Mémoire présenté le 20 juin 2018:

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

Par : Damien LANDON

Titre : Construction d'une échelle bonus-malus pour la tarification des flottes de véhicules

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

X. MILHAUD

J. VIGNANCOUR

Membres présents du jury de l'ISFA

N. LEBOISNE

Entreprise :

Nom :

Signature :

Directeur de mémoire en entreprise :

Nom :

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

Mots-clés : Flotte de véhicules, garantie responsabilité civile, théorie de la crédibilité, classification, échelle bonus-malus, modèles linéaires généralisés, valeurs extrêmes

Dans le cadre de la révision de son tarif sur l'assurance des flottes de véhicules, une société d'assurance souhaite étudier la prise en compte de la sinistralité antérieure dans la tarification. Les flottes qui nous étudions sont de petite taille, elles sont composées de 5 à 39 véhicules. Leur activité principale est le transport pour propre compte. Elles peuvent être composées de voitures, camionnettes, camions, engins de chantier...

Actuellement, la sinistralité est prise en compte à la création du contrat à l'aide d'un coefficient. Il est basé sur la fréquence des sinistres responsabilité civile responsable des trois dernières années. Celui-ci est fixé une fois pour toute et la revalorisation ne se base plus sur la fréquence mais sur le résultat du contrat. La plupart du temps, il n'est revu qu'à la hausse. L'objectif du travail est d'intégrer une nouvelle variable directement dans la tarification de la garantie responsabilité civile comme la zone ou l'activité.

Les deux variables communes à toutes les catégories de véhicules sont la zone et l'activité. Nous montrons que ces variables sont explicatives de la fréquence calculée sur les trois années précédentes. Nous utilisons la théorie de la crédibilité pour calculer une fréquence crédibilisée pour chaque flotte à partir de la fréquence de son groupe zone-activité. A partir de cette fréquence crédibilisée, nous constituons un indice qui est le rapport entre cette fréquence et la fréquence de la zone-activité. Nous classons alors les indices à l'aide d'une classification ascendante hiérarchique. Nous obtenons neuf classes et nous définissons des règles d'évolution sur cette échelle.

Nous utilisons cette échelle dans la modélisation de la garantie responsabilité civile matérielle. Au préalable, nous avons calculer le niveau où se situe la flotte chaque année. Nous pouvons alors utiliser cette variable comme variable explicative. Nous décomposons les sinistres en sinistres hors-graves, sinistres IDA (indemnisation direct de l'assuré) et sinistres graves. Nous modélisons chaque partie par un modèle linéaire généralisé en introduisant le niveau sur l'échelle. La nouvelle variable est bien explicative des fréquences des sinistres hors-grave et IDA.

Abstract

Keywords : Fleets of vehicles, civil liability guarantee, credibility theory, classification, bonus-malus scale, generalizes linear model, extreme value theory

In order to revise the tariff of fleets of vehicles, a mutual insurance group wants to study and change the use of past claim counts in pricing. The fleets we are studying are small. They count between 5 and 39 vehicles. Their main activity is transport for own account. They can be composed of cars, trucks, construction machines...

Currently, the loss ratio is taken into account when the contract is created using a coefficient. It is based on the frequency of civil liability claims responsible for the last three years. It is calculated once. The revalorization is no longer based on the frequency but on the result of the contract. Most of the time, it is only revised upwards. The aim of this report is to integrate a new variable directly into the tariff of the liability guarantee, as area or activity.

Area and activity are the two variables common to all vehicle categories. We show that these variables are explanatory of the frequency wich is calculated over the previous three years. We use the credibility theory to calculate a credibilized frequency for each fleet from the frequency of its area-activity group. From this credibilized frequency, we build an index which is the ratio between this frequency and the frequency of the area-activity. Then we classify this indices using a Hierarchical cluster classification. We obtain nine classes and we define rules of evolution on this scale.

We use this scale in the model of the liability insurance. First, we calculate the level of the fleet each year. Then we can use this variable as an explanatory variable. We divide claims into standard claims, IDA claims (direct compensation of the insured) and large claims. We model each part by a generalized linear model by introducing the level on the scale. The new variable give a good explanation of the claim frequency of the standard and IDA claims.

Table des matières

Introduction 7

1 Cadre de l'étude **9**
1.1 Les flottes automobiles 9
1.1.1 Les catégories de véhicule 10
1.1.2 Les garanties 11

2 Construction de la base **15**
2.1 Les bases disponibles 15
2.2 Processus de construction 16
2.3 Base finale 18
2.4 Sinistres non rattachés 19

3 Tarification actuelle du produit Flotte Auto sur les segments S1-S2 **21**
3.1 Description du portefeuille Flotte Auto sur les segments S1-S2 21
3.2 Le coefficient de personnalisation technique 23
3.3 Définition et calcul du CPT 23
3.4 Biais dans le calcul de la fréquence 24
3.5 Première constations et limites du CPT 25
3.5.1 Introduction à la théorie de la crédibilité 25
3.5.2 Interprétation du CPT 26
3.5.3 Valeur du CPT en fonction du nombre de moteurs 26
3.5.4 Conclusions sur le CPT 29
3.6 Sinistres RC 29
3.6.1 Les conventions IRSA et IRCA 30
3.6.2 Répartition des sinistres 31
3.6.3 Étude des sinistres responsables 33

4 Crédibilité et échelle bonus-malus **35**
4.1 Présentation de la théorie de la crédibilité 35
4.2 Le modèle de Bühlmann-Straub 36
4.2.1 Théorie de Bühlmann-Straub 36

4.2.2 Exemple . . . 37
4.3 Application du modèle de Bühlmann-Straub . . . 38
4.3.1 Étude de la fréquence des sinistres RC responsables . . . 38
4.3.2 Résultats par croisement zone et activité . . . 42
4.3.3 Élaboration d'une échelle . . . 49
4.3.4 Autre proposition d'échelle . . . 52
4.3.5 Test de la première échelle . . . 55
4.4 Règle de mouvement sur l'échelle . . . 57

5 Modélisation de la garantie RC matériel des véhicules de catégorie 1 61
5.1 Distribution des montants . . . 61
5.2 Détermination du seuil de grave . . . 62
5.3 Distribution des sinistres hors-graves . . . 64
5.4 Distribution des fréquences . . . 66
5.4.1 Fréquence des sinistres hors-graves et hors-IDA . . . 66
5.4.2 Fréquence des sinistres IDA . . . 67
5.5 Statistiques descriptives . . . 67
5.6 Modélisation . . . 70
5.6.1 Modélisation de la fréquence des sinistres hors-graves . . . 70
5.6.2 Modélisation de la fréquence des sinistres IDA . . . 73
5.6.3 Modélisation du coût moyen des sinistres hors-graves . . . 74
5.6.4 Sinistres graves . . . 74
5.7 Etude de l'impact de la variable niveau . . . 74
Conclusion . . . 77

Bibliographie 79

A Loi Badinter 81

B Macro CAH 83

C Résultats de la macro CAH 85

D Relevé d'informations 87

Introduction

Le marché des particuliers étant saturé, un des principaux pôles de développement pour les sociétés d'assurance dans le domaine de l'IARD est le secteur des professionnels et des entreprises. Le marché des flottes de véhicules est un des principaux marchés de ce secteur. Un groupe d'assurance à travers sa société souhaite continuer son développement sur ce secteur en proposant un tarif adapté et compétitif.

Sur les produits actuels, les flottes sont tarifées au véhicule. Cependant, un coefficient majorateur/minorateur, appelé CPT, est appliqué sur les principales garanties. Ce coefficient dépend de la sinistralité antérieure de la flotte lors de la création du contrat chez cette société d'assurance. Il permet de prendre en compte la sinistralité globale de la flotte. Il est continue entre 0,55 et 1,5 et il n'évolue pas au cours du temps. L'objet de ce mémoire est de proposer des évolutions pour ce coefficient. Ce travail est le début d'un projet qui sera poursuivi en 2017 où le changement du CPT devrait être décidé.

Les flottes peuvent regrouper plusieurs types de véhicules comme des voitures, des camions ou des engins de chantier. L'idée est de mesurer la sinistralité sur l'ensemble de ces véhicules et ainsi de donner un niveau de risque à la flotte. En effet, il n'existe pas d'échelle réglementaire comme le bonus-malus pour l'assurance des particuliers. Nous devrons donc tenir compte du fait que le système ne s'appliquera qu'au sein des assurés de cette société d'assurance.

Dans les flottes, un véhicule n'est pas forcément attribué à un conducteur. Il n'est donc pas intéressant de descendre au niveau du véhicule pour la sinistralité. Par contre, le chef d'entreprise doit être en mesure d'influer sur la sinistralité à travers son recrutement et la formation des conducteurs. Il nous parait donc intéressant de créer une variable qui mesure la sinistralité de la flotte à son entrée dans le portefeuille mais qui soit aussi capable d'évoluer avec sinistralité.

Dans une première partie, nous présenterons le cadre de l'étude : la société d'assurance, sa place au sein du groupe et la marché des flottes. Ensuite nous présenterons l'étape de construction de la base qui est un passage obligé et qui aura pris pas mal de temps. Dans une troisième partie, nous décrirons comment est tarifé le produit Flotte Auto et nous présenterons en particulier le coefficient de personnalisation technique. Ce coefficient ressemble à un coefficient bonus-malus et une des principales missions était de le comprendre et de le revoir. Dans la quatrième partie, nous utiliserons la théorie de la crédibilité de Bühlmann-

Straub afin de calculer une fréquence crédibilisée par contrat. Nous utilisons cette fréquence crédibilisée pour construire un indicateur. Nous créons alors une échelle bonus-malus à l'aide d'une classification ascendante hiérarchique. Dans une cinquième et dernière partie, nous inclurons cette échelle dans la modélisation de la fréquence de différents types de sinistre. Nous utiliserons pour cela des modèles linéaires généralisés pour obtenir un tarif de la partie matériel de la garantie responsabilité civile.

Chapitre 1

Cadre de l'étude

Dans cette partie, nous présentons l'entreprise où a été effectuée cette étude puis nous décrivons la branche de l'assurance des professionnels et des entreprises avec ses spécificités sur laquelle portera la modélisation.

Ayant été obligé de retirer le nom de la société de mon mémoire, cette partie est allégée par rapport à la version d'origine.

1.1 Les flottes automobiles

Une flotte automobile correspond en général à un parc d'au moins cinq véhicules. Pour certaines activités, la société d'assurance accepte d'assurer des parcs de moins de cinq véhicules appelés micro-flotte. L'assurance de flotte automobile permet de regrouper sous un seul contrat l'ensemble des véhicules de la flotte. Le parc peut être composé de différents types de véhicules (voitures, motos, camions, engins de chantier ...) qui peuvent avoir différents usages (transport pour propre compte, transport de personnes ou de marchandises ...).

Les activités sur le marché des flottes automobiles sont réparties en cinq grandes familles produits :

- **TPPC** *Transport pour propre compte.* Il existe deux produits différents :
  - **Flotte Auto** : permet de garantir tous les risques liés à la circulation des véhicules immatriculés en France et utilisés pour les besoins de l'entreprise (hors transport routier de marchandises et de voyageurs)
  - **Flotte Mission** : permet de couvrir la responsabilité de l'employeur lorsque ses collaborateurs utilisent leur véhicule personnel pour les besoins de l'activité professionnelle pour les dommages causés à des tiers par un collaborateur en mission ou au véhicule du collaborateur en mission
- **TRM** *Transport routier de marchandises* : les contrats **Flotte Transport** couvrent l'assurance circulation pour le transport routier de marchandises pour le compte de tiers

- **TRV** *Transport public de voyageur* : le produit **Flotte Via** permet de garantir tous les risques liés à la circulation des véhicules utilisés pour le transport routier de voyageurs
- **LCD** *Location de courte durée* le produit **LCD** permet de garantir des véhicules appartenant à l'assuré et donnés en location sans chauffeur à des tiers pour une durée de moins de douze mois
- **LLD** *Location de longue durée* même produit que pour la LCD mais pour une durée de plus de douze mois.

Le produit Flotte Auto est de loin le produit le plus important de la gamme que ce soit en nombre ou en montant. Il représente a lui seul plus de 70% de la part des cotisations et plus de 75% des contrats. Notre étude portera sur ce produit qui est le plus important pour l'entreprise.

La tarification sur Flotte Auto se fait véhicule par véhicule et garantie par garantie.

1.1.1 Les catégories de véhicule

Les véhicules sont répartis en cinq catégories dans lesquelles nous distinguons les véhicules moteurs et non-moteurs :

- **1ere catégorie** : véhicules à moteur de PTAC (*poids total à charge*) inférieur ou égal à 3,5T
- **2ième catégorie** : véhicules non agricoles de PTAC supérieur à 3,5T (camion, tracteur routier, remorque, semi-remorque)
- **3ième catégorie** : véhicules à deux roues, tricycles, quadricycles, voitures sans permis
- **4ième catégorie** : véhicules utilisés en tant qu'outils et remorques d'un poids inférieur ou égal à 3,5 tonnes
- **5ième catégorie** : véhicules de plus de 9 places, affectés au transport de personnes (auto-car, mini-car)

Par la suite, nous utilisons la notion de véhicule moteur et non-moteur. Tous les véhicules de 1ere, 2ième, 3ième et 5ième catégorie sont des véhicules moteurs. Dans la 4ième catégorie, nous trouvons les deux types de véhicules. Pour simplifier, un véhicule non-moteur est un engin de chantier qui ne peut pas rouler sur la route.

Nous définissons la notion de moteur-année et de véhicule-année qui indique le temps de présence d'un véhicule dans le portefeuille. Si un véhicule moteur est présent en portefeuille du 01/01/2012 au 30/06/2012, il comptera pour 0,5 véhicule-année (VA) et 0,5 moteur-année (MA). Si un véhicule non-moteur est présent en portefeuille du 01/01/2012 au 30/06/2012, il comptera pour 0,5 véhicule-année (VA) et 0 moteur-année (MA).

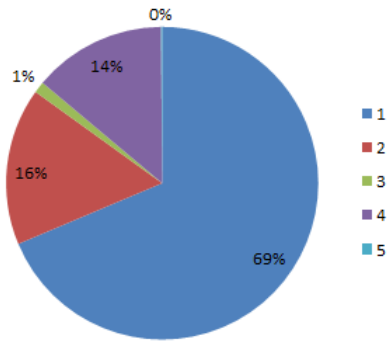


FIGURE 1.1 – Répartition des cotisations par catégorie en 2015

Les véhicules de 1ère catégorie sont largement majoritaires et représentent près de 69% des cotisations totales pour Flotte Auto. Les véhicules de 2ième catégorie et 4ième catégorie représentent chacun environ 15% des cotisations alors que les véhicules de 3ième et 5ième catégorie ne représentent que 1% des cotisations.

1.1.2 Les garanties

Les garanties classiques de Flotte Auto sont les suivantes :

- **responsabilité civile automobile** (RC) : elle couvre l'assuré contre les dommages causés par son véhicule à autrui. La loi prévoit que "tout propriétaire d'un véhicule terrestre à moteur doit l'assurer pour pouvoir le faire circuler". Cette garantie est donc obligatoire. Nous distinguons la responsabilité civile corporelle (RCC) et la responsabilité civile matérielle (RCM),
- **dommages tous accidents** (DTA) : elle prévoit l'indemnisation du propriétaire assuré pour tous les dommages subis par son véhicule, quel que soit le type d'accident ou la responsabilité du conducteur
- **bris de glaces** (BDG) : elle couvre les dommages fait au pare-brise, aux vitres latérales, arrières et de toit, verres de blocs optiques des phares et aux rétroviseurs extérieurs,
- **vol et tentative de vol** : elle couvre les dommages aux véhicules lors d'un vol ou d'une tentative de vol,
- **incendie** : elle couvre les dommages aux véhicules lors d'un incendie,
- **dommages corporels du conducteur** : elle permet une indemnisation si le conducteur fautif est blessé pendant un accident ou s'il n'y a aucun responsable désigné,
- **défense et recours** :elle permet d'assurer la défense du conducteur en garantissant une aide juridique et financière si l'autre personne impliquée dans l'accident intente

une action judiciaire à son encontre. Elle permet également de formuler un recours judiciaire dans le but d'être dédommagé du préjudice causé par l'autre automobiliste, en fournissant une aide juridique et financière.

- **évenements naturels et catastrophes naturelles** : elle couvre les dommages directs aux biens meubles et immeubles (toiture emportée par une tempête, matériel endommagé lors d'une inondation par exemple) ou aux personnes, mais aussi des dommages consécutifs causés à des tiers (chute de tuiles sur des véhicules en stationnement, d'un arbre sur la maison du voisin...) par l'action des forces de la nature.
- **Indemnisation plus** : elle permet de majorer l'indemnisation du véhicule suite à un sinistre dommage.
- **Assistance** (ASS) : elle permet de porter secours en cas d'accident, de vol ou de panne mécanique du véhicule.

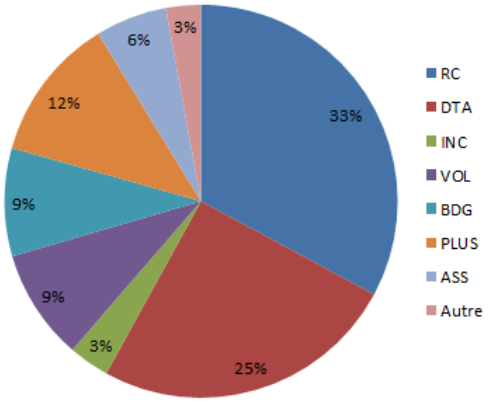


FIGURE 1.2 – Répartition des cotisations par garantie en 2015

Les garanties principales, en nombre et en cotisation, sont les garanties RC et DTA. Notre étude ne portera que sur la garantie RC.

Une flotte ne sera pas tarifée de la même façon si elle a 5 ou 100 véhicules. Dans le premier cas, nous nous baserons principalement sur les caractéristiques des véhicules alors que pour les grandes flottes, nous tiendrons compte davantage de la sinistralité antérieure. Nous distinguons cinq segments sur le produit Flotte Auto :

- **Segment 0 ou micro-flotte** : flotte de moins de 5 véhicules moteurs (réservé à quelques activités particulières comme les ambulances)
- **Segment 1** : flotte de 5 à 10 véhicules moteurs,
- **Segment 2** : flotte de 11 à 39 véhicules moteurs,
- **Segment 3** : flotte de 40 à 99 véhicules moteurs,

- **Segment 4** : flotte de plus de 100 véhicules moteurs.

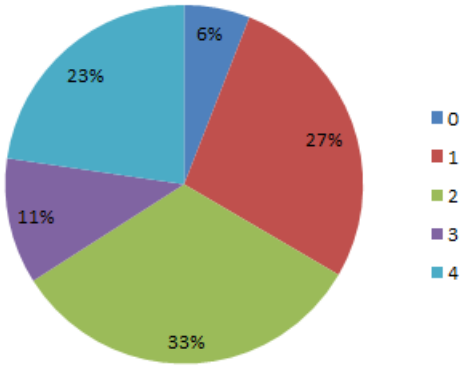


FIGURE 1.3 – Répartition des cotisations par segment en 2015

Les segments 1 et 2 représentent à eux deux 60% des cotisations. Ils représentent donc un enjeu majeur pour l'entreprise et c'est sur ces deux segments que portera l'étude. Ces deux segments sont tarifés de la même façon alors qu'il existe deux autres tarifications pour les segments 3 et 4.

Chapitre 2

Construction de la base

Avant d'effectuer une étude, nous devons construire la base de données qui nous servira pour faire des statistiques descriptives puis des modèles mathématiques. Cette étape est très importante puisque nos résultats dépendront de la qualité des données. Les données doivent donc être les plus fiables possible.

2.1 Les bases disponibles

Les données nécessaires à la construction de la base sont réparties dans cinq bases différentes :

- la **base 1** est au **niveau contrat**. Elle contient le numéro de contrat, le mode de fractionnement, la date d'échéance principale, la date de résiliation...Cette base est corrigée régulièrement. Elle sert à avoir la liste des contrats correspondants au périmètre d'étude
- la **base 2** est également au **niveau contrat** Nous y retrouvons les caractéristiques du contrat comme l'activité, la zone, le nombre de moteurs, le coefficient de personnalisation technique. La base contient une ligne par situation d'un contrat
- la **base 3** est au **niveau véhicule**. Les données ne sont pas toujours corrigées mais c'est la seule base au niveau véhicule dont nous disposons. Elle est découpée en situation de véhicule (une ligne par situation et par véhicule). Chaque véhicule est identifié par son numéro d'immatriculation mais celui-ci n'est pas toujours bien renseigné (NONID1, NONID2, ...). Nous y trouvons les caractéristiques techniques du véhicule comme sa catégorie, sa puissance, son poids, sa valeur, son groupe/classe. Nous y retrouvons également le montant des cotisations par garantie. La base contient les situations de l'année N-4 à l'année N. Nous n'avons donc que les cinq dernières années à notre disposition
- la **base 4** est une base **sinistres hors-assistance**. Chaque ligne correspond à un sinistre et une garantie. Si un sinistre touche deux garanties différentes, nous retrouve-

rons deux lignes dans la base. Nous trouvons la date du sinistre, la garantie touchée, l'immatriculation du véhicule la charge, le montant du recours, le lieu ...

– la **base 5** est une base **sinistres assistance**. Cette base est similaire à la précédente mais ne concerne que les sinistres touchant la garantie assistance

2.2 Processus de construction

Le processus de construction de la base se décompose en six principales étapes données dans le schéma de la figure 2.1 :

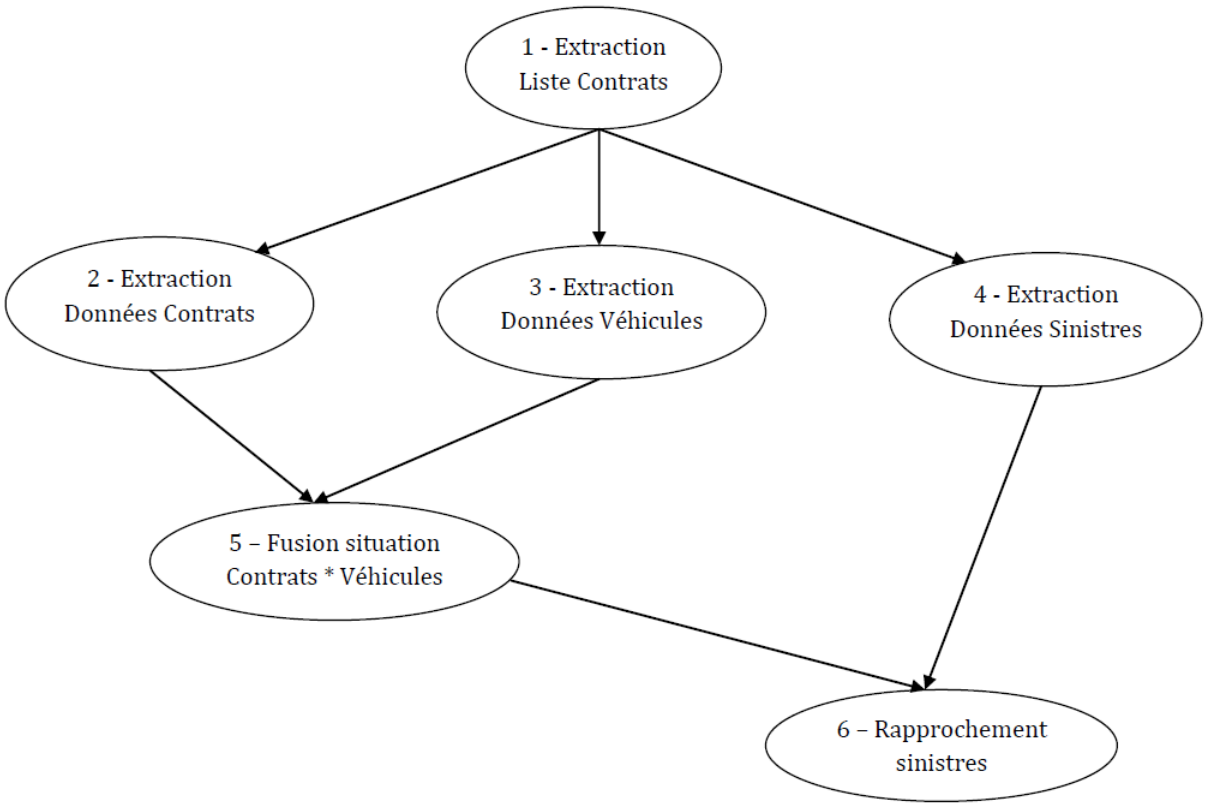


FIGURE 2.1 – Processus de construction de la base de données

Étape 1 Extraction des numéros de contrats du produit Flotte Auto ayant vécu au moins un jour sur la période 2012-2015 dans la base 1. Nous nous limitons à cette période étant donné que la base véhicule ne contient des informations que sur les années de 2012 à 2016. L'année 2016 étant incomplète, nous l'excluons du périmètre d'étude.

