

**Mémoire présenté le :**

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

Par : Sorelle Sandra BEKO NEABO

Titre Modélisation individuelle des sinistres RC Auto corporels graves selon  
une approche hybride dans le cadre d'un modèle interne

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 :*

*Nom : Groupama Assurances Mutuelles*

*Signature :*

*Membres présents du jury de l'ISFA*

*Directeur de mémoire en entreprise :*

*Nom : KUNTZ Sébastien*

*Signature : *

*Invité :*

*Nom :*

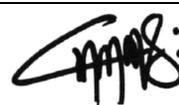
*Signature :*

***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





---

# RÉSUMÉ

Le sujet de ce mémoire est d'évaluer la charge individuelle de sinistres (brute et nette de réassurance) Auto corporels graves et de calculer le SCR de primes associé. L'objectif in fine de cette modélisation est de challenger la modélisation actuelle, tout en corrigeant les limites de celle-ci.

Pour évaluer cette charge de sinistres, nous partirons de notre base de rente initiale contenant les caractéristiques biométriques de chaque victime (sexe, âge, taux d'AIPP), que nous simulerons conditionnellement (en gardant le même profil de rentiers) pour obtenir une base de rentes futures. Par la suite, nous estimerons le montant d'arrérage de chaque rentier simulé en fonction du sexe. Nous utiliserons ensuite une approche classique *fréquence*  $\times$  *sévérité* afin d'estimer le nombre sinistres et le coût total de sinistre. Ce coût se décompose en rentes + capital + frais de soins et est donc fonction des caractéristiques biométriques de la victime. Nous ferons une modélisation distincte en fonction du type d'indemnisation (rentes et hors rentes) et le montant des charges dépendra également du taux d'actualisation utilisé (courbe des taux EOIPA pour les sorties en rentes et taux du barème Gazette du Palais pour les sorties en capital). En plus de l'inflation classique (FGAO), les sinistres subissent une inflation hors monétaire propre aux sinistres Auto corporels, qui est de 5% dans notre étude.

Une fois ces charges estimées, nous sélectionnerons les sinistres dit graves : il s'agit des sinistres dont la charge dépasse le seuil des graves défini par le modèle interne, qui est de 1,2M€. Dans une optique de calculer le SCR de primes brut de réassurance, nous ferons M tirages aléatoires de ces sinistres afin de calculer le quantile à 99,5%.

Les sinistres Auto corporels graves, étant une branche à développement long et dont le coût de sinistres est élevé, les assureurs ont souvent recours à la réassurance afin de protéger leurs fonds propres. Nous appliquerons donc les traités de réassurance à nos sinistres afin de déterminer la prime de la réassurance et le SCR net de réassurance.

Mots clés : Sinistres auto corporels graves, indemnisation en rente, indemnisation en capital, taux d'AIPP (Atteinte à l'Intégrité Physique et Psychique), inflation corporelle, FGAO (Fonds de Garanties des Assurances Obligatoires de dommages), arrérages, modèle de régression linéaire, modèle de régression log-linéaire, modèle GLM, arbres CART, SCR (Capital de Solvabilité Requis), risques vie.

---

# ABSTRACT

The subject of this thesis is to evaluate the individual cost of claims (gross and net of reinsurance) for serious bodily injuries and to calculate the associated SCR of premiums. The ultimate objective of this modeling is to challenge the current modeling, while correcting its limits.

To evaluate this loss burden, we will start from our initial annuity base containing the biometric characteristics of each victim (gender, age, physical and psychic integrity impairment rate), which we will simulate conditionally (keeping the same annuitant profile) to obtain a base for future annuities. Subsequently, we will estimate the amount of arrears for each simulated annuitant based on gender. We will then use a classical *frequency*  $\times$  *severity* approach to estimate the number of claims and the total cost of claim. This cost is broken down into annuities + capital + care costs and is therefore a function of the victim's biometric characteristics. We will do a separate modeling according to the type of compensation (annuities and non annuities) and the amount of the expenses will also depend on the discount rate used (EOIPA rate curve for annuities and Gazette du Palais rate scale for capital outlays). In addition to conventional inflation (FGAO), claims are subject to non-monetary inflation specific to bodily injury claims, which is 5% in our study.

Once these expenses have been estimated, we will select the so-called **serious** claims : these are claims whose expense exceeds the threshold of serious claims defined by the internal model, which is €1.2M. In order to calculate the SCR of gross reinsurance premiums, we will make M random draws of these claims in order to calculate the 99.5% quantile.

Serious Bodily Injury Auto claims, being a long development line with high claim costs, insurers often use reinsurance to protect their equity. We will therefore apply reinsurance treaties to our claims in order to determine the estimated reinsurance premium and the SCR net of reinsurance.

Keywords : Critical bodily injury auto claims, annuity compensation, capital compensation, physical and psychic integrity impairment rate, bodily inflation, FGAO (Mandatory Third Party Liability Insurance Guarantee Fund) , arrears, linear regression model, log-linear regression model, GLM model, CART trees (Classification And Regression Trees), SCR (Solvency Capital Requirement), life risk.

---

# SYNTHÈSE

L'évaluation de la charge de sinistres liée aux dommages corporels en Automobile représente un grand enjeu pour les assureurs du fait de leur faible fréquence d'occurrence mais d'un coût très élevé et dont l'indemnisation peut aller sur plusieurs années, en fonction de la victime. De ce fait, certains assureurs ont recours à un modèle interne afin de mieux capter leur risque et avoir un modèle plus précis. Le sujet de ce mémoire est donc d'évaluer la charge individuelle des sinistres Auto corporels graves selon une approche hybride IARD/vie dans le cadre d'un modèle interne partiel et d'en déduire le SCR de primes associé. L'objectif in fine de cette modélisation est de challenger la modélisation actuelle, tout en corrigeant les limites de celle-ci.

Un sinistre auto corporel représente les atteintes plus ou moins importantes constatées sur une victime d'un accident de la route. Celles-ci peuvent être de nature physique ou psychique. Le coût du sinistre étant fonction de la gravité, l'indemnisation peut se faire sous forme de rente et/ou de capital. Les sinistres sortis en rente devront être revalorisés. Avant 2013, cette majoration était prise en charge par le Fonds de garantie des Assurances Obligatoires (FGAO). Mais depuis le 1<sup>er</sup> Janvier 2013, elle est à la charge des assureurs. Ainsi, ceux-ci devront provisionner cet engagement par la prise en compte d'un taux d'inflation.

Seuls les sinistres considérés comme **graves** selon le modèle interne Groupama (seuil à 1,2M€ défini par le modèle interne) seront étudiés dans ce mémoire. Les assureurs ont de ce fait recours à la réassurance afin de protéger leurs fonds propres et leur bilan. Cependant, la modélisation actuelle faite dans le modèle interne partiel présente certaines limites. Elle estime la charge de sinistres sans distinction de la forme d'indemnisation et donc utilise un taux d'actualisation fixe appelé le taux social, qui représente 60% du TME moyen sur les 24 derniers mois + 10 pb : ce modèle n'est donc pas dynamique car il ne capte pas les fluctuations liées à un mouvement des taux. De plus, il ne prend pas en compte les caractéristiques biométriques de la victime (sexe, âge, taux d'AIPP, ...) et par conséquent, n'intègre pas des risques vie auxquels les sinistres Auto corporels sont associés. Enfin, l'application des traités de réassurance est faite de manière différente de celle des réassureurs, ce qui implique de gros écarts sur les résultats nets de réassurance. Afin de palier à ces manquements, nous allons proposer une modélisation qui sera plus fine et précise.

Sur notre base de rente initiale contenant les caractéristiques biométriques de chaque victime, nous avons fait des statistiques descriptives afin de nettoyer et comprendre le comportement de nos variables. La variable taux d'AIPP étant très volatile et fonction de l'âge et du sexe de la victime, nous avons créé des classes d'âges homogènes par sexe selon le critère de la moyenne. Par la suite, nous avons cherché à établir une relation linéaire entre les variables âge, sexe et taux d'AIPP et le montant de la rente (arrérage). Pour ce faire, nous avons fait des tirages aléatoires des victimes en simulant conditionnellement les variables sexe, âge et taux d'AIPP : en effet, il s'agit de simuler M victimes tout en gardant le même profil de rentiers pour obtenir une base de rentes futures. Une fois les lois conditionnelles connues, nous avons estimé le montant d'arrérage de chaque rentier : nous avons testé plusieurs modèles de régression tels que la régression linéaire multiple, la régression log-linéaire, les modèles GLM, les arbres CART. Nous avons évalué leur qualité prédictive et nous avons utilisé les critères RMSE et MAE afin de choisir le meilleur modèle : chez les hommes, la régression log-linéaire est plus adaptée et chez les femmes, le modèle GLM avec fonction de lien log.

Une fois le montant d'arrérage prédit, nous avons modélisé la sinistralité future en utilisant une approche classique *fréquence × sévérité*. Pour modéliser la fréquence annuelle de sinistres, nous avons testé la loi de poisson et la loi binomiale négative et nous avons retenu la loi binomiale négative (selon les tests statistiques) comme distribution de fréquence. Quant à la sévérité, elle est fonction des caractéristiques biométriques de la victime et du type d'indemnisation. La charge de sinistres se décompose en rentes + capital + frais de soins. Chacune de ses composantes a été modélisée indépendamment des autres. Les provisions de rente / capital correspondent à l'actualisation des cashflows à l'aide de la courbe des taux EOIPA pour les rentes et le taux du barème Gazette du Palais pour les sinistres en capital et est revalorisée par l'inflation classique (2%). En

---

plus de celle-ci, les sinistres subissent une inflation hors monétaire propre aux sinistres corporels, qui est de 5% dans notre étude. Les frais de soins ont été calibré en utilisant une distribution de loi log-normale et ont été ajoutés on-top des provisions.

Une fois la charge annuelle individuelle estimée, nous sélectionnerons les sinistres graves et nous ferons  $M=10000$  tirages aléatoires de ces sinistres afin de calculer la charge annuelle moyenne de sinistres ainsi que le SCR de primes.

NB : Dans le but de comparer nos résultats avec ceux du modèle actuel, nous avons modélisé une 2<sup>eme</sup> charge de sinistre de chaque victime en utilisant le taux social comme taux d'actualisation et nous avons déduit le SCR de primes associé.

En utilisant le taux social comme taux d'actualisation dans le nouveau modèle, nous avons constaté une hausse du coût moyen de sinistres de +45% environ par rapport à celui du modèle actuel. Cette augmentation peut s'expliquer par le fait que le modèle actuel ne capte ni l'effet biométrie ni l'effet non linéaire dans les variables explicatives sur la partie FGAO. Cet impact est le même sur le SCR brut de réassurance. Cependant, en utilisant la courbe des taux S2, nous avons constaté que l'impact du taux n'est pas très grand en terme de coût moyen (8% d'écart environ). Par contre, sur le SCR brut de réassurance, l'écart est de 19% environ : 8% de cet écart peut s'expliquer par le coût moyen, le reste est dû à l'effet taux d'actualisation.

Certes notre modélisation est plus précise, détaillée et permet un gain de SCR (ce qui est très important pour les assureurs), le modèle actuel reste tout de même prudent malgré ses limites car les écarts observés ne sont pas si importants.

En appliquant nos traités de réassurance en Excédent de sinistres (3M XS 12M, -1 XS 15M), nous constatons que les primes estimées avec le nouveau modèle sont très proches de celles réellement cédées au réassureur. De plus, nous avons un gain de SCR net de réassurance de +16% par rapport à celui du modèle actuel. Ce résultat est satisfaisant car l'objectif pour un assureur de contracter une réassurance est de pouvoir limiter l'immobilisation de son capital.

---

# SYNTHESIS

The evaluation of the burden of claims related to bodily injury in the automobile industry represents a major challenge for insurers due to their low frequency of occurrence but very high cost and for which compensation can be paid over several years, depending on the victim. Therefore, some insurers use an internal model to better capture their risk and have a more accurate model. The subject of this thesis is therefore to evaluate the individual burden of serious bodily injury claims using a hybrid PC/Life approach within the framework of a partial internal model and to deduce the associated premium SCR. The ultimate objective of this modeling is to challenge the current modeling, while correcting its limitations.

A bodily injury claim represents the more or less important damages observed on a victim of a road accident. These damages can be physical or psychological. The cost of the claim being a function of the seriousness, the compensation can be made in the form of an annuity and/or capital. Claims that are paid out as an annuity will have to be upgraded. Before 2013, this increase was paid by the Fonds de garantie des Assurances Obligatoires (FGAO). But since the 1st January 2013, it is the responsibility of insurers. Thus, they will have to provide for this commitment by taking into account an inflation rate.

Only claims considered as "serious" according to the Groupama internal model (threshold of €1.2M defined by the internal model) will be studied in this report. Insurers therefore resort to reinsurance in order to protect their equity and their balance sheet. However, the current modeling done in the partial internal model has some limitations. It estimates the cost of claims without distinguishing the form of compensation and therefore uses a fixed discount rate called the social rate, which represents 60% of the average TME over the last 24 months + 10 bp : this model is therefore not dynamic because it does not capture the fluctuations linked to a movement in rates. Moreover, it does not take into account the biometric characteristics of the victim (gender, age, AIPP rate, ...) and consequently does not integrate the life risks to which auto bodily injury claims are associated. Finally, the application of reinsurance treaties is different from that of the reinsurers, which implies large discrepancies in the net reinsurance results. In order to remedy these shortcomings, we will propose a model that will be more refined and precise.

On our initial database containing the biometric characteristics of each victim, we performed descriptive statistics to clean up and understand the behavior of our variables. Since the AIPP rate variable is highly volatile and depends on the age and sex of the victim, we created homogeneous age classes by sex according to the mean criterion. Then, we tried to establish a linear relationship between the variables age, sex and AIPP rate and the amount of the pension (arrears). To do this, we made random draws of the victims by conditionally simulating the variables sex, age and PIA rate : in fact, the idea is to simulate M victims while keeping the same profile of annuitants to obtain a base of future annuities. Once the conditional laws were known, we estimated the amount of arrears for each annuitant : we tested several regression models such as multiple linear regression, log-linear regression, GLM models and CART trees. We evaluated their predictive quality and used the RMSE and MAE criteria to choose the best model : for men, the log-linear regression is better suited and for women, the GLM model with log link function.

Once the arrears amount was predicted, we modeled future claims experience using a classical approach. To model the annual frequency of claims, we tested the poisson distribution and the negative binomial distribution and chose the negative binomial distribution (according to statistical tests) as the frequency distribution. As for the severity, it is a function of the biometric characteristics of the victim and the type of compensation. The cost of claims is broken down into annuities + capital + care costs. Each of these components has been modeled independently of the others. The annuity / capital provisions correspond to the discounting of cash flows using the EOIPA rate curve for annuities and the Gazette du Palais rate for capital claims, and are revalued by traditional inflation (2%). In addition to this, claims are subject to non-monetary inflation specific to bodily injury claims, which is 5% in our study. The care costs have been calibrated using a lognormal distribution and have been added on-top of the reserves.

---

Once the individual annual burden is estimated, we will select the serious claims and make M=10000 random draws of these claims in order to calculate the average annual burden of claims and the premiums SCR.

Note : In order to compare the results with the current model, we have modeled a 2<sup>nd</sup> claims burden for each victim using the social rate as the discount rate and we have deduced the associated premium SCR

Using the social rate as the discount rate in the new model, we found an increase in the average cost of claims of about +45% compared to the current model. This increase can be explained by the fact that the current model does not capture either the biometric effect or the non-linear effect in the explanatory variables on the FGAO part. This impact is the same on the gross reinsurance SCR. However, using the S2 rate curve, we found that the impact of the rate is not very large in terms of average cost (about 8% difference). On the other hand, on the gross reinsurance SCR on the other hand, the difference is about 19% : 8% of this difference can be explained by the average cost, the rest is due to the discount rate effect.

Admittedly, our model is more precise and detailed and allows for a gain in SCR (which is very important for insurers), but the current model is still prudent despite its limitations, since the differences observed are not so significant.

By applying our reinsurance treaties in Excess of Losses (3M XS 12M, -1 XS 15M), we see that the premiums estimated with the new model are very close to those actually ceded to the reinsurer. Moreover, we have a gain in SCR net of reinsurance of +16% compared to the current model. This result is satisfactory because the objective for an insurer to contract reinsurance is to be able to limit the immobilization of its capital.

---

# Sommaire

<b>Résumé</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>Synthèse</b>	<b>3</b>
<b>Synthesis</b>	<b>5</b>
<b>Remerciements</b>	<b>9</b>
<b>Introduction générale</b>	<b>10</b>
<b>I Enjeux de modélisation pour les assureurs</b>	<b>12</b>
<b>1 Périmètre de l'étude</b>	<b>13</b>
1.1 Contexte . . . . .	13
1.2 Solvabilité 2 et gestion des risques non-vie . . . . .	13
1.3 Approche modèle interne et risques associés . . . . .	19
1.4 Les engagements en Automobile . . . . .	22
<b>2 Indemnisation des sinistres corporels Automobile</b>	<b>25</b>
2.1 La loi Badinter . . . . .	25
2.2 Nomenclature DINTILHAC . . . . .	25
2.3 Mise en place de l'indemnisation . . . . .	26
2.4 Les différents postes de préjudice indemnisés . . . . .	26
<b>3 Principes de gestion de sinistres corporels Auto Groupama</b>	<b>29</b>
3.1 Mode de gestion . . . . .	29
3.2 La revalorisation des sinistres . . . . .	31
3.3 Les sources de volatilité de la rente . . . . .	31
3.4 La réassurance . . . . .	32
<b>II Modélisation des rentes futures</b>	<b>35</b>
<b>1 La base de rente RC Auto</b>	<b>36</b>
1.1 Analyse descriptive de la base de rentes . . . . .	36
1.2 Modélisation des variables explicatives et leur dépendance . . . . .	44
<b>2 Méthodes de prédiction de l'arrérage de base</b>	<b>50</b>
2.1 Les variables explicatives . . . . .	50
2.2 Les différents modèles . . . . .	50
2.3 Choix du meilleur modèle . . . . .	65
<b>3 Modélisation stochastique de la sinistralité future</b>	<b>68</b>
3.1 Modélisation de la fréquence de sinistres des rentes . . . . .	68
3.2 Modélisation du coût du sinistre : sévérité . . . . .	69
3.3 Calcul du SCR de Primes brut de réassurance . . . . .	73

---

<b>III</b>	<b>Analyse des modèles et résultats</b>	<b>75</b>
<b>1</b>	<b>Comparaison du modèle développé avec le modèle existant</b>	<b>76</b>
1.1	Approche de modélisation des sinistres Auto corporels graves par le modèle interne . . . . .	76
1.2	Comparaison des résultats . . . . .	77
<b>2</b>	<b>La réassurance</b>	<b>79</b>
2.1	De la sinistralité brute à la sinistralité nette de réassurance . . . . .	79
2.2	Application . . . . .	80
<b>3</b>	<b>Impact du FGAO</b>	<b>81</b>
<b>4</b>	<b>Efficience de la réassurance</b>	<b>82</b>
4.1	Modification de la priorité du traité . . . . .	83
4.2	Modification de la portée du traité . . . . .	83
<b>5</b>	<b>Tests de sensibilité</b>	<b>85</b>
<b>IV</b>	<b>Calcul du SCR Vie sur les rentes Auto Corporels</b>	<b>86</b>
<b>1</b>	<b>Méthodologie de calcul</b>	<b>87</b>
1.1	Le choc de longévité . . . . .	87
1.2	Le choc de frais de gestion (expenses) . . . . .	88
1.3	Le choc de révision . . . . .	88
<b>2</b>	<b>Calcul du SCR vie sur les rentes</b>	<b>89</b>
	<b>Conclusion</b>	<b>90</b>
<b>A</b>	<b>ANNEXES</b>	<b>93</b>
A.1	Définition des postes de préjudice permanents selon la nomenclature Dintilhac . . . . .	93
A.2	Tableau de synthèse des postes de préjudices de la nomenclature DINTILHAC . . . . .	95
A.3	Résultats de la modélisation de l'arrérage chez les femmes . . . . .	96
A.4	Statistiques comparatives des variables explicatives après simulation par sexe . . . . .	99
A.5	Table de mortalité TH/TF 00-02 . . . . .	99
A.6	Courbe des taux d'inflation permettant de calculer les provisions BE . . . . .	100
A.7	Table de mortalité TH/TF 00-02 choquée de -20% . . . . .	101
A.8	Courbe de taux économiques Formule standard, incluant la volatility adjustment (VA) . . . . .	101
A.9	Courbes des taux économiques choquées à la hausse et à la baisse, Formule standard, incluant la volatility adjustment (VA) . . . . .	102
	<b>Bibliographie</b>	<b>102</b>

---

# REMERCIEMENTS

Je souhaite tout d'abord remercier le groupe Groupama et en particulier la Direction Actuariat Groupe de m'avoir accueillie et donné l'opportunité d'effectuer ma seconde année d'alternance dans cette direction.

Je tiens ensuite à adresser ma profonde reconnaissance et mes remerciements à Sébastien KUNTZ (mon tuteur), Yann CUEFF, Romain GUILBON, Damien JOUBERT, Hervé DUMONTROT, Nathalie BOUGNON, Nicolas STIGAS et Camille GATELLIER pour le temps qu'ils m'ont accordé, leur expertise, leur gentillesse et la richesse des échanges que nous avons pu avoir sur les problématiques, les concepts et les méthodes présentés dans ce rapport.

Je remercie également tout le service modèle interne non-vie ainsi que les membres de la Direction Actuariat Groupe, pour leur accueil chaleureux, leur bonne ambiance et leur bonne humeur quotidienne (qu'ils m'ont transmis par ailleurs).

J'adresse enfin toute ma gratitude à l'ensemble des professeurs de l'ISFA qui m'ont permis de confirmer mon intérêt pour le métier d'actuaire en me proposant une formation non seulement théorique mais également pratique.

---

# Introduction générale

Toute personne possédant un véhicule terrestre à moteur doit obligatoirement souscrire à une assurance automobile. Elle permet de couvrir les dommages subis lors d'un accident de la circulation. Dans le cas où le dommage est causé à un tiers, c'est la responsabilité civile de l'assuré (le conducteur) qui indemnise la victime. Cette garantie est donc obligatoire. Le dommage causé pouvant être matériel ou corporel, nous nous intéressons dans le cadre de ce mémoire uniquement aux dommages corporels.

Un dommage corporel est une atteinte plus ou moins importante de l'intégrité physique ou psychique du corps d'une victime à la suite d'un accident. L'évaluation du montant de l'indemnisation de la victime d'un dommage corporel sera fait au moyen du **taux d'AIPP** (taux d'atteinte à l'intégrité physique et psychique). C'est un pourcentage qui permet de décrire la gravité du dommage sur la capacité physique ou psychique d'une victime d'accident. Il varie entre 1 et 100%. Plus il est élevé, plus le dommage corporel est important et plus conséquent sera le montant de l'indemnisation.

L'indemnisation peut se faire sous forme de rentes<sup>1</sup> (viagères ou temporaires) ou sous forme de capital. A cette rente et/ ou capital, s'ajoute une partie frais de soins, pour constituer le coût total prévisionnel d'un sinistre. Si celui-ci est supérieur ou égale à 1 200 000€, alors ce sinistre est considéré comme grave dans le modèle interne Groupama. Dans le cadre de ce mémoire, nous nous intéressons aux sinistres corporels Auto graves.

La modélisation actuelle des sinistres corporels Auto graves dans le modèle interne non-vie de Groupama ne tient pas compte le fait que les sinistres peuvent être potentiellement des rentes futures et donc les pertes liées à ces sinistres seront exposées à un risque supplémentaire, comme le risque de longévité (risque que la victime vive plus longtemps que ce que l'assureur avait prévu lors du calcul de ses provisions) , le risque révision (augmentation du montant de la rente suite à un changement de situation ou une dégradation de l'état de santé de la victime) ou le risque de frais de gestion (sous-estimation du montant des frais de gestion dans les contrats). Le SCR de primes ainsi calculé (correspond au capital économique nécessaire pour faire face au risque non-vie de prime à horizon 1 an) ne prend donc pas en compte cette spécificité.

De plus, ces sinistres corporels Auto étant gérés comme des sinistres dommages classiques, le modèle ne prend pas en compte les caractéristiques biométriques de la victime (l'âge, le sexe, la richesse, l'indice de gravité, ...) et d'autres facteurs tels que la table de mortalité et le taux d'actualisation. Le modèle actuel est également non dynamique car il ne capte pas les fluctuations liées à un mouvement des taux. En effet, une hausse ou une baisse des taux impliquerait respectivement une hausse ou une baisse du capital constitutif, qui correspond à une somme actualisée de prestations futures.

Le législateur prévoit que la victime ne devrait pas supporter la charge de l'inflation économique pour les sinistres qui seront versés en rentes. Ainsi, le montant de la rente étant revalorisé chaque année (tant que l'assurée est en vie), cette revalorisation était prise en charge par le Fonds de Garantie des Assurances Obligatoires de dommages (FGAO). Depuis le 1<sup>er</sup> Janvier 2013, elle est à la charge des assureurs : ils doivent donc les inclure dans le calcul de leurs provisions. Ces sinistres sont revalorisés à un taux d'inflation de 2%. Le modèle interne calcule cette revalorisation de façon agrégée et simpliste : les résultats sont donc moins précis. Il est important de bien modéliser cette revalorisation car non seulement elle fait augmenter les comptes mais aussi, elle n'est pas prise en charge par les réassureurs.

Enfin, dans le but de protéger leurs fonds propres et donc d'immobiliser moins de capital, les assureurs ont très souvent recours à la réassurance. D'une manière simpliste, la réassurance se définit comme «l'assurance de l'assurance». Cependant, pour évaluer la prime de réassurance ainsi que le SCR de primes net de réassurance, le modèle interne applique les traités sur les charges sinistres dont le taux d'actualisation est différent de celui réellement utilisé par les réassureurs.

---

1. suite de versements périodiques versés à la victime tant qu'il est en vie.

---

**Le présent mémoire a pour objectif de proposer une nouvelle approche de modélisation individuelle des sinistres corporels Auto graves qui permettrait de challenger et d'apporter des corrections aux limites du modèle actuel.**

La problématique du mémoire étant énoncée, les travaux s'étaleront sur quatre grandes parties. La première qui regroupe le périmètre de notre étude, les formes d'indemnisation des sinistres Auto corporels et le principe de gestion de ces sinistres au sein de Groupama. Dans la deuxième partie, il s'agira de présenter l'approche de modélisation des sinistres et la méthode de calcul du SCR de primes. La partie 3 est réservée à l'analyse de résultats : comparaison des résultats du nouveau modèle et ceux du modèle actuel, l'application de la réassurance, l'impact du FGAO sur les rentes futures et les tests de sensibilité. La partie 4 quant à elle permettra de calculer les risques vies auxquelles sont associées les rentes potentielles futures et in fine, déduire le SCR vie.

Enfin, nous verrons que notre modélisation présente quelques limites et nous proposerons des pistes d'améliorations pour les prochaines années.

---

Première partie

**Enjeux de modélisation pour les  
assureurs**

---

# 1. Périmètre de l'étude

Les modèles développés dans ce rapport s'inscrivent dans le cadre de la réforme réglementaire européenne Solvabilité 2. Cette partie vise d'une part à présenter brièvement les enjeux et les exigences portés par cette directive, ainsi que leur traduction en termes de calcul des provisions techniques et de gestion des risques. Elle présente aussi les bases de l'assurance automobile.

## 1.1 Contexte

Au regard des risques qui entourent l'activité d'assurance, une supervision de cette activité via des règles juridiques et comptables est indispensable pour préserver les intérêts des assurés. Les enseignements tirés lors des diverses crises économiques ont encouragé les autorités de contrôle européennes à revoir les directives européennes mises en place afin d'inciter les compagnies d'assurance et de réassurance à mieux prendre en compte les risques auxquels elles doivent faire face dans leur calcul de solvabilité.

La crise de 2008 a révélé de nombreuses défaillances dans le contrôle et la réglementation des institutions financières. Ce contexte particulier a fait naître au niveau européen la norme Solvabilité 2. Cette norme vient apporter des réponses axées sur 3 piliers, pour une gestion saine, prudente et efficace de l'activité des organismes d'assurance et une harmonisation des pratiques au niveau européen.

En matière de gouvernance, Solvabilité 2 définit une approche visant à responsabiliser les acteurs. En effet, elle définit un cadre réglementaire dont les organismes doivent s'approprier et l'ajuster en respectant des principes de proportionnalité et de prudence. Ce cadre met en avant une structure organisationnelle, constituée de personnes compétentes et honorables, appuyées par des responsables de fonctions dites **clés**. Les régulateurs et États membres sont également mis à contribution dans le contrôle et le soutien des organismes d'assurance.

La réglementation des assurances au niveau européen vise deux objectifs : protéger l'activité en renforçant les fonds propres (qui doivent prendre en compte les risques encourus par l'assureur) afin d'améliorer la compétitivité de l'industrie de l'assurance et de favoriser l'émergence de sociétés multinationales.

## 1.2 Solvabilité 2 et gestion des risques non-vie

L'assurance est soumise à un régime prudentiel entré en vigueur au 1er janvier 2016 issu de la directive européenne 2009/138/CE appelée **Solvabilité 2** remplaçant ainsi les directives des années 1970 : **Solvabilité 1**. Cette dernière bien qu'ayant été révisée en 2002, ne semblait plus être adaptée aux enjeux économiques, financiers et juridiques de l'assurance. Elle a montré quelques faiblesses à l'occasion de la crise financière de 2008, notamment des insuffisances qualitatives et quantitatives dans l'évaluation de la solvabilité des compagnies d'assurance, un manque de sensibilité aux risques (par exemple, les nouveaux risques), un manque d'harmonisation des normes et des pratiques au niveau européen. La commission européenne a donc fait le choix de réformer le «système de contrôle de la solvabilité des assureurs européens» : la directive-cadre de Solvabilité 2 a été adoptée.

Celle-ci vient renforcer les directives de solvabilité 1 et modifier en profondeur la réglementation. Elle a pour enjeux d'harmoniser et de soutenir le marché de l'assurance européen, en garantissant la solvabilité des assureurs, c'est-à-dire leur aptitude de faire face à leurs engagements envers les assurés. Elle introduit une approche nouvelle par rapport à la Directive Solvabilité 1, sur la manière de valoriser le bilan (valorisation en valeur de marché) et la mise en place d'un système global de gestion des risques.

La directive Solvabilité 2 s'applique à tous les assureurs et réassureurs de l'Union Européenne, y compris les entreprises dérivées, à quelques exceptions près. Certaines compagnies d'assurance de plus petite taille ne relèvent pas du champ d'application de la directive, mais pourront néanmoins demander une autorisation au titre de la directive Solvabilité 2.

---

### 1.2.1 Les objectifs de Solvabilité 2

Solvabilité 2 est une directive de l'Union Européenne dont l'objectif premier est de permettre l'harmonisation des règles européennes de solvabilité s'appliquant au secteur de l'assurance en créant un cadre prudentiel unique au sein de l'Union Européenne. Le but étant l'harmonisation des principes et des pratiques de contrôles, pour rendre les assureurs et réassureurs comparables au niveau international (concurrence équitable).

Cette ambition est complétée par l'objectif de renforcer la protection des assurés. Le durcissement des règles en termes de solvabilité doit ainsi améliorer la capacité des assureurs à honorer leurs engagements vis-à-vis de leurs clients. Le nouveau cadre prudentiel défini par la réforme se veut mieux adapté aux risques réels pris par les sociétés d'assurance à l'actif et au passif de leur bilan.

Solvabilité 2 vise également à introduire la gestion des risques au coeur du pilotage stratégique des compagnies d'assurance en opérant un changement de paradigme en termes de gouvernance des risques. La réforme prône ainsi une vision globale de la solvabilité de la compagnie, une évaluation économique du bilan en cohérence avec les marchés financiers ou encore une approche intégrée des différents risques engagés. Elle permet de prendre en compte les risques inhérents à l'activité de l'assureur.

Cette réforme vient créer une notion de groupe prudentiel mieux contrôlé et plus réglementé. Les exigences de fonds propres se retrouvent plus élevées et proportionnées au risque.

Enfin, la nouvelle directive européenne a pour but d'améliorer la cohérence entre les réglementations impactant les secteurs de l'assurance et de la banque.

Plusieurs acteurs interviennent dans la réforme Solvabilité 2 :

- La Commission Européenne : elle a un rôle législatif de rédaction et pilotage de la directive en collaboration avec les états membres
- EIOPA : l'European Insurance and Occupational Pensions Authority, anciennement le CEIOPS (Committee of European Insurance and Occupational Pensions Supervisors), est un comité regroupant les autorités de contrôle des états membres. L'EIOPA a mené des études d'impacts afin de mesurer les conséquences des nouveautés sur le marché assurantiel.
- ACPR : l'Autorité de Contrôle Prudentiel et de Résolutions a pour objectif de vérifier la bonne application de la nouvelle réglementation par les organismes d'assurance en France. Elle a le pouvoir de sanction en cas de non-respect de ces règles.

La directive européenne a apporté de nombreux changements et dont les principaux sont :

- Les exigences quantitatives du pilier 1 s'appliquent aux groupes de la même manière que pour les entités individuelles ;
- Un système de gouvernance renforcé ;
- Mise en place de fonctions clés et exigences de compétences et d'honorabilité ;
- Organisation d'un contrôle interne niveau groupe ;
- Exigences au niveau de la transmission des informations les plus importantes.

### 1.2.2 Les trois piliers

La réforme Solvabilité 2 s'organise en trois piliers complémentaires, chacun ayant une exigence bien précise : exigences qualitatives, exigences quantitatives et exigences d'information envers les autorités. Nous allons les expliciter dans les lignes suivantes. Néanmoins, nous pouvons les résumer par la figure 1.1.1 suivante :

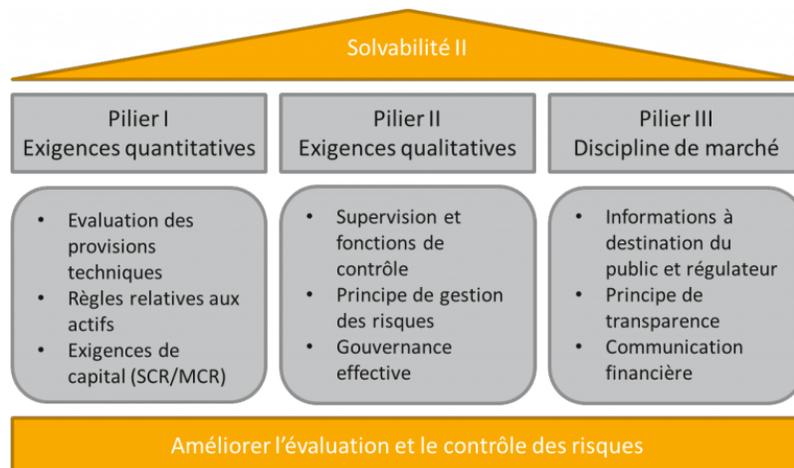


FIGURE 1.1.1: Les trois piliers de la directive Solvabilité 2

### Pilier 1 : Les règles de calcul des risques et de la solvabilité

Le pilier 1 concerne les exigences quantitatives en capital; il définit les normes de calcul des provisions techniques, des fonds propres et l'évaluation des actifs et des passifs d'assurance et les règles d'investissement. De nouvelles normes quantitatives d'évaluation des provisions techniques exigent en effet de décomposer celles-ci en une meilleure estimation des provisions, excluant toute hypothèse prudente, et en une marge pour risque. S'ajoutent à ces normes, le calcul des capitaux économiques réglementaires : SCR et MCR :

- Le **MCR** (*Minimum Capital Requirement*) est le niveau minimal de fonds propres qu'un assureur doit posséder, en dessous duquel l'autorité de contrôle interviendra de manière systématique, et pourra décider de placer sous tutelle l'assureur, voire de lui retirer son agrément
- Le **SCR** (*Solvency Capital Requirement*) est le niveau de fonds propres nécessaires pour absorber un choc provoqué par un risque majeur. Il est défini comme le quantile à 99,5% des fonds propres. Il correspond au capital économique nécessaire pour ne pas être en ruine à horizon 1 an avec une probabilité de 99.5% : ce qui correspond à une ruine tous les 200 ans. Si l'entité ne couvre pas par ses fonds propres ce SCR, le régulateur exigera un plan de redressement pour que l'entité réajuste son taux de couverture.

Ces deux quantités sont mesurées à horizon d'un an : Le SCR est non seulement plus exigeant que le MCR, mais aussi est plus complexe à calculer.

