

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

**L'apport de la théorie de la crédibilité à la tarification dans
les compagnies d'assurance**

Filière

Mathématiques Appliquées

Spécialité

Actuariat

Par

ARAFA Ghada

DIRECTEUR DE THÈSE:

METIRI Farouk

MCA. U.B.M. Annaba

CO-DIRECTEUR DE THÈSE : REMITA Mohamed Riad

Prof. ENSIA. Alger

Devant le jury

PRESIDENT:

BOURAS Mohammed Cherif

Prof. U.B.M. Annaba

EXAMINATEUR : EZZEBSA Abdelali

MCA. U.8 mai 1945. Guelma

EXAMINATEUR : DJEBAR Ahlem

MCA. U.B.M. Annaba

EXAMINATEUR: BELHAMRA Thara

MCA. U.B.M. Annaba

TABLE DES MATIÈRES

1	Remerciement	3
2	Résumé	4
3	Abstract	5
4	Introduction générale	7
1	Généralités sur l'assurance et la théorie de la crédibilité	9
1.1	Définition et Types de l'assurance	9
1.2	La compagnie d'assurance	11
1.3	La tarification en assurances	11
1.4	La théorie de la crédibilité	12
1.4.1	Crédibilité totale	16
1.4.2	Crédibilité partielle	17
1.4.3	Prime individuelle et prime collective	18
1.4.4	Prime bayésienne	19
1.4.5	Le modèle de crédibilité de Bühlmann	20
1.4.6	Le Modèle de crédibilité de Bühlmann-Straub	27
1.5	Conclusion	31
2	La modélisation de la dépendance entre les sinistres d'assurance	33
2.1	Les mesures de dépendance	34

2.2	Modèle à effets communs à deux niveaux de la dépendance aux sinistres	36
2.2.1	Hypothèses du modèle	37
2.2.2	Application d'effets communs pour déterminer les primes de crédibilité	41
2.2.3	Modèle avec dépendance simultanée	48
2.3	Dépendance et copules	51
2.3.1	Définition d'une copule	52
2.3.2	Propriétés des copules	55
2.3.3	Théorème de sklar	56
2.3.4	Densité d'une copule	58
2.3.5	Les copules paramétriques	62
2.3.6	La dépendance de queue	67
2.4	Conclusion	69
3	Sur l'estimateur de la prime de crédibilité des quantiles avec une corrélation identique entre les risques	71
3.1	Formulation du modèle et Préliminaires	72
3.1.1	Définition des quantiles	72
3.1.2	Hypothèses	73
3.2	La dérivation de la prime de crédibilité des quantiles non homogène	82
3.3	La dérivation de la prime de crédibilité des quantiles homogène	85
3.4	Application numérique	90
3.5	Perspectives	94
	Conclusion générale	95
	Annexes	96
	Bibliographie	102

1 Remerciement

Avant tout propos, je voudrais exprimer ma profonde gratitude envers Dieu Tout-Puissant, qui m'a accordé la santé, la patience et la persévérance pour mener à bien ce travail.

Cette thèse est l'aboutissement d'un parcours de longue durée, rendu possible grâce au soutien précieux de nombreuses personnes, que je tiens à remercier sincèrement en clôturant ce chapitre important de ma vie.

Je tiens à remercier chaleureusement mon encadrant Dr METIRI Farouk, pour ses conseils précieux et sa bienveillance tout au long de l'élaboration de cette thèse. Son expertise et son soutien ont été essentiels pour surmonter les défis rencontrés.

Je souhaite également adresser mes remerciements au Pr REMITA Mohamed Riad pour l'intérêt qu'il a porté au déroulement de ce travail académique, ses observations constructives et son expertise ont été d'une grande valeur.

Mes sincères remerciements vont aux membres du jury, Pr BOURAS Mohammed Cherif, Dr DJEBAR Ahlem et Dr BELHAMRA Thara de l'université Badji Mokhtar Annaba et Dr EZZEBSA Abdelali de l'université 8 Mai 1945 Guelma, pour avoir accepté de lire et de juger ma thèse, leurs efforts et leurs observations pertinentes qui ont permis d'améliorer la qualité de ce manuscrit, leur expertise et leur engagement ont apporté une contribution précieuse à ce travail.

Je tiens également à remercier mes collègues de laboratoire, dont les échanges et discussions ont enrichi ma réflexion et contribué à instaurer un environnement propice à la recherche.

Enfin, je tiens à exprimer ma gratitude envers ma famille, en particulier mes parents dont le soutien infaillible, tant moral que matériel, a été une source de force et de motivation. Je remercie aussi mes amies pour leur encouragement et leur présence rassurante durant ces moments cruciaux.

2 Résumé

Dans cette thèse, nous présentons des détails additionnels concernant la théorie de la crédibilité et la modélisation de la dépendance des revendications en exposant un modèle à effets communs appelé à deux niveaux dans le cadre du calcul des primes de crédibilité, qui est vaguement basé sur le modèle bayésien. Ces effets communs sont ajoutés afin de prendre en considération les éventuelles dépendances qui peuvent se manifester chez les assurés et en même temps à travers les différentes périodes.

À cet égard, nous exploiterons un outil couramment utilisé dans d'autres secteurs de travail qui impliquent la dépendance afin de générer des primes de crédibilité. En particulier, nous employons les copules afin de représenter la dépendance dans le temps pour un risque individuel ou un groupe de risque homogène. Ensuite, avec l'utilisation des quantiles et la méthode de la projection orthogonal nous développons un modèle qui permettra d'identifier les estimateurs de crédibilité non homogènes et homogènes avec une structure de corrélation identique entre les risques d'assurance.

Mots-clés : risque individuel, hétérogénéité, dépendance des risques, effet commun, modèle bayésien, prime de crédibilité, copules, estimateurs de crédibilité non homogène et homogène, quantiles.

3 Abstract

In this thesis, we provide additional insights into credibility theory and claim dependence modeling by presenting a common effects model called two-level in the context of calculating credibility premiums, which is loosely based on the model Bayesian. These common effects are added in order to take into consideration possible dependencies that may appear among policyholders and at the same time across different periods.

In this regard, we will leverage a tool commonly used in other work sectors that involve dependency in order to generate credibility bonuses. In particular, we use copulas to represent dependence over time for an individual risk or a homogeneous risk group. Then, with the use of quantiles and the orthogonal projection method we develop a model which will identify non-homogeneous and homogeneous credibility estimators with an equal correlation structure between insurance risks.

Keywords : individual risk, heterogeneity, risk dependence, common effect, Bayesian model, credibility premium, copulas, homogeneous and inhomogeneous credibility estimators, quantiles.

ملخص

في هذه الأطروحة، نقدم معلومات إضافية حول نظرية المصادقية ونمذجة التبعية بين المطالبات من خلال عرض نموذج بتأثيرات مشتركة يسمى بالنموذج ذي المستويين في إطار حساب أقساط المصادقية، والذي يعتمد بشكل فضفاض على النموذج البايزي. يتم إضافة هذه التأثيرات المشتركة من أجل أخذ بعين الاعتبار التبعيات المحتملة التي قد تظهر لدى المستفيدين من التأمين، وأيضاً عبر الفترات الزمنية المختلفة.

في هذا السياق، سنستغل أداة تُستخدم بشكل شائع في قطاعات عمل أخرى تتضمن التبعية لتوليد أقساط المصادقية. على وجه الخصوص، نستخدم "Copules" لتمثيل التبعية عبر الزمن بالنسبة لخطر فردي أو لمجموعة من المخاطر المتجانسة. ثم، باستخدام quantiles وطريقة الإسقاط العمودي، نطور نموذجاً يسمح بتحديد مقدري المصادقية غير المتجانسين والمتجانسين مع هيكل ترابط متطابق بين مخاطر التأمين.

الكلمات الرئيسية: المخاطر الفردية، عدم التجانس، الاعتماد على المخاطر، التأثير المشترك، النموذج بايزي، علاوة المصادقية، Copules، مقدرات المصادقية غير المتجانسة والمتجانسة، القيم المئينية "Quantiles".

4 Introduction générale

L'assurance joue un rôle fondamental dans la gestion des risques, offrant une protection financière contre divers aléas de la vie quotidienne. Cependant, pour que les compagnies d'assurance puissent maintenir leur viabilité financière tout en offrant des primes compétitives, il est indispensable de posséder des procédures de tarification exactes et justes.

Pour déterminer les primes, l'analyse des réclamations antérieures donne une vue inestimable sur les traits inobservables des risques individuels. Par ailleurs, il est fréquent de considérer l'expérience des sinistres passés de l'assuré lorsqu'on prévoit les sinistres et calcule les primes. Cet exercice est connu sous le nom de tarification par expérience -*experience rating*- et est généralement effectué dans le but d'atteindre un taux de prime d'assurance juste et équitable.

La théorie de la crédibilité, développée au milieu du XXe siècle, s'est rapidement imposée comme un outil central dans ce domaine, permettant d'ajuster les primes et traiter la prévision des risques futurs en utilisant des données passées des assurés, tout en tenant compte de l'hétérogénéité des risques au sein des portefeuilles d'assurance, pour établir une tarification qui peut être exprimée comme un équilibre entre l'expérience individuelle d'un assuré et l'expérience collective du portefeuille.

Autrement dit, la théorie de la crédibilité constitue un ensemble d'approches employées par les actuaires afin d'évaluer la prime d'un contrat dans une variété de portefeuilles. Il s'agit d'un outil de tarification en assurance non vie, particulièrement répandu en Amérique du Nord. Elle repose sur le principe de limiter l'analyse aux estimateurs linéaires par rapport aux observations disponibles.

Le présent travail s'inscrit dans cette dynamique d'amélioration continue des techniques actuarielles, en s'intéressant à l'apport de la théorie de la crédibilité à la tarification dans les compagnies d'assurance. Cette thèse se propose d'explorer les fondements de cette théorie, ses développements historiques, ainsi que ses applications et ses extensions contemporaines.

Dans le premier chapitre, nous entreprendrons une rétrospective de la théorie de la crédibilité, en nous focalisant sur les modèles traditionnels de Bühlmann et

Bühlmann-Straub, qui ont instauré les fondements du recours à l'expérience passée pour la détermination des primes d'assurance. Ces modèles seront analysés non seulement dans leur cadre théorique, mais également à travers des applications pratiques illustrant leur utilité dans des contextes réels.

Pendant longtemps, l'existence de la dépendance entre les risques d'assurance était ignorée par les actuaires, pour qui l'indépendance entre les risques était une nécessité méthodologique. Cependant, pour mieux refléter la réalité, il est nécessaire de se concentrer sur des modèles de crédibilité incluant l'hypothèse de dépendance entre les risques. En réalité, les risques d'assurance et les sinistres peuvent être influencés par plusieurs facteurs qui ne sont pas indépendants les uns des autres, tels que les conditions météorologiques et économiques. Alors, les compagnies d'assurance peuvent mieux évaluer et gérer les risques en intégrant les dépendances dans leurs modèles, ce qui conduit à des décisions de tarification et de couverture plus précises.

Le deuxième chapitre élargit cette perspective en abordant les mesures de dépendance entre les risques d'assurance et les sinistres. Il est essentiel de considérer les dépendances dans un contexte où les risques ne sont pas indépendants, ce qui est souvent le cas en assurance dommages. Nous explorerons ici l'effets communs et les copules, des outils statistiques puissants qui permettent de modéliser ces dépendances de manière plus fine et de mieux comprendre les interactions entre les risques.

Enfin, dans le troisième chapitre, nous présenterons un nouveau modèle de crédibilité qui est l'estimateur de crédibilité de quantile avec une corrélation identique entre les risques. Ce modèle innovant est conçu pour répondre aux limitations des approches traditionnelles, en offrant une méthode plus robuste pour estimer les primes dans des situations où les distributions de sinistres sont asymétriques ou corrélées de manière complexe. Une application contenant des données réelles a été incorporée et une comparaison des résultats de notre modèle par rapport aux résultats de Wang et al.(2020) et Pitselis (2013) a été réalisée.

En somme, l'objectif de cette thèse est d'illustrer la valeur et la modernité de la théorie de la crédibilité dans le calcul des coûts d'assurance, tout en suggérant des avancées méthodologiques qui répondent aux défis posés par la complexité croissante des risques modernes.

CHAPITRE 1

Généralités sur l'assurance et la théorie de la crédibilité

La gestion du risque est au coeur de l'assurance, ce qui requiert des outils statistiques robustes afin d'évaluer de manière précise les primes d'assurance. La théorie de la crédibilité joue un rôle essentiel parmi ces outils, fournissant un cadre strict pour adapter les estimations fondées sur l'expérience passée tout en prenant en compte l'incertitude inhérente à toute prévision. Initialement conçue par les experts en assurance pour ajuster les primes en fonction des données historiques, la théorie de la crédibilité est désormais un fondement clé de la tarification en assurance.

1.1 Définition et Types de l'assurance

L'assurance est un système qui assure la protection d'une personne, d'une association ou d'une entreprise contre les conséquences financières et économiques liées à la survenue d'un risque spécifique. Les organismes d'assurance utilisent une méthode pour les protéger contre ce risque en les associant à une communauté de personnes (les assurés), qui s'engage à cotiser afin de pouvoir indemniser ceux parmi eux qui pourraient subir des dommages matériels ou corporels en cas de risque. L'assurance est donc un système de gestion des risques fondé sur la solidarité, puisque c'est l'ensemble de la communauté des assurés qui prend en charge matériellement les

dommages subis par ses membres lorsqu'un risque est réalisé.

Sous l'angle juridique, l'assurance représente un contrat, c'est à dire un accord légal entre l'assuré et l'assureur par lequel l'assuré se voit garantir et assurer, contre une prime, avec une garantie de l'assureur en cas de danger spécifique pour lui ou un tiers établi.

Il existe plusieurs types d'assurance, généralement répartis en deux catégories principales : les assurances de personnes et les assurances de dommages.

► **Assurances de personnes** : Ces assurances couvrent les risques liés à la vie et à la santé des individus, comme :

- Assurance décès : Similaire à l'assurance vie, mais elle ne doit être versée que si l'assuré décède pendant une période spécifique.

- Assurance vie : Dans le cas de la mort ou de la survie de l'assuré, elle permet le versement d'un capital ou une rente aux bénéficiaires désignés.

- Assurance maladie : Elle prend en charge les frais médicaux en cas de maladie ou d'accident.

- Assurance dépendance : Elle intervient en cas de perte d'autonomie, en aidant à financer les soins et services nécessaires.

- Assurance invalidité/incapacité : Elle couvre le risque de perte de revenu dû à une invalidité permanente ou temporaire.

► **Assurances de dommages** : concerne l'assurance des biens et l'assurance de responsabilités civiles, comme :

- Assurance automobile : Elle est obligatoire dans de nombreux pays, elle couvre les dommages causés à autrui en cas d'accident, ainsi que les dommages au véhicule.

- Assurance habitation : Couvre les dommages subis par un logement (vol, incendie, dégâts des eaux,...etc).

- Assurance professionnelle : Destinée aux entreprises et aux indépendants, elle couvre les risques liés à leur activité (responsabilité civile, pertes d'exploitation,...etc).

- Assurance responsabilité civile : Couvre les dommages que l'assuré pourrait causer à un tiers, que ce soit matériel ou corporel.

- Assurance multirisque : Elle combine plusieurs garanties pour couvrir divers types de risques (ex. assurance multirisque habitation ou entreprise).

1.2 La compagnie d'assurance

Une entreprise d'assurance offre des services d'assurance à des clients qui se transforment en assurés. Il s'agit d'une société anonyme (SA) à capital partagé. Son domaine d'activité est la création, la gestion et la vente de produits d'assurance destinés aux clients qui deviennent des assurés. Elle est une entité juridique dont les activités sont soumises aux règles du code des assurances. Une entreprise d'assurance offre des garanties financières en cas de risques variés (détérioration de bien, vol, décès...). C'est une entreprise de services qui offre ses services à toute personne, association ou entreprise qui désire prévoir la couverture d'un risque en versant régulièrement des cotisations. L'activité de l'assurance, similaire à celle des autres intervenants du marché, implique le transfert d'un risque entre un assuré et l'agence d'assurance.

Les contrats d'assurance sont distribués par plusieurs intermédiaires : des employés, un réseau d'agents généraux, qui sont alors libres de gérer leur clientèle, mais sont exclusifs et ne peuvent proposer des produits que de la compagnie dont ils sont mandataires, des associations ou des courtiers.

Les principales compagnies d'assurances à travers le monde sont Axa (France) et Allianz (Allemagne), Chartis Insurance (Etats-Unis), China Life (Chine).

1.3 La tarification en assurances

La tarification désigne le processus utilisé par les compagnies d'assurance pour évaluer si une personne ou un groupe de personnes est éligible à une couverture d'assurance. Les tarificateurs analysent un grand nombre de renseignements recueillis, qu'ils soient médicaux ou d'une autre nature, afin d'évaluer l'éligibilité et décider d'approuver ou de rejeter la proposition.

La tarification consiste à évaluer immédiatement les risques, ce qui nécessite une adaptation de la prime en fonction du risque supporté par les assurés. En assurance on distingue deux types de tarification, la première, dite une tarification a priori permet de déterminer le prix lors de la souscription du contrat en se basant sur les caractéristiques de l'assuré. La seconde est la tarification a posteriori, elle repose sur

la compréhension ultérieure des expériences passées en matière de sinistre, ce qui permet de rectifier la prime a priori, notamment à travers le système de "bonus-malus".

Dans la tarification a priori on utilise des variables dites explicatives pour attribuer un assuré à une classe tarifaire. Par exemple en assurance automobile, des facteurs tels que l'âge, le sexe, le lieu de résidence et la puissance du véhicule sont utilisés pour prédire, a priori, le montant des sinistres pour l'année à venir. Le principe de la tarification repose sur la loi des grands nombres, selon laquelle on fixe une prime égale à ce que l'on s'attend à payer en moyenne pour chaque classe de risque, ce qui permet à la compagnie d'assurance de maintenir un équilibre financier. Cependant, il est également possible de recourir à une tarification a posteriori qui s'appuie sur des informations passées.

La tarification est essentielle pour déterminer l'éligibilité de l'assuré à une couverture d'assurance et les coûts qui y sont associés. Le prix de l'assurance doit correspondre au niveau de risque que représente chaque assuré. La tarification permet d'évaluer ce risque et de fixer un tarif approprié pour la couverture. Les personnes qui présentent un risque plus élevé paieront généralement une prime supérieure à celle présentant un risque plus faible. Les critères de tarification varieront également selon le type et le montant de la couverture demandée, ainsi que les informations personnelles fournies, telles que l'âge, l'état de santé, le mode de vie et la profession.

En d'autres termes, si le souscripteur est jeune et que le capital assuré est peu élevé, les exigences de tarification peuvent être plus simples à remplir. En revanche, les tarificateurs demanderont généralement davantage de renseignements pour les souscripteurs plus âgés et pour des montants de couverture plus élevés.

1.4 La théorie de la crédibilité

La théorie de la crédibilité a été développée aux États-Unis dans les années 1910, dans le cadre des accidents du travail. À cette époque, des entreprises comme la multinationale General Motors (GM) et le petit constructeur indépendant Tucker étaient couvertes par Allstate, ainsi que d'autres fabricants automobiles, contre les

risques d'accidents du travail. Un taux moyen était calculé en se basant sur l'expérience de l'ensemble des constructeurs, puis attribué à chaque entreprise. Cependant, (GM) disposant d'une vaste expérience en matière de sécurité, a réalisé que le taux qui lui était appliqué était plus élevé que celui qu'elle aurait pu calculer elle-même. Constatant que son propre taux aurait été inférieur sur plusieurs années grâce à une expérience plus favorable que celle du groupe, (GM) a demandé à Allstate de lui appliquer son propre taux, soulignant que son grand nombre d'employés garantissait une stabilité suffisante des résultats au fil du temps.

Les actuaires d'Allstate partagent instinctivement le point de vue de (GM), mais ils hésitent face à un problème délicat : si le nombre d'employés de (GM) est suffisamment élevé pour qu'on puisse se fier à son expérience, tandis que celui de Tucker est trop faible pour permettre une telle chose, comment établir la limite entre un effectif considéré comme fiable et un autre qui ne l'est pas ?

En 1914, Mowbray fut le premier à répondre à cette question en instaurant la théorie de la crédibilité, basée sur la stabilité ou la limitation des fluctuations. Cela soulève une autre interrogation : que faire pour les entreprises de plus petite taille qui voudraient proposer un seuil déterminant à partir duquel leur effectif serait jugé adéquat ?

Il définit une prime pure "fiable" (dependable) comme étant "une prime ayant une forte probabilité de ne pas s'écarter de la vraie prime au-delà d'une limite arbitraire."

En termes mathématiques, cela se traduit par la condition :

$$\Pr [(1 - K) E[S] \leq S \leq (1 + K) E[S]] \geq P \quad (1.1)$$

où

K est petit, habituellement 5%,

P est près de 1, habituellement 0.90, 0.95 ou 0.99,

S désigne l'expérience d'un contrat, qu'elle soit sous une forme ou sous une autre.

Quelques années plus tard, Whitney (1918) évoque l'importance, pour l'assuré, de faire la balance entre l'expérience collective et l'expérience individuelle. Quatre facteurs sont retenus par Whitney dans son article pour déterminer la valeur à ac-

corder à l'expérience individuelle, il est important de prendre en compte l'exposition, le niveau de risque, la crédibilité de la prime collective et l'homogénéité du groupe. Dès lors, à l'aide d'une méthode mathématique, il peut donner une expression pour la prime individuelle (P) de la forme

$$P = ZX + (1 - Z)C. \tag{1.2}$$

Où X représente l'expérience personnelle et C représente l'expérience collective. Le facteur de crédibilité Z est utilisé pour pondérer ces valeurs, ce qui donne à Whitney une formule de la forme $n/(n + K)$.

