta}(t_i)^{1-d_i} \text{ avec } d_i = \mathbb{1}_{X_i < C}, D = \sum_{i=1}^n d_i \text{ et } t_i = \begin{cases} x_i & \text{si } x_i < C \\ C & \text{si } x_i \geq C \end{cases}$$

3.1.1 Loi de Pareto

Soit $X \sim \text{Pareto}(\alpha, A)$, étudions les trois situations envisagées :

Cas illimité

Le cas d'une distribution de Pareto illimitée a déjà été expliqué dans la section 2.3, il s'agit d'une distribution utilisée actuellement.

Cas tronqué à droite

$$\begin{aligned}
* \quad \mathbb{E}_u^T &= \frac{1}{\left(\frac{A}{b}\right)^\alpha - \left(\frac{A}{u}\right)^\alpha} \int_u^b \left(\left(\frac{A}{b}\right)^\alpha - \left(\frac{A}{x}\right)^\alpha \right) dx = \frac{1}{b^{-\alpha} - u^{-\alpha}} \left((b-u) \frac{1}{b^\alpha} - \left[\frac{1}{(1-\alpha)x^{\alpha-1}} \right]_u^b \right) \\
&= \frac{(ub)^\alpha}{u^\alpha - b^\alpha} \left((b-u) \frac{1}{b^\alpha} - \frac{1}{1-\alpha} \left(\frac{1}{b^{(\alpha-1)} - u^{(\alpha-1)}} \right) \right) \Rightarrow \boxed{\mathbb{E}_u^T = \frac{u^\alpha (b\alpha + u(1-\alpha)) - ub^\alpha}{u^\alpha - b^\alpha}}
\end{aligned}$$

$$\begin{aligned}
* \quad \mathcal{L}(\alpha) &= \prod_{i=1}^n \frac{\alpha A^\alpha}{x_i^{\alpha+1} \left(1 - \left(\frac{A}{b}\right)^\alpha\right)} = \frac{\alpha^n A^{n\alpha}}{\left(1 - \left(\frac{A}{b}\right)^\alpha\right)^n} \frac{1}{\left(\prod_{i=1}^n x_i\right)^{\alpha+1}} \\
\Rightarrow \ell(\alpha) &= n \ln(\alpha) + n\alpha \ln(A) - n \ln \left(1 - \left(\frac{A}{b}\right)^\alpha\right) - (\alpha+1) \sum_{i=1}^n \ln(x_i) \\
\Rightarrow \frac{\partial \ell(\alpha)}{\partial \alpha} &= \frac{n}{\alpha} - \sum_{i=1}^n \ln \left(\frac{x_i}{A}\right) + n \frac{\ln \left(\frac{A}{b}\right) \left(\frac{A}{b}\right)^\alpha}{\left(1 - \left(\frac{A}{b}\right)^\alpha\right)}
\end{aligned}$$

$$\text{Et donc, } \frac{\partial \ell(\alpha)}{\partial \alpha} = 0 \Leftrightarrow \boxed{\frac{1}{\hat{\alpha}} + \frac{\ln \left(\frac{A}{b}\right)}{\left(\frac{b}{A}\right)^{\hat{\alpha}} - 1} = \frac{\sum_{i=1}^n \ln \left(\frac{x_i}{A}\right)}{n}}$$

En simulant 2000 observations issues d'une loi de Pareto tronquée à droite, le Mean-Excess plot va posséder la forme suivante :

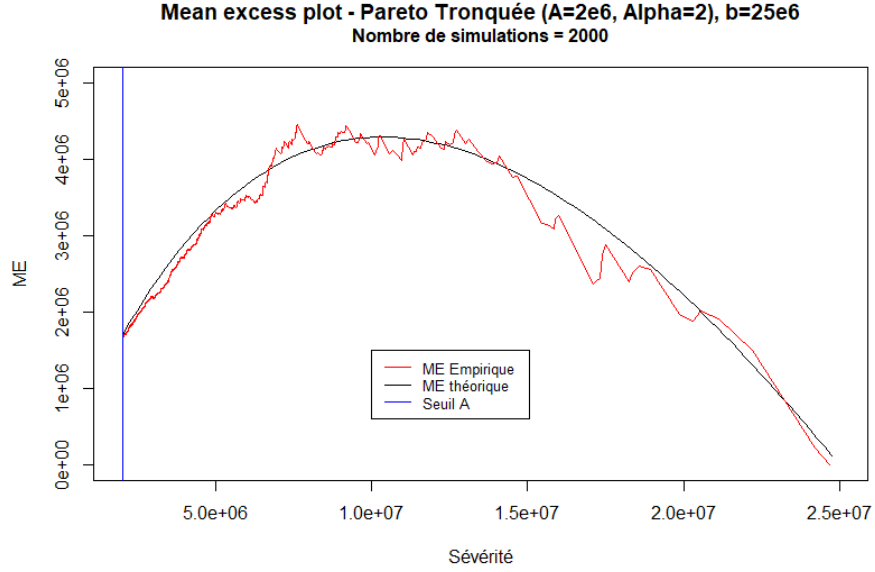


FIGURE 12 – Mean-Excess plot - Loi de Pareto tronquée à droite

Cas censuré à droite

$$* \quad \mathbb{E}_u^C = \frac{1}{\left(\frac{A}{u}\right)^\alpha} \int_u^C \left(\frac{A}{x}\right)^\alpha dx = u^\alpha \left[-\frac{1}{(\alpha-1)x^{(\alpha-1)}} \right]_u^C \Rightarrow \boxed{\mathbb{E}_u^C = \frac{u^\alpha}{(\alpha-1)} \left(\frac{1}{u^{(\alpha-1)}} - \frac{1}{C^{(\alpha-1)}} \right)}$$

$$* \quad \mathcal{L}(\alpha) = \prod_{i=1}^n \left(\frac{\alpha A^\alpha}{t_i^{\alpha+1}} \right)^{d_i} \left(\left(\frac{A}{t_i} \right)^\alpha \right)^{1-d_i} = \alpha^D A^{n\alpha} \frac{1}{\left(\prod_{i=1}^n t_i \right)^\alpha} \frac{1}{\prod_{i=1}^n t_i^{d_i}}$$

$$\Rightarrow \ell(\alpha) = D \ln(\alpha) + n\alpha \ln(A) - \alpha \sum_{i=1}^n \ln(t_i) - \ln \left(\prod_{i=1}^n t_i^{d_i} \right)$$

$$\Rightarrow \frac{\partial \ell(\alpha)}{\partial \alpha} = \frac{D}{\alpha} + n \ln(A) - \sum_{i=1}^n \ln(t_i) = \frac{D}{\alpha} - \sum_{i=1}^n \ln \left(\frac{x_i}{A} \right)$$

$$\text{Et donc, } \frac{\partial \ell(\alpha)}{\partial \alpha} = 0 \Leftrightarrow \hat{\alpha} = \frac{D}{\sum_{i=1}^n \ln \left(\frac{t_i}{A} \right)} \Rightarrow \boxed{\hat{\alpha} = \frac{D}{\sum_{i=1}^n \ln \left(\frac{x_i}{A} \right) - (n-D) \ln \left(\frac{C}{A} \right)}}$$

En simulant 2000 observations issues d'une loi de Pareto censurée à droite, le Mean-Excess plot va posséder la forme suivante :

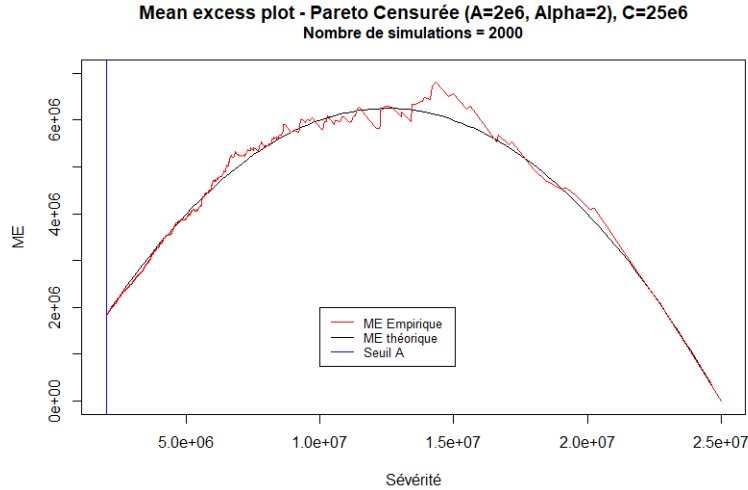


FIGURE 13 – Mean-Excess plot - Loi de Pareto censurée à droite

Conclusion

Les alternatives Pareto censurée à droite et Pareto tronquée à droite présentent un Mean-Excess plot avec une forme très similaire, en cloche, partant d'une valeur initiale non nulle et finissant en zéro au point de censure/troncature. Les différents paramètres peuvent se retrouver de la manière suivante :

- Les paramètres de censure (C) et de troncature (b) sont impossibles à identifier par maximum de vraisemblance, mais ils sont cependant identifiables à l'extrémité droite du Mean-Excess plot, là où la courbe retombe à zéro. En effet, $\mathbb{E}_u^C(C) = \mathbb{E}_u^T(b) = 0$.
- Le seuil A à partir duquel on identifie les lois peut être assimilé au début du Mean-Excess plot lorsqu'une forme en cloche est reconnaissable.
- Le paramètre α : pour la loi censurée, il suffit de reprendre son expression trouvée par maximum de vraisemblance ; et pour la loi tronquée, comme son expression n'est pas linéaire, un algorithme de résolution numérique doit être utilisé (type Newton-Raphson par exemple).

Cependant, plusieurs problèmes en pratique semblent apparaître :

- Pour obtenir les valeurs de C et b , en pratique on a peu d'observations dans les valeurs extrêmes, ce qui rend l'estimation de ces paramètres très compliquée, et cette estimation est fixée uniquement par les dernières observations (les plus grandes).
- Pour l'estimation du seuil A : si l'on tire le seuil vers le bas, on observera probablement toujours une forme de cloche pour le Mean-Excess plot, alors que plus l'on tend vers des sinistres plus faibles plus le comportement des sinistres n'a a priori aucune raison d'être similaire.
- De plus, même si l'on observe une forme en cloche, on ne peut pas précisément définir si l'on a un comportement censuré ou tronqué.

3.1.2 Loi Exponentielle

Soit $X \sim \text{Exponentiel}(\lambda)$, étudions les trois situations envisagées :

Cas illimité

Le cas d'une distribution Exponentielle illimitée a déjà été expliqué dans la section 2.3, il s'agit d'une distribution utilisée actuellement.

Cas tronqué à droite

$$\begin{aligned}
 * \quad \mathbb{E}_u^T &= \frac{1}{e^{-\lambda b} - e^{-\lambda u}} \int_u^b (e^{-\lambda b} - e^{-\lambda x}) dx = \frac{1}{e^{-\lambda b} - e^{-\lambda u}} \left((b-u)e^{-\lambda b} - \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_u^b \right) \\
 &= \frac{1}{e^{-\lambda b} - e^{-\lambda u}} \left((b-u)e^{-\lambda b} + \frac{e^{-\lambda b} - e^{-\lambda u}}{\lambda} \right) = \frac{1}{\lambda} + \frac{(b-u)e^{-\lambda b}}{e^{-\lambda b} - e^{-\lambda u}} \Rightarrow \boxed{\mathbb{E}_u^T = \frac{1}{\lambda} + \frac{(b-u)}{1 - e^{-\lambda(u-b)}}}
 \end{aligned}$$

$$* \quad \mathcal{L}(\lambda) = \prod_{i=1}^n \frac{\lambda e^{-\lambda x_i}}{1 - e^{-\lambda b}} = \left(\frac{\lambda}{1 - e^{-\lambda b}} \right)^n e^{-\lambda \sum_{i=1}^n x_i} \Rightarrow \ell(\lambda) = n \ln(\lambda) - n \ln(1 - e^{-\lambda b}) - \lambda \sum_{i=1}^n x_i$$

$$\Rightarrow \frac{\partial \ell(\lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i - n \frac{\lambda e^{-\lambda b}}{1 - e^{-\lambda b}}$$

$$\text{Et donc, } \frac{\partial \ell(\lambda)}{\partial \lambda} = 0 \Leftrightarrow \boxed{\frac{1}{\hat{\lambda}} = \frac{\hat{\lambda} e^{-\hat{\lambda} b}}{1 - e^{-\hat{\lambda} b}} - \frac{\sum_{i=1}^n x_i}{n}}$$

En simulant 2000 observations issues d'une loi Exponentielle tronquée à droite, le Mean-Excess plot va posséder la forme suivante :

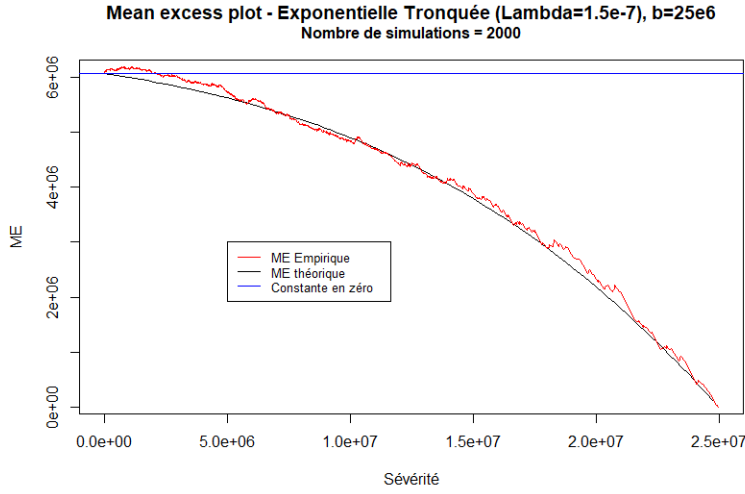


FIGURE 14 – Mean-Excess plot - Loi Exponentielle Tronquée à droite

Cas censuré à droite

$$* \quad \mathbb{E}_u^C = \frac{1}{e^{-\lambda u}} \int_u^C e^{-\lambda x} dx = e^{\lambda u} \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_u^C \Rightarrow \boxed{\mathbb{E}_u^C = \frac{1 - e^{-\lambda(C-u)}}{\lambda}}$$

$$* \quad \mathcal{L}(\lambda) = \prod_{i=1}^n \left(\lambda e^{-\lambda t_i} \right)^{d_i} \left(e^{-\lambda t_i} \right)^{1-d_i} = \lambda^D e^{-\lambda \sum_{i=1}^n t_i} \Rightarrow \ell(\lambda) = D \ln(\lambda) - \lambda \sum_{i=1}^n t_i \Rightarrow \frac{\partial \ell(\lambda)}{\partial \lambda} = \frac{D}{\lambda} - \sum_{i=1}^n t_i$$

Et donc, $\frac{\partial \ell(\lambda)}{\partial \lambda} = 0 \Leftrightarrow \hat{\lambda} = \frac{D}{\sum_{i=1}^n t_i} \Rightarrow \hat{\lambda} = \frac{D}{\sum_{i=1}^n x_i + (n - D)C}$

En simulant 2000 observations issues d'une loi Exponentielle censurée à droite, le Mean-Excess plot va posséder la forme suivante :

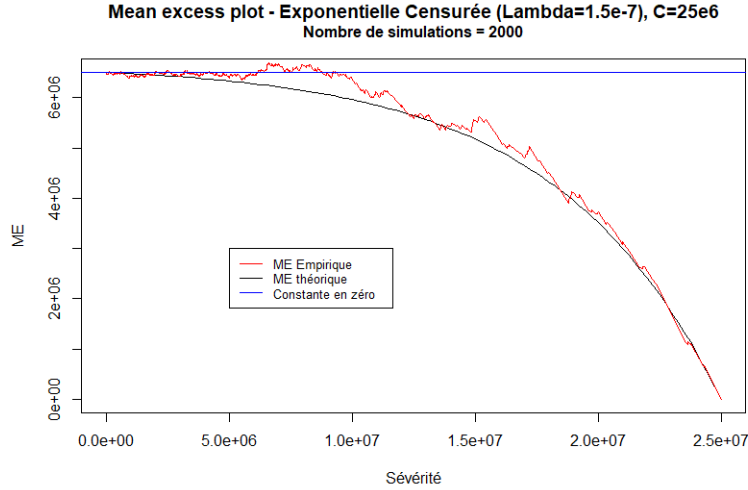


FIGURE 15 – Mean-Excess plot - Loi Exponentielle Censurée à droite

Conclusion

De manière analogue au cas des distributions de Pareto, les alternatives Exponentielle censurée à droite et Exponentielle tronquée à droite présentent un Mean-Excess avec une forme très similaire, qui est décroissante, partant d'une valeur initiale non nulle et finissant en zéro au point de censure/troncature. Cependant, le Mean-Excess dans le cas tronqué a une décroissance plus rapide que dans le cas censuré. Comme pour le cas Pareto, les différents paramètres peuvent se retrouver de la manière suivante :

- Les paramètres de censure (C) et de troncature (b) sont impossibles à identifier par maximum de vraisemblance, mais ils sont cependant identifiables à l'extrémité droite du Mean-Excess plot, là où la courbe retombe à zéro. En effet, $\mathbb{E}_u^C(C) = \mathbb{E}_u^T(b) = 0$.
- Le seuil A à partir duquel on identifie les lois peut être assimilé au début du Mean-Excess plot lorsqu'une forme de "demi-parabole" est reconnaissable.
- Le paramètre λ : pour la loi censurée, il suffit de reprendre son expression trouvée par maximum de vraisemblance; et pour la loi tronquée, comme son expression n'est pas linéaire, un algorithme de résolution numérique doit être utilisé (type Newton-Raphson par exemple).

Nous avons exactement les mêmes problèmes que pour le cas d'une distribution de Pareto. La seule différence, c'est que le seuil semble plus identifiable. En effet, comparé au cas précédent, ici, le seuil semblerait plus simple à percevoir, on ne peut pas tirer plus le seuil vers le bas sans modifier la forme de "demi-parabole" du Mean-Excess plot.

3.1.3 Loi de Weibull

Soit $X \sim \text{Weibull}(k, \lambda)$, étudions les trois situations envisagées :

Cas illimité

$$* \quad \mathbb{E}_u^I = \frac{1}{e^{-\left(\frac{u}{\lambda}\right)^k}} \int_u^\infty e^{-\left(\frac{x}{\lambda}\right)^k} dx = e^{\left(\frac{u}{\lambda}\right)^k} \int_{\left(\frac{u}{\lambda}\right)^k}^\infty \frac{\lambda}{k} e^{-v} v^{\frac{1}{k}-1} dv \Rightarrow \boxed{\mathbb{E}_u^I = \frac{\lambda}{k} e^{\left(\frac{u}{\lambda}\right)^k} \Gamma\left(\frac{1}{k}, \left(\frac{u}{\lambda}\right)^k\right)}$$

où $\Gamma(a, \cdot)$ est la fonction gamma incomplète de paramètre a , telle que $\forall x \in \mathbb{R}^+ : \Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$ et sa fonction complémentaire est définie comme : $\gamma(a, x) = \Gamma(a) - \Gamma(a, x)$ avec $\Gamma(\cdot)$ la fonction gamma.

$$* \quad \mathcal{L}(k, \lambda) = \prod_{i=1}^n \left(\frac{k}{\lambda}\right) \left(\frac{x_i}{\lambda}\right)^{(k-1)} e^{-\left(\frac{x_i}{\lambda}\right)^k} = \frac{k^n}{\lambda^{nk}} \left(\prod_{i=1}^n x_i\right)^{(k-1)} e^{-\sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k}$$

$$\Rightarrow \ell(k, \lambda) = n \ln(k) - nk \ln(\lambda) + (k-1) \sum_{i=1}^n \ln(x_i) - \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k$$

$$\left\{ \begin{array}{l} \frac{\partial \ell(k, \lambda)}{\partial k} = \frac{n}{k} + \sum_{i=1}^n \ln\left(\frac{x_i}{\lambda}\right) - \sum_{i=1}^n \ln\left(\frac{x_i}{\lambda}\right) \left(\frac{x_i}{\lambda}\right)^k \\ \frac{\partial \ell(k, \lambda)}{\partial \lambda} = -\frac{nk}{\lambda} + \frac{k}{\lambda} \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \frac{\partial \ell(k, \lambda)}{\partial \lambda} = 0 \Leftrightarrow \boxed{\hat{\lambda} = \left(\frac{\sum_{i=1}^n x_i}{n}\right)^{\frac{1}{\hat{k}}}} \\ \frac{\partial \ell(k, \lambda)}{\partial k} = 0 \Leftrightarrow \boxed{\frac{n}{\hat{k}} = \sum_{i=1}^n \ln\left(\frac{x_i}{\hat{\lambda}}\right) \left(\left(\frac{x_i}{\hat{\lambda}}\right)^{\hat{k}} - 1\right)} \end{array} \right.$$

En simulant 2000 observations issues d'une loi de Weibull illimitée, le Mean-Excess plot va posséder la forme suivante :

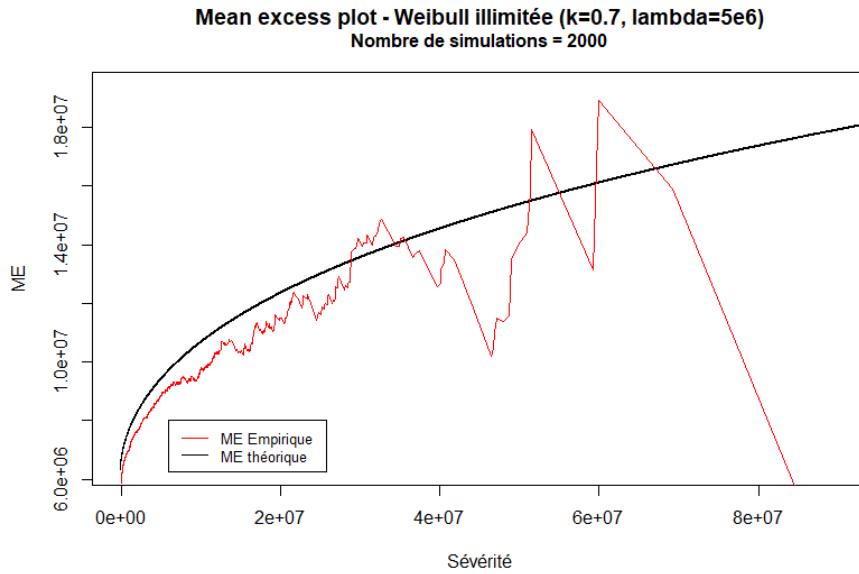


FIGURE 16 – Mean-Excess plot - Loi de Weibull Illimitée

Cas tronqué à droite

$$* \quad \mathbb{E}_u^T = \frac{1}{e^{-\left(\frac{u}{\lambda}\right)^k} - e^{-\left(\frac{b}{\lambda}\right)^k}} \int_u^b \left(e^{-\left(\frac{x}{\lambda}\right)^k} - e^{-\left(\frac{b}{\lambda}\right)^k} \right) dx = \frac{1}{e^{-\left(\frac{u}{\lambda}\right)^k} - e^{-\left(\frac{b}{\lambda}\right)^k}} \left((u-b)e^{-\left(\frac{b}{\lambda}\right)^k} + \int_u^b e^{-\left(\frac{x}{\lambda}\right)^k} dx \right)$$

$$= \frac{1}{e^{-\left(\frac{u}{\lambda}\right)^k} - e^{-\left(\frac{b}{\lambda}\right)^k}} \left((u-b)e^{-\left(\frac{b}{\lambda}\right)^k} + \int_{\left(\frac{u}{\lambda}\right)^k}^{\left(\frac{b}{\lambda}\right)^k} \frac{\lambda}{k} e^{-v} v^{\frac{1}{k}-1} dv \right)$$

$$\Rightarrow \mathbb{E}_u^T = \frac{1}{e^{-\left(\frac{u}{\lambda}\right)^k} - e^{-\left(\frac{b}{\lambda}\right)^k}} \left((u-b)e^{-\left(\frac{b}{\lambda}\right)^k} + \frac{\lambda}{k} \left(\Gamma\left(\frac{1}{k}\right) - \Gamma\left(\frac{1}{k}, \left(\frac{u}{\lambda}\right)^k\right) - \gamma\left(\frac{1}{k}, \left(\frac{b}{\lambda}\right)^k\right) \right) \right)$$

$$* \quad \mathcal{L}(k, \lambda) = \prod_{i=1}^n \frac{\left(\frac{k}{\lambda}\right) \left(\frac{x_i}{\lambda}\right)^{(k-1)} e^{-\left(\frac{x_i}{\lambda}\right)^k}}{1 - e^{-\left(\frac{b}{\lambda}\right)^k}} = \frac{k^n}{\lambda^{nk} \left(1 - e^{-\left(\frac{b}{\lambda}\right)^k}\right)^n} \left(\prod_{i=1}^n x_i \right)^{(k-1)} e^{-\sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k}$$

$$\Rightarrow \ell(k, \lambda) = n \ln(k) - nk \ln(\lambda) + (k-1) \sum_{i=1}^n \ln(x_i) - \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k - n \ln\left(1 - e^{-\left(\frac{b}{\lambda}\right)^k}\right)$$

$$\Rightarrow \left\{ \begin{array}{l} \frac{\partial \ell(k, \lambda)}{\partial k} = \frac{n}{k} + \sum_{i=1}^n \ln\left(\frac{x_i}{\lambda}\right) - \sum_{i=1}^n \ln\left(\frac{x_i}{\lambda}\right) \left(\frac{x_i}{\lambda}\right)^k + n \frac{\ln\left(\frac{b}{\lambda}\right) \left(\frac{b}{\lambda}\right)^k e^{-\left(\frac{b}{\lambda}\right)^k}}{1 - e^{-\left(\frac{b}{\lambda}\right)^k}} \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial \ell(k, \lambda)}{\partial \lambda} = -\frac{nk}{\lambda} + \frac{k}{\lambda} \sum_{i=1}^n \left(\frac{x_i}{\lambda}\right)^k - n \frac{\frac{k}{\lambda} \left(\frac{b}{\lambda}\right)^k e^{-\left(\frac{b}{\lambda}\right)^k}}{1 - e^{-\left(\frac{b}{\lambda}\right)^k}} \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{l} \frac{\partial \ell(k, \lambda)}{\partial \lambda} = 0 \Leftrightarrow \frac{1}{\hat{\lambda}} + \frac{\left(\frac{b}{\hat{\lambda}}\right)^{\hat{k}}}{e^{\left(\frac{b}{\hat{\lambda}}\right)^{\hat{k}}} - 1} = \frac{\sum_{i=1}^n \left(\frac{x_i}{\hat{\lambda}}\right)^{\hat{k}}}{n} \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial \ell(k, \lambda)}{\partial k} = 0 \Leftrightarrow \frac{n}{\hat{k}} = \sum_{i=1}^n \ln\left(\frac{x_i}{\hat{\lambda}}\right) \left(\left(\frac{x_i}{\hat{\lambda}}\right)^{\hat{k}} - 1\right) - n \frac{\ln\left(\frac{b}{\hat{\lambda}}\right) \left(\frac{b}{\hat{\lambda}}\right)^{\hat{k}}}{e^{\left(\frac{b}{\hat{\lambda}}\right)^{\hat{k}}} - 1} \end{array} \right.$$

En simulant 2000 observations issues d'une loi de Weibull tronquée à droite, le Mean-Excess plot va posséder la forme suivante :

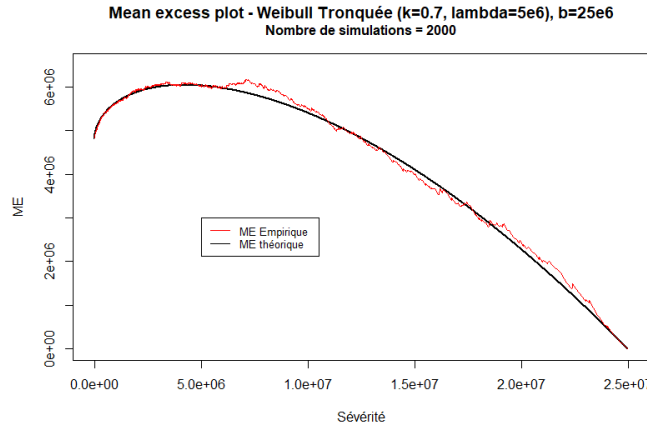


FIGURE 17 – Mean-Excess plot - Loi de Weibull Tronquée à droite

Cas censuré à droite

$$* \quad \mathbb{E}_u^C = \frac{1}{e^{-\left(\frac{u}{\lambda}\right)^k}} \int_u^C e^{-\left(\frac{x}{\lambda}\right)^k} dx = e^{\left(\frac{u}{\lambda}\right)^k} \int_u^C e^{-\left(\frac{x}{\lambda}\right)^k} dx = e^{\left(\frac{u}{\lambda}\right)^k} \int_{\left(\frac{u}{\lambda}\right)^k}^{\left(\frac{C}{\lambda}\right)^k} \frac{\lambda}{k} e^{-v} v^{\frac{1}{k}-1} dv$$

$$\Rightarrow \mathbb{E}_u^C = \frac{\lambda}{k} e^{\left(\frac{u}{\lambda}\right)^k} \left(\Gamma\left(\frac{1}{k}\right) - \Gamma\left(\frac{1}{k}, \left(\frac{C}{\lambda}\right)^k\right) - \gamma\left(\frac{1}{k}, \left(\frac{u}{\lambda}\right)^k\right) \right)$$

$$* \quad \mathcal{L}(k, \lambda) = \prod_{i=1}^n \left(\left(\frac{k}{\lambda}\right) \left(\frac{t_i}{\lambda}\right)^{(k-1)} e^{-\left(\frac{t_i}{\lambda}\right)^k} \right)^{d_i} \left(e^{-\left(\frac{t_i}{\lambda}\right)^k} \right)^{1-d_i} = \frac{k^D}{\lambda^{Dk}} \left(\prod_{i=1}^n t_i^{(k-1)d_i} \right) e^{-\sum_{i=1}^n \left(\frac{t_i}{\lambda}\right)^k}$$

$$\Rightarrow \ell(k, \lambda) = D \ln(k) - Dk \ln(\lambda) + (k-1) \sum_{i=1}^n d_i \ln(t_i) - \sum_{i=1}^n \left(\frac{t_i}{\lambda}\right)^k$$

$$\left\{ \begin{array}{l} \frac{\partial \ell(k, \lambda)}{\partial k} = \frac{D}{k} - D \ln(\lambda) + \sum_{i=1}^n \left(d_i \ln(t_i) - \ln\left(\frac{t_i}{\lambda}\right) \left(\frac{t_i}{\lambda}\right)^k \right) \\ \frac{\partial \ell(k, \lambda)}{\partial \lambda} = -\frac{Dk}{\lambda} + \frac{k}{\lambda} \sum_{i=1}^n \left(\frac{t_i}{\lambda}\right)^k \end{array} \right. \left\{ \begin{array}{l} \frac{\partial \ell(k, \lambda)}{\partial \lambda} = 0 \Leftrightarrow \hat{\lambda} = \left(\frac{\sum_{i=1}^n x_i}{D} \right)^{\frac{1}{k}} \\ \frac{\partial \ell(k, \lambda)}{\partial k} = 0 \Leftrightarrow \frac{1}{\hat{k}} = \frac{\sum_{i=1}^n t_i^{\hat{\lambda} \ln(t_i)}}{\sum_{i=1}^n t_i} - \sum_{i=1}^n \frac{d_i \ln(t_i)}{D} \end{array} \right.$$

En simulant 2000 observations issues d'une loi de Weibull censurée à droite, le Mean-Excess plot va posséder la forme suivante :

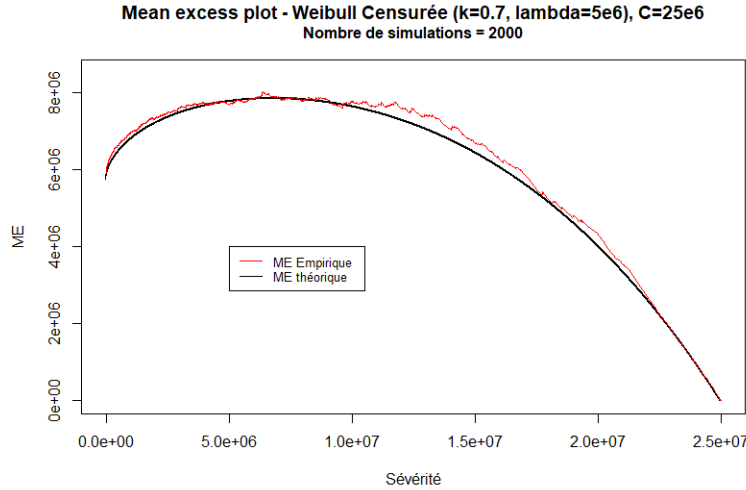


FIGURE 18 – Mean-Excess plot - Loi de Weibull Censurée à droite

Conclusion

La loi de Weibull illimitée possède un Mean-Excess plot très intéressant, sa forme pourrait être en pratique facilement identifiée comme linéaire (ce qui n'est théoriquement pas exactement le cas). L'expression de son Mean-Excess est assez compliquée et les paramètres doivent également être identifiés à l'aide d'un solveur.