Le SCR deviendra l'indicateur principal des autorités de contrôle. Il a comme spécificité d'être fondé sur l'exposition aux risques, en incorporant tous les risques liés à l'activité de la compagnie, c'est-à-dire principalement, le risque de souscription non-vie, le risque de crédit, le risque opérationnel, le risque de liquidité et le risque de marché.

Le pilier 1 présente les règles de valorisation du bilan (actifs, passifs et fonds propres) en valeur de marché, pour construire un bilan en vision économique. La comparaison des comptes réglementaires vus sous les régimes Solvabilité 1 et Solvabilité 2, illustrée en figure 1.1.2, résume les différentes exigences quantitatives auxquelles le pilier 1 soumet les sociétés d'assurance.

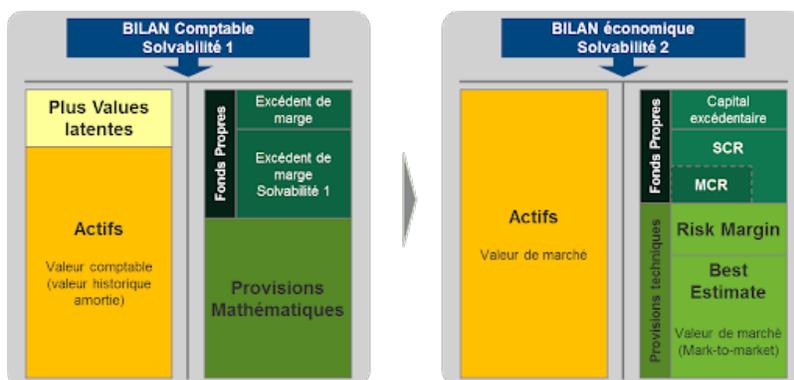


FIGURE 1.1.2: Comparaison des bilans de Solvabilité 1 et 2

Ce pilier définit également deux grandes catégories d'évaluation de la solvabilité : l'approche par la formule standard ou l'approche modèle interne (confère sous section ??).

### Pilier 2 : Contrôles internes, gestion des risques et principes de gouvernance

Le second pilier regroupe les exigences qualitatives, c'est-à-dire les règles de gouvernance et de gestion interne des risques, dont l'évaluation propre des risques de la solvabilité. Il a pour objectifs d'optimiser l'environnement de contrôle et de gestion des risques et d'améliorer la gouvernance à court et moyen terme. Elle définit également comment l'autorité de contrôle devrait exercer son pouvoir.

Il permet d'assurer une meilleure sécurité pour les assurés européens, d'harmoniser les pratiques en Europe, et de mettre la réglementation et le contrôle des entreprises d'assurance aux normes par rapport aux autres réglementations financières ainsi qu'aux progrès en matière de gestion des risques.

Il définit le cadre du contrôle prudentiel effectué par l'autorité de contrôle : elle aura le pouvoir de contrôler la qualité des données et des procédures d'estimation, ainsi que des systèmes mis en place pour mesurer les risques au cas où ils se produisaient.

Concernant le système de gouvernance, il comprend des obligations en matière de contrôle interne, de communication interne et externe, des règles de conformité et insiste sur l'implication du conseil d'administration dans la mise en oeuvre de la gestion de risques. La directive définit la mise en oeuvre d'une structure organisationnelle transparente, adéquate, avec séparation claire et appropriée des responsabilités, ainsi qu'un dispositif efficace de transmission des informations. Les politiques en matière de gestion des risques, contrôle interne, audit interne et sous-traitance doivent être écrites, avec une révision au moins annuelle.

Le système de gouvernance mis en place devrait comprendre à minima :

- Un organe d'administration (AMSB en anglais). Il a pour rôle de déterminer les objectifs stratégiques de l'entreprise et de veiller à leur mise en oeuvre. Les états membres veillent à ce qu'il assume la responsabilité finale du respect des dispositions législatives, réglementaires et administratives adoptées par la directive.
- Des Dirigeants effectifs : Conformément à l'article L. 356-2, alinéa 2 et 3 du Code des Assurances, chaque entreprise devrait disposer à minima de deux dirigeants effectifs (principe des « quatre yeux »), nommés par l'AMSB et qui respectent les conditions d'honorabilité et de compétence. Ceux-ci sont désignés en fonction de leurs compétences, de leurs expériences et de leurs qualités estimées nécessaires au poste ;
- Des exigences de compétence et d'honorabilité : ils s'appliquent à tous les dirigeants effectifs ou aux personnes qui occupent des fonctions clés. Ceci dans le but de garantir que toutes les personnes exerçant des fonctions importantes dans l'entreprise sont dûment qualifiées.

- 
- Du contrôle interne : Il permet de s'assurer que la compagnie est en mesure de calculer et maîtriser ses risques (financiers, techniques, opérationnels) auxquels elle est confrontée et qu'elle est bien capitalisée. Ce contrôle passe par la mise en oeuvre de :

- l'ORSA (*Own Risk Solvency Assessment*), qui est un ensemble de processus et de procédures permettant le suivi continu des risques de la compagnie et l'évaluation des fonds propres nécessaires au maintien de sa solvabilité. Les assureurs sont tenus de s'assurer de leur solvabilité selon leur propre vision des risques.

- Fonctions clés : Il en existe quatre :

La fonction de gestion des risques : Ce système de gestion devra permettre d'identifier, mesurer, contrôler, gérer et déclarer les risques, aux niveaux individuel et agrégé, auxquels les entreprises sont ou pourraient être exposées ainsi que les interdépendances entre ces risques ((démarche ERM : *Entreprise Risk management*). Il permet également de décrire le lien entre l'ORSA, les exigences réglementaires de capital et les limites de tolérance au risque de l'entreprise. Il couvre les domaines de la souscription et provisionnement, de la gestion actif-passif, des investissements, de la gestion de risque de liquidité et de concentration, de la gestion de risque opérationnel, la réassurance et les autres techniques d'atténuation du risque. Pour les entreprises utilisant un modèle interne partiel ou intégral (préalablement approuvé par l'ACPR), la fonction gestion de risques recouvre les tâches supplémentaires suivantes : conception et mise en oeuvre du modèle interne ; test et validation du modèle interne ; suivi documentaire du modèle interne et de toute modification qui lui est apportée ;

La fonction de vérification de la conformité : Elle a pour but de maintenir une veille et une interprétation de la transposition des normes et réglementations dans l'organisation. Elle permet également d'évaluer l'impact possible de tout changement de l'environnement juridique sur les opérations de l'organisme.

La fonction actuarielle : Elle définit des compétences et liste des rôles et fonctions des personnes exerçant la fonction d'actuaire. La responsabilité de cette fonction est exercée par des personnes qui ont une connaissance et une pratique des statistiques, mathématiques actuarielles et financières. En fonction de la nature et de la complexité des risques inhérents à l'activité, cette personne devra démontrer une expérience pertinente et une bonne compétence notamment sur les normes professionnelles et autres normes applicables.

La fonction d'audit interne : Elle doit être efficace, régulière et indépendante des autres fonctions opérationnelles. Elle permet d'évaluer l'adéquation et l'efficacité du dispositif de contrôle interne

### **Pilier 3 : Communication - Reporting**

Les entreprises soumises à la directive solvabilité doivent respecter la discipline de marché : les assureurs doivent mettre à la disposition du grand public (et du superviseur) les informations précises et détaillées sur leur activité. L'ensemble des assureurs européens doit renseigner les mêmes types de modèles, permettant une meilleure transparence du marché et une comparabilité simplifiée entre les acteurs. De plus, le reporting permet de faire le parallèle entre les différentes gouvernances, de confronter la qualité des assureurs et leurs pratiques. Les assureurs et réassureurs auront donc à fournir les informations clés (vérifiables) nécessaires à la détermination de leur exigence de capital.

Les assureurs et réassureurs doivent transmettre les documents suivants :

- Des rapports « narratifs », qui donnent une vision descriptive de la politique prudentielle mise en place par l'assureur. Deux rapports sont à produire : le SFCR (Solvency Financial and Condition Report) destiné au public, et faisant état des lieux de la solvabilité ; et le RSR (Regulator Supervisory Report) , destiné exclusivement au superviseur ;

- 
- Des reportings quantitatifs, appelés QRT(Quantitative Report Template). Il s'agit de tableaux de bord dont les données ont été définies par l'EIOPA et qui couvrent les principaux domaines d'activité d'un assureur.

Ces rapports sont structurés de manière à fournir un nombre d'informations substantielles relativement aux risques de l'entreprise et à la manière de les gérer, aux efforts menés pour garantir le pilotage de l'organisme et sa solvabilité, ainsi qu'un descriptif de son fonctionnement et de son organisation (composition de l'actionnariat, composition du conseil d'administration, des comités spécifiques, la direction générale et les principaux cadres de direction, stratégie de l'entreprise, sa politique sociale, la politique de rémunération. . . .).

### 1.2.3 Zoom sur les provisions techniques

Le premier pilier de Solvabilité 2 marque un changement profond au sein des directions techniques et de gestion des risques des sociétés d'assurance en définissant un nouveau cadre de calcul des provisions techniques et d'évaluation des besoins en capital. Les éléments du bilan sont désormais évalués en valeur économique et non comptable : en effet, les actifs et les passifs doivent être évalués à leur juste valeur, définie de la manière suivante :<sup>2</sup>"Montant pour lequel un actif ou un passif peut être échangé entre deux parties informées et consentantes, dans des conditions normales de concurrence." Les nouvelles normes quantitatives introduites imposent tout d'abord la décomposition des provisions techniques en deux quantités :

- une Meilleure Estimation (ou Best Estimate) ;
- une Marge pour Risque.

#### La provision Best Estimates

Dans le cadre de l'arrêté des comptes, chaque entité réalise une estimation de ses provisions en vision «**best estimate** » à partir des triangles bruts de réassurance. Les méthodes mises en œuvre sont des méthodes actuarielles classiques (Chain-Ladder avec éventuellement retraitement de coefficients de passage, Bornhuetter Fergusson, dire d'expert, . . . ) appliquées aux triangles de paiements ou de charges selon le type de risque.

Le **Best Estimates**(BE) correspond à la meilleure estimation des provisions à faire pour régler un sinistre déjà survenu. En Solvabilité 2, les provisions « Best Estimate » correspondent à l'actualisation de tous les flux probables futurs (cotisations, prestations, frais, fiscalité, . . . ) actualisés avec une courbe des taux sans risques. Ces provisions Best Estimate sont non escomptées et ne contiennent pas les frais de gestion associés.

Le calcul de la meilleure estimation est fondé sur des informations actualisées et crédibles ainsi que sur des hypothèses réalistes. Ce calcul fait appel à des méthodes actuarielles et statistiques adéquates, applicables et pertinentes. L'objectif de la segmentation des obligations de l'assureur est d'obtenir une évaluation de provisions techniques la plus précise possible. Les lignes d'activités choisies doivent représenter au mieux les caractéristiques des risques sous-jacents. La directive exige une segmentation au minimum par ligne d'activité.

Les provisions Best Estimate doivent être calculés brut de réassurance et doit inscrit au passif du bilan. Sa formule est la suivante :

$$BE = \mathbb{E} \left[ \frac{\sum_{k=1}^N CF_k}{(1 + i_k)^k} \right] \quad (1.1.1)$$

avec :

- CF : cash flows
- $k = \{1, \dots, N\}$  avec N le nombre d'années de développement

---

2. Source : cours de Réglementation de l'Assurance, de Franck Levallois

- 
- $i$  est le taux d'actualisation : par défaut le taux sans risque fourni par la Commission Européenne (courbe des taux swap).

Les flux de BE connaissent des incertitudes provenant de : l'incertitude sur le comportement des assurés ; la variation des indices utilisés pour déterminer les montants de sinistres ; les variations des temps nécessaires pour payer un sinistre ; ...

### Le Risk Margin

Encore appelé "marge pour risque", le **risk margin** est la partie des provisions techniques additionnelle aux provisions BE. Il est défini comme le montant de provisions supplémentaire au BE, qu'un assureur tiers réclamerait afin de reprendre à sa charge le passif d'assurance de l'entreprise. Il représente un niveau de prudence supplémentaire dans le cas où les provisions seraient cédées. Le calcul du risk margin se fait de manière à garantir que la valeur des provisions techniques corresponde au montant que les entreprises d'assurance et de réassurance attendraient pour reprendre et honorer les engagements d'assurance et de réassurance.

Ajouté au Best estimate, le risk margin représente la valeur de marché du passif d'une compagnie d'assurance. Il est calculé selon une approche de coût du capital : cette méthode permet de déterminer le coût que représente la mobilisation d'un montant de fonds propres éligibles au SCR nécessaire pour faire face aux engagements d'assurance et de réassurance sur toute la durée de vie de ceux-ci.

Au sein d'une entreprise, le risk margin et le best estimate «doivent être évalués séparément».

Les provisions techniques doivent suivre plusieurs règles :

- couvrir les engagements de l'assureur vis-à-vis des assurés
- être évaluées selon le principe de l'Exit value<sup>3</sup>
- actualiser le montant des cash flows<sup>4</sup> futurs afin que ce soit le montant des provisions
- appuyer les hypothèses sur les informations apportées par les marchés financiers
- être calculées de manière «prudente, fiable, et objective»

La directive Solvabilité 2 laisse aux sociétés d'assurance le choix entre trois approches pour évaluer leur SCR :

- La formule standard, définie et calibrée par la réforme. Elle est identique pour l'ensemble des assureurs, qui se fonde sur des méthodes de calcul simplifiées définies dans les études quantitatives d'impact ;
- Le modèle interne complet, fondé sur la structure de risque spécifique à l'entreprise. Il est soumis à l'approbation de l'autorité de contrôle, qui prend précisément en considération les données propres à la compagnie et doit constituer un outil d'aide à la décision et de pilotage stratégique des risques ;
- Le modèle interne partiel (MIP), combinant les deux approches précédentes selon les risques, ou les branches de l'entreprise.

## 1.3 Approche modèle interne et risques associés

Approche alternative à la formule standard, la mise en place d'un modèle interne, intégral ou partiel, est un choix que font la plupart des compagnies d'assurance influentes sur le marché, Groupama en particulier, afin d'adopter une gestion plus précise de leurs risques. Il est naturellement plus complexe à mettre en œuvre que la formule standard lors de l'évaluation des capitaux économiques requis par Solvabilité 2.

---

3. valeur que l'assureur devrait payer pour transférer les engagements à un autre assureur.

4. ensemble des flux de liquidités générés par les activités d'une société. C'est un bon moyen d'appréhender la solvabilité et la pérennité d'une entreprise.

---

### 1.3.1 Sa structure

Un modèle interne se caractérise par un ensemble de modèles mathématiques, différents d'un risque à l'autre, s'appuyant sur des données historiques, des paramètres extérieures, des informations du marché, des avis d'experts ou encore sur l'expérience métier de l'assureur. Les données dont dispose la compagnie permettent de calibrer les paramètres des lois que le modèle suppose suivies par les différentes variables aléatoires d'intérêt, dans le but d'effectuer par la suite un grand nombre de simulations.

Un modèle interne se fonde ainsi sur des méthodes stochastiques, permettant la projection des profits et pertes futurs engendrés par les différents risques. Il intègre alors la volatilité des phénomènes sous-jacents dans les résultats qu'il produit et permet ainsi une meilleure quantification de leur incertitude. En pratique, pour un risque donné, un grand nombre de scénarios économiques sont générés afin d'obtenir une distribution de la perte future probable à horizon 1 an. Ces scénarios prennent en compte les facteurs de risques propres à la compagnie.

Le modèle interne de GMA modélise deux business units : caisses métropolitaines agrégées (constituées de 9 caisses régionales) et GAN Assurances. Cette modélisation peut se faire soit selon un périmètre solo (modélisation individuelle de chacune des caisses régionales et de GAN Assurances) ou un périmètre Groupe (modélisation séparée de chacune des business units mais groupée pour l'ensemble des caisses régionales). Dans le cadre de ce rapport, notre modélisation sera faite d'un point de vu Groupe, uniquement pour les caisses.

### 1.3.2 Les objectifs d'un modèle interne

L'objectif premier d'une compagnie d'assurance choisissant de gérer ses risques à l'aide d'un modèle interne est le calcul d'un **montant de SCR de souscription non-vie** qui soit le plus adapté possible aux risques spécifiques sur lesquels elle est engagée. Le gain significatif de précision sur le calcul du SCR que peut apporter un modèle interne permet éventuellement de bénéficier d'une diminution du montant de capital obtenu. La réduction de ce capital permet ainsi à la compagnie d'immobiliser une quantité moindre de fonds propres et donc d'être en mesure de distribuer davantage de dividendes à ses actionnaires. Une amélioration du ratio de solvabilité de la compagnie, est également recherchée par l'adoption d'un modèle interne.

Une approche modèle interne permet par ailleurs aux compagnies d'obtenir une vision plus réaliste de leur profil de risque et de fournir aux investisseurs des indicateurs économiques se voulant davantage en adéquation avec leur santé financière. La prise en compte plus précise des risques encourus au moyen de données propres à la compagnie se traduit en pratique par une meilleure compréhension des risques et de leurs éventuels impacts sur le résultat des différentes activités. Ainsi, un modèle interne contribue au pilotage stratégique des risques de la compagnie et permet ainsi une amélioration de sa politique de souscription et un gain significatif de la compétitivité des produits qu'elle commercialise.

Au sein de Groupama, les modèles développés pour le calcul de SCR sont effectués selon une approche modèle interne partiel. Ceux-ci couvrent le risque de souscription non vie et santé non SLT.

Ci-dessous, les modules de risques inclus dans le MIP.

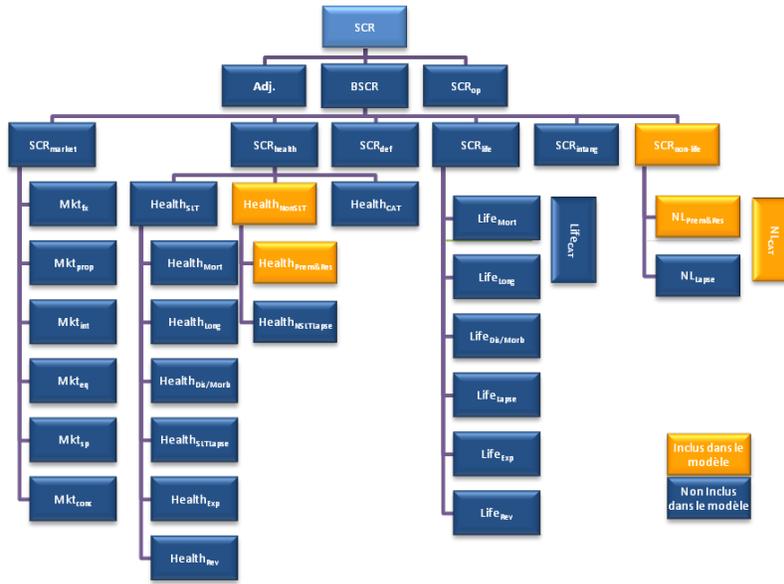


FIGURE 1.1.3: Modules de risques inclus dans le MIP groupe

Dans le cadre de ce rapport, nous intéresserons à certains sous-modules du risque vie (risque de longévité, risque de révision, risque de frais de gestion) et au risque de souscription non-vie (particulièrement au risque de primes) suivants :

- Risque de primes ou de tarification : il résulte de la mauvaise tarification pour les sinistres futurs à horizon 1 an ;
- Risque de longévité : il résulte de la sous-estimation de la durée de vie des assurés, estimée à partir d'une table de mortalité bien précise ;
- Risque de révision : il résulte d'un changement réglementaire ou de l'état de santé des assurés ;
- Risques de frais de gestion : il résulte de la variation des frais de gestion des contrats d'assurance ou de réassurance.

### 1.3.3 Quantification du risque de souscription non vie

Le risque de souscription non-vie reflète le risque découlant des engagements d'assurance non vie, compte tenu des sinistres couverts et des méthodes appliquées dans l'exercice. Il prend également en compte l'incertitude résultant des résultats des compagnies d'assurance et de réassurance dans le cadre de leurs engagements vis-à-vis de leurs assurés, ainsi que des nouveaux contrats souscrits qui sont attendus l'année suivante.

Ce risque est associé au **SCR non vie**, défini comme étant l'opposée de la *Value-at-Risk* (VaR) à 0,5% de la variation de situation nette (notée SN) à 1 an, la situation nette étant définie comme la différence entre les actifs et les passifs en valeur économique, avec une nouvelle année de souscription. La mesure de risque choisie est ainsi la même que celle de la Formule Standard, au même niveau de risque (probabilité de ruine de 99.5%) sur le même horizon (1 an).

Par définition, la Value at Risk de niveau  $\alpha$  de la distribution de pertes  $L$  est le quantile d'ordre  $\alpha$  de  $L$ . Pour  $\alpha = 99,5\%$ , nous avons donc :

$$\text{VaR}_{99,5\%}(L) = \inf\{x \in \mathbb{R}, P(L > x) \leq 0.005\} \quad \text{et} \quad P(L \leq \text{VaR}_{99,5\%}(L)) = 0.995 \quad (1.1.2)$$

Ci-dessous, une illustration graphique de la VaR à 99.5%.

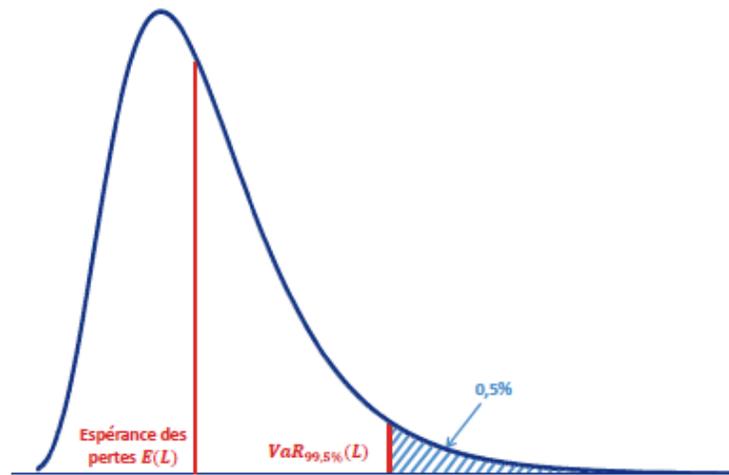


FIGURE 1.1.4: Illustration de la Value At Risk

La VaR n'est pas une mesure de risque « cohérente » au sens des travaux de Artzner et al<sup>5</sup>. En particulier, elle n'est pas sous additive. La sous-additivité permet la diversification du risque, c'est-à-dire que la mesure globale de plusieurs risques combinés est inférieure à la somme des mesures individuelles de ces risques. La diversification des activités est souvent recherchée par les assureurs afin de mutualiser leurs risques. La Var est une mesure de risque souvent retenue par les assureurs, notamment pour les raisons suivantes :

- Elle est homogène en termes de mesure de risque utilisée pour le MIP (modèle interne partiel) et le hors MIP calculé en Formule Standard. Ceci permet l'agrégation du MIP et du hors MIP selon la même mesure de risque.
- Bien que cohérente, sa « concurrente », la Tail Var<sup>6</sup> est souvent plus difficile à expliquer au sein de l'entreprise.

## 1.4 Les engagements en Automobile

En France, toute personne, ayant un véhicule terrestre à moteur est obligée de le faire assurer : il s'agit d'un contrat d'assurance automobile. Il a pour principal objectif de couvrir les dommages que le véhicule pourra occasionner à autrui. Ce contrat comporte plusieurs garanties dont celle obligatoire (d'après l'article L211-1 du code des assurances) est la garantie responsabilité civile automobile (RC Auto). C'est même la seule garantie indispensable pour être autorisé à circuler sur les voies publiques et c'est sur elle que sont bâtis les contrats dits « au tiers ». En cas d'accident de la route, elle couvre les conséquences financières dont l'assuré pourrait avoir la charge en cas de dommages matériels et/ou corporels subis par des tiers (passager, autre conducteur, piéton, ...).

Il existe deux types de sinistres : les sinistres **corporels** et les sinistres **matériels**. Contrairement aux dommages matériels, où une entreprise peut fixer un montant maximal d'indemnisation, les dommages corporels n'ont jamais de limite d'indemnisation. Le coût du dommage dépend de plusieurs facteurs. Ainsi, après un accident de la route, l'assuré doit bien entendu déclarer le sinistre à son assureur dans un délai fixé au contrat. S'il est tenu responsable de l'accident, sa responsabilité civile devra réparer le/les dommage(s) causé à/aux la/les victimes : dommage du véhicule ainsi que dommage corporel. Dans le cadre ce mémoire, nous nous intéresserons uniquement aux sinistres corporels.

5. Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent Measures of Risk. *Mathematical Finance*, 9(3) :203–228, July 1999.

6. Mesure de risque qui permet de quantifier la valeur attendue de la perte étant donné qu'un événement situé en dehors d'un niveau de probabilité donné s'est produit.

---

### 1.4.1 Les sinistres corporels Auto

Par définition, un accident corporel de la circulation désigne tout accident impliquant au moins un véhicule routier en mouvement, survenant sur une voie ouverte à la circulation publique, et dans lequel au moins une personne est atteinte physiquement ou moralement : blessures, troubles de mémoires... , entraînant ainsi une incapacité, invalidité ou même la mort. La personne jugée responsable de cet accident devra donc indemniser le/les victime(s).

Selon la nature des dommages corporels, l'assureur informe ensuite l'assuré de ses droits en matière d'indemnisation (vers la victime) et du suivi du dossier. L'offre d'indemnisation peut être définitive si l'état de santé de la victime est stable, ou prévisionnelle s'il est amené à évoluer. Le plus souvent, il s'agit d'un forfait d'ouverture correspondant à un coût moyen ultime ou, si le dossier est déjà bien étayé par un rapport médical, une évaluation plus détaillée. Le procès verbal, établi par la police ou la gendarmerie permet de déterminer les circonstances de l'accident, les responsabilités des acteurs et donc le taux retenu pour l'indemnisation. Le dénouement du dossier peut se faire soit par transaction, c'est-à-dire avec un accord amiable avec la partie adverse, soit par voie judiciaire, c'est-à-dire que le juge détermine les indemnités à verser.

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

### 1.4.2 Les formes d'indemnisation

Sur la base des l'expertise médicale, l'assureur du conducteur fait une offre d'indemnisation à la victime, de sorte que le montant permette de couvrir intégralement les dommages qu'elle a subis. Celle-ci peut se faire de deux manières : par versement d'un **capital** sur la base d'un barème de capitalisation ou autre fondement à l'appréciation des parties ou du juge ; ou part versement d'une **rente** à la victime, ou à ses ayants droit, jusqu'à un terme prévu. Une rente est une suite de versements périodiques versés à la victime tant qu'il est en vie. Pour un individu d'âge  $x$ , sa probabilité de vie noté  ${}_t p_x$  est définie comme suit :

$${}_t p_x = P(T > x + t | T > x) \quad (1.1.3)$$

Où  $T$  est une variable aléatoire représentant la durée de vie à la naissance.

Chacune de ces formes d'indemnisation possède un avantage et un inconvénient tant pour l'assureur que pour la victime :

<b>Rentes : Avantages</b> Réduction du coût du sinistre si la victime décède prématurément (assureur) Protection viagère de la victime (victime)	<b>Capital : Avantages</b> Liberté de gestion et de placement (victime) Non imposable (victime)
<b>Rentes : Inconvénients</b> Rente imposable (Victime) Mettre en place un suivi permanent de la rente (assureur) Dossier ouvert sur une longue durée (assureur) Paiement de la revalorisation de la rente pour les sinistres survenus après 2013 (assureur)	<b>Capital : Inconvénients</b> Pas de réduction du coût du capital en cas de décès prématuré de la victime (assureur) Négociations sur le choix du barème de capitalisation (assureur) Incertitude sur le barème de capitalisation (victime)

Le versement d'une rente est un mode d'indemnisation utilisé pour les sinistres importants et pour les victimes lourdement handicapées. Il peut soit s'étaler sur toute la vie de la victime, appelés rentes viagères

---

ou pendant une durée limitée, rentes temporaires. La conversion d'une rente en capital est possible et dépend la table de mortalité déterminée par voie réglementaire fondée sur un taux d'intérêt.

Il existe deux types d'indemnisation en rentes :

- Indemnisation **contractuelle** : est celle qui dépend des modalités qui se trouvent dans le contrat.
- Indemnisation **de droit commun** : est celle qui serait obtenue s'il fallait passer devant un juge. Il repose sur un principe de réparation intégrale du préjudice causé au tiers. Les dossiers gérés en droit commun au sein de GROUPAMA sont la RC Auto et la RC non Auto. Nous nous attarderons sur cette forme d'indemnisation et particulièrement sur la RC Auto dans la suite de notre étude (confère loi Badinter : section 2.1).

Notre étude porte essentiellement sur les sinistres indemnisés en rentes pour les raisons suivantes :

- Le montant de l'indemnisation étant élevé, le recours à la réassurance est possible ;
- La revalorisation des rentes n'étant plus à la charge du FGAO depuis Janvier 2013, les assureurs doivent désormais inclure cette revalorisation dans le calcul de leur provisions.

---

## 2. Indemnisation des sinistres corporels Automobile

Le montant de l'indemnisation des sinistres corporels pouvant être important, de nombreux litiges existaient entre les victimes d'accidents et les organismes payeurs. Ces litiges ont conduit le législateur à édicter progressivement des règles et des procédures spécifiques visant particulièrement la défense des intérêts de la victime : réparation intégrale des préjudices visibles et invisibles, économiques et non économiques.

### 2.1 La loi Badinter

La gestion des sinistres corporels Auto est fortement encadrée d'un point de vue législatif, par la loi n° 85-677 du 5 juillet 1985 dite "loi Badinter". C'est une loi qui crée un cadre d'indemnisation rapide et convenable aux victimes d'accidents de la circulation. Elle a pour l'objectif de protéger, de faciliter et d'accélérer la prise en charge de leur préjudice corporel. Cette loi repose sur un principe de **réparation intégrale du préjudice** causé à un tiers.

#### *Cadre d'application de la loi Badinter*

- Types d'accidents de la route : tout accident dans lequel est impliqué un véhicule terrestre à moteur (bus, moto, deux-roues, voiture, ...) en mouvement ou non ;
- Potentiels bénéficiaires : les piétons, cyclistes, passagers d'une voiture accidentée (sauf le conducteur), . Ils ont un droit d'indemnisation à 100% même s'ils sont en tort. Cependant, si la victime a volontairement causé son accident ou a commis une faute inexcusable ( intention suicidaire ou suicide par exemple), elle ne sera pas indemnisée. Il existe toute fois des victimes dits "super privilégiées". Il s'agit des victimes soit âgées de moins de 16 ans ou de plus de 70 ans, soit des personnes invalides à plus de 80% au moment de la survenance du sinistre. Celles-ci ne sont soumises à aucune condition d'indemnisation ;
- Conducteur du véhicule : s'il est une victime de l'accident, son droit à l'indemnisation dépend du degré de gravité de leur faute (article 4).La faute du conducteur victime s'apprécie par rapport à son seul comportement. Plus il est responsable de l'accident, moins sera son indemnisation. S'il n'a aucune responsabilité dans l'accident, il sera indemnisé à 100%. Cependant, s'il est jugé responsable, il peut être indemnisé s'il a souscrit à une garantie supplémentaire.

Cette loi est très contraignante dans le sens où elle oblige l'assureur du conducteur ayant causé le dommage à proposer une offre d'indemnisation à la victime dans un délai maximum de huit mois, à compter de la date de l'accident (conformément à l'article 12). La victime peut soit refuser l'offre, soit l'accepter et dispose d'un délai de rétractation de 15 jours.

### 2.2 Nomenclature DINTILHAC

<sup>7</sup>La loi du 21 décembre 2006 a imposé aux professionnels (juge, expert, avocat, assurance, sécurité sociale) le recours à une méthodologie pour déterminer les préjudices. Cette méthodologie doit mettre en avant les préjudices à indemniser. Ces préjudices sont énumérés par la nomenclature DINTILHAC . Elle a pour objectif de permettre aux victimes d'accidents corporels de lister l'étendue de ses préjudices afin de décider quelle suite indemnitaire choisir.

---

7. Source : jmp-avocat-indemnisation.fr

---

En effet, cette nomenclature a été élaborée en 2005 par un groupe de travail, présidé par Monsieur Jean-Pierre DINTILHAC, Président de la deuxième chambre civile de la Cour de Cassation préjudices corporels. Elle s'appuie sur la jurisprudence antérieure, en vue d'améliorer les conditions d'indemnisation et favoriser l'harmonisation des critères d'indemnisation" du préjudice corporel. L'idée est donc d'évaluer poste par poste le préjudice subi par une victime afin de lui proposer une meilleure indemnisation pour son dommage.<sup>8</sup> Les propositions contenues dans le rapport éponyme ne s'imposent pas : *il ne s'agit ni d'une loi, ni d'une norme réglementaire.*

#### *Détail de la nomenclature*

La nomenclature DINTILHAC distingue les trois cas suivants :

- Les préjudices de la victime directe et les préjudices de la victime indirecte ;
- Les préjudices patrimoniaux et les préjudices extrapatrimoniaux avant et après consolidation ;<sup>9</sup>
- Les préjudices temporaires et les préjudices permanents.

En annexe A1, un tableau de synthétisant les différents postes de préjudices de la nomenclature DINTILHAC.

## 2.3 Mise en place de l'indemnisation

A la suite d'un accident, la/les victime(s) doit/vent déclarer à leurs assureurs au plus tard dans les 5 jours ouvrés. Le conducteur du véhicule devra faire pareil et un dossier victime sera ouvert par le gestionnaire sinistres. L'assureur du conducteur devra se rapprocher de la victime pour avoir toutes ses informations de manière détaillée : dommages corporels, la mutuelle de la victime, . En parallèle, celui-ci rassemble les éléments pour vérifier les garanties mentionnées dans le contrat et évalue la part de responsabilité de l'assuré (sur la base du procès-verbal établi par les services de police ou de gendarmerie) afin de déterminer s'il prendra en charge l'indemnisation du sinistre ou pas.

Par la suite, une évaluation des dommages corporels est faite afin d'évaluer le montant de l'indemnisation. Ce coût est estimé soit sur la base d'un avis médical (examen des pièces médicales par le médecin de l'assureur pour les dommages légers), soit sur la base d'une expertise médicale (pour les dommages plus importants). Cette première provision est appelée le **capital constitutif** .

Il correspond à la somme totale actualisée des montants qui seront versés au bénéficiaire (appelé **arrérage**) pendant toute la durée de service de celle-ci. Ces versements peuvent être versés mensuellement, trimestriellement, semestriellement ou annuellement selon la décision judiciaire prise ou la transaction amiable.

Pour une victime d'âge  $x$ , le capital constitutif est calculé de la manière suivante :

$$Capital_{constitutif} = Arrerage \times \sum_{i=1}^T \frac{p_{x,x+i}}{(1 + taux_{actu})^i} = Arrerage \times E(x) \quad (1.2.1)$$

Avec  $T$  la durée de la rente en année,  $p_{x,x+i}$  la probabilité de survie de l'individu d'âge  $x$  dans  $i$  années,  $taux_{actu}$  le taux d'actualisation et  $E(x)$  l'euro de rente.

Les probabilités de survie sont calculées à partir d'une table de mortalité bien précise.

## 2.4 Les différents postes de préjudice indemnisés

Lorsqu'une victime d'un accident corporel a droit à une indemnisation de droit commun, elle doit être indemnisée de ses dommages corporels et préjudices économiques en résultant. Il s'agit donc du principe de

---

8. <https://www.cottineau.net/nomenclature-dintilhac/>

9. la consolidation étant la stabilisation de votre état.

---

réparation intégrale du préjudice. Les différents postes de préjudice qui peuvent être indemnisés sous forme de rente sont :

### **Assistance Tierce Personne (ATP)**

<sup>10</sup>Ces frais concernent les dépenses liées à l'assistance par tierce personne pour aider la victime à effectuer les actes de la vie quotidienne, préserver sa sécurité et contribuer à restaurer sa dignité et suppléer à sa perte d'autonomie. Il peut y avoir des heures de tierces personnes **actives** (personne de substitution totale ou partielle qui est chargée de réaliser un acte que la victime est dans l'incapacité totale ou partielle d'effectuer) et **passives** (personne de surveillance dont la présence est nécessaire pour assurer la sécurité d'une victime gravement handicapée aux fonctions vitales très lourdement déficitaires).

L'indemnisation est variable selon le nombre d'heures nécessaires et le type de tierce personne requise, l'aide d'une tierce personne spécialisée étant parfois plus coûteuse. Une même victime peut avoir besoin de plusieurs types d'aides au cours de la journée. Ainsi, une fois que le médecin expert aura déterminé les types de tierce personne et le nombre journalier d'heures nécessaires, ce dernier sera multiplié par le nombre de jours de travail et de congés payés de l'année et par le salaire annuel de l'aide, ce qui permettra d'obtenir le montant d'arréage annuel.

