Mémoire présenté le :

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

Par : Antonin Cadiot

Titre : Une approche du risque de modèle lors de la tarification en réassurance non-vie

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

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

Membre présents du jury de l'Institut des Actuaires

signature

Entreprise : SCOR

Nom : Antonin Cadiot

Signature :

Directeur de mémoire en entreprise :

Nom : Emmanuel Figuereau

Signature :

Invité :

Nom :

Signature :

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

Signature du responsable entreprise

Signature du candidat

Membres présents du jury de l'ISFA

Table des matières

Résumé **4**

Abstract **5**

Remerciements **6**

Introduction **7**

I Domaine d'étude **8**

1 Le marché français de la réassurance automobile **9**
1.1 Les chiffres 9
1.2 Le type de programme étudié : l'excédent de sinistre (XS) 12
1.3 Le risque : le dommage corporel grave 14

2 La tarification en réassurance non-vie non-proportionnelle **19**
2.1 Principes généraux 19
2.2 Principaux modèles de tarification 21
2.2.1 Tarification sur expérience 21
2.2.2 Modèle probabiliste 22
2.2.3 La méthode mixte 24

II Risque d'estimation des paramètres de la méthode mixte **25**

3 Une autre façon d'estimer l'inflation du passé **27**
3.1 Un premier essai par une méthode de changement de variable 27
3.1.1 Données utilisées 27
3.1.2 Description de la méthode 28
3.1.3 Résultats 30
3.2 Limites de la méthode 31
3.2.1 Quand l'inflation est invisible 31
3.2.2 Explication 32
3.2.3 Illustration 33
3.3 Un second essai par une méthode de comptage 34
3.3.1 Description de la méthode 34
3.3.2 Résultats 36

4 Une autre façon de déterminer les facteurs de développement **38**
4.1 Première étape : estimer le développement des sinistres au premier euro 38
4.1.1 Description de la méthode 38
4.1.2 Résultats 40
4.2 Deuxième étape : en déduire le développement des charges de réassurance 42
4.2.1 Le concept d'accélération 42
4.2.2 Application 47

III Risque de spécification du modèle de tarification **49**

5 Étude de la sévérité **51**
5.1 Théorie des valeurs extrêmes 53
5.1.1 Éléments de théorie 53
5.1.2 Application 53
5.2 Étude de lois appartenant au domaine d'attraction de Fréchet 55
5.2.1 La loi de Pareto 55
5.2.2 La loi Pareto Généralisée 55
5.2.3 La loi Exponentielle Pareto 56
5.2.4 Calibrage 57

6 Changement de modèle de tarification **61**
6.1 Le risque de modèle 61
6.1.1 Définitions et problématiques 61
6.1.2 Éléments de théorie 62
6.2 Application 64
6.2.1 Description de la méthode de tarification 64
6.2.2 Tarification de plusieurs scénarios 64
6.2.3 Vraisemblance et robustesse des différents modèles 66
6.2.4 Quantification du risque de spécification du modèle 69

Conclusion **70**

Bibliographie **70**

Annexes **75**
A Résumés statistiques sur les victimes d'un accident corporel grave 75
B Triangle de fonctions de répartition 77
C Résultats de la partie 3.1.3 78
D Résultats de la partie 3.2.3 81
E Mesure de l'incertitude du processus de liquidation 83
E.1 Le modèle de Mack 83
E.2 Le bootstrap 83

Résumé

Les méthodes de tarification en réassurance peuvent être particulièrement complexes, les hypothèses des modèles ne semblent pas toujours parfaitement validées et les données sont parfois en très faible quantité. En conséquence, l'actuaire peut avoir un manque de visibilité sur l'impact de ses décisions en termes de modèle et de calibrage. Cela représente un challenge pour les compagnies de réassurance car la prime pure calculée par les actuaires tarificateurs a des conséquences aussi bien dans la gestion, la stratégie que dans la politique de risque de l'entreprise.

C'est pourquoi la directive Solvabilité 2 encourage la prise d'initiatives autour de la problématique du risque de modèle. C'est dans ce contexte que s'inscrit ce mémoire. L'objectif est d'éclairer les choix de l'actuaire au cours du processus de tarification. Pour cela, le cadre de cette étude et notamment la méthode mixte - qui servira de modèle de référence - seront tout d'abord présentés. Puis une illustration du risque d'estimation des paramètres de cette méthode, via la proposition d'estimateurs discriminants, sera proposée. Ensuite, le questionnement d'une hypothèse courante - à savoir que la sévérité au-delà d'un seuil suit une loi de Pareto - aboutira à la présentation d'autres modèles de tarification. Enfin une proposition de quantification du risque de choisir la méthode mixte comme modèle sera présentée.

Mots clés :
Réassurance non-vie, tarification, risque de modèle, méthode mixte, estimation des paramètres, changement de variable, effet de levier, inflation invisible

Abstract

Pricing models in reinsurance can be particularly complicated, hypotheses don't always seem perfectly validated and data might be scarce. Therefore, the actuary might lack visibility on the impact of his decisions in terms of model and calibration. This represents a challenge for reinsurance companies as the pure premium given by pricing actuaries impacts the management, as well as the strategy and the politics of the company.

This is why the Solvency II directive encourages leadership actions regarding the model risk problematic. It is in this context that this study was carried out. The goal is to lighten the actuary's choices regarding the pricing process. In order to do that, the context of this study and notably the mixed method - which will serve as the reference model - will be presented. Then, an illustration of the risk of estimating the parameters of that method, by the proposal of discriminating estimators, will be proposed. The questionning of a common hypothesis - that the severity above a threshold follows a Pareto distribution - will then lead to the presentation of other pricing models. Finally a proposal of the quantification of the risk of chosing the mixed method as a model will be presented.

Key words :
Non-life reinsurance, pricing, model risk, parameters estimation, variable change, gearing effect, invisible inflation

Remerciements

Je tiens tout d'abord à exprimer ma reconnaissance à toute l'équipe pédagogique de l'Institut de Science Financière et d'Assurances et à mon tuteur Christian Robert, pour m'avoir apporté les connaissances ainsi qu'une ouverture d'esprit nécessaires afin de mieux comprendre ce qu'est l'Actuariat.

Je tiens également à remercier tout particulièrement Alex-Michel Ngningha pour la formation qu'il m'a donnée. La liberté d'entreprendre qu'il m'a accordée, associée à son aptitude à me challenger tout au long de ces douze mois, a fortement contribué au bon déroulement de mon alternance.

Je remercie également Emmanuel Figuereau pour ses idées qui sont à l'origine de ce mémoire, et pour son aide. J'ai été impressionné par sa manière très à l'écoute de manager et par sa capacité de permettre à son équipe de prendre du recul sur les méthodes utilisées.

Merci aussi à Valérie Klein pour son expertise dans le domaine de la tarification en réassurance et pour son ouverture d'esprit au regard de mes propositions de nouveaux outils d'études.

Mes remerciements vont également à François Blanchet pour sa patience et son habileté à expliquer les subtilités du monde juridique.

Merci également à toute l'équipe SCOR pour m'avoir fait partager leurs connaissances et aussi pour les nombreux échanges d'idées qui ont été essentiels à la réalisation de ce mémoire, rendant ainsi cette expérience très enrichissante.

Merci enfin à mes amis et à ma famille, et avant tout à ma mère, pour m'avoir soutenu.

Introduction

Lors de la tarification de traités de réassurance responsabilité civile automobile, l'actuaire utilise en général un modèle qu'il calibre. Il y a ainsi quelques méthodes, majoritairement utilisées, qui servent de référence. L'une d'elles repose sur le modèle fréquence-coût Poisson-Pareto, calibré à l'aide d'une tarification sur expérience dite *burning cost*. Cette façon de tarifer un contrat est appelée "méthode mixte", ou encore "méthode extrapolation Pareto" [1]. L'idée de ce mémoire est, dans le cas de la tarification d'un traité en excédent de sinistre, de s'interroger sur les conséquences du choix de ce modèle, plutôt que d'un autre, et également de s'intéresser à son calibrage.

Les principales problématiques de notre étude peuvent se résumer ainsi : Existe-t-il d'autres modèles aussi pertinents que la méthode mixte ? Et si oui, l'estimateur de la prime pure serait-il vraiment différent avec de tels modèles ?

Pour apporter des éléments de réponse à ces questions, le contexte et notamment le risque sous-jacent à cette étude - à savoir le corporel grave - seront tout d'abord présentés. Puis la tarification en réassurance non-vie et en particulier la méthode mixte - le modèle de référence pour cette étude - seront exposées. Dans une seconde partie, le risque d'estimation des paramètres de la méthode mixte, via la proposition d'autres méthodes pour estimer la valeur de ses principaux paramètres, sera illustrée. Enfin, la troisième partie ira plus loin en présentant une alternative à la méthode mixte. Une quantification du risque de choisir cette dernière comme modèle sera alors proposée.

1. Source : BLONDEAU J. et PARTRAT C. [2003] « La Réassurance Approche technique » , 40-41

Première partie

Domaine d'étude

Chapitre 1

Le marché français de la réassurance automobile

La branche assurance automobile se décompose en deux parties : la responsabilité civile, obligatoire dans tous les pays européens (1958 en France) et la garantie dommages subis par le véhicule ou par le conducteur. Cette séparation repose sur l'idée que l'on ne peut être responsable du dommage causé à soi-même. Les dommages causés à autrui relèvent de la responsabilité civile. Cependant on note que cette distinction est remise en cause par le projet de réforme de la Responsabilité Civile Automobile (RCA).

1.1 Les chiffres

Les principaux chiffres selon la Fédération Française de l'Assurance [**9**] sont les suivants :

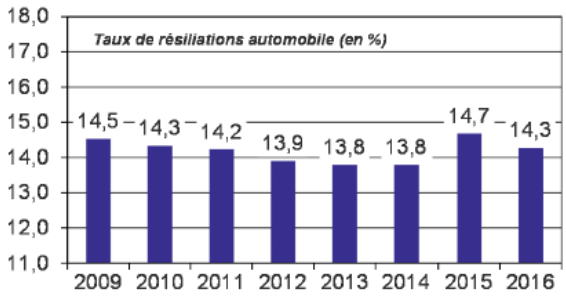
Le parc des véhicules de première catégorie hors flottes est estimé en 2016 à 41,591 millions de véhicules dont 40,467 millions assurés en métropole.

(*En milliers*)	2016	Variation 2016/2015
Véhicules assurés en mono contrat	**48 306**	**+ 1,6 %**
Véhicules de 1re catégorie	**41 591**	**+ 1,8 %**
Dont assurés en métropole	40 467	+ 1,8 %
Véhicules de 3e catégorie	**3 915**	**- 0,4 %**
Autres véhicules assurés en mono contrats(1)	**2 800**	**+ 0,9 %**

(1) Véhicules agricoles, véhicules de deuxième catégorie, missions, remorques, caravanes, engins spéciaux, polices garages, etc…

Table 1.1 – Nombre de véhicules assurés au 31 décembre 2016

Le taux de couverture des contrats, qui est le rapport entre le taux d'affaires nouvelles et le taux de résiliations, est de 113,6 % en 2016.



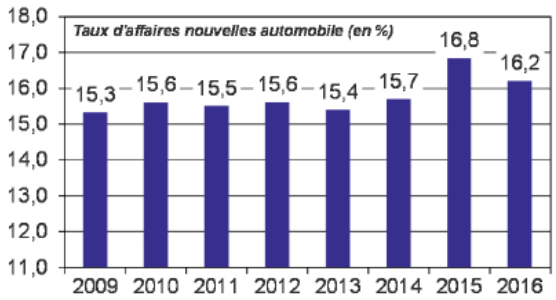


FIGURE 1.1 – Taux de résiliations et taux d'affaires nouvelles en assurance automobile

Le chiffre d'affaires de l'assurance automobile, avec 20,8 milliards d'euros en 2016 représente 39% de l'ensemble des cotisations des assurances de dommages aux biens et de responsabilité civile.

(En milliards d'euros)	2015	2016 (p)	Répartition en 2016	Variation 2016/2015
Ensemble automobile	**20,4**	**20,8**	**100 %**	**+ 1,8 %**
Véhicules assurés hors flottes :	18,5	18,8	90 %	+ 1,8 %
- dont véhicules de 1re catégorie	16,4	16,7	80 %	+ 1,7 %
- dont véhicules de 3e catégorie	0,9	0,9	4 %	+ 0,8 %
- dont autres contrats (1)	1,2	1,2	6 %	Ns
Contrats flottes (hors missions)	2,0	2,0	10 %	+ 1,5 %

(p) provisoire Ns : non significatif

(1) Véhicules agricoles, véhicules de deuxième catégorie, missions, remorques, caravanes, engins spéciaux, polices garages, etc….

TABLE 1.2 – Cotisations selon le type de contrat et le type de véhicule assuré

La prime moyenne d'un véhicule assuré au tiers (RC, vol, incendie et bris de glaces) en 2016 est estimée à 237 euros. Pour un contrat tous risques, la prime moyenne est de 487 euros.

	Prime moyenne 2016	Variation 2016/2015	Variation 2015/2014
Ensemble Véhicules 1re catégorie	400 €	0 %	0 %
Véhicules assurés au tiers (1)	237 €	- 0,1 %	+ 0,3 %
Dont Responsabilité civile	146 €	0 %	+ 0,3 %
Tous risques (hors assistance)	487 €	- 0,1 %	+ 0,4 %

(1) Véhicules assurés en RC, Vol, Incendie et Bris de glaces

TABLE 1.3 – Niveau de prime moyenne selon le type de contrat

En 2016, le ratio combiné comptable net de réassurance s'établit à 105 % des primes pour l'ensemble du marché automobile.

Les sinistres corporels au titre de la responsabilité civile représentent la charge la plus importante :

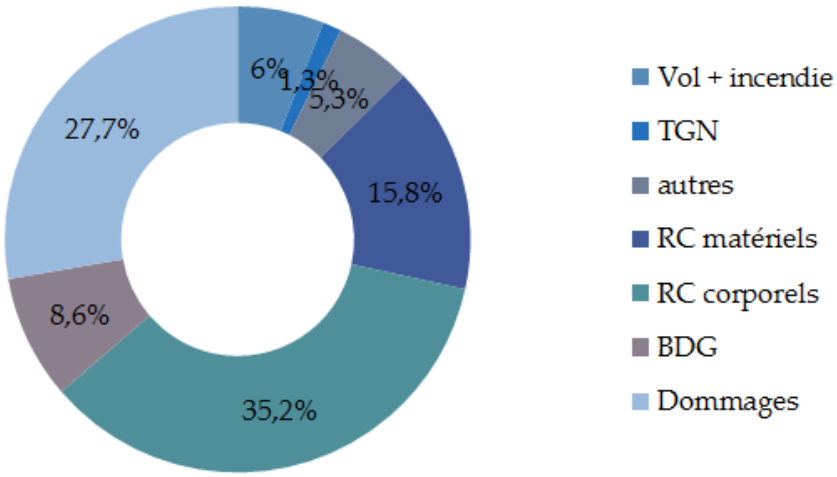


FIGURE 1.2 – Répartition de la charge des sinistres selon la garantie

La fréquence au titre de la responsabilité civile des accidents corporels, le risque sous-jacent à l'étude présentée dans ce mémoire, a augmenté en 2016.

	Fréquence (1) 2016	Variation 2016/2015	Coût moyen 2016	Variation 2016/2015
RC	**37,6 ‰**	- 0,3 %	///	///
- dont RC corporels	**3,7 ‰**	+ 2,7 %	///	///
- dont RC matériels	**33,9 ‰**	- 0,6 %	**1 405 €**	+ 2,7 %
Vol	**5,1 ‰**	- 9,5 %	**3 270 €**	+ 6,3 %
Bris de glaces	**66,1 ‰**	- 3,9 %	**430 €**	+ 4,4 %
Dommages tous accidents	**89,5 ‰**	+ 0,7 %	**1 470 €**	+ 3,6 %

(1) La fréquence représente le nombre de sinistres avec suite pour 1 000 véhicules assurés pour cette garantie

TABLE 1.4 – Évolution des fréquences et des coûts moyens suivant les garanties souscrites

L'indemnité totale moyenne d'un blessé léger sans AIPP est d'environ 1 500 euros en 2015. Avec AIPP, l'indemnité moyenne d'un blessé est estimée en 2015 à 74 900 euros. Ainsi, le coût moyen d'une victime blessée se situe aux alentours de 25 600 euros en 2015.

Remarque : l'atteinte permanente à l'intégrité physique et/ou psychique (AIPP, aussi appelé taux d'AIPP ou encore DFP pour Déficit Fonctionnel Permanent) est une évaluation sur une échelle de 0 à 100 qui indique la réduction du potentiel physique, psychosensoriel ou intellectuel dont est atteinte la victime. Plus le chiffre est élevé, plus le traumatisme est important.

Les coûts des postes de préjudices "assistance par tierce personne" et "préjudices professionnels" sont en augmentation sur la période 2005 à 2015, respectivement de 11,4% et 11,7% par an en moyenne. Ces deux postes de préjudices représentent une part majeure du coût des sinistres de la base de données utilisée dans cette étude.

	Coût moyen 2014	Coût moyen 2015	*Variation du coût moyen 2015/2014*	Variation moyenne annuelle	
				Sur la période 2010 à 2015	Sur la période 2005 à 2015
Ensemble des indemnités	**73 400 €**	**74 900 €**	**+ 2,0 %**	**+ 4,1 %**	**+ 6,3 %**
Dont :					
Dépenses de santé actuelles et futures	16 800 €	17 100 €	+ 1,7 %	+ 2,0 %	+ 3,3 %
Préjudices professionnels (1)	14 950 €	15 110 €	+ 1,1 %	+ 7,3 %	+ 11,7 %
Déficit Fonctionnel Permanent	11 980 €	12 180 €	+ 1,7 %	+ 1,5 %	+ 2,1 %
Préjudice personnels (2)	6 580 €	6 590 €	+ 0,1 %	+ 1,9 %	+ 1,9 %
Assistance par tierce personne	14 900 €	15 700 €	+ 5,7 %	+ 10,2 %	+ 11,4 %
Autres indemnités (*)	8 220 €	8 140 €	- 0,9 %	- 0,1 %	+ 11,3 %

(1) *Pertes de gains professionnels actuelles et futures, Incidence professionnelle et Préjudice scolaire & universitaire*
(2) *Souffrances endurées, préjudice esthétique temporaire et permanent et préjudice d'agrément*
(*) *Les autres indemnités comportent les postes de préjudice non détaillés ci-dessus*

TABLE 1.5 – Évolution des indemnités des blessés avec AIPP des principaux postes de préjudices en 2015

1.2 Le type de programme étudié : l'excédent de sinistre (XS)

Sur le marché français, la responsabilité civile automobile se réassure principalement avec des traités en excédent de sinistre. La partie ci-dessous présente ce type de traité. Si le lecteur est familier avec son fonctionnement et ses différentes clauses, il pourra alors directement aller à la partie 1.3.

Un tel traité fonctionne ainsi, pour chaque montant de sinistre, la cédante supporte une franchise, que l'on appelle priorité. Le réassureur prend à charge la portion du sinistre qui dépasse la priorité, jusqu'à une certaine limite, que l'on appelle la portée. Une illustration est proposée ci-dessous dans le cas d'un traité XS ayant une priorité de 250 et une portée de 500.

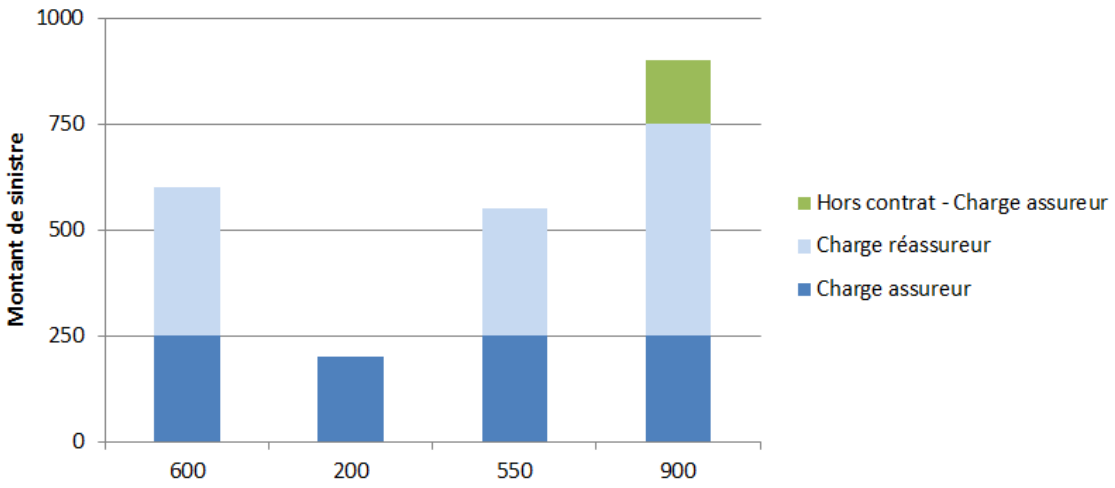


TABLE 1.6 – Illustration du fonctionnement d'un traité en excédent de sinistre par risque

En général sur le marché français un tel traité est résumé dans la note de couverture, appelée *slip* en anglais. Il comporte au titre de la garantie responsabilité civile :

- La période de couverture
- Les risques couverts
- Les exclusions
- La priorité
- La portée
- Les clauses
- La prime de réassurance
- Les informations sinistres
- Le traitement des rentes
- Le courtage
- L'assiette de prime
- Les informations diverses

Nous allons maintenant présenter quatre des principales clauses présentes dans les traités de réassurance couvrant la responsabilité civile automobile.

- Clause de stabilisation

Cette clause s'utilise dans les branches à développement long, c'est-à-dire lorsqu'il s'écoule plusieurs années entre la survenance du sinistre et les différents paiements. Elle a pour objectif de garder le partage du sinistre dans les mêmes proportions qu'à la signature du contrat, en prenant en compte l'inflation.

Comme expliqué dans BLONDEAU et PARTRAT [2003] [**2**] la clause fonctionne ainsi : pour chaque sinistre dont le paiement total jusqu'à la date de calcul t dépasse la priorité, on applique aux paiements partiels annuels P_n un facteur I_0/I_n, défini par un indice I, actualisant ce paiement en valeur initiale de l'année de survenance du sinistre. On obtient alors les paiements ajustés. En faisant le rapport de la somme des paiements annuels par la somme des paiements annuels ajustés, on obtient un coefficient supérieur à 1. Ce coefficient est alors multiplié, selon le contrat, soit à la priorité soit à la portée, soit aux deux. C'est ainsi que l'on obtient les nouvelles priorités et limites du contrat.

On note toutefois que le taux contractuel défini dans la clause de stabilisation ne reflète pas parfaitement l'inflation du sinistre. En effet, une part d'inflation sociale et juridique n'est pas prise en compte et est entièrement prise en charge par la réassurance.

- Clause de rachat

Cette clause concerne les sinistres donnant lieu au versement d'une rente. Dans ce cas le montant du capital constitutif de chaque rente sera calculé à partir, notamment, de la table et du taux précisés dans la clause de rachat. Le montant ainsi obtenu sera en pratique exempt de tout chargement. Le réassureur participera alors à l'indemnisation par le versement de sa part au capital constitutif, calculé en général après application de la clause de stabilité. Le réassureur sera alors libéré de ses engagements sous plusieurs réserves définies contractuellement. Par exemple en cas de modification de l'état de santé et/ou d'aggravation de la situation de la victime avant un certain délai, alors le montant à charge du réassureur pourra être revu.

Remarque : On supposera par la suite que nous sommes dans le cas d'un traité de réassurance ayant une clause de rachat avec la table TD 88-90 et un taux de 1,5 %.

- Clause de franchise annuelle

Cette clause, aussi appelée AAD pour *Annual Aggregate Deductible*, est présente lorsque la cédante souhaite garder une partie de ses risques cédés. Elle a pour effet de baisser le coût de la réassurance. Dans ce cas, le traité s'écrit Portée XS Priorité XS AAD.

Ainsi dans le cas où la tranche de réassurance est touchée, on a :

$$\text{Charge}_{\text{Réass}} = \sum_{i=1} \min(\max(S_i - \text{Priorité}, 0), \text{Portée}) - \text{AAD}$$

où S_i représente le montant du i-ième sinistre.

Une attention doit être prise au niveau de la volatilité de la prime pure qui augmente avec cette clause. En effet avec un AAD le prix de la réassurance diminue, car l'espérance de la variable aléatoire égale aux paiements futurs du réassureur diminue. Néanmoins la variance de cette même variable aléatoire augmente, ce qui augmente le chargement ajoutée à la prime pure. Il y a donc ce qu'on appelle un dilemme biais-variance.

- Clause de plafond annuel

Cette clause, aussi appelée AAL pour *Annual Aggregate Limit*, est présente lorsque le réassureur souhaite mettre un plafond à sa charge. Elle a pour effet d'augmenter le prix de la réassurance.

On écrit :

$$\text{Charge}_{\text{Réass}} = \min[\sum_{i=1} \min(\max(S_i - \text{Priorité}, 0), \text{Portée}), \text{AAL}]$$

Dans le cas d'un contrat avec les deux clauses AAD et AAL, la charge du réassureur s'écrit :

$$\text{Charge}_{\text{Réass}} = \min[\sum_{i=1} \min(\max(S_i - \text{Priorité}, 0), \text{Portée}) - \text{AAD}, \text{AAL}]$$

Pour plus de détails sur les autres types de clause le lecteur pourra se référer à BLONDEAU et PARTRAT [2003] [**2**].