En 1945, Bailey fait une première mention de la crédibilité dans un cadre qui semble proche de l'approche non paramétrique étudiée par Bühlmann. Cependant, une notation confuse rend son texte difficile à comprendre, ce qui en limite l'impact. En revanche, son article de 1950 a eu un effet plus significatif sur les fondements de la théorie de la crédibilité. Dans un contexte bayésien, Bailey montre que la diminution de l'erreur quadratique conduit à un estimateur qui correspond précisément à la prime de crédibilité, notamment pour les distributions binomiale/bêta, Poisson/gamma et normale/normale.

En 1967, Hans Bühlmann a formalisé la théorie de la crédibilité moderne avec son article "Experience Rating and Credibility". Bühlmann a développé une méthode pour combiner les données historiques d'un risque particulier avec des données générales pour obtenir une meilleure estimation des primes. Il est important de souligner que la contribution de Bühlmann (1967-1969) ne sera pas suffisante pour régler ce problème de tenir compte de l'expérience en tarification.

En 1970, Bühlmann, en collaboration avec son étudiant de doctorat « Erwin Straub », renforçait son modèle en y intégrant des poids pour les données et en définissant des estimateurs pour les paramètres structurels.

Dans les années 70, plusieurs modèles de précision de crédibilité ont vu le jour, alors que les années 80 ont été des années d'étude des estimateurs des paramètres structurels. Au milieu des années 80 et au début des années 90, la recherche en théorie de crédibilité a connu un ralentissement significatif.

Nous observons depuis quelques années un renouveau d'intérêt pour la recherche de nouveaux estimateurs de paramètres de structure, plus performants et dont les propriétés seraient mieux maîtrisées.

Notions

Étant donné que la loi des grands nombres n'est pas applicable à un seul individu, il est clair que chaque assuré, pris isolément, se verra attribuer une prime basée sur les résultats obtenus pour l'ensemble du portefeuille, qui regroupe plusieurs contrats évalués jusqu'à la date t .

Supposons que les montants annuels des sinistres x_1, x_2, \dots, x_n aient été enregistrés par l'individu j au cours des n périodes d'observation. Alors, la prime pure est

$$\bar{P}_n = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (1.3)$$

L'assureur pourrait décider de demander à cet assuré un montant \bar{P}_n , cependant, que faire des assurés qui n'ont jamais déclaré des sinistres (ceux pour lesquels $x_1 = x_2 = \dots = x_n = 0$), devrait-on les exempter du paiement de la prime tout en garantissant leur couverture ?

L'assureur est donc confronté à un dilemme : soit il poursuit sa demande d'un montant $P_{collectif}$ équivalent à tous les assurés, mais cela pourrait mécontenter les « bons » assurés qui ressentent une certaine gêne et sont donc susceptibles de faire concurrence, ou il se laisse tenter par \bar{P}_n .

La confrontation entre ces deux positions extrêmes a poussé les actuaires américains à solliciter une prime dont le montant constituerait un compromis.

L'idée consiste à tarifier le contrat n_j pour la $(t + 1)^{ème}$ période par

$$Pr = (1 - Z) P_{collectif} + Z \bar{P}_n \quad (1.4)$$

Où

Z est le facteur de crédibilité.

\bar{P}_n correspond à la moyenne des sinistres des années antérieures, en relation avec la prime individuelle de l'assuré.

$P_{collectif}$ est la prime collective correspond à la moyenne des sinistres encourus par l'ensemble des assurés.

Pr est la prime pure.

- Le risque individuel et son évolution dans le temps déterminent la grandeur Z . Son importance témoigne de la confiance accordée aux données concernant le risque lui-même par rapport aux données applicables à tous les groupes.

Il existe trois cas possibles :

$$Z = 1$$

$$Z = 0$$

$$Z \in]0; 1[$$

- Si $Z = 1$, Nous évoquons de la crédibilité totale, ce qui signifie que la prime pure correspond à la prime individuelle : les tarifs du contrat sont établis uniquement en fonction de l'expérience personnelle de l'assuré. La mise en œuvre de la crédibilité totale n'est possible que si la taille de la société est appropriée.

- Si $Z = 0$, La prime pure crédible est identique à la prime collective. Nous accordons peu d'importance à l'expérience individuelle et nous nous appuyons exclusivement sur les données des sinistres de l'ensemble du portefeuille.

- Si $Z \in]0; 1[$, dans cette situation, nous évoquons la crédibilité partielle, où la prime constituera le point central et de la prime collective et de la prime individuelle.

1.4.1 Crédibilité totale

La crédibilité totale est évoquée lorsque la prime de la $(t + 1)^{ème}$ période repose uniquement sur la sinistralité propre à l'assuré observée pendant les t premières périodes.

Considérons que le montant des sinistres engendrés par une compagnie soit $S = \sum_{k=1}^N X_k$ ou $N \sim P(\lambda)$ et (X_k) est une suite de variables aléatoires indépendantes et distribuées de manière uniforme, avec une moyenne μ et une variance σ^2 indépendante de N , qui représentent les frais associés aux sinistres. Si la probabilité de l'écart entre la charge de sinistres (S) et la prime pure $E(S)$ est assez faible (c%), nous accorderons

une crédibilité totale à l'entreprise, si :

$$\begin{aligned}
 P[(1-c)E(S) < S < (1+c)E(S)] &= 1 - \beta & (1.5) \\
 \implies P \left[\frac{-cE(S)}{\sqrt{Var(S)}} < \frac{S - E(S)}{\sqrt{Var(S)}} < \frac{cE(S)}{\sqrt{Var(S)}} \right] &= 1 - \beta
 \end{aligned}$$

où β et c sont des constantes fixées a priori (par exemple : $\beta = 5\%$ et $c = 5\%$).

La variable aléatoire $\frac{S-E(S)}{\sqrt{Var(S)}}$ est approximée par la loi normale centrée réduite de moyenne 0 et variance 1. Donc $\frac{cE(S)}{\sqrt{Var(S)}} = x_\beta$.

$$E(S) = E(E(S | N)) = E \left(E \left(\sum_{k=1}^N X_k | N \right) \right) = E(X) \times E(N) = \lambda\mu \quad (1.6)$$

$$\begin{aligned}
 Var(S) &= E(Var(S | N)) + Var(E(S | N)) \\
 &= E(N) \times Var(X) + Var(N) \times E(X)^2 \\
 &= \lambda\sigma^2 + \mu^2\lambda = \lambda(\sigma^2 + \mu^2) & (1.7)
 \end{aligned}$$

Donc

$$x_\beta = \frac{c\lambda\mu}{\sqrt{\lambda(\sigma^2 + \mu^2)}} \implies \lambda = \frac{x_\beta^2}{c^2} \left[1 + \left(\frac{\sigma}{\mu} \right)^2 \right] = \lambda_H \quad (1.8)$$

La valeur de λ_H doit être interprétée comme une mesure de la quantité de volume que la société doit atteindre pour lui attribuer une crédibilité totale. Dans l'expression de λ_H , les sinistres sont calculés en utilisant leur coefficient de variation $\frac{\sigma}{\mu}$.

1.4.2 Crédibilité partielle

Quand un assuré ne possède pas le niveau de risque λ_H , nous évoquons la crédibilité partielle. Dans ce cas, la prime sera un élément central de la prime collective et la prime individuelle.

Nous trouvons le facteur de crédibilité Z tel que

$$P[(ZE(S) - cE(S)) < ZS < (ZE(S) + cE(S))] = 1 - \beta \quad (1.9)$$

Equivalent à

$$P\left[\frac{-cE(S)}{Z\sqrt{Var(S)}} < \frac{S - E(S)}{\sqrt{Var(S)}} < \frac{cE(S)}{Z\sqrt{Var(S)}}\right] = 1 - \beta \quad (1.10)$$

où β et c sont des constantes fixées a priori (par exemple : $\beta = 5\%$ et $c = 5\%$).

La variable aléatoire $\frac{S-E(S)}{\sqrt{Var(S)}}$ est approximée par la loi normale centrée réduite de moyenne 0 et variance 1. Donc $\frac{cE(S)}{Z\sqrt{Var(S)}} = x_\beta$.

Donc

$$x_\beta = \frac{c\lambda\mu}{Z\sqrt{\lambda(\sigma^2 + \mu^2)}} \implies Z = \frac{\sqrt{\lambda}}{\frac{x_\beta}{c}\sqrt{1 + \left(\frac{\sigma}{\mu}\right)^2}} = \sqrt{\frac{\lambda}{\lambda_H}}, (\lambda < \lambda_H) \quad (1.11)$$

Avec

λ_H : la quantité minimale de sinistres prévue pour la prochaine période afin de pouvoir mettre en œuvre la crédibilité totale.

$\frac{\sigma}{\mu}$: le coefficient de variation.

λ : l'espérance du nombre des sinistres.

1.4.3 Prime individuelle et prime collective

Dans le modèle de crédibilité simplifié on a :

Définition 1.1 La prime individuelle pour un assuré ayant un profil de risque θ est

$$\mu(\theta_i) = E[S_{it} | \Theta = \theta_i] = \int_0^{+\infty} xf(x | \theta_i) dx \quad (1.12)$$

Définition 1.2 La prime collective est

$$m = E[\mu(\Theta)] = \int_{-\infty}^{+\infty} \mu(\theta) u(\theta) d\theta \quad (1.13)$$

on a

$$m = E[E[S_{it} | \Theta]] = E[S_{it}] \quad (1.14)$$

Cette prime représente le montant moyen de sinistre prévu par risque sur l'ensemble du portefeuille.

1.4.4 Prime bayésienne

La prime bayésienne est la prévision optimale de $S_{i,n+1}$ que nous pouvons déterminer. La prime bayésienne, tout comme la prime collective, est une moyenne pondérée des primes de risque, cependant, elle se base sur la répartition a posteriori de Θ au lieu de celle a priori, cette prime est donné par

$$B_{i,n+1} = \int_{-\infty}^{+\infty} \mu(\theta) u(\theta | S_{i1}, \dots, S_{in}) d\theta \quad (1.15)$$

En revanche, on peut considérer la prime collective comme une sorte de prime bayésienne pour la première année, lorsque aucune donnée expérience n'est encore disponible.

Deux problèmes pratiques se posent en matière de crédibilité bayésienne :

1. La prime bayésienne ne sert de prime de crédibilité qu'à des fins particulières.
2. La prime repose sur des suppositions hautement subjectives en ce qui concerne les distributions de Θ_i et $S_{it} | \Theta_i$.

Afin de traiter le premier problème, Bühlmann (1967) a restreint l'évaluation de la prime de risque à des fonctions linéaires des observations, soit sous la forme

$$c_0 + \sum_{t=1}^n c_t S_{it} \quad (1.16)$$

On verra que la prime de crédibilité est la meilleure approximation.

Quant au second problème, il sera résolu en utilisant une approche non paramétrique pour déterminer la prime de crédibilité, tel que proposé par Bühlmann (1969).

1.4.5 Le modèle de crédibilité de Bühlmann

Au milieu du 20^{ème} siècle, le modèle de crédibilité développé par Bühlmann a fait ses débuts en tant que premier modèle de crédibilité dans le domaine de la tarification en assurances. À la base, le modèle cherchait à affiner le processus d'estimation des primes simplement en intégrant les expériences individuelles des assurés et les données collectives plus larges sur les risques.

Jusqu'à présent, la tarification d'un risque spécifique a été étudiée. Maintenant, nous allons prendre en compte un portefeuille de risques comparables.

On note $X_i = (X_{i1}, \dots, X_{in})'$ le vecteur des observations lié au risque i sera noté, tandis que Θ_i représentera son profil de risque. Les risques I sont semblables car chaque risque i est conforme aux Hypothèses (H1) et (H2). Il s'agit donc du contexte défini par Bühlmann [1967].

Hypothèses de Modèle de Bühlmann

(B1) Les X_{ij} ($j = 1, \dots, n$) sont des variables aléatoires, conditionnellement à $\Theta_i = \theta$, indépendantes et distribuées de manière identique, selon une loi F_θ avec les moments conditionnels.

$$\begin{aligned}\mu(\theta) &= E[X_{ij} \mid \Theta_i = \theta] \\ \sigma^2(\theta) &= Var[X_{ij} \mid \Theta_i = \theta]\end{aligned}$$

(B2) Les couples $(\Theta_1; X_1), \dots, (\Theta_I; X_I)$ sont indépendants et identiquement distribués.

Le niveau de (B2) indique une diversité dans le portefeuille : les profils de risque $1, \dots, I$ sont indépendants et présentent la même répartition de structure $U(\theta)$. Ce-

pendant, ces risques partagent le fait qu'on ne peut les distinguer à première vue.

Notations

On définit la notation suivante

$$\begin{aligned}\mu &= E[\mu(\Theta_i)] \\ a &= Var [E [S_{it} | \Theta_i]] = Var [\mu (\Theta_i)] \\ S^2 &= E [Var [S_{it} | \Theta_i]] = E [\sigma^2 (\Theta_i)]\end{aligned}$$

Formule de l'estimateur de crédibilité de Bühlmann

Suite à la résolution de problème de minimisation suggérée par Bühlmann en 1967, sous forme linéaire (voir [5]), ce dernier a obtenu un estimateur de crédibilité pour la prime d'assurance $\hat{\mu}(\Theta_i)$ tel que

$$\hat{\mu}(\Theta_i) = Z_i \bar{X}_i + (1 - Z_i) \mu \quad (1.17)$$

Où

$\hat{\mu}(\Theta_i)$: Estimation finale de la prime de crédibilité.

μ : Moyenne globale des sinistres pour la totalité de la population.

\bar{X}_i : Moyenne des sinistres observés.

Z : Facteur de crédibilité, qui est calculé par : $Z = \frac{n}{n + \frac{S^2}{a}}$ tel que n est le nombre d'observations.

Application de modèle de crédibilité de Bühlmann en assurance non-vie

Dans cette étude, on va utiliser le modèle de Bühlmann pour calculer les primes d'assurance à court terme en exploitant les données authentique des contrats d'assurance générale publiées dans les rapports annuels de l'autorité de régulation des assurances dans un pays africain, couvrant la période 2017 à 2022 pour 13 contrats. Les contrats sélectionnés comprenaient par ordre : l'aviation, l'ingénierie, l'incendie domestique, l'incendie industriel, la responsabilité civile, la marine, l'automobile privée, l'automobile commerciale, les accidents corporels, le vol, l'indemnisation des accidents du travail, l'assurance de matériel médicale et les assurances diverses.

Les données globales sont indiquées ci dessous

Risques	2017	2018	2019	2020	2021	2022
1	48571	28336	9285	43775	3515	5880
2	431756	411244	597785	697881	592670	708591
3	309981	477500	414559	394133	421323	205136
4	1100373	1141362	825219	1149283	979632	1486378
5	340822	597548	399152	462135	636608	513688
6	725376	669398	669464	598869	1043536	748793
7	14181895	14301162	16635527	15418425	18015197	19585676
8	12646816	13618583	15304543	14705102	17884591	17686508
9	1081310	747352	464566	481601	1063712	294008
10	969900	1304641	887812	722603	860395	694374
11	3126587	2374487	1548271	1910300	1752873	2291780
12	20632162	20515402	20424313	20975226	25941119	32583822
13	653554	740981	781086	752126	647599	878928

Tableau 1.1. Données de portefeuille (montant $\times 10^3$).

Maintenant, pour appliquer le modèle de Bühlmann, notre première étape consiste à calculer le coût moyen des sinistres pour chaque contrat. Nous procédons ensuite au calcul du coût moyen global des sinistres pour 13 contrats d'assurance non-vie, qui joue un rôle essentiel dans l'évaluation de la variance intra et inter-contractuelle. Enfin, nous déterminons le facteur de crédibilité. Collectivement, ces étapes constituent une estimation des paramètres sous le modèle Bühlmann.

Estimation des paramètres sous le modèle de Bühlmann

Soit X_{ij} le nombre de sinistres pour le i^{eme} risque, où $i = 1, 2, \dots, N = 13$, au cours de l'année j , où $j = 1, 2, \dots, n = 6$.

$m(\Theta_i)$, qui est la prime de crédibilité pour la prochaine période immédiate donnée X obtenue en obtenant d'abord les estimations de $E[m(\Theta)]$, $var[m(\Theta)]$ et $E[s^2(\Theta)]$ sous réserve des hypothèses suivantes :

- (a) La distribution de chaque X_j dépend d'un paramètre Θ , qui est fixé.
- (b) Étant donné la valeur fixe de Θ , les X_j^s sont présumés indépendants et distribués identiquement.
- (c) Les niveaux d'activité variables demeurent constants tout au long de chaque contrat.

Par conséquent, les estimateurs non biaisés peuvent être obtenus comme suit

$$\mu = E[m(\Theta)] = \bar{X} = \frac{\sum_{i=1}^N \bar{X}_i}{N} \quad (1.18)$$

$$var[m(\Theta)] = \frac{1}{N-1} \sum_{i=1}^N (\bar{X}_i - \bar{X})^2 - \frac{1}{Nn} \sum_{i=1}^N \left[\frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 \right] \quad (1.19)$$

$$Z = \frac{n}{n + \frac{E[s^2(\Theta)]}{var[m(\Theta)]}} \quad (1.20)$$

La prime de crédibilité est donnée par

$$\hat{\mu}(\Theta_i) = Z_i \bar{X}_i + (1 - Z_i) \mu \quad (1.21)$$

Résultats et discussions

Dans cette partie, nous exposons le modèle de crédibilité de Bühlmann pour calculer les primes de crédibilité. En outre, il estime les paramètres $E[m(\Theta)]$, $E[s^2(\Theta)]$ et $var[m(\Theta)]$, qui seront essentiels pour estimer la crédibilité primes pour les contrats d'assurance sur la base des données du tableau (1.1). Les calculs impliquaient l'application des équations (1.18) à (1.21) ainsi que l'utilisation des statistiques récapit-

tulatives présentées dans le tableau (1.2) (dérivées du tableau (1.1)).

Risques	\bar{X}_i	$\frac{1}{5} \sum_{j=1}^6 (X_{ij} - \bar{X}_i)^2$	$\frac{1}{12} \sum_{i=1}^{13} (\bar{X}_i - \bar{X})^2$
1	23227000	$3.949002476 \times 10^{11}$	2.325771×10^{19}
2	573321166.66	1.621913×10^{13}	1.825452×10^{19}
3	370438666.66	9.507894×10^{12}	2.002933×10^{19}
4	1113707833.33	4.845867×10^{13}	1.39289×10^{19}
5	491658833.33	1.29778×10^{13}	1.8959×10^{19}
6	742572666.66	2.44527×10^{13}	1.68369×10^{19}
7	16356313666.66	4.618003×10^{15}	1.324907×10^{20}
8	15307690500	4.51913×10^{15}	1.094501×10^{20}
9	688758166.66	1.094249×10^{14}	1.728143×10^{19}
10	906620833.33	4.876524×10^{13}	1.551754×10^{19}
11	2167383000	3.199298×10^{14}	7.174197×10^{18}
12	23512007333.33	2.428813×10^{16}	3.484253×10^{20}
13	742379000	7.42279×10^{12}	1.683849×10^{19}
Somme	62996078667	3.402282×10^{16}	7.584442×10^{20}
Moyenne	4845852205	2.61714×10^{15}	6.320368×10^{19}

Tableau 1.2.statistique récapitulatives

Où

$$\begin{aligned}
 E [s^2(\Theta)] &= \frac{1}{N} \sum_{i=1}^N \left[\frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 \right] \\
 &= \frac{1}{13} \sum_{i=1}^{13} \left[\frac{1}{5} \sum_{j=1}^6 (X_{ij} - \bar{X}_i)^2 \right] \\
 &= \frac{1}{13} \times 3.402282 \times 10^{16} \\
 &= 2.61714 \times 10^{15} \approx 2.62 \times 10^{15}
 \end{aligned}$$

$$\mu = E[m(\Theta)] = \bar{X} = \frac{\sum_{i=1}^N \bar{X}_i}{N} = \frac{62996078667}{13} = 4845852205$$

$$\begin{aligned} \text{var}[m(\Theta)] &= \frac{1}{N-1} \sum_{i=1}^N (\bar{X}_i - \bar{X})^2 - \frac{1}{Nn} \sum_{i=1}^N \left[\frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 \right] \\ &= \frac{1}{12} (7.584442 \times 10^{20}) - \frac{1}{6} E[s^2(\Theta)] \\ &= \frac{1}{12} (7.584442 \times 10^{20}) - \frac{1}{6} (2.61714 \times 10^{15}) \\ &= 6.320324714333333 \times 10^{19} \approx 6.32 \times 10^{19} \end{aligned}$$

Et

$$Z = \frac{n}{n + \frac{E[s^2(\Theta)]}{\text{var}[m(\Theta)]}} = \frac{6}{6 + \frac{2.62 \times 10^{15}}{6.32 \times 10^{19}}} = 0.999993$$

Les estimations des paramètres structurels et des facteurs de crédibilité sont donc présentées dans le tableau (1.3).

Paramètre	Valeur
$E[m(\Theta)]$	4845852205
$E[s^2(\Theta)]$	2.62×10^{15}
$\text{var}[m(\Theta)]$	6.32×10^{19}
Z	0.999993

Tableau 1.3. Paramètres de structure et facteur de crédibilité

Après avoir appliqué le modèle de crédibilité de Bühlmann et utilisé les paramètres structurels comme composants, les valeurs de prime de crédibilité résultantes pour

la prochaine période immédiate sont présentées dans le tableau (1.4).