### **Perte de Gain Professionnel futur (PGPF)**

<sup>11</sup>Il s'agit ici d'indemniser la victime de la perte ou de la diminution de ses revenus consécutive à l'incapacité permanente à laquelle elle est désormais confrontée dans la sphère professionnelle à la suite du dommage. De plus, Il s'agit d'indemniser une invalidité spécifique partielle ou totale qui entraîne une perte ou une diminution directe de ses revenus professionnels futurs à compter de la date de consolidation. Cette perte ou diminution des gains professionnels peut provenir soit de la perte de son emploi par la victime, soit de l'obligation pour celle-ci d'exercer un emploi à temps partiel à la suite du dommage consolidé. En outre, concernant les jeunes victimes ne percevant pas à la date du dommage de gains professionnels, il conviendra de prendre en compte pour l'avenir la privation de ressources professionnelles engendrée par le dommage en se référant à une indemnisation par estimation. Ce poste est très souvent indemnisé sous forme de rente lorsque la victime est jeune.

### **Incidence Professionnelle (IP)**

<sup>12</sup> Ce poste est le complément du poste PGPF. Il a pour objet d'indemniser non la perte de revenus liée à l'invalidité permanente de la victime, mais les incidences périphériques du dommage touchant à la sphère professionnelle comme le préjudice subi par la victime en raison de sa dévalorisation sur le marché du travail, de sa perte d'une chance professionnelle, ou de l'augmentation de la pénibilité de l'emploi qu'elle occupe imputable au dommage ou encore du préjudice subi qui a trait à sa nécessité de devoir abandonner la profession qu'elle exerçait avant le dommage au profit d'une autre qu'elle a du choisir en raison de la survenance de son handicap.

### **Dépenses de Santé Futures (DSF)**

<sup>13</sup>Les dépenses de santé futures sont les frais hospitaliers, médicaux, paramédicaux, pharmaceutiques et assimilés, même occasionnels mais médicalement prévisibles, rendus nécessaires par l'état pathologique de la victime après la consolidation. Ces frais futurs ne se limitent pas aux frais médicaux au sens strict : ils incluent, en outre, les frais liés soit à l'installation de prothèses pour les membres, les dents, les oreilles ou les yeux, soit à la pose d'appareillages spécifiques qui sont nécessaires afin de suppléer le handicap physiologique permanent qui demeure après la consolidation. Pour ce préjudice, il y a autant de rentes constituées sur justificatifs distinctes qu'il y a d'appareillages ou de matériels médicaux à renouveler dès lors que la fréquence de renouvellement est différente.

---

10. Extrait de la nomenclature Dintilhac et dont la définition complète se trouve en annexe A.1

11. Extrait de la nomenclature Dintilhac et dont la définition complète se trouve en annexe A.1

12. Extrait de la nomenclature Dintilhac et dont la définition complète se trouve en annexe A.1

13. Extrait de la nomenclature Dintilhac et dont la définition complète se trouve en annexe A.1

---

### **Frais de Logement Adapté (FLA)**

<sup>14</sup> Les frais de logement adapté concernent les frais que doit déboursier la victime directe à la suite du dommage pour adapter son logement à son handicap et bénéficier ainsi d'un habitat en adéquation avec ce handicap. Ce poste de préjudice inclut non seulement l'aménagement du domicile préexistant, mais éventuellement celui découlant de l'acquisition d'un domicile mieux adapté prenant en compte le sur-coût financier engendré par cette acquisition. En outre, il est possible d'inclure au titre de l'indemnisation de ce poste de préjudice les frais de déménagement et d'emménagement, ainsi que ceux liés à un surcoût de loyer pour un logement plus grand découlant des difficultés de mobilité de la victime devenue handicapée.

### **Frais de Véhicule Adapté (FVA)**

<sup>15</sup> Ils correspondent aux dépenses nécessaires pour procéder à l'adaptation d'un ou de plusieurs véhicules aux besoins de la victime atteinte d'un handicap permanent. Il convient d'inclure dans ce poste de préjudice frais de véhicule adapté le ou les surcoût(s) lié(s) au renouvellement du véhicule et à son entretien. En outre, ce poste frais de véhicule adapté doit inclure non seulement les dépenses liées à l'adaptation d'un véhicule, mais aussi le surcoût d'achat d'un véhicule susceptible d'être adapté. Enfin, il est également possible d'assimiler à ces frais d'adaptation du véhicule les surcoûts en frais de transport rendus nécessaires à la victime en raison de ses difficultés d'accessibilité aux transports en commun survenues depuis le dommage.

### **Déficit Fonctionnel Permanent (DFP)**

<sup>16</sup> Ce poste préjudice cherche à indemniser un préjudice extra-patrimonial découlant d'une incapacité constatée médicalement qui établit que le dommage subi a une incidence sur les fonctions du corps humain de la victime. Il s'agit ici de réparer les incidences du dommage qui touchent exclusivement à la sphère personnelle de la victime. Il convient d'indemniser, à ce titre, non seulement les atteintes aux fonctions physiologiques de la victime, mais aussi la douleur permanente qu'elle ressent, la perte de la qualité de vie et les troubles dans les conditions d'existence qu'elle rencontre au quotidien après sa consolidation.

Il est évalué à la consolidation dans le cadre de l'expertise médicale finale dans une échelle de 0% à 100%. Ainsi, le DFP est constitué de trois éléments : l'atteinte aux fonctions physiques et à l'intégrité morale, les souffrances endurées après la consolidation, la perte de qualité de vie et les troubles ressentis dans les conditions d'existence personnelles, familiales et sociales de la victime.

---

14. Extrait de la nomenclature Dintilhac et dont la définition complète se trouve en annexe A.1

15. Extrait de la nomenclature Dintilhac et dont la définition complète se trouve en annexe A.1

16. Extrait de la nomenclature Dintilhac et dont la définition complète se trouve en annexe A.1

---

## 3. Principes de gestion de sinistres corporels Auto Groupama

En assurance non-vie, nous calculons souvent deux types de provisions : les PSAP<sup>17</sup>, lorsque la rente est en attente (décision judiciaire ou amiable non arrêtée) et les PM<sup>18</sup> lorsque la rente est déjà effective (en service).

Dans cette section, nous présenterons les principes de gestion des sinistres corporels Auto qui seront indemnisés sous forme de rente.

### 3.1 Mode de gestion

Comme dit précédemment, l'indemnisation de certains postes de préjudice d'une victime de dommage corporel peut se faire en capital ou en rente. Le règlement en rente des prestations futures est privilégié pour les victimes lourdement handicapées. Il se retrouve donc essentiellement dans les dossiers lourds : il s'agit des dossiers cogérés dont le coût total prévisionnel (CTP) est supérieur ou égal à 2 M€.

#### 3.1.1 La coût total présumé

Le CTP tient compte du droit à indemnisation de chaque victime, tel qu'il résulte des dispositions de la Loi du 5 juillet 1985 et du Droit Commun, ainsi que des règles établies dans le cadre de la Convention IRCA (convention d'Indemnisation et de Recours Corporel Automobile) et du Protocole d'Accord Organismes Sociaux / Entreprises d'Assurances. Et donc le coût du sinistre (évalué dossier dossier) n'est pas seulement égal à la provision mais dépend aussi d'autres composantes et dont la somme donne le CTP :

$$CTP = \text{Règlement} + \text{Provision} - \text{Recours} - \text{Provisions de recours} \quad (1.3.1)$$

Lorsque des réserves sont formulées sur le plan médical quant à la survenance ultérieure d'une aggravation, le provisionnement de cette aggravation doit être apprécié en fonction de sa probabilité, de son délai prévisible de survenance et selon avis médical. De ce fait la provision sera réévaluée et le CTP peut être modifié.

#### 3.1.2 Règles de provisionnement

Pour les dossiers RC Auto gérés en rentes de droit commun, les règles de provisionnement varient d'un poste de préjudice à un autre :

- Assistance Tierce Personne : Ce poste est provisionné en rente et celle-ci est revalorisée dans les dossiers RC auto à compter du 1er janvier 2013. La table de mortalité à utiliser est celle utilisée par Groupama, le taux d'actualisation est le taux social (calculé sur la moyenne du TME au cours des 24 derniers mois + 10 pb).
- Pertes de gains professionnels futurs : Ce poste est provisionné en rente lorsque l'âge de la victime au moment de l'accident est inférieur ou égal à 25 ans. Cette rente est revalorisée dans les dossiers RC auto à compter du 1er janvier 2013. La table de mortalité à utiliser est celle utilisée par Groupama, le taux d'actualisation est le taux social (calculé sur la moyenne du TME au cours des 24 derniers mois + 10 pb).
- Dépenses de santé futures : Ce poste peut être indemnisé sous forme de rente viagère et cela concerne les frais médicaux, pharmaceutiques et hospitaliers (FMPH) après liquidation et appareillage viager. **Cette**

---

17. Provisions pour Sinistres à Payer

18. Provision mathématique

**rente n'est jamais revalorisée.** La table de mortalité à utiliser est celle utilisée par Groupama, le taux d'actualisation est le taux social (calculé sur la moyenne du TME au cours des 24 derniers mois + 10 pb).

NB : La table de mortalité à utiliser pour le calcul des provisions de rentes est la table réglementaire TD 88-90. Cependant, Groupama utilise la table de mortalité TF-TH 00-02, qui est plus prudente que la table réglementaire.

### 3.1.3 Barème de capitalisation

Le point de départ de la capitalisation étant la date prévisible de liquidation, pour les comptes sociaux<sup>19</sup>, les provisions relatives aux **rentes de droit commun non-vie** (en service et en attente) sont actualisées selon les principes suivants :

- Table de mortalité : Table prospective TF-TH 00-02 (elle est prudente que la table réglementaire TD 88-90) ;
- Taux d'actualisation : calculé en utilisant la formule 60% du TME des 24 derniers mois + 10 bp. (égal 0.37% au 31/12/2019).
- Taux de revalorisation : Taux à 2% au 31/12/2019.

Pour les postes de préjudice qui seront gérés sous forme de **capital**, ils sont capitalisés sur la base du Barème Gazette du Palais 2018 : taux d'actualisation à 0.50% pour les postes frais de véhicule endommagés, pertes de revenus et PGPF si la victime a plus de 25 ans lors de l'accident.

Le barème des organismes sociaux à 1,29% est appliqué pour les dépenses de santé futures si un règlement en capital est prévu.

Le figure 1.3.1 résume les règles de provisionnement et principes de capitalisation énoncées précédemment :

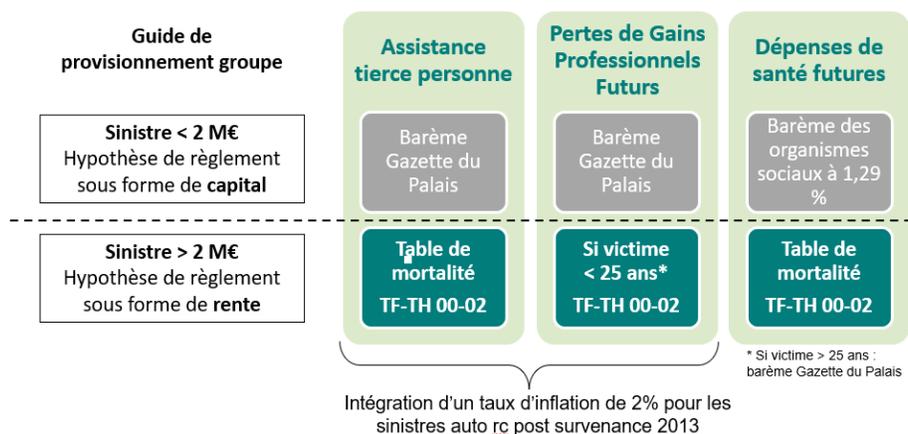


FIGURE 1.3.1: Récapitulatif du guide de provisionnement groupe

### 3.1.4 Calcul de la provision mathématique

Le montant de la provision d'une rente est le produit du montant d'arrérages à verser par l'euro de rente correspondant aux caractéristiques de la victime (âge, terme, sexe, ...). Nous allons ci-dessous détailler la méthode de provisionnement d'une rente viagère avec taux de revalorisation.

19. normes françaises : règlement ANC n°2015-11

---

On note :

- $C$  : le montant de la rente versée chaque année ;
- $x$  : l'âge de l'assuré au moment du calcul de la provision. Il est égale à :  $x = N - N_{naissance} + \delta$  avec  $N$  la date de calcul de la provision,  $N_{naissance}$  sa date de naissance et  $\delta$  qui vaut 1 si la naissance du rentier a eu lieu dans le premier semestre, 0 si non ;
- $tx_{revalo}$  : le taux revalorisation ;
- $tx_{actu}$  : le taux d'actualisation ;
- $p_{x,x+i}$  : la probabilité qu'un individu d'âge  $x$  vive au moins  $i$  années, calculé à partir la table de mortalité ;  $p_{x,x+i} = \frac{l_{x+i}}{l_x}$
- $\omega$  : l'âge maximal de la table de mortalité.

On a :

$$Provision = C \times \sum_{i=1}^{w-x} \left[ \frac{(1 + tx_{revalo})^i}{(1 + tx_{actu})^i} \times p_{x,x+i} \right] = C \times a_x^{revalo} \quad (1.3.2)$$

Avec  $a_x^{revalo}$  qui est le prix d'un euro de rente.

Les probabilités de survie sont calculées à partir d'une table visée à l'article A-331-10 du Code des assurances, soit sur la table de capitalisation TF-TH 00-02 fourni par le groupe. De plus, le taux d'actualisation utilisé est le taux social, calculé de la manière suivante : 60% TME<sup>20</sup> moyen sur 24 derniers mois + 10bp.

Sauf information médicale ou judiciaire, les capitaux constitutifs de rente ainsi évalués ne changent pas dans le système de gestion.

## 3.2 La revalorisation des sinistres

Pour les sinistres survenus avant le 1<sup>er</sup> Janvier 2013, la revalorisation des rentes Auto corporels étaient pris en charge par le Fonds de Garantie des Assurances Obligatoires de dommages (FGAO). Il a été créé en 1951 dans le but d'indemniser les victimes d'accidents de la circulation provoqués par des personnes non assurées ou non identifiées. Au fil du temps, ses missions se sont élargies : prise en charge des victimes d'accidents survenus à l'étranger. Ils fournissent également une assistance aux propriétaires de maisons endommagées par les activités minières, et proposent une indemnisation aux personnes blessées qui ont signé des contrats de responsabilité automobile ou une indemnisation en cas de défaillance d'une compagnie d'assurance nationale ou européenne.

Depuis 2004, les capitaux propres du FGAO sont épuisés. Ainsi, en décembre 2012, le bureau des finances rectificatives a voté une loi visant à régler ce problème de financement. A partir du 1<sup>er</sup> Janvier 2013, ces revalorisations de rentes futures sont désormais à la charge de l'assureur. Ils doivent donc les inclure dans le calcul de leurs provisions. Ces revalorisations représentent une grosse charge pour les assureurs et ne sont pas pris en charge par les réassureurs. La revalorisation du FGAO ne concerne que les sinistres automobiles dont l'indemnisation se fait sous forme de rente.

## 3.3 Les sources de volatilité de la rente

Les régleurs de sinistres évaluent les montants annuels de rentes victime par victime, en fonction des éléments de l'expertise médicale. Ils dépendent des caractéristiques de chaque victime : situation personnelle et professionnelle et de leurs besoins. Ces éléments peuvent donc créer de la différence entre les montants des rentes des victimes. Pour deux victimes présentant les mêmes caractéristiques (âges, sexe, ...) et ayant subi les mêmes dommages, les sources de volatilité des rentes peuvent provenir de :

---

20. Taux Moyen d'emprunt d'Etat

- 
- La localisation des victimes : le coût de la tierce personne sera plus élevé pour une victime vivant dans une zone isolée que pour une autre vivant dans une zone accessible ;
  - Du taux d'invalidité : deux victimes peuvent avoir des rentes différentes, dépendant du membre touché lors de l'accident, des soins particuliers dont l'un ou l'autre peut avoir besoin ;
  - Le besoin et la situation familiale de chacune des victimes ;
  - De la décision judiciaire : En fonction du tribunal et/ou du juge, le montant de l'indemnisation peut varier. Par exemple, pour le poste ATP, les honoraires de l'intervenant peuvent varier selon le tribunal.

## 3.4 La réassurance

Le risque RC corporel Auto, étant une branche à développement long, il existe une forte incertitude sur les charges finales que l'assureur sera devra payer, compte tenu de la durée de vie des sinistres. Ainsi, certaines provisions pourront être sous-estimées et les assureurs devront immobiliser du capital additionnel pour faire face à la dérive des provisions.

Dans le but de limiter leurs éventuelles pertes techniques et de protéger leurs fonds propres, les assureurs ont souvent recours à **la réassurance** : «Une opération de réassurance est un contrat sur lequel un réassureur (dit cessionnaire) vis-à-vis d'un assureur professionnel (dit cédant) qui répond seul et intégralement vis-à-vis des assurés des risques par lui assurés, prend en charge moyennant rémunération tout ou partie des sommes dues ou versées aux assurés à titre de sinistres.» (Picard et Besson)

Comme pour le contrat d'assurance, l'assureur est libre d'avoir tous ses contrats chez le même réassureur ou d'avoir différents contrats chez différents réassureurs. Il est assez rare qu'une cédante ait recours à un seul réassureur sur un de ses contrats de réassurance, elle préfère en général élargir son panel de réassureurs afin de limiter son risque crédit ; de même, les réassureurs ont en général ce que l'on appelle une « capacité de souscription » qui correspond à l'engagement maximal qu'ils sont autorisés à avoir sur un contrat donné. Du fait de cette limitation, ils n'ont souvent pas la possibilité de prendre le risque à 100% et doivent alors le partager avec d'autres réassureurs.

Il existe deux types de traités de réassurance : les traités proportionnels et les traités non-proportionnels.

### 3.4.1 Les traités proportionnels

Dans ce traité, les deux parties (la cédante et le cessionnaire) se partagent la prime ainsi que les sinistres selon un taux de cession fixe. Par exemple, le réassureur reçoit 30% de la prime d'un risque et s'engage en contrepartie à couvrir 30% des sinistres affectant ce même risque. De plus, les frais de gestion et d'acquisition du portefeuille étant entièrement assumés par la Cédante, le contrat de réassurance proportionnelle prévoit le versement par le réassureur d'une « commission de Réassurance » visant à couvrir sa part dans ses différents frais. Deux structures de réassurances sont à distinguer : Le Quote Part et l'Excédent de Plein.

#### Le traité en Quote Part (QP)

Dans ce traité, un taux de cession identique  $x\%$  est défini pour tous les risques dans le portefeuille, quelque soit la somme assurée. De plus, le réassureur verse une commission à la cédante afin de participer aux frais de gestion à la charge de l'assureur.

L'inconvénient de ce traité réside dans le fait que tous les sinistres (petits ou grands) sont réassurés de la même manière. Ainsi, tous les risques acceptés par la cédante doivent être partagés proportionnellement avec le réassureur.

#### L'Excédent de Plein

Encore appelé *Surplus*, il s'agit d'un traité dans lequel l'assureur peut décider de ne garder que les petits sinistres et d'en céder ceux qui dépassent un certain seuil. On définit ainsi le « plein de rétention » : il s'agit

du montant maximal que la cédante souhaite payer par sinistre. L'assureur et le réassureur se partagent la prime et les sinistres selon un taux de cession  $x_i$ . Ce taux est différent pour chaque risque.

Ainsi, soit  $n$  le nombre de risques de l'assureur pour un portefeuille et  $i$  le  $i^{me}$  sinistre. l'assureur conserve  $\sum_{i=1}^n x_i P_i$  de primes et en cède  $\sum_{i=1}^n (1 - x_i) P_i$  au réassureur. De la même manière, ils se partagent les sinistres, soit  $\sum_{i=1}^n x_i S_i$  pour l'assureur et  $\sum_{i=1}^n (1 - x_i) S_i$  pour le réassureur.

Le fonctionnement est le même que pour le Quote Part, si ce n'est que le taux de cession différent pour chaque risque. Le calcul de ce taux de cession se fait selon la formule suivante :

$$x_i = \min \left( \max \left( \frac{K_i - R}{K_i}, 0 \right), \frac{C}{K_i} \right)$$

Avec :

- $K_i$  : le capital assuré du risque  $i$
- $C$  : la capacité de l'Excédent de Plein
- $R$  : le plein de rétention de la Cédante

Ainsi, si  $K_i < R$ , le taux de cession est nul et la cédante conserve les risques dont le capital assuré est inférieur au plein de rétention.

### 3.4.2 Les traités non-proportionnels

Ici, il n'y a pas de relation de proportionnalité entre la prime cédée et la participation du réassureur dans les sinistres. Le réassureur s'engage à payer les sinistres dépassant un certain seuil défini contre le paiement d'une prime, évalué en fonction de la part de sinistres à sa charge. Ce type de traité a pour but d protéger la cédante contre la survenance d'évènements catastrophiques. On distingue les traités en excédent de sinistres (excess of loss) et les traités en excédents de perte annuelle (stop loss).

#### Excédent de sinistres (XS) ou excess of loss (XL)

Dans ce traité, une franchise est définie appelé « priorité », qui est le seuil à partir duquel le traité est déclenché. Le cessionnaire s'engage donc à payer tous les sinistres dépassant cette franchise et dans une limite appelé « portée ». Il paiera donc au plus le montant de la portée pour chaque sinistre. Dans la pratique, la notation d'un traité XS est la suivante : « portée XS priorité »

Graphiquement, un traité XS se présente comme suit :

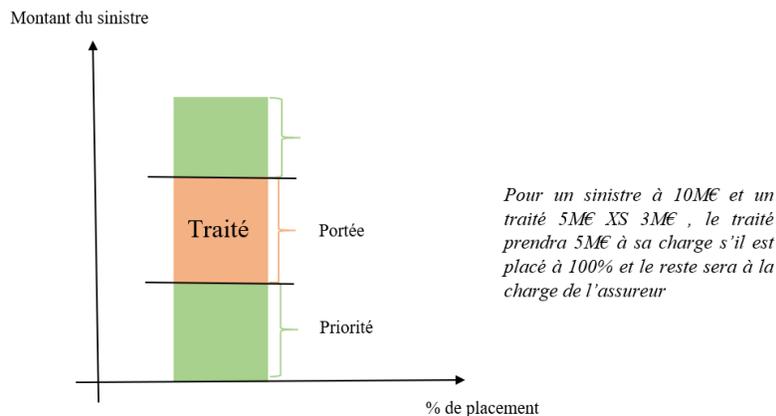


FIGURE 1.3.2: Exemple d'un traité XS

---

Soit  $X$  le montant d'un sinistre couvert par un traité XS, le montant à la charge du réassureur sera :

$$Cession = \begin{cases} 0 & \text{si } X \leq \textit{priorite} \\ X - \textit{priorite} & \text{si } \textit{priorite} \leq X \leq \textit{priorite} + \textit{portee} \\ \textit{portee} & \text{si } \textit{priorite} + \textit{portee} \leq X \end{cases}$$

Si l'on note  $S$  le sinistre à charge du réassureur, on a :

$$S = \min(\max(X - \textit{priorite}, 0), \textit{portee}) \quad (1.3.3)$$

Par exemple, pour un sinistre de 4M€ et un traité 2M€XS 3M€, la cédante prendra à sa charge 3M€ et le réassureur 1M€. Par contre, si le sinistre coûte 6M€, le réassureur paiera la portée 2M€ et l'assureur prendra à sa charge le reste, soit 4M€.

### Excédent de pertes ou Stop loss (SL)

Le principe d'un traité SL est le même que celui d'un traité XS, sauf que le SL protège le ratio S/P de l'année. Dans ce traité, la portée et la priorité sont exprimés en pourcentage de la prime directe. Ainsi, si l'on note  $X_i, i = 1; \dots; n$  les sinistres de la cédante durant la période de couverture, la charge du Stop Loss sera :

$$L_{\text{charge du Stop Loss}} = \begin{cases} 0 & \text{si } \frac{\sum X_i}{P} \leq \textit{priorite} \\ (\frac{\sum X_i}{P} - \textit{priorite}) * P & \text{si } \textit{priorite} \leq \frac{\sum X_i}{P} \leq \textit{priorite} + \textit{portee} \\ \textit{portee} * P & \text{si } \textit{priorite} + \textit{portee} \leq \frac{\sum X_i}{P} \end{cases}$$

Si l'on note  $S$  le sinistre à charge du réassureur, on a :

$$S = \min(\max(\frac{\sum X_i}{P} - \textit{priorite}, 0), \textit{portee}) * P \quad (1.3.4)$$

Le contrat de réassurance est ici très important. En effet il doit détailler clairement les différents éléments pris en compte dans le calcul du ratio S/P ainsi que la période sur laquelle il est calculé (année comptable ou exercice de souscription par exemple).

---

Deuxième partie

Modélisation des rentes futures

---

# 1. La base de rente RC Auto

## 1.1 Analyse descriptive de la base de rentes

Notre base est constituée des rentes AUTO de l'ensemble des caisses régionales. Sur le total des rentes Auto, 95% sont en cours de paiement ou temporairement suspendues, et 5% sont des rentes closes du fait de la nature temporaire de la rente, ou du fait du décès de son bénéficiaire. Elle contient 2906 rentiers dont les informations suivantes y sont renseignées : les caractéristiques biométriques de l'assuré (son identifiant, son année de naissance, son sexe, son taux d'AIPP, le type de bénéficiaire : orphelin, conjointe, victime...), les caractéristiques du sinistre (l'année de survenance, le numéro de sinistre) et les caractéristiques de la rente (son état : en service ou clos, le montant de la rente, la date de création de la rente, le date de fin, le type de rente : viagère ou temporaire, le fractionnement de la rente : qui est annuel dans notre cas).

Notre base contient en tout 25 variables explicatives mais nous n'utiliserons en réalité que les variables suivantes :

- DT\_NAISS : correspond à la date de naissance du rentier ;
- SURVENANCE : correspond à la date de survenance du sinistre du rentier ; ces deux premières variables nous permettront de calculer l'âge du rentier à la survenance du sinistre.
- SEXE : correspond au sexe du rentier (1 : homme, 2 : femme) ;
- TAUX\_AIPP : correspond à un indicateur de la gravité des conséquences de l'accident sur l'état de santé de la victime ;
- PMR\_BASE : correspond au montant de l'arrérage de base du sinistre. Il est donné par le juge.

Les variables sont analysées différemment selon leur nature : quantitative ou qualitative. Les variables quantitatives sont résumées sous forme d'indicateurs (moyenne, écart-type, ...), et sont présentées graphiquement sous forme d'histogramme et de boîtes à moustache ou box-plot.

### 1.1.1 Tri à plat

Il s'agit ici d'étudier le comportement individuel des variables explicatives précédentes :

#### Sexe

Dans notre base, 63% des rentiers sont des hommes et 37% des femmes.

#### Année de survenance

Sur la figure 2.1.1 ci-dessous, nous pouvons remarquer que le nombre de rentiers visible dans la base pour les années de survenance récentes est faible, du fait du "délai de mise en effet"<sup>21</sup> nécessaire avant le début d'indemnisation de la rente. Bien que le plus ancien des sinistres de la base soit survenu en 1921, le nombre de rentiers Auto de GROUPAMA n'est devenu significatif qu'à partir des années 1960 et décroît très vite autour des années 2006.

---

21. délai de consolidation de l'état de santé de la victime et durée de la procédure judiciaire ou de la négociation à l'amiable

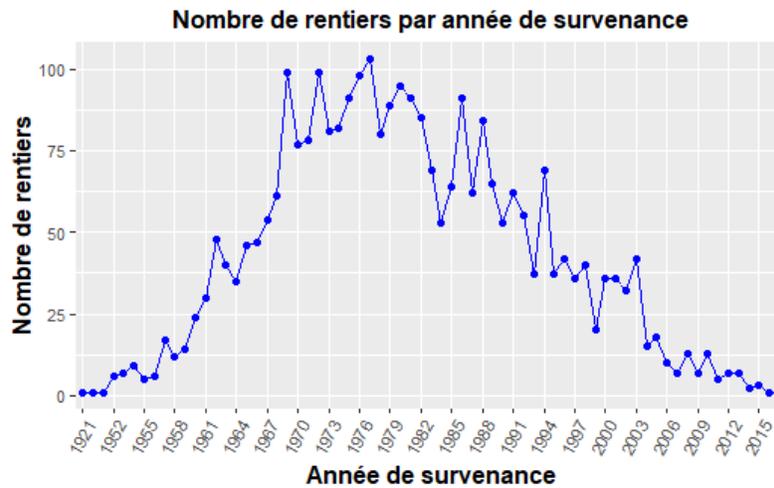


FIGURE 2.1.1: Nombre de rentiers par année de survenance

### Âge à la survenance

Nos rentiers ont des âges compris entre 0 à 82 ans. L'âge moyen de survenance des sinistres de la base est de 28 ans, sa valeur médiane est de 26 ans, et son troisième quartile est de 36 ans. Cette variable, que l'on appellera *Age\_surv* est calculée par l'arrondi de la différence entre l'année de survenance du sinistre et l'année de naissance de la victime.

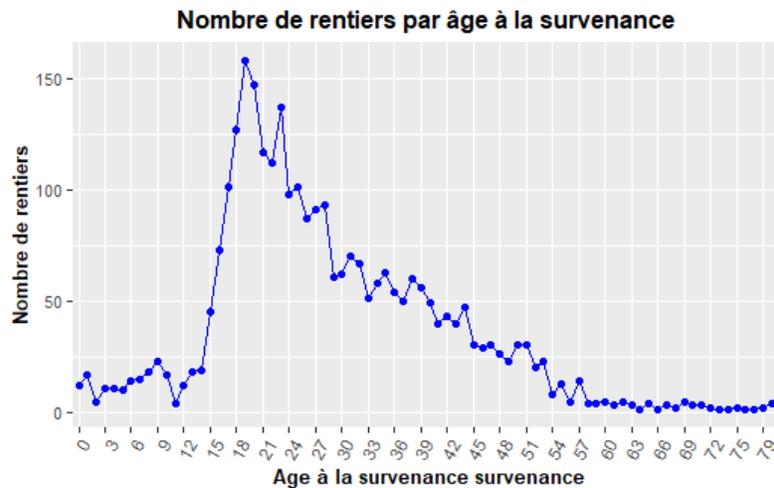


FIGURE 2.1.2: Nombre de rentiers par âge à la survenance

Nous remarquons sur la figure ci-dessus qu'il y'a très peu de rentiers âgés de moins de 18 ans et de plus de 50 ans. De plus, nous constatons un pic du nombre de rentiers aux alentours de 18 ans et puis qui décroît aux environs de 24 ans.

### Taux d'AIPP

L'indice d'AIPP est l'indice de gravité des conséquences de l'accident sur l'état de santé de la victime.

C'est une variable numérique, déterminée par des experts médicaux, et qui prend la valeur d'un entier entre 1 et 100.

Dans notre base, nous avons 1483 individus dont le taux d'AIPP est nul : nous ne saurions dire si ces valeurs ont été remplacées par des 0 ou s'il s'agit des erreurs de saisies. Dans un premier temps, nous les garderons dans notre base.

Min	1er quartile	Médiane	Moyenne	3eme quartile	Max	Variance
0	0	0	21.33	30	100	1073.478

TABLE 1.1: Moments de la variable TAUX\_D'AIPP

Nous constatons que la médiane et la variance sont éloignées de la moyenne : les valeurs sont trop dispersées.

#### Arrérage de base

Il correspond au montant annuel de la rente. Sur l'ensemble des rentes, nous avons :

Min	1er quartile	Médiane	Moyenne	3eme quartile	Max	Ecart-type
0	472	1 984	9 210	8 232	203 472	21 079

TABLE 1.2: Moments de la variable PMR\_BASE

Le 3<sup>eme</sup> quartile étant très proche de la moyenne, cela signifie que 75% de notre base a un montant d'arrérage inférieur ou égale à l'arrérage moyen. De plus, la médiane est très faible par rapport à la moyenne, ce qui traduit une forte dispersion des valeurs. C'est confirmé par la valeur de l'écart-type, proche de 2 fois la moyenne et par les quartiles.

Graphiquement, la distribution de l'arrérage est :

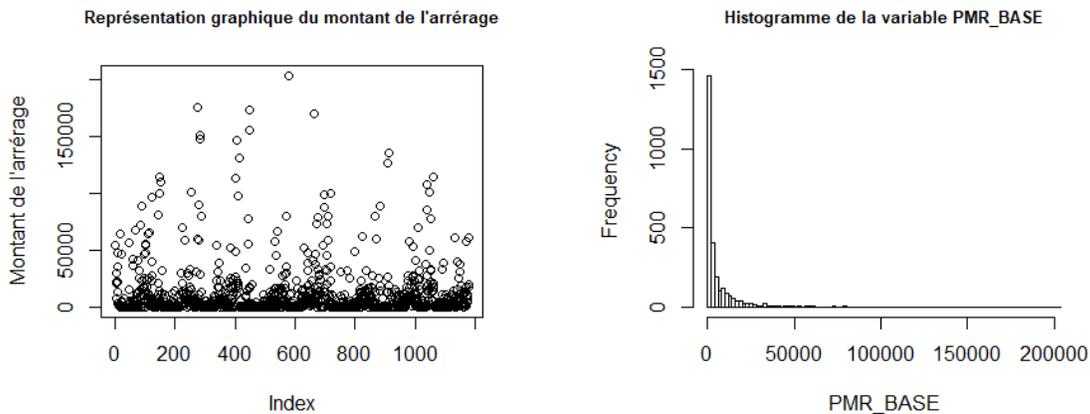


FIGURE 2.1.3: Distribution du montant de l'arrérage

La distribution de l'arrérage étant asymétrique à droite, nous pouvons dire à première vue que cette variable suit une loi de Gamma.

Faisons les tests d'adéquation à cette loi pour confirmer ou infirmer notre intuition.

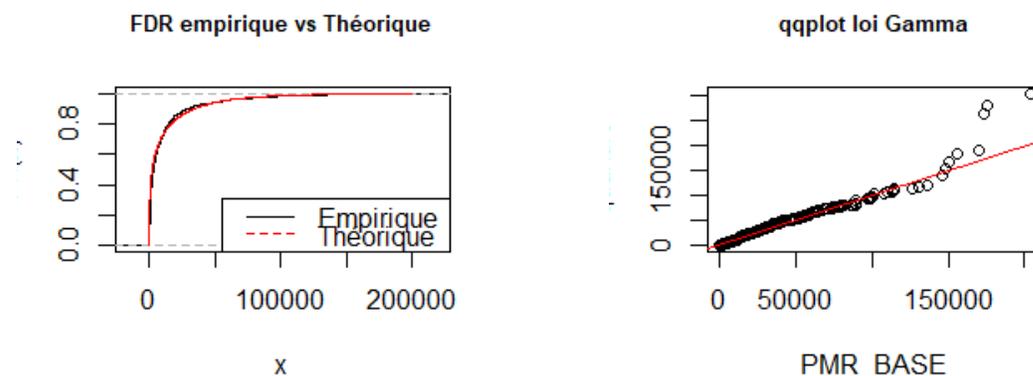


FIGURE 2.1.4: Ajustement de la variable `PMR_BASE` à la loi Gamma

Sur la figure 2.1.4, nous constatons qu'il y'a une adéquation à la loi Gamma en se basant sur les fonctions de répartition (empiriques et théoriques) et que la droite rouge d'équation  $y = x$  s'ajuste bien sur le qqplot, sauf vers la fin, ce qui peut être dû à une faible quantité de données. Nous pouvons donc dire que le montant d'arrérage suit une loi de Gamma. Ainsi la régression log-linéaire ou Gamma pourraient être adaptée pour notre modélisation. Nous verrons dans la suite de notre étude, si cette théorie est vérifiée.

### 1.1.2 Tri croisé

Il s'agit dans cette partie d'étudier la corrélation entre les variables, ceci dans le but de connaître les variables qui sont liées ou celles qui ont une influence sur une autre. Nous rappelons que notre base d'étude finale n'est constituée que des rentiers dont le taux d'AIPP  $> 1$  et le montant d'arrérage non nul.

En effet, nous cherchons à expliquer le montant d'arrérage en fonction des variables explicatives et ensuite le modéliser (dans la section suivante).