Dans cette partie le fonctionnement des traités XS et ses principales clauses ont été présentés. Une présentation du principal risque couvert par ce type de traité dans le cadre de l'assurance automobile va maintenant être proposée.

1.3 Le risque : le dommage corporel grave

La responsabilité civile automobile est ce que l'on appelle une "branche longue", c'est à dire qu'il peut s'écouler plusieurs années entre la survenance du sinistre et l'estimation de son coût final,

puis de son règlement. Une illustration de la vie d'un tel sinistre est proposée ci-dessous :

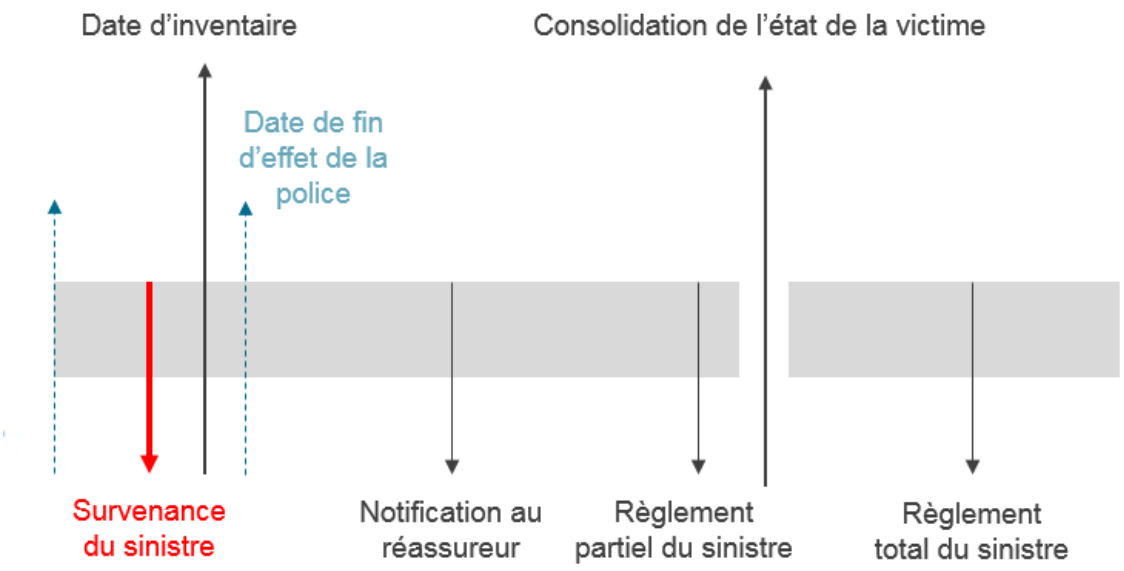


FIGURE 1.3 – Vie d'un sinistre grave

La consolidation de l'état de la victime est une notion clé autour de laquelle s'articule la réparation des dommages. En droit, la réparation doit être entière, le traumatisme de la victime doit donc être parfaitement connu. Néanmoins, les dommages de la victime peuvent évoluer dans le temps ce qui rend impossible l'estimation du coût du sinistre. C'est là qu'intervient la date de consolidation, qui correspond au moment à partir duquel un médecin expert estime que l'état de la victime n'évoluera plus. En outre avant cette date l'indemnisation de la victime concerne des postes de préjudices temporaires, au sens de la nomenclature Dinthillac. Ce n'est en effet qu'après la consolidation de la victime que l'on peut définir le coût des postes de préjudices permanents.

Plusieurs résumés statistiques extraits d'une étude d'un réassureur [**17**] sur les victimes d'un accident corporel grave, qui représente la très grande majorité des sinistres de la base de données de cette étude, sont proposés en annexe A.

Nous allons maintenant présenter brièvement les enjeux auxquels sont soumis les assureurs face à l'indemnisation du corporel grave en France. Les principaux sont la définition de la place de la personne en situation de handicap au sein de la société française, l'augmentation du coût des dossiers, l'absence de règles d'indemnisation unifiées, le projet de réforme de la responsabilité civile du conducteur et l'arrivée de nouveaux véhicules terrestres à moteur.

- Place de la personne en situation de handicap au sein de notre société

La place de la personne en situation de handicap au sein de notre société est un enjeu car le montant d'un sinistre au titre de la responsabilité civile automobile est défini en fonction du principe de réparation intégrale du préjudice. En effet, l'assureur s'engage à replacer la personne dans l'état le plus proche qu'elle aurait été si elle n'avait pas eu l'accident. C'est à ce titre que l'on a besoin de définir la place de la personne en situation de handicap au sein de notre société.

Les principaux textes de lois y faisant référence sont :

- La loi du 11 février 2005 qui porte sur l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées.

- Le Livre Blanc sur l'Indemnisation du Dommage Corporel publié par les assureurs en avril 2008. Il propose des réponses adaptées aux attentes particulières des victimes lourdement handicapées.

- Augmentation du coût des dossiers

Les dommages corporels couverts par la RC représentent une part majeure des prestations versées aux assurés. Or le coût du corporel grave ne cesse d'augmenter. En effet, d'après une étude sur la responsabilité civile automobile **[16]** on constate une hausse très importante chaque année du coût du corporel grave. De plus cette évolution semble déconnectée des indicateurs économiques.

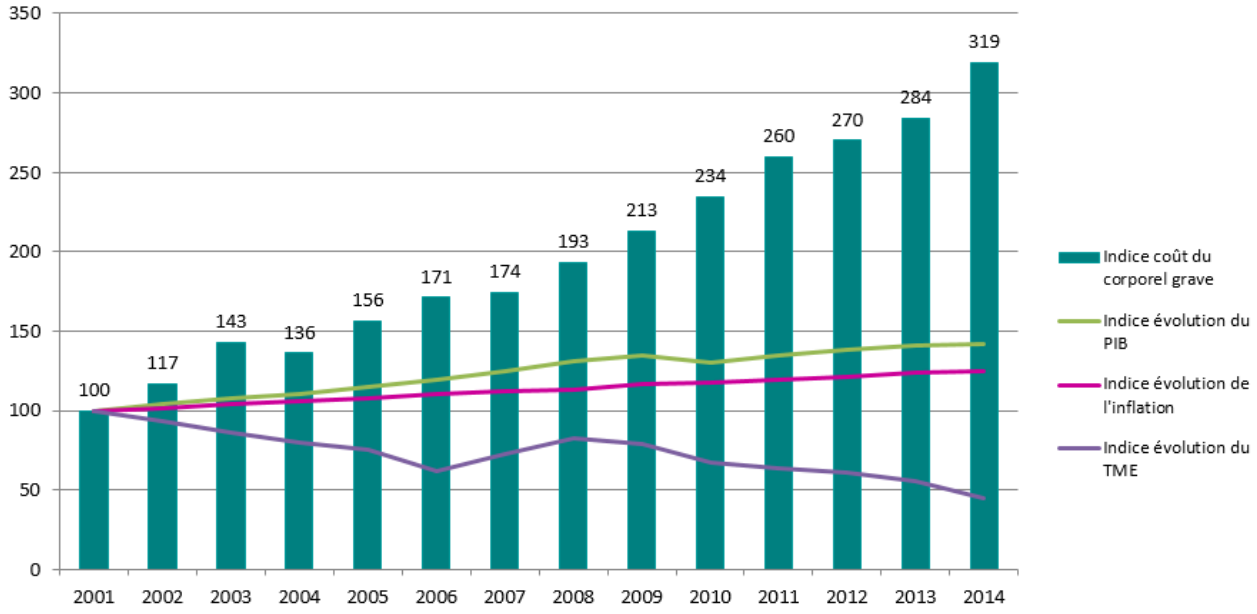


FIGURE 1.4 – Évolution du PIB et de l'inflation par rapport au coût moyen d'un sinistre corporel grave

- Absence de règles d'indemnisation unifiées

En France, contrairement en Belgique et en Espagne par exemple, il n'y a pas de texte restrictif qui encadre l'indemnisation du corporel grave. Seule la nomenclature Dinthillac, en place depuis 2006, harmonise les postes de préjudices mais n'est qu'indicative.

On résume ci-dessous les principaux postes de préjudices définis par cette nomenclature :

	Patrimoniaux	Personnels
Permanents	Dépenses de santé futures Frais de logement adapté Frais de véhicule adapté Assisstance par tierce personne Pertes de gains professionnels futurs Incidence professionnelle Préjudice scolaire, universitaire ou de formation	Déficit fonctionnel permanent Préjudice d'agrément Préjudice esthétique permanent Préjudice sexuel Préjudice d'établissement Préjudices permanents exceptionnels
Temporaires	Dépenses de santé actuelles Frais divers Pertes de gains professionnels actuels	Déficit fonctionnel temporaire Souffrances endurées Préjudice esthétique temporaire

FIGURE 1.5 – Nomenclature des principaux préjudices corporels proposée par la commission Dinthillac

- Réforme de la responsabilité civile automobile

Un texte sur un projet de réforme de la responsabilité civile est à l'ordre du jour depuis 2017. Ce projet semble à ce jour en veille puisqu'il n'est pas inscrit au calendrier parlementaire. Néanmoins la Chancellerie reste déterminée à changer le droit à l'indemnisation du conducteur. En outre ce changement est voulu rétroactif ce qui forcerait les assureurs à indemniser des victimes relatives à des risques qu'ils n'ont pas provisionnés. C'est pourquoi la Fédération Française de l'Assurance travaille sur une alternative à proposer.

- Les nouveaux véhicules terrestres à moteur

L'article 1 de la loi du 5 juillet 1985, aussi appelée loi Badinter, énonce que :

Article 1 : "Les dispositions du présent chapitre s'appliquent, même lorsqu'elles sont transportées en vertu d'un contrat, aux victimes d'un accident de la circulation dans lequel est impliqué un véhicule terrestre à moteur ainsi que ses remorques ou semi-remorques, à l'exception des chemins de fer et des tramways circulant sur des voies qui leur sont propres."

Ainsi la loi Badinter repose sur 4 notions clés. La première est la notion de victime, la seconde est la notion d'accident de la circulation, la troisième est la notion d'implication et la dernière est la notion de véhicule terrestre à moteur. Pour cette sous-partie on va se concentrer sur cette dernière notion.

On définit un véhicule terrestre à moteur comme un engin circulant sur le sol, équipé d'un moteur lui conférant une force cinétique propre susceptible de servir à un transport même s'il ne transporte que le conducteur lui-même.

On remarque que ces trois critères semblent remplis aujourd'hui par des véhicules qui ne roulaient pas sur les routes en 1985. C'est le cas par exemple des voitures autonomes ou des *segway* qui

soulèvent plusieurs questions. Sont-ils vraiment des véhicules terrestres à moteur ? La question se pose tout particulièrement pour les *driving wheels*. En outre, comment se passe l'indemnisation en cas de cause de dommages à un tiers par le conducteur fautif d'un tel véhicule, une trottinette électrique par exemple ? À ce titre on cite l'article 2 de la loi du 5 juillet 1985 :

Article 2 : "Les victimes, y compris les conducteurs, ne peuvent se voir opposer la force majeure ou le fait d'un tiers par le conducteur ou le gardien d'un véhicule mentionné à l'article 1er."

Ainsi même s'il a été jugé responsable et même s'il représente la cause exclusive de l'accident, l'indemnisation des dommages doit rester prise en charge par l'assureur du responsable.

En conclusion, on voit que le risque relatif au corporel grave est en pleine évolution et que la jurisprudence va devoir continuer à intervenir.

Chapitre 2

La tarification en réassurance non-vie non-proportionnelle

Un rappel des différents modèles de tarification en réassurance non-proportionnelle, la problématique sous-jacente à ces modèles et leurs caractéristiques principales, issus de BLONDEAU et PARTRAT [2003] [**2**], vont maintenant être proposés.

Les principaux modèles utilisés en réassurance non-proportionnelle sont applicables soit dans un cadre marché avec des statistiques globales, soit à partir des données individuelles propres à une cédante en particulier.

2.1 Principes généraux

Le prix du risque, ou prime pure, vient de la sinistralité moyenne attendue au-delà de la franchise. On détaille les principaux facteurs dont dépendent cette sinistralité :

- Environnement socio-économique

- Comportement du marché relativement au risque
- Comportement des assurés vis-à-vis de l'assurance
- Contexte juridique

- Politique de souscription de la cédante

- Objectif de souscription
- Appétence au risque
- Composition du portefeuille
- Moyens préventifs mis en œuvre

- Profil de risque

- Volume de risque
- Primes

- Évolution du contrat

- Changement de garanties
- Franchises
- Exclusions

Afin de prendre en compte le comportement d'un portefeuille au regard des facteurs mentionnés, le réassureur se base généralement sur la sinistralité historique. La question qu'il se pose est : Quelle va être la charge future compte tenu de l'historique sinistre ?

Une idée pour répondre à cette question est de faire l'hypothèse que la sinistralité du passé sera représentative de la sinistralité future. Il faut alors calculer combien coûteraient aujourd'hui les sinistres ayant précédemment eu lieu dans le passé, s'ils venaient à se reproduire.

Les différentes composantes d'une mise en *as-if* permettent d'apporter des éléments de réponse à cette problématique. On va maintenant détailler cette étape qui se décompose en 3 parties : l'actualisation de la statistique, le redressement de la statistique et la revalorisation de la statistique.

- L'actualisation de la statistique

Cette étape consiste à effacer l'impact des changements monétaires. L'idée est donc de se mettre sur la base d'une unité monétaire constante.

- Le redressement de la statistique

Cette seconde étape consiste à prendre en compte l'évolution des profils de risques (valeurs assurées, nombre de polices...) et les éventuels changements sur la nature des garanties accordées (ajout ou retrait de clause, exclusions de garantie...). Il s'agit donc de se positionner sur une base de portefeuille homogène dans le temps.

- La revalorisation de la statistique

Cette dernière étape doit prendre en compte l'évolution du coût du risque, par exemple dans le cas qui nous intéresse l'évolution des coûts du corporel grave en responsabilité civile automobile. Mais également les changements juridiques, dans le cas qui nous intéresse dans ce mémoire cela peut être par exemple la réforme du Fond de Garantie des Assurances Obligatoires (FGAO) en 2013 ou la potentielle future réforme de la responsabilité civile automobile évoquée partie 1.3. Le but ici est donc de se replacer sur la base d'un environnement homogène.

Dans la pratique, l'actualisation, le redressement et la revalorisation de la statistique se traduisent en général par la mise en place de deux indices :

- Un indice prime : il est censé refléter l'évolution de la prime originale moyenne.

- Un indice sinistre : il est censé refléter l'évolution du coût des sinistres.

Remarque : L'indice sinistre peut en pratique prendre la forme de 2 taux, un pour l'inflation du passé et un autre pour l'inflation du futur. L'objectif du chapitre suivant sera de proposer une méthode de détermination de l'estimateur de l'inflation du passé.

2.2 Principaux modèles de tarification

Les premières questions à se poser pour choisir un modèle de tarification sont :

- Quelle est la nature de la réassurance considérée : Traité ou Facultative ?
- Quelle est la nature de la branche considérée : Courte ou Longue ?
- Quelle est la nature de la tranche considérée : Travaillante ou non touchée par la sinistralité historique ?

D'après BLONDEAU et PARTRAT [2003] [**2**], les modèles de tarification peuvent se résumer ainsi :

Catégorie	Nom du modèle
Modèles basés sur l'historique sinistre	Tarification sur expérience Modèle probabiliste Tarification par simulations
Autres modèles	Tarification sur exposition

FIGURE 2.1 – Vue d'ensemble des modèles

2.2.1 Tarification sur expérience

Le principe de cette méthode est de calculer un taux pur appelé *burning cost*. Ce taux est une moyenne, pondérée par l'exposition, de ratios sinistres sur primes. Les sinistres en question sont en général revalorisés, l'objet de la partie 2 sera de proposer des estimateurs nécessaires à cette revalorisation. En effet, le chapitre 3 étudiera l'indexation des sinistres et le chapitre 4 étudiera la liquidation des sinistres.

Une illustration de l'impact sur le montant de sinistre original de cette revalorisation est maintenant proposée :

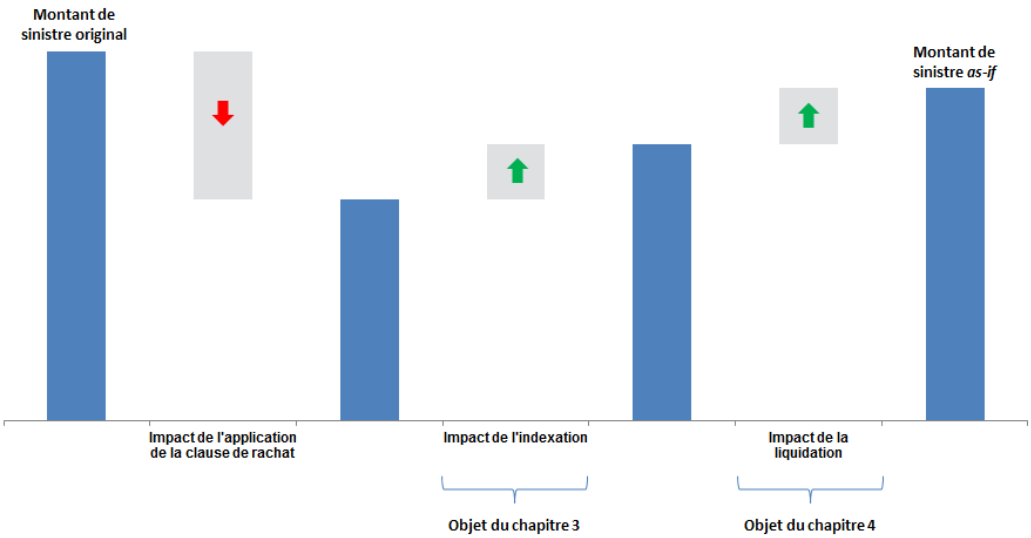


FIGURE 2.2 – Impact des différentes étapes de la mise en *as-if* d'un sinistre

Formellement le *burning cost* se calcule ainsi :

- x_i^j le montant *as if* du i-ième sinistre relatif à l'année j.
- n^j désigne le nombre de sinistres observés à l'année j.
- f désigne la franchise, aussi appelée priorité
- p désigne la portée
- $c_{as\ if}^j = \sum_{i=1}^{n^j}[(x_i^j - f) \wedge p]\ 1_{(x^j \geq f)}$ désigne la part des montants revalorisés, à charge de l'XS, l'année j.
- a^j l'assiette de prime relative à l'année j.
- s désigne le nombre d'années d'historique.

Le taux pur est alors donné par la formule :

$$BC_{pur} = \frac{1}{s} \sum_{j=1}^{s} \frac{c_{asif}^j}{a_j}$$

Pour plus de détails le lecteur pourra se diriger vers BLONDEAU et PARTRAT [2003] [**2**].

2.2.2 Modèle probabiliste

Le principe de cette méthode est de modéliser le coût des sinistres (X) et leur nombre (N), on en déduit la modélisation des montants à charge du réassureur (S_N). Le prix de l'XS est alors l'espérance de la variable aléatoire S_N. Classiquement la loi composée Poisson-Pareto est utilisée pour modéliser S_N.

— Modélisation du montant d'un sinistre individuel

Les deux lois les plus couramment utilisées en réassurance sont la loi de Pareto et la loi LogNormale.

Distribution et paramètres	Fonction de répartition
Pareto Pa(t, α) t > 0, $\alpha > 0$, x $>$ t	$F(x) = 1 - (\frac{t}{x})^{\alpha}$
Log Normale LN(t, m, σ) t > 0, $\sigma > 0$, x $>$ t	F(x) $= \Phi(\frac{ln(x-t)-m}{\sigma})$ Φ fonction de répartition de la loi $N(0,1)$

Pour l'estimation des paramètres, on choisit généralement la méthode du maximum de vraisemblance, la méthode des moments ou une méthode de minimisation des écarts.

Pour la loi de Pareto $Pa(t, \alpha)$ on retient :

- l'estimateur du maximum de vraisemblance de t :

$$\hat{t} = \min_i(x_i)$$

- l'estimateur du maximum de vraisemblance de α :

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} ln(\frac{x_i}{t})}$$

Pour la loi Log Normale $LN(t, m\sigma)$ on retient :

- l'estimateur de t :

$$\hat{t} = \min_i(x_i)$$

- l'estimateur de m :

$$\hat{m} = \frac{1}{n} \sum_{i=1}^{n} ln(\frac{x_i}{\hat{t}})$$

- l'estimateur de σ :

$$\hat{\sigma} = [\frac{1}{n} \sum_{i=1}^{n} (ln(\frac{x_i}{\hat{t}}))^2 - (\frac{1}{n} \sum_{i=1}^{n} ln(\frac{x_i}{\hat{t}}))^2]^{\frac{1}{2}}$$

— Modélisation du montant de la fréquence

En plus de modéliser la loi du montant d'un sinistre individuel, le modèle probabiliste nécessite une modélisation de la loi de fréquence des sinistres.

Si on note N la variable aléatoire qui représente le nombre total annuel de sinistres, alors on suppose en général que N suit une loi de Poisson.

Distribution et paramètre	Loi de probabilité	Moments
Poisson $P(\lambda)$, $\lambda > 0$	$P(N = n) = \frac{e^{-\lambda}\lambda^n}{n!}$	$E(N) = Var(N) = \lambda$

Lorsque la moyenne empirique est trop éloignée de la variance empirique, la loi Binomiale Négative est aussi souvent utilisée. D'autres lois plus fines peuvent également être utilisées, c'est le cas des lois Poisson Inverse Gaussienne, Poisson Inverse Gaussienne Généralisée ou Poisson Gamma Généralisée.

La fréquence et la sévérité étant définies, on peut donc calculer la prime pure de l'excédent de sinistre qui correspond à l'espérance mathématique de S_N.

Avec les mêmes notations et hypothèses on a :

$$E[S_N] = E[\sum_{i=1}^{N} Y_i\ ] = E[N]\ E[Y]$$

$$E[Y] = l\ P(X > l) - f\ P(X \geq f) + E[X\ 1_{f \leq X \leq l}]$$

Par exemple, dans le cas de la loi de Pareto Pa(t, α) de paramètre $\alpha > 1$, on obtient :

- Si l est infinie :

$$E[S_N] = \lambda \, \frac{t^\alpha}{\alpha - 1} \, f^{1-\alpha}$$

- Si l est finie :

$$E[S_N] = \lambda \, \frac{t^\alpha}{\alpha - 1} \, (f^{1-\alpha} - l^{1-\alpha})$$

2.2.3 La méthode mixte

Il existe une méthode, très utilisée en pratique, appelée la méthode mixte ou méthode "extrapolation Pareto" [1]. Cette méthode calibre un modèle probabiliste de façon à ce que la prime pure obtenue par le modèle probabiliste sur la tranche travaillante corresponde à la prime pure donnée par le *burning cost*. En particulier, ce sont les choix relatifs à cette méthode que l'on va challenger par la suite.

Formellement, on suppose a priori que la loi du montant total à charge de l'XS est une loi composée Poisson-Pareto et on note :

- f_t la franchise travaillante, elle correspond au montant du plus petit sinistre.
- l_t la limite travaillante, elle correspond au montant du plus gros sinistre.
- $p_t = l_t - f_t$ désigne la portée travaillante.
- P_{tt}^{BC} = Prime pure de l'XS P_t XS f_t obtenue avec la méthode *burning cost*.
- $\hat{\lambda}$ est l'estimateur du paramètre de la loi de Poisson.

Comme on l'a vu précédemment, la méthode probabiliste avec une loi composée Poisson-Pareto donne, avec $\alpha > 1$ et une limite l finie :

$$\mathrm{E}[S_N] = \lambda \, \frac{t^\alpha}{\alpha - 1} \, (f^{1-\alpha} - l^{1-\alpha})$$

Le $\hat{\alpha}$ retenu pour cette méthode sera alors le α solution de l'équation :

$$P_{tt}^{BC} = \hat{\lambda} \, \frac{f_t^\alpha}{\alpha - 1} \, (f_t^{1-\alpha} - l_t^{1-\alpha})$$

On suppose donc que la sévérité d'un sinistre individuel suit une loi de Pareto Pa (f_t, $\hat{\alpha}$).

La cotation se base donc sur un modèle probabiliste calibré avec un modèle par expérience *burning cost*.

D'autres méthodes peuvent également être utilisées, c'est le cas de la tarification par simulation et de la tarification sur exposition.

1. Source : BLONDEAU et PARTRAT [2003] "La Réassurance Approche Technique", 40-41

Deuxième partie

Risque d'estimation des paramètres de la méthode mixte

Préambule :

Lors de l'exercice de tarification d'un traité de réassurance en excédent de sinistre, avec la méthode mixte, pour une branche longue, il y a en général deux étapes majeures où l'actuaire doit estimer des paramètres :

- La première étape est l'indexation du triangle de sinistralité. En pratique, cette étape peut-être effectuée avec deux taux, un pour l'inflation annuelle du passé, et un autre pour l'inflation annuelle du futur. L'objet du chapitre 3 sera d'étudier une méthode qui utilise exclusivement les données fournies par la cédante, afin de déterminer un estimateur du taux annuel de l'inflation du passé.

- La seconde étape est la liquidation. En pratique, la méthode Chain-Ladder est parfois utilisée avec des facteurs de développement dits "marchés". C'est-à-dire avec des facteurs de développement obtenus avec une étude portant sur un ensemble de cédantes d'un même marché. L'objet du chapitre 4 sera d'étudier une autre méthode qui utilise exclusivement les données de la cédante, afin de déterminer ces estimateurs.