Les alternatives Weibull censurée à droite et Weibull tronquée à droite présentent un Mean-Excess plot avec une forme très similaire, comme un début de spirale logarithmique, partant d'une valeur initiale non nulle et finissant en zéro au point de censure/troncature. Comme pour le cas Pareto et Exponentielle, les différents paramètres peuvent se retrouver de la manière suivante :

- Les paramètres de censure (C) et de troncature (b) sont impossibles à identifier par maximum de vraisemblance, mais ils sont cependant identifiables à l'extrémité droite du Mean-Excess plot, là où la courbe retombe à zéro. En effet, $\mathbb{E}_u^C(C) = \mathbb{E}_u^T(b) = 0$.
- Le seuil A à partir duquel on identifie les lois peut être assimilé au début du Mean-Excess plot lorsqu'une forme de "spirale logarithmique" est reconnaissable.
- Les paramètres pour les lois censurées et tronquées : un algorithme de résolution numérique doit être utilisé (type Newton-Raphson par exemple).

Ces alternatives impliquent de travailler avec la fonction gamma incomplète qui est très dure à identifier. À la fois la forme et l'expression du Mean-Excess sont en pratique impossibles à observer, notamment avec très peu de sinistres. La forme de "demi-spirale" logarithmique peut empiriquement très bien être identifiée à une "cloche" et donc au cas d'une Pareto censurée ou tronquée, ou bien être assimilée à une "demi-parabole" et donc au cas Exponentiel censuré ou tronqué. De plus, une distribution de Weibull tronquée à droite ou bien censurée à droite, fait intervenir plus de paramètres que les autres lois, et leur expression trouvée par maximum de vraisemblance, nécessite l'utilisation d'un solveur.

3.1.4 Conclusion

Les lois censurées à droite et tronquées à droite, rajoutent des paramètres difficilement identifiables. En pratique, nous n'avons que très peu de données dans les valeurs extrêmes, l'identification du paramètre de censure et de troncature devient alors très compliquée, puisque de manière quasiment systématique le Mean-Excess plot nous indique que la dernière valeur observée constitue notre point de censure ou bien de troncature. Ainsi, nous abandonnerons les lois censurées à droite et tronquée à droite pour modéliser la sévérité des sinistres.

De plus, la loi de Weibull entraîne des expressions compliquées dont on ne maîtrise pas bien le comportement, c'est pourquoi dans le reste de notre étude, nous abandonnerons le choix de cette famille de distribution.

3.2 Estimation pour des sinistres à valeur non-fixe

Considérer que l'on puisse avoir des sinistres à valeur non-fixe permettrait d'améliorer l'estimation de la sévérité des sinistres en prenant en compte l'information sous-jacente à chacun des sinistres. Actuellement, chaque sinistre est estimé par une valeur ponctuelle. En prenant comme hypothèse que l'on ait des valeurs non-fixes pour les sinistres, et donc que chaque sinistre puisse prendre ses valeurs dans un intervalle donné, on envisagera deux possibilités :

- Chaque valeur d'un sinistre sera remplacée par différentes valeurs possibles, chacune ayant une probabilité de survenir. D'une valeur initiale estimée pour un sinistre, on aura ici plusieurs situations envisagées avec chacune une pondération discrète.
- Chaque valeur d'un sinistre sera remplacée par une variable aléatoire continue suivant une distribution donnée. C'est-à-dire que le sinistre peut prendre une infinité de valeurs possibles dans le domaine de définition de sa distribution. Les sinistres sont alors représentés par des "intervalles continus".

La situation la plus complète est la deuxième, qui semble envisager plus de possibilités, et généralise la première situation. Le but avec ces différentes possibilités est de modéliser une loi pour chaque sinistre en intégrant le plus d'information possible.

Sur base de ces deux situations, nous allons étudier la fonction de vraisemblance pour la loi Exponentielle et la loi de Pareto afin de déterminer la modélisation finale de la sévérité de l'ensemble des sinistres. On considérera que les sinistres sont définis au-delà de différents seuils. En effet, prenons une cédante qui donne ses sinistres au-delà de 1M, parmi ces sinistres figurent deux sinistres ayant le même seuil, sauf que l'un est survenu en 2010 et l'autre en 2011. Après indexation, les deux sinistres ne seront pas définis à partir du même seuil. Le sinistre survenu en 2010 sera défini au-delà d'un seuil plus grand que celui survenu en 2011 (puisque'il sera indexé par un indice plus grand). Ainsi, si l'on considère un seuil fixe, il est probable selon le seuil fixé que l'on supprime le sinistre survenu en 2011 ; mais si on considère que chaque sinistre à son seuil au-delà duquel il est défini alors on conserve les deux sinistres. C'est pourquoi les lois seront ajustées au-delà de différents seuils, et ceci pour les deux situations expliquées précédemment.

3.2.1 La fonction de vraisemblance

Afin d'estimer les paramètres d'une distribution théorique qui pourrait représenter des observations, on a tendance à considérer la méthode d'estimation par maximum de vraisemblance de manière assez systématique.

Prenons le cas extrême où l'on aurait que peu d'observations et une fonction de vraisemblance possédant une forme générale de dromadaire, avec deux bosses. Il est très probable que le maximum de la vraisemblance soit plus proche d'une des deux bosses alors que la médiane de la vraisemblance sera proche du milieu entre les deux bosses très probablement dans le "no man's land". Il serait ainsi peut-être préférable de considérer la médiane de vraisemblance plutôt que le maximum de vraisemblance. Cela permettrait d'avoir une équiprobabilité de sous-estimer ou de sur-estimer le paramètre. En prenant la médiane, on prendrait certes une valeur d'un paramètre moins probable que la valeur du paramètre obtenu par maximum de vraisemblance, mais qui répartirait les probabilités de sous-estimation ou de sur-estimation du paramètre.

Lemme

On introduit ici un lemme sur la loi Gamma nécessaire à la compréhension de la suite de l'étude.

Soit $\Lambda \sim \text{Gamma}(k = K_1 + 1, \theta = \frac{1}{K_2})$ avec $(K_1, K_2) \in \mathbb{N}^{*2}$.

Et "*Median*" la médiane de la vraisemblance du modèle (notée \mathcal{L}).

Par récurrence on peut montrer que : $\forall (a, b) \in \mathbb{R}^+$ avec $a < b$: $\int_a^b e^{-\lambda} \lambda^{K_1} d\lambda = - \left[e^{-\lambda} \sum_{i=0}^{K_1} \lambda^i \frac{K_1!}{i!} \right]_a^b$.

On définit : $\forall \lambda \in \mathbb{R}^+, \mathbb{P}_\Omega(\Lambda = \lambda) = \frac{\mathcal{L}(\lambda)}{\int_0^\infty \mathcal{L}(\mu) d\mu}$ et $\mathbb{P}_\Omega(\Lambda \leq \lambda) = \int_0^\lambda p_\Omega(\Lambda = \mu) d\mu$.

$$* \quad \mathbb{P}_\Omega(\Lambda = \lambda) = \frac{e^{-\lambda K_2} \lambda^{K_1}}{\int_0^\infty e^{-\mu K_2} \mu^{K_1} d\mu} = \frac{e^{-\lambda K_2} \lambda^{K_1}}{\frac{1}{K_2} \int_0^\infty e^{-\mu_0} \left(\frac{\mu_0}{K_2}\right)^{K_1} d\mu_0} = \frac{K_2 e^{-\lambda K_2} (\lambda K_2)^{K_1}}{- \left[e^{-\mu_0} \sum_{i=0}^{K_1} (\mu_0)^i \frac{K_1!}{i!} \right]_0^\infty} = \frac{K_2 e^{-\lambda K_2} (\lambda K_2)^{K_1}}{K_1!}$$

$$* \quad \mathbb{P}_\Omega(\Lambda \leq \lambda) = \int_0^\lambda \frac{K_2 e^{-\mu K_2} (\mu K_2)^{K_1}}{K_1!} d\mu = \int_0^{\lambda K_2} \frac{e^{-\mu_0} (\mu_0)^{K_1}}{K_1!} d\mu_0 = - \frac{1}{K_1!} \left[e^{-\mu_0} \sum_{i=0}^{K_1} (\mu_0)^i \frac{K_1!}{i!} \right]_0^{\lambda K_2}$$

$$= 1 - e^{-K_2 \lambda} \sum_{i=0}^{K_1} \frac{(K_2 \lambda)^i}{i!}$$

$$\begin{aligned}
* \quad \mathbb{E}_\Omega(\Lambda) &= \int_0^\infty \lambda K_2 \frac{e^{-\lambda K_2} (\lambda K_2)^{K_1}}{K_1!} d\lambda = \frac{1}{K_2 K_1!} \int_0^\infty e^{-\lambda_0} (\lambda_0)^{K_1+1} d\lambda_0 = \frac{1}{K_2} \left[e^{-\lambda_0} \sum_{i=0}^{K_1+1} (\lambda_0)^i \frac{(K_1+1)!}{i!} \right]_0^\infty \\
&= \frac{1}{K_2} \frac{(K_1+1)!}{K_1!} \Rightarrow \boxed{\mathbb{E}_\Omega(\Lambda) = \frac{K_1+1}{K_2}}
\end{aligned}$$

$$* \quad \int_0^{Median} p_\Omega(\Lambda = \lambda) d\lambda = 1 - e^{-K_2 Median} \sum_{i=0}^{K_1} \frac{(K_2 Median)^i}{i!} = 0.5$$

Cette expression nécessite l'utilisation d'un solveur, mais en pratique, en utilisant l'approximation introduite par CHOI [1994]⁹, on pourra se contenter de l'approximation : $\boxed{Median = \frac{K_1 + \frac{2}{3}}{K_2}}$.

3.2.2 Cas avec des poids discrets

Nous sommes ici dans le cas où chaque sinistre est représenté par différentes valeurs possibles, chacune ayant une probabilité de survenir (situation 1 de la section 3.2). Et nous allons étudier l'ajustement de la loi Exponentielle et de la loi de Pareto avec des sinistres définis au-delà de différents seuils.

Supposons que l'on ait en notre possession n observations (X_i, T_i, w_i) avec :

- X_i le montant de la i-ème observation.
- T_i la valeur au-dessus de laquelle nous savons qu'elle se trouve.
- w_i le poids discret que nous voulons donner à la i-ème observation.

Loi Exponentielle avec des sinistres définis au-delà de différents seuils

Supposons que nous voulons ajuster une distribution Exponentielle de paramètre d'intensité λ au-dessus d'un certain seuil A. Nous pouvons écrire la fonction de vraisemblance, avec $A_i = \max(A, T_i)$ et $w_i^A = w_i \mathbb{1}_{X_i > A}$ le poids de la i-ème observation au-delà de A, de la manière suivante :

$$* \quad \mathcal{L}(\lambda) = \prod_{i=1}^n \left(\lambda e^{-\lambda(x_i - T_i)} \right)^{w_i \mathbb{1}_{x_i > A}} = \lambda^{\sum_{i=1}^n w_i^A} e^{-\lambda \sum_{i=1}^n w_i^A (x_i - T_i)}$$

$$\text{Ainsi, } \mathcal{L}(\lambda) \sim \text{Gamma} \left(k = 1 + \sum_{i=1}^n w_i^A, \theta = \frac{1}{\sum_{i=1}^n w_i^A (x_i - T_i)} \right)$$

$$\Rightarrow \ell(\lambda) = \ln(\lambda) \sum_{i=1}^n w_i^A - \lambda \sum_{i=1}^n w_i^A (x_i - T_i)$$

$$\text{Et donc, } \frac{\partial \ell(\lambda)}{\partial \lambda} = \frac{\sum_{i=1}^n w_i^A}{\lambda} - \sum_{i=1}^n w_i^A (x_i - T_i) \text{ Et finalement, } \frac{\partial \ell(\lambda)}{\partial \lambda} = 0 \Leftrightarrow \boxed{\hat{\lambda} = \frac{\sum_{i=1}^n w_i^A}{\sum_{i=1}^n w_i^A (x_i - T_i)}}$$

$$* \quad \text{Moyenne de Vraisemblance : } \mathbb{E}(\Lambda) = k\theta = \frac{\sum_{i=1}^n w_i^A + 1}{\sum_{i=1}^n w_i^A (x_i - T_i)}$$

9. CHOI [1994] - cf Bibliographie [9]

$$* \quad \text{Approximation de la Médiane de Vraisemblance : } Median = \frac{\sum_{i=1}^n w_i^A + \frac{2}{3}}{\sum_{i=1}^n w_i^A (x_i - T_i)}$$

Loi de Pareto avec des sinistres définis au-delà de différents seuils

Supposons que nous voulons ajuster une distribution de Pareto de paramètre de forme α au-dessus d'un certain seuil A . Nous pouvons écrire la fonction de vraisemblance, avec $A_i = \max(A, T_i)$ et $w_i^A = w_i \mathbb{1}_{X_i > A}$ le poids de la i -ème observation au-delà de A , de la manière suivante :

$$* \quad \mathcal{L}(\alpha) = \prod_{i=1}^n \left(\frac{\alpha A_i^\alpha}{x_i^{\alpha+1}} \right)^{w_i \mathbb{1}_{x_i > A}} = \alpha^{\sum_{i=1}^n w_i^A} e^{-\alpha \sum_{i=1}^n w_i^A \ln \left(\frac{x_i}{A_i} \right)} \frac{1}{\prod_{i=1}^n x_i^{w_i^A}}$$

$$\text{Ainsi, } \mathcal{L}(\alpha) \sim \text{Gamma} \left(k = 1 + \sum_{i=1}^n w_i^A, \theta = \frac{1}{\sum_{i=1}^n w_i^A \ln \left(\frac{x_i}{A_i} \right)} \right)$$

$$\Rightarrow \ell(\alpha) = \ln(\alpha) \sum_{i=1}^n w_i^A - \alpha \sum_{i=1}^n w_i^A \ln \left(\frac{x_i}{A_i} \right)$$

$$\text{Et donc, } \frac{\partial \ell(\alpha)}{\partial \alpha} = \frac{\sum_{i=1}^n w_i^A}{\alpha} - \sum_{i=1}^n w_i^A \ln \left(\frac{x_i}{A_i} \right) \text{ Et finalement, } \frac{\partial \ell(\alpha)}{\partial \alpha} = 0 \Leftrightarrow \hat{\alpha} = \frac{\sum_{i=1}^n w_i^A}{\sum_{i=1}^n w_i^A \ln \left(\frac{x_i}{A_i} \right)}$$

$$* \quad \text{Moyenne de Vraisemblance : } \mathbb{E}(\alpha) = k\theta = \frac{\sum_{i=1}^n w_i^A + 1}{\sum_{i=1}^n w_i^A \ln \left(\frac{x_i}{A_i} \right)}$$

$$* \quad \text{Approximation de la Médiane de Vraisemblance : } Median = \frac{\sum_{i=1}^n w_i^A + \frac{2}{3}}{\sum_{i=1}^n w_i^A \ln \left(\frac{x_i}{A_i} \right)}$$

3.2.3 Cas avec des poids continus

Nous sommes ici dans le cas où chaque sinistre est représenté par une variable aléatoire continue suivant une distribution donnée (situation 2 de la section 3.2). Et nous allons étudier l'ajustement de la loi Exponentielle et de la loi de Pareto avec des sinistres définis au-delà de différents seuils.

Reprenons les mêmes notations que précédemment, mais cette fois-ci en donnant un "poids continu" à chaque sinistre. C'est-à-dire que : $w_i = \int_0^\infty w_i(x_i) dx_i$ constitue le poids donné au i -ème sinistre.

Loi Exponentielle avec des sinistres définis au-delà de différents seuils

Supposons que nous voulons ajuster une distribution Exponentielle de paramètre d'intensité λ au-dessus d'un certain seuil A . Nous pouvons écrire le log de la fonction de vraisemblance, avec $A_i = \max(A, T_i)$ et $w_i^A = \int_A^\infty w_i(x_i) dx_i$ le poids de la i -ème observation au-delà de A , de la manière suivante :

$$\begin{aligned}
\ell(\lambda) &= \sum_{i=1}^n \int_0^{\infty} \ln \left((f_{A_i}(x_i))^{w_i(x_i) \mathbb{1}_{(x_i > A_i)}} \right) dx_i = \sum_{i=1}^n \int_A^{\infty} \mathbb{1}_{(x_i > T_i)} w_i(x_i) \ln(f_{A_i}(x_i)) dx_i \\
&= \sum_{i=1}^n \int_A^{\infty} \mathbb{1}_{(x_i > T_i)} w_i(x_i) (\ln(\lambda) - \lambda(x_i - A_i)) dx_i \\
&= \ln(\lambda) \sum_{i=1}^n \int_A^{\infty} \mathbb{1}_{(x_i > T_i)} w_i(x_i) dx_i - \lambda \sum_{i=1}^n \int_A^{\infty} \mathbb{1}_{(x_i > T_i)} w_i(x_i) (x_i - A_i) dx_i
\end{aligned}$$

En excluant du "set" des données initiales tous les cas de sinistres en dessous de leur seuil sur la dernière situation connue (Last Known Situation : LKS), les indicatrices $\mathbb{1}_{(x_i > T_i)}$ seront toujours de 1. On peut donc simplifier l'expression de la manière suivante :

$$\ell(\lambda) = \ln(\lambda) \sum_{i=1}^n \int_A^{\infty} w_i(x_i) dx_i - \lambda \sum_{i=1}^n \int_A^{\infty} w_i(x_i) (x_i - A_i) dx_i \Rightarrow \boxed{\ell(\lambda) = \ln(\lambda) K_1^A - K_3^A \lambda}$$

avec K_1^A et K_3^A explicités ci-dessous.

$$* \quad \mathcal{L}(\lambda) = \lambda^{K_1^A} e^{-\lambda K_3^A} \sim \text{Gamma}(k = 1 + K_1^A, \theta = \frac{1}{K_3^A})$$

$$\text{Et nous avons donc, } \frac{\partial \ell(\lambda)}{\partial \lambda} = 0 \Leftrightarrow \frac{K_1^A}{\lambda} - K_3^A = 0 \Leftrightarrow \boxed{\hat{\lambda} = \frac{K_1^A}{K_3^A}}$$

* Expression de K_1^A et K_3^A :

$$\star \quad K_1^A = \sum_{i=1}^n \int_A^{\infty} w_i(x_i) dx_i = \sum_{i=1}^n w_i \int_0^{\infty} \mathbb{1}_{(x_i > A)} \frac{w_i(x_i)}{w_i} dx_i = \sum_{i=1}^n w_i \mathbb{E}_{w_i}(\mathbb{1}_{(x_i > A)}) = \sum_{i=1}^n w_i \mathbb{P}_{w_i}(X_i > A)$$

$$\Rightarrow \boxed{K_1^A = \sum_{i=1}^n w_i^A} \text{ avec } w_i^A = \int_A^{\infty} w_i(x_i) dx_i \text{ On retrouve ici la somme des poids des observations au-delà de } A.$$

$$\begin{aligned}
\star \quad K_3^A &= \sum_{i=1}^n \int_A^{\infty} (x_i - A_i) w_i(x_i) dx_i = \sum_{i=1}^n \int_A^{x_A^i} (x_i - A) w_i(x_i) dx_i + \sum_{i=1}^n \int_{x_A^i}^{\infty} (x_i - T_i) w_i(x_i) dx_i \\
&= \sum_{i=1}^n \int_A^{\infty} (x_i - A) w_i(x_i) dx_i - \sum_{i=1}^n \int_{x_A^i}^{\infty} (x_i - A) w_i(x_i) dx_i + \sum_{i=1}^n (1 - k_i) \int_{x_A^i}^{\infty} x_i w_i(x_i) dx_i
\end{aligned}$$

Si l'on a un modèle multiplicatif de développement, où k_i est le rapport entre la LKS du sinistre i et son seuil en LKS, dépendant uniquement de i , alors :

$$\begin{aligned}
K_3^A &= \sum_{i=1}^n \int_A^{\infty} (x_i - A) w_i(x_i) dx_i - \sum_{i=1}^n \int_{x_A^i}^{\infty} (x_i - A + x_A^i - x_A^i) w_i(x_i) dx_i + \sum_{i=1}^n (1 - k_i) \int_{x_A^i}^{\infty} x_i w_i(x_i) dx_i \\
&= \sum_{i=1}^n \int_A^{\infty} (x_i - A) w_i(x_i) dx_i - \sum_{i=1}^n \int_{x_A^i}^{\infty} (x_i - x_A^i) w_i(x_i) dx_i - \sum_{i=1}^n \int_{x_A^i}^{\infty} (x_A^i - A) w_i(x_i) dx_i + \sum_{i=1}^n (1 - k_i) \int_{x_A^i}^{\infty} x_i w_i(x_i) dx_i \\
&= \sum_{i=1}^n \int_A^{\infty} (x_i - A) w_i(x_i) dx_i + \sum_{i=1}^n x_A^i \int_{x_A^i}^{\infty} w_i(x_i) dx_i - \sum_{i=1}^n \int_{x_A^i}^{\infty} (x_A^i - A) w_i(x_i) dx_i - \sum_{i=1}^n k_i \int_{x_A^i}^{\infty} x_i w_i(x_i) dx_i
\end{aligned}$$

$$K_3^A = \sum_{i=1}^n w_i^A \mathbb{E}_{w_i}(X_i - A | X_i > A) - \sum_{i=1}^n w_i^{x_A^i} \mathbb{E}_{w_i}(X_i - x_A^i | X_i > x_A^i) - \sum_{i=1}^n (x_A^i - A) w_i^{x_A^i} + \sum_{i=1}^n (1 - k_i) w_i^{x_A^i} \mathbb{E}_{w_i}(X_i | X_i > x_A^i)$$

* Moyenne de Vraisemblance : $\mathbb{E}(\lambda) = k\theta = \frac{K_1^A + 1}{K_3^A}$

* Approximation de la Médiane de Vraisemblance : $Median = \frac{K_1^A + \frac{2}{3}}{K_3^A}$

Loi de Pareto avec des sinistres définis au-delà de différents seuils

Supposons que nous voulons ajuster une distribution de Pareto de paramètre de forme α au-dessus d'un certain seuil A . Nous pouvons écrire le log de la fonction de vraisemblance, avec $A_i = \max(A, T_i)$ et $w_i^A = \int_A^\infty w_i(x_i) dx_i$ le poids de la i -ème observation au-delà de A , de la manière suivante :

$$\begin{aligned} \ell(\alpha) &= \sum_{i=1}^n \int_0^\infty \ln \left((f_{A_i}(x_i))^{w_i(x_i) \mathbb{1}_{(x_i > A_i)}} \right) dx_i = \sum_{i=1}^n \int_A^\infty \mathbb{1}_{(x_i > T_i)} w_i(x_i) \ln(f_{A_i}(x_i)) dx_i \\ &= \sum_{i=1}^n \int_A^\infty \mathbb{1}_{(x_i > T_i)} w_i(x_i) \left(\ln(\alpha) - \ln(x_i) + \alpha \ln\left(\frac{A_i}{x_i}\right) \right) dx_i \\ &= \ln(\alpha) \sum_{i=1}^n \int_A^\infty \mathbb{1}_{(x_i > T_i)} w_i(x_i) dx_i - \sum_{i=1}^n \int_A^\infty \mathbb{1}_{(x_i > T_i)} w_i(x_i) \ln(x_i) dx_i + \alpha \sum_{i=1}^n \int_A^\infty \mathbb{1}_{(x_i > T_i)} w_i(x_i) \ln\left(\frac{A_i}{x_i}\right) dx_i \end{aligned}$$

De manière analogue, en excluant du "set" des données initiales tous les sinistres en dessous de leur seuil sur la LKS, les indicatrices $\mathbb{1}_{(x_i > T_i)}$ seront toujours de 1. On peut donc simplifier l'expression de la manière suivante :

$$\begin{aligned} \ell(\alpha) &= \ln(\alpha) \sum_{i=1}^n \int_A^\infty w_i(x_i) - \sum_{i=1}^n \int_A^\infty w_i(x_i) \ln(x_i) dx_i + \alpha \sum_{i=1}^n \int_A^\infty w_i(x_i) \ln\left(\frac{A_i}{x_i}\right) dx_i \\ \Rightarrow \ell(\alpha) &= \ln(\alpha) K_1^A - K_2^A - K_3^A \alpha \end{aligned}$$

* $\mathcal{L}(\alpha) = \alpha^{K_1^A} e^{-\alpha K_3^A} e^{-K_2^A} \sim \text{Gamma} \left(k = 1 + K_1^A, \theta = \frac{1}{K_3^A} \right)$

Et nous avons donc, $\frac{\partial \ell(\alpha)}{\partial \alpha} = 0 \Leftrightarrow \frac{K_1^A}{\alpha} - K_3^A = 0 \Leftrightarrow \hat{\alpha} = \frac{K_1^A}{K_3^A}$.

* Expression de K_1^A , K_2^A et K_3^A :

* $K_1^A = \sum_{i=1}^n \int_A^\infty w_i(x_i) dx_i = \sum_{i=1}^n w_i \int_0^\infty \mathbb{1}_{(x_i > A)} \frac{w_i(x_i)}{w_i} dx_i = \sum_{i=1}^n w_i \mathbb{E}_{w_i}(\mathbb{1}_{(x_i > A)}) = \sum_{i=1}^n w_i \mathbb{P}_{w_i}(X_i > A)$

$\Rightarrow K_1^A = \sum_{i=1}^n w_i^A$ avec $w_i^A = \int_A^\infty w_i(x_i) dx_i$

On retrouve ici aussi la somme des poids des observations au-delà de A .

$$\begin{aligned} \star \quad K_2^A &= \sum_{i=1}^n \int_A^\infty w_i(x_i) \ln(x_i) dx_i = \sum_{i=1}^n w_i \int_0^\infty \mathbb{1}_{(x_i > A)} \ln(x_i) \frac{w_i(x_i)}{w_i} dx_i = \sum_{i=1}^n w_i \mathbb{E}_{w_i}(\ln(X_i) \mathbb{1}_{(x_i > A)}) \\ &= \sum_{i=1}^n w_i \mathbb{P}_{w_i}(X_i > A) \mathbb{E}_{w_i}(\ln(X_i) | X_i > A) \end{aligned}$$

$$\Rightarrow \boxed{K_2^A = \sum_{i=1}^n w_i^A \mathbb{E}_{w_i}(\ln(X_i) | X_i > A)}$$

$$\begin{aligned} \star \quad K_3^A &= \sum_{i=1}^n \int_A^\infty w_i(x_i) \ln\left(\frac{A}{X_i}\right) dx_i = \sum_{i=1}^n \int_A^{x_A} w_i(x_i) \ln\left(\frac{A}{X_i}\right) dx_i + \sum_{i=1}^n \int_{x_A}^\infty w_i(x_i) \ln\left(\frac{T_i}{X_i}\right) dx_i \\ &= \sum_{i=1}^n \int_A^{x_A} w_i(x_i) \ln\left(\frac{A}{X_i}\right) dx_i + \sum_{i=1}^n \ln\left(\frac{T}{X}\right) \int_{x_A}^\infty w_i(x_i) dx_i, \text{ pour un modèle multiplicatif de développement,} \end{aligned}$$

et avec x_A montant au delà duquel $T_i \geq A$.

$$K_3^A = \sum_{i=1}^n \int_A^\infty \ln\left(\frac{A}{X_i}\right) w_i(x_i) dx_i + \ln\left(\frac{T}{X}\right) \sum_{i=1}^n w_i^{x_A}$$

$$\Rightarrow \boxed{K_3^A = \sum_{i=1}^n w_i^A \mathbb{E}_{w_i}\left(\ln\left(\frac{A}{X_i}\right) | X_i > A\right) - \sum_{i=1}^n w_i^{x_A} \mathbb{E}_{w_i}\left(\ln\left(\frac{A}{X_i}\right) | X_i > x_A\right) + \ln\left(\frac{T}{X}\right) \sum_{i=1}^n w_i^{x_A}}$$

$$\star \quad \text{Moyenne de Vraisemblance : } \mathbb{E}(\alpha) = k\theta = \frac{K_1^A + 1}{K_3^A}$$

$$\star \quad \text{Approximation de la Médiane de Vraisemblance : } \text{Median} = \frac{K_1^A + \frac{2}{3}}{K_3^A}$$

3.2.4 Biais des estimateurs

Les estimateurs trouvés par maximum, médiane ou moyenne de vraisemblance ont une expression assez compliquée, dont on ne peut pas déterminer de façon simple et littérale s'ils sont biaisés ou non. Afin de déterminer ce biais s'il existe, nous allons travailler sur base de simulations, dans le cas le plus général possible, c'est-à-dire avec une pondération continue autour de chaque sinistre ainsi que dans le cas moins général, avec une pondération discrète pour chaque valeur possible de chaque sinistre. Pour cela, nous allons prendre peu de sinistres afin de mieux mettre en évidence ce biais s'il existe.

