

CNAM
- Master « Science de Gestion, Finance de Marché »
Spécialité « Actuariat » -

**« Flottes automobiles :
Un nouveau modèle de tarification.
Impact de la conservation sur la distribution du ratio
sinistres à primes »**

- Mémoire confidentiel -

Mémoire présenté le 8 juillet 2008 par

Mademoiselle NGUYEN Thi To-Vong

Jury

M. FROMENTEAU Michel
M. MOLINA Eric
M. PETAUTON Pierre
M. WARIN Patrick
M. WEISS François

Responsable du suivi dans l'entreprise
M. MONGIS Gilles

Mot-clés : grosses flottes automobiles, garantie responsabilité civile, tarification, système bonus malus, conservation, ratio sinistres à primes, mesures de risque, modélisation des graves, lois mélanges

Résumé :

Ce travail concerne les grosses flottes automobiles de plus de cent moteurs. La tarification actuelle, d'approche simple, de la garantie RC bénéficie d'une forte personnalisation : il s'agit d'une crédibilité totale pour la charge sinistre écrêtée jusqu'à un certain seuil, et d'une mutualisation pour la sur crête. Un nouveau modèle de tarification, dont l'approche est plus élaborée, est proposée : la sinistralité est modélisée sur le portefeuille des grosses flottes automobiles selon le modèle composé fréquence/coût d'un sinistre. Les sinistres graves sont distingués des sinistres de fréquence. Un modèle de loi mélange de deux lois lognormales est adopté pour le coût d'un sinistre grave. Enfin l'antériorité sinistre est prise en compte sous forme d'un coefficient bonus malus adapté aux grosses flottes pour obtenir la prime pure a posteriori. La seconde partie de nos travaux traite de l'impact du système de conservation (pratique courante chez les grosses flottes automobiles consistant pour l'entreprise à garder à sa charge une franchise annuelle) sur la distribution du ratio sinistres à primes. En effet, l'application d'une conservation a pour conséquence une augmentation de la volatilité des rapports sinistres à primes comparativement au cas où aucune conservation n'est appliquée. Nous souhaitons remédier au risque accru d'obtenir un ratio S/P très défavorable en instaurant une compensation (un taux de chargement) pour cause de conservation. L'approche consiste en un raisonnement en terme d'iso valeurs : nous faisons des hypothèses de chargement en fonction du niveau de conservation appliqué, de manière à ramener le risque d'avoir un ratio S/P très défavorable au même niveau de celui sans conservation au sens d'un indicateur de risque donné.

Keywords: large motor fleets, automobile liability, ratemaking, bonus malus system, annual deductible, loss ratio, risk measures, modelling large claims, mixed distributions

Abstract:

This dissertation deals with large car fleets of one hundred motors and more. The current ratemaking procedure for the automobile liability insurance which approach is simple, benefits of a strong personalization: there is a full credibility for the amount below a truncation point, and a mutualisation for the amount over that certain point. A new ratemaking procedure which approach is more elaborated is presented: the loss is modelled from the portfolio of large car fleets according to the collective risk model. We will distinguish the attritional claims from the large claims. Large claims will be modelled by a mixed distribution of two lognormal distributions. Also the claim history is taken into account through a bonus malus coefficient adapted to the fleets to get the a posteriori rating.

The second part of this work will deal with the impact of an annual deductible (a common practice among large car fleets) on the distribution of the loss ratio. Indeed the application of a deductible implies an increased volatility of the loss ratio compared to the case where none is applied. Our goal is to remedy the increased risk of having a very unfavourable loss ratio by setting a compensation (a charging rate) due to the presence of deductible. The approach consists in reasoning in terms of "isovalues": hypothesis of charging rate are made according to a level of applied deductible so that the risk to have a very unfavourable loss ratio is at the same level than the one without deductible for a given risk indicator.

REMERCIEMENTS

Je remercie chaleureusement mes collègues, qui ont su être présents et m'encourager...

Un grand merci particulièrement à Gilles Mongis pour son encadrement et ses conseils de qualité...

Merci à Claude Jupin pour sa patience et ses commentaires pertinents...

Sommaire

INTRODUCTION	1
 CHAPITRE 1 - PRESENTATION DE COVEA FLEET, DU MARCHE DES FLOTTES AUTOMOBILES ET DU SYSTEME DE CONSERVATION	5
I. Covéa Fleet et le marché des flottes d'entreprises.....	6
II. Les grosses flottes et le système des conservations.....	8
III. Les données	10
 CHAPITRE 2 - MODELISATION DES SINISTRES.....	12
I. Préliminaires	13
1. Le modèle individuel.....	13
2. Le modèle collectif.....	14
II. Modélisation des sinistres RC dits « de fréquence »	15
1. Modélisation du nombre de sinistres.....	16
2. Modélisation des coûts de sinistres	23
III. Modélisation des sinistres RC graves.....	30
1. Translation.....	31
2. « Grands » sinistres et lois mélanges	31
IV .Tests d'adéquation	35
1. Test du χ^2	36
2. Test de Kolmogorov Smirnov	36
3. Test d'Anderson Darling.....	37
4. Test d'adéquation à une famille de lois de probabilité	37
V. Applications numériques.....	38
1. Modélisation des sinistres « de fréquence ».....	39
a) Modélisation de la fréquence	39
b) Modélisation du coût.....	41

2. Modélisation des sinistres graves	44
a) Fréquence d'un sinistre grave	44
b) Modélisation du coût d'un sinistre grave	45
CHAPITRE 3 - ESTIMATION DE LA PRIME PURE D'UNE FLOTTE : DIVERSES APPROCHES	48
I. Détermination de la prime RC Covéa Fleet d'une flotte	49
II. Modèle Bonus-Malus et application aux flottes	51
1. Système de Bonus-Malus appliqué à la famille des lois de Poisson-mélange	52
a) Distribution prédictive conditionnelle $N_{t+1} / N_1 = n_1, \dots, N_t = n_t$	52
b) Prime pure a posteriori et indice de fréquence	53
2. Application du Bonus Malus à une flotte de n véhicules	55
III. Théorie de la fluctuation limitée	57
1. Crédibilité totale	58
2. Application à différentes distributions	58
3. Crédibilité partielle	59
4. Remarques sur la crédibilité américaine	60
IV. Applications numériques	60
1. Quelques précisions	61
2. Cas étudiés	63
3. Commentaires	74
CHAPITRE 4 - IMPACT DE LA CONSERVATION SUR LA DISTRIBUTION DU RATIO S/P	77
I. Problématique	78
II. Les Mesures de risque	81
1. Les mesures de risque retenues	81
a) L'écart type	81
b) La Value-at-Risk (VaR)	81
c) La Tail Value-at-Risk (TVaR)	82
d) Un ratio de performance : le ratio Oméga	82
2. Commentaires sur les mesures de risque	83
III. Méthodes de simulation	83
1. Loi exacte et moments de la charge sinistre	84
2. Application aux distributions de Poisson et Binomiale Négative	85
3. Méthodes d'approximation basées sur les moments	86

IV. Applications numériques.....	89
1. Préambule.....	89
2. Cas étudiés.....	91
a) Value at Risk	91
b) Tail Value at Risk	95
c) Ratio Oméga.....	97
3. Conclusions	100
CONCLUSION.....	101
ANNEXES.....	104
BIBLIOGRAPHIE	107

INTRODUCTION

Présentation des travaux

Dans le cadre de notre étude, nous nous intéressons particulièrement aux grosses flottes automobiles de plus de cent véhicules moteurs.

A la différence des flottes de taille inférieure qui bénéficient d'un tarif segmenté et dont l'approche du risque est collective, les grosses flottes sont traitées de façon plus personnalisée et étudiées cas par cas lors de réunions périodiques par le « comité de souscription et de renouvellement » (collectif regroupant des responsables commerciaux, techniques, de la souscription et d'une cellule dédiée aux gros contrats).

La tarification actuelle de ces grosses flottes repose sur une approche simple bénéficiant d'une forte personnalisation ; elle est basée sur l'observation de la charge sinistre propre de la flotte sur les années antérieures pour les sinistres écrêtés à 30k euros, et sur une mutualisation de la partie des sinistres excédant le seuil d'écèlement sous forme de provision pour graves.

Notre travail consiste à adopter une approche plus élaborée afin de proposer un nouveau modèle de tarification en nous appuyant sur la base de données de l'ensemble du portefeuille des grosses flottes automobiles. Nous modéliserons la fréquence de sinistralité et le coût d'un sinistre pour un véhicule afin d'évaluer la prime pure a priori d'une flotte. Nous apporterons ensuite une personnalisation a posteriori de ces flottes par l'application d'un coefficient qui sera déterminé grâce à la théorie du bonus malus.

Une comparaison entre ces deux approches pour le calcul de la prime pure sera ensuite effectuée dans l'optique d'une consolidation de chacune de ces méthodes.

Précisons que dans notre étude, nous nous concentrerons sur le risque Responsabilité Civile, qui constitue le noyau dur des garanties proposées en flottes (près de 70% des primes).

Nous porterons ensuite une attention particulière au système de conservation qui s'avère être une pratique courante dans le milieu de l'assurance des grosses flottes automobiles. Il s'agit d'un montage dont l'avantage est essentiellement fiscal ; il permet une réduction du budget pour le client par l'exonération des prélèvements obligatoires appliqués aux primes RC automobiles sur la partie conservation gardée à leur charge par les entreprises. Le montage s'interprète comme une décomposition de la prime pure « au premier euro » en :

- une partie conservation correspondant à une avance pour sinistres ; elle est supposée couvrir des sinistres de survenance certaine,
- une partie prime qui prend en charge l'aléa.

L'application de ces conservations a pour conséquence une augmentation de la volatilité des rapports sinistres à primes comparativement au cas où aucune conservation n'est appliquée.

Nous nous proposons d'étudier de plus près l'impact de la mise en place de différents niveaux de conservation sur la distribution du ratio sinistres à primes de la flotte, plus particulièrement sur la queue de distribution, et de normer la prise en compte de cet aléa plus élevé.

Plan de l'étude

Dans une première partie, nous commençons par présenter l'entreprise Covéa Fleet et le marché des flottes sur lequel elle intervient. Nous définissons également quelques notions spécifiques aux flottes automobiles.

Nous détaillons ensuite le système des conservations pour les grosses flottes automobiles et présentons les données qui seront utilisées dans la suite de nos études.

Dans un deuxième temps, nous nous focalisons sur la modélisation de la sinistralité pour un véhicule moteur concernant la garantie Responsabilité Civile dont nous distinguerons la RC Corporelle et la RC Matérielle pour les sinistres de fréquence.

Les variables aléatoires fréquence de sinistres et montant individuel de sinistres sont supposées indépendantes. Il s'agit de déterminer les paramètres et les types de distributions de ces variables aléatoires à partir de la base de données du portefeuille des grosses flottes automobiles.

Pour la modélisation du montant individuel de sinistre, une attention particulière doit être prêtée à la forme de la queue de distribution. Aussi, nous distinguons les sinistres de petite et moyenne intensités auxquels une loi Gamma (où la queue est légère) est ajustée, et ceux de grande intensité où un mélange de lois (de deux lois Lognormales notamment) s'avère plus adéquat.

Les paramètres des distributions des variables aléatoires fréquence et coût déterminés permettent dans la suite d'évaluer la charge sinistre d'une flotte automobile.

Dans la troisième partie, après un rappel de l'approche adoptée par Covéa Fleet pour tarifier une grosse flotte automobile, nous souhaitons modéliser la charge sinistre associée à une flotte automobile. Pour ce faire, nous procédons aux deux étapes suivantes :

- nous nous basons en premier lieu sur les paramètres évalués précédemment pour déterminer la prime pure a priori,
- ensuite, la prise en compte de l'antériorité sinistres de la flotte grâce à la théorie du bonus malus appliquée aux flottes nous permet d'obtenir la prime pure a posteriori

Finalement, nous confrontons les deux approches qui permettent d'obtenir une évaluation de la prime pure d'une flotte automobile connaissant sa sinistralité.

Dans cette partie, nous nous intéressons également au poids à donner à l'expérience dans la tarification et donnons quelques pistes en nous appuyant sur la théorie de la crédibilité de stabilité.

Dans la dernière partie, après avoir exposé la problématique liée à la présence d'une conservation, nous présenterons les différentes mesures de risque que nous serons amenés à utiliser dans nos applications pour évaluer l'impact de la conservation sur la distribution du ratio sinistres à primes. L'objectif dans cette partie est de remédier au risque accru d'obtenir un ratio S/P très défavorable en instaurant une compensation pour cause de conservation. Plus

concrètement, le but recherché est d'évaluer le taux de chargement nécessaire à appliquer à la partie prime en fonction du niveau de conservation retenu.

L'approche consiste en un raisonnement en terme d'iso valeurs : nous faisons des hypothèses de chargement en fonction de la conservation appliquée, de manière à ramener le risque d'avoir un ratio S/P très défavorable au même niveau de celui sans conservation au sens d'un indicateur de risque donné.

Cela suppose les étapes successives suivantes :

- la simulation de la charge sinistre d'une flotte par la méthode de Monte Carlo,
- l'obtention de la distribution du rapport sinistres à primes,
- la détermination des mesures de risque retenues à des seuils définis,
- l'application d'un niveau de conservation et l'obtention de la nouvelle distribution du rapport sinistres à primes : ensuite détermination des valeurs des mêmes mesures de risque aux seuils identiques,
- l'évaluation du taux de chargement à appliquer à la prime pour retrouver les valeurs initiales des mesures de risque.

CHAPITRE 1

-

PRESENTATION DE COVEA FLEET, DU MARCHE DES FLOTTES AUTOMOBILES ET DU SYSTEME DE CONSERVATION

I. Covéa Fleet et le marché des flottes d'entreprises

Le marché des flottes automobiles constitue une composante importante de l'assurance automobile. A la fin de l'année 2006, les flottes représentaient 37,22% du nombre des immatriculations qui était de 2 440 530. Côté assurance, en 2006, le montant global des cotisations automobiles a atteint 17,8 Md d'euros dont 3,2 Md pour les flottes.

Le marché de l'assurance flotte est jugé extrêmement concurrentiel, ce qui pèse sur les prix. En 2006, les grosses flottes ont fait l'objet de nombreux appels d'offres et des affaires ont changé de mains. Depuis 2004, année où le marché est entré dans un cycle baissier, les acteurs se livrent une guerre des tarifs acharnée. Les marges de manœuvre se sont réduites pour les assureurs. La situation est ainsi propice à exercer une pression pour des montages avec conservation pour les grosses flottes automobiles, de façon à proposer une tarification plus compétitive aux clients.

Covéa Fleet est un acteur incontournable du secteur d'activité des flottes automobiles. Créée en 2001, elle résulte du rapprochement de MMA, MAAF Assurances, SMABTP (depuis 2002), l'Auxiliaire (depuis 2005) et AZUR (dont le portefeuille sera transféré en 2008). C'est la seule société en France dédiée à l'assurance des flottes d'entreprises, du transport et des marchandises transportées.

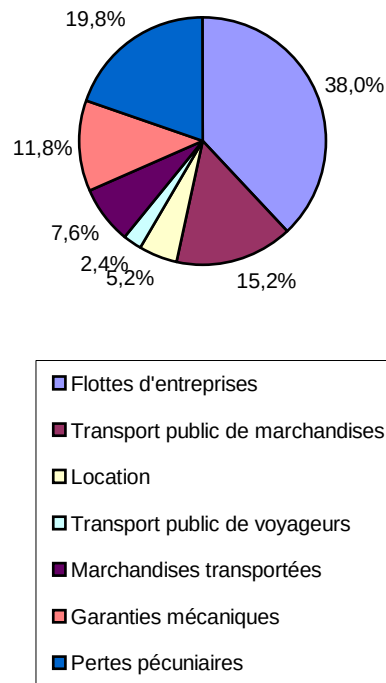
De 2001 à 2006, le chiffre d'affaires de Covéa Fleet n'a cessé d'augmenter, passant de 245 millions d'euros à 573 millions d'euros. Cette augmentation est due, en grande partie, à l'intégration des portefeuilles des partenaires.

Covéa Fleet propose une large gamme de produits qui s'articule autour de trois pôles :

- Flottes d'entreprises, Covéa Fleet assure les risques de circulation des flottes d'entreprises ainsi que ceux liés aux transports publics de marchandises et de voyageurs,
- Marchandises transportées, qui concerne les entreprises propriétaires de leurs marchandises ainsi que les professionnels transportant des marchandises pour compte de tiers et ce en tous points du globe,
- Financement et garanties mécaniques, qui répond aux besoins particuliers de clients grands comptes, établissements financiers, loueurs longue durée, professionnels de l'automobile et autres groupements.

Le chiffre d'affaires de Covéa Fleet se répartit selon les différents segments de marché dans les proportions suivantes en 2006 :

**Ventilation du chiffre d'affaires
par segments de marché**



Notion de flottes d'entreprises

Deux types de flottes sont généralement distingués :

- Les flottes naturelles sont constituées d'un ensemble de véhicules appartenant ou exploités par un même propriétaire ou entité juridique ayant souscrit un contrat collectif pour la couverture globale de son parc. Ainsi tous les véhicules sont soumis aux mêmes règles tarifaires, les primes de chacun d'eux sont recouvrées en une seule fois et les conditions du contrat sont applicables indistinctement à tous.
- Les flottes artificielles correspondent au regroupement « mutualisé » de contrats automobiles couvrant des clients distincts d'un prescripteur ayant les mêmes besoins en terme d'assurance, chacun acquittant la prime relative à son véhicule.

A la différence d'une flotte naturelle, la flotte artificielle relève plus d'un simple rassemblement de mono véhicules. Nous nous intéressons dans la suite uniquement aux flottes naturelles, la tarification des flottes artificielles se rapprochant plus de ceux relatifs au marché des particuliers.

Seules les flottes d'entreprise comprenant au moins cinq moteurs (sauf pour le TPM, TPV et la location) constituent la cible de Covéa Fleet. Nous distinguons quatre segments de flottes (tailles de flottes) à Covéa Fleet pour les activités suivantes :

	TPPC (transport privé pour propre compte)	TPM (transport public de marchandises)	TPV (transport public de voyageurs)	Location
Segment 1	de 5 à 10 véhicules moteurs	de 1 à 5 véhicules moteurs	de 1 à 5 véhicules moteurs	de 1 à 5 véhicules moteurs
Segment 2	de 11 à 39 véhicules moteurs	de 6 à 19 véhicules moteurs	de 6 à 19 véhicules moteurs	de 6 à 19 véhicules moteurs
Segment 3	de 40 à 100 véhicules moteurs	de 20 à 100 véhicules moteurs	de 20 à 100 véhicules moteurs	de 20 à 100 véhicules moteurs
Segment 4	> 100 véhicules moteurs	> 100 véhicules moteurs	> 100 véhicules moteurs	> 100 véhicules moteurs

Le segment 4 qui nous intéresse dans ce projet regroupe les grosses flottes automobiles de plus de cent véhicules moteurs.

II. Les grosses flottes et le système des conservations

La conservation annuelle fixe est un montage spécifique aux grosses flottes automobiles largement pratiqué sur le marché pour une raison essentiellement fiscale. En effet, le législateur n'a pas décidé de répercuter la totalité des prélèvements obligatoires appliqués aux primes RC automobiles (33,1% pour les voitures particulières et les véhicules utilitaires légers, 15,10% pour les plus de 3,5T) sur les franchises RC conservées à leur charge par les entreprises. Ces dernières restent néanmoins soumises à la contribution au fonds de garantie automobile (FGA) dont le taux s'élève à 5%. Dans la pratique ce règlement se fait directement par l'assuré au FGA.

En réduisant la partie de la prime soumise à la fiscalité pesant sur l'assurance automobile, il est donc proposé à l'entreprise un budget plus compétitif.

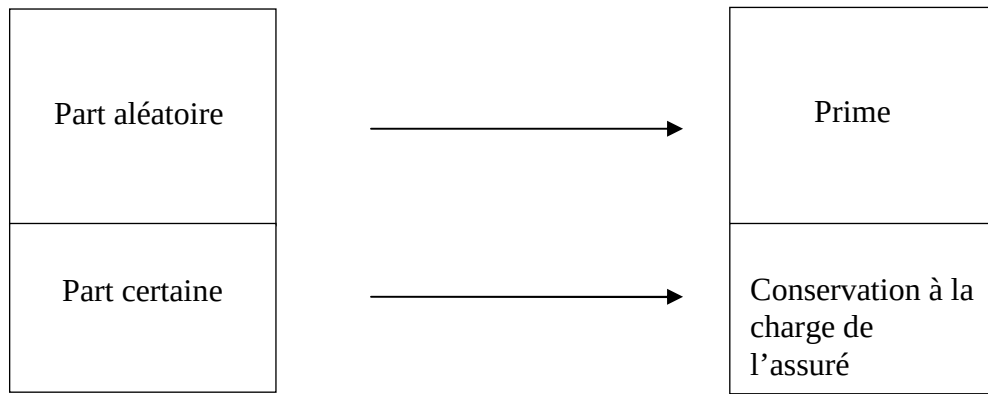
L'instauration de cette conservation permet également une moralisation du risque. L'entreprise a un intérêt direct à agir sur sa sinistralité car en cas de baisse de sa charge sinistre, elle peut récupérer une partie de la conservation. Du point de vue de l'assureur, il aura intérêt bien évidemment à ce que le niveau de rétention soit fixé assez bas pour que cette dernière corresponde bien à la part certaine du risque de façon à ne pas avoir à rétrocéder la conservation non consommée.

Nous considérons qu'à partir d'une certaine taille de flotte il y a un seuil de sinistralité incompressible et certain que l'assuré, pour les raisons fiscales évoquées ci-dessus, peut souhaiter garder à sa charge.

Ainsi le montage se décompose en :

- la part certaine de sinistralité qui sera à la charge de l'assuré : la conservation
- la partie aléatoire du risque qui correspond à la prime

D'un point de vue schématique, le montage consiste en :



« Charge de sinistre annuelle »

« Budget annuel »

Le budget total de l'assuré se compose donc d'une rétention à sa charge et d'une prime qui permet de couvrir la sinistralité excédentaire aléatoire.

Après études des statistiques des flottes (le procédé de tarification actuel sera détaillé ultérieurement au chapitre 3) lors des réunions périodiques du collectif pour la souscription et le renouvellement, un seuil de rétention peut être retenu et les règles de mise en place d'une conservation sont les suivantes :

- La prime pure totale RC doit être supérieure à 75 000 euros,
- Le montant de la conservation doit être significativement inférieur à la sinistralité récurrente,
- Le montant de la conservation doit être égal au maximum à la moitié de la prime pure totale.

Concernant la fréquence de l'instauration d'une conservation, sur l'ensemble des contrats segments 4, nous pouvons constater qu'elle est appliquée plus d'une fois sur trois.

Dans la pratique, la gestion de la conservation est incombée totalement à l'assureur qui perçoit des frais de gestion. Une convention de gestion relevant du droit commercial est établie entre l'assuré et l'assureur et décrit les prestations contractuelles de la conservation instaurée.

Notons qu'une conservation peut de la même manière être appliquée aux garanties dommages. La logique est identique à celle de la garantie RC, bien que son instauration soit moins « profitable » en raison d'un taux de taxe moindre que celui de la RC.

Exemple d'une flotte avec conservation :

Nous présentons ci-dessous un exemple de tarification d'une flotte de véhicules de première catégorie (véhicules de moins de 3,5T) pour montrer l'intérêt pour le client d'une conservation. Soit une flotte dont la prime pure totale est de 200 000 euros.

Une tarification sans conservation se décompose ainsi, en supposant par exemple les taux suivants (taux que nous conserverons dans la suite de l'étude) :

- Frais généraux : 19%
- Commissions : 15%

Sa cotisation H.T. s'élève alors à : $200\,000 / 0,66 = 303\,030$ euros

Sa cotisation TTC est de : $303\,030 * 1,331 = \mathbf{403\,333}$ euros

Une tarification avec une conservation de 100 000 euros se décompose en :

- une cotisation H.T. égale à $100\,000 / 0,66 = 151\,515$ euros
- une cotisation TTC de $151\,515 * 1,331 = 201\,666$ euros
- une conservation = 100 000 euros
- les honoraires pour la gestion de la conservation : $100\,000 * 0,14 = 14\,000$
- la contribution au FGA : $100\,000 * 0,05 = 5\,000$

d'où le budget TTC s'élève à **320 666** euros, soit un budget minoré de 20,5% par rapport à une tarification sans conservation.