Risques	\bar{X}_i	Z_i	Primes de crédibilité
1	23227000	0.999993	23260758
2	573321166.66	0.999993	573351074
3	370438666.66	0.999993	370469995
4	1113707833.33	0.999993	1113733958
5	491658833.33	0.999993	491689313
6	742572666.66	0.999993	742601390
7	16356313666.66	0.999993	16356233093
8	15307690500	0.999993	15307617267
9	688758166.66	0.999993	688787266
10	906620833.33	0.999993	906648408
11	2167383000	0.999993	2167401749
12	23512007333.33	0.999993	23511876670
13	742379000	0.999993	742407724

Tableau 1.4. Estimations de la prime de crédibilité

Le tableau (1.4) démontre une uniformité du facteur de crédibilité pour tous les contrats. Cette uniformité découle de l'hypothèse faite par le modèle de crédibilité de Bühlmann, qui considère que les volumes de risque sont identiques. De plus, on peut observer que les primes de crédibilité pour chaque contrat présentent des variations. Cette variance est un reflet direct des différentes expériences de sinistres.

Enfin, il est clair que tous des données ont un poids réduit, comme l'indique la variance globale plus élevée présentée dans le tableau (1.3). Cela souligne une plus grande dépendance aux données du risque individuel par rapport à l'expérience du groupe.

L'objectif principal de cette étude était de calculer les coûts moyens des sinistres pour chaque type de contrat d'assurance générale, d'établir les facteurs de crédibilité correspondants pour chacun et de déduire la prime de crédibilité appropriée pour

chaque assuré en fonction des sinistres globaux de chaque contrat. Au cours de cette application, le modèle de crédibilité de Bühlmann a été employé pour obtenir les coûts moyens associés aux sinistres, les facteurs de crédibilité et les primes d'assurance. Il convient de souligner que la fixation des prix c'est à dire la tarification est essentielle dans le fonctionnement des compagnies d'assurance, car les primes collectées constituent généralement leur principale source de revenus. Par conséquent, il est essentiel de disposer de méthodes efficaces pour déterminer les niveaux de primes appropriés pour les clients. Il est donc recommandé aux principaux acteurs du secteur de l'assurance (assureurs et régulateur) d'envisager, d'adopter cette approche tout en intégrant d'autres facteurs qui influencent l'estimation des primes. Ce faisant, ils peuvent améliorer la précision de la tarification des produits d'assurance générale, cela conduit à une amélioration générale des pratiques au sein du secteur.

1.4.6 Le Modèle de crédibilité de Bühlmann-Straub

Le premier modèle de Bühlmann (1967) avait un désavantage car il supposait à priori que tous les contrats étaient identiques, sans prendre en compte leur taille. A cause de ce défaut, Bühlmann et son étudiant Straub ont amélioré le premier modèle dans leur article en 1970, Afin d'inclure les données concernant la taille du contrat. Prenons l'exemple d'un ensemble de contrats d'assurance non vie qui nécessitent une tarification. Une segmentation a été effectuée sur cet ensemble de contrats, cela a conduit à la création de catégories de risque dans lesquelles on ne peut pas préalablement distinguer les risques. La théorie de la crédibilité basée sur le passé sinistre, permettra d'améliorer les informations a priori pour chaque danger.

Dans ce contexte, le modèle de Bühlmann utilise une Hypothèse peu réaliste :

$$\sigma^2(\theta) = Var[X_{it} | \Theta_i = \theta] \quad (1.22)$$

Il est tout à fait normal que cette variance conditionnelle diminue avec l'exposition au risque i . Dans ce contexte, le modèle de Bühlmann-Straub généralise le modèle de Bühlmann en renforçant cette supposition par l'ajout de poids. Ce modèle est

couramment employé dans le domaine de l'assurance non-vie, de l'assurance vie ou de la réassurance.

Ensuite, nous étudierons un ensemble de I risques et évaluerons X_{it} , qui indique le taux de sinistre (ou la fréquence des dommages, la taille moyenne des dommages ou tout autre indicateur) du danger i au cours de l'année t , tandis que ω_{it} indique le poids lié.

Hypothèses de (Modèle de Bühlmann-Straub)

Ce modèle propose des hypothèses semblables à celles de Bühlmann, mais inclut une variable de volume ω_{it} additionnelle qui reflète l'importance de la police i durant l'année t dans le portefeuille. Le poids de chaque variable X_{it} était identique dans le modèle précédent, tandis que dans ce modèle, X_{it} est associée au poids ω_{it} . Pour chaque contrat $i = 1, \dots, I$, X_{it} est la moyenne de groupe de ω_{it} contrats, où ω_{it} représente le poids de groupe i à l'année t , avec $t = 1, \dots, n$:

$$X_{it} = \frac{1}{\omega_{it}} \sum_{r=1}^{\omega_{it}} X_{it}^r, \quad t = 1, \dots, n \text{ et } i = 1, \dots, I \quad (1.23)$$

Le risque i est précisé par une réalisation de Θ_i , c'est à dire le profil de risque θ_i .

(Hp1) Les X_{it} ($t = 1, \dots, n$) sont des variables aléatoires, conditionnellement à Θ_i , indépendantes avec les moments conditionnels :

$$\begin{aligned} \mu(\Theta_i) &= E[X_{it} | \Theta_i] \\ \text{Var}[X_{it} | \Theta_i] &= \frac{\sigma^2(\Theta_i)}{\omega_{it}} \\ \text{Cov}(X_{it}, X_{is} | \Theta_i) &= \frac{\delta_{ts}}{\omega_{it}} \sigma^2(\Theta_i) \end{aligned}$$

(Hp2) Les $(\Theta_1, X_1), \dots, (\Theta_I, X_I)$ sont indépendants et les $\Theta_1, \dots, \Theta_I$ sont indépendants et identiquement distribués.

Paramètres de structure

La prime collective est $\mu_0 = E(\mu(\Theta_i))$ avec la prime individuelle $\mu(\Theta_i)$.

1.4. La théorie de la crédibilité

La variance intra contrat $\sigma^2 = E(\sigma^2(\Theta_i))$ mesure la capacité d'un contrat à varier au cours du temps.

La variance inter contrat $\tau^2 = Var(\mu(\Theta_i))$ mesure l'hétérogénéité du portefeuille.

Notation

$$\begin{aligned}\omega_{\bullet\bullet} &= \sum_{i=1}^I \omega_{i\bullet} = \sum_{i=1}^I \sum_{t=1}^n \omega_{it} \\ Z_i &= \frac{\omega_{i\bullet}}{\omega_{i\bullet} + \frac{\sigma^2}{\tau^2}} \text{ et } Z_{\bullet} = \sum_{i=1}^I Z_i \\ X_{i\omega} &= \sum_{t=1}^n \frac{\omega_{it}}{\omega_{i\bullet}} X_{it} \\ X_{\omega\omega} &= \sum_{i=1}^I \frac{\omega_{i\bullet}}{\omega_{\bullet\bullet}} X_{i\omega} = \sum_{i=1}^I \sum_{t=1}^n \frac{\omega_{it}}{\omega_{\bullet\bullet}} X_{it} \\ X_{z\omega} &= \sum_{i=1}^I \frac{Z_i}{Z_{\bullet}} X_{i\omega}\end{aligned}$$

L'estimateur de crédibilité du modèle "Bühlmann-Straub" est

$$\widehat{S}_i = Z_i X_{i\omega} + (1 - Z_i) \widehat{\mu}_0 \quad (1.24)$$

Avec $Z_i = \frac{\omega_{i\bullet}}{\omega_{i\bullet} + \frac{\sigma^2}{\tau^2}}$. (Voir preuve de Bühlmann-Straub (1970)).

Application de modèle de crédibilité de Bühlmann-Straub

Soit une entreprise d'assurance qui compte deux groupes d'assurés A et B, le tableau ci-dessous présente les coûts totalisés des trois premières années de sinistre.

Groupes	Année	1	2	3	4
A	Montant du coût total	20000	24000	19000	?
	Nombre de personnes	100	120	115	95
B	Montant du coût total	8000	11000	15000	?
	Nombre de personnes	40	50	70	75

Tableau 1.5. Coûts totales des trois premières années

Maintenant, nous allons calculer le montant du coûts totales de la 4^{ème} année pour chaque groupe. Dans cet application, nous disposons deux groupes ($I = 2$) chacun possédant expérience de trois années ($n_1 = n_2 = 3$), le poids (ω_{it}) est le nombre de personnes dans le groupe i , pour la $t^{\text{ème}}$ année. Le tableau suivant montre les coûts moyens X_{it} avec le poids ω_{it} pour chaque groupe dans les trois années (X_{it} est la division du montant du coût total par le nombre de personnes).

Groupes	Année	1	2	3	4
A	X_{1t}	200	200	165.22	?
	ω_{1t}	100	120	115	95
B	X_{2t}	200	220	214.29	?
	ω_{2t}	40	50	70	75

Tableau 1.6. Coûts moyens des trois premières années

Les tailles totales et les moyennes totales du coût sont obtenues comme suit

La taille totale de groupe A	La taille totale de groupe B	Le poid total
$\omega_{1\bullet} = \sum_{t=1}^3 \omega_{1t} = 335$	$\omega_{2\bullet} = \sum_{t=1}^3 \omega_{2t} = 160$	$\omega_{\bullet\bullet} = \sum_{i=1}^2 \sum_{t=1}^3 \omega_{it} = 495$
La moyenne totale du coût pour groupe A	La moyenne totale du coût pour groupe B	La moyenne totale des deux groupes
$X_{1\omega} = 188.06$	$X_{2\omega} = 212.5$	$X_{\omega\omega} = 195.96$

La valeur de l'espérance de la moyenne a priori est

$$\hat{\mu}_0 = X_{\omega\omega} = \sum_{i=1}^I \frac{\omega_{i\bullet}}{\omega_{\bullet\bullet}} X_{i\omega} = \sum_{i=1}^I \sum_{t=1}^n \frac{\omega_{it}}{\omega_{\bullet\bullet}} X_{it} = 195.96$$

La valeur de l'espérance de la variance a priori est

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^I \sum_{t=1}^n \omega_{it} (X_{it} - X_{i\omega})^2}{I(n-1)} = 25160.58$$

1.5. Conclusion

La valeur de la variance de la moyenne a priori est

$$\widehat{\tau}^2 = \frac{\omega_{\bullet\bullet} \left(\sum_{i=1}^I \omega_{i\bullet} (X_{i\omega} - X_{\omega\omega})^2 - \widehat{\sigma}^2 (I - 1) \right)}{\omega_{\bullet\bullet}^2 - \sum_{i=1}^I \omega_{i\bullet}^2} = 182.48$$

Les valeurs de facteurs de crédibilité sont respectivement

$$Z_1 = \frac{\omega_{1\bullet}}{\omega_{1\bullet} + \frac{\sigma^2}{\tau^2}} = \frac{335}{335 + 137.88} = 0.708$$
$$Z_2 = \frac{\omega_{2\bullet}}{\omega_{2\bullet} + \frac{\sigma^2}{\tau^2}} = \frac{160}{160 + 137.88} = 0.537$$

Selon le modèle de Bühlmann-Straub, la sinistralité moyenne (coût moyen) est de

$$\widehat{S}_1 = Z_1 X_{1\omega} + (1 - Z_1) \widehat{\mu}_0 = 190.37$$
$$\widehat{S}_2 = Z_2 X_{2\omega} + (1 - Z_2) \widehat{\mu}_0 = 204.84$$

Donc on obtient le montant du coût total de sinistres prévu par ces deux groupes pour l'année 4 comme suit

$$\widehat{C}_1 = \omega_{14} \times \widehat{S}_1 = 95 \times 190.37 = 18085$$
$$\widehat{C}_2 = \omega_{24} \times \widehat{S}_2 = 75 \times 204.84 = 15363$$

1.5 Conclusion

Dans ce chapitre, nous avons exploré les bases théoriques de la crédibilité en assurance, en se concentrant sur les modèles de Bühlmann et Bühlmann-Straub. Ces modèles se distinguent par leur capacité à fournir des ajustements précis des primes d'assurance en intégrant à la fois les données spécifiques des groupes et les variations inter-groupes. Cependant, une limite importante de ces modèles est leur hypothèse

d'indépendance des risques d'assurance, une condition qui ne s'applique pas toujours dans la pratique.

Alors que les modèles de Bühlmann et Bühlmann-Straub offrent des outils puissants pour la tarification des risques sous l'hypothèse d'indépendance, ils peuvent ne pas refléter correctement les situations où les risques sont interdépendants. Dans le chapitre suivant, nous explorerons les dynamiques de dépendance entre les risques d'assurance, en discutant des modèles et méthodes qui permettent de détecter et d'étudier ces dépendances.

CHAPITRE 2

La modélisation de la dépendance entre les sinistres d'assurance

Dans la littérature actuelle en actuariat et en assurance, de nombreux articles traitent de la notion de la dépendance entre les sinistres sous différents aspects. Les articles de Yeo et Valdez (2006), Wen et al.(2009) et Ebrahimzadeh et al.(2013), proposent l'utilisation d'effets communs pour prendre en compte la dépendance aux sinistres et fournir les primes de crédibilité. L'article de Wen et Deng (2011) généralise le modèle de crédibilité de Bühlmann en considérant qu'il existe une structure de corrélation égale sur les risques et donne une estimation des paramètres structurels. De plus, Zhang et al.(2018) donne une version étendue de l'idée de Wen et Deng (2011) en ce qui concerne une fonction de perte équilibrée, il suppose l'existence d'un taux d'intérêt constant et d'une structure de corrélation identique sur les risques et du temps pour obtenir de nouvelles primes pour le modèle de crédibilité de Bühlmann. D'autres modèles ont exploité les recherches menées par Bolancé et al.(2003), Purcaru et Denuit (2002, 2003), Frees et coll.(1999, 2001) et les références qui y figurent où la dépendance conditionnelle des réclamations au temps ont été étudiées.

En statistique et en probabilité, la recherche de la relation entre deux ou plusieurs variables permet d'évaluer la possibilité d'une connexion entre ces variables.

Le coefficient de corrélation linéaire est souvent employé pour mesurer la dépendance entre deux ou plusieurs variables. Cette mesure de dépendance est efficace lorsque la dimension est réduite, mais elle pourrait se révéler inefficace lorsque celle-ci grandit.

Habituellement, trois mesures de dépendance sont utilisées dans la littérature. Nous les présentons ci-après.

2.1 Les mesures de dépendance

Plus Classiquement, Les mesures couramment utilisées (d'Agostino et Stephens, 1986, Joe,1997, Genest et Favre,2007) pour la caractérisation de la dépendance sont :

1) Le coefficient de corrélation linéaire de pearson

Définition 2.1 Soient X et Y deux variables aléatoires avec des variances finies, le coefficient de corrélation linéaire entre X et Y est donné par

$$r(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X).Var(Y)}} \quad (2.1)$$

2) Le coefficient de corrélation de kendall et le rho de spearman

a) Notation de la concordance

Définition 2.2 Notons (X_1, Y_1) et (X_2, Y_2) deux observations d'un vecteur aléatoire continu (X, Y) , alors (X_1, Y_1) et (X_2, Y_2) sont dites

- Concordantes si

$$[(x_1 - x_2)(y_1 - y_2)] > 0 \Leftrightarrow (x_1 < x_2 \text{ et } y_1 < y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 > y_2)$$

- Discordantes si

$$[(x_1 - x_2)(y_1 - y_2)] < 0 \Leftrightarrow (x_1 < x_2 \text{ et } y_1 > y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 < y_2)$$

Généralement soit $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ un échantillon de n observations d'un couple (X, Y) . il existe $C_n^2 = \frac{n!}{2!(n-2)!}$ paires de distributions de couples $(x_i, y_i), (x_j, y_j)$ qui sont, soit concordantes, soit discordantes.

2.1. Les mesures de dépendance

Soit (X_1, Y_1) et (X_2, Y_2) deux couples de variables aléatoires continues, ayant des lois conjointes H_1 et H_2 mais partageant des marges communes F pour X_1 et X_2 et G pour Y_1 et Y_2 .

Définition 2.3 La fonction de concordance entre les couples (X_1, Y_1) et (X_2, Y_2) est définie par

$$Q = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \quad (2.2)$$

b) Le taux de kendall

Définition 2.4 (X_1, Y_1) et (X_2, Y_2) sont deux vecteurs aléatoires continus (iid) de fonction de répartition conjointe H, de fonctions marginales F pour $(X_1$ et $X_2)$ et G pour $(Y_1$ et $Y_2)$, le taux de kendall noté τ est défini par

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \quad (2.3)$$

Il est possible d'estimer le taux de kendall à partir d'un échantillon $\{(x_1, y_1), \dots, (x_n, y_n)\}$ de (X, Y) de la façon suivante

$$\hat{\tau} = \frac{2}{n(n-1)} \sum_{j=2}^n \sum_{i=1}^{j-1} \text{signe}(x_j - x_i)(y_j - y_i) \quad (2.4)$$

Où

$$\text{signe}(z) = \begin{cases} 1 & \text{si } z \geq 0 \\ -1 & \text{si } z < 0 \end{cases}$$

c) Le coefficient de corrélation des rangs "rho de spearman"

Définition 2.5 Soit (X, Y) un vecteur aléatoire avec deux copies vecteurs (\tilde{X}, \tilde{Y}) et (X', Y') , alors, le rho de Spearman est donné par

$$\rho_s(X, Y) = 3 \left[P \left\{ (X - \tilde{X})(Y - Y') > 0 \right\} - P \left\{ (X - \tilde{X})(Y - Y') < 0 \right\} \right] \quad (2.5)$$

On peut écrire le rho de Spearman en utilisant le coefficient de corrélation de

Pearson $\rho_s(X, Y) = \rho(F_X(x), F_Y(y))$ ou F_X et F_Y sont les fonctions de répartition de X et Y .

La construction d'un estimateur de *rho* de Spearman à partir d'un échantillon $\{(x_1, y_1), \dots, (x_n, y_n)\}$ de (X, Y) , est la suivante :

$$\hat{\rho}_s(X, Y) = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}} \quad (2.6)$$

avec

R_i : est un statistique de rang.

S_i : est un statistique de signe.

2.2 Modèle à effets communs à deux niveaux de la dépendance aux sinistres

Le modèle à effets communs à deux niveaux de la dépendance aux sinistres peut être utilisé dans de nombreuses formes d'assurance (habitation, automobile, etc...). Prenons l'exemple d'une assurance habitation, qui assure la maison contre le feu et le vol, ainsi que d'autres catastrophes potentielles comme des cyclones ou des inondations. Les deux niveaux d'effets communs peuvent être observés, avec les individus affectés par des événements ordinaires, non catastrophiques, qui sont généralement individualistes dans la nature, et ceux affectés par des événements probablement moins fréquents, mais catastrophiques, qui affectent généralement une partie du portefeuille d'assureurs.

Nous allons exposer dans cette partie un modèle qui repose sur l'approche bayésienne. Avec ce modèle de dépendance, il sera possible de compter à la fois sur l'expérience des sinistres parmi les risques individuels et de compter sur un risque individuel particulier au fil du temps. Ainsi, nous présentons les hypothèses sélectionnées pour ce modèle.

2.2.1 Hypothèses du modèle

Nous allons nous intéresser à l'estimation des réclamations de la période suivante pour chaque individu $X_{i,T+1}$, sur la base de toutes les réclamations observées X_1, X_2, \dots, X_I . Ainsi, notre objectif est d'évaluer $E(X_{i,T+1} \mid X_1, X_2, \dots, X_I)$ sur la base du modèle de dépendance d'effets communs à deux niveaux. En effet, on finit par évaluer la distribution de la variable aléatoire $X_{i,T+1} \mid X_1, X_2, \dots, X_I$.

La dépendance entre les risques individuels sera décrite par un effet commun noté Λ qui est une variable aléatoire où la fonction de probabilité sera supposée connue et notée par $f_\Lambda(\lambda)$. Les réalisations de cet effet commun sont notées λ . Sous cet effet commun, les vecteurs aléatoires X_i sont indépendants. Comme Λ est un effet courant parmi tous les risques, il définira la structure de dépendance entre les risques, et il pourra soit être une variable discrète, continue ou mixte. Maintenant, pour un individu fixe disons i , la dépendance des sinistres dans le temps peut être décrite par une autre variable aléatoire à effet commun notée Θ_i dont la fonction de densité de probabilité est censée être connue et notée $f_{\Theta_i}(\theta_i)$, avec θ_i désignant sans ambiguïté ses réalisations.

À la lumière de tous ceux-ci, nous donnons les hypothèses de ce modèle :

H1) La fonction de densité de probabilité de la variable aléatoire, connue $f_\Lambda(\lambda)$.

H2) Les vecteurs aléatoires $X_i \mid \Lambda$ où $X_i = (X_{i,1}, X_{i,2}, \dots, X_{i,T})'$ avec $i = 1, 2, \dots, I$, sont conditionnellement indépendants.

H3) Pour $i = 1, 2, \dots, I$, la variable aléatoire Θ_i a une fonction de densité de probabilité connue $f_{\Theta_i}(\theta_i)$.

H4) Les variables aléatoires $\Theta_1, \Theta_2, \dots, \Theta_I$ sont indépendantes deux à deux c'est à dire Θ_i est indépendant de Θ_j pour tout $i \neq j$ avec $i, j = 1, 2, \dots, I$.

H5) Pour $i = 1, 2, \dots, I$ la variable aléatoire Θ_i est indépendante de Λ .

H6) Pour $i = 1, 2, \dots, I$ et $t = 1, 2, \dots, T$, il y a une variable aléatoire conditionnelle $X_{i,t} \mid \Theta_i, \Lambda$ avec une fonction de densité de probabilité notée par $f_{X_{i,t} \mid \Theta_i, \Lambda}(x_{i,t} \mid \theta_i, \lambda)$.

H7) Les vecteurs aléatoires $X_{i,t} \mid \Theta_i, \Lambda$ avec $i = 1, 2, \dots, I$ sont conditionnellement indépendants.