En résumé, l'objet de cette partie est de proposer une illustration du risque d'estimation des paramètres de la méthode mixte, via l'étude d'estimateurs discriminants.

Chapitre 3

Une autre façon d'estimer l'inflation du passé

3.1 Un premier essai par une méthode de changement de variable

Dans un premier temps, une autre façon de déterminer l'estimateur du taux annuel de l'inflation du passé sera étudiée. Pour cela une approche basée sur des triangles de fonctions de répartition, originalement développée dans le document FIGUEREAU [1992] [**10**], sera proposée.

3.1.1 Données utilisées

Les données utilisées proviennent des sinistres d'une cédante importante du marché automobile français. On utilise tous les sinistres dont l'assureur a eu connaissance, en cours et clos, ayant un jour été supérieurs au seuil de communication et survenus entre le 1er janvier 2002 et le 31 juin 2017. Cette base comporte 1176 sinistres au titre de la responsabilité civile automobile, qui sont en très grande majorité relatifs à des accidents corporels graves.

Voici la liste des informations fournies par la cédante et utilisées dans cette étude :

- L'historique du portefeuille, c'est-à-dire l'évolution du nombre de polices au cours du temps.
- Les assiettes de primes historiques.
- Le triangle des sinistres en base gestion, avec le détail pour chaque sinistre, des payés et des réserves à chaque arrêté comptable.
- la valeur du seuil de communication, au-dessous duquel la cédante ne communique pas ses sinistres.
- Le détail sur le calcul des montants des rentes, avec notamment la table, le taux, le fractionnement, le terme (échu ou à échoir) et l'âge de la victime.

Cette dernière information est nécessaire car cette étude se place dans le cas d'un traité de réassurance ayant une clause de rachat, basée sur la table TD 88-90 avec un taux de 1,5 %. Ainsi pour les sinistres ayant des postes de préjudices, au sens de la nomenclature Dinthillac, ayant été provisionnés en rente et non en capital, on procède à un retraitement avec la table et le taux de la clause de rachat.

Une illustration des données utilisées est proposée ci-dessous :

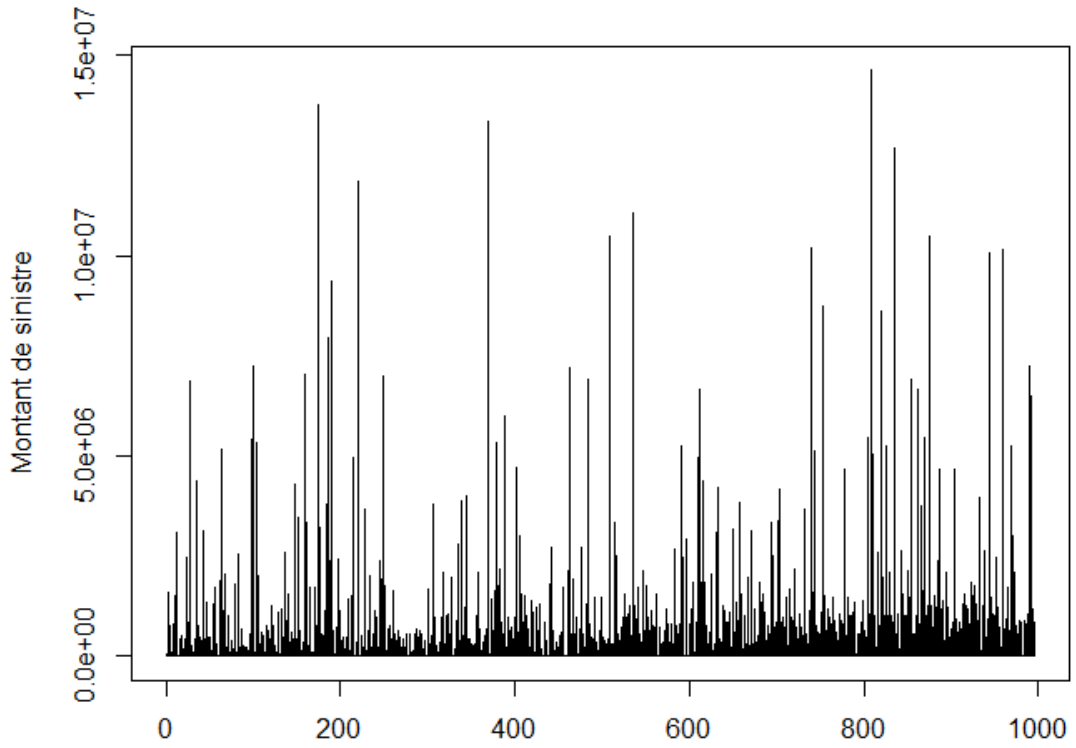


FIGURE 3.1 – Échantillon des montants de 1000 sinistres extrait de notre base de données

Remarque : pour des raisons de confidentialité les données ont été anonymisées.

3.1.2 Description de la méthode

Premièrement, on obtient les graphiques des fonctions de répartition relatives à nos données. Chaque ligne correspond à une année de survenance, et chaque colonne correspond à une vision comptable.

Deuxièmement, on sélectionne les données qui nous intéressent. Dans la partie 3.1 de cette étude on ne s'intéresse qu'à la vision en 2017 des sinistres survenus avant 2013, ces données sont illustrées dans le rectangle rouge de la figure 3.2 ci-dessous. La première raison est que la responsabilité civile automobile est une branche longue, il semble donc nécessaire de ne pas prendre les années récentes afin de pouvoir faire l'hypothèse que l'estimation des montants ultimes est robuste. La seconde raison est qu'il y a eu une réforme majeure en assurance automobile avec prise d'effet le 1er janvier 2013. A partir de cette date le Fond de Garantie des Assurances Obligatoires (FGAO) ne prend plus en charge la revalorisation des rentes pour les accidents automobiles. On souhaite en effet avoir des données les plus homogènes possible et donc se placer dans un environnement juridique le moins changeant possible.

FIGURE 3.2 – Exemple de triangle de fonctions de répartition

Remarque : Ce triangle est à titre illustratif, il est issu de données fictives. Une illustration partielle du triangle obtenu à partir de nos données est disponible en annexe B.

Ensuite, on va s'intéresser à l'évolution des ultimes pour une année de survenance n à une année de survenance n+1, pour n allant de 2002 à 2011. Concrètement, on regarde les 11 cases à l'intérieur du rectangle rouge ci-dessus, et on se demande comment passer d'une case à celle immédiatement dessous.

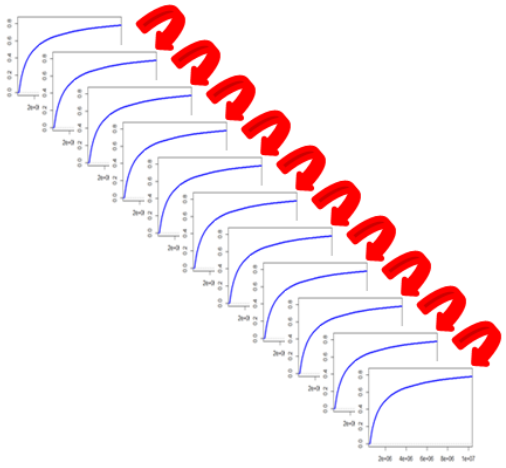


FIGURE 3.3 – Évolution des visions en 2017 des fonctions de répartition des montants de sinistres survenus entre 2002 et 2012

Le but sous-jacent à cet exercice est de pouvoir répondre à la question suivante : supposons qu'un sinistre survenu une année n se reproduise l'année de cotation, combien ce sinistre coûterait-il de plus par rapport au sinistre original ?

Pour mesurer cette différence on va utiliser un changement de variable. On projette la distribution des ultimes d'une année de survenance donnée et on la compare avec la distribution des ultimes de l'année de survenance suivante.

Remarque : on fait un abus de langage puisque la vision dite "ultime" ici correspond en réalité à la vision en 2017.

On note :

- $X_{i,j}$ l'ensemble des sinistres survenus l'année i, vu l'année comptable j.

- $\hat{\tau}_i$ l'estimateur du taux d'inflation que l'on suppose pour l'année i.

La méthode consiste d'abord à tracer sur un même graphe la fonction de répartition de l'échantillon $(1 + \hat{\tau}_i) * X_{i,2017}$ et la fonction de répartition de l'échantillon $X_{i+1,2017}$. Puis on mesure l'adéquation avec des diagrammes quantile-quantile. On répète le processsus pour j = 2002, ..., 2011.

3.1.3 Résultats

Pour plus de clarté, on illustre ici uniquement le cas où i = 2002. Le lecteur peut toutefois voir l'ensemble des calculs en annexe C.

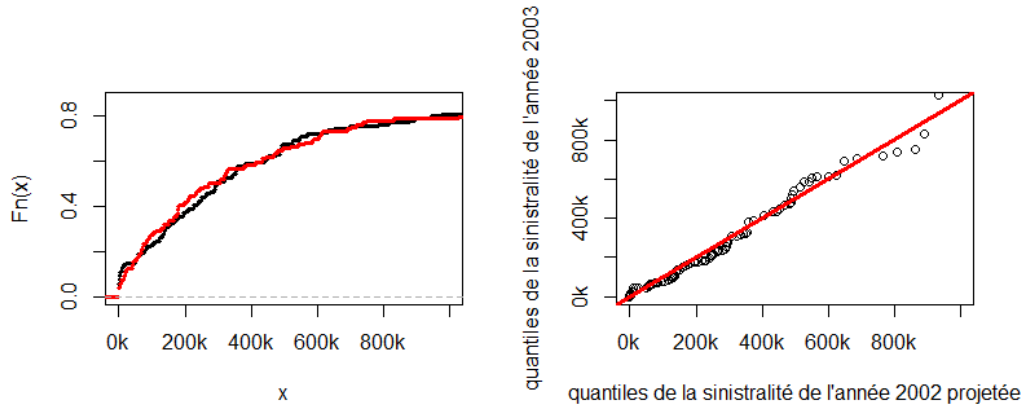


FIGURE 3.4 – Comparaison des deux fonctions de répartitions ainsi obtenues et mesure de l'adéquation par QQ plot

On résume ci-dessous les valeurs des estimateurs des inflations annuelles obtenues :

Année	Estimateur de la valeur du taux d'inflation	test visuel QQ-Plot
2003	20%	accepté
2004	0%	accepté
2005	10%	accepté
2006	10%	accepté
2007	10%	rejetté
2008	10%	accepté
2009	10%	rejetté
2010	20%	accepté
2011	-5%	rejetté
2012	5%	accepté

TABLE 3.1 – Tableau des estimateurs de la valeur des taux d'inflation annuelle

La moyenne géométrique des estimateurs des taux d'inflation annuelle, dont le test par diagramme quantile-quantile a été accepté, est de 10,7 %. Ce résultat est à prendre avec du recul car la méthode est imprécise. Néanmoins, on remarque que ce chiffre n'est pas très éloigné de

ceux du corporel grave présentés dans le premier chapitre.

3.2 Limites de la méthode

Dans cette partie, inspirée de BRAZAUSKAS et al. [2009] **[3]**, on s'intéresse aux limites de la méthode proposée précédemment.

3.2.1 Quand l'inflation est invisible

Plusieurs challenges font face à un réassureur lorsqu'il souscrit des polices qui couvrent un risque au-delà d'un seuil. Le réassureur ne connaît rien des montants de sinistres qui sont au-dessous d'un certain pourcentage de la priorité, appelé seuil de communication de la cédante, il peut ainsi avoir des difficultés à modéliser le risque. Cela peut même le conduire à des observations paradoxales, c'est ce que nous allons illustrer dans cette partie.

On simule des montants de sinistres correspondants à 8 années successives, numérotées de 1 à j, à l'aide d'un modèle fréquence-coût. Pour la fréquence on utilise une loi de Poisson de paramètre 2000. Pour être proche des hypothèses de sévérité prises en tarification de traités responsabilité civile automobile, on utilise pour l'année j une loi de Pareto de paramètre de seuil θ =100 000 * 1.1^{j-1} et de paramètre de forme $\alpha = 2.5$. On suppose que la priorité du traité est fixée à 2 millions et que la couverture est illimitée.

On peut maintenant observer les diagrammes en boite de tous les montants simulés et les montants observés par le réassureur :

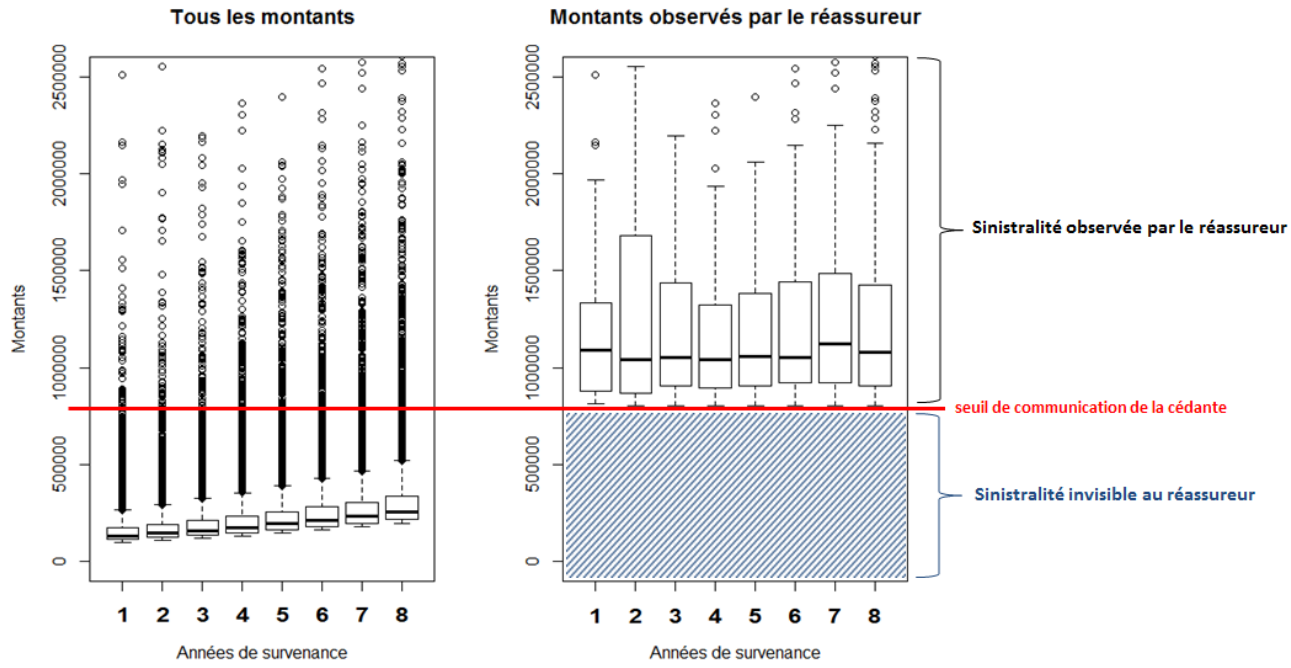


FIGURE 3.5 – À gauche, diagrammes en boite de tous les montants et à droite, diagrammes en boite des montants observés par le réassureur

Ces résultats sont peu intuitifs. Bien qu'il y ait sans aucun doute possible une inflation des montants de sinistres, bien visible dans les huit diagrammes de la partie gauche, elle semble totalement invisible au réassureur sur les huit diagrammes de droite. Nous allons expliquer ce phénomène dans la partie suivante.

3.2.2 Explication

Mathématiquement, si le réassureur avait accès à toutes les données alors il aurait des échantillons des réalisations des variables aléatoires suivantes :

$$\left\{X_{j,1}, ..., X_{j,N_j}\right\} \sim \text{Pareto } (\theta_j, \alpha), \text{ j = 1,..., J}$$

où :

- J représente le nombre d'années d'historique à la disposition du réassureur

- N_j le nombre de sinistres l'année j.

Or, il ne dispose que des sinistres dont le montant est supérieur au seuil d de communication de la cédante :

$$\left\{Y_{j,1}, ..., Y_{j,M_j}\right\} \sim \text{Pareto (d, } \alpha), \text{ j = 1,..., J}$$

où :

- $Y_j = X_j \mid X_j > d$

- M_j est le nombre de sinistres dont le montant est supérieur au seuil de communication l'année j.

La loi que suivent les variables aléatoires $\left\{Y_{j,1}, ..., Y_{j,M_j}\right\}$, pour j allant de 1 à J, est bien une loi de Pareto de même paramètre de forme α. C'est la propriété de stabilité par seuil de la loi de Pareto, et plus généralement des lois *Generalized Pareto Distribution* aussi appelées loi GPD. En revanche le paramètre de seuil change, il est maintenant égal à d le montant à partir duquel la cédante communique ses sinistres.

Par conséquent, on a :

$$\begin{aligned} P[Y_j > x] &= P[X_j > x \mid X_j > d] \\ &= \frac{P[X_j > x]}{P[X_j > d]} \\ &= (\frac{d}{x})^{\alpha} \end{aligned}$$

Cette distribution ne dépend donc pas de j, ce qui est une particularité de la loi de Pareto et la raison pour laquelle l'inflation semble invisible au réassureur par la méthode du diagramme en boite.

3.2.3 Illustration

On vient de voir que théoriquement la méthode vue partie 3.1.2 ne devrait pas permettre de voir une inflation. Vérifions-le avec un test sur les données fictives de la partie précédente.

Pour plus de clarté, on illustre ici uniquement le dernier cas. Le lecteur peut toutefois voir l'ensemble des calculs en annexe D.

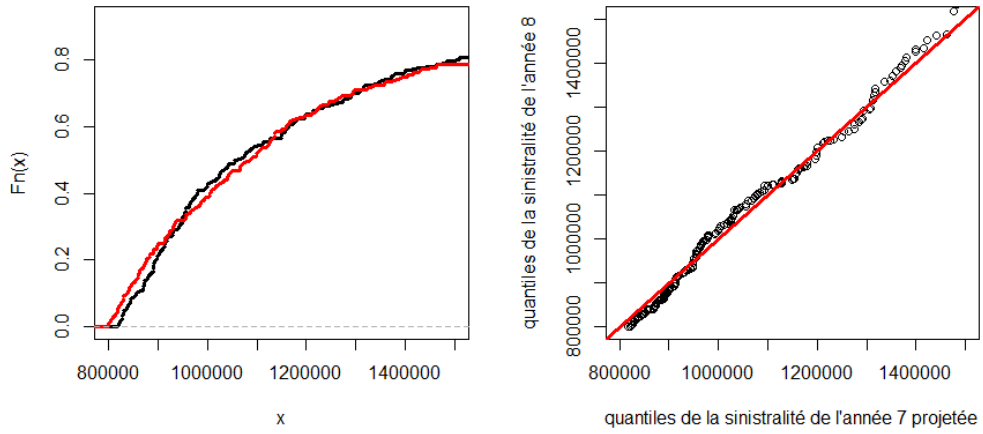


FIGURE 3.6 – Comparaison des deux fonctions de répartitions ainsi obtenues et mesure de l'adéquation par QQ plot

On observe les résultats pour l'inflation annuelle ci-dessous :

Année	Estimateur de la valeur du taux d'inflation	test visuel QQ-Plot
1	-8%	accepté
2	5%	rejetté
3	5%	accepté
4	0%	rejetté
5	2%	accepté
6	-2%	accepté
7	2%	accepté

TABLE 3.2 – Tableau des taux d'inflation annuels supposés

La moyenne géométrique des taux d'inflation annuels, dont les tests par diagramme quantile-quantile ont été acceptés, est de -0,2 %. La méthode basée sur les fonctions de répartition donne donc une estimation de la valeur de l'inflation presque nulle lorsque la sinistralité suit une loi de Pareto, alors que l'inflation annuelle sur ces données simulées est de 10%. On a donc bien confirmé que la méthode vue partie 3.1.2 ne marche donc pas dans ce cas précis. Dans la suite, on va reprendre la méthode proposée par BRAZAUSKAS et al. [2009] [**3**] pour proposer une solution.

3.3 Un second essai par une méthode de comptage

3.3.1 Description de la méthode

On va chercher dans cette partie à retrouver le taux d'inflation de 10% utilisé pour la simulation de nos données de la partie 3.2., en supposant que l'on a à disposition uniquement les montants au-dessus d'un certain seuil comme illustré figure 3.5. On se met donc à la place d'un réassureur qui, exclusivement à partir des données que la cédante lui donne, essaie de déterminer un taux annuel d'inflation.

Mathématiquement, on cherche à estimer r_j tel que :

$$\begin{aligned}\frac{E[X_j]}{E[X_{j-1}]} &= \frac{\theta_j}{\theta_{j-1}} \\ &= 1 + r_j\end{aligned}$$

Pour cela on remarque que bien que l'inflation n'ait pas d'impact sur le coût moyen observé, comme vu précédemment figure 3.5, elle a un impact sur le nombre de sinistres observables par le réassureur :

	Toutes les données		Données à disposition du réassureur	
Année	Coût moyen	Nombre de sinistres total	Coût moyen observé	Nombre de sinistres observés
1	167 908	2 029	1 187 234	12
2	182 980	1 966	1 151 980	13
3	203 050	2 055	1 445 912	17
4	218 365	1 921	1 372 130	12
5	247 254	1 979	1 125 018	33
6	272 525	1 964	1 409 518	39
7	394 176	2 001	1 324 633	59
8	324 140	2 021	1 361 460	52
	Inflation simulée de 10% par an	Fréquence simulée stable	L'inflation est invisible, ce qu'on a étudié partie 3.2	L'inflation est visible, ce qu'on va étudier partie 3.3

FIGURE 3.7 – Sommaire des données simulées

Explication de la figure ci-dessus :

On a simulé partie 3.2 des données au-delà d'un seuil sur 8 années avec une fréquence stable mais avec une augmentation de la sévérité de 10 % par an. Et ce qui à première vue semble paradoxal, c'est que le réassureur voit l'inverse. En effet, il voit une sinistralité moyenne stable mais une augmentation de la fréquence.

C'est pour cette raison que l'on a supposé dans la partie 3.2 que la méthode proposée partie 3.1.2, qui utilise des triangles de fonctions de répartition et des changements de variables, ne peut pas fonctionner dans le cas d'une sinistralité qui suit une loi de Pareto. On va donc dans cette partie 3.3 proposer une autre méthode pour palier à ce problème, basée sur le nombre de sinistres observés par le réassureur.

Cette nouvelle méthode repose sur deux étapes que l'on va bientôt détailler. La première est la détermination d'un estimateur du paramètre de forme $\hat{\alpha}$ de la loi de sévérité. Pour cela on va utiliser les $Y_{j,k}$, c'est à dire la 4e colonne du tableau de la figure 3.8. Une fois cet estimateur déterminé, nous allons utiliser les M_j, c'est-à-dire la 5e colonne du tableau de la figure 3.8, pour proposer un estimateur du taux d'inflation utilisé.

Formellement, M_j est la variable aléatoire qui correspond à la fréquence de sinistres au-dessus du seuil de communication de la cédante l'année j.

$$M_j = \sum_{k=1}^{N_j} 1_{X_{j,k}>d}$$

où :

- Chaque $X_{j,k}$ correspond au montant du k-ième sinistre de l'année j.

- Chaque N_j est la variable aléatoire qui correspond à la fréquence de sinistres de la cédante l'année j.

- d est le seuil de communication de la cédante, toujours inférieur à la priorité du traité.

On a supposé n'avoir que les montants auxquels aurait accès le réassureur. On suppose alors que N_1,..., N_J sont des variables aléatoires indépendantes qui suivent une loi de Poisson. Pour chaque j, le paramètre de la loi de Poisson correspondant est λ_j, tel que $\lambda_j = \lambda e_j$. Le deuxième facteur e_j est une mesure de l'exposition, proportionnel aux engagements de l'assureur l'année j et le premier facteur λ est le taux de sinistre par unité d'exposition.

On s'intéresse maintenant à la distribution des M_j. On admet que si les N_j suivent une loi de Poisson alors les M_j aussi. Pour démontrer cela il y a plusieurs façons de faire, l'une d'entre elles passe par l'utilisation des fonctions génératrices des moments qui caractérisent la loi.

Ainsi chaque M_j, j = 1,..., J, suit une loi de Poisson de paramètre $\lambda\ e_j\ (\frac{\theta_j}{d})^{\alpha}$, et on a :

$$E[M_j] = \lambda\ e_j\ (\tfrac{\theta_j}{d})^{\alpha}$$

Ainsi, en passant au log on obtient :

$$\begin{aligned} log E[M_j] &= log(\lambda) + log(e_j) - \alpha\ log(d) + \alpha\ log(\theta_j) \\ &= log(\lambda) + log(e_j) - \alpha\ log(d) + \alpha\ log(\theta) + \{\alpha\ log(1+r)\}\,(j-1) \end{aligned}$$

donc :

$$\log(\frac{E[M_j]}{e_j}) = log(\lambda) - \alpha\ log(d) + \alpha\ log(\theta) + \{\alpha\ log(1+r)\}\,(j-1)$$

La partie droite de cette équation est une fonction linéaire en j de pente $\alpha\ log(1+r)$. On va donc dans un premier temps estimer α, on rappelle que c'était notre première étape décrite plus haut. Puis dans un second temps, on peut faire une régression linéaire sur l'échantillon de points (j, $\log(m_j/e_j)$), j = 1,..., J, et utiliser le coefficient directeur comme estimateur de l'inflation r.