Nous prendrons 15 observations issues d'une loi de Pareto (respectivement d'une loi Exponentielle) et pour chacune des observations, nous simulerons un bruit gaussien autour de ces observations. Pour le cas avec des poids continus, on prendra l'intervalle continu défini par l'observation ainsi que le bruit l'entourant, alors que pour le cas avec des poids discrets, on ne prendra qu'une dizaine de valeurs issue l'observation avec son bruit gaussien, ces valeurs constitueront l'ensemble des valeurs possibles de chaque sinistre.

Sur base de ces simulations, nous calculerons l'estimateur obtenu par maximum de vraisemblance, celui par la moyenne de vraisemblance, celui par approximation de la médiane de vraisemblance ainsi que l'estimateur finalement retenu.

Afin de déterminer un estimateur qui serait le moins biaisé possible, on comparera nos estimateurs avec les estimateurs classiques en prenant une seule valeur possible pour chaque sinistre dans son "range" de possibilités et en calculant la valeur du paramètre obtenue ainsi après maximisation classique de la vraisemblance (voir section 2.3). En réitérant cette opération un nombre de fois suffisamment grand (ici 10.000 réitérations), on envisage ainsi toutes les possibilités qui auraient pu arriver.

L'estimateur "amélioré" présenté ci-dessous a été trouvé de manière intuitive, en "dé-biaisant" l'estimateur comme dans le cas classique où l'on a des sinistres avec des valeurs fixes (voir section 2.3). Pour rappel, nous avons avec les notations précédentes :

- L'estimateur par maximum de vraisemblance $\frac{K_1^A}{K_3^A}$
- L'estimateur par la médiane de vraisemblance : $\frac{K_1^A + \frac{2}{3}}{K_3^A}$
- L'estimateur par la moyenne de vraisemblance : $\frac{K_1^A + 1}{K_3^A}$
- L'estimateur amélioré : $\frac{K_1^A - 1}{K_3^A}$

Les graphiques suivants représentent "la fonction de répartition empirique" de nos observations simulées avec leur bruit ainsi que le seuil à partir duquel nous souhaitons ajuster une loi (Pareto ou Exponentielle). À partir de ce seuil, on représente la fonction de répartition de la loi obtenue pour les différents estimateurs explicités précédemment. L'enveloppe bootstrap est une enveloppe convexe qui représente les fonctions de répartition des lois (Pareto ou Exponentielle) obtenues pour toutes les estimations du paramètre dans le cas classique, en prenant une seule valeur possible pour chaque sinistre dans son "range" de possibilités.

Ajustement d'une loi de Pareto, pour des sinistres ayant des poids continus

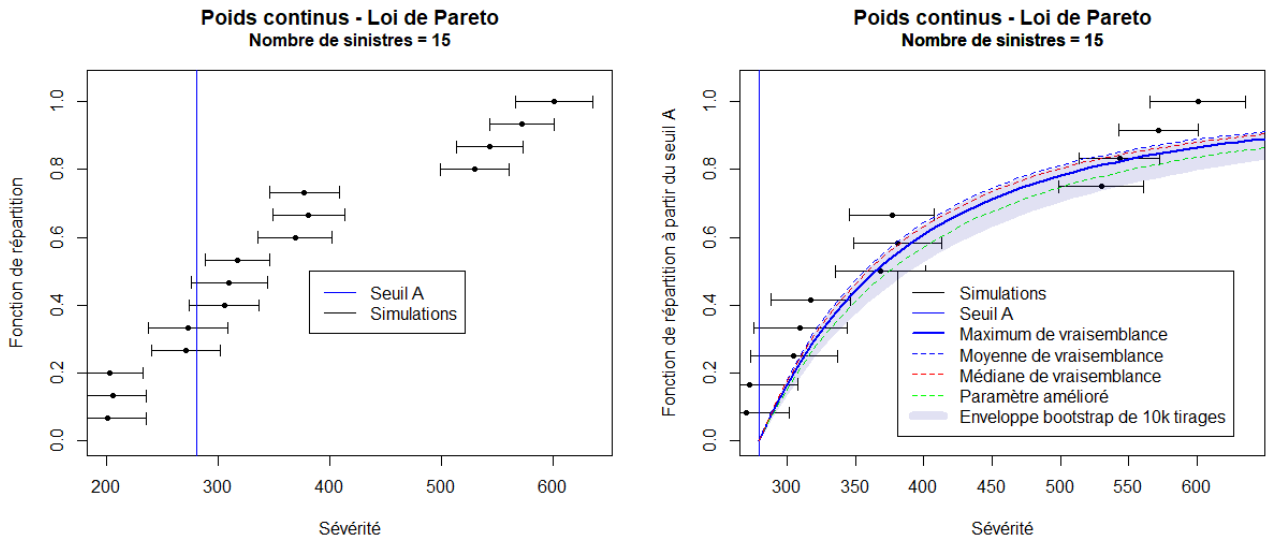


FIGURE 19 – Mise en lumière du biais - Poids continus - Loi de Pareto

Ajustement d'une loi Exponentielle, pour des sinistres ayant des poids continus

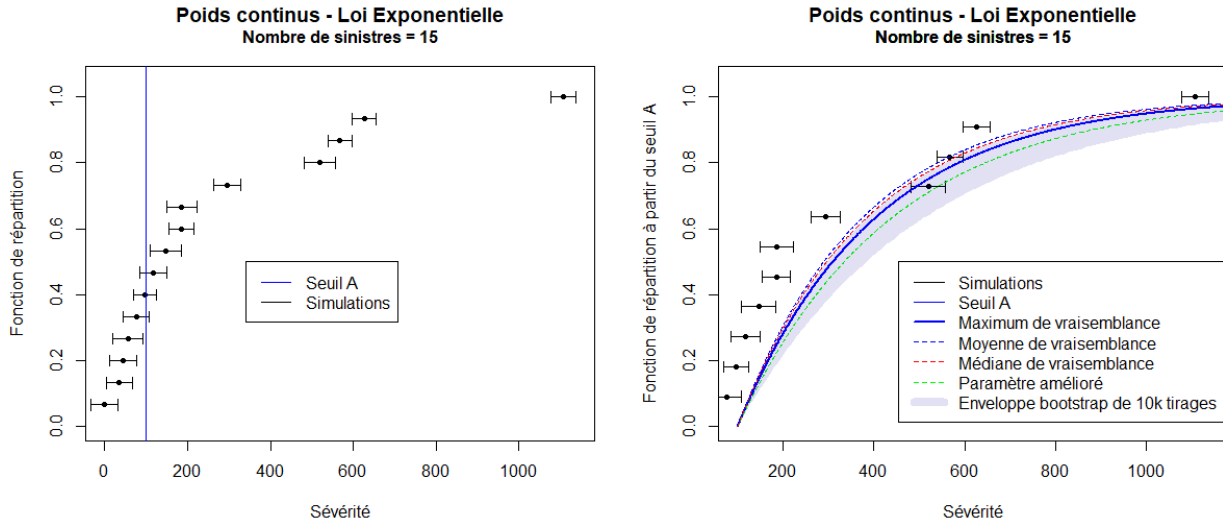


FIGURE 20 – Mise en lumière du biais - Poids continus - Loi Exponentielle

D'après le graphique, il apparaît qu'il faille appliquer cette correction intuitive avec l'estimateur "amélioré" (fonction de répartition en verte sur le graphique), puisque la fonction de répartition de la loi avec l'estimateur trouvé par maximum de vraisemblance (en bleu foncé) semble toujours et de manière systématique être dans la partie supérieure de l'enveloppe bootstrap constituée de nos 10k tirages. Les fonctions de répartition avec les autres estimateurs (moyenne et médiane) possèdent un paramètre plus grand que celui obtenu par maximum de vraisemblance et ont donc une fonction de répartition encore moins sévère. On aurait donc une sur-estimation du paramètre puisque les fonctions de répartition avec les estimateurs trouvés précédemment sont moins sévères en moyenne que les fonctions de répartition de nos échantillons bootstrap. De manière analogue avec des poids discrets pour chaque valeur possible pour chaque sinistre, on obtient un résultat identique :

Ajustement d'une loi de Pareto, pour des sinistres ayant des poids discrets

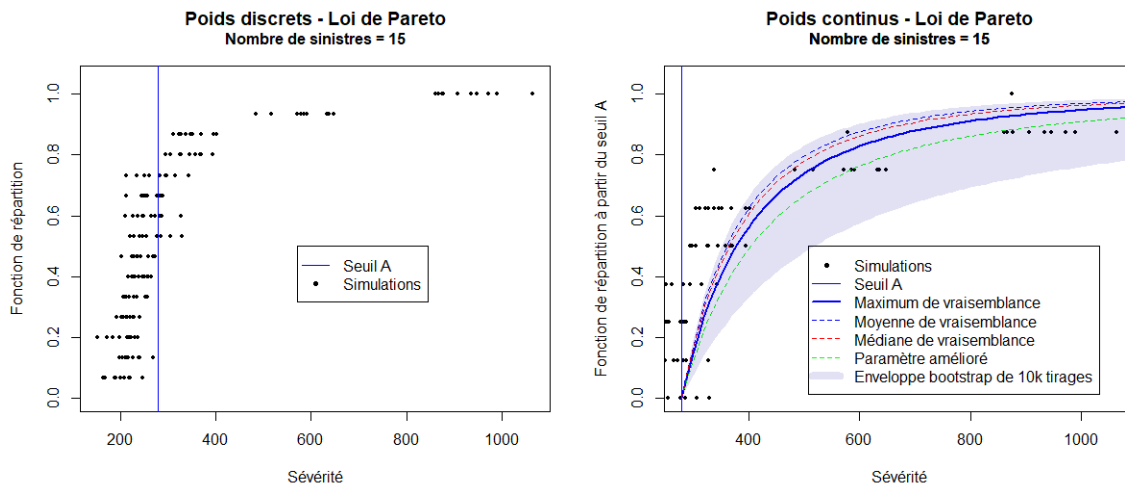


FIGURE 21 – Mise en lumière du biais - Poids discrets - Loi de Pareto

Ajustement d'une loi Exponentielle, pour des sinistres ayant des poids discrets

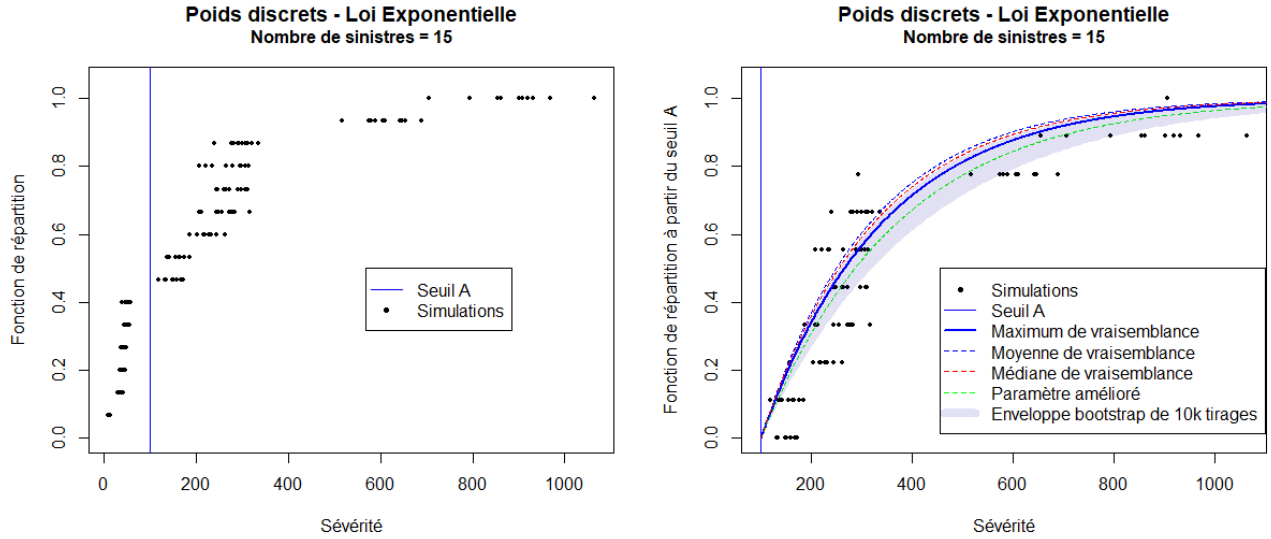


FIGURE 22 – Mise en lumière du biais - Poids - discrets - Loi Exponentielle

Bien évidemment, ici le nombre d'observations choisies est très faible afin de mettre en évidence le biais existant. Si l'on a plus de sinistres qui constituent la Tail, le biais devient négligeable (en supposant que ces sinistres soient bien dans la Tail et que l'on n'a pas descendu le seuil de séparation entre le Body et la Tail trop bas). Et dans ce cas-là, envisager nos différentes méthodes d'estimation n'aurait qu'un très faible impact sur le choix final du paramètre. Prenons par exemple un ajustement avec une distribution de Pareto sur 100 sinistres ayant des poids continus, on obtient alors :

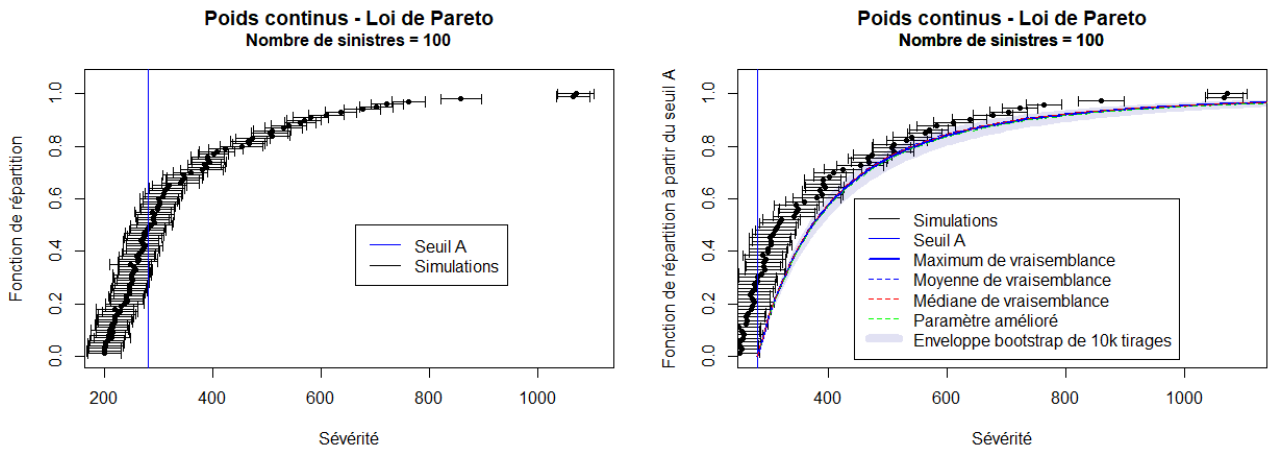


FIGURE 23 – Réduction du biais en augmentant le nombre de sinistres

On voit bien que les courbes des fonctions de répartition des estimateurs sont très proches l'une de l'autre. L'enveloppe bootstrap s'est également réduite, ce qui est logique puisque nous avons ici plus d'observations qui impactent les paramètres obtenus, en les rendant ainsi plus robustes et moins volatils.

3.2.5 Adaptation du Alpha-Plot et Lambda-Plot

Comme dans le cas classique, avec des sinistres fixes, il faut développer des outils permettant de déterminer le seuil, ainsi que les paramètres de la loi à ajuster.

Pour cela, il est repris comme dans le modèle classique avec des sinistres à valeur fixe, la même idée qu'avec le Mean-Excess plot ainsi que le Alpha plot pour une loi de Pareto ou bien le Lambda plot pour une loi Exponentielle (expliqués en section 2.3). Il sera traité ici le cas le plus général possible avec des poids continus ; le cas avec des poids discret découlera simplement de la forme plus générale.

Pour rappel, dans le cas dit classique, le graphique pour obtenir le paramètre de la loi à ajuster est évalué à chacune des observations. Ici, nous avons des sinistres avec des poids continus, ces sinistres sont donc représentés par des intervalles continus. Afin de procéder à l'évaluation du paramètre, il faut donc faire l'évaluation pour une infinité de valeurs pour chaque sinistre. Cela n'étant en pratique pas réalisable, nous discrétiserons les sinistres avec un pas suffisamment petit pour procéder à cette évaluation (on se ramène donc au cas où les sinistres ont des poids discrets et prennent un nombre fini de valeurs dans un intervalle donné). Ainsi, pour chaque sinistre compris dans un intervalle, on va discrétiser ses valeurs avec un pas suffisamment petit. Supposons que l'on ait m sinistres (X_1, \dots, X_m) représentés par des intervalles continus, on prend un pas de discrétisation permettant d'avoir 100 valeurs possibles par sinistre, nos observations deviennent alors : $(u_1, \dots, u_{100m}) = (u_1, \dots, u_n)$ avec $n = 100m$.

De manière analogue au cas avec des sinistres fixes, on va afficher l'estimateur du paramètre de la loi à ajuster à partir du seuil de chacune de nos observations discrétisées, et ceci avec l'estimateur amélioré explicité précédemment. L'analogie est donc faite de la manière suivante :

Soit (u_1, \dots, u_n) les valeurs discrétisées ordonnées par ordre croissant, qui sont les seuils au-delà desquels nous allons faire une estimation.

$$X_{n-k} \longrightarrow \int_0^{\infty} x_{n-k} w_{n-k}(x_{n-k}) dx_{n-k} \text{ le } n\text{-k ème sinistre avec un poids de } \int_0^{\infty} w_{n-k}(x_{n-k}) dx_{n-k}$$

Ajustement d'une loi de Pareto, pour des sinistres ayant des poids continus

$$\forall k \in 1, \dots, n-1, \hat{\alpha}_k = \frac{\sum_{i=1}^k \int_{u_{n-k}}^{\infty} w_{n-i}(x_{n-i}) dx_{n-i} - 1}{\sum_{j=1}^k \ln \left(\int_{u_{n-k}}^{\infty} x_{n-j+1} w_{n-j+1}(x_{n-j+1}) dx_{n-j+1} \right) - \ln \left(\int_{u_{n-k}}^{\infty} x_{n-k} w_{n-k}(x_{n-k}) dx_{n-k} \right)}$$

Ajustement d'un loi Exponentielle, pour des sinistres ayant des poids continus

$$\forall k \in 1, \dots, n-1, \hat{\lambda}_k = \frac{\sum_{i=1}^k \int_{u_{n-k}}^{\infty} w_{n-i}(x_{n-i}) dx_{n-i} - 1}{\sum_{j=1}^k \int_{u_{n-k}}^{\infty} x_{n-j} w_{n-j}(x_{n-j}) dx_{n-j}}$$

Il ne reste plus qu'à tracer les points : $(u_{n-k}, \hat{\alpha}_k)$ et $(u_{n-k}, \hat{\lambda}_k)$, pour $k=1, \dots, n-1$.

Si la forme du graphique est constante à partir d'un certain seuil, alors il semblerait qu'on puisse ajuster la loi correspondante à partir de ce seuil.

Illustrons cet outil avec l'ajustement d'une loi de Pareto :

Simulons des observations issues d'une loi de Pareto et pour chacune des observations, nous simulerons un bruit gaussien autour de ces observations. Afin de comparer l'outil avec celui utilisé actuellement dans le cas de sinistres à valeur fixe, on regardera le graphique classique d'estimation des paramètres, en prenant comme valeur fixe de chaque sinistre son observation simulée sans le bruit gaussien qui l'entoure (qui est égal environ à sa moyenne sur son intervalle puisque le bruit simulé autour de l'observation est centré).

Soit (X_1, \dots, X_n) les valeurs simulées des sinistres, ordonnées par ordre croissant.

Loi de Pareto, pour des sinistres à valeur fixe

$$\forall k \in \{1, \dots, n-1\}, \hat{\alpha}_{k,\text{fixe}} = \frac{k-1}{\sum_{j=1}^k \ln\left(\frac{X_{n-j+1}}{X_{n-k}}\right)}$$

Et dans le cas de l'ajustement d'une loi Exponentielle, on aurait :

Loi Exponentielle, pour des sinistres à valeur fixe

$$\forall k \in \{1, \dots, n-1\}, \hat{\lambda}_{k,\text{fixe}} = \frac{k-1}{\sum_{j=1}^k X_{n-j}}$$

D'après les graphiques suivants, pour des observations issues d'une loi de Pareto (avec un bruit gaussien autour des observations), on observe un Alpha plot approximativement constant, avec bien évidemment une certaine volatilité surtout pour les valeurs extrêmes. Et la valeur d'Alpha pour des sinistres à valeur non-fixe reste bien cohérente avec la valeur d'Alpha ($\hat{\alpha}_{k,\text{max}}, \forall k \in \{1, \dots, n-1\}$) dans le cas où les sinistres sont représentés par une valeur fixe. Par exemple, pour des observations issues d'une loi différente, comme une loi Uniforme (avec un bruit gaussien autour des observations), le Alpha plot n'est pas du tout constant ce qui nous indique que nous n'avons pas une loi de Pareto.

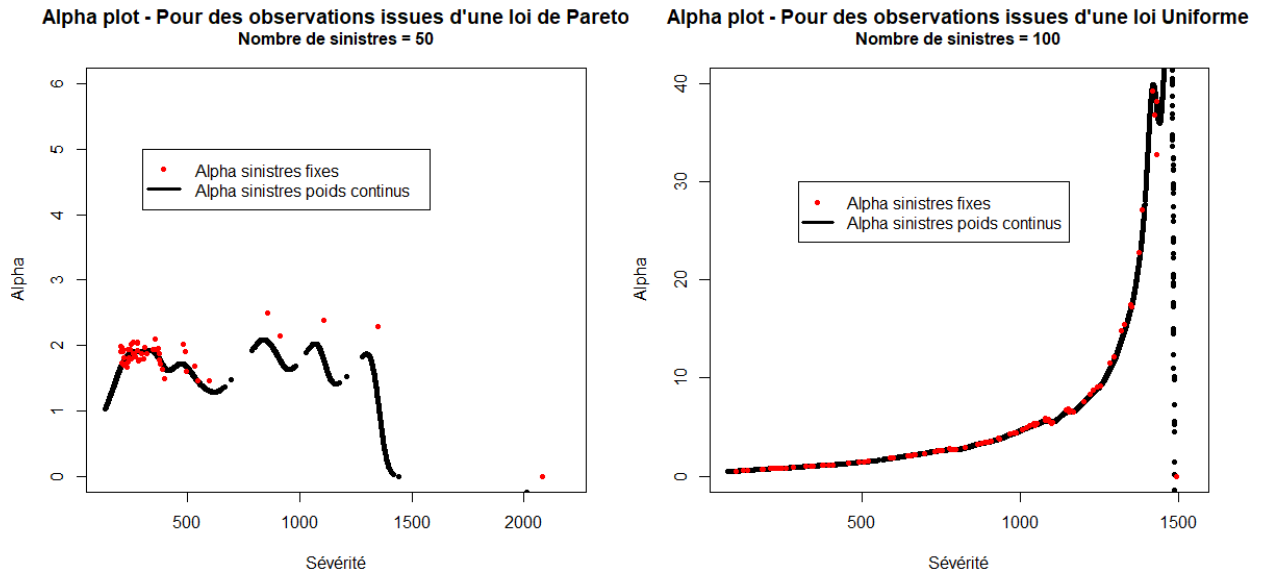


FIGURE 24 – Alpha plot

3.2.6 Adaptation du Mean-Excess Plot

De manière analogue, nous allons adapter le Mean-Excess plot. En pratique, il était pris l'estimateur empirique suivant : $\forall k \in \{1, \dots, n-1\}$, $E_k = \hat{e}_n(X_{n-k}) = \frac{1}{k} \sum_{i=1}^k X_{n-i+1} - X_{n-k}$, qui devient donc désormais, en reprenant les mêmes notations, et la même discrétisation que précédemment :

$$E_k = \frac{\sum_{i=1}^k \int_{u_{n-k}}^{\infty} x_{n-i+1} w_{n-i}(x_{n-i}) dx_{n-i}}{\sum_{j=1}^k \int_{u_{n-k}}^{\infty} w_{n-j}(x_{n-j}) dx_{n-j}} - \int_{u_{n-k}}^{\infty} x_{n-k} w_{n-k}(x_{n-k}) dx_{n-k}$$

Afin de comparer l'outil avec celui utilisé actuellement, on regardera également le Mean-Excess plot classique dans le cas où les sinistres sont représentés par des valeurs fixes, en reprenant comme valeur fixe de chaque sinistre son observation simulée sans le bruit gaussien qui l'entoure :

$$\forall k \in \{1, \dots, n-1\}, E_{k,\text{fixe}} = \hat{e}_n(X_{n-k}) = \frac{1}{k} \sum_{i=1}^k X_{n-i+1} - X_{n-k}$$

Simulons des observations issues d'une loi de Pareto (respectivement issues d'une loi Exponentielle) et pour chacune des observations, on simulera un bruit gaussien autour de l'observation afin de représenter les sinistres par des intervalles. Les graphiques des Mean-Excess plot obtenus sont les suivants :

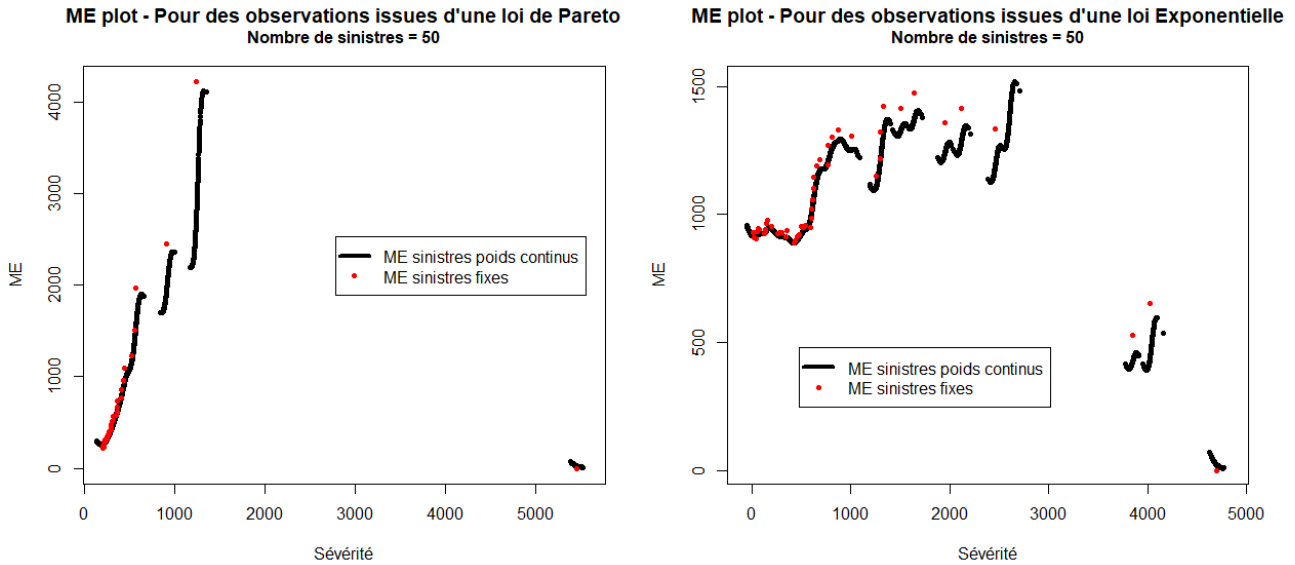


FIGURE 25 – Mean-Excess plot

On retrouve une pente linéaire croissante pour les observations issues d'une loi de Pareto et une courbe constante pour des observations issues d'une loi Exponentielle, avec bien évidemment une certaine volatilité surtout pour les valeurs extrêmes.

Pour le graphique d'obtention des paramètres des lois ou bien le Mean-Excess plot, le cas où l'on n'aurait pas des poids continus, mais des poids discrets pour chaque sinistre se traite de la même manière. En effet, cela revient à avoir un pas de discrétisation plus gros que dans le cas continu. L'évaluation de la distribution des excès et du paramètre des lois est simplement faite aux points de chacune des valeurs possibles des sinistres.

3.2.7 Comparaison des prix pour un contrat Excess of loss (XS)

Considérer des sinistres non-fixes permet de ne pas être trop affecté par le choix du seuil choisi afin d'ajuster la loi de la Tail. Un cas extrême serait par exemple une Tail possédant que très peu de sinistres. Imaginons que l'on ne soit pas capable en regardant les outils graphiques mit à disposition de fixer le seuil de façon précise, passer de 2 à 3 sinistres dans la Tail affecterait énormément la loi à ajuster. Alors qu'avec des sinistres non-fixes pondérés avec des poids continus ou discrets, on capterait l'ensemble des comportements possibles avec par exemple 2.5 sinistres appartenant à la Tail. Il en est de même pour le prix final qui devrait logiquement capter l'ensemble des comportements possibles des prix.