H8) Les $X_{i,t} \mid \Theta_i, \Lambda$ avec $i = 1, 2, \dots, I$ et $t = 1, 2, \dots, T$, sont conditionnellement

indépendantes.

Remarque 2.1 L'hypothèse (H4) stipule que les paramètres de risque individuels sont indépendants l'un à l'autre. Ceci est raisonnable en raison de l'existence de Λ , la variable aléatoire Λ prend déjà en compte toutes les caractéristiques communes à chaque risque individuel de sorte que les Θ_i soient des caractéristiques individuelles vraiment particulières. L'hypothèse (H5) stipule également que le paramètre de risque global est indépendant de tous les paramètres des risques individuels.

Remarque 2.2 L'hypothèse (H7) ci-dessus affirme simplement que compte tenu du paramètre de risque global, tous les individus sont indépendants. L'hypothèse (H8) stipule qu'étant donné la connaissance de λ et θ_i , l'expérience du risque de chaque individu à n'importe quelle période de temps est indépendante de toute autre expérience, y compris celles de la même personne, mais à une période de temps différente. C'est encore raisonnable comme connaissance des deux paramètres λ et θ_i , supprime les sources de dépendance entre les risques et dans le temps pour le même risque.

Lemme 2.1 Considérons le modèle à effets communs à deux niveaux qui satisfait les hypothèses de (H1) jusqu'à (H8). La densité conjointe des risques individuels X , le paramètre du risque individuel Θ et le paramètre du risque global Λ peuvent être exprimés comme

$$f_{X,\Theta,\Lambda}(x, \theta, \lambda) = f_{\Lambda}(\lambda) \times \prod_{i=1}^I f_{\Theta_i}(\theta_i) \times \prod_{i=1}^I f_{X_i|\Theta_i,\Lambda}(x_i | \theta_i, \lambda) \quad (2.7)$$

Preuve En utilisant la définition de la densité conditionnelle $X_1, \dots, X_I | \Theta, \Lambda$ on a

$$f_{X|\Theta,\Lambda}(x | \theta, \lambda) = \frac{f_{X,\Theta,\Lambda}(x, \theta, \lambda)}{f_{\Theta,\Lambda}(\theta, \lambda)} \quad (2.8)$$

En se basant sur (H5), on a

$$f_{X|\Theta,\Lambda}(x | \theta, \lambda) = \frac{f_{X,\Theta,\Lambda}(x, \theta, \lambda)}{f_{\Theta,\Lambda}(\theta, \lambda)} = \frac{f_{X,\Theta,\Lambda}(x, \theta, \lambda)}{f_{\Lambda}(\lambda) \times f_{\Theta}(\theta)}$$

Alors

$$f_{X,\Theta,\Lambda}(x, \theta, \lambda) = f_{X|\Theta,\Lambda}(x | \theta, \lambda) \times f_{\Lambda}(\lambda) \times f_{\Theta}(\theta)$$

D'après (H4) et (H7), les $X_{i,t} | \theta_i, \Lambda$ avec $i = 1, 2, \dots, I$, sont conditionnellement indépendants et Θ_i est indépendant de Θ_j pour tout $i \neq j$, alors on peut en déduire que

$$f_{X,\Theta,\Lambda}(x, \theta, \lambda) = f_{\Lambda}(\lambda) \times \prod_{i=1}^I f_{\Theta_i}(\theta_i) \times \prod_{i=1}^I f_{X_i|\Theta_i,\Lambda}(x_i | \theta_i, \lambda)$$

Remarque 2.3 En utilisant la définition de la densité conditionnelle, nous avons

$$f_{\Theta,\Lambda|X}(\theta, \lambda | x) = \frac{f_{X,\Theta,\Lambda}(x, \theta, \lambda)}{f_X(x)} \quad (2.9)$$

D'après le lemme précédent, on a

$$f_{\Theta,\Lambda|X}(\theta, \lambda | x) = \beta \times f_{\Lambda}(\lambda) \times \prod_{i=1}^I f_{\Theta_i}(\theta_i) \times \prod_{i=1}^I f_{X_i|\Theta_i,\Lambda}(x_i | \theta_i, \lambda) \quad (2.10)$$

Avec

$$\beta = \frac{1}{f_X(x)} = [f_X(x)]^{-1} = \left[\int \int \dots \int f_{X,\Theta,\Lambda}(x, \theta, \lambda) \left(\prod_{i=1}^I d\theta_i \right) d\lambda \right]^{-1} \quad (2.11)$$

Où $\prod_{i=1}^I d\theta_i = d\theta_1 \times d\theta_2 \times \dots \times d\theta_I$. Nous indiquons maintenant ce résultat pour la densité conditionnelle souhaitée de $X_{j,T+1} | X$.

Théorème 2.1 Considérons le modèle à effets communs à deux niveaux qui satisfait les hypothèses de (H1) à (H8) ci-dessus. Pour un individu fixe $j = 1, 2, \dots, I$,

la densité conditionnelle de $X_{j,T+1} | X$ peut s'écrire

$$f_{X_{j,T+1}|x}(x_{j,T+1} | X) = \int \int \dots \int f_{X_{j,T+1}|\Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) \times f_{\Theta, \Lambda|X}(\theta, \lambda | x) \left(\prod_{i=1}^I d\theta_i \right) d\lambda \quad (2.12)$$

Preuve

Considérons un individu fixe j avec $1 \leq j \leq I$. D'après la définition de la densité conditionnelle, on a

$$f_{X_{j,T+1}|x}(x_{j,T+1} | X) = \frac{f_{X_{j,T+1}}(x_{j,T+1}, X)}{f_X(x)} \quad (2.13)$$

On note que le numérateur $f_{X_{j,T+1}}(x_{j,T+1}, X)$ peut s'écrire comme suit

$$\begin{aligned} f_{X_{j,T+1}}(x_{j,T+1}, X) &= \int \int \dots \int f_{X_{j,T+1}, X|\Theta_j, \Lambda}(x_{j,T+1}, X | \theta_j, \lambda) \times f_{\Theta, \Lambda}(\theta, \lambda) (d\theta_1, \dots, d\theta_I) d\lambda \\ &= \int \int \dots \int f_{X_{j,T+1}|\Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) \times f_{X|\Theta, \Lambda}(x | \theta, \lambda) \times f_{\Theta, \Lambda}(\theta, \lambda) \\ &\quad \times (d\theta_1, \dots, d\theta_I) d\lambda \end{aligned}$$

En se focalisant sur l'hypothèse (H6), on a

$$\begin{aligned} f_{X_{j,T+1}, X|\Theta, \Lambda}(x_{j,T+1}, X | \theta, \lambda) &= f_{X_{j,T+1}, X_j|\Theta_j, \Lambda}(x_{j,T+1}, X_j | \theta_j, \lambda) \prod_{\substack{i=1 \\ i \neq j}}^I f_{X_i|\Theta, \Lambda}(x_i | \theta, \lambda) \\ &= f_{X_{j,T+1}|\Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) \prod_{\substack{i=1 \\ i \neq j}}^I f_{X_i|\Theta, \Lambda}(x_i | \theta, \lambda) \\ &= f_{X_{j,T+1}|\Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) \times f_{X|\Theta, \Lambda}(x | \theta, \lambda) \end{aligned}$$

En remplaçant $f_{X_{j,T+1}}(x_{j,T+1}, X)$ par sa valeur dans l'équation (2.13), on obtient

$$\begin{aligned} f_{X_{j,T+1}|x}(x_{j,T+1} | X) &= \int \int \dots \int f_{X_{j,T+1}|\Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) \left(\frac{f_{X|\Theta, \Lambda}(x | \theta, \lambda) \times f_{\Theta, \Lambda}(\theta, \lambda)}{f_X(x)} \right) \\ &\quad \times (d\theta_1, \dots, d\theta_I) d\lambda \end{aligned}$$

Toujours en utilisant (H6), on peut déduire que

$$f_{\Theta,\Lambda|X}(\theta, \lambda | x) = \frac{f_{X|\Theta,\Lambda}(x | \theta, \lambda) \times f_{\Theta,\Lambda}(\theta, \lambda)}{f_X(x)}$$

D'où

$$f_{X_{j,T+1}|x}(x_{j,T+1} | X) = \int \int \dots \int f_{X_{j,T+1}|\Theta_j,\Lambda}(x_{j,T+1} | \theta_j, \lambda) \times f_{\Theta,\Lambda|X}(\theta, \lambda | x) \left(\prod_{i=1}^I d\theta_i \right) d\lambda$$

Remarque 2.4

Le théorème mentionné précédemment découle d'une expression explicite de la densité conditionnelle en fonction de toutes les informations disponibles. En premier lieu.

La densité conditionnelle est calculée en intégrant plusieurs fois (ou en sommant dans le cas discret) le produit de la densité conditionnelle $f_{X_{j,T+1}|\Theta_j,\Lambda}(x_{j,T+1} | \theta_j, \lambda)$ avec la densité conditionnelle $f_{\Theta,\Lambda|X}(\theta, \lambda | x)$, obtenue à l'aide de l'hypothèse (H6).

2.2.2 Application d'effets communs pour déterminer les primes de crédibilité

Dans cette partie, nous allons évaluer la prime de crédibilité lorsque les effets communs sont répartis de manière normale (effets communs normaux). En d'autres termes, nous allons recevoir des formulations intéressantes relatives à la somme de la réclamation anticipée ou la prime de crédibilité.

Formule de la prime de crédibilité

Notre modèle sera désigné sous le nom de modèle à effets communs à deux niveaux. Notre intérêt premier est de prévoir le sinistre dans la période qui suit, pour chaque individu en se basant de toutes les réclamations constatées X_1, X_2, \dots, X_I ou $X_i = (X_{j,1}, X_{j,2}, \dots, X_{j,T})'$. Cela est présenté par $E(X_{j,T+1} | X_1, X_2, \dots, X_I)$ qui est le

meilleur estimateur de $X_{j,T+1}$ sur la base de toutes les sinistres observés X_1, X_2, \dots, X_I au sens des moindres carrés.

Avant de considérer des cas particuliers de l'hypothèse d'effets communs normaux, considérons le cas général où nous avons des personnes assurées et où les effets communs ont des variances qui ne sont pas nécessairement unitaires. Pour réaliser la dérivation, nous faisons les hypothèses suivantes :

HS1) Les variables aléatoires $X_{i,t} \mid \theta_i, \Lambda$ sont des distributions normales de moyenne $(\theta_i + \lambda)$ et de variance σ_x^2 , c'est-à-dire $X_{i,t} \mid \theta_i, \Lambda \rightsquigarrow N(\theta_i + \lambda, \sigma_x^2)$, pour $i = 1, \dots, I$.

HS2) Les effets communs " individuels " θ_i sont des distributions normales de moyenne commune μ_θ et de variance commune σ_θ^2 , c'est à dire $\Theta_j \rightsquigarrow N(\mu_\theta, \sigma_\theta^2)$.

HS3) Les effets communs " généraux " λ sont des distributions normales de moyenne μ_λ et de variance σ_λ^2 , c'est à dire $\Lambda \rightsquigarrow N(\mu_\lambda, \sigma_\lambda^2)$.

Par la suite on a

$$f_{X_{i,t} \mid \Theta_i, \Lambda}(x_{i,t} \mid \theta_i, \lambda) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{x_{i,t} - (\theta_i + \lambda)}{\sigma_x} \right]^2} \quad (2.14)$$

$$f_{\Theta_i}(\theta_i) = \frac{1}{\sigma_\theta \sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{\theta_i - \mu_\theta}{\sigma_\theta} \right]^2} \quad (2.15)$$

$$f_{\Lambda}(\lambda) = \frac{1}{\sigma_\lambda \sqrt{2\pi}} e^{-\frac{1}{2} \left[\frac{\lambda - \mu_\lambda}{\sigma_\lambda} \right]^2} \quad (2.16)$$

Premièrement, dérivons la densité de $X_{j,T+1} \mid X_1, X_2, \dots, X_I$ en utilisant le théorème précédent. Pour ce faire, on peut supposer $j = 1$ dont les expressions sont similaires pour tous les autres individus. Alors pour $j = 1$ on a $X_{1,T+1} \mid X_1, X_2, \dots, X_I$ est une distribution normale de moyenne $\mu_{1,T+1}$ et de variance $\sigma_{1,T+1}^2$ c'est-à-dire $X_{1,T+1} \mid X_1, X_2, \dots, X_I \rightsquigarrow N(\mu_{1,T+1}, \sigma_{1,T+1}^2)$.

L'annexe page (98) montre également les expressions détaillées de la moyenne $\mu_{1,T+1}$ et de variance $\sigma_{1,T+1}^2$.

$$E(X_{1,T+1} | X_1, X_2, \dots, X_I) = \mu_{1,T+1} = \omega_1 \overline{X_1} + \omega_{i \neq 1} \overline{\overline{X_{i \neq 1}}} + \omega_{\theta, \lambda} (\mu_\theta + \mu_\lambda) \quad (2.17)$$

où $\overline{X_1} = \frac{1}{T} \sum_t X_{1,t}$ est la moyenne de l'échantillon observé de l'individu concerné pour $j = 1$ et $\overline{\overline{X_{i \neq 1}}} = \frac{1}{(I-1)T} \sum_{i \neq 1} \sum_t X_{i,t}$ est la moyenne de l'échantillon du reste des individus. Ainsi, d'après (2.17) on peut déduire que la prime de crédibilité est une moyenne pondérée des moyennes d'échantillons observées ainsi que les moyennes agrégées des effets communs. Comme démontré dans l'annexe de [40], les poids $\omega_1, \omega_{i \neq 1}$ et $\omega_{\theta, \lambda}$ ont les expressions suivantes :

1) Le poids lié à l'expérience propre de l'individu :

$$\omega_1 = \frac{T [(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T + \sigma_x^2 (\sigma_\theta^2 + \sigma_\lambda^2)]}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)}$$

2) Le poids lié à l'expérience du reste de portefeuille :

$$\omega_{i \neq 1} = \frac{T(I-1) \sigma_\lambda^2 \sigma_x^2}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)}$$

3) Le poids lié à l'expérience du passé :

$$\omega_{\theta, \lambda} = \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2)}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)}$$

Avec les poids, il devient simple de montrer que la prime de crédibilité en (2.17) est en fait une moyenne pondérée dont les pondérations totalisent un, c'est-à-dire

$$\omega_1 + \omega_{i \neq 1} + \omega_{\theta, \lambda} = 1$$

Bien que cela ait très peu de valeur pour notre objectif, à des fins d'exhaustivités, donnons la variance de $X_{1,T+1} | X_1, X_2, \dots, X_I$. En se basant sur l'annexe de [4], on a :

$$\text{Var}(X_{1,T+1} \mid X_1, X_2, \dots, X_T) = \sigma_{1,T+1}^2 \quad (2.18)$$

et

$$\sigma_{1,T+1}^2 = \frac{\sigma_x^2 [(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + T [\sigma_\lambda^2 I (\sigma_x^2 + \sigma_\theta^2) + \sigma_\theta^2 (2\sigma_x^2 + \sigma_\theta^2)] + \sigma_x^2 (\sigma_x^2 + \sigma_\theta^2 + \sigma_\lambda^2)]}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)} \quad (2.19)$$

De plus, il est important de noter que la prime de crédibilité est en fait un estimateur non biaisé pour la réclamation de la période suivante. On a

$$\begin{aligned} E(\mu_{1,T+1}) &= E \left[\omega_1 \overline{X}_1 + \omega_{i \neq 1} \overline{X}_{i \neq 1} + \omega_{\theta, \lambda} (\mu_\theta + \mu_\lambda) \right] \\ &= \omega_1 E(\overline{X}_1) + \omega_{i \neq 1} E(\overline{X}_{i \neq 1}) + \omega_{\theta, \lambda} (\mu_\theta + \mu_\lambda) \end{aligned}$$

Alors

$$E(\mu_{1,T+1}) = \mu_\theta + \mu_\lambda \quad (2.20)$$

Propriétés asymptotiques

On peut extraire quelques propriétés asymptotiques intéressantes des formules de prime de crédibilité développées dans la sous-section précédente. Pour étudier ces caractéristiques asymptotiques, nous allons nous intéresser à la formule (2.17), où la prime est la somme pondérée des moyennes de l'échantillon observé d'un individu et du reste des individus. Il est possible de faire des observations similaires intéressantes en se concentrant sur la formule de la prime de crédibilité dans (2.17). On peut résumer ces propriétés asymptotiques de la manière suivante :

P1) En termes d'expérience en matière de réclamations, si $T \rightarrow 0$, c'est-à-dire que l'expérience passée est manquante pour tous les individus, alors $\omega_1 \rightarrow 0$, $\omega_{i \neq 1} \rightarrow 0$ et $\omega_{\theta, \lambda} \rightarrow 1$. Plus l'assureur a d'expérience disponible pour évaluer l'expérience future des sinistres, plus il a un impact sur le poids à priori.

P2) La richesse du passé de l'expérience personnelle. À la différence de la propriété précédente, si $T \rightarrow \infty$, c'est-à-dire qu'il existe une abondance du passé de

l'expérience pour tous les individus, on obtient $\omega_1 \rightarrow 1$, $\omega_{i \neq 1} \rightarrow 0$ et $\omega_{\theta, \lambda} \rightarrow 0$.

P3) La richesse de l'expertise du groupe. Lorsque nous laissons $I \rightarrow \infty$, c'est-à-dire qu'il y a une grande quantité d'expérience de groupe, nous observons que

$\omega_1 \rightarrow \frac{\sigma_\theta^2 T}{\sigma_x^2 + \sigma_\theta^2 T}$, $\omega_{i \neq 1} \rightarrow \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2 T}$ et $\omega_{\theta, \lambda} \rightarrow 0$. Dans ce cas, notre prime de crédibilité devient

$$\mu_{1, T+1} = \frac{\sigma_\theta^2 T}{\sigma_x^2 + \sigma_\theta^2 T} \bar{X}_1 + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2 T} \bar{\bar{X}}_{i \neq 1} \quad (2.21)$$

Avec le facteur de crédibilité

$$Z = \frac{\sigma_\theta^2 T}{\sigma_x^2 + \sigma_\theta^2 T} = \frac{T}{T + \left(\frac{\sigma_x^2}{\sigma_\theta^2} \right)} \quad (2.22)$$

Cette prime de crédibilité est autonome par rapport à la diversité des événements communs « globaux ». En effet, les échanges globaux communs ont été éliminés en raison de l'abondance de l'expérience de groupe. En outre, si d'une part, $\sigma_\theta^2 \rightarrow \infty$, c'est-à-dire que les effets communs « individuels » deviennent très variables, alors nous devrions simplement accorder une crédibilité totale à l'individu. En d'autres termes, $\omega_1 \rightarrow 1$ lorsque $\sigma_\theta^2 \rightarrow \infty$ et $I \rightarrow \infty$. En revanche, si $\sigma_x^2 \rightarrow \infty$ avec $I \rightarrow \infty$, c'est-à-dire que la variabilité des réclamations de l'individu est trop importante pour pouvoir faire une prédiction significative, nous n'ajouterions que l'expérience des autres individus du groupe sera prise en compte.

P4) Une large diversité de demandes individuelles : Il est observé que la variabilité des réclamations individuelles augmente, c'est-à-dire que $\sigma_x^2 \rightarrow \infty$, ce qui signifie que $\omega_1 \rightarrow 0$, $\omega_{i \neq 1} \rightarrow 0$, $\omega_{\theta, \lambda} \rightarrow 1$ et que la prime de crédibilité devient $\mu_{1, T+1} = \mu_\theta + \mu_\lambda$.

P5) Le paramètre de risque individuel présente une grande variabilité : Si $\sigma_\theta^2 \rightarrow \infty$, c'est-à-dire que le paramètre de risque individuel présente une plus grande variabilité, nous avons $\omega_1 \rightarrow 1$, $\omega_{i \neq 1} \rightarrow 0$ et $\omega_{\theta, \lambda} \rightarrow 0$.

Toutes les charges relèvent de l'expérience personnelle de l'individu. Cela est intuitif, car un θ très variable réduit la fiabilité de $\bar{\bar{X}}_{i \neq 1}$ et de μ_θ en tant que mesure de l'étendue du risque causé par l'individu 1.

P6) Un important changement dans le paramètre de risque global : S'il y a une

plus grande variabilité du paramètre de risque global, c'est-à-dire $\sigma_\lambda^2 \rightarrow \infty$, nous avons

$$\omega_1 \rightarrow \frac{\sigma_\theta^2 IT + \sigma_x^2}{\sigma_\theta^2 IT + \sigma_x^2 I}, \omega_{i \neq 1} \rightarrow \frac{(I-1)\sigma_x^2}{\sigma_\theta^2 IT + \sigma_x^2 I} \text{ et } \omega_{\theta, \lambda} \rightarrow 0$$

i) Si en plus, $T \rightarrow \infty$ ou $\sigma_\theta^2 \rightarrow \infty$, Alors $\omega_1 \rightarrow 1$ et $\omega_{i \neq 1} \rightarrow 0$.

ii) Si en plus, $I \rightarrow \infty$, alors $\omega_1 \rightarrow \frac{\sigma_\theta^2 T}{\sigma_x^2 + \sigma_\theta^2 T}$ et $\omega_{i \neq 1} \rightarrow \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2 T}$.

iii) Si en plus, $\sigma_x^2 \rightarrow \infty$, alors $\omega_1 \rightarrow \frac{1}{I}$ et $\omega_{i \neq 1} \rightarrow \frac{I-1}{I}$.

Il est évident que la prime de crédibilité $\mu_{1, T+1} = \overline{\overline{X}}$, qui correspond à une moyenne globale.

*Voir l'annexe pour les démonstrations de ces différentes propriétés.

