وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

BADJI MOKHTAR-ANNABA UNIVERSITY

UNIVERSITE BADJI MOKHTAR ANNABA

جامعة باجي مختار

-عنابة-

Faculté des Sciences

Département de Mathématiques **Année : 2024/2025**

THÈSE

Présentée en vue de l'obtention du diplôme de Doctorat

Tarification et Calcul de Prime en Assurance Non-Vie

Filière
Mathématiques Appliquées

Spécialité
Actuariat

Par
OUCHEN Imene

DIRECTEUR DE THÈSE : REMITA Mohamed Riad **Prof. ENSIA - Alger**

Devant le jury

PRESIDENT:	BOUTABIA Hacène	**Prof.**	U.B.M. - Annaba
EXAMINATEUR:	BELHAMRA Thara	**MCA.**	U.B.M. - Annaba
EXAMINATEUR:	DJEBAR Ahlem	**MCA.**	U.B.M. - Annaba
EXAMINATEUR:	HADDARI Allaeddine	**MCA.**	U. Batna 2

Dédicaces

Du profond de mon coeur, je dédie ce travail à tous ceux qui me sont chers,

À la douce mémoire de mon père. Aucun mot ne saurait exprimer mon grand chagrin en ton absence. Puisse Dieu, le tout puissant, t'avoir en sa sainte miséricorde.

À ma mère, ma source de lumière et mon amour éternel.

À ma famille et mes proches, à mon frère et sa femme, à ma soeur et son mari, à mes chers neuveux et nièces adorés.

À tous mes collègues et mes amis qui m'ont toujours encouragé, et à qui je souhaite plus de succès.

À tous ceux que j'aime, ce travail est dédié pour vous.

Remerciements

Vidal Sassoon nous rappelle que le succès nécessite du travail acharné : « Le seul endroit où le succès vient avant le travail, c'est dans le dictionnaire. » Tandis que Desmond Tutu nous affirme que « L'espoir, c'est être capable de voir la lumière malgré les ténèbres. »

Après avoir parcouru un chemin semé d'efforts intenses et de luttes acharnées entre l'éclat de l'espoir et l'ombre de la fatigue, après avoir fait preuve de patience et de persévérance, c'est avec une immense joie et un profond honneur que j'ai atteint mon but. Ces lignes sont tracées aujourd'hui pour exprimer ma gratitude envers tous ceux qui ont participé à la création de ce projet, fruit d'une existence dédiée au labeur et annonciateur d'un nouveau chapitre.

C'est à Dieu tout puissant que j'en dois tout. Je remercie ALLAH de m'avoir donné force, volonté et courage jusqu'à cette réussite.

Je tiens à exprimer mes sincères remerciements envers mon directeur de thèse ***Pr. Remita M.R*** pour m'avoir confié le sujet de cette thèse. Je vous remercie pour votre soutien, votre assistance et votre sagesse.

J'exprime ma gratitude au ***Pr. Boutabia H.*** pour avoir accepté de présider le jury de la soutenance, ainsi qu'aux ***Dr Haddari A., Dr Belhamra T.*** et ***Dr Djebar A.*** qui ont eu l'amabilité d'agir en tant qu'examinateurs pour cette thèse.

Je tiens à exprimer ma sincère, profonde et éternelle reconnaissance envers ***Dr. Saadoun A***. Votre disponibilité, votre patience, votre sagesse et vos conseils éclairés ont été la boussole qui m'a guidé à travers les eaux parfois tumultueuses de la recherche. Je vous suis particulièrement reconnaissante d'avoir été la source de lumière dans les moments d'obscurité de ce parcours. Je souhaite également exprimer ma gratitude envers ***Pr. Zeghdoudi H.*** et ***Dr Metiri F.*** pour leurs précieuses assistances tout au long de cette période. Mes remerciements s'adressent également à mes amis et à toutes les personnes qui m'ont apporté leurs soutiens, de près ou de loin, dans cette longue traversée.
J'ai délibérément réservé la partie la plus significative pour la conclusion. Les termes me font défaut pour témoigner de ma profonde gratitude et de ma reconnaissance envers ma mère, qui a toujours été et demeurera à jamais l'élément central de mon existence et mon mentor. Sans elle, ma présence en ce jour serait inconcevable. Je tiens également à exprimer ma reconnaissance envers tous les membres de ma famille pour leur affection, leur soutien et leur assistance.

Ce travail a été réalisé grâce à la contribution de chacun d'entre vous et est dédié à vous tous.

OUCHEN Imene.

Résumé

Les systèmes traditionnels de bonus-malus établis sur le nombre de sinistres uniquement peuvent être injustes. En réalité, un conducteur qui a soumis une réclamation pour des dommages mineurs est sanctionné de la même manière qu'un conducteur confronté à des dommages significatifs. Afin de corriger cette lacune, on ajuste la prime en considérant en même temps la fréquence et le montant des sinistres.Dans cette étude, deux approches seront présentées. La première hypothèse est que les sinistres suivent une distribution poisson-Akash en fréquence et une distribution Inverse-Gamma Lindley en montant. La dernière méthode utilise la distribution Poisson-New XLindley pour la fréquence des sinistres et la distribution Inverse-Gamma (2, λ) Lindley pour les montants des sinistres. La prime est calculée avec une approche bayésienne ainsi que la fonction de perte linex. Ensuite, des exemples concrets d'application sont présentés avec des données réelles analysées via le logiciel R. Finalement, on comparera les primes basées sur les différentes distributions mentionnées.

Mots clés : Système bonus-malus, Distribution Poisson-Akash, Distribution Poisson-New XLindley, Distribution Inverse-Gamma Lindley, Distribution Inverse-Gamma(2, λ) Lindley, linex, Prime bayésienne.

Abstract

Traditional bonus-malus systems determine insurance premiums solely based on the number of claims, which may lead to inequities. In reality, a driver who files a claim for minor damages faces the same penalties as one who has suffered significant damage. To remedy this shortcoming, the premium adjustment takes into account both the frequency and the cost of claims.This study presents two distinct approaches. The first approach assumes that the number of claims follows a Poisson-Akash distribution, while the amount of claims is described by a new distribution called the Inverse-Gamma Lindley. The second method uses the Poisson-New XLindley distribution to model the number of claims and the one-parameter Inverse-Gamma(2, λ) Lindley distribution to represent the severity of claims. We calculate the premium using a Bayesian approach based on the linex loss function. Concrete examples of application are then provided using real data analyzed with the R software. Finally, we compare the premiums derived from the various previously mentioned distributions.

Keywords : bonus-malus System, Poisson Akash distribution, Poisson-New XLindley distribution, Inverse-Gamma Lindley distribution, Inverse-Gamma (2, λ) Lindley distribution, Bayesian premium, linex loss function.

ملخص

تُحدد أنظمة المكافأة والعقوبة التقليدية (BMS) الأقساط التأمينية استنادًا فقط إلى عدد المطالبات، مما قد يؤدي إلى عدم العدالة. في الواقع، يواجه السائق الذي يقدم مطالبة عن أضرار طفيفة نفس العقوبات التي يتعرض لها من تكبد أضرار جسيمة. لمعالجة هذا القصور، من أجل تحقيق التوازن في المحفظة بين السائقين الجيدين والسيئين، يتم احتساب جزء جديد من القسط كحل وسط. وللقيام بذلك، يتم تعريف وإدخال شدة المطالبات لتوفير قسط عادل.

تقدم هذه الدراسة نهجين متميزين. يفترض النهج الأول أن عدد المطالبات يتبع توزيع Poisson-Akash، بينما يُوصف مبلغ المطالبات بتوزيع جديد يُسمى توزيع Inverse-Gamma Lindley. أما الطريقة الثانية فتستخدم توزيع Poisson-New XLindley لنمذجة عدد المطالبات وتوزيع Inverse-Gamma (2, λ) Lindley بمعلمة واحدة لتمثيل شدة المطالبات.

نحسب القسط باستخدام منهج بايزي استنادًا إلى دالة الخسارة linex. تُعرض أمثلة تطبيقية ملموسة باستخدام بيانات فعلية تم تحليلها بواسطة برنامج R. أخيرًا، نقارن بين الأقساط الناتجة عن التوزيعات المختلفة المذكورة سابقًا.

كلمات مفتاحية: نظام المكافآت، توزيع Poisson-Akash، توزيع Poisson-New XLindley، توزيع Inverse-Gamma (2,λ) Lindley، توزيع Inverse-Gamma (2, λ) Lindley، النهج البايزي، دالة الخسارة Linex.

Table des matières

Introduction

L'assurance est un domaine essentiel de la gestion des risques et de la protection financière, jouant un rôle crucial dans la stabilité économique et sociale. L'assurance revêt une importance primordiale dans la protection financière et la gestion des risques, étant un acteur clé de la stabilité économique et sociale. À travers diverses formes, telles que l'assurance-vie, l'assurance maladie, l'assurance automobile et l'assurance habitation, ce secteur permet de transférer les risques individuels et collectifs à des entités spécialisées, appelées assureurs. Les origines de l'assurance remontent à l'antiquité, mais c'est à l'ère moderne que ce secteur a véritablement pris son essor, notamment avec la révolution industrielle et l'urbanisation croissante. Aujourd'hui, l'assurance est omniprésente et constitue un pilier fondamental du système financier mondial. Elle permet non seulement de protéger les biens et les personnes contre des événements imprévus, mais elle joue également un rôle clé dans le financement de l'économie par les investissements des compagnies d'assurance. L'assurance repose sur des principes fondamentaux tels que la mutualisation des risques, la loi des grands nombres et la diversification. En mutualisant les risques, les assureurs peuvent prévoir plus précisément les pertes potentielles et établir des primes justes et viables pour les assurés. La loi des grands nombres permet d'atténuer les variations statistiques, tandis

que la diversification des risques à travers différents types d'assurances et de régions géographiques réduit l'impact des sinistres majeurs. L'assurance automobile est un pilier fondamental de l'assurance non-vie. Cette assurance est apparue au début du XXe siècle, avec la croissance de l'utilisation des véhicules motorisés. Elle est devenue une exigence dans de nombreux pays afin de garantir la protection des conducteurs, des passagers, des piétons et des biens matériels contre les conséquences financières des accidents de la route. Le défi des actuaires en assurance automobile consiste à facturer une prime juste et équitable pour chaque assuré. La tarification en ce domaine est un processus complexe qui vise à déterminer les primes que les assurés doivent payer pour être couverts contre divers risques. Ce processus repose sur des principes actuariels et statistiques rigoureux pour garantir l'équilibre financier des compagnies d'assurance tout en offrant des prix compétitifs et justes aux consommateurs. En terme de tarification, on distingue plusieurs méthodes, telles que les GLM, la crédibilité ainsi que le système bonus-malus.
La théorie des modèles linéaires généralisés (GLM), formulée par John Nelder et Robert Wedderburn [34] au milieu du 20e siècle, offre un cadre unificateur pour divers modèles statistiques, comme la régression linéaire, la régression logistique et la régression de Poisson. Par ailleurs, Ils ont proposé une approche itérative nommée méthode des moindres carrés repondérés itérativement pour estimer les paramètres du modèle en utilisant le maximum de vraisemblance. Cette méthode demeure fréquemment adoptée et constitue l'option standard dans une multitude de logiciels statistiques. En effet, les GLM sont largement reconnus et fréquemment employés dans le domaine des statistiques, en particulier dans l'actuariat. Il existe beaucoup d'illustration de l'application des GLM en assurance, par exemple : [8, 18, 19, 40].
La théorie de la crédibilité, quant à elle, révèle le fait que les sinistres d'une

personne devraient refléter, au moins en partie, son niveau de risque réel. En effet, les assurés ne sont pas uniformément exposés au risque, certains affichant un profil de risque plus élevé que d'autres. Par conséquent, il peut sembler injuste de facturer à chacun une prime du même montant, car cela conduirait inévitablement à ce que certains assurés soient surévalués et que ces primes supplémentaires soient utilisées pour compenser les pertes causées par les personnes les plus à risque. Pour garantir la justice, Whitney (1918) [52] stipule que les expériences collectives d'une part et les expériences individuelles d'autre part doivent être pondérées. Le problème fondamental de l'équilibre entre l'expérience collective et individuelle a été redéfini par Hans Bühlmann. La proposition à l'origine de la théorie contemporaine de la crédibilité a été introduite par Buhlmann (1967 [3], 1969 [4]), qui a développé un coefficient de crédibilité de type $Z = n/(n + K)$ dans un contexte non paramétrique en forçant la prime bayésienne à être linéaire, avec une formule basique et universelle pour K. Pour des explications détaillées sur la théorie de la crédibilité, veuillez vous référer aux ouvrages cités [5, 15, 35, 36, 53].

La dernière méthode de tarification est le système bonus-malus. Cette tarification représente un mécanisme fondamental dans l'assurance automobile, conçu pour inciter les conducteurs à adopter des comportements de conduite prudents tout en ajustant équitablement les primes d'assurance en fonction du risque individuel. En vigueur dans de nombreux pays, ce système joue un rôle crucial dans la tarification des contrats d'assurance automobile, influençant à la fois les assureurs et les assurés. Le concept du système bonus-malus repose sur l'ajustement des primes d'assurance en fonction du comportement routier passé de l'assuré. Les conducteurs vertueux, qui n'ont pas eu d'accidents responsables sur une période donnée, bénéficient d'un bonus, c'est-à-dire d'une réduction de leur prime d'assurance. À l'inverse, les conducteurs ayant été responsables des sinistres voient leur prime augmenter, ce qui correspond au malus. Historiquement, ce système a été mis en place afin de renforcer la sécurité routière et de promouvoir la responsabilisation des conducteurs. Il est envisagé de diminuer les primes d'assurance pour les conducteurs prudents. De ce fait, les assureurs créent une incitation financière directe à adopter des comportements plus sûrs au volant. De plus, en augmentant les coûts pour les conducteurs plus à risque, le système contribue à une répartition plus juste des coûts liés aux sinistres. On emploie fréquemment la distribution Poisson pour modéliser le nombre de sinistres en matière d'assurance automobile. Cependant, une simple distribution de poisson s'avère insuffisante pour offrir une représentation précise du nombre de réclamations au sein d'un portefeuille d'assurance. Ainsi, Lemaire (1995) [26] a démontré qu'une distribution de Poisson mixte, caractérisée par une queue plus épaisse, permet d'obtenir de meilleurs résultats en termes de fréquence des sinistres dans le cadre d'un système bonus-malus optimal, comme l'a souligné Tremblay (1992) [48] et Walhin (1999) [51]. La distribution de Poisson mixte est employée afin de représenter la fréquence des accidents dans le domaine de l'assurance automobile. Dans cette thèse, on présentera et utilisera deux différentes distributions mixtes : Poisson-Akash [47] et Poisson-New XLindley [33]. Cependant, le système bonus-malus n'est pas exempt de critiques et de défis. Certains critiques soulignent que le système peut pénaliser excessivement les conducteurs après un seul incident. Pour remédier à cette lacune, une nouvelle approche au système bonus-malus a été introduite dans cette thèse. Une nouvelle partie a été ajoutée dans la prime en prenant compte du nombre de sinistres d'un côté et du montant de chaque sinistre d'un autre côté (voir [13, 29, 31]). Cela permettra d'avoir une prime plus raisonnable et équilibrée pour les souscripteurs et la compagnie d'assurance. Dans ce travail, on va introduire et utiliser

deux différentes distributions pour le montant de sinistres, la distribution Inverse-Gamma Lindley et la distribution à un paramètre Inverse-Gamma (2, λ) Lindley.
En outre, ce travail introduit la fonction de perte asymétrique linex de Varian (1975) [49]. Cette fonction de perte est largement utilisée pour déterminer les estimateurs bayésiens et calculer les primes bayésiennes, comme indiqué dans les références [1, 30, 44].

La thèse est structurée en cinq chapitres :
Le premier établit les fondements en présentant le concept du système bonus-malus, le deuxième traite de diverses techniques d'estimation, en exposant succinctement le concept de maximum de vraisemblance et en approfondissant celui de l'approche bayésienne ainsi que quelques fonctions de perte. Le troisième a été consacré à la nouvelle approche du système bonus malus qui se base sur la fréquence et le montant des sinistres. On a employé la distribution Poisson-Akash pour modéliser le nombre de sinistres, tandis que le montant des sinistres est déterminé à partir d'une distribution novatrice mixte introduite dans ce chapitre, la distribution Inverse-Gamma Lindley. Le quatrième chapitre représente une application numérique des résultats obtenus en utilisant des données réelles à l'aide du logiciel R. Le chapitre final présente les primes bayésiennes du système bonus-malus qui font appel à la distribution Poisson-New XLindley pour modéliser la fréquence des sinistres, ainsi qu'à la distribution à un paramètre Inverse-Gamma(2,λ) Lindley pour évaluer le montant des sinistres. Une application numérique sera réalisée en se servant du logiciel R et des données réelles afin d'estimer les primes d'assurance. En dernier lieu, une comparaison sera établie entre les primes calculées dans ce chapitre et celles établies dans le chapitre précédent.