Afin d'illustrer cela, sur base de simulations, nous reprendrons des observations issues d'une loi de Pareto, et pour chacune des observations, nous simulerons un bruit gaussien autour de ces observations. Les lois ici sont donc ajustées sur des sinistres avec des poids continus, mais cette fois-ci avec 40 observations et trois seuils d'ajustement différents. En reprenant les mêmes légendes et notations que précédemment, on appellera "fonction de répartition empirique" ce qui est en réalité plutôt un "range" de fonction de répartition pour des sinistres représentés sous forme d'intervalle, les lois ajustées sont ainsi représentées de la manière suivante :

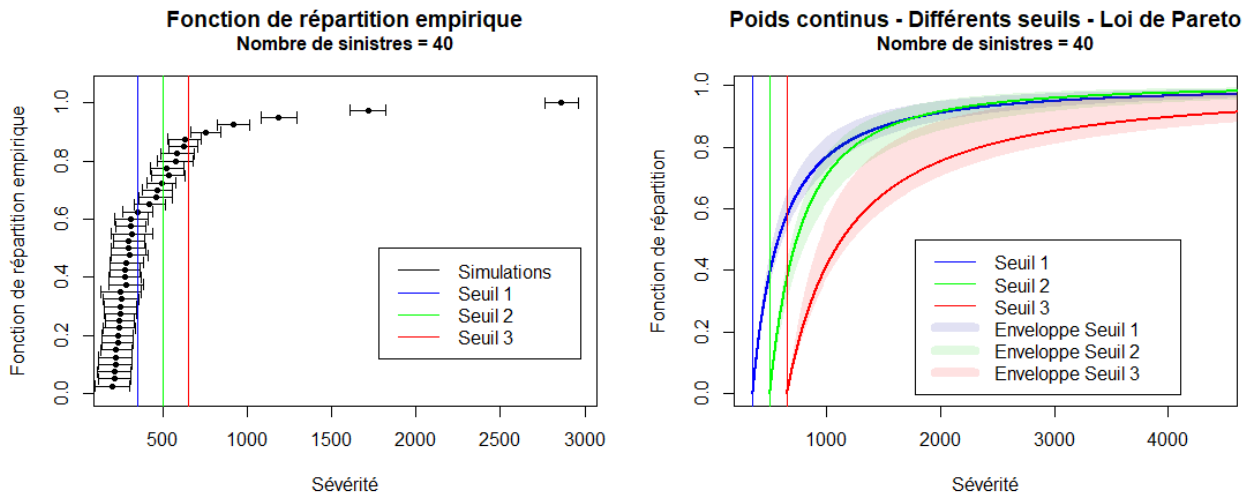


FIGURE 26 – Fonctions de répartition avec différents seuils d'ajustement

Le nombre de sinistres au-delà de chacun des seuils d'ajustement des lois, lorsque l'on a des sinistres sous forme d'intervalle, n'est pas une valeur entière. Ce qui fait une très grande différence avec l'enveloppe bootstrap constituée des 10k tirages, qui elle possède des cas extrémaux avec beaucoup moins ou beaucoup plus de sinistres dépassant le seuil d'ajustement. Pour chacun des seuils précédents, on a ainsi ces fréquences au-delà du seuil :

	Seuil.1	Seuil.2	Seuil.3
Nombre de sinistres au delà du seuil avec des poids continus	15.73769	11.12539	5.429128
Minimum de sinistres au delà du seuil pour l'enveloppe bootstrap	14.00000	9.00000	5.000000
Maximum de sinistres au delà du seuil pour l'enveloppe bootstrap	19.00000	13.00000	8.000000

FIGURE 27 – Nombre de sinistres au-delà du seuil

Pour les trois différents seuils d'ajustement, nous allons regarder les prix pour une tranche d'un Excess of loss de priorité le seuil d'ajustement des lois et de limite infinie. On s'intéresse à la modélisation des grands sinistres, et la tarification des Excess of loss est une matérialisation directe de l'impact de la modélisation des grands sinistres. Il a été choisi des tranches illimitées afin de rester le plus possible dans un cadre général. De plus, il est facile de se ramener à une tranche bornée en calculant la différence de prix de deux tranches illimitées.

Ces tranches illimitées vont nous permettre de comparer les prix obtenus grâce aux sinistres représentés par des intervalles et les prix de l'ensemble des situations envisagées représentées par les enveloppes bootstrap. Pour chacun des seuils d'ajustement, l'histogramme des prix représente la répartition des différentes possibilités des prix des enveloppes bootstrap, et la barre verticale de couleur représente le prix obtenu dans le cas où les sinistres sont représentés par des intervalles :

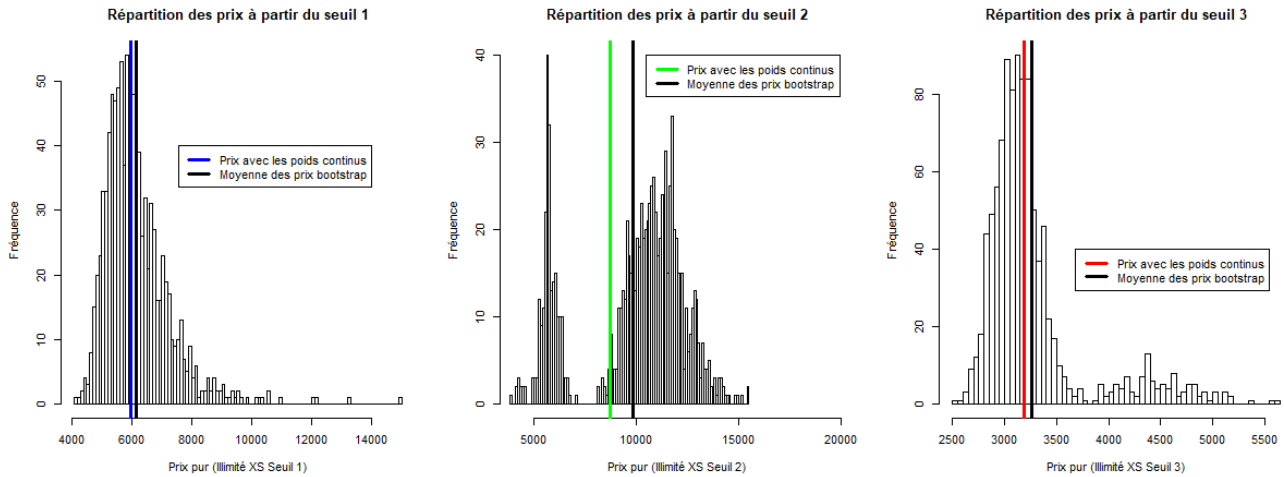


FIGURE 28 – Répartition des prix Excess of loss (+Infini xs Seuil d'ajustement)

Ce qu'il est très important de remarquer, c'est que la méthode capture très bien l'ensemble des comportements possibles. Dans le deuxième graphique, le prix obtenu avec les sinistres représentés avec des poids continus se trouve dans le "no man's land", cette situation représente bien le problème du choix du seuil d'ajustement de la loi de la Tail. En reprenant l'exemple cité dans les paragraphes précédents, ces deux bosses représentent les deux choix possibles lorsque l'on aurait dû choisir entre prendre deux ou bien trois sinistres dans la Tail. La sévérité aurait soit baissé soit augmenté, sans prendre en compte les deux situations possibles. Il en est de même pour le prix qui aurait en quelque sorte été choisi de manière binaire en tranchant pour telle ou telle situation. La méthode développée dans ce chapitre permet ainsi de ne pas avoir à prendre de décision drastique pouvant avoir de fortes répercussions sur le prix pur final mais au contraire, elle permet de prendre en compte l'ensemble des situations possibles. De plus, d'après les deux autres graphiques le prix obtenu avec les poids continus n'est guère loin de la moyenne de l'ensemble des prix, et reflète très bien l'ensemble des situations possibles.

3.2.8 Conclusion

À travers l'étude de la vraisemblance, il a été développé des estimateurs des paramètres des lois (Pareto et Exponentielle) à ajuster à partir d'un certain seuil, et ceci dans le cas où les sinistres seraient représentés par un intervalle. Ces estimateurs que l'on a ajustés en corrigeant le biais asymptotique semblent être des estimations suffisantes des paramètres des lois à ajuster. De plus des outils graphiques ont été élaborés afin d'identifier la loi à ajuster ainsi que le seuil à partir duquel modéliser cette loi. Le but était de prendre en compte des sinistres à valeur variable afin d'avoir une meilleure estimation de la réelle sévérité des sinistres ; et une étude des prix nous a montré que non seulement, il était possible de tarifer sur base de sinistre compris dans un intervalle, mais que de plus, cela permettait de capturer l'ensemble des comportements possibles et réglait ainsi les problématiques du choix du seuil de séparation entre la Tail et le Body. Il ne reste plus qu'à modéliser le reste de la distribution, qui est appelé le "Body".

4 Modélisation du Body de la distribution de sévérité

4.1 Principe et base de donnée

Comme il a été expliqué en section 2.2, en pratique pour le Body de la distribution, les sinistres empiriques sont repris dans leur forme la plus brute. La fonction de répartition utilisée pour une tarification est donc une fonction de répartition empirique, en escalier. Les différents inconvénients de cette modélisation sont nombreux : gros sauts dans la distribution, manque d'observations, irrégularité d'information selon les endroits où l'on se situe dans la distribution empirique du Body... (Voir section 2.2 pour plus de précisions). C'est ce qui amène à établir une nouvelle modélisation afin de pallier ces défauts. Les deux principaux objectifs sont :

- De remplacer cette distribution empirique par une distribution théorique : d'une courbe en escalier, on établirait ainsi une courbe lissée.
- Comme pour la Tail, de considérer que l'on puisse avoir des sinistres à valeur non-fixe et ainsi d'avoir une meilleure estimation de la réelle sévérité des sinistres en prenant en compte l'information sous-jacente à chacun des sinistres. En prenant l'hypothèse que l'on ait des valeurs comprises dans un intervalle de possibilités pour chacun des sinistres, il faudrait établir une méthode de représentation de la distribution de la sévérité du Body.

Dans ce chapitre, il est réalisé une étude théorique pour des sinistres à valeurs fixes afin de mieux comprendre tous les enjeux de la modélisation proposée. Le cas plus général avec des sinistres compris dans des intervalles sera abordé succinctement afin de réaliser une modélisation complète (Tail et Body) de la sévérité de l'ensemble des sinistres sur base de données réelles développées par MOURIER [2020]¹⁰ lors du dernier chapitre de ce mémoire (5).

Les données utilisées ici, sont issues d'une base de données pour une certaine branche d'activités. En annexe A, figure un tableau récapitulatif de cette base de données. Les données auraient très bien pu être simulées, en effet, ce qui nous intéresse ici, c'est d'établir un modèle pouvant s'adapter à n'importe quel type de base de données représentant les sinistres d'un Body quelconque. Ces données seront tronquées à droite et à gauche afin de sélectionner uniquement les sinistres appartenant au Body de la distribution, il a été fait un choix arbitraire de conserver les données partant du quantile à 40% jusqu'au quantile à 95%. De plus, il a été fait le choix de sélectionner une base de données ayant un nombre de sinistres relativement faible (de l'ordre d'une centaine) afin de refléter la problématique du manque de données en réassurance ; et donc de développer une méthode qui pourrait être utilisée à la fois pour une étude de marché complète (avec plusieurs cédantes), mais également pour une seule et unique cédante présentant peu de sinistre. En sélectionnant peu d'observations et une répartition non-uniforme des données, il est plus aisé de tester la robustesse du modèle qui sera envisagé.

4.2 Les mélanges d'Erlangs

Afin de modéliser le Body de la distribution de sévérité des sinistres, il a été choisi d'étudier les mélanges d'Erlangs, il s'agit d'un mélange de lois Gamma avec un paramètre de forme entier.

On considère un mélange de $M \in \mathbb{N}^*$ distributions d'Erlangs avec les notations suivantes :

- $\theta > 0$ le paramètre d'échelle commun à toutes les Erlangs
- $\mathbf{R}=(r_1, \dots, r_M) \in \mathbb{N}^{*M}$ avec $r_1 < \dots < r_M$ les paramètres de forme des distributions d'Erlangs
- $\mathbf{A} = (\alpha_1, \dots, \alpha_M) \in \mathbb{R}^{*+M}$ vérifiant $\sum_{i=1}^M \alpha_i = 1$ les poids de pondération de chaque distribution d'Erlang dans le mélange

10. MOURIER [2020] - cf Bibliographie [1]

La densité de ce mélange peut s'écrire sous la forme :

$$\forall x \in \mathbb{R}^+, f(x; R, \alpha, \theta) = \sum_{i=1}^M \alpha_i \frac{x^{r_i-1} e^{-\frac{x}{\theta}}}{\theta^{r_i} (r_i - 1)!} \mathbb{1}_{(x>0)} = \sum_{i=1}^M \alpha_i f(x; r_i, \theta)$$

Avec des données tronquées, on peut se ramener à un mélange d'Erlangs avec une pondération différente, en effet, avec :

- $\Gamma = (\gamma_1, \dots, \gamma_M) \in \mathbb{R}^{*+}$ vérifiant $\sum_{i=1}^M \gamma_i = 1$ les nouveaux poids de pondération de chaque distribution d'Erlang tronquée
- $(a, b) \in \mathbb{R}^{*+2}$ où $a < b$, les paramètres respectivement de troncature gauche et de troncature droite
- Θ l'ensemble des paramètres (R, θ)
- F la fonction de répartition pour la densité f

La nouvelle fonction de densité peut ainsi s'écrire sous la forme :

$$\begin{aligned} f(x; \Theta, a, b) &= \frac{f(x; \Theta)}{F(b; \Theta) - F(a; \Theta)} = \sum_{i=1}^M \alpha_i \frac{f(x; r_i, \theta)}{F(b; \Theta) - F(a; \Theta)} \\ &= \sum_{i=1}^M \left(\alpha_i \frac{F(b, r_i, \theta) - F(a, r_i, \theta)}{F(b; \Theta) - F(a; \Theta)} \right) \frac{f(x; r_i, \theta)}{F(b, r_i, \theta) - F(a, r_i, \theta)} = \sum_{i=1}^M \gamma_i f(x, b, a, r_i, \theta) \quad (*) \end{aligned}$$

$$\text{Avec } \forall i \in \{1, \dots, M\}, \gamma_i = \alpha_i \frac{F(b, r_i, \theta) - F(a, r_i, \theta)}{F(b; \Theta) - F(a; \Theta)}$$

Pour diverses raisons, cette modélisation a été envisagée :

- Les mélanges d'Erlangs appartiennent à une classe de lois très flexibles, notamment concernant leur forme de distribution très ajustable à des données empiriques. Ces mélanges peuvent se scinder en deux parties de paramétrisation différente : les densités dans le mélange sont spécifiées de manière paramétrique comme des Erlangs, tandis que les poids associés à chaque Erlang forment la partie non-paramétrique. On peut ainsi considérer les Erlangs possédant une pondération significative comme des paramètres de lissage, modulant la forme de la distribution.
- Le mode et donc la valeur où la densité d'une distribution d'Erlang admet un maximum vaut $(r_i - 1)\theta$ ce qui intuitivement, lorsque l'on fixe le paramètre d'échelle commun à toutes les Erlangs, permet de jouer uniquement sur les paramètres de forme afin de représenter les éventuelles bosses de la distribution. On peut donc voir r_i comme un "pas" qui est facilement accessible et réglable très intuitivement.
- Les mélanges d'Erlangs sont denses dans l'espace des distributions positives. Il existe toujours une série de mélanges d'Erlangs qui convergent (au minimum faiblement) vers n'importe quelle distribution positive. Et dans notre cas, la classe des mélanges de distributions d'Erlang ayant une échelle commune est également fermée sous mélange et convolution. De plus par la transformée de Laplace, de nombreuses distributions peuvent être écrites sous la forme d'un mélange d'Erlangs. Une démonstration intuitive de cette densité est faite par TIJMS [1994]¹¹ :

Soit f une fonction de densité d'une distribution positive (et F sa fonction de répartition), montrons qu'elle peut se mettre sous la forme :

$$\forall x \in \mathbb{R}^+, \text{ et } \theta \in \mathbb{R}^+, f_\theta(x) = \sum_{k=1}^{\infty} (F(k\theta) - F((k-1)\theta)) \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k (k-1)!} \mathbb{1}_{(x>0)}$$

11. TIJMS [1994] - cf Bibliographie [10]

Sa fonction caractéristique est donnée par : $\forall x > 0$,

$$\begin{aligned}\psi_\theta(x) &= \sum_{k=1}^{\infty} (F(k\theta) - F((k-1)\theta))(1-i\theta x)^{-k} = \sum_{k=1}^{\infty} \int_{(k-1)\theta}^{k\theta} (1-i\theta x)^{-k} f(u) du \\ &= \sum_{k=1}^{\infty} \int_{(k-1)\theta}^{k\theta} (1-i\theta x)^{-\lceil \frac{u}{\theta} \rceil} f(u) du = \int_0^{\infty} (1-i\theta x)^{-\lceil \frac{u}{\theta} \rceil} f(u) du \text{ avec } \lceil x \rceil \text{ la partie entière supérieure de } x.\end{aligned}$$

Or $|\theta x| < 1$ et $(1-i\theta x)^{-\lceil \frac{u}{\theta} \rceil} \xrightarrow{\theta \rightarrow 0} e^{ixu}$.

$$\text{Donc } \psi_\theta(x) \xrightarrow{\theta \rightarrow 0} \int_0^{\infty} e^{ixu} f(u) du = \psi(x) \text{ fonction caractéristique de } f \Rightarrow \boxed{f_\theta(x) \xrightarrow{\theta \rightarrow 0} f(x)}.$$

4.3 L'algorithme de maximisation de l'espérance (EM)

Afin de modéliser le Body de la sévérité des sinistres, on retient les observations allant de la première observation jusqu'à un certain seuil, on considère ainsi une troncature de l'ensemble de nos observations. On fixe la troncature droite comme le seuil où il a été décidé de modéliser la Tail de la distribution, et la troncature gauche comme le seuil à partir duquel nous avons des observations fournies par les cédantes. De plus, comme l'on souhaite considérer des sinistres prenant des valeurs dans un intervalle donné, il se peut que certaines observations ne soient pas complètement incluses entre la troncature droite et la troncature gauche, ainsi certains sinistres peuvent être censurés soit à gauche soit à droite ou bien des deux côtés dans le cas où un sinistre serait représenté par un très grand intervalle commençant avant la troncature gauche et terminant après la troncature droite.

Pour chaque sinistre $X_i, \forall i \in \{1, \dots, n\}$, on notera l_i la borne minimale de son intervalle et u_i la borne maximale de son intervalle (lorsque l'intervalle est compris entre les deux points de troncature). Les différents cas pourront être explicités de la manière suivante :

- Cas non censuré : $a \leq l_i = u_i \leq b$
- Cas censuré à gauche : $a = l_i < u_i < b$
- Cas censuré à droite : $a < l_i < u_i = b$
- Cas intervalle censuré : $a < l_i < u_i < b$

Prenons par exemple un Body tronqué à : $[a; b] = [10; 20]$. Illustrons un exemple de chacune des situations possibles : une observation non censurée pourra être notée $(l_i; u_i) = (2; 2)$ (un sinistre avec une seule valeur fixe) ; une observation censurée à gauche pourra être notée $(l_i; u_i) = (10; 15)$, une observation censurée à droite pourra être notée $(l_i; u_i) = (16; 20)$ et un intervalle censuré pourra être noté $(l_i; u_i) = (12; 15)$.

Pour l'étude du Body, on ne tiendra compte que de la borne supérieure et inférieure de chaque sinistre contrairement à la Tail où il avait été retenu l'ensemble de la distribution dans l'intervalle de chaque sinistre. En effet, nous avons ici plus d'observations que dans la Tail, on supposera que l'on peut se contenter de deux points pour chaque sinistre. Étant donné la complexité de l'ajustement du mélange d'Erlangs explicité dans les prochains paragraphes, retenir uniquement deux points par sinistre semble plus réaliste pour ajuster le mélange sur nos données.

Le set d'observations devient donc : $\mathcal{X} = \{(l_i; u_i) \text{ pour } i = 1, \dots, n\}$. Notons U, le set des données non censurées et C le set des données censurées. Le logarithme de la vraisemblance du mélange peut s'écrire :

$$\begin{aligned} \ell(\Theta, \mathcal{X}) = & \sum_{i \in U} \ln \left(\sum_{j=1}^M \alpha_j f(x_i, r_j, \theta) \right) + \sum_{i \in C} \ln \left(\sum_{j=1}^M \alpha_j (F(u_i, r_j, \theta) - F(l_i, r_j, \theta)) \right) \\ & - \sum_{i=1}^n \ln \left(\sum_{j=1}^M \alpha_j (F(a, r_j, \theta) - F(b, r_j, \theta)) \right) \quad (**) \end{aligned}$$

Cette expression est très compliquée à optimiser compte tenu du nombre important de paramètres à déterminer. C'est pourquoi l'algorithme de Maximisation de l'Espérance (EM) sera utilisé ici. Une étude complète a été menée par ANTONIO et al. [2014]¹² pour ce type de mélange avec des données censurées et tronquées.

Dans ce mémoire, nous n'explicitons que l'idée sous-jacente à la méthode ainsi que les étapes importantes afin de comprendre comment optimiser ce type de mélange à l'aide de l'algorithme EM.

Le nombre important de paramètres à optimiser amène l'algorithme EM à compléter l'information contenue dans la vraisemblance afin de pallier le manque d'information. On considère l'échantillon censuré comme étant incomplet, les observations non censurées $\mathbf{x} = (x_1, \dots, x_n)$ sont associées ainsi à des variables indicatrices $\mathbf{z} = (z_1, \dots, z_n)$ avec :

$$z_{ij} = \begin{cases} 1 & \text{si l'observation } x_i \text{ appartient à la } j\text{-ème densité } f(x, r_j, \theta) \\ 0 & \text{sinon} \end{cases}$$

Les variables z_1, \dots, z_n sont considérés comme des valeurs observables des variables aléatoires, indépendantes et identiquement distribuées Z_1, \dots, Z_n où $Z_1, \dots, Z_n \sim \text{Multinomial}(1, \alpha)$ avec $\sum_{i=1}^M \alpha_i = 1$.

Le vecteur des données complétées, $\mathcal{Y} = (x_1, \dots, x_n, z) = \{(x_i; z_i) \text{ pour } i = 1, \dots, n\}$ contient ainsi toutes les observations non censurées x_i et leur variable indicatrice z_i correspondant.

En reprenant l'expression (*) en section 4.2, les fonctions des densités $f(x_i, a, b, r_j, \theta)$ sont des versions tronquées des fonctions des densités originales $f(x_i, r_j, \theta)$. Les poids γ_j sont des versions re-pondérées des poids originaux α_j à l'aide de la probabilité que la composante correspondante se trouve dans l'intervalle de troncature. Le logarithme de la vraisemblance de l'échantillon complet \mathcal{Y} devient donc :

$$\ell(\Theta, \mathcal{Y}) = \sum_{i=1}^n \sum_{j=1}^M z_{ij} \ln (\gamma_j f(x_i, a, b, r_j, \theta))$$

Cette expression est plus simple que le logarithme de la vraisemblance incomplète de la formule (**) car elle ne contient pas de logarithmes de sommes. Ensuite, deux étapes sont réalisées pour l'algorithme EM (pour plus de détails sur ces étapes, le lecteur pourra regarder l'étude complète de ce type de mélange¹³). L'idée principale est :

- Une étape appelée "E" (pour "Espérance") qui consiste à calculer l'espérance de la vraisemblance en prenant en compte les variables complétées.
- Une étape appelée "M" (pour "Maximisation") qui consiste à maximiser l'espérance de la vraisemblance trouvée à l'étape E en estimant par maximum de vraisemblance les paramètres.

Les paramètres obtenus à l'étape M sont alors réutilisés afin de débiter une nouvelle estimation en procédant une nouvelle fois à l'étape E, et ainsi de suite. On fixe un critère d'arrêt selon la précision d'estimation souhaitée et la boucle est réitérée jusqu'à ce que la vraisemblance du modèle n'évolue plus par rapport au critère d'arrêt.

12. ANTONIO et al. [2014] - cf Bibliographie [11]

13. ANTONIO et al. [2014] - cf Bibliographie [11]

Initialisation

Il faut d'abord fixer le nombre M d'Erlangs constituant le mélange. Les paramètres de forme (r_1, \dots, r_M) sont fixés initialement avec $r_j = j$ puis selon la disparité qu'on souhaite leur donner, on applique un facteur de dispersion noté s de la manière suivante $r_j = s * j$. Cette méthode a été introduite par LEE S. et LIN S., un lecteur intéressé pourra se référer à leur article sur les mélanges d'Erlangs¹⁴. Un meilleur ajustement des paramètres de forme sera expliqué dans la partie suivante.

Pour l'initialisation du paramètre d'échelle commun θ_0 , on le prend tel que $\theta_0 * r_M = \max(\text{l'ensemble des observations})$. Et les paramètres de pondération $(\alpha_1, \dots, \alpha_M)$ sont initialisés en prenant $\forall j \in \{1, \dots, M\}, \alpha_j$ comme la fréquence relative du nombre d'observations dans l'intervalle $[r_{j-1}\theta_0; r_j\theta_0]$. L'idée est assez intuitive, en effet, on a vu que le mode et donc la valeur où la densité d'une distribution Erlang admet un maximum vaut $(r-1)\theta$. Cette initialisation permet ainsi intuitivement de partir d'un mélange le plus représentatif que possible de nos observations.

Ajustement des paramètres de forme

Dans les étapes E et M, l'algorithme EM n'ajuste pas les paramètres de forme (r_1, \dots, r_M) . En effet, si l'on avait rajouté ces paramètres à optimiser, même en complétant l'information contenue dans la vraisemblance du modèle, le nombre de paramètres aurait été trop grand pour pouvoir tous les optimiser. On fixe donc ces paramètres de formes comme cela est expliqué lors de l'initialisation, puis l'on applique l'algorithme EM en ajustant le paramètre d'échelle et les paramètres de pondération ; et l'on va ensuite faire varier les paramètres de forme et comparer à chaque fois le logarithme de la vraisemblance du modèle afin de conserver le modèle le plus vraisemblable. Cette variation des paramètres de forme est effectuée de la manière suivante (introduit par LEE S. et LIN S. [2010]¹⁵) :

- On exécute l'algorithme EM avec les paramètres de forme $(r_1, \dots, r_M + 1)$. On réitère l'algorithme tant que la vraisemblance du modèle continue de s'améliorer. Puis l'on applique cette procédure à la $(M-1)$ -ème composante, on appliquant ainsi l'algorithme EM avec les paramètres de forme $(r_1, \dots, r_{M-1} + 1, r_M^*)$ où r_M^* est le nouveau paramètre obtenu précédemment. Et ainsi de suite jusqu'à ce que tous les paramètres de forme soient traités.
- Puis l'on refait exactement la même opération mais cette fois-ci en appliquant la correction $(r_1 - 1, \dots, r_M)$, et en décalant la procédure ensuite à la 2ème composante.
- On répète les deux étapes que l'on vient d'expliquer jusqu'à ce que la vraisemblance du modèle ne s'améliore plus (en comparaison au critère d'arrêt fixé au préalable).

Par cet algorithme, on atteint un maximum local de la vraisemblance du modèle, ce qui nous indique que continuer à faire varier les paramètres de forme n'améliorera pas la vraisemblance du mélange.

4.4 Ajustement avec différents nombres d'Erlangs dans le mélange

La base de données utilisée comporte 183 observations fixes, et une fois la troncature appliquée allant du quantile à 40% au quantile à 95%, il ne reste plus que 90 observations qui vont constituer le Body de la distribution.

Nous n'avons initialement pas d'informations sur le nombre d'Erlangs optimal afin de modéliser les observations, c'est pourquoi dans toute cette partie les différentes études menées sur l'ajustement du mélange d'Erlangs seront faites avec des mélanges d'Erlangs comportant différents nombres d'Erlangs. De plus, afin

14. LEE S. et LIN S. [2010] - cf Bibliographie [12]

15. LEE S. et LIN S. [2010] - cf Bibliographie [12]

de bien comprendre tous les enjeux et aboutissements de cette méthode d'ajustement, les observations utilisées ici ne sont pas représentées par des intervalles et ne sont donc pas censurées. Une application avec des observations représentées par des intervalles sera faite sur des données réelles en section 5.