Deux risques individuels avec écart d'unité

Deux risques individuels ($I = 2$) sont maintenant pris en compte, avec des effets communs à la fois individuels et globaux et des écarts unitaires ($\sigma_\lambda^2 = \sigma_\theta^2 = \sigma_x^2 = 1$). Il est facile de démontrer que dans cette situation, la prime de crédibilité se résume à la formule suivante

$$E(X_{1, T+1} | X_1, X_2, \dots, X_I) = \mu_{1, T+1} = \omega_1 \overline{X_1} + \omega_2 \overline{X_2} + \omega_{\theta, \lambda} (\mu_\theta + \mu_\lambda) \quad (2.23)$$

Où $\overline{X_1}$ et $\overline{X_2}$ indiquent les moyennes d'échantillons observés à partir du premier et du deuxième individu, respectivement. Les poids de crédibilité sont

$$\omega_1 = \frac{T(3T+2)}{(3T+1)(T+1)} = 1 - \frac{1}{2(T+1)} - \frac{1}{2(3T+1)}$$

$$\omega_2 = \frac{T}{(3T+1)(T+1)} = \frac{1}{2(T+1)} - \frac{1}{2(3T+1)}$$

$$\omega_{\theta, \lambda} = \frac{T+1}{(3T+1)(T+1)} = \frac{1}{(3T+1)}$$

De plus, on peut exprimer la variance de la manière suivante

$$\text{Var}(X_{1,T+1} | X_1, X_2, \dots, X_I) = \sigma_{1,T+1}^2 = \frac{3T^2 + 7T + 3}{(3T + 1)(T + 1)} = 1 + \frac{3}{2(3T + 1)} + \frac{1}{2(T + 1)} \quad (2.24)$$

Ces résultats montrent que la prime de crédibilité est constante en $\overline{X_1}$, $\overline{X_2}$ et $(\mu_\theta + \mu_\lambda)$. Les poids sont également réduits à l'unité, c'est-à-dire $\omega_1 + \omega_2 + \omega_{\theta,\lambda} = 1$. Il convient de souligner que lorsque l'expérience personnelle est élevée, c'est-à-dire lorsque $T \rightarrow \infty$, on a $\omega_1 \rightarrow 1$, $\omega_2 \rightarrow 0$, $\omega_{\theta,\lambda} \rightarrow 0$, ce qui nous amène à une dépendance croissante envers l'expérience personnelle au détriment de l'ignorance des deux connaissances antérieures : l'expérience des personnes touchées et l'expérience d'une autre personne. Par ailleurs, si $T \rightarrow 0$, on a $\omega_1 \rightarrow 0$, $\omega_2 \rightarrow 0$, $\omega_{\theta,\lambda} \rightarrow 1$, ce qui signifie que les connaissances préalables des revendications sont de plus en plus prises en compte et que les expériences individuelles sont négligées.

Plusieurs risques individuels avec écart d'unité

Dans le cas particulier où $I > 2$ individus, avec toujours des variances unitaires ($\sigma_\lambda^2 = \sigma_\theta^2 = \sigma_x^2 = 1$), on peut prouver que la prime de crédibilité peut se représenter de la façon suivante

$$E(X_{1,T+1} | X_1, X_2, \dots, X_I) = \mu_{1,T+1} = \omega_1 \overline{X_1} + \omega_{i \neq 1} \overline{\overline{X_{i \neq 1}}} + \omega_{\theta,\lambda} (\mu_\theta + \mu_\lambda)$$

Où $\overline{X_1}$ et $\overline{\overline{X_{i \neq 1}}}$ = $\frac{1}{(I-1)T} \sum_{i=2}^I \sum_{t=1}^T X_{i,t}$ sont les moyennes habituelles des échantillons observés. Les poids de crédibilité sont

$$\omega_1 = \frac{T[(I+1)T+2]}{[(I+1)T+1](T+1)} = 1 - \frac{(I-1)}{I(T+1)} - \frac{1}{I[(I+1)T+1]}$$

$$\omega_{i \neq 1} = \frac{T(I-1)}{[(I+1)T+1](T+1)} = \frac{(I-1)}{I(T+1)} - \frac{I-1}{I[(I+1)T+1]}$$

$$\omega_{\theta,\lambda} = \frac{T+1}{[(I+1)T+1](T+1)} = \frac{1}{[(I+1)T+1]}$$

En outre, on peut exprimer la variance de la manière suivante

$$Var(X_{1,T+1} | X_1, X_2, \dots, X_I) = \sigma_{1,T+1}^2 = \frac{(I+1)T^2 + (2I+3)T + 3}{[(I+1)T+1](T+1)} \quad (2.25)$$

Comme précédemment, la prime de crédibilité est linéaire en $\overline{X_1}$, $\overline{\overline{X_{i \neq 1}}}$ et $(\mu_\theta + \mu_\lambda)$. Les poids ω_1 , $\omega_{i \neq 1}$ et $\omega_{\theta,\lambda}$ sont représentés par l'unité ($\omega_1 + \omega_{i \neq 1} + \omega_{\theta,\lambda} = 1$). Il est important de noter que lorsqu'il y a une grande expérience individuelle, c'est-à-dire lorsque $T \rightarrow \infty$, on a $\omega_1 \rightarrow 1$, $\omega_{i \neq 1} \rightarrow 0$, $\omega_{\theta,\lambda} \rightarrow 0$, ce qui signifie qu'il y a une dépendance croissante à l'expérience individuelle en raison des coûts liés à la fois à la connaissance antérieure de l'expérience des sinistres et à l'expérience d'un autre individu. En revanche, lorsque l'expérience individuelle $T \rightarrow 0$, on observe $\omega_1 \rightarrow 0$, $\omega_{i \neq 1} \rightarrow 0$, $\omega_{\theta,\lambda} \rightarrow 1$, ce qui signifie que l'expérience individuelle est de plus en plus dépendante du savoir passé sur les sinistres et que l'expérience individuelle est négligée.

Par ailleurs, lorsque le nombre de risques individuels augmente et tend vers l'infini ($I \rightarrow \infty$), on a $\omega_1 \rightarrow \frac{T}{T+1}$, $\omega_{i \neq 1} \rightarrow \frac{1}{T+1}$ et $\omega_{\theta,\lambda} \rightarrow 0$, ce qui nous amène à une dépendance croissante envers l'expérience individuelle, notamment celle de l'individu, et à l'ignorance des connaissances préalables des revendications.

2.2.3 Modèle avec dépendance simultanée

Prenons l'exemple d'un ensemble de contrats d'assurance J où chaque contrat individuel possède un vecteur de réclamations observables à des périodes T notées par

$$X'_j = (X_{j,1}, X_{j,2}, \dots, X_{j,T}, X_{j,T+1}) \text{ pour } j = 1, 2, \dots, J$$

Dans tous les contrats, le vecteur $X' = (X_1, X_2, \dots, X_J)$ est défini comme le vec-

teur de réclamations observable. X peut être considéré comme une matrice aléatoire d'ordre $J * T$.

Si l'on considère un seul risque individuel j , il y a une corrélation entre les réclamations à différentes périodes. Dans cette partie, on propose de développer ces constructions lorsqu'il existe également une dépendance entre les individus. La structure de dépendance entre les individus sera modélisée en utilisant des effets communs similaires conditionnés par Θ . Pour cela, nous supposerons que les X_j , sont indépendants, ce qui nous permettra d'obtenir

$$f_X(X) = \int_{-\infty}^{+\infty} f_{\Theta}(\theta) \times f_{X|\Theta}(X | \theta) d\theta = \int_{-\infty}^{+\infty} f_{\Theta}(\theta) \prod_{j=1}^J f_{X_j|\Theta}(X_j | \theta) d\theta \quad (2.26)$$

La structure de dépendance est également présentée par une fonction de copule notée C pour le même individu j à différentes périodes de temps, $t = 1, \dots, T$. Les copules ont la capacité de capter une variété et diverses formes de dépendance. Dans cette optique, l'ajout de copules à notre modèle peut améliorer la méthode d'éta-lonnage et conduire à des prédictions plus précises. Afin d'être plus pratique, il est possible de supposer que la même copule C est valable pour tous les individus afin d'obtenir

$$F_{X_j|\Theta}(X_j | \theta) = C(u_{j|\theta}), \text{ pour } j = 1, 2, \dots, J$$

où $u_{j|\theta} = (u_{j,1|\theta}, u_{j,2|\theta}, \dots, u_{j,T|\theta})'$ et $u_{j,t|\theta} = F_{X_{j,t}|\Theta}(x_{j,t} | \theta)$. maintenant, en réalisant des dérivées partielles par rapport à $x_{j,1}, x_{j,2}, \dots, x_{j,T}$, on trouve la densité conditionnelle associée suivante

$$f_{X_j|\Theta}(X_j | \theta) = C(u_{j|\theta}) \prod_{t=1}^T f_{X_{j,t}|\Theta}(x_{j,t} | \theta) \quad (2.27)$$

$C(u_{j|\theta})$ est la densité conditionnelle du copule C avec $C(u_{j|\theta}) = \frac{\partial^j C(u_{j|\theta})}{\partial u_{j|\theta}}$ pour $j = 1, 2, \dots, J$, on a

$$f_X(X) = \int_{-\infty}^{+\infty} f_{\Theta}(\theta) \times \prod_{j=1}^J \left(C(u_{j|\theta}) \prod_{t=1}^T f_{X_{j,t}|\Theta}(x_{j,t} | \theta) \right) d\theta \quad (2.28)$$

Examinons le calcul de la réclamation prévue $E(X_{j,T+1} \mid X = x)$ pour chaque contrat $j = 1, 2, \dots, J$. Observons le vecteur aléatoire ci-dessous

$$X_j^{(T+1)'} = (X_{j,1}, X_{j,2}, \dots, X_{j,T}, X_{j,T+1}) \text{ pour } j = 1, 2, \dots, J$$

Quand la réclamation observable augmente au cours de la période suivante, le montant total correspond au vecteur aléatoire total est

$$X^{(j,T+1)'} = (X_1, X_2, \dots, X_j^{T+1}, \dots, X_J) \text{ pour } j = 1, 2, \dots, J$$

Par la suite, en suivant des étapes semblables à celles décrites dans le développement de l'expression (2.28), on observe que pour un individu spécifique, on a

$$f_{X^{(i,T+1)}}(x^{(i,T+1)}) = \int_{-\infty}^{+\infty} f_{\Theta}(\theta) C(u_{j|\theta}) \times \prod_{t=1}^T f_{X_i|\Theta}(x_i \mid \theta) \times \prod_{\substack{j=1 \\ j \neq i}}^J \left(C(u_{j|\theta}) \prod_{t=1}^T f_{X_j|\Theta}(x_j \mid \theta) \right) d\theta \quad (2.29)$$

En se basant sur les résultats intermédiaires mentionnés précédemment, on peut donc représenter la densité prédictive des sinistres comme le rapport entre les équations (2.29) et (2.28)

$$f_{X_{i,T+1}|X}(x_{i,T+1} \mid x) = \frac{f_{X^{(i,T+1)}}(x^{(i,T+1)})}{f_X(X)} \quad (2.30)$$

Ainsi, nos réclamations prédictives peuvent être évaluées avec l'utilisation de l'espérance conditionnelle.

$$E(X_{i,T+1} \mid X = x) = \int_{-\infty}^{+\infty} x_{i,T+1} \times f_{X_{i,T+1}|X}(x_{i,T+1} \mid x) dx_{i,T+1}$$

En remplaçant $f_{X_{i,T+1}|X}(x_{i,T+1} \mid x)$ par sa valeur, on obtient

$$E(X_{i,T+1} \mid X = x) = \int_{-\infty}^{+\infty} x_{i,T+1} \left(\frac{f_{X^{(i,T+1)}}(x^{(i,T+1)})}{f_X(X)} \right) dx_{i,T+1} \quad (2.31)$$

En raison de l'intégration dans (2.29), il n'est pas possible d'isoler la densité marginale de $X_{i,T+1}$ dans (2.31) pour pouvoir l'écrire comme une espérance conditionnelle, comme on l'a vu précédemment.

2.3 Dépendance et copules

La théorie des copules a été développée par Abe Sklar en 1959, à partir des travaux de F Rechet dans les années 50. Les copules représentent des fonctions de répartition de vecteurs aléatoires multidimensionnelles avec des marginales uniformes sur $[0;1]$. Les fonctions de dépendance sont également désignées par Deheuvels en 1979. Dans les applications statistiques à grande échelle, les copules sont largement employés car elles offrent la possibilité de modéliser et d'estimer aisément la répartition de vecteurs aléatoires en estimant les marginaux et les copules individuellement. En effet, cette théorie a prouvé sa validité dans la modélisation des distributions à plusieurs variables dans le secteur de la finance et de l'assurance.

Les copules permettent d'analyser la dépendance entre les variables aléatoires, en excluant toute information provenant des lois marginales de ces variables.

Dans ce travail, nous incorporons la structure de dépendance de copule dans l'estimateur des réclamations, ce qui peut aboutir à d'autres formes de formules de crédibilité explicites (linéaires). Il convient de noter que la prime de crédibilité peut ne plus être linéaire par rapport au sinistre passé, mais plutôt être de forme non linéaire. Par conséquent, il n'est peut-être pas possible d'avoir la forme de crédibilité linéaire à laquelle nous avons été habitués. Il est bien connu en statistique que le meilleur prédicteur de Z_{T+1} basé sur toutes les affirmations observées de Z_1, \dots, Z_T , au sens des moindres carrés est l'espérance conditionnelle

$$E(Z_{T+1} \mid Z_T = z_T) \tag{2.32}$$

En d'autres termes, le prédicteur dans (2.32) est la fonction requise $g(Z_T)$ qui

minimise l'erreur de prédiction quadratique moyenne suivante :

$$E[Z_{T+1} - g(Z_T)]^2 \quad (2.33)$$

Voir, par exemple, Shao (2003) à la page 40 pour une belle preuve. Ce même résultat est bien connu dans la littérature actuarielle et d'assurance, par exemple, voir Dannenburg, Kaas et Goovaerts (1996).

En procédant maintenant, la densité conditionnelle $f_{Z_{T+1}|Z_T}(z_{T+1} | Z_T)$ est souhaitée car l'espérance conditionnelle en (2.32) donne notre meilleure estimation des réclamations de la période suivante et donne également notre prime de crédibilité souhaitée. La densité conditionnelle $f_{Z_{T+1}|Z_T}(z_{T+1} | Z_T)$ a été appelée dans Frees et Wang (2005) la densité prédictive et nous supposons qu'elle existe. Les réclamations estimées peuvent être commodément exprimées par

$$E(Z_{T+1} | Z_T = z_T) = \int_{-\infty}^{+\infty} z_{T+1} \times f_{Z_{T+1}|Z_T}(z_{T+1} | z_T) dz_{T+1} \quad (2.34)$$

où l'intégrale est l'intégrale de Riemann-Stieltjes. Notez que les limites de l'intégrale sont la ligne réelle entière.

L'objectif de ce travail est de construire des prédicteurs pour la réclamation de la période suivante d'un contrat individuel sur la base de toutes ses réclamations observées Z_1, \dots, Z_T , pour laquelle la réclamation de la période suivante doit être désignée par la variable aléatoire Z_{T+1} . Comme indiqué précédemment, nous supposons qu'il n'y a de dépendance que sur des périodes de temps pour le même individu et cette structure de dépendance est spécifiée avec une fonction de copule.

2.3.1 Définition d'une copule

Définition 2.6 On appelle une copule, chaque fonction de répartition à plusieurs variables C_T dans l'hypercube $[0, 1]^T$ avec des marginales de loi continue uniforme sur $[0, 1]$.

Définition 2.7 Soit un vecteur a dans $[0, 1]^T$ et C_T une fonction de $[0, 1]^T$ vers

2.3. Dépendance et copules

$[0, 1]$. C_T est une copule si elle vérifie les trois propriétés suivantes :

1. $C_T(a_1, \dots, a_T)$ est croissante au sens large en chaque composante a_i .
2. Pour tout vecteur a dans $[0, 1]^T$, $C_T(a) = 0$ si une (ou plus) des composantes de a est égale à 0 et $C_T(a) = a_i$ si les composantes de a sont toutes égales à 1, excepté a_i .
3. Pour tous (a_1^1, \dots, a_T^1) et $(a_1^2, \dots, a_T^2) \in [0, 1]^T$ tels que $a_i^1 \leq a_i^2$ avec $i \in \{1, \dots, T\}$, alors on a

$$\sum_{k_1=1}^2 \dots \sum_{k_T=1}^2 (-1)^{k_1+\dots+k_T} \times C_T(a_1^{k_1}, \dots, a_T^{k_T}) \geq 0 \quad (2.35)$$

Définition 2.8 Une copule bivarié (2-dimensions) est une fonction $C_2 : [0, 1]^2 \rightarrow [0, 1]$ telle que

a)

$$\forall (a, b) \in [0, 1]^2 \quad C_2(a, 0) = C_2(0, b) = 0 \quad (2.36)$$

b)

$$\forall (a, b) \in [0, 1]^2 \quad \begin{cases} C_2(a, 1) = a \\ C_2(1, b) = b \end{cases} \quad (2.37)$$

c) $\forall a_1, a_2, b_1, b_2 \in [0, 1]$ avec $a_2 \geq a_1$ et $b_2 \geq b_1$, on a

$$C_2(a_2, b_2) - C_2(a_2, b_1) - C_2(a_1, b_2) + C_2(a_1, b_1) \geq 0 \quad (2.38)$$

Remarque 2.5

Prenons A_1 et A_2 comme deux variables aléatoires uniformes sur $[0, 1]$, on a :

$$C_2(a, b) = P(A_1 \leq a, A_2 \leq b) \quad \forall (a, b) \in [0, 1]^2 \quad (2.39)$$

Une copule peut également être déterminée à l'aide d'une loi multivariée préexistante. Le théorème de Sklar précise alors le lien établi par la copule C_T , entre les fonctions de répartition marginales univariées F_1, F_2, \dots, F_T et la distribution complète multivariée F .

Exemple 2.1

Soit une copule produit notée \prod et définie par :

$$C_2(a, b) = \prod(a, b) = a.b \quad \forall (a, b) \in [0, 1]^2 \quad (2.40)$$

-Vérifions la condition (a) :

$$C_2(a, 0) = \prod(a, 0) = a.0 = 0 \quad \text{et} \quad C_2(0, b) = \prod(0, b) = 0.b = 0$$

-Vérifions la condition (b) :

$$C_2(a, 1) = \prod(a, 1) = a.1 = a \quad \text{et} \quad C_2(1, b) = \prod(1, b) = 1.b = b$$

-Vérifions pour l'équation (c) :

$$\begin{aligned} & C_2(a_2, b_2) - C_2(a_2, b_1) - C_2(a_1, b_2) + C_2(a_1, b_1) \geq 0 \\ \Rightarrow & \prod(a_2, b_2) - \prod(a_2, b_1) - \prod(a_1, b_2) + \prod(a_1, b_1) \geq 0 \\ \Rightarrow & a_2.b_2 - a_2.b_1 - a_1.b_2 + a_1.b_1 \geq 0 \\ \text{si } a_2 \geq a_1 \Rightarrow & \begin{cases} a_2.b_2 \geq a_1.b_2 \\ a_2.b_1 \geq a_1.b_1 \end{cases} \end{aligned}$$

$$a_2.b_2 + a_2.b_1 \geq a_1.b_2 + a_1.b_1 \Rightarrow a_2.b_2 - a_2.b_1 - a_1.b_2 + a_1.b_1 \geq 0$$

D'où \prod est une copule.

Exemple 2.2

$\forall a, b \in [0, 1]$, la fonction $M(a, b) = \min(a, b)$ est une copule.

En effet :

$$- \forall (a_1, a_2) \in [0, 1]^2, \min(a_1, 0) = \min(0, a_2) = 0 \Rightarrow M \text{ vérifie (1).}$$

$$- \forall (a_1, a_2) \in [0, 1]^2, \min(a_1, 1) = a_1, \min(1, a_2) = a_2 \Rightarrow M \text{ vérifie (2).}$$

- $\forall (a_1, b_2) \in [0, 1]^2, (a_1, b_2) \in [0, 1]^2$ avec $a_1 \leq a_2$ et $b_1 \leq b_2$, nous avons les cas suivantes :

Si $a_1 \leq a_2 \leq b_1 \leq b_2$ on a

$$\min(a_2, b_2) - \min(a_2, b_1) - \min(a_1, b_2) + \min(a_1, b_1) = a_2 - a_2 - a_1 + a_1 = 0$$

Si $b_1 \leq b_2 \leq a_1 \leq a_2$ on a

$$\min(a_2, b_2) - \min(a_2, b_1) - \min(a_1, b_2) + \min(a_1, b_1) = b_2 - b_1 - b_2 + b_1 = 0$$

Si $a_1 < b_1 < a_2 < b_2$ on a

$$\min(a_2, b_2) - \min(a_2, b_1) - \min(a_1, b_2) + \min(a_1, b_1) = a_2 - b_1 - a_1 + a_1 \geq 0$$

Donc, $M(a, b)$ est une copule.

2.3.2 Propriétés des copules

Théorème 2.2 (Continuité)

- Cas multivarié

Soit C_T une copule, $\forall a, b \in [0, 1]^T$

$$|C_T(a) - C_T(b)| \leq \sum_{k=1}^T |a_k - b_k| \quad (2.41)$$

- Cas bivarié

C est une copule à 2-dimensions, pour tout $a_1, a_2, b_1, b_2 \in [0, 1]$ avec $a_2 \geq a_1$ et $b_2 \geq b_1$, on a :

$$|C_2(a_1, a_2) - C_2(b_1, b_2)| \leq |a_2 - a_1| + |b_2 - b_1| \quad (2.42)$$

Théorème 2.3 (Différentiabilité) C est une copule $\forall b \in [0, 1]$, il existe sûrement une dérivé partielle $\frac{\partial C}{\partial a}$ pour tout $a \in [0, 1]$, alors on a :

$$0 \leq \frac{\partial}{\partial a} C_2(a, b) \leq 1 \quad (2.43)$$

Et pour chaque $a \in [0, 1]$, la dérivé partielle $\frac{\partial C}{\partial b}$ existe sûrement pour tout $b \in [0, 1]$, alors on a :

$$0 \leq \frac{\partial}{\partial b} C_2(a, b) \leq 1 \quad (2.44)$$

En outre, les fonctions $a \rightarrow \frac{\partial}{\partial b} C_2(a, b)$ et $b \rightarrow \frac{\partial}{\partial a} C_2(a, b)$ sont croissantes quasiment partout sur $[0, 1]$.