Chapitre

1 Tarification à l'aide du système bonus malus

Sommaire

Les premiers dispositifs du système bonus-malus ont été introduits dans le domaine de l'assurance automobile dès le début de 20e siècle en Angleterre, et ont été rapidement adoptés en 1930 par le Canada (voir Lemaire (1995) [26]). Ces mécanismes prévoyaient, par exemple, une réduction de 10% en cas d'absence de sinistre au cours d'une année. Mais en cas de sinistre déclaré, aucune pénalité n'était appliquée. Depuis lors, les systèmes de bonus-malus ont connu une évolution significative, et une approche basée sur les chaînes de Markov a été développée pour en améliorer l'analyse. Ces dispositifs sont largement utilisés dans le secteur de l'assurance automobile, en raison de la perception générale selon laquelle les conducteurs ont une certaine maîtrise sur leur fréquence d'accidents.

1.1 Fonctionnement du système bonus malus

1.1.1 Principes de base

Le système bonus malus repose sur la modulation de la prime d'assurance automobile en fonction du comportement du conducteur. Plus précisément, il s'agit d'un coefficient multiplicateur appliqué à la prime de base. Si l'assuré n'a enregistré aucun sinistre dans l'année, il obtiendra un bonus, et donc une réduction dans la prime de l'année suivante. Par contre, si l'assuré a été responsable d'un sinistre ou plus, sa prime augmentera pour l'année prochaine.

1.1.2 Modèle théorique

En assurance automobile, l'assureur doit mettre en place un système de tarification qui lui permet d'être compétitif tout en gérant le risque qu'il prend en charge. Nous noterons le risque à tarifer comme $(X_t)_{t\in\mathbb{N}}$ et la classe de tarif d'un risque comme $(C_t)_{t\in\mathbb{N}}$. À l'aide de cette notation, clarifions la terminologie.

Définition 1.1.1.

La classe de tarif (C_t) *est liée à la prime à payer au temps* t *pour couvrir le risque survenant pendant la période* $[t, t+1]$. *Ce processus* $(C_t)_{t\in\mathbb{N}}$ *décrit comment la classe de tarif évolue au fil du temps.*

En général, dans la plupart des modèles de tarification, on suppose que le processus $(X_t)_{t\in\mathbb{N}}$ (représentant le risque) est indépendant du processus $(C_t)_{t\in\mathbb{N}}$ (représentant la classe de tarif). Cela signifie que la prime à payer ne dépend pas directement de la classe tarifaire. Cependant, il est crucial de noter que la classe de tarif doit tenir compte du risque pour établir un système de tarification cohérent et équitable.

Définition 1.1.2.

Dans le domaine de la tarification, la variable a posteriori se définit comme une mesure dont la valeur ne peut être établie qu'après l'observation effective du risque. Autrement dit, si l'on considère $(Y_t)_{t\in\mathbb{N}}$ *comme la variable a posteriori associée au risque modélisé par* $(X_t)_{t\in\mathbb{N}}$, *alors* Y_{t+1} *est précisément déterminée dès que* X_{t+1} *est observé, pour tout* $t \in \mathbb{N}$. *Cette démarche permet d'intégrer les informations réelles recueillies sur le risque dans le processus de tarification, ce qui est fondamental pour tout système tarifaire.*

En effet, les variables a posteriori jouent un rôle essentiel dans l'estimation du risque dans les systèmes de tarification. Les indicateurs étudiés incluent notamment le nombre de réclamations, le nombre d'accidents pour lesquels l'assuré est tenu pour responsable ainsi que le décompte des infractions commises au code de la route. Plusieurs travaux, tel que l'étude réalisée par Lemaire en 1977 [25], ont démontré que les variables a posteriori sont plus efficaces pour évaluer le risque que les variables a priori. Ainsi, il est crucial d'intégrer ces variables dans la conception de la règle de décision u d'un système de tarification. Par exemple, dans un système bonus-malus, le nombre d'accidents responsables est représenté par Y_t, tandis que la classe de tarif de la période précédente est notée A_t.

Définition 1.1.3.

(i) *Dans un système bonus-malus, le mécanisme de tarification se déploie en deux étapes. Initialement, au début de chaque période, un risque se voit attribuer une classe tarifaire C_t. Par la suite, en fin de période, ce risque est réévalué et reclassé dans une nouvelle catégorie C_{t+1} à l 'aide d'une règle de décision u (voir Krahnen (2001) [22]). Précisément, cette règle intègre l'information de la classe tarifaire antérieure C_t ainsi que le nombre d'accidents dont l 'assuré est responsable, noté Y_{t+1}, observés durant la période précédente.*

$$C_{t+1} = u\left(C_t, Y_{t+1}\right) \tag{1.1}$$

En temps $t = 0$, i_0 représente la valeur de C_0.

(ii) *Les classes de tarifs $(C_t)_{t\in\mathbb{N}}$ peuvent être choisis parmi les différentes classes possibles l. La classe 1 se caractérise par l'octroi du bonus le plus élevé, alors que la classe l se distingue par l'attribution du malus maximal.*

(iii) *Chaque classe de tarif i est associée à un pourcentage de la prime de base b_i, avec la propriété que $b_1 \leqslant b_2 \leqslant ... \leqslant b_i$.*

Dans le cadre d'un système bonus-malus, les variables préalablement introduites, telles que l'âge ou le type de véhicule, servent à établir la prime de base d'un nouvel assuré. La prime finale est ainsi déterminée en multipliant ladite prime par le coefficient b_i correspondant à la classe i. Par ailleurs, la définition initiale du système bonus-malus classique ouvre la voie à une généralisation du modèle par la modification de la règle de décision u. À titre d'exemple, outre l'observation du nombre d'accidents dont l'assuré est jugé responsable Y_{t+1}, il serait également possible de faire dépendre les classes de tarifs du type d'accident Y'_{t+1}, comme indiqué ci-dessous :

$$C_{t+1} = u\left(C_t, Y_{t+1}, Y'_{t+1}\right).$$

Exemple 1.1. ***Système bonus-malus classique***

Le système bonus-malus en Thaïlande, tel que présenté par Lemaire (1994) [28], illustre un exemple typique d'un système classique de bonus-malus. Il comporte sept classes avec des niveaux de prime correspondants $(b_1, ..., b_7) =$(60%, 70%, 80%, 100%, 120%, 130%, 140%)*,* $C_0 = 4$ *représente la classe de départ. Ce système utilise la règle de transition suivante :*

$$u(i.k) = \begin{cases} \max(1, i-1), k = 0 \ \ et\ 1 \leqslant i \leqslant 7 \\ 4, k = 1 \ \ et\ i < 4 \\ 5, k > 1 \ \ et\ i < 4 \\ min(7, i+1), k \neq 0 \ \ et\ i \geqslant 4 \end{cases}$$

Cela est illustré de manière concise dans le tableau 1.1 ci-dessous.

Classe	*Niveau de prime*	$k = 0$	$k = 1$	$k > 1$
7	140	6	7	7
6	130	5	7	7
5	120	4	6	6
4	100	3	5	5
3	80	2	4	5
2	70	1	4	5
1	60	1	4	5

Table. 1.1 - *Classe attribuée après k réclamations*

Exemple 1.2. ***Système bonus-malus de la SAAQ***

Dans la SAAQ au Québec, le système de points d'inaptitude est utilisé pour surveiller le comportement des conducteurs sur la route. Lorsqu'un conducteur ou un passager est déclaré coupable d'une infraction, des points d'inaptitude sont ajoutés à son dossier. Le nombre de points varie en fonction de la nature de l'infraction enregistrée. Ces points sont enregistrés dans le dossier du conducteur pendant une durée de deux ans. La société de l'Assurance Automobile du Québec présente un système bonus-malus en cinq classes. Le tableau 1.2 ci-dessous présente les diverses catégories du système, voir [45].

Classe	*Points d'inaptitude*	*Niveau de prime*
5	> 15	796%
4	$[12, 14]$	572%
3	$[8, 11]$	348%
2	$[4, 7]$	200%
1	$[0, 3]$	100%

Table. 1.2- *Classe attribuée selon le nombre de points d'inaptitude*

Ce système diffère d'un bonus-malus classique, car les catégories sont établies en fonction de la gravité relative des violations du code de la route, plutôt que de leur fréquence. Néanmoins, il conserve la structure d'un système bonus-malus, illustrant ainsi la distinction entre les sanctions appliquées aux infractions mineures et majeures. Par conséquent, il vise à encourager une conduite plus sécuritaire en sanctionnant les comportements fautifs sur la route. Ainsi, le modèle de tarification de la SAAQ remédie à l'une des lacunes des systèmes bonus-malus identifiée par divers auteurs, tels que Lemaire (2004) [27].

1.1.3 Structure markovienne

Il est couramment supposé que $(Y_t)_{t\in\mathbb{N}}$ est constituée de variables aléatoires indépendantes et identiquement distribuées. Cette hypothèse implique que les compétences de conduite d'un assuré demeurent constantes dans le temps, c'est-à-dire que les conducteurs n'acquièrent pas de nouvelles compétences à partir de leurs expériences. Nous adhérerons également à cette hypothèse. Comme nous l'examinerons dans la remarque 1.1.1, les systèmes de bonus-malus intègrent un mécanisme visant à pallier les limites de cette hypothèse.

Proposition 1.1.
Le processus de classe de tarif $(C_t)_{t\in\mathbb{N}}$ constitue une chaîne de Markov homogène.

Démonstration. Soit i_t la valeur de la classe tarifaire à l'instant t. En utilisant l'équation 1.1, nous obtenons :

$$\begin{aligned}
&P(C_{t+1} = i_{t+1} \mid C_t = i_t, ..., C_0 = i_0) \\
&\qquad = P(u(i_t, Y_{t+1}) = i_{t+1} \mid u(i_{t-1}, Y_t) = i_t, ..., C_0 = i_0) \\
&\qquad = P(u(i_t, Y_{t+1}) = i_{t+1}) \qquad (1.2) \\
&\qquad = P(u(i_t, Y_{t+1}) \mid C_t = i_t) = P(C_{t+1} = i_{t+1} \mid C_t = i_t).
\end{aligned}$$

□

L'indépendance de $u(i_t, Y_{t+1})$ des variables $C_0, ..., C_t$ définies par $Y_1, \dots, Y_t$ a été utilisée dans 1.2. La probabilité conditionnelle $P(C_{t+1} = i_{t+1} \mid C_t = i_t)$ exprimée par $P(u(i_t, Y_{t+1}) = i_{t+1})$ reste indépendante de C_t du fait que les variables aléatoires Y_t sont toutes distribuées de manière identique. Ainsi, le processus des classes de tarif $(C_t)_{t\in\mathbb{N}}$ constitue une chaîne de Markov homogène.

Remarque 1.1.1.

L'équation 1.1 peut être interprétée comme une équation récursive stochastique. Comme souligné par Rolski et al. (1998) [43], le processus $(C_t)_{t\in\mathbb{N}}$ forme automatiquement une chaîne de Markov. La démonstration précédente est conforme à celle exposée dans la référence citée.

La distribution de probabilité commune à la suite $(Y_t)_{t\in\mathbb{N}}$ est donnée par $\{p_k\}_{k\in\mathbb{N}}$ où p_k=P(Y_t =k). Pour obtenir la matrice de transition Q associée à la suite $(C_t)_{t\in\mathbb{N}}$, il convient de considérer les éléments individuels q_{ij} tels que $Q = (q_{ij})_{i,j=1,\ldots,l}$.

Proposition 1.2.

La probabilité de transition q_{ij}, représentant le passage de la classe i à la classe j, est définie par :

$$q_{ij} = E\left(1_j\left(u\left(i, Y_{t+1}\right)\right)\right) = \sum_{k=0}^{\infty} p_k 1_j\left(u\left(i, Y_{t+1}\right)\right).$$

Démonstration.

Étant donné que $(C_t)_{t\in\mathbb{N}}$ constitue une chaîne de Markov, en utilisant l'équation 1.1, nous parvenons au résultat souhaité :

$$\begin{aligned} q_{ij} &= P(C_{t+1} = j \mid C_t = i) \\ &= P(u(i, Y_t) = j \mid C_t = i) \\ &= E\left(1_j\left(u\left(i, Y_{t+1}\right)\right)\right) \end{aligned}$$

□

Remarque 1.1.2. *Il est parfois courant de représenter la quantité* $1_j\ u\,(i, k)$ *sous la forme d'une règle de transition* $t_{ij}(k)$*, où :*

$$t_{ij}(k) = \begin{cases} 1, \text{ si la police passe de la classe } i \text{ à la classe } j \\ \text{lorsque } k \text{ réclamations surviennent.} \\ 0, \textit{autrement} \end{cases}$$

Cela permet de construire une matrice de transition $T(k) = (t_{ij}(k))_{i,j=1,\ldots,l}$, où $T(k)$ est une matrice binaire avec exactement un élément égal à 1 dans chaque ligne. L'utilisation de la notation $1_j\,(u\,(i, k))$ facilite l'analyse, comme démontré dans la proposition 1.1, tandis que la notation $t_{ij}(k)$ simplifie la présentation des règles de décision, comme illustré dans l'exemple 1.2. Le tableau résumant cet exemple correspond en fait à une représentation concise de la matrice $T(k)$. Le choix entre l'une ou l'autre des notations est déterminé par le contexte.

Dans la pratique, on suppose couramment que la fréquence des sinistres Y_t suit une loi Poisson avec le paramètre Λ (Lemaire 1994 [28]). La caractérisation de Λ est établie à travers sa fonction de densité g, connue sous le nom de fonction de structure. La distribution Gamma est souvent privilégiée pour modéliser Λ. Par conséquent, la fréquence des réclamations suit une distribution Binomiale Négative.

1.2 Efficacité du système bonus malus

L'efficacité d'un système est définie par sa capacité à atteindre ses objectifs. Dans le cas d'un système de tarification, trois objectifs distincts sont poursuivis. Tout d'abord, en influençant directement les primes facturées, un système de tarification vise à maintenir les réserves de l'assureur à un niveau optimal. Cet objectif ne peut être atteint que si le système évalue de manière adéquate le risque encouru par l'assureur. Enfin, les ajustements apportés aux primes par le système ne doivent pas être excessivement sévères, sous peine de compromettre la capacité du système à proposer aux assurés une couverture d'assurance attrayante.

Définition 1.2.1. *L'efficacité d'un système de tarification se mesure par sa capacité à garantir :*

(i) *La stabilité financière du système.*

(ii) *Une estimation précise du risque.*

(iii) *Le respect du principe de transfert du risque en assurance.*

La définition ci-dessus, ainsi que la partie suivante, sont conformes à l'approche proposée par Lemaire pour évaluer l'efficacité d'un système bonus-malus (Lemaire [26, 27]). Cette partie précise la définition des composantes sur lesquelles repose l'efficacité d'un système de tarification et présente les mesures applicables dans le cadre d'un système bonus-malus. À cette fin, nous ferons appel au processus de surplus $(U_t)_{t\in\mathbb{N}}$, dont voici la définition précise.

Définition 1.2.2. *Le montant des réclamations dans la période* $[t, t+1]$ *est représenté par* $X_{\ t+1}$*, tandis que* $\pi(C_t)$ *désigne la prime facturée en début de période pour la classe* C_t*, et* u *correspond au niveau de réserve initial. En évaluant le niveau de surplus* U_{t+1} *à la fin de la période, nous pouvons définir le processus de surplus* $(U_t)_{t\in\mathbb{N}}$ *comme :*

$$U_{t+1} = u + \sum_{s=0}^{t} \pi(C_s) - \sum_{s=1}^{t+1} X_s = u + \sum_{s=0}^{t} \pi(u\,(C_{s-1}, Y_s)) - \sum_{s=1}^{t+1} X_s$$

Sans porter atteinte à la généralité, il est convenu de fixer la prime de base à 1 et d'échelonner le montant des réclamations à une échelle unitaire. Cette approche permet de concentrer l'analyse sur l'influence des niveaux de primes $\{b_j\}$ spécifiques aux systèmes bonus-malus.

$$\pi(C_t) = \sum_{j=1}^{l} q_{i_0 j}^{(t)} b_j.$$
$$E(X_t) = 1, \forall t \in \mathbb{N}.$$

Remarque 1.2.1. *Pour résumer brièvement la démarche de calibration d'un système bonus-malus, nous commençons par calculer les $(b_j)_{j=1,\ldots,l}$ à partir de données et d'outils statistiques. Ensuite, nous choisissons les valeurs de u et $(b_j)_{j=1,\ldots,l}$ de manière à équilibrer les critères d'efficacité mentionnés dans la définition 1.2.1. Pour atteindre cet équilibre, il peut être nécessaire d'ajuster les $(b_j)_{j=1,\ldots,l}$. Lemaire (1995) [26] a détaillé cette approche, tandis que l'exemple concret fourni par Denuit (2003) [10] offre un éclairage pratique sur le sujet.*

1.2.1 Stabilité financière

Un système de tarification efficace doit engendrer une grille de primes permettant d'assurer la stabilité financière de la compagnie d'assurance. Les réductions accordées par ce système ne doivent en aucun cas entraîner une sous-évaluation des tarifs.