Étape 2 Extraction des données techniques du contrat dans la base 2 à l'aide des numéros de contrat extrait lors de l'étape 1. Si une situation s'étale sur plusieurs exercices, elle est découpée.

Étape 3 Extraction dans la base 3 des situations de véhicules avec les données techniques des contrats extraits lors de l'étape 1.

Étape 4 Extraction dans les bases 4 et 5 des sinistres survenus entre 2012 et 2015 et correspondant aux contrats de l'étape 1.

Étape 5 Jointure entre les tables crées aux étapes 2 et 3 en utilisant le numéro de contrat et les dates de situation des contrats et des véhicules : nous agrégeons les situations dont l'intersection entre les périodes est non-vide. Il faut alors recalculer les cotisations et les véhicules-année pour la nouvelle période. Nous avons une ligne par situation contrat*véhicule.

Étape 6 Rapprochement des sinistres par le numéro de contrat, le numéro d'immatriculation ou le numéro d'ensemble et la date de survenance.

Les étapes 1 à 5 ne posent pas particulièrement de problèmes. L'étape 6 présentent plus de difficultés qui proviennent de deux sources :

- le mode de gestion du contrat
- la mauvaise saisie du numéro d'immatriculation

Les contrats peuvent être gérés suivant trois modes :

- **au mouvement** : l'assuré signale immédiatement à l'assurance tous les mouvements sur sa flotte. Le changement entraîne une modification immédiate de la prime
- **périodique** : l'assuré signale immédiatement à l'assurance tous les mouvements sur sa flotte mais la facturation n'est faite qu'à l'échéance
- **demi-différence** : l'assuré déclare son parc à chaque échéance. Le parc est amené à fluctuer sans que l'assureur ne connaisse exactement les dates d'entrée et de sortie du véhicule

Le mode demi-différence pose des problèmes pour le rattachement d'un sinistre à une situation contrat*véhicule : la date de survenance du sinistre peut ne pas appartenir à une période d'une situation contrat*véhicule. Dans le cas des contrats gérés en demi-différence, nous choisissons de rattacher un sinistre à la situation correspondante la plus proche.

La mauvaise saisie du numéro d'immatriculation peut intervenir sous deux formes :

- l'immatriculation n'est pas bien saisie au moment de la déclaration du véhicule
- erreur lors de la déclaration du sinistre (interversion de chiffres ou lettres, erreur d'un chiffre ou d'une lettre).

Dans ces deux cas, nous pouvons utiliser un autre identifiant qui est le numéro de l'ensemble du véhicule. Tous les véhicules d'une flotte appartienne à un ensemble. Tous les véhicules présentant les même caractéristiques appartiennent au même ensemble. Si un ensemble ne

contient qu'un seul véhicule, nous pouvons alors utiliser ce numéro d'ensemble pour rattacher un sinistre avec une situation contrat*véhicule.

2.3 Base finale

La base finale contient l'ensemble des situations contrat*véhicule sur au plus un exercice. Si une modification est faite sur un contrat ou un véhicule, une nouvelle situation est créée. Si la situation contrat*véhicule est impactée par plusieurs sinistres ou par un sinistre touchant plusieurs garanties, la situation sera dupliquée autant de fois qu'il y a de sinistres ou de garanties sur cette situation. Nous aurons sur chaque ligne le numéro de sinistre et la garantie. Le couple numéro de sinistre et garantie est unique. Une vision de la base finale est donnée dans la figure 2.2.

Num contrat	date début situ	date fin situ (coupée en exercice)	Activité	Zone	Groupe/ classe	Autres critères techniques	Immat	PresRC	Franchise RC	sin RC_Mat	charges RC_Mat	sin RC_Corp	charges RC_Corp	Pres DTA	Franchise DTA	sin DTA	charge DTA	,,,
cont A	01/01/2014	31/12/2014	Commerçant	2	GC2	,,,	immat 1	0	0	0	0	0	0	0	0	0	0	
cont B	01/01/2014	15/05/2014	Collect.	3	GC2	,,,	immat 2	1	200	0	0	0	0	0	0	0	0	
cont B	16/05/2014	30/06/2014	Collect.	3	GC2	,,,	immat 2	1	200	0	0	0	0	1	400	0	0	
cont B	01/07/2014	31/12/2014	Collect.	2	GC2	,,,	immat 2	1	200	1	100	0	0	1	400	0	0	
cont C	01/01/2014	30/03/2014	Artisans	1	GC3	,,,	immat 3	1	0	1	200	0	0	0	0	0	0	
cont C	01/01/2014	30/03/2014	Artisans	1	GC3	,,,	immat 3	1	0	1	400	0	0	0	0	0	0	
cont C	31/03/2014	31/12/2014	Artisans	1	GC3	,,,	immat 3	1	200	0	0	1	600	0	0	0	0	
cont C	01/01/2014	30/03/2014	Artisans	1	GC3	,,,	immat 4	1	0	0	0	0	0	0	0	0	0	
cont C	31/04/2014	31/12/2014	Artisans	1	GC3	,,,	immat 4	1	200	1	100	0	0	0	0	0	0	
cont C	31/04/2014	31/12/2014	Artisans	1	GC3	,,,	immat 4	1	200	1	500	0	0	0	0	0	0	
cont C	31/04/2014	31/12/2014	Artisans	1	GC3	,,,	immat 4	1	200	1	1700	0	0	0	0	0	0	
cont D	01/11/2013	31/12/2013	Commerçant	3	GC2	,,,	immat 5	1	200	0	0	0	0	1	200	0	0	
cont D	01/01/2014	30/04/2014	Commerçant	3	GC2	,,,	immat 5	1	200	1	200	0	0	1	200	0	0	
cont D	01/01/2014	30/04/2014	Commerçant	3	GC2	,,,	immat 5	1	200	0	0	1	1300	1	200	0	0	
cont D	01/05/2014	31/12/2014	Commerçant	3	GC2	,,,	immat 5	1	0	1	680	0	0	1	200	0	0	
cont D	01/05/2014	31/12/2014	Commerçant	3	GC2	,,,	immat 5	1	0	0	0	0	0	1	200	1	70	
cont D	01/11/2013	31/12/2013	Commerçant	3	GC2	,,,	immat 6	1	200	0	0	0	0	1	200	0	0	
cont D	01/01/2014	30/04/2014	Commerçant	3	GC2	,,,	immat 6	1	200	0	0	0	0	1	200	0	0	
cont D	01/05/2014	31/12/2014	Commerçant	3	GC2	,,,	immat 6	1	0	0	0	0	0	1	200	1	200	
cont D	01/11/2013	31/12/2013	Commerçant	3	GC3	,,,	immat 7	1	200	0	0	0	0	1	200	0	0	
cont D	01/01/2014	30/04/2014	Commerçant	3	GC3	,,,	immat 7	1	200	0	0	0	0	1	200	0	0	
cont D	01/05/2014	31/12/2014	Commerçant	4	GC3	,,,	immat 7	1	0	0	0	0	0	1	200	0	0	

FIGURE 2.2 – Exemple simplifié de la base finale

Si nous regardons le contrat C, nous observons quatre situations contrat*véhicule distinctes qui concernent deux véhicules différents. Pour les deux véhicules, il y a deux situations pour l'année 2014. Une nouvelle situation est créée suite à la modification de la franchise. La situation du contrat C et de l'immatriculation Immat3 du 01/01/2014 au 30/03/2014 est dupliquée car elle est touchée par deux sinistres RC matériel différents.

2.4 Sinistres non rattachés

Nous avons un certain nombre de sinistres que nous n'avons pas réussi à associer avec une situation contrat*véhicule.

	Non rattaché	Rattaché	% Non rattaché
Nombre de sinistres hors assistance	26 889	386 592	6,50 %
Charge hors assistance	16, 7 M€	630 M€	2,58 %
Nombre de sinistres assistances	8 624	78 052	9,95 %
Charge assistance	1, 7 M€	15,8 M€	9,74 %
Nombre de sinistres hors assistance segment 1-2	9 146	224 942	3,91 %
Charge hors assistance segment 1-2	14, 5 M€	367,6 M€	3,80 %

TABLE 2.1 – Statistiques sur les sinistres non-rattachés

Dans le tableau 2.1, nous voyons que les sinistres hors assistance sont mieux rattachés que les sinistres assistances. Pour les sinistres assistances, nous n'avons pas le numéro d'ensemble qui permet de rattacher davantage de sinistres. Dans les sinistres assistances, les sinistres sont mieux rattachés en montant qu'en nombre. Nous avons rattaché à la main quelques sinistres graves dont un de plus de 9 millions d'euros. Ceci permet d'améliorer facilement la proportion de la charge sinistre rattachée mais pas le nombre.

Notre étude portera sur les segments 1 et 2. Sur ces segments, la proportion en nombre et en montant est similaire et en dessous de 4%. Nous n'avons pas trouvé d'autre méthode pour rattacher les sinistres restants, sans devoir les traiter un par un à la main. Nous avons considéré que cette proportion était acceptable. Les sinistres de ces segments sont moins importants en montant ce qui explique la perte au niveau de la proportion de la charge rattachée. Par contre, ces contrats sont souvent gérés au mouvement et il est plus facile de les rattacher en nombre.

En descendant au niveau garantie, nous avons les résultats suivants pour les garanties qui nous intéresserons dans ce mémoire :

- en RCC, $1,5\%$ des sinistres ne sont pas rattachés en nombre pour 3% de la charge
- en RCM, $1,1\%$ des sinistres ne sont pas rattachés en nombre pour $4,8\%$ de la charge

Ces résultats sont bons pour l'étude de la fréquence mais moins bons pour l'étude du coût moyen.

Chapitre 3

Tarification actuelle du produit Flotte Auto sur les segments S1-S2

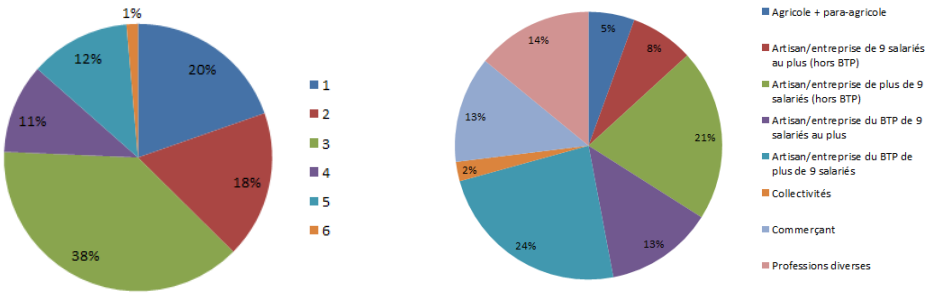
Dans cette partie, nous présentons la dernière tarification qui est actuellement utilisée. Cela permet de comprendre comment est construit le tarif et de dégager les atouts et les inconvénients de celui-ci afin de proposer des évolutions. L'étude se concentre sur l'activité TPPC (transport pour propre compte) pour les flottes des segments 1 et 2 (5 à 39 véhicules moteurs). Nous commençons par présenter les critères utilisés pour la tarification de la garantie que nous souhaitons faire évoluer : la garantie RC.

3.1 Description du portefeuille Flotte Auto sur les segments S1-S2

Lors de la dernière tarification, les variables utilisées pour la garantie RC sont les suivantes :

- l'activité qui correspond à l'activité principale de l'entreprise. Cette variable comporte huit classes :
  - agricole + para-agricole
  - artisan/entreprise du BTP de plus de 9 salariés
  - artisan/entreprise du BTP de 9 salariés au plus
  - artisan/entreprise de 9 salariés au plus (hors BTP)
  - artisan/entreprise de plus de 9 salariés (hors BTP)
  - professions diverses
  - commerçant
  - collectivités
- la zone qui comporte six classes, numérotée de 1 à 6, de la moins à la plus risquée

- les caractéristiques du véhicule suivant sa catégorie :
  - pour la catégorie 1, le groupe/classe qui est une combinaison entre la puissance et la valeur du véhicule
  - pour la catégorie 2, le PTAC
  - pour la catégorie 3, la cylindrée
  - pour la catégorie 4, le type d'engin
  - pour la catégorie 5, le nombre de places
- l'âge du véhicule avec deux classes suivant que le véhicule à plus ou moins de 5 ans
- la présence de multi-garantie RC et DTA.
- la franchise
- le mode de gestion
- le fractionnement de la prime : annuel, semestriel ou trimestriel



(a) Zone

(b) Activité

FIGURE 3.1 – Répartition du nombre de contrats passés dans le portefeuille entre 2012 et 2015 par zone et par activité

Sur la figure 3.1(a), la zone 3 est la plus représentée en nombre de contrats avec 38% des affaires nouvelles. Nous trouvons ensuite la zone 1 puis la zone 2. Ces trois zones regroupent plus de 75% des contrats. Les zones 4 et 5 sont correctement représentées avec 11% et 12%. La zone 6 ne représentant que 1% des contrats, il faudra surveiller les valeurs que nous obtenons pour cette zone.

Au niveau des activités (figure 3.1(b)), deux activités représentent plus de 20% des contrats : artisan/entreprise de plus de 9 salariés (hors BTP) et artisan/entreprise du BTP de plus de 9 salariés. Trois activités sont peu présentes dans le portefeuille avec chacune

moins de 10% : les collectivités, l'agricole et le para-agricole et artisan/entreprise de moins de 9 salariés (hors BTP). Pour le cas des activités, il est plus difficile de faire des rapprochements comme regrouper les collectivités et l'agricole qui sont des activités très différentes. Nous pourrions regrouper tous les artisan/entreprise (hors BTP) en une seule famille qui représenterait 29% des affaires ou revoir le seuil de 9 salariés. En revoyant ce nombre à la baisse, nous aurions deux familles de taille supérieure à 10%. Il faudra contrôler plus particulièrement les résultats obtenus pour les trois familles sous-représentées.

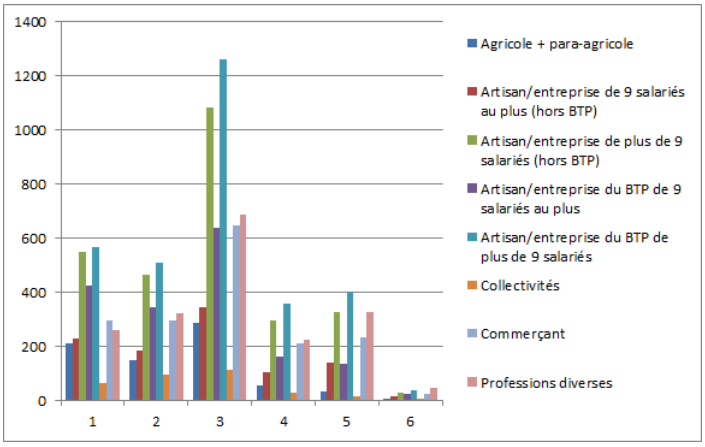


FIGURE 3.2 – Répartition des affaires nouvelles entre 2012 et 2015

En croisant la zone et l'activité, nous obtenons la distribution de la figure 3.2 où nous retrouvons ce que nous avons vu sur chaque variable séparément. Nous avons 48 classes pour 13 222 contrats, 200 contrats représentent donc $1,5\%$ du nombre total. Sur les zones 4 à 6, peu d'activités sont bien représentées. Sur les zones 1 à 3, toutes les activités sont assez bien représentées hormis les collectivités.

3.2 Le coefficient de personnalisation technique

3.3 Définition et calcul du CPT

Le coefficient de personnalisation technique (CPT) est calculé à l'affaire nouvelle une fois pour toute. Il est calculé à partir de la sinistralité des 36 derniers mois. Pour obtenir le tarif final, nous multiplions le tarif théorique par le CPT.

Ce coefficient permet de prendre en compte la sinistralité globale de la flotte afin de l'utiliser pour la tarification. Sans ce coefficient, chaque véhicule serait tarifé de façon indépendante et nous ne tiendrions pas compte de son appartenance à une flotte. Ce coefficient permet donc de lier tous les véhicules d'une même flotte.

Nous calculons une fréquence propre à l'entreprise en divisant le nombre de sinistres sur les trois dernières années par le nombre de moteurs-année sur les trois dernières années. Cette fréquence est décorrélée de l'effet de la zone et de l'activité puis crédibilisée suivant le nombre de moteurs de la flotte. Le CPT est donnée par la formule suivante :

$$\text{CPT} = C \times \left[\frac{Freq\_flotte}{Freq(zone, activite)} * a + (1 - a) \right] \tag{3.1}$$

oùe

- a dépend du nombre de véhicules moteurs n

$$a = 1 - \frac{4}{n},$$

- $Freq\_flotte$ est la fréquence sur les 36 derniers mois (au moins 30 derniers mois)(les sinistres pris en compte sont les sinistres responsabilité civile automobile (dommages corporels et matériels) responsables et les sinistres réglés au titre de la Loi Badinter),
- $Freq(zone, activite)$ est une fréquence qui dépend de la zone et de l'activité
- C est une constante

Ce coefficient est borné inférieurement à 0,55 et supérieurement à 1,5.

Si l'entreprise a été crée depuis moins de 30 mois, le CPT est fixé à 1. Il paraitrait plus cohérent que le CPT prenne dans ce cas la valeur C. Celle-ci étant actuellement plus petite que 1, les entreprises récentes sont pénalisées par rapport aux entreprises ayant plus d'ancienneté. D'ailleurs, nous observons de bons résultats en terme de S/C sur cette catégorie.

3.4 Biais dans le calcul de la fréquence

Au moment de la souscription d'un nouveau contrat, la compagnie quittée fournit un relevé d'informations (RI) sur lequel figurent les sinistres RC corporels et matériels responsables et loi Badinter (voir annexe A). Des exemples de RI sont donnés en annexe D. Il n'y a pas de relevé type et chaque compagnie peut utiliser son format. Ceci ne facilite pas la lecture et l'extraction des informations. Le client donne également le nombre de moteurs de son entreprise sur les trois années précédant la souscription.

Le nombre de moteur est purement déclaratif. Il ne figure pas sur le relevé d'information. C'est un nombre entier qui ne correspond pas exactement au nombre de moteurs-année du contrat. Ceci peut entraîner un biais sur le calcul de la fréquence de la flotte.

Les sinistres ne sont pas répertoriés dans le fichier de l'AGIRA (*Association pour la Gestion des Informations sur le Risque en Assurance*) comme c'est le cas pour les voitures particulières. L'assureur n'a donc pas de moyen pour vérifier les données figurant sur le relevé d'informations. Nous avons donc un deuxième biais dans le calcul de la fréquence de la flotte.

3.5 Première constations et limites du CPT

Le CPT est calculé lors de la souscription de l'affaire et reste figé tout au long de la vie du contrat. Chaque année le montant de la prime est revalorisé à partir de la prime de l'année précédente et du S/C calculé sur les résultats des dernières années. Si le S/C est bon, il n'y a pas ou alors une faible majoration. Si le S/C est élevé, nous majorons la prime d'un pourcentage pouvant aller jusqu'à 15%. Avec les bornes actuelles du CPT, la prime peut varier du simple au triple, soit une variation de plus 172% entre la prime la plus basse et la prime la plus élevée. La prime est rarement revue à la baisse.

Prenons l'exemple de deux entreprises similaires A et B (même véhicules, même zone, même activité) qui arriveraient en même temps dans le portefeuille de la société d'assurance. L'entreprise A a plus de trois années d'existence et aucun sinistre RC responsable sur les trois dernières années et l'autre vient d'être créée. Notons P la prime de référence. La première année, l'entreprise A paiera $0,55P$ et l'entreprise B paiera P. L'entreprise A a ensuite plusieurs années de suite un mauvais S/C et sa prime est augmentée de 10% chaque année (cela correspond à une forte augmentation) alors que l'entreprise B n'a pas de sinistres RC responsables et sa prime reste la même. Les deux primes seront égales pour $-\ln(0,55)/\ln(1,1) > 6$ années. Il faudra attendre la septième année avant que l'entreprise A ne paie une prime supérieure à l'entreprise B alors que pendant ce temps l'entreprise A aura coûté plus cher que l'entreprise B. Or, le CPT n'existe que chez cette société d'assurance. Ce n'est pas un coefficient réglementaire comme le coefficient bonus-malus pour l'assurance automobile des particuliers. L'entreprise B pourra peut-être trouver des tarifs plus attractifs à la concurrence.

De plus, la valeur du CPT dépend fortement de la valeur de la fréquence $Freq(flotte, activite)$. Or cette valeur n'est pas mise à jour régulièrement alors que cette fréquence a une nette tendance à la baisse ces quatre dernières années, comme nous le verrons dans la section 3.6.3.