Les différents mélanges d'Erlangs sont pris pour $M \in \{2, 4, 6, 10, 20, 50\}$. Ces valeurs ont été choisies arbitrairement afin d'étudier l'ajustement pour des mélanges d'Erlangs très différents. Sur base des observations, les ajustements des mélanges obtenus grâce à l'algorithme EM, affichent les densités suivantes :

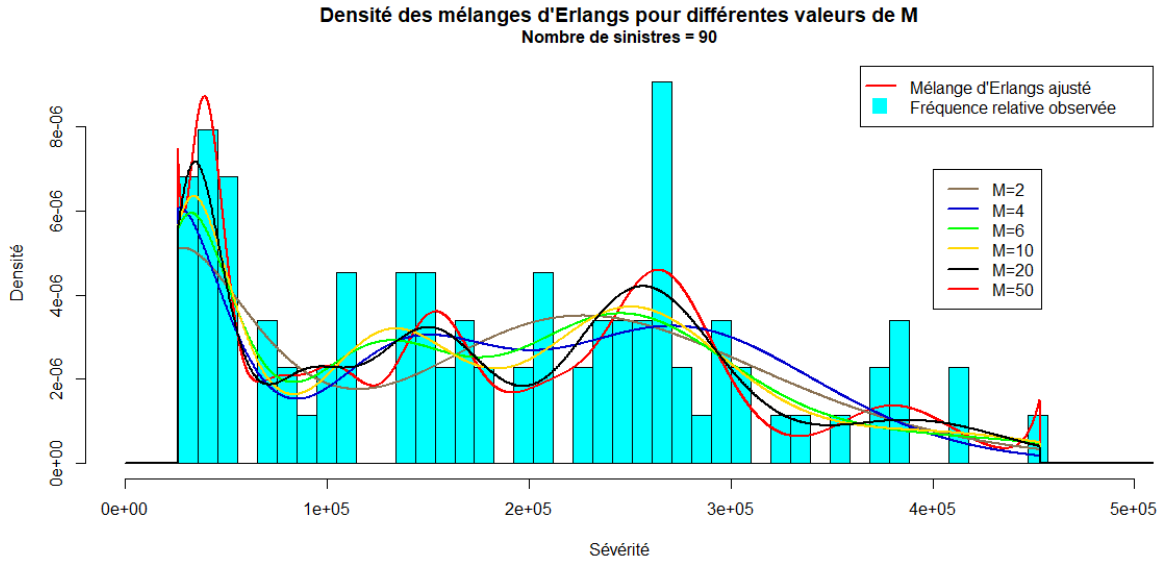


FIGURE 29 – Densité des mélanges d'Erlangs pour différentes valeurs de M

Sur base des mêmes observations, les fonctions de répartition de ces mélanges affichent la forme suivante :

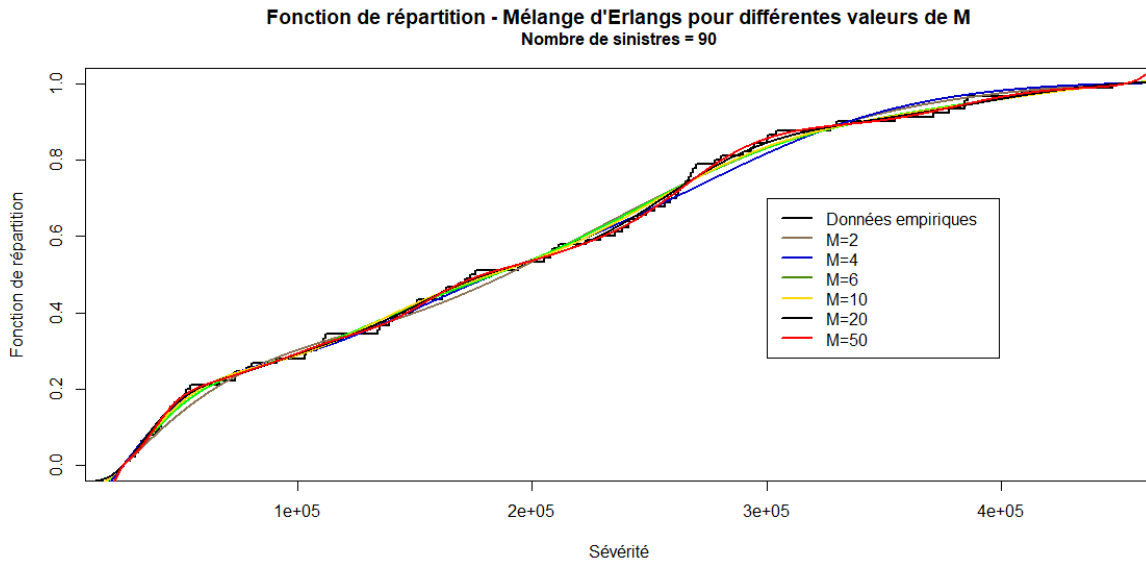


FIGURE 30 – Fonctions de répartition des mélanges d'Erlangs avec différentes valeurs de M

Plusieurs bosses apparaissent dans la répartition des observations. Comme attendu, plus le nombre d'Erlangs dans le mélange est important, plus la courbe a tendance à épouser la courbe empirique représentative des données. Il faudra faire attention à ne pas avoir de sous-ajustement ou de sur-ajustement. Lorsque le nombre d'Erlangs dans un mélange est inférieur au nombre de bosses de la répartition des données, les courbes ne semblent pas très bien représenter l'empirique, en lissant probablement un peu trop la forme générale. Il semblerait donc important à première vue de ne pas sous-estimer le nombre d'Erlangs nécessaire, en écrasant toute l'information contenue dans la représentation empirique, ni de sur-estimer ce nombre en prenant une courbe à peine lissée et beaucoup trop proche de l'empirique. Un sur-ajustement ou un sous-ajustement ne résoudrait pas les problèmes exposés lorsque l'on conserve l'empirique pour modéliser cette partie de la distribution (voir section 2.2 et 2.3).

4.4.1 Temps d'exécution

Le but final est de remplacer la distribution empirique par une distribution de mélange d'Erlangs. Nous avons vu que l'algorithme EM mis en place pour ce type de mélange est complexe et demande beaucoup d'optimisation sur de nombreux paramètres. Dans un contexte d'entreprise, le temps d'exécution de cet algorithme doit être raisonnable afin de permettre aux Analystes d'utiliser cette méthode pour n'importe quelle tarification (que ce soit pour une étude de marché globale ou bien juste pour une cédante). Pour chaque mélange sur notre jeu de données, avec comme nombre respectif d'Erlangs : $M=2, 4, 6, 10, 20$ et 50 ; le temps d'exécution de l'algorithme EM complet en optimisant chacun des paramètres est de :

	M. 2	M. 4	M. 6	M. 10	M. 20	M. 50
Temps d'exécution	3.522514 secs	3.681936 secs	4.907485 secs	8.525589 secs	21.27087 secs	1.34956 mins

FIGURE 31 – Temps d'exécution de l'algorithme EM

L'ordre de grandeur des temps d'exécution est tout à fait convenable, et permettrait l'utilisation de l'algorithme EM dans un cadre d'entreprise. De plus, le nombre d'Erlangs dans un mélange n'est pas destiné à dépasser 50, ici un cas extrême a été conservé seulement afin d'envisager un plus grand nombre de possibilités. Ajuster un mélange avec trop d'Erlangs n'aurait pas vraiment de sens, il y a énormément de paramètres à ajuster en comparaison d'un très faible nombre de sinistres. Les deux paramètres ayant un impact sur le temps d'exécution sont le nombre d'observations ainsi que le nombre d'Erlangs dans le mélange. Et, il est assez intuitif de penser que même en augmentant le nombre d'observations, le temps d'exécution resterait convenable puisque l'on aurait alors une distribution empirique plus robuste et moins d'Erlangs seraient alors nécessaires pour ajuster un mélange.

4.4.2 Paramètres obtenus

Paramètres de pondération $(\alpha_1, \dots, \alpha_M)$

Pour les différents mélanges d'Erlangs, regardons la valeur des paramètres de pondération. Dans le tableau suivant sont répertoriés toutes les valeurs de $(\alpha_1, \dots, \alpha_M)$, avec un seul chiffre significatif, pour les différentes valeurs de M :

	M=2	M=4	M=6	M=10	M=20	M=50
Theta	28622.54	13325.32	10854.02	8418.788	4942.102	2199.701

FIGURE 34 – Paramètre d'échelle commun

Plus on augmente le nombre d'Erlangs dans un mélange, plus θ diminue, ce qui est cohérent avec les valeurs des paramètres de forme trouvées dans la section précédente. En effet, en augmentant le nombre d'Erlangs dans un mélange, on permet une augmentation des valeurs des paramètres de forme; et comme le mode d'une distribution d'Erlang vaut $(r - 1)\theta$; lorsque les paramètres de formes augmentent, la valeur de θ diminue afin de conserver le même mode. Augmenter le nombre d'Erlangs permet ainsi d'avoir un "pas d'ajustement" identifié $\forall j \in \{1, \dots, M\}$ à $(r_j - 1)\theta$ beaucoup plus fin.

Étude des modes

Pour rappel, le mode d'un ensemble de valeurs de données est la valeur qui est la plus souvent représentée dans les données. Lorsque Y est une variable aléatoire continue, le mode est la valeur "y" pour laquelle la densité de probabilité prend sa valeur maximale. Pour les différents mélanges d'Erlangs, regardons la valeur des modes de chaque Erlang dans chaque mélange.

Dans le tableau suivant donc répertorié la valeur de $(r_j - 1)\theta, \forall j \in \{1, \dots, M\}$, pour les différentes valeurs de M en gardant 3 chiffres significatifs :

M. 2	28600	229000																	
M. 4	26700	93300	147000	280000															
M. 6	32600	76000	130000	141000	250000	412000													
M. 10	33700	58900	84200	135000	160000	244000	253000	261000	295000	404000									
M. 20	34600	54400	84000	93900	114000	148000	153000	173000	193000	213000	252000	257000	272000	292000	311000	331000	351000	390000	
M. 50	6600	37400	39600	50600	59400	68200	74800	85800	94600	106000	112000	139000	154000	156000	174000	191000	209000	211000	
M. 2																			
M. 4																			
M. 6																			
M. 10																			
M. 20	435000																		
M. 50	216000	227000	235000	251000	262000	264000	266000	279000	293000	304000	323000	350000	376000	381000	383000	403000	411000	598000	

FIGURE 35 – Modes des mélanges d'Erlangs

En rouge, apparaissent les modes pour lesquels on a jugé que la pondération était significative. Pour résumer ces modes valent pour les différentes valeurs de M :

M. 2	28600	229000			
M. 4	26700	147000	280000		
M. 6	32600	130000	250000		
M. 10	33700	135000	253000		
M. 20	34600	153000	257000	390000	
M. 50	39600	154000	264000		

FIGURE 36 – Modes significatifs des mélanges d'Erlangs

Hormis pour M=2, les modes significatifs de chaque mélange sont très proches. Il apparaît ainsi que 2 Erlangs dans un mélange n'est pas suffisant pour représenter les 3 bosses qui apparaissent dans la répartition des observations. Il faut donc au moins N Erlangs dans un mélange pour bien représenter une distribution possédant N bosses, et plus l'on augmente le nombre d'Erlangs dans le mélange plus l'on affine le modèle et plus l'on détecte d'autres bosses. Il faut cependant faire attention, détecter d'autres bosses qui ne sont pas réellement représentatives des observations n'est pas une bonne chose, on serait dans le cas d'un sur-ajustement.

Si l'on prend la moyenne des modes significatifs obtenus (sans prendre en compte le cas où $M=2$), on obtient à trois chiffres significatifs près : 33.400, 144.000 et 261.000. Plaçons ces moyennes des modes significatifs sur la répartition empirique des données :

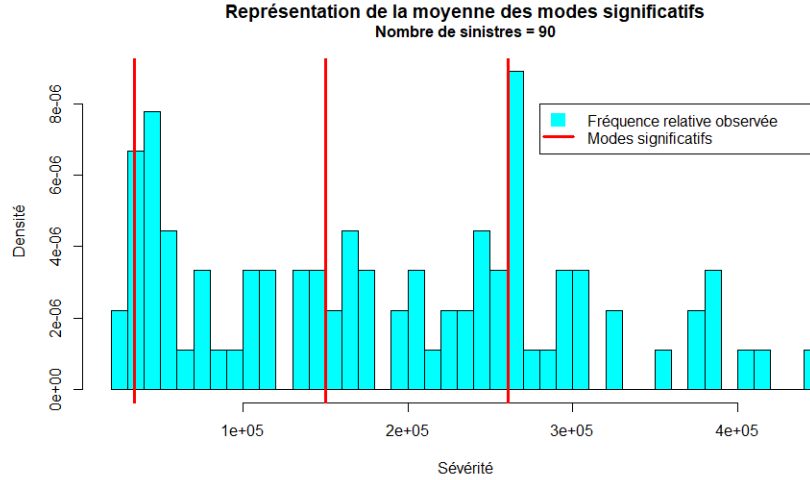


FIGURE 37 – Moyenne des modes significatifs des mélanges d'Erlangs

Ces moyennes se situent bien aux niveaux des bosses de la distribution. Les "modes significatifs" sont donc très représentatifs de la distribution empirique.

Conclusion

D'après l'étude des différents paramètres, il faut donc regarder le nombre de bosses de la distribution empirique que l'on souhaite représenter dans l'ajustement de notre mélange d'Erlangs et prendre ce nombre comme le nombre minimum d'Erlangs composant le mélange. Plus l'on va augmenter le nombre d'Erlangs, plus l'on va détecter "d'autres bosses". Cependant, certaines bosses considérées comme trop petite ne sont pas forcément utiles à représenter. En effet avec très peu de sinistre, on pourrait très bien considérer qu'une petite bosse est uniquement due au hasard et n'est pas représentative. Il ne faut donc pas non plus sur-évaluer le nombre d'Erlangs à prendre dans un mélange mais en conserver un nombre nécessaire et suffisant.

4.4.3 Espérance et Variance

L'ajustement d'un mélange d'Erlangs est obtenu grâce à l'algorithme EM. Cet algorithme repose sur la maximisation de l'espérance, on doit ainsi avoir approximativement l'égalité entre l'espérance empirique et l'espérance théorique du mélange d'Erlangs. La distribution étant tronquée en $(a, b) \in \mathbb{R}^+$ avec $a < b$, l'égalité des espérances se traduit par :

$$\mathbb{E}_{\text{théorique}} = \mathbb{E}_{\text{empirique}} \Leftrightarrow \int_a^b \left(1 - \frac{F(u) - F(a)}{F(b) - F(a)} \right) du = \frac{\sum_{i=1}^N X_i}{N}$$

avec N le nombre de sinistres composant les observations empiriques X_i , et F la fonction de répartition du mélange d'Erlangs.

Il n'existe pas d'expression formelle de la variance et de l'espérance pour une loi Gamma à paramètre entier (si et seulement si une loi d'Erlang) tronquée à droite et à gauche. Pour résoudre cela, on va utiliser la

relation pratique suivante ainsi que RStudio afin de calculer l'espérance et la variance :

Soit $X \sim \text{Gamma}(k, \frac{1}{\theta})$, avec $k \in \mathbb{N}^*$, $\theta \in \mathbb{R}^+$ et F_k sa fonction de répartition. On a alors :

$$\mathbb{E}(X) = \frac{1}{F_k(b) - F_k(a)} \int_a^b x \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} dx = \frac{\theta k}{F_k(b) - F_k(a)} \int_a^b \frac{x^{(k+1)-1} e^{-\frac{x}{\theta}}}{\theta^{k+1} \Gamma(k+1)} dx = \theta k \frac{F_{(k+1)}(b) - F_{(k+1)}(a)}{F_k(b) - F_k(a)}$$

$$\mathbb{E}(X^2) = \frac{1}{F_k(b) - F_k(a)} \int_a^b x^2 \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} dx = \frac{\theta^2 k(k+1)}{F_k(b) - F_k(a)} \int_a^b \frac{x^{(k+2)-1} e^{-\frac{x}{\theta}}}{\theta^{k+2} \Gamma(k+2)} dx = \theta^2 k(k+1) \frac{F_{(k+2)}(b) - F_{(k+2)}(a)}{F_k(b) - F_k(a)}$$

En compilant ces expressions sous RStudio, on obtient pour les différents mélanges d'Erlangs les valeurs de l'espérance et de la variance suivantes (à quatre chiffres significatifs près) :

	Empirique	M=2	M=4	M=6	M=10	M=20	M=70
Espérance	1.863e+05	1.863e+05	1.863e+05	1.863e+05	1.863e+05	1.863e+05	1.863e+05
Variance	1.263e+10	1.196e+10	1.221e+10	1.270e+10	1.278e+10	1.275e+10	1.266e+10

FIGURE 38 – Espérance et variance des mélanges d'Erlangs

Comme attendu, les espérances des différents mélanges et celle de l'empirique sont très proches. Plus le nombre d'Erlangs dans le mélange est grand, plus la variance a tendance à se rapprocher de la variance empiriquement observée. Et plus ce nombre est petit, plus la variance du mélange diminue. Pour un nombre d'Erlangs faible, l'ajustement va très fortement lisser la distribution empirique, or nous avons une grande bosse au début de la distribution (et ce, pour n'importe quelle distribution de Body), il faut donc faire attention à ne pas écraser cette information.

4.5 Tests de sensibilité

Les observations représentent la valeur ultime des sinistres, mais ces sinistres peuvent sensiblement évoluer d'une année sur l'autre. Le but de cette modélisation du Body est d'obtenir la distribution la plus stable possible, et pallier ainsi à la forte variation que les observations empiriques peuvent subir d'une année sur l'autre. Afin de tester la stabilité de l'ajustement par un mélange d'Erlangs, nous choisirons un mélange avec un nombre d'Erlangs qui semble convenable au regard des analyses faites précédemment (ici, on prendra $M=10$), et, nous allons faire deux tests différents de sensibilité :

- Un test de sensibilité point par point à différents endroits stratégiques, notamment aux valeurs extrêmes de la distribution du Body.
- Un test de sensibilité plus général en perturbant de manière aléatoire chacune des observations.

4.5.1 Test de sensibilité aux valeurs extrêmes

Rappelons que le Body est composé ici de 90 sinistres. L'idée est de rajouter 5 sinistres à chaque fois à différentes positions de la répartition des observations. On va rajouter les 5 sinistres à 1/20 ème, puis à 2/20 ème, ..., puis à 20/20 ème du maximum des observations du Body et dans chacun des cas regarder ainsi l'ajustement obtenu par le mélange d'Erlangs.

Il faudra porter une attention toute particulière lorsque l'on rajoute les observations aux endroits extrêmes de la distribution, ce sont en effet eux qui pourraient potentiellement le plus modifier le mélange d'Erlangs ajusté. De plus, comme pour la modélisation pour la Tail de la distribution, la méthode doit permettre de régler la problématique du choix du seuil de séparation entre la Tail et le Body. Il faut donc que lorsque l'on fait varier le nombre de sinistres à l'extrémité droite du Body, le mélange d'Erlangs ajusté ne soit pas trop modifié.

Pour le rajout des 5 sinistres à chacune des positions indiquées, le graphique suivant illustre la fonction de répartition de l'ajustement des mélanges d'Erlangs qui est affiché en vert ; ainsi que la fonction de répartition obtenue par le mélange d'Erlangs lorsque aucun sinistre n'est rajouté à la distribution qui est affichée en rouge :

Lecture du graphique :

- L'axe en largeur (ou abscisse) représente la sévérité.
- L'axe en profondeur (ou ordonnée) représente la position du rajout des sinistres dans la distribution ; position allant de 1/20 ème à 20/20 ème du maximum des observations.
- L'axe en hauteur (ou cote) représente la valeur de la fonction de répartition.

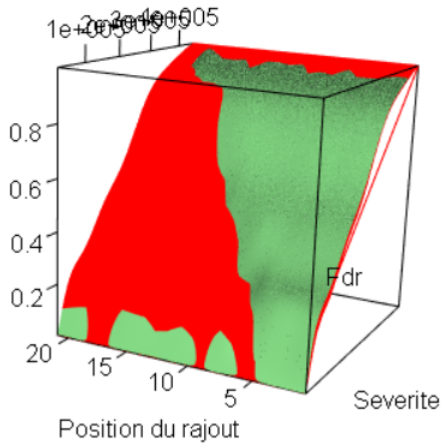


FIGURE 39 – Fonction de répartition - Vue 1

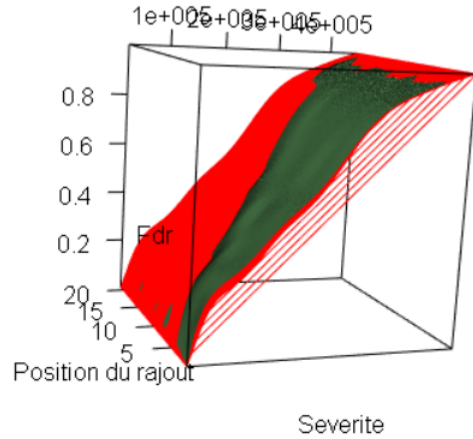


FIGURE 40 – Fonction de répartition - Vue 2

On voit qu'un "triangle" apparaît, séparant la courbe rouge et la courbe verte. La fonction de répartition où les données ont été modifiées (en vert), passe progressivement sous la fonction de répartition avec les données non modifiées (en rouge). En effet, plus l'on rajoute les sinistres tardivement dans la distribution, plus l'on va aplatir la fonction de répartition jusqu'à la position du rajout où la fonction de répartition augmente d'un coup. Plus les données ont été rajoutées loin dans la distribution, plus la courbe verte passe sous la courbe rouge tardivement. L'ajustement du mélange d'Erlangs traduit ainsi exactement ce qui se passe empiriquement.

En reprenant les mêmes notations pour les axes ainsi que le même code couleur, regardons ce qui se passe pour les densités des mélanges :

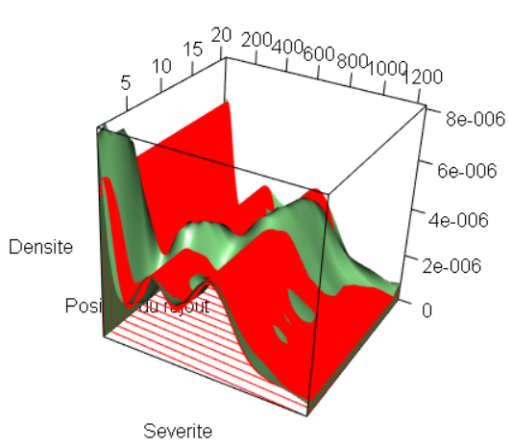


FIGURE 41 – Densité - Vue 1

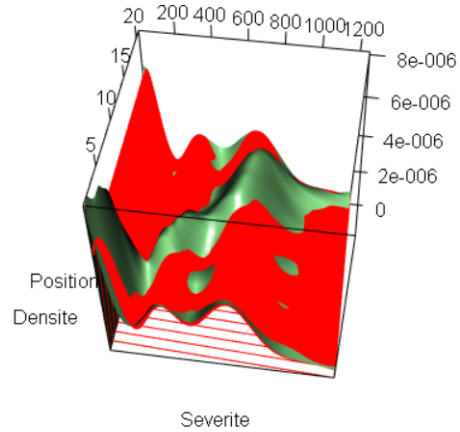


FIGURE 42 – Densité - Vue 2

On observe également ici une "diagonale", où la courbe verte de la densité avec des observations rajoutées dépasse la courbe rouge de la densité sans les observations rajoutées. Cette diagonale correspond bien à la localisation des observations qui ont été rajoutées.

Il est important de noter, qu'à l'extrémité droite de la distribution, l'ajout des 5 sinistres ne modifie pas énormément l'ajustement du mélange d'Erlangs. En effet, la courbe verte passe dessus celle en rouge, mais le dépassement est du même ordre de grandeur qu'aux autres endroits de la distribution. À cause d'un sur-ajustement probable aux extrémités, on aurait pu s'attendre à une explosion de la densité de l'ajustement du mélange aux extrémités du Body. Avec un nombre d'Erlangs dans le mélange jugé convenable (ici, $M=10$), nous n'avons donc pas de problème de "sur-estimation"; mais en augmentant le nombre d'Erlangs dans le mélange, l'ajustement aurait eu tendance à vouloir trop épouser la distribution empirique.

Il semblerait ainsi que la modélisation proposée soit peu dépendante de l'ajout d'observations même si elles sont localisées en fin de distribution. Le choix du seuil de séparation entre la Tail et le Body est donc beaucoup moins impactant sur la distribution du Body que dans le cas classique.

4.5.2 Test de sensibilité générale

Effectuons maintenant un test de sensibilité plus général. Les mêmes données sont reprises et le nombre d'Erlangs constituant le mélange est toujours fixé à $M=10$.

D'une année sur l'autre, les observations des sinistres survenus peuvent sensiblement changer. Chaque année, une nouvelle évaluation de ces valeurs ultimes est faite. Afin de tester la sensibilité de l'ajustement du mélange d'Erlangs, on va reproduire cette variation sur nos observations. En appliquant cette variation, l'idée est de comparer la différence entre l'évolution des observations empiriques et l'évolution induite sur l'ajustement du mélange d'Erlangs.

Afin de modéliser cette variation, on reprend les observations empiriques sur lesquelles on applique un coefficient correcteur multiplicatif qui suit une loi Log-Normale centrée en 1 de variance 0.1. Ainsi chaque observation issue de l'ensemble de la distribution, x_i devient $x_i * y_i$ avec y_i une observation issue d'une loi $\mathcal{LN}(1, 0.1)$. Nous avons choisi une variation qui semblait raisonnable en comparaison aux réels mouvements des sinistres. En appliquant cette correction, certaines observations se trouvant dans l'intervalle du Body de la distribution peuvent en sortir et inversement, le nombre d'observations se trouvant après cette correction dans le Body peut donc être modifié (on peut avoir plus ou moins d'observations que les 90 observations initiales).

L'expérience faite en appliquant une correction issue d'une loi simulée sur chacune des observations sera faite 10 fois, et à chaque fois, on regardera l'ajustement par les mélanges d'Erlangs obtenu.

Le graphique suivant illustre les 10 expériences faites. La fonction de répartition de l'ajustement des mélanges d'Erlangs est affichée en verte ; et la fonction de répartition obtenue par le mélange d'Erlangs lorsque aucune observation n'est modifiée est affichée en rouge. Les axes sont les mêmes que précédemment, sauf pour l'axe de profondeur (ou ordonnée) qui représente donc les 10 expériences réalisées :

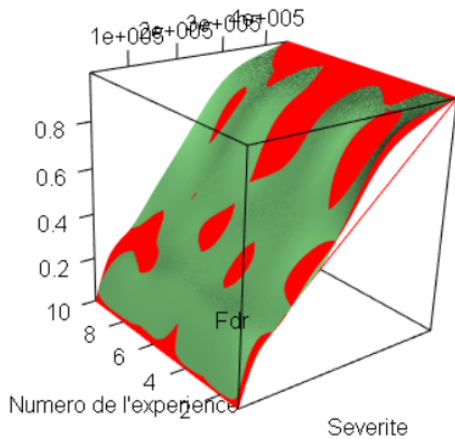


FIGURE 43 – Fonction de répartition - Vue 1

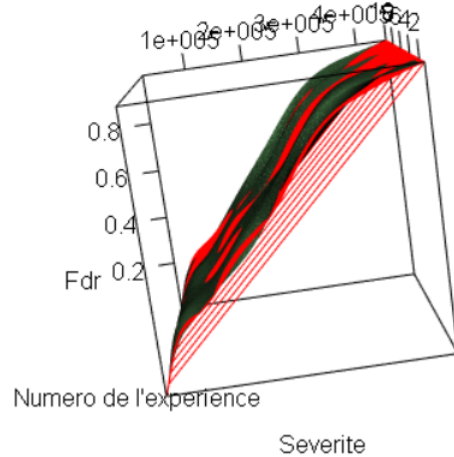


FIGURE 44 – Fonction de répartition - Vue 2

En reprenant les mêmes notations pour les axes ainsi que le même code couleur, regardons ce qui se passe pour la densité des mélanges :

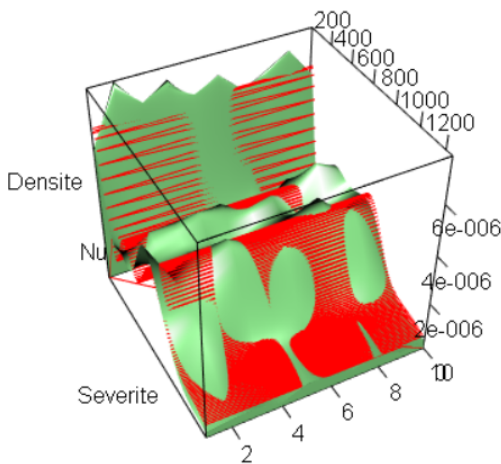


FIGURE 45 – Densité - Vue 1

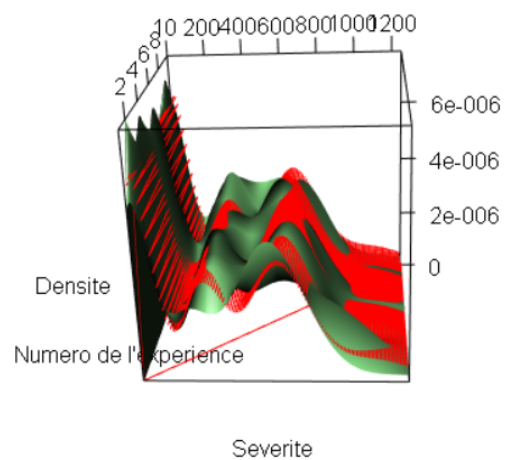


FIGURE 46 – Densité - Vue 2

La variation des ajustements des mélanges d'Erlangs semble raisonnable. Il y a en effet un mouvement apparent de l'ajustement, mais la forme de la densité ainsi que de la fonction de répartition ne semble pas réellement bouger en comparaison avec l'ajustement du mélange d'Erlangs lorsque les observations ne sont pas modifiées.

Lorsque la correction est appliquée, regardons maintenant la différence entre l'évolution de l'empirique et l'évolution de l'ajustement par les mélanges d'Erlangs. En prenant une des expériences réalisées, les deux évolutions sont représentées ci-dessous, grâce à un graphique affichant les densités :

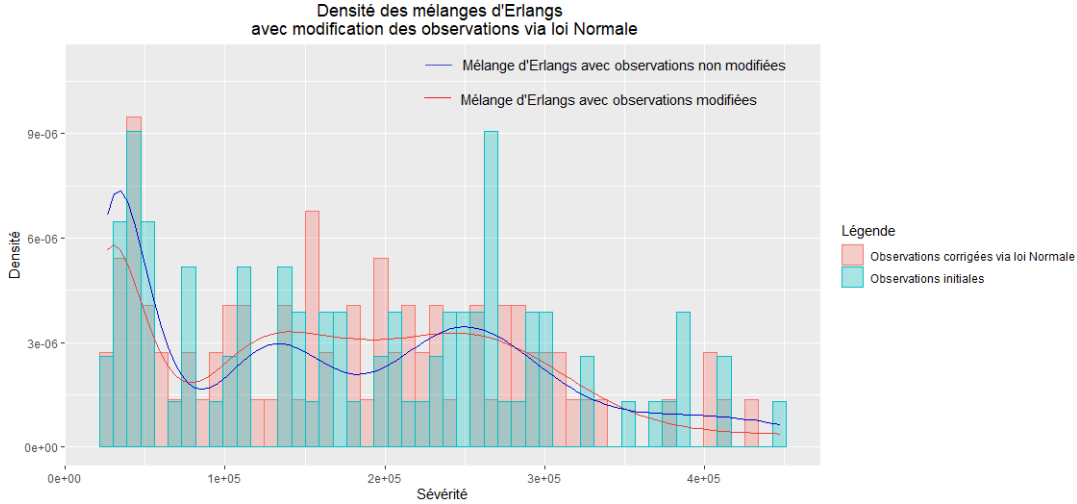


FIGURE 47 – Évolution de la densité des mélanges d'Erlangs

Suite à la correction appliquée sur les observations, l'ajustement du mélange d'Erlangs a bien évidemment changé. Mais en regardant l'histogramme ci-dessus, pour cette expérience, l'empirique a très forte été modifié, alors que l'ajustement est resté dans l'ensemble plutôt stable. De plus, la forme générale des deux ajustements, avant et après modification des données, est la même, avec le même nombre de bosses et la même localisation des modes locaux. L'ajustement des mélanges d'Erlangs permet ainsi de ne pas être trop sensible aux variations que peuvent subir les sinistres d'une année sur l'autre ; variations qui ne sont en réalité pas forcément représentatives de ce qui va se passer dans les années à venir.