Définition 1.2.3. *Dans le cadre des systèmes de bonus-malus, on évalue la stabilité financière en se basant sur le pourcentage stationnaire attendu de la prime de base, noté par :*

$$b' = \sum_{j=1}^{l} a_j b_j.$$

où a_j représente la probabilité stationnaire d'être dans j.
c' représente la classe stationnaire espérée

$$c' = \sum_{j=1}^{l} a_j j,$$

et $NSRE$ c'est le niveau de stationnarité relatif :

$$NSRE = \frac{b' - b_1}{b_l - b_1}.$$

Un taux stationnaire attendu b' supérieur à 1 indique que le système accorde principalement des malus. En revanche, un taux inférieur à 1 suggère que le système accorde principalement des bonus. Cette interprétation s'applique également au $NSRE$. Un faible niveau de $NSRE$ indique une forte concentration d'assurés dans les classes à forts bonus, tandis qu'un niveau élevé de $NSRE$ suggère une répartition plus équilibrée des assurés à travers les différentes classes. D'autres indicateurs tels que la probabilité de faillite ou le montant du déficit en cas de faillite peuvent également être mobilisés. Cependant, dans le domaine des systèmes de bonus-malus, l'accent est mis sur le paramètre $b^{'}$ en raison de la nature markovienne de ces systèmes.

Proposition 1.3. *Un système bonus-malus est considéré comme stable financièrement lorsque :*

(i) *Il est transparent.*

$$b' = 1.$$

(ii) *Une fois que la stationnarité est atteinte, l'assuré n'est pas affecté à une classe extrême.*

$$0 < NSRE < 1.$$

1.2.2 Transfert adéquat du risque

Un mécanisme de tarification a pour objectif d'assurer une contribution équitable de chaque assuré au financement des sinistres. Cependant, des malus excessifs pourraient dissuader les assurés de déposer des réclamations, les privant ainsi des prestations auxquelles ils ont droit, voire les inciter à résilier leur contrat après avoir causé des dommages à des tiers. Afin d'éviter ces écueils, il est essentiel que les ajustements de prime liés aux bonus et malus demeurent raisonnables, faute de quoi le système ne parviendrait pas à transférer efficacement le risque. La définition qui suit propose des indicateurs permettant d'évaluer l'impact des variations de prime engendrées par ce mécanisme.

Définition 1.2.4. *L'évaluation de l'impact des systèmes bonus-malus sur les primes se fait en analysant le coefficient de variation des primes*

$$\rho_t = \frac{\sqrt{Var(\pi(C_t))}}{E\left(\pi(C_t)\right)},$$

et la rétention optimale moyenne

$$\varepsilon^* = maxE(R_t(x,\varepsilon)),$$

où $R_t(x,\varepsilon)$ s'agit de la compensation effective engendrée par le système en cas de rétention de ε.

$$R_{t+1}(x,\varepsilon) = \sum_{s=t+1}^{\infty}\left(X_S \mid X_t = x\right) - \sum_{s=t+1}^{\infty}\left[\pi\left(C_S \mid X_t = x-\varepsilon\right) - \pi\left(C_S \mid X_t = x\right)\right].$$

La notion de compensation réelle implique la prise en compte du montant versé pour les réclamations ainsi que de la perte de bonus liée à la déclaration d'une réclamation. Cette approche requiert une considération exhaustive de l'ensemble des réclamations et des bonus futurs, comme souligné par Holtan (2001) [16]. La rétention, symbolisée par ε, peut être assimilée à une franchise que l'assuré choisit afin d'optimiser sa compensation réelle.

Un faible coefficient de variation de la prime ρ_t indique que les variations de prime dues aux bonus ou malus sont peu importantes, favorisant ainsi un transfert efficace du risque. De même, une faible rétention moyenne optimale suggère que le système assure un transfert adéquat du risque.

Proposition 1.4. *Un système bonus-malus est conforme au principe de transfert du risque lorsque :*

(i) *La prime de l'assuré ρ_t présente un coefficient de variation raisonnable.*

(ii) *La valeur moyenne optimale de la rétention, notée ε^*, est jugée raisonnable.*

1.2.3 Estimation adéquate du risque

Dans le domaine de l'assurance automobile, un mécanisme de tarification est mis en place dans le but d'évaluer de manière appropriée le risque associé à chaque assuré, en lui attribuant une prime correspondante. Cette approche individualisée vise à répartir équitablement les coûts des sinistres entre les assurés. Ainsi, la différence entre la prime facturée $\pi(C_t)$ et le risque effectivement encouru $E(X_{t+l})$ est utilisée comme indicateur de la précision de l'estimation du risque par le système. Il est également essentiel d'analyser l'évolution de cette différence dans le temps, ainsi que la capacité du système à s'adapter à des variations dans la distribution de X_t.

Proposition 1.5. *Un système bonus-malus permet d'obtenir une estimation appropriée du risque $(X_t)_{t\in N}$ lorsque :*

(i) *En régime stationnaire, une faible erreur quadratique de tarification espérée, notée ET, indique que le niveau de prime stationnaire est étroitement lié à l'espérance du montant des réclamations.*

$$ET = E((b' - E(x))^2)$$

(ii) *Il parvient à obtenir un coefficient d'élasticité η proche de 1 pour x appartenant à l'intervalle $[a_1, a_2]$, où $[a_1, a_2]$ représente l'intervalle de valeurs typiques pour l'espérance des classes de tarifs.*

$$\eta = \frac{dlnb'}{dlnx} = \frac{bd'}{dlnx} / \frac{dx}{x}$$

(iii) *Les probabilités de transition convergent rapidement vers l'état stationnaire lorsque la variation totale VT_n diminue rapidement avec le temps.*

$$VT_n = \sum_{j=1}^{l} |q_{i_0 j}^{n} - a_j|$$

L'élasticité η est un indicateur de la capacité du système à faire face à des modifications dans la distribution de Y. La conception d'un système capable de réagir à tous les changements possibles dans la distribution de Y est souvent impossible, ce qui explique pourquoi il est limité à un intervalle de valeurs communes. Lorsque le système évalue correctement le risque, il est capable de séparer les conducteurs compétents des conducteurs inexpérimentés le plus rapidement possible. Cependant, cette sélection des conducteurs est terminée une fois que les probabilités de transition atteignent la stationnarité. C'est la raison pour laquelle un système qui atteint rapidement la stationnarité atteint également son objectif d'évaluer correctement le risque.

Chapitre 2 Méthodes d'estimation

Sommaire

Les méthodes d'estimation occupent une place centrale en statistique et en actuariat, car elles permettent de déterminer les paramètres inconnus des modèles à partir des données observées. L'objectif est de fournir des estimateurs fiables qui traduisent au mieux la réalité étudiée, tout en tenant compte des contraintes théoriques et pratiques.

2.1 La méthode des moments

La méthode des moments, introduite par Pearson en 1892 [38], est une technique simple pour estimer les paramètres en statistiques, utilisée dès les débuts de cette discipline. Elle implique l'estimation des paramètres en égalisant certains moments théoriques (dépendants de ces paramètres) avec leurs homologues empiriques. Cette égalisation est légitimée par la loi des grands nombres, ce qui implique que la moyenne empirique peut tendre vers l'espérance mathématique. Ainsi, la résolution d'un système d'équations est nécessaire. En d'autres termes, lorsque $\theta = E(X)$, l'estimation de θ obtenue par la méthode des moments (EMM) est la suivante :

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

2.2 La méthode du maximum de vraisemblance

La méthode du maximum de vraisemblance, introduite par Fisher en 1922 [11] et en 1925 [12], est une technique d'estimation générale qui présente des propriétés statistiques remarquables, notamment dans des situations asymptotiques, comme l'ont souligné Cox et Hinkley (1974) [7]. Cette méthode repose sur le principe fondamental de choix des valeurs qui maximisent la probabilité d'observer les données réellement collectées parmi toutes les valeurs possibles. En d'autres termes, la définition d'une fonction de vraisemblance vise à identifier les paramètres qui maximisent la probabilité des données observées. Le vecteur des estimations du maximum de vraisemblance (noté $\widehat{\theta}$) correspond alors au maximum de cette fonction.

Considérons $x_1, x_2, ..., x_n$ comme étant n observations indépendantes et identiquement distribuées. Dans ce contexte, la fonction de vraisemblance s'exprime comme suit :

$$L(\Theta; x_1, x_2, ..., x_n) = \prod_{i=0}^{n} f(x_i; \Theta).$$

Définition 2.2.1. *On exprime la fonction de vraisemblance de la manière suivante :*

$$\theta \to L(\theta; x_1, x_2, ..., x_n),$$

et on estime le maximum de vraisemblance du paramètre θ en examinant chaque point maximal de ladite fonction :

$$\hat{\theta}_{mv} = arg\, max_{\theta} L(\theta; x_1, x_2, ..., x_n).$$

Le tableau suivant représente des estimateurs pour quelques distributions en utilisant le maximum de vraisemblance (voir [20, 41]).

Distribution	Paramètre	Estimateur
Loi uniforme sur $[0, a]$	a	$Sup(x_i)$
Loi binomiale (k nombre de succés en n épreuves)	p	$\widehat{p} = \frac{k}{n}$
Loi Poisson	λ	$\widehat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Loi normale	m	$\widehat{m} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
	σ^2	$\widehat{\sigma^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$

Tableau 2.1 – Estimateurs obtenus par la méthode du maximum de vraisemblance.

2.3 La méthode bayésienne

2.3.1 Le principe bayésien

L'inférence bayésienne est une approche statistique permettant d'évaluer la probabilité d'un événement en se fondant sur les probabilités d'événements antérieurs, en se référant essentiellement au théorème de Bayes.

Théorème 2.3.1. *(voir [2]). A et B sont deux événements et $P(B) \neq 0$, alors $P(A|B)$ et $P(B|A)$ sont reliés par :*

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\ &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})} \end{aligned}$$

Définition 2.3.1. *Un modèle statistique bayésien est défini comme un modèle statistique paramétrique, noté $f(x|\theta)$, accompagné d'une distribution a priori pour les paramètres, Θ.*

Contrairement aux approches antérieures qui présupposent une distribution de probabilité fixe, notée par θ, pour chaque échantillon et attribuent la variation des données à des différences entre les échantillons, l'estimation bayésienne postule une variabilité de la distribution de la population, et donc du paramètre θ. Cette méthode se base sur les données observées pour évaluer la probabilité qu'un paramètre prenne une valeur donnée. Par conséquent, au lieu de fournir une valeur déterministe pour un paramètre, l'estimation bayésienne produit sa distribution.

Il est nécessaire de clarifier plusieurs concepts en premier lieu.

- **La distribution a priori**, généralement symbolisée par $\pi(\theta)$, représente la distribution d'un paramètre hypothétique avant toute observation.
- **La distribution a posteriori** représente la distribution estimée du paramètre, fondée sur les données observées $x = (x_1, x_2, ..., x_k)$. Plus précisément, il s'agit de la probabilité conditionnelle de θ sachant x, notée $\pi(\theta|x)$.
- **La distribution modèle**, également connue sous le nom de distribution d'échantillonnage, illustre la distribution de la variable aléatoire sous-jacente X pour une valeur spécifique du paramètre. Elle est symbolisée par $f(\theta|x)$, en référence à la fonction de vraisemblance définie antérieurement.
- **La distribution marginale** la distribution marginale se réfère à la distribution de mélange de la variable aléatoire X, lorsque le paramètre est traité comme ayant une distribution a priori avec sa fonction de densité de probabilité :

$$\pi(x) = \int f(x|\theta)\pi(\theta)d\theta$$

Le principe de cette approche s'appuie sur le théorème de Bayes. Pour obtenir de plus amples informations, veuillez consulter les références [39, 41],

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta}.$$

Comparée à d'autres méthodes, l'approche bayésienne se distingue par sa flexibilité, offrant ainsi la possibilité au modèle de s'adapter de manière dynamique aux données observées. Néanmoins, en raison de l'implication d'intégrales ou de sommes, la complexité des calculs s'est accrue, rendant

parfois difficile l'obtention d'une forme explicite.
En pratique, il est généralement requis de fournir une estimation du paramètre. Pour obtenir une telle estimation, il est courant de chercher à minimiser la différence par rapport à la vraie valeur. Une fonction de perte quadratique $l(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$ est fréquemment employée dans ce contexte. Par conséquent, l'estimation bayésienne correspond à la moyenne du paramètre, et son raisonnement est étroitement similaire à celui de la méthode des moindres carrés utilisée dans les analyses de régression linéaire.
On consacre la section suivante à l'exposition de quelques fonctions de perte.

2.4 Théorie de la décision et les fonctions de pertes

La théorie de la décision est un domaine d'étude pluridisciplinaire qui se situe à l'interface de l'économie, des mathématiques et des statistiques. Son domaine d'étude se focalise sur les processus de prise de décision, qu'ils soient individuels ou collectifs, en proposant des cadres analytiques et des modèles permettant d'évaluer les différentes alternatives et leurs conséquences. Depuis ses débuts au début du 20e siècle, marqués par les travaux pionniers de chercheurs tels que Abraham Wald [50], la théorie de la décision a progressé en intégrant diverses perspectives sophistiquées. Dans un contexte marqué par l'incertitude et la complexité croissante, la théorie de la décision revêt une importance capitale en guidant les acteurs dans des environnements où les données sont fréquemment lacunaires et les issues incertaines. Que ce soit dans le domaine de la gestion d'entreprise, de la finance ou des activités quotidiennes, une maîtrise approfondie des fondements de la théorie de la décision peut non seulement améliorer la pertinence des décisions prises, mais aussi optimiser les résultats obtenus. Les fonctions de perte jouent un rôle fondamental dans la théorie de la décision, notamment dans le contexte des prises de décision en situation d'incertitude. Elles servent à évaluer de manière quantitative le coût lié à une décision spécifique et à ses retombées éventuelles. Grâce à ces fonctions, il est possible d'organiser les concepts de risque et de préférence d'assigner des valeurs numériques aux différents résultats potentiels d'une décision.

Néanmoins, il est pertinent d'analyser l'incidence d'une décision sur un paramètre par rapport à sa valeur réelle inconnue. La théorie de la décision offre un cadre pour mener une telle analyse de validation d'un estimateur. Le critère de perte est établi par l'ensemble des choix envisageables, désigné par D, qui constitue l'espace de décision. La plupart des exemples théoriques se focalisent sur le cas où $\mathrm{D} = \Theta$, correspondant au cadre d'estimation standard.

D'un point de vue mathématique, une fonction de perte $L(d, \theta)$ est utilisée pour quantifier les pertes encourues lors de l'estimation du paramètre θ par d. Une distinction est toujours observée entre l'estimateur et le paramètre. Si l'on note d l'estimateur du paramètre θ, la perte est généralement exprimée en termes de la différence $d - \theta$ ou du rapport $\frac{d}{\theta}$. Dans le cas d'un seul paramètre, en fixant θ, différentes valeurs de d peuvent être obtenues en tant qu'estimations de θ. Lorsque $d = \theta$, il n'y a aucune perte observée. Ainsi, il est possible d'établir une fonction de perte $L(d, \theta)$ dépendant des variables d et θ.

Définition 2.4.1. *Une fonction de perte est toute fonction L de $\Theta \times D$ dans $[0\,;+\infty)$.*

La fonction de perte $L(d, \theta)$ est définie comme une fonction à valeur réelle satisfaisant :

(i) $L(d, \theta) \geqslant 0$ pour tous les estimateurs possibles d et toutes θ sous la

population choisie.

(ii) $L(d, \theta) = 0$ pour $d = \theta$.

En science actuarielle, les fonctions de perte sont employées dans le domaine des assurances afin de représenter les paiements effectués en plus des primes. Cette section expose diverses fonctions de perte et les estimateurs bayésiens correspondants.

2.4.1 Fonction de perte quadratique

La fonction de perte quadratique, proposée par Legendre (1805) [24], Gauss (1809) [14] et Laplace (1810) [23], est certainement le critère d'évaluation le plus couramment utilisé. Elle est définie de la manière suivante :

$$L(d, \theta) = (d - \theta)^2$$

Dans son article, Gauss a mis en évidence la nature aléatoire de la perte quadratique et l'a justifiée par souci de simplicité. Ces critiques demeurent pertinentes de nos jours. Néanmoins, malgré ces critiques, la perte quadratique demeure largement utilisée en raison de sa capacité à conduire généralement à des solutions bayésiennes acceptables, c'est-à-dire des solutions obtenues par une inférence non-décisionnelle basée sur la densité a posteriori. En effet, les estimateurs de Bayes liés à la perte quadratique correspondent aux moyennes a posteriori.