3.5.1 Introduction à la théorie de la crédibilité

Nous reviendrons plus en détail sur la théorie de la crédibilité dans la partie 4.1 mais nous présentons rapidement son principe. Nous souhaitons utiliser par exemple la fréquence pour calculer un tarif. Prenons le cas d'une flotte de n véhicules. Nous connaissons la fréquence de la flotte sur les trois dernières années, nous la notons $freq_{flotte}$. Nous connaissons également la fréquence de tout notre portefeuille sur les trois dernières années, nous la notons $freq_{portefeuille}$. La fréquence de la flotte est calculée à partir d'un nombre beaucoup plus petit de véhicules que la fréquence du portefeuille. Elle est donc plus volatile et nous ne pouvons pas nous baser uniquement sur l'expérience de la flotte. L'idée de la crédibilité est de dire que la fréquence devrait se trouver entre la fréquence de la flotte et la fréquence du portefeuille. Cette fréquence crédibilisée doit être d'autant plus proche de la fréquence de la flotte que la flotte est de grande taille. Nous pouvons résumer cela par l'égalité :

$$freq_{credibilise} = z(n)freq_{flotte} + (1 - z(n))freq_{portefeuille} \tag{3.2}$$

où n désigne la taille de la flotte et z est une fonction croissante à valeur dans l'intervalle $[0, 1]$

3.5.2 Interprétation du CPT

Le tarif est calculé à partir de modèles linéaires généralisés (MLG) sur la fréquence et le coût moyen. Nous obtenons pour la fréquence une valeur

$$frequence = \exp\left(\beta_0 + \beta_{activite} + \beta_{zone} + \beta_{group/classe} + \beta_{MULTIGAR}\right) \tag{3.3}$$

Les variables activité et zone sont les plus explicatives et interviennent dans toutes les catégories de véhicules. Nous utilisons la fréquence propre de la flotte pour personnaliser la fréquence de chaque contrat. En utilisant la théorie de la crédibilité, nous remplaçons la fréquence due à l'activité et la zone $\exp(\beta_0 + \beta_{activite} + \beta_{zone})$ par $a \times freq(flotte) + (1 - a)\exp(\beta_0 + \beta_{activite} + \beta_{zone})$ où a est le facteur de crédibilité et $freq(flotte)$ est la fréquence de la flotte. Nous obtenons alors

$$PP = \left(a\frac{freq(flotte)}{\exp(\beta_0 + \beta_{activite} + \beta_{zone})} + (1 - a)\right) frequence \times CM \tag{3.4}$$

Nous réécrivons le quotient

$$\frac{freq(flotte)}{\exp(\beta_0 + \beta_{activite} + \beta_{zone})} = \frac{freq(flotte)\exp(\beta_0)}{\exp(\beta_0 + \beta_{activite})\exp(\beta_0 + \beta_{zone})} \tag{3.5}$$

La valeur $\exp(\beta_0)$ correspond à la fréquence avec les modalités de référence qui sont les plus représentées. Cette valeur doit être proche de la valeur moyenne de la fréquence RC (cela devra être vérifié). De la même façon, les valeurs de $\exp(\beta_0 + \beta_{activite})$ et $\exp(\beta_0 + \beta_{zone})$ peuvent être interprétées comme la fréquence moyenne de l'activité et de la zone. Nous retrouvons l'expression actuelle du CPT à une constante multiplicative près. La valeur de la constante a pu être ajustée pour compenser les approximations faites et avoir l'équilibre entre le montant de prime modélisé et le montant effectif des sinistres.

Ceci est une version simplifié du tarif. La modélisation du tarif RC est plus complexe que la modélisation de la fréquence et du coût moyen puisque nous distinguons les sinistres attritionnels, les sinistres graves, les sinistres IRSA responsables et non responsables.

3.5.3 Valeur du CPT en fonction du nombre de moteurs

Dans cette section, nous considérons une version simplifiée du CPT pour étudier ses variations et ces valeurs en fonction du nombre de moteurs n sur une année. Nous supposons

que ce nombre est constant sur les trois années prises en compte pour le calcul du CPT. Nous regardons ensuite la valeur du CPT, à fréquence fixée, comme une fonction de n, CPT $= f(n)$ où f est la fonction donnée par :

$$f(n) = C \times \left[4 \times \left(1 - \frac{Freq\_flotte}{Freq(flotte, activite)} \right) \frac{1}{n} + \frac{1}{Freq(flotte, activite)} \right] \tag{3.6}$$

Cette fonction dépend donc de la fréquence de la zone et du coefficient d'activité. Nous allons considérer la moyenne des CPT sur l'ensemble des valeurs $Freq(flotte, activite)$. Nous obtenons alors le tableau 3.1. Les valeurs des fréquences correspondent respectivement à : aucun sinistre, un sinistre, deux sinistres et trois sinistres sur trois ans pour une flotte de 5 véhicules moteurs.

		Nombre de moteurs-année			
		5	10	20	40
Fréquence	0	0,64	0,32	0,16	0,08
	0,067	0,78	0,73	0,71	0,7
	0,133	0,92	1,15	1,26	1,32
	0,2	1,05	1,56	1,82	1,94

TABLE 3.1 – Valeur moyen du CPT non borné en fonction du nombre de moteurs-année

Pour une fréquence nulle, c'est à dire s'il n'y a eu aucun sinistre RC responsable sur les trois dernières années, le CPT est linéaire en $1/n$ et nous devrions avoir une différence de 88% entre le CPT pour une flotte de 40 moteurs par rapport à une flotte de 5 moteurs. Dans ce cas cependant, le CPT prend la valeur de la borne inférieure $0,55$. La taille de la flotte n'intervient donc pas sur la valeur du CPT.

Pour une fréquence de $0,067$, les valeurs du CPT sont très proches. Nous n'observons qu'une baisse de 10% du CPT pour 40 moteurs par rapport au CPT pour 5 moteurs. Pour les fréquences plus grandes, la différence est plus grande entre les deux extrêmes. Elle est de plus 43% pour une fréquence de $0,133$ et devrait être de plus 85% pour une fréquence de $0,2$. Cependant, la majoration du CPT à $1,5$ réduit la différence à 43%.

Dans le cas où il n'y a pas de sinistre pour 5 moteurs, le CPT est donc baissé à $0,55$ alors que sa valeur théorique est de $0,64$. Le fait de baisser cette valeur augmente la différence entre zéro sinistre et un sinistre. Théoriquement, il devrait y avoir une augmentation de 22% entre les deux cas. En prenant comme valeur $0,55$, nous doublons presque la différence puisque nous avons alors une augmentation de 42%.

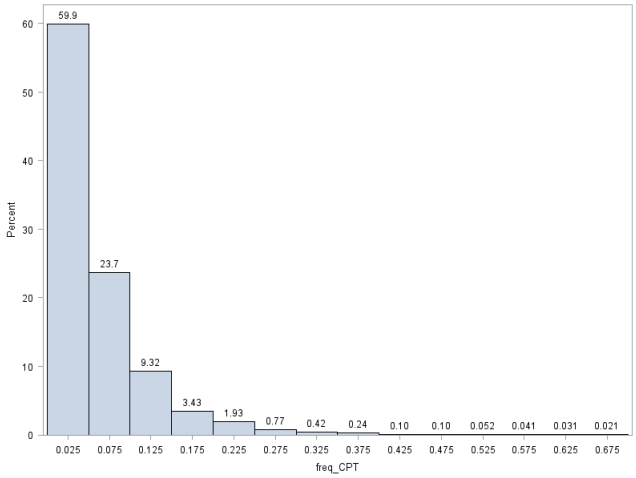


FIGURE 3.3 – Distribution de la fréquence utilisée pour le calcul du CPT sur les affaires nouvelles de 2012 à 2015

La figure 3.3 donne la répartition des fréquences utilisées pour le calcul du CPT. Près de 60% des nouveaux contrats ont une fréquence inférieure à $0,05$ dont 45% de fréquence nulle. Un peu moins de 85% des contrats ont une fréquence inférieure à $0,01$. C'est autour de cette valeur que s'inverse l'ordre des CPT : au-delà de cette valeur, le CPT pour un nombre de moteurs plus élevé devient plus grand que le CPT pour un petit nombre de moteurs.

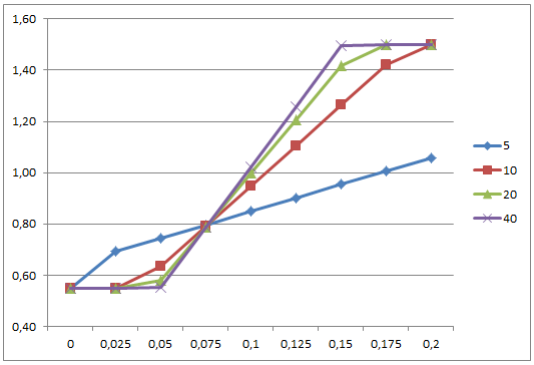


FIGURE 3.4 – CPT en fonction de la fréquence pour la zone 3 et l'activité professions diverses pour quatre valeurs du nombre de moteurs

Le graphique de la figure 3.4 illustre le changement d'ordre des valeurs du CPT. La zone 3 et l'activité professions diverses ont une valeur moyenne de la fréquence zone et du coefficient d'activité. L'égalité a lieu pour une fréquence valant $freq\_egal = Freq\_zone \times Coef\_Activite$ (ici $0,076$). La fréquence moyenne étant de 0.056% Nous pouvons voir que

le tarif avantage nettement les petites flottes dès que la fréquence dépasse $freq\_egal$. Cela peut en partie expliquer le petit nombre de contrats dès que la fréquence dépasse 0,1. L'écart maximum entre le CPT d'une petite flotte et d'une grande flotte dépasse les 55% alors que l'écart maximum pour les petites fréquences est de 25%. Nous pensons qu'il faut revoir le coefficient de crédibilisation qui déséquilibre le tarif en faveur des petites flottes. L'écart est d'autant plus important que le CPT est calculé une fois pour toute au début du contrat et va donc impacter la prime tout au long de la vie du contrat.

De plus le coefficient de crédibilisation vaut déjà 90% pour 40 moteurs. Ceci laisse peut de marge de manœuvre pour le segment 3 qui correspond aux flottes de 40 à 99 moteurs. Actuellement le coefficient de crédibilisation prend des valeurs inférieurs à 90% sur le début du segment 3.

3.5.4 Conclusions sur le CPT

Cette présentation du CPT nous permet de mettre en évidence les défauts du coefficient actuel :

- il existe deux biais dans le calcul de la fréquence de la flotte causée par l'imprécision sur le nombre de moteurs et l'impossibilité de vérifier les sinistres figurants au relevé d'informations
- il est fixe tout au long du contrat et peut varier du simple au triple
- il pénalise les jeunes entreprises
- il "favorise" les petites flottes
- le coefficient de crédibilisation croît trop vite
- le coefficient n'est pas intégré à la modélisation
- la fréquence qui dépend de la zone et de l'activité, $Freq(zone, activite)$ n'est pas mise à jour régulièrement.

Il existe peu de littérature sur la tarification des flottes de véhicules. Nous pouvons citer les travaux fait au Canada [ADD04]. Cela concerne des flottes homogènes qui sont composées de camions. Les services responsables d'étudier ce qui ce fait à la concurrence n'ont pas pu nous donner des renseignements sur la prise en compte de la sinistralité.

3.6 Sinistres RC

Avant de décrire les sinistres RC, nous décrivons deux conventions entre assureurs qui impactent le montant des sinistres.

3.6.1 Les conventions IRSA et IRCA

Les assureurs ont mis en place des conventions pour permettre aux assurés, dans la plupart des cas, d'être dédommagés plus rapidement et directement par leur propre assureur :

Quels que soient la typologie de l'accident de la circulation, la nature et le montant des dommages, les sociétés adhérentes s'obligent, préalablement à l'exercice de leurs recours, à indemniser elles-mêmes leurs assurés, dans la mesure de leur droit à réparation, déterminé selon les règles du droit commun.

L'assureur établit lui-même la responsabilité de son assuré et l'indemnise directement des dommages et préjudices subis avant de présenter un recours à l'assureur adverse.

Dans le domaine de l'assurance automobile, il existe deux types de convention suivant que le sinistre soit matériel ou corporel :

- la convention IRSA (convention d'indemnisation directe de l'assuré et de recours entre les sociétés d'assurance automobile) concerne l'indemnisation des dommages matériels
- la convention IRCA (convention d'indemnisation et de recours corporel automobile) concerne les dommages corporels.

L'état, qui est son propre assureur, a également signé une convention avec les sociétés d'assurances.

Le principe de l'indemnisation directe : chaque assureur de responsabilité civile indemnise directement son assuré dans la mesure où celui-ci a un droit à indemnisation. Non fautif dans la réalisation de l'accident, il est remboursé en totalité ; en partie fautif, il ne reçoit qu'une indemnité partielle. Dans le cas d'une responsabilité totale et en l'absence de garantie « dommages », aucune indemnité n'est versée.

En cas d'accident causé par un ensemble constitué d'un véhicule tracteur et d'une remorque garantis en responsabilité civile par des assureurs différents, la demande d'indemnisation peut être adressée soit à l'assureur du véhicule tracteur, soit à l'assureur de la remorque.

2012	1 236 €
2013	1 242 €
2014	1 276 €
2015	1 308 €

TABLE 3.2 – Évolution du montant du forfait IDA

Dans le cas de la convention IRSA, si le montant des dommages est inférieur au plafond de 6 500 €, fixé par la convention, le recours est forfaitaire dans la limite du forfait IDA (indemnisation directe de l'assuré). Les montants de ce forfait sont donnés dans la table 3.2. Le recours exercé est proportionnel au niveau de responsabilité de l'auteur des dommages. Si

le montant des dommages est supérieur à 6 500 €, le recours est réel. Ainsi pour un sinistre non responsable dans le cadre de la convention IRSA d'un montant de 2 000 € en 2015. L'assuré sera indemnisé totalement par son assureur à qui il restera 692 € à charge après recours. Dans le même cas mais pour un montant de 1 000 €, l'assureur gagnera 308 €. Nous aurons alors dans la base un sinistre de charge négative égale à -308 €.

Nous n'avons pas de montants pour les forfaits des sinistres de la convention IRCA. Nous identifions ces sinistres à partir des pics dans la distribution. Nous comptons le nombre de fois qu'apparaît chaque montant et nous considérons qu'au delà de 10 fois, celui-ci doit être un forfait. Nous remarquons que nous retrouvons les montants du forfait IDA parmi eux.

Montant	759	1050	1236	1242	1276	1308	1397	1490	1518	2726	2760	2794	2826	6000
Occurence	19	19	36	46	38	40	12	12	93	44	74	109	100	133

TABLE 3.3 – Montants possibles des forfaits pour la convention IRCA

Le montant de 6 000 € qui apparaît le plus souvent correspond au montant du forfait RC corporel léger. Nous n'avons pas identifié à quoi correspondent les autres montants.

3.6.2 Répartition des sinistres

Le montant des forfaits dans le cas des conventions IRSA et IRCA entraîne des pics dans la distribution du coût des sinistres. Pour pouvoir modéliser cette distribution, nous devrons traiter séparément ces sinistres afin d'obtenir une distribution plus lisse pour laquelle il sera plus facile d'ajuster une loi.

Pour la modélisation des sinistres RC, nous distinguons les sinistres RC matériel (RCM) et RC corporels (RCC). Les sinistres RCM sont plus fréquents que les sinistres RCC mais le coût moyen d'un sinistre RCC est beaucoup plus important que le coût moyen d'un sinistre RCM.

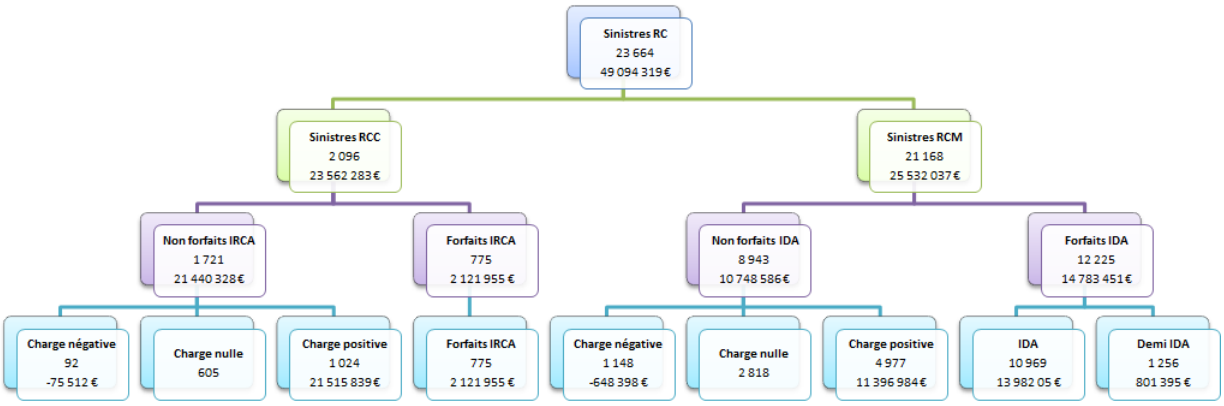


FIGURE 3.5 – Répartition des sinistres RC

Nous décrivons l'arbre de la figure 3.5 de bas en haut et de gauche à droite. Les sinistres RCC ne représentent que 9% des sinistres RC mais près de 50% de la charge sinistre. Ainsi le coût moyen pour un sinistre RCC est de 11 241 € alors qu'il n'est que de 1 206 € pour un sinistre RCM.

Pour les sinistres RCC, nous distinguons les sinistres relevant de la convention IRCA des autres. Ces sinistres représentent 37% des sinistres pour 9% de la charge totale. Étant donné le petit nombre de sinistres et le nombre relativement important de montants, il sera difficile d'en faire une modélisation.

Pour les sinistres RCM, nous distinguons les sinistres dont le montant correspond au montant du forfait IDA (taux de responsabilité 100%) ou au forfait demi-IDA (taux de responsabilité 50%). Les sinistres IDA représentent 58% du nombre pour 58% de la charge des sinistres RCM.

Pour les sinistres ne relevant pas de la convention IRSA, nous distinguons les sinistres suivant le signe de la charge. Les sinistres de charge négative sont peu nombreux et représentent un montant négligeable par rapport à la charge sinistre de cette catégorie. Les sinistres sans suite représentent 35% des sinistres. Nous ne considérerons dans la modélisation que les 1024 sinistres de charge positive qui ont un coût moyen de 21 011 € .

Pour les sinistres RCM ne relevant pas de la convention IRSA, la charge négative n'est pas négligeable et ces sinistres devront être pris en compte pour une modélisation. Les sinistres sans-suite seront mis de côté et nous modéliserons les sinistres de montant positif. Pour les sinistres relevant du forfait IDA, le montant est connu et nous prendrons le montant de la dernière année : 1 308 € pour le forfait IDA et 654 € pour le forfait demi-IDA. Pour ces sinistres, seule la fréquence est à modéliser.

3.6.3 Étude des sinistres responsables

Dans cette section, nous ne considérons que les sinistres responsables RC. Si un même sinistre est à la fois un sinistre corporel et matériel, nous compterons alors deux sinistres puisque nous avons deux montants différents qui seront réglés. Le CPT étant basé sur la fréquence de la zone et de l'activité, nous nous intéressons à l'évolution de ces fréquences au cours du temps.

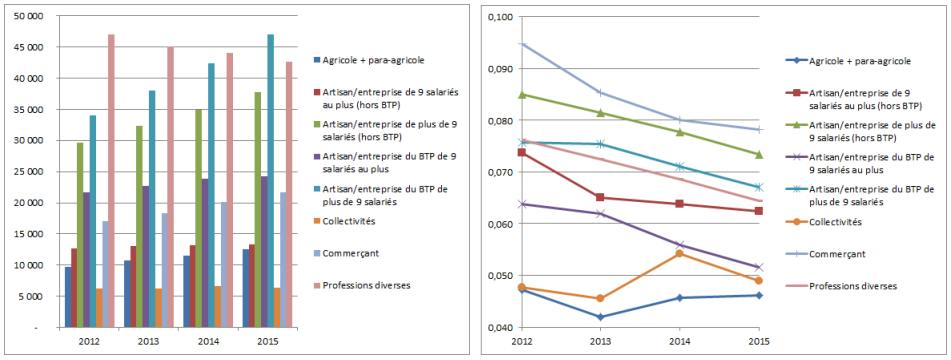


FIGURE 3.6 – Répartition des moteurs-année (MA) et fréquences RC responsable par année et par activité

Les graphiques de la figure 3.6 mettent en évidence deux points :

- le nombre croissant du nombre de véhicules moteurs en portefeuille au cours du temps à part pour les professions diverses

- la tendance à la baisse de la fréquence au cours de ces quatre dernières années pour les activités. Les deux exceptions sont les activités "collectivité" et "agricole + para-agricole". Elles représentent un petit volume, ce qui peut expliquer le comportement moins régulier.

Ces résultats montrent qu'il faut actualiser régulièrement les valeurs du coefficient d'activité prises en compte dans le calcul du CPT en prenant, par exemple, la fréquence moyenne sur les trois dernières années.

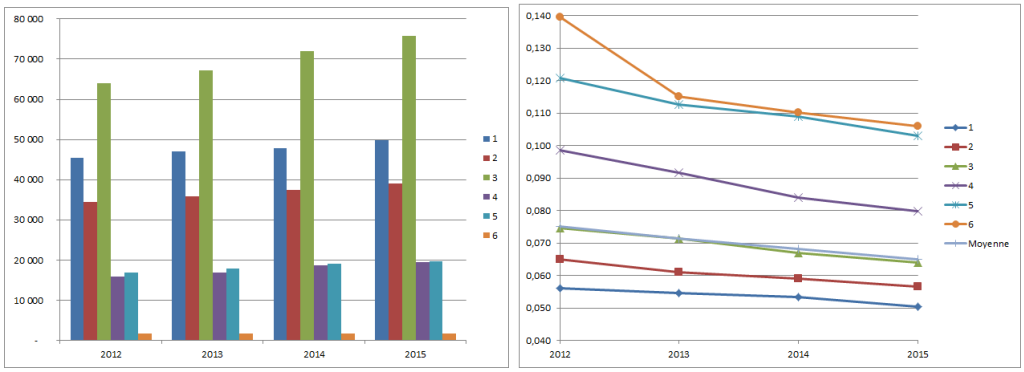


FIGURE 3.7 – Répartition des MA et fréquences RC responsable par année et par zone

Nous observons sur la figure 3.7 le développement du nombre de moteurs à la hausse et la même tendance à la baisse des fréquences par zone ces quatre dernières années. Les fréquences entre les zones 1 à 5 sont bien distinctes. Par contre, la différence entre les zones 5 et 6 est très faible. Il pourrait être envisagé de fusionner ces deux zones. La zone 6 étant peu représentée, cela permettrait d'avoir des résultats plus robustes sur les modélisations.

Chapitre 4

Crédibilité et échelle bonus-malus

4.1 Présentation de la théorie de la crédibilité

La théorie de la crédibilité [BG05, DMPW07, DC15, VG10] permet d'ajuster le niveau de prime d'un assuré suivant son historique de sinistralité. Dans notre cas, celui-ci est donné par le relevé d'informations fourni par la compagnie d'assurance quittée. Nous pouvons nous baser sur le coût moyen ou sur la fréquence. Nous nous baserons dans un premier temps sur la crédibilisation de la fréquence, le coût moyen n'étant pour l'instant pas renseigné par les souscripteurs. Nous pourrons étudier ultérieurement la possibilité d'utiliser le coût moyen par rapport à la charge de travail supplémentaire que cela représente pour les souscripteurs. Il est plus facile et plus rapide de compter le nombre de sinistres que de calculer leur coût moyen.

Dans le relevé d'informations, nous disposons du nombre de sinistres et de leur montant sur les trois dernières années. Nous ne tiendrons compte que du nombre. Sur le relevé d'information, nous ne disposons pas de l'exposition au risque de l'assuré, c'est à dire du nombre de véhicule-année sur les trois dernières années. Le nombre de moteurs pris en compte actuellement est purement déclaratif et constitue une première source d'erreur sur l'estimation de la fréquence de la flotte.

Nous disposons du nombre de sinistres RC responsables et du nombre de véhicules-année sur les années N, $N-1$ et $N-2$. Les nombres de sinistres sont notés n_N, n_{N-1} et n_{N-2} et le nombre de véhicules-année VA_N, VA_{N-1} et VA_{N-2}. Nous pouvons alors estimer la fréquence moyenne de l'assuré sur les trois dernières années par

$$\bar{f} = \frac{n_N + n_{N-1} + n_{N-2}}{VA_N + VA_{N-1} + VA_{N-2}} \tag{4.1}$$

L'assureur pourrait se baser uniquement sur cette fréquence de l'assuré pour calculer sa prime mais cela irait contre le principe de mutualisation. De plus, l'assuré n'ayant pas eu de sinistres sur les trois dernières années aurait une prime nulle.