#### L'évolution de l'âge moyen à la survenance

Sur la Figure ci-dessous, nous pouvons observer que l'âge moyen à la survenance est globalement plus faible pour les sinistres les plus anciens. Cela veut dire qu'à cet époque, les rentes n'étaient versées qu'aux jeunes sinistrés. Par contre, sur les années récentes, la tendance a changé. De plus, nous constatons que l'âge moyen par année de survenance est compris entre 20 et 30 ans. Nous remarquons également que l'âge moyen à la survenance est globalement plus élevé pour les sinistres les plus récemment survenus.

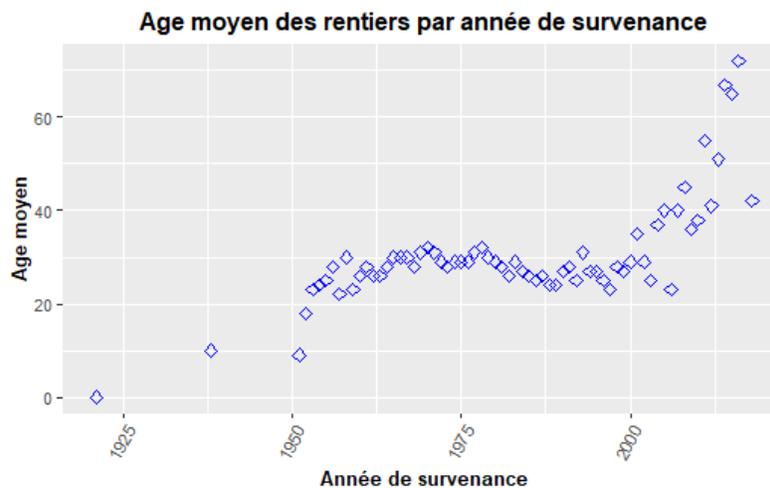


FIGURE 2.1.5: Évolution de l'âge moyen à la survénance

Sexe  $\times$  Âge à la survénance

À partir de la figure ci-dessous, nous constatons que l'âge moyen de survénance des femmes est de 31 ans et celui des hommes est de 27 ans. De plus, nous constatons que les hommes de moins de 45 ans ont plus de sinistres que les femmes, qui quant à elles, en ont plus autour de 45 ans. Par ailleurs, aux âges élevés, nous pouvons constater qu'il y'a autant de femmes que d'hommes.

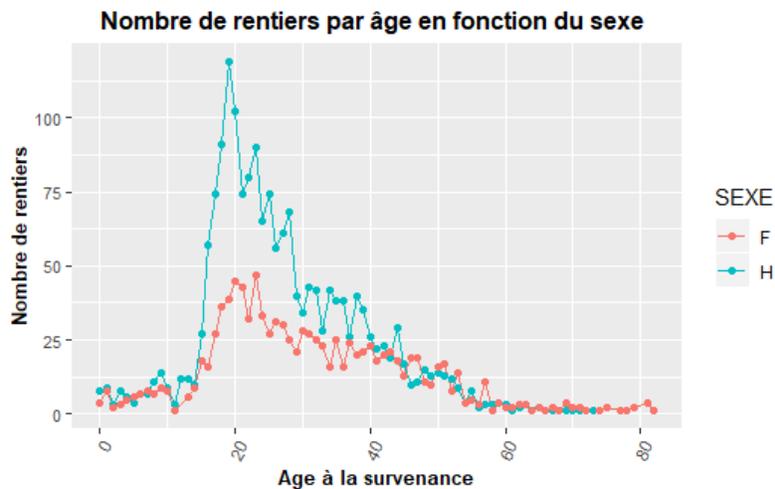


FIGURE 2.1.6: Nombre de rentiers par âge en fonction du sexe

Âge à la survénance  $\times$  AIPP Moyen

L'indice d'AIPP moyen calculé par âge de survénance est représenté sur la figure ci-dessous. Nous remarquons qu'il n'existe pas de relation linéaire entre les deux variables (le coefficient de corrélation de Pearson nous le confirme : -0.11). Le taux d'AIPP étant très volatile (comme présenté à la sous section 1.1.1, tableau 1.1), nous représenterons plutôt l'indice moyen. Sur la figure 2.1.7 ci-dessous, nous constatons qu'aux âges

élevés et faibles, il y'a une grande dispersion (dû probablement aux faibles données), et les taux d'AIPP moyens sont très élevés. Ce taux baisse jusqu'à 35 ans environ, se stabilise entre 35 et 45 ans, pour ensuite croître : il existe donc une certaine tendance en fonction de l'âge.

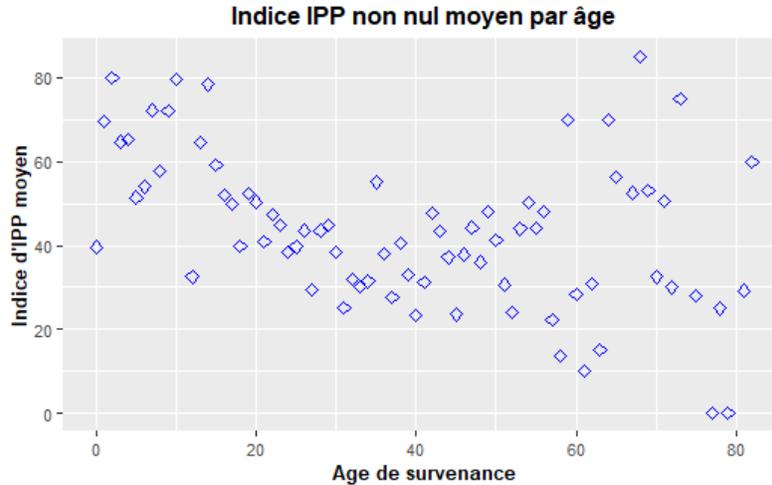


FIGURE 2.1.7: Indice AIPP moyen par âge à la survenance

Regardons cette tendance de près : créons les classes d'âges (par pas de 10 ans) et observons le comportement du taux d'AIPP moyen par sexe.

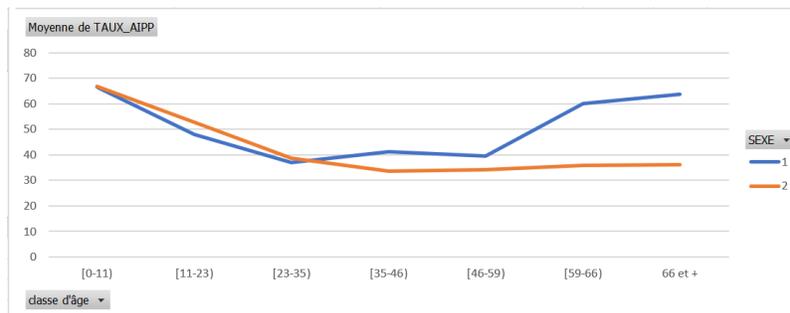


FIGURE 2.1.8: Évolution du taux d'AIPP moyen par classe d'âge en fonction du sexe

Ce graphique nous montre que la gravité moyenne des accidents est plus élevée chez les femmes de moins de 35 ans (allure décroissante), se stabilise entre 35 et 45 ans et croît légèrement après 45 ans. Par contre chez les hommes, il y'a une chute du taux d'AIPP moyen jusqu'à 35 ans puis une stabilisation et ensuite une hausse.

Ainsi, nous pouvons dire qu'en moyenne, il existe une relation (non pas linéaire mais par morceaux) entre le sexe, âge et le taux d'AIPP.

#### Année de survenance × Arrérage moyen

Sur la figure 2.1.9 ci-dessous, nous constatons que l'arrérage augmente avec les années. Cela semble être logique du fait de l'effet d'inflation. Pour des années récentes, nous observons une forte dispersion sur les années récentes, due à une faible quantité de données.

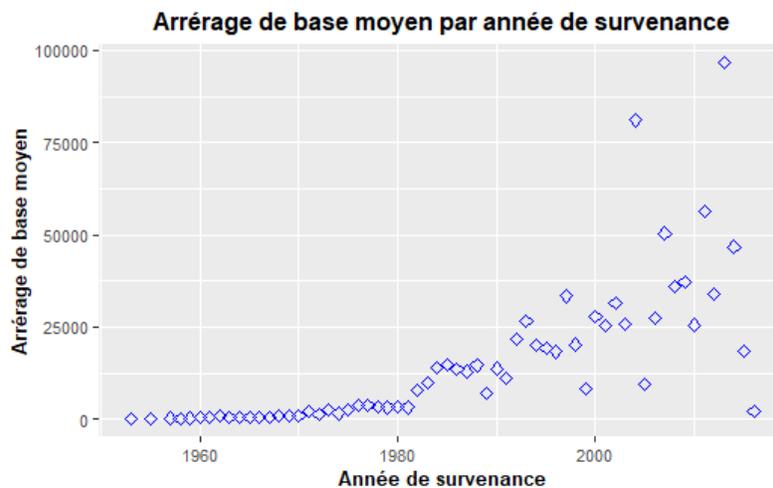


FIGURE 2.1.9: Montant d'arrérage moyen par année de survénance

SEXE × Arrérage moyen

Sur la figure 2.1.10 ci-dessous, nous constatons que le montant moyen d'arrérage des hommes et des femmes est très proche et est situé sur le 3<sup>e</sup> quartile : nous pouvons dire que 3/4 des rentiers (hommes/femmes) ont un montant d'arrérage inférieur ou égale à l'arrérage moyen par sexe. Nous pouvons donc dire que le montant de l'arrérage ne dépend en aucun cas du sexe. Nous constatons néanmoins quelques points atypiques.

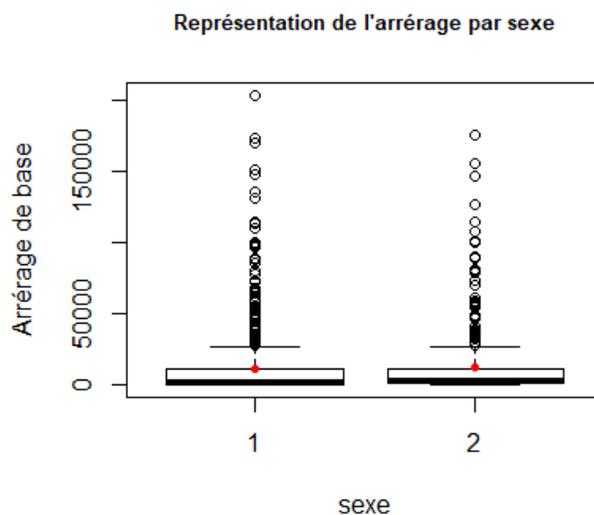


FIGURE 2.1.10: Distribution de l'arrérage par sexe

Âge à la survénance × Arrérage moyen

Nous constatons sur le graphe 2.1.11 ci-dessous que le montant de l'arrérage est relativement élevé pour les jeunes (moins de 18 ans) et ensuite décroît puis tend à se stabiliser. Cela peut s'expliquer par le fait

que pour les jeunes, plusieurs postes de préjudice peuvent être impactés, notamment la PGPF, qui obligera l'assureur de rembourser environ ce que le jeune gagnerait s'il n'avait pas eu d'accident. De plus, nous observons quelques points atypiques chez les vieux, ceux qui ont réellement été affectés par l'accident et qui auront besoin d'assistance : le poste tiers personnes (active et passive) est donc déclenché et ce pour un bon moment car l'arrérage est très élevé.

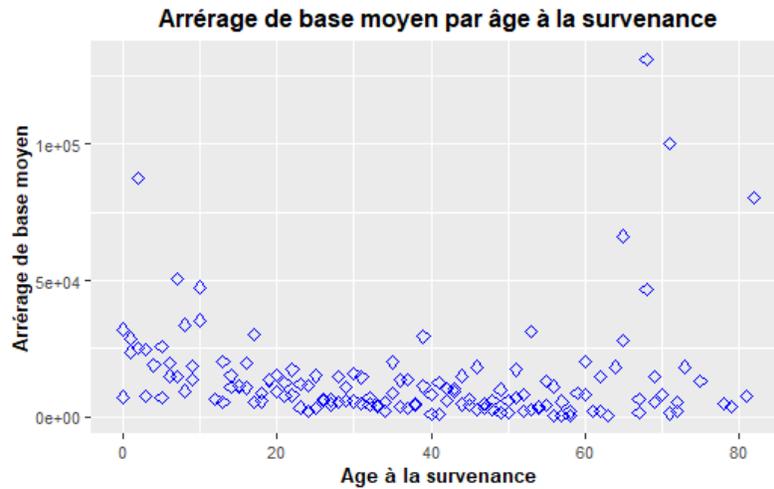


FIGURE 2.1.11: Montant d'arrérage moyen par âge à la survenance

#### Taux d'AIPP × Arrérage moyen

Nous constatons d'après la figure 2.1.12 ci-après que l'arrérage augmente progressivement avec l'indice d'AIPP : les deux variables sont donc positivement corrélées. Cela peut être justifié par le fait que plus l'AIPP est important, plus les séquelles sont importantes et plus le montant de la rente sera important. Nous constatons néanmoins des cas atypiques.

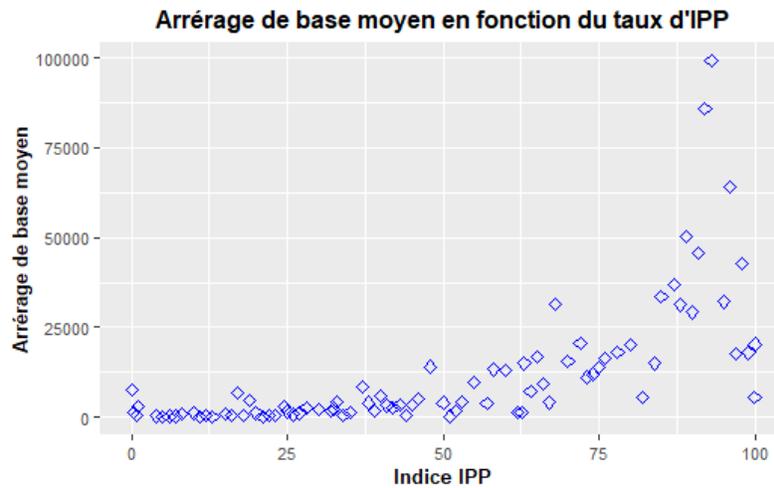


FIGURE 2.1.12: Montant d'arrérage moyen par en fonction du taux d'AIPP

En fonction de l'âge et du sexe de la victime, on a le graphique suivant :

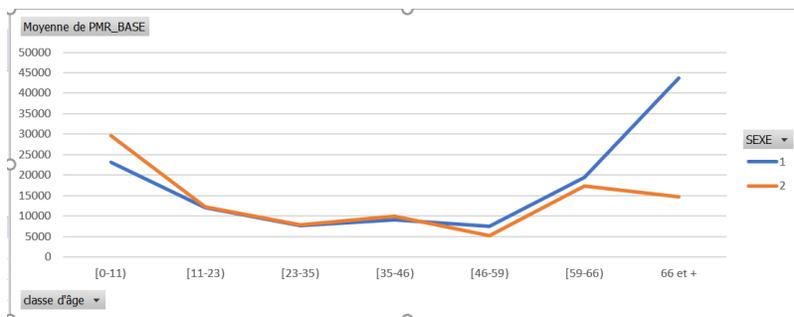


FIGURE 2.1.13: Évolution de l'arrérage moyen par classe d'âge en fonction du sexe

Nous pouvons constater qu'en moyenne, le coût du sinistre est plus élevé chez les hommes âgés que chez les femmes. Par contre, le coût du sinistre est plus élevé chez les femmes de moins de 23 ans.

**Conclusion** : Nous pouvons donc dire que le sexe est un facteur important pour la suite de notre modélisation. Nous ferons donc deux modèles.

Notre base d'étude finale est celle dont les rentiers ont un taux d'AIPP  $> 1$ . De plus, vu que nous cherchons à modéliser l'arrérage, seuls les montants non nuls sont nécessaires à l'étude.

Faire une analyse graphique des variables est juste, mais une analyse statistique est encore meilleure.

## 1.2 Modélisation des variables explicatives et leur dépendance

Cette section a pour but de confirmer/infirmier les résultats et conclusions graphiques qu'on a pu observer dans la section précédente.

Il sera question d'étudier la dépendance entre nos différentes variables. Cette étape est très importante car dans l'élaboration du modèle : nous ne devons pas mettre des variables dépendantes car non seulement, l'une des variables n'apporte aucune information pertinente, mais aussi, elle risque de biaiser notre étude.

### 1.2.1 Création des classes d'âges homogènes

Nous avons vu dans la sous-section 1.1.1 que l'indice de gravité de la victime était très volatile : c'est la raison pour laquelle nous avons décidé de croiser cette variable avec une autre afin de mieux l'étudier. Dans la sous-section 1.1.2, figure 2.1.7, nous l'avons croisé avec l'âge : nous avons constaté une sorte de tendance en fonction de l'âge. Nous avons donc approfondi en créant des classes d'âges par pas de 10 ans et par sexe (figure 2.1.8) : nous sommes arrivés à la conclusion qu'il existe bel et bien une relation non-linéaire entre ces trois variables. De plus, dans le but d'améliorer la qualité de nos variables de base pour les rendre plus efficaces pour mon modèle prédictif, nous avons décidé de créer des classes d'âges homogènes selon le critère de la **moyenne**. Cette démarche a pour objectif de diminuer la volatilité du taux d'AIPP et construire un modèle fin et robuste.

En effet, nous allons créer des classes d'âges de telle sorte que la moyenne du taux d'AIPP entre les classes soit relativement différente et la variance intra-classe soit faible (seuil à 10%). Pour ce faire, nous allons utiliser les arbres de décision : **l'algorithme CART**.

L'algorithme CART ( Classification And Regression Trees ) est une méthode de construction d'arbres binaires de décision, modélisant une discrimination ou une régression. Un arbre CART se décompose en :

- Branches : elles constituent les règles de segmentation de la population ;
- Racine : il s'agit de toute la population étudiée ;
- Feuilles : il s'agit des sous populations finales et elles fournissent une estimation de la quantité d'intérêt.

---

Il consiste à séparer de manière itérative les individus en deux régions selon un critère de division et à modéliser la variable continue par la moyenne de la valeur de la variable sur chacune des classes créées. A chaque itération, le critère de division permettant de créer les deux classes, sera déterminé à partir d'une des variables explicatives qui sera la variable de segmentation. La variable de segmentation peut être différente à chaque itération.

Le processus de segmentation est le suivant :

- Il existe une covariable, soit elle est :
  - Quantitative (numérique, ordonnée) : les partitionnements possibles se situent entre 2 valeurs successives observées pour la variable ;
  - Qualitative (catégorielle) : les partitionnements possibles sont toutes les combinaisons possibles entre les modalités ;
- L'arbre teste les partitionnements et calcule un critère d'homogénéité (minimisation de l'erreur quadratique moyenne) basé sur la variable réponse ;
- L'arbre choisit le partitionnement qui conduit à la plus grande homogénéité globale dans les deux sous espaces créés ;
- Les étapes 1 à 3 sont répétées pour les variables explicatives qu'on dispose : nous obtenons  $k$  mesures d'homogénéité maximales, chacune correspondant à une covariable ;
- L'arbre sélectionne la variable et son partitionnement qui maximise globalement l'homogénéité.

L'objectif de l'arbre CART est de découper l'espace des variables explicatives en  $K$  régions, notées  $R_1, \dots, R_K$ , qui correspondent aux feuilles de l'arbre. Ce découpage doit permettre de minimiser la quantité suivante :

$$RSS = \sum_{j=1}^K \sum_{i \in R_j} (y_i - \hat{y}_j)^2 \quad (2.1.1)$$

Minimiser le  $RSS$  revient à minimiser :

$$RSS = \frac{1}{n} \sum_{j=1}^K n_j V_j \quad (2.1.2)$$

Avec  $V_j = \frac{1}{n_j} \sum_{i \in R_j} (y_i - \bar{y}_j)^2$ , qui est la variance intra-classe.

Dans notre cas, la variable réponse est le  $TAUX\_D'AIPP$ , qui est une variable continue : il s'agit donc des arbres de régression CART. Cette méthode vérifie récursivement pour chaque nœud si une séparation est possible sur la base de la mesure choisie : **minimisation de la variance intra-classe**.

Les critères d'arrêts de l'algorithme sont :

- un nombre d'observation  $q$  minimale dans les feuilles ;
- une réduction du critère d'erreur d'un seuil minimal.

## Application

A l'aide du package "*rpart*" du logiciel R, sa fonction "*rpart*" permet de détailler l'algorithme en précisant les paramètres de tuning suivants : la complexité (noté *cp* sous R), le nombre minimum d'observations qui doit exister dans un nœud pour segmenter (noté *minsplitt*), et le nombre minimum d'observations dans un nœud terminal ou feuille (*minbucket*). Ainsi, la fonction "*rpart.plot*" de ce package nous permet d'avoir la segmentation suivante pour les hommes :

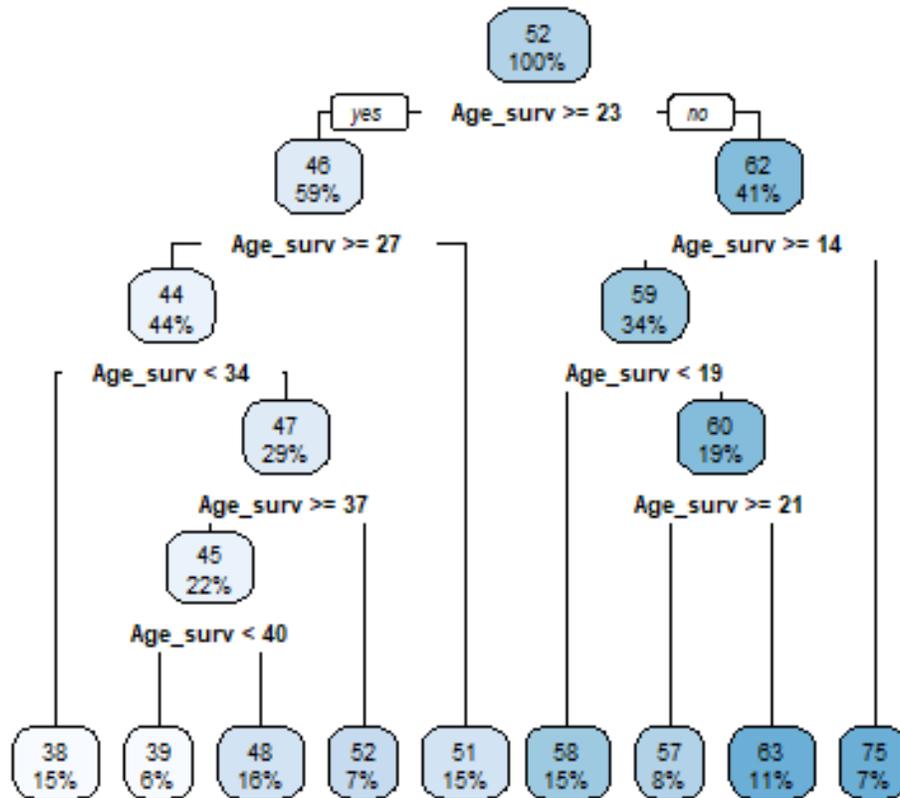


FIGURE 2.1.14: Arbre de régression de l'âge à la survénance

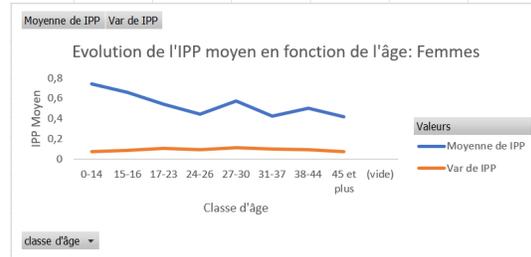
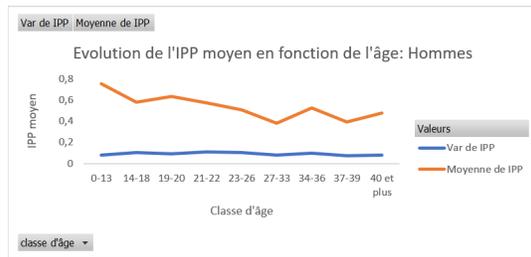
A chaque segmentation de l'arbre, la classe de gauche rassemble les rentes vérifiant le critère de division ("yes"), et la classe de droite rassemble les rentes ne le vérifiant pas ("no"). Dans chaque classe de l'arbre, sont indiquées la proportion de rentiers dans la classe d'âge et la moyenne d'âge dans chaque classe.

Et Donc les classes que nous retenons pour les hommes sont :  $[0, 13], [14, 18], [19, 20], [21, 22], [23, 26], [27, 33], [34, 36], [37, 39], [40, \infty]$ .

Et pour les femmes :  $[0, 14], [15, 16], [17, 23], [24, 26], [27, 30], [31, 37], [38, 44], [45, \infty]$ .

Faisons le **test de l'analyse de la variance à un facteur** pour comparer les classes. L'ANOVA permet de voir si une variable numérique a des valeurs différentes en fonction de plusieurs groupes. C'est une généralisation du test de Student permettant de comparer plus de deux groupes. En utilisant la fonction "aov" de R, nous avons obtenu une valeur de la p-value très faible ( inférieur au seuil 5%) : le test est donc significatif. Par contre, nous ne savons pas quelles classes sont concernées. Il faudra donc faire un **test de Student non-apparié** pour comparer les classes deux par deux en utilisant la moyenne. A l'aide de la fonction "t.test()" de R, nous avons obtenu les p-values significatifs : et donc nos classes sont homogènes.

Graphiquement, nous avons :



Nous pouvons constater que la variance intra-classe est plutôt faible et que les taux d'AIPP moyens par classe d'âge sont distincts : nos classes d'âges sont homogènes (chez les hommes comme chez les femmes).

## 1.2.2 Test corrélation entre variables quantitatives

Trois statistiques peuvent être calculées :

### Le coefficient de corrélation de Pearson

En statistiques, mesurer la corrélation entre variables aléatoires est très souvent réduit à la corrélation linéaire, c'est-à-dire l'ajustement d'une variable par rapport à l'autre par une relation affine obtenue par régression linéaire. Pour cela, nous calculons un coefficient de corrélation linéaire, obtenu par le quotient de leur covariance par le produit de leurs écarts types. Ce coefficient ne permet pas de mesurer l'intensité de la liaison mais permet de quantifier l'évolution de la réponse en fonction de la variable prédictive. La valeur absolue du coefficient est toujours comprise entre 0 et 1.

*NB : Un coefficient nul n'implique pas indépendance, car d'autres types de corrélation peuvent exister.*

### Le rho de Spearman

La corrélation de Spearman est une mesure de dépendance statistique non paramétrique entre deux variables numériques. Elle permet d'évaluer la force d'une liaison monotone entre deux variables. Elle consiste à trouver un coefficient de corrélation souvent noté  $\rho$ , non pas entre les valeurs prises par les deux variables mais entre les rangs de ces valeurs. La valeur de ce coefficient est compris entre  $-1$  et  $1$  : plus la valeur absolue du coefficient est proche de 1, plus la relation monotone entre les deux variables est forte (parfaite).

La corrélation de Spearman est définie comme suit :  $\rho = \frac{cov(r_{gX}, r_{gY})}{\sigma_{r_{gX}} \cdot \sigma_{r_{gY}}}$

Où :

- $cov(r_{gX}, r_{gY})$  est la covariance des variables de rang
- $\sigma_{r_{gX}}$  et  $\sigma_{r_{gY}}$  sont les écarts-type des variables de rang.

*NB : Cette définition correspond à la corrélation de Pearson des variables de rang.*

### Le tau de Kendall

La corrélation de Kendall est une autre méthode non paramétrique utilisée pour évaluer la corrélation monotone de deux variables. Cette méthode n'est pas basée sur le rang des observations, mais sur la paire de concordantes et discordantes. Dans la plupart des cas, les résultats de corrélation de Spearman et Kendall sont très similaires. Par conséquent, la corrélation Kendall peut être utilisée à la corrélation Spearman.

### **Résultats :**

Sur notre base d'étude, nous avons trois variables quantitatives : *Age\_surv*, *TAUX\_AIPP*, *PMR\_BASE*.

Correlation de Pearson				Correlation de Spearman				Correlation de Kendall			
	Age_surv	Taux_IPP	PMR_BASE		Age_surv	Taux_IPP	PMR_BASE		Age_surv	Taux_IPP	PMR_BASE
Age_surv	1			Age_surv	1			Age_surv	1		
Taux_IPP	-0.2220147	1		Taux_IPP	-0.2413285	1		Taux_IPP	-0.1634349	1	
PMR_BASE	-0.1068878	0.4192508	1	PMR_BASE	-0.2217252	0.6453714	1	PMR_BASE	-0.148246	0.4653377	1

FIGURE 2.1.15: Tableau récapitulatif des tests de corrélation entre nos variables quantitatives

Nous constatons que les variables *TAUX\_AIPP* et *PMR\_BASE* sont positivement corrélées. De plus, il existe une dépendance plutôt forte (64%) : donc le taux d'AIPP permet d'expliquer à 64% le montant de l'arrérage. Ce qui confirme le résultat graphique 2.1.12 que nous avons vu dans la sous-section 1.1.2. Cependant, les variables *Age\_surv*, *TAUX\_AIPP* et *Age\_surv*, *PMR\_BASE* sont négativement et faiblement corrélées. Donc l'âge du rentier n'a pas une grande influence ni sur le montant d'arrérage ni sur l'indice de gravité du sinistre.

### 1.2.3 Test d'indépendance entre variables qualitatives

Dans la base d'étude, plusieurs variables sont qualitatives, mais deux seules nous intéressent pour notre modélisation : la variable *SEXE* et la variable représentant les classes d'âges (précédemment créées), notée *Age\_survNew*.

Pour mesurer la dépendance entre ces deux variables, nous avons utilisé deux statistiques :

Test de Khi-2 d'indépendance :  $\chi^2$

Il permet de savoir s'il existe un effet entre les deux variables qualitatives croisées. Nous définissons l'hypothèse nulle "H0 : est-ce que le fait de connaître la classe d'âge d'un individu permet de supposer son sexe?".

La fonction "*chisq.test*" nous retourne les statistiques suivantes :

- Df : le nombre de degré de libertés ;
- p-value : donne la probabilité de validation de H0, c'est à dire la probabilité de ne voir aucun lien entre les variables. Plus p-value est petite (inférieur au seuil de 5%), plus il y a un lien entre les critères (et donc pas d'indépendance) ;
- X-squared : permet de retrouver manuellement la p-value en s'aidant d'un tableau disponible dans tout bon livre de statistique.

Pour nos variables, nous avons obtenu une p-value  $< 2.2e - 16$ , nous concluons qu'il y'a bien une liaison entre les variables. Cependant, nous avons également tester la dépendance entre l'âge (variable continue) et le sexe et nous obtenons une p-value de 0.004563. Donc nous pouvons dire que ce n'est pas le fait d'avoir fait les classes d'âges par sexe qui a permis de relier ces variables : cette liaison existait déjà.

V de cramer

Le test V de Cramer permet de comparer l'intensité du lien entre les deux variables étudiées. Le V de Cramer est la racine carrée du  $\chi^2$  divisée par le  $\chi_{max}^2$ . Ce  $\chi_{max}^2$  théorique est égal à l'effectif multiplié par le plus petit côté du tableau (nombre de lignes ou de colonnes) moins 1. Mathématiquement,

$$V = \sqrt{\frac{\chi^2}{\chi_{max}^2}} = \sqrt{\frac{\chi^2}{n \times [\min(l, c) - 1]}} \quad (2.1.3)$$

La statistique V prend des valeurs entre  $-1$  et  $1$  : plus V est proche de 0, moins les variables étudiées sont dépendantes.

Pour nos deux variables,  $V = 1$ , et donc nos deux variables sont complètement dépendantes.

---

En conclusion, Le Khi-2 ici nous indique donc qu'il existe une liaison entre les deux variables; le V de Cramer nous indique que cette liaison est très forte par sa valeur élevée.

#### 1.2.4 Test d'indépendance entre variables quantitatives et qualitatives

Pour tester la dépendance entre nos variables quantitatives (*PMR\_BASE*, *TAUX\_D'AIPP*) et nos variables qualitatives (*Age\_SurvNew*, *SEXE*), nous avons utilisé le test de Khi-2 une fois de plus. La même interprétation est valable ici :

	PMR_BASE & Age_SurvNew	PMR_BASE & SEXE	TAUX_D'AIPP & SEXE	TAUX_D'AIPP & Age_SurvNew
p-value	0.404	0.5	9.977e-08	< 2.2e-16

FIGURE 2.1.16: Tableau récapitulatif du test de Khi-2

Nous constatons que le montant de l'arrérage ne dépend ni du sexe ni de l'âge de la victime, mais de l'indice de gravité. Cependant, l'âge et le taux d'AIPP n'étaient pas corrélés (voir section 2.1.15), mais lorsque l'âge est transformé en variable catégorielle, ils le sont. Cela confirme toutes les analyses univariées et bivariées que nous avons fait précédemment.

Nous savons dès à présent comment nos variables se comportent (individuellement ou croisées). Dans la section suivante, nous allons tester quelques modèles pour prédire nos arrérages et ensuite choisir le meilleur modèle en fonction de certains critères : AIC, BIC,  $R^2$ ,  $R^2$  ajusté....

---

## 2. Méthodes de prédiction de l'arrérage de base

Afin de simuler des rentes futures, nous avons tout d'abord essayé d'expliquer le montant d'arrérage annuel de la rente à partir des caractéristiques biométriques de l'assuré : l'âge, le sexe et l'indice de gravité. Dans notre base de rentes Auto, nous avons 14 rentes dont le montant de l'arrérage est nul, nous les supprimons de la base d'étude : et donc notre base d'étude finale contient toutes les rentes dont le montant d'arrérage et le taux d'AIPP sont non nuls.

Nous avons utilisé et comparé plusieurs modèles : la régression linéaire, la régression log-linéaire, les modèles GLM, et enfin les arbres de prédiction.

### 2.1 Les variables explicatives

Dans la section 1.1.2, nous avons vu que le sexe et le montant de l'arrérage sont indépendants mais le taux d'AIPP était différent selon le sexe et l'âge : ces trois variables sont donc corrélées. De plus, dans les sections 1.2.3 et 1.2.4, nous avons vu que la variable sexe est corrélée aux deux autres variables : elle ne sera donc pas comptée parmi nos variables explicatives si nous ne faisons qu'un modèle. Enfin, en utilisant le sexe comme variable explicative dans un modèle, nous remarquons que l'erreur quadratique moyenne (RMSE) est plus élevée, tandis qu'elle est plus faible lorsque nous appliquons deux modèles (par sexe). Pour toutes ces raisons, nous avons décidé de faire des modèles par sexe.

Nos variables explicatives sont donc :  $Age\_SurvNew$  et  $TAUX\_AIPP$  et la variable réponse est  $PMR\_BASE$ , qui correspond l'arrérage de base.

### 2.2 Les différents modèles

Les différentes méthodes présentées dans cette section sont appliquées tant pour les hommes que pour les femmes. Les résultats présentés dans cette section sont ceux des hommes. Ceux des femmes se retrouvent en annexe A.3.

#### 2.2.1 La régression linéaire

Un modèle de régression linéaire est un modèle qui consiste à établir une relation linéaire qui peut exister entre une variable à expliquer  $Y$  (le montant de l'arrérage), mesurée sur une échelle continue et les variables explicatives  $X$  (âge, taux d'AIPP). Ces variables explicatives peuvent être continues ou catégorielles. La variable  $Age\_SurvNew$  étant catégorielle, nous avons donc 9 modalités pour les hommes et 8 chez les femmes. Chacune d'elle sera considérée comme une variable explicative dans notre modèle. Ainsi, pour chacun de nos deux modèles (par sexe), nous aurons  $p$  paramètres  $(\beta_0, \dots, \beta_p)$  à estimer : soit 10 chez les hommes et 9 chez les femmes.

#### Modèle théorique

Supposons qu'on dispose d'un échantillon de  $n$  observations de variables aléatoires  $(y_i, x_{i1}, \dots, x_{ip})$ . Le modèle théorique de régression linéaire s'écrit pour chaque individu  $i$  sous la forme :

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{i,j} + \epsilon_i \quad \forall i = 1, \dots, n \quad (2.2.1)$$

Avec :

- $p$  est le nombre de paramètres et varie en fonction du sexe ;
- $\beta = (\beta_0, \dots, \beta_p)$  inconnus : vecteur de paramètres à estimer à partir des observations ;
- $y_i$  et  $x_{i,j}$  sont fixes ;

- $\epsilon_i, i = 1, \dots, n$  représente les résidus observés pour chaque individu  $i$ , c'est-à-dire l'écart entre la valeur de  $Y$  observée et celle prédite par le modèle  $\hat{Y}$ . Dans le modèle, les variables  $\epsilon_i$  ne sont pas observées.