Proposition 2.1. *L'estimateur de Bayes δ^π est défini comme l'espérance a posteriori, c'est-à-dire la valeur qui minimise la perte quadratique en prenant en compte la distribution a priori π. En d'autres termes, il s'agit de l'estimateur qui combine l'information a priori et les données observées pour obtenir la meilleure estimation possible :*

$$\delta^\pi = E^\pi[\theta|x] = \frac{\int_\theta \theta f(x|\theta)\pi(\theta)d\theta}{\int_\theta f(x|\theta)\pi(\theta)d\theta} = \delta^\pi(x).$$

Démonstration. On a :

$$E^{\pi}[(\theta-\delta)^2|x] = E^{\pi}[\theta^2|x] - 2\delta E^{\pi}[\theta|x] + \delta^2.$$

C'est un polynôme de second degré qui arrive à son minimum en :
$\delta^{\pi}(x) = E^{\pi}[\theta|x]$. □

2.4.2 Fonction de perte linex

Le terme " linex " attribué à la fonction de perte découle de sa capacité à être appliquée à la fois dans des contextes linéaires et exponentiels. Introduite par Varian (1975) [49], cette fonction présente une croissance presque exponentielle d'un côté de zéro et une linéarité approximative de l'autre côté. Elle est couramment utilisée dans l'analyse des problèmes d'estimation statistique et de prédiction, comme en témoignent les travaux de Rojo (1989) [42] et Nassar (2004) [32]. Une expression courante de la fonction de perte linex est la suivante, sous l'hypothèse que la perte minimale est obtenue pour $\tilde{u} = u$

$$(\Delta) \propto e^{(a\Delta)} - a\Delta - 1, \quad a \neq 0, \tag{2.1}$$

avec $\Delta = (\tilde{u} - u)$, où $\tilde{u}$ est une estimation de la variable u. Le signe et la norme du paramètre a indiquent respectivement la direction et l'intensité de l'asymétrie ($a > 0$ signifie que la surestimation est plus préjudiciable que la sous-estimation, et vice versa). Lorsque le paramètre a tend vers zéro, la fonction de perte linex peut être approximée par la fonction de perte quadratique ; la formule(2.1) devient :

$$E_u(L(\tilde{u}-u)) \propto e^{a\tilde{u}} E_u(e^{-au}) - a(\tilde{u} - E_u(u)) - 1, \tag{2.2}$$

$E_u(.)$ désigne l'espérance conditionnelle par rapport à la densité postérieure de la variable u. $\tilde{u}_L$ est l'estimateur de Bayes de $\tilde{u}$ selon la fonction de perte linex qui correspond à la valeur de $\tilde{u}$ qui optimise (2.2). Pour déterminer cet estimateur, il convient de dériver l'équation (2.2) par rapport à $\tilde{u}$, ce qui nous donne :

$$\frac{d}{d\tilde{u}}(E_u(L(\tilde{u}-u))) = ae^{-a\tilde{u}}E_u(e^{-au}) - a.$$

En annulant cette expression, nous parvenons au résultat suivant :

$$ae^{-a\tilde{u}}E_u(e^{-}au) = a,$$

d'où

$$e^{a\tilde{u}} = E_u(e^{-au}).$$

On introduit le logarithme à la formule précédente, on obtient :

$$-a\tilde{u} = \log E_u(e^{-}au).$$

Ainsi, $\tilde{u}_L$ l'estimateur bayésien de $\tilde{u}$ sous linex est :

$$\tilde{u}_L = -\frac{1}{a}\log(E_u(e^{-au}))E_u(e^{au}),$$

Étant donné que l'espérance $E_u(e^{-au})$ existe et est finie.

2.4.3 Fonction de perte entropie

Dans de multiples applications pratiques, il est souvent plus approprié de quantifier la perte en utilisant le rapport $\frac{\widehat{\theta}}{\theta}$. Une fonction de perte asymétrique couramment employée dans ce contexte est la perte d'entropie. Cette fonction est dérivée de linex, laquelle a été introduite par Calabria et Pulcini (1994) [6]. Elle se définit de la manière suivante :

$$L_E(\theta, d) \propto (\frac{d}{\theta})^p - p\ln(\frac{d}{\theta}) - 1.$$

La fonction précédente atteint son minimum lorsque $d = \theta$.

Sous la fonction de perte entropie, l'estimateur de Bayes pour le paramètre θ

est défini comme suit :

$$\delta(x) = (E_\theta(\theta^{-p}))^{-1/p}.$$

Chapitre 3 Primes bayésiennes du système bonus-malus en assurance automobile

Sommaire

Ce chapitre, ainsi que le chapitre suivant « Application numérique », font l'objet d'une publication internationale intitulée « On Bayesian bonus-malus Premium Under linex Loss Function With Application » dans la revue de classe B « Studies in Engineering and Exact Sciences », voir [17].

Dans les systèmes de bonus-malus, évaluer la prime uniquement sur le nombre de sinistres peut sembler injuste. En effet, un sinistre de faible montant ne devrait pas pénaliser autant qu'un sinistre entraînant des dommages considérables. Pour combler cette lacune, on suggère d'ajouter le montant des sinistres au calcul de la prime (voir [13, 31]). Dans ce chapitre, on calculera d'abord les primes du système bonus-malus en se basant uniquement sur le nombre de sinistres utilisant la distribution Poisson-Akash dans la première section. Ensuite, dans la dernière section, on présentera les primes du système bonus-malus dépendantes du nombre de sinistres avec la distribution Poisson-Akash et du montant des sinistres qui est établi sur la nouvelle distribution mixte Inverse-Gamma Lindley afin de créer une prime plus équitable et équilibrée entre les risques du portefeuille.

3.1 La fréquence de sinistres en utilisant la distribution Poisson-Akash

Dans cette section, nous allons utiliser la distribution Poisson-Akash pour calculer les primes du système bonus-malus en nous basant uniquement sur la fréquence de sinistres.

3.1.1 Distribution Poisson-Akash

Le nombre de sinistres l est supposé suivre une distribution Poisson avec le paramètre θ et la fonction de masse de probabilité suivante :

$$P(l|\theta) = \frac{e^{-\theta}\theta^{l}}{l!}, l = 0, 1, 2, ..., \theta > 0.$$

L'espérance de la variable aléatoire Poisson est :

$$E[l|\theta] = Var[l|\theta] = \theta.$$

Nous supposons que θ suit la distribution Akash avec le paramètre γ. Ainsi, la fonction de densité de probabilité (pdf) de θ peut être exprimée comme suit :

$$\pi(\theta) = \frac{\gamma^3}{\gamma^2+2}(1+\theta^2)e^{-\gamma\theta}.$$

Par conséquent, la représentation de la distribution Poisson mélangée à la distribution Akash (voir Shanker(2016) [47]) définie sur $\Re^+$ est :

$$\begin{aligned}
f(l) &= \int_0^\infty P(l|\theta)\pi(\theta)d\theta \\
&= \int_0^\infty \frac{e^{-\theta}\theta^l}{l!}\left(\frac{\gamma^3}{\gamma^2+2}\right)\left(1+\theta^2\right)e^{-\gamma\theta}d\theta \\
&= \frac{\gamma^3}{\gamma^2+2}\frac{1}{l!}\left[\int_0^\infty \theta^l e^{-\theta(1+\gamma)}d\theta + \int_0^\infty \theta^{l+2}e^{-\theta(1+\gamma)}d\theta\right] \\
&= \frac{\gamma^3}{\gamma^2+2}\frac{1}{l!}\left[\frac{\Gamma(l+1)}{(1+\gamma)^{l+1}} + \frac{\Gamma(l+3)}{(1+\gamma)^{l+3}}\right] \\
&= \frac{\gamma^3}{\gamma^2+2}\left[\frac{(1+\gamma)^2+(l+2)(l+1)}{(1+\gamma)^{l+3}}\right] \\
f(l) &= \frac{\gamma^3}{\gamma^2+2}\frac{\gamma^2+2\gamma+l^2+3l+3}{(1+\gamma)^{l+3}} \qquad (3.1)
\end{aligned}$$

3.1.2 Estimateur du Maximum de Vraisemblance

Dans le cadre de l'estimation des paramètres du modèle, la méthode couramment utilisée est l'estimation du maximum de vraisemblance (MLE). L'approche fondamentale consiste à maximiser la fonction de vraisemblance, qui vise à estimer la valeur du paramètre permettant d'obtenir les données d'observation les plus probables.

$l_1, l_2, ..., l_n$ est un échantillon aléatoire de la distribution Poisson-Akash avec une fonction de densité de probabilité dans 3.1, où la taille de l'échantillon est n.

Afin d'obtenir la valeur la plus probable du paramètre γ, la fonction de vraisemblance L doit être maximisée comme suit :

$$\begin{aligned}
L(\gamma, l_i) &= \prod_{i=1}^{n} f(l_i, \gamma) \\
&= \prod_{i=1}^{n} \frac{\gamma^3}{\gamma^2+2} \frac{l_i^2+3l_i+(\gamma^2+2\gamma+3)}{(1+\gamma)^{l_i+3}} \\
&= \left(\frac{\gamma^3}{\gamma^2+2}\right)^n \prod_{i=1}^{n} \frac{l_i^2+3l_i+(\gamma^2+2\gamma+3)}{(1+\gamma)^{l_i+3}}
\end{aligned}$$

La fonction log de vraisemblance est :

$$\begin{aligned}
\ln L(\gamma, l_i) &= n \ln\left(\frac{\gamma^3}{\gamma^2+2}\right) - \sum_{l=1}^{n} \ln(\gamma+1)^{l_i+3} \\
&\quad + \sum_{l=1}^{n} \ln\left[l_i^2+3l_i+\left(\gamma^2+2\gamma+3\right)\right] \\
&= 3n\ln\gamma - n\ln\left(\gamma^2+2\right) - \ln(1+\gamma)\sum_{i=1}^{n}(l_i+3) \\
&\quad + \sum_{i=1}^{n} \ln\left(\gamma^2+2\gamma+l_i^2+3l_i+3\right)
\end{aligned}$$

La solution de l'équation $\frac{\partial \ln L(\gamma;x_i)}{\partial\gamma} = 0$, donne l'estimateur $\widehat{\gamma}$ du paramètre γ. Ainsi

$$\begin{aligned}
\frac{\partial \ln L(\gamma; l_i)}{\partial \gamma} &= 0, \\
\frac{\partial}{\partial\gamma}\Bigg[3n\ln\gamma - n\ln\left(\gamma^2+2\right) - \ln(1+\gamma)\sum_{i=1}^{n}(l_i+3) & \\
+ \sum_{i=1}^{n}\ln\left(\gamma^2+2\gamma+l_i^2+3l_i+3\right)\Bigg] &= 0, \\
\frac{3n}{\gamma} - \frac{2n\gamma}{\gamma^2+2} - \frac{\sum_{l=1}^{n}(l_i+3)}{\gamma+1} + \sum_{l=1}^{n}\frac{2(\gamma+1)}{(\gamma^2+2\gamma+l_i^2+3l_i+3)} &= 0.
\end{aligned}$$

où $\overline{x}$ est la moyenne de l'échantillon.

3.1.3 Méthode bayésienne

$l_{1,}l_2, ..., l_t$ est un échantillon de taille t. $N = \sum_{i=1}^{t} l_i$ est le nombre total de réclamations effectuées par un assuré en t ans, où l_i désigne le nombre de réclamations effectuées par cet assuré durant la i-ème année, pour $i = 1, 2, \ldots, t$. La fonction de vraisemblance est :

$$\begin{aligned} L(\theta; l_{1,}l_2, ..., l_n) &= \prod_{i=1}^{t} \frac{e^{-\theta}\theta^{l_i}}{l!} \\ &= \frac{1}{\prod l!} e^{-t\theta}\theta^{\sum l_i} \propto e^{-t\theta}\theta^N. \end{aligned}$$

La distribution a priori est :

$$\begin{aligned} \pi(\theta) &= \frac{\gamma^3}{\gamma+2}(1+\theta^2)e^{-\gamma\theta} \\ &\propto (1+\theta^2)e^{-\gamma\theta}. \end{aligned}$$

Ainsi, pour calculer la fonction de distribution a posteriori pour un assuré ayant un historique de sinistres de $l_{1,}l_2, ..., l_t$, nous appliquons le théorème de Bayes. La fonction de distribution a posteriori est obtenue en calculant le produit de la fonction de vraisemblance et de la distribution a priori comme suit :

$$\begin{aligned} \pi^*(\theta | l_1, ..., l_n) &\propto P(l_1, ..., l_n | \theta)\pi(\theta) \\ &= e^{-\theta t}\theta^N(1+\theta^2)e^{-\gamma\theta} \\ &= \theta^N(1+\theta^2)e^{-\theta(t+\gamma)}. \end{aligned}$$

Considérons

$$\int_0^\infty \pi^*(\theta | l_1, ..., l_n)d\theta \propto \int_0^\infty e^{-\theta(t+\gamma)}(\theta^2+1)\theta^N d\theta.$$

On a

$$\int_0^\infty Ae^{-\theta(t+\gamma)}(\theta^2+1)\theta^N d\theta = 1,$$

où A est une constante obtenue comme suit :

$$\begin{aligned}
A[\int_0^\infty \theta^{N+2}e^{-(t+\gamma)\theta}d\theta + \int_0^\infty \theta^N e^{-(t+\gamma)\theta}d\theta] &= 1, \\
A[\frac{\Gamma(n+3)}{(t+\gamma)^{n+3}} + \frac{\Gamma(n+1)}{(t+\gamma)^{n+1}}] &= 1, \\
A[\frac{\Gamma(n+3)+(t+\gamma)^2\Gamma(n+1)}{(t+\gamma)^{n+3}}] &= 1.
\end{aligned}$$

Alors

$$A = \frac{(t+\gamma)^{n+3}}{\Gamma(n+1)[(n+2)(n+1)+(t+\gamma)^2]}.$$

Ainsi, la fonction de distribution a posteriori pour la fréquence des sinistres est exprimée par :

$$\begin{aligned}
\pi^*(\theta|l_1,...,l_n) &= \frac{(t+\gamma)^{n+3}}{\Gamma(n+1)[(n+2)(n+1)+(t+\gamma)^2]} \\
&\quad * \theta^N(1+\theta^2)e^{-\theta(t+\gamma)}.
\end{aligned}$$

3.1.4 Calcul de prime

Dans cette partie, nous allons calculer la prime nette qui est la prime de base, définie comme la moyenne du nombre de sinistres de chaque assuré. Supposons que l'historique des sinistres d'un assuré soit $l_1, l_2, ..., l_t$, alors la moyenne de la fonction de distribution a posteriori pour la distribution Poisson-Akash (ou le nombre attendu de sinistres de cet assuré) sera :

$$\begin{aligned}
\hat{\theta}_{t+1} &= E[\theta|l_1,...,l_t] = E[l_1,...,l_t|\theta] \\
&= \int_0^\infty \theta\pi^*(\theta_i|l_1,...,l_t) \\
&= \frac{(t+\gamma)^{n+3}}{\Gamma(n+1)[(n+2)(n+1)+(t+\gamma)^2]}
\end{aligned}$$

$$
\begin{aligned}
& \times \left[\frac{\Gamma(n+4)}{(t+\gamma)^{n+4}} + \frac{\Gamma(n+2)}{(t+\gamma)^{n+2}} \right] \\
= \; & \frac{(t+\gamma)^{n+3}}{\Gamma(n+1)[(n+2)(n+1)+(t+\gamma)^2]} \\
& \times \left[\frac{\Gamma(n+4)+(t+\gamma)^2\Gamma(n+2)}{(t+\gamma)^{n+4}} \right] \\
= \; & \frac{(n+1)\Gamma(n+1)[(n+3)(n+2)+(t+\gamma)^2]}{\Gamma(n+1)[(t+\gamma)^2+(n+2)(n+1)](t+\gamma)} \\
\hat{\theta}_{t+1} = \; & \frac{(n+1)(n+2)(n+3)+(n+1)(t+\gamma)^2}{(t+\gamma)^3+(t+\gamma)(n+2)(n+1)}.
\end{aligned}
$$

La prime initiale au temps $t = 0$ est supposée être égale à 100. Alors, la prime au temps $t + 1$ est :

$$
\begin{aligned}
Premium_{t+1} = \; & 100\frac{\gamma(\gamma^2+2)}{\gamma^2+6} \\
& \times \frac{(n+1)(n+2)(n+3)+(n+1)(t+\gamma)^2}{(t+\gamma)^3+(t+\gamma)(n+2)(n+1)}.
\end{aligned} \tag{3.2}
$$

3.1.5 Prime sous la fonction de perte linex

Dans cette partie, nous utiliserons la fonction de perte asymétrique linex [49] pour calculer les primes du système bonus-malus.

$$L(\hat{\theta}, \theta) = \exp\left(a\left(\hat{\theta} - \theta\right)\right) - a\left(\hat{\theta} - \theta\right) - 1, \qquad a \neq 0,$$

où $\hat{\theta}$ est l'estimateur de θ sous la fonction de perte linex qui minimise l'équation ci-dessus (voir Zellner (1986) [54]), $\hat{\theta}$ est donné par :

$$\hat{\theta} = -\frac{1}{a}\ln(E[e^{-a\theta}|X]).$$

On a

$$\begin{aligned}
E[e^{-a\theta}|l_1,...,l_t] &= \int_0^\infty e^{-a\theta}\pi^*(\theta_i|l_1,...,l_t)d\theta \\
&\propto \int_0^\infty \theta^{N+2}e^{-(t+\gamma+a)\theta}d\theta + \int_0^\infty \theta^N e^{-(t+\gamma+a)\theta}d\theta \\
&= \left[\frac{\Gamma(n+3)}{(t+\gamma+a)^{n+3}} + \frac{\Gamma(n+1)}{(t+\gamma+a)^{n+1}}\right] \\
&\quad * \frac{(t+\gamma)^{n+3}}{\Gamma(n+1)(n+2)(n+1)+(t+\gamma)^2} \\
E[e^{-a\theta}|l_1,...,l_n] &= \frac{(n+2)(n+1)+(t+\gamma+a)^2}{(t+\gamma+a)^{n+3}} \\
&\quad * \frac{(t+\gamma)^{n+3}}{(n+2)(n+1)+(t+\gamma)^2}.
\end{aligned}$$

Enfin

$$\hat{\theta} = -\frac{1}{a}\ln\left(\frac{(t+\gamma)^{n+3}}{(t+\gamma+a)^{n+3}}\frac{(n+2)(n+1)+(t+\gamma+a)^2}{(n+2)(n+1)+(t+\gamma)^2}\right).$$

Au temps $t=0$, la prime initiale est supposée être égale à 100. La prime au temps $t+1$ est alors donnée comme suit :

$$\begin{aligned}
Premium_{t+1} &= 100\frac{\gamma(\gamma^2+2)}{\gamma^2+6} \\
&\quad * \left[-\frac{1}{a}\ln(\frac{(t+\gamma)^{n+3}}{(t+\gamma+a)^{n+3}}\right. \\
&\quad \left. * \frac{(n+2)(n+1)+(t+\gamma+a)^2}{(n+2)(n+1)+(t+\gamma)^2})\right].
\end{aligned} \tag{3.3}$$

3.2 Le montant de sinistres en utilisant la distribution Inverse-Gamma Lindley

Dans cette section, nous allons introduire une nouvelle distribution mixte afin de l'utiliser pour calculer la partie du montant des sinistres dans les primes du système bonus-malus. Cette distribution est l'Inverse-Gamma Lindley.