Avec la méthode du maximum de vraisemblance pour l'estimation de α puis la méthode des moindres carrés pour estimer r. On obtient les estimateurs suivants :

$$\hat{\alpha} = \frac{\sum_{j=1}^{J} m_j}{\sum_{j=1}^{J}\sum_{k=1}^{m_j} log(x_{j,k}/d)}$$

$$\hat{r} = \exp\left\{\frac{12\sum_{j=1}^{J} j\ log(m_j/e_j) - 6(J+1)\sum_{j=1}^{J} log(m_j/e_j)}{\hat{\alpha}J(J^2-1)}\right\}$$

Le lecteur intéressé par plus de détails sur les calculs peut se reporter à BRAZAUSKAS et al. [2009] [**3**].

3.3.2 Résultats

L'application numérique de la méthode donne les résultats suivants :

seuil	$\hat{\alpha}$	écart avec α	$\hat{r}$	écart avec r
200 000	2,4934	0%	0,0982	2%
300 000	2,4846	1%	0,0949	5%
400 000	2,5654	3%	0,0973	3%
500 000	2,6022	4%	0,1054	5%
600 000	2,3929	4%	0,0904	10%
700 000	2,6168	5%	0,0813	19%
800 000	2,3692	5%	0,1193	19%
900 000	2,7224	9%	0,0749	25%
1 000 000	2,2008	12%	0,1395	40%
1 500 000	1,8718	25%	0,1501	50%

TABLE 3.3 – Application numérique pour différents seuils

L'analyse du tableau ci-dessus nous permet de conclure que cette méthode semble potentiellement très imprécise dans le cas où le seuil est très haut. Néanmoins, dans le cas où la cédante communique ses sinistres à partir d'un seuil relativement bas, elle semble bien permettre d'avoir une vision de l'inflation quand celle-ci semble invisible comme dans la figure 3.5.

Pour élargir l'analyse de cette partie 3.3, qui repose sur la simulation d'une sinistralité au delà d'un seuil via un modèle Poisson-Pareto, on remarque l'importance pour le réassureur d'avoir des cédantes qui communiquent leurs sinistres à partir d'un montant le plus bas possible. En effet outre la connaissance que cela apporte sur les sinistres extrêmes de moindres montants, la

connaissance des "petits sinistres extrêmes" semble permettre de mieux estimer l'inflation des "gros sinistres extrêmes" comme on le voit sur la table 3.3.

Pour résumé ce chapitre, dont l'objectif était d'illustrer le risque d'estimation des paramètres de la méthode mixte via l'étude d'un estimateur non classique du taux annuel de l'inflation du passé ; on a premièrement mis en application, sur les données d'une cédante du marché français, une méthode basée sur des triangles de fonctions de répartition et sur des changements de variables. Ensuite, on a pointé ce qui apparaît comme une des limites de cette méthode, en étudiant un cas où elle ne fonctionne apparemment pas, car sous un certain point de vue l'inflation semble invisible. Enfin, nous avons proposé une solution dans le cadre de la problématique de la partie 3.2, via une méthode de comptage des sinistres au-delà d'un seuil.

Chapitre 4

Une autre façon de déterminer les facteurs de développement

On souhaite s'inspirer de la méthode précédemment vue partie 3.1.2 pour proposer un estimateur, exclusivement avec les données de la cédante que l'on a décrit partie 3.1.1, des facteurs de développement issus de la méthode Chain-Ladder nécessaires à l'utilisation de la méthode mixte. On étudiera le cas du développement des sinistres au premier euro, mais également le développement de charge de réassurance. On proposera donc dans un premier temps une méthode d'obtention de la vision du développement des sinistres au premier euro. Puis dans un second temps, on verra comment traduire cette vision de l'augmentation des prix des sinistres au premier euro en vision de l'augmentation de la charge de la réassurance, grâce à la théorie autour du concept d'accélération.

4.1 Première étape : estimer le développement des sinistres au premier euro

4.1.1 Description de la méthode

On s'inspire de la méthode vue précédemment pour trouver un estimateur des facteurs de développement Chain-Ladder de sinistres au premier euro. Ces facteurs reflètent notamment l'évolution de l'état de santé de la victime et donc l'évolution du coût du sinistre au cours du temps. On note que pour plus de simplicité, on ne s'intéresse qu'à la détermination des 3 premiers facteurs de développement de Chain-Ladder.

Illustrons directement notre méthode avec l'exemple de la détermination du premier facteur de développement de Chain-Ladder. Pour déterminer le premier facteur de développement de Chain-Ladder, on s'intéresse aux deux premières colonnes de notre triangle de fonctions de répartition illustré ci-dessous. Chaque ligne correspond à une année de survenance, et chaque colonne correspond à une année de développement.

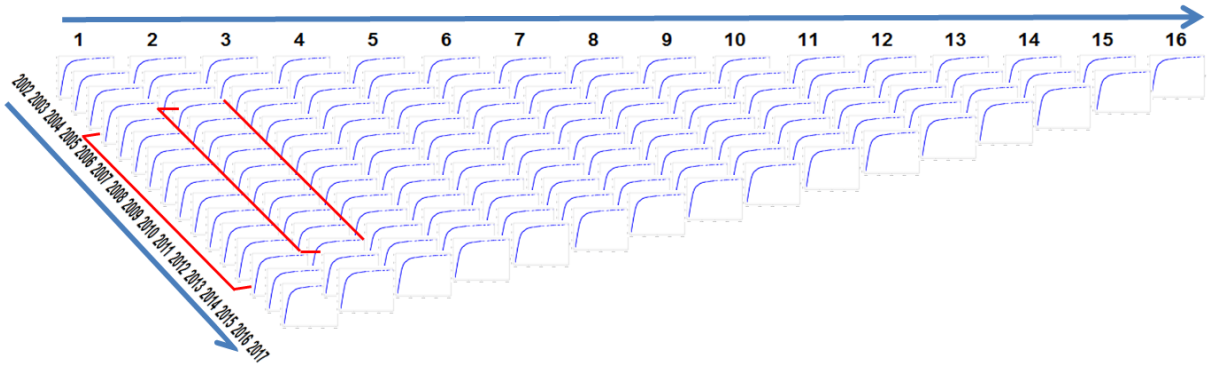


FIGURE 4.1 – Illustration d'un triangle de fonctions de répartition

Remarques :

- Ce triangle est à titre illustratif, il est issu de données simulées. Une illustration du triangle de fonctions de répartition utilisé dans notre étude est disponible en annexe B.

- On ne s'intéresse qu'aux années de survenance de la période 2005-2014. En effet on ne peut pas prendre les années postérieures à 2014 car on n'a pas suffisamment de vision pour déterminer les 3 premiers facteurs de développement. On ne prend pas non plus les années de 2002 à 2004 car on fait l'hypothèse que 10 années d'historique suffisent à notre étude.

Ensuite, on trace les deux fonctions de répartition empiriques des sinistres survenus sur la période 2005-2014 vus en année 1, puis vus en année 2.

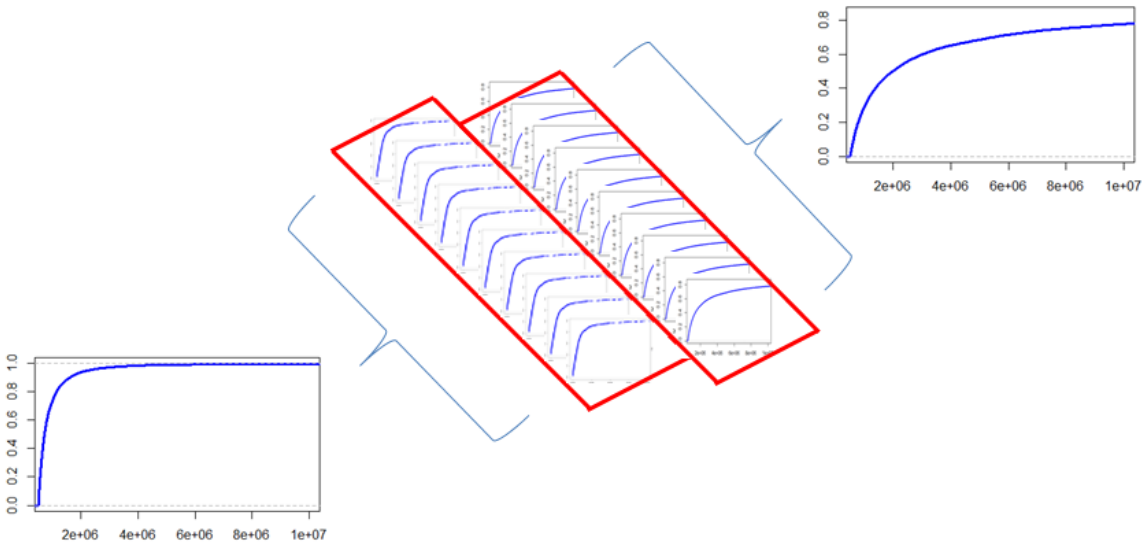


FIGURE 4.2 – Fonctions de répartition empiriques des sinistres survenus sur la période 2005-2014 vus en année 1, puis vus en année 2

On s'interroge alors sur la différence entre ces deux fonctions de répartition.

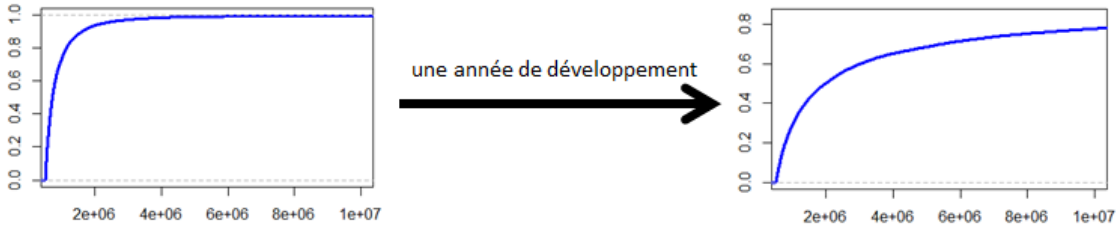


FIGURE 4.3 – Impact du développement des sinistres entre la première et la deuxième année sur la fonction de répartition issue des sinistres survenus sur la période 2005-2014

Finalement, on va utiliser la méthode de changement de variable, décrit partie 3.1.2, pour déterminer notre premier facteur de développement :

Soit $X_{i,j}$ l'ensemble des sinistres survenus l'année i, vu après j années de développement.

soit $\hat{f}_j$ l'estimateur de la valeur du facteur de développement entre la j-ième année et la j+1-ième année.

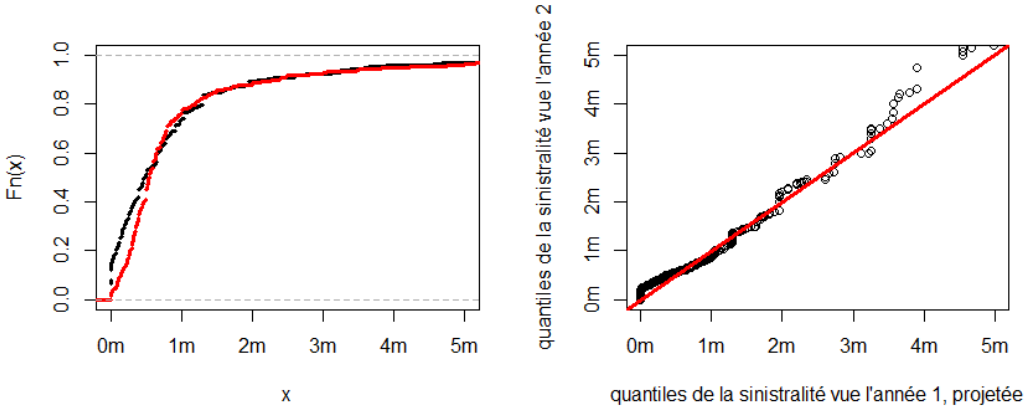
La méthode consiste à :

- Tracer sur un même graphe la fonction de répartition de l'échantillon $\cup_{i=2005}^{2014} X_{i,j}$ auquel on a appliqué le facteur multiplicatif $(1+\hat{f}_j)$ et la fonction de répartition de l'échantillon $\cup_{i=2005}^{2014} X_{i,j+1}$.

- Mesurer l'adéquation avec des diagrammes quantile-quantile.

- Répéter le processus pour j = 1,2,3.

4.1.2 Résultats



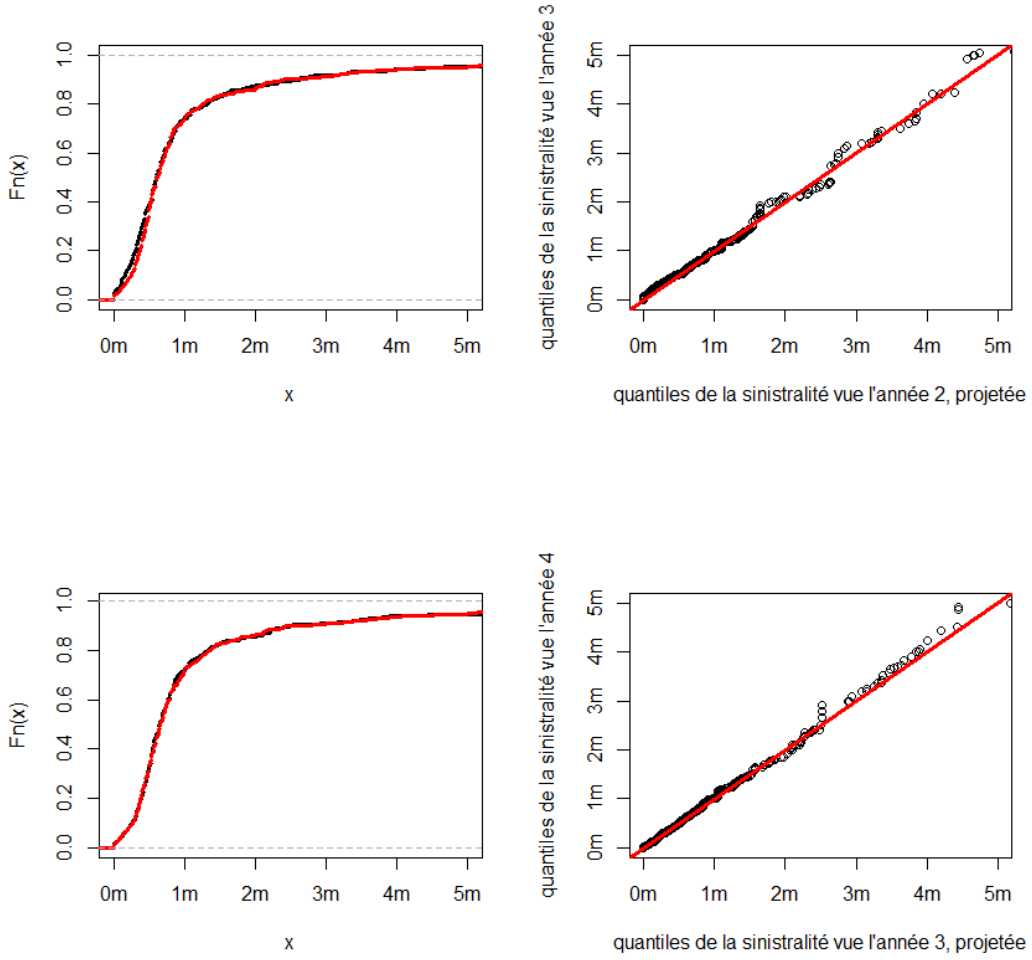


FIGURE 4.4 – Projections de fonctions de répartition par changement de variable et mesure de l'adéquation par QQ-plot

On résume ci-dessous les estimateurs des valeurs des facteurs de développement obtenus :

Années	Facteur de développement
1 à 2	1,20
2 à 3	1,05
3 à 4	1,02

TABLE 4.1 – Facteurs de développement au premier euro

Une question se pose alors : comment peut-on traduire cette inflation au premier euro en inflation à la charge du Traité XS ?

4.2 Deuxième étape : en déduire le développement des charges de réassurance

4.2.1 Le concept d'accélération

- Premier exemple

On considère les hypothèses suivantes :

- Priorité = 1 000 000 €
- Portée = 3 000 000 €
- Sinistre au premier euro = 2 500 000 €
- Dégradation du sinistre = 10 %

Le sinistre à charge de réassurance est de :

Sinistre au premier euro - Priorité = 1 500 000 €

Le sinistre dégradé à charge de réassurance est de :

Sinistre au premier euro . (1 + dégradation du sinistre) - Priorité = 1 750 000 €

Ainsi, pour une dégradation au premier euro de 10 % on constate une dégradation de la charge de réassurance de :

(1 750 000 / 1 500 000) - 1 = **16,7** %

Pour une inflation de 10% au premier euro on a donc une inflation de la charge de réassurance 1,67 fois plus importante.

- Deuxième exemple

On considère les hypothèses suivantes :

- Priorité = 1 000 000 €
- Portée = 3 000 000 €
- Sinistre au premier euro = 1 200 000 €
- Dégradation du sinistre = 10 %

Le sinistre à charge de réassurance est de :

Sinistre au premier euro - Priorité = 200 000 €

Le sinistre dégradé à charge de réassurance est de :

Sinistre au premier euro . (1 + dégradation du sinistre) - Priorité = 320 000 €

Ainsi, pour une dégradation au premier euro de 10 % on constate une dégradation de la charge de réassurance de :

(320 000 / 200 000) - 1 = **60** %

Pour une inflation de 10% au premier euro on a donc une inflation de la charge de réassurance, non plus 1.67 fois plus importante comme dans l'exemple précédent, mais 6 fois plus importante.

On observe donc qu'une dégradation d'un sinistre au premier euro provoque une dérive relative du sinistre à charge de réassurance plus importante que celle vue sur le sinistre au premier euro mais très variable.

- Le concept d'accélération

On va maintenant s'intéresser plus en détails au rapport entre la dérive relative vue sur les sinistres à charge de réassurance et la dérive relative du sinistre au premier euro. On appellera ce rapport l'**accélération**. Par exemple l'accélération dans les deux exemples précédents était respectivement de 1,67 et de 6.

Remarque : les sinistres devenant à charge ont une accélération infinie, la démonstration est immédiate puisqu'avant inflation la charge de réassurance était de 0.

Hypothèse : On se place dans un cadre simplifié où la portée est illimitée et où il n'y a pas de clause de stabilité, ni de clause de franchise annuelle.

On peut alors définir l'accélération ainsi :

$$\begin{aligned}\text{accélération} &= \frac{\Delta\text{relative du Sinistre à charge}}{\Delta\text{relative du Sinistre au premier euro}} \\ &= \frac{\frac{\text{Charge du sinistre dégradé}}{\text{Charge initiale du sinistre}} - 1}{\frac{\text{Montant du sinistre dégradé au premier euro}}{\text{Montant initial du sinistre au premier euro}} - 1}\end{aligned}$$

En développant le calcul, on obtient rapidement :

$$\text{accélération} = \frac{\text{Montant initial du sinistre au premier euro}}{\text{Montant initial du sinistre au premier euro - Priorité}}$$

L'accélération a donc la propriété d'être indépendante de la dégradation.

On va maintenant introduire la notion de **ratio**, qui représente le rapport entre le sinistre au premier euro et la priorité. Ainsi par calcul on obtient une seconde formule pour l'accélération :

$$\text{accélération} = \frac{\text{ratio}}{\text{ratio}-1}$$

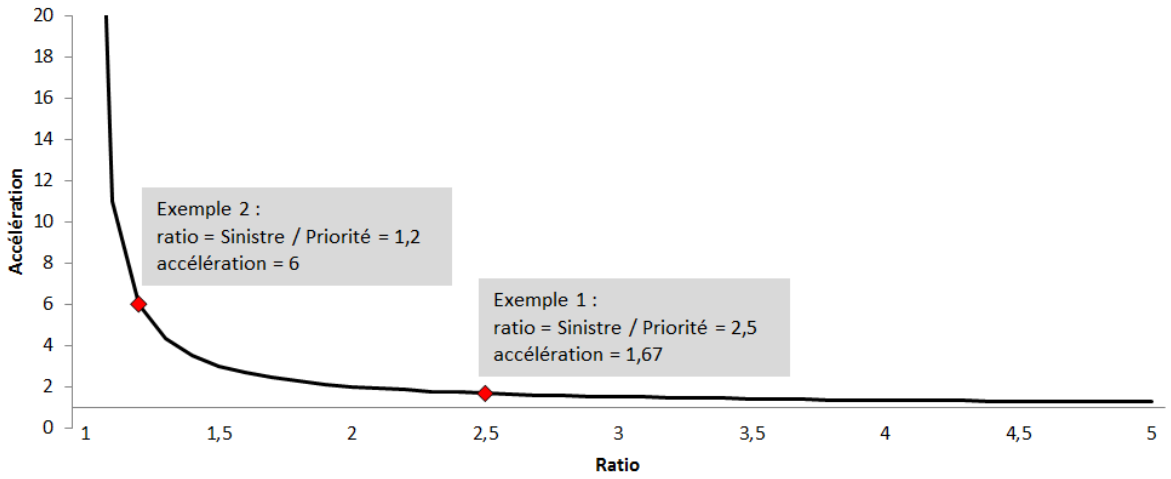


FIGURE 4.5 – L'accélération en fonction du ratio

Ce graphe nous permet de voir que plus le montant du sinistre est proche de la priorité, plus l'accélération est importante. Inversement, plus le montant du sinistre est éloigné de la priorité plus l'accélération se rapproche de 1.

- L'accélération globale et l'accélération moyenne

On propose maintenant d'étudier l'accélération globale à l'échelle de notre jeu de données. On commence par définir l'accélération globale et l'accélération moyenne :

L'accélération globale, que l'on note a_G, est égale au rapport entre la charge initiale brute de réassurance sur la charge initiale de réassurance.

L'accélération moyenne, que l'on note a_M, est l'accélération qui correspond au ratio moyen. On admet que l'accélération moyenne est égale à l'accélération globale.

Ainsi on peut calculer l'accélération globale de notre programme :

$$a_G = \frac{\sum_{i=1}^{N} S_i}{\sum_{i=1}^{N} S_i - N * \text{Priorité}}$$

On résume ci-dessous l'accélération globale de notre jeu de données, relativement à différents niveaux de priorité :

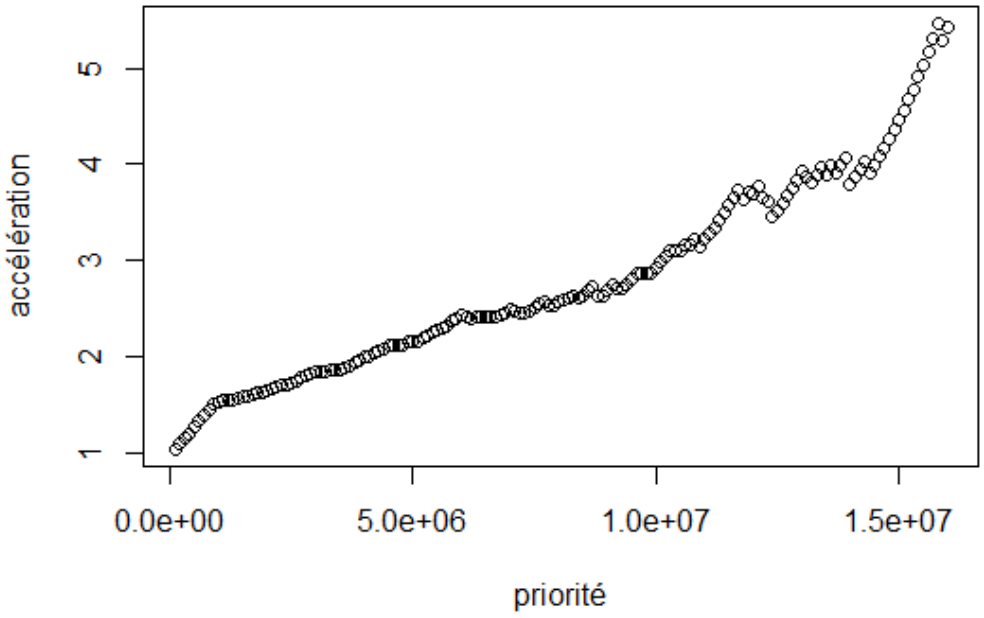


FIGURE 4.6 – Accélération selon différents niveaux de priorité

Ces résultats sont cohérents avec la théorie. On vérifie qu'à jeu de données constant, l'accélération augmente avec la priorité.

Remarque : Il pourrait être intéressant pour l'actuaire, en amont d'une cotation, d'évaluer l'accélération du programme. En effet, il aurait alors une vision globale de l'impact qu'aurait une inflation des sinistres au premier euro. Par exemple, comme cela pourrait-être le cas avec la possible réforme de la RC automobile.

- Mesures d'impact

On va maintenant illustrer le concept d'accélération précédemment défini. Pour cela, on étudie l'impact du taux d'indexation des sinistres sur la prime pure.

On souhaite illustrer l'importance pour le réassureur de bien comprendre l'inflation des sinistres au cours du temps, et donc de bien comprendre l'effet seuil. En effet, une hypothèse d'inflation annuelle constante de 10% conduit dans l'exemple visible ci-dessous de la figure 4.8 à une augmentation du prix de plus de 60% par rapport à une hypothèse d'inflation annuelle constante de 5%. En effet, avec un taux annuel d'inflation à 5% la prime pure est proche de 5,5% de l'assiette de prime alors qu'avec un taux annuel de 10% la prime pure est proche de 9% de l'assiette de prime.