Matriciellement, l'équation 2.2.1 s'écrit :

$$Y = X\beta + \epsilon \quad (2.2.2)$$

$$\text{Avec } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

La première colonne de la matrice  $X$  sert à indiquer que la régression est effectuée avec constante.

L'estimation de  $\beta$  à partir des données observées se fait par la méthode des moindres carrés ordinaires (MCO) : elle consiste à estimer  $\beta$  en minimisant la somme des carrés des résidus ( $SSR$ ), telle que :

$$\hat{\beta}_{MCO} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \|y - X\beta\|_{I_n}^2 \quad (2.2.3)$$

Où  $\hat{\beta}$  est le vecteur des paramètres estimés.

L'utilisation d'un modèle linéaire requiert le calibrage du modèle et la vérification des hypothèses du modèle, qui sont :

- **(H1)** : Non colinéarité des variables explicatives ;
- **(H2)** : Les erreurs sont centrées, c'est à dire que  $\mathbb{E}(\epsilon|X_j) = 0$  ;
- **(H3)** : Indépendance des erreurs, c'est à dire que  $\forall i \neq j, \operatorname{cov}(\epsilon_i, \epsilon_j) = 0$  ;
- **(H4)** : L'échantillon  $(Y_j, X_{1,j}, \dots, X_{p,j}) ; j = 1, \dots, n$  est aléatoire, c'est-à-dire que les variables expliquées et explicatives sont indépendantes et identiquement distribuées entre les individus ;
- **(H5)** : Homoscédasticité du modèle, c'est à dire la variance des erreurs est constante. Si l'hypothèse précédente est vérifiée, on note  $\mathbb{E}(\epsilon_i^2|x_i) = \sigma^2$  ;
- **(H6)** : Normalité des résidus, c'est à dire qu'on suppose que les résidus suivent une loi normale centrée, de variance  $\sigma^2$  soit  $\epsilon_i|x_i \sim N(0, \sigma^2)$ .

Sous ces hypothèses, l'estimateur MCO, noté  $\hat{\beta}$  ou  $\hat{\beta}_{MCO}$  est

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2.2.4)$$

Les hypothèses de Gauss-Markov<sup>22</sup> et les hypothèses de normalité (décrites ci-dessus) garantissent des propriétés particulièrement intéressantes des estimateurs des coefficients de régression : estimateurs sans biais, convergents et qui ont la plus petite variance.

L'estimateur MCO coïncide avec l'estimateur du maximum de vraisemblance si la variable à expliquer  $Y$  suit une loi normale. Dans le cas où les résidus sont indépendants et identiquement distribués, l'estimateur des moindres carrés ordinaires est le plus efficace des estimateurs linéaires sans biais.

### Mise en oeuvre

Comme annoncé plus haut, tous les modèles présentés dans cette section sont fait en double et séparément : pour les hommes et pour les femmes. Nous n'afficherons que les résultats obtenus en utilisant la base des

<sup>22</sup> Le théorème de Gauss-Markov énonce que, parmi tous les estimateurs linéaires non-biaisés, l'estimateur par moindres carrés présente une variance minimale.

hommes. La même méthodologie et analyse sont faites chez les femmes et les résultats se trouvent en annexe A2.

Nous avons créé un échantillon d'apprentissage (80% de la base des hommes) et un échantillon test (20%). En utilisant la fonction *lm* du logiciel R sur les données d'apprentissage, nous avons régressé linéairement le montant de l'arrérage en fonction des classes d'âge des rentiers et de leur taux d'AIPP. A l'aide de la fonction *summary*, nous obtenons les résultats suivants :

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      238.70    3405.19   0.070  0.9441
Age_survNew[14-18] -8684.95    3473.46  -2.500  0.0127 *
Age_survNew[19-20]  -645.58    3702.92  -0.174  0.8617
Age_survNew[21-22]   910.90    3994.35   0.228  0.8197
Age_survNew[23-26] -4299.31    3530.95  -1.218  0.2239
Age_survNew[27-33] -4136.52    3530.29  -1.172  0.2418
Age_survNew[34-36] -7633.93    4264.94  -1.790  0.0740 .
Age_survNew[37-39] -4593.85    4331.10  -1.061  0.2893
Age_survNew40 et + -3896.36    3511.04  -1.110  0.2676
TAUX_AIPP          274.99      26.02  10.568  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19150 on 584 degrees of freedom
Multiple R-squared:  0.203,    Adjusted R-squared:  0.1907
F-statistic: 16.52 on 9 and 584 DF,  p-value: < 2.2e-16

```

FIGURE 2.2.1: Résultats de la régression linéaire chez les hommes

La figure ci-dessus nous donne en sortie :

- Les coefficients estimés et leur écart-type ;
- La valeur de la statistique *t de Student* : Test de nullité des coefficients des variables  $H_0 : \beta_i = 0$ . Cette hypothèse nulle signifie que cette variable  $X_i$  n'a aucun impact sur la variable à expliquer, et n'est donc pas significative ;
- La p-value du test associée à chaque coefficient : probabilité que le coefficient soit significativement différent de zéro. C'est le plus petit niveau à partir duquel on rejette l'hypothèse nulle du test ;
- La statistique *F de Fisher* (testant la significativité globale des variables), son degré de liberté, et la p-value associée ;
- Les coefficients de détermination  $R^2$  et  $R^2$  ajusté : permettant de tester la qualité du modèle. C'est la part de variance (de la réponse), expliquée par le modèle de régression.

Les étoiles permettent de préciser le niveau de significativité de chaque variable pour un seuil précis. De plus, le signe du coefficient associé à une variable indique le sens de l'effet de cette variable (lorsqu'elle est significative) sur la variable à expliquer.

Nous constatons que les seules variables significatives sont le taux d'AIPP et la classe d'âge de [14 – 18] ans (respectivement à 99% et à 95%). Toutes choses égales par ailleurs, nous pouvons dire que l'augmentation d'une unité de taux d'AIPP entraîne une augmentation de l'arrérage de 275€ environ et donc le montant de l'arrérage croît avec l'indice d'AIPP (résultat trouvé à la section 1.1.2 et même 1.2.2).

De plus, la p-value associée à la *F – statistic* (confère figure 2.2.1) étant très faible, nous rejetons fortement l'hypothèse  $H_0$ , à savoir le modèle est globalement significatif.

### Qualité prédictive du modèle

Afin d'évaluer la qualité de l'ajustement linéaire, nous allons calculer et analyser le coefficient de détermination le  $R^2/R^2$  ajusté. Par la suite, nous allons analyser les résidus afin de mieux comprendre ce résultat.

Le  $R^2$  est un critère qui permet de mesurer la part de variance de l'échantillon qui est expliquée par les variables explicatives du modèle. Il est toujours compris entre 0 et 1 : plus sa valeur est proche de 1, et plus l'adéquation entre le modèle et les données observées va être forte. Si elle vaut 1, alors les variables explicatives expliquent parfaitement la réponse, et 0, les variables X et Y ne sont pas linéairement corrélées.

La décomposition de la variation totale des observations s'écrit :  $SCT = SCR + SCE$  avec

$SCT = \sum_{i=1}^n (y_i - \bar{y})^2$ , représente la variation totale des observations  $y_i$  autour de leur moyenne  $\bar{y}$  ;

$SCR = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ , représente la variance résiduelle ou non expliquée (avec  $\hat{y}_i$  la valeur prédite) ;

$SCE = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ , représente la variance expliquée par la régression.

Ainsi, le  $R^2$  est défini par :

$$R^2 = \frac{SCE}{SCT} \quad (2.2.5)$$

Cependant, le  $R^2$  peut être trompeur pour évaluer la qualité d'un modèle car il augmente avec le nombre de prédicteurs. Nous allons de ce fait utiliser le  $R^2$  ajusté, qui est une version modifiée du  $R^2$ . Il est ajusté pour tenir compte du nombre de prédicteurs dans le modèle. Il est toujours inférieur au  $R^2$ .

Le  $R^2$  ajusté est défini par :

$$R^2_{ajust} = \frac{\frac{SCE}{n-p-1}}{\frac{SCT}{n-1}} = 1 - \frac{\frac{SCR}{n-p-1}}{\frac{SCT}{n-1}} \quad (2.2.6)$$

Avec notre régression, nous obtenons un  $R^2$  ajusté de 19%, ce qui signifie que seulement 19% de la variance de notre échantillon est expliquée par le modèle, et donc que celui-ci n'est peut être pas adapté à nos données.

Analysons les résidus : les hypothèses  $(H_1)$ ,  $(H_2)$  et  $(H_4)$  étant vérifiées pour nos données, testons celles restantes.

#### Test d'indépendance des résidus

Nous constatons sur la figure 2.2.2 que les résidus ne sont pas répartis de manière indépendante de part et d'autre de la droite d'équation  $y = 0$  (la moyenne des résidus est nulle dans notre cas), alors les résidus sont corrélés.

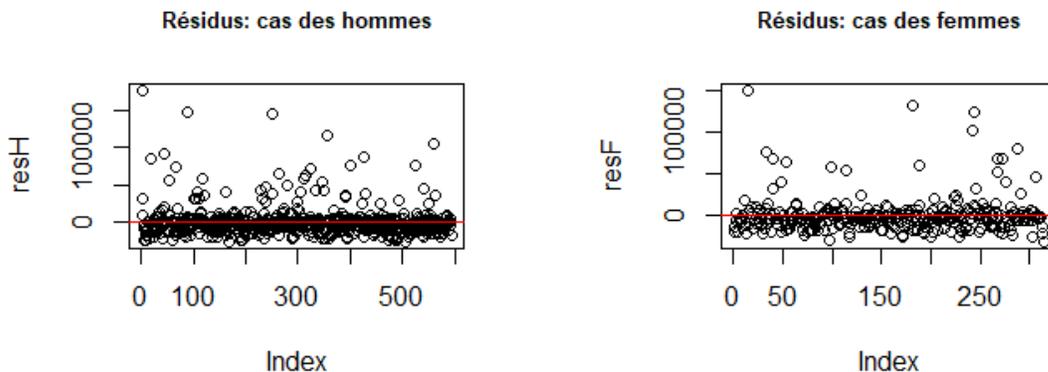


FIGURE 2.2.2: Graphe des résidus

---

### Test de normalité des résidus

Le QQ-plot schématisé à la figure 2.2.3 permet de représenter les quantiles de nos résidus observés et réduits, en fonction des quantiles de la loi normale centrée réduite.

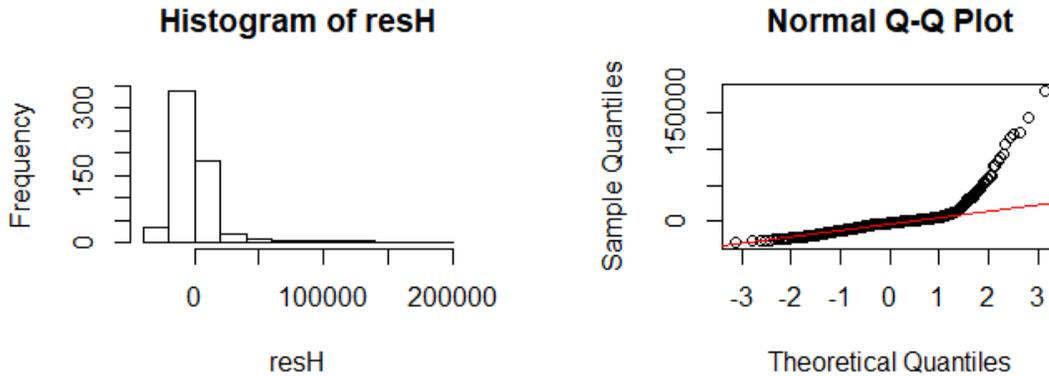


FIGURE 2.2.3: Histogramme et Normal Q-Q-plot des résidus : cas de la régression linéaire chez les hommes

Nous constatons que beaucoup de points ne sont pas alignés sur la première bissectrice (droite rouge) et donc les quantiles de notre échantillon après réduction, ne sont pas égaux aux quantiles de la loi normale centrée réduite. Nous avons effectué les tests de Shapiro-Wilk et de Jarque-Bera, dont l'hypothèse nulle est  $H_0$  : les résidus sont gaussiens. Les p-values de chacun de ces tests, étant inférieures à  $2.2e^{-16}$ , nous rejetons donc  $H_0$  à tous les niveaux de significativité.

Nous avons effectué les mêmes tests sur le modèle linéaire des femmes et nous en sommes arrivés à la même conclusion que celle précédente.

### Test d'homoscédasticité des résidus

Il s'agit ici de tester l'hypothèse ( $H_5$ ) requise pour un modèle linéaire. Si cette hypothèse est vérifiée, la variance des résidus estimée est indépendante des valeurs d'arrérages prédites par le modèle.

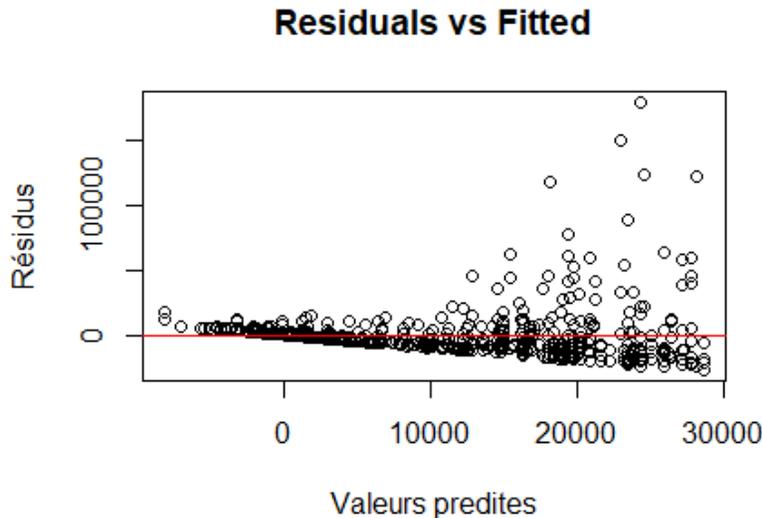


FIGURE 2.2.4: Résidus estimés par le modèle en fonction des valeurs prédites

Nous constatons sur la figure 2.2.4 que la dispersion verticale des points autour de 0 n'est pas la même pour toutes les valeurs de l'arrérage prédites, avec une variance qui croît avec la valeur prédite de l'arrérage. L'hypothèse de constance de la variance des résidus n'est donc pas vérifiée.

#### Adéquation de la distribution de l'arrérage à une loi Normale

Nous avons fait le test de Kolmogorov-Smirnov : c' est un test d'ajustement utilisé pour déterminer si un échantillon suit bien une loi continue donnée , ou bien si deux échantillons suivent la même loi. Ce test est basé sur les fonctions de répartition empiriques.

A partir de la fonction *ks.test* de R, nous avons effectué un test d'ajustement de notre variable à expliquer à une loi normale. Nous avons obtenu une p-value  $< 2.2e^{-16}$  : le montant de l'arrérage ne suit pas une loi normale.

**Conclusion** : Trois raisons importantes nous permettent de dire que la régression linéaire n'est pas adaptée pour notre modélisation :

- La principale et la plus importante est que la distribution de la variable à expliquer n'est pas compatible avec le modèle linéaire présenté ci-dessus. Exemple : notre modèle prédit des valeurs négatives d'arrérage. Par ailleurs, considérer la variable à expliquer (et son erreur) comme suivant une loi normale suppose une distribution symétrique autour de la moyenne, alors que ce n'est très probablement pas le cas ;
- L'autre raison est que, dans le modèle linéaire, les variables prédictives ont un effet linéaire sur la variable mesurée, or ces effets ne sont peut-être pas linéaires en réalité ;
- Enfin, les hypothèses  $(H_3)$ ,  $(H_5)$  et  $(H_6)$  ne sont pas vérifiées.

Afin d'améliorer les résultats, la solution la plus répandue consiste à trouver une transformation mathématique de la variable à expliquer.

---

## 2.2.2 La régression log-linéaire

Les montants de l'arrérage étant positifs, nous allons utiliser la régression log-linéaire, puisqu'elle permettrait d'obtenir des valeurs prédites strictement positives (contrairement à la régression linéaire). Elle est utilisée pour modéliser la relation entre une variable réponse et une ou plusieurs variables explicatives, en supposant que la variable réponse s'écrit comme le **logarithme** d'une fonction affine des variables explicatives.

Mathématiquement, soit  $f$  une fonction positive (qui est ici la fonction  $\log$ ). Nous allons régresser  $f(Y)$  sur la matrice des observations  $X$ , soit :

$$f(Y) = X\beta + \epsilon \quad (2.2.7)$$

Logiquement, les prédictions de l'arrérage seront positives, vu qu'elles seront composées par la fonction inverse de  $f$  : la fonction exponentielle dont la valeur est dans  $]0; \infty[$  :

$$\hat{Y} = f^{-1}(X\hat{\beta}) = \exp(X\hat{\beta}) \quad (2.2.8)$$

NB : La fonction  $\log$  permet d'avoir un effet croisé des variables explicatives sur la variable réponse (nous passons d'un modèle additif à un modèle multiplicatif). On a donc :

$$\hat{Y}_j = \exp(\beta_0) \times \prod_{i=1}^p \exp(X_{i,j}\hat{\beta}_i) \quad \forall j = 1, \dots, n \quad (2.2.9)$$

L'équation de régression du modèle log-linéaire s'écrit :

$$\log(Y) = X\beta + \epsilon \quad (2.2.10)$$

Posons  $Z = \log(Y)$ , nous pouvons donc écrire  $\mathbb{E}(Z|X) = X\hat{\beta}$  (d'après l'équation 2.2.10). Le problème d'estimation se réduit à estimer  $\beta$ , non plus par MCO mais par la méthode dite de « maximum de vraisemblance ».

Les résidus d'une régression linéaire étant supposés gaussiens,  $Z|X \sim N(\mu, \sigma^2)$  et donc  $Y|X$  suivent une loi log-normale de paramètres  $(\mu, \sigma^2)$ , dont la densité par rapport à la mesure de Lebesgue est donnée par :

$$f(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log(y) - \mu)^2}{2\sigma^2}\right), \quad \forall y \in \mathbb{R}_+, \quad \forall \mu, \sigma \in \mathbb{R} \quad (2.2.11)$$

Nous savons que :

$$\begin{cases} \mathbb{E}(Z|X) = X\hat{\beta} = \mu & (1) \\ \mathbb{E}(Y|X) = e^{\mu + \frac{\sigma^2}{2}} & (2) \end{cases}$$

En remplaçant (1) dans (2), on a  $\mathbb{E}(Y|X) = \mathbb{E}(\exp(Z)|X) = e^{X\hat{\beta} + \frac{\sigma^2}{2}}$  qui est bien évidemment **différent** de  $\exp(\mathbb{E}(Z|X))$ .

Une fois nos prédicteurs linéaires  $X\hat{\beta}$  obtenus, nous devons non seulement les passer à l'exponentielle mais aussi multiplier cette exponentielle par  $\exp(\frac{\sigma^2}{2})$  pour avoir nos valeurs d'arrérages prédites  $\mathbb{E}(Y|X) = e^{X\hat{\beta} + \frac{\sigma^2}{2}}$ , où  $\sigma^2$  est la variance estimée de  $\log(Y)$ .

### Mise en oeuvre

Nous conservons les mêmes variables explicatives que dans le modèle précédent (l'âge de survenance découpé en classe d'âge et le taux d'AIPP). Nous avons régressé  $\log(Y)$  sur l'ensemble des variables explicatives et nous obtenons les estimateurs convergents suivants :

---

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.289388   0.228407  27.536 < 2e-16 ***
Age_survNew[14-18] -0.952512   0.232986  -4.088 4.95e-05 ***
Age_survNew[19-20] -0.531553   0.248377  -2.140 0.03276 *
Age_survNew[21-22] -0.451104   0.267925  -1.684 0.09278 .
Age_survNew[23-26] -0.945670   0.236842  -3.993 7.36e-05 ***
Age_survNew[27-33] -0.721616   0.236798  -3.047 0.00241 **
Age_survNew[34-36] -0.817523   0.286076  -2.858 0.00442 **
Age_survNew[37-39] -0.805093   0.290513  -2.771 0.00576 **
Age_survNew40 et + -0.413611   0.235506  -1.756 0.07957 .
TAUX_AIPP      0.041814   0.001745  23.957 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.284 on 584 degrees of freedom
Multiple R-squared:  0.5489,    Adjusted R-squared:  0.5419
F-statistic: 78.95 on 9 and 584 DF,  p-value: < 2.2e-16

```

FIGURE 2.2.5: Résultats de la régression log-linéaire chez les hommes

Nous constatons que presque toutes les variables sont significatives, contrairement au modèle linéaire, où la variable taux d'AIPP était la seule variable significative. De plus, l'arrérage décroît avec l'âge et croît avec le taux d'AIPP (résultats que nous avons obtenu à la section 1.1.2).

### Qualité prédictive du modèle

Le  $R^2$  ajusté est de 54%, ce qui est bien supérieur à celui obtenu avec le modèle linéaire classique (19%) : le modèle log-linéaire est meilleure que le modèle linéaire (selon ce critère).

Cependant, nous ne pouvons pas comparer la régression linéaire classique à la régression log-linéaire en utilisant le critère d'information d'Akaike (en anglais AIC) car la variable à prédire n'est pas la même.

Le modèle log-linéaire, étant la modélisation linéaire de la variable  $\log(Y)$ , vérifions les hypothèses du modèle linéaire.

#### Test d'indépendance des résidus

Traçons les résidus issus de la régression.

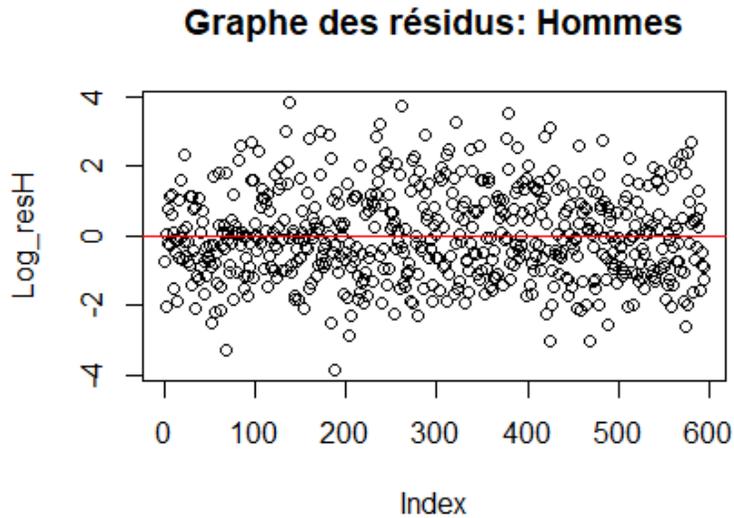


FIGURE 2.2.6: Résidus de la régression log-linéaire chez les hommes

D'après la figure 2.2.6 ci-dessus, nous constatons que les résidus sont indépendants, malgré le fait que certains points semblent corrélés entre eux.

#### Test de normalité des résidus

La moyenne de nos résidus est de  $-6.160605e-17$  : les résidus sont donc bien centrés. Sur la figure 2.2.7, nous avons schématisé la densité de nos résidus en utilisant un histogramme ainsi que le QQplot (permet de représenter les quantiles observés de nos résidus réduits en fonction des quantiles de la loi normale centrée réduite).

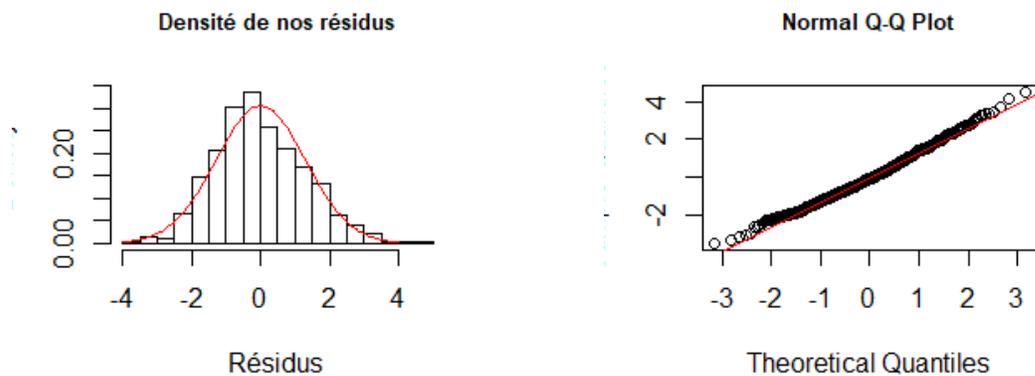


FIGURE 2.2.7: Histogramme et Normal QQ-plot des résidus : cas de la régression log-linéaire chez les hommes

Nous constatons que la densité théorique (courbe en rouge) s'ajuste bien à la densité empirique de nos résidus. De plus, presque tous les points sont situés sur la première bissectrice (droite rouge) : nos résidus sont donc gaussiens.

---

### Test d'homoscédasticité des résidus

Schématisons les résidus estimés en fonction des valeurs de  $\log(\text{arrérage})$  prédites par le modèle.

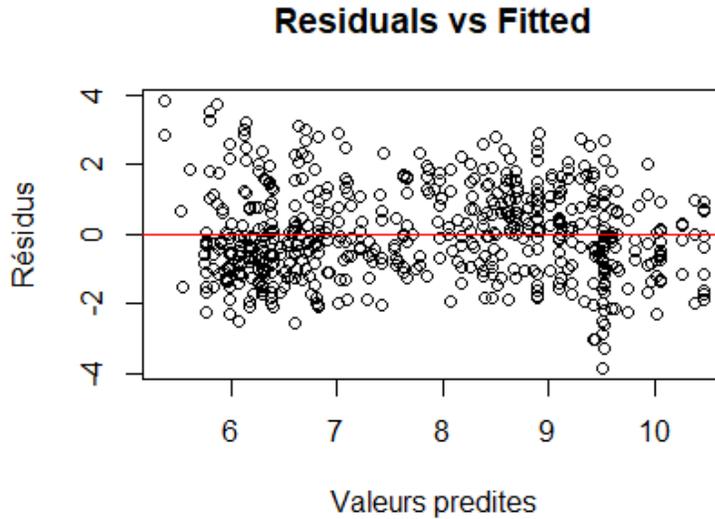


FIGURE 2.2.8: Résidus estimés en fonction des valeurs prédites

Nous constatons que les points sont quasiment repartis autour de 0 : l'hypothèse de constance des résidus est donc vérifiée.

La section suivante a pour objectif de trouver un modèle qui nécessite peu d'hypothèses et une loi plus adaptée à nos données.

### 2.2.3 Les Modèles Linéaires Généralisés (GLM for Generalized Linear Models)

Nous avons vu que le modèle linéaire général repose sur une hypothèse forte : **le terme d'erreur suit une loi normale centrée, de même variance** et que la distribution de la variable dépendante est normale. Mais, ce n'est toujours pas le cas. Les modèles GLM sont donc une extension du modèle linéaire classique : ils introduisent dans la structure du modèle un degré de non linéarité. Ils permettent de modéliser des distributions de résidus non-gaussiens.

L'idée des modèles GLM est d'utiliser une transformation mathématique sur la variable réponse  $Y$  en tenant compte cette fois-ci de la véritable distribution de la variable à expliquer. On considère que pour une variable aléatoire à expliquer  $Y$ , il existe une relation de la forme :

$$g(\mathbb{E}[Y|x_1, \dots, x_p]) = \sum_{j=1}^p \beta_j X_{ij} \quad (2.2.12)$$

entre les  $p$  variables explicatives  $X_i (i = 1, \dots, p)$  et l'espérance conditionnelle de  $Y$ . La fonction  $g$ , strictement monotone, dérivable et défini de  $[0, 1]$  vers  $\mathbb{R}$ , est appelée **fonction de lien** du modèle. Elle permet de déterminer la relation qui existe entre le prédicteur linéaire  $X\beta$  et l'espérance conditionnelle  $\mathbb{E}(Y|X)$ . Elle a pour but de contraindre les valeurs prédites à être dans l'échelle des valeurs observées. Pour obtenir la prédiction moyenne, il est nécessaire d'appliquer la fonction de lien inverse :

$$\mu_y = g^{-1}(\eta) \tag{2.2.13}$$

Où  $\mu_y = \mathbb{E}[Y|x_1, \dots, x_p]$  et  $\eta = \sum_{j=1}^p \beta_j X_{ij}$

Il reste donc à déterminer la distribution de la variable réponse. Les lois les plus couramment utilisées dans les GLM sont : la loi Normale, la loi de Poisson, la loi Binomiale, et la loi Gamma. Ainsi, on impose donc à  $Y$  d'appartenir à la famille de loi dite **exponentielle**, dont la densité s'écrit :

$$f_{\beta, \phi}(y) = \exp\left(\frac{y\beta - b(\beta)}{\phi} + c(y, \phi)\right) \tag{2.2.14}$$

Avec  $b()$  une fonction définie sur  $\mathbb{R}$ , deux fois dérivable et de dérivée première injective,  $c()$  une fonction définie sur  $\mathbb{R}^2$  (c'est à dire indépendante de  $\beta$ ),  $\beta$  le paramètre d'intérêt et  $\phi$  le paramètre de dispersion. Les lois précédemment citées appartiennent bien à la famille exponentielle (lorsque les fonctions  $a(), b(), c()$  et les paramètres  $\phi, \beta$  sont bien choisis).

Pour choisir un modèle GLM, il faut :

- choisir la loi de  $Y|X = x$  dans la famille exponentielle des GLM ;
- choisir une fonction de lien inversible  $g$ .

Détermination de la loi de la variable réponse

Nous avons vu à la section 1.1 que notre variable d'intérêt suivait une loi Gamma, qui décrit les variables continues positives ou nulles. L'espérance et la variance d'une variable aléatoire  $Y$  qui suit une famille exponentielle s'écrivent :

$$\begin{cases} \mathbb{E}(Y) = \mu = b'(\beta) & (1) \\ Var(Y) = b''(\beta)\phi & (2) \end{cases}$$

De l'équation de la moyenne (1), on a :  $\beta = b'^{-1}(\mu)$ . En remplaçant  $\beta$  dans l'équation de la variance (2), on obtient que  $Var(Y) = b''(b'^{-1}(\mu)) \times \phi$

Posons  $V(\mu) = b''(b'^{-1}(\mu))$ . Cette fonction  $V()$  appelée **fonction de variance**, ne dépend que de l'espérance de la variable réponse et permet de passer de l'espérance à la variance. Ainsi, à partir de la forme de la variance de la variable d'intérêt, on peut identifier la loi GLM la plus appropriée pour la modélisation.

Le tableau ci-dessous nous donne les fonctions de variance des lois classiques de la famille exponentielle :

Loi	Fonction de variance
Normale	1
Poisson	$\mu$
Gamma	$\mu^2$
Binomiale	$\mu * (1 - \mu)$

TABLE 2.1: Fonctions de variance des lois classiques de la famille exponentielle

Notons  $\mu_j = \mathbb{E}(Y_j|X_{.,j})$  la moyenne de  $Y$  pour une observation  $j$ . Nous pouvons donc écrire la variance de chaque observation  $Var(Y_j|X_{.,j}) = \phi \times \frac{V(\mu_j)}{w_j}$ , avec  $w_j$  le poids de chaque observation  $j$ .

Nous cherchons ainsi à exprimer la variance du montant de l'arrérage en fonction de son espérance. En supposant que  $w_j = 1, \forall j = 1, \dots, n$ , on a donc :

$$Var(Y|X) = \phi \times \mathbb{E}(Y|X)^r \iff \ln(Var(Y|X)) = \ln(\phi) + r \ln(\mathbb{E}(Y|X)) \tag{2.2.15}$$

Par application, nous obtenons une estimation du coefficient  $r = 2$  et la constante  $\phi = 3.952675$ . Nous avons l'équation  $Var(Y|X) = 3.952675 \times \mathbb{E}(Y|X)^2$ .

De plus, selon la littérature, la loi Gamma est la plus adaptée pour modéliser la sévérité de sinistres.

**On va donc modéliser l'arrérage par une fonction de loi Gamma.**

*Choix de la fonction de lien*

Le choix de la fonction de lien est l'étape la plus délicate et fondamentale d'un GLM, car celle-ci permet de définir la structure du modèle. Elle doit être adaptée en fonction du domaine de définition de  $Y$ . Contrairement aux modèles linéaires classiques, les valeurs prédites par le prédicteur linéaire ne correspondent pas à la prédiction moyenne d'une observation, mais à la transformation (par une fonction mathématique) de celle-ci.

Chacune des loi usuelles appartenant à la famille exponentielle possède une fonction de lien spécifique, appelée **fonction de lien canonique**. Elle permet de relier l'espérance  $\mu_y$  au paramètre naturel  $\beta$  tel que  $f(\mu_y) = \beta$ . Or,  $\mu_y = b'(\beta) \implies \beta = b'^{-1}$ . Les fonctions de lien canonique des lois classiques telles que la loi Normale, la loi de Poisson, la loi Binomiale, la loi Gamma sont :

Loi	Lien naturel
$N(\mu, \sigma^2)$	Identité : $\eta = \mu$
$B(\mu)$	logit : $\eta = \ln\left(\frac{\mu}{1-\mu}\right)$
$P(\mu)$	log : $\eta = \ln(\mu)$
$G(\mu, \beta)$	inverse : $\eta = \frac{1}{\mu}$
$IN(\mu, \lambda)$	inverse <sup>2</sup> : $\eta = \frac{1}{\mu^2}$

TABLE 2.2: Fonctions de lien des lois classiques de la famille exponentielle

Toutefois, il est possible de choisir n'importe quelle fonction de lien bijective  $g$  vérifiant  $g(\mu) = X\beta$ . Nous testerons donc la fonction de lien **inverse**, qui est le lien naturel de la loi Gamma mais aussi la fonction de lien **log**, puisque les valeurs de notre variable réponse sont strictement positives.

**Mise en oeuvre**

Nous continuons d'utiliser les mêmes variables sélectionnées plus haut. Cependant, supposer que la loi de la variable réponse n'est plus gaussienne implique que les coefficients de régression ne sont plus estimés par la méthode des MCO mais par la méthode dite de « maximum de vraisemblance » (qui ne coïncide pas avec la méthode des MCO dans le cas non gaussien).

En maximisant la log-vraisemblance du modèle avec un modèle GLM de loi Gamma et la fonction de lien inverse, nous obtenons des estimateurs dont la valeur et la significativité sont représentées dans la figure 2.2.9 ci-dessous :

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.398e-04  3.323e-05  10.226 <2e-16 ***
Age_survNew[14-18] 2.179e-05  1.736e-05   1.255  0.210
Age_survNew[19-20] -6.155e-06  1.195e-05  -0.515  0.607
Age_survNew[21-22] -6.450e-06  1.158e-05  -0.557  0.578
Age_survNew[23-26] 1.917e-05  1.833e-05   1.046  0.296
Age_survNew[27-33] 1.852e-05  2.558e-05   0.724  0.470
Age_survNew[34-36] 3.013e-05  3.072e-05   0.981  0.327
Age_survNew[37-39] 4.109e-05  5.444e-05   0.755  0.451
Age_survNew40 et + -5.561e-06  1.151e-05  -0.483  0.629
TAUX_AIPP      -3.231e-06  3.290e-07  -9.822 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 3.062459)

Null deviance: 1817.7 on 593 degrees of freedom
Residual deviance: 1253.5 on 584 degrees of freedom
AIC: 11642

```

FIGURE 2.2.9: Résultats du modèle GLM Gamma avec fonction de lien inverse

Nous constatons que seule la variable *TAUX\_AIPP* est significative. Les modalités de la variable *Age\_surv* ne sont pas du tout significatives. Cela signifie que seule la variable *TAUX\_AIPP* permet d'expliquer la variable d'intérêt.

Par contre, en utilisant la fonction de lien log dans le modèle GLM Gamma, nous obtenons les résultats suivants :

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    7.346633   0.281523  26.096 < 2e-16 ***
Age_survNew[14-18] -0.788756   0.287166  -2.747  0.00621 **
Age_survNew[19-20] -0.216882   0.306137  -0.708  0.47895
Age_survNew[21-22] -0.113746   0.330231  -0.344  0.73064
Age_survNew[23-26] -0.730537   0.291919  -2.503  0.01260 *
Age_survNew[27-33] -0.709525   0.291865  -2.431  0.01536 *
Age_survNew[34-36] -1.028558   0.352602  -2.917  0.00367 **
Age_survNew[37-39] -0.815424   0.358071  -2.277  0.02313 *
Age_survNew40 et + -0.452987   0.290273  -1.561  0.11917
TAUX_AIPP      0.035701   0.002151  16.595 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 2.506335)

Null deviance: 1721.3 on 593 degrees of freedom
Residual deviance: 990.6 on 584 degrees of freedom
AIC: 11467

```

FIGURE 2.2.10: Résultats du modèle GLM Gamma avec fonction de lien log

Avec cette modélisation, nous constatons que certaines modalités de la variable *Age\_surv* sont devenues significatives (certaines à un moindre niveau que d'autres). À priori, ce modèle est meilleur que le précédent. Cependant, pour valider un modèle GLM, on ne s'appuie pas uniquement sur la significativité des coefficients mais aussi sur d'autres paramètres tels que la déviance, le critère d'AIC, l'allure des résidus ou par confrontation "modélisé / empirique".