4.6 Coût moyen

Pour les différentes valeurs de M constituant le mélange d'Erlangs, nous avons besoin d'un indicateur du coût moyen de la sinistralité sur le Body de la distribution. Cet indicateur de coût moyen va nous permettre selon le différent nombre d'Erlangs dans un mélange de comparer l'écart en terme d'espérance moyenne entre la situation empirique et l'ajustement des mélanges. En effet, nous avons vu que pour les différentes valeurs de M , les espérances des mélanges d'Erlangs ajustés sur l'ensemble de la distribution du Body avaient la même valeur que l'espérance empirique, mais cela n'est pas vrai en tout point de la distribution.

Pour rappel, la distribution du Body est tronquée aux points $(a, b) \in \mathbb{R}^+$ avec $a < b$.

Le rapport de coût moyen ainsi considéré est $\frac{\mathbb{E}(X, u)}{\mathbb{E}(X, b)}, \forall u \in [a, b]$.

Pour les n observations empiriques du Body (x_1, \dots, x_n) ce rapport s'explique de la manière suivante :

$$\frac{\mathbb{E}(X, u)}{\mathbb{E}(X, b)} = \frac{\sum_{i=1}^n x_i \mathbb{1}_{(a \leq x_i \leq u)}}{\sum_{i=1}^n x_i \mathbb{1}_{(a \leq x_i \leq b)}} \frac{\sum_{i=1}^n \mathbb{1}_{(a \leq x_i \leq b)}}{\sum_{i=1}^n \mathbb{1}_{(a \leq x_i \leq u)}}$$

Et pour les ajustements des mélanges d'Erlangs avec différents nombres d'Erlangs, ce rapport s'exprime de la manière suivante :

Soit $X \sim$ Mélange de M Erlangs, et F_{r_i} la fonction de répartition de la loi d'Erlang de paramètre de forme $r_i \in \mathbb{N}^*$, de paramètre de pondération dans le mélange α_i , et de paramètre d'échelle commun $\theta \in \mathbb{R}^+$.

$$\frac{\mathbb{E}(X, u)}{\mathbb{E}(X, b)} \text{ avec, } \forall u \in [a, b] :$$

$$\mathbb{E}(X, u) = \frac{\int_a^{\min(u, b)} x \sum_{i=1}^M \alpha_i \frac{x^{r_i-1} e^{-\frac{x}{\theta}}}{\theta^{r_i} \Gamma(r_i)} dx}{\sum_{i=1}^M \alpha_i (F_{r_i}(b) - F_{r_i}(a))} = \frac{\theta \sum_{i=1}^M \alpha_i r_i (F_{(r_i+1)}(\min(u, b)) - F_{(r_i+1)}(a))}{\sum_{i=1}^M \alpha_i (F_{r_i}(b) - F_{r_i}(a))}$$

Pour les différentes valeurs de M , comparons les courbes obtenues entre le rapport de coût moyen empirique et le rapport de coût moyen des ajustements :

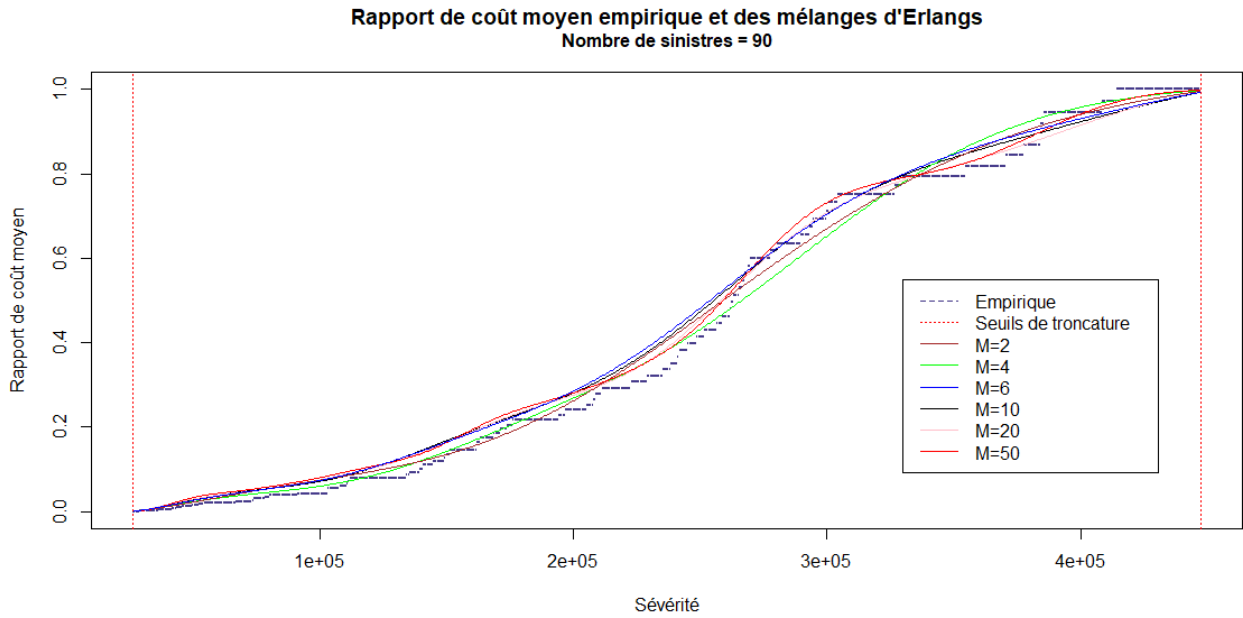


FIGURE 48 – Rapport de coût moyen

Comme attendu théoriquement, au point de troncature droite, le rapport vaut bien 1, et au point de troncature gauche le rapport vaut bien 0. Plus le nombre d'Erlangs dans un mélange est grand, plus la courbe de l'ajustement se rapproche de la courbe empirique. Afin d'avoir une idée des écarts entre l'empirique et les ajustements, affichons, pour chacun des ajustements cet écart relatif en valeur absolue :

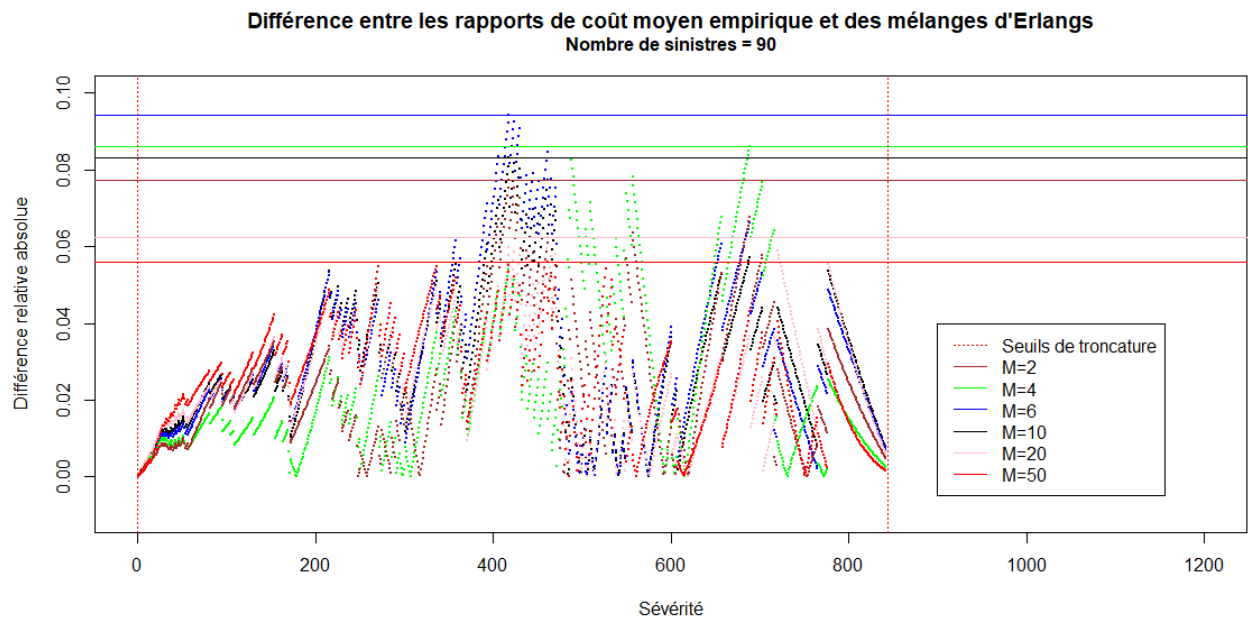


FIGURE 49 – Écart entre les rapports de coût moyen

Pour les différents mélanges d'Erlangs, les lignes horizontales représentent l'écart maximal atteint entre la courbe de l'ajustement du mélange et celle de l'empirique. On voit ici également que plus le nombre d'Erlangs est grand, plus l'écart maximal atteint est faible. Ces écarts maximaux pour les différents mélanges valent :

	M=2	M=4	M=6	M=10	M=20	M=50
Maximum	0.0774206	0.08599485	0.0943312	0.08311887	0.06243051	0.05589461

FIGURE 50 – Écart maximal entre les rapports de coût moyen

Si l'on ne souhaite pas avoir un écart trop important entre le coût moyen de la sévérité de l'empirique et celui de l'ajustement, cet indicateur de coût moyen sur l'ensemble du Body constitue un bon critère pour choisir le nombre d'Erlangs dans le mélange. En effet, on voit ici que pour une faible valeur de M , par exemple $M=6$, l'écart maximal absolu en relatif atteint presque 10% ce qui est vraiment beaucoup, un Analyste voulant opter pour une distribution plus proche de l'empirique et réduire la variabilité du coût moyen sur le Body pourra décider d'augmenter le nombre d'Erlangs dans le mélange.

4.7 Critères d'information

Nous allons utiliser trois critères d'information afin d'affiner la sélection du modèle. Les critères AIC, BIC et AICc sont utilisés pour choisir le meilleur modèle à partir d'un ensemble de modèles plausibles. On considère qu'un modèle est meilleur qu'un autre si le critère (AIC, BIC ou AICc) est plus petit. Ces trois critères s'expriment de la manière suivante :

- $AIC = 2k - 2 \ln(\mathcal{L})$
- $AICc = AIC + \frac{2k(k+1)}{n-k-1}$
- $BIC = \ln(n)k - 2 \ln(\mathcal{L})$

Avec k le nombre de paramètres à estimer du modèle, \mathcal{L} la vraisemblance du modèle et n la taille de l'échantillon.

À savoir que pour un mélange de M Erlangs, nous avons le nombre de paramètres suivant à estimer :

- M paramètres de pondération $(\alpha_1, \dots, \alpha_M)$
- M paramètres de forme (r_1, \dots, r_M)
- 1 paramètre d'échelle commun θ

De plus, la relation $\sum_{i=1}^M \alpha_i = 1$ enlève un paramètre à évaluer, et au final, $M+M+1-1=2M$ paramètres doivent être estimés. Selon la valeur de M , ce nombre de paramètres à évaluer peut être énorme au regard du nombre d'observations constituant le Body de la distribution. Or d'après les expressions des critères précédents, si l'on souhaite pénaliser plus fortement le nombre de paramètres, il serait plutôt préférable de prendre le critère BIC par rapport au critère AIC ou AICc.

Avec les différents mélanges d'Erlangs, pour ces trois critères on obtient les valeurs suivantes :

	BIC	AIC	AICc
M=2	2360.083	2350.083	2350.554
M=4	2350.604	2330.606	2332.384
M=6	2361.295	2331.297	2335.349
M=10	2392.869	2342.873	2355.047
M=20	2468.968	2373.975	2440.914
M=50	2616.988	2437.001	NA

FIGURE 51 – Critères d'information

Il apparaît que le mélange minimisant le BIC posséderait entre 4 et 5 Erlangs, et que celui minimisant les critères AIC et AICc posséderait entre 4 et 6 Erlangs. Ces critères d'information constituent également un bon indicateur sur le nombre d'Erlangs à conserver dans un mélange, en effet on sait que plus ce nombre est important plus l'ajustement est précis, mais aussi que les critères deviennent de plus en plus mauvais. On évite ainsi le sur-ajustement en éliminant les mélanges pour lesquels les critères d'information ont des valeurs trop grandes.

4.8 Méthode de discrimination

Suite à toute cette étude sur l'ajustement d'un mélange d'Erlangs sur des observations, nous devons désormais mettre en place une méthode de discrimination au regard des différents modèles afin de choisir le nombre d'Erlangs le plus adapté pour représenter les observations.

Tout d'abord, les temps d'exécution pour l'ajustement d'un mélange d'Erlangs sont tout à fait raisonnables afin d'utiliser cette méthode pour n'importe quelle tarification (que ce soit pour une étude de marché globale ou bien juste pour une cédante). Afin de choisir le bon nombre d'Erlangs dans le mélange, voici une méthode prenant en compte toutes variables significatives pouvant influencer l'ajustement :

- Regarder le nombre d'observations qui constituent le Body, et convenir pour ce nombre, d'un nombre maximum d'Erlangs. En effet pour un mélange de M Erlangs, on a vu qu'il faut estimer 2^*M paramètres. Considérer un nombre d'Erlangs trop important n'aurait pas vraiment de sens au vu du faible nombre d'observations (par exemple, si l'on a N Erlangs et 2^*N observations, on a au final autant de paramètres à estimer que d'observations, cela pose problème).
→ On convient d'un $M_{\max 1}$.
- Regarder le nombre de bosses dans la distribution qui semblent réellement significatives. On a vu dans la section 4.4.2 avec l'étude des modes qu'il faut prendre pour le nombre d'Erlangs du mélange au moins

le nombre de bosses de la distribution.

→ On convient d'un $M_{\min 1}$.

- Regarder le critère BIC et juger pour quelles valeurs de M , ce critère est minimal. Afin de ne pas négliger certains modèles pas forcément optimaux par rapport au critère BIC mais tout à fait envisageables, on conservera plusieurs modèles avec différentes valeurs de M qui ont le plus faible critère BIC et pas uniquement le modèle qui minimise exactement le critère.

→ On convient d'un $M_{\min 2}$ et d'un $M_{\max 2}$.

- On a vu en section 4.5 en étudiant la sensibilité des ajustements, que les mélanges d'Erlangs sont plutôt stables. Cependant, par mesure de précaution, on pourra étudier la sensibilité de l'ajustement pour les différentes valeurs de M .

- Regarder l'écart maximal et minimal que l'on souhaite avoir du coût moyen sur la distribution du Body entre l'ajustement du mélange et les observations empiriques.

→ On convient d'un $M_{\min 3}$ et d'un $M_{\max 3}$.

- Toutes les variables qui nous ont semblé significatives ont ainsi été abordées, on conserve donc un nombre d'Erlangs minimum valant : $M_{\min} = \max(M_{\min 1}, M_{\min 2}, M_{\min 3})$, et un nombre maximal d'Erlangs valant : $M_{\max} = \min(M_{\max 1}, M_{\max 2}, M_{\max 3})$.

On a donc finalement un nombre d'Erlangs constituant le mélange appartenant à $[M_{\min}, M_{\max}]$, l'Analyste pourra décider dans cet intervalle la valeur qui lui semble la plus appropriée à refléter les observations au regard des différentes études menées.

4.9 Conclusion

À travers ce chapitre, nous avons pu développer une proposition de modélisation du Body de la distribution de la sévérité des sinistres grâce à un mélange d'Erlangs. Ici, la base de données avait été choisie avec volontairement peu de sinistres afin de mieux comprendre tous les enjeux d'une telle modélisation. De plus, par souci de simplification, les observations n'étaient pas représentées par des intervalles, mais uniquement par des valeurs fixes non censurées. La méthode proposée permet de traiter le cas général avec des sinistres représentés par des intervalles en considérant la valeur minimale et maximale de chaque sinistre, en effet l'algorithme EM s'adapte très bien à ce cas-là (voir section 4.3).

On travaillait jusqu'à présent sur des données simulées ou bien choisies de manière aléatoire. Nous allons désormais faire une application réelle sur une base de données concrètes, développée par MOURIER [2020]¹⁶ au sein de QBE où à la fois pour la Tail, mais également pour le Body de la distribution, les sinistres seront représentés par des intervalles.

16. MOURIER [2020] - cf Bibliographie [1]

5 Mise en application sur des données réelles

Nous allons dans cette partie, appliquer la méthode proposée dans les parties précédentes afin de modéliser une distribution de sévérité concrète. L'idée est de comparer les résultats obtenus par la nouvelle méthode et ceux de l'ancienne méthode de modélisation lorsque l'on avait des sinistres à valeur fixe (voir section 2.3).

5.1 Base de données

Afin de réaliser cette mise en pratique, nous prendrons les sinistres développés dans le mémoire de MOURIER [2020]¹⁷ au sein de QBE. Initialement, les sinistres étaient estimés par une unique valeur fixe. Dans ce mémoire, l'auteur a développé une méthode permettant d'estimer chaque sinistre par différentes valeurs, chacune ayant une probabilité de survenance, et cela, en prenant en compte le plus d'information possible contenue dans chacun des sinistres (Année de survenance, cadence de règlement, taille du sinistre, etc.). En annexe B, figure un tableau récapitulatif des informations ayant aidé à développer ces sinistres comme des sinistres à valeurs dans des intervalles. Dorénavant, d'une valeur initialement fixe, chaque sinistre peut en réalité se trouver dans un intervalle de possibilité. Nous allons modéliser la distribution finale de la sévérité de ces nouveaux sinistres.

La base de données utilisée est caractérisée par :

- 1717 sinistres issus d'un portefeuille s'étalant sur 20 ans, de 2000 à 2019, et indexés vers l'année l'actuelle 2020.
- Les sinistres sont issus d'une même branche d'activité pour différentes cédantes.
- Chaque sinistre peut avoir de 1 à 1707 valeurs possibles, c'est-à-dire qu'un sinistre ayant une seule valeur possible est un point fixe et un sinistre ayant 1707 valeurs possibles se trouve dans un intervalle de possibilités très grand.
- Chaque valeur d'un sinistre est pondérée par une probabilité de survenance, et la somme des probabilités de survenance de chaque valeur possible pour un seul sinistre vaut environ 0.8. En effet, les sinistres ont été développés avec un certain indice de confiance, ne conservant que 80% des valeurs les plus probables pour chaque sinistre.

Nous avons également la base originelle, comportant également 1717 sinistres, où les sinistres sont représentés par des points fixes. Cette base nous servira d'élément de comparaison pour comparer les résultats obtenus entre la nouvelle méthode de modélisation avec des sinistres à valeurs dans un intervalle et l'ancienne méthode de modélisation avec des sinistres à valeur fixe.

Pour les deux situations (sinistres à valeurs dans un intervalle et sinistres à valeur fixe), nous allons appliquer la modélisation adaptée en reprenant chaque étape théorique étudiée dans les parties précédentes : séparation de la Tail et du Body, modélisation de la Tail et modélisation du Body. De plus, afin de connaître l'impact final sur la tarification de contrats de réassurance, nous considérerons des contrats "Excess of loss" et regarderons la différence des prix obtenus entre les deux méthodes.

Dans toute cette partie, comme chaque sinistre peut prendre jusqu'à 1707 valeurs possibles, afin de rendre les graphiques plus lisibles, chaque sinistre sera représenté par un intervalle. L'intervalle correspondant débute de la plus petite valeur possible du sinistre et se termine à sa plus grande valeur possible. Ainsi un sinistre ayant les valeurs probables 10, 11, 17, 22, 34, et 56 ; avec chaque valeur une probabilité de survenance, sera modélisé graphiquement par un intervalle allant de 10 à 56. Mais en réalité, la méthode d'estimation prend en compte chacune des valeurs possibles des sinistres avec sa probabilité de survenance correspondante et non uniquement la borne maximale et la borne minimale de l'intervalle (sauf pour la modélisation du Body, comme cela a été expliqué en section 4.3).

17. MOURIER [2020] - cf Bibliographie [1]

La densité et la fonction de répartition des deux situations sont représentées de la manière suivante :

Situation originelle avec des sinistres à valeur fixe :

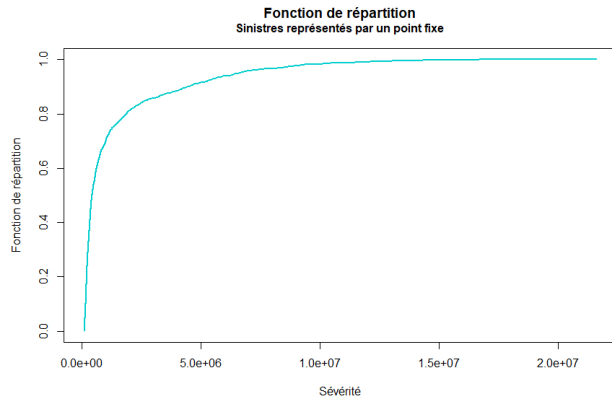


FIGURE 52 – Fonction de répartition

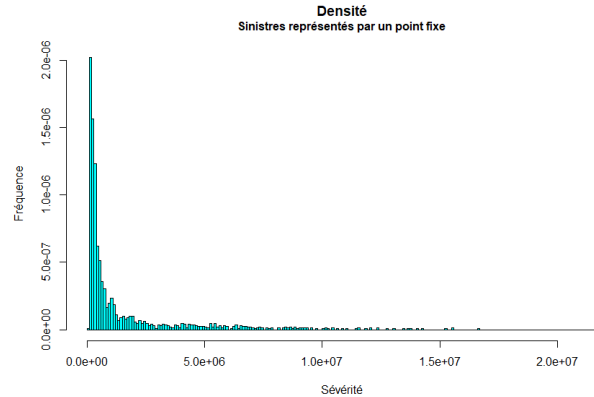


FIGURE 53 – Densité

À cause du nombre important de sinistres, visuellement, la fonction de répartition apparaît comme une fonction lissée, mais il s'agit bien en réalité d'une fonction en escalier.

Nouvelle situation avec des sinistres à valeurs dans un intervalle de possibilités :

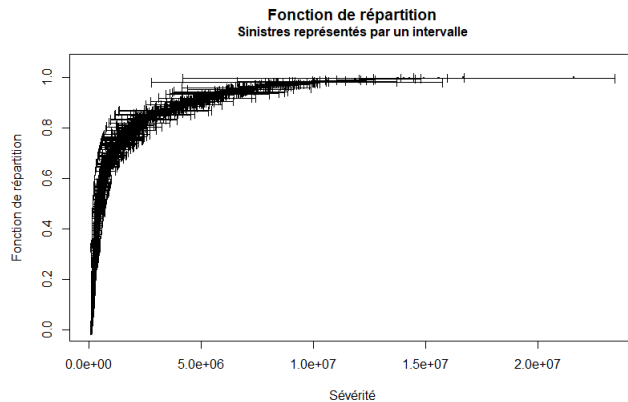


FIGURE 54 – Fonction de répartition

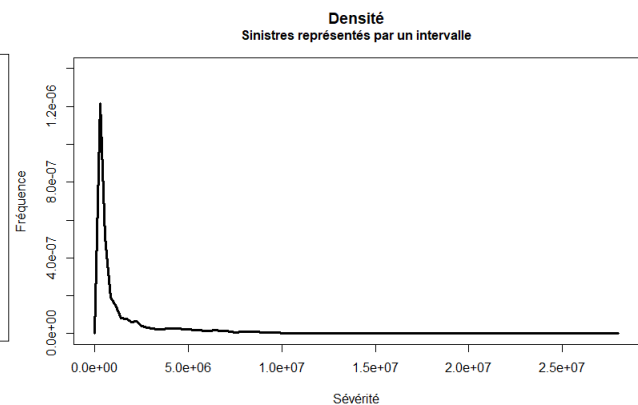


FIGURE 55 – Densité

Les sinistres étant ici représentés par des intervalles, la densité a été créée en prenant en compte la probabilité de survenance de chaque possibilité de valeur de chaque sinistre, c'est pourquoi elle n'est pas affichée ici sous forme d'histogramme. Pour la "fonction de répartition", comme les sinistres sont représentés par des intervalles, il s'agit plutôt en réalité d'un "range" de fonctions de répartition.

En comparant les deux situations, il apparaît que la situation avec des sinistres à valeur fixe représente en quelque sorte la moyenne pour chacun des sinistres sur son intervalle. En superposant ces deux situations, on voit bien une grande similitude entre les formes des distributions :

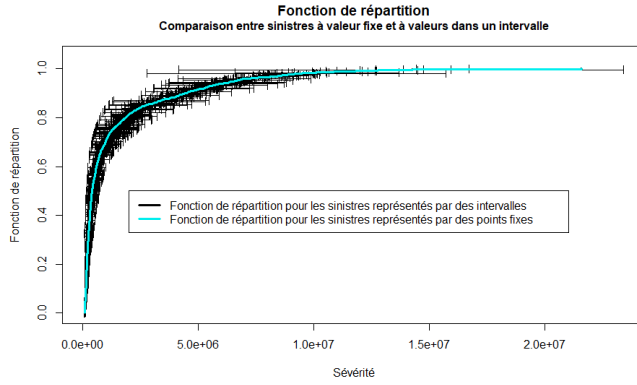


FIGURE 56 – Fonction de répartition

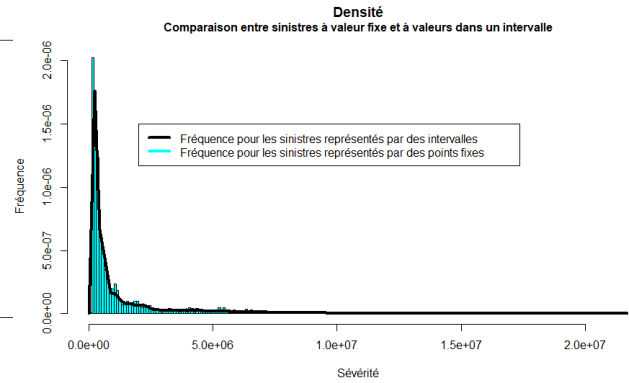


FIGURE 57 – Densité

C'est sur base de ces sinistres que nous allons dans les prochaines parties modéliser la distribution de la sévérité à chaque fois dans les deux situations, afin de montrer l'impact de la nouvelle modélisation proposée.

5.2 Séparation de la Tail et du Body

Afin de déterminer le seuil de séparation entre la Tail et le Body de la distribution, nous reprendrons les outils explicités précédemment qui sont le Mean-Excess plot ainsi que le graphique d'estimation du paramètre de la loi de la Tail. Ces deux outils sont expliqués en section 2.1.4 et 2.3 pour le cas des sinistres à valeur fixe ; et en section 3.2.5 et 3.2.6 pour les sinistres à valeurs dans un intervalle.

5.2.1 Situation originelle avec des sinistres à valeur fixe

Voici le Mean-Excess plot pour cette situation :

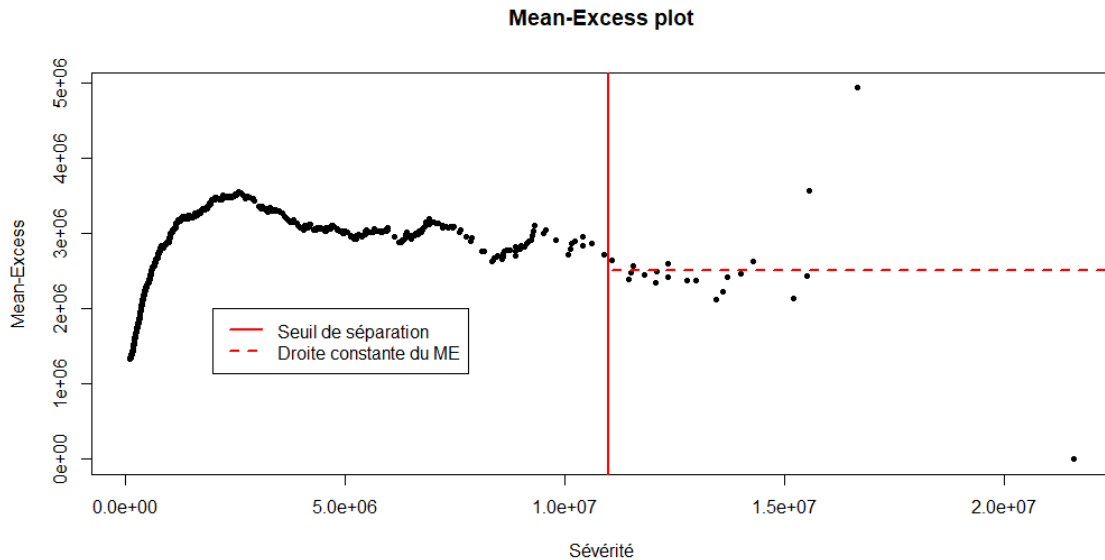


FIGURE 58 – Mean-Excess plot - Sinistres fixes

À partir du seuil $A=11 \times 10^6$, le Mean-Excess plot a une forme relativement constante, on pourrait donc ajuster une loi exponentielle pour modéliser la Tail de la distribution à partir de ce seuil. Vérifions le graphique "Lambda plot" (expliqué section 2.3) afin d'estimer le paramètre de cette loi Exponentielle :

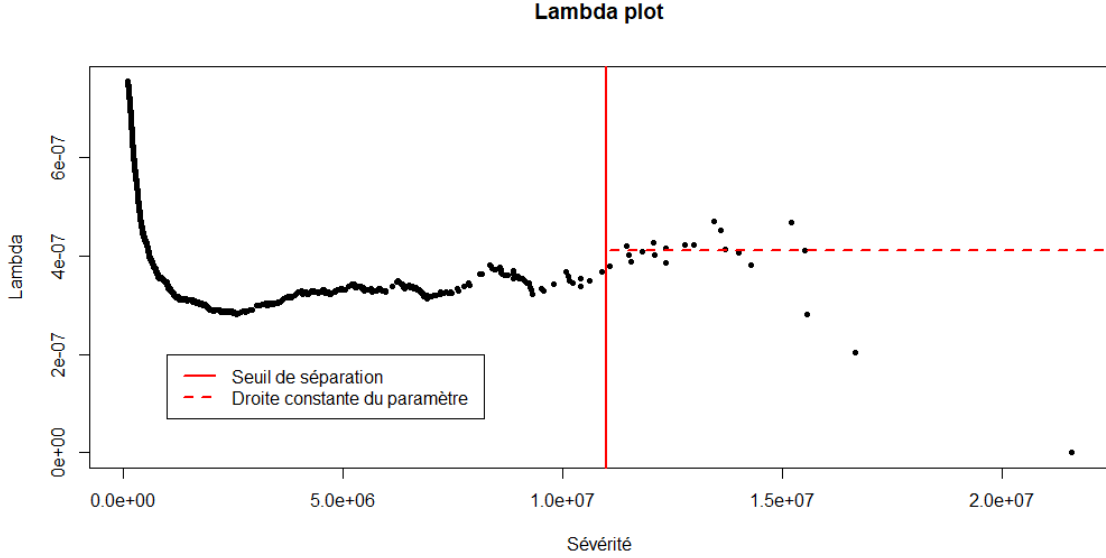


FIGURE 59 – Lambda plot - Sinistres fixes

À partir du seuil de 11×10^6 , le graphique est également approximativement constant, modéliser une loi Exponentielle à partir de ce seuil semble ainsi être un bon choix. On retiendra donc les choix suivants :

- Seuil de séparation entre le Body et la Tail de la distribution : 11×10^6
- Loi de la Tail au-delà du seuil de séparation : loi Exponentielle de paramètre $\lambda = 4.1 \times 10^{-7}$ (ou bien $\theta = \frac{1}{\lambda} \simeq 4.44 \times 10^6$)
- Modélisation du Body : modélisation empirique sans ajustement de loi (voir section 2.2 pour plus d'explications)
- Cette séparation implique que la Tail soit composée des 21 plus grands sinistres et que le Body soit constitué des 1696 autres sinistres.