Nous soulignons que les enjeux financiers relatifs aux conservations sont importants. En effet, sur l'année 2006, l'encaissement des conservations s'élevait à 42 millions d'euros pour Covéa Fleet (dont 33 millions provenaient de MMA/MAAF et 9 millions de la SMABTP).

III. Les données

Lorsque la base de données de notre étude a été constituée dans le courant de l'année 2005, nous avions à notre disposition uniquement les données issues des systèmes d'informations (SI) correspondant aux enseignes MMA et MAAF.

Nous, avons par la suite, restreint notre extraction de données aux :

- flottes automobiles du segment de marché TPPC (Transport Pour Propre Compte) ; en effet notre étude porte sur ce segment de marché qui représente environ les 3/4 de la totalité des contrats du segment 4 dans notre portefeuille
- parcs désignés (les immatriculations de tous les véhicules de la flotte sont connues) afin de pouvoir associer à chaque véhicule les sinistres correspondant dans la partie modélisation

Nous porterons également notre intérêt plus particulièrement aux flottes de 1^{ère} catégorie de façon à travailler sur une classe homogène :

ainsi notre partie modélisation des sinistres se focalisera uniquement sur la sinistralité RC d'un véhicule de 1^{ère} catégorie utilisé dans le cadre d'une activité de transport pour propre compte.

Des précisions supplémentaires sur les données sur lesquelles les modélisations se baseront, seront données ultérieurement lors des applications numériques, dans chaque partie concernée.

CHAPITRE 2

-

MODELISATION DES SINISTRES

Dans ce chapitre, après avoir présenté le modèle collectif, nous nous intéressons à la modélisation des données sinistres pour déterminer les lois s'adaptant le mieux à la fréquence de sinistralité et au coût des sinistres du risque RC.

Nous distinguerons les sinistres « de fréquence » dont les montants sont de faible intensité et les sinistres graves dont l'évènement est plus rare mais l'impact beaucoup plus lourd en terme de coût.

Nous présenterons notamment la modélisation des sinistres graves par une loi mélange de deux distributions.

Différents tests statistiques qui seront appliqués pour mesurer la qualité de l'ajustement du modèle retenu à ces données sont présentés.

Les applications numériques qui consistent en la détermination des paramètres des lois de distribution pour la garantie RC Corporelle, la garantie RC Matérielle et les sinistres graves RC clôtureront ce chapitre.

I. Préliminaires

Il existe deux principales approches en assurance dommage pour modéliser la sinistralité d'un ensemble de contrats (éventuellement réduit à un unique contrat) selon le type d'information disponible et recherché par l'assureur : l'approche individuelle et l'approche collective.

Nous présenterons ci-dessous les deux modèles, l'approche collective étant le point de vue adopté dans la suite de ce rapport.

Les résultats, qui seront obtenus à la fin de ce chapitre suite à l'application des modèles de fréquence et du coût présentés dans cette partie, seront utilisés dans les chapitres suivants pour modéliser et simuler la charge sinistre annuelle d'une flotte dans le cadre de ce modèle dit composé.

1. Le modèle individuel

Considérons une classe C de K contrats. Soit S la v.a.r. charge sinistres de la classe C au cours de la période considérée (souvent annuelle).

Pour $k = 1, \dots, K$, S_k désigne la v.a.r. montant cumulé des sinistres du $k^{\text{ième}}$ contrat pendant la période considérée.

Dans le modèle individuel, S est donnée par :

$$S = \sum_{k=1}^K S_k$$

La sinistralité de chaque contrat est donc connu individuellement en fin de période et la distribution et les principales caractéristiques de la charge sinistre totale S sont déduites à partir de celles de $(S_1, \dots, S_k)$ supposées connues.

2. Le modèle collectif

Dans l'approche collective, les contrats ne sont pas distingués et les charges sinistres individuelles ne sont pas forcément connues. La charge sinistres S de la classe C est déterminée à l'aide du nombre total de sinistres de la classe et des montants de chacun de ces sinistres. Il s'agit du modèle composé. C'est le modèle de base de l'assurance non-vie.

Le montant cumulé des sinistres S de la classe C s'écrit donc :

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

où :

- N est la v.a.r. (à valeurs entières) nombre ou fréquence de sinistres survenant à la classe C lors de la période considérée,
- $(X_i)_{i \geq 1}$ est la suite des v.a.r. montants individuels de sinistre (dans l'ordre chronologique),
- $X_0=0$ par notation.

Nous ferons également les hypothèses suivantes :

- i. Indépendance de la fréquence et des coûts.

Les v.a.r. N et $(X_i)_{i \geq 1}$ sont indépendantes.

- ii. Indépendance et équidistribution des coûts de sinistre.

Les coûts des sinistres successifs $X_1, X_2, \dots, X_i, \dots$ sont indépendants et de même loi que la v.a.r. parente X .

Sous ces hypothèses et si N et X admettent des montants d'ordre 1, nous avons :

$$E(S) = E(N).E(X)$$

II. Modélisation des sinistres RC dits « de fréquence »

Nous étudierons dans cette partie les sinistres de petite et moyenne intensité que nous différencions des sinistres lourds dont le traitement sera effectué dans la partie suivante. L'entreprise a pour tradition d'utiliser un seuil de 75K euros pour différencier ces deux types de sinistres (pour le calcul des réserves, pour la réassurance). Nous garderons ce même seuil de séparation pour nos études.

La fréquence de sinistres, qui est la première composante du montant cumulé des sinistres S dans le modèle composé, est analysée en premier lieu. Ensuite la deuxième composante de S , le coût de sinistres, sera étudiée. Les lois usuelles pour la modélisation de chacune de ces composantes sont présentées dans la suite de ce chapitre. Le lecteur pourra consulter l'ouvrage de Besson et Partrat (2005) pour de plus amples détails sur ces aspects.

Définition et Notation

Pour une v.a.r. X à valeurs entières, le moment factoriel d'ordre k est défini par :

$$\mu_{(k)} = E[X(X-1)\dots(X-k+1)]$$

Nous notons le moment simple d'ordre k d'une v.a.r. X : $m_k = E(X^k)$.

Les moments simples s'expriment en fonction des moments factoriels, ainsi nous avons :

$$m_1 = \mu_{(1)}$$

$$m_2 = \mu_{(2)} + \mu_{(1)}$$

$$m_3 = \mu_{(3)} + 3\mu_{(2)} + \mu_{(1)}$$

$$m_4 = \mu_{(4)} + 6\mu_{(3)} + 7\mu_{(2)} + \mu_{(1)}$$

Nous appelons moment centré d'ordre k de la v.a.r. X et nous notons μ_k :

$$\mu_k = E[(X - E(X))^k]$$

Quelques moments particuliers :

Les moments centrés d'ordre k se déduisent des moments simples par les expressions suivantes :

m_1 est la moyenne.

$\mu_2 = m_2 - m_1^2$ est la variance.

$$\mu_3 = m_3 - 3m_1m_2 + 2m_1^3$$

$$\mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4$$

μ_2, μ_3 et μ_4 sont utilisés pour caractériser la forme d'une distribution.

Nous définissons :

le coefficient d'asymétrie (skewness) : $\gamma_1 = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}$

le coefficient d'aplatissement (kurtosis) : $\gamma_2 = \frac{\mu_4}{(\mu_2)^2} - 3$

1. Modélisation du nombre de sinistres

Nous présentons ci-dessous les caractéristiques des trois lois usuelles pour modéliser une fréquence de sinistres, en faisons quelques commentaires sur leurs utilisations dans les cas pratiques et exposons les méthodes pour déterminer les paramètres de ces lois.

- Distribution Binomiale $B(n, p)$

Elle a pour paramètres n dans N^* et $p \in]0;1[$; nous poserons $q = 1 - p$.

Loi :

$$p_k = P(N = k) = C_n^k p^k q^{n-k} \text{ pour } k = 0, 1, \dots, n \text{ et } P(N = k) = 0 \text{ sinon}$$

Moments :

$$\mu_{(k)} = n(n-1)\dots(n-k+1)p^k$$

ce qui donne

$$m = E(N) = np, \quad \sigma^2 = V(N) = npq \quad (< E(N)),$$

$$\gamma_1 = \frac{q-p}{\sqrt{npq}},$$

$$\gamma_2 = \frac{1-6pq}{npq}$$

Semi additivité :

Soient deux v.a.r. indépendantes N_1 et N_2 suivant respectivement les lois $B(n_1, p)$ et $B(n_2, p)$, alors la v.a.r. $N_1 + N_2$ suit la loi $B(n_1 + n_2, p)$.

- Distribution de Poisson $P(\lambda)$

Elle a pour paramètre $\lambda > 0$

Loi :

$$p_n = P(N = n) = e^{-\lambda} \frac{\lambda^n}{n!} \text{ pour } n \in \mathbb{N}$$

Moments :

$\mu_{(k)} = \lambda^k$ d'où pour les premiers moments

$$m = E(N) = \lambda, \quad \sigma^2 = V(N) = \lambda \quad (= E(N)),$$

$$\gamma_1 = \frac{1}{\sqrt{\lambda}},$$

$$\gamma_2 = \frac{1}{\lambda}$$

Additivité :

Soient deux v.a.r. indépendantes N_1 et N_2 suivant respectivement les lois $P(\lambda_1)$ et $P(\lambda_2)$ alors la v.a.r. $N_1 + N_2$ suit la loi $P(\lambda_1 + \lambda_2)$.

- Distribution Binomiale Négative $BN(r, p)$

Elle a pour paramètres $r > 0$ et $p \in]0,1[$; nous poserons $q = 1 - p$.

Loi :

$$p_n = P(N = n) = \frac{\Gamma(r+n)}{\Gamma(r)n!} p^r q^n, \text{ pour } n \in \mathbb{N}$$

avec Γ la fonction Gamma définie par $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ vérifiant :

$$\frac{\Gamma(r+n)}{\Gamma(r)} = (r+n-1)\dots r \text{ si } n \geq 1 \text{ et } 1 \text{ si } n = 0$$

Moments :

$$\mu_{(k)} = r(r+1)\dots(r+k-1) \left(\frac{q}{p}\right)^k \text{ d'où nous en déduisons}$$

$$m = E(N) = r \frac{q}{p}, \quad V(N) = r \frac{q}{p^2} \quad (> E(N)),$$

$$\gamma_1 = \frac{1+q}{\sqrt{rq}},$$

$$\gamma_2 = \frac{p^2 + 6q}{rq}$$

Semi additivité :

Soient deux v.a.r. indépendantes N_1 et N_2 suivant respectivement les lois $BN(r_1, p)$ et $BN(r_2, p)$ alors la v.a.r. $N_1 + N_2$ suit la loi $BN(r_1 + r_2, p)$.

Rappel :

La loi Binomiale Négative est un cas particulier de la famille des lois de Poisson mélange.

Soit la v.a.r. N fréquence de sinistres d'une police appartenant à une classe de risques, elle est distribuée selon une loi de Poisson $P(\lambda)$, λ étant la réalisation d'une v.a.r. $\Lambda > 0$ dite variable de structure de la classe.

Soit h la densité de Λ , dépendant d'un paramètre $\theta \in \Theta$ ouvert de $\Re^q$, la loi de N est donnée par :

$$P(N = n) = \int_0^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} h(\lambda) d\lambda \quad (n \text{ entier})$$

N suit une loi de Poisson-Gamma si la variable de structure Λ est distribuée selon une loi Gamma $\gamma(r, \alpha)$ de densité, avec $r, \alpha > 0$:

$$h(\lambda) = \frac{\alpha^r}{\Gamma(r)} \lambda^{r-1} e^{-\lambda\alpha} \text{ si } (\lambda > 0),$$

$$\text{avec } E(\Lambda) = \frac{r}{\alpha}, \quad V(\Lambda) = \frac{r}{\alpha^2}$$

la loi de N est alors une loi Binomiale Négative $BN\left(r, \frac{\alpha}{\alpha+1}\right)$

- Commentaires sur l'utilisation des différentes lois

La distribution binomiale $B(n, p)$ est appliquée le plus souvent lorsque le paramètre n (qui borne la v.a.r. fréquence de sinistres) est donné. Le fait que $E(N) > V(N)$ rend l'utilisation de cette loi plus difficile car le plus souvent c'est l'inégalité inverse qui est rencontrée dans les données empiriques en assurance non vie.

La loi de Poisson $P(\lambda)$ est la loi fondamentale de l'assurance non vie pour tous les avantages qu'elle présente dans l'aspect pratique : unique paramètre facilement compréhensible et estimable, elle est à la base des distributions Poisson-mélange et de l'élaboration d'un système Bonus-Malus, ...

Elle est cependant moins flexible en raison de son unique paramètre, elle présente une queue de distribution légère, et l'égalité $E(N) = V(N)$ peut être contraignante.

Dans le cas de la distribution Binomiale Négative $BN(r, p)$, $V(N) > E(N)$, ce qui la rend facilement utilisable dans la pratique. Elle est donc plus flexible que la loi de Poisson grâce à ses deux paramètres. Enfin, elle permet de modéliser naturellement la distribution de la fréquence de sinistres d'un contrat d'une classe hétérogène de risques.

- Procédures d'estimation des paramètres

Nous présentons trois méthodes d'estimation que nous utiliserons dans nos applications.

Soient X une v.a.r. de loi $P_X \in P_0 = \{P_0^\theta : \theta \in \Theta\}$ de paramètre naturel $\theta \in \Theta$ ouvert de $\mathbb{R}^p$, $\underline{X} = (X_1, \dots, X_n)$ un n -échantillon de X prenant ses valeurs dans l'espace des observations X et $\pi(\theta)$ un paramètre q -vectoriel à estimer.

Estimer le paramètre $\pi(\theta)$ consiste pour chaque réalisation $\underline{x}$ dans X à donner une valeur estimée de la vraie valeur inconnue $\pi(\theta_v)$ de ce paramètre.

Méthode des Moments

Il s'agit d'exprimer le plus simplement possible $\pi(\theta)$ en fonction des moments simples théoriques de X soit :

$$\pi(\theta) = \varphi(m_1(\theta), \dots, m_r(\theta))$$

L'estimateur de $\pi(\theta)$ donné par la méthode des moments (emm) noté $T(\underline{X})$ est tel que :

$$T(\underline{X}) = \varphi(M_1, \dots, M_r) \text{ avec } M_k = \frac{1}{n} \sum_{i=1}^n X_i^k \text{ le moment empirique d'ordre } k$$

Méthode du Maximum de Vraisemblance

C'est une méthode fondamentale pour la recherche d'estimateurs.

Soient X une v.a.r. de loi $P_X \in P_0 = \{P_0^\theta : \theta \in \Theta\}$ où Θ est un ouvert de $\mathbb{R}^p$, et $\underline{X} = (X_1, \dots, X_n)$ un n -échantillon de X prenant ses valeurs dans l'espace des observations X .

La fonction de vraisemblance associée à l'échantillon $\underline{X}$ est définie comme étant la fonction :

$$L : X \times \Theta \rightarrow \mathbb{R}_+ \\ (\underline{x}; \theta) \rightarrow L(\underline{x}; \theta) = \prod_{i=1}^n P_0^\theta(X = x_i)$$

Dans le cas où X est une v.a.r. absolument continue de densité $f_0^\theta(x)$, la fonction de vraisemblance s'écrit :

$$L(\underline{x}, \theta) = \prod_{i=1}^n f_0^\theta(x_i)$$

L'idée est de déterminer l'estimateur $\hat{\theta}$ pour lequel la vraisemblance avec laquelle est observé $\underline{x}$ soit le maximum.

Définition :

On appelle estimateur du maximum de vraisemblance (emv) de θ un estimateur $\hat{\theta}(\underline{X})$ tel que pour $\underline{x} \in X$

$$L(\underline{x}; \hat{\theta}(\underline{x})) \geq L(\underline{x}; \theta) \quad \text{pour tout} \quad \theta \in \Theta$$

Lorsque pour $\underline{x} \in X$, $L(\underline{x}; \theta)$ admet des dérivées partielles premières, l'emv $\hat{\theta}$ de θ est solution des équations de vraisemblance :

$$\frac{\partial \ln L(\underline{x}; \theta)}{\partial \theta_k} = 0, \quad k = 1, \dots, p$$

Minimisation de la distance de Cramer-von Mises

Une autre approche pour l'estimation des paramètres de lois est l'utilisation d'une mesure de distance qu'il faut minimiser. Nous retenons plus particulièrement la distance de Cramer-von Mises.

Soient une f.r. G donnée, continue, et la f.r. F de C . F étant cependant inconnue, la fonction empirique F_n est substituée à F . La distance de CVM entre F_n et G est définie par :

$$d_{wF_n}(G) = \int_0^{+\infty} [F_n(x) - G(x)]^2 dG(x)$$

Hogg et Klugman (1984) suggèrent d'utiliser l'expression suivante. L'avantage est d'avoir une pondération plus lourde aux bords de la distribution.

L'expression opérationnelle est :

$$\sum_{i=1}^n w_i [F_n(x_i) - G(x_i)]^2$$

avec $w_i = \frac{n}{\frac{n}{n+1} F_n(x_i) \left(1 - \frac{n}{n+1} F_n(x_i) \right)}$

Applications : formules pour les emm et emv

- Loi Binomiale $B(n, p)$

Soit $(n_1, \dots, n_k)$ la réalisation d'un k -échantillon $(N_1, \dots, N_k)$ de N .

n est généralement connu, et les estimations de p par la méthode des moments et la méthode du maximum de vraisemblance donnent le même estimateur :

$$\hat{p} = \frac{1}{kn} \sum_{i=1}^k n_i$$

- Loi Poisson $P(\lambda)$

Soit $(n_1, \dots, n_k)$ la réalisation d'un k -échantillon $(N_1, \dots, N_k)$ de N .

La méthode des moments et la méthode du maximum de vraisemblance fournissent le même estimateur $\hat{\lambda} = \bar{n}$ de λ (avec $\bar{n}$ moyenne empirique de l'échantillon).

- Loi Binomiale Négative $BN(r, p)$

Soit $(n_1, \dots, n_k)$ la réalisation d'un k -échantillon $(N_1, \dots, N_k)$ de N .

La méthode des moments donne les estimateurs suivants (nous notons $\hat{\sigma}^2$ la variance empirique de l'échantillon) :

$$\tilde{r} = \frac{\bar{n}^2}{\hat{\sigma}^2 - \bar{n}} \quad \text{et} \quad \tilde{p} = \frac{\bar{n}}{\hat{\sigma}^2} \quad \text{avec} \quad \hat{\sigma}^2 > \bar{n}$$

La méthode du maximum de vraisemblance donne le résultat suivant :

Soient $l = \max\{n_1, \dots, n_k\}$, $k_j =$ nombre des n_i égaux à j , avec $j = 0, \dots, l$,

$f_j = \frac{k_j}{k}$ et $\sum_{\alpha=0}^j f_\alpha$ la proportion des n_i inférieurs ou égaux à j .

Proposition :

Si $\hat{\sigma}^2 > \bar{n}$, alors l'emv de (r, p) noté $(\hat{r}, \hat{p})$ est donné par $\hat{p} = \frac{\hat{r}}{\hat{r} + \bar{n}}$

et $\hat{r}$ est l'unique solution positive de l'équation en r

$$\sum_{j=1}^l f_j \left(\frac{1}{r+j-1} + \dots + \frac{1}{r} \right) - \ln \left(1 + \frac{\bar{n}}{r} \right) = 0$$

La démonstration de cette proposition est détaillée dans Simonsen (1979)

Cette équation se résout numériquement avec une valeur initiale de r fournie par la méthode des moments (voir partie applications numériques).

2. Modélisation des coûts de sinistres

Il s'agit de modéliser le montant individuel de sinistre, la deuxième composante du modèle collectif. Nous présentons les distributions continues usuelles pour la modélisation des coûts et faisons quelques commentaires sur leurs utilisations.

Les procédures d'estimation des paramètres sont identiques à celles de la modélisation de la fréquence, elles ne seront par conséquent pas présentées à nouveau dans ce paragraphe, mais nous indiquerons les formules de détermination des paramètres pour chacune des lois.

Nous présenterons également le cas de mélange de distributions qui est utile pour modéliser un montant individuel de sinistre lorsque différents types de sinistres sont en présence.

Dans la suite nous appelons C la v.a.r. positive coût de sinistre.

- Distribution exponentielle $\mathcal{E}(\theta)$

Elle dépend d'un unique paramètre $\theta > 0$.

Fonction de répartition et densité :

$$f(x) = \theta e^{-\theta x} \text{ et } F(x) = 1 - e^{-\theta x}$$

Moments :

$$m_k = \frac{k!}{\theta^k} \text{ d'où}$$

$$m = E(C) = \frac{1}{\theta}, \quad V(C) = \frac{1}{\theta^2},$$

$$\gamma_1 = 2,$$

$$\gamma_2 = 6$$

- Distribution LogNormale $\text{LogN}(\mu, \sigma^2)$

Elle a pour paramètres $\mu \in \mathbb{R}$ et $\sigma > 0$.

La v.a.r. coût C suit une loi LogNormale $\text{LogN}(\mu, \sigma^2)$ si et seulement si la v.a.r. $\ln C$ suit la loi Normale $N(\mu, \sigma^2)$.

Fonction de répartition et densité :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(\ln x - \mu)^2} \quad \text{et} \quad F(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$$

avec Φ fonction de répartition de la loi normale centrée réduite.

Moments :

$$m_k = e^{k\mu + \frac{k}{2}\sigma^2} \quad \text{ainsi,} \quad \text{en posant} \quad \omega = e^{\sigma^2}$$

$$m = E(C) = e^{\mu + \frac{\sigma^2}{2}} = e^{\mu} \omega^{1/2}, \quad V(C) = e^{2\mu} \omega(\omega - 1),$$

$$\gamma_1 = (\omega - 1)^{1/2}(\omega + 2),$$

$$\gamma_2 = \omega^4 + 2\omega^3 + 3\omega^2 - 6$$

- Distribution Gamma $\gamma(\alpha, \beta)$

Elle a pour paramètres α (la forme) et β (l'échelle) strictement positifs.

Fonction de répartition et densité :

$$f(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad \text{et} \quad F(x) = \int_0^x \beta^\alpha x^{\alpha-1} e^{-\beta x} dx / (\Gamma(\alpha)) \quad \text{avec la fonction Gamma } \Gamma \text{ définie par}$$

$$\Gamma(\nu) = \int_0^{+\infty} e^{-t} t^{\nu-1} dt \quad \text{pour } \nu > 0.$$

Moments :

$$m_k = \frac{\alpha(\alpha+1)\dots(\alpha+k-1)}{\beta^k}, \quad \text{ainsi}$$

$$m = E(C) = \frac{\alpha}{\beta}, \quad V(C) = \frac{\alpha}{\beta^2},$$

$$\gamma_1 = \frac{2}{\sqrt{\alpha}},$$

$$\gamma_2 = \frac{6}{\alpha}$$

Semi additivité :

Soient deux v.a.r. indépendantes C_1 et C_2 suivant respectivement les lois $\gamma(\alpha_1, \beta)$ et $\gamma(\alpha_2, \beta)$ alors la v.a.r. $C_1 + C_2$ suit la loi $\gamma(\alpha_1 + \alpha_2, \beta)$.

- Distribution Weibull $W(\tau, \alpha)$

Elle a pour paramètres τ (la forme) et α (l'échelle) strictement positifs.

Fonction de répartition et densité :

$$f(x) = \frac{\tau}{\alpha^\tau} x^{\tau-1} e^{-\left(\frac{x}{\alpha}\right)^\tau} \text{ et } F(x) = 1 - e^{-\left(\frac{x}{\alpha}\right)^\tau}$$

Moments :

$$m_k = \alpha^k \Gamma\left(\frac{k}{\tau} + 1\right), \text{ ainsi}$$

$$m = E(C) = \alpha \Gamma\left(\frac{1}{\tau} + 1\right), \quad V(C) = \alpha^2 \left[\Gamma\left(\frac{2}{\tau} + 1\right) - \Gamma^2\left(\frac{1}{\tau} + 1\right) \right]$$

- Distribution Pareto $P(a, \alpha)$

Elle a pour paramètres a (souvent donné, niveau d'une franchise par exemple) et α strictement positifs.