### Qualité prédictive des deux modèles GLM

Il s'agit ici de mesurer la qualité de chacun des modèles GLM présentés ci-dessus et de choisir le modèle le plus pertinent.

---

### Critère AIC

Le critère d'information d'Akaike (en anglais AIC) permet d'évaluer la qualité d'un modèle en se basant sur le maximum de vraisemblance. Il permet de pénaliser les modèles en fonction du nombre de paramètres afin de satisfaire le critère de parcimonie.  $AIC = 2 * \ln L(\beta) + 2k$  où  $k$  est le nombre de paramètres à estimer du modèle et  $\ln L(\beta)$  est la log vraisemblance du modèle. Le meilleur modèle correspond à celui avec un AIC faible.

L'AIC du modèle Gamma avec fonction de lien inverse est de 11 642 tandis que celui avec fonction de lien log est de 11 467. Selon ce critère, le modèle Gamma avec fonction de lien log est le plus pertinent.

NB : Nous ne pouvons pas comparer l'AIC du modèle Gamma et l'AIC de la régression log-linéaire car la variable prédite par chacun de ces modèles est différentes ( $Y$  pour le modèle Gamma et  $\log(Y)$  pour le modèle log-linéaire). Par contre, nous pouvons comparer l'AIC du modèle Gamma avec celui du du modèle linéaire classique = 13 510. Nous confirmons donc que le modèle Gamma est mieux adapté à nos données.

### Estimation des paramètres de déviance

La déviance représente l'écart entre la log-vraisemblance du modèle estimé et la log-vraisemblance d'un modèle saturé c'est-à-dire le modèle possédant autant de paramètres que d'observations. Par définition, elle est égale à :  $D = -2 \times (\ln L(Y|Y) - \ln L(\hat{\mu}|Y))$ .  $D$  est positif et « petit » pour un modèle de bonne qualité.

La déviance du modèle GLM Gamma avec fonction de lien inverse est de 1 253 tandis que celle avec fonction de lien log est de 990. Le modèle GLM Gamma avec fonction de lien log est plus pertinent.

Cet indicateur global est en pratique complété par une analyse observation par observation ; cette analyse se base souvent sur l'analyse des résidus.

En se basant sur le critère AIC et sur la déviance, nous pouvons dire que le modèle GLM Gamma log est plus performant que le modèle GLM Gamma inverse.

## 2.2.4 Les arbres de régression de type CART

Dans cette sous-partie, nous allons utiliser l'une des méthodes non paramétriques pour prédire la valeur de la variable réponse  $Y$  en fonction des valeurs explicatives  $X$  : les arbres de régression de type CART. Contrairement à la régression linéaire (ou non-linéaire) qui nécessite de formuler des hypothèses sur la structure et la distribution des données, les arbres ne formulent qu'une seule hypothèse : les données à prédire  $Y$  sont générées de façon identique et indépendante par un processus  $P$  à partir du vecteur des variables explicatives  $X$ .

Comme nous l'avons expliqué à la section 1.2.1, CART est un algorithme de construction d'arbres binaires qui effectue un partitionnement récursif des données, puis estime un modèle très simple dans chaque élément de la partition (feuilles de l'arbre). A chaque noeud, la condition de coupure permettant de créer les deux classes, sera déterminée à partir d'une des variables explicatives, appelée variable de segmentation. Celle-ci peut être différente à chaque itération.

Pour prédire  $Y$  pour un individu dont on connaît les valeurs  $X_1, \dots, X_p$ , on procède étape par étape de la manière suivante : en partant de la racine, à chaque noeud, on vérifie si la condition de coupure est vérifiée ou pas : si la condition est vérifiée, on se dirige vers la branche associée à la réponse "Oui" (répondant à la question implicite "Est-ce la condition est vérifiée?"), sinon, on se dirige vers la branche associée à la réponse "Non". À la dernière étape, on aboutit alors à une seule feuille de l'arbre. Ainsi, une valeur plausible de  $Y$  est la moyenne des valeurs de  $Y$  associées aux individus de cette feuille.

Les classes sont telles que leurs moyennes soient les plus éloignées possible et les valeurs de la variable à prédire soient les plus proches possible : il s'agit donc soit de maximiser la variance inter-classe ou de minimiser la variance intra-classe.

L'algorithme s'arrête soit lorsqu'on a atteint le nombre d'observation minimal dans les feuilles ou lorsque

---

le critère d'erreur minimal est respecté.

### Mise en oeuvre

Nous cherchons à régresser le montant de l'arrérage en fonction des variables explicatives. A l'aide de la fonction *"rpart"* de R, nous avons tracé dans un premier temps l'arbre maximal. Cependant, cet arbre maximal possède une grande variance et donc son erreur de généralisation est très grande. De plus, le nombre de feuilles terminales est trop grand et inexploitable. Dans le but d'obtenir plus de précision et de réduire l'erreur de prédiction, nous allons élaguer l'arbre maximal.

L'élagage consiste à créer une suite de sous-arbres à l'aide d'un critère pénalisé. Nous allons donc utiliser le paramètre de complexité, noté *cp*. Nous allons tracer le taux d'erreur de prédiction  $\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$  en fonction du *cp* et nous allons choisir le *cp* qui minimise ce taux d'erreur.

La figure 2.2.11 illustre l'erreur relative en fonction du *cp* (cas des hommes). Nous constatons que le paramètre de complexité qui minimise cette erreur est de 0.076.

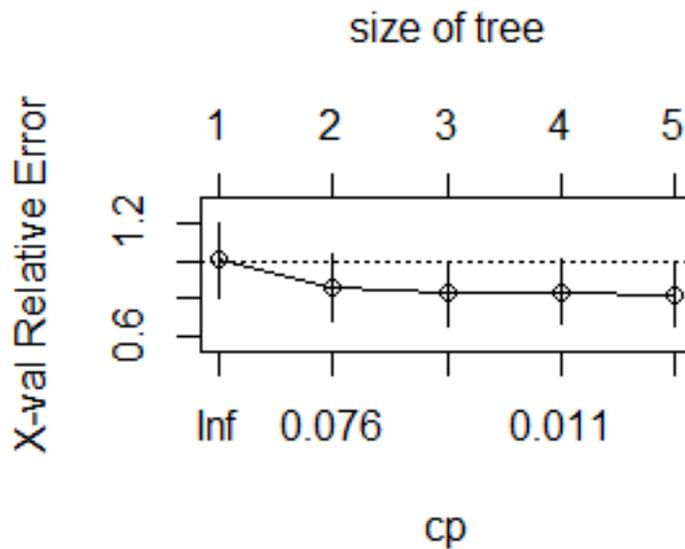


FIGURE 2.2.11: Erreur relative en fonction du paramètre de complexité

Nous allons donc utiliser ce *cp* pour tracer l'arbre optimal (figure 2.2.12), qui ne contient que 2 noeuds.

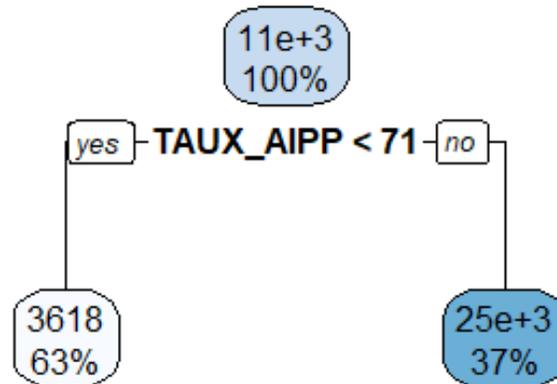


FIGURE 2.2.12: Arbre de régression de l'arrérage

Dans chaque classe de l'arbre, sont indiqués la proportion de rentes respectant le critère de division et le montant moyen d'arrérage sur ces rentes. Il nous permet d'avoir une idée du nombre de rentes sur lequel est modélisé l'arrérage et les variables explicatives utilisées pour cette modélisation. Ainsi, nous constatons que le montant de l'arrérage ne dépend pas de l'âge mais juste du taux d'AIPP. Il ne fait de distinction qu'entre un taux d'AIPP inférieur ou supérieur à 71 pour estimer le montant de l'arrérage. Hors, nous avons vu sur le graphique 2.1.12 que l'arrérage croît avec le taux d'AIPP. De plus, nous savons que plus un individu est jeune et a un taux d'AIPP élevé, plus élevé sera son indemnisation. Nous pouvons déjà dire que l'arbre de régression sera inadapté pour modéliser notre montant d'arrérage.

## 2.3 Choix du meilleur modèle

Nous avons vu que le modèle linéaire classique n'était pas adapté pour la modélisation du montant l'arrérage car non seulement les hypothèses d'un modèle linéaire ne sont pas vérifiées mais aussi ce modèle prédit des valeurs négatives de l'arrérage. Par la suite, nous avons effectué une modélisation de  $\log(\text{arrerage})$  sur nos variables explicatives et nous vu que celle-ci était plus adaptée que la précédente (hypothèses d'un modèle linéaire vérifiées et plus grande part de variance expliquée par ce modèle).

Cependant, notre variable d'intérêt n'étant pas gaussienne, nous avons opté pour des modèles GLM, qui ont la particularité d'utiliser la réelle distribution de la variable réponse. Pour nos données, nous avons trouvé que la loi Gamma est celle qui s'ajuste bien aux montants d'arrérage. Toutefois, nous avons testé ce modèle GLM Gamma avec des fonctions de lien Inverse et Log. Nous sommes arrivés à la conclusion selon laquelle le modèle GLM avec fonction de lien Log est plus pertinent que le modèle GLM avec fonction de lien inverse (AIC et déviance plus faibles dans le premier cas ainsi que l'allure des résidus plus satisfaisants). Nous avons également constaté que les variables explicatives étaient plus significatives dans le modèle log-linéaire que dans les modèles GLM.

Nous avons par la suite testé les arbres de régression de type CART. Ils ont l'avantage (devant les modèles GLM) de permettre de détecter et modéliser des structures de dépendance entre les variables, en posant comme unique hypothèse le fait que les valeurs à prédire doivent être générées de façon indépendante et identique par un processus  $P$  à partir du vecteur des variables explicatives.

L'objectif de cette section est d'évaluer la qualité prédictive de nos différents modèles et d'en sélectionner le meilleur selon les critères que nous présenterons ci-après. Pour que chacun de ces critères soit pertinent,

nous avons scindé notre base en deux échantillons : un échantillon d'apprentissage (constitué de 80% de la base initiale tirée de manière aléatoire) sur lequel nous effectuons notre modélisation et un échantillon test, sur lequel nous testons nos modèles.

- **L'erreur quadratique moyenne (RMSE)** : il s'agit de calculer la racine carrée de la moyenne arithmétique des écarts entre les valeurs d'arrérages prédites par notre modèle et celles observées. Elle est défini comme suit :

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^n (\widehat{ar\grave{e}rage}_i - ar\grave{e}rage_i)^2} \quad (2.2.16)$$

Avec  $n$  la taille de notre échantillon et  $\widehat{ar\grave{e}rage}_i$  qui représente les valeurs d'arrérages prédites par le modèle considéré.

- **L'erreur absolue moyenne (MAE)** : il s'agit de calculer la moyenne arithmétique des valeurs absolues des écarts entre les valeurs d'arrérages prédites par notre modèle et celles observées sur un échantillon choisi. Mathématiquement, l'erreur moyenne de prédiction est défini comme suit :

$$MAE = \frac{1}{n} \times \sum_{i=1}^n |\widehat{ar\grave{e}rage}_i - ar\grave{e}rage_i| \quad (2.2.17)$$

## Résultats

Les tableaux ci-dessous présentent les résultats des RMSE et des MAE calculés sur l'échantillon d'apprentissage et l'échantillon test et classés du meilleur au moins bon.

### Cas des hommes

Modèle	Ech. test	Ech. App	Modèle	Ech. test	Ech. App
Régression log-linéaire	19 318	20 504	Régression log-linéaire	9 572	8 873
GLM avec lien log	19 952	20 615	GLM avec lien log	11 800	9 613
GLM avec lien inverse	20 426	20 811	Arbre CART	12 661	10 219
Arbre CART	20 707	21 720	GLM avec lien inverse	13 136	10 365
Régression linéaire	21 580	22 912	Régression linéaire	13 302	11 038

TABLE 2.3: Résultats du RMSE vs MAE des hommes

### Cas des femmes

Modèle	Ech. test	Ech. App	Modèle	Ech. test	Ech. App
GLM avec lien log	26 984	19 506	GLM avec lien log	10 251	11 688
Régression log-linéaire	27 673	19 561	Régression log-linéaire	10 873	12 029
Arbre CART	28 467	19 750	Arbre CART	11 067	12 443
GLM avec lien inverse	30 589	22 350	GLM avec lien inverse	11 506	12 504
Régression linéaire	32 481	25 245	Régression linéaire	12 038	13 189

TABLE 2.4: Résultats du RMSE vs MAE des femmes

---

Nous constatons que le modèle le moins bon pour estimer le montant d'arrérage tant chez les rentiers hommes que femmes est la régression linéaire classique car le RMSE (ainsi que le MAE) sur l'échantillon test est la plus élevée. Par ailleurs, le meilleur modèle pour les rentiers hommes est la régression log-linéaire tandis que chez les femmes, le modèle GLM avec fonction de lien log est meilleur. Nous allons retenir ces modèles dans la suite de notre étude.

#### *Limite de ces modèles*

Nos modèles ne peuvent prédire qu'une faible part de la variance de l'arrérage car nous ne disposons pas de toutes les informations qu'utilise le régleur pour estimer le montant d'arrérage de chaque victime. De nombreux éléments sont traités et pris en compte au cas par cas en fonction de la situation personnelle, professionnelle et des besoins de la victime. Bien qu'avec les dernières réglementations le cadre d'indemnisation est plus précis, les variables nécessaires ne sont pas disponibles. De plus, il y'a l'aléa juridique qui vient également justifier la limite nos modèles.

---

## 3. Modélisation stochastique de la sinistralité future

Il est question dans cette partie de simuler les rentes futures, rattachées à la survenance prochaine, pour les sinistres Auto corporels graves. Nous utiliserons une approche *fréquence*  $\times$  *sévérité*, en se basant sur les caractéristiques biométriques de la victime : son sexe, son âge et son indice de gravité. Cette méthode consiste à modéliser indépendamment le nombre annuel de rentes (la fréquence), et la charge annuelle de rentes (la sévérité).

Cette partie a pour but de mesurer in fine le risque de prime auquel sont exposés les sinistres Auto corporels graves.

### 3.1 Modélisation de la fréquence de sinistres des rentes

Comme nous l'avons vu sur la figure 2.1.1, le nombre de rentiers par année de survenance a connu beaucoup de fluctuations. Celui-ci n'est devenu significatif qu'entre les années  $t_1 = 1962$  et  $t_2 = 2004$ . Pour les années de survenance très récentes, la plupart des sinistres n'est peut être pas encore passée en rente. Et donc, le plus judicieux pour nous est de considérer les sinistres entre  $t_1$  et  $t_2$ .

Le nombre annuel moyen de sinistres sur notre intervalle d'années de survenance se calcule comme suit :

$$\mathbb{E}(N) = \frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} N_t \quad (2.3.1)$$

Où  $N_t$  est le nombre de sinistres survenus au cours de l'année  $t$ .

Sur notre portefeuille de rentes, nous avons en moyenne 24 sinistres par an.

Cherchons la loi de la fréquence de sinistres. Nous allons tester les lois de comptage suivantes : loi de Poisson et la loi Binomiale Négative.

#### Loi de Poisson

La loi de Poisson est une loi de probabilité discrète qui prend des valeurs positives et permet de décrire le comportement du nombre d'événements se produisant dans un intervalle de temps fixé, si ces événements se produisent avec une fréquence moyenne  $\lambda$ , et indépendamment du temps écoulé depuis l'événement précédent. Celle-ci ne dépend que du seul paramètre  $\lambda$ , qui est un réel strictement positif. La fonction de masse d'une variable aléatoire  $X$  qui suit une loi de Poisson s'exprime par :

$$p(k) = \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad (2.3.2)$$

La loi de Poisson est une loi équi-dispersée car  $\mathbb{E}(X) = V(X) = \lambda$ .

#### Loi Binomiale Négative

La loi binomiale négative est une distribution de probabilité discrète qui prend des valeurs positives. Elle permet de décrire une expérience qui consiste en une série de tirages aléatoires indépendants, donnant un « succès » avec probabilité  $p$  et un « échec » avec une probabilité complémentaire  $q = 1 - p$ . Cette expérience se poursuit jusqu'à l'obtention d'un nombre donné  $n$  de succès.

La loi binomiale négative dépend de 2 paramètres :  $n$  qui est entier naturel non nul et  $p$ , une probabilité comprise entre 0 et 1. La fonction de densité d'une variable aléatoire  $X$  distribuée selon une loi Binomiale Négative de paramètres  $(n, p)$  est définie comme suit :

$$f(k; n; p) = \binom{k+n-1}{k} p^n q^k \quad (2.3.3)$$

L'espérance et la variance de cette variable  $X$  s'écrivent :

$$\begin{cases} \mathbb{E}(X) = \frac{nq}{p} \\ V(X) = \frac{nq}{p^2} \end{cases}$$

Par la méthode des moments, nous estimons les paramètres de cette loi à partir des estimateurs empiriques de l'espérance et de la variance :

$$\begin{cases} \hat{n} = \frac{\widehat{\mathbb{E}(X)}^2}{\widehat{V(X)} - \widehat{\mathbb{E}(X)}} \\ \hat{p} = \frac{\widehat{\mathbb{E}(X)}}{\widehat{V(X)}} \end{cases} \quad (2.3.4)$$

Avec  $\widehat{\mathbb{E}(X)} = \bar{X}$  et  $\widehat{V(X)} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$

Notons par ailleurs que les estimateurs du maximum de vraisemblance coïncident avec les estimateurs des moments.  $p$  étant compris entre 0 et 1, la variance d'une loi binomiale négative est strictement supérieure à sa moyenne : cette loi est dite "sur-dispersée".

Sur notre échantillon de rentiers, nous obtenons les estimateurs  $\hat{n} = 8.845$  et  $\hat{p} = 0.266$ . La variance du nombre annuel de sinistres est 91.375. Celle-ci étant bien plus élevée que sa moyenne, la loi binomiale négative est adaptée pour modéliser le nombre annuel de rentes.

Pour confirmer ou infirmer cette théorie, nous avons effectué le test de Kolmogorov-Smirnov (à l'aide de la fonction *ks.test* de R) qui permet de tester l'ajustement de nos données à l'une des lois décrites ci-dessus. Nous avons obtenu une p-value égale à 0.6267 pour la loi Binomiale Négative et 0.02646 pour la loi de Poisson. Au seuil de  $\alpha = 5\%$ , nous concluons en la faveur de la loi Binomiale Négative.

### Mise en oeuvre

Le modèle interne non-vie Groupama utilise également la loi binomiale négative pour modéliser les sinistres corporels Auto et dont les caractéristiques sont :  $\mathbb{E}(X) = 38$  et  $V(X) = 91.97$ .

Sur notre échantillon de rentiers, nous avons en moyenne 24 sinistres qui sortiront en rentes (soit 63% environ de sinistres du modèle interne). Et donc pour être cohérent et comparable avec le modèle interne non-vie, nous allons supposer que 14 sinistres en moyenne sortiront en capital (qui représente environ 37% des sinistres). Le nombre de sinistres qui sortiront en rente sera modélisé par la loi binomiale négative de paramètres  $\hat{n} = 8.845$  et  $\hat{p} = 0.266$ . La différence entre la fréquence de sinistre du modèle interne et celle des sorties en rentes, représentera la fréquence des sinistres hors rente.

## 3.2 Modélisation du coût du sinistre : sévérité

Afin de modéliser le montant d'arrérage annuel de chaque rente, à partir duquel nous calculerons le montant de provision mathématique de chaque rente ( $PM_i^{rentes}$ ), nous avons utilisé les modèles : GLM avec fonction de lien log (pour les femmes) ; régression log-linéaire (pour les hommes) , et nous procéderons comme suit :

- Calibrer la loi de chacune des variables explicatives : sexe, âge et taux d'AIPP ;
- Prédire le montant d'arrérage à partir des modèles de régression trouvés à la section 2.3 et des variables explicatives simulées ;

---

— Calculer la charge annuelle de sinistres.

Comme évoqué à la section 1, nous ne travaillons qu’avec les rentes viagères (soit 95% des rentes de notre base) d’arrérage non nul. Nous faisons une modélisation tête par tête, pour avoir un modèle plus fin et détaillé.

### 3.2.1 Simulation des variables explicatives

Il est question dans cette partie de créer un générateur de lois qui va nous permettre simuler des victimes en gardant les mêmes caractéristiques que notre portefeuille initial. Pour ce faire, nous allons simuler chacune des trois variables explicatives conditionnellement à celle(s) précédemment simulée(s). Dans les faits, nous simulons d’abord le sexe, ensuite nous simulons l’âge en fonction du sexe et enfin le taux d’AIPP en fonction de l’âge et du sexe. Nous avons généré 300 000 victimes.

#### Calibration de la loi du sexe

Sur notre échantillon, nous avons 35% de femmes et 65% d’hommes. Nous avons donc utilisé une loi de Bernouilli de paramètre  $p$ , qui vaut 1 lorsque le rentier est un homme et 0 s’il s’agit d’une femme.  $p$  est alors estimé par la proportion empirique d’hommes dans notre échantillon, soit 65%.

#### Calibration de la loi de l’âge en fonction du sexe

Il s’agit de construire une loi de distribution des âges, conditionnellement au sexe. Pour ce faire, nous avons décidé de nous baser sur les effectifs de chaque modalité d’âges. Nous avons calculé les fréquences empiriques de chaque modalité d’âges. Celles-ci représentent les probabilités d’occurrences de chaque modalité.

A l’aide de ces probabilités, nous avons construit la fonction de répartition de la variable âge. Pour simuler un rentier d’âge  $x$  et de sexe  $y$ , nous faisons un tirage aléatoire de loi uniforme sur  $[0,1]$  et en fonction du résultat, nous cherchons le quantile associé. Ce quantile correspond l’âge de cet individu.

#### Calibration de la loi du taux d’AIPP en fonction du sexe et de l’âge

Connaissant le sexe et l’âge d’un individu, nous pouvons générer le taux d’AIPP correspondant. Nous procédons comme pour la simulation des âges : nous calculons les fréquences empiriques de chaque modalité de taux d’AIPP par âge et par sexe ; ensuite nous construisons la fonction de répartition de la variable taux d’AIPP. Pour un individu de sexe  $y$  et d’âge  $x$ , nous faisons un tirage aléatoire de loi uniforme sur  $[0,1]$  et en fonction du résultat, nous cherchons le quantile associé. Ce quantile correspond à l’indice de gravité de cet individu.

### 3.2.2 Simulation du montant d’arrérage

Une fois nos variables explicatives simulées, nous obtenons un triplet de variables aléatoires corrélées (sexe, âge, AIPP) qu’on va mettre dans le modèle de prédiction de l’arrérage qui va nous permettre d’estimer le montant de l’arrérage. Pour ce faire, nous créons tout d’abord les classes d’âges (avec les âges simulés), ensuite nous régressons l’arrérage observé sur l’ensemble des variables explicatives simulées. Nous utilisons les modèles retenus à la section 2.3.

Nous effectuons ensuite des tests de convergence (test de moyenne), pour se rassurer qu’en moyenne, notre portefeuille simulé a les mêmes caractéristiques que le portefeuille initial. Par exemple, nous comparons la moyenne et la variance des variables suivantes dans les deux portefeuilles (initial et simulés) : âge, taux d’AIPP, l’arrérage moyen par classe d’âges. Nous faisons cette comparaison par sexe.

Nous avons constaté que les valeurs sont très proches : les résultats sont plutôt satisfaisants. Nous constatons également que les montants d’arrérages simulés augmentent avec les taux d’AIPP simulés, comme nous l’avions vu sur la figure 2.1.12.

Sur la base totale (sans distinction de sexe), nous obtenons les statistiques moyennes suivantes :

	Valeurs moyennes simulées	Valeurs moyennes observées
Age	27.8952	27.8776
Taux d'AIPP	52.9794	52.9738
Arrérage	11 842	12 807

TABLE 3.1: Statistiques comparatives entre les valeurs simulées et les valeurs observées

Notre modélisation étant faite par sexe, nous avons également fait ces comparaisons par sexe. Les résultats se trouvent dans les tableaux A.1 et A.2, se trouvant en annexe A.4

Une fois les montants d'arrérages simulés, nous pouvons dès à présent calculer la charge annuelle de sinistres tête par tête. Dans les faits, elle se décompose en trois parties : une partie rente (pour les sinistres qui sortiront en rente), une partie capital (pour les sinistres qui sortiront en capital) et une partie frais de soins. Dans le cadre de ce rapport, nous allons créer deux parties, que nous allons modéliser de façon indépendante :

- Une partie rente, à laquelle nous rajouterons la partie frais de soins ;
- Une partie hors rente, à laquelle nous rajouterons la partie frais de soins modélisée par sa loi de distribution.

### 3.2.3 Simulation de la charge annuelle de rente

Il s'agit de calculer la charge annuelle des sinistres qui sortiront en rentes. Comme dit plus haut, nous calculerons dans un premier temps le BE de rentes, ensuite nous modéliserons la partie frais de soins. La somme de ces deux montants constituera la charge annuelle de rente.

#### Modélisation de la provision Best Estimates de rente

Sous Solvabilité 2, la provision BE correspond à l'actualisation des cash flows futurs de paiement de la rente à l'instant  $t=0$  (instant de calcul de la rente). Elle est estimée sur la base de la courbe des taux sans risque pertinent au  $31/12/N$  (courbe de l'EIOPA) représenté en annexe A11 et des tables d'évolutions démographiques. Vu que nous modélisation les rentes, nous devons revaloriser le montant de la rente.

Pour un individu d'âge  $x$  à  $t=0$ , le BE  $y$  compris FGAO se calcule comme suit :

$$BE_x = Arrerage \times (1 + \text{taux}_{gestion}) \times \sum_{i=1}^{\infty} \frac{p_{x,x+i} \times (1 + \text{taux}_{revalo})^i}{(1 + \text{taux}_{actu,i})^i} = Arrerage \times E_x \quad (2.3.5)$$

Avec *Arrerage*, le montant qui lui est versé annuellement (puisque nous simulons uniquement des rentes fractionnées annuellement) ;  $p_{x,x+i} = \frac{l_{x+i}}{l_x}$ , la probabilité qu'un individu d'âge  $x$  soit vivant dans  $i$  années, estimée à partir d'une table de mortalité bien définie ;  $\text{taux}_{revalo}$  le taux de revalorisation des rentes ;  $\text{taux}_{actu,i}$  le taux d'actualisation des flux versés la  $i^{eme}$  et  $\text{taux}_{gestion}$  le taux de frais de gestion.

Pour le calcul de cette provision, nous faisons les hypothèses suivantes :

- Les tables de mortalité utilisées sont les tables TH et TF 00-02 (annexe A.7) ;
- Le taux de frais de gestion de 1.09% ;
- La revalorisation future des rentes a été intégrée dans les calculs à un taux de 2% par décret suite à la réforme du FGAO (confère section 3.2) ;
- La courbe des taux utilisée pour l'actualisation est la courbe sans risque (courbe EOIPA). Ces taux économiques comprennent la *volatility adjustment*.

---

— Fractionnement de la rente : Annuel

### Modélisation de la partie frais de soins

Il s'agit dans cette sous section de déterminer le poids du montant de frais de soins de chaque sinistre (en fonction du sexe et de l'âge de la victime) sorti en rente dans le coût total et ensuite l'appliquer à nos rentes simulées afin de déterminer la charge totale annuelle de rente par sinistre .

Pour ce faire, nous disposons d'une base de rentes futures contenant les principales informations suivantes : âge de la victime, sexe, montant des PM par poste de préjudice et le CTP (coût total prévisionnel). Pour chaque sinistre, nous déterminons la partie hors rente (qui est la différence entre le CTP et la PM totale de rente) ainsi que sa proportion dans le CTP (ratio partie hors rente sur CTP).

Nous utilisons par la suite ces informations pour créer un tableau à double entrée : classe d'âge  $\times$  sexe, à partir duquel nous créons des classes de montants de rentes (en utilisant les arbres CART). Ainsi, pour un sexe, une classe d'âge et une classe de montant de rente donnée, nous calculons le poids moyen (en pourcentage) de la partie fixe, que nous appliquons ensuite à nos rentiers simulés afin de déterminer la charge annuelle :

$$CTP = PM + \text{frais\_de\_soins} = PM + x\%.CTP \Leftrightarrow CTP(1 - x\%) = PM \Rightarrow \boxed{CTP = \frac{PM}{1 - x\%}}$$

Ainsi, la partie frais de soins simulée est la différence entre la charge annuelle de rente et la PM simulée.

Ainsi, pour chaque sinistre simulé, nous connaissons les caractéristiques de la victime, son montant d'ar-rérage annuel et la charge totale annuelle de rente. Cette charge inclut la revalorisation par l'inflation (au taux de 2%).

### 3.2.4 Simulation de la charge annuelle hors rente

Il s'agit de calculer la charge annuelle des sinistres qui sortiront en capital. Comme dit plus haut, nous calculerons dans un premier temps le capital constitutif, ensuite nous modéliserons la partie frais de soins. La somme de ces deux montants constituera la charge annuelle de la partie hors rente.

#### Modélisation du capital constitutif

Il dépend des caractéristiques biométriques de la victime est estimée sur la base d'une table d'évolution démographiques , d'un taux d'actualisation (différent de celui utilisé dans le calcul de la PM de rente). Ces sinistres sont non revalorisés.

Pour un individu d'âge  $x$  à  $t=0$ , le capital constitutif (CC) est calculé de la manière suivante :

$$CC = Arrerage \times \sum_{i=1}^T \frac{p_{x,x+i}}{(1 + \text{taux}_{actu})^i} = Arrerage \times E(x) \quad (2.3.6)$$

Avec *Arrerage*, le montant qui lui est versé annuellement ;  $p_{x,x+i}$ , la probabilité qu'un individu d'âge  $x$  soit vivant dans  $i$  années, estimée à partir d'une table de mortalité bien définie et  $\text{taux}_{actu}$  le taux d'actualisation des flux versés.

Pour le calcul de ce CC, nous faisons les hypothèses suivantes :

- Les tables de mortalité utilisées sont les tables TH et TF 00-02 (annexe A.7) ;
- Le taux d'actualisation est de 0.50% (Barème Gazette du Palais) ;
- Fractionnement de la rente : Annuel

---

### Modélisation de la partie frais de soins

Nous avons utilisé la même base de rentes futures que nous avons utilisé pour modéliser la partie frais de soins des rentes. Nous avons déterminé la loi de la partie hors rente (par classe d'âge et par sexe) : elle correspond à la loi log-normale (de paramètres différents). Le test de Kolmogorov-Smirnov nous a confirmé qu'il y'a bien adéquation entre la distribution de la variable hors rente et la loi log-normale. Nous avons ensuite ajouté ce montant de frais de soins au montant de CC. La somme de ces deux montants constitue la charge annuelle des sorties en capital.

NB : Vu que nous modélisons les sinistres futurs, nous inflatons de 5% la charge de chaque sinistre (rente et hors rente).

#### 3.2.5 Base des sinistres graves Auto corporels

Vu que nous voulons modéliser les sinistres Auto corporels graves, nous avons dans un premier temps, simulés 300 000 rentiers (âge, sexe, AIPP, arrérage). Ensuite nous avons scindé cette base en deux : une base rente (représente 63% de la base globale) et une base hors rente. Dans la base de rente, nous avons calculé la charge annuelle de rente tête par tête (y compris FGAO et hors FGAO) et dans la base de hors rente, nous avons calculé la charge annuelle hors rente ligne à ligne (qui est non revalorisée). Pour chacune de ces bases, nous avons constitué une base de sinistres graves en sélectionnant les sinistres dont la charge annuelle est supérieure ou égale au seuil des sinistres graves Auto du modèle interne, soit 1 200 000€.

Cette base nous servira d'appui pour le calcul du SCR de primes.

### 3.3 Calcul du SCR de Primes brut de réassurance

Le SCR de primes correspond au capital économique nécessaire pour faire face au risque non-vie de prime à horizon 1 an. Pour calculer ce SCR, nous avons simulé  $M= 10\,000$  scénarios de sinistres futurs à partir de la base de sinistres graves (rentes et hors rentes). Pour chaque scénario, la charge totale annuelle des sinistres graves pour chaque année  $i$  est égale à :

$$S_i = \sum_{k=1}^{N_i^R} X_{i,k}^R + \sum_{k=1}^{N_i^{HR}} X_{i,k}^{HR} \quad (2.3.7)$$

Avec :

- $N_i^R$  (respectivement  $N_i^{HR}$ ), la variable aléatoire représentant le nombre de sinistres graves de rentes (respectivement hors rentes) qui surviendront pendant l'année  $i$  ;
- $X_{i,k}^R$  (respectivement  $X_{i,k}^{HR}$ ), la variable aléatoire de la charge du sinistre de rente (respectivement hors rentes)  $k$ , de l'année  $i$  brut de réassurance vu à l'ultime, y compris inflation future.

De plus, nous supposons que  $N$  et  $X_i$  sont indépendantes et les  $X_i$  indépendantes entre elles.

La charge annuelle moyenne des sinistres (pour tous les  $M$  scénarios) est donc égale à :

$$S_{moy} = \frac{1}{M} \sum_{i=1}^M S_i \quad (2.3.8)$$

Avec  $S_i$  la charge annuelle totale du scénario  $i$ .

Le SCR de prime brut de réassurance, d'un point de vu modèle interne non-vie Grouopama se calcule de la manière suivante :

$$SCR_{primes} = VaR_{99,5\%} \quad (2.3.9)$$

---

Où VaR est la la Value-at-Risk, définie par :

$$VaR_\alpha = \inf\{x \in \mathbb{R}, P(L > x) \leq 1 - \alpha\}$$

C'est le quantile d'ordre  $\alpha$  de la distribution de perte L.

---

Troisième partie

Analyse des modèles et résultats

---

# 1. Comparaison du modèle développé avec le modèle existant

Les deux approches étant différentes, nous allons tout d'abord présenter l'approche de modélisation du modèle interne, par la suite nous allons comparer les résultats obtenus avec les deux modélisations. Nous appliquerons ensuite les traités de réassurance à ces sinistres, afin d'obtenir la charge annuelle nette de réassurance ainsi que le SCR de primes net de réassurance. Puis, nous ferons une sensibilité en modifiant ces traités de base afin de choisir le meilleur en fonction de plusieurs indicateurs. Nous verrons également l'impact du FGAO sur les provisions et nous ferons enfin quelques tests de sensibilité.

## 1.1 Approche de modélisation des sinistres Auto corporels graves par le modèle interne

Pour modéliser les sinistres corporels Auto graves, le modèle interne non-vie Groupama utilise une approche *fréquence*  $\times$  *sévérité*, calibré comme suit :

### 1.1.1 Calibration de la loi de sévérité

Le calibrage de la sévérité grave est réalisé sur les coûts individuels «As If»<sup>23</sup> survenance 2019 des sinistres graves retenus des survenances 1996 à 2018. Un sinistre est retenu si sa charge dépasse le seuil as if 2019 de 1.2 M€(seuil calculé à partir du seuil as if N-1 inflaté de l'inflation corporelle, qui est de 5%). Ces charges sont estimées en utilisant comme taux d'actualisation, le **taux social** : 60% du TME sur les 24 derniers mois + 10 pb, soit 0.37% en 2019.

Les distributions suivantes sont testées sur ces sinistres graves : la loi Bêta, la loi Gamma, la loi Lognormale, la loi de Pareto à deux paramètres, la loi de Pareto simple et la loi de Weibull. Toutes ces distributions sont ajustées par maximum de vraisemblance, à l'exception de la distribution Bêta qui est ajustée par méthode des moments. Au cours du calibrage, les statistiques d'ajustement suivantes sont produites : la statistique d'Akaike, la statistique des moindres carrés, la statistique de Kolmogorov-Smirnov, la statistique de Kuiper, la statistique d'Anderson-Darling. La distribution retenue pour la modélisation sera celle qui satisfait aux critères suivants :

- Elle doit présenter, au sens statistique, une bonne qualité d'ajustement sur la partie centrale de la distribution, c'est-à-dire des moments, des mesures de statistiques calculées et des QQ-plot ;
- Elle doit également refléter une démarche prudentielle, c'est-à-dire qu'elle doit permettre d'intégrer une marge de prudence de la sinistralité calibrée dans sa queue de distribution.