5.2.2 Nouvelle situation avec des sinistres à valeurs dans un intervalle de possibilités

Pour rappel, nous sommes ici dans le cas de sinistres ayant plusieurs valeurs possibles avec une pondération discrète. Dans ce cas, le Mean-Excess plot est exprimé de la manière suivante (voir section 3.2.6) :

Soit n le nombre de valeurs possibles pour de l'ensemble des sinistres, X_i le montant de cette i -ème valeur possible, et w_i le poids discret attribué à cette valeur. Le Mean-Excess plot revient à tracer les points (E_k, X_{n-k}) où (X_1, \dots, X_n) sont ordonnés par ordre croissant, avec :

$$\forall k \in 1, \dots, n-1, E_k = \frac{\sum_{i=1}^k X_{n-i+1} w_{n-i}}{\sum_{j=1}^k w_{n-j}} - X_{n-k}$$

Voici le Mean-Excess plot pour cette situation :

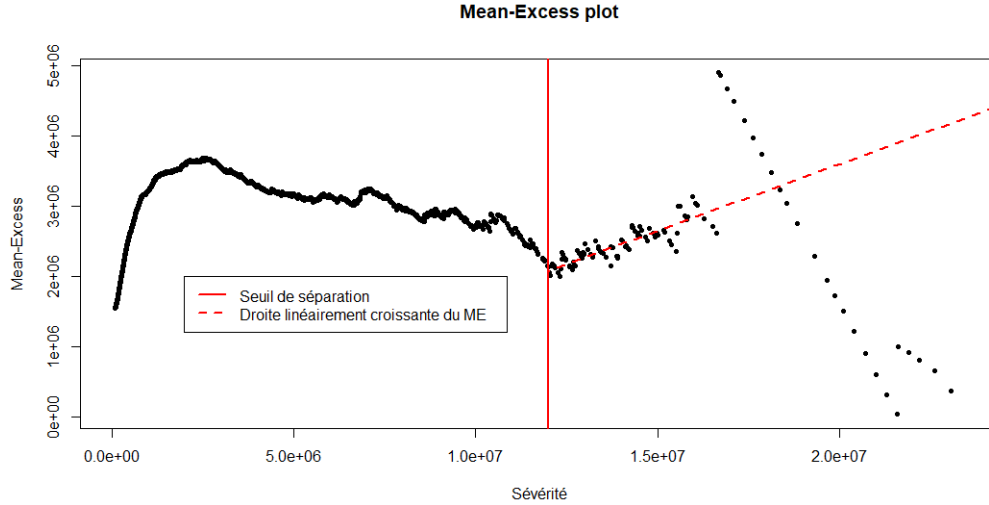


FIGURE 60 – Mean-Excess plot - Sinistres sous forme d'intervalle

La dernière diagonale sur le graphique aux alentours de 20×10^7 représente en réalité uniquement un seul sinistre, l'ensemble des points sur cette diagonale représente l'évaluation du Mean-Excess à chacune des valeurs possibles du sinistre correspondant. À partir du seuil $A=12 \times 10^6$, hormis cette grande diagonale et de la plus petite juste après (représentant seulement 2 sinistres), le Mean-Excess plot a une forme approximativement linéairement constante. On pourrait donc ajuster une loi de Pareto pour modéliser la Tail de la distribution à partir de ce seuil. Vérifions maintenant le graphique "Alpha plot" (expliqué section 3.2.5) afin d'estimer le paramètre de cette loi de Pareto :

Rappelons également que dans ce cas, l'estimateur du paramètre α est donné par (voir section 3.2.5) :

$$\forall k \in 1, \dots, n-1, \hat{\alpha}_k = \frac{\sum_{i=1}^k w_{n-i} - 1}{\sum_{j=1}^k \ln \left(X_{n-j+1} w_{n-j+1} \right) - \ln \left(X_{n-k} w_{n-k} \right)} \quad (***)$$

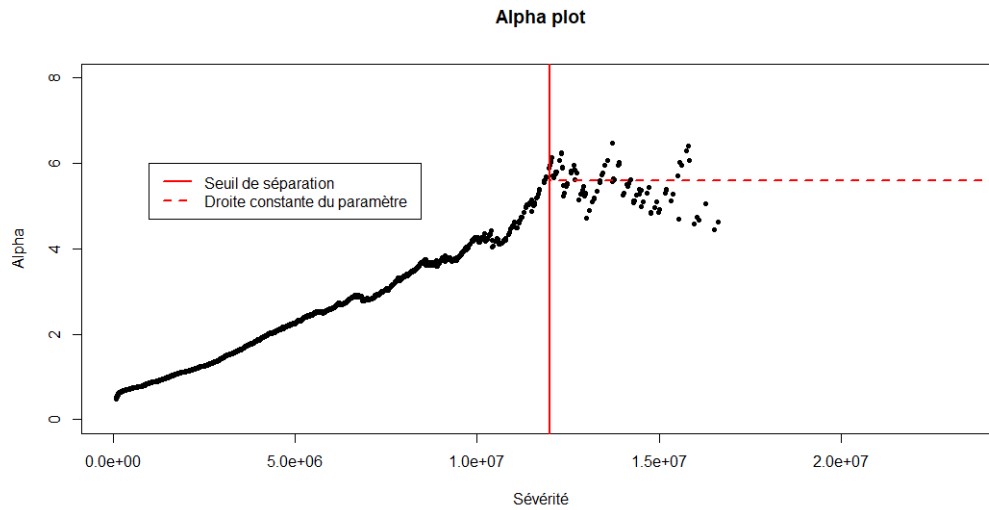


FIGURE 61 – Alpha plot - Sinistres sous forme d'intervalle

À partir du seuil de 12×10^6 , le graphique "Alpha plot" est approximativement constant, modéliser une loi de Pareto à partir de ce seuil semble ainsi être un bon choix. On retiendra les choix suivants :

- Seuil de séparation entre le Body et la Tail de la distribution : 12×10^6
- Loi de la Tail au-delà du seuil de séparation : loi de Pareto de paramètre α environ égal à 6, sa valeur exacte sera explicitée dans la partie suivante.

5.3 Modélisation de la Tail

5.3.1 Situation originelle avec des sinistres à valeur fixe

Section 5.2, il a été établi que dans cette situation, nous modéliserons la Tail de la distribution de sévérité des sinistres par une loi Exponentielle de paramètre $\lambda = 4.1 \times 10^{-7}$ et que cette partie de la distribution contient les 21 plus grands sinistres. Cette loi ajuste ainsi les données empiriques à partir du seuil 11×10^6 de la manière suivante :

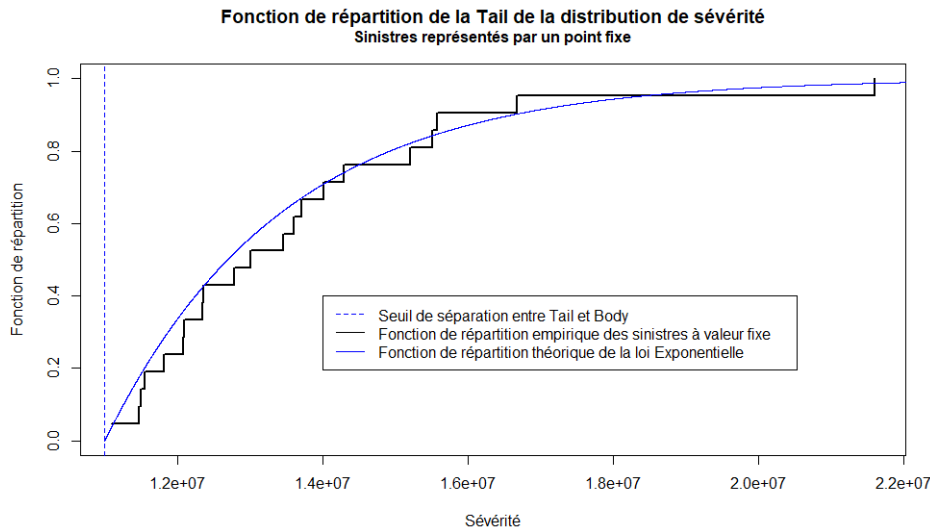


FIGURE 62 – Loi de la Tail - Sinistres à valeur fixe

La distribution Exponentielle ainsi ajustée à partir du seuil de 11×10^6 aux sinistres empiriques représente plutôt bien la distribution empirique.

5.3.2 Nouvelle situation avec des sinistres à valeurs dans un intervalle de possibilités

En reprenant l'estimateur de l'expression (***) ci-dessus, la loi théorique que l'on modélisera à partir du seuil $A=12 \times 10^6$, est une loi de Pareto de paramètre de forme $\alpha \simeq 6.13$. Ce seuil de séparation implique que nous avons ainsi 16.81 sinistres constituant la Tail de la distribution et 1700.19 sinistres représentant le Body de la distribution. En effet, à cause de la représentation sous forme d'intervalle des sinistres, en séparant la distribution globale en deux (Tail et Body), le nombre de sinistres dans chacune des parties n'est pas un nombre entier. Cette loi ajuste les données empiriques de la manière suivante :

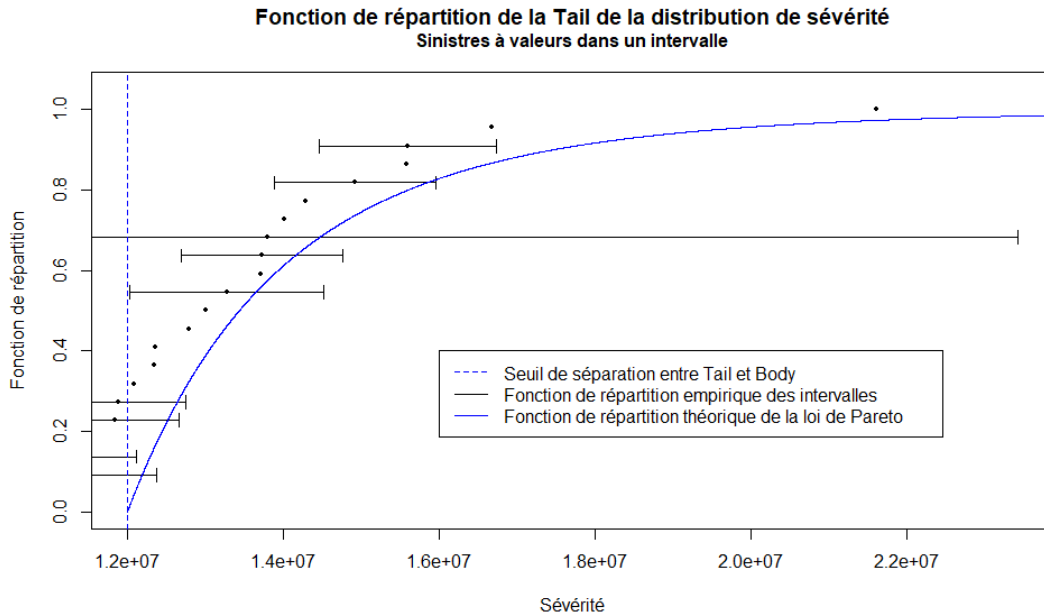


FIGURE 63 – Loi de la Tail de distribution - Sinistres sous forme d'intervalle

Les points (en noir) non compris dans un intervalle représentent les cas où les sinistres auraient dû être représentés par un intervalle avec 80% d'exactitude, mais que l'on a estimé avec le même taux de confiance que l'on pouvait garder une valeur fixe pour ces sinistres. Tous les autres sinistres sont bien représentés par des intervalles, les points compris dans un intervalle représentent uniquement le milieu de cet intervalle. C'est ces points qui permettent de créer "une fonction de répartition empirique" qui est en réalité plutôt un "range" de fonctions de répartition.

En comparant la représentation empirique des sinistres sous forme d'intervalle et la loi théorique ajustée, on remarque que cette distribution colle relativement bien aux sinistres observés au-delà du seuil de 12×10^6 .

5.4 Modélisation du Body

5.4.1 Situation originelle avec des sinistres à valeur fixe

Comme indiqué en section 2.2, pour les sinistres à valeur fixe, le Body de la distribution de la sévérité des sinistres qui est retenu est la distribution empirique. Ici, nous avons beaucoup d'observations (1696), ce qui implique que la distribution empirique est assez robuste en comparaison avec une distribution possédant peu de sinistres (pour une seule cédante par exemple).

Ici également, à cause du nombre important de sinistres, la fonction de répartition apparaît comme une fonction lissée, mais il s'agit en réalité d'une fonction en escalier. Cette distribution empirique du Body est représentée par :

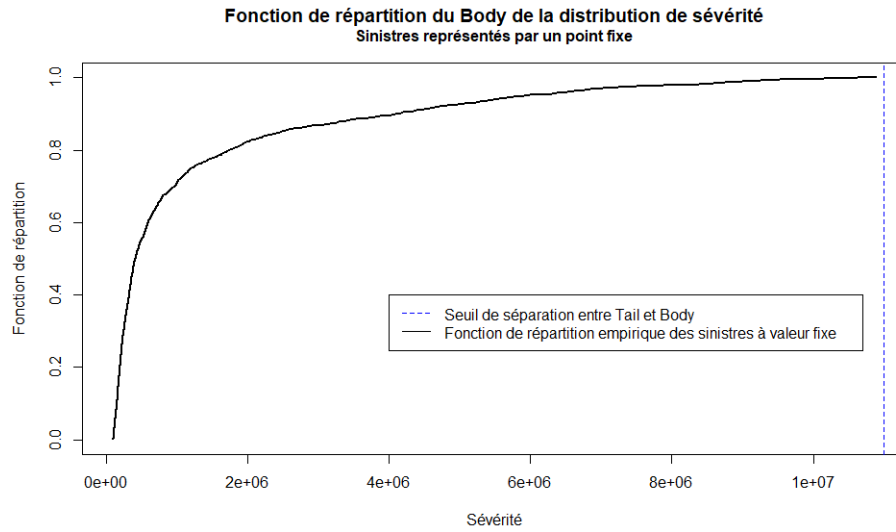


FIGURE 64 – Fonction de répartition du Body - Sinistres à valeur fixe

Et la répartition des sinistres appartenant au Body de la distribution est représentée par :

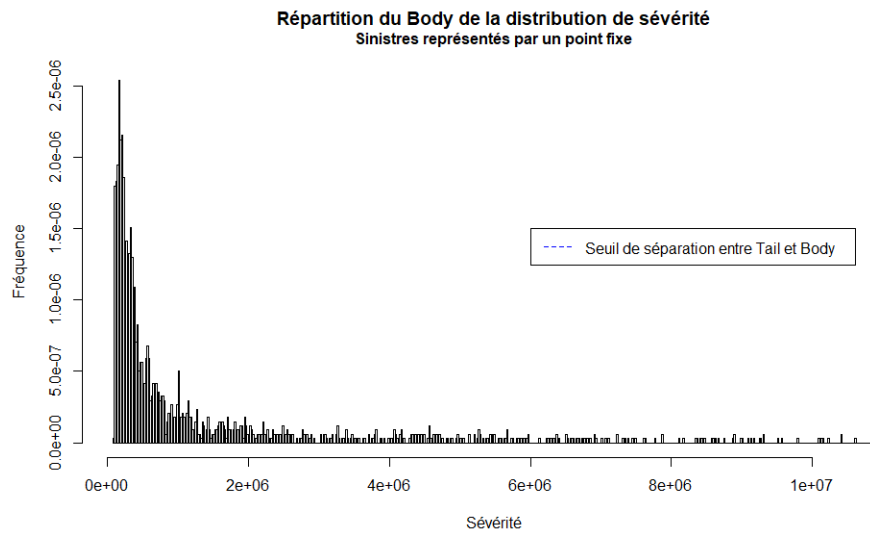


FIGURE 65 – Répartition du Body - Sinistres à valeur fixe

5.4.2 Nouvelle situation avec des sinistres à valeurs dans un intervalle de possibilités

Afin de modéliser le Body de la distribution de la sévérité dans cette situation avec des sinistres à valeurs dans un intervalle, nous allons utiliser l'ajustement des mélanges d'Erlangs. Ne sachant pas le nombre nécessaire et suffisant d'Erlangs à utiliser dans le mélange, nous allons donc utiliser la méthode de discrimination du nombre d'Erlangs dans un mélange (expliquée en section 4.8).

Nombre d'observations

Ce Body est constitué de 1700.19 sinistres. Nous avons vu que pour un mélange de M Erlangs, il faudra estimer $2*M$ paramètres. Compte tenu du nombre important de sinistres, nous nous permettrons de faire correspondre au moins 60 sinistres pour chaque paramètre à estimer. Ainsi, on ne doit pas avoir plus de : $\frac{1696}{60} * \frac{1}{2} \simeq 14.13$ Erlangs constituant le mélange \rightarrow On prendra donc $M_{\max 1} = 14$.

Nombre de bosses dans la distribution du Body

Ces sinistres sont répartis de la manière suivante :

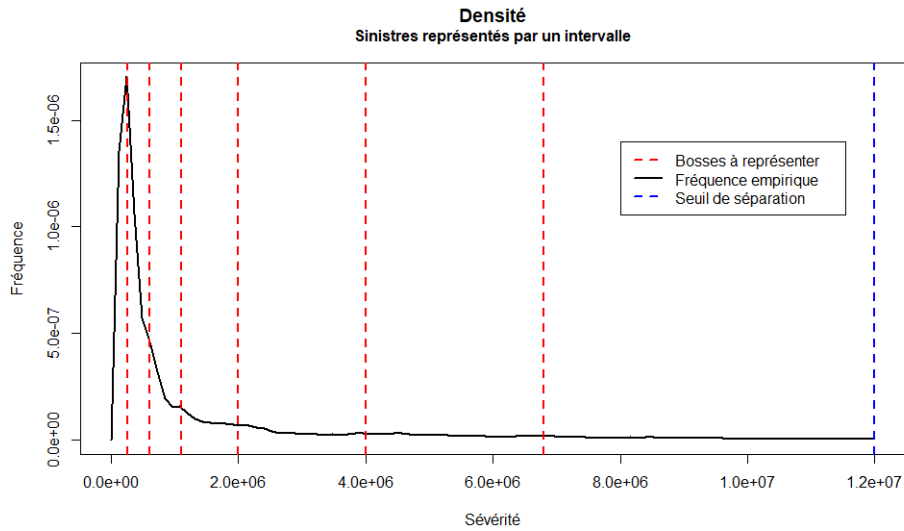


FIGURE 66 – Bosses du Body - Sinistres à valeurs dans un intervalle

Les droites verticales en rouge, représentent les "bosses" à modéliser de la répartition des données. Puisque nous avons 6 bosses, d'après les analyses faites précédemment, il faudra donc compter au moins 6 Erlangs dans le mélange \rightarrow On prendra donc $M_{\min 1} = 6$.

Temps d'exécution

Nous avons donc un nombre d'Erlangs composant le mélange compris entre 6 et 14. Pour ces différentes valeurs, regardons les temps d'exécution des ajustements obtenus grâce à l'algorithme EM :

	M.6	M.7	M.8	M.9	M.10	M.11	M.12	M.13	M.14
Temps d'exécution	2.15 mins	2.73 mins	4.08 mins	4.34 mins	5.68 mins	7.62 mins	9.35 mins	11.9 mins	12.6 mins

FIGURE 67 – Temps d'exécution

Ces temps sont assez faibles, et le temps total d'exécution de l'algorithme pour l'ensemble de ces valeurs est d'environ une heure. L'algorithme EM développé est ainsi assez rapide pour pouvoir ajuster ce type de mélange.

Critères d'information

Regardons maintenant la valeur des différents critères d'information pour les différents mélanges d'Erlangs :

	BIC	AIC	AICc
M=6	16355.14	16289.76	16289.95
M=7	15988.77	15912.49	15912.74
M=8	15692.65	15605.48	15605.80
M=9	15441.32	15354.15	15354.55
M=10	15546.00	15445.97	15446.48
M=11	15820.57	15706.15	15706.79
M=12	15965.44	15837.19	15837.95
M=13	16127.90	15985.37	15986.29
M=14	16143.70	15999.86	16000.93

FIGURE 68 – Critères d'information

Il apparaît que le nombre d'Erlangs minimisant les trois critères (BIC, AIC et AICc) est de 9. Comme la méthode de discrimination le prévoit, afin de ne pas négliger certains modèles pas forcément optimaux par rapport à ces critères, mais tout à fait envisageables, nous ne retiendrons pas uniquement le modèle qui minimise exactement les critères. Plusieurs modèles avec différents nombres d'Erlangs qui ont les plus faibles valeurs pour ces critères seront retenus. On conserve ainsi les valeurs $M = 8, 9$ et 10 .

Stabilité de l'ajustement

Nous avons vu que l'ajustement par les mélanges d'Erlangs n'est pas trop dépendant de l'ajout de sinistres même s'ils sont localisés en fin de distribution. De plus, le choix du seuil de séparation entre la Tail et le Body est beaucoup moins impactant sur les distributions obtenues qu'avec la méthode traditionnelle de tarification. L'ajustement des mélanges d'Erlangs permet de ne pas être trop sensible aux variations que peuvent subir les sinistres d'une année sur l'autre. En effet, en réalité ces variations ne sont pas forcément représentatives de ce qui va se passer dans les années à venir, elles n'ont donc aucune raison d'être reproduites dans l'ajustement du mélange.

De plus, les sinistres sont issus de différentes cédantes et représentent l'étude d'un marché complet pour une certaine branche d'activités. Le grand nombre de sinistres constituant le Body (170.19) permet que l'ajustement soit beaucoup moins sensible que lors de l'analyse réalisée avec peu de sinistres (voir section 4.5). On ne jugera ainsi pas nécessaire de refaire une analyse de la sensibilité pour ces données.

Écart de coût moyen

Pour les différentes valeurs de M constituant le mélange d'Erlangs, qui ont été pour l'instant retenues (8, 9 et 10), nous avons besoin d'un indicateur du coût moyen sur le Body de la distribution. Pour cela, nous reprendrons la méthode utilisant le coût moyen sur l'ensemble du Body de la distribution, expliquée en section 4.6.

Pour les différentes valeurs de M , comparons les courbes obtenues entre le rapport de coût moyen empirique et le rapport de coût moyen des ajustements. L'empirique est pris comme la valeur moyenne de chaque sinistre sur son intervalle :

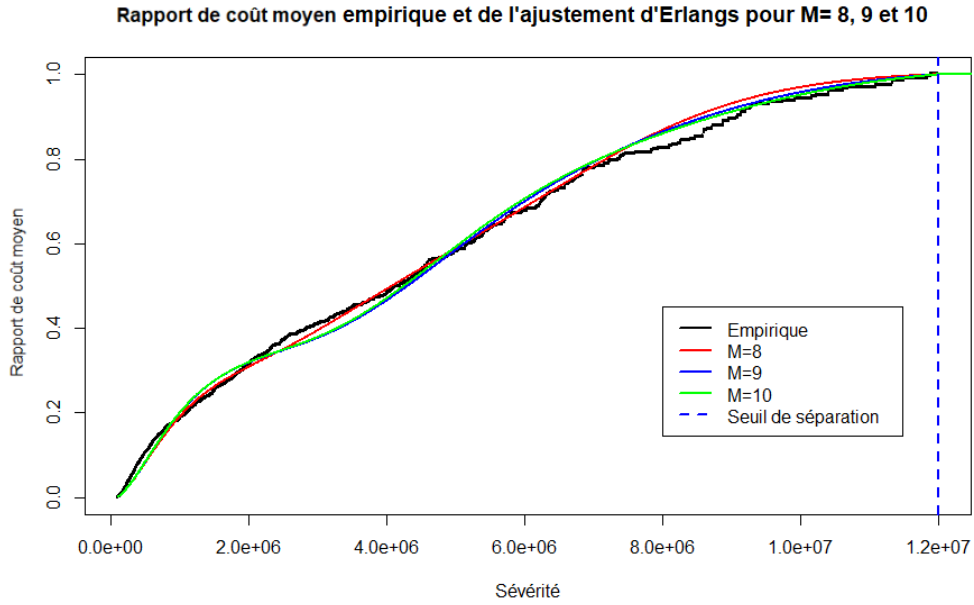


FIGURE 69 – Courbes d'exposition

Afin d'avoir une idée des écarts entre l'empirique et les ajustements, affichons, pour chacun des ajustements cet écart relatif en valeur absolue :

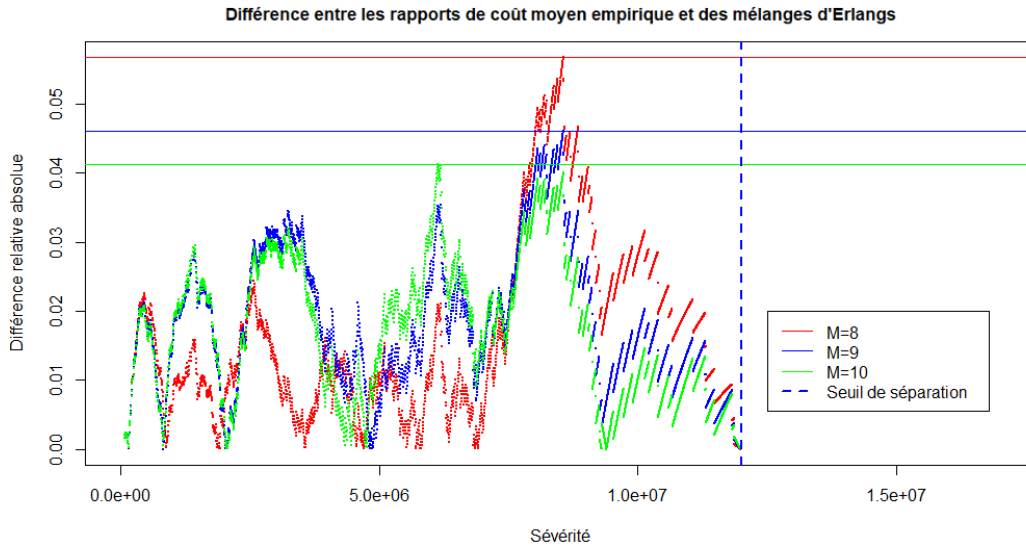


FIGURE 70 – Écart entre les courbes d'exposition

Les barres horizontales représentent pour les différents mélanges d'Erlangs, l'écart maximal atteint entre la courbe de l'ajustement du mélange et celle de l'empirique. Comme attendu, plus le nombre d'Erlangs est grand, plus l'écart maximal atteint est faible. Cependant, la différence de cet écart pour les différentes valeurs de M est très petite. Même si l'écart maximal atteint pour $M=8$ est légèrement plus grand que pour $M=9$ ou 10 , sur le reste de la distribution l'écart pour $M=8$ reste plus petit que pour $M=9$ ou 10 . De plus, la différence entre les trois mélanges d'Erlangs étant très petite, afin de ne pas complexifier le modèle inutilement, nous déciderons donc de retenir $M=8$ pour le mélange d'Erlangs.