3.2.1 Distribution Inverse-Gamma Lindley

La taille des sinistres de chaque assuré est représentée par une variable aléatoire X. Nous supposons que X suit une distribution Inverse-Gamma avec une fonction de densité donnée par :

$$f(x|\lambda) = \frac{\lambda^{\alpha} x^{-\alpha-1}}{\Gamma(\alpha)} e^{-\frac{\lambda}{x}}.$$

L'espérance de la variable aléatoire Inverse-Gamma est :

$$E[X] = \tfrac{\lambda}{\alpha-1}.$$

Supposant que λ suit la distribution Lindley avec le paramètre β. Par conséquent, la fonction de densité de probabilité de λ est donnée ci-dessous

$$\pi(\lambda) = \frac{\beta^2}{\beta+1}(\lambda+1)e^{-\beta\lambda}.$$

Par conséquent, la distribution Inverse-Gamma mélangée avec la distribution Lindley définie sur $\Re^+$ est obtenue comme suit :

$$\begin{aligned}
f(x) &= \int_0^{\infty} \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\frac{\lambda}{x}} \frac{\beta^2}{\beta+1}(\lambda+1)e^{-\beta\lambda} d\lambda \\
&= \frac{x^{-\alpha-1}\beta^2}{\Gamma(\alpha)(\beta+1)} \int_0^{\infty} \lambda^{\alpha}(\lambda+1)e^{-\lambda(\beta+\frac{1}{x})} d\lambda \\
&= \frac{x^{-\alpha-1}\beta^2}{\Gamma(\alpha)(\beta+1)} \left[\int_0^{\infty} \lambda^{\alpha+1} e^{-\lambda(\beta+\frac{1}{x})} d\lambda + \int_0^{\infty} \lambda^{\alpha} e^{-\lambda(\beta+\frac{1}{x})} d\lambda\right] \\
&= \frac{x^{-\alpha-1}\beta^2}{\Gamma(\alpha)(\beta+1)} \left[\frac{\Gamma(\alpha+2)}{(\beta+\frac{1}{x})^{\alpha+2}} + \frac{\Gamma(\alpha+1)}{(\beta+\frac{1}{x})^{\alpha+1}}\right] \\
&= \frac{x^{-\alpha-1}\beta^2\Gamma(\alpha+1)}{\Gamma(\alpha)(\beta+1)(\beta+\frac{1}{x})^{\alpha+2}} \left[\alpha+1+\beta+\frac{1}{x}\right] \\
f(x) &= \frac{x^{-\alpha-1}\beta^2(\alpha+1)}{(\beta+1)(\beta+\frac{1}{x})^{\alpha+2}} \left[\alpha+1+\beta+\frac{1}{x}\right]. \qquad (3.4)
\end{aligned}$$

3.2.2 Estimateur du Maximum de Vraisemblance

$X = (x_1, x_2, \ldots, x_n)^T$ est un vecteur d'observations indépendantes et identiquement distribuées pour la distribution Inverse-Gamma Lindley avec une fonction de densité de probabilité dans (3.4). La valeur la plus probable du paramètre β qui maximise la fonction de vraisemblance est obtenue comme suit.

La fonction de vraisemblance est :

$$\begin{aligned} L(\beta, x_i) &= \prod_{i=1}^{n} f(x_i, \beta_i) \\ &= \prod_{i=1}^{n} \frac{x_i^{-\alpha-1}\beta^2(\alpha+1)}{(\beta+1)(\beta+\frac{1}{x_i})^{\alpha+2}}\left[\alpha+1+\beta+\frac{1}{x_i}\right] \\ L(\beta, x_i) &= \frac{\beta^{2n}(\alpha+1)^n}{(\beta+1)^n}\prod_{i=1}^{n}\frac{x_i^{-\alpha-1}}{(\beta+\frac{1}{x_i})^{\alpha+2}}\left[\alpha+1+\beta+\frac{1}{x_i}\right]. \end{aligned}$$

La fonction log de vraisemblance est exprimée comme suit :

$$\begin{aligned} \ln L(\beta, x_i) &= 2n\ln\beta + n\ln(\alpha+1) - n\ln(\beta+1) + \sum_{i=1}^{n}(-\alpha-1)\ln x_i \\ &\quad -(\alpha+2)\sum_{i=1}^{n}\ln(\beta+\frac{1}{x_i}) + \sum_{i=1}^{n}\ln(\alpha+1+\beta+\frac{1}{x_i}) \\ &= n\left[-\ln(\beta+1) + 2\ln\beta + \ln(\alpha+1)\right] - (\alpha+1)\sum_{i=1}^{n}\ln x_i \\ &\quad -(\alpha+2)\sum_{i=1}^{n}\ln(\beta+\frac{1}{x_i}) + \sum_{i=1}^{n}\ln(\alpha+1+\beta+\frac{1}{x_i}). \end{aligned}$$

Par conséquent, l'estimateur $\widehat{\beta}$ du paramètre β est :

$$\begin{aligned} \frac{\partial \ln L(\beta; x_i)}{\partial \beta} &= 0, \\ \frac{2n}{\beta} - \frac{n}{\beta+1} - (\alpha+2)\sum_{i=1}^{n}\left(\frac{1}{\beta+\frac{1}{x_i}}\right) + \sum_{i=1}^{n}\frac{1}{\alpha+1+\beta+\frac{1}{x_i}} &= 0. \end{aligned}$$

L'estimateur n'est pas sous sa forme explicite, des solutions numériques peuvent être appliquées pour le déterminer.

3.2.3 Méthode bayésienne

$N = \sum_{i=1}^{t} l_i$ est le nombre total de sinistres effectués par un assureur sur t ans. Soit X la taille du sinistre l pour $l = 1, 2, ..., N$. La fonction de vraisemblance est :

$$\begin{aligned} L(\lambda|x_{1,...,}x_n) &= f(x_{1,...,}x_n|\lambda) = \prod_{i=1}^{n} \frac{\lambda^{\alpha} x^{-\alpha-1}}{\Gamma(\alpha)} e^{-\lambda \frac{1}{x_i}} \\ &= \frac{\lambda^{n\alpha}}{(\Gamma(\alpha))^n} \prod_{i=1}^{n} x^{-\alpha-1} e^{-\lambda \sum_{i=1}^{n} \frac{1}{x_i}} \\ &\propto \lambda^{n\alpha} e^{-\lambda \sum_{i=1}^{n} \frac{1}{x_i}}. \end{aligned}$$

La distribution a priori est :

$$\pi(\lambda) \propto (\lambda + 1) e^{-\beta\lambda}.$$

Le théorème de Bayes est appliqué pour obtenir la fonction de distribution a posteriori comme suit :

$$\begin{aligned} \pi^*(\lambda|x_{1,...,}x_n) &\propto f(x_{1,...,}x_n|\lambda)\pi(\lambda) \\ &\propto \lambda^{n\alpha}(1+\lambda) e^{-\lambda\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)}. \end{aligned}$$

On considère que

$$\int_0^{\infty} \pi^*(\lambda|x_{1,...,}x_n) d\lambda \propto \int_0^{\infty} \lambda^{n\alpha}(1+\lambda) e^{-\lambda\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)} d\lambda.$$

Alors

$$\int_0^{\infty} \pi^*(\lambda|x_{1,...,}x_n) d\lambda = \int_0^{\infty} B\lambda^{n\alpha}(1+\lambda) e^{-\lambda\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)} d\lambda = 1,$$

où B est une constante obtenue comme suit :

$$\begin{aligned} B\left[\frac{\Gamma(\alpha n + 2)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n + 2}} + \frac{\Gamma(\alpha n + 1)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n + 1}}\right] &= 1, \\ B\left[\frac{\Gamma(\alpha n + 1)\left(\alpha n + 1 + \gamma + \sum_{i=1}^{n} \frac{1}{x_i}\right)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n + 2}}\right] &= 1, \end{aligned}$$

Alors

$$B = \frac{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n+2}}{\Gamma(\alpha n+1)\left(\alpha n + 1 + \beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)}.$$

Ainsi, la fonction de distribution a posteriori pour le montant est donnée par :

$$\begin{aligned}\pi^*(\lambda|x_{1,\dots,}x_n) &= \frac{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n+2}}{\Gamma(\alpha n+1)\left(\alpha n + 1 + \beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)} \\ &\quad \times \lambda^{n\alpha}(1+\lambda)e^{-\lambda\left(\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}.\end{aligned} \tag{3.5}$$

3.2.4 Calcul de prime

Dans cette partie, comme dans la fréquence des sinistres, nous utiliserons la prime nette.

L'espérance de la formule précédente (3.5) pour la distribution Inverse-Gamma Lindley est :

$$\begin{aligned}\widehat{\lambda}_{t+1} &= E\left[\lambda|x_1,\dots,x_n\right] \\ &= \int_0^{\infty} \lambda\pi^*(\lambda|x_1,\dots,x_n)d\lambda \\ &= \frac{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n+2}}{\Gamma(\alpha n+1)\left(\alpha n + 1 + \beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)} \\ &\quad \times \left[\frac{\Gamma(\alpha n+2)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n+2}} + \frac{\Gamma(\alpha n+3)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n+3}}\right] \\ &= \frac{(\Gamma(\alpha n+2))}{\Gamma(\alpha n+1)\left(\alpha n + 1 + \beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)} \\ &\quad \times \left[\frac{(\alpha n + 2 + \beta + \sum_{i=1}^{n} \frac{1}{x_i})}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{\alpha n+3}}\right] \\ \widehat{\lambda}_{t+1} &= \frac{(\alpha n+1)(\alpha n + 2 + \beta + \sum_{i=1}^{n} \frac{1}{x_i})}{(\beta + \sum_{i=1}^{n} \frac{1}{x_i})(\alpha n + 1 + \beta + \sum_{i=1}^{n} \frac{1}{x_i})}.\end{aligned}$$

À partir de

$$E\left[\lambda|x_1,...,x_n\right] \;=\; \widehat{\lambda}.$$

Alors

$$E\left[x_1,...,x_n|\lambda\right] \;=\; \frac{\widehat{\lambda}}{\alpha-1}.$$

Donc

$$E\left[x_1,...,x_n|\lambda\right] = \frac{(\alpha n+1)(\alpha n+2+\gamma+\sum_{i=1}^{n}\frac{1}{x_i})}{(\alpha-1)\left(\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)(\alpha n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i})}.$$

Finalement, la prime finale basée sur la distribution Poisson-Akash pour la fréquence des sinistres et l'Inverse-Gamma Lindley pour le montant des sinistres que l'assuré doit payer est :

$$\begin{aligned} Premium_{t+1} &= \frac{(n+1)(n+2)(n+3)+(n+1)(t+\gamma)^2}{(t+\gamma)^3+(t+\gamma)(n+2)(n+1)} \\ &\times\frac{(\alpha n+1)(\alpha n+2+\beta+\sum_{i=1}^{n}\frac{1}{x_i})}{(\alpha-1)\left(\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)(\alpha n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i})}. \end{aligned} \tag{3.6}$$

Maintenant, en appliquant la distribution Poisson-Akash sous la fonction de perte linex pour la fréquence des sinistres et la distribution Inverse-Gamma Lindley pour le montant des sinistres, la prime finale que l'assuré doit payer est :

$$\begin{aligned} Premium_{t+1} &= -\frac{1}{a}\ln(\frac{(n+2)\left(n+1\right)+\left(t+\gamma+a\right)^2}{\left(t+\gamma+a\right)^{n+3}} \\ &\times\frac{\left(t+\gamma\right)^{n+3}}{(n+2)\left(n+1\right)+\left(t+\gamma\right)^2}) \\ &\times\frac{(\alpha n+1)(\alpha n+2+\beta+\sum_{i=1}^{n}\frac{1}{x_i})}{(\alpha-1)\left(\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)(\alpha n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i})}. \end{aligned} \tag{3.7}$$

Chapitre 4 Application numérique

Sommaire

Pour l'application numérique, nous avons utilisé un ensemble de données réelles pour comparer les primes basées uniquement sur le nombre de sinistres avec celles basées à la fois sur la fréquence et le montant des sinistres. Nous avons examiné l'ensemble de données basé sur des polices d'assurance automobile d'un an souscrites en 2004 ou 2005. Cet ensemble de données est accessible sur le site internet de la faculté de commerce et d'économie de l'université Macquarie (Sydney, Australie), voir également de Jong et Heller (2008) [9]. Au total, le portefeuille comprend 67 856 contrats dont 4 624 présentent au moins un sinistre. Parmi ces polices, 4 333 ont fait une réclamation, 271 ont fait deux réclamations, 18 ont fait trois réclamations et 2 ont fait quatre réclamations.

4.1 Les primes basées uniquement sur la fréquence des sinistres

Pour calculer les primes basées exclusivement sur la fréquence des sinistres, nous utilisons, pour le premier tableau, la distribution Poisson-Akash introduite plus haut dans (3.1). $\hat{\gamma} = 14.0125$ est l'estimateur du maximum de vraisemblance du paramètre de cette distribution. Les primes établies uniquement sur la fréquence des sinistres sont calculées à partir de (3.3), les résultats sont présentés dans le tableau suivant :

t	nombre de sinistres				
	0	1	2	3	4
0	100				
1	93.10335	187.7328	283.7319	380.8901	478.9653
2	87.10775	175.4826	265.01	355.5334	446.8701
3	81.84590	164.755	248.6416	333.388	418.8542
4	77.18976	155.28	234.2057	313.8769	394.1853
5	73.03964	146.8484	221.3763	296.5537	372.2957
6	69.31671	139.2954	209.8972	281.0674	352.7389
7	65.95779	132.4893	199.564	267.1384	335.1592

TABLEAU 4.1 – Les primes en utilisant la distribution Poisson-Akash basées uniquement sur la fréquence des sinistres.

D'après le tableau 4.1, on remarque que les primes augmentent en cas de sinistre et diminuent lorsque l'assuré a une année sans sinistre. Selon les résultats de ce tableau, un assuré qui a eu un sinistre la première année devra payer un malus égal à 87,73 % de la prime de base. Un bonus de 6,89% de la prime de base sera ajouté à la prime pour la première année sans sinistre.

Dans le tableau suivant, pour calculer les primes, nous utilisons la distribution Poisson-Akash sous la fonction de perte linex en (3.4). Nous avons pris a égal à 1.1 (surestimation). Le tableau 4.2 montre les résultats des primes.

t	nombre de sinistres				
	0	1	2	3	4
0	96.13338				
1	89.74638	180.8718	273.2438	366.6832	460.9817
2	84.16501	169.4815	255.8516	343.1409	431.1913
3	79.24454	159.4601	240.5731	322.4814	405.0633
4	74.87319	150.5725	227.0418	304.2025	381.9597
5	70.96317	142.6349	214.9718	287.9125	361.3823
6	67.44461	135.5014	204.1363	273.3010	342.9363
7	64.26111	129.0547	194.3536	260.1196	326.3050

Tableau 4.2 – Les primes en considérant la distribution Poisson-Akash sous linex (surestimation) basées uniquement sur la fréquence des sinistres.

Comme nous pouvons le voir, la prime à $t = 0$ est 96.13338. Un sinistre effectué par un assuré la première année entraînera un paiement de malus de 80.87 % de la prime de base. Un bonus de 10.25 % de la prime de base sera ajouté à la prime pour la première année sans sinistre.
Pour le tableau suivant, les primes sont calculées à l'aide de la distribution Poisson-Akash sous la fonction de perte linex en (3.4). Nous avons pris, cette fois, a égal à -0,3 (sous-estimation). Le tableau 4.3 présente les résultats des primes.

t	nombre de sinistres				
	0	1	2	3	4
0	101.1294				
1	94.07916	189.7294	286.7865	385.0298	484.2065
2	87.95956	177.2214	267.6655	359.1284	451.4195
3	82.59612	166.2833	250.9721	336.5396	422.8403
4	77.85570	156.6341	236.2678	316.6628	397.7068
5	73.63484	148.0569	223.2142	299.0345	375.4296
6	69.85195	140.3807	211.5459	283.2909	355.5462
7	66.44175	133.4695	201.0515	269.1430	337.6885

Tableau 4.3 – Les primes en utilisant la distribution Poisson-Akash sous linex (sous-estimation) basées uniquement sur la fréquence des sinistres.