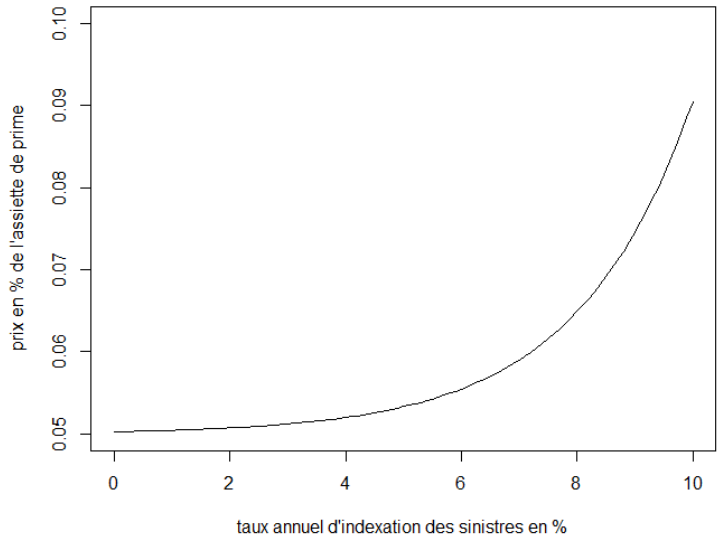


FIGURE 4.7 – Illustration de l'évolution de la prime pure en fonction du taux d'indexation des sinistres

Dans un second temps on s'intéresse à l'impact du couple (taux d'indexation des sinistres, priorité) :

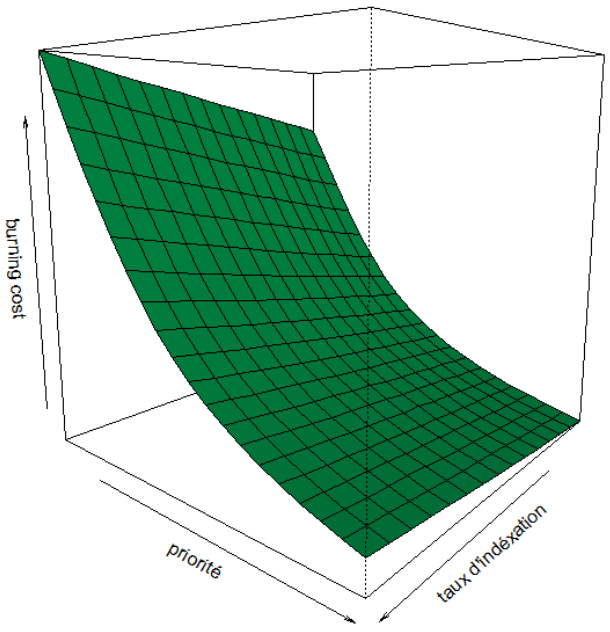


FIGURE 4.8 – Illustration de l'évolution de la prime pure en fonction de la priorité et du taux d'indexation

Ce graphique est difficile à interpréter. Néanmoins on a l'impression que plus la priorité augmente, moins l'impact de l'indexation semble important. Cela est peut-être dû au fait que plus la priorité est basse plus il y a de sinistres qui, une fois inflatés, touchent la priorité du traité.

Remarque : Le graphique ci-dessus demande la réalisation de 10 000 tarifications différentes d'un traité XS avec notamment des étapes qui peuvent être relativement complexes en branche longue telles que l'indexation, la liquidation et l'application de la clause de stabilité. Ces calculs ont été réalisés avec un outil R créé personnellement. Outre le fait de permettre la réalisation de tests de sensibilité en plusieurs dimensions comme l'illustre la figure ci-dessus, cet outil a pour vocation d'élargir les possibilités techniques de l'actuaire (boostrap, changement de méthode, mesure de la variance des estimateurs) grâce à un gain de flexibilité et de temps.

4.2.2 Application

Suite à l'obtention des résultats de la partie 4.1.2, une question se pose alors : comment peut-on traduire cette inflation au premier euro en inflation à la charge du Traité XS ?

Dans le cadre où la sinistralité ne suit pas une loi de Pareto, il peut être extrêmement difficile de répondre à cette question. Néanmoins grâce à la théorie de la partie 4.2.1, on va voir que si la sinistralité suit une loi de Pareto de paramètre de forme α alors une augmentation au premier euro de x % conduit à une inflation de la charge de réassurance de $(1 + x\%)^{\alpha}$. Autrement dit, l'accélération est égale au paramètre de forme de la loi de Pareto, vérifions cela :

On suppose que les sinistres à charge suivent une loi de Pareto Pareto (θ, α).

L'espérance du montant d'un sinistre est : E[S] = $\frac{\theta\alpha}{\theta-1}$

Ainsi, le ratio moyen est : $\frac{E[S]}{\alpha}$

Par un calcul rapide on peut montrer que : $a_M = a_G = \theta$

On vérifie donc bien que, dans le cas où la sinistralité suit une loi de Pareto, l'accélération globale est égale au paramètre de forme.

Dans la partie suivante, on vérifiera que l'on peut faire l'hypothèse, qu'au-delà d'un certain seuil, notre sinistralité suit une loi de Pareto de paramètre 1,56. Sous ces hypothèses on obtient alors les facteurs de développement de charges suivants :

Années	Facteur de développement
1 à 2	1,329
2 à 3	1,079
3 à 4	1,031

TABLE 4.2 – Facteurs de développement de charges au Traité XS

Les estimateurs des facteurs de développement ainsi obtenus sont à prendre avec du recul. En

effet ces estimateurs, tout comme ceux fournis par la méthode de Chain-Ladder, ont une variance qui peut être importante. On précise que l'étude de la variance de nos estimateurs n'entre pas dans le cadre de ce mémoire. Néanmoins, une courte présentation de deux méthodes classiques de mesure de la variance d'un estimateur dans le cadre du provisionnement sont disponibles en annexe E.

On précise que l'objet de ce document n'est pas de prouver la validité théorique de cette façon de faire, mais d'illustrer le risque d'estimation des paramètres en proposant une autre méthode de détermination des 3 premiers facteurs de Chain-Ladder qui utilise les données de la cédante et non des données marché.

Pour résumer ce chapitre, on a fait l'hypothèse que l'information nécessaire sur les 3 premiers facteurs de développement est déjà contenue dans les données cédantes. Puis, on a proposé une méthode de détermination des facteurs de développement basée sur les montants de sinistres au premier euro, non agrégés de la cédante. Grâce à la théorie autour du concept d'accélération, on a pu en déduire des estimateurs des facteurs de développement de la charge du réassureur.

Troisième partie

Risque de spécification du modèle de tarification

Préambule :

Dans la partie précédente nous avons étudié des méthodes d'estimation de la valeur des paramètres utilisés pour l'indexation et la liquidation des montants de sinistres. Dans cette partie nous souhaitons aller plus loin, en étudiant non plus le risque d'estimation des paramètres d'un modèle défini, mais en challengeant le modèle lui-même. Plus particulièrement, nous souhaitons étudier l'adéquation de la loi de Pareto pour modéliser la sévérité dans le processus de tarification. Cela va nous conduire à proposer d'autres lois et donc à remettre en question la méthode mixte, aussi appelée méthode "extrapolation Pareto", dont l'utilisation classique repose sur cette loi. On va ainsi étudier l'impact sur la prime pure de tarifer avec un modèle fréquence-coût plutôt qu'avec la méthode mixte.

Nous allons principalement nous intéresser à la modélisation de la sévérité et non de la fréquence. On rappelle également que nous étudions la tarification des traités en excédent de sinistres (XS) responsabilité civile automobile (RCA), où les couvertures offertes par les réassureurs sont potentiellement illimitées. Pour étudier la sévérité nous allons donc avoir besoin de poser les bases de la théorie des valeurs extrêmes.

Chapitre 5

Étude de la sévérité

En premier lieu la question suivante se pose : pourquoi remettre en question la méthode "mixte", ou méthode appelée "extrapolation Pareto", définie partie 2.2.3 ?

Une des réponses possibles est que son utilisation classique impose le choix de la loi de Pareto pour modéliser la sévérité, ce qui peut représenter parfois une contrainte, comme l'illustre l'exemple suivant :

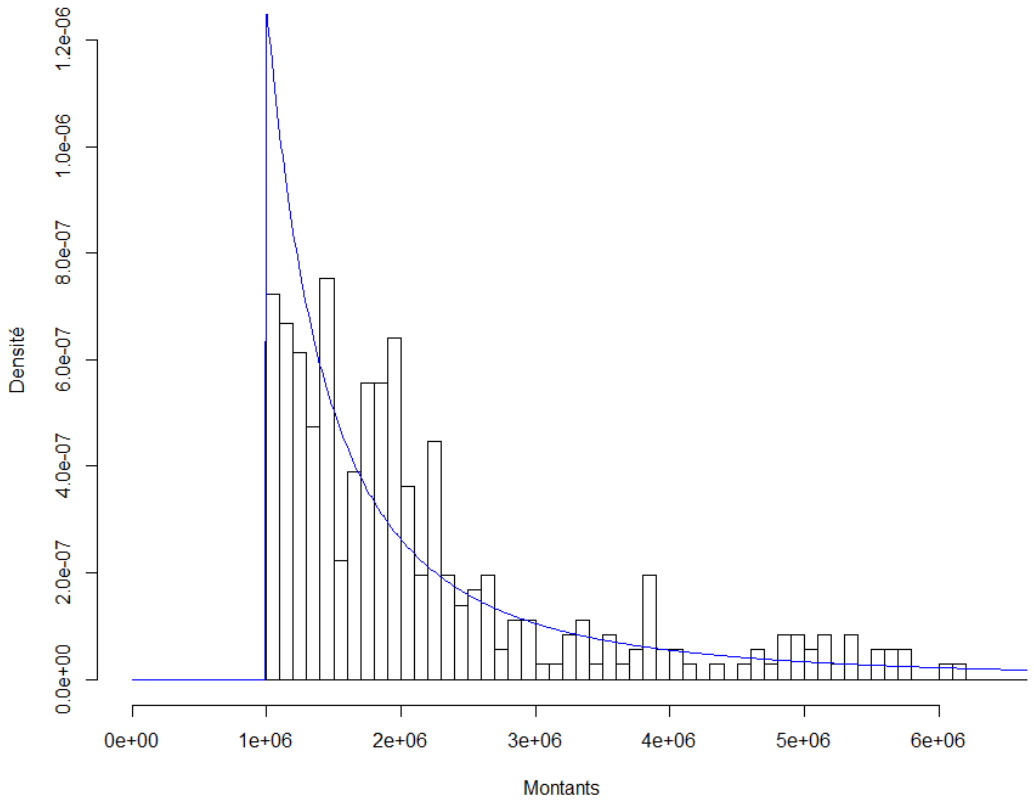


FIGURE 5.1 – Illustration du possible manque de flexibilité de la loi de Pareto pour modéliser la sinistralité

Cet exemple est typique des challenges que rencontre l'actuaire tarificateur. Bien souvent en effet,

la loi de Pareto semble permettre un très bon calibrage sur la queue de distribution mais pas sur la partie gauche de la distribution. Le cas inverse est également possible, où le calibrage semble bon sur la partie gauche mais pas sur la partie droite. Bien que l'impact de cette incertitude soit souvent minimisé en pratique par l'utilisation d'une courbe de marché au-delà d'un certain seuil, cela reste néanmoins un problème qui semble pouvoir être résolu avec une approche plus fine. Par exemple via l'utilisation d'autres lois plus flexibles telle que la loi Exponentielle-Pareto que nous étudierons par la suite, comme le montre la figure 5.2 :

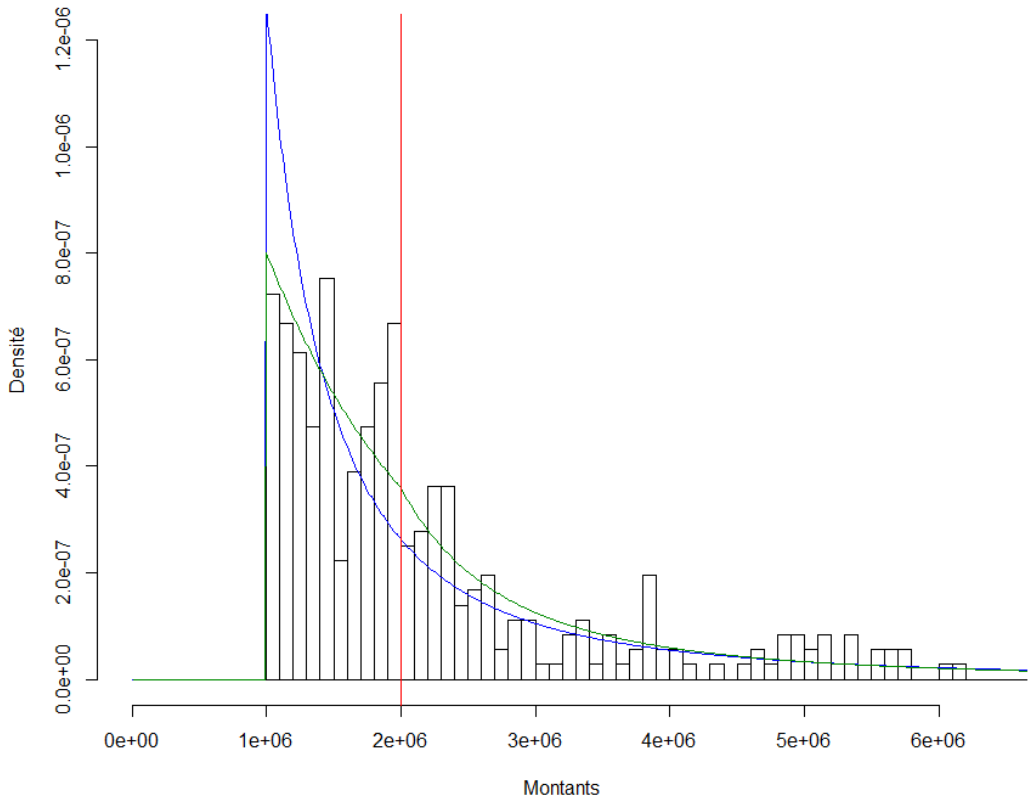


FIGURE 5.2 – Illustration du possible gain de flexibilité apportée par la loi Exponentielle-Pareto à 4 paramètres

Pour des montants inférieurs à 2 millions, la sinistralité semble être correctement modélisée par une loi exponentielle, alors que pour la queue de distribution la loi de Pareto semble mieux adaptée. L'ajout de deux paramètres semble donc permettre à la loi Exponentielle-Pareto d'être plus flexible que la distribution de Pareto car elle semble mieux correspondre aux données.

On peut toutefois s'interroger si le gain gagné en précision n'est pas au détriment de la simplicité. En effet, l'utilisation d'une autre loi pour modéliser la sévérité implique de renoncer à la méthode mixte, ou méthode dite "extrapolation Pareto", telle que présentée partie 2.2.3.

5.1 Théorie des valeurs extrêmes

5.1.1 Éléments de théorie

Pour pouvoir utiliser la méthode mixte il est nécessaire de se demander tout d'abord si la sévérité des sinistres suit une loi de Pareto. Pour le savoir il faut d'abord se poser une question plus générale : quel est le type de loi avec lequel nous pouvons modéliser notre sinistralité extrême ? Et ainsi se demander : comment définir la sinistralité extrême ? La théorie des valeurs extrêmes permet de donner des réponses à ces questions grâce à diverses méthodes graphiques d'estimation de seuil et grâce à la théorie autour des domaines d'attraction.

Parmi les méthodes graphiques d'estimation du seuil de sinistralité extrême les plus classiques, nous pouvons citer la *mean excess function* et le *Hill-plot*.

Nous ne détaillerons pas la théorie sous-jacente. Pour plus de détails le lecteur peut se diriger vers DEPREY et GODZINSKI [2007] [**7**].

Pour déterminer le domaine d'attraction on peut utiliser la théorie sur l'analyse des extrêmes univariés dans un cadre statique. À ce titre on rappelle quelques définitions et théorèmes.

Définition :

Une fonction L mesurable et positive sur $]0, \infty[$ est à *variations lentes*, si, pour tout x > 0, $\lim_{t\to\infty}$ L(tx) / L(t) = 1.

Théorèmes :

- Une distribution F appartient au domaine d'attraction de la loi de Fréchet Φ_α ($\alpha > 0$), si et seulement si, S(x) = $x^{-\alpha}$ L(x), pour une fonction à variations lentes L.

- Une distribution F appartient au domaine d'attraction de la loi de Gumbel Λ, si et seulement si il existe une fonction positive g telle que :

$$\lim_{x\to x^F} \frac{S(x+tg(x))}{S(x)} = e^{-t},\ t \in \mathbb{R}$$

- Une distribution F appartient au domaine d'attraction de la loi de Weibull Ψ_α, si et seulement si $x^F < \infty$ et $S(x^F - x^{-1}) = x^{-\alpha}L(x)$ pour une fonction à variations lentes L.

Remarque : pour déterminer le domaine d'attraction on peut également utiliser des méthodes visuelles à l'aide des diagrammes quantile-quantile.

5.1.2 Application

L'étude de la *mean excess function* vue précédemment ne fait pas apparaître clairement une droite. La difficulté à appliquer la théorie est en effet une problématique majeure en réassurance où on a parfois très peu de données. C'est pourquoi on va toutefois supposer que la sinistralité au-delà de 1,5 millions d'euros semble avoir un même comportement. On remarque que l'on

pourrait aussi choisir visuellement le seuil de 3,5 millions d'euros, malheureusement il faudrait renoncer à utiliser une bonne partie des données à notre disposition. On verra toutefois par la suite avec la loi Exponentielle Pareto qu'il est possible de choisir un seuil de 3,5 millions d'euros tout en utilisant quand même la sinistralité entre 1,5 millions d'euros et 3,5 millions d'euros.

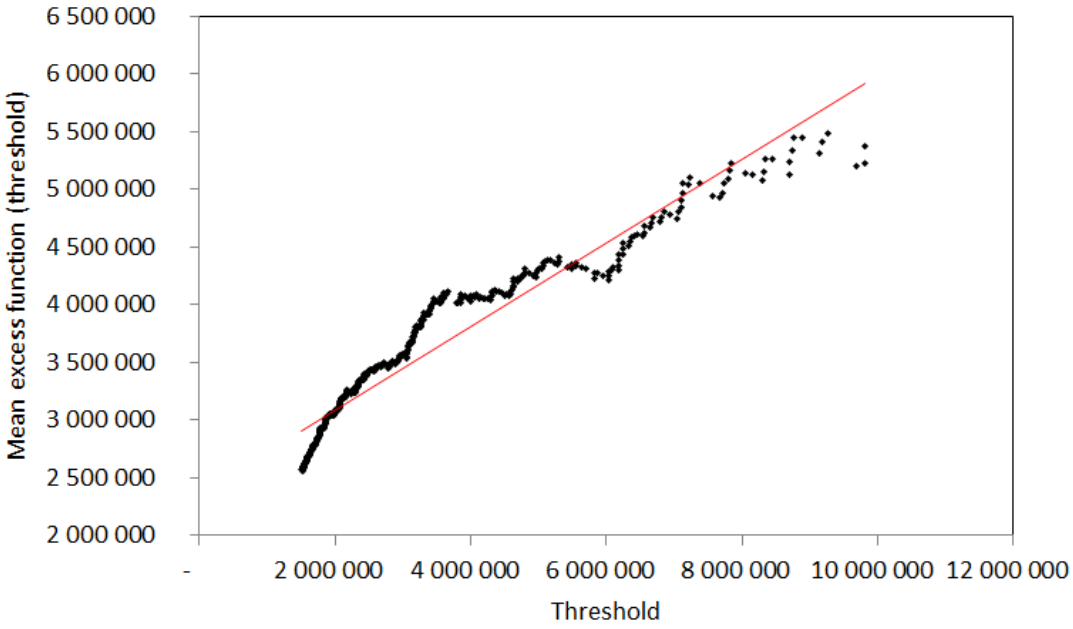


FIGURE 5.3 – Mean excess function appliquée à notre sinistralité

Pour faire le lien avec la partie précédente, on remarque que cela illustre une nouvelle fois la contrainte que peut représenter le fait d'utiliser la loi de Pareto pour modéliser la sévérité. En effet avec des données simulées avec une loi de Pareto, la *mean excess function* fait bien apparaître une droite :

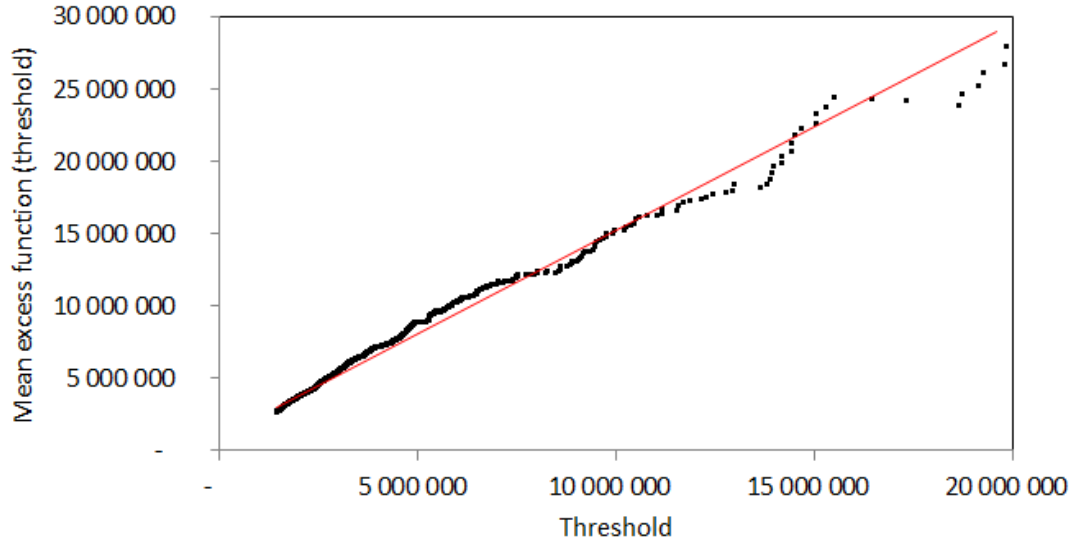


FIGURE 5.4 – Mean excess function appliquée à un échantillon simulé avec une loi de Pareto

On va toutefois supposer que le graphe fait apparaître une droite à partir de 1,5 millions et ainsi supposer premièrement que notre sinistralité extrême commence à partir de ce seuil et deuxièmement qu'elle appartient au domaine d'attraction de Fréchet, auquel appartient la loi de Pareto.

On va donc maintenant étudier plusieurs lois appartenant à ce domaine d'attraction. En effet, après avoir défini notre sinistralité extrême et supposé que les lois qui nous intéressent sont celles appartenant au domaine d'attraction de la loi de Fréchet, il faut encore choisir quelle loi sélectionner au sein de cette famille.

5.2 Étude de lois appartenant au domaine d'attraction de Fréchet

On va donc maintenant présenter quelques distributions appartenant au domaine d'attraction de la loi de Fréchet. On se restreint volontairement à 3 lois : la loi de Pareto, la loi de Pareto généralisée et la loi Exponentielle Pareto. Ces 3 lois ont respectivement 2, 3 et 4 paramètres, dont un paramètre de seuil (deux pour le cas de la loi Exponentielle Pareto) .

5.2.1 La loi de Pareto

Soit une variable aléatoire X qui suit une loi de Pareto de paramètre de seuil $\theta > 0$ et de paramètre de forme $\alpha > 0$.

Alors sa densité est de la forme :

$$\mathrm{f(x)} = \frac{\alpha\ \theta^\alpha}{x^{\alpha+1}} 1_{[\theta,+\infty[}\ (x)$$

Sa fonction de répartition est :

$$\mathrm{F(x)} = 1 - (\frac{\alpha}{x})^\theta$$

Et ses deux premiers moments sont :

$$E[X] = \frac{\alpha\theta}{\theta-1}$$
$$V[X] = \frac{\alpha^2\theta}{(\theta-2)(\theta-1)^2}$$

5.2.2 La loi Pareto Généralisée

Soit une variable aléatoire X qui suit une loi de Pareto généralisée de paramètre d'échelle β et de paramètre de forme ϵ.

Alors sa densité est de la forme :

$$f_{\epsilon,\beta} = \begin{cases} f_{\epsilon,\beta} = \frac{1}{\beta}(1 + \frac{\epsilon x}{\beta})^{-1/\epsilon-1} \text{ si } \epsilon \neq 0 \\ \frac{1}{\beta} e^{-x/\beta} \text{ si } \epsilon = 0 \end{cases}$$

Sa fonction de répartition est :

$$F_{\epsilon,\beta} = \begin{cases} 1 - [1 + \epsilon(x/\beta)]_+^{-1/\epsilon} \text{ si } \epsilon \neq 0 \\ 1 - e^{-x/\beta} \text{ si } \epsilon = 0 \end{cases}$$

Ses moments d'ordre r sont définis si $\epsilon < 1/\text{r}$, $\text{r} \in \mathbb{N}$, alors :

$$E[X^r] = \frac{\beta^r}{\epsilon^{r+1}} \frac{\Gamma(\epsilon^{-1} - r)}{\Gamma(1+\epsilon^{-1})} r!$$

Remarque : On peut considérer un troisième paramètre μ, appelé paramètre de localisation, en remplaçant x par (x-μ).

5.2.3 La loi Exponentielle Pareto

Dans un article de HURLIMANN [2014] **[12]** il est proposé deux lois pour la tarification prudente de traités de réassurance en excédent de sinistre. L'une d'entre elle est la distribution Exponentielle Pareto. Cette loi est le mélange entre une loi exponentielle tronquée et une loi de Pareto pour la queue de distribution.