2.3.3 Théorème de sklar

Le théorème est nommé en l'honneur de l'inventeur du concept des copules en 1959. Ce théorème explique la manière dont la fonction de répartition bivariée et les lois marginales peuvent être liées au concept de copule.

Définition 2.9 (inverse généralisé) Supposons que F soit une fonction de répartition à une seule variable z , le quantile de F est une fonction symbolisé F^{-1} appartient au domaine de définition $]0, 1[$, telle que :

$$F^{-1}(t) = \inf\{z \in \mathbb{R} : F(z) \geq t\} \quad (2.45)$$

Proposition 2.1

F^{-1} est une fonction croissante et continue à gauche. telle que $\forall z \in \mathbb{R}$ et $m \in [0, 1]$ on a

$$F(z) \geq m \Rightarrow z \geq F^{-1}(m)$$

- Supposons que Z soit une variable aléatoire dont les valeurs de la fonction de répartition F sont dans \mathbb{R} . Lorsque F est continue, la loi de $F(Z)$ est uniforme sur $[0, 1]$.

Théorème 2.4 (sklar) F est une fonction de répartition (à T-dimensions) avec des marginales F_1, F_2, \dots, F_T , il existe une copule C_T dont pour tout z de \mathbb{R}^T ,

$$F(z_1, \dots, z_T) = C_T(F_1(z_1), F_2(z_2), \dots, F_T(z_T)). \quad (2.46)$$

De plus si les F_i sont continues alors C_T est unique et on dit que C_T est la copule de (z_1, \dots, z_T) ou de F .

Théorème 2.5 (sklar bivarié) F est une fonction de répartition (à 2-dimensions) de marges F_1 et F_2 . La copule liée à F s'écrit, pour tout $(a_1, a_2) \in [0, 1]^2$ et pour tout $(z_1, z_2) \in \mathbb{R}^2$, supposons $a_1 = F_1(z_1)$ et $a_2 = F_2(z_2)$ on a :

$$\begin{aligned} C_2(a_1, a_2) &= C_2(F_1(z_1), F_2(z_2)) \\ C_2(a_1, a_2) &= F(F_1^{-1}(a_1), F_2^{-1}(a_2)) \\ C_2(a_1, a_2) &= F(z_1, z_2) \end{aligned}$$

Et

$$F(z_1, z_2) = F(F_1(z_1), F_2(z_2)) \quad (2.47)$$

C est unique lorsque les marges F_1 et F_2 sont continues, au contraire, elle est déterminée de façon distincte sur $Im(F_1) \times Im(F_2)$. En revanche, si la copule C_2 est associée aux fonctions de répartition F_1 et F_2 à une seule variable, la fonction F dans (2.47) représente la fonction de répartition associée aux marginales F_1 et F_2 .

Preuve. Pour tout $(a_1, a_2) \in [0, 1]^2$ et pour tout $(z_1, z_2) \in \mathbb{R}^2$, d'après la proposition précédente, $F_1(z_1)$ et $F_2(z_2)$ sont uniformes sur $[0, 1]$ Lorsque F_1 et F_2 sont continues. De plus, F est la fonction de répartition de $F_1(z_1)$ et $F_2(z_2)$. En appliquant le principe de l'équation (2.39), on trouve :

$$\begin{aligned} C_2(a_1, a_2) &= P(F_1(z_1) \leq a_1, F_2(z_2) \leq a_2) \\ C_2(a_1, a_2) &= p(z_1 \leq F_1^{-1}(a_1), z_2 \leq F_2^{-1}(a_2)) \\ C_2(a_1, a_2) &= F(F_1^{-1}(a_1), F_2^{-1}(a_2)) \\ C_2(a_1, a_2) &= F(z_1, z_2) \end{aligned}$$

D'où

$$F(z_1, z_2) = F(F_1(z_1), F_2(z_2))$$

Corollaire 2.1 Soient $F : \{F_1, \dots, F_T\}$, et C_T les fonctions définies dans le théorème de Sklar et soient $F_1^{-1}, \dots, F_T^{-1}$ les inverses généralisées de F_1, \dots, F_T respectivement. Alors

$$C_T(a_1, \dots, a_T) = F(F_1^{-1}(a_1), \dots, F_T^{-1}(a_T)) \quad \forall a \in [0, 1] \quad (2.48)$$

Avec les lois marginales inverses généralisées, comme indiquées par (2.45).

Théorème 2.6 (Bornes de Fréchet)

Supposons que F soit une fonction de répartition jointe d'un couple aléatoire (U, V) de fonctions de répartition marginales F_1 et F_2 . Pour toute copule à 2-dimensions (bivariée) C liée à F et $\forall (a, b) \in [0, 1]^2$, on a

$$W(a, b) = \max(a + b - 1, 0) \leq C(a, b) \leq \min(a, b) = M(a, b) \quad (2.49)$$

Preuve

$$\text{On a } \forall a, b \in [0, 1] \quad \begin{cases} C(a, b) \leq C(a, 1) = a \\ C(a, b) \leq C(1, b) = b \end{cases} \Rightarrow C(a, b) \leq \min(a, b) = M(a, b)$$

D'après la définition (2.7) de copules et $\forall a, b \in [0, 1]$, on a

$$C(a, b) \geq C(a, 1) + C(1, b) - C(1, 1) \Rightarrow C(a, b) \geq (a + b - 1)$$

Où $\forall a, b \in [0, 1]$, et $C(a, b) \geq 0$

$$\Rightarrow C(a, b) \geq \max(a + b - 1, 0) = W(a, b)$$

Donc

$$W(a, b) \leq C(a, b) \leq M(a, b)$$

2.3.4 Densité d'une copule

Définition 2.10 La densité c_T d'une copule C_T , est définie par

$$c_T(a_1, \dots, a_T) = \frac{\partial^T}{\partial a_1 \dots \partial a_T} C_T(a_1, \dots, a_T) \quad (2.50)$$

2.3. Dépendance et copules

Où $c : [0.1]^T \longrightarrow \mathbb{R}^+$

Selon le théorème de Sklar, la distribution jointe F est absolument continue. On peut déterminer la densité d'une copule aléatoire (Z_1, \dots, Z_T) en se basant sur ses marginales f_1, f_2, \dots, f_T comme suit

$$f(z_1, \dots, z_T) = c_T(F_1(z_1), \dots, F_T(z_T)) \cdot f_1(z_1), f_2(z_2), \dots, f_T(z_T) \quad (2.51)$$

Exemple 2.3 Soit une copule produit notée Π et définie par

$$C_2(a, b) = \Pi(a, b) = a.b \quad \forall (a, b) \in [0.1]^2$$

En appliquant l'équation (2.50) dans le cas à 2-dimensions, on obtient

$$c_2(a, b) = \frac{\partial^2}{\partial a \partial b} C_2(a, b) = \frac{\partial^2}{\partial a \partial b} \Pi(a, b) = 1$$

Comme $C_2(a, b) = 1$, d'après la formule (2.51), la densité de copule est

$$f(z_1, z_2) = f_1(z_1), f_2(z_2)$$

Par conséquent, la copule produit est la caractéristique de l'indépendance entre deux variables aléatoires.

Proposition 2.2

L'espérance conditionnelle de $Z_{T+1} \mid Z_T$ peut être exprimée par :

$$E(Z_{T+1} \mid Z_T = z_T) = \int_{-\infty}^{+\infty} z_{T+1} \frac{c_{T+1}(a_{T+1})}{c_T(a_T)} dF_{T+1}(z_{T+1}) \quad (2.52)$$

où $c_{T+1}(a_{T+1})$ et $c_T(a_T)$ sont respectivement définies à l'aide de l'équation (2.50) et F_{T+1} est la fonction de distribution marginale de Z_{T+1} .

Preuve

En se basant sur la définition classique de la densités conditionnelle on a :

$$f_{Z_{T+1}|Z_T}(z_{T+1} | z_T) = \frac{h_{T+1}(z_{T+1})}{h_T(z_T)}$$

avec

$$h_{T+1}(z_{T+1}) = \frac{\partial^{T+1} C_{T+1}(a_{T+1})}{\partial a_1 \dots \partial a_T \partial a_{T+1}} \left(\prod_{t=1}^{T+1} f_t(a_t) \right) = c_{T+1}(a_{T+1}) \times \left(\prod_{t=1}^{T+1} f_t(a_t) \right)$$

de la même manière on trouve que

$$h_T(z_T) = c_T(a_T) \times \left(\prod_{t=1}^T f_t(a_t) \right)$$

Calculons maintenant l'espérance conditionnelle

$$\begin{aligned} E(Z_{T+1} | Z_T = z_T) &= \int_{-\infty}^{+\infty} z_{T+1} \times f_{Z_{T+1}|Z_T}(z_{T+1} | z_T) dz_{T+1} \\ &= \int_{-\infty}^{+\infty} z_{T+1} \times \frac{c_{T+1}(a_{T+1}) \times \left(\prod_{t=1}^{T+1} f_t(a_t) \right)}{c_T(a_T) \times \left(\prod_{t=1}^T f_t(a_t) \right)} dz_{T+1} \\ &= \int_{-\infty}^{+\infty} z_{T+1} \times \frac{c_{T+1}(a_{T+1})}{c_T(a_T)} \times f_{T+1}(a_{T+1}) dz_{T+1} \end{aligned}$$

On sait que $f_{T+1}(a_{T+1}) dz_{T+1} = dF_{T+1}(z_{T+1})$

D'où le résultat de l'équation (2.52) qui est

$$E(Z_{T+1} | Z_T = z_T) = \int_{-\infty}^{+\infty} z_{T+1} \frac{c_{T+1}(a_{T+1})}{c_T(a_T)} dF_{T+1}(z_{T+1})$$

A la lumière de la proposition précédente, on observe que le rapport $\frac{c_{T+1}(a_{T+1})}{c_T(a_T)}$ induit par un changement de mesure de probabilité correspondant à l'observation Z_{T+1} dans la période suivante. L'espérance conditionnelle de l'équation(2.52) peut-être réécrite comme une espérance inconditionnelle par rapport à la nouvelle densité

marginale.

$$dF_{T+1}^Q(z_{T+1}) = \frac{c_{T+1}(a_{T+1})}{c_T(a_T)} dF_{T+1}(z_{T+1}) \quad (2.53)$$

telle que

$$\begin{aligned} E(Z_{T+1} | Z_T = z_T) &= \int_{-\infty}^{+\infty} z_{T+1} dF_{T+1}^Q(z_{T+1}) \\ E(Z_{T+1} | Z_T = z_T) &= \int_{-\infty}^{+\infty} z_{T+1} f_{T+1}^Q(z_{T+1}) dz_{T+1} \end{aligned}$$

Alors

$$E(Z_{T+1} | Z_T = z_T) = E^Q(Z_{T+1}) \quad (2.54)$$

où $E^Q(Z_{T+1})$ est l'espérance de la nouvelle densité marginale.

A noter que le rapport d'une densité de copule est la densité conditionnelle :

$$c_{T+1}(a_{T+1} | a_T) = \frac{c_{T+1}(a_{T+1})}{c_T(a_T)} \quad (2.55)$$

Ce rapport est également appelé le basculement induit par la copule à densité conditionnelle.

En cas d'indépendance, ce rapport vaut 1 c'est à dire $\frac{c_{T+1}(a_{T+1})}{c_T(a_T)} = 1$, l'espérance conditionnelle devient également $E(Z_{T+1} | Z_T = z_T) = E^Q(Z_{T+1})$. Ainsi, la réclamation de la période suivante n'est pas influencée par les revendications précédentes.

L'utilisation des copules pour modéliser la dépendance temporelle des sinistres pour un seul risque individuel a une structure similaire à celle expliquée dans Frees et Wang (2005). Néanmoins, il existe quelques différences entre les deux œuvres respectives. La différence la plus importante réside dans la représentation de l'espérance conditionnelle. Dans ce travail, nous montrons que l'espérance conditionnelle peut être interprétée comme un changement de mesure et elle peut être réécrite comme une espérance inconditionnelle sous une densité marginale révisée. Cette révision de la densité marginale peut être interprétée comme une "re-pondération" de la fonction de densité après avoir observé les réclamations précédentes.

2.3.5 Les copules paramétriques

Copules archimédiennes

Définition 2.11 (cas multivarié) Soit φ une fonction convexe, continue et strictement décroissante de $[0, 1]$ dans $[0, \infty]$, la copule archimédienne de générateur φ s'écrit comme suit

$$C(a_1, \dots, a_T) = \begin{cases} \varphi^{-1}(\varphi(a_1) + \dots + \varphi(a_T)) & \text{si } \sum_{i=1}^T \varphi(a_i) \leq \varphi(0) \\ 0 & \text{sinon} \end{cases} \quad (2.56)$$

Il est nécessaire de sélectionner le générateur φ de classe C^2 afin que $\varphi(1) = 0$, $\varphi'(a) \leq 0$ et $\varphi''(a) > 0$. Les copules les plus connues de cette famille sont :

► **Copule de Gumbel (1960)** : La copule de Gumbel, définie par le générateur $\varphi_\beta(t) = (-\ln t)^\beta$, de paramètre $\beta \geq 1$, s'écrit

$$C_\beta^{Gum}(a_1, \dots, a_T) = \exp \left\{ - \left[\sum_{i=1}^T (-\log(a_i))^\beta \right]^{\frac{1}{\beta}} \right\}, \quad \beta \geq 1 \quad (2.57)$$

► **Copule de Clayton(1978)** : Cette copule est définie par

$$C_\beta^{Cl}(a_1, \dots, a_T) = \left[1 + \sum_{i=1}^n (a_i^{-\beta} - 1) \right]^{-\frac{1}{\beta}}, \quad \beta > 0 \quad (2.58)$$

avec $\varphi_\beta(t) = \frac{1}{\beta}(t^{-\beta} - 1)$ de pseudo inverse $\varphi_\beta^{-1}(t) = (\beta t + 1)^{-\frac{1}{\beta}}$ et de paramètre de dépendance $\beta \in [-1, 0[\cup]0, +\infty]$ (si $\beta \rightarrow 0$, on aura une copule d'indépendance).

► **Copule de Frank(1979)** : Cette copule est définie par

$$C_\beta^{Fr}(a_1, \dots, a_T) = -\frac{1}{\beta} \log \left\{ 1 + \frac{\prod_{i=1}^T (\exp(-\beta a_i) - 1)}{\exp(-\beta) - 1} \right\}, \quad \beta > 0 \quad (2.59)$$

avec $\varphi_\beta(t) = -\ln\left(\frac{\exp(-\beta t)-1}{\exp(-\beta)-1}\right)$ et le paramètre de dépendance $\beta \in [-1, 0[\cup]0, +\infty]$.

Copules elliptiques

Les copules elliptiques sont définies en fonction des lois de distribution elliptique. Ce qui suit définit la distribution elliptique et présente deux catégories de copules appartenant à cette famille, la copule Gaussienne et la copule de Student.

► La copule de student (cas multivarié)

Définition 2.12 Soit $\rho \in [0, 1]$, la fonction de répartition de Student à v degré de liberté est donnée par

$$t_{\rho,v}(y, z) = \int_{-\infty}^y \int_{-\infty}^z \frac{1}{2\pi\sqrt{1-\rho}} \left(1 + \frac{s^2 + t^2 - 2\rho st}{v(1-\rho^2)}\right)^{\frac{-v+2}{2}} ds dt \quad (2.60)$$

Définition 2.13 La copule de Student se définit en fonction du coefficient de corrélation linéaire ρ et du degré de liberté v respectivement. Par

$$\begin{aligned} C_{\rho,v}(a_1, \dots, a_T) &= t_{\rho,v}(t_v^{-1}(a_1), \dots, t_v^{-1}(a_T)) \\ &= \int_{-\infty}^{t_v^{-1}(a_1)} \dots \int_{-\infty}^{t_v^{-1}(a_T)} \frac{1}{2\pi\sqrt{1-\rho}} \left(1 + \frac{s^2 + t^2 - 2\rho st}{v(1-\rho^2)}\right)^{\frac{-v+2}{2}} ds dt \end{aligned} \quad (2.61)$$

Quand $v \rightarrow \infty$, alors la copule de Student converge vers la copule gaussienne, ce qui rend la distinction entre ces deux copules très difficile.

► La copule Gaussienne (cas multivarié)

La copule gaussienne revêt une grande importance car elle est essentielle à la distribution normale multivariée.

Définition 2.14 Soit $Z = (Z_1, Z_2)$ un vecteur aléatoire gaussien ($Z \sim N(0; \Sigma)$), avec Σ une matrice de covariance. Notons par ϕ_Σ et ϕ les fonction de distribution Z et Z_i , $i = 1, 2, \dots, T$ respectivement. Le vecteur aléatoire Gaussien dispose d'une copule connue sous le nom de copule Gaussienne qui est déterminée par

$$C_T(a_1, \dots, a_T) = \Phi_{\Sigma_T}(\Phi^{-1}(a_1), \dots, \Phi^{-1}(a_T)) \quad (2.62)$$

Avec Φ_{Σ_T} est une fonction de distribution normale de dimension T , Φ^{-1} fait référence à la fonction quantile de la distribution normale standard unidimensionnelle et Σ_T est la matrice de corrélation avec une représentation $\Sigma_T = (\rho_{ij})$ avec $\rho_{ij} = 1$ pour $i = j$ et $i, j = 1, \dots, T$.

Définition 2.15 (Densité d'une copule gaussienne)

La copule c_T est dite densité de la copule gaussienne si et seulement si c_T peut s'écrire sous cette forme :

$$c_T(a_T) = \frac{\exp\left(-\frac{1}{2}\zeta_T' \Sigma_T^{-1} \zeta_T\right)}{\sqrt{(2\pi)^T |\Sigma_T|} \prod_{k=1}^T \phi(\Phi^{-1}(a_k))} \quad (2.63)$$

avec $\zeta_T = (\phi^{-1}(a_1), \dots, \phi^{-1}(a_T))$ et $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$, la densité de la loi normale centrée réduite.

Proposition 2.3 Soit Z_T un vecteur de réclamations de sinistres avec une structure de copule normale comme spécifié dans l'équation(2.62), le rapport de densité de la copule normale peut-être exprimé par :

$$\frac{c_{T+1}(a_{T+1})}{c_T(a_T)} = \frac{1}{\sigma_{K.T+1}} \times \frac{\phi(\Phi^{-1}(\tilde{a}_{T+1}))}{\phi(\Phi^{-1}(a_{T+1}))} \quad (2.64)$$

avec $\Phi^{-1}(\tilde{a}_{T+1}) = \frac{\Phi^{-1}(a_{T+1}) - \mu_{K.T+1}}{\sigma_{K.T+1}}$, $\mu_{K.T+1} = \rho'_{T+1,T} \sum_{T+1}^{-1} \zeta_T$ et $\sigma_{K.T+1}^2 = 1 - \rho'_{T+1,T} \sum_{T+1}^{-1} \rho_{T+1,T}$.

Preuve En utilisant équation(2.63), on peut conclure que

$$\frac{c_{T+1}(a_{T+1})}{c_T(a_T)} = \frac{\exp\left[-\frac{1}{2}\left(\zeta'_{T+1} \sum_{T+1}^{-1} \zeta_{T+1} - \zeta'_T \sum_T^{-1} \zeta_T\right)\right]}{\sqrt{2\pi \frac{|\Sigma_{T+1}|}{|\Sigma_T|}} \phi(\Phi^{-1}(a_{T+1}))}$$

2.3. Dépendance et copules

D'après la référence [62], on a

$$\zeta'_{T+1} \sum_{T+1}^{-1} \varsigma_{T+1} - \zeta'_T \sum_T^{-1} \varsigma_T = \left(\frac{\Phi^{-1}(a_{T+1}) - \mu_{K.T+1}}{\sigma_{K.T+1}} \right)^2$$

En utilisant les propriétés du déterminant on a

$$\frac{|\sum_{T+1}|}{|\sum_T|} = 1 - \rho'_{T+1,T} \sum_T^{-1} \rho_{T+1,T} = \sigma_{K.T+1}^2$$

puis en remplaçant les quantités par leurs valeurs, on aura :

$$\begin{aligned} \frac{c_{T+1}(a_{T+1})}{c_T(a_T)} &= \frac{\exp \left[-\frac{1}{2} \left(\zeta'_{T+1} \sum_{T+1}^{-1} \varsigma_{T+1} - \zeta'_T \sum_T^{-1} \varsigma_T \right)^2 \right]}{\sqrt{2\pi \frac{|\sum_{T+1}|}{|\sum_T|}} \phi(\Phi^{-1}(a_{T+1}))} \\ &= \frac{\exp \left[-\frac{1}{2} \left(\frac{\Phi^{-1}(a_{T+1}) - \mu_{K.T+1}}{\sigma_{K.T+1}} \right)^2 \right]}{\sigma_{K.T+1} \cdot \sqrt{2\pi} \phi(\Phi^{-1}(a_{T+1}))} \end{aligned}$$

alors on peut en déduire que

$$\frac{c_{T+1}(a_{T+1})}{c_T(a_T)} = \frac{1}{\sigma_{K.T+1}} \times \frac{\phi(\Phi^{-1}(\tilde{a}_{T+1}))}{\phi(\Phi^{-1}(a_{T+1}))}$$

Corollaire 2.2

Soit Z_T un vecteur de réclamations de sinistres possédant une structure de copule Normale comme spécifié dans l'équation(2.62), alors les réclamations futures de la période suivante peuvent être exprimées par

$$E(Z_{T+1} | Z_T = z_T) = E_K \left\{ F_{T+1}^{-1} \left[\Phi(\mu_{K.T+1} + K \cdot \sigma_{K.T+1}) \right] \right\} \quad (2.65)$$

où l'espérance du côté droit est calculée pour une variable aléatoire Normale standard K .