L'assureur va alors se baser sur la fréquence de l'assuré mais aussi sur une fréquence collective f_{coll} d'un sous-groupe de son portefeuille ayant un risque homogène. La fréquence attribuée à l'assuré aura une valeur qui sera comprise entre la fréquence de l'assuré et la fréquence collective :

$$f = z\bar{f} + (1-z)f_{coll} \tag{4.2}$$

où z est le coefficient de crédibilité. Il est compris entre 0 et 1. Il est d'autant plus proche de 1 que nous accordons de crédit à l'information donnée par la fréquence propre de l'assuré. Si $z = 0$, l'estimation de la fréquence est basée uniquement sur la fréquence collective alors que si $z = 1$ la fréquence est basée uniquement sur l'expérience de la flotte. Nous parlons de crédibilité totale.

Il existe plusieurs méthodes pour estimer la valeur du facteur de crédibilité. Nous utiliserons l'approche de Bühlmann-Straub qui est la plus utilisée dans le monde de l'assurance.

4.2 Le modèle de Bühlmann-Straub

Le modèle de Bühlmann-Straub est une généralisation du modèle de Bühlmann. Il permet d'affaiblir l'hypothèse $\sigma^2(\Theta_i) = \mathrm{Var}[X_{ij}|\Theta_i]$ en introduisant des poids qui mesurent l'exposition.

4.2.1 Théorie de Bühlmann-Straub

Nous traitons un portefeuille de I risques (ou classes de risques) et nous noterons X_ij le ratio de sinistres (ou la fréquence de sinistres ou la taille moyenne des sinistres ou tout autre indicateur) du risque i pendant l'année j et w_{ij} le poids associé.

Pour chaque risque i, nous cherchons l'estimateur de crédibilité $\widehat{\widehat{\mu(\Theta_i)}}$ de son ratio de sinistres individuels $\mu(\Theta_i)$.

Théorème 4.2.1 *L'estimateur homogène de crédibilité du modèle de Bühlmann-Straub est donné par*

$$\widehat{\widehat{\mu(\Theta_i)}}^{hom} = \alpha_i X_i + (1-\alpha_i)\,\widehat{\widehat{\mu_0}}\,, \tag{4.3}$$

où

$$X_i = \sum_j \frac{w_{ij}}{w_{i\bullet}} X_{ij},$$

$$\widehat{\widehat{\mu_0}} = \sum_i \frac{\alpha_i}{\alpha_\bullet} X_i,$$

$$w_{i\bullet} = \sum_j w_{ij},$$

$$\alpha_i = \frac{w_{i\bullet}}{w_{i\bullet} + \frac{\sigma^2}{\tau^2}}$$

Théorème 4.2.2 *Sous les hypothèses du modèle de Bühlmann-Straub, nous avons l'erreur quadratique moyenne de l'estimateur homogène de crédibilité :*

$$\mathbb{E}\left[\left(\widehat{\widehat{\mu(\Theta_i)}}^{hom} - \mu(\Theta_i)\right)^2\right] = \tau^2(1-\alpha_i)\left(1 + \frac{1-\alpha_i}{\alpha_\bullet}\right) \tag{4.4}$$

La formule de l'estimateur de crédibilité nécessite la connaissance des trois paramètres μ_0, σ^2 et τ^2. Un estimateur de μ_0 est déjà donné dans le théorème 4.2.1. Il reste à déterminer des estimateurs pour σ^2 et τ^2.

Théorème 4.2.3 *Les estimateurs*

$$\widehat{\sigma^2} = \frac{1}{I}\sum_{i=1}^{I}\frac{1}{n-1}\sum_{j=1}^{n} w_{ij}(X_{ij} - X_i)^2 \tag{4.5}$$

$$\widehat{\widehat{\tau^2}} = \frac{w_{\bullet\bullet}}{w_{\bullet\bullet} - \sum_{i=1}^{I} w_{i\bullet}^2}\left(\sum_{i=1}^{I} w_{i\bullet}(X_i - \bar{X})^2 - (I-1)\,\widehat{\sigma^2}\right) \tag{4.6}$$

sont sans biais et convergents (avec la condition $\sum_i \frac{w_{i\bullet}}{w_{\bullet\bullet}} \to 0$ *quand* $I \to 0$ *pour* $\widehat{\widehat{\tau^2}}$*).*

Remarque 1 *L'estimateur* $\widehat{\widehat{\tau^2}}$ *peut prendre des valeurs négatives. Cela signifie que nous ne pouvons pas faire de différences entre les risques. Dans ce cas, nous posons alors* $\widehat{\tau^2} = 0$ *et l'estimateur homogène de crédidibilité est alors* $\widehat{\widehat{\mu_0}}$*. Finalement, nous utilisons l'estimateur*

$$\widehat{\tau^2} = \max\left(\widehat{\widehat{\tau^2}}, 0\right) \tag{4.7}$$

Cette estimateur n'est plus sans biais.

4.2.2 Exemple

Pour se familiariser avec ce modèle, prenons l'exemple de deux flottes, A et B pour lesquelles nous souhaitons connaître la fréquence crédibilisée. Nous connaissons la sinistralité et le nombre de moteurs sur les trois années précédentes par le relevé d'informations.

Flotte		N	N+1	N+2
A	Nombre de sinistres	4	3	5
	Nombre de moteurs	30	29	32
	fréquence	0,13	0,1	0,15
B	Nombre de sinistres	4	1	2
	Nombre de moteurs	40	38	36
	fréquence	0,1	0,026	0,056

TABLE 4.1 – Observations sur trois année de deux contrats flotte

Les poids des contrats A et B, ω_A et ω_B, sont donnés par

$$\omega_A = 91 \quad \text{et} \quad \omega_B = 114$$

Nous obtenons pour les fréquences moyennes sur les trois années pour les deux contrats A et B, f_A et f_B, et la fréquence globale f :

$$f_A = 12/91, \quad f_B = 7/114 \quad \text{et} \quad f = 19/205$$

L'estimateur sans biais de la variance pour chaque contrat nous donne

$$\hat{\sigma}_A^2 = 0,021 \quad \text{et} \quad \hat{\sigma}_B^2 = 0,054$$

Les estimateurs de $\hat{\sigma^2}$ et $\hat{\hat{\tau^2}}$ sont donnés par

$$\widehat{\sigma^2} = \hat{\sigma}_A^2 + \hat{\sigma}_B^2 = 0,0375$$

$$\widehat{\widehat{\tau^2}} = \frac{\omega}{\omega^2 - (\omega_A^2 + \omega_B^2)}(\omega_A(f_A - f)^2 + \omega_B(f_B - f)^2 - \widehat{\sigma^2}) = 0,0021$$

4.3 Application du modèle de Bühlmann-Straub

4.3.1 Étude de la fréquence des sinistres RC responsables

Dans cette section, nous étudions la fréquence des sinistres RC responsables par contrat. Nous souhaitons vérifier que les variables zone et activité sont bien des variables explicatives de la fréquence. Nous avons additionné les sinistres responsables RCM et RCC par contrat et par année. La modélisation ne tenant pas compte des caractéristiques véhicules, nous avons récupéré les sinistres non rattachés et nous les avons joints à l'aide du numéro de contrat et de l'année de survenance du sinistre. Dans la base, nous obtenons 315 contrats (pour 799 situations) qui n'ont que des véhicules non-moteurs.

Dans un premier temps, nous étudions les statistiques univariées pour chaque variable. Les variables choisies correspondent aux variables communes à tous les véhicules du contrat :

la zone et l'activité. La variable présence de multi-garantie RC et DTA peut varier d'un véhicule à l'autre et ne peut donc pas être prise en compte. Nous prenons également l'année sinon il faudrait retravailler les données pour avoir une fréquence constante sur les quatre années.

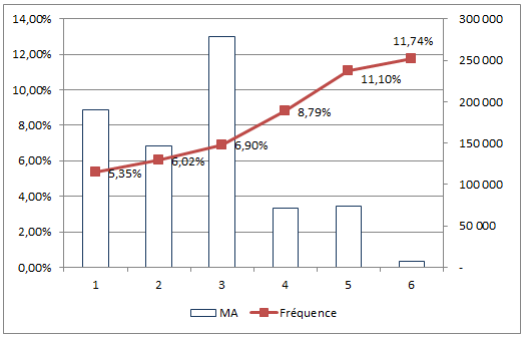


FIGURE 4.1 – Fréquences RC responsable et MA par zone

La figure 4.1 donne la répartition du nombre de moteurs-année (MA) et la fréquence RC responsable par zone. Nous retrouvons que les zones 1 à 3 sont les plus représentées et totalisent près de 80% des véhicules et 36% uniquement pour la zone 3. La zone 6 a très peu de véhicules. Nous observons une fréquence croissante avec la zone ce qui est cohérent avec la construction du zonier, la zone 1 étant la moins risquée et la zone 6 la plus risquée. La différence entre chaque zone est nette avec une augmentation de plus 10% entre chaque zone adjacente.

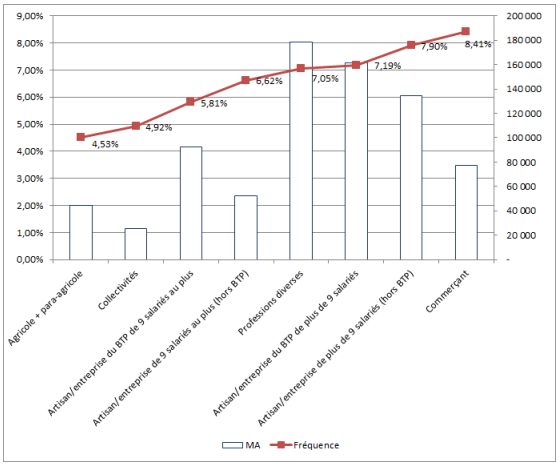


FIGURE 4.2 – Fréquences RC responsable et MA par activité

La figure 4.2 donne la répartition du nombre de moteurs-année et la fréquence RC responsable par activité. Les activités ont été triées par ordre de fréquence. Nous observons des fréquences bien différentes des autres pour chaque activité hormis les activités "professions diverses" et "Artisan/entreprise de 9 salariés au plus (hors BTP)". Nous observons une différence pour les activités où nous distinguons si l'entreprise a plus ou moins de 9 salariés. Cela justifie le fait de distinguer les deux cas même si les entreprises de 9 salariés au plus ont une exposition moindre.

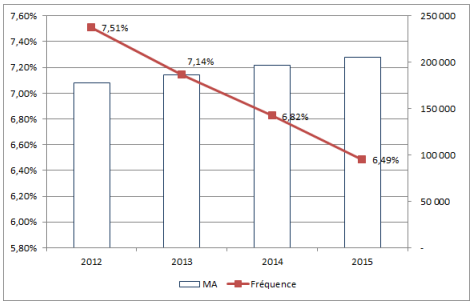


FIGURE 4.3 – Fréquences RC responsable et MA par année

La figure 4.3 donne la répartition du nombre de moteurs-année et la fréquence RC responsable par année. Nous retrouvons la tendance à la baisse de la fréquence des sinitres RC au cours de ces quatre dernières années. Nous avons une diminution de près de 5% chaque année. Nous observons une augmentation du nombre de moteurs en portefeuille. Chaque année, ce nombre augmente de près de 5%. Le nombre de sinistres est stable dans le temps (autour de 13 300) et c'est l'augmentation du nombre de moteurs en portefeuille qui conduit à la baisse de la fréquence. Cela montre l'importance de mettre l'année comme variable explicative.

Pour confirmer les tendances données par les statistiques univariées, nous modélisons le nombre de sinistres par la zone, l'activité et l'année en mettant le logarithme du nombre de moteurs-année en variable offset à l'aide des modèles linéaires généralisés. Commençons par calculer l'espérance et la variance de la distribution du nombre de sinistres par contrat et par année. Soit N ce nombre de sinistres. Nous obtenons les valeurs suivantes :

$$\mathbb{E}[N] = 0,62 \quad \text{et} \quad \text{Var}(N) = 1,33 \tag{4.8}$$

La variance est beaucoup plus grande que l'espérance. Étant donnée cette sur-dispersion, la loi de Poisson ne sera pas en adéquation avec cette distribution. Une loi binomiale négative sera plus appropriée. En utilisant la commande *goodfit* dans R, nous obtenons le graphique de la figure 4.4.

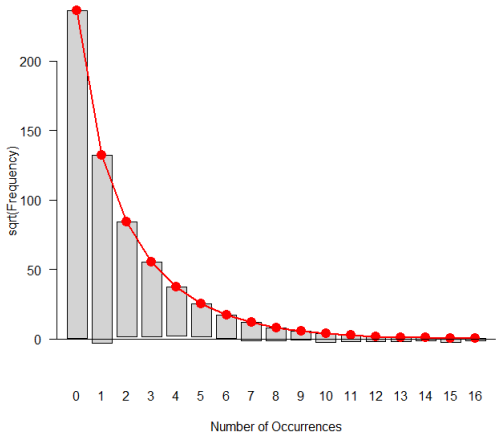


FIGURE 4.4 – Adéquation de la distribution du nombre de sinistres par contrat et par année avec une loi binomiale négative

La distribution ne satisfait pas le test d'adéquation du χ^2 avec une p-value petite. Cependant, l'échantillon étant de grande taille et cette loi étant la plus proche de notre distribution, nous utilisons cette loi pour réaliser un modèle linéaire généralisé sous SAS avec la procédure *GENMOD*.

Un modèle linéaire généralisé est une extension du modèle linéaire. Il permet d'étudier la dépendance entre une variable à expliquer Y et un ensemble de variables explicatives $X_1, \ldots, X_n$. La variable Y est une variable aléatoire qui suit une loi de probabilité appartenant à la famille exponentielle. Une fonction lien g permet d'établir la relation de dépendance de $\mathbb{E}[g(Y)]$ en fonction d'une combinaison linéaire des variables X_i. Ce modèle étant paramétrique. Le résultat dépend donc du choix de la loi modélisant la distribution de Y et de la fonction lien. Le modèle peut se mettre sous la forme :

$$g\left(\mathbb{E}[g(Y)]\right) = \beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n \tag{4.9}$$

En général, la fonction lien est la fonction ln qui permet d'obtenir un modèle multiplicatif. Pour plus de détail sur la théorie des modèles linéaires généralisés, on pourra se référer à [DMPW07, DC15, DV84].

Nous expliquons le nombre de sinistres par les variables zone, activité et année.

Source	DDL	Khi-2	Pr > Khi-2
Zone	5	1770,47	<0,0001
Activité	7	617,29	<0,0001
Année	3	135,24	<0,0001

TABLE 4.2 – Statistiques de type 3

Les trois variables sont bien explicatives du nombre de sinistres. Nous obtenons de bons résultats pour toutes les modalités de chaque variable à part pour l'activité "Artisan/entreprise de 9 salariés au plus (hors BTP)" . Nous ne nous attardons pas sur ces résultats, l'objectif ici étant de vérifier que les variables zone et activité expliquent bien le nombre de sinistres. D'après le tableau 4.2, la variable la plus explicative est la zone, ensuite nous avons l'activité et en troisième l'année. Nous remarquons de nouveau la tendance à la baisse de la fréquence des sinistres de 2012 à 2015. Cette tendance pourrait être prise en compte en utilisant le modèle de Hachemeister.

La variable zone étant plus explicative que la variable activité, nous pourrons étudier le cas où nous ne tenons compte que de la variable zone pour regrouper les contrats.

4.3.2 Résultats par croisement zone et activité

Nous cherchons à calculer le facteur de crédibilité pour chaque croisement de zone par l'activité. Nous divisons donc le portefeuille en 48 groupes qui correspondent au croisement des 6 zones avec les 8 classes d'activité. Pour chaque groupe, nous appliquons la théorie de la crédibilité de Bühlmann-Straub à l'aide de la fonction *cm* du package *actuar* de R. Nous utilisons les fréquences des années 2013 à 2015 pour estimer la fréquence de 2016.

Les zones sont déjà numérotées de 1 à 6 et les activités sont numérotées par ordre alphabétique de 1 à 8. Nous obtenons ainsi 48 bases *BaseBSij* qui seront traitées une par une. Toutes les bases sont traitées en une seule fois à l'aide de deux boucles. Voici un exemple du code pour une base :

```
> table<-read.table("Y:/Alternance/6_Programme_R/baseBS33.csv",sep=";",header=TRUE)

> cm(~NU_AFFA, table, ratios=freq_2013:freq_2015, weights=MA_2013:MA_2015)

Call:
cm(formula = ~NU_AFFA, data = table, ratios = freq_2013:freq_2015,
    weights = MA_2013:MA_2015)
```

```
Structure Parameters Estimators

Collective premium: 0.06989901

Between NU_AFFA variance: 0.001834874
Within NU_AFFA variance: 0.07686865
```

Nous obtenons la fréquence collective pour la zone 3 et l'activité 3 :

$$f_{col} = 0,06989901$$

Les valeurs des estimateurs sont données par

$$\widehat{\sigma^2} = 0,07686865 \quad \text{et} \quad \widehat{\widehat{\tau^2}} = 0,001834874$$

Ce qui nous donne une valeur de K égale à

$$K = \frac{\widehat{\sigma^2}}{\widehat{\widehat{\tau^2}}} = 41,89$$

Avec cette valeur de K le coefficient de crédibilisation $z = \frac{n_{MA}}{n_{MA}+K}$ varie entre 0,26 pour 15 moteurs sur les 3 ans et 0,74 pour 117 moteurs sur les 3 ans.

La fréquence crédibilisée aura alors la forme

$$f_{cred} = \frac{n_{MA}}{n_{MA}+K} f_{flotte} + \frac{K}{n_{MA}+K} f_{col}$$

Si nous prenons par exemple un assuré qui arriverait avec la sinistralité suivante :

Année	2013	2014	2015
Nombre de MA	2	1	1
Nombre de sinistres	28	26	27

La fréquence de la flotte est alors égale à $f_{flotte} = 4/81 \simeq 0,049$, le facteur de crédibilisation $z = n_{MA}/(n_{MA}+K) = 111/(111+41,89) \simeq 0,726$ et la fréquence crédibilisée est alors

$$f_{cred} \simeq 0,726 \times 0,049 + (1-0,726) \times 0,07 \simeq 0,055.$$

Nous obtenons l'ensemble des valeurs des tables 4.3 et 4.4 pour l'ensemble des groupes

Zone	Nom Activité	freq col	K	MA	Nb contrats
1	Agricole + para-agricole	0,04062	38	11 646	528
1	Artisan/entreprise de 9 salariés au plus (hors BTP)	0,0527	57	11 968	631
1	Artisan/entreprise de plus de 9 salariés (hors BTP)	0,06089	5	25 691	932
1	Artisan/entreprise du BTP de 9 salariés au plus	0,04335	50	20 830	1 081
1	Artisan/entreprise du BTP de plus de 9 salariés	0,0504	65	32 975	1 062
1	Collectivités	0,04034	∞	4 745	163
1	Commerçant	0,06351	29	13 979	625
1	Professions diverses	0,04681	33	28 186	1 190
2	Agricole + para-agricole	0,03453	27	8 426	372
2	Artisan/entreprise de 9 salariés au plus (hors BTP)	0,05183	53	8 376	455
2	Artisan/entreprise de plus de 9 salariés (hors BTP)	0,06864	32	17 757	710
2	Artisan/entreprise du BTP de 9 salariés au plus	0,04586	38	14 931	742
2	Artisan/entreprise du BTP de plus de 9 salariés	0,05844	49	24 357	794
2	Collectivités	0,06168	20	5 133	169
2	Commerçant	0,06934	54	10 602	521
2	Professions diverses	0,05484	29	27 334	1 098
3	Agricole + para-agricole	0,04398	20	12 373	549
3	Artisan/entreprise de 9 salariés au plus (hors BTP)	0,06391	37	14 035	802
3	Artisan/entreprise de plus de 9 salariés (hors BTP)	0,0699	42	42 715	1 677
3	Artisan/entreprise du BTP de 9 salariés au plus	0,05493	42	26 611	1 380
3	Artisan/entreprise du BTP de plus de 9 salariés	0,06658	72	48 248	1 733
3	Collectivités	0,0478	58	7 022	220
3	Commerçant	0,0762	34	22 252	1 042
3	Professions diverses	0,06366	22	48 005	1 994

TABLE 4.3 – Fréquence collective, valeur de K, nombre de MA sur trois ans et nombre de contrats par groupe de zone et activité

Zone	Nom Activité	freq col	K	MA	Nb contrats
4	Agricole + para-agricole	0,05316	42	2 241	90
4	Artisan/entreprise de 9 salariés au plus (hors BTP)	0,07742	18	2 705	178
4	Artisan/entreprise de plus de 9 salariés (hors BTP)	0,08418	27	11 597	429
4	Artisan/entreprise du BTP de 9 salariés au plus	0,07369	39	5 656	312
4	Artisan/entreprise du BTP de plus de 9 salariés	0,09195	50	12 023	445
4	Collectivités	0,03668	40	1 404	51
4	Commerçant	0,08953	26	7 471	306
4	Professions diverses	0,07661	19	14 603	590
5	Agricole + para-agricole	0,08222	31	651	30
5	Artisan/entreprise de 9 salariés au plus (hors BTP)	0,10338	14	3 102	178
5	Artisan/entreprise de plus de 9 salariés (hors BTP)	0,11212	17	11 442	444
5	Artisan/entreprise du BTP de 9 salariés au plus	0,10915	26	4 233	244
5	Artisan/entreprise du BTP de plus de 9 salariés	0,11205	63	13 397	483
5	Collectivités	0,0535	13	934	30
5	Commerçant	0,10859	28	7 165	315
5	Professions diverses	0,09197	21	17 897	702
6	Agricole + para-agricole	0,06629	∞	136	7
6	Artisan/entreprise de 9 salariés au plus (hors BTP)	0,10963	43	682	40
6	Artisan/entreprise de plus de 9 salariés (hors BTP)	0,08835	32	739	38
6	Artisan/entreprise du BTP de 9 salariés au plus	0,0976	112	613	31
6	Artisan/entreprise du BTP de plus de 9 salariés	0,09314	196	1 386	49
6	Collectivités	0,0301	∞	100	7
6	Commerçant	0,14334	22	582	26
6	Professions diverses	0,10038	12	1 146	54

TABLE 4.4 – Fréquence collective, valeur de K, nombre de MA sur trois ans et nombre de contrats par groupe de zone et activité (suite)

Pour trois groupes (collectivités en zone 1, agricole + para-agricole en zone 6, collectivités en zone 6), l'estimateur τ^2 est négatif. Pour les activités de la zone 6, nous n'avons que 7 contrats et une centaine de moteurs. Dans ce cas, nous prenons 0 pour la valeur de $\widehat{\widehat{\tau^2}}$. K prend alors une valeur infinie et nous prenons pour la fréquence crédibilisée la valeur de la fréquence collective. Nous trouvons une valeur très petite de $K = 5$ pour les Artisan/entreprise de plus de 9 salariés (hors BTP) de la zone 1. Pour cette valeur de K, nous accordons une grande crédibilité à la fréquence de la flotte. A l'opposé, nous obtenons deux grandes valeurs de K pour les "Artisan/entreprise du BTP de 9 salariés au plus" et les "Artisan/entreprise du BTP de plus de 9 salariés" de la zone 6. Dans ce cas, le facteur de crédibilité est petit et nous prenons davantage en considération la valeur de la fréquence du groupe.