Fonction de répartition et densité :

$$f(x) = \frac{\alpha a^\alpha}{x^{\alpha+1}} 1(x)_{]a, +\infty[} \text{ et } F(x) = \left[1 - \left(\frac{a}{x} \right)^\alpha \right] 1(x)_{]a, +\infty[}$$

Moments :

$$m_k = \frac{\alpha a^k}{\alpha - k} \text{ si } \alpha > k \quad (m_k = +\infty \text{ sinon})$$

$$m = \frac{\alpha a}{\alpha - 1} \text{ si } \alpha > 1, \quad V(C) = \frac{\alpha a^2}{(\alpha - 1)^2 (\alpha - 2)} \text{ si } \alpha > 2,$$

$$\gamma_1 = \frac{2(\alpha + 1)}{\alpha - 3} \sqrt{\frac{\alpha - 2}{\alpha}} \text{ si } \alpha > 3$$

- Distribution Normale $N(\mu, \sigma^2)$

La symétrie de cette loi ne lui permettra pas de modéliser les montants individuels de sinistre, mais elle nous sera utile lors de la modélisation de la charge sinistre.

Elle a pour paramètres $\mu \in \mathbb{R}$ et $\sigma > 0$.

Fonction de répartition et densité :

$f(x) = \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right)$ avec φ la densité Normale standard $N(0,1)$ définie par :

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \text{ et}$$

$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$ où Φ est la fonction de répartition de la loi Normale standard d'expression intégrale :

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

Moments :

$$\mu_k = (k-1)(k-3)\dots 3.1.\sigma^2 \text{ si } k \text{ est pair, } 0 \text{ si } k \text{ est impair}$$

$$m = \mu, \quad V(C) = \sigma^2,$$

$$\gamma_1 = \gamma_2 = 0$$

- Commentaires sur l'utilisation des différentes lois

La loi exponentielle $\varepsilon(\theta)$ est très simple d'utilisation mais présente une queue de distribution très légère, ce qui ne lui permet pas la modélisation des coûts de la plupart des garanties en assurance non-vie.

La distribution LogNormale $\text{LogN}(\mu, \sigma^2)$ est la plus utilisée en assurance non-vie et en réassurance. Nous l'utiliserons notamment pour la modélisation des sinistres graves sous la forme d'un mélange de deux lois.

Une autre distribution classique en assurance non-vie est la Gamma $\gamma(\alpha, \beta)$ qui s'adapte efficacement à la modélisation des lois dont la queue de distribution n'est pas trop épaisse.

La loi Weibull $W(\tau, \alpha)$ est en revanche moins utilisée en assurance non-vie à cause des expressions des moments à l'aide de la fonction Γ et une estimation des paramètres qui s'avère plus difficile.

La distribution de Pareto $P(a, \alpha)$ est la loi privilégiée en réassurance en raison de sa queue de distribution lourde. Cependant l'existence des moments n'est pas certaine.

La loi Normale, comme dit précédemment, ne présente pas d'utilité lorsque la modélisation du montant individuel de sinistre est souhaitée en raison de sa symétrie, mais sera utilisée pour la modélisation de la charge sinistre.

- Détermination des paramètres : formules des emm et emv

Les procédures d'estimation des paramètres sont les mêmes que celles décrites dans le paragraphe précédent. Nous donnons pour chacune des lois les formules des emm et emv.

- Loi exponentielle $\varepsilon(\theta)$

Soit $(c_1, \dots, c_n)$ la réalisation d'un n -échantillon $(C_1, \dots, C_n)$ de C . Les méthodes des moments et du maximum de vraisemblance donnent le même estimateur $\hat{\theta} = \frac{1}{\bar{c}}$ de θ (avec $\bar{c}$ moyenne empirique de l'échantillon).

- Loi LogNormale $\text{LogN}(\mu, \sigma^2)$

Soit $(c_1, \dots, c_n)$ la réalisation d'un n -échantillon $(C_1, \dots, C_n)$ de C . Les estimations de (μ, σ^2) par la méthode des moments et la méthode du maximum de vraisemblance donnent les mêmes estimateurs :

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \ln c_i \quad \text{et} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln c_i - \hat{\mu})^2$$

- Loi Gamma $\gamma(\alpha, \beta)$

Soit $(c_1, \dots, c_n)$ la réalisation d'un n -échantillon $(C_1, \dots, C_n)$ de C . La méthode des moments donne immédiatement :

$\tilde{\alpha} = \frac{\bar{c}^2}{\hat{\sigma}^2}$ et $\tilde{\beta} = \frac{\bar{c}}{\hat{\sigma}^2}$ pour $\bar{c}$ et $\hat{\sigma}^2$ moyenne et variance empiriques de l'échantillon.

La méthode du maximum de vraisemblance donne, en posant la fonction

Digamma $\Psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$, le résultat suivant :

L'emv $(\hat{\alpha}, \hat{\beta})$ est défini par :

- $\hat{\alpha}$ solution de l'équation : $\ln \alpha - \Psi(\alpha) = \ln \frac{\bar{c}}{g}$

où $g = \left(\prod_{i=1}^n c_i \right)^{1/n}$ moyenne géométrique de l'échantillon.

- $\hat{\beta} = \frac{\hat{\alpha}}{\bar{c}}$

- Loi Weibull $W(\tau, \alpha)$

Soit $(c_1, \dots, c_n)$ la réalisation d'un n -échantillon $(C_1, \dots, C_n)$ de C . L'emv $\tilde{\tau}$ de τ est solution de l'équation en τ :

$$\ln \Gamma\left(\frac{2}{\tau} + 1\right) - 2 \ln \Gamma\left(\frac{1}{\tau} + 1\right) = \ln\left(1 + \frac{\hat{\sigma}^2}{\bar{c}^2}\right)$$

L'emv $\tilde{\alpha}$ de α vaut $\tilde{\alpha} = \frac{\bar{c}}{\Gamma\left(\frac{1}{\tilde{\tau}} + 1\right)}$

La méthode du maximum de vraisemblance donne le résultat suivant :

$\hat{\tau}$ est solution de l'équation

$$\frac{1}{\tau} = \frac{\sum_{i=1}^n c_i^{\tau} \ln c_i}{\sum_{i=1}^n c_i^{\tau}} - \frac{1}{n} \sum_{i=1}^n \ln c_i \quad \text{et}$$

$$\hat{\alpha} = \left(\frac{1}{n} \sum_{i=1}^n c_i^{\hat{\tau}} \right)^{1/\hat{\tau}}$$

- Loi Pareto $P(a, \alpha)$

Soit $(c_1, \dots, c_n)$ la réalisation d'un n -échantillon $(C_1, \dots, C_n)$ de C .

a est supposé donné. La méthode des moments n'est pas conseillée car l'existence des moments n'est pas assurée.

L'emv de α est :

$$\hat{\alpha} = \frac{1}{\ln g / a}$$

où $g = \left(\prod_{i=1}^n c_i \right)^{1/n}$ moyenne géométrique de l'échantillon.

- Loi Normale $N(\mu, \sigma^2)$

Soit $(c_1, \dots, c_n)$ la réalisation d'un n -échantillon $(C_1, \dots, C_n)$ de C . Les estimations de (μ, σ^2) par la méthode des moments et la méthode du maximum de vraisemblance donnent les mêmes estimateurs :

$$\hat{\mu} = \bar{c} \quad \text{et} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (c_i - \bar{c})^2$$

- Mélange de distributions

En présence de k différents types de sinistres, nous modéliserons le montant d'un sinistre par une distribution mélange de f.r. F :

$$F(x) = \sum_{j=1}^k p_j F_j(x) \quad \text{avec} \quad \sum_{j=1}^k p_j = 1$$

où F_j est la f.r. du montant d'un sinistre de type j et p_j est la probabilité qu'un sinistre soit de type j .

F est le mélange des lois F_j avec pondérations p_j t.q. $\sum_{j=1}^k p_j = 1$.

Nous en verrons les applications pour les modélisations des montants de sinistres RC hors graves, ainsi que des sinistres RC graves.

III. Modélisation des sinistres RC graves

Nous qualifions de sinistres « graves » les sinistres de montant supérieur à 75K euros. Nous considérerons que la probabilité de survenance d'un sinistre de grande intensité est d'autant plus importante (plus faible) que la probabilité de survenance d'un sinistre RC hors graves est élevée (faible).

Remarquons qu'il s'agit d'un raisonnement similaire adopté par Covéa Fleet pour le calcul de la provision pour graves qui est déterminée en fonction d'une sinistralité écrêtée par application d'un taux.

Nous admettrons ainsi que l'espérance de la v.a. fréquence de sinistres graves est égale à un taux de l'espérance de la v.a. fréquence de sinistres RC hors graves.

Nous supposerons que la v.a. fréquence de sinistres graves est modélisée par une loi de Poisson dont le paramètre est, par conséquent, égal au produit de ce taux par l'espérance de la v.a. fréquence de sinistres RC hors graves (ceci sera détaillé ultérieurement dans la partie applications numériques sur les sinistres graves).

En ce qui concerne le coût des sinistres graves, nous souhaitons distinguer deux types de sinistres : ceux qui sont « ordinaires » et ceux qui sont « non ordinaires ».

Nous modélisons alors le coût par une loi mélange de deux distributions dont la première correspond aux sinistres « graves ordinaires » et la deuxième correspond aux sinistres « graves non ordinaires ». Une telle approche est basée sur le travail de Benabbou et Partrat (1994).

Nous commençons par rappeler une approche pour la modélisation d'un montant au-delà d'un certain seuil, la translation, dont nous en appliquerons le principe pour modéliser le coût des graves ; nous présentons ensuite les grandes lignes de la procédure qui permettent de déterminer les paramètres de la loi mélange considérée.

1. Translation

Il s'agit de modéliser une v.a.r. C montant d'un sinistre au-delà d'un certain seuil u par une loi dont la densité est f_C . Nous modélisons $C-u$ dont le support a pour origine 0 par la loi dont nous nommons la densité f et nous avons :

$$f_C(x) = f(x-u)1_{[u; +\infty[}(x)$$

Dans la suite, nous appliquerons cette translation à deux reprises :

- d'une part pour la modélisation des sinistres « graves » de montant supérieur à 75K euros par une loi mélange,
- d'autre part en raison de l'intervention d'un seuil x_0 (à choisir) qui est tel que tout sinistre de montant inférieur à x_0 est « grave ordinaire » et tout sinistre de montant supérieur à x_0 est un mélange de sinistre « grave ordinaire » et de sinistre « grave non ordinaire ».

Nous en présentons les détails dans la suite.

2. « Grands » sinistres et lois mélanges

Nous présentons assez succinctement les points essentiels qui permettront la mise en pratique de l'approche décrite dans l'article pour déterminer les paramètres de la distribution mélange modélisant le coût des sinistres graves. Pour plus de détails, le lecteur pourra se reporter à l'article de Benabbou et Partrat (1994). Dans leurs travaux, les auteurs s'intéressent à la modélisation de l'ensemble des sinistres qui touchent un portefeuille, en différenciant les grands sinistres des sinistres ordinaires. Nous précisons que notre approche consistera à reprendre ce même raisonnement en l'appliquant uniquement aux sinistres graves d'un portefeuille.

Le modèle

La v.a.r. X montant d'un sinistre a pour f.r.

$$F(x, \lambda, \theta_1, \theta_2) = \lambda F_1(x, \theta_1) + (1 - \lambda) F_2(x, \theta_2) \text{ où}$$

F_1 est la distribution des sinistres « ordinaires » dépendant d'un paramètre $\theta_1 \in \Theta_1$ ouvert de $\mathfrak{R}^{q_1}$,

F_2 est la distribution des sinistres « non ordinaires » dépendant d'un paramètre $\theta_2 \in \Theta_2$ ouvert de $\mathfrak{R}^{q_2}$ vérifiant pour tout θ_2

$F_2(x, \theta_2) = 0$ pour $x \leq x_0$ (x_0 seuil donné),
 $\lambda \in]0;1]$ est la probabilité qu'un sinistre soit « ordinaire ».

Nous noterons $\theta = (\theta_1, \theta_2) \in \Theta = \Theta_1 \times \Theta_2$.

Estimation des paramètres

Nous souhaitons déterminer $(\hat{\lambda}_n, \hat{\theta}_{1n}, \hat{\theta}_{2n})$ l'estimateur du maximum de vraisemblance de $(\lambda, \theta_1, \theta_2)$ basé sur la réalisation $(x_i)_{i=1, \dots, n}$ d'un échantillon $(X_i)_{i=1, \dots, n}$ de la v.a.r. X de densité

$$f(x, \lambda, \theta) = \lambda f_1(x, \theta_1) + (1 - \lambda) f_2(x, \theta_2)$$

Soit la log-vraisemblance associée $\ln L = \sum_{i=1}^n \ln f(x_i, \lambda, \theta)$, alors $(\hat{\lambda}_n, \hat{\theta}_n)$ sous réserve d'existence est solution du système :

$$(S) \quad \begin{cases} \frac{\partial \ln L}{\partial \lambda} = \sum_{i=1}^n \frac{f_1(x_i, \theta_1) - f_2(x_i, \theta_2)}{f(x_i, \lambda, \theta)} = 0 \\ \frac{\partial \ln L}{\partial \theta_1^{(k)}} = \lambda \sum_{i=1}^n \frac{\partial f_1(x_i, \theta_1) / \partial \theta_1^{(k)}}{f(x_i, \lambda, \theta)} = 0 \quad k = 1, \dots, q_1 \\ \frac{\partial \ln L}{\partial \theta_2^{(l)}} = (1 - \lambda) \sum_{i=1}^n \frac{\partial f_2(x_i, \theta_2) / \partial \theta_2^{(l)}}{f(x_i, \lambda, \theta)} = 0 \quad l = 1, \dots, q_2 \end{cases}$$

L'algorithme E.M. permet de déterminer la solution de ce système.

Nous posons (en supposant $q_1 = q_2 = 1$ pour alléger les notations) :

$$t_1(x) = \frac{\lambda f_1(x, \theta_1)}{f(x, \lambda, \theta)} \quad \text{et} \quad t_2(x) = \frac{(1 - \lambda) f_2(x, \theta_2)}{f(x, \lambda, \theta)} = 1 - t_1(x)$$

Le système (S) devient alors :

$$(S') \quad \begin{cases} \lambda = \frac{1}{n} \sum_{i=1}^n t_1(x_i) \\ \sum_{i=1}^n t_1(x_i) \frac{\partial \ln f_1(x_i, \theta_1)}{\partial \theta_1} = 0 \\ \sum_{i=1}^n t_2(x_i) \frac{\partial \ln f_2(x_i, \theta_2)}{\partial \theta_2} = 0 \end{cases}$$

Une fois la valeur initiale $(\lambda^{(0)}, \theta^{(0)})$ déterminée, le processus est itératif. Nous avons :

Etape E :

$$t_1^{(m)}(x) = \frac{\lambda^{(m)} f_1(x, \theta_1^{(m)})}{f(x, \lambda^{(m)}, \theta^{(m)})} \quad \text{et} \quad t_2^{(m)}(x) = \frac{(1 - \lambda^{(m)}) f_2(x, \theta_2^{(m)})}{f(x, \lambda^{(m)}, \theta^{(m)})}$$

Etape M :

$$\lambda^{(m+1)} = \frac{1}{n} \sum_{i=1}^n t_1^{(m)}(x_i)$$

$$\theta_1^{(m+1)} \text{ est solution de } \sum_{i=1}^n t_1^{(m)}(x_i) \frac{\partial \ln f_1(x_i, \theta_1)}{\partial \theta_1} = 0$$

$$\theta_2^{(m+1)} \text{ est solution de } \sum_{i=1}^n t_2^{(m)}(x_i) \frac{\partial \ln f_2(x_i, \theta_2)}{\partial \theta_2} = 0$$

Estimation de la valeur initiale $(\lambda^{(0)}, \theta^{(0)})$

Un bon choix de la valeur initiale s'avère nécessaire pour que la suite $(\lambda^{(n)}, \theta^{(n)})$ converge vers $(\hat{\lambda}_n, \hat{\theta}_n)$. Sa détermination est basée sur l'identité $F(x, \lambda, \theta) = \lambda F_1(x, \theta)$ pour $x \leq x_0$. Ainsi la loi F tronquée sur $[0; x_0]$ coïncide avec la loi F_1 tronquée sur $[0; x_0]$:

$$F_{[0; x_0]}(x, \lambda, \theta) = \frac{F(x, \lambda, \theta)}{F(x_0, \lambda, \theta)} = \frac{F_1(x, \theta_1)}{F_1(x_0, \theta_1)} = F_{1[0; x_0]}(x, \theta_1)$$

Une estimation $\tilde{\theta}_{1n}$ de θ_1 est déterminée par les procédures classiques d'estimation par le maximum de vraisemblance de paramètres de lois tronquées à droite (COHEN, 1991).

Dans nos applications numériques, nous modéliserons les sinistres graves par un mélange de lois Lognormales. Aussi nous détaillons ci-dessous les formules permettant de déterminer les paramètres correspondants.

Les deux composantes de la loi mélange $\text{LogN}(\mu_1, \sigma_1^2)$ et $\text{LogN}(\mu_2, \sigma_2^2)$ ont pour densité :

$$f_1(x, \mu_1, \sigma_1) = \frac{1}{\sqrt{2\pi}\sigma_1 x} \exp\left(-\frac{1}{2\sigma_1^2} (\ln x - \mu_1)^2\right)$$

$$f_2(x, \mu_2, \sigma_2) = \frac{1}{\sqrt{2\pi}\sigma_2 (x - x_0)} \exp\left(-\frac{1}{2\sigma_2^2} (\ln(x - x_0) - \mu_2)^2\right) 1_{[x_0; +\infty[}(x)$$

Si n_0 est le nombre de $x_i \leq x_0$, l'emv de (μ_1, σ_1^2) basé sur la réalisation $(x_i)_{i=1, \dots, n_0}$ de la loi $\text{LogN}(\mu_1, \sigma_1^2)$ tronquée à droite de x_0 s'obtient par :

$$\text{Posons } \bar{y} = \frac{1}{n_0} \sum_{i=1}^{n_0} \ln x_i, \quad s^2 = \frac{1}{n_0} \sum_{i=1}^{n_0} (\ln x_i - \bar{y})^2 \quad \text{et} \quad w = \frac{s^2}{(\ln x_0 - \bar{y})^2}$$

Soient les fonctions $W(x) = \frac{\varphi(x)}{\phi(x)}$, où φ et ϕ sont respectivement la densité et la f.r. de la loi normale $N(0,1)$, et $Q(w) = \frac{W(\hat{u})}{W(\hat{u}) + \hat{u}}$ où $\hat{u}$ est l'unique solution de l'équation en u :

$$\frac{1 - W(u)[W(u) + u]}{[W(u) + u]^2} = w$$

alors nous avons :

$$\tilde{\mu}_{1n} = \bar{y} + Q(w)(\ln x_0 - \bar{y})$$

$$\tilde{\sigma}_{1n}^2 = s^2 + Q(w)(\ln x_0 - \bar{y})^2$$

Pour $w < 0,57081$, nous avons l'approximation rationnelle suivante (SCHNEIDER, 1986) :

$$Q(w) = \begin{cases} 0 & \text{si } w \leq 0.06246 \\ \frac{P_4(w)}{P_3(w)} & \text{si } 0.06246 < w < 0.57081 \end{cases}$$

où

$$P_4(w) = -0.00374615 + 0.17462558w - 2.87168509w^2 + 17.48932655w^3 - 11.91716546w^4$$

$$P_3(w) = 1 + 5.74050101w - 13.53427037w^2 + 6.88665552w^3$$

Une fois déterminée $(\tilde{\mu}_{1n}, \tilde{\sigma}_{1n}^2)$ l'estimation initiale de (μ_1, σ_1^2) , nous en déduisons les estimations initiales $\tilde{\lambda}_n$ et $(\tilde{\mu}_{2n}, \tilde{\sigma}_{2n}^2)$ de λ et (μ_2, σ_2^2) .

$$\tilde{\lambda}_n = \frac{F_n^*(x_0)}{F_1(x_0, \tilde{\mu}_{1n}, \tilde{\sigma}_{1n}^2)} \text{ avec } F_n^* \text{ la fonction de répartition empirique associée à}$$

$$\text{l'échantillon } (X_i)_{i=1, \dots, n} : F_n^*(x) = \frac{1}{n} \sum_{i=1}^n 1_{[0; x]}(X_i)$$

$(\tilde{\mu}_{2n}, \tilde{\sigma}_{2n}^2)$ sont solutions de l'équation suivante en (μ_2, σ_2^2) , où v est une valeur réelle

$$\text{choisie dans }]x_0; +\infty[: F_2(v, \mu_2, \sigma_2) = \frac{F_n^*(v) - \tilde{\lambda}_n F_1(v, \tilde{\mu}_{1n}, \tilde{\sigma}_{1n}^2)}{1 - \tilde{\lambda}_n}$$

Et à l'étape $(m+1)$, les paramètres des deux lois Lognormales sont obtenus par les relations :

$$\mu_1^{(m+1)} = \frac{\sum_{i=1}^n t_1^{(m)}(x_i) \ln x_i}{\sum_{i=1}^n t_1^{(m)}(x_i)}, \quad \sigma_1^{(m+1)^2} = \frac{\sum_{i=1}^n t_1^{(m)}(x_i) (\ln x_i - \mu_1^{(m)})^2}{\sum_{i=1}^n t_1^{(m)}(x_i)}$$

$$\mu_2^{(m+1)} = \frac{\sum_{\substack{i=1 \\ x_i > x_0}}^n t_2^{(m)}(x_i) (\ln(x_i - x_0))}{\sum_{\substack{i=1 \\ x_i > x_0}}^n t_2^{(m)}(x_i)}, \quad \sigma_2^{(m+1)^2} = \frac{\sum_{\substack{i=1 \\ x_i > x_0}}^n t_2^{(m)}(x_i) (\ln(x_i - x_0) - \mu_2^{(m)})^2}{\sum_{\substack{i=1 \\ x_i > x_0}}^n t_2^{(m)}(x_i)}$$

IV .Tests d'adéquation

Le modèle choisi et le paramètre estimé, il convient d'évaluer la qualité de l'ajustement aux données de façon à retenir une adéquation « satisfaisante ». Notre problématique consiste à tester l'adéquation des observations à une famille de lois dépendant d'un paramètre que nous

aurons estimé à partir de cet échantillon (il s'agit d'un cas différent de celui où l'adéquation à une loi de probabilité entièrement spécifiée est testée).

Plusieurs tests d'adéquation sont souvent utilisés, basés sur : la statistique du χ^2 , la distance de Kolmogorov Smirnov et l'écart d'Anderson Darling. Chacun présente ses avantages et ses inconvénients, et selon que la v.a.r. est discrète ou continue, un test spécifique pourra être appliqué.

1. Test du χ^2

Ce test universel largement utilisé s'applique généralement à une variable discrète. C'est celui que nous utilisons pour l'adéquation des modèles de fréquence. Il peut aussi être appliqué à une variable continue (le coût d'un sinistre par exemple), mais cela implique un regroupement des données en classes et par conséquent une perte d'information.

La statistique χ^2 se définit comme suit (nous la définissons dans le cas d'une variable continue ; le cas d'une variable discrète s'en déduit de façon immédiate) :

$$\chi^2 = \sum_{i=1}^K \frac{(N_i - E_i)^2}{E_i}$$

où

K représente le nombre d'intervalles

N_i représente le nombre de valeurs observé dans le $i^{\text{ème}}$ intervalle

E_i représente le nombre de valeurs attendu dans le $i^{\text{ème}}$ intervalle

2. Test de Kolmogorov Smirnov

Pour une v.a.r. continue telle que le coût de sinistres, un test approprié (plus puissant que le test du χ^2) est celui de Kolmogorov Smirnov qui fait partie de la classe des tests EDF (Empirical Distribution Function).

Soit $(C_1, \dots, C_n)$ un n -échantillon de la v.a.r. coût C de fonction de répartition F et F_n sa fonction de répartition empirique. Par définition de F_n , on a :

$$F_n(x) = \frac{i}{n} \text{ pour } C_{(i)} < x \leq C_{(i+1)} \text{ avec les } (C_{(i)})_{i=1, \dots, n} \text{ ordonnés}$$

Nous souhaitons tester l'adéquation de la f.r. F de la v.a.r. C à une f.r. G continue donnée. Comme F est inconnue, nous lui substituons la f.r. empirique F_n .