Les résultats du tableau 4.3 montrent que la prime initiale est de 101,1294. Un assuré aura 89.72% de malus à payer sur la prime de base s'il fait un sinistre au cours de la première année. Néanmoins, l'absence de sinistre durant la première année se traduira par l'application d'un bonus équivalent à 5.92 % de la prime initiale.
À titre de comparaison, nous présentons dans le tableau suivant les primes bayésienne du système bonus-malus établies sur la distribution Poisson-Lindley, comme le montre le tableau 2 en [31].

t	nombre de sinistres				
	0	1	2	3	4
0	100				
1	93.26	185.92	278.08	369.81	461.17
2	87.37	174.23	260.67	346.74	432.50
3	82.17	163.92	245.30	326.37	407.17
4	77.56	154.75	231.63	308.24	384.61
5	73.43	146.55	219.40	292.01	364.41
6	69.72	139.18	208.39	277.39	346.21
7	66.37	132.50	198.42	264.16	329.74

Tableau 4.4 – Les primes en considérant la distribution Poisson-Lindley basées uniquement sur la fréquence des sinistres.

D'après les résultats du tableau 4.4, on remarque qu'un assuré qui a fait un sinistre la première année aura 85,92 % de malus à payer de la prime de base. Cependant, l'absence de sinistre la première année sera indiquée dans la prime sous forme de bonus à 6,74% de la prime de base.
En comparant les tableaux 1 à 4, nous remarquons que les primes déterminées à partir de la distribution Poisson-Akash sous la fonction de perte de linex en cas de surestimation ($a = 1.1$) pour la fréquence uniquement sont les plus souples envers les bons conducteurs. Par contre, les primes fondées sur la même distribution en cas de sous-estimation ($a = -0.3$) sont les plus sévères pour les mauvais conducteurs.
De plus, les primes augmentent rapidement lorsque le nombre de sinistres augmente. Ces primes diminuent lentement avec le temps lorsqu'il n'y a pas de sinistre.

Remarque 4.1.1. *Nous avons remarqué que lorsque la valeur de a tend vers*

0, les primes basées sur la distribution Poisson-Akash sous la fonction de perte linex (surestimation et sous-estimation) pour la fréquence des sinistres seront exactement les mêmes primes que dans le tableau 4.1 pour les primes basées sur la distribution Poisson-Akash.

4.2 Les primes en utilisant la distribution Lognormal-Gamma pour le montant des sinistres

Dans cette section, nous allons calculer les primes du système bonus-malus en utilisant la distribution Lognormal-Gamma pour le montant des sinistres et deux différentes distributions pour la fréquence des sinistres. Le tableau suivant représente la taille de la réclamation correspondant au nombre de réclamations et à la taille totale des réclamations. Cette information servira de base a la création des tableaux qui suivront dans cette étude.

nombre de sinistres	0	1	2	3	4
montant des sinistress		235	471	706	942
Montant total des sinistres		235	706	1412	2354

Tableau 4.5 – Le montant d'un sinistre et le montant total des sinistres correspondant au nombre de sinistres.

Comme on peut le voir, le montant d'un seul sinistre sera de 235, et le montant de deux sinistres sera le cumul du premier sinistre (235) et du second sinistre (471), soit un total de 706, et ainsi de suite.
Le tableau suivant présente les primes en utilisant la distribution Poisson-

Akash pour la fréquence des sinistres et la distribution Lognormal-Gamma pour le montant des sinistres.

t	nombre de sinistres				
	0	1	2	3	4
0	592.9014				
1	552.0111	1052.479	1518.483	1963.454	2396.876
2	516.4631	983.8018	1418.287	1832.743	2236.263
3	485.2655	923.6596	1330.686	1718.585	2096.064
4	457.6592	870.5402	1253.428	1618.008	1972.613
5	433.0531	823.2705	1184.767	1528.708	1863.072
6	410.9797	780.9264	1123.333	1448.877	1765.204
7	391.0647	742.7698	1068.031	1377.075	1677.230

Tableau 4.6 – Les primes en utilisant la distribution Poisson-Akash pour la fréquence des sinistres.

Comme on peut le voir, la prime initiale est de 592,9014. Si un assuré a une année sans sinistre, la prime sera de 552,0111 pour l'année suivante. Mais si cet assuré a un sinistre au cours de la première année, la prime qui devra être payée pour l'année suivante sera de 1052.479. Cette prime atteindra 2396.876 en cas de réalisation de quatre sinistres durant la première année. Ce qui indique que les primes augmentent rapidement lorsque le nombre de sinistres augmente. Par contre, les primes diminuent lentement avec les années où il n'y a pas de sinistre.

Le tableau suivant montre les primes fondées sur la distribution Poisson-Akash sous la fonction de perte linex pour la fréquence des sinistres et la distribution Lognormal-Gamma pour le montant des sinistres. Nous avons pris $a = 1.1$ (surestimation).

t	nombre de sinistres				
	0	1	2	3	4
0	569.9762				
1	532.1075	1014.015	1462.353	1890.219	2306.882
2	499.0155	950.1578	1369.272	1768.861	2157.802
3	469.8420	893.9753	1287.505	1662.363	2027.050
4	443.9242	844.1493	1215.088	1568.137	1911.433
5	420.7417	799.6490	1150.491	1484.163	1808.458
6	399.8800	759.6565	1092.502	1408.843	1716.149
7	381.0050	723.5143	1040.146	1340.893	1632.921

TABLEAU 4.7 – Les primes en utilisant la distribution Poisson-Akash sous linex (surestimation) pour la fréquence des sinistres.

Les résultats montrent que la prime initiale est de 569.9762 ; pour un assuré qui a eu une année sans sinistre, la prime sera de 532.1075 pour l'année suivante. Cependant, si cet assuré a un sinistre au cours de la première année, la prime qui devra être payée pour l'année suivante sera de 1014.015, mais s'il fait quatre réclamations la première année, il paiera 2306.882.

On remarque que lorsqu'il n'y a pas de réclamation, les primes diminuent lentement avec le temps. Cependant, les primes augmentent avec la hausse du nombre de sinistres.

Le tableau suivant montre les primes en utilisant la distribution Poisson-Akash sous la fonction de perte linex pour la fréquence des sinistres avec $a = -0.3$ (sous-estimation) et la distribution Lognormal-Gamma pour le montant des sinistres.

t	nombre de sinistres				
	0	1	2	3	4
0	599.5978				
1	557.7966	1063.673	1534.831	1984.794	2423.105
2	521.5134	993.5497	1432.499	1851.275	2259.030
3	489.7136	932.2276	1343.158	1734.832	2116.011
4	461.6075	878.1322	1264.464	1632.369	1990.236
5	436.5820	830.0456	1194.603	1541.496	1878.755
6	414.1532	787.0109	1132.156	1460.339	1779.253
7	393.9341	748.2649	1075.992	1387.408	1689.888

Tableau 4.8 – Les primes en utilisant Poisson-Akash sous linex (sous-estimation) pour la fréquence des sinistres.

D'après les résultats du tableau 4.8, nous observons que la prime initiale est de 599,5978. Si un assuré fait un sinistre, la prime de l'année prochaine sera de 1063,673. Cette prime continuera d'augmenter pour atteindre 2423.105 pour quatre sinistres. Si un assuré ne fait plus des sinistre, la prime sera de 1689.888 au temps $t = 7$. Les résultats montrent que les primes augmentent rapidement lorsque le nombre de sinistres augmente. Ces primes diminuent lentement lorsqu'il n'y a pas de sinistres pour les prochaines années.
Pour but de comparaison, nous introduisons dans le tableau suivant les primes basées sur la distribution Poisson-Lindley pour la fréquence du sinistre et la distribution Lognormale-Gamma pour le montant des sinistres, représenté comme le tableau 4 dans [31].

t	nombre de sinistres				
	0	1	2	3	4
0	592.53				
1	552.60	1041.67	1487.29	1905.13	2306.30
2	517.67	976.18	1394.19	1786.31	2162.93
3	486.90	918.40	1311.99	1681.35	2036.22
4	459.56	867.05	1238.89	1587.96	1923.43
5	435.12	821.10	1173.46	1504.34	1822.40
6	413.14	779.76	1114.56	1429.04	1731.40
7	393.26	742.37	1061.26	1360.88	1649.00

Tableau 4.9 – Les primes en employant la distribution Poisson-Lindley pour la fréquence des sinistres.

La prime initiale pour un assuré dans le tableau 4.9 est de 592.53. Un assuré qui n'a pas eu de sinistres dans l'année paiera 552.60 pour l'année suivante. Cependant, si cet assuré fait une réclamation, la prime qui devra être payée pour l'année suivante sera de 1041.67.

D'après les tableaux [4.6, 4.9], on observe que le modèle basé sur la distribution Poisson Akash sous la fonction de perte linex en cas de surestimation ($a = 1.1$) pour la fréquence des sinistres et Lognormal-Gamma pour le montant des sinistres est le plus généreux envers les bons conducteurs. Cependant, les primes en utilisant le même modèle en cas de sous-estimation ($a = -0.3$) sont plus strictes pour les mauvais conducteurs.

De plus, les primes augmentent rapidement lorsque le nombre de sinistres augmente. Ces primes diminuent lentement avec le temps lorsqu'il n'y a pas de sinistre.

Par ailleurs, nous remarquons que les primes en utilisant la distribution Poisson-Akash sous la fonction de perte linex en cas de surestimation et les primes en utilisant la loi Poisson-Lindley sont légèrement différentes pour les mauvais conducteurs, en particulier lorsque l'assuré fait quatre réclamations.

4.3 Les primes en considérant la distribution Inverse-Gamma Lindley pour le montant des sinistres

Dans cette section, nous présentons les primes du système bonus-malus en utilisant la distribution Inverse-Gamma Lindley présentée précédemment pour le montant des sinistres avec différentes distributions de fréquence des sinistres. Les paramètres de la distribution Inverse-Gamma Lindley dans (3.4) sont : $\hat{\alpha}$=1,08 et $\hat{\beta}$=0,001766.

Le tableau suivant présente les primes basées sur la distribution Poisson-Akash pour la fréquence des sinistres et la distribution Inverse Gamma Lindley pour le montant des sinistres en utilisant la formule (3.5).

t	nombre de sinistres				
	0	1	2	3	4
0	515.3173				
1	479.7776	590.1701	1001.840	1537.207	2183.010
2	448.8813	551.6597	935.7339	1434.872	2036.728
3	421.7660	517.9354	877.9382	1345.497	1909.038
4	397.7722	488.1491	826.9660	1266.754	1796.603
5	376.3859	461.6429	781.6661	1196.840	1696.836
6	357.2010	437.8988	741.1339	1134.340	1607.701
7	339.8919	416.5028	704.6480	1078.125	1527.576

Tableau 4.10 – Les primes en utilisant la distribution Poisson-Akash pour la fréquence des sinistres.

D'après les résultats, nous remarquons que la prime initiale pour la distribution Poisson-Akash est de 515.3173. Cette prime diminuera à 479.7776 si un assuré a une année sans sinistre. La prime passera à 590.1701 si cet assuré fait un sinistre au cours de la première année.

Cette prime continuera d'augmenter rapidement par l'augmentation du nombre de sinistres pour atteindre 2183.010 lorsque l'assuré fait quatre sinistres. Si l'assuré ne fait aucun sinistre pour les prochaines années, la prime diminuera lentement pour atteindre 1527.576.

Le tableau ci-dessous décrit les primes basées sur la distribution Poisson-Akash sous la fonction de perte linex pour la fréquence des sinistres et la distribution Inverse-Gamma Lindley pour le montant des sinistres en appliquant la formule (3.6) avec $a = 1.1$ (surestimation).

t	nombre de sinistres				
	0	1	2	3	4
0	495.3919				
1	462.4786	568.6013	964.8070	1479.871	2101.046
2	433.7168	532.7940	903.3960	1384.858	1965.268
3	408.3608	501.2901	849.4486	1301.480	1846.183
4	385.8345	473.3505	801.6707	1227.710	1740.882
5	365.6855	448.3973	759.0520	1161.966	1647.095
6	347.5537	425.9718	720.7927	1102.996	1563.023
7	331.1486	405.7054	686.2506	1049.798	1487.221

Tableau 4.11 – Les primes en utilisant la distribution Poisson-Akash sous la fonction de perte linex (surestimation) pour la fréquence des sinistres.

La prime initiale indiquée dans le tableau 4.11 est de 495.3919, la prime monte à 568.6013 lorsque l'assuré fait une réclamation au cours de l'année. Si l'assuré a une année sans sinistre, la prime pour l'année suivante sera de 462.4786.

D'après les résultats de ce tableau, on observe que lorsque le nombre de sinistres augmente, les primes augmentent rapidement pour atteindre 2101.046 si un assuré fait quatre sinistres. Ces primes diminueront lentement avec le temps.

Le tableau suivant décrit les primes basées sur la distribution Poisson-Akash sous la fonction de perte linex pour la fréquence des sinistres et la distribution Inverse-Gamma Lindley pour le montant des sinistres en prenant $a = -0.3$ (sous-estimation).

t	nombre de sinistres				
	0	1	2	3	4
0	521.1374				
1	484.8061	596.4468	1012.625	1553.915	2206.899
2	453.2708	557.1258	945.1103	1449.381	2057.463
3	425.6321	522.7398	886.1668	1358.216	1927.206
4	401.2038	492.4062	834.2470	1277.997	1812.653
5	379.4530	465.4420	788.1557	1206.852	1711.119
6	359.9591	441.3106	746.9554	1143.314	1620.495
7	342.3858	419.5841	709.9004	1086.215	1539.104

TABLEAU 4.12 – Les primes en utilisant la distribution Poisson-Akash sous linex (sous-estimation) pour la fréquence des sinistres.

Le tableau 4.12 montre que la prime initiale est de 521.1374. La prime s'élèvera à 596.4468 si l'assuré fait une réclamation, cette prime continuera d'augmenter pour atteindre 2206.899 lorsque l'assuré fait quatre réclamations. Si cet assuré ne fait aucun sinistre pour les années suivantes, la prime au temps t=7 sera de 1539.104. Cela indique que les primes augmentent rapidement avec l'augmentation du nombre de sinistres. Ces primes diminuent lentement lorsqu'il n'y a pas de sinistre.

Le tableau suivant représente les primes basées sur la distribution Poisson-Lindley de [31] pour la fréquence des sinistres et la distribution Inverse-Gamma Lindley pour le montant des sinistres.

t	nombre de sinistres				
	0	1	2	3	4
0	514.9946				
1	480.2882	584.1088	981.2601	1491.545	2100.599
2	449.9433	547.3892	919.8378	1398.526	1970.016
3	423.1892	514.9887	865.6050	1316.350	1854.603
4	399.4262	486.1908	817.3748	1243.235	1751.876
5	378.1808	460.4284	774.2070	1177.768	1659.861
6	359.0742	437.2474	735.3475	1118.812	1576.973
7	341.7996	416.2795	700.1844	1065.447	1501.923

Tableau 4.13 – Les primes fondées sur la distribution Poisson-Lindley pour la fréquence des sinistres.

Les résultats illustrés dans le tableau 4.13 montrent que la prime à t=0 est de 514.9946. Si l'assuré ne fait pas de réclamation au cours de la première année, la prime sera réduite à 480.2882. Sinon, la prime passera à 584.1088 s'il y'a un sinistre.

Ces primes continueront d'augmenter lentement avec la hausse du nombre de sinistres atteignant 2100.599 pour quatre sinistres effectués au cours de l'année. Si l'assuré ne fait aucune réclamation au cours des prochaines années, la prime diminuera rapidement pour atteindre 1501.923.

En comparant les tableaux [4.10, 4.13], on remarque que les primes en utilisant la distribution Poisson-Akash sous la fonction de perte linex en cas de surestimation ($a = 1.1$) pour la fréquence des sinistres et la distribution Inverse Gamma Lindley pour le montant des sinistres sont les plus souples pour les mauvais conducteurs. Cependant, les primes en utilisant le même modèle en cas de sous-estimation ($a = -0,3$) sont plus sévères pour les mauvais conducteurs.

De plus, nous observons que les primes augmentent rapidement lorsque le nombre de sinistres augmente et diminuent lentement lorsqu'il n'y a pas de sinistre. Ce qui signifie que les conducteurs sont pénalisées plus sévèrement lorsque le nombre de sinistres est grand.

Chapitre 5
Primes en utilisant un nouveau modèle à un paramètre pour le montant des sinistres

Sommaire

Ce chapitre fait l'objet d'une publication internationale sous le titre « Bayesian bonus-malus Premium of Novel One Parameter Model for the Severity Component Case » dans la revue de classe B « Studies in Science of Science » (voir [37]).

Ce chapitre débutera par le calcul des primes en se fondant exclusivement sur le nombre de sinistres, en utilisant la distribution Poisson-New XLindley dans la première partie. Par la suite, la deuxième section exposera les primes qui varient en fonction de la fréquence des sinistres selon la distribution Poisson-New XLindley, ainsi que du montant des sinistres établi selon la distribution Inverse-Gamma(2, λ) Lindley. Cette approche vise à instaurer une prime plus juste et équilibrée en tenant compte des différents risques du portefeuille. La troisième partie expose une application numérique qui fait usage du logiciel R et de données réelles afin d'illustrer les résultats théoriques. La dernière section compare les primes de ce chapitre avec celles du précédent.

5.1 La fréquence des sinistres en utilisant la distribution Poisson-New XLindley

Dans cette partie, nous allons nous servir de la distribution Poisson-New XLindley pour déterminer les primes dans le cadre du système bonus-malus, en tenant compte exclusivement de la fréquence des sinistres.