Preuve Comme $E(Z_{T+1} | Z_T = z_T) = \int_{-\infty}^{+\infty} z_{T+1} \frac{c_{T+1}(a_{T+1})}{c_T(a_T)} dF_{T+1}(z_{T+1})$ et en uti-

lisant la proposition précédente on obtient

$$E(Z_{T+1} | Z_T = z_T) = \int_{-\infty}^{+\infty} z_{T+1} \frac{1}{\sigma_{K.T+1}} \times \frac{\phi(\Phi^{-1}(\tilde{a}_{T+1}))}{\phi(\Phi^{-1}(a_{T+1}))} dF_{T+1}(z_{T+1})$$

En faisant une transformation $k = \frac{\Phi^{-1}(a_{T+1}) - \mu_{K.T+1}}{\sigma_{K.T+1}}$ alors on obtient

$$E(Z_{T+1} | Z_T = z_T) = \int_{-\infty}^{+\infty} F_{T+1}^{-1} [\Phi(\mu_{K.T+1} + K \cdot \sigma_{K.T+1}) \cdot \phi(k) dk]$$

D'où on peut en déduire que

$$E(Z_{T+1} | Z_T = z_T) = E_K \{ F_{T+1}^{-1} [\Phi(\mu_{K.T+1} + K \cdot \sigma_{K.T+1})] \}$$

Exemple 2.4 (La distribution normale multivariée)

Considérons le cas particulier où Z_{T+1} a une distribution normale multivariée, on peut écrire la matrice de corrélation comme suit :

$$\Sigma_T = \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \dots & \rho & 1 \end{pmatrix} = (1 - \rho) \left(I_T + \frac{\rho}{1 - \rho} e_T \cdot e_T' \right)$$

où $e_T' = (1, 1, \dots, 1)$ est un vecteur de dimension T et I_T une matrice identité d'ordre $T \times T$.

L'inverse de la matrice de corrélation peut-être écrite comme

$$\Sigma_T^{-1} = (1 - \rho)^{-1} \left(I_T + \frac{\rho}{1 - \rho + \rho T} e_T \cdot e_T' \right)$$

Supposons que les marginaux sont distribués de manière identique avec comme moyenne μ et variance σ^2 , ainsi on a

$$\mu_{K.T+1} = \frac{\rho}{1 - \rho + \rho T} e_T' \zeta_T = \frac{\rho}{1 - \rho + \rho T} \sum_{i=1}^T \left(\frac{z_i - \mu}{\sigma} \right)$$

L'estimateur futur des réclamations

$$E(Z_{T+1} | Z_T = z_T) = E_K \left\{ F_{T+1}^{-1} \left[\Phi \left(\mu_{K.T+1} + K \cdot \sigma_{K.T+1} \right) \right] \right\} = \mu_{K.T+1} \sigma + \mu$$

et cette expression peut-être écrite sous la forme linéaire de crédibilité qui est

$$E(Z_{T+1} | Z_T = z_T) = \frac{\rho T}{1 - \rho + \rho T} \bar{z} + \frac{1 - \rho}{1 - \rho + \rho T} \mu$$

D'où l'expérience future peut être écrite comme étant une forme de crédibilité.

Dans le cas ou $T = 1$, on a

$$E(Z_2 | Z_1 = z_1) = \rho z_1 + (1 - \rho) \mu$$

qui est donc équivalent à la prime de crédibilité du modèle classique.

2.3.6 La dépendance de queue

On utilise le terme « dépendance de queue » pour désigner la dépendance dans les queues de distribution. Il s'agit d'une mesure locale, contrairement au rho de Spearman et au tau de Kendall qui jugent la dépendance globale sur toute la distribution.

► La dépendance de queue supérieure

Définition 2.16

Considérons deux variables aléatoires continues Y et Z , qui ont respectivement des fonctions de distribution F_Y et F_Z . La limite Λ_S est utilisée pour déterminer le coefficient de dépendance de la queue supérieure entre Y et Z , comme suit

$$\Lambda_S = \lim_{a \rightarrow 1^-} P(Y > F_Y^{-1}(a) | Z > F_Z^{-1}(a)) \quad (2.66)$$

- La quantité Λ_S est une fonction de copule lorsque $\Lambda_S \in]0, 1]$ existe, ce qui implique qu'elle demeure invariable au fur et à mesure que la transformation progresse.

- On dit que Y et Z sont dépendantes asymptotiquement au niveau supérieur de la queue de distribution, lorsque $\Lambda_S \in]0, 1]$.

- On dit que Y et Z sont indépendantes asymptotiquement au niveau supérieur de la queue de distribution, lorsque $\Lambda_S = 0$.

► **La dépendance de queue inférieure**

Définition 2.17

Considérons deux variables aléatoires continues Y et Z , qui ont respectivement des fonctions de distribution F_Y et F_Z . La limite Λ_I est utilisée pour déterminer le coefficient de dépendance de la queue inférieure entre Y et Z , comme suit

$$\Lambda_I = \lim_{a \rightarrow 0^+} P(Y < F_Y^{-1}(a) \mid Z < F_Z^{-1}(a)) \quad (2.67)$$

- On dit que Y et Z sont dépendantes asymptotiquement au niveau inférieur de la queue de distribution, lorsque $\Lambda_I \in]0, 1]$.

- On dit que Y et Z sont indépendantes asymptotiquement au niveau supérieur de la queue de distribution, lorsque $\Lambda_I = 0$.

Proposition 2.4

Prenons deux variables aléatoires Y et Z de copule $C(Y, Z)$, on a

$$\Lambda_S = \lim_{a \rightarrow 1^-} \frac{1 - 2a + C(a, a)}{1 - a} \quad (2.68)$$

$$\Lambda_I = \lim_{a \rightarrow 0^+} \frac{C(a, a)}{a} \quad (2.69)$$

Preuve

Par définition, on a :

$$\begin{aligned}
 \Lambda_S &= \lim_{a \rightarrow 1^-} P(Y > F_Y^{-1}(a) \mid Z > F_Z^{-1}(a)) \\
 &= \lim_{a \rightarrow 1^-} \frac{P(Y > F_Y^{-1}(a), Z > F_Z^{-1}(a))}{P(Z > F_Z^{-1}(a))} \\
 &= \lim_{a \rightarrow 1^-} \frac{1 - P(Y \leq F_Y^{-1}(a) \text{ ou } Z \leq F_Z^{-1}(a))}{1 - P(Z \leq F_Z^{-1}(a))} \\
 &= \lim_{a \rightarrow 1^-} \frac{1 - [P(Y \leq F_Y^{-1}(a)) + P(Z \leq F_Z^{-1}(a)) - P(Y \leq F_Y^{-1}(a) \text{ et } Z \leq F_Z^{-1}(a))]}{1 - P(Z \leq F_Z^{-1}(a))} \\
 &= \lim_{a \rightarrow 1^-} \frac{1 - (a + a - C(a, a))}{1 - a} \\
 &= \lim_{a \rightarrow 1^-} \frac{1 - 2a + C(a, a)}{1 - a}
 \end{aligned}$$

On peut également écrire

$$\begin{aligned}
 \Lambda_I &= \lim_{a \rightarrow 0^+} P(Y < F_Y^{-1}(a) \mid Z < F_Z^{-1}(a)) \\
 &= \lim_{a \rightarrow 0^+} \frac{P(Y < F_Y^{-1}(a), Z < F_Z^{-1}(a))}{P(Z < F_Z^{-1}(a))} \\
 &= \lim_{a \rightarrow 0^+} \frac{C(a, a)}{a}
 \end{aligned}$$

2.4 Conclusion

Dans ce chapitre, nous avons exploré différentes approches pour traiter la dépendance entre les sinistres. Nous avons appliqué un modèle à effets communs à deux niveaux, mais ce modèle présente des limites. En effet, certaines opérations d'intégration complexes rendent ce modèle généralement impraticable, cela entrave l'obtention de solutions claires dans plusieurs cas. Il est donc nécessaire d'examiner d'autres techniques offrant une plus grande diversité de structures de dépendance. Pour répondre à ces besoins, nous avons envisagé d'employer les copules pour modéliser la structure de dépendance des réclamations décrites dans le modèle à effets

communs à deux niveaux. Grâce à l'introduction des copules, nous avons pu constater que la distribution prédictive peut être exprimée comme une espérance sous une nouvelle mesure de probabilité (un changement de mesure) reflétant le rapport des densités des copules relatives aux revendications historiques.

CHAPITRE 3

Sur l'estimateur de la prime de crédibilité des quantiles avec une corrélation identique entre les risques

Le concept de quantiles devient progressivement un élément important du cadre de crédibilité. L'application de quantiles pour fournir des primes de crédibilité a été suggérée pour la première fois dans Pitselis (2013). Les travaux de Pitselis (2016) et Pitselis (2017) étudient l'utilisation de la crédibilité quantile dans la mesure du risque. Depuis, d'autres généralisations et modèles alternatifs de quantiles de crédibilité ont suivi, notamment Bozikas et Pitselis (2020) et Pitselis (2020), abordant leur impact sur le cadre de régression. De plus, la référence Wang et al.(2020) modélise le modèle de crédibilité quantile sous des effets communs. Notre étude apporte des nouvelles perspectives sur la modélisation des dépendances des sinistres, pour établir des primes de crédibilité. Nous nous concentrons sur la modélisation de la dépendance éventuelle présente chez les assurés, en incorporant les quantiles et l'hypothèse d'une corrélation identique entre les risques.

L'utilisation d'une hypothèse de corrélation identique dans la modélisation du risque d'assurance est plus réaliste dans certaines situations en raison de sa simplicité, de sa facilité d'étalonnage et de sa traçabilité analytique. Elle permet de gagner du temps et réduit la complexité impliquée dans l'estimation d'un grand nombre

de corrélations par paires entre les aléas. En postulant que tous les risques sont également corrélés entre eux, cela facilite l'analyse des données et simplifie la structure des dépendances et les modèles mathématiques utilisés dans la tarification des assurances. Ces facteurs en font un choix pratique pour les assureurs qui doivent trouver un équilibre entre gérabilité et précision dans leurs évaluations des risques. En revanche, bien que l'hypothèse des effets communs puisse fournir une représentation plus détaillée et potentiellement plus précise des dépendances, elle introduit une complexité significative, nécessitant davantage de données et de ressources informatiques. Dans le secteur de l'assurance, une corrélation identique entre les risques peut survenir dans divers scénarios, par exemple, considérons une compagnie d'assurance qui propose une assurance habitation pour un groupe de maisons d'habitation situées dans un même quartier. Ces maisons peuvent être exposées au même risque d'incendie si elles sont construites de la même manière, disposent de mesures de protection contre l'incendie comparables et sont situées dans la même région. Dans ce cas, la corrélation entre les risques associés à l'assurance de ces maisons contre les dommages causés par le feu pourrait être considérée comme relativement identique. Ainsi, la compagnie d'assurance évalue et garantit les risques liés à l'assurance de plusieurs logements en fonction de leurs caractéristiques communes et de leur susceptibilité aux dangers similaires. Cette approche permet aux assureurs de gérer efficacement leurs portefeuilles, d'établir des primes adaptées et de maintenir une couverture suffisante pour les assurés. Ce travail développe une prime de crédibilité qui peut être représentée sous une forme linéaire. Nous comparons les résultats numériques de notre modèle avec ceux de Pitselis (2013) et de Wang et al (2020).

3.1 Formulation du modèle et Préliminaires

3.1.1 Définition des quantiles

Définition 3.1 Soit X_j une variable aléatoire avec sa fonction de distribution cumulative $F_j(x)$, telque ($F_j(x)$ est continue partout et différentiable), et p un nombre réel compris entre 0 et 1, c'est-à-dire $0 < p < 1$. On peut écrire la fonction p-quantile

ξ_p comme suit

$$\xi_{pj} = F_j^{-1}(p) = \inf \{x : F_j(x) \geq p\} \quad (3.1)$$

Etant donné un échantillon X_1, X_2, \dots, X_n d'une variable aléatoire continue X avec les statistiques d'ordres correspondantes $X^{(1)}, X^{(2)}, \dots, X^{(n)}$, soit $\hat{\xi}_{pj}$ qui désigne l'échantillon p -quantile dont

$$-\infty < x < +\infty \text{ et } x \in [0, 1[: F_j(x) \geq p \text{ si et seulement si } \xi_{pj} \leq x$$

Supposons maintenant que le vecteur $\xi_{pj} = (\xi_{p1}, \xi_{p2}, \dots, \xi_{pk})'$ contien les $p^{\text{ième}}$ quantiles pour les K types de risques correspondants. Lorsque la fonction de distribution de X est inconnue, le $p^{\text{ième}}$ quantile empirique $\hat{\xi}_{pj}$, devient l'estimateur de $p^{\text{ième}}$ quantile ξ_{pj} . Par conséquent, les quantiles empiriques sont maintenant regroupés dans le vecteur $\hat{\xi}_{pj} = (\hat{\xi}_{p1}, \hat{\xi}_{p2}, \dots, \hat{\xi}_{pk})'$.

Ainsi, la fonction de distribution empirique peut être obtenue comme suit

$$\hat{\xi}_{pj} = n \left(\frac{i}{n} - p \right) X_{j(i-1)} + n \left(p - \frac{i-1}{n} \right) X_{ji} \text{ for } \frac{i-1}{n} \leq p \leq \frac{i}{n}, \quad i = 1, 2, \dots, n \quad (3.2)$$

Pour plus de détails, voir **Parzen (1979)** ou **Herbert and Nagaraja (2003)**.

3.1.2 Hypothèses

Tout au long de ce chapitre, nous allons considérer un portefeuille qui contient k assurés. Chaque individu de ce portefeuille a une expérience de réclamations X_{ij} qui sont des variables aléatoires où $j = 1, 2, \dots, n_i$ désigne la période. Ici, K désigne le nombre de contrats dans le portefeuille et n_i est la période pendant laquelle les réclamation les réclamations historiques sont disponibles.

Après avoir observer les risques pour n_i périodes d'expérience, une séquence d'expériences de sinistres observés $X_i = (X_{i1}, X_{i2}, \dots, X_{in_i})'$, $i = 1, 2, \dots, k$ est contribué pour chaque assuré, nous voulons estimer $\Xi_p(\Theta_i)$ qui peut être interprétée comme la

prime de risque pour le $p^{ième}$ quantile. En utilisant presque les mêmes hypothèses de **Pitselis (2013)**, nous pouvons construire un modèle de crédibilité des quantiles caractérisé par une structure de corrélation identique entre des risques. Les hypothèses sont énoncées comme suit :

Hypothèse 1 Sachant θ_i , les observations X_{ij}, \dots, X_{in_i} sont uniformément dépendantes de la fonction de distribution commune $F(X | \theta)$. Ensuite, étant donné θ , la fonction de distribution conditionnelle correspondante au p -quantile est définie comme $F(\xi_{pj} | \theta)$.

Hypothèse 2 ϕ_i est une variable aléatoire où les paramètres $\theta_1, \theta_2, \dots, \theta_k$ sont également corrélés, avec la même fonction de distribution de structure $\Pi(\theta)$ qui est appelée distribution à priori en actuariat. De manière équivalente, nous supposons qu'il existe une corrélation identique entre les primes individuelles $\Xi_p(\theta_1), \Xi_p(\theta_2), \dots, \Xi_p(\theta_k)$ expliquée par le même coefficient de corrélation ρ , c'est-à-dire. $Corr(\Xi_p(\theta_i), \Xi_p(\theta_s)) = \rho, i \neq s$.

Hypothèse 3

Notons

$$E(\hat{\xi}_i^p | \theta_i) = \Xi_p(\theta_i), \quad Var(\hat{\xi}_i^p | \theta_i) = V_p(\theta_i) = \frac{\omega_p(\theta_i)}{n_i},$$

$$Cov(\hat{\xi}_{is}^p, \hat{\xi}_{it}^p | \theta_i) = \eta(\theta_i), \quad \text{avec } i = 1, 2, \dots, k, \quad j = 1, 2, \dots, n_i. \text{ et } s \neq t$$

Remarque 3.1 Le terme $\omega_p(\theta_i)$ peut être estimé à partir des données par plusieurs méthodes, (**voir section 3.2.2 de Pitselis (2013)**).

Les paramètres de structure sont définis par :

La prime collective est

$$E[\Xi_p(\theta_i)] = \Xi_p$$

La variabilité moyenne du portefeuille est

$$E[V_p(\theta_i)] = \sigma_p^2$$

La variance entre les moyennes est

$$Var(\Xi_p(\theta_i)) = \Psi_p$$

La covariance moyenne entre quantiles est

$$E(\eta(\theta_i)) = \eta$$

Pour faciliter la présentation, nous introduisons la notation suivante :

$$d_i = \frac{n_i}{\sigma_p^2 + (n_i - 1)\eta + n_i(1 - \rho)\Psi_p}$$

$$D = \sum_{i=1}^k d_i$$

La moyenne des quantiles est

$$\overline{\hat{\xi}^p} = \frac{\sum_{i=1}^K \hat{\xi}_i^p}{K}$$

La moyenne pondérée des quantiles est

$$\overline{\hat{\xi}^p}^d = \frac{\sum_{i=1}^K d_i \hat{\xi}_i^p}{D}$$

L'objectif est maintenant d'estimer la prime future à payer pour le contrat i correspondant à des estimateurs de crédibilité des quantiles non homogène et homogène de la prime de risque $\Xi_p(\theta_i)$. Comme proposé par Buhlmann (1967), les problèmes de minimisation suivants avec une forme linéaire et un constant (noté par α dans notre modèle), doivent être résolus :

$$\min_{\alpha_0, \alpha_i} E \left[\Xi_{i, n_{i+1}}^p - \alpha_0 - \sum_{i=1}^k \alpha_i' \hat{\xi}_i^p \right]^2 \quad (3.3)$$

avec

$$E(\Xi_{i,n_i+1}^p) = E\left(\alpha_0 + \sum_{s=1}^k \alpha'_s \hat{\xi}_s^p\right), \quad \text{Où } \alpha_0 \in \mathbb{R} \text{ and } \alpha_i \in \mathbb{R}^{n_i}$$

La solution de (3.3) est définie comme la prime de crédibilité des quantiles non homogènes de Ξ_{i,n_i+1} , qui est notée $\widehat{\Xi}_{i,n_i+1}^p$ ^{non hom}.

Et

$$\min_{\alpha_i} E \left[\widehat{\Xi}_{i,n_i+1}^p - \sum_{i=1}^k \alpha'_i \hat{\xi}_i^p \right]^2 \quad (3.4)$$

avec

$$E(\Xi_{i,n_i+1}^p) = E\left(\sum_{s=1}^k \alpha'_s \hat{\xi}_s^p\right), \quad \text{Où } \alpha_i \in \mathbb{R}^{n_i}$$

La solution de (3.4) est définie comme la prime de crédibilité des quantiles homogènes de Ξ_{i,n_i+1} , qui est notée $\widehat{\Xi}_{i,n_i+1}^p$ ^{hom}.

On note

$$L\left(\hat{\xi}^p, 1\right) := \left\{ \alpha_0^j + \sum_{i=1}^k \alpha_i^{j'} \hat{\xi}_i^p, \alpha_0 \in \mathbb{R}, \alpha_i \in \mathbb{R}^{n_i}, \text{ avec } E\left[\alpha_0 + \sum_{i=1}^k \alpha'_i \hat{\xi}_i^p\right] = E(\Xi_{i,n_i+1}) \right\} \quad (3.5)$$

Un sous-espace fermé de toutes les fonctions carrées intégrables \mathcal{L} .

La classe $L\left(\hat{\xi}^p, 1\right)$ désigne toutes les solutions dans un système non homogène d'équations linéaires, dans lequel le terme constant n'est pas nécessairement égal à 0.

Et

$$L_E\left(\hat{\xi}^p\right) := \left\{ \sum_{i=1}^k \alpha_i^{j'} \hat{\xi}_i^p, \alpha_i \in \mathbb{R}^{n_i}, \text{ avec } E\left[\sum_{i=1}^k \alpha'_i \hat{\xi}_i^p\right] = E(\Xi_{i,n_i+1}) \right\} \quad (3.6)$$

3.1. Formulation du modèle et Préliminaires

La classe de tous les estimateurs collectivement sans biais pour $\Xi_p(\theta_i)$ qui est un sous-espace affine fermé de \mathcal{L}^2 . Où $\hat{\xi}^p = (\hat{\xi}_1^{p'}, \hat{\xi}_2^{p'}, \dots, \hat{\xi}_k^{p'})'$.

La classe $L_E(\hat{\xi}^p)$ désigne toutes les solutions dans un système homogène d'équations linéaires, dans lequel le terme constant est égal à 0.

D'après **Bühlmann et Gisler (2005)**, les estimateurs homogènes ne contiennent pas de terme constant contrairement à $\widehat{\Xi}_{i,n_i+1}^{p, non\ hom}$, mais s'il est nécessaire que $\widehat{\Xi}_{i,n_i+1}^{p, hom}$ soit sans biais sur le collectif, il n'est pas nécessaire que l'estimateur soit sans biais pour une valeur individuelle de θ , cette condition est satisfaite pour l'estimateur de crédibilité non homogène.