La distance de Kolmogorov Smirnov est définie en G par :

$$D_n(G) = \sup_x |F_n(x) - G(x)| = \max[D_n^+(G), D_n^-(G)]$$

avec

$$D_n^+(G) = \sup_x [F_n(x) - G(x)]$$

$$D_n^-(G) = \sup_x [G(x) - F_n(x)]$$

Dans la pratique nous utilisons la proposition suivante :

Proposition

$$D_n^+(G) = \max_{1 \leq i \leq n} \left[\frac{i}{n} - G(C_{(i)}) \right]$$

$$D_n^-(G) = \max_{1 \leq i \leq n} \left[G(C_{(i)}) - \frac{i-1}{n} \right]$$

L'avantage de ce test sur celui du χ^2 est qu'il ne perd pas d'information car nul regroupement en classes n'est nécessaire. Son inconvénient est de ne pas bien évaluer les éventuelles divergences en queue de distribution (il est en fait précis au niveau de l'écart maximal entre deux distributions). Le test suivant permet de remédier à ce défaut.

3. Test d'Anderson Darling

Le choix de l'écart d'Anderson Darling est pertinent dans les cas où la queue de distribution doit être bien prise en compte car elle y est sur pondérée.

Nous l'utiliserons notamment pour mesurer l'ajustement des distributions des sinistres graves.

L'expression opérationnelle de A_n^2 est :

$$A_n^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \{ \ln F(C_{(i)}) + \ln [1 - F(C_{(n-i+1)})] \}$$

où $C_{(i)}$ désigne la $i^{\text{ième}}$ plus petite réalisation de la v.a.r. coût de sinistre C dans l'échantillon.

4. Test d'adéquation à une famille de lois de probabilité

Nous utiliserons les approches spécifiques suivantes lors de nos applications :

- cas du test du χ^2 : cas d'une famille de lois à deux paramètres (loi Binomiale Négative par exemple)

Si $\chi^2_{obs} = \sum_{i=1}^K \frac{(N_i - n\hat{p}_i)^2}{n\hat{p}_i}$ représente la statistique du χ^2 observée sur l'échantillon regroupé en K classes avec :

N_i le nombre d'observations appartenant à la classe i

$n\hat{p}_i$ le nombre attendu de valeurs de la classe i (avec $\hat{p}_i$ déduits de l'estimateur $\hat{\theta}$)

alors l'hypothèse d'adéquation n'est pas refusée si

$\chi^2_{obs} \leq \text{seuil}$ avec $\text{seuil} = \chi^2_{K-1-2}$ (lu dans la table du χ^2 , deux degrés de liberté sont enlevés en raison des deux paramètres estimés)

- cas d'un test de type EDF

Les lois limites des statistiques EDF diffèrent de celles du cas où les lois sont entièrement spécifiées. Pour certaines familles de lois (loi exponentielle, loi normale par exemple), des tables de quantiles qui leur sont spécifiques permettent de déterminer l'adéquation de l'échantillon à ces dernières (pour plus de détails, se reporter par exemple à D'Agostino et Stephens, 1986). Dans les situations où des tests EDF à paramètres estimés ne sont pas disponibles (par exemple les cas de la loi Gamma, ainsi que des lois mélange), nous aurons recours à la méthode du demi-échantillon. Elle consiste à extraire de manière aléatoire un sous-échantillon de taille $(n/2)$ de l'échantillon initial et d'estimer θ à l'aide d'un estimateur $\hat{\theta}$ asymptotiquement efficace. Soit ensuite une statistique EDF retenue pour le test, calculée sur la totalité de l'échantillon, elle a alors pour loi limite celle du cas entièrement spécifiée (en pratique cette technique est utilisable dès que $n \geq 30$).

V. Applications numériques

Nous appliquons les méthodes décrites ci-dessus pour modéliser les fréquence et coût de sinistres de la RC Corporelle, la RC Matérielle et la RC Graves.

Pour chaque modélisation, les étapes seront les suivantes :

- Choix d'un modèle de distribution
- Estimation des paramètres du modèle
- Validation du choix du modèle et des paramètres estimés par un test d'adéquation

Nous commençons par donner quelques précisions sur le poids des sinistres graves dans notre portefeuille.

Nous présentons dans le tableau suivant le poids des sinistres supérieurs à 75000 euros de l'exercice 2004 où l'évaluation est vue à fin 2005.

	NB	Charge (K€)
RC >= 75 K€	116	62296
RC TOTAL	57349	135378
% par rapport au total	0,20%	46,02%

Nous remarquons que la part des graves au niveau du nombre est infime, mais qu'ils représentent une part importante pour la charge.

1. Modélisation des sinistres « de fréquence »

Il s'agit des sinistres dont le montant est inférieur à 75K euros. Nous distinguerons les sinistres RC Corporels, les sinistres RC Matériels hors IDA et les sinistres Matériels IDA. La séparation des sinistres à forfait IDA (qui constitue une majorité des sinistres RC Matériels) permet une meilleure modélisation du coût des sinistres matériels.

Rappel :

La convention IDA est une convention de règlement de sinistres entre assureurs concernant la RC automobile et qui prévoit que si la partie « matériel » du sinistre a un montant inférieur au plafond IDA (6500 euros en 2006), le conducteur non responsable est directement indemnisé par son assureur ; ce dernier exercera un recours forfaitaire (1204 euros pour l'année 2006) auprès de l'assureur du conducteur responsable. Le montant de ce forfait vise à une compensation globale par exercice.

Pour la modélisation de la fréquence, les données sur lesquelles se base notre étude sont les parcs désignés du segment 4 sur l'exercice 2004. La base de données retenue comprend un nombre de 13458 véhicules de 1^{ère} catégorie présents les 366 jours. Le nombre de sinistres RC hors graves associés s'élève à 1113.

Pour la modélisation des coûts, nous nous intéressons aux sinistres de tous les segments sur l'exercice 2004 concernant les véhicules de 1^{ère} catégorie pour augmenter la base de données de montants. Nous supposons qu'à la différence de la fréquence, la variable coût n'a pas une distribution « spécifique » au segment 4.

a) Modélisation de la fréquence

Nous modélisons la v.a. nombre de sinistres RC hors graves N_{rc} par une loi Binomiale Négative. En effet, c'est la distribution la mieux adaptée pour traduire l'hétérogénéité de la fréquence de sinistres dans les grosses flottes automobiles (également justifiée par les valeurs observées de la moyenne et de la variance empiriques – variance > moyenne -).

Soient :

λ_{rcc} le taux de corporels (probabilité d'un sinistre RC hors graves d'être un corporel)

λ_{rcm_NI} la probabilité d'un sinistre RC hors graves d'être un matériel non IDA

λ_{rcm_I} la probabilité d'un sinistre RC hors graves d'être un matériel IDA

avec $\lambda_{rcc} + \lambda_{rcm_NI} + \lambda_{rcm_I} = 1$

alors la v.a. coût d'un sinistre RC hors graves a une distribution mélange de fonction de répartition :

$$F(x) = \lambda_{rcc} F_{rcc}(x) + \lambda_{rcm_NI} F_{rcm_NI}(x) + \lambda_{rcm_I} F_{rcm_I}(x)$$

où F_{rcc} est la f.r. du montant d'un sinistre RC corporelle,

F_{rcm_NI} est la f.r. du montant d'un sinistre RC matérielle non IDA,

et F_{rcm_I} est la f.r. du montant d'un sinistre RC matérielle IDA.

Nous déterminons les paramètres (r, p) de la loi Binomiale Négative modélisant le nombre de sinistres RC hors graves par la méthode du Maximum de Vraisemblance et en utilisant le solveur d'Excel.

Les valeurs initiales utilisées lors de cette résolution numérique sont naturellement les emm qui auront été calculés auparavant. Nous testons ensuite leurs adéquations aux données avec la statistique du χ^2 .

Les valeurs de λ_{rcc} , λ_{rcm_NI} , λ_{rcm_I} sont déterminées empiriquement à partir des données.

Garantie RC hors graves :

Les moyenne et variance empiriques observées sont respectivement :

$$\bar{n} = 0,08270174 \quad \text{et} \quad \overline{\sigma^2} = 0,0928141$$

Les paramètres déterminés sont présentés dans le tableau suivant ; suivra le tableau résumant le résultat du test d'adéquation du χ^2 .

Paramètres de la loi Binomiale Négative pour la RC hors graves

r	p
0,65425239	0,887778987

Test du χ^2

	Test χ^2
	0,06579
seuil à 5%	3,841

Au vu du résultat de ce test, nous constatons que l'hypothèse de la loi Binomiale Négative est largement acceptée pour nos données à un niveau de signification 0,05 (la P-valeur est égale à 0,797).

(la statistique du χ^2 est déterminée pour n=4 et deux paramètres estimés)

Les taux de corporels, matériels non IDA et IDA sont donnés dans le tableau ci-dessous.

λ_{rcc}	0,10703409
λ_{rcm_NI}	0,21662923
λ_{rcm_I}	0,67633668

b) Modélisation du coût

Pour chacune des deux « garanties » RCC et RCM (hors IDA), nous souhaitons modéliser le coût de sinistre par une des lois classiques présentées précédemment. Les distributions qui paraissent répondre au mieux à cette modélisation sont la loi lognormale et la loi gamma dont les utilisations sont usuelles en assurance non-vie, et dont les queues de distribution ne présentent pas une trop grande épaisseur comme la loi Pareto par exemple. Nous présentons les résultats des modélisations avec ces deux distributions. Nous verrons que la loi retenue est la loi Gamma qui s'avèrera être une distribution bien adaptée ; elle ajustera en effet de façon satisfaisante les queues de distribution des sinistres de petite et moyenne intensité tels que ceux étudiés ici.

Nous déterminons les paramètres des distributions par une méthode de minimisation de la distance de Cramer- von Mises pondérée. Nous utiliserons le solveur d'Excel pour déterminer les paramètres. La méthode du demi échantillon est utilisée dans les modélisations qui suivent pour la loi Gamma (un demi échantillon est donc extrait de l'échantillon d'origine de manière aléatoire et les paramètres seront estimés).

Les valeurs initiales seront les emm qui auront été calculés préalablement à partir des données empiriques.

Nous testons ensuite l'adéquation des distributions associées à ces paramètres aux données initiales par la statistique de Kolmogorov Smirnov.

Garantie RC Corporelle :

Loi Gamma

Les paramètres de la loi Gamma déterminés sont présentés dans le tableau suivant :

α	β
0,84826913	0,00021074

Le test de Kolmogorov Smirnov donne le résultat ci-dessous :

	Test KS
	0,077334
seuil à 5%	0,085

La statistique de Kolmogorov Smirnov observée D_n (égale à 0,077) est calculée à partir des paramètres déterminés grâce au demi-échantillon tiré aléatoirement de l'ensemble des données (avec $n=256$ le nombre de valeurs de l'échantillon initial) et est comparée au quantile d'ordre 0,95 de la statistique KS pour une loi entièrement spécifiée. Elle est inférieure à ce quantile égal à 0,085, ainsi la modélisation des données RCC par une loi Gamma n'est pas rejetée au niveau de signification 0,05.

La moyenne empirique vaut $\bar{C} = 4166$ euros.

L'espérance de la distribution Gamma déterminée est $E(C) = 4025,13$ euros.

Loi Lognormale

Les paramètres de la loi Lognormale déterminées sont donnés dans le tableau ci-dessous :

μ	σ
7,78233226	1,22181322

et les résultats du test de Kolmogorov Smirnov sont :

	Test KS
	0,08482172
seuil à 5%	0,0559375

La valeur de la statistique de KS observée est supérieure au quantile d'ordre 0,95 de la statistique KS pour une loi Lognormale dont les paramètres sont estimés, ainsi nous rejetons le modèle Lognormale pour l'ajustement de nos données au niveau de signification 0,05.

Nous retiendrons ainsi la loi Gamma comme la distribution modélisant les coûts de la RC Corporelle.

Garantie RC Matérielle Non IDA :

Loi Gamma

Nous présentons ci-dessous les paramètres calculés de la loi Gamma modélisant le coût de la garantie RCM non IDA :

α	β
1,46433013	0,00109156

Le test de Kolmogorov Smirnov donne le résultat suivant :

	Test KS
	0,038192
seuil à 5%	0,061882

La statistique de Kolmogorov Smirnov observée D_n (égale à 0,038) est inférieure au quantile d'ordre 0,95 de la statistique KS pour une loi entièrement spécifiée (pour $n=483$) qui est de 0,062, ainsi nous retenons la loi Gamma pour l'ajustement des données de la RC Matérielle (hors IDA) au niveau de signification 0,05. (la P-valeur est $> 0,25$).

La moyenne empirique vaut $\bar{C} = 1554$ euros.

L'espérance de la distribution Gamma déterminée est $E(C) = 1341,5$ euros

Loi Lognormale

Les paramètres calculés de la loi Lognormale sont indiqués dans le tableau suivant :

μ	σ
6,89823997	0,93745679

Les résultats du test de Kolmogorov Smirnov sont :

	Test KS
	0,04880865
seuil à 5%	0,04072391

Nous rejeterons donc également la loi Lognormale pour la modélisation des coûts de la RC Matérielle hors IDA au niveau de signification 0,05.

Garantie RC Matérielle IDA :

Le coût d'un forfait IDA est constant et s'élève à 1204 euros pour l'année 2004.

La fonction de répartition d'un coût de sinistre RC hors graves s'écrit finalement :

$$F(x) = 0.10703409.F_{rc}(x) + 0.21662923.F_{rcm_NI}(x) + 0.67633668.F_{rcm_I}(x)$$

2. Modélisation des sinistres graves

Pour l'étude des sinistres graves nous ne distinguons pas les catégories de véhicules pour des raisons de volumétrie.

Les données sur lesquelles nous effectuons l'étude sont, pour la modélisation de la fréquence, l'ensemble du parc tppc de segment 4 dont nous connaissons les nombres de véhicules des flottes. Nous nous intéresserons à l'analyse de la fréquence de sinistralité sur l'exercice 2004.

Pour la modélisation du coût des sinistres graves, nous étendrons notre sélection de données à tous les segments (et non uniquement le segment 4) de façon à obtenir une volumétrie plus importante. Les sinistres sont ceux survenus lors de l'exercice 2004 et les coûts sont vus à la fin de l'année 2006. A cette date là, notons que les données MMA et MAAF ont pu être complétées avec celles de SMABTP.

Par ce choix de périmètre des données, nous considérons que les flottes de plus de cent moteurs ont une fréquence de sinistralité qui leur est spécifique alors que la taille d'une flotte n'influerait pas sur la structure du coût de ses sinistres graves. Rappelons également que pour l'étude des graves, nous avons pris en compte tous les moteurs de toutes les catégories sans restriction aux véhicules de 1^{ère} catégorie, à la différence du cas de l'analyse des sinistres de fréquence.

a) Fréquence d'un sinistre grave

Nous modélisons la v.a. fréquence de sinistres RC graves par une loi de Poisson (dont le choix est naturel pour la modélisation de la survenance d'un événement rare) dont le paramètre λ est égal au produit entre :

- le taux λ_1 égal au rapport de l'espérance de la fréquence de sinistres graves pour un véhicule moteur sur l'espérance de la fréquence de sinistres hors graves pour un véhicule moteur qui a été déterminé empiriquement et est indiqué ci-dessous

$$\lambda_1 = 0,00232813$$

En effet sur la totalité des flottes (désignées et non désignées) du segment 4 dont il a été possible d'évaluer le nombre de moteurs (par demi différence), il a été observé 28 sinistres graves en 2004 pour une totalité de 145 424 véhicules moteurs.

- l'espérance de la v.a. fréquence de sinistres RC hors graves

Cette approche nous permettra dans la suite de pouvoir prendre en compte la sinistralité antérieure d'une flotte (hors grave) dans l'évaluation de la sinistralité grave a posteriori. Nous le verrons plus en détail dans le chapitre suivant.

b) Modélisation du coût d'un sinistre grave

Nous souhaitons modéliser le coût d'un sinistre grave par une loi mélange de deux distributions lognormales. Nous élaborons un outil Excel en nous basant sur les résultats de l'article Benabbou et Partrat (1994). Cet outil nous permettra de déterminer les cinq paramètres de la loi mélange.

Un sinistre grave est distingué par son niveau de montant supérieur à 75K euros. Le but est donc de modéliser une v.a.r. C montant d'un sinistre au-delà de ce seuil de 75K. La fonction de répartition de la v.a.r. C s'écrit alors :

$$F_C(x) = k.F_1(x - 75000). \frac{I}{x > 75000} + (1 - k).F_2(x - (x_0 + 75000)). \frac{I}{x > (x_0 + 75000)}$$

avec

F_1 la distribution des sinistres « graves ordinaires » de paramètres (μ_1, σ_1^2) ,

F_2 la distribution des sinistres « graves non ordinaires » de paramètres (μ_2, σ_2^2) ,

k la probabilité pour qu'un sinistre grave soit « ordinaire », et

x_0 tel que $(x_0 + 75000)$ est le seuil en dessous duquel un sinistre grave sera considéré comme « grave ordinaire »

Pour le choix de x_0 , dont la valeur ne semble pouvoir être fixée initialement de façon pertinente (quel est le seuil entre un sinistre grave « ordinaire » et un sinistre grave « non ordinaire » ?), nous avons été amenés à la déterminer de la manière suivante :

Dans nos applications, différentes valeurs de x_0 sont testées, et pour chacune d'entre elles, lorsque les cinq paramètres de la loi mélange sont déterminés, nous testons l'adéquation de cette dernière distribution aux données.

Finalement nous retenons le seuil $(x_0 + 75000)$ qui donne le meilleur ajustement.

Un seuil de 1 000 000 euros (c'est-à-dire $x_0 = 925\,000$) donne des résultats légèrement meilleurs au niveau des tests d'adéquation par rapport à d'autres seuils, mais la différence est en réalité très peu sensible.

Nous présentons dans les tableaux qui suivent les paramètres déterminés ainsi que les résultats des différents tests d'adéquation pour une valeur $x_0 = 925\,000$.

Valeurs des paramètres de la loi mélange

k	μ_1	σ_1	μ_2	σ_2
0,98706511	10,9750274	1,91848535	14,9489309	0,177151

Tests d'adéquation de Kolmogorov Smirnov et d'Anderson Darling (nous rappelons que la méthode du demi échantillon a été utilisée, le nombre de valeurs initiales étant n=209)

KS		AD	
Dn	P-valeur	An	P-valeur
0,0514523	P>=0,25	0,60266562	P>=0,25

Nous pouvons constater qu'une loi mélange de deux Lognormales s'avère être un très bon ajustement aux coûts des sinistres graves au vu des statistique KS et statistique AD.

Notons que l'espérance de la v.a.r. C montant d'un sinistre grave s'écrit :

$$E(C) = 75000 + k \cdot \exp\left(\mu_1 + \frac{\sigma_1^2}{2}\right) + (1-k) \cdot \left(\exp\left(\mu_2 + \frac{\sigma_2^2}{2}\right) + 925000\right)$$

ce qui donne $E(C) = 490831$ et la moyenne empirique des coûts de sinistres graves $\bar{C} = 421902$

Remarque :

L'approche usuelle pour modéliser les sinistres graves jusqu'à présent à Covéa Fleet fait appel à « une simple » loi lognormale.

Ci-dessous nous présentons les résultats de cette modélisation : les paramètres déterminés et les résultats des tests d'adéquation.

Valeurs des paramètres de la loi lognormale

μ	σ
11,055826	1,77439206

Les tests d'adéquation de Kolmogorov Smirnov et Anderson Darling donnent les résultats suivants :

KS		AD	
Dn	P-valeur	An	P-valeur
0,05200267	$P \geq 0,15$	0,74288138	$P \geq 0,05$

La loi Lognormale s'ajuste, bien sur, également aux données du coût d'un sinistre grave, mais la lecture des P-valeurs indique que la qualité de l'ajustement par un mélange de deux lois Lognormales s'avère être meilleure. (Rappelons que les P-valeurs du mélange de lois correspondent à ceux d'une loi entièrement spécifiée, et que pour celles de la loi Lognormale, les paramètres ont été estimés)

Dans ce chapitre, nous avons modélisé la sinistralité RC automobile en distinguant deux populations de sinistres :

- les sinistres de faible et moyenne intensité dont nous avons modélisé la RC en faisant la distinction entre la RC Corporelle, la RC Matérielle hors IDA et la RC Matérielle IDA,
- les sinistres graves, dont le coût a été modélisé par une approche particulière, un mélange de deux lois lognormales.

Nous avons adopté le modèle composé : la fréquence de sinistres et le coût d'un sinistre ont été modélisés.

Nous avons ainsi à notre disposition les paramètres de toutes ces différentes distributions qui serviront de socle aux chapitres suivants pour la détermination de la prime pure d'une flotte, la simulation de la charge sinistre d'une flotte, et ensuite pour évaluer l'impact du niveau de la conservation appliqué sur la distribution du ratio S/P .

CHAPITRE 3

-

**ESTIMATION DE LA PRIME PURE
D'UNE FLOTTE : DIVERSES
APPROCHES**

Dans ce chapitre nous nous intéressons à l'estimation de la charge sinistre totale d'une garantie RC d'un ensemble de véhicules sur une période donnée par deux approches différentes.

En premier lieu, nous présentons la logique de l'approche qui est actuellement utilisée à Covéa Fleet pour déterminer la prime pure RC d'une flotte de segment 4.

Dans un second temps nous utilisons d'une part les résultats obtenus à partir de la modélisation du chapitre précédent pour déterminer la charge sinistre a priori d'une flotte, et d'autre part, pour prendre en compte la sinistralité antérieure de la flotte, nous faisons appel à la théorie du modèle Bonus Malus pour calculer la charge sinistre a posteriori. Après avoir présenté les bases de cette théorie s'appliquant à un véhicule, nous effectuons une généralisation à une flotte de véhicules.

Dans une troisième partie, nous présentons la théorie de la crédibilité de stabilité, grâce à laquelle les paramètres estimés lors de l'étape de modélisation permettent d'apporter quelques éléments de réponse quant à la taille minimale d'une flotte pour que sa tarification puisse être basée uniquement sur l'expérience.

Enfin, dans la dernière partie, nous appliquons les différents concepts théoriques présentés dans ce chapitre et donnons les résultats concernant la comparaison entre les primes pures déterminées par les deux différentes approches : celle, actuelle, de Covéa Fleet, et celle résultant de notre modélisation. Les applications numériques sont effectuées sur un certain nombre de contrats sélectionnés dans le portefeuille.

I. Détermination de la prime RC Covéa Fleet d'une flotte

Les grosses flottes sont traitées de façon personnalisée pour la détermination de leur prime RC lors des réunions du « comité de souscription et de renouvellement ». Chaque contrat est décrit dans un classeur Excel comprenant un onglet d'informations globales et un onglet regroupant les détails des sinistres (type des sinistres, montant des sinistres, date de survenance,...).

Sur l'onglet « informations globales », les informations de la flotte sont données pour chaque partie suivante : description du parc (et garanties souscrites), sinistralité, et calcul de la cotisation.

- description du parc : les nombres de véhicules sur les exercices N, N-1, N-2 et éventuellement N-3 sont donnés.
- sinistralité :

i. Nombre de sinistres

Pour chacune des garanties RC Matérielle et RC Corporelle, nous avons le nombre de sinistres survenus sur les exercices N-1, N-2 et éventuellement N-3.

Pour l'année N en cours, les nombres de sinistres sont extrapolés selon le nombre de mois d'observation de cette période pour obtenir le nombre de sinistres ultimes sur l'année N.

ii. Charge sinistre

Les montants de sinistres sont écrêtés à 30K euros.

Les charges sinistres reportées dans le tableau de la partie sinistralité tiennent compte de cet écrêtement.

Les informations concernant la charge sinistre sont fournies de façon similaire au nombre de sinistres :

Les charges sinistres pour les garanties RC Matérielle et RC Corporelle sont indiquées pour chaque année (et sont bien entendu actualisés pour prendre en compte l'augmentation du coût moyen des sinistres au fil des ans), et sont extrapolés pour l'année en cours.

Une espérance de la charge sinistres écrêtée est ensuite calculée pour le prochain exercice : elle correspond à la charge moyenne par véhicule multipliée par le parc de l'année N.

S'il y a eu un ou des sinistres supérieur(s) à 30K euros, le montant est reporté sur la fiche en tant qu'observation complémentaire.

Le traitement de la charge pour graves est pris en compte sous forme d'une provision pour sinistres graves et est détaillé dans la suite.

- calcul de la cotisation :

Prime Pure

Le calcul de la prime pure de la garantie RC se décompose en une partie « Prime pure RC écrêtée » et « Provision pour sinistres graves ».

- la Prime Pure RC écrêtée à 30K euros : il s'agit de la somme des espérances des charges sinistres définies ci-dessus pour le parc de l'année N des garanties RC Matérielle et RC Corporelle.
- la Provision pour Sinistres graves (part des sinistres au-dessus de 30 K) : elle est calculée à partir de la prime pure écrêtée et consiste en un taux de provision appliqué sur cette dernière. Il s'agit du taux correspondant à la part de la charge sinistre RC excédant 30K euros déterminée à partir d'une étude historique.