Les autres valeurs sont comprises entre 12 et 72. Nous ne voyons pas de tendance de

la valeur de K que ce soit par zone ou activité. Il faudrait donc avoir un coefficient de crédibilisation par croisement de zone et d'activité.

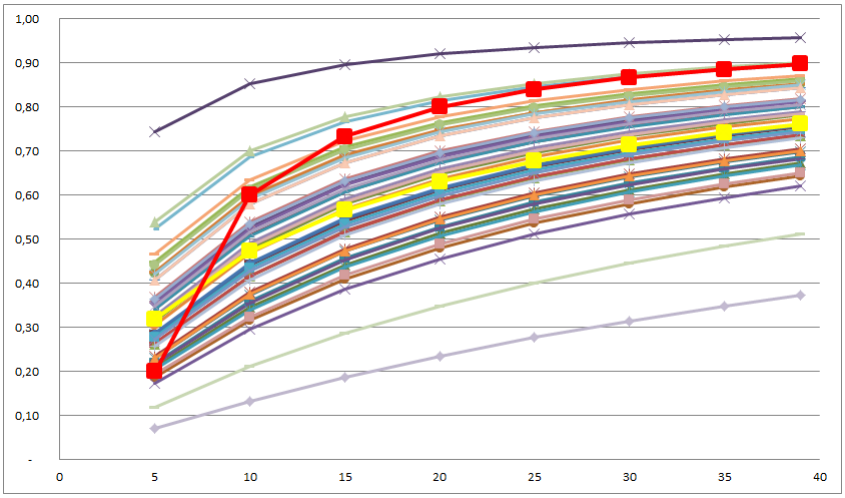


FIGURE 4.5 – Représentation de l'évolution des 45 coefficients de crédibilisation en fonction du nombre de moteurs sur une année. En rouge le coefficient de crédibilisation actuel et en jaune le coefficient moyen

Sur le graphique 4.5, nous distinguons les courbes des trois valeurs atypiques de K. Les autres courbes sont plus rapprochées. En comparant l'allure du facteur de crédibilité dans le CPT avec les différents coefficients que nous venons de calculer, nous observons une plus petite amplitude entre la valeur pour 5 moteurs et la valeur pour 39 moteurs. La croissance est donc moins rapide et cela crée moins d'écart entre les petites et les grosses flottes.

Un des problèmes est la non stabilité de la valeur de K dans le temps. Nous ne disposons que de 4 ans sur la base. Nous avons comparé les résultats obtenus en utilisant sur les années 2013-2015 avec ceux de 2012-2014. Mis à part pour la zone 3, nous obtenons des valeurs de K qui peuvent être très différentes surtout sur les groupes où il n'y a pas beaucoup de contrats et de moteurs.

Pour palier ce manque de stabilité, nous avons choisi de ne prendre qu'une seule valeur de K. Cette valeur correspond au K qui permet d'être proche de la courbe moyenne en enlevant les trois valeurs atypiques. La valeur retenu est $K = 33$. Il faut ensuite recalculer l'estimation de la fréquence collective pour les 48 groupes. Les écarts avec la fréquence calculée avec des valeurs de K différentes est faible. Cinq groupes seulement ont un écart de plus ou moins de 5%. Les valeurs obtenues sont données dans le tableau 4.5.

	1	2	3	4	5	6
Agricole + para-agricole	0,041	0,036	0,045	0,055	0,082	0,058
Artisan/entreprise de 9 salariés au plus (hors BTP)	0,054	0,054	0,066	0,080	0,108	0,116
Artisan/entreprise de plus de 9 salariés (hors BTP)	0,061	0,070	0,072	0,085	0,115	0,092
Artisan/entreprise du BTP de 9 salariés au plus	0,044	0,048	0,057	0,078	0,114	0,104
Artisan/entreprise du BTP de plus de 9 salariés	0,051	0,060	0,068	0,093	0,116	0,096
Collectivités	0,041	0,062	0,048	0,038	0,055	0,031
Commerçant	0,066	0,071	0,079	0,092	0,112	0,152
Professions diverses	0,047	0,056	0,066	0,080	0,094	0,111

TABLE 4.5 – Estimation des fréquences collectives par zone et activité pour un facteur de crédibilisation commun égal à n/(n+K)

Nous ne pouvons pas utiliser ces données brutes. En effet, nous observons dans le tableau 4.5 que les valeurs de la zone 6 ne sont pas cohérentes dans l'ensemble : pour une activité donnée, la fréquence de la zone 6 est inférieure à celle de la zone 5. Pour corriger cela, nous avons deux possibilités :

- corriger à la main les fréquences de la zone 6 pour qu'elles soient supérieures à celles de la zone 5
- regrouper la zone 6 avec la zone 5, la distinction entre ces deux zones n'étant pas nette et le nombre de contrats pour ces zones étant petit

Pour deux activités, "Agricole + para-agricole" et "collectivités", la fréquence n'est pas croissante avec la zone. Ces deux activités représentent de faibles volumes, ce qui peut expliquer le comportement de la fréquence. Nous avions déjà observé ce comportement sur le graphique 4.2. Dans ces deux cas, il est difficile de faire un regroupement cohérent avec une autre activité pour avoir davantage de volume. Ces fréquences doivent être corrigées à la main. Nous pourrons nous baser sur les fréquences de la zone 3 où il y a le plus de contrats et où l'estimation est la plus fiable. Dans le cas de l'activité "Agricole + para-agricole", seule la fréquence de la zone 2 est à modifier. Nous avons un peu de volume sur les zones 1 et 3. Nous fixons une valeur entre ces deux valeurs pour la zone 2, 0,043 par exemple. Pour l'activité "collectivités", nous conservons la fréquence de la zone 1 et 3 et nous extrapolons

pour les autres zones (voir tableau 4.6).

Zone	1	2	3	4	5
Fréquence	0,041	0,044	0,048	0,060	0,088

TABLE 4.6 – Fréquences corrigées pour l'activité "collectivités"

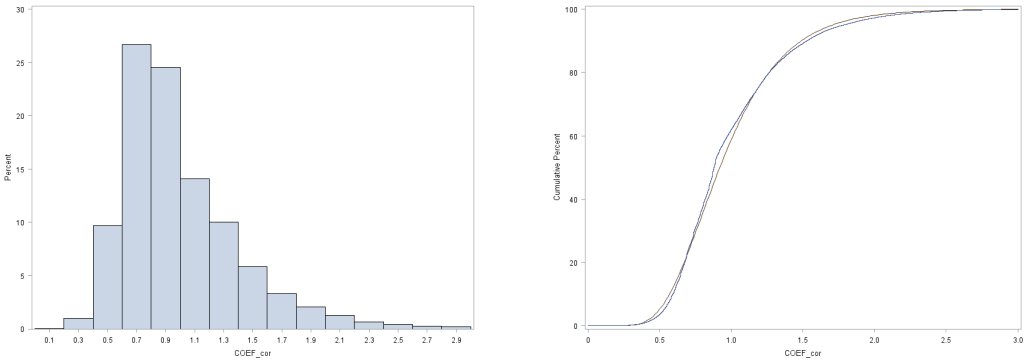
Finalement nous choisissons d'affecter la même fréquence à la zone 6 qu'à la zone 5 pour une activité donnée, comme c'est le cas actuellement dans la fréquence zone du CPT.

Nous pouvons maintenant calculer la fréquence crédibilisée pour chaque contrat à partir de la fréquence sur les trois dernières années et de la fréquence de la zone et de l'activité. En divisant cette fréquence par la fréquence zone nous obtenons un indicateur qui nous informe si la fréquence crédibilisée de la flotte est inférieure ou supérieure à la fréquence de la zone et de l'activité.

$$freq\_cred = \frac{n_{MA}}{n_{MA}+33} freq\_flotte + \left(1 - \frac{n_{MA}}{n_{MA}+K}\right) freq\_za \tag{4.10}$$

$$IND = \frac{n_{MA}}{n_{MA}+33} \frac{freq\_flotte}{freq\_za} + \left(1 - \frac{n_{MA}}{n_{MA}+33}\right) \tag{4.11}$$

Nous obtenons les distributions de la figure 4.6(a) et 4.6(b). La moyenne des coefficients est égale à 1. Ceci peut laisser à penser que si nous appliquons ce coefficient à un tarif modélisé, nous devrions retrouver au total le même montant de prime que sans le coefficient.



(a) (b)

FIGURE 4.6 – Distribution et fonction de répartition de la variable IND

Nous avons écrêté les valeurs de l'indicateur IND à 3 pour plus de lisibilité des graphiques. Nous avons donc enlevé 95 contrats sur les 25 549 étudiés, soit moins de $0,4\%$. Nous observons une distribution dissymétrique avec une queue de distribution assez longue. Elle est proche d'une loi lognormale.

Pour la suite, nous voyons deux approches possibles pour calculer un nouveau tarif en affaire nouvelle :

- utiliser la fréquence crédibilisée brute. Elle sera juste bornée pour ne pas sous ou sur tarifer. Ces bornes seront à ajuster de façon à garder une moyenne constante sur l'ensemble des fréquences. Elle sera utilisée en variable offset dans la modélisation classique fréquence×coût moyen. Les variables zone et activité ne seront pas remises dans le modèle de fréquence puisqu'elles ont déjà été utilisées pour calculer la fréquence crédibilisée.

- utiliser l'indicateur IND pour classer les contrats par groupe. Chaque groupe se verra attribuer un coefficient minorateur ou majorateur. Par exemple, si la valeur de IND est proche de 1, le groupe aura un coefficient égal à 1. Ceci nous donnera une valeur d'entrée sur une échelle bonus-malus et il faudra alors déterminer les règles pour changer de "barreau". Le coefficient sera indépendant de la tarification à priori fréquence×coût moyen.

Nous avons déjà vu qu'il pouvait y avoir des imprécisions sur le calcul de la fréquence utilisée pour le calcul du CPT à cause du nombre de moteurs-année qui est déclaratif. Il n'est donc pas intéressant d'avoir un coefficient continu étant donné l'imprécision que nous pouvons avoir. Nous avons donc choisi de faire des classes et nous abordons donc uniquement la deuxième méthode.

4.3.3 Élaboration d'une échelle

N'ayant pas d'idée à priori sur les seuils à utiliser et sur le nombre de classes à faire, nous avons commencé par faire une classification ascendante hiérarchique [GP11]. Nous cherchons à ce que les contrats regroupés au sein d'une même classe (homogénéité intra-classe) soient le plus semblables possibles tandis que les classes soient le plus dissemblables (hétérogénéité inter-classe). Nous utilisons la méthode de Ward qui repose sur un gain minimum d'inertie intra-classe à chaque agrégation.

Nous avons utilisé la macro cah de Philippe Besse sous SAS. Le détail du code est donné dans l'annexe B.

Pour choisir le nombre optimal de classes, nous utilisons la courbe des pseudos T^2 et la courbe des R^2 semi-partiels. La courbe des R^2 nous indique la proportion de la variance expliquée par les classes.

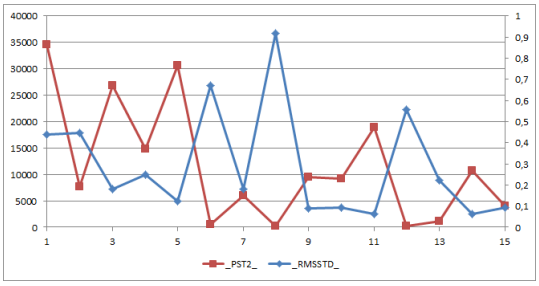


FIGURE 4.7 – Représentation graphique des courbes des pseudos T^2 et des R^2 semi-partiels

Nous devons identifier les points où coïncident un minimum pour la courbe des pseudos T^2 et un maximum pour la courbe des R^2 semi-partiels. Les points du graphique qui vérifient ces conditions sont pour 6, 8 et 12 classes. Les valeurs du R^2 correspondantes sont 0,907, 0,946 et 0,973. A six classes, la proportion de variance expliquée est donc déjà grande.

Quand nous augmentons le nombre de classes, nous observons (voir annexe C) que les classes divisées sont celles où il y a le moins de contrats et où la valeur de IND est élevée. Nous choisissons alors d'enlever les contrats dont la valeur du IND est supérieur à 2. Cela correspond à moins de 5% des contrats. Nous recommençons la recherche du nombre optimal de classes. Le meilleur nombre de classes est 8 pour lequel nous avons une valeur du R^2 de 97%.

Cluster	Nombre contrats	Moyenne	Ecart-type	Minimum	Maximum
5	2 432	0,51	0,08	0,05	0,60
3	3 229	0,65	0,03	0,60	0,70
1	4 438	0,77	0,04	0,70	0,84
2	3 183	0,88	0,02	0,84	0,93
8	3 864	1,04	0,06	0,93	1,15
6	2 645	1,24	0,06	1,15	1,35
4	1 704	1,45	0,06	1,35	1,57
7	1 412	1,75	0,12	1,57	2,00

TABLE 4.7 – Huit classes pour IND borné à 2

Nous obtenons les huit classes données dans le tableau 4.7. Nous observons que le nombre de contrats est assez bien réparti dans toutes les classes et qu'il n'y a pas de classe sous représentée (la plus petite classe représente plus de 5% du nombre de contrats). La première proposition d'échelle est donnée dans le tableau 4.8. Pour cette échelle, nous avons gardé

pour les valeurs prises par la variable IND les résultats données par la CAH. Pour chaque intervalle, nous attribuons pour cette classe la valeur maximal du coefficient (nous aurions aussi pu prendre la valeur moyenne).

Niveau	Valeur COEF	Coefficient BM	augmentation
1	<0,6	0,6	
2	[0,6 ; 0,7[	0,7	17%
3	[0,7 ; 0,84[	0,84	20%
4	[0,85 ; 0,93[	0,93	11%
5	[0,93 ; 1[	1	8%
6	[1 ;1,15[	1,15	15%
7	[1,15 ; 1,35[	1,35	15%
8	[1,35 ; 1,57[	1,57	16%
9	$\geq$ 1,57	2	27 %

TABLE 4.8 – Première proposition d'échelle BM à 9 niveaux

La colonne de droite intitulée "augmentation" du tableau 4.8 donne le pourcentage d'augmentation lorsqu'un contrat passe du niveau $n-1$ au niveau n. Ce pourcentage est en général supérieur à 10%. Ceci peut poser des problèmes d'un point de vu commercial : il est difficile de faire passer de telles hausses auprès des clients et des agents. Pour remédier à ce problème, nous pouvons agir dans deux directions :

– augmenter le nombre de niveaux de l'échelle

– réduire l'amplitude de l'échelle

Pour réduire l'amplitude de l'échelle, nous pouvons utiliser la transformation t qui permet de passer du segment $[a, b]$ au segment $[c, d]$:

$$t(x) = \frac{x-a}{b-a} \times (d-c) + c \tag{4.12}$$

Nous estimons que l'échelle proposée dans le tableau 4.8 devrait être réduite. Nous choisissons, par exemple, de réduire l'intervalle à $[0,8; 1,3]$.

Niveau	Valeur COEF	Nouveau coefficient	Augmentation
1	$\leq 0,6$	0,80	
2	]0,6 ; 0,7]	0,84	6%
3	]0,7 ; 0,84]	0,89	8%
4	] 0,85 ; 0,93]	0,92	5%
5	] 0,93 ; 1]	1,01	4%
6	]1 ;1,15]	1,08	8%
7	]1,15 ; 1,35]	1,19	10%
8	$>$ 1,57	1,30	10%

TABLE 4.9 – Deuxième proposition d'échelle BM à 8 niveaux

Le changement d'échelle permet bien de réduire la différence entre deux niveaux consécutifs. Nous observons une différence plus petite entre les premiers niveaux de l'échelle, niveaux 1 à 5, qu'entre les niveaux les plus élevés, niveaux 6 à 9. Ceci peut permettre d'avoir un tarif plus compétitif et une plus grande segmentation sur les bons risques. Ce changement permet de conserver un coefficient proche de 1 pour le niveau 5, comme c'était le cas dans l'échelle précédente.

4.3.4 Autre proposition d'échelle

Pour l'échelle précédente, nous avions utilisé un coefficient de crédibilisation $n/(n + K)$ avec K=33. Cela correspondait au coefficient moyen obtenu pour les différents couples zone/activité. Avec cette valeur de K, le coefficient de crédibilisation varie entre 0,31 pour une flotte de 5 moteurs et 0,78 pour une flotte de 39 moteurs. La valeur pour 39 moteurs laisse peut de marge pour le segment 3 (39 à 99 moteurs) si nous voulons avoir un tarif cohérent.

Actuellement, il existe une autre formule du CPT pour le segment 3. Ce coefficient a été ajusté pour avoir une cohérence dans le tarif. Si nous calculons la cotisation pour 40 moteurs avec le tarif su segment 2 et 3, nous ne devons pas observer de grandes différences. C'est le cas pour les tarifs actuels. Cependant, le coefficient de crédibilisation est égal à 0,9 pour 40 moteurs pour le segment 2 et il est inférieur pour 40 moteurs pour le segment 3. L'idée de prendre un coefficient de crédibilisation plus petit pour 39 moteurs pourrait permettre de garder une cohérence avec le segment 3, voir d'avoir le même coefficient pour les deux segments.

Si nous gardons la même forme pour le coefficient de crédibilisation, nous pouvons choisir le coefficient le plus petit du graphique 4.5. Cela correspond à une valeur de K égal à 85. Ainsi nous aurons un coefficient de crédibilisation égale à 0,15 pour 5 moteurs et à 0,58 pour

39 moteurs. La figure 4.8 permet de comparer les deux coefficients pour des flottes de 5 à 39 moteurs.

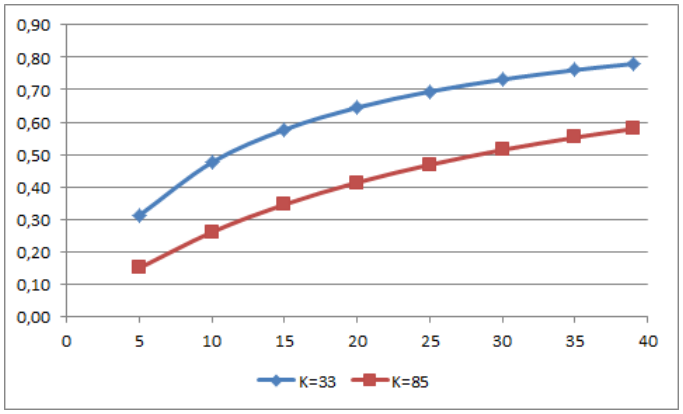


FIGURE 4.8 – Comparaison des coefficients de crédibilisation $n/(n+K)$ pour $K = 33$ et $K = 85$

Nous reprenons la même démarche que dans la section 4.3.3 pour élaborer une nouvelle échelle. Le nombre optimal de classes proche de 8 correspond à 7 ou 9 classes. Pour 7 classes, l'échelle n'a que deux classes pour lesquelles la valeur du coefficient est inférieure à 0,89. Nous choisissons donc de faire 9 classes.

Cluster	Nombre contrats	Moyenne	Ecart-type	Minimum	Maximum
7	1057	0,66	0,07	0,11	0,72
3	3561	0,79	0,03	0,72	0,83
1	3396	0,86	0,02	0,83	0,89
2	7094	0,94	0,03	0,89	1,02
5	2729	1,06	0,02	1,02	1,10
4	3048	1,16	0,04	1,10	1,25
6	1091	1,29	0,03	1,25	1,35
8	996	1,43	0,06	1,35	1,55
9	541	1,70	0,12	1,55	2,00

TABLE 4.10 – Neuf classes pour COEF borné à 2

Si nous comparons le tableau 4.7 avec le tableau 4.10, nous observons que le nombre de contrats est plus faible dans les deux classes extrêmes pour cette dernière classification. La

classe qui contient 1 est de loin la plus représentée et son effectif est plus important que dans la section précédente. Ceci s'explique par la diminution du coefficient de crédibilisation qui entraîne une plus grande concentration des valeurs autour de la valeur théorique et donc davantage de coefficients proches de 1.

En appliquant le même principe que dans la section précédente, nous obtenons l'échelle suivante :

Niveau	Valeur COEF	Coefficient BM	augmentation
1	<0,72	0,72	
2	[0,72 ; 0,83[	0,83	15%
3	[0,83 ; 0,89[	0,89	7%
4	[0,89 ; 1,02[	1,02	14%
5	[1,02 ; 1,1[	1,1	8%
6	[1,1 ;1,25[	1,25	14%
7	[1,25 ; 1,35[	1,35	8%
8	[1,35 ; 1,55[	1,55	14%
9	$\geq$ 1,55	2	29 %

TABLE 4.11 – Proposition d'échelle BM à 9 niveaux

Nous observons que l'augmentation entre chaque niveau n'est pas régulière. En ramenant cette échelle sur la même amplitude que dans la section précédente, nous obtenons l'échelle du tableau 4.12.

Niveau	Valeur COEF	Coefficient BM	augmentation
1	<0,72	0,8	
2	[0,72 ; 0,83[	0,84	5,3%
3	[0,83 ; 0,89[	0,86	2,7%
4	[0,89 ; 1,02[	0,91	5,7%
5	[1,02 ; 1,1[	0,95	3,5%
6	[1,1 ;1,25[	1,01	6,2%
7	[1,25 ; 1,35[	1,05	4%
8	[1,35 ; 1,55[	1,12	7,3%
9	$\geq$ 1,55	1,3	15,8 %

TABLE 4.12 – Deuxième proposition d'échelle BM à 9 niveaux

Nous remarquons que ce changement décale la valeur 1 qui n'est plus au centre de l'échelle. Il faudrait pour cela augmenter la borne supérieur de l'intervalle. Nous observons également davantage de disparités au niveau de l'augmentation lors du passage dans la classe supérieure.

Les échelles présentées sont des propositions qui n'ont pas encore été discutées et validées.

4.3.5 Test de la première échelle

Nous allons tester la première échelle proposée dans le tableau 4.9. Nous étudions l'affectation d'un contrat suivant sa fréquence et son nombre de moteurs. Dans un premier temps, nous prenons le cas d'un contrat en professions diverses et en zone 3. La fréquence collective est alors égale à 0,066, ce qui représente la fréquence collective moyenne sur l'ensemble des zones et des activités.

Fréquence	Niveau pour 5 moteurs	Niveau pour 10 moteurs	Niveau pour 20 moteurs	Niveau pour 30 moteurs	Niveau pour 40 moteurs	Ecart maximal
0	2	1	1	1	1	1
0,01	3	1	1	1	1	2
0,02	3	2	1	1	1	2
0,03	3	3	2	2	1	2
0,04	4	3	3	3	2	2
0,05	4	4	4	3	3	1
0,06	5	5	5	5	4	1
0,07	6	6	6	6	6	0
0,08	6	6	6	7	7	1
0,09	6	7	7	7	7	1
0,1	7	7	7	7	7	0
0,11	7	7	7	7	7	0
0,12	7	7	7	8	8	1
0,13	7	7	8	8	8	1
0,14	7	7	8	8	8	1
0,15	7	8	8	8	8	1
0,16	7	8	8	8	8	1
0,17	7	8	8	8	8	1
0,18	7	8	8	8	8	1
0,19	8	8	8	8	8	0
0,2	8	8	8	8	8	0

TABLE 4.13 – Niveau dans l'échelle bonus-malus donnée par le tableau 4.9 suivant la fréquence et le nombre de moteurs de la flotte

Dans le tableau 4.13, nous observons d'abord que pour une fréquence nulle, les petites flottes ne partent pas du même niveau que les autres. Elles commenceront au niveau 2 alors que les autres partent du niveau le plus petit. Le CPT était corrigé pour que tout le monde commence à 0,55 mais il n'y a pas de raison claire pour faire cette correction. Pour une même fréquence, l'écart maximal est de deux niveaux. Pour une fréquence de 0,01, nous avons les niveaux 1 et 3, ce qui représente un écart de tarif de 15%. Pour une fréquence de 0,04, nous avons les niveaux 2 à 4, ce qui représente un écart maximal de tarif de 14%. Sinon, nous avons au plus un niveau d'écart, soit moins de 10% d'écart tarifaire. En testant de la même façon avec les fréquences collectives extrêmes, nous n'avons pas d'écart de plus de deux niveaux dans tous les cas.