On considère donc une loi de 4 paramètres, notée EP (α, β, T, γ), de fonction de répartition :

$$F(x) = \begin{cases} 1 - exp\left(-\frac{x-\alpha}{\beta}\right) & , \ \alpha \leq \ x \leq \ T \\ 1 - exp\left(-\frac{T-\alpha}{\beta}\right)\left(\frac{x}{T}\right)^{-\gamma} & , \ x \geq T \end{cases}$$

Ce choix est motivé par le possible manque de flexibilité de la loi de Pareto. En effet, sur certains jeux de données la loi de Pareto permet un très bon calibrage sur la queue de distribution mais pas sur les montants moins extrêmes, comme on l'a vu en début de partie. Cette loi de probabilité semble donc particulièrement convenir à notre étude. On doit toutefois vérifier qu'elle appartient bien au domaine d'attraction de la loi de Fréchet.

Pour cela on fait un bref rappel de théorie des valeurs extrêmes :

On considère une variable aléatoire absolument continue et on note f sa densité. On peut définir l'inverse de la fonction de hasard par :

$$h(x) = \frac{1 - F(x)}{f(x)}, \ x_F < x < x^F$$

où x_F et x^F sont les valeurs qui définissent le support de la variable aléatoire.

On peut alors caractériser les conditions d'appartenance à un domaine d'attraction. En effet, une distribution GEV(0,1,ϵ) est obtenue comme limite si :

$$h'(y) \to \epsilon \text{ quand y} \to x^F$$

On va donc maintenant utiliser cette caractérisation pour trouver le domaine d'appartenance de la distribution Exponentielle Pareto.

On suppose x $>$ T, alors :

$$F(x) = 1 - exp(-\frac{T-\alpha}{\beta}).(\frac{x}{T})^{-\gamma}$$

Par dérivation :

$$f(x) = exp(-\frac{T-\alpha}{\beta}).\frac{\gamma}{T}(\frac{x}{T})^{-\gamma^{-1}}$$

On définit maintenant la fonction h comme l'inverse de la fonction de hasard :

$$\begin{aligned} h(x) &= \frac{1-F(x)}{f(x)} \\ &= \frac{x}{\gamma} \end{aligned}$$

ainsi :

$$h'(x) = \gamma - 1 > 0$$

Par conséquent, la loi Exponentielle Pareto appartient bien au domaine d'attraction de la loi de Fréchet.

5.2.4 Calibrage

L'étape suivante consiste en l'estimation des paramètres de nos lois puis à la mise en œuvre de tests d'adéquation.

Pour estimer les paramètres de nos lois on utilise une méthode de minimisation des carrés des écarts entre la fonction de répartition empirique et la fonction de répartition des lois présentées dans la partie précédente.

On obtient les paramètres estimés suivants :

Distribution Pareto	
seuil	1 500 000
alpha	1,56

Distribution Pareto Généralisée	
seuil	1 500 000
epsilon	0,42
beta	1 197 524

Distribution Exponentielle Pareto	
alpha	1 500 000
beta	1 509 237
T	3 500 000
gamma	2,51

FIGURE 5.5 – Estimation des paramètres des 3 lois proposées

On regarde comment ces résultats se comparent avec la fonction de répartition empirique de notre échantillon de montants de sinistres extrêmes :

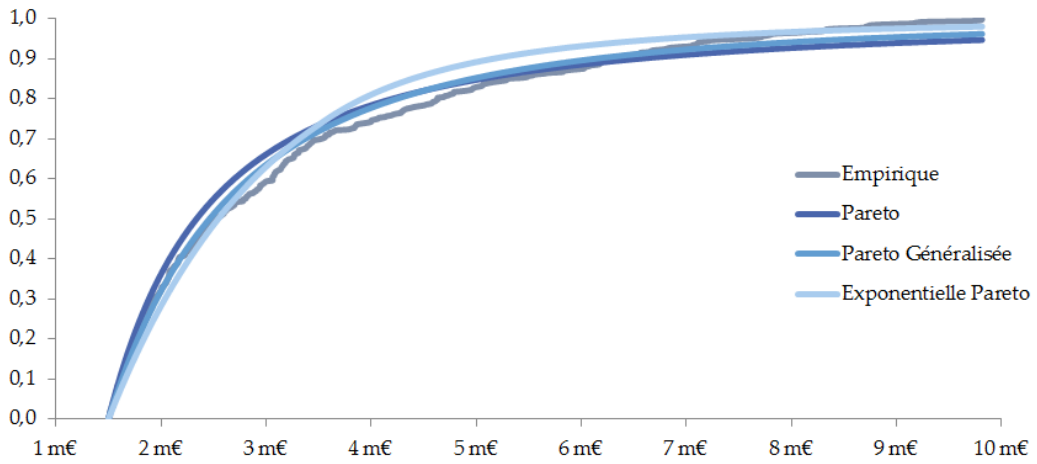


FIGURE 5.6 – Comparaison des fonctions de répartition obtenues avec la fonction de répartition empirique de notre échantillon

On peut observer que les paramètres de forme de nos lois de Pareto, Pareto Généralisée et Exponentielle Pareto ne proposent pas tous la même épaisseur pour la queue de distribution. En effet, la loi de Pareto semble trop prudente par rapport à la loi Pareto généralisée et la loi Exponentielle Pareto. Cela s'explique par le manque de flexibilité de la loi de Pareto qui, avec un seul paramètre (hors paramètre de seuil), peut difficilement convenir à la fois sur le bas et le haut de notre distribution empirique. On en conclut que la méthode mixte, qui repose exclusivement sur la loi de Pareto pour modéliser la sévérité, peut-être remise en question sur ce point.

On va maintenant procéder à des tests d'adéquations pour vérifier nos estimations de lois et de paramètres. Pour cela, on utilise les graphes quantile-quantile :

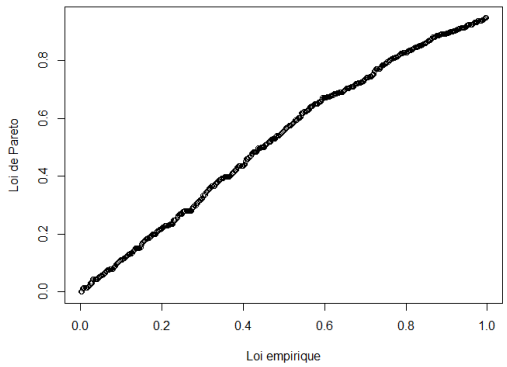


FIGURE 5.7 – QQ-plot de notre échantillon empirique avec la loi de Pareto

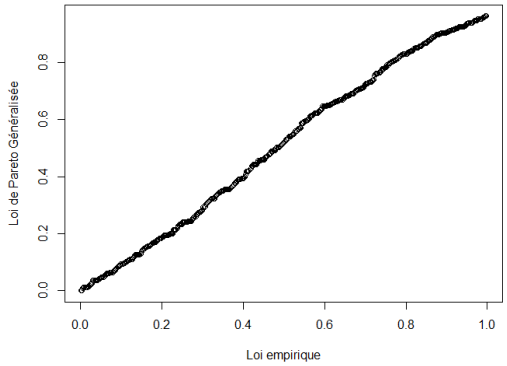


FIGURE 5.8 – QQ-plot de notre échantillon empirique avec la loi de Pareto Généralisée

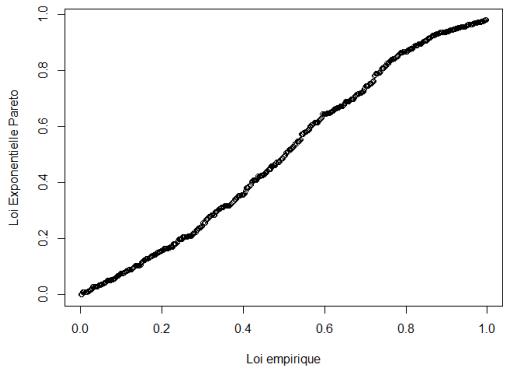


FIGURE 5.9 – QQ-plot de notre échantillon empirique avec la loi Exponentielle Pareto

L'étude des figures 5.7, 5.8 et 5.9 semblent illustrer que la loi de Pareto n'est pas la seule loi adéquate pour modéliser la sévérité dans le cas étudié. La loi Pareto Généralisée, dont la loi de Pareto est un cas particulier semble l'être également, tout comme la loi Exponentielle Pareto.

Remarque : On aurait aussi pu utiliser les tests de Kolmogorov-Smirnov et Anderson-Darling avec une modification pour prendre en compte le fait que l'on travaille avec des lois tronquées.

Pour plus de détails le lecteur peut se diriger vers DEPREY et GODZINSKI [2007] [**7**].

En résumé, dans cette partie nous avons étudié l'adéquation de la loi de Pareto pour modéliser la sévérité dans le processus de pricing. Nous nous sommes d'abord interrogés sur le seuil à partir duquel notre sinistralité pouvait être considérée comme extrême. Nous avons ainsi fait l'hypothèse que les sinistres de notre échantillon de données qui sont supérieurs à 1,5 million d'euros représentent notre sinistralité extrême. Nous avons ensuite retenu le domaine d'attraction de la loi de Fréchet. Enfin, nous avons étudié et calibré des lois appartenant à ce domaine. Cela nous a permis de nous rendre compte que la méthode mixte, qui repose sur la loi de Pareto, n'est peut-être pas l'unique modèle pertinent.

Chapitre 6

Changement de modèle de tarification

6.1 Le risque de modèle

L'étude de la sévérité de la sinistralité extrême, évoquée au cours du chapitre précédent, a fait apparaître le fait que la loi de Pareto n'est peut-être pas l'unique loi pertinente dans ce cas. C'est pourquoi dans ce chapitre nous souhaitons challenger la méthode mixte qui, dans son utilisation usuelle, repose sur la loi de Pareto. C'est ainsi que nous allons proposer un changement de modèle de tarification. Plus généralement, nous souhaitons proposer une quantification du risque de spécification du modèle. Autrement dit, quel est l'impact de choisir un tel modèle plutôt qu'un autre dans le processus de tarification ?

6.1.1 Définitions et problématiques

Un modèle peut se définir comme étant une méthode qui applique des théories scientifiques afin de quantifier un phénomène à partir de données.

Ayant défini ce qu'est un modèle, on peut maintenant définir le risque de modèle. Une récente étude, voir DAVESNE [2015] [**5**], défini le risque de modèle comme l'ensemble des risques engendrés par la création d'un modèle mathématique et propose la décomposition suivante sur la figure 6.1 :

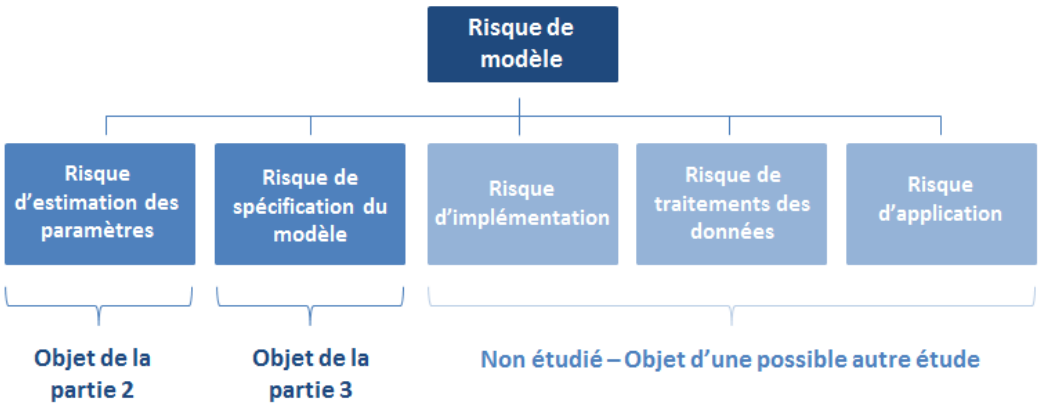


FIGURE 6.1 – Possible décomposition du risque de modèle

Les régulateurs demandent de faire preuve de prudence par rapport au risque de modèle, notamment via le processus ORSA, *Own Risk and Solvency Assessment*, qui demande une cartographie des risques quantitatifs et qualitatifs. Pour plus de détails le lecteur pourra se diriger vers GRIGOROV [2012] [**11**]. Néanmoins aucune méthodologie claire n'est proposée. On va donc voir quelles peuvent être les problématiques auxquelles un actuaire peut avoir à faire face.

Lors d'une démarche actuarielle avec des données, un modèle et une donnée d'intérêt à quantifier, on identifie trois sources d'incertitudes auxquelles l'actuaire est confronté :

- L'incertitude relative à la nature stochastique du risque.

- L'incertitude relative à la valeur des paramètres du modèle.

- L'incertitude relative au modèle choisi en lui-même.

Au sein de la dernière incertitude, on distingue trois formes d'incertitudes plus fines :

- Le vrai modèle est supposé être un individu, fixe mais inconnu, d'une classe générale de modèles. Par exemple, si on suppose que la sévérité d'un sinistre suit une loi de Pareto de paramètre (x_m, α)

- La classe de modèles considérée est constituée de modèles approximatifs, qui simplifie une réalité plus complexe sur lequel l'actuaire à une bonne base de connaissance.

- La classe de modèles considérée est constituée de modèles approximatifs, qui peuvent servir à appréhender une réalité plus complexe sur lequel l'actuaire n'a pas ou peu de connaissances.

Le risque de modèle englobe donc plusieurs notions différentes. Toutefois, d'après CAIRNS [2000] [**4**], les actuaires se concentrent souvent uniquement sur la structure d'un modèle pris pour acquis avec ses paramètres fixes. Ce n'est pas ce que nous avons fait dans cette étude. En effet avant de challenger le modèle nous avons d'abord souhaité étudier la façon dont ses paramètres étaient choisis. C'est l'objet de la partie 2 sur le risque d'estimation des paramètres de la méthode mixte.

Pour finir sur une citation, une autre croyance répandue et source d'incertitudes, est le principe de parcimonie. D'après JEFFREYS [1961], les gens préfèrent les explications simples aux explications complexes. Il déclare : "la loi la plus simple est choisie car elle est la plus susceptible de donner des prédictions correctes".

Nous venons de brièvement exposer le contexte autour du risque de modèle, nous allons maintenant présenter quelques éléments de théorie dont nous aurons besoin en fin d'étude.

6.1.2 Éléments de théorie

On présente maintenant des éléments de théorie dont on va chercher à s'inspirer par la suite pour proposer une quantification du risque de spécification du modèle.

- La notion de mesure de risque

Soit (Ω, $\mathcal{A}$, $\mathbb{P}$) un espace de probabilité. Une mesure de risque $\mathcal{R}$ est une fonction définie sur l'espace des variables aléatoires, et prenant ses valeurs dans $\mathbb{R}$:

$$\mathcal{R} \ : \ \Omega \ \rightarrow \ \mathbb{R}$$

$$\mathrm{X} \ \mapsto \ \mathcal{R}(\mathrm{X})$$

De plus, une mesure de risque vérifie six propriétés :

- Invariance en loi : $\mathrm{X} \sim \mathrm{Y} \Rightarrow \ \mathcal{R}\ (\mathrm{X}) = \ \mathcal{R}\ (\mathrm{Y})$
- Invariance par translation : $\forall \lambda \in \mathbb{R}$, $\mathcal{R}\ (\mathrm{X} + \mathrm{k}) = \ \mathcal{R}\ (\mathrm{X}) + \mathrm{k}$
- Homogénéité positive : $\forall \lambda \in \mathbb{R}^{+}$, $\mathcal{R}\ (\lambda\ \mathrm{X}) = \ \lambda\ \mathcal{R}\ (\mathrm{X})$
- Croissance : $\mathrm{X} \geq \mathrm{Y} \Rightarrow \mathcal{R}\ (\mathrm{X}) \geq \ \mathcal{R}\ (\mathrm{Y})$
- Sous-additivité : $\mathcal{R}\ (\mathrm{X} + \mathrm{Y}) \geq \ \mathcal{R}\ (\mathrm{X}) \ + \ \mathcal{R}\ (\mathrm{Y})$
- Convexité : $\forall \beta \in [0,1]$, $\mathcal{R}\ (\beta\ \mathrm{X} + (1\text{-}\beta)\ \mathrm{Y}) \geq \ \beta\ \mathcal{R}\ (\mathrm{X}) + (1\text{-}\beta)\ \mathcal{R}\ (\mathrm{Y})$

Ce mémoire n'a pas pour objet d'aller en profondeur dans la théorie. Pour plus de détails sur la notion de mesure de risque, le lecteur pourra se diriger vers DENUIT et CHARPENTIER [2004] [**6**].

- L'approche *worst case*

L'approche *worst case*, proposée par KHERKOF et al [2002] [**13**] mesure l'écart entre le modèle de référence et le modèle qui s'en éloigne le plus.

On définit ainsi la mesure *worst case* ρ :

$$\rho(M_{base}) = sup_{M_i \in M}\ \mathcal{R}(M_i)\ \text{-}\ \mathcal{R}(M_{base})$$

Cette approche est très simple mais peut s'avérer trop prudente car elle est totalement soumise aux extrêmes. Encore une fois, nous ne cherchons pas ici à présenter en détail la théorie mais plutôt à présenter des éléments sur lesquels s'inspirer par la suite.

De la même façon, on s'intéresse à une autre approche, l'approche bayésienne

- L'approche bayésienne

L'approche bayésienne, proposée par DRAPER [1995] [**8**] est plus fine que l'approche *worst case* car elle introduit des pondérations. En effet, on définit ainsi la mesure bayésienne ρ :

$$\rho(M_{base}) = \textstyle\sum_i \omega_i(\mathrm{R}(\mathrm{M}_i)\ \text{-}\ \mathcal{R}(M_{base})\ )$$

Avec ω_i qui reflète la vraisemblance du modèle i.

On suppose que cette approche permet de quantifier le risque de spécification du modèle, en étant moins sensible aux valeurs extrêmes que l'approche *worst case*. Néanmoins elle nécessite

de trouver les pondérations ω_i, ce qui peut s'avérer extrêmement difficile en pratique.

Pour plus de détails sur l'approche théorique pour évaluer l'incertitude, voir DRAPER [1995] [**8**].

Maintenant que nous avons présenté les éléments de théorie dont nous souhaitons nous inspirer, nous allons proposer une quantification du risque de modèle dans le cadre de nos données décrites partie 3.1.1.

6.2 Application

Tout d'abord, nous allons définir les scénarios que l'on va étudier. Puis, nous allons définir de nouveaux modèles fréquences-coût adaptés à nos données. Nous allons également proposer une mesure de la vraisemblance de nos modèles, avec l'utilisation notamment d'un bootstrap ligne à ligne. Ces étapes préliminaires nous permettrons finalement de proposer une quantification du risque de spécification du modèle, dans le cas de plusieurs scénarios.

Cette partie a demandé la création personnelle d'un outil sur R. L'outil comprend notamment l'indexation des sinistres, l'application de la clause de stabilité et la liquidation des charges *as-if*. La création d'un tel outil était en effet nécessaire afin de pouvoir faire des calculs de sensibilités rapidement, et en plusieurs dimensions comme par exemple dans la partie 4.2.1, mais également pour pouvoir utiliser la théorie du bootstrap que l'on va aborder dans cette partie.

6.2.1 Description de la méthode de tarification

Dans la partie 6.2, une autre méthode de tarification que la méthode mixte sera abordée, à savoir la méthode fréquence-coût classique.

La première étape de la méthode de tarification avec un modèle fréquence-coût est la modélisation de la sinistralité brute *as-if*. Il y a 3 étapes : le retraitement des données, et notamment l'application de la clause de rachat, l'indexation des sinistres, et la liquidation. Cette dernière étape est aussi appelée mise à l'ultime des sinistres. Une illustration de l'impact de ces étapes est proposée partie 2.2.1.

C'est à partir de la sinistralité brute as-if que sont calibrés les modèles fréquence-coût. La seconde étape est alors l'application du traité XS avec clause de stabilisation (voir partie 1.2 pour plus de détails) et ainsi obtenir la prime pure.

6.2.2 Tarification de plusieurs scénarios

On définit 8 scénarios différents que l'on obtient en modifiant 3 variables : le nombre d'années d'historique retenues, le type de données utilisées pour estimer l'inflation passée et les facteurs de développement (voir partie 2 pour plus de détails) et le niveau de priorité. Tous les autres inputs (couverture, taux prospectif de la clause de stabilité, indice de revalorisation des primes etc...) sont les mêmes pour tous les scénarios.

Scénario	1	2	3	4	5	6	7	8
Historique	8 ans	8 ans	12 ans	12 ans	8 ans	8 ans	12 ans	12 ans
Données utilisées pour l'inflation passée et la liquidation	marché	cédante	marché	cédante	marché	cédante	marché	cédante
Priorité	5 000 000	5 000 000	5 000 000	5 000 000	3 000 000	3 000 000	3 000 000	3 000 000

TABLE 6.1 – Tableaux des différents scénarios utilisés

Pour chaque scénario on propose 4 tarifications, La première avec la méthode mixte, la seconde avec un modèle fréquence-coût Poisson-Pareto, la troisième avec un modèle fréquence-coût Poisson-Pareto Généralisée et la quatrième avec un modèle fréquence-coût Poisson-Exponentielle Pareto.

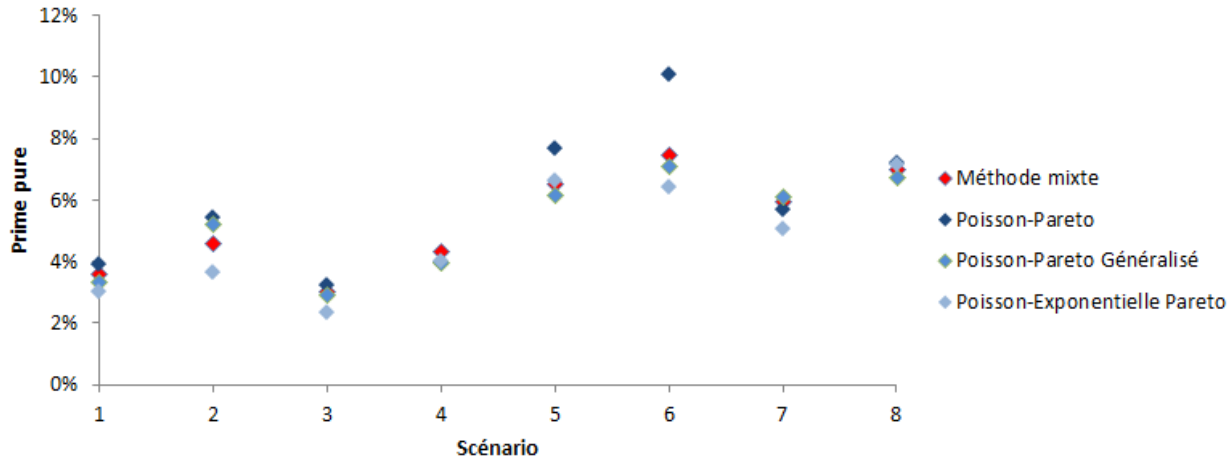


FIGURE 6.2 – Synthèse des différences de prix par scénario et par méthode de tarification utilisée

Premièrement, on remarque que le modèle Poisson-Pareto classique semble propoer des primes pures relativement plus élevées que la méthode mixte, qui elle aussi est un modèle Poisson-Pareto mais calibré d'une manière spécifique, voir partie 2.2.3. Ces résultats illustrent donc une nouvelle fois, après l'avoir vu dans la partie 2, l'importance non seulement du modèle utilisé mais du calibrage des paramètres.

Deuxièmement, on voit que le modèle Poisson-Pareto Généralisée semble proposer des prix relativement beaucoup moins élevés que le modèle Poisson-Pareto, et également un peu moins élevés que la méthode mixte. Cela peut s'expliquer par une épaisseur de la queue de distribution plus forte pour la loi de Pareto que pour la loi de Pareto Généralisée.

Enfin, on voit que le modèle Poisson-Exponentielle Pareto semble proposer des prix relativement moins élevés que la méthode mixte. Cela peut s'expliquer une nouvelle fois par le fait que la loi de Pareto semble parfois manquer de flexibilité pour modéliser entièrement la sévérité.