L'addition de ces deux montants nous donne la prime pure RC totale.

Nous remarquerons que dans cette approche, seule la charge sinistre intervient dans le calcul de la prime pure et que l'information nombre de sinistres n'est pas utilisée même si elle est fournie. Mais cette notion nous servira dans la suite pour notre approche modélisation lors de l'estimation de la charge sinistre a posteriori.

Rappelons que dans l'approche actuelle, la distinction entre les sinistres de fréquence et les sinistres graves est au niveau de 30K euros (le seuil d'écèlement), alors que dans notre modélisation, le seuil est fixé à 75K euros.

Observons que l'approche actuelle consiste à se baser uniquement sur l'expérience de la flotte pour les sinistres de montant inférieur à 30K euros ; la sur crête fait l'objet d'une mutualisation par application d'un taux sur la prime pure écèlement. Pour cette approche, le modèle de crédibilité se porte donc sur la partie écèlement de la sinistralité.

Conservation et Cotisation TTC

La prime pure totale déterminée, un niveau de conservation (auquel un taux d'honoraire sera appliqué) est éventuellement décidé, et il est appliqué à la prime pure restante (prime pure hors conservation) différents taux dont les taxes, ce qui donnera la cotisation TTC. (des précisions plus détaillées ont été données auparavant dans le chapitre 1 de présentation)

II. Modèle Bonus-Malus et application aux flottes

Dans la partie ci-dessus, la tarification de la flotte est basée (presque) uniquement sur l'expérience propre de la flotte, avec une prise en compte mutualisée sous forme de provision pour sinistres graves de la sinistralité lourde.

Dans le chapitre précédent, la modélisation des sinistres permet de déterminer une prime « a priori » basée uniquement sur certaines caractéristiques communes à une même classe de risque. Cette tarification « a priori » seule ne prendrait ainsi aucune expérience propre à la flotte en compte.

Dans la partie qui suit, nous nous proposons de combiner à cette modélisation « a priori » une « personnalisation » liée à la flotte qui prendrait le nombre de sinistres RC observés en considération sous la forme d'un Bonus-Malus. Le système Bonus-Malus est largement utilisé en Europe dans l'assurance de véhicules pour prendre en compte l'expérience dans la tarification.

Nous rappelons en premier lieu les bases théoriques du modèle Bonus-Malus et généralisons le concept à une flotte de véhicules.

1. Système de Bonus-Malus appliqué à la famille des lois de Poisson-mélange

Nous considérons un risque dont la fréquence annuelle de sinistres est distribuée selon une loi de Poisson-mélange. Notons N_i la fréquence pour la i^{e} année pour tout i entier et $(n_1, \dots, n_t)$ est la réalisation du vecteur aléatoire $(N_1, \dots, N_t)$ qui correspond à l'observation de ce risque pendant t années.



Nous souhaitons à la fin de la t^{e} année déterminer la loi prédictive de N_{t+1} sachant $(N_1 = n_1, \dots, N_t = n_t)$. Nous supposons que la classe à laquelle appartient le risque considéré possède une variable de structure Λ de fonction de répartition h . Nous faisons les hypothèses suivantes :

(H1) : les v.a.r. $N_1, \dots, N_t, N_{t+1}$ sont conditionnellement en Λ indépendantes

(H2) : la loi conditionnelle de $N_i / \Lambda = \lambda$ est une loi de Poisson $P(\lambda)$ pour tout $i = 1, \dots, t+1$

a) Distribution prédictive conditionnelle $N_{t+1} / N_1 = n_1, \dots, N_t = n_t$

Nous commençons par déterminer la loi a posteriori de Λ sachant $N_1 = n_1, \dots, N_t = n_t$.

Distribution a posteriori de $\Lambda / N_1 = n_1, \dots, N_t = n_t$

Le théorème de Bayes permet d'obtenir la densité de la loi a posteriori de $\Lambda / N_1 = n_1, \dots, N_t = n_t$:

$$h^{n_1, \dots, n_t}(\lambda) = \frac{h(\lambda) P(N_1 = n_1, \dots, N_t = n_t / \Lambda = \lambda)}{\int_0^{+\infty} P(N_1 = n_1, \dots, N_t = n_t / \Lambda = \lambda) h(\lambda) d\lambda}$$

Par les hypothèses (H1) et (H2) et en faisant entrer les constantes dans le signe proportionnel à nous obtenons :

$$h^{n_1, \dots, n_t}(\lambda) \propto e^{-\lambda t} \lambda^{\sum_{i=1}^t n_i} h(\lambda)$$

Observons que cette loi ne dépend de $(n_1, \dots, n_t)$ que par $\sum_{i=1}^t n_i$

Loi prédictive de $N_{t+1} / N_1 = n_1, \dots, N_t = n_t$

Par les hypothèses (H1) et (H2) nous avons également :

$$P(N_{t+1} = m / N_1 = n_1, \dots, N_t = n_t) = \int_0^{+\infty} \frac{e^{-\lambda} \lambda^m}{m!} h^{n_1, \dots, n_t}(\lambda) d\lambda$$

qui est une loi de Poisson-mélange $P(\Lambda')$ où la v.a.r. Λ' a pour densité $h^{n_1, \dots, n_t}(\lambda)$.

Notons que cette probabilité ne dépend de la sinistralité antérieure que par le nombre total des sinistres au cours des t dernières années $\sum_{i=1}^t n_i$.

Autrement dit :

$$P(N_{t+1} = m / N_1 = n_1, \dots, N_t = n_t) = P\left(N_{t+1} = m / \sum_{i=1}^t N_i = n\right) \text{ pour tout t-uplet } (n_1, \dots, n_t)$$

tel que $n_1 + \dots + n_t = n$.

Cas particulier : loi Poisson-Gamma

Si Λ suit une loi $\gamma(r, \alpha)$ de densité $h(\lambda) = \frac{\alpha^r}{\Gamma(r)} e^{-\alpha\lambda} \lambda^{r-1}$ pour $\lambda > 0$, nous en déduisons aisément que :

$$h^{n_1, \dots, n_t}(\lambda) \propto \lambda^{r + \sum_{i=1}^t n_i - 1} e^{-(\alpha + t)\lambda}$$

qui est la densité d'une loi $\gamma\left(r + \sum_{i=1}^t n_i, \alpha + t\right)$.

La loi prédictive de N_{t+1} sachant $N_1 = n_1, \dots, N_t = n_t$ est alors une loi Binomiale Négative

$$BN\left(r + \sum_{i=1}^t n_i, \frac{\alpha + t}{\alpha + t + 1}\right).$$

b) Prime pure a posteriori et indice de fréquence

Dans le modèle composé (sous les hypothèses de base), le montant cumulé du risque considéré pour la $(t+1)^{i\text{e}}$ année est :

$S_{t+1} = \sum_{j=0}^{N_{t+1}} X_j^{(t+1)}$ avec $(X_j^{(t+1)})_{j \geq 1}$ la suite des montants de sinistres individuels de la $(t+1)^{i\grave{e}}$ année. La prime pure s'écrit :

$$E(S_{t+1}) = E(N_{t+1}) \cdot E(X^{(t+1)})$$

avec $X^{(t+1)}$ la v.a.r. parente montant d'un sinistre en $(t+1)^{i\grave{e}}$ année.

De la même manière, la prime pure a posteriori sachant $N_1 = n_1, \dots, N_t = n_t$ s'écrit :

$$E^{N_1=n_1, \dots, N_t=n_t}(S_{t+1}) = E^{N_1=n_1, \dots, N_t=n_t}(N_{t+1}) \cdot E(X^{(t+1)})$$

en faisant l'hypothèse que les fréquences de la sinistralité antérieure n'influe pas sur le coût moyen de la $(t+1)^{i\grave{e}}$ année.

Définition :

Nous définissons l'indice de fréquence de la $(t+1)^{i\grave{e}}$ année

$$I_{t+1}(n_1, \dots, n_t) = \frac{E^{N_1=n_1, \dots, N_t=n_t}(N_{t+1})}{E(N_{t+1})}$$

(qui n'est autre que le rapport de la prime pure a posteriori $E^{N_1=n_1, \dots, N_t=n_t}(S_{t+1})$ et la prime pure a priori $E(S_{t+1})$).

En conditionnant par Λ et par les hypothèses (H1) et (H2) :

$$E(N_{t+1}) = E[E(N_{t+1} / \Lambda)] = E(\Lambda) \text{ et}$$

$$E^{N_1=n_1, \dots, N_t=n_t}(N_{t+1}) = E^{N_1=n_1, \dots, N_t=n_t} [E^{N_1=n_1, \dots, N_t=n_t}(N_{t+1} / \Lambda)] = E^{N_1=n_1, \dots, N_t=n_t}(\Lambda)$$

d'où

$$\boxed{I_{t+1}(n_1, \dots, n_t) = \frac{E^{N_1=n_1, \dots, N_t=n_t}(\Lambda)}{E(\Lambda)}},$$

c'est la formule que nous utiliserons en pratique.

Nous obtenons la prime pure de la $(t+1)^{i\grave{e}}$ année sachant $N_1 = n_1, \dots, N_t = n_t$:

$$E^{N_1=n_1, \dots, N_t=n_t}(S_{t+1}) = I_{t+1}(n_1, \dots, n_t) \cdot E(S_{t+1})$$

avec $I_{t+1}(n_1, \dots, n_t)$ ne dépendant que de t et $\sum_{i=1}^t n_i$ et

$E(S_{t+1})$ la prime pure de référence du risque de la $(t+1)^{i\grave{e}}$ année.

Application à la loi de Poisson-Gamma :

Si Λ suit une loi $\gamma(r, \alpha)$ alors nous avons

$$E(\Lambda) = \frac{r}{\alpha} \text{ et } E^{N_1=n_1, \dots, N_t=n_t}(\Lambda) = \frac{r + \sum_{i=1}^t n_i}{\alpha + t}$$

(car la loi de densité de $\Lambda / N_1 = n_1, \dots, N_t = n_t$ est une loi $\gamma\left(r + \sum_{i=1}^t n_i, \alpha + t\right)$),

d'où

$$I_{t+1}(n_1, \dots, n_t) = \frac{1 + \frac{\sum_{i=1}^t n_i}{r}}{1 + \frac{t}{\alpha}}$$

remarque :

Avec les paramètres de la loi Binomiale Négative $BN(r, p)$ la formule de l'indice de fréquence de la $(t+1)^{i\grave{e}}$ année se réécrit en remplaçant α par $\frac{p}{1-p}$.

2. Application du Bonus Malus à une flotte de n véhicules

Notre objet d'étude étant les flottes, nous souhaitons appliquer ce modèle de Bonus Malus à un parc de véhicules et déterminer un indice de fréquence pour la $(t+1)^{i\grave{e}}$ année en fonction de sa sinistralité observée.

Soit une flotte avec un nombre constant n de véhicules supposés « identiques et indépendants » sur les différents exercices. La survenance des sinistres pour chacun d'entre

eux suit une loi Binomiale Négative $BN(r, p)$. Nous déduisons facilement la distribution de la fréquence des sinistres au sein de cette flotte par semi-additivité : elle suit une loi Binomiale Négative $BN(n.r, p)$.

La formule de l'indice de fréquence de la $(t+1)^{i\grave{e}}$ année s'écrit alors :

$$I_{t+1}(n_1, \dots, n_t) = \frac{1 + \frac{\sum_{i=1}^t n_i}{n.r}}{1 + \frac{t(1-p)}{p}}$$

avec n_i le nombre de sinistres observés de la flotte l'année i , pour $i = 1, \dots, t$.

Hors en réalité, les grosses flottes ont un parc fluctuant (le nombre de véhicules n n'est pas constant sur les différents exercices), et le nombre de mois d'observation de l'année t peut être inférieur à 12 mois.

La pratique dans ce cas est d'extrapoler la sinistralité de l'année en cours pour obtenir une estimation du nombre de sinistres sur l'année complète.

Nous adopterons ensuite l'approche suivante dans notre étude des cas pratiques.

Nous faisons alors les hypothèses suivantes :

- n (le nombre constant de véhicules) dans le terme du numérateur de l'expression ci-dessus devient n' le nombre moyen de véhicules sur les t exercices et est déterminé par la relation :

$$n' = \frac{mot_1 + \dots + mot_{t-1} + mot_t * K}{t}$$

- le nombre de sinistres total sur les t années s'écrit :

$$n_1 + \dots + n_{t-1} + n_t * K$$

avec

$$K = \frac{\text{nombre de mois année } t}{12},$$

n_i le nombre de sinistres observés de la flotte l'année i , pour $i = 1, \dots, t$, avec n_t le nombre extrapolé de sinistres sur l'année complète t

mot_i le nombre de moteurs de l'année i , pour $i = 1, \dots, t$.

La formule se réécrit :

$$I_{t+1} = \frac{1 + \frac{t}{r} \frac{n_1 + n_2 + \dots + n_t * K}{mot_1 + mot_2 + \dots + mot_t * K}}{1 + t \frac{1-p}{p}}$$

III. Théorie de la fluctuation limitée

La théorie de la fluctuation limitée ou crédibilité américaine (appelée aussi crédibilité de stabilité) a été fondée par Mowbray (1914) en 1914 puis Whitney (1918) en 1918 ; elle donne des éléments de réponse à la question de l'importance à accorder à l'expérience. Cette crédibilité augmente avec les volumes (expositions, nombres de sinistres).

Nous en rappelons les fondements théoriques et donnons quelques résultats de son application à des flottes de véhicules. Lors des applications numériques dans le paragraphe suivant, nous en tirerons des indications sur la pertinence de l'approche actuelle de Covéa Fleet vis à vis des grosses flottes automobiles.

Le principe de la crédibilité est de tarifier un contrat pour la $(t+1)^{i\grave{e}}$ période par :

$$P = z_t \cdot \overline{p_t} + (1 - z_t) \cdot \hat{p}$$

avec :

z_t est appelé le facteur de crédibilité ($0 \leq z_t \leq 1$),

$\overline{p_t}$ le montant moyen des sinistres observés sur les t périodes,

$\hat{p}$ le montant moyen estimé des sinistres (il s'agira d'un montant moyen estimé à partir d'un groupe englobant le risque étudié).

Deux situations sont possibles :

- $z_t = 1$, la crédibilité est totale : l'expérience acquise suffit à elle seule pour la tarification,
- $z_t \in]0;1[$, il s'agit de crédibilité partielle.

Notations :

Nous gardons les mêmes notations utilisées dans les chapitres précédents, soient

S la v.a. montant cumulé des sinistres d'un risque donné sur une période fixée,
 $(X_i)_{i \geq 1}$ la suite des v.a. montants individuels d'un sinistre,
 N la v.a. nombre de sinistres du risque au cours de la période considérée.

1. Crédibilité totale

La crédibilité totale sera accordée si la probabilité p que le montant total des sinistres soit dispersé faiblement autour de la moyenne à un niveau choisi k , est suffisamment élevée.

De façon plus formelle, soient :

k est le facteur de dispersion autour de la moyenne,

p est la probabilité que le montant total des sinistres soit à $\pm k\%$ de la moyenne

si la relation suivante est vérifiée :

$$P\left(\left|\frac{S - E(S)}{E(S)}\right| \leq k\right) \geq p$$

nous parlerons alors de crédibilité complète d'ordre (k, p) et la tarification reposera uniquement sur l'expérience propre du risque.

Supposons qu'il y a un grand nombre de sinistres (cas des grosses flottes), le théorème central-limite permet alors d'approcher la distribution de la v.a. $\frac{S - \mu_s}{\sigma_s}$ par une loi normale centrée réduite $N(0,1)$. Nous obtenons une approximation de la relation précédente par :

$$\frac{E(S)^2}{V(S)} \geq \left(\frac{\Phi^{-1}\left(\frac{p+1}{2}\right)}{k} \right)^2$$

avec Φ la fonction de répartition de la loi normale centrée réduite.

2. Application à différentes distributions

Nous énonçons ci-dessous les conditions pour avoir une pleine crédibilité d'ordre (k, p) selon la distribution soit de la v.a. montant total des sinistres S , soit de la v.a. nombre de sinistres N .

Cas de la loi Binomiale

S suit une loi Binomiale de paramètres (n, π) .

La crédibilité est totale d'ordre (k, p) si n vérifie :

$$n \geq \left(\frac{\Phi^{-1}\left(\frac{p+1}{2}\right)}{k} \right)^2 \times \frac{1-\pi}{\pi}$$

Cas de la loi Poisson

Nous supposons que la v.a. nombre de sinistres N suit une loi de Poisson de paramètre λ .

Nous avons une crédibilité complète d'ordre (k, p) si λ vérifie :

$$\lambda \geq \left(\frac{\Phi^{-1}\left(\frac{p+1}{2}\right)}{k} \right)^2 \times \left[1 + \frac{V(X)}{E(X)^2} \right]$$

Cas de la loi Binomiale Négative

Nous supposons que la v.a. nombre de sinistres N suit une distribution Binomiale Négative de paramètres (r, π) .

Nous avons une pleine crédibilité d'ordre (k, p) si r vérifie :

$$r \geq \left(\frac{\Phi^{-1}\left(\frac{p+1}{2}\right)}{k} \right)^2 \times \left(\frac{1}{1-\pi} + \frac{\pi}{1-\pi} \times \frac{V(X)}{E(X)^2} \right)$$

3. Crédibilité partielle

Lorsque le volume de risque requis n'est pas atteint, la prime de l'assuré est alors une pondération entre la prime d'expérience et la prime collective (coût estimé).

Nous avons :

$$P_i = z_i \cdot \bar{X}_i + (1 - z_i) \cdot E_i$$

avec

z_i le facteur de crédibilité partielle,

P_i la prime de l'assuré i ,

$\bar{X}_i$ le coût actuel des sinistres de l'assuré i et

E_i le coût estimé.

Les formules suivantes sont proposées pour le facteur de crédibilité partielle. Soit n_0 le seuil de pleine crédibilité vu précédemment (que nous noterons respectivement λ_0 dans le cas de la loi Poisson et r_0 dans le cas de la loi Binomiale Négative), alors :

- $z = \min\left(\sqrt{\frac{n}{n_0}}, 1\right)$ qui est la formule la plus utilisée,

- $z = \min\left(\left(\frac{n}{n_0}\right)^{2/3}, 1\right)$ et

- $z = \frac{n}{n + K}$, formule de Whitney où K est une constante fixée telle que les variations de primes soient limitées.

4. Remarques sur la crédibilité américaine

L'intérêt évident de cette théorie est sa simplicité de mise en œuvre. Les inconvénients sont néanmoins d'une part l'approximation par le théorème central-limite et le choix un peu aléatoire des deux paramètres k et p (pour l'ordre de crédibilité (k, p)) pour lequel la justification est délicate.

IV. Applications numériques

L'application de la théorie de la stabilité nous suggère de nous restreindre à l'étude des flottes de taille « suffisamment » grande afin d'avoir une crédibilité totale pour un certain seuil retenu. Nous donnons plus de détails sur ce point ci-après.

Ainsi pour les applications numériques, nous sélectionnons dix contrats de différentes tailles (nombre de véhicules moteurs compris entre 339 et 3237 véhicules) et de différentes fréquences de sinistralité afin d'obtenir un échantillon reflétant la diversité du portefeuille de grosses flottes automobiles.

Nous souhaitons ensuite, pour chaque contrat comparer la Prime Pure calculée par Covéa Fleet avec la charge sinistre a posteriori, que nous nommerons Prime Pure modélisée, évaluée à partir des paramètres déterminés dans le chapitre précédent et de la sinistralité observée de la flotte.

1. Quelques précisions

Concernant la modélisation de la charge sinistre RC d'une flotte :

Les paramètres des lois de fréquence de sinistralité déterminés dans le chapitre 2 se rapportent à un véhicule moteur. Pour nos contrats, les paramètres correspondants seront bien entendu calculés par rapport au nombre de véhicules moteurs des flottes par simple additivité ou semi additivité des lois de Poisson et Binomiale Négative.

- Les sinistres RC hors graves

Nous en déduisons aisément l'espérance de la charge sinistre RC Hors Graves a priori à partir de la connaissance des paramètres des distributions des variables aléatoires fréquence N et coût de sinistre X , à laquelle ensuite nous appliquons le coefficient bonus malus correspondant. Ce dernier est déterminé en fonction de la sinistralité antérieure connue sur 3 (ou éventuellement 4) périodes.

- Les sinistres RC graves

Le nombre espéré de sinistres graves est calculé à partir de l'espérance du nombre de sinistres hors graves de la flotte à laquelle nous appliquons le taux λ_1 décrit dans le chapitre précédent qui correspond au rapport de l'espérance de la fréquence de sinistres graves sur l'espérance de la fréquence de sinistres hors graves pour un véhicule. De ce fait, la prise en compte de la sinistralité antérieure est donc implicite dans le calcul de la charge sinistre RC graves.

Concernant le choix des flottes retenues pour les applications numériques :

Nous nous sommes restreints aux flottes dont la taille permet une crédibilité totale à un certain seuil en appliquant les formules déterminées dans le paragraphe précédent. Nous

avons retenu un ordre de crédibilité de (20%;80%) pour lequel nous souhaitons que la flotte puisse être tarifiée par sa propre expérience.

Le choix de cet ordre de crédibilité résulte du fait qu'un seuil supérieur réduirait de façon très considérable le nombre de contrats qui « resteraient » concernés par l'approche actuelle de Covéa Fleet (ainsi les flottes de taille supérieure à 300 moteurs constituent environ 35% des grosses flottes totales).

Nous nous basons sur les paramètres déterminés dans le cadre de notre étude lors de la modélisation dans le second chapitre, pour évaluer la taille minimum à laquelle doit répondre les flottes pour une crédibilité totale.

Bien entendu, ces paramètres correspondent à une distinction des sinistres de fréquence et des sinistres graves au seuil de 75K euros, ce qui diffère d'un écrêtement actuel de 30K. Les résultats que nous obtenons nous permettent cependant d'avoir une indication sur l'ordre de grandeur de la taille à retenir.

Applications numériques

La v.a. fréquence de sinistres N pour un moteur suit une distribution Binomiale Négative de paramètres (0,65425239;0,887778987).

Si la flotte comprend au minimum :

- soit 300 moteurs avec 4 années d'expérience
- soit 400 moteurs avec 3 années d'expérience

alors la crédibilité est totale d'ordre (20%;80%).

En choisissant des ordres de crédibilité autres, nous avons :

ordre de crédibilité (k;p)	nbre moteurs si 3 ans exp	nbre moteurs si 4 ans exp
(20%;90%)	660	495
(10%;80%)	1600	1200
(10%;90%)	2640	1980

Au vu de ces résultats nous pouvons conclure que le principe actuel de tarification de Covéa Fleet est valide pour les flottes comportant un minimum de 300 moteurs si l'expérience sur 4 années est connue, sinon un minimum de 400 moteurs pour une sinistralité connue sur les 3 années précédentes. En fait ces résultats renforcent une idée qui était déjà connue et qui tendrait à relever la taille minimum actuelle d'une flotte pour une approche personnalisée (de la charge sinistre écrêtée).

Dans la suite nous confrontons la prime pure déterminée par Covéa Fleet avec la prime pure déterminée par notre modélisation.

2. Cas étudiés

Les contrats étudiés sont des contrats provenant de notre portefeuille TPPC segment 4 qui sont passés en comité de souscription et de renouvellement lors de l'année 2005. Nous avons retenu les flottes ne comportant que des véhicules de 1^{ère} catégorie.

Les flottes sont dans la suite présentées par ordre croissant de leur taille. Pour chaque contrat, les nombres de moteurs ainsi que la sinistralité observée sur les 3 ou 4 périodes (selon le cas) sont indiqués dans un tableau.

Le nombre de mois pour l'année en cours est également donné pour l'extrapolation de la sinistralité.

La fréquence moyenne de sinistralité et le coefficient bonus malus associé à chaque flotte sont calculés à partir de ces informations.

Les charges sinistres espérées avant **et** après application du coefficient bonus malus déterminées par notre modélisation sont indiquées.

Nous notons Prime pure modélisée la charge sinistre espérée après application du coefficient bonus malus.