Cette échelle pourrait sembler satisfaisante. Cependant, la fréquence d'une flotte n'est pas une variable continue. Ainsi pour une flotte de 5 moteurs, nous avons les fréquences suivantes :

- pour 0 sinistre sur trois ans, fréquence 0
- pour 1 sinistre sur trois ans, fréquence $1/15 \simeq 0,067$
- pour 2 sinistres sur trois ans, fréquence $2/15 \simeq 1,33$

Ainsi, une flotte de 5 moteurs ne pourrait occuper que les niveaux 2, 7 et 8. Pour 10 moteurs, seuls les niveaux 1, 3, 6, 7 et 8 pourraient être atteints. A partir de 15 moteurs, tous les niveaux peuvent être atteints.

L'échelle ne paraît donc pas adaptée aux petites flottes. Il faudra utiliser une autre méthode pour ces flottes. Nous pourrons peut-être nous inspirer de ce qui est fait en transport routier de marchandise. Pour les petites flottes, il existe un tableau qui donne le coefficient suivant le nombre de sinistres et la taille de la flotte.

Il n'est pas évident de choisir entre ces deux propositions. Pour la suite, nous avons choisi de retenir le cas $K = 33$ pour le coefficient de crédibilisation. Pour cette valeur, le nombre de contrats est plus homogène dans chaque classe et la variation plus régulière entre deux classes adjacentes. Nous allons définir des règles d'évolution sur l'échelle définie dans le tableau 4.8.

4.4 Règle de mouvement sur l'échelle

La plupart du temps, les échelles bonus-malus sont faites indépendamment de la modélisation de la fréquence. Des méthodes permettent l'étude de ces échelles par une modélisation par des chaînes de Markov. Des critères donnent la performance de ces échelles. Notre but ici est d'intégrer directement l'échelle dans la modélisation par des modèles linéaires généralisés. Par construction de notre échelle, nous l'utiliserons uniquement dans le modèle de fréquence.

Pour un nouveau contrat, nous nous basons sur le relevé d'informations pour définir son niveau d'entrée sur l'échelle. Nous calculons une fréquence crédibilisé que nous comparons à la fréquence moyenne du groupe zone/activité. Nous nous fixons comme règle que tous les ans l'indicateur IND sera recalculé. Nous aurons alors trois cas suivant la valeur :

- si elle appartient à l'intervalle du niveau où se trouve le contrat, alors le contrat restera au même niveau
- si elle est supérieure à la borne supérieure du niveau actuel, alors le contrat passera au niveau supérieur
- si elle est inférieure à la borne inférieure du niveau actuel, alors de contrat passera au niveau inférieur

Ces règles sont assez simples. Elles permettent de réajuster le niveau tous les ans sans avoir de grosses variations de primes pour que ce changement soit acceptable pour le client.

Il nous reste à décider quelle fréquence nous devons utiliser pour calculer la valeur de l'indicateur IND. Jusqu'à présent, nous utilisions les résultats sur trois années ce qui permet

de diminuer la volatilité de la fréquence. Nous avons vu que les valeurs prises pour calculer la fréquence à l'affaire nouvelle n'était pas d'une grande précision. Cependant elle nous donne quand même une tendance. Nous pourrions nous baser uniquement sur les résultats depuis l'entrée du contrat dans le portefeuille mais la première année pourrait ne pas être significative. Une possibilité est d'utiliser des poids pour chaque année. Nous pouvons ainsi donner plus d'importance à la dernière année, un peu moins à l'année précédente et encore moins à l'année d'avant. Nous choisissons de donner à la dernière année un poids de 3, à l'année d'avant un poids de 2 et l'année encore d'avant de 1. Ainsi la dernière année a autant de poids que les deux années précédentes. La fréquence est donc donnée par la formule :

$$\bar{f} = \frac{3n_N + 2n_{N-1} + n_{N-2}}{3VA_N + 2VA_{N-1} + VA_{N-2}} \tag{4.13}$$

Nous regardons ce que cela aurait donné au niveau de la répartition par niveau pour les contrats entrés en 2012, 2013 et 2014. Nous ne gardons que ces trois années car pour les contrats entrés avant 2011, nous n'avons pas tous les contrats dans la base. Nous n'utilisons pas les contrats entrées en 2015 : la base ayant été constituée début 2016, nous n'avons pas connaissance de l'ensemble des sinistres de 2015.

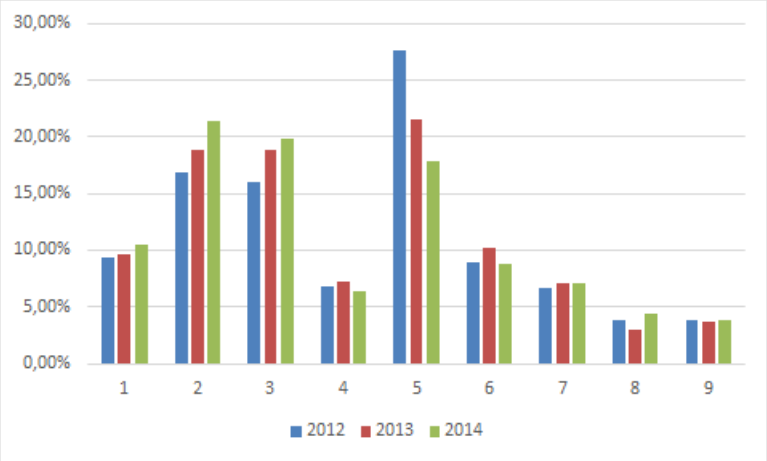


FIGURE 4.9 – Comparaison de la répartition des contrats par niveau la première année

Nous observons sur la figure 4.9 une stabilité sur la proportion des contrats par niveau. Le niveau où il y a le plus d'écart est le niveau 5 qui correspond en particulier au niveau d'entrée pour les nouvelles entreprises. Nous avons une différence de 10% entre les années 2012 et 2014 et la proportion décroit au cours des trois ans sur ce niveau. En parallèle, la proportion augmente au cours du temps pour les niveaux 1, 2 et 3.

Regardons à présent l'évolution au cours du temps de la répartition par niveau pour les contrats entrés en 2012.

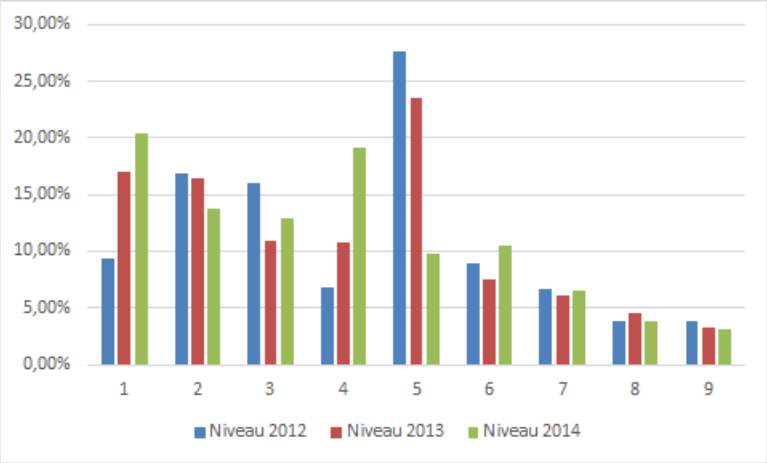


FIGURE 4.10 – Evolution de la répartition des contrats par niveau pour les contrats entrés en 2012

Les proportions sont exprimées à partir du nombre de contrats toujours en vigueur durant l'année de calcul. Ainsi il y a eu $5,8\%$ de contrats résiliés entre 2012 et 2013 et $10,7\%$ entre 2013 et 2014. Nous observons une stabilité pour les niveaux 7 à 9. Il y a une nette décroissance sur le niveau 5. Les contrats pour les créations d'entreprises ne restent pas sur ce niveau par défaut et une grande partie se retrouve sur les niveaux inférieurs. Nous observons davantage d'irrégularités sur les niveaux inférieurs.

Chapitre 5

Modélisation de la garantie RC matériel des véhicules de catégorie 1

En assurance automobile, il est courant de distinguer les sinistres matériels et corporels pour la modélisation. En effet, les montants sont beaucoup plus élevés pour les sinistres corporels mais ils sont beaucoup moins fréquents que les sinistres matériels. Dans ce mémoire, nous ne traiterons que les sinistres matériels mais les sinistres corporels peuvent se traiter suivant la même méthode.

Chaque catégorie de véhicules est tarifée séparément, les caractéristiques entre une voiture, un camion et un engin de chantier étant très différents. Nous ne traiterons que les véhicules de catégorie 1 parmi les flottes des segments un et deux. Souhaitant intégrer l'indice que nous avons construit au chapitre précédent, nous devons restreindre la base à l'ensemble des véhicules de catégorie un pour les contrats ayant été souscrits à partir de 2012.

Nous commençons par étudier la distribution du montant des sinistre afin de mettre en évidence les particularités des indemnisations en RC matériel automobile que nous avons déjà évoquées dans la partie 3.6.1.

5.1 Distribution des montants

Sur la figure 5.1, nous pouvons voir trois particularités :

- des montants de sinistres négatifs
- un grand nombre de sinistres d'un faible montant
- une queue de distribution épaisse pour des montants plus élevés

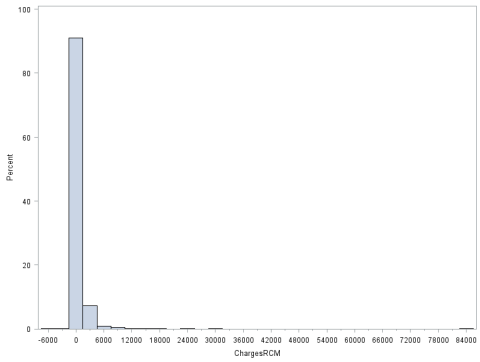


FIGURE 5.1 – Répartition des coûts moyen pour les véhicules de catégorie 1

Nous avons déjà expliqué la raison des sinistres négatifs. Nous verrons comment les traiter dans une prochaine section. Nous traitons d'abord la queue de distribution. Etant donné que nous allons modéliser la prime pure par un modèle de fréquence par coût moyen à l'aide de modèles linéaires généralisés, nous aurons à modéliser la densité par une loi. La queue épaisse entraînera une moins bonne adéquation des lois ou lognormal par exemple. Nous utilisons la théorie des valeurs extrêmes pour déterminer le montant des sinistres graves.

5.2 Détermination du seuil de grave

Il n'existe pas de méthode donnant une valeur fixe d'un seuil. Les méthodes utilisées sont des méthodes graphiques [DMPW07, DC15]. Le seuil est ensuite déterminé par l'interprétation de ces graphiques. Le seuil pourra donc être différent d'une personne à l'autre. Nous décrivons deux méthodes permettant de déterminer le seuil du montant des sinistres graves :

- la fonction de dépassement moyen
- la méthode de Hill

Soit X une variable aléatoire. La fonction de dépassement moyen est la fonction de $f(u)$ donnée par :

$$f(u) = \mathbb{E}\left(X - u | X > u\right) \tag{5.1}$$

Si la queue de distribution suit une loi de Pareto généralisée à partir d'un seuil u, alors le graphique de la courbe de la fonction f doit être linéaire au delà de ce seuil

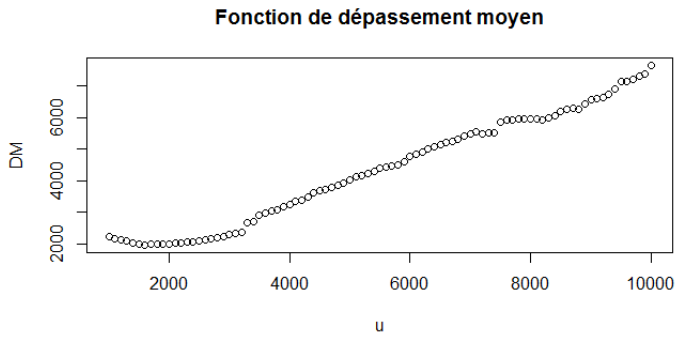


FIGURE 5.2 – Fonction de dépassement moyen

Sur le graphique 5.2, nous observons que la courbe est linéaire à partir de 3500 euros puis nous voyons un changement de pente à 7500 euros.

Pour les distributions de Fréchet, nous pouvons utiliser l'estimateur de Hill qui est définit par

$$\xi_{k,n} = \frac{1}{n}\sum_{i=1}^{n} \ln X_{(i)} - \ln X_{(k+1)} \tag{5.2}$$

où $(X_{(n)})$ désigne l'échantillon ordonné de (X_n). La courbe doit se stabiliser autour de la valeur de ξ, paramètre de la distribution de Fréchet.

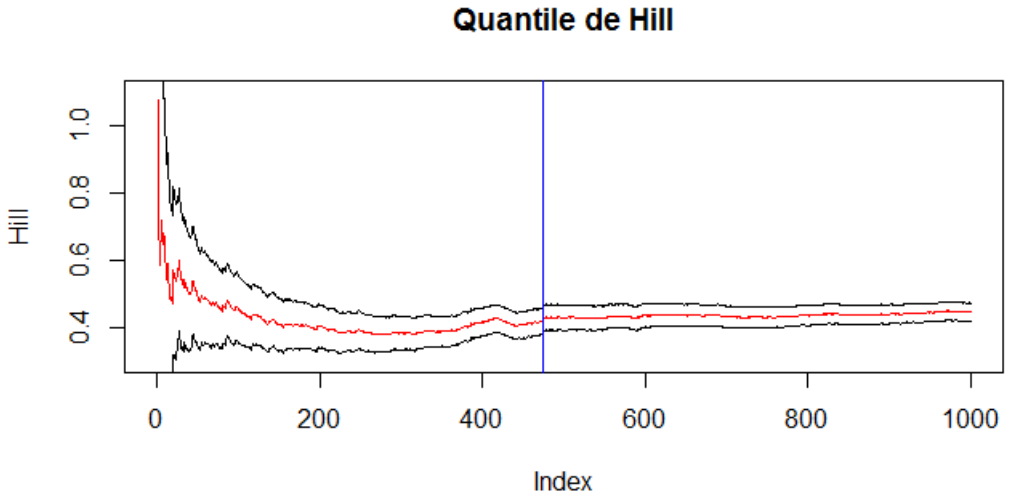


FIGURE 5.3 – Quantile de Hill

Notre interprétation du graphique 5.3 nous donne un seuil autour de l'indice 475 qui correspond à 7356 euros.

Nous retenons alors le montant de 7500 euros pour les montants translatés, soit 6000 euros pour le montant initial. Nous testons l'adéquation de la queue de distribution à une loi de Pareto généralisée. La comparaison des fonctions de répartitions empirique et théoriques, directement et avec un QQ-plot nous conforte dans notre choix.

5.3 Distribution des sinistres hors-graves

Nous enlevons les sinistres graves. Ensuite nous négligeons les quatre sinistres de moins de -1500 euros qui sont isolés. Pour modéliser par une loi usuelle, nous avons besoin d'une distribution à support positif. Nous translatons alors la distribution de 1500 euros.

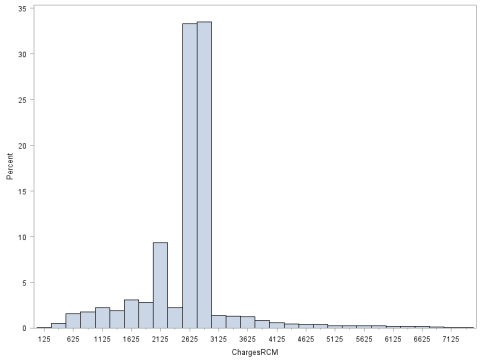


FIGURE 5.4 – Répartition des coûts moyen pour les véhicules de catégorie 1 hors grave et translater de 1500 euros

Nous observons deux grandes irrégularités dans la distribution, une autour de 2100 euros et une autre autour de 2800 euros. Ces montants correspondent au montant des sinistres IDA et demi-IDA responsables translatés de 1500. Afin d'avoir une distribution plus régulière, nous retirons les sinistres de ces montants. Nous obtenons alors une distribution qui pourrait être proche d'une loi gamma ou d'une loi log-normale.

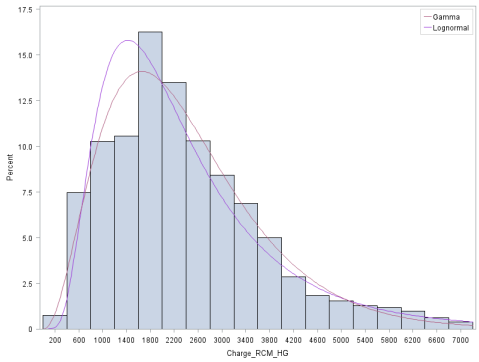


FIGURE 5.5 – Répartition des coûts moyen pour les véhicules de catégorie 1 hors graves et hors IDA et translater de 1500 euros

Pour choisir entre ces deux lois, nous utilisons les tests statistiques. Nous pouvons les illustrer le test de Kolmogorov-Smirnov et de Cramer von Mises à l'aide des fonctions de répartitions. Le test de de Kolmogorov-Smirnov se base sur la distance entre la courbe empirique et la courbe théorique alors que le test de Cramer von Mises se base sur l'aire entre les deux courbes.

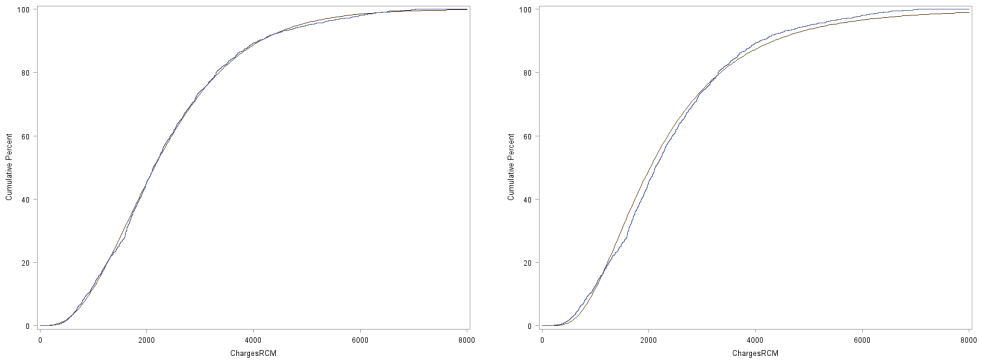


FIGURE 5.6 – Comparaison de la fonction de répartition

Visuellement, nous voyons sur le graphique 5.6 que la distribution des coûts est plus proche d'une loi gamma que d'une loi log-normale. Nous allons confirmer l'impression visuelle à l'aide des test statistiques.

Goodness-of-Fit Tests for Gamma Distribution				
Test	Statistique		P-value	
Kolmogorov-Smirnov	D	0.03136321	Pr > D	<0.001
Cramer-von Mises	W-Sq	0.17453161	Pr > W-Sq	0.014
Anderson-Darling	A-Sq	1.23634400	Pr > A-Sq	0.004

Goodness-of-Fit Tests for Lognormal Distribution				
Test	Statistique		P-value	
Kolmogorov-Smirnov	D	0.0622327	Pr > D	<0.010
Cramer-von Mises	W-Sq	1.7317825	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	10.1311629	Pr > A-Sq	<0.005

FIGURE 5.7 – Comparaison des résulats des tests d'adéquation à une loi

Les trois tests rejettent l'hypothèse d'adéquation à l'une ou l'autre loi. Nous travaillons ici sur un échantillon de grande taille et ces test ne donnent pas toujours de bons résultats dans ce cas. Si nous comparons les valeurs des tests, nous observons qu'elles sont toujours plus petites pour le loi gamma. Cela veut dire que l'écart entre la distribution empirique et la loi est plus petite à chaque fois pour la loi gamma. Nous retenons donc la loi gamma pour modéliser la distribution.

5.4 Distribution des fréquences

5.4.1 Fréquence des sinistres hors-graves et hors-IDA

Nous dénombrons 2 282 sinistres hors-graves sur les trois années d'études pour les contrats entrés en vigueur entre 2012 et 2014. Il se produit entre aucun et quatre sinistres par situation de véhicule qui est d'au plus un an. L'espérance est égale à $0,0145243$ et la variance à $0,01693118$. La variance n'étant pas égale à l'espérance, la loi de poisson a peu de chance de convenir pour modéliser le nombre de sinistres par situation. Nous testons la loi de Poisson et la loi binomiale négative à l'aide de la fonction goofit du package vcd de R avec l'estimation des paramètres par la méthode du maximum de vraisemblance.

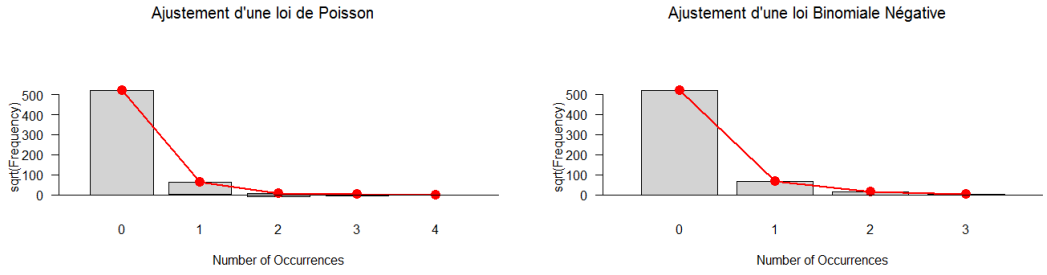


FIGURE 5.8 – Comparaison de la distribution du nombre de sinistres avec une loi de poisson et une loi binomiale négative

La figure 5.8 illustre l'adéquation de la distribution du nombre de sinistre à une loi de Poisson et à une loi binomiale négative. Visuellement la loi binomiale négative semble meilleure. Les deux lois ne satisfont pas le test du khi-deux mais la p-value est plus élevée

pour la loi binomiale négative. De plus, les valeurs modélisées sont meilleures pour la loi binomiale négative. Nous retenons donc cette loi pour le modèle linéaire généralisé.

5.4.2 Fréquence des sinistres IDA

Nous dénombrons 4 720 sinistres IDA sur les trois années d'études pour les contrats entrés en vigueur entre 2012 et 2014. Nous comparons de la même façon que pour la fréquence des sinistres hors-graves. La variance et l'espérance sont respectivement égales à $0,01846804$ et $0,01715512$. Il y a un écart de 10%. La loi de Poisson ne devrait donc pas convenir pour modéliser cette distribution.