On résume ci-dessous l'écart de prix des différents modèles fréquence-coût par rapport à la méthode mixte :

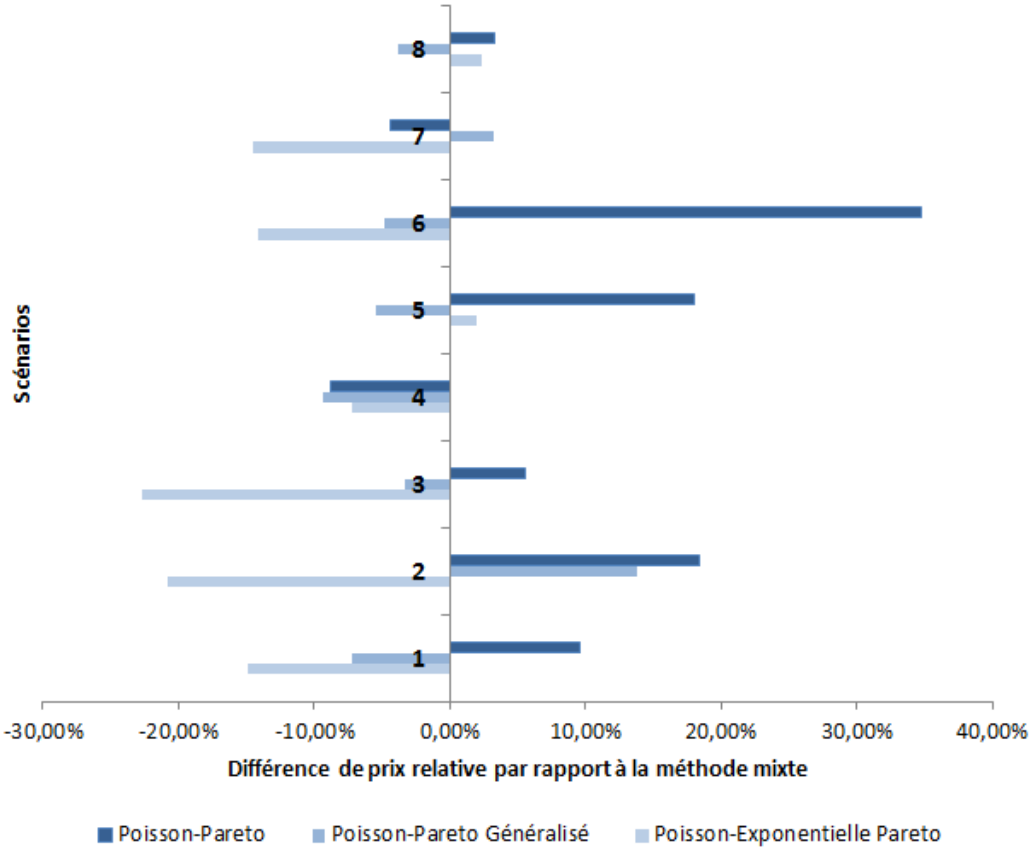


FIGURE 6.3 – Écarts de prix des modèles fréquence-coût par rapport à la méthode mixte

On ne peut toutefois pas conclure que telle ou telle méthode donne un prix plus compétitif que telle autre. En effet, les choix de paramètre pour chacune de ces méthodes sont très subjectifs et dépendent fortement du jugement de l'actuaire qui les utilise.

On précise qu'une idée pour réduire ce biais pourrait être de demander à plusieurs actuaires de faire cette étude de manière indépendante.

6.2.3 Vraisemblance et robustesse des différents modèles

Maintenant que l'on a les écarts de prix, on va chercher à les pondérer par la vraisemblance de chacun des modèles. Pour cela, on fait l'hypothèse que plus un modèle est vraisemblable alors plus il est robuste. Cela semble être une exigence dans la pratique pour le souscripteur de produit de réassurance. En effet à portefeuille, environnement monétaire et juridiction constants, une petite variation de la sinistralité au cours d'une année ne produit idéalement pas un changement significatif du prix de la réassurance.

On cherche donc à mesurer la robustesse de nos modèles. Pour cela on propose de "bootstraper" n fois notre sinistralité et de tarifier n fois le traité de réassurance considéré avec chacun des échantillons *bootstrap*. On supposera alors que notre mesure de robustesse sera l'inverse du coefficient de variation empirique ainsi obtenu.

Remarque : Cela n'est cependant possible que pour les modèles Poisson-Pareto et Poisson-Pareto

généralisée pour lesquels on dispose de moyens simples pour estimer les paramètres. Ce n'est en revanche pas le cas pour le modèle Poisson-Exponentielle Pareto.

On illustre le cas du modèle Poisson-Pareto :

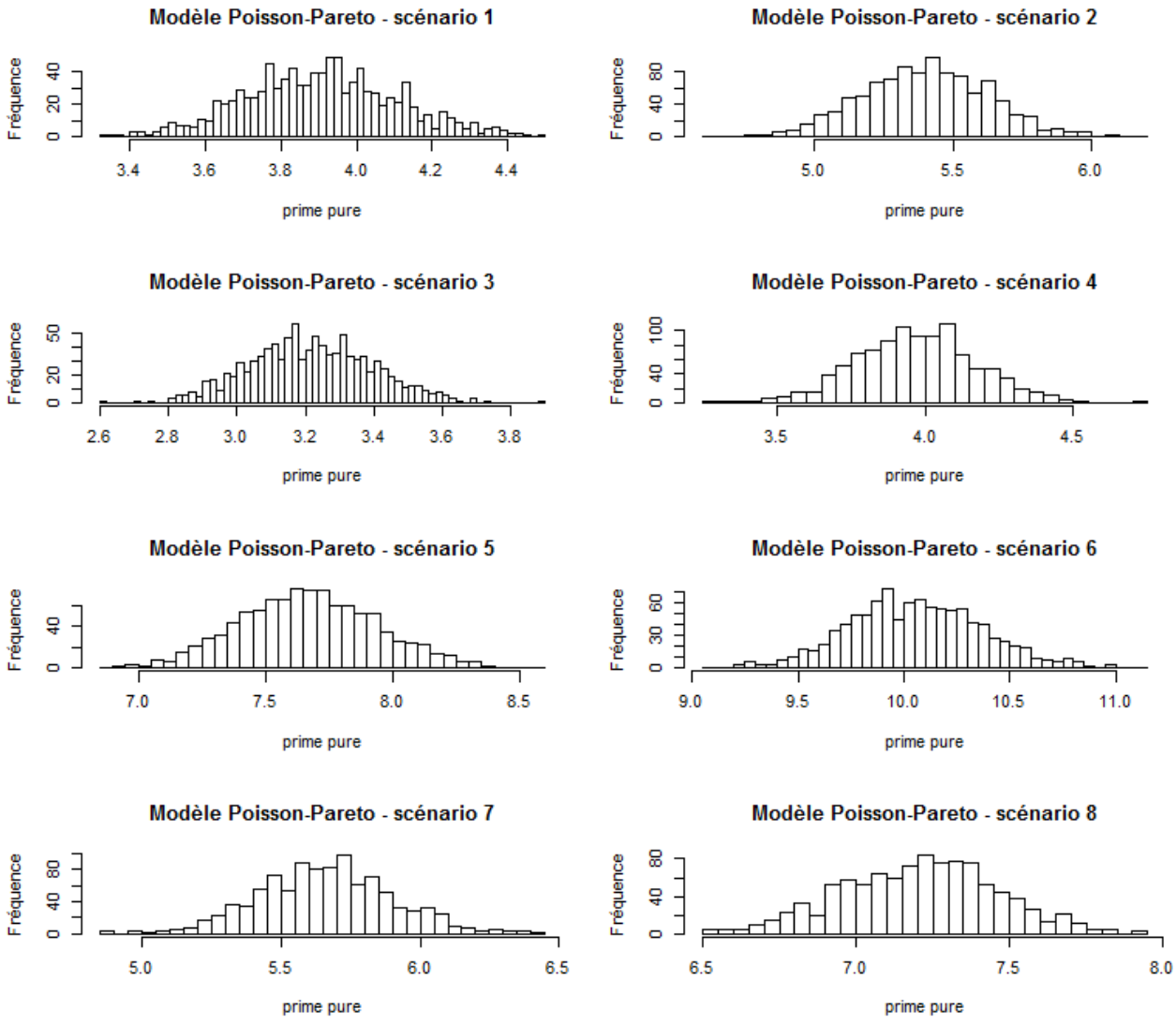


FIGURE 6.4 – Illustration des distributions obtenues par bootstrap ligne à ligne

Cela nous permet d'obtenir les coefficients de variation empiriques :

Scénario	1	2	3	4	5	6	7	8
Coefficient de variation empirique obtenu avec le modèle Poisson-Pareto	5,54	4,75	5,33	4,58	3,97	3,18	4,63	3,55
Coefficient de variation empirique obtenu avec le modèle Poisson-Pareto Généralisée	6	4,96	6,15	5,21	4,14	3,62	3,69	3,67

TABLE 6.2 – Coefficients de variations empiriques des distributions de la figure 6.4

Ces résultats ne nous permettent pas de valider ou non le choix de ces modèles, ce n'est pas l'objet de cette étude. Néanmoins, ces résultats nous éclairent pour l'attribution des pondérations dans le cadre de l'approche bayésienne. En effet on fait l'hypothèse, dans le cadre de l'approche bayésienne, que chacune des pondérations relatives à un modèle correspond à une probabilité. Cette probabilité doit refléter, pour un scénario donné, la vraisemblance qu'un modèle a d'être le vrai modèle, sachant que l'on fait l'hypothèse que le vrai modèle se trouve parmi les modèles proposés. C'est ainsi qu'en fonction de la robustesse des modèles que l'on a essayé de mesurer et par jugement d'expert, que l'on obtient les pondérations suivantes :

Scénario	1	2	3	4	5	6	7	8
Poisson-Pareto	32%	35%	37%	39%	25%	24%	26%	29%
Poisson-Pareto Généralisé	38%	35%	33%	31%	35%	36%	34%	31%
Poisson-Exponentielle Pareto	30%	30%	30%	30%	40%	40%	40%	40%

FIGURE 6.5 – Pondérations choisies, dont on fait l'hypothèse qu'elles correspondent à la probabilité que le modèle considéré est le vrai modèle, sous l'hypothèse que le "bon modèle" est contenu dans les 3 modèles proposés

Pour les 8 modèles Poisson-Exponentielle Pareto on a fait un jugement d'expert sur la vraisemblance d'être le bon modèle. Notre jugement d'expert a été guidé par l'hypothèse que plus la priorité est basse, plus la sinistralité est hétérogène et donc plus on a besoin de flexibilité. Ainsi on attribue une vraisemblance plus importante aux scénarios où la priorité est à 3 millions d'euros.

Pour les modèles Poisson-Pareto et Poisson-Pareto Généralisée les pondérations choisies sont obtenues via les formules suivantes :

$$\omega_{i,\text{Pareto}} = (1 - \omega_{i,\text{Exponentielle Pareto}}) * \frac{\rho_{i,\text{Pareto}}}{\rho_{i,\text{Pareto}} + \rho_{i,\text{Pareto Généralisée}}}$$

$$\omega_{i,\text{Pareto Généralisée}} = (1 - \omega_{i,\text{Exponentielle Pareto}}) * \frac{\rho_{i,\text{Pareto Généralisée}}}{\rho_{i,\text{Pareto}} + \rho_{i,\text{Pareto Généralisée}}}$$

où :

- i représente le numéro du scénario considéré
- $\omega_{i,\text{Pareto}}$ représente la pondération accordé au modèle Poisson-Pareto pour le scénario i
- $\omega_{i,\text{Pareto Généralisée}}$ représente la pondération accordé au modèle Poisson-Pareto Généralisée pour le scénario i
- $\omega_{i,\text{Exponentielle Pareto}}$ représente la pondération accordée au modèle Poisson-Exponentielle Pareto, accordé par jugement d'expert, pour le scénario i
- $\rho_{i,\text{Pareto}}$ représente l'inverse du coefficient de variation empirique de la figure 6.8 correspondant
- $\rho_{i,\text{Pareto Généralisée}}$ représente l'inverse du coefficient de variation empirique de la figure 6.8 correspondant

6.2.4 Quantification du risque de spécification du modèle

On propose ci-dessous une quantification du risque de spécification du modèle. Premièrement en s'inspirant de l'approche *worst-case* puis en s'inspirant de l'approche bayésienne, vues partie 6.1.2.

— Quantification avec l'approche *worst-case*

Selon cette approche on considère que l'erreur de spécification de modèle est la valeur absolue de l'écart de prix relatif entre le prix donné par la méthode mixte et le prix donné par le modèle fréquence-coût qui s'en écarte le plus.

On obtient ainsi les résultats suivants :

Scénario	1	2	3	4	5	6	7	8
Erreur de spécification du modèle	14,9%	20,8%	22,6%	9,4%	18,0%	34,7%	14,5%	3,8%

TABLE 6.3 – Erreurs de modèle proposées avec l'approche *worst-case* pour chacun des 8 scénarios

— Quantification avec l'approche bayésienne

Selon cette approche on considère que l'erreur de spécification de modèle est la moyenne de la valeur absolue des écarts de prix obtenus par les modèles fréquence-coût et le prix obtenu avec la méthode mixte, pondéré par la mesure de vraisemblance de la figure 6.8.

On obtient ainsi les résultats suivants :

Scénario	1	2	3	4	5	6	7	8
Erreur de spécification du modèle	10,3%	17,5%	9,9%	8,5%	7,2%	15,6%	8,0%	3,1%

TABLE 6.4 – Erreurs de modèle proposées avec l'approche bayésienne pour chacun des 8 scénarios

Les erreurs de modèle données par la seconde proposition sont moins importantes car l'approche considérée est plus prudente, en raison de l'introduction de pondérations. Il est très difficile de pousser l'analyse des résultats plus loin. En outre, ce n'est pas l'objectif de ce mémoire. En effet, l'objectif de ce chapitre était de proposer un changement de modèle et de proposer une quantification du risque de spécification du modèle. Autrement dit mesurer l'impact de choisir un modèle en particulier plutôt qu'un autre ; en l'occurrence, l'impact de choisir la méthode mixte pour la tarification plutôt qu'un modèle fréquence-coût classique.

Conclusion

Cette étude visait à apporter des éléments de réponses à plusieurs questions, notamment :

La méthode mixte, dans son utilisation classique, est-elle l'unique modèle pertinent ? En supposant son utilisation, comment peut-on le calibrer et quel est l'impact du choix de calibration ? Pour aller plus loin, puis-je avoir une mesure quantifiée de l'impact du choix de ce modèle plutôt que de tel autre ?

L'objectif de ce mémoire était donc de proposer de nouvelles méthodes, de nouveaux indicateurs et d'illustrer l'impact de ces nouveaux choix.

C'est ainsi que le chapitre 3 propose un autre estimateur de l'inflation du passé, qui a permis de mettre en évidence certaines propriétés de la loi de Pareto. En effet, dans le cas où la sévérité suit une loi de Pareto, alors, sous certains points de vue, l'inflation semble invisible. L'explication mathématique du phénomène a alors permis la mise en évidence des limites de la méthode initiale, et d'en proposer une solution, grâce à l'utilisation d'une méthode de comptage.

Dans un second temps, l'étude du chapitre 4 illustre une alternative à Chain-Ladder, qui parait particulièrement intéressante sous l'hypothèse que la sévérité suit une loi de Pareto. En effet, la théorie autour du phénomène d'effet de seuil en réassurance permet, dans ce cas, d'obtenir très simplement un estimateur des facteurs de développement de la charge du réassureur.

Enfin, la troisième partie a permis de voir que dans le cas étudié, la loi de Pareto ne semble pas être l'unique choix pertinent pour modéliser la sévérité. En effet, il existe d'autres lois qui possèdent également des propriétés très intéressantes. L'étude de ces lois a permis de proposer de nouveaux modèles de tarification.

L'objet du chapitre 6 a alors été la construction de nouveaux estimateurs, d'abord de la robustesse des modèles, puis de leur vraisemblance et enfin du risque de spécification du modèle. Cette partie permet donc d'avoir une mesure quantifiée du risque de choisir la méthode mixte comme modèle.

Pour finir, ce mémoire peut être élargi à un ensemble de cédantes. Une telle étude aurait un intérêt aussi bien pour la construction d'une courbe de marché, que dans le cadre de l'ORSA qui incite à quantifier le risque de modèle de manière globale. Une approche plus fine de la méthode de la partie 2, par exemple à l'échelle de chaque poste de préjudice, représente aussi un axe de développement. Enfin, une étude sur la variance des estimateurs proposés peut également être envisagée.

Bibliographie

[1] BERNARDO J., SMITH A. [1994] « Bayesian Theory »

[2] BLONDEAU J., PARTRAT C. [2003] « La Réassurance Approche Technique »

[3] BRAZAUSKAS V., JONES B., ZITIKIS R. [2009] « When Inflation Causes No Increase in Claim Amounts »

[4] CAIRNS A. [2000] « A Discussion of Parameter and Model Uncertainty in Insurance »

[5] DAVESNE C. [2015] « Etude du risque de modèle dans le cadre d'un modèle interne », *Mémoire d'actuariat*

[6] DENUIT M., CHARPENTIER A. [2004] « Mathématiques de l'assurance non-vie », Tome I

[7] DEPREY E., GODZINSKI A. [2007] « Problématique de seuil dans la modélisation de la sinistralité en Réassurance Non Vie », *Mémoire d'actuariat*

[8] DRAPER. D. [1995] « Assessment and Propagation of Model Uncertainty »

[9] Fédération Fançaise de l'Assurance [2017] « Le Marché de l'Assurance Automobile en 2016 »

[10] FIGUEREAU. E [1992] « Utilisation de la méthode de l'échantillon saturé en assurances I.A.R.D. », *Mémoire d'actuariat*

[11] GRIGOROV V. [2012] « ORSA and model risk Focus on Life insurance », *Conférence-débat de l'Institut des Actuaires*

[12] HURLIMANN W. [2014] « Pareto Type Distribution and Excess-of-Loss Reinsurance »

[13] KERKHOF J., MELENBERG B., SCHUMACHER H. [2002] « Model Risk and Regulatory Capital », *Available at SSRN 301531*

[14] LACOUR [2018] « Problématique de seuils de gravité en provisionnement non-vie mémoire d'actuaire », *Mémoire d'actuariat.*

[15] LALLEMENT T., NISIPASU E., TOPUZU M., ELBAHTOURI L. [2014] « Le risque de modèle »

[16] Newsletter Technique SCOR [2015] « Responsabilité civile automobile », *Analyse de l'indemnisation du corporel grave pour la période 2012 à 2014 au sein du marché français*, numéro 28

[17] QBE Re (Europe) Ltd – Secura Branch [2016] « Sinistres corporels graves en France »

Table des figures

1.1 Taux de résiliations et taux d'affaires nouvelles en assurance automobile 10
1.2 Répartition de la charge des sinistres selon la garantie 11
1.3 Vie d'un sinistre grave 15
1.4 Évolution du PIB et de l'inflation par rapport au coût moyen d'un sinistre corporel grave 16
1.5 Nomenclature des principaux préjudices corporels proposée par la commission Dinthillac 17

2.1 Vue d'ensemble des modèles 21
2.2 Impact des différentes étapes de la mise en *as-if* d'un sinistre 21

3.1 Échantillon des montants de 1000 sinistres extrait de notre base de données . . . 28
3.2 Exemple de triangle de fonctions de répartition 29
3.3 Évolution des visions en 2017 des fonctions de répartition des montants de sinistres survenus entre 2002 et 2012 29
3.4 Comparaison des deux fonctions de répartitions ainsi obtenues et mesure de l'adéquation par QQ plot 30
3.5 À gauche, diagrammes en boite de tous les montants et à droite, diagrammes en boite des montants observés par le réassureur 31
3.6 Comparaison des deux fonctions de répartitions ainsi obtenues et mesure de l'adéquation par QQ plot 33
3.7 Sommaire des données simulées 34

4.1 Illustration d'un triangle de fonctions de répartition 39
4.2 Fonctions de répartition empiriques des sinistres survenus sur la période 2005-2014 vus en année 1, puis vus en année 2 39
4.3 Impact du développement des sinistres entre la première et la deuxième année sur la fonction de répartition issue des sinistres survenus sur la période 2005-2014 . . 40
4.4 Projections de fonctions de répartition par changement de variable et mesure de l'adéquation par QQ-plot 41
4.5 L'accélération en fonction du ratio 44
4.6 Accélération selon différents niveaux de priorité 45
4.7 Illustration de l'évolution de la prime pure en fonction du taux d'indexation des sinistres 46
4.8 Illustration de l'évolution de la prime pure en fonction de la priorité et du taux d'indexation 46

5.1 Illustration du possible manque de flexibilité de la loi de Pareto pour modéliser la sinistralité 51

5.2 Illustration du possible gain de flexibilité apportée par la loi Exponentielle-Pareto à 4 paramètres 52
5.3 Mean excess function appliquée à notre sinistralité 54
5.4 Mean excess function appliquée à un échantillon simulé avec une loi de Pareto . . 54
5.5 Estimation des paramètres des 3 lois proposées 58
5.6 Comparaison des fonctions de répartition obtenues avec la fonction de répartition empirique de notre échantillon 58
5.7 QQ-plot de notre échantillon empirique avec la loi de Pareto 59
5.8 QQ-plot de notre échantillon empirique avec la loi de Pareto Généralisée 59
5.9 QQ-plot de notre échantillon empirique avec la loi Exponentielle Pareto 59

6.1 Possible décomposition du risque de modèle 61
6.2 Synthèse des différences de prix par scénario et par méthode de tarification utilisée 65
6.3 Écarts de prix des modèles fréquence-coût par rapport à la méthode mixte 66
6.4 Illustration des distributions obtenues par bootstrap ligne à ligne 67
6.5 Pondérations choisies, dont on fait l'hypothèse qu'elles correspondent à la probabilité que le modèle considéré est le vrai modèle, sous l'hypothèse que le "bon modèle" est contenu dans les 3 modèles proposés 68
6.6 Répartition par classe d'âge des victimes d'un accident corporel grave 75
6.7 Répartition par statut social des victimes d'un accident corporel grave 75
6.8 Répartition par profession des victimes d'un accident corporel grave 76
6.9 Délai de consolidation par type de séquelles 76
6.10 Illustration de la 12e colonne, sur les 16, de notre triangle de fonctions de répartition 77
6.11 Comparaison des fonctions de répartitions ainsi obtenues et mesure de l'adéquation par QQ plot 80
6.12 Comparaison des fonctions de répartitions ainsi obtenues et mesure de l'adéquation par QQ plot 82

Liste des tableaux

1.1 Nombre de véhicules assurés au 31 décembre 2016 9
1.2 Cotisations selon le type de contrat et le type de véhicule assuré 10
1.3 Niveau de prime moyenne selon le type de contrat 10
1.4 Évolution des fréquences et des coûts moyens suivant les garanties souscrites . . . 11
1.5 Évolution des indemnités des blessés avec AIPP des principaux postes de préjudices en 2015 12
1.6 Illustration du fonctionnement d'un traité en excédent de sinistre par risque . . . 12

3.1 Tableau des estimateurs de la valeur des taux d'inflation annuelle 30
3.2 Tableau des taux d'inflation annuels supposés 33
3.3 Application numérique pour différents seuils 36

4.1 Facteurs de développement au premier euro 41
4.2 Facteurs de développement de charges au Traité XS 47

6.1 Tableaux des différents scénarios utilisés 65
6.2 Coefficients de variations empiriques des distributions de la figure 6.4 67
6.3 Erreurs de modèle proposées avec l'approche *worst-case* pour chacun des 8 scénarios 69
6.4 Erreurs de modèle proposées avec l'approche bayésienne pour chacun des 8 scénarios 69

Annexes

A Résumés statistiques sur les victimes d'un accident corporel grave

.

Ces illustrations sont extraites du document [**17**].

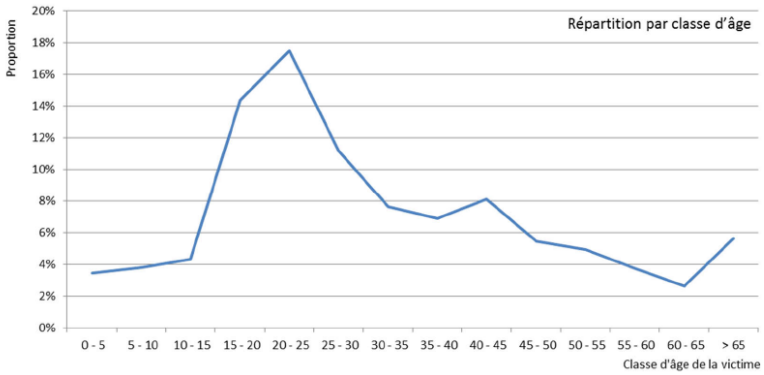


FIGURE 6.6 – Répartition par classe d'âge des victimes d'un accident corporel grave

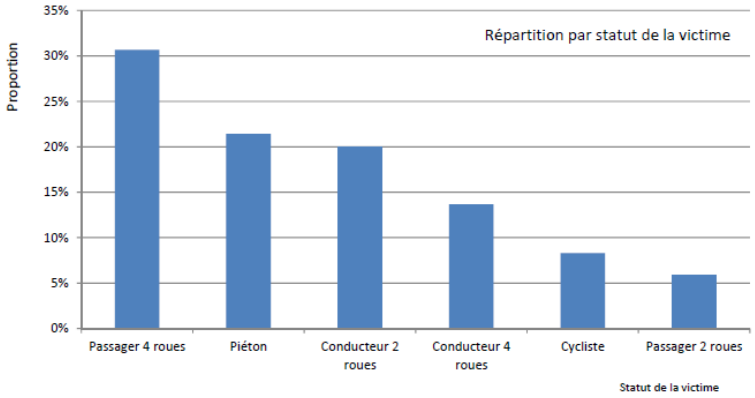


FIGURE 6.7 – Répartition par statut social des victimes d'un accident corporel grave

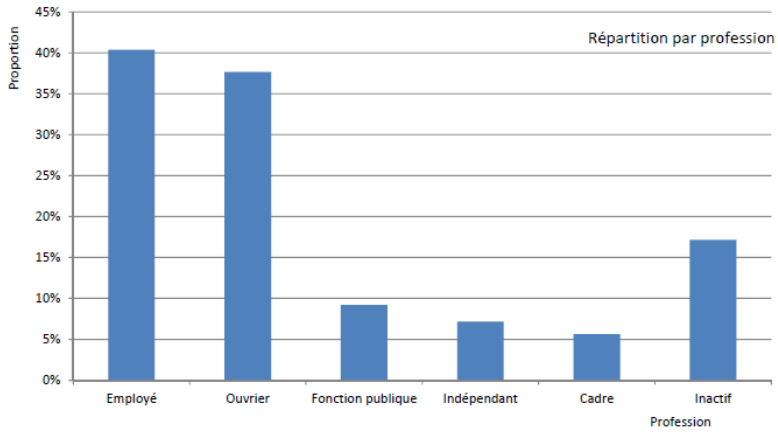


FIGURE 6.8 – Répartition par profession des victimes d'un accident corporel grave

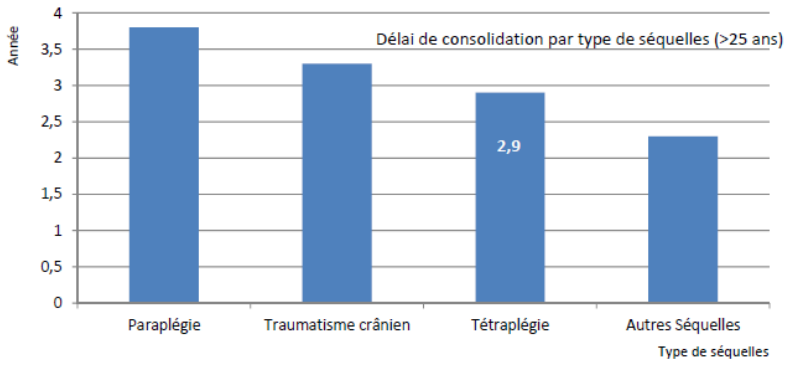


FIGURE 6.9 – Délai de consolidation par type de séquelles

B Triangle de fonctions de répartition

Pour les parties 3.1, on travaille sur un triangle de fonction de répartition de 16 lignes et 16 colonnes. On présente ici à titre d'exemple la 12e colonne de ce triangle.