Les écarts tarifaires par rapport aux primes pures Covéa Fleet correspondantes sont également calculés :

$$\text{Ecart tarifaire} = (\text{Prime pure modélisée} - \text{Prime pure Covéa}) / \text{Prime pure Covéa}$$

Contrat 1

Notre première flotte étudiée comprend 339 véhicules pour l'année en cours. Nous notons que pour l'année en cours N les statistiques ne sont connues que pour 9 mois. Sa sinistralité pour les années antérieures est indiquée dans le tableau ci-dessous :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>	411	372	361	339	4	9
<i>nb sin RC hg</i>	39	35	30	38		

La prime pure Covéa Fleet associée à ce contrat s'élève à 80377 euros.

La fréquence moyenne de sinistralité et le coefficient bonus malus correspondant à cette flotte sont indiqués dans le tableau suivant :

Fréqce moy	coeff BM
9,48%	1,05

Il s'agit d'une vraie fréquence moyenne dont l'année n est pondérée par son nombre de mois. Les charges sinistres attendues de cette flotte avant et après application du coefficient bonus malus sont déterminées et sont présentées dans la table ci-dessous. Nous indiquons également les écarts tarifaires respectifs avec la prime pure Covéa Fleet.

		écart avec PP comité
Espérance de la charge sans BM	74904	-6,81%
Espérance de la charge avec BM	78675	-2,12%

L'application du coefficient bonus malus à la prime espérée donne une meilleure estimation de la charge sinistre réellement attendue compte tenu de l'antériorité sinistre de la flotte. La comparaison entre cette espérance de la charge sinistre avec bonus malus et la prime pure déterminée par Covéa Fleet donne un écart tarifaire de - 2,12% , ce qui est un très bon résultat et nous permet donc de considérer que les deux approches aboutissent à une prime quasi-équivalente.

Contrat 2

Cette flotte comprend 356 moteurs.

Les informations concernant le nombre de moteurs sur les années précédentes et sa sinistralité antérieure sont présentées ci-dessous :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>	365	368	346	356	4	5
<i>nb sin RChg</i>	27	24	23	20		

La prime pure déterminée par Covéa Fleet pour cette flotte est de 67249 euros.

Nous indiquons la fréquence moyenne de sinistres de cette flotte et le coefficient bonus malus associé à ce contrat dans le tableau suivant :

Fréqce moy	coeff BM
6,71%	0,94

Les charges sinistres moyennes avant et après application du coefficient bonus malus, ainsi que les écarts tarifaires respectifs avec la prime pure de Covéa Fleet, sont donnés ci-après.

		écart avec PP comité
Espérance de la charge sans BM	78632	16,93%
Espérance de la charge avec BM	73730	9,64%

La charge espérée de la flotte sans prise en compte de sa sinistralité antérieure est nettement plus élevée que la prime pure Covéa. Après application du coefficient bonus malus, l'écart tarifaire se réduit et s'élève à 9,64%.

L'impact du coefficient bonus malus permet d'obtenir un résultat acceptable.

Contrat 3

Nous avons une flotte constituée de 451 véhicules moteurs, dont le détail du parc et la sinistralité sur les 3 dernières années sont reportés dans le tableau suivant :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>	455	455	482	451	4	8
<i>nb sin RC hg</i>	51	46	45	53		

La prime pure Covéa Fleet est de 107614 euros.

Les valeurs suivantes pour la fréquence moyenne de sinistres et le coefficient bonus malus ont été déterminées :

Fréqce moy	coeff BM
10,48%	1,09

Les charges sinistres espérées avant et après application du coefficient bonus malus de cette flotte et les écarts tarifaires associés entre ces dernières et la prime pure déterminée par Covéa Fleet sont indiqués ci-dessous :

		écart avec PP comité
Espérance de la charge sans BM	99985	-7,09%
Espérance de la charge avec BM	108953	1,24%

De la même manière, l'application du coefficient bonus malus permet d'obtenir un écart tarifaire très faible entre la prime modélisée et la prime pure.

Les deux approches pour l'obtention de la prime pure de la flotte donne ainsi des résultats très similaires.

Contrat 4

La taille de cette flotte est de 589 véhicules moteurs.

Nous indiquons le détail de son parc automobile et sa sinistralité passée sur les précédentes périodes dans le tableau ci-après :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>		662	581	589	3	10
<i>nb sin RC hg</i>		60	43	50		

La prime pure Covéa associée à cette flotte s'élève à 125403 euros.

La fréquence moyenne de sinistralité de cette flotte et le coefficient bonus malus associé sont indiqués dans le tableau ci-après :

Fréqce moy	coeff BM
8,34%	1,00

Le tableau suivant présente les charges sinistres attendues avant et après application du coefficient bonus malus, ainsi que les écarts tarifaires respectifs par rapport à la prime pure Covéa Fleet.

		écart avec PP comité
Espérance de la charge sans BM	<i>130489</i>	<i>4,06%</i>
Espérance de la charge avec BM	<i>130841</i>	<i>4,34%</i>

Cette flotte présente une sinistralité très proche de celle qui a été modélisée à partir du portefeuille. Son coefficient bonus malus est donc égal à 1, et l'écart tarifaire avec la prime pure Covéa est relativement faible et s'élève à 4,34%.

Contrat 5

Il s'agit d'une flotte comportant 611 véhicules moteurs.

Son parc de véhicules et sa sinistralité sur les périodes précédentes sont rappelés dans le tableau ci-dessous :

	n-3	n-2	n-1	n	t	rbmoisN
<i>rbmoteurs</i>		585	592	611	3	8
<i>rb sin RChg</i>		59	70	88		

La prime pure Covéa de ce contrat est de 141190 euros.

Nous indiquons la fréquence moyenne de sinistres de la flotte et le coefficient bonus malus déterminé dans le tableau suivant :

Fréqce moy	coeff BM
11,85%	1,12

L'espérance de la charge sinistre et l'écart tarifaire entre cette dernière et la prime pure Covéa sont présentés dans le tableau qui suit, avant et après application du coefficient bonus malus.

		écart avec PP comité
Espérance de la charge sans BM	<i>135234</i>	-4,22%
Espérance de la charge avec BM	<i>151407</i>	7,24%

Cette flotte présente une sinistralité plus élevée que la moyenne empirique du portefeuille, l'application du coefficient bonus malus qui est égal à 1,12 donne une prime modélisée qui est plus élevée que la prime Covéa de 7,24%.

Contrat 6

Il s'agit ici d'une flotte de 813 véhicules moteurs.

Dans le tableau ci-dessous sont repris le détail du parc et la sinistralité antérieure de ce contrat lors des années précédentes :

	n-3	n-2	n-1	n	t	rbmisN
<i>rbmoteurs</i>		752	752	813	3	8
<i>rb sinRChg</i>		53	37	95		

La prime pure Covéa de cette flotte s'élève à 161935 euros.

Nous avons dans le tableau suivant la fréquence moyenne de sinistres de la flotte, ainsi que le coefficient bonus malus déterminé à partir de sa sinistralité :

Fréqce moy	coeff BM
7,49%	0,97

Les charges sinistres espérées avant et après application du coefficient bonus malus et les écarts tarifaires associés sont donnés ci-dessous :

		écart avec PP comité
Espérance de la charge sans BM	179973	11,14%
Espérance de la charge avec BM	175407	8,32%

L'application du coefficient bonus malus égal à 0,97 permet de réduire l'écart tarifaire entre la prime pure Covéa et la prime modélisée. Cette dernière est plus élevée que la prime pure Covéa et l'écart s'élève alors à 8,32%. Nous notons que ce contrat présente une bonne fréquence de sinistralité.

Contrat 7

La flotte est constituée de 1071 véhicules.

Son parc automobile et sa sinistralité antérieure sur les années précédentes sont indiqués dans le tableau qui suit :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>		848	926	1071	3	7
<i>nb sin RC hg</i>		73	129	162		

La prime pure Covéa pour ce contrat s'élève à 275408 euros.

Nous présentons ci-dessous la fréquence moyenne de la flotte et le coefficient bonus malus associé :

Fréqce moy	coeff BM
12,36%	1,14

L'espérance du montant cumulé des sinistres avant et après application du coefficient bonus malus, ainsi que les écarts tarifaires correspondants sont indiqués dans le tableau suivant.

		écart avec PP comité
Espérance de la charge sans BM	237252	-13,85%
Espérance de la charge avec BM	269615	-2,10%

La flotte présente une fréquence de sinistralité élevée et sans application du coefficient bonus malus, l'écart avec la prime pure Covéa s'élève à -13,85%. L'application du coefficient bonus malus permet de réduire nettement l'écart tarifaire qui n'est plus que de -2,10%, ce qui réconcilie les deux primes Covéa et modélisée.

Les deux approches pour l'obtention de la prime pure donnent des résultats similaires.

Contrat 8

La flotte comprend 1409 véhicules moteurs.

Nous indiquons son parc et sa sinistralité antérieure dans le tableau suivant :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>		1131	1279	1409	3	8
<i>nb sin RC hg</i>		146	168	207		

La prime pure Covéa Fleet s'élève à 401186 euros.

Nous présentons ci-dessous la fréquence moyenne de la sinistralité de la flotte et le coefficient bonus malus déterminé à partir de cette sinistralité :

Fréqce moy	coeff BM
13,50%	1,17

Les espérances de la charge sinistre avant et après application du coefficient bonus malus sont indiqués dans le tableau qui suit. Nous présentons également les écarts tarifaires par rapport à la prime pure Covéa Fleet.

		écart avec PP comité
Espérance de la charge sans BM	<i>312156</i>	-22,19%
Espérance de la charge avec BM	<i>366510</i>	-8,64%

La fréquence de sinistralité de cette flotte est très élevée, et nous constatons que la charge sinistre espérée avant application du coefficient bonus malus est très en-dessous de la prime pure Covéa. L'application du coefficient bonus malus va donner une meilleure estimation de la charge sinistre réellement attendue compte tenu de l'antériorité sinistre de la flotte : l'écart tarifaire avec la prime Covéa se réduit et s'élève à -8,64%.

La comparaison entre les primes des deux approches donne un résultat assez satisfaisant après intervention du coefficient bonus malus.

Contrat 9

La flotte est constituée de 1805 véhicules moteurs.

Ses antécédents sinistres et son parc automobile antérieur sur les périodes précédentes sont rappelés dans le tableau qui suit :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>	1858	1805	1805	1805	4	7
<i>nb sin RC hg</i>	214	243	222	230		

La prime pure déterminée par le comité de souscription est de 477476 euros.

Nous présentons la fréquence moyenne de sinistralité de la flotte et son coefficient bonus malus ci-après :

Fréqce moy	coeff BM
12,47%	1,17

Le tableau suivant donne les valeurs des charges sinistres espérées avant et après application du coefficient bonus malus, ainsi que les écarts tarifaires (par rapport à la prime Covéa) correspondants.

		écart avec PP comité
Espérance de la charge sans BM	399939	-16,24%
Espérance de la charge avec BM	468307	-1,92%

Sans prise en compte de sa sinistralité, la charge sinistre attendue de la flotte est très inférieure à la prime pure Covéa : l'écart tarifaire est de -16,24% !

Nous constatons clairement que l'application du coefficient bonus malus associé à la flotte permet de réduire nettement l'écart tarifaire avec la prime pure Covéa : il s'élève à -1,92%. Cet écart très faible réconcilie de nouveau l'approche actuelle avec notre modélisation.

Contrat 10

Cette flotte est la plus grande de notre échantillon et est constituée de 3237 véhicules moteurs. Le tableau suivant rappelle son parc moteur et sa sinistralité sur les périodes précédentes :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>		3074	3074	3237	3	8
<i>nb sin RC hg</i>		366	376	401		

La prime pure évaluée de Covéa Fleet est de 793467 euros.

Ci-dessous, sont présentés la fréquence moyenne de la flotte ainsi que le coefficient bonus malus associé :

Fréqce moy	coeff BM
12,15%	1,13

Les charges sinistres espérées avant et après application du coefficient bonus malus et les écarts tarifaires associés sont indiqués dans le tableau :

		écart avec PP comité
Espérance de la charge sans BM	<i>717519</i>	-9,57%
Espérance de la charge avec BM	<i>810227</i>	2,11%

La flotte présente une fréquence de sinistralité élevée et sans application du coefficient bonus malus, la charge espérée est plus faible que la prime pure Covéa : l'écart avec la prime pure Covéa s'élève à -9,57%. L'application du coefficient bonus malus permet de réduire nettement l'écart tarifaire qui n'est plus que de 2,11%, ce qui réconcilie les deux primes Covéa Fleet et modélisée.

Les deux approches pour l'obtention de la prime pure donnent donc des résultats très identiques.

3. Commentaires

La concordance des primes pures évaluées par le « comité de souscription » et par notre modélisation permet donc de « valider » :

- d'une part l'approche actuelle même de Covéa Fleet,
- d'autre part notre approche modélisation, plus sophistiquée que la précédente, mais donnant des résultats similaires

pour la détermination de la charge sinistre attendue d'une flotte, compte tenu de son antériorité sinistre.

Nous récapitulons dans le tableau ci-dessous pour chaque flotte la taille de son parc, sa fréquence moyenne de sinistralité, le coefficient bonus malus associé, et l'écart tarifaire entre la prime modélisée et la prime Covéa Fleet.

Contrat	taille flotte	Fqce Moy	Coeff BM	écart
1	339	9,48%	1,0503	-2,12%
2	356	6,71%	0,9376	9,64%
3	451	10,48%	1,0897	1,24%
4	589	8,34%	1,0027	4,34%
5	611	11,85%	1,1196	7,24%
6	813	7,49%	0,9746	8,32%
7	1071	12,36%	1,1364	-2,10%
8	1409	13,50%	1,1741	-8,64%
9	1805	12,47%	1,1709	-1,92%
10	3237	12,15%	1,1292	2,11%

Nous remarquons que l'écart tarifaire entre la prime pure Covéa Fleet et la prime modélisée est très variable selon la flotte, et semble être impacté par la fréquence moyenne de sinistralité de cette dernière.

Dans un deuxième tableau les flottes sont ré-ordonnées selon leur fréquence moyenne de sinistralité et nous constatons aisément des divergences croissantes au niveau des écarts tarifaires lorsque les fréquences moyennes de sinistralité deviennent « extrêmes ». Ainsi les divergences sont plus élevées pour les flottes dont la sinistralité est très bonne ou inversement très mauvaise.

Contrat	taille flotte	Fqce Moy	Coeff BM	écart
2	356	6,71%	0,9376	9,64%
6	813	7,49%	0,9746	8,32%
4	589	8,34%	1,0027	4,34%
1	339	9,48%	1,0503	-2,12%
3	451	10,48%	1,0897	1,24%
5	611	11,85%	1,1196	7,24%
10	3237	12,15%	1,1292	2,11%
7	1071	12,36%	1,1364	-2,10%
9	1805	12,47%	1,1709	-1,92%
8	1409	13,50%	1,1741	-8,64%

Nous avons ainsi un écart tarifaire de 9,64% pour une très bonne fréquence de sinistralité de 6,71% et un écart négatif de -8,64% pour une très mauvaise fréquence égale à 13,50%.

Les divergences observées aux bords ne nous surprennent pas ; elles s'expliquent par des approximations faites dans nos approches, consécutives aux compromis entre la théorie et la mise en pratique (par exemple : un ordre de crédibilité de (20%;80%) pour l'approche actuelle de Covéa Fleet, une segmentation au niveau de la modélisation pas assez fine ?, une personnalisation basée uniquement sur la fréquence ne prenant pas en compte le coût, ...)

Dans ce chapitre, l'utilisation des résultats du chapitre 2 ainsi que l'application de la théorie du bonus malus aux flottes nous a permis de comparer et de « valider » deux approches pour la détermination de la prime pure d'une grosse flotte automobile en prenant en considération sa sinistralité. Nous rappelons ces approches ci-après :

- l'approche basée uniquement sur l'expérience pour les sinistres écrêtés à 30K euros avec une mutualisation de la sur crête qui fournit la prime pure Covéa Fleet
- l'approche « modélisation » qui utilise les paramètres déterminés par la modélisation de la sinistralité (en distinguant les sinistres de fréquence des sinistres graves) pour un véhicule et prend en compte la sinistralité observée d'une flotte par l'application d'un coefficient bonus malus

De surcroît, l'application de la théorie de la crédibilité de stabilité permet de donner un ordre de grandeur concernant la taille minimum d'une flotte pour qu'elle soit tarifée uniquement à partir de son expérience propre avec un ordre de crédibilité de (20%;80%).

Les écarts tarifaires entre la prime pure Covéa et la prime modélisée étant suffisamment faibles, bien que tendant à diverger aux bords lorsque la fréquence de sinistralité est très mauvaise ou très bonne, nous réconcilions les deux approches qui permettent de déterminer de façon satisfaisante la charge sinistre attendue de la flotte.

L'étape suivante consiste alors à simuler la charge sinistre RC d'une flotte et de déterminer sa distribution. Différents niveaux de conservation seront appliqués aux contrats, et dans la suite de notre étude, nous évaluerons les taux de chargement respectifs à appliquer à la prime pure pour couvrir le risque accru d'avoir un ratio S/P très défavorable dû à la présence d'une conservation.

CHAPITRE 4

-

IMPACT DE LA CONSERVATION SUR LA DISTRIBUTION DU RATIO S/P

Les résultats des chapitres précédents permettent la simulation de la charge sinistre RC d'une flotte avec prise en compte de ses antécédents. Nous obtenons la distribution de la variable aléatoire charge sinistre RC, et en déduisons les distributions des variables aléatoires rapports sinistres à primes dans le cas où il n'y a pas de conservation ainsi que dans le cas où un certain niveau de conservation est appliqué. Nous nous intéressons à l'impact de l'application d'une conservation sur la queue de distribution du rapport sinistres à primes et nous proposons de déterminer le taux de chargement à appliquer sur la prime pure de façon à remédier à l'augmentation du risque d'avoir un ratio S/P très défavorable.

Dans une première partie, nous commençons par la présentation de la problématique liée à la présence d'une conservation.

Nous introduisons dans une seconde partie différents indicateurs de risque que nous utiliserons pour étudier l'impact de la conservation. Ces mesures de risque présentent l'intérêt d'être d'un concept accessible à l'utilisateur non spécialiste, et de mise en œuvre relativement simple.

Nous rappelons ensuite les méthodes de simulation directe du montant cumulé des sinistres S à partir de la connaissance des premiers moments de ses composantes variables aléatoires fréquence N et coût X . Les bases théoriques de cette approche sont présentées et les cas particuliers où le modèle de fréquence suit une distribution de Poisson ainsi qu'une loi Binomiale Négative sont abordés. Pour une mise en œuvre pratique, des méthodes numériques efficaces telles que les Approximations Normal Power et Wilson Hilferty sont présentées.

Finalement, dans la partie applications numériques, nous donnerons les résultats de ces études sur les 10 cas étudiés dans le chapitre précédent. Des tableaux finaux récapituleront les taux de chargement à retenir selon le niveau de conservation appliqué et la mesure de risque choisie.

I. Problématique

Nous nous intéressons dans notre étude au ratio sinistres à primes S/P dont l'utilisation est largement répandue auprès du management comme étant le principal outil de mesure de performance.

Que ce soit dans une branche automobile sans ou avec conservation, le ratio S/P reste l'indicateur de référence pour juger de la qualité d'un contrat, et plus globalement, d'un portefeuille.

Lorsque le ratio S/P est calculé sur l'ensemble des grosses flottes automobiles (dont une proportion admet une conservation) ou même plus généralement sur l'ensemble du portefeuille, il n'est pas fait de différenciation entre les flottes avec conservation de celles sans conservation.

Hors la mise en place d'une conservation dans une grosse flotte automobile modifie la forme de la distribution du rapport sinistres à primes, que ce soit son propre ratio S/P ou plus globalement celui du portefeuille auquel il appartient.

Premières observations :

Intéressons-nous au ratio S/P d'une grosse flotte automobile avec conservation.

Notons le ratio sinistres à primes avec présence d'une conservation

$$\left(\frac{S}{P}\right)_c = \frac{S - C}{P - f(C)} \text{ lorsque la conservation est atteinte, sinon le ratio } \left(\frac{S}{P}\right)_c \text{ est égal à } 0$$

avec $f(C)$ la valeur de la prime déduite correspondante pour cette conservation C , P la prime H.T., S la charge sinistre et PP la prime pure.

Nous avons :

- si $S < PP$, alors $0,66 > \frac{S}{P} > \left(\frac{S}{P}\right)_c$ et nous avons également

$$E\left(\frac{S}{P} / \frac{S}{P} < 0,66\right) > E\left(\left(\frac{S}{P}\right)_c / \left(\frac{S}{P}\right)_c < 0,66\right)$$

- si $S = PP$, alors $\frac{S}{P} = \left(\frac{S}{P}\right)_c = 0,66$

- si $S > PP$, alors $\frac{S}{P} < \left(\frac{S}{P}\right)_c$ et nous avons également

$$E\left(\frac{S}{P} / \frac{S}{P} > 0,66\right) < E\left(\left(\frac{S}{P}\right)_c / \left(\frac{S}{P}\right)_c > 0,66\right)$$

C'est le dernier cas de figure qui nous intéressera plus particulièrement, notamment l'impact de la conservation sur la queue de distribution du ratio S/P .

Examinons maintenant d'un point de vue plus formel l'impact d'une conservation notée C sur la distribution du rapport sinistres à primes d'une grosse flotte automobile.

Espérance du ratio S/P

L'idée qui est implicitement adoptée par le « comité de souscription et de renouvellement » dans son approche de la tarification suppose que la conservation est toujours atteinte et que l'espérance de la variable aléatoire ratio S/P est invariant avec l'instauration d'une conservation.

Nous avons donc :

$E\left[\left(\frac{S}{P}\right)_c\right] = E\left(\frac{S}{P}\right) = 0,66$ en reprenant les différents taux cités auparavant qui sont appliqués à la prime pure pour obtenir la prime H.T. (0,66 est le ratio S/P d'équilibre).

Remarque :

Notons que dans la pratique il est constaté que pour moins de un contrat sur vingt, une restitution d'une partie de la conservation pour non consommation totale est effectuée à l'assuré ; la totalité des remboursements s'élève à nettement moins de 1% du montant total des conservations encaissées (0,23% en 2006). Les contrats concernés sont en réalité des vieux contrats dans le portefeuille ayant bénéficié des « largesses » du passé où les taux de conservation appliqués pouvaient atteindre 80% ; les taux ont été ramenés au fur et à mesure à des niveaux « plus sains ».

Variance du ratio S/P

Intéressons nous maintenant à la variance du rapport $\left(\frac{S}{P}\right)_c$ par rapport à celle du ratio S/P (sans conservation).

A la différence de l'espérance, nous montrons que la variance du ratio sinistres à prime avec conservation $\left(\frac{S}{P}\right)_c$ croît avec le niveau de cette dernière.

En effet elle s'écrit :

$$V\left[\left(\frac{S}{P}\right)_c\right] = V\left[\frac{S}{P}\right] \times \frac{P^2}{(P - f(C))^2}$$

Pour une conservation C_1 supérieure à une conservation C_2 nous avons alors :

$$V\left[\left(\frac{S}{P}\right)_{C_1}\right] > V\left[\left(\frac{S}{P}\right)_{C_2}\right]$$

La mise en place d'une conservation implique donc une volatilité du ratio S/P d'autant plus élevé que le niveau de conservation l'est.

Hormis cet indicateur de risque qu'est l'écart type (ou plus exactement la variance), nous souhaitons dans notre étude retenir d'autres mesures de risque, qui nous permettront notamment de nous focaliser sur la queue de distribution du ratio S/P . En effet du point de vue du manager, il paraît naturel d'adopter une approche prudentielle en se « prémunissant » des scénarii les plus défavorables.

Notre démarche consistera, *au sens d'une mesure de risque choisie*, à faire un raisonnement en terme d'iso valeurs ; pour un niveau de conservation appliquée nous faisons des

hypothèses de chargement sur la partie prime, de manière à retrouver une valeur de la mesure de risque équivalente à la valeur initiale lorsque aucune conservation n'est appliquée.

Il s'agit donc de rendre cohérent l'indicateur ratio S/P entre les flottes avec et sans conservation *au sens d'une mesure de risque* par l'instauration d'une compensation.

II. Les Mesures de risque

Nous présentons des mesures de risques qui sont généralement utilisées dans le domaine financier. Nous nous intéressons à une distribution d'une v.a. X . Cette v.a. X , dans notre étude, correspond naturellement au ratio S/P .