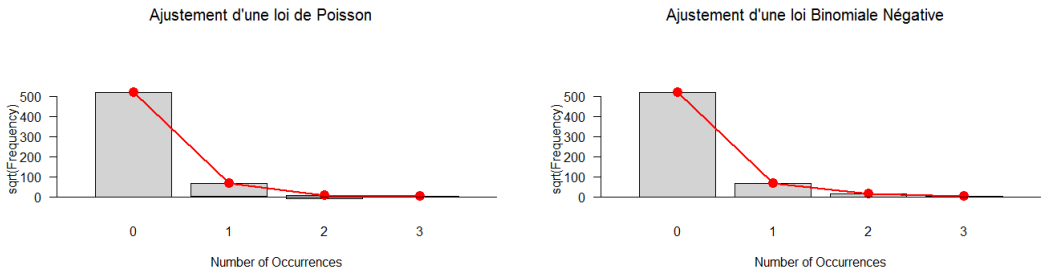


FIGURE 5.9 – Comparaison de la distribution du nombre de sinistres IDA avec une loi de poisson et une loi binomiale négative

Visuellement sur la figure 5.9 la loi binomiale négative semble mieux modéliser la distribution testée. Dans ce cas, la loi binomiale négative satisfait le test du khi-deux avec une p-value de $0,08356572$ contrairement à la loi de Poisson. Nous retenons donc la loi binomiale négative.

5.5 Statistiques descriptives

Nous observons les statistiques de répartition des véhicules-année, des fréquences, des coûts moyens et des primes pures pour différentes caractéristiques des véhicules et de la flotte. Nous commençons par la variable que nous avons créée : le niveau de l'échelle.

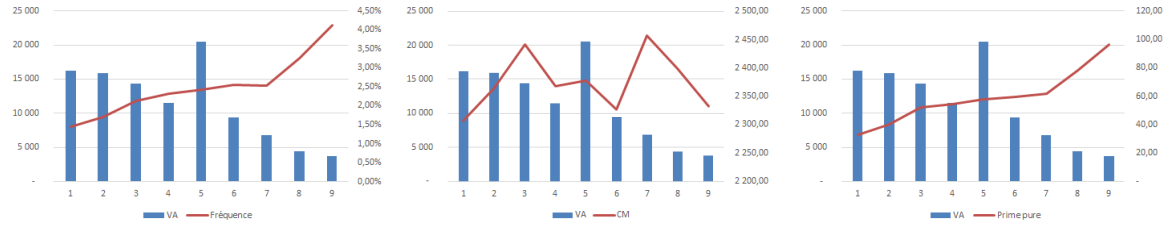


FIGURE 5.10 – Représentation de la fréquence, du coût moyen et de la prime pure par niveau

Nous observons, sur les graphiques de la figure 5.10, un plus grand nombre de véhicules-année sur les niveaux 1 à 5 avec un pic sur le niveau 5 qui correspond au niveau d'entrée d'une entreprise sans ou avec peu d'ancienneté. Ce nombre décroit du niveau 6 au niveau 9 qui est le niveau le plus risqué. Il y avait un tri sur les CPT élevés donc cela paraît cohérent d'avoir moins de véhicules-année sur les niveaux élevés, le risque étant plus élevé sur ces niveaux. Au niveau de la fréquence, nous observons une augmentation de la fréquence entre chaque niveau sauf du niveau 6 au niveau 7 où il y a un palier. L'augmentation est plus prononcée entre les niveaux 1 à 3 et 6 à 9 mais il n'y a pas de trop forte augmentation entre deux niveaux successifs. La variable niveau semble une bonne candidate en tant que variable explicative de la fréquence. Sur le coût moyen, il ne se dégage pas de tendance et nous avons un écart de seulement 7% entre le plus petit et le plus grand coût moyen. La variable niveau n'est donc pas une variable qui permette de modéliser le coût moyen. La courbe de la prime pure, qui est le produit de la fréquence et du coût moyen, a la même allure que celle de la fréquence.

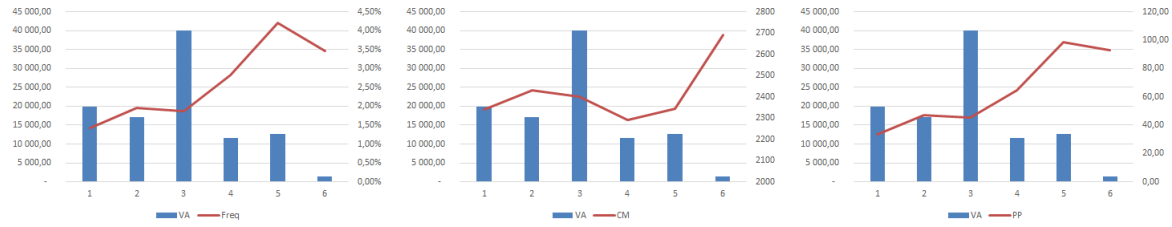


FIGURE 5.11 – Représentation de la fréquence, du coût moyen et de la prime pure par zone

Sur la figure 5.11, les zones 1 et 2 représentent chacune un peu moins de 20% des véhicules-année. La zone 3 est celle qui compte le plus de véhicules-année avec prêt de 40%. Dans les zones 4 et 5, nous trouvons un peu plus de 10% des véhicules-année et seulement 1% pour la zone 6. Nous pouvons nous demander l'intérêt d'avoir une zone 6 pour si peu de véhicules-année. Pour la modélisation, la zone 6 sera intégrée à la zone 5 son poids étant trop faible par rapport à l'ensemble des zones. Son coefficient pourra ensuite être ajusté si nous voulons

conserver un tarif plus élevé pour la zone 6. La fréquence est globalement croissante de la zone 1 à la zone 5. Il y a cependant un palier entre le niveau 2 et 3. La zone pourra être testée sur la modélisation de la fréquence. Le coût moyen varie peu d'une zone à l'autre : 6% d'écart entre le minimum et le maximum pour les zones 1 à 5. Ce coût est beaucoup plus élevé sur la zone 6. Cette variable n'est donc pas significative pour le coût moyen. La prime pure a le même comportement que la fréquence ce qui nous conforte dans le choix de ne retenir cette variable que pour modéliser la fréquence.

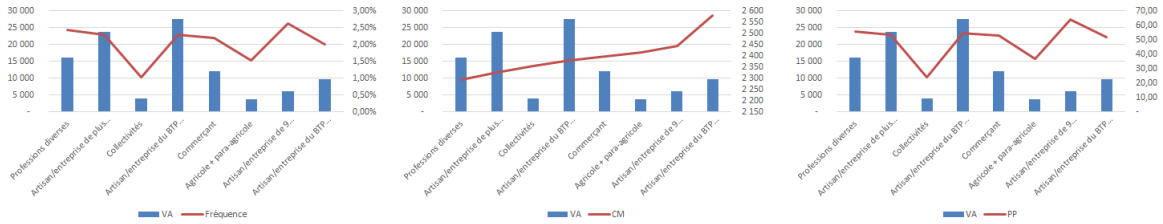


FIGURE 5.12 – Représentation de la fréquence, du coût moyen et de la prime pure suivant l'activité

Nous avons neuf activités différentes. La figure 5.12 met en évidence la disparité tant au niveau des poids des véhicules-année que des fréquences ou des coûts moyens. Cette variable peut aussi bien expliquer la fréquence que le coût moyen. Etant donné que l'allure de la courbe de la prime pure est plus proche de celle de la fréquence que du coût moyen, cette variable devrait avoir plus d'impact sur la fréquence. Nous avons seulement 12% d'écart entre le plus petit et le plus grand coût moyen alors qu'il y a plus de 150% entre la plus petite et la plus grande fréquence. Il n'est pas évident de voir quels regroupements pourront être faits afin de diminuer le nombre de modalités tout en restant cohérent.

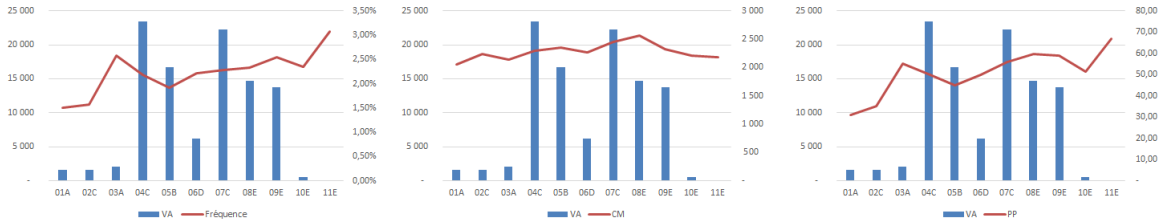


FIGURE 5.13 – Représentation de la fréquence, du coût moyen et de la prime pure suivant le groupe/classe

Pour le groupe-classe, nous avons cinq modalités qui sont bien représentées au niveau des véhicules-année avec entre 13% et 22% du nombre total de véhicules-année. Toutes les autres modalités représentent moins de 7% du total. En laissant cette variable en l'état, les

résultats ne seront sans doute pas probants au niveau du modèle linéaire généralisé. Nous observons un coût-moyen et une fréquence globalement croissants avec le groupe-classe. La prime-pure a un comportement plus proche de celui de la fréquence que du coût moyen sans cependant être aussi similaire que pour les autres variables. Le coût moyen a donc un impact non négligeable sur la prime pure.

FIGURE 5.14 – Représentation de la fréquence pour les sinistres IDA pour les quatre variables niveau, zone, activité et groupe/classe

Nous observons sur les graphiques de la figure 5.14 des comportements des fréquences pour les sinistres IDA différents de celui des sinistres hors-graves. Les fréquences augmentent bien avec le niveau et la zone si nous négligeons la zone 6. L'augmentation est régulière sans qu'il y ait de gros sauts entre deux niveaux ou zones successifs. Les fréquences pour les activités sont assez proches mise à par la fréquence pour l'activité "agricole+para-agricole" qui est beaucoup plus basse que les autres. Il ne se dégage pas de comportement bien identifiable pour les fréquences des différents groupe-classe.

5.6 Modélisation

5.6.1 Modélisation de la fréquence des sinistres hors-graves

La distribution de la fréquence des sinistres hors-graves est modélisée par une loi binomiale-négative. Nous utilisons la fonction de lien ln. Nous entrons en variable offset le logarithme du nombre de véhicules-année. Nous avons vérifié au préalable que toutes les variables étaient

décorrelées à l'aide des V de cramer (voir [GP11]). Toutes les valeurs étant plus petites que 0,4, nous pouvons supposer que toutes les variables ne sont pas corrélées. Etant donné le découpage que nous avons utilisé, le nombre de véhicules-année est plus petit que 1 pour chaque situation. Nous avons tester le modèle avec plusieurs variables et le meilleur modèle que nous ayons obtenu se base sur trois variables explicatives.

TABLE 5.1 – Statistiques de type 3

Source	DF	Chi_deux	Pr >Chi_deux
Niveau	6	248,40	<0,0001
Zone	4	522,62	<0,0001
Groupe-classe	2	19,06	<0,0001

Les statistiques de type III du tableau 5.1 montrent que les variables niveau, zone et groupe-classe sont bien explicatives de la fréquence. Pour obtenir ces résultats, nous avons du faire des regroupements. Pour la variable niveau, nous avons regroupé :

- les niveaux 3 et 4
- les niveaux 6 et 7.

Pour la zone, nous avons regroupé les zones 5 et 6. Pour le groupe-classe, il ne reste plus que trois groupes qui regroupent les 11 modalités :

- premier groupe : 01A-02C-03A
- deuxième groupe : 04C-05B-06D
- troisième groupe : 07C-08E-09E-10E-11E.

TABLE 5.2 – Résultat de la régression binomiale-négative pour la fréquence des sinistres hors-graves

Paramètre	Valeur	Coeff β	Ecart type	Intervalle de confiance à 95%		Chi-deux	Pr >Chi-deux
Intercept		-3,4307	0,0443	-3,5175	-3,344	6007,63	<,001
niveau	1	-0,3425	0,0621	-0,4642	-0,2208	30,43	<,001
niveau	2	-0,2463	0,0599	-0,3637	-0,129	16,92	0,0085
niveau	3-4	0	0	0	0	.	.
niveau	5	0,1283	0,0488	0,0327	0,2239	6,92	<,001
niveau	6-7	0,2063	0,0513	0,1059	0,3068	16,21	<,001
niveau	8	0,4805	0,0749	0,336	0,6273	41,13	<,001
niveau	9	0,7073	0,0736	0,563	0,8517	92,3	<,001
zone	1	-0,251	0,0554	-0,3596	-0,1424	20,51	<,001
zone	2	-0,1241	0,0516	-0,2254	-0,0229	5,78	0,0162
zone	3	0	0	0	0	.	.
zone	4	0,5087	0,0524	0,4061	0,6113	94,41	<,001
zone	5-6	0,8588	0,0449	0,7708	0,9467	366,29	<,001
groupe-classe	01A-02C-03A	-0,2345	0,0845	-0,4002	-0,0689	7,7	0,0055
groupe-classe	04C-05B-06D	-0,1308	0,0344	-0,1982	-0,0633	14,45	0,0001
groupe-classe	07C-08Eà11E	0	0	0	0	.	.

Dans le modèle dont les résultats sont donnés dans le tableau 5.2, les modalités de référence sont celles où nous observons le plus de véhicules-année :

- le niveau 3-4
- la zone 3
- le groupe-classe 07C-08E-09E-10E-11E

Ainsi la fréquence modélisée pour un véhicule du niveau 3 ou 4, de la zone 3, et du groupe-classe 07C ou 08E ou 09E ou 10E ou 11E est donnée par :

$$\exp(\beta_0) = \exp(-3,4307) = 3,24\%$$

La fréquence la plus basse est obtenue par un véhicule du niveau 1, de la zone 1 et du groupe-classe 01A ou 02C ou 03A et vaut :

$$\exp(\beta_0 + \beta_{niveau1} + \beta_{zone1} + \beta_{01A-02C-03A}) = \exp(-3,4307-0,3425-0,251-0,2345) = 1,41\%$$

La fréquence la plus élevée est obtenue par un véhicule du niveau 9, de la zone 5 ou 6 et du groupe-classe 07C ou 08E ou 09E ou 10E ou 11E et vaut :

$$\exp(\beta_0 + \beta_{niveau9} + \beta_{zone5-6}) = \exp(-3,4307 + 0,7073 + 0,8588) = 15,5\%,$$

soit plus de 10 fois la fréquence la moins élevée.

Nous observons que les valeurs des β, au sein d'une même variable, est cohérente avec leur construction. En effet, pour le niveau et pour la zone, les β sont croissants avec le niveau et la zone.

5.6.2 Modélisation de la fréquence des sinistres IDA

Nous procédons de la même façon pour la fréquence des sinistres demi-IDA. Nous avons vu que la loi qui modélisait le mieux leur distribution était également le loi binomial-négative. Le modèle ne donne que deux variables explicatives.

TABLE 5.3 – Statistiques de type 3

Source	DF	Chi_deux	Pr >Chi_deux
Zone	4	185,16	<0,0001
Niveau	5	596,76	<0,0001

Dans le tableau 5.3, nous voyons que les variables zone et niveau sont bien explicatives de la fréquence des sinistres IDA. Pour obtenir ce résultat, nous avons du regrouper des modalités. Pour la zone, nous avons juste regroupé les zones 5 et 6. Pour la variable niveau, nous avons du regroupé les niveaux 3, 4, 5 et 6.

TABLE 5.4 – Résultat de la régression binomiale-négative pour la fréquence des sinistres IDA

Paramètre	Valeur	DF	Coeff β	Ecart Type	Intervalle de confiance à 95%		Chi-deux	Pr >Chi-deux
Intercept		1	-3,1192	0,0283	-3,1747	-3,0638	12174,92	0,0000
Zone	1	1	-0,1735	0,0458	-0,2632	-0,0838	14,37	0,0002
Zone	2	1	-0,0621	0,0463	-0,1527	0,0286	1,80	0,1797
Zone	3	0	0,0000	0,0000	0,0000	0,0000		
Zone	4	1	0,2621	0,0477	0,1686	0,3555	30,22	0,0000
Zone	5-6	1	0,4355	0,0425	0,3522	0,5188	104,97	0,0000
Niveau	1	1	-0,3520	0,0499	-0,4499	-0,2541	49,65	0,0000
Niveau	2	1	-0,1507	0,0462	-0,2412	-0,0602	10,66	0,0011
Niveau	3-4-5-6	0	0,0000	0,0000	0,0000	0,0000		
Niveau	7	1	0,2530	0,0551	0,1451	0,3609	21,12	0,0000
Niveau	8	1	0,2980	0,0661	0,1685	0,4275	20,33	0,0000
Niveau	9	1	0,4896	0,0665	0,3593	0,6200	54,20	0,0000

Dans le tableau de résultats 5.4, les modalités de référence sont la zone 3 et le niveau qui regroupe les niveaux 3-4-5-6. Ainsi la fréquence de référence est donnée par :

$$\exp(\beta_0) = \exp(-3,1192) = 4,41\%$$

La fréquence la plus basse est obtenue pour un véhicule de la zone 1 et du niveau 1 et vaut

$$\exp(\beta_0 + \beta_{zone1} + \beta_{niveau1}) = \exp(-3,1192 - 0,1735 - 0,3520) = 2,61\%$$

La fréquence la plus élevée est obtenue pour un véhicule de la zone 5 ou 6 et du niveau 9 et vaut

$$\exp(\beta_0 + \beta_{zone5-6} + \beta_{niveau9}) = \exp(-3,1192 + 0,4355 + 0,4896) = 11,15\%$$

Nous obtenons une amplitude plus faible que pour la fréquence des sinistres hors-graves entre les deux fréquences extrêmes. Nous observons de nouveau que la variation des β est cohérente avec la construction des variables zone et niveau. La fréquence augmente bien quand la zone ou le niveau augmente.

Pour les sinistres demi-IDA, le nombre est trop faible pour pouvoir modéliser la fréquence par un modèle linéaire généralisé. Nous avons 508 sinistres. Nous retenons le montant du forfait demi-IDA de 2015 qui est égal à 638 euros. Cela représente 324 104 euros. Nous répartissons cette somme de façon uniforme sur les $102888,88$ véhicules-année, nous obtenons une prime pure de $3,15$ euros pour chaque véhicule.

5.6.3 Modélisation du coût moyen des sinistres hors-graves

Les tentatives de modélisation du coût moyen par modèle linéaire généralisé ne donne pas de résultats concluants, même en regroupant les variables. L'objectif n'étant pas d'avoir une tarification complètement aboutie mais d'étudier le tarif en introduisant la variable niveau, nous prendrons la valeur moyenne de tous les sinistres. Cette variable étant construite à partir de la fréquence, elle n'aura pas d'impact sur le coût moyen. Nous pourrions envisager de prendre tous les sinistres de l'ensemble du portefeuille et pas seulement sur les contrats entrés entre 2012 et 2015. Nous retenons le montant de $875,10$ euros pour le coût moyen d'un sinistre hors-grave.

5.6.4 Sinistres graves

Nous avons 62 sinistres graves de plus de 7500 euros pour un montant total de $889153,75$ euros. Nous ne pouvons envisager de modélisation pour un si petit nombre de sinistres que ce soit au niveau de la fréquence comme du coup moyen. Le montant n'est pas négligeable. Nous choisissons de la répartir sur l'ensemble des véhicules. Nous avons $102888,88$ véhicules-année sur l'ensemble des situations de l'étude. Nous obtenons une prime pure de $8,65$ euros.

5.7 Etude de l'impact de la variable niveau

Nous avons modélisé la fréquence des sinistres hors-graves et avons retenu pour la fréquence une fréquence moyenne de 875,1 euros. Nous avons modéliser la fréquence des sinistres

IDA et nous connaissons le montant d'un sinistre IDA. Nous retenons le montant de 2014 : 1276 euros. Pour les sinistres demi-IDA et graves, nous avons retenu un montant forfaitaire qui est égal respectivement à 3,15 et 8,65 euros. Finalement, la prime pure pour un véhicule sera donnée par la formule :

$$\text{PP} = \text{freq}_{HG} \times 875,1 + \text{freq}_{IDA} \times 1276 + 3,15 + 8,65$$

Nous voulons étudier l'impact de la variable niveau. Nous nous fixons donc une zone et un groupe classe. Nous choisissons arbitrairement la zone 3, qui est la plus représentée, et la classe E. Ces variables étant fixés, nous étudions la prime pure et son évolution en fonction du niveau. Etant donnés les résultats des modèles linéaires généralisés, nous savons que la prime pure sera croissante avec le niveau. Ce qui nous intéresse particulièrement est la vitesse à laquelle elle va croître et notamment l'écart entre deux niveaux successifs.

Niveau	Prime pure	Augmentation entre les niveaux i-1 et i
1	71,57	
2	82,44	15%
3	96,51	17%
4	96,51	0%
5	100,39	4%
6	103,00	3%
7	119,23	16%
8	133,55	12%
9	161,26	21%

TABLE 5.5 – Evolution de la prime pure pour un véhicule de la zone 3 et de classe E

Dans le tableau 5.5, nous donnons la prime pure obtenue pour chaque niveau ainsi que le pourcentage d'augmentation entre le niveau i-1 et le niveau i, i prenant les valeurs entières de 2 à 9. Si nous prenons les deux primes pures extrêmes, nous observons un rapport de 1 à 2,25 entre la prime la plus élevée et la prime la plus basse. Cela est inférieur au rapport actuel qui vaut 2,72. La variation entre deux niveaux successifs n'est pas régulière : nous avons entre 0% d'augmentation entre le niveau 3 et 4 et 21% entre le niveau 8 et 9. L'augmentation entre le niveau 8 et 9 n'est pas choquante puisque nous sommes ici sur des flottes qui ont une fréquence très élevée et il paraît cohérent de sanctionner financièrement des flottes qui ont une forte sinistralité. L'absence d'augmentation entre le niveau 3 et 4 s'explique par les regroupements que nous avons dû faire pour que la valeur de tous les β soit significative.

Niveau	β_{IDA}	β_{HG}	Prime pure	Augmentation entre les niveaux i et i+1
1	-0,3520	-0,3425	71,57	
2	-0,1507	-0,2463	82,44	15%
3	-0,05	-0,05	92,38	12%
4	0,0000	0,05	97,96	6%
5	0,05	0,1283	103,28	5%
6	0,1000	0,2063	108,93	5%
7	0,2530	0,2063	119,23	9%
8	0,2980	0,4805	133,55	12%
9	0,4896	0,7073	161,26	21%

TABLE 5.6 – Evolution de la prime pure retravaillée pour un véhicule de la zone 3 et de classe E

Dans le tableau 5.6, nous proposons une modification de la valeur des β pour séparer les niveaux où nous avons fait des regroupements. Nous prenons des valeurs qui permettent de garder une cohérence avec les niveaux supérieurs et inférieurs. Ainsi le β_{IDA} pour les niveaux 3 à 6 doit être compris entre -0,1507 et 0,253. Il doit également être proche de 0. Nous choisissons d'attribuer les valeurs -0,05, 0, 0,05 et 0,1 pour, respectivement les niveaux 3, 4 5 et 6. Nous faisons de même pour le β_{HG}. Les modifications ne se font donc que pour les niveaux 3 à 6. Nous observons dans ce cas que les variations sont plus homogènes au centre de l'échelle de niveau. Les écarts sont plus grand entre les petits niveaux ou les grands niveaux. Ainsi une flotte peut naviguer au centre de l'échelle sans que son tarif ne subissent trop de variation mais dès qu'elle atteint les extrémités, sa prime sera davantage impactés. Par construction de l'échelle, si une flotte arrive au niveau inférieur ou supérieur, elle aura parcouru les niveaux intermédiaires et nous saurons, par sa sinistralité antérieures, si la flotte a eu de bons ou mauvais résultats par le passé. Ainsi nous pouvons récompenser les flottes que nous savons prudentes et au contraire pénaliser plus fortement les flottes dont la sinistralité est très élevée depuis plusieurs années.