La distribution retenue pour calibrer les sinistres Auto corporels graves est la loi de **Pareto Simple**.

### 1.1.2 Calibration de la loi de fréquence

La loi des nombres de sinistres graves est déterminée par business unit (caisses régionales et GAN Assurances), en procédant par étape comme suit :

**Étape 1** : Projection à l'ultime des nombres de sinistres graves historiques

Les triangles des nombres de sinistres graves sont construits à partir des coûts de sinistres ultimes as if survenance 2019. Les coefficients de passage utilisés sont les coefficients de Chain-Ladder classiques.

**Étape 2** : Détermination des nombres ultimes de sinistres graves modélisés

---

23. c'est-à-dire de retraiter les données pour les rapporter à la situation économique et à l'exposition de l'année de cotation

---

A partir des nombres de sinistres historiques, vus à fin 2018, nous appliquons les facteurs de développement déterminés plus haut, afin d'en déduire les nombres ultimes de sinistres graves.

Ces ultimes sont ensuite pondérés par des indices d'exposition, afin d'en déduire des ultimes réajustés du volume de portefeuille. Le calcul de ces indices est basé sur l'indice d'évolution tarifaire quand l'information est disponible, afin d'appréhender l'inflation spécifique au risque considéré. A défaut, le calcul se base sur l'indice d'inflation historique retenu. Les volumes de primes sont également pris en compte pour le calcul de ces indices d'exposition.

Enfin, nous réalisons une sélection différenciée des survenances sur lesquelles sont déterminés les moyennes et les écarts-types des nombres ultimes de la survenance de modélisation.

### Étape 3 : Le calibrage en lui-même

Il est réalisé en deux étapes : la sélection de la distribution de modélisation et la détermination des paramètres de cette loi.

En effet, vu qu'à l'étape précédente nous avons obtenu les moyennes et écarts-types, nous faisons un ratio variance/moyenne. S'il est supérieur à 1, alors la loi Binomiale-Négative est choisie. Et par la suite, nous pourrions déterminer les paramètres de cette loi (soit par maximum de vraisemblance, soit par la méthode des moments).

Le modèle interne utilise ainsi la loi **Binomiale Négative** pour calibrer la loi de fréquence.

## 1.2 Comparaison des résultats

L'approche de modélisation présentée dans ce mémoire permet de challenger celle du modèle interne actuel, tout en corrigeant leurs limites. Le point commun de ces deux modèles est l'approche de modélisation : approche *fréquence × sévérité*, mais modélisé différemment sur les points suivants :

- Méthode de simulation des variables explicatives sexe, âge et taux d'AIPP : elles sont simulées de manière indépendantes dans le modèle actuel tandis que dans notre modélisation, ces variables sont corrélées ;
- Mode d'indemnisation des sinistres : dans le nouveau modèle, nous distinguons les sinistres qui sortiront en rentes et ceux qui sortiront en capital. Tandis que dans le modèle actuel, aucune distinction n'est faite ;
- Méthode de calibration de la loi de sévérité : le modèle interne calibre la loi de sévérité des sinistres graves en se basant sur l'historique et par la suite fait une adéquation de loi. Tandis que dans cette étude, le coût de sinistre est une fonction des caractéristiques biométriques de la victime ;
- Taux d'actualisation : le modèle interne utilise comme taux d'actualisation le taux social pour modéliser les charges de sinistres, tandis que ce nouveau modèle utilise la courbe des taux EOIPA pour les sinistres qui sortiront en rente et le taux barème gazette du palais pour ceux qui sortiront en capital.

Cependant, pour être comparable avec le modèle actuel et déterminer les sources de volatilité entre les modèles, nous avons décidé de modéliser une deuxième charge en utilisant le taux social comme taux d'actualisation. Le tableau ci-dessous nous donne un récapitulatif des résultats obtenus par le modèle interne et par notre étude :

	Modèle développé		Modèle actuel
	Utilisation du taux social* pour l'actualisation	Utilisation du taux S2** / taux BGP*** pour l'actualisation	Utilisation du taux social* pour l'actualisation
Coût moyen de sinistres	8 306 346	5 254 512	5 703 677
Coût moyen annuel de sinistres	319 700 440	202 239 338	219 865 270
SCR brut	724 935 125	424 754 653	506 132 608
SCR net	596 607 966	299 882 855	345 614 270

\* : taux social = 60% TME moyen sur les 24 derniers mois + 10 pb

\*\* : taux S2= courbe des taux EOIPA

\*\*\* : taux GBP = taux du Barème Gazette du Palais

FIGURE 3.1.1: Résultats obtenus via le nouveau modèle vs ceux du modèle actuel

En utilisant le taux social comme taux d'actualisation dans le nouveau modèle, nous constatons une augmentation de +45% environ sur le coût moyen de sinistres par rapport à celui du modèle actuel. Cette hausse peut s'expliquer par le fait que le modèle actuel ne capte ni l'effet biométrie ni l'effet non linéaire dans les variables explicatives sur la partie FGAO. Cet impact est le même sur le SCR brut de réassurance. Cette absence de linéarité provient des tirages aléatoires dépendants que nous avons effectués sur nos variables explicatives.

Cependant, à taux d'actualisation différents, les charges moyennes de sinistres ne sont pas très éloignées : l'impact du taux n'est pas très grand en terme de coût moyen (8% d'écart environ). Par contre, sur le SCR brut de réassurance, l'écart est de 19% environ : 8% de cet écart peut s'expliquer par le coût moyen, le reste est probablement dû à la dispersion des charges. Ci-dessous le tableau représentant la volatilité des charges.

	Modèle développé		Modèle actuel
	Taux social	Taux S2 / taux BGP	Taux social
Volatilité du coût	11 343 330	6 105 454	7 311 292

FIGURE 3.1.2: Dispersion des charges en fonction du modèle et du taux utilisé

Nous pouvons constater que la volatilité des coûts est plus élevée lorsque nous utilisons le taux social comme taux d'actualisation dans le nouveau modèle : de ce fait, la dispersion constatée n'est pas liée aux sinistres en eux-même, mais plutôt au taux d'actualisation utilisé lors de la modélisation.

Certes notre modélisation est plus précise, détaillée et permet un gain de SCR (ce qui est très important pour les assureurs d'ailleurs), le modèle actuel reste tout de même prudent malgré ses limites car les écarts observés ne sont pas si importants.

Comme nous l'avons dit à la section 3.4, le risque RC corporel Auto, étant une branche à développement long, l'une des solutions pour limiter l'immobilisation de capital est la **réassurance**. D'une manière simpliste, on pourrait définir la réassurance comme « l'assurance de l'assurance ». Dans le chapitre suivant, nous allons présenter les traités de réassurance utilisés au sein du modèle interne. Nous pourrions in fine calculer la charge moyenne cédée ainsi que le prix de la réassurance et le SCR net de réassurance.

---

## 2. La réassurance

Le montant d'indemnisation de la garantie RC corporelle étant très souvent illimité en France, les programmes de réassurance utilisés sont pour la plupart du temps des traités non proportionnels, particulièrement les traités XS (individuels et/ou agrégés).

Un seul traité prend rarement en charge la protection de toute une année de survenance : il est généralement décomposé en plusieurs tranches avec une possibilité d'avoir des tranches avec cessions illimitées. De plus, la notion de « **reconstitution** » (en abrégé REC) est très importante en réassurance : «<sup>24</sup> Chaque fois qu'un sinistre met en jeu l'engagement du réassureur en tout ou en partie, le contrat de réassurance peut prévoir que le montant de la garantie ainsi épuisé sera automatiquement reconstitué, avec ou sans paiement d'une prime de réassurance additionnelle. »

### 2.1 De la sinistralité brute à la sinistralité nette de réassurance

Le modèle interne applique les traités XS individuels et agrégés sur les charges de sinistres. Afin d'être comparable avec eux, nous allons également appliquer les mêmes traités.

Nous appliquerons les traités suivants :

- Le traité XS individuel : composé des tranches T1 : 3M€ XS 12M€ avec 9 REC gratuites, T2 : Illimité XS 15M€ avec REC illimitées et gratuites. Une clause de rachat est prévu pour ce traité au taux de 1,50%.
- Le traité XS Aggregate : il s'agit d'un traité XS classique mais qui s'applique sur la charge annuelle de sinistres. Ce traité est le suivant : 105M€ XS 195M€. Il comporte les clauses suivantes :
  - une franchise de 2 000 000€ pour chaque sinistre ;
  - un placement du traité au maximum à 90%.

#### Mise en oeuvre

Comme dit précédemment, la charge annuelle de sinistres brute de réassurance est obtenue en actualisant les cash-flows soit avec la courbe des taux EOIPA pour les sinistres en rentes , soit avec le taux du barème de la Gazette du Palais pour les sinistres en capital et en y rajoutant une partie frais de soins. Pour calculer la charge annuelle nette de réassurance, nous allons procéder en deux étapes :

#### Une vision sinistre par sinistre

Nous calculons les éléments suivants tête par tête :

- Les charges de sinistres non revalorisées : calculées en utilisant le taux de réassurance qui vaut 1,50% comme taux d'actualisation ;
- Les charges de sinistres revalorisées au taux de 2% : calculées en utilisant le taux social qui vaut 0,37% comme taux d'actualisation ;
- Les charges cédées sur chacune des tranches : Soit  $X_i$  la charge du  $i^{eme}$  sinistre calculée avec le taux de réassurance, la charge cédée au réassureur noté  $S_i$  pour ce sinistre  $i$  est  $S_i = \min(\max(X_i - \text{priorite}, 0), \text{portee})$  ;
- La charge cédée nette XS : elle correspond à la différence entre la charge de sinistre actualisée au taux social et revalorisée et les cessions XS des deux tranches cumulées ;

---

24. source : [www.argusdelassurance.com](http://www.argusdelassurance.com)

- 
- La charge de sinistre individuelle qui servira pour appliquer le traité XS Aggregate : nous l'appellerons *Aliment Aggregate individuel*. Cette charge correspond à la charge cédée nette XS avec la franchise appliquée.

### Une vision annuelle

Pour chaque année simulée, nous allons calculer les éléments suivants :

- La charge de sinistre annuelle à partir de la charge *Aliment Aggregate individuel* : il s'agit d'agréger la charge individuelle de sinistre par année ;
- La charge annuelle cédée pour le traité Aggregate ;
- Les charges annuelles cédées sur chacune des tranches ;
- La charge annuelle de sinistres nette de réassurance : elle correspond à la la différence entre la charge annuelle brute de réassurance et les charges annuelles cédées sur chacun des traités.

Chacune des sorties de la vision annuelle est sous forme de vecteur à  $M = 10\,000$  lignes. Nous pourrons ainsi calculer :

- Le prix théorique de la réassurance pour chaque traité (sauf pour le traité XS Aggregate) : elle correspond à la prime pure du traité ajoutée d'un chargement de sécurité. Mathématiquement, on a :

$$PR = PP + \alpha\% \times \sigma \quad (3.2.1)$$

Avec  $PR$  le prix théorique de la réassurance,  $PP$  la prime pure (correspond à la moyenne de la charge annuelle cédée du traité considéré) ,  $\alpha$  le taux de chargement de sécurité ( $\alpha = 25\%$  par hypothèse) et  $\sigma$  , l'écart-type de la charge annuelle cédée du traité considéré.

- Le SCR de primes net de réassurance, qui correspond au quantile à 99,5% du vecteur de charge annuelle de sinistres nette de réassurance ;

## 2.2 Application

Sur notre base de victimes simulées, nous avons appliqué les traités de réassurance précédents et nous avons obtenu les primes de réassurance suivantes :

	Primes estimées	Primes réellement cédées
Traité 3 XS 12	3 569 391	3 896 265
Traité -1 XS 15	5 749 316	8 319 240

TABLE 2.1: Comparaison des résultats : Prix de la réassurance

Nous constatons que les primes estimées avec le nouveau modèle sont très proches de celles réellement cédées au réassureur. Les charges individuelles de sinistres, brutes de réassurance, estimées par notre modèle sont correctes : notre modèle est fiable.

Par ailleurs, nous obtenons un SCR net de réassurance de 299 882 855€ tandis que celui du modèle interne 345 614 270€ : nous avons donc un gain de capital d'environ 16%. Ce résultat est satisfaisant car l'objectif pour un assureur de contracter une réassurance est de pouvoir limiter l'immobilisation de son capital.

---

### 3. Impact du FGAO

Comme nous l'avons dit à la section 3.2, la revalorisation par l'inflation sur les rentes est désormais à la charge de l'assureur pour les sinistres survenus à partir du 1er Janvier 2013. Cette réforme entraîne de lourdes conséquences financières pour les assureurs automobiles en France. A partir de notre base de rentes pour les sinistres corporels Auto graves précédemment simulés, nous allons estimer l'impact de la réforme FGAO sur les provisions de rentes.

Sur notre portefeuille de rentes, nous obtenons les résultats suivants :

	Y compris FGAO	Hors FGAO	Impact
CM de sinistres	7 105 662	4 254 481	67,02%
CM annuel de sinistres	170 421 484	102 039 042	67,02%
SCR de primes	424 754 653	252 798 434	68,02%

TABLE 3.1: Résultats des simulations des rentes futures

Avec CM représentant le coût moyen.

En faisant la différence entre le coût moyen de rentes y compris FGAO et hors FGAO, que l'on divise par le coût moyen hors FGAO, nous constatons que le FGAO augmente de 67% environ le coût moyen de rentes chez les assureurs. Il en est de même sur la charge annuelle moyenne de rente.

En appliquant la même démarche sur le SCR de primes, nous constatons une augmentation du SCR de primes de 68% environ. Ainsi, les assureurs devront immobiliser plus de capital pour faire face au risque de sous tarification de la sinistralité future.

---

## 4. Efficience de la réassurance

Il est question dans ce chapitre de modifier les structures de réassurance afin de mesurer l'efficacité de la réassurance à travers les indicateurs suivants :

- Le prix théorique de la réassurance ;
- Le SCR net de réassurance ;
- Le ROE (Return On Equity), calculé comme suit :

$$ROE = \frac{RT_{cede} \text{ net d'impot}}{|SCR_{brut} - SCR_{net}|} \quad (3.4.1)$$

Avec  $RT_{cede}$  qui représente le résultat technique cédé.

Ce ratio est calculé hors primes futures car nous ne disposons pas du SCR brut comprenant les primes futures. Cet indicateur est important pour les assureurs car il leur permet de déterminer le meilleur programme de réassurance : plus il est faible, plus la réassurance est efficace (soit en réduisant le résultat cédé au réassureur, soit en permettant une économie de fonds propres importante).

### Méthodologie de calcul du ROE

Vu que nous cherchons à tester l'efficience de la réassurance, nous allons effectuer à tour de rôle une sensibilité de la portée et de la priorité de notre programme de réassurance initial 3M XS 12M (que nous appellerons **scénario central**) et nous calculerons pour chaque scénario secondaire les indicateurs cités plus haut.

Pour calculer les indicateurs des scénarios secondaires, nous calculerons tout d'abord ceux du scénario central :

- Calcul de la charge moyenne cédée de sinistres notée  $S$  : elle correspond à la prime pure du traité central ;
- Calcul du prix du traité central notée  $P$  : confère sous-section 2.1, équation 3.2.1 ;
- Calcul du  $S/P$  de cession ;
- Calcul du résultat technique cédé net d'impôt noté  $RT_{cede}$  :

$$RT_{cede} = (P - S) \times (1 - \alpha\%) \quad (3.4.2)$$

Avec  $\alpha = 28\%$  qui correspond au taux d'imposition des sociétés en 2020

- Calcul du SCR net de réassurance
- Calcul du ROE

NB : Le SCR brut de réassurance est une constante du modèle.

Les indicateurs du scénario central étant calculés, nous pouvons calculer ceux de chaque traité des scénarios secondaires. La méthodologie de calcul est similaire à celle du scénario central, à la seule différence que nous devons retarifier la prime de chaque nouveau traité en gardant le  $S/P$  de cession constant (celui du traité central).

Pour chaque traité  $i$ , on refait les étapes suivantes :

- Calcul de la charge moyenne cédée de sinistres notée  $S_i$
- Calcul du prix du traité  $P_i$  :

$$P_i = S_i / (S/P)$$

- Calcul du résultat technique cédé net d'impôt noté  $RT_{cede,i}$  :
- Calcul du SCR net de réassurance
- Calcul du ROE

À portée (respectivement priorité) fixée, nous allons modifier la priorité (respectivement la portée), afin d'étudier l'impact de ces modifications sur les indicateurs ci-dessus et choisir le meilleur programme de réassurance.

## 4.1 Modification de la priorité du traité

Dans cette section, nous ferons varier la priorité du traité initial 3M XS 12M et ce en fixant la portée.

La figure 3.4.1 permet de représenter l'évolution : du prix de réassurance, du SCR net de réassurance et du ROE en fonction de la priorité.

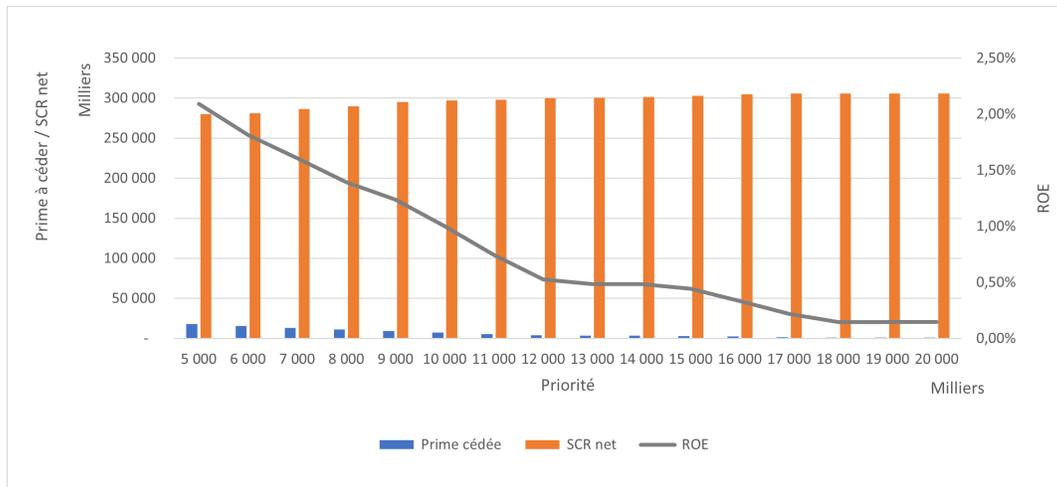


FIGURE 3.4.1: Effizienz der Rückversicherung in Abhängigkeit der Priorität

Nous pouvons constater que le prix du traité décroît lorsque la priorité augmente; tandis-que le SCR net augmente avec la priorité. Ces résultats sont logiques car une priorité élevée signifie que l'assureur cédera moins de sinistres et donc la prime de réassurance associée à ces sinistres sera moins élevée. Les cessions étant faibles, le SCR net de réassurance sera de ce fait élevé car  $SCR_{net} = SCR_{brut} - cessions$ .

Le ROE étant un ratio, dont le numérateur et le dénominateur diminuent respectivement lorsque la priorité augmente, nous obtenons donc un ROE décroissant avec la priorité. Un ROE faible signifie que l'assureur a obtenu un bon rendement de la réassurance : il a cédé de la marge et en contrepartie, il a fait une économie de capital importante. Ce gain de capital représente la différence entre les SCR brut et net.

Conclusion : Nous pouvons donc dire que plus la priorité augmente (pour une portée fixe), meilleur est le programme de réassurance

## 4.2 Modification de la portée du traité

Dans cette section, nous maintenons la priorité à 12M€ et nous ferons varier la portée entre 1M€ et 16M€, par pas de 1M€.

La figure 3.4.2 permet de représenter l'évolution de chacun des indicateurs cités ci-dessus.

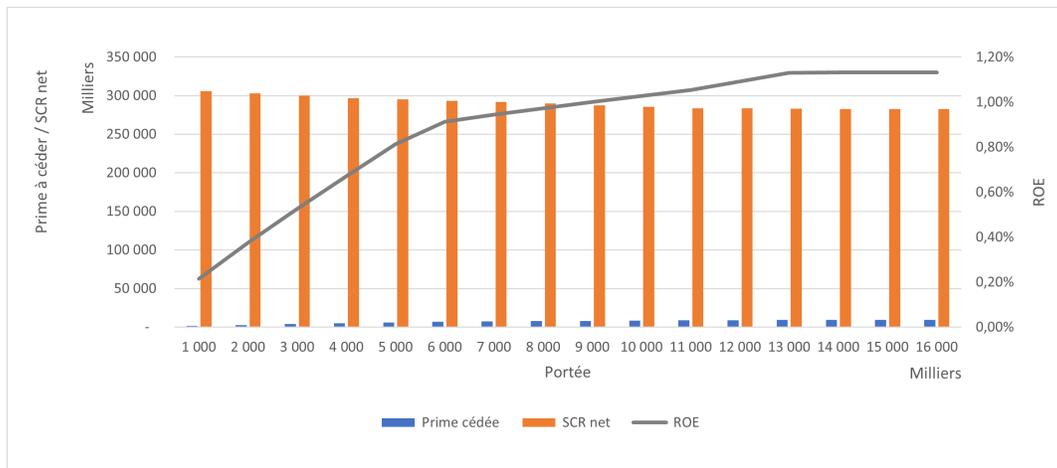


FIGURE 3.4.2: Efficacité de la réassurance en fonction de la portée

Nous constatons dans un premier temps que le prix de la réassurance et le SCR net n'évoluent pas dans le même sens : le SCR net décroît avec la portée tandis que le prix croît avec celle-ci. Ces résultats peuvent s'expliquer par le fait que plus la portée est grande, plus le réassureur prend à sa charge des sinistres à coûts élevés et donc réclamera une prime de réassurance plus élevée. Ce qui implique une hausse des charges cédées par l'assureur au réassureur et donc une baisse du SCR net.

Le ROE quant à lui augmente avec la portée : les sinistres et primes cédées augmentent avec la portée, ce qui implique une augmentation du résultat technique cédé. Le SCR net diminue, on a donc une différence de SCR croissante.

Conclusion : Nous pouvons donc dire que plus la portée augmente (pour une priorité fixe), moins bon est le programme de réassurance.

## 5. Tests de sensibilité

Il est question dans ce chapitre de tester la sensibilité du modèle suite à un changement d'hypothèses. Pour ce faire, nous allons appliquer un choc +/-100 bps sur la courbe des taux sans risque EOIPA : nous obtenons ainsi une courbe des taux à la hausse et une autre à la baisse (annexe A.9).

Nous avons refait la modélisation en appliquant, chacun à son tour, ces nouvelles courbes de taux lors du calcul des provisions mathématiques de rentes. Le tableau ci-dessous présente les résultats obtenus en appliquant les nouveaux taux d'actualisation :

	Taux EOIPA	Taux à la hausse	Taux à la baisse
CM de sinistres	5 254 512	4 713 015	5 914 18
CM annuel de sinistres	202 239 338	181 463 786	227 772 237
SCR brut de réassurance	424 754 653	362 345 714	487 353 122

TABLE 5.1: Sensibilité du modèle suite à une modification de la courbe des taux

Nous constatons que choquer la courbe des taux à la hausse diminue la charge moyenne de sinistre de -11% environ, soit une baisse de SCR brut de -17% environ. Tandis qu'un choc à la baisse produit l'effet inverse : augmentation de +11% environ du coût moyen de sinistres et de +12% environ du SCR brut.

Ces résultats nous semblent logiques car un dénominateur grand (dans une fraction) implique un résultat faible. À contrario, un dénominateur faible entraîne un résultat élevé.

### Impact sur la réassurance

La modification de la courbe des taux n'a eu aucun effet (ou presque pas) sur le prix de réassurance (confère tableau 5.2 ci-dessous) : les primes de réassurance sont relativement stables, peu importe la courbe des taux utilisée. Ce résultat s'explique par le fait que la réassurance s'applique sur la charge annuelle de sinistre calculée en utilisant comme taux d'actualisation, le taux de réassurance 1.50% (confère 2.1). Ainsi, les charges de sinistres cédées resteront plus ou moins inchangées.

Certes, les cessions ne changent pas, mais les charges annuelles de sinistres (pour chaque victime) varient en fonction de la courbe de taux : d'où la baisse de SCR net de réassurance de -27% environ lorsque nous modélisons avec la courbe des taux à la hausse et l'augmentation de +17%, avec la courbe des taux à la baisse.

	Taux EOIPA	Taux à la hausse	Taux à la baisse
Prix du traité XS 12	3 569 391	3 502 017	3 464 291
Prix du traité XS Illimité	5 749 316	5 737 380	5 425 553
SCR net de réassurance	299 882 855	234 458 144	362 190 314

TABLE 5.2: Sensibilité du modèle suite à une modification des taux d'actualisation : Impact sur la réassurance

*Conclusion :* Le changement de courbe de taux d'actualisation aura un effet significatif sur les charges annuelles de sinistres brutes de réassurance, les SCR bruts et nets de réassurance mais n'en aura aucun sur les prix de réassurance.

---

Quatrième partie

Calcul du SCR Vie sur les rentes Auto  
Corporels

---

# 1. Méthodologie de calcul

Les rentes Auto corporels, pouvant être versées durant toute la durée de vie de la victime, elles sont de ce fait soumises aux risques vie suivants : le risque de longévité, le risque de frais de gestion (expenses) et le risque de révision. Ces risques nous permettront de calculer le SCR Vie de notre portefeuille de rentes Auto.

Le SCR vie correspond à une agrégation via une matrice de corrélation des SCR des risques vie ( $SCR_{longevite}$ ,  $SCR_{expenses}$ ,  $SCR_{revision}$ ). Ces trois SCR, calculés individuellement, sont obtenus par différence entre le BE du dit risque et le BE central.

Soit  $i$  un individu d'âge  $x$ , qui perçoit  $A_i$  d'arrérage annuel. Le cash flow versé à cet individu à l'année  $t$  est :

$$CF_{i,t} = A_i \times p_{x,x+t} \times (1 + tx_{frais} \times tx_{inflationcumulee,t}) \times (1 + tx_{revalo})^t \quad (4.1.1)$$

Avec  $tx_{frais}$  le taux de frais de gestion,  $tx_{revalo}$  le taux de revalorisation, et  $tx_{inflationcumulee,t}$  le taux d'inflation cumulée à l'année  $t$ . En effet, nous disposons d'une courbe d'inflation provenant de la Direction Financière de Groupama (annexe A9), que nous cumulons par la suite sur chaque année :

$$tx_{inflationcumulee,t} = tx_{inflationcumulee,t-1} \times (1 + tx_{inflation,t}).$$

## 1.1 Le choc de longévité

Les assureurs vie se questionnent toujours sur la durée de vie d'une population donnée afin de calculer de manière précise leurs provisions. Ainsi, le risque de longévité correspond au risque que les rentiers vivent en moyenne plus longtemps que ce que l'assureur avait prévu lors du calcul de ses provisions : il s'agit donc d'une erreur d'estimation de la table de mortalité. Les assureurs supposent donc que la table de mortalité utilisée sous-estime la mortalité des rentiers et appliquent des chocs de longévité à la hausse afin de capter cette incertitude.

Les équipes vie de GMA calibrent un niveau de choc de longévité de -20%. Choquer la longévité de -20% revient à diminuer la probabilité de décès  $q_x = \frac{l_{x+1}}{l_x}$  dans l'année d'un individu d'âge  $x$  de -20%. On construit alors une table de mortalité choquée en appliquant la formule :

$$\begin{cases} l_{x+1, choc} = l_{x, choc} \times (1 - q_{x, choc}) & \text{pour } x = 1, \dots, 114 \text{ ans} \\ l_{0, choc} = l_0 = 100000 \end{cases}$$

Avec  $l_{x, choc}$  le nombre de de rentiers vivants d'âge  $x$  dans la table choquée et  $q_{x, choc}$  la probabilité de décès choquée de ces rentiers. En annexe A10, la table de mortalité choquée utilisée à Groupama.

Ce choc de longévité permet d'augmenter la probabilité de survie des rentiers. Ce qui implique donc une augmentation des cash-flows versés à ceux-ci, calculés somme suit :

$$CF_{longev,i,t} = A_i \times p_{x,x+t}^{choc} \times (1 + tx_{frais} \times tx_{inflationcumulee,t}) \times (1 + tx_{revalo})^t$$

Pour chaque rentier de notre base, nous allons calculer le BE central ainsi que le BE stressé après un choc de longévité par les formules suivantes :

$$BE_{central,i} = \sum_{t=0}^T CF_{i,t} \times (1 + tx_{actu,t})^{-t} \quad (4.1.2)$$

avec  $tx_{actu,t}$  le taux d'actualisation provenant de la courbe des taux EOIPA (annexe A11).

$$BE_{longev,i} = \sum_{t=0}^T CF_{longev,i,t} \times (1 + tx_{actu,t})^{-t} \quad (4.1.3)$$

Le SCR longévité est ainsi égal au capital économique nécessaire pour absorber ce choc de longévité. Il est donc déduit de la formule :

$$SCR_{longev} = BE_{longev} - BE_{central} \quad (4.1.4)$$

## 1.2 Le choc de frais de gestion (expenses)

La provision pour frais de gestion permet de couvrir les frais de gestion des sinistres déjà survenus. Le risque de frais de gestion correspond ainsi à l'impact d'une évaluation insuffisante des frais supportés par l'assureur. Choquer les frais, revient à augmenter les frais de 10% et la courbe d'inflation de 1% sur chaque période.

Nous obtenons les cash-flows suivants :

$$CF_{expense,i,t} = A_i \times p_{x,x+t} \times (1 + tx_{frais}^{choc} \times tx_{inflationcumulee,t}^{choc}) \times (1 + tx_{revalo})^t$$

Avec

$$\begin{cases} tx_{frais}^{choc} = tx_{frais} \times (1 + 10\%) \\ tx_{inflationcumulee,t}^{choc} = tx_{inflationcumulee,t-1} \times (1 + tx_{inflation,t} + 1\%) \end{cases}$$

Nous avons obtenus par la suite le BE de frais :

$$BE_{expense,i} = \sum_{t=0}^T CF_{expense,i,t} \times (1 + tx_{actu,t})^{-t} \quad (4.1.5)$$

Ainsi, le capital économique nécessaire pour absorber ce choc de frais de gestion se déduit facilement :

$$SCR_{expense} = BE_{expense} - BE_{central} \quad (4.1.6)$$

## 1.3 Le choc de révision

Le risque de révision correspond à une possible augmentation de montant d'arrérage, suite à une aggravation de l'état de santé de la victime. Au sein de Groupama, il s'agit d'augmenter de 4% le montant annuel d'arrérage de chaque rente, ce qui conduit à une augmentation de 4% des cash-flows annuels :

$$CF_{revision,i,t} = A_i^{choc} \times p_{x,x+t} \times (1 + tx_{frais} \times tx_{inflationcumulee,t}) \times (1 + tx_{revalo})^t$$

Avec  $A_i^{choc} = A_i \times 1.04$

Le BE de révision ainsi que le SCR de révision se déduisent respectivement des formules suivantes :

$$BE_{revision,i} = \sum_{t=0}^T CF_{revision,i,t} \times (1 + tx_{actu,t})^{-t} \quad (4.1.7)$$

$$SCR_{revision} = BE_{revision} - BE_{central} \quad (4.1.8)$$

---

## 2. Calcul du SCR vie sur les rentes

Les risques vie étant corrélés entre eux (c'est dire que si l'un des chocs se réalise, il existe une probabilité non nulle qu'un autre choc se réalise), nous allons les agréger via une matrice de corrélation  $M_c$  (issue de la formule standard) afin d'obtenir le SCR Vie sur les rentes Auto corporels. Cette matrice  $M_c$  est la suivante :

	Longévit�	Expense	R�vision
Long�vit�	1	0.25	0.25
Expense	0.25	1	0.5
R�vision	0.25	0.5	1

TABLE 2.1: Matrice de cor lation issue de la formule standard

Cette matrice signifie qu'une augmentation de la long vit  entra nera une augmentation du montant d'arr rage avec un coefficient de cor lation de 25%.

Le SCR Vie sur les rentes se d duit de la formule :

$$SCR_{vie} = \sqrt{(SCR_{longev}, SCR_{expense}, SCR_{revision}) \times M_c \times (SCR_{longev}, SCR_{expense}, SCR_{revision})^t} \quad (4.2.1)$$

### Application et r sultats

Sur notre base de rentiers simul e et projet e sur  $M = 10\,000$  ann es , nous avons appliqu  la m thodologie pr c demment pr sent e et nous avons obtenus  $M$  valeurs de SCR Vie.

Dans le tableau 2.2 ci-dessous, nous pouvons constater que la valeur m diane est tr s proche de la valeur moyenne : il n'existe pas une grande dispersion entre nos diff rentes ann es. De plus, le 3 me quartile  tant proche de la moyenne, cela signifie que 75% de notre base a un SCR Vie inf rieur ou  gal au SCR vie moyen. Il existe n anmoins quelques ann es atypiques, caract ris es ici par des valeurs tr s faibles ou tr s  lev es.

Min	1er quartile	M�diane	Moyenne	3eme quartile	Max
249 600	2 193 849	2 750 812	3 140 892	3 899 129	9 909 311

TABLE 2.2: Statistiques sur nos  $M$  valeurs de SCR Vie

Nous pouvons ainsi conclure que nos ann es se ressemblent   quelque chose pr s.

---

# Conclusion

Le but de ce mémoire était de proposer une modélisation individuelle des pertes futures associées aux sinistres Auto corporels graves. Celle-ci prendra en compte les aspects suivants (qui ne sont pas pris en compte dans le modèle actuel) :

- les risques de longévité, de frais et de révision liés aux rentes futures ;
- les caractéristiques biométriques de la victime ;
- le FGAO, modélisé non pas de manière agrégée mais individuelle ;
- le taux d'actualisation  $S2$  et la table de mortalité ;
- Application de la réassurance sur les sinistres actualisés au bon taux de réassurance

La modélisation des sinistres corporels futurs que nous avons présentée dans ce rapport est faite tête par tête et utilise les caractéristiques biométriques de la victime (âge, sexe, AIPP, sortie en rentes ou capital,  $\dots$ ). Elle est basée sur une approche *fréquence  $\times$  coût*, estimée comme suit :

- La fréquence de sinistres : elle dépend du mode d'indemnisation de la victime. Pour les sorties en rente, la fréquence de sinistres est modélisée par une loi Binomiale négative, de paramètres  $\hat{n} = 8.845$  et  $\hat{p} = 0.266$  et pour les sorties en capital, elle est modélisée par la différence entre le nombre total de sinistres corporels Auto graves du modèle interne et le nombre de sinistre de rentes ;
- Le coût du sinistre est modélisé de manière stochastique : il est fonction des caractéristiques biométriques de la victime. Après avoir simulé les caractéristiques de la victime, nous avons régressé le montant d'arérage sur ces caractéristiques. Le choix du modèle de régression s'est fait à partir du critère de minimisation de l'erreur moyenne de prédiction calculée sur les différents modèles suivants : régression linéaire simple, régression log-linéaire, modèle GLM de loi Gamma avec fonction de lien inverse, modèle GLM de loi Gamma avec fonction de lien log et arbre de régression CART. Par la suite, en fonction du mode d'indemnisation (rente ou capital), nous avons calculé le montant de la provision (en utilisant la courbe des taux EOIPA pour les rentes et le taux du barème Gazette du Palais : 0.50% pour les sorties en capital). Les provisions de rentes sont revalorisées au taux d'inflation de 2%, tandis que celles des sorties en capital ne sont pas revalorisées. Nous avons ensuite modélisé les frais de soins pour chaque partie (rentes/capital), que nous avons ajouté à la provision : nous obtenons ainsi une charge de sinistres tête par tête.