Choix final

Grâce à l'ensemble des analyses menées pour l'ajustement du mélange d'Erlangs sur les sinistres à valeurs dans un intervalle, nous avons décidé de retenir 8 Erlangs pour constituer le mélange. Pour ce mélange, les paramètres de la distribution obtenus grâce à l'algorithme EM sont :

- Les paramètres de forme : $(r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8) = (1, 8, 9, 16, 20, 24, 39, 387)$
- Les paramètres de pondération : $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8) = (8.37 \times 10^{-1}, 5.02 \times 10^{-76}, 1.21 \times 10^{-1}, 2.93 \times 10^{-2}, 4.56 \times 10^{-20}, 6.35 \times 10^{-35}, 2.01 \times 10^{-168}, 1.28 \times 10^{-2})$, à trois chiffres significatifs près
- Le paramètre d'échelle commun à chaque Erlang : $\theta = 519342.4$

Et la distribution obtenue peut être représentée par :

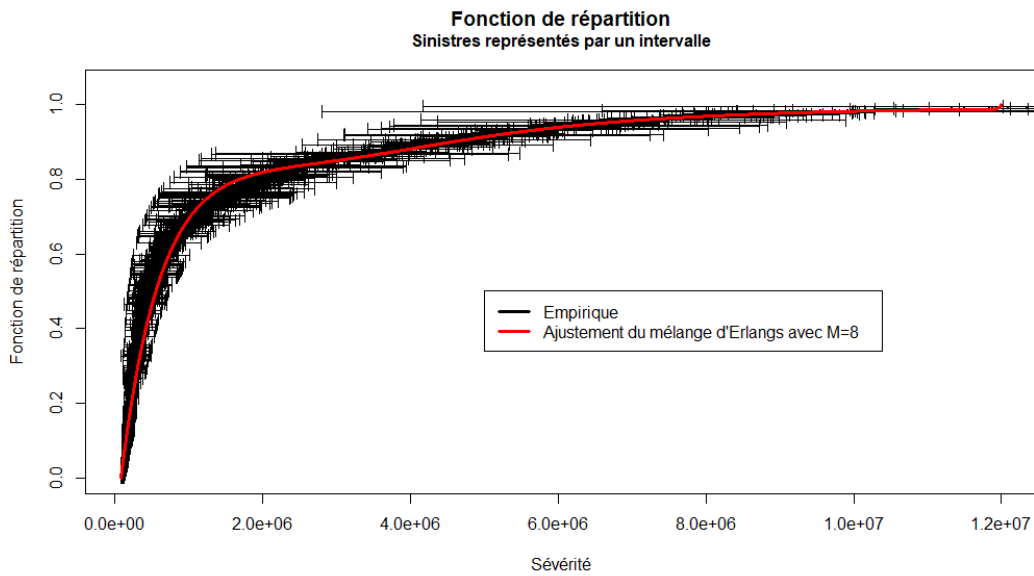


FIGURE 71 – Ajustement du mélange d'Erlangs sur le Body

5.5 Impact tarifaire

5.5.1 Modélisation globale retenue

Résumons les différentes distributions choisies pour modéliser la sévérité des sinistres dans les deux situations possibles :

- Pour les sinistres à valeur fixe :
 - Le Body de la distribution est le Body empirique constitué des sinistres observés.
 - Le seuil de séparation entre le Body et la Tail est de 11×10^6 .
 - La loi de la Tail de la distribution est une loi Exponentielle de paramètre $\lambda = 4.1 \times 10^{-7}$.
- Pour les sinistres à valeurs dans un intervalle :
 - Le Body de la distribution est un mélange de 8 Erlangs de paramètres de forme $(1, 8, 9, 16, 20, 24, 39, 387)$, de paramètres de pondération $(8.37 \times 10^{-1}, 5.02 \times 10^{-76}, 1.21 \times 10^{-1}, 2.93 \times 10^{-2}, 4.56 \times 10^{-20}, 6.35 \times 10^{-35}, 2.01 \times 10^{-168}, 1.28 \times 10^{-2})$, et de paramètre d'échelle commun à chaque Erlang : $\theta = 519342.4$.

- Le seuil de séparation entre le Body et la Tail est de 12×10^6 .
- La loi de la Tail de la distribution est une loi de Pareto de paramètre $\alpha = 6.13$.

Nous avons ainsi représenté la distribution de la sévérité dans les deux situations, nous allons désormais quantifier l'impact sur le prix final de réassurance de ces deux méthodes. Pour cela, nous considérerons les contrats classiquement utilisés en réassurance qui sont les "Excess of Loss" (noté XS).

5.5.2 Prix Excess of Loss (XS)

La tarification de tels contrats a été expliquée en section 1.2.2. Il faut savoir que n'importe quelle tranche (C xs P) de ce type de contrat peut être obtenue par la différence de deux tranches illimitée. En effet en considérant les tranches illimitées suivantes (+Infini xs P) et (+Infini xs L=C+P) nous avons la relation suivante : (C xs P) \longleftrightarrow (+Infini xs P) - (+Infini xs L=C+P).

Nous allons donc modéliser le prix pour des tranches illimitées ayant une priorité P, et cela, pour n'importe quelle priorité P. Le coût moyen des sinistres au-delà du seuil P est borné, mais plus la priorité P est élevée, plus la fréquence du nombre de sinistres dépassant la priorité est faible et tend vers 0. C'est pourquoi lorsque la priorité P tend vers l'infini, le prix du contrat (+Infini xs P) va décroître.

En reprenant la modélisation de la sévérité des deux situations envisagées avec des sinistres à valeur fixe et des sinistres à valeurs dans un intervalle, les prix de ces contrats (+Infini xs P) sont calculés de la manière suivante :

$$\text{Prix XS} = \mathbb{E}(X - P | X > P) \times \mathbb{E}(\text{fréquence au-delà de } P)$$

avec X la variable aléatoire représentant la sévérité des sinistres.

Ce qui donne dans les deux situations :

Sinistres à valeur fixe :

Soit (x_1, \dots, x_n) les n sinistres, S1 le seuil de séparation entre le Body et la Tail ($=11 \times 10^6$), et $\lambda = 4.1 \times 10^{-7}$ le paramètre de la loi Exponentielle, on a alors :

Si $P \in [0, S1]$:

$$\begin{aligned} \mathbb{E}(X - P | X > P) &= \left[\left(\frac{\sum_{i=1}^n x_i \mathbb{1}_{(P < x_i < S1)}}{\sum_{i=1}^n \mathbb{1}_{(P < x_i < S1)}} \right) \sum_{i=1}^n \mathbb{1}_{(P < x_i < S1)} + \frac{\int_{S1}^{\infty} x \lambda e^{-\lambda x} dx}{e^{-\lambda S1}} \sum_{i=1}^n \mathbb{1}_{(x_i > S1)} - P \sum_{i=1}^n \mathbb{1}_{(x_i > P)} \right] \left(\sum_{i=1}^n \mathbb{1}_{(x_i > P)} \right)^{-1} \\ &= \left[\sum_{i=1}^n x_i \mathbb{1}_{(P < x_i < S1)} + \frac{\int_{S1}^{\infty} x \lambda e^{-\lambda x} dx}{e^{-\lambda S1}} \sum_{i=1}^n \mathbb{1}_{(x_i > S1)} \right] \left(\sum_{i=1}^n \mathbb{1}_{(x_i > P)} \right)^{-1} - P \end{aligned}$$

Et, $\mathbb{E}(\text{fréquence au-delà de } P) = \sum_{i=1}^n \mathbb{1}_{(x_i > P)}$

Si $P > S1$:

$$\mathbb{E}(X - P | X > P) = \frac{\int_P^{\infty} x \lambda e^{-\lambda x} dx}{e^{-\lambda P}} - P, \text{ et, } \mathbb{E}(\text{fréquence au-delà de } P) = \sum_{i=1}^n \mathbb{1}_{(x_i > P)}$$

Sinistres à valeurs dans un intervalle :

Soit $S2$ le seuil de séparation entre le Body et la Tail ($=12 \times 10^6$), $\alpha = 6.13$ le paramètre de forme de la loi de Pareto modélisée à partir de $S2$, (x_1, \dots, x_N) les N valeurs possibles de l'ensemble des sinistres, et (w_1, \dots, w_N) les poids correspondants à chacune de ces valeurs possibles, on a alors :

Si $P \in [0, S2]$:

En reprenant l'expression notée en section 4.6 de l'espérance pour un mélange d'Erlangs; avec F_{r_i} la fonction de répartition de la loi d'Erlang de paramètre de forme $r_i \in \mathbb{N}^*$, de paramètre de pondération dans le mélange α_i , et de paramètre d'échelle commun $\theta \in \mathbb{R}^+$, on a alors :

$$\begin{aligned} \mathbb{E}(X - P | X > P) &= \left[\left(\frac{\theta \sum_{i=1}^M \alpha_i r_i (F_{r_i+1}(S2) - F_{r_i+1}(P))}{\sum_{i=1}^M \alpha_i (F_{r_i}(S2) - F_{r_i}(P))} \right) \sum_{i=1}^N w_i \mathbb{1}_{(P < x_i < S2)} + \frac{\int_{S2}^{\infty} x \alpha \frac{S2^\alpha}{x^{\alpha+1}} dx}{\left(\frac{S2}{S2}\right)^\alpha} \sum_{i=1}^N w_i \mathbb{1}_{(x_i > S2)} \right. \\ &\quad \left. - P \sum_{i=1}^N w_i \mathbb{1}_{(x_i > P)} \right] \left(\sum_{i=1}^N w_i \mathbb{1}_{(x_i > P)} \right)^{-1} \\ &= \left[\left(\frac{\theta \sum_{i=1}^M \alpha_i r_i (F_{r_i+1}(S2) - F_{r_i+1}(P))}{\sum_{i=1}^M \alpha_i (F_{r_i}(S2) - F_{r_i}(P))} \right) \sum_{i=1}^N w_i \mathbb{1}_{(P < x_i < S2)} + \frac{\alpha S2}{\alpha - 1} \sum_{i=1}^N w_i \mathbb{1}_{(x_i > S2)} \right] \left(\sum_{i=1}^N w_i \mathbb{1}_{(x_i > P)} \right)^{-1} - P \end{aligned}$$

Et, $\mathbb{E}(\text{fréquence au-delà de } P) = \sum_{i=1}^N w_i \mathbb{1}_{(x_i > P)}$

Si $P > S2$:

$$\mathbb{E}(X - P | X > P) = \frac{\int_{P}^{\infty} x \alpha \frac{S2^\alpha}{x^{\alpha+1}} dx}{\left(\frac{S2}{P}\right)^\alpha} \sum_{i=1}^N w_i \mathbb{1}_{(x_i > S2)} - P = \frac{P}{\alpha - 1}, \text{ et, } \mathbb{E}(\text{fréquence au-delà de } P) = \sum_{i=1}^N w_i \mathbb{1}_{(x_i > P)}$$

Ces expressions sont à utiliser avec le mélange d'Erlangs que nous avons ajusté dans l'étude du Body de la distribution, c'est-à-dire avec le mélange suivant :

- $M=8$
- Les paramètres de forme : $(r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8) = (1, 8, 9, 16, 20, 24, 39, 387)$
- Les paramètres de pondération : $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8) = (8.37 \times 10^{-1}, 5.02 \times 10^{-76}, 1.21 \times 10^{-1}, 2.93 \times 10^{-2}, 4.56 \times 10^{-20}, 6.35 \times 10^{-35}, 2.01 \times 10^{-168}, 1.28 \times 10^{-2})$, à trois chiffres significatifs près
- Le paramètre d'échelle commun à chaque Erlang : $\theta = 519342.4$

Comparaison des deux situations :

Traçons le prix de ces XS de tranche illimitée, comme fonction de la priorité P , pour les deux situations envisagées :

Prix pour un XS (+Infini xs priorité)

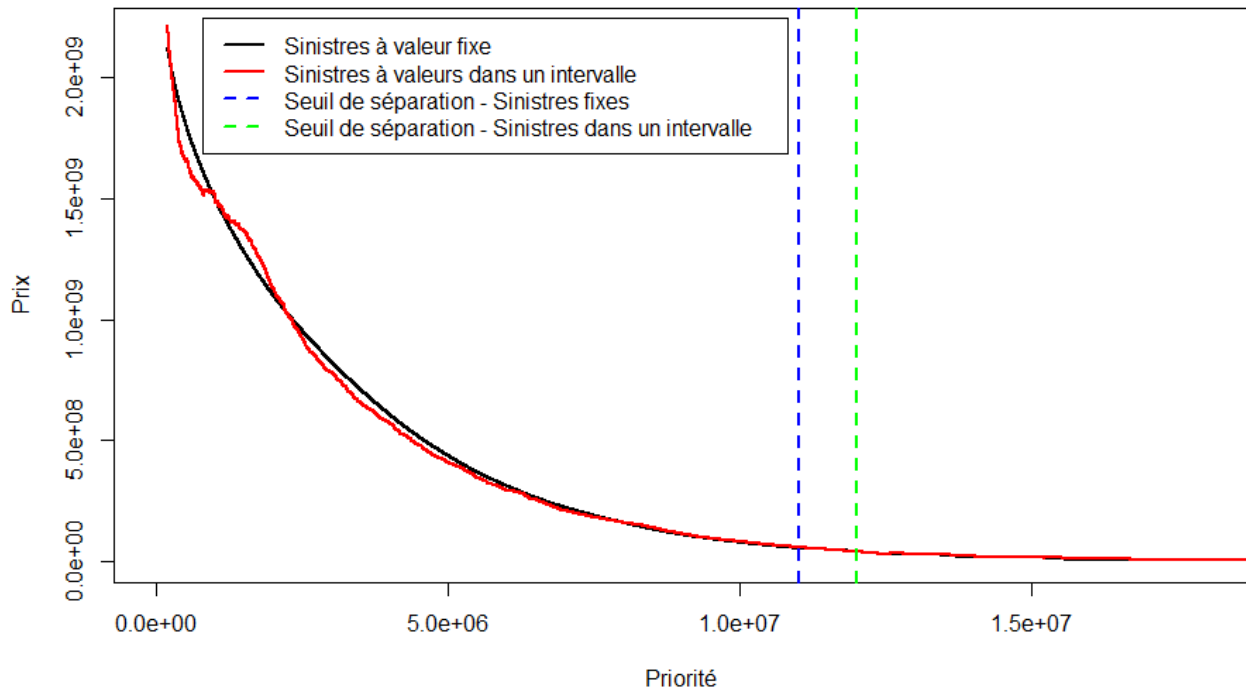


FIGURE 72 – Prix d'un XS (+Infini xs priorité)

Ces courbes de prix présentent les caractéristiques suivantes :

- Les prix sont très similaires d'une méthode à l'autre.
- Nous avons une décroissance exponentielle des prix lorsque la priorité augmente.
- Comme prévu, lorsque la priorité tend vers +Infini, alors les prix tendent vers 0.
- L'écart entre les prix des deux méthodes lorsque la priorité dépasse les seuils de séparation entre Tail et Body devient presque nul. Ce qui est cohérent puisque les prix tendent vers 0. La différence principale vient donc du Body.
- Nous n'avons pas une méthode qui sur-estime ou sur-estime les prix de manière systématique par rapport à l'autre méthode, puisque les courbes de prix se croisent à plusieurs reprises. Ainsi, selon les tranches considérées, ce n'est pas toujours la même méthode qui sera plus avantageuse du côté du réassureur (ou bien du côté assureur).

Regardons désormais les écarts relatifs de prix entre ces deux courbes (si et seulement si, entre ces deux méthodes). Il s'agit ici de l'écart relatif de la méthode avec les sinistres à valeurs dans un intervalle par rapport à la méthode avec les sinistres à valeur fixe :

- Un écart positif signifie que le prix de la méthode avec les sinistres à valeurs dans un intervalle est plus cher.
- Un écart négatif signifie que le prix de la méthode avec les sinistres à valeur fixe est plus cher.

Voici ces écarts relatifs entre les deux courbes de prix :

Écart relatif de prix pour un XS (+Infini xs priorité)
Écart relatif entre sinistres à valeur fixe et sinistres à valeurs dans un intervalle

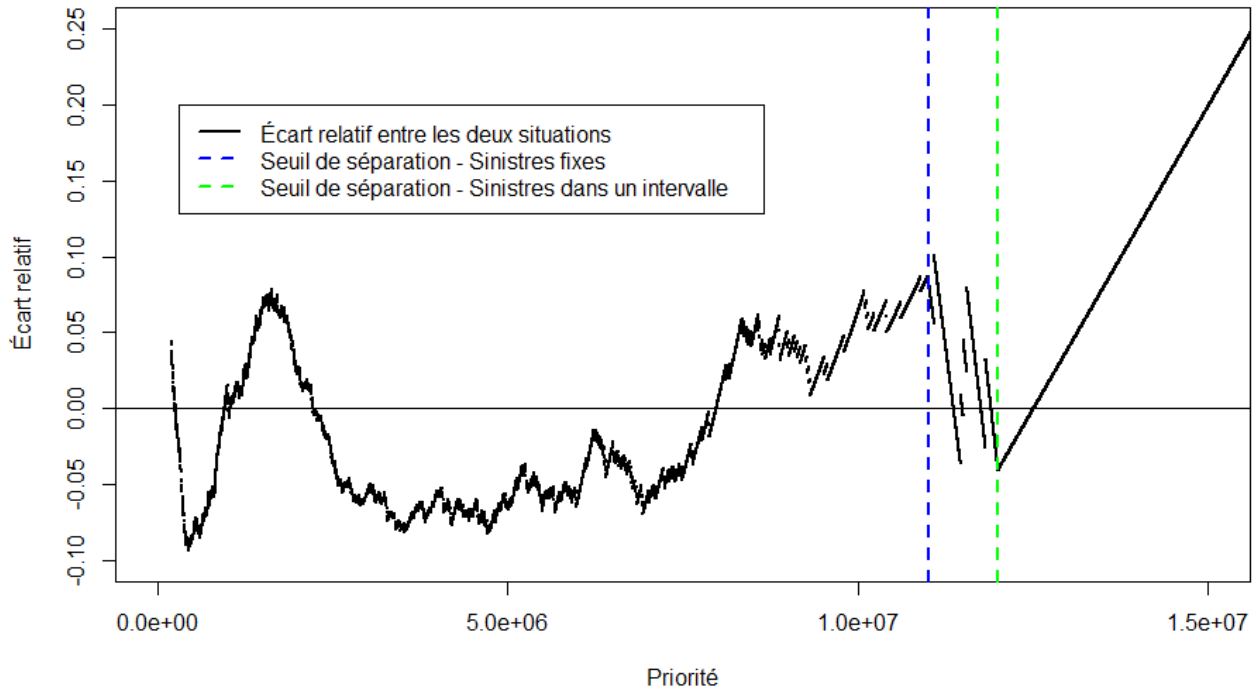


FIGURE 73 – Écart relatif des prix d'un XS (+Infini xs priorité)

Notons qu'après les seuils de séparation entre Tail et Body, les écarts relatifs sont très élevés, cependant ces écarts ne sont pas très significatifs. En effet, pour une priorité dans cette région de la distribution, les prix sont très proches de 0, c'est pourquoi en relatif les écarts sont très grands, mais il s'agit en réalité de très petits prix. Pour cette partie-là, il est cohérent que la méthode avec les sinistres à valeur fixe soit moins chère que la nouvelle méthode, puisque la Tail de la distribution de la sévérité des sinistres avec cette méthode a été modélisée par une loi Exponentielle, qui est une loi à queue plus légère que la loi de Pareto utilisée pour modéliser la Tail avec la méthode où les sinistres sont à valeurs dans des intervalles. En dehors de cette région "non-significative", l'écart relatif maximal éteint est d'environ 10%, ce qui est très important. La nouvelle méthode proposée avec des sinistres à valeurs dans un intervalle impacte ainsi fortement les prix de réassurance des "Excess of loss". Il n'y a cependant pas de méthode qui a tendance à hausser ou diminuer les prix, puisque la baisse ou la hausse de prix peut très bien être en faveur de chacune des méthodes.

6 Conclusion

À travers ce mémoire, nous avons été initiés au fonctionnement général de la Réassurance et nous avons pu retracer les différentes méthodes de tarification utilisées dans ce secteur de l'assurance.

Afin de modéliser la sévérité des sinistres, nous avons vu que pour le Body de la distribution il est d'usage de conserver la distribution empirique, et nous avons vu les différents estimateurs théoriques découlant de la Théorie des Valeurs extrêmes, pouvant être utilisés afin de modéliser la Tail de la distribution. Sauf que la théorie sous-jacente à ces différents estimateurs est une théorie asymptotique, et par définition, plus l'on s'approche de la Tail de la distribution moins l'on a de données. La réassurance est un secteur de l'assurance où l'on manque cruellement de données, en pratique nous utilisons deux outils qui sont le Mean-Excess plot ainsi que le graphique d'estimation du paramètre de la loi modélisant la Tail de la distribution.

Nous avons vu que cette modélisation faite en pratique présente de nombreux inconvénients, c'est pour pallier ces défauts que ce mémoire s'est penché sur des solutions alternatives afin de modéliser la sévérité des sinistres.

Tout d'abord, nous avons étudié de nouvelles lois pour modéliser la Tail de la distribution. En effet, il est d'usage de toujours considérer que la loi de la Tail est une loi à support illimité, alors qu'en réalité ce n'est pas forcément le cas. Ainsi, nous avons envisagé des lois Exponentielles, de Pareto et de Weibull dans le cas où ces lois pourraient être censurées à droite ou bien tronquées à droite. Cependant, l'ajustement de ces lois grâce aux outils actuels de modélisation (Mean-Excess plot et graphique d'estimation du paramètre) n'est en pratique pas facile à appliquer et nécessiterait l'utilisation certainement de plusieurs autres outils. C'est pourquoi ces lois ont donc été écartées pour modéliser la Tail de la distribution de la sévérité des sinistres.

Dans un second temps, nous avons envisagé que les sinistres puissent être représentés par des intervalles et non par des points fixes. En prenant cela pour acquis, nous avons cherché à développer une modélisation de la sévérité pour ce type de forme de sinistres. Nous avons ainsi considéré deux situations possibles pour la forme des sinistres : chaque valeur d'un sinistre sera remplacée par différentes valeurs possibles chacune ayant une probabilité de survenir, ou alors chaque valeur d'un sinistre sera remplacée par une variable aléatoire continue suivant une distribution donnée.

Sur base de ces formes de sinistres, nous avons regardé comment trouver par maximum de vraisemblance les paramètres issus d'une loi Exponentielle ou d'une loi de Pareto, et nous avons pu trouver des estimateurs de ces paramètres les moins biaisés asymptotiquement que possible. En adaptant les outils actuels, nous avons développé un équivalent du Mean-Excess plot ainsi que du graphique d'obtention du paramètre d'une loi. Grâce à ces outils développés, nous avons donc pu trouver le seuil de séparation entre la Tail et le Body, ainsi que la loi à modéliser pour la Tail de la distribution à partir de ce seuil. De plus, pour le Body de la distribution, nous avons considéré que comme le nombre de données était plus conséquent que pour la Tail, nous pouvons nous permettre de ne conserver que la borne minimale ainsi que la borne maximale de chaque sinistre afin d'ajuster un mélange d'Erlangs sur le Body. L'ajustement du mélange d'Erlangs est réalisé à l'aide d'un algorithme de maximisation de l'espérance, et grâce à l'étude de l'ajustement de ce type de mélange, il est apparu que les mélanges d'Erlangs représentent des distributions très stables et assez flexibles pour bien s'adapter à un jeu de données.

Ensuite, en prenant les sinistres développés sous forme d'intervalle dans le mémoire de MOURIER.P, nous avons pu mettre en pratique la nouvelle méthode de modélisation de la sévérité des sinistres tout en la comparant à l'ancienne méthode qui se sert des sinistres à valeur fixe.

Nous avons vu que la nouvelle méthode s'adaptait très bien aux sinistres sous forme d'intervalle, qu'elle était très simple à utiliser, et que malgré un temps d'exécution plus long que l'ancienne méthode, elle était tout à fait envisageable pour réaliser une tarification complète sur base des sinistres à valeurs dans un intervalle.

Nous avons donc considéré des contrats de Réassurance "Excess of loss" et nous avons vu que la nouvelle méthode proposée pouvait impacter fortement les prix de ce type de contrats, même si en moyenne ni la nouvelle méthode, ni l'ancienne ne sur-évaluait ou ne sous-évaluait la sévérité des sinistres par rapport à l'autre méthode. En prenant pour acquis que les sinistres développés sous forme d'intervalles représentent une meilleure estimation de la véritable valeur des sinistres, la méthode de modélisation proposée permettrait de capter cette nouvelle estimation. De plus, la nouvelle méthode, ne capte ainsi pas uniquement la valeur moyenne du sinistre, mais sa distribution sur tout son intervalle de distribution et permet une modélisation adéquate à cette forme de sinistres.

7 Bibliographie

- [1] MOURIER P. [2020] *"Ultimates as intervals in reinsurance"*.
- [2] WALHIN J.F. [2012] *"La réassurance, Editions Larcier"*.
- [3] DELACROIX A. [2019] *"Réassurance Non Vie"*.
- [4] LOLJEEH R. [2012] *"Estimation de courbes d'exposition dans le cas de données incomplètes"*.
- [5] GILLIER V. [2018] *"Calibration avancée du modèle de tarification à l'exposition dans le cadre de la réassurance incendie par risque"*.
- [6] DESMEDT S., SNOUSSI M., CHENUT X., WALHIN J. F. [2012] *"Experience and Exposure Rating for Property Per Risk Excess of Loss Reinsurance Revisited"*.
- [7] ROBERT C. Y. [2019] *"Théorie des Valeurs Extrêmes"*.
- [8] LADOUCETTE S. [2011] *"Theory and estimation"*.
- [9] CHOI K. P. [1994] *"On the medians of gamma distributions and an equation of Ramanujan"*.
- [10] TIJMS H. C. [1994] *"Stochastic models : an algorithmic approach - Wiley"*.
- [11] ANTONIO K., BADESCU A., GONG L., LIN S., VERBELEN R. [2014] *"Fitting mixtures of Erlangs to censored and truncated data using the EM algorithm"*.
- [12] LEE S., LIN S. [2010] *"Modeling and Evaluating Insurance Losses Via Mixtures of Erlang Distributions"*. Article in North American Actuarial Journal.

8 Annexe

A.

Index Valeur

1	0	31	3817	61	9481	91	52253	121	194300	151	300087	181	1900000
2	0	32	3902	62	9598	92	52656	122	196900	152	300785	182	2385755
3	0	33	4000	63	9651	93	54375	123	205215	153	304179	183	3109530
4	0	34	4043	64	9758	94	66963	124	208145	154	327000		
5	0	35	4152	65	10237	95	73543	125	208872	155	329961		
6	0	36	4230	66	10531	96	73787	126	211374	156	354669		
7	0	37	4241	67	11000	97	78233	127	222922	157	370998		
8	18	38	4256	68	11651	98	80590	128	229218	158	377875		
9	108	39	4448	69	12643	99	91512	129	235210	159	384318		
10	216	40	4663	70	16733	100	103480	130	238626	160	384526		
11	249	41	5022	71	17000	101	103623	131	241268	161	385862		
12	368	42	5057	72	22833	102	108087	132	241844	162	409000		
13	519	43	5162	73	24000	103	110924	133	245338	163	414428		
14	600	44	5164	74	25736	104	111187	134	248816	164	447464		
15	793	45	5298	75	26652	105	112026	135	252046	165	453692		
16	817	46	5624	76	28865	106	134440	136	256742	166	492378		
17	1092	47	5650	77	30901	107	135292	137	258863	167	493435		
18	1404	48	5794	78	32171	108	139167	138	261795	168	508143		
19	1434	49	6017	79	33326	109	140675	139	262282	169	589108		
20	1479	50	6232	80	33621	110	144839	140	263298	170	605442		
21	1623	51	6564	81	36575	111	149199	141	265563	171	647909		
22	1642	52	6744	82	39968	112	150080	142	266684	172	859876		
23	1692	53	6885	83	41000	113	151378	143	267898	173	921000		
24	1730	54	7281	84	41888	114	161958	144	268766	174	972338		
25	1738	55	7342	85	42752	115	162021	145	269954	175	1027691		
26	1749	56	7568	86	43216	116	163614	146	277933	176	1038370		
27	2000	57	7865	87	45078	117	169022	147	280632	177	1047000		
28	2197	58	8687	88	46000	118	171584	148	290128	178	1347519		
29	2609	59	8858	89	47957	119	173676	149	293561	179	1377888		
30	3583	60	9427	90	50036	120	176182	150	294685	180	1742002		



Sinistres conservés pour représenter le Body

Sinistres non conservés pour représenter le Body

90 Nombre de sinistres dans le Body

93 Nombre de sinistres hors du Body

FIGURE 74 – Base de données - Section 4

B. Section 5.1 : Définition des nouvelles variables

Note du mémoire de MOURIER [2020] : "Afin de créer des catégories de sinistres, de nouvelles variables sont créées. L'objectif de ces variables est de réussir à catégoriser les comportements des sinistres. Ces variables sont nommées « Métriques », car elles ont pour objectif de mesurer certains aspects du comportement du sinistre. Pour certaines variables, la notion de « montant significatif » est utilisée, ce montant vaut 100.000€. Cela veut dire que le sinistre sera pris en compte uniquement s'il a dépassé la valeur de 100.000€ au moins une fois. Ainsi, nous évitons les problèmes posés par les sinistres déclarés tardivement et pour lesquels le coût estimé du sinistre pour les premières années de développement est très faible. Voici un détail de chacune des métriques et la raison sous-jacente à leur création :"

Nom de la variable	Calcul	Description
$Incurred_n$	Donnée par les cédantes	Dernier montant total du sinistre estimé par la cédante pour l'année n actuelle.
$Incurred_0$	Donnée par les cédantes	Montant de la première évaluation significative du sinistre par la cédante.
Incurred evolution	$\frac{Incurred_n}{Incurred_{n-1}}$	Évolution du montant estimé du sinistre par rapport à l'estimation de l'année précédente.
% Min	$\frac{\min_{k=0 \dots n} (Incurred_k)}{Incurred_n}$	Cette métrique permet de comparer le montant minimum significatif estimé depuis l'année de survenance du sinistre par rapport à l'estimation actuelle du montant du sinistre.
% Max	$\frac{Incurred_n}{\max_{k=0 \dots n} (Incurred_k)}$	Cette métrique permet de comparer le montant actuel estimé par rapport à l'estimation maximum du montant du sinistre depuis son année de survenance.
Incurred moyen	$\frac{\sum_{k=0}^n Incurred_k}{n}$	Cette métrique représente la moyenne des estimations du montant du sinistre.
Moyenne des trois derniers Incurred	$\frac{\sum_{k=n-2}^n Incurred_k}{3}$	Moyenne des trois dernières estimations du montant du sinistre.
Somme des incurred relatif	$\frac{\sum_{k=0}^n Incurred_k}{Incurred_n}$	Somme des situations passées par rapport à la situation actuelle
Grande variation	$\begin{cases} 1 \text{ si } Incurred \text{ evolution} > 1,5 \\ 1 \text{ si } Incurred \text{ evolution} < 0,5 \\ 0 \text{ sinon} \end{cases}$	Métrique valant 1 si le sinistre a beaucoup évolué depuis la dernière situation ($\pm 50\%$)
Nombre de très grandes variations	$\sum_{k=0}^n \text{Grande variation}_k$	Nombre total d'années où le sinistre a connu une grande variation
Nombre de changements de sens de la variation		Nombre total d'années où le sinistre a changé de sens de variation : passage de la hausse à la baisse (ou inversement) entre deux années de développement
Nombre d'années où le sinistre était évalué à un montant insignifiant	$\sum_{k=0}^n 1_{Incurred_k < 100000}$	Cette métrique donne une indication sur le nombre d'années où le sinistre était évalué à un montant en dessous de la plupart des seuils de réassurance.
Paid	Donnée par les cédantes	Montant total du sinistre payé en année n.
Proportion Paid	$\frac{Paid_n}{Incurred_n}$	Proportion du sinistre payé par rapport à l'estimation du montant
Paid evolution	$\frac{Paid_n}{Paid_{n-1}}$	Évolution de la Proportion Paid entre l'année précédente et l'année actuelle

FIGURE 75 – Métriques - Section 5