5.1.1 Distribution Poisson-New Xlindley

On suppose que le nombre de sinistres l suit une distribution de Poisson avec le paramètre θ, ce qui se traduit par la fonction de masse de probabilité suivante :

$$P(l|\theta) = \frac{e^{-\theta}\theta^{l}}{l!}, l = 0, 1, 2, ..., \theta > 0.$$

On suppose que le paramètre θ est distribué selon la loi New XLindley avec le paramètre γ (voir Khodja [21]). Par conséquent, la représentation de la fonction de densité de probabilité (pdf) pour θ peut être formulée comme ceci :

$$\pi(\theta) = \frac{\theta}{2}(1 + \theta\ x)e^{-\theta\ x}, \quad x, \theta > 0.$$

Ainsi, l'expression de la distribution Poisson mélangée à la distribution New XLindley (voir [33, 46]) définie sur $\Re^+$ est

$$\begin{aligned} P_{PNXLD}(x,\theta) &= \int_0^\infty P(l|\theta)\pi(\theta)d\theta \\ &= \int_0^\infty \frac{e^{-\theta}\theta^l}{l!}\frac{\theta}{2}(1+\theta\ x)e^{-\theta\ x}d\theta \\ &= \frac{\theta(\theta x + 2\theta + 1)}{2(1+\theta)^{x+2}}, \quad x,\theta > 0. \end{aligned} \tag{5.1}$$

Le r^{ieme} moment de la nouvelle distribution Poisson-New XLindley est donné par (voir Seghier [46]).

$$\mu'_{(r)} = \frac{(r+2)r!}{2\theta^r}, r = 1, 2, \ldots$$

5.1.2 Estimateur du Maximum de Vraisemblance

L'échantillon aléatoire $l_1, l_2, ..., l_n$ est extrait de la distribution Poisson-New XLindley avec une fonction de densité de probabilité dans 5.1, où n représente la taille de l'échantillon.

Pour déterminer la valeur la plus probable du paramètre γ, il convient de maximiser la fonction de vraisemblance L de la manière suivante :

La fonction de vraisemblance est :

$$\begin{aligned} L(\theta, l_i) &= \prod_{i=1}^n f(l_i, \theta_i) \\ &= \prod_{i=1}^n \frac{\theta(\theta l_i + 2\theta + 1)}{2(1+\theta)^{l_i+2}} \\ L(\theta, l_i) &= \left(\frac{\theta}{2}\right)^n \prod_{i=1}^n \frac{(\theta l_i + 2\theta + 1)}{(1+\theta)^{l_i+2}}. \end{aligned}$$

$\ln L(\theta, l_i)$ ci-dessous représente la fonction de log-vraisemblance :

$$\begin{aligned}\ln L(\theta, l_i) &= n(ln\theta - ln2) + \sum_{i=1}^{n}[ln(\theta l_i + 2\theta + 1) - (l_i + 2)ln(1+\theta)], \\ &= nln\theta - 2nln(1+\theta) - ln(1+\theta)\sum_{i=1}^{n} l_i + \sum_{i=1}^{n} ln(\theta l_i + 2\theta + 1).\end{aligned}$$

Ainsi, l'estimateur $\hat{\theta}$ du paramètre θ est :

$$\begin{aligned}\frac{\partial \ln L(\theta; l_i)}{\partial \theta} &= 0, \\ \frac{n}{\theta} + \sum_{i=1}^{n} \frac{l_i + 2}{\theta l_i + 2\theta + 1} - \frac{1}{1+\theta}\sum_{i=1}^{n}(l_i + 2) &= 0.\end{aligned}$$

L'estimateur n'est pas exprimé de manière explicite, des méthodes numériques peuvent être employées pour le calculer.

5.1.3 Méthode bayésienne

Soit $l_1, l_2, ..., l_t$ un échantillon de taille t. Le nombre total des sinistres déclarés par un assuré sur t années est $n = \sum_{i=1}^{t} l_i$. l_i est le nombre de sinistres qu'un assuré à déclarer en i année, avec $i = 1, 2, ..., t$.

La fonction de vraisemblance est :

$$L(\theta | \, l_1, \ldots, l_t) = \prod_{i=1}^{t} \frac{e^{-\theta}\theta^{l_i}}{l_i!} \propto e^{-t\theta}\theta^{N}$$

La distribution a priori est :

$$\pi(\theta) = \frac{\theta}{2}(1+\theta\gamma)e^{-\theta\gamma} \propto \theta(1+\theta\gamma)e^{-\theta\gamma}.$$

Le théorème de Bayes est appliqué pour obtenir la fonction de distribution a posteriori comme suit :

$$\pi^*(\theta | \, l_1, \ldots, l_t) \propto \; f(l_1, \ldots, l_t | \theta)\, \pi(\theta) \propto \theta^{n+1}(1+\theta\gamma)e^{-\theta(t+\gamma)}$$

On considère que

$$\int_0^{\infty} \pi^*(\theta | \, l_1, \ldots, l_t)\, d\theta = \int_0^{\infty} A\theta^{n+1}(1+\theta\gamma)e^{-\theta(t+\gamma)} d\theta = 1,$$

où A est une constante obtenue comme suit :

$$\begin{aligned} A\left[\frac{\Gamma(n+2)}{(t+\gamma)^{n+2}}+\frac{\gamma\Gamma(n+3)}{(t+\gamma)^{n+3}}\right] &= 1, \\ A\left[\frac{\Gamma(n+2)(t+3\gamma+n\gamma)}{(t+\gamma)^{n+3}}\right] &= 1. \end{aligned}$$

Alors

$$A=\frac{(t+\gamma)^{n+3}}{\Gamma(n+2)(t+3\gamma+n\gamma)}.$$

Par conséquent, la fonction de distribution a posteriori pour la fréquence des sinistres est donnée par :

$$\pi^*(\theta|\ l_1,\ldots,l_t)=\frac{(t+\gamma)^{n+3}}{\Gamma(n+2)(t+3\gamma+n\gamma)}\ \theta^{n+1}(1+\theta\gamma)e^{-\theta(t+\gamma)}. \tag{5.2}$$

5.1.4 Calcul de prime

Notre objectif est de définir la prime nette. Considérons que l'historique des réclamations est $l_1, l_2, ..., l_t$. On obtient la moyenne de la fonction de distribution a posteriori pour la distribution Poisson-New XLindley par :

$$\begin{aligned} \hat{\theta}_{t+1} &= E[\theta|\ l_1,\ldots,l_t]=E[l_1,\ldots,l_t|\theta]=\int_0^\infty \theta\pi^*(\theta|l_1,\ldots,l_t)\,d\theta \\ &= \frac{(t+\gamma)^{n+3}}{\Gamma(n+2)(t+3\gamma+n\gamma)}\left[\frac{\Gamma(n+2)}{(t+\gamma)^{n+2}}+\frac{\gamma\Gamma(n+3)}{(t+\gamma)^{n+3}}\right] \\ \hat{\theta}_{t+1} &= \frac{(t+\gamma)(n+2)+\gamma(n+2)(n+3)}{(t+\gamma)(t+n\gamma+3\gamma)}. \end{aligned}$$

Prenons en compte que 100 représente la prime initiale à l'instant t = 0. Par conséquent, la prime au moment t + 1, basée uniquement sur la fréquence des sinistres, est définie comme suit :

$$\begin{aligned} Premium_{t+1} &= 100\frac{\hat{\theta}_{t+1}}{E[x]} \\ &= 100\frac{2\theta}{3}\frac{(t+\gamma)(n+2)+\gamma(n+2)(n+3)}{(t+\gamma)(t+n\gamma+3\gamma)}. \end{aligned} \tag{5.3}$$

5.2 Le montant des sinistres en utilisant la distribution Inverse Gamma (2, λ)-Lindley

Dans cette partie, nous allons présenter une distribution à un seul paramètre pour l'appliquer au calcul de la partie du montant des sinistres dans les primes du système bonus-malus. Il s'agit de la distribution Inverse-Gamma(2, λ) Lindley.

5.2.1 Distribution Inverse Gamma (2, λ)-Lindley

X est une variable aléatoire qui symbolise le montant des réclamations de chaque assuré. On suppose que X suit une distribution Inverse Gamma (2, λ), avec une fonction de densité de probabilité donnée par :

$$f(x|\lambda) = \lambda^2 x^{-3} e^{-\frac{\lambda}{x}}.$$

L'espérance de la variable aléatoire Inverse Gamma (2, λ) est :

$$E[X] = \lambda.$$

Considérons que la variable aléatoire λ suit une loi de Lindley de paramètre β. Ainsi, la fonction de densité pour λ est illustrée ci-dessous :

$$\pi(\lambda) = \frac{\beta^2}{\beta+1}(\lambda+1)e^{-\beta\lambda}.$$

Donc, nous obtenons une distribution Inverse Gamma (2, λ) mélangée à la distribution Lindley de la manière suivante :

$$\begin{aligned} f(x) &= \int_0^\infty \frac{\lambda^2}{\Gamma(\alpha)} x^{-3} e^{-\frac{\lambda}{x}} \frac{\beta^2}{\beta+1}(\lambda+1)e^{-\beta\lambda} d\lambda \qquad (5.4) \\ &= \frac{3x^{-3}\beta^2}{(\beta+1)\left(\beta+\frac{1}{x}\right)^4}\left[3+\beta+\frac{1}{x}\right]. \end{aligned}$$

5.2.2 Estimateur du Maximum de Vraisemblance

Soit $X = (x_1, x_2, \ldots, x_n)^T$ un vecteur d'observations indépendantes et identiquement distribuées selon la loi Inverse-Gamma(2, λ) Lindley, dont la fonction de densité est définie par l'équation (5.4). La valeur du paramètre β qui maximise la fonction de vraisemblance, c'est-à-dire l'estimateur du maximum de vraisemblance, est alors déterminée de la manière suivante :
La fonction de vraisemblance est :

$$\begin{aligned}
L(\beta, x_i) &= \prod_{i=1}^{n} f(x_i, \beta_i) \\
&= \prod_{i=1}^{n} \frac{3x_i^{-3}\beta^2}{(\beta+1)\left(\beta+\frac{1}{x_i}\right)^4}\left[3+\beta+\frac{1}{x_i}\right] \\
L(\beta, x_i) &= \frac{\beta^{2n}3^n}{(\beta+1)^n}\prod_{i=1}^{n} \frac{x_i^{-3}}{\left(\beta+\frac{1}{x_i}\right)^4}\left[3+\beta+\frac{1}{x_i}\right].
\end{aligned}$$

La fonction de log-vraisemblance est déterminée comme suit :

$$\begin{aligned}
\ln L(\beta, x_i) &= 2n\ln\beta + n\ln(3) - n\ln(\beta+1) - 3\sum_{i=1}^{n}\ln x_i \\
&\quad -4\sum_{i=1}^{n}\ln\left(\beta+\frac{1}{x_i}\right) + \sum_{i=1}^{n}\ln\left(3+\beta+\frac{1}{x_i}\right) \\
&= n\left[-\ln(\beta+1) + 2\ln\beta + \ln(3)\right] - 3\sum_{i=1}^{n}\ln x_i \\
&\quad -4\sum_{i=1}^{n}\ln\left(\beta+\frac{1}{x_i}\right) + \sum_{i=1}^{n}\ln\left(3+\beta+\frac{1}{x_i}\right).
\end{aligned}$$

Ainsi, l'estimateur du paramètre β, noté $\hat{\beta}$, s'exprime de la manière suivante :

$$\begin{aligned}
\frac{\partial \ln L(\beta; x_i)}{\partial\beta} &= 0, \\
\frac{2n}{\beta} - \frac{n}{\beta+1} - 4\sum_{i=1}^{n}\left(\frac{1}{\beta+\frac{1}{x_i}}\right) + \sum_{i=1}^{n}\frac{1}{3+\beta+\frac{1}{x_i}} &= 0.
\end{aligned}$$

L'estimateur ne possède pas une forme explicite ; il peut être déterminé à l'aide de méthodes numériques adaptées.

5.2.3 Méthode bayésienne

Soit X le montant des sinistres pour $l = 1, 2, ..., n$. La fonction de vraisemblance est :

$$\begin{aligned} L(\lambda |\, x_{1,...,}x_n) &= f(x_{1,...,}x_n|\lambda) = \prod_{i=1}^{n} \lambda^2 x_i^{-3} e^{-\lambda \frac{1}{x_i}} \\ &= \lambda^{2n} \prod_{i=1}^{n} x_i^{-3} e^{-\lambda \sum_{i=1}^{n} \frac{1}{x_i}} \\ &\propto \lambda^{2n} e^{-\lambda \sum_{i=1}^{n} \frac{1}{x_i}} \end{aligned}$$

La distribution a priori est :

$$\pi(\lambda) \propto (\lambda + 1) e^{-\beta\lambda}.$$

On applique le théorème de Bayes pour obtenir la fonction de distribution a posteriori comme suit :

$$\begin{aligned} \pi^*(\lambda |\, x_{1,...,}x_n) &\propto f(x_{1,...,}x_n|\lambda)\, \pi(\lambda) \\ &\propto \lambda^{2n} (1 + \lambda) e^{-\lambda\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)}. \end{aligned}$$

On considère que

$$\int_0^\infty \pi^*(\lambda |\, x_{1,...,}x_n)\, d\lambda \propto \int_0^\infty \lambda^{2n} (1 + \lambda) e^{-\lambda\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)} d\lambda.$$

Alors

$$\int_0^\infty \pi^*(\lambda |\, x_{1,...,}x_n)\, d\lambda = \int_0^\infty B\lambda^{2n} (1 + \lambda) e^{-\lambda\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)} d\lambda = 1,$$

où B est une constante obtenue comme suit :

$$\begin{aligned} B\left[\frac{\Gamma(2n+2)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{2\ n+2}} + \frac{\Gamma(2n+1)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{2n+1}}\right] &= 1, \\ B\left[\frac{\Gamma(2n+1)\left(2n + 1 + \gamma + \sum_{i=1}^{n} \frac{1}{x_i}\right)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{2n+2}}\right] &= 1 \quad . \end{aligned}$$

Alors

$$B = \frac{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{2n+2}}{\Gamma(2n+1)\left(2n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}.$$

Donc, la fonction de distribution a posteriori pour le montant des sinistres est donnée par :

$$\pi^*(\lambda|\ x_{1,\dots,}x_n) = \frac{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{2n+2}}{\Gamma(2n+1)\left(2n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)} \lambda^{2n}(1+\lambda)\, e^{-\lambda\left(\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}. \tag{5.5}$$

5.2.4 Calcul de prime

Dans cette partie, tout comme pour la fréquence des sinistres, nous ferons usage de la prime de base.

L'espérance de la formule précédente (5.5) pour la distribution Inverse Gamma (2, λ) Lindley s'exprime comme suit :

$$\begin{aligned}
\widehat{\lambda_{t+1}} &= E[\lambda|\ x_1,\dots,x_n] = \int_0^\infty \lambda\pi^*(\lambda|\ x_1,\dots,x_n)\, d\lambda \\
&= \frac{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{2n+2}}{\Gamma(2n+1)\left(2n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)} \\
&\quad \times \left[\frac{\Gamma(2n+2)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{2n+2}} + \frac{\Gamma(2n+3)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{2n+3}}\right] \\
\widehat{\lambda_{t+1}} &= \frac{(2n+1)\left(2n+2+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)\left(2n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}.
\end{aligned}$$

On a

$$E[\lambda|\ x_1,\dots,x_n] = E[x_1,\dots,x_n|\lambda] = \hat{\lambda}.$$

Par conséquent

$$E[x_1,\dots,x_n|\lambda] = \frac{(2n+1)\left(2n+2+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}{\left(\beta + \sum_{i=1}^{n} \frac{1}{x_i}\right)\left(2n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}.$$

La prime finale basée sur la distribution Poisson-New XLindley pour la fréquence des sinistres et la distribution Inverse Gamma(2, λ) Lindley pour le montant des sinistres que l'assuré doit payer est :

$$\begin{aligned} Premium_{t+1} &= \frac{(t+\gamma)(N+2)+\gamma(N+2)(N+3)}{(t+\gamma)(t+N\gamma+3\gamma)} \\ &\times \frac{(2n+1)\left(2n+2+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}{\left(\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)\left(2n+1+\beta+\sum_{i=1}^{n}\frac{1}{x_i}\right)}. \end{aligned} \quad (5.6)$$

5.3 Application numérique

Pour l'application numérique, nous allons faire usage des mêmes données réelles employées dans le chapitre précédent (voir de Jong et Heller [9]).

5.3.1 Les primes du système bonus-malus basées uniquement sur la fréquence des sinistres

Pour calculer les primes en ce basant uniquement sur de fréquence des sinistres, nous utilisons la distribution Poisson-New XLindley qui a été présentée précédemment dans ce travail. Dans l'équation (5.1), l'estimateur du maximum de vraisemblance du paramètre θ de la distribution Poisson-New XLindley est représenté par $\hat{\theta}$=14.2.

Le tableau ci-dessous présente les primes basées uniquement sur la fréquence des sinistres, effectué à l'aide de la formule (5.3).

t	nombre de sinistres				
	0	1	2	3	4
0	100				
1	92.36186	165.1296	232.7445	298.2554	362.7020
2	85.73208	154.0829	217.6451	279.2130	339.7610
3	79.92890	144.3560	204.3245	262.4010	319.4994
4	74.81120	135.7292	192.4884	247.4510	301.4751
5	70.26780	128.0287	181.9040	234.0716	285.3381
6	66.20999	121.1153	172.3844	222.0291	270.8080
7	62.56618	114.8762	163.7783	211.1337	257.6570

Tableau 5.1 – primes en utilisant la distribution Poisson-New XLindley.