A partir de la charge de sinistre modélisée, nous avons constitué notre base de sinistres graves en utilisant le seuil des graves du modèle interne, qui est de 1 200 000€. Ainsi, les sinistres graves sont ceux dont la charge est supérieure ou égale à ce seuil. Dans une optique de calcul du SCR de primes, nous avons fait 10 000 scénarios, que nous agrégeons ensuite par scénario, pour obtenir la charge annuelle totale de sinistre. Nous obtenons ainsi 10 000 valeurs de pertes annuelles totales et nous déduisons le SCR de primes brut de réassurance. Il correspond au quantile 99,5% du vecteur de charges annuelles. Nous avons ensuite appliqué les traités XS de réassurance à ces sinistres dont la charge a été calculée en utilisant comme taux d'actualisation le taux de réassurance qui vaut 1,50%. Nous avons pu déduire le SCR net de réassurance et nous avons calculé le prix théorique de la réassurance en supposant que le chargement de sécurité vaut 25%.

En comparant les résultats obtenus par le modèle actuel et ceux obtenus par la nouvelle modélisation, nous avons constaté que si nous utilisons le taux social comme taux d'actualisation, comme fait dans le modèle interne, le coût moyen de sinistre augmentait de +45%. Cette différence est due à l'effet biométrie et à la façon de simuler les variables explicatives du modèle. Par contre, quand nous utilisons les taux  $S2$ , le coût moyen baisse de -8% uniquement. En terme de SCR brut de réassurance, il baisse de -19% (à taux différents bien sûr) : différence liée à l'effet coût moyen et à la dispersion des charges. Ainsi, le modèle interne reste tout de même prudent dans sa modélisation.

---

Lorsque nous appliquons la réassurance avec le bon taux, nous obtenons un prix de la réassurance (pour chaque traité XS) qui se rapproche de celui réellement cédé au réassureur. En net de réassurance, nous obtenons une baisse de SCR de -16% par rapport au modèle interne. Notre modèle est donc satisfaisant.

Le FGAO étant désormais à la charge des assureurs depuis 2013, nous avons également mesuré l'impact du FGAO sur les rentes futures. Nous avons constaté qu'il augmente de 67% environ la charge annuelle moyenne de rentes et de 68% le SCR de primes brut de réassurance chez les assureurs : ils devront donc immobiliser plus de capital. Cette nouvelle réforme impacte fortement les provisions des assureurs.

En conclusion, ce nouveau modèle est dynamique, plus détaillé et précis par rapport à la modélisation actuelle. Elle a permis de réaliser un gain de SCR brut d'environ 19%, ce qui n'est pas négligeable pour les assureurs. Ainsi, ceux-ci pourront avoir une meilleure vision de leur risque. Cependant, ce modèle comporte tout de même des limites. Par exemple, nous ne disposons pas de tous les éléments dont dispose le juge pour déterminer le montant de l'arrérage. De plus, nous avons émis des hypothèses lors de notre modélisation : sur le chargement de sécurité, sur le taux de frais de gestion, sur la table de mortalité.

*Pistes d'améliorations :*

Nous avons certes déterminé les fréquences de sinistres dont les sorties sont des rentes et du capital, mais elles sont déterministes : nous nous sommes appuyé sur la fréquence du modèle interne. Cependant, il serait intéressant de modéliser de manière stochastique la probabilité de passer en rente ou capital et ainsi déterminer leurs taux de sortie. De plus, il serait intéressant de modéliser le seuil des graves à partir des données et non se baser sur l'existant.

---

# ANNEXES

---

# A. ANNEXES

## A.1 Définition des postes de préjudice permanents selon la nomenclature Dintilhac

### Assistance Tierce Personne (ATP)

Ces dépenses de tierce personne visent à indemniser le coût pour la victime de la présence nécessaire, de manière définitive, d'une tierce personne à ses côtés pour l'assister dans les actes de la vie quotidienne, préserver sa sécurité, contribuer à restaurer sa dignité et suppléer sa perte d'autonomie. Ces dépenses de tierce personne constituent des dépenses permanentes qui ne se confondent pas avec les frais temporaires que la victime peut être amenée à déboursier durant la maladie traumatique, lesquels sont déjà susceptibles d'être indemnisés au titre du poste "Frais divers".

Les indemnités perçues en droit commun par la victime d'un accident sont de 10 à 12 euros de l'heure, sans les charges, capitalisées sur le nombre de temps nécessaire afin d'assister la victime dans les actes de la vie courante. Les indemnités perçues en droit commun suite à une erreur médicale, un aléa thérapeutique, un accident de la route ou un accident du travail se capitalisent suivant le temps requis à l'auxiliaire de vie ou la tierce personne (famille) pour effectuer la tâche lors du calcul de l'indemnité avec l'assurance ou le fond de garantie.

### Perte de Gains Professionnels Futurs (PGPF)

Il s'agit ici d'indemniser la victime de la perte ou de la diminution de ses revenus consécutive à l'incapacité permanente à laquelle elle est désormais confrontée dans la sphère professionnelle à la suite du dommage. Ne pas confondre avec l'incidence professionnelle.

Il s'agit d'indemniser une invalidité spécifique partielle ou totale qui entraîne une perte ou une diminution directe de ses revenus professionnels futurs à compter de la date de consolidation. Cette perte ou diminution des gains professionnels peut provenir soit de la perte de son emploi par la victime, soit de l'obligation pour celle-ci d'exercer un emploi à temps partiel à la suite du dommage consolidé. Ce poste perte ou diminution des gains professionnels n'englobe pas les frais de reclassement professionnel, de formation ou de changement de poste qui ne sont que des conséquences indirectes du dommage. En outre, concernant les jeunes victimes ne percevant pas à la date du dommage de gains professionnels, il conviendra de prendre en compte pour l'avenir la privation de ressources professionnelles engendrée par le dommage en se référant à une indemnisation par estimation.

### Dépenses de Santé Futures (DSF)

Les dépenses de santé futures sont les frais hospitaliers, médicaux, paramédicaux, pharmaceutiques et assimilés, même occasionnels mais médicalement prévisibles, rendus nécessaires par l'état pathologique de la victime après la consolidation. Elles ne se limitent pas aux frais médicaux au sens strict : ils incluent, en outre, les frais liés soit à l'installation de prothèses pour les membres, les dents, les oreilles ou les yeux, soit à la pose d'appareillages spécifiques qui sont nécessaires afin de suppléer le handicap physiologique permanent qui demeure après la consolidation. Elles sont postérieurs à la consolidation de la victime, dès lors qu'ils sont médicalement prévisibles, répétitifs et rendus nécessaires par l'état pathologique permanent et chronique de la victime après sa consolidation définitive (frais liés à des hospitalisations périodiques dans un 33 établissement de santé, à un suivi médical assorti d'analyses, à des examens et des actes périodiques, des soins infirmiers, ou autres frais occasionnels, etc.).

### L'incidence professionnelle (IP)

Ce poste d'indemnisation l'incidence professionnelle vient compléter celle déjà obtenue par la victime au titre du poste "pertes de gains professionnels futurs" sans pour autant aboutir à une double indemnisation

---

du même préjudice. Cette incidence professionnelle à caractère définitif a pour objet d'indemniser non la perte de revenus liée à l'invalidité permanente de la victime, mais les incidences périphériques du dommage touchant à la sphère professionnelle comme le préjudice subi par la victime en raison de sa dévalorisation sur le marché du travail, de sa perte d'une chance professionnelle, ou de l'augmentation de la pénibilité de l'emploi qu'elle occupe imputable au dommage ou encore du préjudice subi qui a trait à sa nécessité de devoir abandonner la profession qu'elle exerçait avant le dommage au profit d'une autre qu'elle a du choisir en raison de la survenance de son handicap. Il convient, en outre, de ranger dans ce poste de préjudice incidence professionnelle les frais de reclassement professionnel, de formation ou de changement de poste assumés par la sécurité sociale et / ou par la victime elle-même qui sont souvent oubliés, alors qu'ils concernent des sommes importantes. Il s'agit des frais déboursés par l'organisme social et / ou par la victime elle-même immédiatement après que la consolidation de la victime soit acquise afin qu'elle puisse retrouver une activité professionnelle adaptée une fois sa consolidation achevée : elle peut prendre la forme d'un stage de reconversion ou d'une formation. Là encore, le pragmatisme doit conduire à ne pas retenir une liste limitative de ses frais spécifiques incidence professionnelle, mais à l'inverse à inclure dans ce poste de préjudice patrimonial tous les frais imputables au dommage nécessaires à un retour de la victime dans la sphère professionnelle. Ce poste de préjudice incidence professionnelle cherche également à indemniser la perte de retraite que la victime va devoir supporter en raison de son handicap, c'est à dire le déficit de revenus futurs, estimé imputable à l'accident, qui va avoir une incidence sur le montant de la pension auquel pourra prétendre la victime au moment de sa prise de retraite. Comme pour l'indemnisation du poste précédent, il convient de noter que si les pertes de gains professionnels peuvent être évaluées pour des victimes en cours d'activité professionnelle, elles ne peuvent cependant qu'être estimées pour les enfants ou les adolescents qui ne sont pas encore entrés dans la vie active. Une fois encore, la liste des préjudices à intégrer dans ce poste est indicative. Ainsi, il peut, par exemple, être prévu une indemnisation, au titre de ce poste, de la mère de famille sans emploi pour la perte de la possibilité, dont elle jouissait avant l'accident, de revenir sur le marché du travail.

#### **Frais de véhicule adapté (FVA)**

Les frais de véhicule adapté suite à un accident de la route ou suite à un accident médical correspondent aux dépenses nécessaires pour procéder à l'adaptation d'un ou de plusieurs véhicules aux besoins de la victime atteinte d'un handicap permanent. Il convient d'inclure dans ce poste de préjudice frais de véhicule adapté le ou les surcoût(s) lié(s) au renouvellement du véhicule et à son entretien. En revanche, les frais liés à l'adaptation, à titre temporaire, du véhicule avant la consolidation de la victime ne sont pas à intégrer, car ils sont provisoires et déjà susceptibles d'être indemnisés au titre du poste "Frais divers". En outre, ce poste frais de véhicule adapté doit inclure non seulement les dépenses liées à l'adaptation d'un véhicule, mais aussi le surcoût d'achat d'un véhicule susceptible d'être adapté. Enfin, il est également possible d'assimiler à ces frais d'adaptation du véhicule les surcoûts en frais de transport rendus nécessaires à la victime en raison de ses difficultés d'accessibilité aux transports en commun survenues depuis le dommage.

#### **Frais de logement adapté (FLA)**

Les frais de logement adapté concernent le remboursement des frais que doit exposer la victime à la suite de sa consolidation, dans la mesure où les frais d'adaptation du logement exposés, à titre temporaire, sont déjà susceptibles d'être indemnisés au titre du poste de préjudice "Frais divers". Les frais de logement adapté et son indemnisation intervient sur la base de factures, de devis ou même des conclusions du rapport de l'expert médical sur la consistance et le montant des travaux nécessaires à la victime pour vivre dans son logement. Les frais de logement adapté doivent être engagés pendant la maladie traumatique afin de permettre à la victime handicapée de pouvoir immédiatement retourner vivre à son domicile dès sa consolidation acquise. Les frais de logement adapté incluent non seulement l'aménagement du domicile préexistant, mais éventuellement celui découlant de l'acquisition d'un domicile mieux adapté prenant en compte le surcoût financier engendré par cette acquisition. En outre, il est possible d'inclure au titre de l'indemnisation de ce poste de préjudice les frais de déménagement et d'emménagement, ainsi que ceux liés à un surcoût de loyer pour un logement plus grand découlant des difficultés de mobilité de la victime devenue handicapée. Les frais de logement adapté intègrent également les frais de structure nécessaires pour que la victime handicapée puisse disposer d'un

autre lieu de vie extérieur à son logement habituel de type foyer ou maison médicalisée.

### Déficit Fonctionnel Permanent (DFP)

Ce poste de préjudice cherche à indemniser un préjudice extra-patrimonial découlant d'une incapacité constatée médicalement qui établit que le dommage subi a une incidence sur les fonctions du corps humain de la victime. Il s'agit ici de réparer les incidences du dommage qui touchent exclusivement à la sphère personnelle de la victime. Il convient d'indemniser, à ce titre, non seulement les atteintes aux fonctions physiologiques de la victime, mais aussi la douleur permanente qu'elle ressent, la perte de la qualité de vie et les troubles dans les conditions d'existence qu'elle rencontre au quotidien après sa consolidation. Ce poste peut être défini, selon la Commission européenne à la suite des travaux de Trèves de juin 2000, comme correspondant à « la réduction définitive du potentiel physique, psycho-sensoriel ou intellectuel résultant de l'atteinte à l'intégrité anatomo-physiologique médicalement constatable donc appréciable par un examen clinique approprié complété par l'étude des examens complémentaires produits, à laquelle s'ajoutent les phénomènes douloureux et les répercussions psychologiques, normalement liées à l'atteinte séquelle décrite ainsi que les conséquences habituellement et objectivement liées à cette atteinte dans la vie de tous les jours ». En outre, ce poste de préjudice doit réparer la perte d'autonomie personnelle que vit la victime dans ses activités journalières, ainsi que tous les déficits fonctionnels spécifiques qui demeurent même après la consolidation. En raison de son caractère général, ce déficit fonctionnel permanent ne se confond pas avec le préjudice d'agrément, lequel a pour sa part un objet spécifique en ce qu'il porte sur la privation d'une activité déterminée de loisirs.

## A.2 Tableau de synthèse des postes de préjudices de la nomenclature DINTILHAC

Préjudice de la victime directe	Préjudices patrimoniaux	Préjudice extrapatrimoniaux
Temporaires (avant consolidation)	Dépenses de santé actuelles	DSA Déficit fonctionnel temporaire DFT
	Frais Divers	FD Souffrances endurées SE
	Pertes de gains professionnels actuels	PGPA Préjudice esthétique temporaire PET
Permanents (après consolidation)	Dépenses de santé futures	DSF Déficit fonctionnel permanent DFP
	Frais de logement adapté	FLA Préjudice d'agrément PA
	Frais de véhicule adapté	FVA Préjudice esthétique permanent PEP
	Assistance tierce Personne	ATP Préjudice sexuel PS
	Pertes de gains professionnels futurs	PGPF Préjudice d'établissement PE
	Incidence professionnelle	IP Préjudice permanents exceptionnels PPE
Evolutifs (hors consolidation)	Préjudice scolaire, universitaire ou de formation	PSU
		Préjudices liés à des pathologies évolutives PEV
Préjudice de la victime indirecte	Préjudices patrimoniaux	Préjudice extrapatrimoniaux
En cas de décès de la victime directe	Frais d'obsèques	FO Préjudice d'accompagnement PAC
	Pertes préjudice économique	PR Préjudice d'affection PAF
	Frais divers des proches	FD
En cas de survie de la victime directe	Pertes préjudice économique	PR Préjudice d'affection PAF
	Frais divers des proches	FD Préjudices extra- patrimoniaux exceptionnels PEX

FIGURE A1: Synthèse des postes de préjudices de la nomenclature DINTILHAC

---

## A.3 Résultats de la modélisation de l'arrérage chez les femmes

### Régression linéaire

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    5469.46   4743.10   1.153   0.2498
Age_survNew[15-16] -13600.37  6921.69  -1.965   0.0503 .
Age_survNew[17-23] -9014.23   4351.52  -2.072   0.0392 *
Age_survNew[24-26] -14471.73  6264.68  -2.310   0.0216 *
Age_survNew[27-30] -12889.51  5137.26  -2.509   0.0126 *
Age_survNew[31-37]  -7146.02   5263.55  -1.358   0.1756
Age_survNew[38-44] -3948.05   5550.01  -0.711   0.4774
Age_survNew45 et + -10632.58  4799.72  -2.215   0.0275 *
TAUX_AIPP        300.29    43.04   6.977 1.88e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 22490 on 305 degrees of freedom
Multiple R-squared:  0.2053,    Adjusted R-squared:  0.1845
F-statistic: 9.852 on 8 and 305 DF,  p-value: 3.44e-12
```

FIGURE A2: Résultats de la régression linéaire chez les femmes

Nous n'allons pas présenter les tests faits sur les résidus car nous avons que cette régression n'est pas adaptée pour modélisation (confère section 2.2.1).

### Régression log-linéaire

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    7.247073   0.263874  27.464 < 2e-16 ***
Age_survNew[15-16] -0.772717   0.385076  -2.007  0.04567 *
Age_survNew[17-23] -0.679529   0.242089  -2.807  0.00532 **
Age_survNew[24-26] -1.026879   0.348524  -2.946  0.00346 **
Age_survNew[27-30] -0.928176   0.285802  -3.248  0.00129 **
Age_survNew[31-37] -0.762139   0.292828  -2.603  0.00970 **
Age_survNew[38-44] -0.550125   0.308764  -1.782  0.07579 .
Age_survNew45 et + -0.552316   0.267024  -2.068  0.03944 *
TAUX_AIPP        0.031002   0.002394  12.947 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.251 on 305 degrees of freedom
Multiple R-squared:  0.4356,    Adjusted R-squared:  0.4208
F-statistic: 29.42 on 8 and 305 DF,  p-value: < 2.2e-16
```

FIGURE A3: Résultats de la régression log-linéaire chez les femmes

Nous constatons que la variance expliquée par ce modèle est deux fois plus grande que celle expliquée par la régression linéaire classique. Ce modèle semble plus adapté.

Faisons les tests de normalité des résidus.

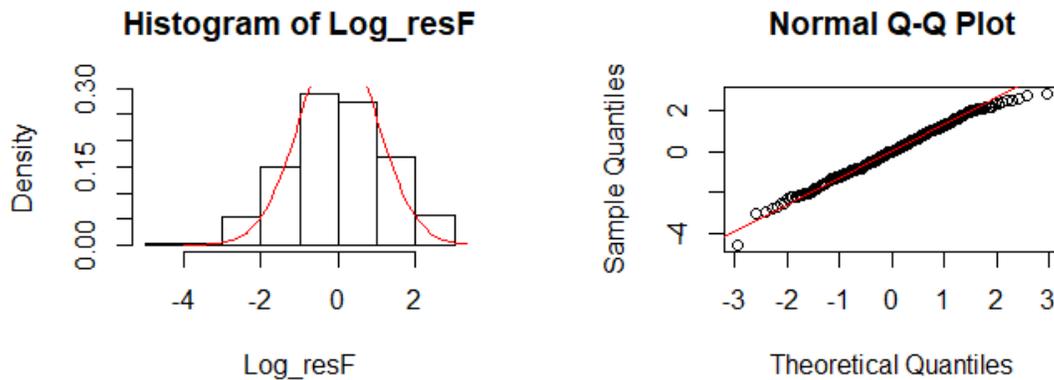


FIGURE A4: Test de normalité des résidus pour les femmes

De plus, les résidus semblent gaussiens. Nous avons également effectué un test de shapiro et nous avons obtenu une p-value de  $0.2163 > 0.05$ . Les résidus sont bien gaussiens.

#### Modèle GLM Gamma avec fonction de lien Inverse

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.609e-04  3.314e-05  7.874 6.08e-14 ***
Age_survNew[15-16]  2.083e-05  3.266e-05  0.638  0.524
Age_survNew[17-23]  8.814e-06  1.252e-05  0.704  0.482
Age_survNew[24-26]  1.680e-04  1.073e-04  1.566  0.118
Age_survNew[27-30]  3.240e-05  2.108e-05  1.537  0.125
Age_survNew[31-37]  7.109e-06  1.597e-05  0.445  0.657
Age_survNew[38-44] -1.555e-06  1.211e-05 -0.128  0.898
Age_survNew45 et +  2.547e-05  2.377e-05  1.071  0.285
TAUX_AIPP        -2.497e-06  3.381e-07 -7.386 1.46e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 2.419825)

Null deviance: 750.75  on 313  degrees of freedom
Residual deviance: 493.29  on 305  degrees of freedom
AIC: 6334.9

```

FIGURE A5: Résultats du modèle GLM Gamma avec fonction de lien Inverse chez les femmes

Nous constatons que la seule variable significative est la variable *TAUX\_AIPP*. C'est le même constat que nous avons eu dans la modélisation chez les hommes.

#### Modèle GLM Gamma avec fonction de lien log

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      7.95118    0.27109  29.331 < 2e-16 ***
Age_survNew[15-16] -0.98087    0.39560  -2.479  0.01370 *
Age_survNew[17-23] -0.55056    0.24871  -2.214  0.02759 *
Age_survNew[24-26] -1.17618    0.35805  -3.285  0.00114 **
Age_survNew[27-30] -0.91316    0.29362  -3.110  0.00205 **
Age_survNew[31-37] -0.76715    0.30083  -2.550  0.01126 *
Age_survNew[38-44] -0.48352    0.31721  -1.524  0.12846
Age_survNew45 et + -0.68024    0.27432  -2.480  0.01369 *
TAUX_AIPP          0.03066    0.00246  12.465 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 1.652109)

Null deviance: 750.75  on 313  degrees of freedom
Residual deviance: 430.82  on 305  degrees of freedom
AIC: 6283.3

```

FIGURE A6: Résultats du modèle GLM Gamma avec fonction de lien log chez les femmes

Par contre, dans ce modèle, presque toutes les variables sont significatives.

En comparant l'AIC des deux modèles GLM, le modèle GLM Gamma avec fonction de lien log est plus adapté car l'AIC est plus faible. Il en est de même lorsque nous comparons les déviances.

Le modèle GLM Gamma avec fonction de lien log est donc plus adapté que le GLM Gamma avec fonction de lien Inverse.

### Arbre de régression CART

Nous obtenons le graphe de l'erreur relative en fonction du cp ainsi que celui de l'arbre de régression après application :

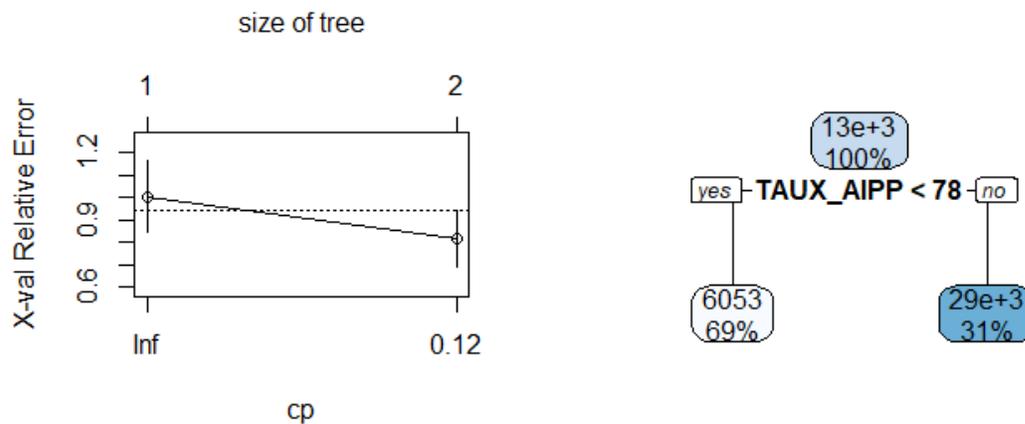


FIGURE A7: Graphe de l'erreur relative et arbre optimal de régression

Pour les mêmes raisons que celles présentées dans la sous section 2.2.4, cette régression semble ne pas être appropriée pour la modélisation.

## A.4 Statistiques comparatives des variables explicatives après simulation par sexe

	Valeurs moyennes simulées	Valeurs moyennes observées
Age	27.02	27.03
Taux d'AIPP	52.76	52.79
Arrérage	11 350	12 416

TABLE A.1: Statistiques comparatives entre les valeurs simulées et les valeurs observées chez les hommes

	Valeurs moyennes simulées	Valeurs moyennes observées
Age	29.48	29.51
Taux d'AIPP	53.36	52.32
Arrérage	12 771	13 853

TABLE A.2: Statistiques comparatives entre les valeurs simulées et les valeurs observées chez les femmes

## A.5 Table de mortalité TH/TF 00-02

Age	lx_F	lx_H	Age	lx_F	lx_H	Age	lx_F	lx_H	Age	lx_F	lx_H	Age	lx_F	lx_H	Age	lx_F	lx_H
0	100000	100000	20	99274	98921	40	98242	96369	60	93329	85538	80	71088	47390	100	3185	663
1	99616	99511	21	99239	98820	41	98130	96141	61	92892	84558	81	68423	44234	101	2171	410
2	99583	99473	22	99205	98716	42	98007	95887	62	92425	83514	82	65478	40946	102	1426	244
3	99562	99446	23	99171	98612	43	97872	95606	63	91923	82399	83	62233	37546	103	900	139
4	99545	99424	24	99137	98509	44	97724	95295	64	91382	81206	84	58680	34072	104	544	75
5	99531	99406	25	99103	98406	45	97563	94952	65	90797	79926	85	54828	30575	105	314	39
6	99519	99390	26	99068	98303	46	97387	94575	66	90164	78552	86	50706	27104	106	172	19
7	99508	99376	27	99033	98198	47	97197	94164	67	89476	77078	87	46362	23707	107	89	9
8	99498	99363	28	98997	98091	48	96993	93720	68	88726	75501	88	41868	20435	108	44	4
9	99488	99350	29	98960	97982	49	96776	93244	69	87907	73816	89	37319	17338	109	20	2
10	99478	99338	30	98921	97870	50	96546	92736	70	87010	72019	90	32821	14464	110	9	1
11	99467	99325	31	98879	97756	51	96304	92196	71	86024	70105	91	28469	11852	111	4	0
12	99456	99312	32	98833	97639	52	96049	91621	72	84941	68070	92	24328	9526	112	1	0
13	99444	99296	33	98782	97517	53	95778	91009	73	83751	65914	93	20444	7498	113	0	0
14	99431	99276	34	98725	97388	54	95489	90358	74	82442	63637	94	16860	5769	114	0	0
15	99415	99250	35	98662	97249	55	95180	89665	75	80998	61239	95	13618	4331	115	0	0
16	99395	99213	36	98593	97100	56	94851	88929	76	79402	58718	96	10750	3166	116	0	0
17	99371	99163	37	98518	96939	57	94501	88151	77	77633	56072	97	8277	2249	117	0	0
18	99342	99097	38	98435	96765	58	94131	87329	78	75671	53303	98	6204	1549	118	0	0
19	99309	99015	39	98343	96576	59	93741	86460	79	73496	50411	99	4516	1032	119	0	0

FIGURE A8: Table de mortalité TH/TF 00-02

---

## A.6 Courbe des taux d'inflation permettant de calculer les provisions BE

Période	Inflation	Période	Inflation	Période	Inflation
1	1,66%	21	0,43%	41	2,08%
2	1,31%	22	0,70%	42	2,08%
3	1,06%	23	0,96%	43	2,08%
4	0,93%	24	1,18%	44	2,07%
5	0,77%	25	1,31%	45	2,06%
6	0,67%	26	1,46%	46	2,09%
7	0,61%	27	1,58%	47	2,07%
8	0,58%	28	1,67%	48	2,05%
9	0,55%	29	1,77%	49	2,06%
10	0,48%	30	1,80%	50	2,08%
11	0,42%	31	1,86%		
12	0,39%	32	1,92%		
13	0,35%	33	1,97%		
14	0,24%	34	1,98%		
15	0,03%	35	1,98%		
16	-0,23%	36	2,02%		
17	-0,35%	37	2,05%		
18	-0,34%	38	2,04%		
19	-0,19%	39	2,05%		
20	0,09%	40	2,07%		

FIGURE A9: Courbe des taux d'inflation économique

## A.7 Table de mortalité TH/TF 00-02 choquée de -20%

Age	lx_F choc	lx_H choc															
0	100000	100000	20	99419	99136	40	98591	97085	60	94628	88259	80	76141	55125	100	6709	2000
1	99693	99609	21	99391	99055	41	98501	96901	61	94273	87450	81	73858	52188	101	5000	1389
2	99666	99578	22	99364	98972	42	98403	96696	62	93894	86586	82	71315	49085	102	3627	939
3	99650	99557	23	99336	98888	43	98294	96470	63	93486	85661	83	68487	45824	103	2557	616
4	99636	99539	24	99309	98806	44	98175	96219	64	93046	84669	84	65359	42432	104	1748	389
5	99625	99525	25	99282	98723	45	98046	95942	65	92570	83602	85	61927	38948	105	1157	240
6	99615	99512	26	99254	98640	46	97904	95637	66	92053	82452	86	58202	35411	106	738	141
7	99606	99501	27	99226	98556	47	97752	95304	67	91491	81214	87	54213	31860	107	453	82
8	99598	99490	28	99197	98470	48	97587	94945	68	90878	79885	88	50009	28342	108	270	45
9	99590	99480	29	99167	98383	49	97413	94559	69	90207	78458	89	45662	24906	109	152	27
10	99582	99470	30	99136	98293	50	97228	94147	70	89470	76930	90	41260	21603	110	85	16
11	99573	99460	31	99102	98201	51	97033	93708	71	88659	75295	91	36883	18482	111	47	3
12	99565	99449	32	99065	98107	52	96827	93241	72	87766	73546	92	32591	15580	112	19	0
13	99555	99437	33	99025	98009	53	96608	92743	73	86783	71683	93	28428	12927	113	4	0
14	99545	99421	34	98979	97905	54	96375	92212	74	85698	69702	94	24441	10542	114	0	0
15	99532	99400	35	98928	97793	55	96126	91646	75	84497	67600	95	20682	8440			
16	99516	99370	36	98873	97674	56	95860	91044	76	83165	65374	96	17197	6624			
17	99497	99330	37	98813	97544	57	95577	90407	77	81683	63017	97	14032	5089			
18	99473	99277	38	98746	97404	58	95278	89733	78	80031	60528	98	11221	3822			
19	99447	99211	39	98672	97252	59	94962	89018	79	78191	57901	99	8778	2801			

FIGURE A10: Table de mortalité TH/TF 00-02 choquée

## A.8 Courbe de taux économiques Formule standard, incluant la volatility adjustment (VA)

Période	Taux_actu														
1	-0,351%	21	0,623%	41	1,882%	61	2,525%	81	2,862%	101	3,066%	121	3,204%	141	3,302%
2	-0,321%	22	0,685%	42	1,927%	62	2,546%	82	2,874%	102	3,075%	122	3,209%	142	3,306%
3	-0,268%	23	0,753%	43	1,970%	63	2,568%	83	2,887%	103	3,083%	123	3,215%	143	3,311%
4	-0,215%	24	0,824%	44	2,011%	64	2,588%	84	2,899%	104	3,090%	124	3,221%	144	3,315%
5	-0,159%	25	0,897%	45	2,051%	65	2,608%	85	2,910%	105	3,098%	125	3,226%	145	3,319%
6	-0,094%	26	0,971%	46	2,089%	66	2,628%	86	2,922%	106	3,106%	126	3,231%	146	3,323%
7	-0,014%	27	1,044%	47	2,126%	67	2,646%	87	2,933%	107	3,113%	127	3,237%	147	3,327%
8	0,052%	28	1,117%	48	2,162%	68	2,665%	88	2,944%	108	3,120%	128	3,242%	148	3,330%
9	0,117%	29	1,188%	49	2,196%	69	2,682%	89	2,955%	109	3,127%	129	3,247%	149	3,334%
10	0,183%	30	1,257%	50	2,229%	70	2,700%	90	2,965%	110	3,134%	130	3,252%	150	3,338%
11	0,234%	31	1,325%	51	2,261%	71	2,716%	91	2,975%	111	3,141%	131	3,257%		
12	0,283%	32	1,390%	52	2,291%	72	2,733%	92	2,985%	112	3,148%	132	3,262%		
13	0,338%	33	1,453%	53	2,321%	73	2,749%	93	2,995%	113	3,155%	133	3,266%		
14	0,391%	34	1,514%	54	2,349%	74	2,764%	94	3,005%	114	3,161%	134	3,271%		
15	0,432%	35	1,573%	55	2,377%	75	2,779%	95	3,014%	115	3,168%	135	3,276%		
16	0,459%	36	1,630%	56	2,404%	76	2,794%	96	3,023%	116	3,174%	136	3,280%		
17	0,479%	37	1,684%	57	2,430%	77	2,808%	97	3,032%	117	3,180%	137	3,285%		
18	0,501%	38	1,736%	58	2,454%	78	2,822%	98	3,041%	118	3,186%	138	3,289%		
19	0,530%	39	1,787%	59	2,479%	79	2,836%	99	3,050%	119	3,192%	139	3,294%		
20	0,570%	40	1,835%	60	2,502%	80	2,849%	100	3,058%	120	3,198%	140	3,298%		

FIGURE A11: Courbe de taux économiques incluant la VA

## A.9 Courbes des taux économiques choquées à la hausse et à la baisse, Formule standard, incluant la volatility adjustment (VA)

Période	Hausse des taux	Baisse des taux	Période	Hausse des taux	Baisse des taux	Période	Hausse des taux	Baisse des taux	Période	Hausse des taux	Baisse des taux
1	0,649%	-0,351%	17	1,479%	0,364%	33	2,453%	1,070%	49	3,196%	1,650%
2	0,679%	-0,321%	18	1,501%	0,380%	34	2,514%	1,116%	50	3,229%	1,678%
3	0,732%	-0,268%	19	1,530%	0,397%	35	2,573%	1,161%			
4	0,785%	-0,215%	20	1,570%	0,425%	36	2,630%	1,204%			
5	0,841%	-0,159%	21	1,623%	0,463%	37	2,684%	1,245%			
6	0,906%	-0,094%	22	1,685%	0,508%	38	2,736%	1,285%			
7	0,986%	-0,014%	23	1,753%	0,557%	39	2,787%	1,324%			
8	1,052%	0,052%	24	1,824%	0,608%	40	2,835%	1,362%			
9	1,117%	0,101%	25	1,897%	0,661%	41	2,882%	1,398%			
10	1,183%	0,148%	26	1,971%	0,715%	42	2,927%	1,434%			
11	1,234%	0,185%	27	2,044%	0,768%	43	2,970%	1,468%			
12	1,283%	0,221%	28	2,117%	0,821%	44	3,011%	1,500%			
13	1,338%	0,263%	29	2,188%	0,873%	45	3,051%	1,532%			
14	1,391%	0,301%	30	2,257%	0,924%	46	3,089%	1,563%			
15	1,432%	0,334%	31	2,325%	0,975%	47	3,126%	1,593%			
16	1,459%	0,350%	32	2,390%	1,023%	48	3,162%	1,622%			

FIGURE A12: Courbe de taux économiques choquée à la hausse et à la baisse de +/- 100 bp, incluant la VA

---

# Bibliographie

- [1] BONTOUX A. “Les rentes en assurance et en réassurance non-vie”. In : *ce qu'il faut savoir, Swiss Re.* (2007).
- [2] ACPR. “<https://acpr.banque-france.fr/>”. In : *Solvabilité II* (2007).
- [3] Intranet ARPEGE. “Indemnisation des dommages corporels : Recueil méthodologique commun”. In : *Préjudice Corporel* (2013).
- [4] Anaïs BELABED. “Modélisation de la sinistralité atypique en RC automobile avec prise en compte des spécificités d’un versement en rente”. In : *Mémoire d’actuariat, ISUP* (nov. 2016).
- [5] CEA. “Les engagements d’assurance vie”. In : *CEA 1ère année – 02/02/17* (Février 2017).
- [6] Christophe CHESNEAU. “Introduction aux arbres de décision (de type CART)”. In : *Master. France. 2019. ffcel-02281064f* (2019).
- [7] Thamar DABI. “Modélisation des sinistres corporels Auto par une approche hybride IARD/Vie”. In : *Mémoire d’actuariat, ISFA* (nov. 2016).
- [8] Antoine MISERAY FRÉDÉRIC PLANCHET. “Tarification IARD”. In : *Introduction aux techniques avancées* (mar. 2017).
- [9] Documentation GROUPAMA. “Calcul du SCR de souscription Non-Vie: risques de réserves et de primes”. In : *Corpus Méthodologiques des risques de Primes et de Reserves* (Août 2017).
- [10] Documentation GROUPAMA. “Règles spécifiques applicables aux sinistres corporels Rc auto/Rc non auto”. In : *Guide de provisionnement* (Août 2018).
- [11] Documentation GROUPAMA. “Rapport de modélisation”. In : *RISQUE DE PRIME ; 04 – MTPL* (Décembre 2019).
- [12] Jean-Christophe HAMERY. “Provisionnement des sinistres corporels en RC Automobile”. In : *Mémoire d’actuariat, CEA* (sept. 2014).
- [13] DINTILHAC J.-P. “Rapport du groupe de travail chargé d’élaborer une nomenclature des préjudices corporels”. In : (2005).
- [14] Julien JACQUES. “Test d’hypothèses et régression linéaire”. In : *Université Lumière Lyon 2* ().
- [15] KPMG. “Modélisation des rentes”. In : *Journées IARD 2018* (mar. 2018).
- [16] Laurent ROUVIÈRE. “Régression logistique avec R”. In : *Université Rennes 2* ().
- [17] OROZCO PENALOZA SERGIO. “Mise en place d’une couverture en réassurance rétrospective sur les dommages corporels de la branche automobile”. In : *Mémoire d’actuariat, ISFA* ().