1. Les mesures de risque retenues

De nombreux indicateurs de risque existent, nous présentons dans la suite les mesures de risque que nous avons retenues pour leur pertinence dans notre étude : elles présentent une mise en œuvre simple, sont relativement faciles à appréhender.

a) L'écart type

L'écart type $\sigma(X)$ est une mesure statistique de la dispersion d'une distribution et est la racine carrée de la variance qui se décrit simplement comme la moyenne des carrés des écarts à la moyenne

Cette mesure de risque prend en compte la distribution entière, donnant aux événements extrêmes plus d'influence. L'utilisateur ne pourra cependant pas s'intéresser uniquement à une partie de la courbe, à la queue de distribution par exemple.

b) La Value-at-Risk (VaR)

$\alpha > 0$ étant fixé (par exemple 0,05 ; 0,005), la Value at Risk (VaR) correspondante notée $VaR_\alpha(X)$ est le quantile d'ordre $(1-\alpha)$ de X , c'est-à-dire la valeur telle que :

$$P(VaR_\alpha(X) \leq X) = \alpha$$

Formellement, elle s'écrit :

$$VaR_\alpha(X) = F_X^{-1}(1-\alpha)$$

$VaR_\alpha(X)$ est donc la valeur qui ne sera dépassée que dans α % des cas.

Cette mesure de risque est utilisée communément dans le secteur bancaire. C'est un concept simple à appréhender, et assez facile à calculer dans la pratique. Elle nous permet de nous focaliser en un point particulier mais elle ne fournit aucune information sur les valeurs au-delà de ce point. La mesure suivante y remédie.

c) La Tail Value-at-Risk (TVaR)

Pour α donné, la Tail Value at Risk est la moyenne des VaR de seuil supérieur à α et s'écrit :

$$TVaR_{\alpha}(X) = E(X / X \geq VaR_{\alpha}(X))$$

Plus formellement :

$$TVaR_{\alpha}(X) = \frac{1}{1-\alpha} \int_{\alpha}^1 VaR_{\zeta}(X) d\zeta$$

Nous avons, de manière immédiate :

$$TVaR_{\alpha}(X) \geq VaR_{\alpha}(X)$$

Cet indicateur de risque est couramment utilisé, assez simple à expliquer et à calculer. La TVaR permet donc une meilleure description du comportement de la queue de distribution de X que la mesure précédente.

d) Un ratio de performance : le ratio Oméga

Il a été introduit par Keating et Shadwick (2002) en 2002. Un avantage de cette mesure est qu'elle prend en compte la distribution entière de la variable aléatoire étudiée.

C'est une mesure qui divise les résultats en deux parties relativement à un seuil fixé L : les « bons résultats » et les « mauvais résultats ». Oméga est une fonction de ce seuil de résultat que le décideur choisit.

L'indicateur Oméga est initialement défini par rapport à une variable « rendement », les « bons résultats » étant donc les rendements élevés et réciproquement les « mauvais résultats » étant les rendements faibles. Dans notre cas, un ratio S/P élevé est un « résultat indésirable » et inversement un ratio S/P faible est un « résultat favorable ». Nous adaptons la définition à notre étude du ratio Oméga :

$$\Omega_F(L) = \frac{E[(L - X)^+]}{E[(X - L)^+]}$$

qui se réécrit :

$$\Omega_F(L) = \frac{F(L) \times [L - E(X / X \leq L)]}{(1 - F(L)) \times [E(X / X \geq L) - L]}$$

où $F(\cdot)$ est la fonction de répartition de la v.a. X et L est le seuil sélectionné.

Le numérateur de Oméga est donc la probabilité d'avoir un résultat favorable multipliée par la différence « attendue » entre le seuil et les bons résultats ; de la même manière, le dénominateur est la probabilité d'avoir un résultat indésirable multipliée par l'espérance de la différence entre les mauvais résultats et le seuil.

Cet indicateur reflète dans une seule valeur les caractéristiques favorable et défavorable de la distribution.

Pour un seuil donné, le ratio Oméga est d'autant plus favorable qu'il est élevé.

La mesure Oméga présente également la propriété suivante :

$$\text{pour } L = E(X), \Omega_F(L) = 1.$$

2. Commentaires sur les mesures de risque

L'écart type est une mesure de risque largement utilisée, facile à déterminer et prend en compte l'entière distribution. Cependant elle est moins sensible aux événements extrêmes que les autres indicateurs. Nous préférons l'utilisation des indicateurs plus élaborés suivants.

La VaR est une mesure de risque très populaire, mais présente l'inconvénient de ne pas prendre en compte le comportement de la queue de distribution de X au-delà du quantile d'ordre $(1 - \alpha)$. Elle se focalise en un unique point de la distribution, aucune information concernant la queue de distribution ou toute autre partie n'est capturée par cette mesure.

L'indicateur de risque TVaR est également assez simple d'utilisation et de concept aisé à appréhender. En ne se focalisant pas sur un unique point, elle permet d'avoir une description plus riche du risque que celle obtenue par la VaR.

La VaR et la TVaR sont des outils très utiles pour l'observation des plus mauvais scénarii. Cependant, à la différence de l'écart type, seule une partie de la distribution influe sur la mesure de risque, les bons résultats par exemple n'ayant aucun impact.

Le ratio Oméga peut être considéré plus comme une mesure de performance que de risque pur, cette notion incorporant également les résultats favorables avec les résultats défavorables.

III. Méthodes de simulation

Nous nous basons sur le modèle composé et nous souhaitons déterminer l'expression de la loi exacte et des caractéristiques de la v.a.r. charge sinistre annuelle notée S .

Nous étudierons les cas particuliers où la v.a.r. N suit une loi de Poisson et une loi Binomiale Négative.

Nous indiquerons également des méthodes numériques types pour la simulation du montant cumulé des sinistres S basée sur l'utilisation des premiers moments de S et de la loi Normale.

Soit N la v.a.r. nombre de sinistres sur la période considérée du risque considéré et $(X_i)_{i \geq 1}$ la suite des v.a.r. coûts individuels de sinistre de v.a.r. parente X .

Le montant cumulé des sinistres S s'écrit :

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

Nous rappelons les hypothèses du modèle :

- i. La v.a.r. N et la suite des v.a.r. $(X_i)_{i \geq 1}$ sont indépendantes.
- ii. Les coûts des sinistres successifs $X_1, X_2, \dots, X_i, \dots$ sont indépendants et de même loi que la v.a.r. parente X .

1. Loi exacte et moments de la charge sinistre

Nous souhaitons déterminer la fonction de répartition F_S de la charge annuelle S . Elle est définie pour tout réel x par : $F_S(x) = P(S \leq x)$

Nous montrons aisément la proposition suivante :

Proposition :

$$F_S(x) = \sum_{n=0}^{\infty} P(N = n) \cdot P\left(\sum_{i=1}^n X_i \leq x\right) \text{ si } x \geq 0, F_S(x) = 0 \text{ si } x < 0$$

Nous obtenons également les relations suivantes :

Proposition :

Les premiers moments de S sont connus à partir des premiers moments de ses composantes fréquence N et coût individuel de sinistre X et sont donnés par les relations suivantes :

$$E(S) = E(N) \cdot E(X)$$

$$V(S) = E(N) \cdot V(X) + E(X)^2 \cdot V(N)$$

$$\mu_3(S) = E(N).\mu_3(X) + 3.V(N).E(X).V(X) + \mu_3(N).E(X)^3$$

2. Application aux distributions de Poisson et Binomiale Négative

Nous supposons que les hypothèses du modèle composé sont toujours vérifiées.

Processus de Poisson :

Dans le cas particulier où N suit une loi de Poisson $P(\lambda)$, soit F la fonction de répartition de X , la variable aléatoire montant cumulé S est dite suivre une loi de Poisson composée que nous notons $PC(\lambda, F)$.

Nous déduisons facilement les moments de S par les données du paramètre λ et par la connaissance de la loi du v.a.r. coût individuel de sinistre X et de ses premiers moments. Nous avons :

$$\begin{aligned} E(S) &= \lambda.m_1, \\ V(S) &= \lambda.m_2 \text{ et} \\ \mu_3(S) &= \lambda.m_3 \end{aligned}$$

avec $\mu_3(S)$ le moment centré d'ordre 3 de la v.a.r. S et $m_k = E(X^k)$ le moment simple d'ordre k de la v.a.r. X pour tout $k > 0$.

Lors de la mise en oeuvre pratique, le coefficient d'asymétrie $\gamma_1(S)$ s'avèrera plus utile, aussi il s'écrit :

$$\gamma_1(S) = \frac{m_3}{\sqrt{\lambda.m_2^{3/2}}}$$

Nous avons également le résultat suivant :

Si S_1 suit une loi de Poisson composée $PC(\lambda_1, F_1)$ et S_2 suit une loi de Poisson composée $PC(\lambda_2, F_2)$ et si S_1 et S_2 sont indépendantes alors $S = S_1 + S_2$ suit une loi de Poisson composée $PC(\lambda, F)$ où λ et F sont donnés par

$$\lambda = \lambda_1 + \lambda_2 \quad \text{et} \quad F = \frac{\lambda_1.F_1 + \lambda_2.F_2}{\lambda_1 + \lambda_2}$$

Processus de Poisson-Gamma :

Nous supposons que N suit une loi de Poisson-Gamma de paramètres (r, α) et soient m_i les moments simples de la v.a.r. coût individuel de sinistre X .

Nous en déduisons les moments de S :

$$E(S) = n.m_1,$$

$$V(S) = n.m_2 + \frac{n^2.m_1^2}{r} \text{ et}$$

$$\gamma_1(S) = \frac{n.m_3 + 3.n^2.m_1.m_2/r + 2.n^3.m_1^3/r^2}{V(S)^{3/2}}$$

$$\text{avec } n = \frac{r}{\alpha}$$

Remarque :

Avec les paramètres de la loi $BN(r, p)$, les moments se réécrivent en remplaçant α par $\frac{p}{1-p}$.

3. Méthodes d'approximation basées sur les moments

La loi Normale joue un rôle primordial dans ces méthodes d'approximation qui se reposent également sur les premiers moments de S .

Nous considérons une variable aléatoire positive S modélisant le montant cumulé annuel des sinistres. Nous désignons par F_S sa fonction de répartition et soient μ_S et σ_S respectivement les moyenne et variance de la v.a. S .

Posons pour tout $x \geq 0$

$$y = \frac{x - \mu_S}{\sigma_S}$$

alors l'approximation normale qui consiste en :

$$F_S(x) \approx \phi(y)$$

est justifiée dans les cas présentés ci-dessous.

Approximation à l'aide d'une loi normale

Soit $(S_\alpha)_{\alpha \geq 1}$ une suite de v.a.r. composées Poissonniennes avec $S_\alpha = \sum_{j=0}^{N_\alpha} X_j^{(\alpha)}$, N_α suivant la loi de Poisson $P(\lambda_\alpha)$ et $X^{(\alpha)}$ étant la v.a.r. parente montant d'un sinistre.

Si $X^{(\alpha)}$ a, pour α suffisamment grand, des moments d'ordre 1 et 2, m_1 et m_2 , indépendants de α , nous avons :

$$E(S_\alpha) = \lambda_\alpha m_1 \text{ et } V(S_\alpha) = \lambda_\alpha m_2.$$

Théorème :

Si, quand $\alpha \rightarrow +\infty$, $\lambda_\alpha \rightarrow +\infty$, nous avons :

$$\frac{S_\alpha - \lambda_\alpha m_1}{\sqrt{\lambda_\alpha m_2}} \xrightarrow[\alpha \rightarrow +\infty]{\ell} N(0,1)$$

Ainsi ce résultat justifie que pour une distribution de Poisson composée S , lorsque la fréquence moyenne de sinistres λ est « grande », nous avons :

$$F_S(x) \cong \Phi\left(\frac{x - \mu_S}{\sigma_S}\right)$$

De la même manière, la proposition suivante pour une distribution Binomiale Négative permet une approximation analogue.

Soient $(S_\alpha)_{\alpha \geq 1}$ une suite de v.a.r. composées, N_α suivant la loi $BN(r_\alpha, p)$ et $X^{(\alpha)}$ admettant des moments d'ordre 1 et 2, m_1 et m_2 , indépendants de α . Nous avons

$$E(S_\alpha) = r_\alpha \frac{q}{p} m_1 \text{ et } V(S_\alpha) = r_\alpha \frac{q}{p} \left(m_2 + \frac{q^2}{p^2} m_1^2 \right).$$

Proposition :

Si, quand $\alpha \rightarrow +\infty$, $r_\alpha \rightarrow +\infty$, nous avons :

$$\frac{S_\alpha - E(S_\alpha)}{\sigma_{S_\alpha}} \xrightarrow[\alpha \rightarrow +\infty]{\ell} N(0,1)$$

Remarque :

Dans la pratique, cette approximation s'avère insuffisante car la queue de distribution est sous estimée. Elle est donc utilisée dans les cas où l'asymétrie de la distribution est très faible.

Pour améliorer les résultats en queue de distribution, une transformation préalable des variables est opérée pour symétriser la distribution avant approximation par la loi normale. Nous présentons les deux types d'approximations ci-après.

Approximation de type Normal Power

L'idée est de transformer la v.a.r. S afin que la nouvelle distribution soit moins dissymétrique que celle initiale de S , ce qui améliorera l'approximation par la loi normale.

Nous obtenons l'approximation Normal Power suivante :

$$F_S(x) \cong \Phi \left[-\frac{3}{\gamma_s} + \sqrt{\frac{9}{\gamma_s^2} + 1} + \frac{6}{\gamma_s} \left(\frac{x - \mu_s}{\sigma_s} \right) \right]$$

$$\text{pour } x \geq \mu_s - \sigma_s \left[\frac{\gamma_s}{6} + \frac{3}{2\gamma_s} \right]$$

Approximation de Wilson Hilferty

Une autre approximation est appliquée à la v.a.r. S et mène à la formule suivante :

$$F_S(x) \cong \Phi \left(c_1 + c_2 \cdot \left[\left(\frac{x - \mu_s}{\sigma_s} \right) + c_3 \right]^{1/3} \right)$$

$$\begin{aligned} c_1 &= \frac{1}{3g} - 3g \\ c_2 &= 3g^{2/3} \\ c_3 &= g \\ g &= \frac{2}{\gamma_s} \end{aligned}$$

avec

IV. Applications numériques

1. Préambule

Nous reprenons les dix contrats présentés dans le chapitre précédent dont nous simulons les charges sinistres RC correspondantes. Le nombre de scénarii effectués est de 10 000 pour chacun des dix cas étudiés.

Deux approches différentes pour les simulations des charges sont utilisées : une approche directe de la charge cumulée pour les sinistres RC hors graves et une approche individuelle pour les sinistres graves.

Nous présentons dans la suite, sans entrer dans les détails, la mise en pratique de ces approches, mais le lecteur trouvera des explications plus détaillées sur un cas d'étude particulier en annexes.

Les sinistres RC hors graves

Nous simulons directement le montant cumulé des sinistres en utilisant l'approximation de Wilson Hilferty qui s'avèrera efficace car dans tous les cas étudiés, nous avons pour leur skewness, $\gamma_s \leq 2$.

Les premiers moments du montant cumulé se déduisent facilement des paramètres de la distribution de la v.a. fréquence d'un sinistre RC hors graves et des premiers moments simples des v.a. coût RC Corporelle, coût RC Matérielle hors IDA et coût RC Matérielle IDA. Les calculs et les simulations ci-dessus sont présentés plus en détail dans la partie annexe.

Le coefficient Bonus Malus est déterminé en fonction de la sinistralité antérieure sur 3 ou 4 exercices selon l'information disponible pour chaque contrat ; il est ensuite appliqué à la v.a. montant cumulé des sinistres pour prendre en compte l'antériorité correspondante.

Les sinistres graves

La réalisation de 10 000 scénarii de la loi de Poisson de paramètre égal au produit de $\lambda_1 = 0,00232813$ par l'espérance de la v.a. nombre de sinistre « a posteriori » RC hors graves de la flotte nous fournit une distribution de la v.a. nombre de sinistres graves attendus pour cette flotte.

Soit n le nombre de réalisations de sinistres graves, alors elles seront constituées de n_0 réalisations de sinistres graves de type « ordinaire » où n_0 est réalisé par une loi Binomiale de paramètres $(n; k)$, avec k la probabilité pour qu'un sinistre grave soit « ordinaire ».

La différence $(n - n_0)$ correspond au nombre de réalisations de sinistres graves de type « non ordinaire ». Les coûts correspondants sont ensuite réalisés pour chaque sinistre.

But :

L'objectif est de ramener le risque qui est accru d'avoir un ratio S/P défavorable par l'application d'une conservation au même niveau de risque initial (sans conservation).

En d'autres termes, il s'agit de déterminer le taux de chargement nécessaire de façon à ramener la valeur de l'indicateur de risque (VaR, TVaR ou Oméga) déterminée sur la nouvelle v.a. rapport S/P avec conservation à la même valeur que celle de l'indicateur de risque de la v.a. S/P initiale.

La distribution de la v.a. rapport S/P est obtenue en faisant 10 000 tirages de la charge sinistre S , et P est la « prime hors taxe ». Pour chaque niveau de conservation appliqué, la distribution du rapport S/P correspondant est déterminée.

Les étapes pour déterminer le taux de chargement nécessaire sont les suivantes :

- pour le contrat étudié, les valeurs des mesures de risque retenues à différents niveaux de confiance (pour la VaR et la TVaR) et à un certain seuil choisi (pour le ratio Oméga) sont déterminées lorsque aucune conservation n'est appliquée.
- selon le niveau de conservation (10% à 50% de la prime pure) appliqué, les valeurs de ces mêmes mesures de risque aux mêmes niveaux de confiance et seuil sont déterminées ; ces valeurs vont traduire un nouveau niveau de risque (supposé plus élevé).
- pour chaque niveau de conservation, le taux de chargement (à appliquer à la « nouvelle » prime pure) est déterminé :
 - pour la value at risk et la tail value at risk, la détermination est immédiate car le taux de chargement est équivalent au taux d'augmentation de la valeur de la mesure de risque concernée.
 - pour le ratio Oméga, le solveur excel est utilisé pour calculer de façon très simple les taux de chargement à appliquer tels que les valeurs du ratio Oméga associées à ces taux soient égales aux valeurs initiales.

Remarque :

Il est commercialement nécessaire pour le client que l'application du taux de chargement pour cause de conservation lui reste favorable en terme de budget.

En gardant les taux de frais et de commission décrits dans le premier chapitre, les taux de chargement maxima tels que leur application reste « favorable » au client sont, en fonction du niveau de conservation :

taux de conservation	taux de chargement maximum
10,00%	4,55%
20,00%	10,25%
30,00%	17,57%
40,00%	27,33%
50,00%	40,99%

Précision :

Dans nos 10 exemples numériques, la prime P utilisée est la prime actuelle Covéa Fleet (et non pas la prime modélisée qui serait dans ce cas égale à l'espérance de la charge sinistre). En effet, au vue de la réconciliation des deux méthodes démontrée précédemment, nous gardons l'approche utilisée actuellement pour le calcul de la prime pure. Ayons donc à l'esprit que la moyenne de la charge sinistre simulée sera légèrement différente de la prime pure retenue.

2. Cas étudiés

Nous présentons sous forme de tableau, dans un premier temps les résultats concernant la mesure de risque VaR, dans un deuxième temps ceux de l'indicateur TVaR et finalement les résultats associés à l'indicateur de performance oméga.

Les dix flottes étudiées sont présentées par ordre de taille croissante. Dans chaque tableau nous indiquons selon le taux de conservation appliqué (de 10% à 50%), le taux de chargement déterminé pour le niveau de confiance ou le seuil retenu.

a) Value at Risk

Nous retenons initialement le niveau de confiance : 95% qui est un niveau usuel utilisé lors d'une approche prudentielle. Nous avons ensuite été amenés à tester des niveaux de confiance inférieurs afin de laisser le budget client minoré par rapport au budget initial sans conservation. Nous avons testé ainsi les niveaux de confiance 94%, 93%, 92%, 91% et 90%. Les tableaux ci-dessous présentent selon un niveau de conservation croissant les taux de chargement à appliquer pour des niveaux de confiance 95% et 94%. Les valeurs de la mesure de risque VaR (0,95) sont également données dans un tableau à part ci-dessous.

Nous observons comme attendu que les taux de chargement augmentent avec le niveau de conservation : plus le niveau de conservation est élevé, plus l'impact sur la queue de distribution du ratio S/P est fort.

<i>taux conservation</i>	10%	20%	30%	40%	50%
contrat 1	4,70%	10,57%	18,12%	28,19%	42,28%
contrat 2	5,43%	12,22%	20,95%	32,60%	48,90%
contrat 3	5,17%	11,64%	19,95%	31,03%	46,55%
contrat 4	5,23%	11,76%	20,16%	31,36%	47,04%
contrat 5	5,52%	12,42%	21,30%	33,13%	49,70%
contrat 6	5,61%	12,61%	21,62%	33,64%	50,45%
contrat 7	5,25%	11,81%	20,25%	31,50%	47,24%
contrat 8	4,83%	10,86%	18,61%	28,95%	43,43%
contrat 9	5,63%	12,66%	21,71%	33,76%	50,65%
contrat 10	6,21%	13,97%	23,96%	37,27%	55,90%

Niveau de confiance = 95% : taux de chargement

<i>taux de conservation</i>	0%	10%	20%	30%	40%	50%
<i>contrat</i>						
1	1,14	1,19	1,26	1,35	1,46	1,62
2	1,29	1,36	1,45	1,56	1,71	1,92
3	1,23	1,29	1,37	1,48	1,61	1,81
4	1,24	1,31	1,39	1,49	1,63	1,83
5	1,31	1,38	1,47	1,59	1,74	1,96
6	1,33	1,4	1,5	1,62	1,78	2
7	1,25	1,31	1,4	1,5	1,65	1,84
8	1,16	1,22	1,29	1,38	1,5	1,67
9	1,33	1,41	1,5	1,62	1,78	2,01
10	1,49	1,59	1,7	1,85	2,05	2,33

Niveau de confiance = 95% : valeurs de VaR(0,95)

<i>taux conservation</i>	10%	20%	30%	40%	50%
contrat 1	3,83%	8,62%	14,77%	22,98%	34,47%
contrat 2	4,26%	9,59%	16,43%	25,56%	38,34%
contrat 3	4,30%	9,67%	16,58%	25,79%	38,68%
contrat 4	4,39%	9,88%	16,94%	26,35%	39,52%
contrat 5	4,57%	10,29%	17,65%	27,45%	41,17%
contrat 6	4,76%	10,71%	18,36%	28,57%	42,85%
contrat 7	4,31%	9,71%	16,64%	25,89%	38,83%
contrat 8	4,05%	9,10%	15,61%	24,28%	36,42%
contrat 9	4,88%	10,98%	18,83%	29,29%	43,94%
contrat 10	5,58%	12,55%	21,51%	33,46%	50,19%

Niveau de confiance = 94% : taux de chargement

Les taux sont élevés, ceci s'explique par le fait que la Value at Risk se focalise sur un point de la zone des plus mauvais scénarii. Cette approche prudentielle coûte chère et se traduit par un chargement d'autant plus élevé que le niveau de confiance l'est : le scénario observé est d'autant plus défavorable que le niveau de confiance est élevé. Ainsi les taux de chargements pour un même niveau de conservation sont supérieurs pour le niveau de confiance 95% par rapport au niveau 94%.

Nous avons, pour un niveau de conservation de 50% :

- le taux de chargement par rapport à la VaR(0,95) se situe entre 42% et 56% selon le contrat,
- le taux de chargement nécessaire pour la VaR au niveau 0,94 est compris entre 34% et 50%.

Pour un niveau de confiance et un niveau de conservation donnés, les taux de chargement sont d'un même ordre de grandeur, ils fluctuent de façon relativement limitée ; nous considérons qu'ils restent assez stables.

Il se déduit de façon immédiate que l'impact du taux de chargement pour le client est défavorable en retenant un niveau de confiance de 95%. Néanmoins, nous pouvons nous restreindre à un niveau de confiance moins prudentiel et garder le bénéfice d'une minoration du budget.

Une proposition sur les taux de chargement à appliquer selon le niveau de conservation instauré, dans l'optique d'une mise en application pratique, est exposée dans le tableau suivant : en considérant que les taux sont relativement stables quelque soit le contrat pour un niveau de confiance 94%, les valeurs dans le tableau qui suit ont été déterminées en calculant la moyenne des valeurs correspondantes sur les dix contrats étudiés.