Les résultats semblent conforter l'idée d'utiliser une variable modélisant la sinistralité de la flotte à l'intérieur du modèle linéaire généralisé. Ces premiers résultats sont encourageants et pousseraient à continuer dans cette voix. L'avantage d'intégrer cette variable dans la modélisation est d'assurée l'équilibre du tarif ce qui ne serait pas le cas si les coefficients étaient fixés au préalable comme nous l'avions fait dans un premier temps.

Conclusion

Le travail présenté est le début d'une étude qui a été poursuivie par un cabinet de conseil pour le compte de cette société d'assurance en 2017. Un nouveau coefficient appelé coefficient d'antériorité a été mis en place.

Nous avons d'abord essayé de comprendre la formule du CPT actuel en montrant ces limites. Le fait qu'il soit figé paraît être le principal inconvénient avec également la crédibilisation très rapide.

Nous avons commencé par revoir le coefficient de crédibilisation en utilisant la théorie de Bühlmann-Straub, en travaillant sur des groupes définis par la zone et l'activité qui sont les deux seules variables communes à tous les types de véhicules. Au préalable, nous avions montré que ces variables sont bien des variables explicatives de la fréquence.

Nous sommes partis sur un coefficient de crédibilisation de la forme $n/(n+K)$ en prenant une seule valeur de K pour tous les couples zone/activité. Nous avons travaillé avec deux valeurs de K, une qui correspond au coefficient de crédibilité moyen et une qui est choisie pour avoir un facteur de crédibilité assez petit pour 40 moteurs.

Pour un coefficient donné, nous avons calculé une fréquence crédibilisé et l'avons comparée à la fréquence théorique de la zone/activité. Nous avons ainsi construit un coefficient que nous avons cherché à diviser en classe à l'aide d'une classification hiérarchique. Ceci nous a permis d'obtenir une première échelle que nous avons ensuite modifiée pour diminuer l'écart entre chaque niveau.

Nous avons défini des modalités d'évolution simple sur cette échelle et nous avons calculer le niveau occupé chaque année par chaque contrat afin de pouvoir inclure cette variable dans la modélisation du tarif par modèle linéaire généralisé. Nous avons montré que cette nouvelle variable était bien explicative des fréquences et avons étudié son impact sur la prime pure. Les résultats obtenus semblent montrer que la sinistralité est intéressante à utiliser.

Cette étude a permis de mettre en évidence l'intérêt de mettre en place une échelle bonus-malus pour la tarification de flottes de véhicules pour la garantie RC sur la partie corporelle. Nous n'obtenons pas un tarif abouti mais les résultats sont encourageant pour continuer dans cette voie. Il faudra étudier son impact sur la partie RC corporel que nous n'avons pas traité. Il paraît difficile de ne changer que le CPT sans changer l'intégralité du tarif. Ainsi le zonier et la classification des activités et des groupes-classe seraient également à revoir. Actuellement, le même coefficient est appliqué aux garanties RC et DTA alors qu'il

n'est construit que sur la sinistralité RC. Une réflexion devra être menée sur la pertinence de l'utilisation d'une même variable explicative pour ces deux garanties. La méthode de classification pourrait être comparée avec par exemple des méthodes de forêt aléatoire ou des *k-means* afin de comparer les résultats.

Bibliographie

[ADD04] J.-F. Angers, D. Dejardins, G. Dionne, *Modèle Bayésien de tarification de l'assurance des flottes de véhicules*, L'actualité économique, vol. 80, n 2-3, http ://id.erudit.org/iderudit/011388ar, 2004 p.253-303.

[BG05] H. Bühlmann, A. Gisler, *A course in credibility theory and its applications*, Springer, 2005.

[DMPW07] M. Denuit, X. Maréchal, S. Pitrebois, J-F. Walhin, *Actuarial modelling of claim counts*, Wiley, 2007.

[DC15] M. Denuit, A. Charpentier, *Mathématiques de l'assurance non-vie, tome II : tarification et provisionnement*, Economica, 2015.

[DV84] G.W. De Wit & J. Van Eeghen, *Rate making and society's sense of fairness*, ASTIN Bulletin 14, 151-163, 1984.

[KGDD08] Rob Kaas, Marc Goovaerts, Jan Dhaene, Michel Denuit, *Modern Actuarial Risk Theory*, Springer, 2008.

[VG10] Vincent Goulet, *Mathématiques actuarielles IARD II (Théorie de la crédibilité)*, Notes de cours, 2010.

[GP11] Gilbert Saporta, *Probabilités, analyse des données et Statistique*, Edition TECNIP, 2011.

Annexe A

Loi Badinter

Loi n 85-677 du 5 juillet 1985 tendant à l'amélioration de la situation des victimes d'accidents de la circulation et à l'accélération des procédures d'indemnisation

Version consolidée au 14 janvier 2010

Chapitre Ier : Indemnisation des victimes d'accidents de la circulation.

Article 1

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

Section I : Dispositions relatives au droit à indemnisation.

Article 2

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

Article 3

Les victimes, hormis les conducteurs de véhicules terrestres à moteur, sont indemnisées des dommages résultant des atteintes à leur personne qu'elles ont subis, sans que puisse leur être opposée leur propre faute à l'exception de leur faute inexcusable si elle a été la cause exclusive de l'accident. Les victimes désignées à l'alinéa précédent, lorsqu'elles sont âgées de moins de seize ans ou de plus de soixante-dix ans, ou lorsque, quel que soit leur âge, elles sont titulaires, au moment de l'accident, d'un titre leur reconnaissant un taux d'incapacité

permanente ou d'invalidité au moins égal à 80 p. 100, sont, dans tous les cas, indemnisées des dommages résultant des atteintes à leur personne qu'elles ont subis. Toutefois, dans les cas visés aux deux alinéas précédents, la victime n'est pas indemnisée par l'auteur de l'accident des dommages résultant des atteintes à sa personne lorsqu'elle a volontairement recherché le dommage qu'elle a subi.

Article 4

La faute commise par le conducteur du véhicule terrestre à moteur a pour effet de limiter ou d'exclure l'indemnisation des dommages qu'il a subis.

Article 5

La faute, commise par la victime a pour effet de limiter ou d'exclure l'indemnisation des dommages aux biens qu'elle a subis. Toutefois, les fournitures et appareils délivrés sur prescription médicale donnent lieu à indemnisation selon les règles applicables à la réparation des atteintes à la personne.
Lorsque le conducteur d'un véhicule terrestre à moteur n'en est pas le propriétaire, la faute de ce conducteur peut être opposée au propriétaire pour l'indemnisation des dommages causés à son véhicule. Le propriétaire dispose d'un recours contre le conducteur.

Article 6

Le préjudice subi par un tiers du fait des dommages causés à la victime directe d'un accident de la circulation est réparé en tenant compte des limitations ou exclusions applicables à l'indemnisation de ces dommages.

Annexe B

Macro CAH

```
%macro cah(dataset,ident,listvar,nc,methode=ward);
%* Classification ascendante hierarchique;
%*   dataset : table contenant les donnees;
%*   ident   : identificateurs des individus;
%*   listvar : liste des variables numeriques;
%*   methode : methode de classification hierarchique utilisee;
%*   nombre desire de classes;
data donnees;
     set work.&dataset (keep=&ident &listvar);
run;
proc cluster data=donnees method=&methode outtree=tree noprint;
     var &listvar;
     copy &ident;
run;
proc sort data=tree;
     by _ncl_;
run;
proc tree data=tree  out=work.chclasse nclusters=&nc noprint;
     copy &ident &listvar;
run;
proc sort data=work.chclasse;
     by cluster;
run;
proc means data=work.chclasse ;
     var &listvar;
     by cluster;
     output out=poles;
run;
```

```
data poles(drop=_type_);
     set poles;
     if _stat_='MEAN';
run;
%mend;

%cah(freq_cred,NU_AFFA,&listvar.,12);


/*Réduction du nombre de classe pour les graphiques*/
data tree;
set tree;
where _NCL_<=15;
run;

proc gplot data=tree;
symbol1 interpol=join value=dot;
plot _PST2_*_NCL_=1 ;
plot _RMSSTD_*_NCL_=1;
plot _RSQ_*_NCL_=1;
run;
quit;
```

Annexe C

Résultats de la macro CAH

Cluster	Nombre contrats	Moyenne	Ecart-type	Minimum	Maximum
2	5661	0,5910349	0,0888777	0,0464135	0,699532
1	7621	0,8157107	0,0639816	0,6996306	0,9294015
5	3864	1,0368624	0,063972	0,9297656	1,1508095
4	3789	1,293856	0,0945255	1,1511619	1,4895176
3	2218	1,7276516	0,180946	1,4901907	2,1556485
6	500	2,6833194	0,6714044	2,1574817	9,3748408

TABLE C.1 – Résultats de la macro CAH pour six classes

Cluster	Nombre contrats	Moyenne	Ecart-type	Minimum	Maximum
2	5661	0,5910349	0,0888777	0,0464135	0,699532
1	7621	0,8157107	0,0639816	0,6996306	0,9294015
5	3864	1,0368624	0,063972	0,9297656	1,1508095
4	3789	1,293856	0,0945255	1,1511619	1,4895176
3	1571	1,6285558	0,0947562	1,4901907	1,8276119
6	647	1,9682688	0,0931462	1,8294038	2,1556485
7	401	2,4375674	0,2234104	2,1574817	2,9648918
8	99	3,6787393	0,918423	2,9873807	9,3748408

TABLE C.2 – Résultats de la macro CAH pour huit classes

Cluster	Nombre contrats	Moyenne	Ecart-type	Minimum	Maximum
6	2432	0,5100877	0,0755177	0,0464135	0,5963757
3	3229	0,6520022	0,030015	0,5966711	0,699532
1	4438	0,7699562	0,0401412	0,6996306	0,8414839
2	3183	0,8795054	0,0237794	0,8416145	0,9294015
8	3864	1,0368624	0,063972	0,9297656	1,1508095
7	2645	1,2415838	0,0554476	1,1511619	1,3482934
5	1144	1,4147127	0,0394673	1,3488198	1,4895176
4	1571	1,6285558	0,0947562	1,4901907	1,8276119
9	647	1,9682688	0,0931462	1,8294038	2,1556485
10	401	2,4375674	0,2234104	2,1574817	2,9648918
11	97	3,5745002	0,5559389	2,9873807	5,5050206
12	2	8,7343332	0,9058145	8,0938256	9,3748408

TABLE C.3 – Résultats de la macro CAH pour douze classes

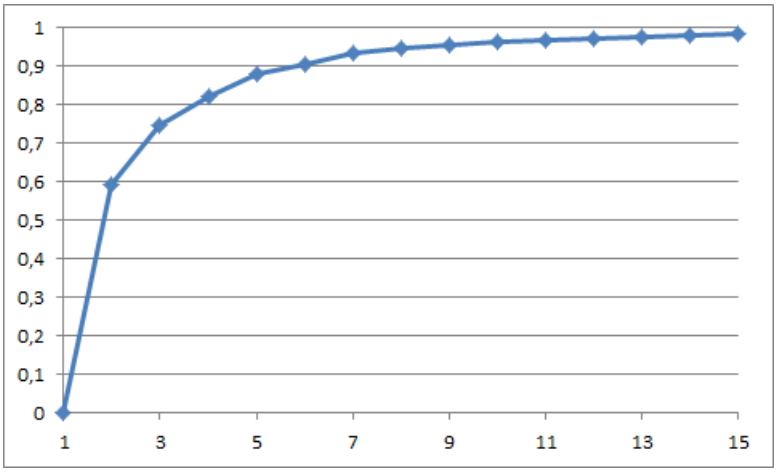


FIGURE C.1 – Courbe des R^2

Annexe D

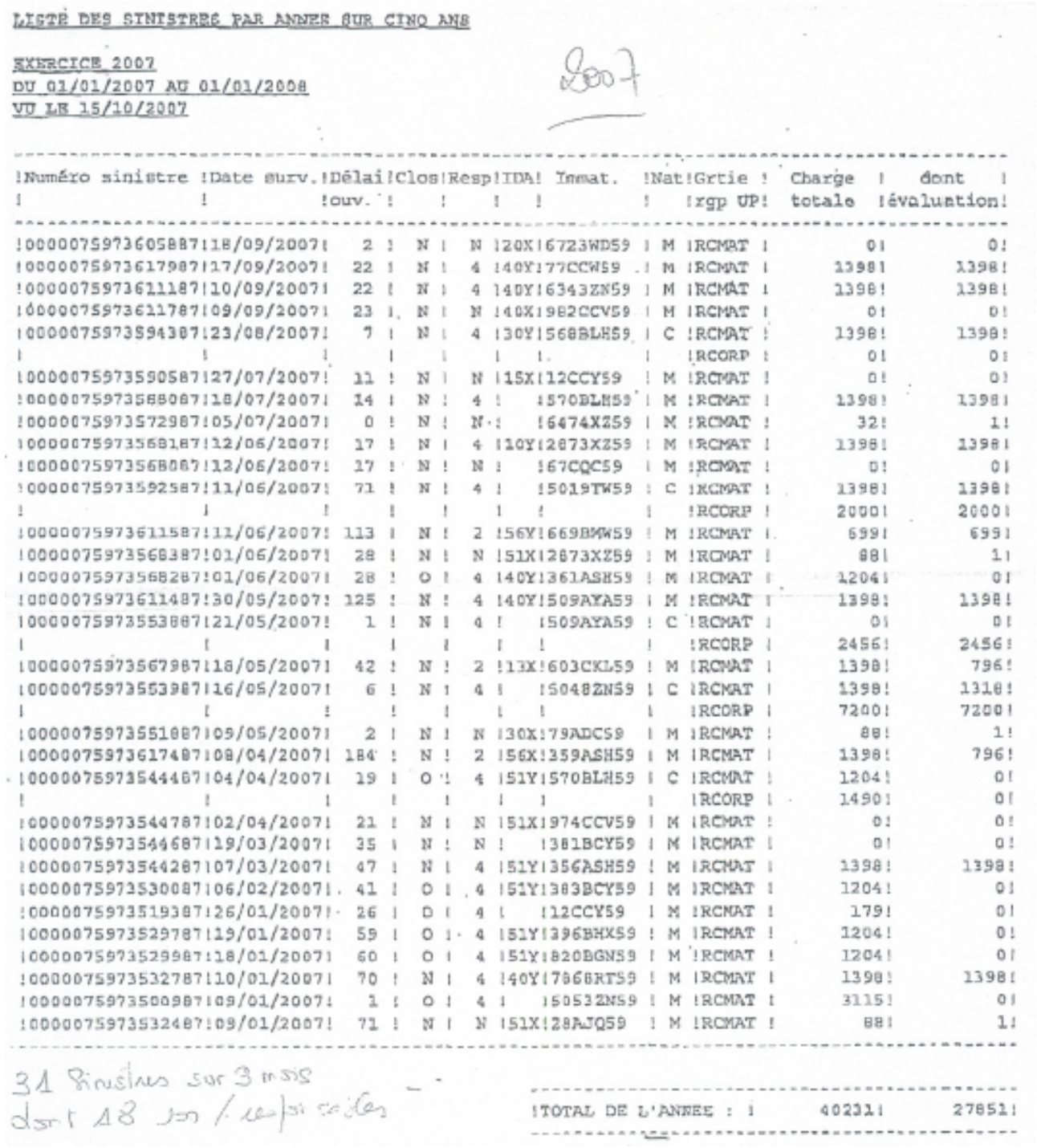
Relevé d'informations

Exercice : 2013 **Période de référence : 01/10/2013 au 30/09/2014** **Nombre sinistres : 14**

FLOTTES

Réf sinistre	N° contrat	Produit	Date sinistre	Date enreg. sinistre	Immatriculation	TEA	Sin corp	Tx resp IRSA	GMJ	Règlements de dépense	Règlements de recours	Provisions de dépense	Provisions de recours	Coût GMJ	Etat GMJ
AL14045358M	0886251Q00006	GROUPAMA FLOTTE	27/08/2014	28/08/2014	BM-391-EH	VEHICULE +3.5T	N	T	DSC	4 768,17	0,00	0,00	0,00	4 768,17	C
AL14045358M	0886251Q00006	GROUPAMA FLOTTE	27/08/2014	28/08/2014	BM-391-EH	VEHICULE +3.5T	N	T	RC	0,00	0,00	0,00	0,00	0,00	C
AL14038907C	0886251Q00006	GROUPAMA FLOTTE	20/07/2014	24/07/2014	DB-511-MX	VEHICULE +3.5T	O	100	DC	2 049,72	0,00	0,00	0,00	2 049,72	C
AL14038907C	0886251Q00006	GROUPAMA FLOTTE	20/07/2014	24/07/2014	DB-511-MX	VEHICULE +3.5T	O	100	RC	1 276,00	0,00	0,00	0,00	1 276,00	C
AL14034277C	0886251Q00006	GROUPAMA FLOTTE	01/07/2014	03/07/2014	DB-511-MX	VEHICULE +3.5T	N	0	DC	7 069,66	6 969,67	0,00	0,00	99,99	C
AL14036973E	0886251Q00006	GROUPAMA FLOTTE	16/06/2014	16/07/2014	AT-257-EV	VEHICULE 4R-3.5	N	E	BDG	55,60	0,00	0,00	0,00	55,60	C
AL14036979L	0886251Q00006	GROUPAMA FLOTTE	31/05/2014	16/07/2014	CV-893-EC	VEHICULE +3.5T	N	E	DSC	0,00	0,00	0,00	0,00	0,00	C
AL14487212A	0886251Q00006	GROUPAMA FLOTTE	29/04/2014	13/05/2014	571ANK29	VEHICULE 4R-3.5	N	NS	BDG	599,79	0,00	0,00	0,00	599,79	C
AL14026243V	0886251Q00006	GROUPAMA FLOTTE	25/04/2014	27/05/2014	DB-511-MX	VEHICULE +3.5T	N	T	RC	101,17	0,00	0,00	0,00	101,17	C
AL14484388M	0886251Q00006	GROUPAMA FLOTTE	14/04/2014	29/04/2014	1292ZH29	VEHICULE 4R-3.5	N	NS	BDG	259,32	0,00	0,00	0,00	259,32	C
AL14479282T	0886251Q00006	GROUPAMA FLOTTE	30/03/2014	14/04/2014	BM-391-EH	VEHICULE +3.5T	N	NS	BDG	908,45	0,00	0,00	0,00	908,45	C
AL14031082E	0886251Q00006	GROUPAMA FLOTTE	27/02/2014	19/06/2014	CS-730-RT	VEHICULE +3.5T	N	T	RC	4 437,24	0,00	0,00	0,00	4 437,24	C
AL14007572D	0886251Q00006	GROUPAMA FLOTTE	14/02/2014	17/02/2014	5460ZM29	REMORQUE / SEMI	N	E	EVN	5 726,30	0,00	0,00	0,00	5 726,30	C
AL13069358H	0886251Q00006	GROUPAMA FLOTTE	18/12/2013	23/12/2013	DA-664-ZM	VEHICULE +3.5T	N	0	DC	4 297,10	1 242,00	0,00	0,00	3 055,10	C
AL13065197R	0886251Q00006	GROUPAMA FLOTTE	17/11/2013	03/12/2013	BL-360-XA	REMORQUE / SEMI	N	100	RC	1 242,00	0,00	0,00	0,00	1 242,00	C
AL13527038W	0886251Q00006	GROUPAMA FLOTTE	16/10/2013	23/10/2013	BD-650-HG	VEHICULE 4R-3.5	N	NS	BDG	494,42	0,00	0,00	0,00	494,42	C
										33 284,94	8 211,67	0,00	0,00	25 073,27	

FIGURE D.1 – Extrait d'un relevé d'information de Groupama Loire Bretagne



LISTE DES SINISTRES PAR ANNEE SUR CINQ ANS

EXERCICE 2007
DU 01/01/2007 AU 01/01/2008
VU LE 15/10/2007

2007

Numéro sinistre	Date surv.	Délai ouv.	Clos	Resp	IDA	Immat.	Nat	Grtie rgp UP	Charge totale	dont évaluation
0000075973605887	18/09/2007	2	N	N	20X	6723WD59	M	RCMAT	0	0
0000075973617987	17/09/2007	22	N	4	40Y	77CCW59	M	RCMAT	1398	1398
0000075973611187	10/09/2007	22	N	4	40Y	6343ZN59	M	RCMAT	1398	1398
0000075973611787	09/09/2007	23	N	N	40X	982CCV59	M	RCMAT	0	0
0000075973594387	23/08/2007	7	N	4	30Y	568BLH59	C	RCMAT	1398	1398
								RCORP	0	0
0000075973590587	27/07/2007	11	N	N	15X	12CCY59	M	RCMAT	0	0
0000075973588087	18/07/2007	14	N	4		570BLH59	M	RCMAT	1398	1398
0000075973572987	05/07/2007	0	N	N		6474XZ59	M	RCMAT	32	1
0000075973568187	12/06/2007	17	N	4	10Y	2873XZ59	M	RCMAT	1398	1398
0000075973568087	12/06/2007	17	N	N		67CQC59	M	RCMAT	0	0
0000075973592587	11/06/2007	71	N	4		5019TW59	C	RCMAT	1398	1398
								RCORP	2000	2000
0000075973611587	11/06/2007	113	N	2	56Y	669BMW59	M	RCMAT	699	699
0000075973568387	01/06/2007	28	N	N	51X	2873XZ59	M	RCMAT	88	1
0000075973568287	01/06/2007	28	O	4	40Y	361ASH59	M	RCMAT	1204	0
0000075973611487	30/05/2007	125	N	4	40Y	509AYA59	M	RCMAT	1398	1398
0000075973553887	21/05/2007	1	N	4		509AYA59	C	RCMAT	0	0
								RCORP	2456	2456
0000075973567987	18/05/2007	42	N	2	13X	603CKL59	M	RCMAT	1398	796
0000075973553987	16/05/2007	6	N	4		5048ZN59	C	RCMAT	1398	1318
								RCORP	7200	7200
0000075973551887	09/05/2007	2	N	N	30X	79ADC59	M	RCMAT	88	1
0000075973617487	08/04/2007	184	N	2	56X	359ASH59	M	RCMAT	1398	796
0000075973544487	04/04/2007	19	O	4	51Y	570BLH59	C	RCMAT	1204	0
								RCORP	1490	0
0000075973544787	02/04/2007	21	N	N	51X	974CCV59	M	RCMAT	0	0
0000075973544687	19/03/2007	35	N	N		381BCY59	M	RCMAT	0	0
0000075973544287	07/03/2007	47	N	4	51Y	356ASH59	M	RCMAT	1398	1398
0000075973530087	06/02/2007	41	O	4	51Y	383BCY59	M	RCMAT	1204	0
0000075973519387	26/01/2007	26	O	4		12CCY59	M	RCMAT	179	0
0000075973529787	19/01/2007	59	O	4	51Y	396BHX59	M	RCMAT	1204	0
0000075973529987	18/01/2007	60	O	4	51Y	820BGN59	M	RCMAT	1204	0
0000075973532787	10/01/2007	70	N	4	40Y	7868RT59	M	RCMAT	1398	1398
0000075973500987	09/01/2007	1	O	4		5053ZN59	M	RCMAT	3115	0
0000075973532487	09/01/2007	71	N	N	51X	28AJQ59	M	RCMAT	88	1

31 Sinistres sur 3 mois
dont 18 non / responsables

TOTAL DE L'ANNEE :	40231	27851

FIGURE D.2 – Extrait d'un relevé d'information d'AXA