Remarque : Pour la partie 4.1, c'est le même triangle mais chaque colonne correspond à une année de développement et non à une année comptable.

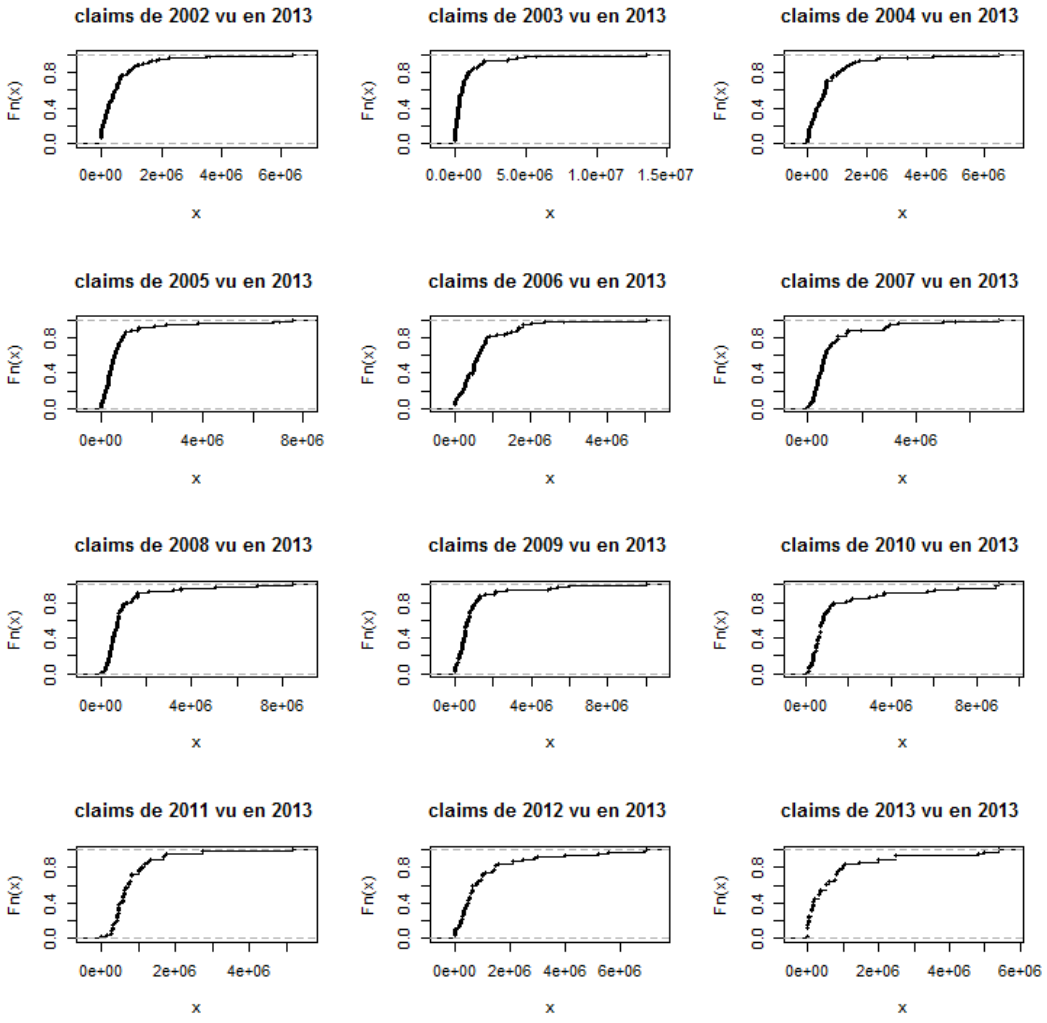
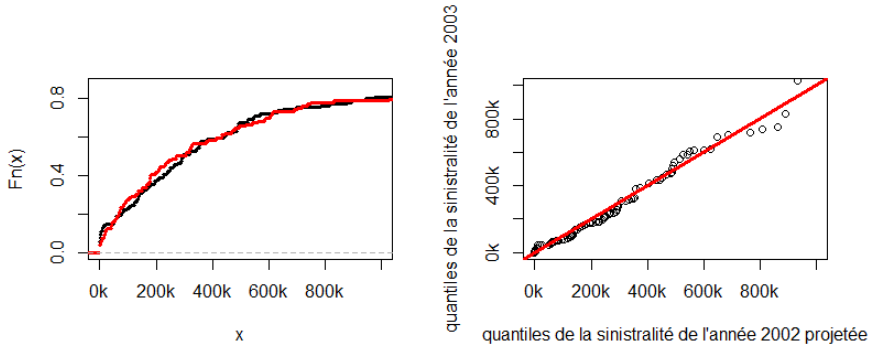
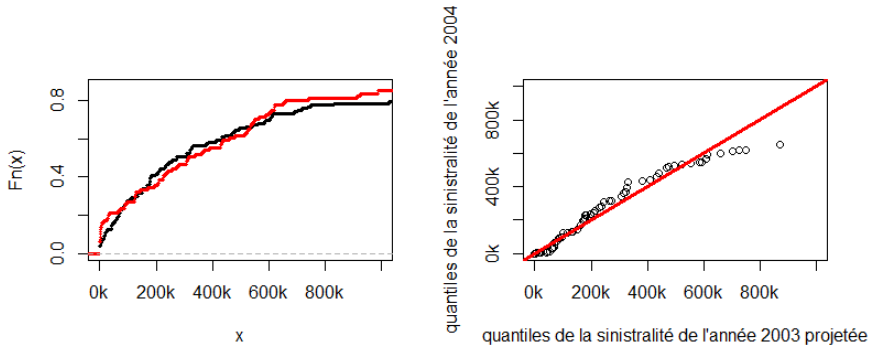
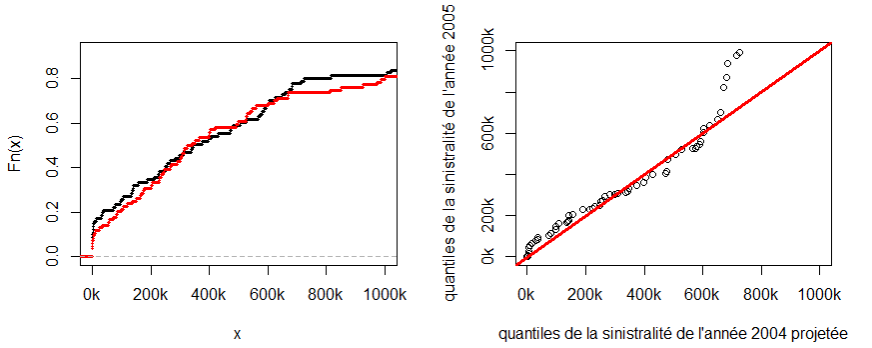


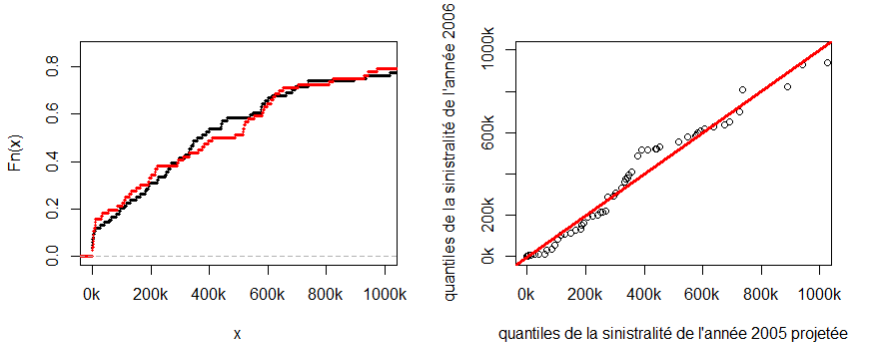
FIGURE 6.10 – Illustration de la 12e colonne, sur les 16, de notre triangle de fonctions de répartition

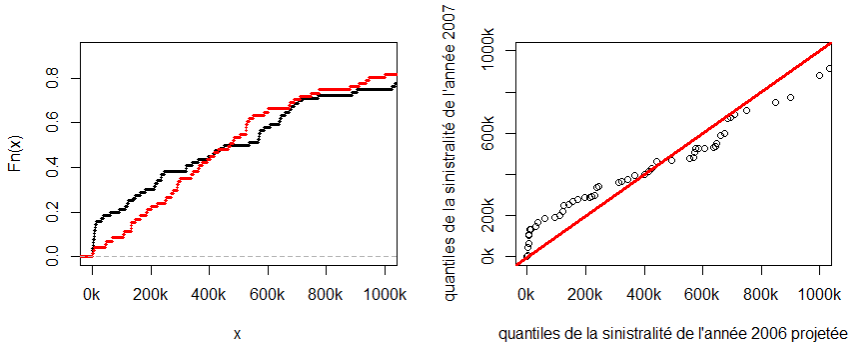
C Résultats de la partie 3.1.3

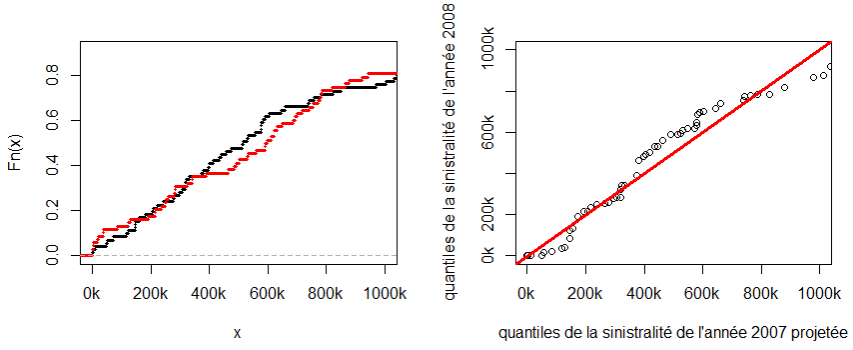


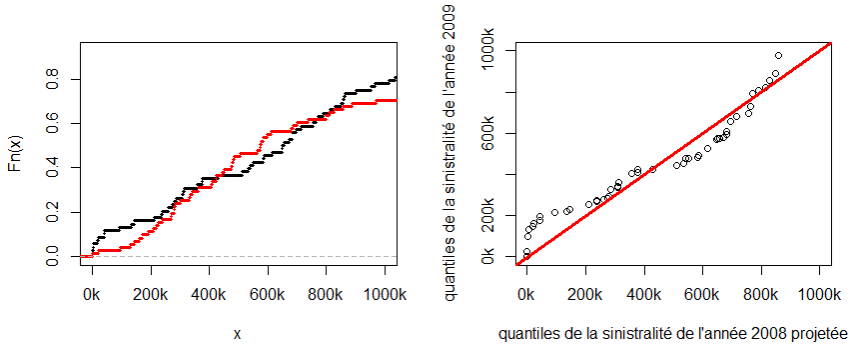


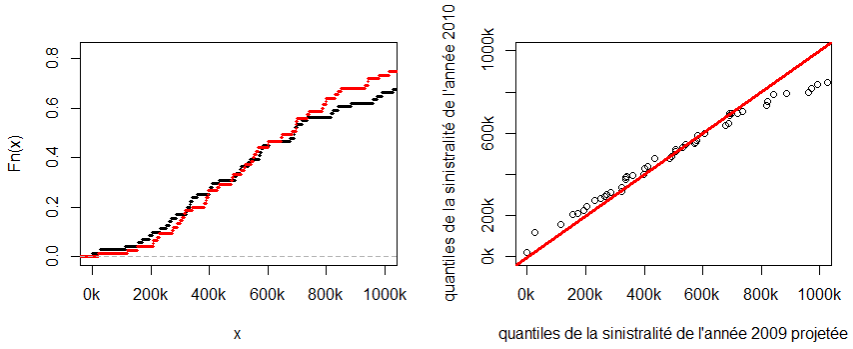





Fn(x)
0.0
0.2
0.4
0.6
0.8
0k
200k
400k
600k
800k
1000k
x
quantiles de la sinistralité de l'année 2007
quantiles de la sinistralité de l'année 2006 projetée


Fn(x)
0.0
0.2
0.4
0.6
0.8
0k
200k
400k
600k
800k
1000k
x
quantiles de la sinistralité de l'année 2008
quantiles de la sinistralité de l'année 2007 projetée


Fn(x)
0.0
0.2
0.4
0.6
0.8
0k
200k
400k
600k
800k
1000k
x
quantiles de la sinistralité de l'année 2009
quantiles de la sinistralité de l'année 2008 projetée


Fn(x)
0.0
0.2
0.4
0.6
0.8
0k
200k
400k
600k
800k
1000k
x
quantiles de la sinistralité de l'année 2010
quantiles de la sinistralité de l'année 2009 projetée

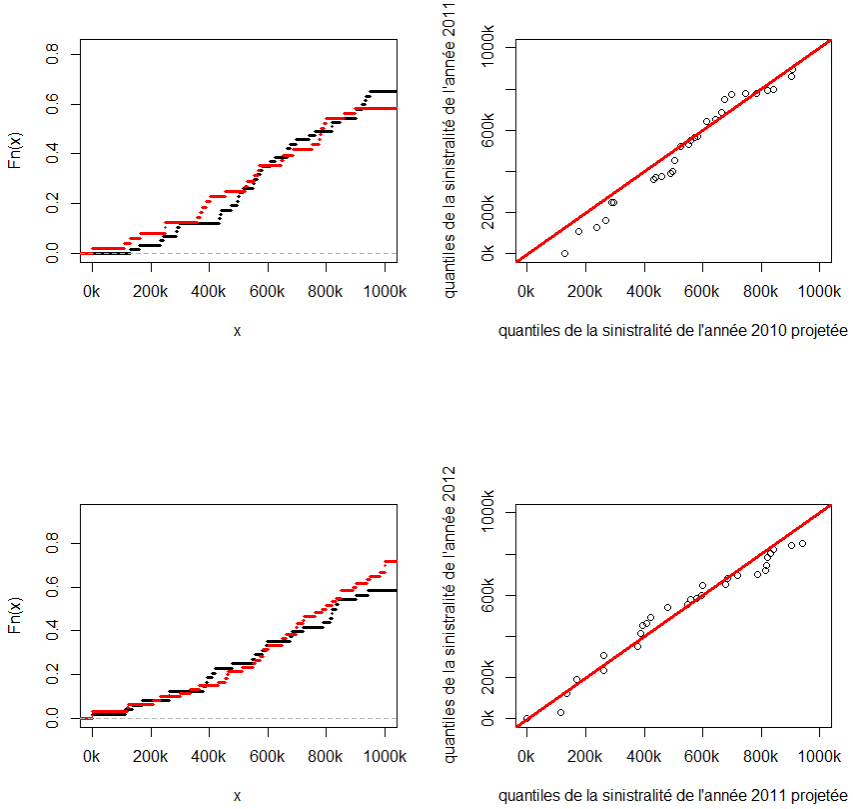
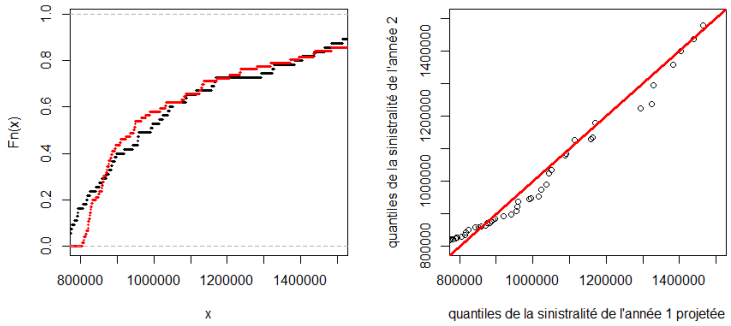
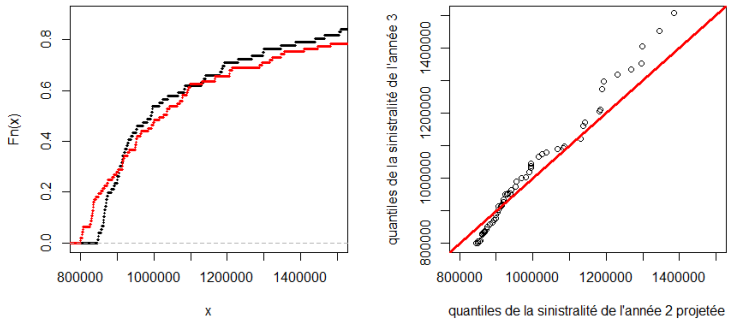
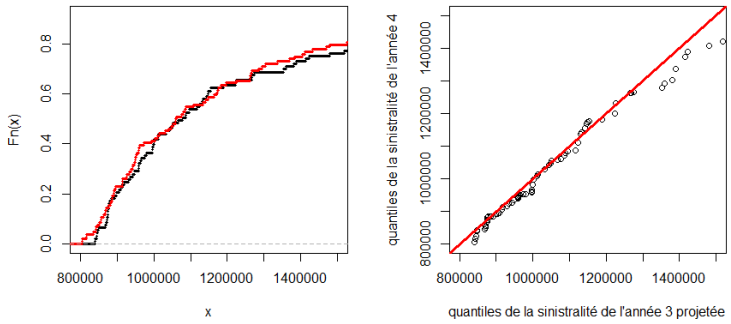


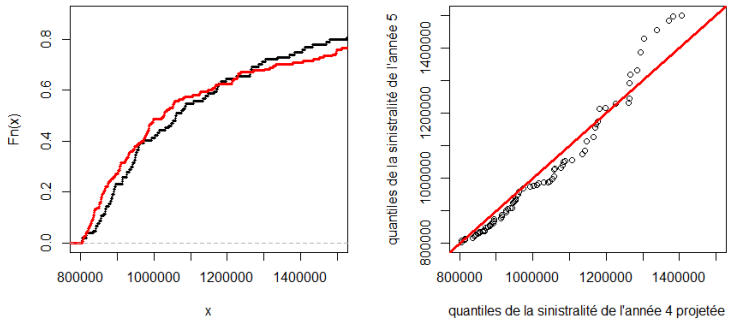
FIGURE 6.11 – Comparaison des fonctions de répartitions ainsi obtenues et mesure de l'adéquation par QQ plot

D Résultats de la partie 3.2.3









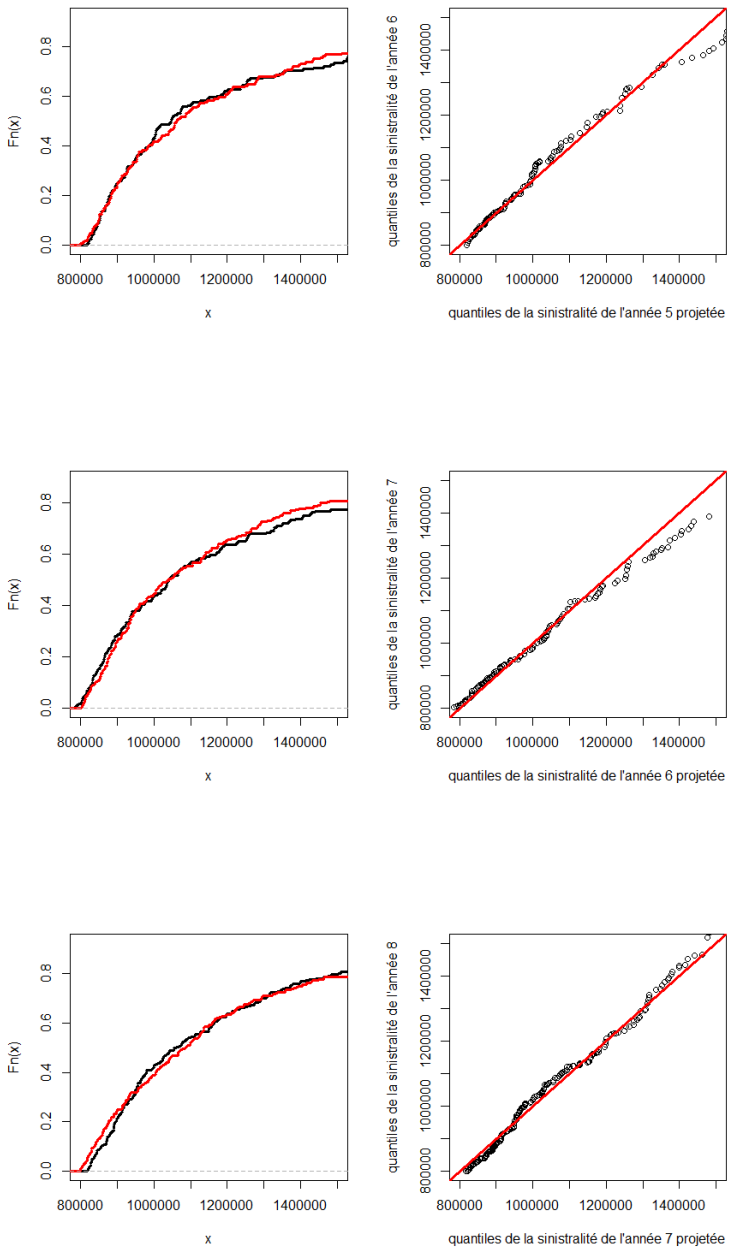


FIGURE 6.12 – Comparaison des fonctions de répartitions ainsi obtenues et mesure de l'adéquation par QQ plot

E Mesure de l'incertitude du processus de liquidation

Pour plus de détails le lecteur pourra se diriger vers le document LACOUR [2018] [**14**], dont cette annexe est inspirée.

E.1 Le modèle de Mack

La formule de Mack permet de calculer l'incertitude des montants de sinistres liquidés avec la méthode de Chain Ladder. Elle décompose cette incertitude en deux termes, le premier correspondant à la variance de l'estimateur et le deuxième à la variance du processus.
Mack définit alors ainsi l'estimateur de la MSEP (Mean Square Error Prediction), pour chaque année de survenance i :

$$\begin{aligned}\widehat{MSEP}_{R_i|D_I}(\widehat{R}_i) &= \widehat{MSEP}_{C_{i,J}|D_I}(\widehat{C}_{i,J}) \\ &= \mathbb{E}[(\widehat{C}_{i,J} - C_{i,J})^2|D_I] \\ &= \underbrace{(\ \widehat{C}_{i,J} - \mathbb{E}[C_{i,J}|D_I]\ )^2}_{erreur\ d'estimation} + \underbrace{Var[C_{i,J}|D_I]}_{erreur\ de\ processus}\end{aligned}$$

avec $\hat{\sigma}_j^2$ et $\hat{f}_j$ les estimateurs selon Mack.

E.2 Le bootstrap

Le bootstrap, en plus de mesurer l'incertitude des montants liquidés comme la formule de Mack, permet de donner une hypothèse sur la distribution des montants liquidés.

On suit les étapes suivantes :

1. Application de la méthode Chain Ladder sur le triangle initial.
2. Reconstruction du triangle supérieur à l'aide des facteurs de développements $\hat{f}_j$ calculés avec la méthode Chain Ladder.
3. Calcul des résidus de Mack ajustés : $r_{i,j} = \sqrt{\theta_j}.\frac{C_{i,j+1} - C_{i,j}.\hat{f}_j}{\sqrt{\hat{\sigma}_j^2.C_{i,j}}}$ avec $\theta_j = \frac{I-j+1}{I-j}$

Ces trois étapes ne sont à effectuer qu'une seule fois. Les étapes suivantes sont à répéter autant de fois que souhaité. Prenons un nombre B d'échantillons bootstrap et $b \in 1, B$.

4. Obtention de nouveaux triangles de résidus $\widehat{r}_{i,j}^b$ en procédant par tirage avec remise des résidus initiaux $\widehat{r}_{i,j}$.
5. Calcul de nouveaux facteurs de développement $\hat{f}_j^b$ par inversion de la formule des résidus en étape 3.
6. Application des nouveaux facteurs de développement $\hat{f}_j^b$ sur la dernière diagonale du triangle initial et projection du triangle inférieur.
7. Simulation des réalisations futures de la charge à l'aide d'une loi Gamma et des moyennes et variances des charges futures théoriques calculées à l'étape 6.