D'après le tableau 5.1, les primes augmentent avec le nombre de sinistres déclarés. À mesure que t augmente, les primes diminuent pour un même nombre de sinistre, suggérant une amélioration des coûts pour les assurés qui conservent un bon historique. Selon les résultats de ce tableau, un assuré qui a eu un sinistre la première année devra payer un malus égal à 65.13 % de la prime de base. Un bonus de 7.64 % de la prime de base sera ajouté à la prime pour la première année sans sinistre.

5.3.2 Les primes en utilisant la distribution Inverse-Gamma(2,λ) Lindley pour le montant des sinistres

Dans cette partie, nous présentons les primes qui prennent en compte la fréquence et le montant des sinistres. La fréquence de sinistre est représentée par la distribution Poisson-New XLindley, tandis que le montant des sinistres est calculé en utilisant la distribution Inverse-Gamma (2, λ) Lindley qui a été présentée précédemment. Le paramètre de la distribution Inverse-Gamma(2, λ) Lindley dans (5.4) est : $\hat{\beta} = 0.002$. Les primes basées sur la fréquence et le montant des sinistres en utilisant la formule (5.6) sont présentées dans le tableau suivant :

t	nombre de sinistres				
	0	1	2	3	4
0	519.3219				
1	482.3742	712.6217	1064.084	1493.103	1992.325
2	450.1049	666.3900	996.1464	1398.664	1867.058
3	421.6908	625.6048	936.1662	1315.255	1756.399
4	396.4903	589.3649	882.8293	1241.056	1657.983
5	373.9957	556.9573	835.0957	1174.626	1569.770
6	353.8004	527.8101	792.1315	1114.811	1490.363
7	335.5749	501.4597	753.2602	1060.674	1418.479

TABLEAU 5.2 – primes en utilisant les distributions Poisson-New XLindley et Inverse-Gamma (2, λ) Lindley.

Comme on peut le voir, la prime initiale est de 519.3219. Si un assuré a une année sans sinistre, la prime sera de 482.3742 pour l'année suivante. Mais si cet assuré a un sinistre au cours de la première année, la prime qui devra être payée pour l'année suivante sera de 712.6217. Pour un assuré qui a eu quatre sinistres dans la même année, la prime sera de 1992.325.
On remarque également que les primes augmentent rapidement lorsque le nombre de sinistres augmente. Par contre, les primes diminuent lentement avec les années où il n'y a pas de sinistre.

5.4 Comparaison des primes

Dans la présente section, nous proposons d'effectuer une analyse comparative entre les primes calculées dans le chapitre 4 précédent et celles déterminées dans ce chapitre. Dans un premier temps, nous réaliserons une comparaison exclusivement axée sur la fréquence des sinistres. Par la suite, nous adopterons une approche combinée intégrant à la fois le montant et la fréquence des sinistres.

L'analyse débutera par l'examen des primes fondées uniquement sur le nombre de sinistres. Nous confronterons ensuite les primes obtenues en modélisant le montant des sinistres à l'aide de la distribution log-Normal Gamma associée à diverses distributions pour la fréquence, aux primes calculées en recourant à la distribution Poisson-New XLindley pour la fréquence des sinistres et à la distribution Inverse-Gamma(2, λ) Lindley pour le montant des sinistres. Enfin, nous comparerons les résultats de ce chapitre aux primes issues de l'utilisation de la distribution Inverse-Gamma Lindley pour modéliser le montant des sinistres, en combinaison avec différents modèles de distribution pour le nombre de sinistres.

5.4.1 Les primes basées uniquement sur la fréquence des sinistres

Nos résultats montrent que, dans le cadre d'une sous-estimation (avec $a = 1.1$), l'application de la distribution Poisson-Akash sous la fonction de perte linex conduit à l'attribution des bonus les plus élevés en l'absence de sinistres durant la première année. D'autre part, il semble que les primes obtenues par la distribution Poisson-New XLindley se distinguent par leur plus grande souplesse pour les conducteurs jugés à risque, ce qui entraîne l'attribution des malus les plus bas. D'où la possibilité pour l'assureur de sélectionner le modèle le plus approprié selon les exigences particulières de sa société d'assurance.

5.4.2 Les primes basées sur la fréquence et le montant des sinistres

Nous avons évalué les primes en combinant la fréquence et le montant des sinistres dans le cadre de cette recherche. Les primes ont d'abord été calculées en liant la distribution Poisson-Akash pour la modélisation du nombre de sinistres à la distribution Log-Normal Gamma pour le montant des sinistres. Nous avons ensuite fait une nouvelle estimation en associant la distribution Poisson-Akash à la distribution Inverse-Gamma Lindley pour le montant des sinistres. Enfin, nous avons estimé les primes grâce à la distribution Poisson-New XLindley pour le nombre de sinistres, liée à la distribution à un paramètre Inverse-Gamma(2, λ) Lindley pour le montant des sinistres. La suite de cette section est consacrée à une discussion des résultats obtenus.

5.4.2.1 Les primes en utilisant la distribution Log-Normal Gamma

Les résultats obtenus avec la distribution Poisson-Akash sous la fonction de perte linex en contexte de sous-estimation indiquent que la méthodologie appliquée conduit à une tarification particulièrement rigoureuse des bonus, se traduisant par des primes élevées pour les conducteurs jugés « bons » et par l'imposition de malus très sévères pour les conducteurs à risque. En revanche, la distribution Poisson-New XLindley produit des primes plus souples, tant pour les bons que pour les mauvais conducteurs. La diversité des modèles de prime suggérés dans cette étude permet d'affirmer que l'assureur peut adapter la tarification de sa clientèle en fonction des objectifs et du positionnement stratégique de sa compagnie.

5.4.2.2 Les primes en utilisant la distribution Inverse-Gamma Lindley

En examinant les tableaux 4.10, 4.11, 4.12, 4.13 et 5.2, on constate que les primes établies sur la base de la distribution Poisson-Akash associée à la fonction de perte linex en sous-estimation pour la fréquence des sinistres, conjuguée à la distribution Inverse-Gamma Lindley pour le montant des sinistres (tableau 4.12), se révèlent être les plus sévères. Ce modèle impose des pénalités particulièrement accroissantes pour les mauvais conducteurs, tout en offrant des réductions limitées pour les bons conducteurs. À l'inverse, les primes résultant du procédé de surestimation (tableau 4.11) se montrent nettement plus flexibles, en octroyant des réductions importantes aux bons conducteurs. Par ailleurs, l'analyse des primes obtenues en utilisant la distribution Poisson-New XLindley pour la fréquence des sinistres conjointement à la distribution Inverse-Gamma $(2, \lambda)$ Lindley pour le montant des sinistres (tableau 5.2) indique que ce modèle offre une meilleure souplesse pour les mauvais conducteurs, se traduisant par des pénalités moins marquées.

Conclusion

Dans cette thèse, nous avons proposé un nouveau modèle pour calculer les primes du système bonus-malus. Ce modèle est basé à la fois sur la fréquence et le montant des sinistres. Nous avons opté pour l'utilisation de la distribution mixte Poisson-Akash pour modéliser la fréquence des sinistres, tandis que la nouvelle distribution mixte Inverse-Gamma Lindley a été choisie pour représenter le montant des sinistres. Les primes ont été déterminées en appliquant la méthode bayésienne. De plus, nous avons présenté la fonction de perte linex dans ce travail et calculé les primes du système bonus-malus sous cette fonction de perte pour la fréquence et le montant des sinistres.

Par ailleurs, nous avons déterminé les primes en utilisant la méthode bayésienne et en se basant sur la distribution Poisson-New XLindley pour la fréquence des sinistres ainsi que la distribution à un paramètre Inverse-Gammal(2, λ) Lindley pour le montant des sinistres. Une application numérique utilisant des données réelles de l'assurance automobile a été utilisée pour illustrer notre modèle.

D'après les résultats de l'application numérique, nous avons conclu qu'en utilisant la distribution Poisson-Akash sous la fonction de perte linex, l'assureur a la possibilité d'adapter les primes en fonction des exigences de l'entreprise en modulant la valeur de a.

La diversité des résultats entre souple stricte et moyenne offre un large choix des modèles et stratégie pour les assureurs afin d'adapter leurs choix avec les besoins de la compagnie d'assurance pour garder la solvabilité de la compagnie d'un côté et pour donnez des primes attirante pour satisfaire les assurés d'un autre côté.
Comme perspectives, nous souhaitons déterminer les primes du système bonus-malus en utilisant de nouvelles distributions ainsi que d'autres fonctions de perte.

Bibliographie

[1] Basu, A.P. and Ebrahimi, N. Bayesian approach to life testing and reliability estimation using asymmetric loss function. *Journal of statistical planning and inference*, 29(1-2) :21–31, 1991.

[2] Bayes, T. LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S. *Philosophical transactions of the Royal Society of London*, (53) :370–418, 1763.

[3] Bühlmann, H. Experience rating and credibility. *ASTIN Bulletin : The Journal of the IAA*, 4(3) :199–207, 1967.

[4] Bühlmann, H. Experience rating and credibility. *ASTIN Bulletin : The Journal of the IAA*, 5(2) :157–165, 1969.

[5] Bühlmann, H. and Gisler, A. *A course in credibility theory and its applications*, volume 317. Springer, 2005.

[6] Calabria, R. and Pulcini, Gs. An engineering approach to bayes estimation for the weibull distribution. *Microelectronics Reliability*, 34(5) :789–802, 1994.

[7] D. Cox and D. Hinkley. Theoretical statistics. *Chapman and Hall*, 1974.

[8] M. David. Auto insurance premium calculation using generalized linear models. *Procedia Economics and Finance*, 20 :147–156, 2015.

[9] P. De Jong and G. Heller. Generalized linear models for insurance data. *Cambridge Books*, 2008.

[10] M. Denuit, S. Pitrebois, and J. Walhin. Marketing et systèmes bonus-malus. *Actu-L*, 3 :89–105, 2003.

[11] R. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, 222(594-604) : 309–368, 1922.

[12] R. Fisher. Theory of statistical estimation. In *Mathematical proceedings of the Cambridge philosophical society*, volume 22, pages 700–725. Cambridge University Press, 1925.

[13] N. Frangos and S. Vrontos. Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance. *ASTIN Bulletin : The Journal of the IAA*, 31(1) : 1–22, 2001.

[14] C. Gauss. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium.* Hamburg : Friedrich Perthes and I. H. Besser, 1809. Translated as *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*, London : Little, Brown, and Company, 1857.

[15] E. Gómez-Déniz and F. J. Vázquez-Polo. Exact credibility reference

bayesian premiums. *Insurance : Mathematics and Economics*, 105 : 128–143, 2022.

[16] J. Holtan. Optimal loss financing under bonus-malus contracts. *ASTIN Bulletin : The Journal of the IAA*, 31(1) :161–173, 2001.

[17] O. Imene, S. Ahmed, M. Farouk, and R. M. Riad. On bayesian bonus-malus premium under linex loss function with applications. *Studies in Engineering and Exact Sciences*, 5(2) :e6226, 2024.

[18] S. Kafková, L. Křivánková, et al. Generalized linear models in vehicle insurance. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 62(2) :383–388, 2014.

[19] H. Kalouti. Pure premium estimation based on generalized linear models. Master's thesis, Marmara Universitesi (Turkey), 2021.

[20] S. Kermoun. *Inférence bayesiènne dans des modèles d'assurance*. PhD thesis, Université Badji Mokhtar Annaba, 2022.

[21] N. Khodja, A. M. Gemeay, H. Zeghdoudi, K. Karakaya, A. M. Alshangiti, M. E. Bakr, O. S. Balogun, A. H. Muse, and E. Hussam. Modeling voltage real data set by a new version of lindley distribution. *IEEE Access*, 11 :67220–67229, 2023.

[22] J. Krahnen and M. Weber. Generally accepted rating principles : A primer. *Journal of Banking & Finance*, 25(1) :3–23, 2001.

[23] P. Laplace. Sur les approximations des formules qui sont fonctions de tres grands nombres et sur leur application aux probabilites. *Œuvres complètes*, 12 :301–345, 1810.

[24] A. Legendre. *Mémoire sur les opérations trigonométriques : dont les résultats dépendent de la figure de la terre*. Number 1. F. Didot, 1805.

[25] J. Lemaire. Selection procedures of regression analysis applied to automobile insurance. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*.

[26] J. Lemaire. *Bonus-malus systems in automobile insurance*, volume 19. Springer science & business media, 1995.

[27] J. Lemaire. Bonus-malus systems : In encyclopedia of actuarial science, 2004.

[28] J. Lemaire and H. Zi. A comparative analysis of 30 bonus-malus systems. *ASTIN Bulletin : The Journal of the IAA*, 24(2) :287–309, 1994.

[29] M. Mert and Y. Saykan. On a bonus-malus system where the claim frequency distribution is geometric and the claim severity distribution is pareto. *Hacettepe Journal of Mathematics and Statistics*, 34(1) :75–81, 2005.

[30] F. Metiri, H. Zeghdoudi, and M. Remita. On bayes estimates of lindley distribution under linux loss function : informative and non informative priors. *Global journal of Putre and Applied Mathematics*, 12 :391–400, 2016.

[31] A. Moumeesri, W. Klongdee, and T. Pongsart. Bayesian bonus-malus premium with poisson-lindley distributed claim frequency and lognormal-gamma distributed claim severity in automobile insurance. *WSEAS Transactions on Mathematics*, 19 :443–451, 2020.

[32] M. M. Nassar and F. H. Eissa. Bayesian estimation for the exponentiated weibull model. *Communications in Statistics-Theory and Methods*, 33 (10) :2343–2362, 2004.

[33] Z. Neche, C. Benatmane, A. Djebar, and H. Zeghdoudi. Around poisson new xlindley process : Comparison and application in actuarial science. *Journal of Computational Analysis and Applications*, 34(2), 2025.

[34] J. A. Nelder and R. W. Wedderburn. Generalized linear models. *Journal of the Royal Statistical Society Series A : Statistics in Society*, 135(3) : 370–384, 1972.

[35] R. Norberg. A credibility theory for automobile bonus systems. *Scandinavian Actuarial Journal*, 1976(2) :92–107, 1976.

[36] R. Norberg. Credibility theory. *Encyclopedia of Actuarial Science*, 1 : 398–406, 2004.

[37] I. Ouchen, M. R. Remita, et al. Bayesian bonus-malus premium of novel one parameter model for the severity component case. *Studies in Science of Science— ISSN : 1003-2053*, 43(4) :183–191, 2025.

[38] K. P. *The Scope and Method of Science [1892]*, pages 66–78. Harvard University Press, Cambridge, MA and London, England, 1949.

[39] E. Parent and J. Bernier. *Le raisonnement bayésien : modélisation et inférence*. Springer Science & Business Media, 2007.

[40] T. Rahmawati, D. Susanti, and R. Riaman. Determining pure premium of motor vehicle insurance with generalized linear models (glm). *International Journal of Quantitative Research and Modeling*, 4(4) :207–214, 2023.

[41] C. Robert. *Le choix bayésien : Principes et pratique.* Springer Science & Business Media, 2005.

[42] J. Rojo. On the admissibility of c $\overline{X}$ + d with respect to the linex loss function. *Communications in Statistics-Theory and Methods*, (16) : 3745–3748, 1989.

[43] T. Rolski, H. Schmidli, V. Schmidt, and J. Teugels. *Stochastic processes for insurance and finance.* John Wiley & Sons, 2009.

[44] A. Saadoun, H. Zeghdoudi, F. Attoui, and M. Remitra. On bayesian premium estimators for gamma lindley model under squared error loss function and linex loss function. *Journal of Mathematics and Statistics*, 284 :291, 2017.

[45] SAAQ. Société de l'assurance automobile du québec, les points d'inaptitude. *www.saaq.gouv.qc/ca/documents/pdf/permis/points.inaptitude.php.*

[46] F. Seghier, M. Ahsan-ul Haq, H. Zeghdoudi, and S. Hashmi. A new generalization of poisson distribution for over-dispersed, count data : mathematical properties, regression model and applications. *Lobachevskii Journal of Mathematics*, 44(9) :3850–3859, 2023.

[47] R. Shanker. On poisson-akash distribution and its applications. *Biometrics & Biostatistics International Journal*, 3(6), 2016.

[48] L. Tremblay. Using the poisson inverse gaussian in bonus-malus systems. *ASTIN Bulletin : The Journal of the IAA*, 22(1) :97–106, 1992.

[49] H. Varian. A bayesian approach to real estate assessment. *North-Holland Pub. Co.*, pages 195–208, 1975.

[50] A. Wald. Statistical decision functions. *The Annals of Mathematical Statistics*, pages 165–205, 1949.

[51] J. Walhin and J. Paris. Using mixed poisson processes in connection with bonus-malus systems1. *ASTIN Bulletin : The Journal of the IAA*, 29(1) :81–99, 1999.

[52] A. Whitney. The theory of experience rating. *Proceedings of the Casualty Actuarial Society*, 4 :275–293, 1918.

[53] Y. Yan and K.-S. Song. A general optimal approach to bühlmann credibility theory. *Insurance : Mathematics and Economics*, 104 :262–282, 2022.

[54] A. Zellner. Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, 81(394) : 446–451, 1986.