<i>Indicateur risque</i>	<i>VaR (0,94)</i>
<i>taux conservation</i>	
10%	4,49%
20%	10,11%
30%	17,33%
40%	26,96%
50%	40,44%

Taux de chargement à appliquer

En reprenant l'exemple présenté dans le premier chapitre, déterminons le « gain » pour le client avec ce taux de chargement :

Le nouveau budget TTC s'élèverait à **402 220** euros, le budget initial sans conservation étant de **403 333** euros. La minoration du budget, avec une telle approche prudentielle, s'élève donc à **0,3%**.

Dans la pratique, il est bien évidemment plus opportun commercialement de retenir un niveau de confiance inférieur à 94% pour obtenir une réduction du budget plus conséquente tout en adoptant une approche jugée suffisamment prudentielle.

Nous reprenons la même approche que celle du niveau de confiance 94% pour l'appliquer aux niveaux de confiance 93%, 92% et 91%. Pour le même exemple que précédemment, les minorations du budget sont les suivantes :

- pour le niveau de confiance 93%, la minoration du budget s'élève à 3,1%
- pour le niveau de confiance 92%, elle est de 6,1%
- pour le niveau de confiance 91%, la minoration est de 9,2%

Finalement, le niveau de confiance 90% semble être un bon compromis entre l'aspect prudentiel et l'aspect commercial ; nous obtenons le tableau suivant (déterminé de la manière décrite auparavant) :

Indicateur risque	VaR (0,9)
taux conservation	
10%	2,08%
20%	4,69%
30%	8,03%
40%	12,50%
50%	18,75%

En reprenant à nouveau le même exemple, la minoration du budget pour le client dans ce cas s'élève à 11,1%.

Pour information, les valeurs de la mesure de risque VaR(0,90) pour les dix contrats exemples sont présentées dans le tableau ci-dessous. Nous pouvons constater qu'effectivement l'influence de la conservation sur l'indicateur est moindre que pour le niveau de confiance 95% .

taux de conservation	0%	10%	20%	30%	40%	50%
contrat						
1	0,69	0,69	0,7	0,7	0,71	0,72
2	0,7	0,71	0,72	0,73	0,74	0,75
3	0,74	0,75	0,77	0,78	0,8	0,83
4	0,76	0,78	0,79	0,81	0,84	0,87
5	0,81	0,83	0,85	0,88	0,91	0,96
6	0,83	0,85	0,88	0,91	0,95	1,01
7	0,78	0,8	0,82	0,84	0,87	0,91
8	0,75	0,76	0,78	0,79	0,81	0,85
9	0,82	0,84	0,87	0,89	0,94	0,99
10	0,92	0,95	0,99	1,04	1,1	1,19

Niveau de confiance = 90% : valeurs de VaR(0,90)

b) Tail Value at Risk

Le niveau de confiance retenu initialement est, toujours pour une approche très prudentielle, 95%. Nous testons également deux autres niveaux de confiance beaucoup moins élevés de 60% et 50% de façon à ce que les taux de chargement qui leur sont associés induisent pour le client un budget final proche ou inférieur au budget initial sans conservation.

Les tableaux ci-dessous présentent, selon la part de conservation appliquée, les taux de chargement déterminés pour les niveaux de confiance 95% et 60%.

Les valeurs de la mesure de risque TVaR (0,95) sont également données dans un tableau à part ci-dessous.

<i>taux conservation</i>	10%	20%	30%	40%	50%
contrat 1	9,83%	22,12%	37,92%	58,98%	88,48%
contrat 2	9,98%	22,45%	38,48%	59,86%	89,78%
contrat 3	9,82%	22,10%	37,89%	58,94%	88,42%
contrat 4	9,81%	22,07%	37,83%	58,85%	88,27%
contrat 5	9,81%	22,06%	37,82%	58,83%	88,25%
contrat 6	9,81%	22,06%	37,82%	58,84%	88,26%
contrat 7	9,57%	21,52%	36,89%	57,39%	86,09%
contrat 8	9,35%	21,04%	36,06%	56,10%	84,15%
contrat 9	9,39%	21,13%	36,22%	56,35%	84,52%
contrat 10	9,24%	20,79%	35,64%	55,45%	83,17%

Niveau de confiance = 95% : taux de chargement

<i>taux de conservation</i>	0%	10%	20%	30%	40%	50%
<i>contrat</i>						
1	5,72	6,29	6,99	7,9	9,1	10,79
2	6,46	7,1	7,91	8,94	10,32	12,26
3	5,59	6,14	6,82	7,7	8,88	10,52
4	5,62	6,17	6,86	7,75	8,93	10,59
5	5,61	6,16	6,85	7,74	8,92	10,57
6	5,62	6,17	6,86	7,74	8,92	10,58
7	4,74	5,19	5,76	6,49	7,46	8,82
8	4,16	4,55	5,04	5,66	6,5	7,66
9	4,26	4,66	5,16	5,81	6,66	7,87
10	3,92	4,28	4,73	5,32	6,09	7,18

Niveau de confiance = 95% : valeurs de TVaR(0,95)

Le niveau de confiance 95% pour la Tail Value at Risk n'est clairement pas applicable dans la pratique en raison des taux de chargement associés beaucoup trop élevés. La prise en compte des 5% des scénarii les plus défavorables est une approche très prudentielle et cela se traduit par un chargement très élevé.

<i>taux conservation</i>	10%	20%	30%	40%	50%
contrat 1	4,81%	10,81%	18,54%	28,84%	43,26%
contrat 2	5,47%	12,30%	21,09%	32,80%	49,20%
contrat 3	4,95%	11,13%	19,08%	29,69%	44,53%
contrat 4	5,00%	11,24%	19,27%	29,97%	44,96%
contrat 5	5,11%	11,50%	19,71%	30,66%	45,98%
contrat 6	5,15%	11,60%	19,88%	30,93%	46,39%
contrat 7	4,43%	9,97%	17,08%	26,57%	39,86%
contrat 8	3,83%	8,63%	14,79%	23,00%	34,50%
contrat 9	4,26%	9,59%	16,43%	25,56%	38,34%
contrat 10	4,39%	9,87%	16,92%	26,32%	39,47%

Niveau de confiance = 60%

Dans ce dernier cas, un niveau de confiance de 60% pour la Tail Value at Risk correspond à la moyenne des 40% des scenarii les plus défavorables.

De même que pour la Value at Risk, les taux de chargement à appliquer sont logiquement d'autant plus élevés que le niveau de confiance l'est, en raison de la focalisation de la mesure de risque sur des scenarii d'autant plus défavorables.

Une proposition sur les taux de chargement à appliquer selon le niveau de conservation, pour une mise en application dans la pratique, est présentée dans le tableau suivant : en considérant que les taux sont relativement stables quelque soit le contrat pour un niveau de confiance 60%, les valeurs dans le tableau qui suit ont été déterminées en calculant la moyenne sur les valeurs correspondantes des dix contrats étudiés.

<i>Indicateur risque</i>	<i>TVaR (0,6)</i>
<i>taux conservation</i>	
10%	4,74%
20%	10,66%
30%	18,28%
40%	28,43%
50%	42,65%

Taux de chargement à appliquer

Nous constatons aisément que ces valeurs sont supérieures aux taux maxima à appliquer de façon à rester favorables au client.

En reprenant l'exemple du premier chapitre, le nouveau budget TTC pour le client avec un taux de chargement égal à 42,65% s'élèverait à **406 677** euros, le budget initial sans conservation étant de **403 333** euros. Nous avons donc une majoration du budget dans ce cas, quasiment nulle puisqu'elle s'élève à **0,8%**.

Néanmoins en prenant un niveau de confiance à 50% pour cette mesure de risque TVaR, nous obtenons de la même manière les valeurs suivantes concernant le chargement :

<i>Indicateur risque</i>	<i>TVaR (0,5)</i>
<i>taux conservation</i>	
10%	3,80%
20%	8,55%
30%	14,66%
40%	22,81%
50%	34,21%

Toujours en reprenant le même exemple, la minoration du budget dans ce cas s'élèverait à **3,4%**.

Il ne paraît pas pertinent de choisir un niveau de confiance inférieur à 50% qui consisterait d'une certaine manière à considérer que plus de la moitié des scénarii sont des scénarii défavorables.

Pour information, les valeurs de la mesure de risque TVaR(0,50) pour les dix contrats exemples sont présentées dans le tableau ci-dessous. Nous pouvons constater qu'effectivement l'influence de la conservation sur l'indicateur est nettement plus faible que pour le niveau de confiance 95% .

<i>taux de conservation</i>	<i>0%</i>	<i>10%</i>	<i>20%</i>	<i>30%</i>	<i>40%</i>	<i>50%</i>
<i>contrat</i>						
1	1	1,04	1,09	1,15	1,23	1,35
2	1,12	1,17	1,24	1,32	1,43	1,58
3	1,01	1,05	1,1	1,16	1,24	1,36
4	1,04	1,08	1,13	1,2	1,29	1,42
5	1,06	1,1	1,16	1,23	1,33	1,46
6	1,07	1,11	1,17	1,24	1,34	1,48
7	0,95	0,99	1,03	1,08	1,15	1,25
8	0,88	0,9	0,93	0,97	1,03	1,1
9	0,94	0,97	1,01	1,06	1,12	1,22
10	0,96	1	1,04	1,09	1,16	1,26

Niveau de confiance = 50% : valeurs de TVaR(0,5)

c) Ratio Oméga

Nous faisons quelques précisions sur le seuil retenu pour l'indicateur de risque (ou plutôt de performance) Oméga. Ce seuil correspond au ratio *S/P* d'équilibre, son choix est naturel et pertinent car il permet la distinction des résultats « favorables » et « défavorables » relativement à ce seuil.

Les taux de chargement déterminés en fonction du niveau de conservation appliqué pour les dix contrats étudiés sont indiqués dans le tableau suivant :

taux conservation	10%	20%	30%	40%	50%
contrat 1	0,00%	0,00%	0,09%	1,32%	6,67%
contrat 2	0,00%	0,00%	0,02%	0,56%	3,71%
contrat 3	0,00%	0,00%	0,02%	0,59%	4,28%
contrat 4	0,00%	0,00%	0,00%	0,23%	2,62%
contrat 5	0,00%	0,00%	0,00%	0,16%	2,03%
contrat 6	0,00%	0,00%	0,00%	0,12%	1,15%
contrat 7	0,00%	0,00%	0,00%	0,08%	0,91%
contrat 8	0,00%	0,00%	0,00%	0,06%	0,75%
contrat 9	0,00%	0,00%	0,00%	0,01%	0,86%
contrat 10	0,00%	0,00%	0,00%	0,00%	0,12%

Indicateur Oméga : seuil = S/P d'équilibre

Nous pouvons faire les premières remarques suivantes sur la différence concernant l'ordre de grandeur des taux de chargement à appliquer selon la mesure de risque : pour l'indicateur Oméga, les taux à appliquer sont très faibles voire nuls selon le niveau de conservation instauré, ce qui est notablement différent des taux élevés constatés précédemment pour la VaR et la TVaR.

En effet, les mesures de risque VaR et TVaR prennent en compte uniquement une partie de la distribution au niveau des scénarii les plus défavorables, alors qu'avec l'indicateur Oméga, les scénarii défavorables sont pondérés par les scénarii favorables et nous pouvons donc nous attendre à un chargement beaucoup moins élevé pour retrouver un niveau de risque équivalent à un contrat sans conservation.

Néanmoins, nous donnons dans la suite une approche plus formelle pour expliquer la faiblesse de ces taux de chargement. Nous rappelons que :

$$\text{pour } L = E(X), \Omega_F(L) = 1,$$

autrement dit, en prenant un seuil L égal à l'espérance du ratio S/P (c'est-à-dire le S/P d'équilibre), l'indicateur $\Omega_F(L) = 1$.

Hors, ce ratio Oméga est supposé être invariant lorsqu'une conservation est appliquée : en effet, l'espérance de la variable aléatoire du ratio S/P est invariante avec l'instauration d'une conservation. En conséquence, le ratio Oméga est aussi égal à 1 lorsqu'il y a conservation, ce qui supposerait des taux de chargement nuls.

Dans les cas étudiés, cependant, les taux de chargement minimes constatés s'expliquent par la raison suivante :

- les conservations mises en place ne sont pas toujours atteintes ; elles sont d'autant plus atteintes que leur niveau est faible et que la taille de la flotte est grande. Cette tendance se reflète dans le tableau des taux de chargement avec des taux « supérieurs » pour les flottes de plus petites tailles, et pour les niveaux de conservation les plus élevés.

Il s'avère donc au vu des résultats obtenus (les taux de chargement minimales) avec ce ratio Oméga, que l'approche actuelle de Covéa Fleet est en accord avec un raisonnement en termes de performance (mesurée par l'indicateur Oméga avec un seuil correspondant au ratio S/P d'équilibre).

La raison citée précédemment inciterait cependant à appliquer ces faibles taux de chargement pour conservation non atteinte, de façon à garder une performance identique au cas où aucune conservation n'est appliquée.

Nous obtenons ci-après le tableau indiquant les taux à appliquer selon la taille de la flotte : nous avons retenu une taille charnière de 500 moteurs en nous basant sur l'observation que les taux de chargement déterminés pour les flottes de taille inférieure à 500 moteurs sont légèrement supérieurs à ceux des plus grosses flottes pour une conservation de 50%. Nous procédons de la même manière que pour les mesures de risques précédentes en faisant la moyenne sur les contrats concernés :

<i>Oméga (s/p d'équil)</i>	<= 500 mot > 500 mot	
<i>taux conservation</i>		
10%	0,00%	0,00%
20%	0,00%	0,00%
30%	0,00%	0,00%
40%	1,00%	0,00%
50%	5,00%	1,00%

Remarques :

- En réalité dans la pratique, le taux de 50% pour la conservation n'est appliqué environ qu'une fois sur deux, particulièrement pour les plus grosses flottes. Dans la moitié des cas restant, il s'agit plutôt de taux de 30% ou 40%. Les chargements ci-dessus n'auraient donc pas vraiment l'occasion d'être instaurés.
- Nous pourrions nous interroger sur la valeur de Oméga si nous retenions un seuil L inférieur au ratio S/P d'équilibre, par exemple $L=0,6$; c'est-à-dire que nous considérerions qu'un S/P inférieur à 0,6 est un bon résultat alors que s'il est supérieur, il serait un mauvais résultat. Dans ce cas la valeur de l'indicateur Oméga pour ce seuil est supérieure lorsqu'il y a une conservation. La conservation entraînerait donc une meilleure performance au sens de Oméga pour ce seuil 0,6. Cela provient du fait que bien que le risque d'avoir un ratio S/P très mauvais soit accru par rapport au cas sans conservation, simultanément, la probabilité d'avoir un ratio S/P très favorable est également accrue avec la présence d'une conservation.

3. Conclusions

L'utilisateur sera en mesure, selon qu'il souhaite se focaliser sur les événements les plus défavorables, ou prendre en compte aussi bien les événements favorables que défavorables, d'opter soit pour les mesures de risque VaR ou TVaR, soit pour l'indicateur Oméga.

L'approche actuelle, bien que se basant sur un raisonnement en terme « d'espérance », s'assimile également, comme nous l'avons vu, à un raisonnement en terme de performance dont la mesure est effectuée par cet indicateur Oméga avec un seuil égal à la valeur d'équilibre du ratio S/P . Il est donc logique que les taux de chargement demandés par rapport à l'indicateur Oméga soient très minimes, voire inexistantes selon le niveau de conservation appliqué.

Pour une approche prudentielle, le taux usuel de 95% pour la VaR et la TVaR s'avère trop élevé pour que l'instauration d'une conservation ait une utilité. En effet le budget pour le client se retrouverait majoré par rapport au budget initial sans conservation.

Pour la Value at Risk, un niveau de confiance de 94%, ou plus vraisemblablement 90%, favorisant le budget du client tout en respectant un niveau prudentiel jugé convenable, peuvent être retenus.

Concernant la Tail Value at Risk, le niveau de confiance de 50% s'avère le seul à avoir éventuellement une utilité pratique, car n'entraînant pas une augmentation du budget client voire permettant une légère diminution.

Le souscripteur pourra juger plus favorable l'approche prudentielle au sens de la mesure de risque Value at Risk qui permettra un meilleur positionnement sur le plan commercial.

CONCLUSION

Dans le cadre de nos travaux, nous avons été amenés à modéliser la charge sinistre d'une grosse flotte automobile afin de proposer un nouveau modèle de tarification dont l'approche est plus élaborée que l'actuelle. Cette modélisation de la charge sinistre d'une flotte va permettre ensuite de déterminer la distribution du ratio sinistres à prime qui lui est associée, dans l'optique d'étudier l'impact de la mise en place d'une conservation sur la distribution du ratio S/P de cette flotte.

La base de notre approche pour la modélisation de la prime pure a consisté en la détermination des paramètres des distributions de la fréquence de sinistralité et du coût d'un sinistre pour un véhicule moteur. Une simple généralisation de ces paramètres à une flotte de n moteurs permet d'obtenir la prime pure a priori qui est associée à cette flotte. Rappelons qu'une distinction a été faite avec un seuil de gravité fixé à 75K euros, entre les sinistres de faible et moyenne intensité, et ceux de grande intensité. Une approche particulière a été adoptée pour la modélisation des sinistres graves en faisant appel à un mélange de deux lois Lognormales.

Une application de la théorie du système de bonus malus aux flottes nous a permis ensuite de prendre en compte l'antériorité d'une flotte pour déterminer sa prime pure a posteriori.

Les principaux résultats qui se dégagent de notre étude permettent, d'une part, d'étayer l'approche actuelle de Covéa Fleet pour la détermination technique de la prime pure d'une grosse flotte, et, d'autre part de consolider le raisonnement actuel (en terme d'espérance) adopté suite à l'instauration d'une conservation :

- Réconciliation de la prime pure Covéa Fleet et de la prime pure modélisée

Dans notre étude nous déterminons pour les dix exemples de flottes automobiles leur prime pure par l'approche « modélisation » pour la comparer à celle de l'approche « Covéa Fleet ». Les résultats obtenus donnent des primes pures très proches, ce qui permet non seulement de consolider une approche dont la mise en œuvre est simple telle que la tarification actuelle, mais aussi d'entériner une approche plus élaborée telle que celle décrite dans ce rapport.

- Stabilité de la performance au sens de l'indicateur Oméga lors de l'instauration d'une conservation

L'approche actuelle pour la tarification d'une flotte avec conservation est basée sur un raisonnement en terme d'iso espérance. Mais elle s'avère être également en concordance avec un raisonnement en terme d'iso performance au sens d'un indicateur sophistiqué tel que Oméga pour un seuil équivalent au ratio S/P d'équilibre.

L'approche précédente relève donc plus d'une notion de performance que de risque pur.

Dans le cas d'une approche prudentielle, des indicateurs tels que la VaR ou la TVaR, qui se focalisent sur les scénarii les plus défavorables, seront plus pertinents. Ils impliquent l'instauration d'une compensation pour cause de conservation sous forme d'un chargement sur la prime pure.

Les taux de chargement à appliquer pour retrouver le même niveau de risque d'avoir un résultat très défavorable à celui où aucune conservation n'est instaurée au sens de ces indicateurs, sont très élevés. En réalité, le client n'aurait pas d'avantage à avoir un montage avec conservation si les taux de chargement calculés pour le niveau de confiance usuel 95% sont appliqués.

Néanmoins en retenant un niveau de prudence plus faible tel que 94% et plus particulièrement 90% pour la VaR, le client pourra obtenir un budget qui lui est favorable.

L'exercice délicat consistera donc à choisir le bon compromis entre niveau de prudence pour le manager et réduction pour le budget du client, qui sera jugé satisfaisant pour les deux parties.

Concernant la TVaR, un niveau de confiance à 50% permettra au client d'avoir une réduction par rapport au budget initial sans conservation, mais cette réduction s'avérant très faible, son attrait paraît moindre que celui lié à l'indicateur VaR qui sera donc privilégié.

ANNEXES

Simulation de la charge sinistre d'une flotte

Dans cette partie, nous présentons de façon plus détaillée les différentes étapes de la simulation de la charge sinistre annuelle RC d'une flotte.

Nous nous intéressons à la plus grosse flotte de notre échantillon qui comporte 3237 véhicules. Deux approches différentes sont utilisées : une approche directe pour la charge de sinistre RC hors graves, une approche individuelle pour les sinistres graves.

Pour la détermination du coefficient bonus malus associé à cette flotte, nous rappelons sa sinistralité constatée dans le tableau suivant :

	n-3	n-2	n-1	n	t	nb mois N
<i>nb moteurs</i>		3074	3074	3237	3	8
<i>nb sin RC hg</i>		366	376	401		

Le coefficient bonus malus associé est égal à 1,13.

Sinistres de fréquence : approche globale directe

Les différentes étapes sont les suivantes :

- Détermination des moments simples d'ordre 1, 2 et 3 pour chacune des variables aléatoires coût rcc, coût rcm non ida, coût rcm ida
- Déduction des moments simples de la v.a. coût d'un sinistre RC hors graves à partir des moments simples ci-dessus
- Détermination de l'espérance du nombre de sinistres RC non graves de la flotte par semi-additivité de la loi binomiale négative.
- Déduction des premiers moments simples de la charge annuelle RC hors graves S de la flotte
- Approximation Wilson Hilferty : calcul des coefficients c_i
- 10 000 tirages de la loi normale centrée réduite
- Application de la formule de Wilson Hilferty et obtention de 10 000 réalisations de la charge sinistre RC hors grave (que nous classerons par ordre décroissant)
- Application du coefficient bonus malus pour obtenir la distribution de la charge sinistre annuelle RC hors graves a posteriori

Sinistres graves : approche individuelle

En raison de la très faible survenance de ce type de sinistres, chacun de ces sinistres graves est simulé individuellement selon les étapes décrites ci-dessous.

Le nombre de sinistres graves de la flotte est simulé selon la loi de Poisson de paramètre λ .

λ est déterminé comme étant le produit des termes suivants :

- le taux égal à 0,00232813 tel que :
Espérance (nombre sin RC graves) = 0,00232813 * Espérance (nombre sin RC hors graves)
- l'espérance du nombre de sinistres RC hors graves a posteriori de la flotte de 3237 véhicules moteur qui est déterminée par la formule suivante :
coefficient bonus malus de la flotte * nombre moteurs de la flotte * Espérance (nombre sin RC hors graves pour un véhicule) = 1,13*3237*0,08270174

Nous souhaitons obtenir 10 000 réalisations de la charge sinistre grave :

- nous effectuons 10 000 tirages de la v.a. nombre de sinistres graves dont la loi est décrite ci-dessus
- pour le tirage i , soit n_i le nombre de sinistres graves
- n_i est constitué de n_{i0} sinistres de type « grave ordinaire » et $(n - n_{i0})$ sinistres de type « grave non ordinaire », où n_{i0} est la réalisation d'une loi Binomiale de paramètres $(n_i; 0,98706511)$
- réalisation de chaque coût de sinistre correspondant selon le type « grave ordinaire » ou « grave non ordinaire »
- pour chaque tirage i , sommation et obtention de la charge sinistre grave annuelle
- classement par ordre décroissant des 10 000 réalisations de la charge sinistre grave annuelle et obtention de la distribution de la charge sinistre grave annuelle

En conservant ce classement par ordre décroissant et en sommant les sinistres RC de fréquence et les sinistres RC graves, nous obtenons finalement la distribution de la charge sinistre RC globale de la flotte.

BIBLIOGRAPHIE

D'AGOSTINO R. B. et STEPHENS M. A. (1986) : « Goodness-of-fit techniques ». Statistics, textbooks and monographs ; vol. 68. Marcel Dekker.

BENABBOU Z. et PARTRAT C. (1994) : « *Grands sinistres et lois mélanges* ». XXVth Astin Colloquium, Cannes.

BESSON J.-L. et PARTRAT C. (2005) : « Assurance non-vie ». Economica.

COHEN A.C. (1991) : “Truncated and censored samples : Theory and Applications”. Marcel Dekker.

HOGG R. V. et KLUGMAN S. A. (1984) : « Loss distribution ». Wiley.

KEATING C. & SHADWICK W. F. (2002) : “A Universal Performance Measure”. The Finance Development Centre London.

MOWBRAY A. H. (1914) : “How extensive a payroll is necessary to give a dependable pure premium?”. Proceedings of the CAS, vol. 1, 24-30.

SCHNEIDER H. (1986) : “Truncated and censored samples from normal populations”. Marcel Dekker.

SIMONSEN W. (1979) : Correction note. (“On the solution of a maximum likelihood equation on the Negative Binomial distribution”). Scand. Act. J., 228-229 (220-231).

WHITNEY A. (1918) : “The theory of experience rating”. Proceedings of the CAS, vol. 4, 274-292.