De plus, la projection orthogonale (représentée par "PROJ" en termes mathématiques) de Ξ_{i,n_i+1} sur les sous-espaces fermés $L(\hat{\xi}_p, 1)$ et $L_E(\hat{\xi}_p)$, nous donne respectivement les primes suivantes $\widehat{\Xi}_{i,n_i+1}^{p, non\ hom}$ and $\widehat{\Xi}_{i,n_i+1}^{p, hom}$, c'est-à-dire :

$$\begin{aligned}\widehat{\Xi}_{i,n_i+1}^{p, non\ hom} &= PROJ\left(\Xi_{i,n_i+1} \mid L(\hat{\xi}_p, 1)\right) \\ \widehat{\Xi}_{i,n_i+1}^{p, hom} &= PROJ\left(\Xi_{i,n_i+1} \mid L_E(\hat{\xi}_p)\right)\end{aligned}$$

Pour déterminer facilement les estimateurs de crédibilité des quantiles, nous citons d'abord un lemme préliminaire sur les projections de variables aléatoires que l'on peut trouver dans **Wen et al.(2009)**.

Lemme 3.1 Les estimateurs de crédibilité des quantiles non homogènes et homogènes sont en fait la projection orthogonale d'une variable aléatoire $Y \in L(X, 1)$ et $L_E(X)$. les formules suivantes sont vraies :

$$PROJ(Y \mid L(X, 1)) = E(Y) + Cov(Y, X) Var(X)^{-1} (X - E(X)) \quad (3.7)$$

Et

$$PROJ(Y \mid L_E(X)) = \left(Cov(Y, X) + \frac{(E(Y) - Cov(Y, X) Var(X)^{-1} E(X)) E(X')}{E(X') Var(X)^{-1} E(X)} \right) Var(X)^{-1} X \quad (3.8)$$

Où $X = (X'_1, X'_2, \dots, X'_k)'$ et $X_i = (X_{i1}, X_{i2}, \dots, X_{in_i})'$.

Afin de faciliter la dérivation de la formule (3.5), nous présentons un autre lemme sur la projection orthogonale. (voir la preuve dans **Bühlmann et Gisler (2005)**).

Lemme 3.2 Pour les deux sous espaces fermés $M \subset M' \subset L^2$, and $Y \in L^2$, L'égalité suivante est vraie :

$$PROJ(Y | M') = PROJ(PROJ(Y | M) | M')$$

Par conséquent, lorsque $L_E(X) \subset L(X, 1)$, on a :

$$PROJ(Y | L_E(X)) = PROJ(PROJ(Y | L(X, 1)) | L_E(X)) \quad (3.9)$$

La formule ci-dessus sera utilisée pour calculer la prime de crédibilité des quantiles dans le cas homogène ” Ξ_{i, n_i+1} ” .

Lemme 3.3 Sous les trois hypothèses mentionnées ci-dessus et les notations précédentes, nous avons :

1) L'espérance de $\hat{\xi}_i^p$ est

$$E(\hat{\xi}_i^p) = \Xi_p 1_{n_i}, \quad E(\Xi_{i, n_i+1}^p) = \Xi_p, \quad \text{and} \quad E(\hat{\xi}^p) = \Xi_p 1_N, \quad \text{where} \quad N = \sum_{i=1}^k n_i \quad (3.10)$$

où $1_{n_i} = (1_{n_1}, 1_{n_2}, \dots, 1_{n_k})'$ est vecteur colonne à n_i dimensions avec 1 dans toutes les n_i entrées.

2) La covariance de $\hat{\xi}_i^p$ est

$$\sum_{XX} = \sum_{\hat{\xi}^p \hat{\xi}^p} = Var(\hat{\xi}^p) = Diag(\Lambda_1, \dots, \Lambda_k) + \rho \Psi_p 1_N 1_N' \quad (3.11)$$

Où $Diag(\dots)$ est la diagonale d'une matrice par bloc où les éléments entre parenthèses sont sur la diagonale et

$$\Lambda_i = (\sigma_p^2 - \eta) I_{n_i} + (\eta + (1 - \rho) \Psi_p) 1_{n_i} 1_{n_i}'$$

3.1. Formulation du modèle et Préliminaires

où I_{n_i} est la matrice identité de dimension n_i .

3) La matrice de covariance entre Ξ_{i,n_i+1}^p et $\hat{\xi}^p$ est

$$\sum_{YX} = \sum_{\Xi_{i,n_i+1}^p \hat{\xi}^p} = cov(\Xi_{i,n_i+1}, \hat{\xi}^p) = (\eta + (1 - \rho) \Psi_p) v_i' \otimes 1'_{n_i} + \rho \Psi_p 1'_{n_i} \quad (3.12)$$

Où v_i représente un vecteur avec 1 dans la $i^{ième}$ entrée et 0 dans les autres entrées et \otimes indique le produit de Kronecker des matrices.

4) L'inverse de la matrice de covariance de $\hat{\xi}^p$ est donnée par

$$\sum_{\hat{\xi}^p \hat{\xi}^p}^{-1} = Diag(\Lambda_1^{-1}, \dots, \Lambda_k^{-1}) - \frac{1}{\frac{1}{\rho \Psi_p} + \sum_{i=1}^k \frac{n_i}{\sigma_p^2 + (n_i - 1)\eta + n_i(1 - \rho)\Psi_p}} \begin{pmatrix} \wedge_1^{-1} 1_{n_1} \\ \wedge_2^{-1} 1_{n_2} \\ \vdots \\ \wedge_k^{-1} 1_{n_k} \end{pmatrix} \left(1'_{n_1} \wedge_1^{-1}, \dots, 1'_{n_k} \wedge_k^{-1} \right) \quad (3.13)$$

Où

$$\Lambda_i^{-1} = \frac{1}{\sigma_p^2 - \eta} \left(I_{n_i} - \frac{\eta + (1 - \rho) \Psi_p}{\sigma_p^2 + (n_i - 1)\eta + n_i(1 - \rho)\Psi_p} 1_{n_i} 1'_{n_i} \right)$$

Preuve.

1) Soit $\Theta = (\Theta_1, \Theta_2, \dots, \Theta_k)$, en utilisant le théorème de la double espérance, on

a

$$E(\Xi_{i,n_i+1}^p) = E[E(\Xi_{i,n_i+1}^p | \Theta)] = E[\Xi_p(\theta_i)] = \Xi_p$$

et

$$E(\hat{\xi}_i^p) = E[E(\hat{\xi}_i^p | \theta_i)] = E(\Xi_p(\theta_i) 1_{n_i}) = \Xi_p 1_{n_i}$$

Alors

$$E \left(\hat{\xi}^p \right) = E \left(\hat{\xi}_1^{p'}, \hat{\xi}_2^{p'}, \dots, \hat{\xi}_k^{p'} \right) = \Xi_p 1_N$$

Alors pour $\Theta = (\Theta_1, \dots, \Theta_k)'$. on a

$$E \left[\hat{\xi}_i^p \right] = E \left[E \left(\hat{\xi}_i^p \mid \theta_i \right) \right] = E \left[\Xi_p(\theta_i) 1_{n_i} \right] = E \left[E \left(\left(\hat{\xi}_1^p, \hat{\xi}_2^p, \dots, \hat{\xi}_k^p \right)' \mid \theta_i \right) \right] = \Xi_p \left[1'_1, \dots, 1'_k \right]'$$

2) En se référant à les hypothèses 2 et 3, c'est apparent que

$$Cov \left(\hat{\xi}_i^p, \hat{\xi}_j^p \mid \Theta_i \right) = \begin{cases} (V_p(\theta_i) - \eta(\theta_i))I_{n_i} + \eta(\theta_i)1_{n_i}1'_{n_j}, & i = j \\ 0 & i \neq j \end{cases} \quad (3.14)$$

et

$$Cov \left(E \left(\hat{\xi}_i^p \mid \Theta_i \right), E \left(\hat{\xi}_j^p \mid \Theta_j \right) \right) = \begin{cases} \Psi_p 1_{n_i} 1'_{n_j}, & i = j \\ \rho \Psi_p 1_{n_i} 1'_{n_j}, & i \neq j \end{cases} \quad (3.15)$$

Ensuite, d'après la formule de la variance conditionnelle, on peut écrire

$$\begin{aligned} Cov \left(\hat{\xi}_i^p, \hat{\xi}_j^p \right) &= E \left[Cov \left(\hat{\xi}_i^p, \hat{\xi}_j^p \mid \Theta \right) \right] + Cov \left[E \left(\hat{\xi}_i^p \mid \Theta \right), E \left(\hat{\xi}_j^p \mid \Theta \right) \right] \\ &= \begin{cases} (\sigma_p^2 - \eta)I_{n_i} + (\eta + \Psi_p)1_{n_i}1'_{n_j}, & i = j \\ \rho \Psi_p 1_{n_i} 1'_{n_j}, & i \neq j \end{cases} \end{aligned} \quad (3.16)$$

Il est simple de vérifier que

$$Cov \left(\hat{\xi}_i^p, \hat{\xi}_i^p \right) = Var \left(\hat{\xi}_i^p \right) = \sum_{\hat{\xi}_i^p \hat{\xi}_i^p} = Diag \left(\Lambda_1, \dots, \Lambda_K \right) + \rho \Psi_p 1_N 1'_N$$

3) Comme $Cov \left(\Xi_{i,n_i+1}^p, \hat{\xi}_i^p \mid \Theta \right) = \eta(\theta_i) v'_i \otimes 1'_{n_i}$. Alors on peut écrire

$$\begin{aligned}
 \sum_{\Xi_{i,n_i+1}^p \hat{\xi}^p} &= E \left[Cov \left(\Xi_{i,n_i+1}^p, \hat{\xi}^p \mid \Theta \right) \right] + Cov \left[E \left(\Xi_{i,n_i+1}^p \mid \Theta \right), E \left(\hat{\xi}_i^p \mid \Theta \right) \right] \\
 &= E \left[\eta (\theta_i) v'_i \otimes 1'_{n_i} \right] + Cov \left[\Xi_p (\theta_i), \Xi_p (\theta_i) 1'_{n_i} \right] \\
 &= (\eta + (1 - \rho) \Psi_p) v'_i \otimes 1'_{n_i} + \rho \Psi_p 1'_N
 \end{aligned}$$

Ainsi, la formule (3.12) a été prouvée.

4) En utilisant la formule suivante d'inversion de la matrice bien connue dans **Rao et Toutenburg (1995)**,

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \quad (3.17)$$

il s'ensuit que

$$\begin{aligned}
 \sum_{\hat{\xi}^p \hat{\xi}^p}^{-1} &= (Diag(\Lambda_1, \dots, \Lambda_K) + \rho \Psi_p 1_{n_i} 1'_{n_i})^{-1} \\
 &= Diag(\Lambda_1^{-1}, \dots, \Lambda_k^{-1}) - \begin{pmatrix} \wedge_1^{-1} 1_{n_1} \\ \wedge_2^{-1} 1_{n_2} \\ \vdots \\ \wedge_k^{-1} 1_{n_k} \end{pmatrix} \left(\frac{1}{\rho \Psi_p} + \sum_{i=1}^k 1'_{n_i} \wedge_i^{-1} 1_{n_i} \right)^{-1} \begin{pmatrix} 1'_{n_1} \wedge_1^{-1} \\ \dots \\ 1'_{n_k} \wedge_k^{-1} \end{pmatrix}
 \end{aligned}$$

Parceque

$$\Lambda_i = (\sigma_p^2 - \eta) I_{n_i} + (\eta + (1 - \rho) \Psi_p) 1_{n_i} 1'_{n_i}$$

Et en utilisant l'équation (3.17) on dérive

$$\begin{aligned}
 \Lambda_i^{-1} &= [(\sigma_p^2 - \eta)I_{n_i} + (\eta + (1 - \rho)\Psi_p)1_{n_i}1'_{n_i}]^{-1} \\
 &= \frac{1}{\sigma_p^2 - \eta} \left(I_{n_i} - I_{n_i} \wedge_i \left(\frac{1}{\eta + (1 - \rho)\Psi_p} + \frac{n_i}{\sigma_p^2 - \eta} \right)^{-1} (1'_{n_i} \frac{1}{\sigma_p^2 - \eta} I_{n_i}) \right) \\
 &= \frac{1}{\sigma_p^2 - \eta} \left(I_{n_i} - \frac{(\eta + (1 - \rho)\Psi_p)d_i}{n_i} 1_{n_i}1'_{n_i} \right)
 \end{aligned}$$

En utilisant la notation d_i mentionnée dans l'hypothèse 3, on obtient

$$\Lambda_i^{-1} = \frac{1}{\sigma_p^2 - \eta} \left(I_{n_i} - \frac{\eta + (1 - \rho)\Psi_p}{\sigma_p^2 + (n_i - 1)\eta + n_i(1 - \rho)\Psi_p} 1_{n_i}1'_{n_i} \right)$$

Notons que $\Lambda_i^{-1}1_{n_i} = \frac{1_{n_i}}{\sigma_p^2 + (n_i - 1)\eta + n_i(1 - \rho)\Psi_p}$, $1'_{n_i}\Lambda_i^{-1} = \frac{1'_{n_i}}{\sigma_p^2 + (n_i - 1)\eta + n_i(1 - \rho)\Psi_p}$, et $1'_{n_i}\Lambda_i^{-1}1_{n_i} = \frac{n_i}{\sigma_p^2 + (n_i - 1)\eta + n_i(1 - \rho)\Psi_p} = d_i$. Par conséquent, on peut facilement dériver la formule de \sum_{xx}^{-1} comme (3.13).

3.2 La dérivation de la prime de crédibilité des quantiles non homogène

Dans cette section, nous trouverons l'estimateur de crédibilité des quantiles non homogène pour la prime Ξ_{i,n_i+1}^p pour $i = 1, 2, \dots, K$ avec une dépendance entre les risques et au temps. le théorème ci-dessous dérive l'estimateur de crédibilité des quantiles dans le cas non homogène .

Théorème 3.1 Sous les hypothèses 1 et 2, l'estimateur de crédibilité non homogène de la future réclamation Ξ_{i,n_i+1} , $i = 1, \dots, K$ est

$$\widehat{\Xi_{i,n_i+1}^p}^{non\ hom} = Z_{i1}\hat{\xi}_i^p + Z_{i2}\hat{\xi}^{p^d} + (1 - Z_{i1} - Z_{i2})\Xi_p \quad (3.18)$$

Où Z_{i1} et Z_{i2} sont les facteurs de crédibilité satisfaisant $Z_{i1} = (\eta + (1 - \rho)\Psi_p)d_i$ et $Z_{i2} = \frac{\rho D\Psi_p[1 - (\eta + (1 - \rho)\Psi_p)d_i]}{\rho D\Psi_p + 1}$.

Preuve. Le problème d'optimisation (3.3) peut être réduit à

3.2. La dérivation de la prime de crédibilité des quantiles non homogène

$$\min_{\alpha_0 \in \mathbb{R}, \alpha_i \in \mathbb{R}^{n_i}} E \left[\left(\Xi_{i,n_i+1} - \alpha_0 - \sum_{i=1}^k \alpha_i' \hat{\xi}_i^p \right)^2 \right]$$

En utilisant le Lemme (3.4), on peut écrire l'estimateur de crédibilité non homogène $\widehat{\Xi}_{i,n_i+1}^p$ comme

$$\widehat{\Xi}_{i,n_i+1}^p \text{ non hom} = PROJ \left(\Xi_{i,n_i+1} \mid L \left(\hat{\xi}^p, 1 \right) \right) = E \left(\Xi_{i,n_i+1} \right) + \sum_{\Xi_{i,n_i+1}} \sum_{\hat{\xi}^p}^{-1} \left(\hat{\xi}^p - \Xi_p \right) \quad (3.19)$$

A noter que $1'_{n_s} \wedge_i^{-1} 1_{n_s} = d_i$, en effectuant le remplacement des equations (3.10), (3.12), et (3.13) dans l'équation (3.19) après quelques calculs, on peut obtenir l'estimateur $\widehat{\Xi}_{i,n_i+1}^p$ non hom comme mentionné dans (3.18) :

$$\begin{aligned} \widehat{\Xi}_{i,n_i+1}^p \text{ non hom} &= \Xi_p + (\eta + (1 - \rho) \Psi_p) 1'_{n_i} \wedge_i^{-1} \left(\hat{\xi}_i^p - \Xi_p 1_{n_i} \right) + \\ &\quad \left(\rho \Psi_p - \frac{\rho \Psi_p \sum_{s=1}^k d_i + (\eta + (1 - \rho) \Psi_p) d_i}{\frac{1}{\rho \Psi_p} + \sum_{i=1}^k d_i} \right) \left(\sum_{i=1}^k 1'_{n_i} \wedge_i^{-1} \left(\hat{\xi}_i^p - \Xi_p 1_{n_i} \right) \right) \\ &= \Xi_p + (\eta + (1 - \rho) \Psi_p) \frac{n_i}{\sigma_p^2 + (n_i - 1)\eta + n_i (1 - \rho) \Psi_p} \left(\hat{\xi}_i^p - \Xi_p \right) + \\ &\quad \frac{\rho \Psi_p [1 - (\eta + (1 - \rho) \Psi_p) d_i]}{d \rho \Psi_p + 1} \sum_{s=1}^k \frac{n_i \left(\hat{\xi}_i^p - \Xi_p \right)}{\sigma_p^2 + (n_i - 1)\eta + n_i (1 - \rho) \Psi_p} \\ &= \Xi_p + (\eta + (1 - \rho) \Psi_p) d_i \left(\hat{\xi}_i^p - \Xi_p \right) + \frac{\rho D \Psi_p [1 - (\eta + (1 - \rho) \Psi_p) d_i]}{\rho D \Psi_p + 1} \left(\overline{\hat{\xi}^p}^d - \Xi_p \right) \\ &= ((\eta + (1 - \rho) \Psi_p) d_i) \hat{\xi}_i^p + \left(\frac{\rho D \Psi_p [1 - (\eta + (1 - \rho) \Psi_p) d_i]}{\rho D \Psi_p + 1} \right) \overline{\hat{\xi}^p}^d + \\ &\quad \left(1 - (\eta + (1 - \rho) \Psi_p) d_i - \frac{\rho D \Psi_p [1 - (\eta + (1 - \rho) \Psi_p) d_i]}{\rho D \Psi_p + 1} \right) \Xi_p \\ &= Z_{i1} \hat{\xi}_i^p + Z_{i2} \overline{\hat{\xi}^p}^d + (1 - Z_{i1} - Z_{i2}) \Xi_p \end{aligned}$$

Remarque 3.2 Dans ce modèle, si on suppose que pour $n_1 = \dots = n_k = n$, avec $\rho = 0$ et $\eta = 0$, ensuite

$$\begin{aligned} d_i &= \frac{n_i}{\sigma_p^2 + (n_i - 1)\eta + n_i(1 - \rho)\Psi_p} = \frac{n}{\sigma_p^2 + n\Psi_p} \\ D &= \sum_{i=1}^k d_i = \frac{kn}{\sigma_p^2 + n\Psi_p} \\ \widehat{\xi}^{p,d} &= \frac{\sum_{i=1}^K d_i \hat{\xi}_i^p}{D} = \frac{\sum_{i=1}^K \hat{\xi}_i^p}{K} \end{aligned}$$

Dans cette situation, les facteurs de crédibilité et les estimateurs de crédibilité peuvent être écrits comme

$$\begin{aligned} Z_{i1} &= \frac{n\Psi_p}{n\Psi_p + \sigma_p^2} = \frac{\Psi_p}{\Psi_p + \frac{\sigma_p^2}{n}}, \quad Z_{i2} = 0, \quad (3.20) \\ \widehat{\Xi}_{i,n_i+1}^{p, non\ hom} &= Z_{i1}\hat{\xi}_i^p + (1 - Z_{i1})\Xi_p \\ &= \frac{\Psi_p}{E[V_p(\theta_i)] + \Psi_p}\hat{\xi}_i^p + \left(1 - \frac{\Psi_p}{E[V_p(\theta_i)] + \Psi_p}\right)\Xi_p =: \widehat{\Xi}_{i,n_i+1}^{p, Pitselis} \end{aligned}$$

Ce qui est appelé l'estimateur de la prime des quantiles (c'est-à-dire, **Pitselis (2013)**). Par conséquent, notre résultat est réduit à la prime de **Pitselis (2013)**.

Remarque 3.3

Si on prend $n_1 = n_2 = \dots = n_k = n$ avec $\eta = 0$, $\sigma_{0p}^2 = n\rho\Psi_p$, $\sigma_{1p}^2 = \sigma_p^2$, et $\sigma_{2p}^2 = n(1 - \rho)\Psi_p$,

Ainsi, l'estimateur de crédibilité $\widehat{\Xi}_{i,n_i+1}^{p, non\ hom}$ pour $i = 1, \dots, K$ dans l'équation du théorème (3.1) devient

$$\widehat{\Xi}_{i,n_i+1}^{p, non\ hom} = Z_1\hat{\xi}_i^p + Z_2\tilde{\xi}^p + (1 - Z_1 - Z_2)\Xi_p$$

Où $Z_1 = \frac{\sigma_{2p}^2}{\sigma_{1p}^2 + \sigma_{2p}^2}$, $Z_2 = \frac{\sigma_{0p}^2\sigma_{1p}^2 K}{(\sigma_{1p}^2 + \sigma_{2p}^2)(\sigma_{1p}^2 + \sigma_{2p}^2 + K\sigma_{0p}^2)}$ et $\tilde{\xi}^p = \frac{\sum_{i=1}^K \hat{\xi}_i^p}{K}$.

Qui ne sont que des formules du théorème (3.1) dans **Wang et al.(2020)**. évidemment, les estimateurs de crédibilité des quantiles à effets communs aléatoires généreront également des corrélations identiques et ils représentent un cas particulier

de notre modèle.

Pour les estimateurs de crédibilité quantile non homogène mentionnés dans cette section, la valeur de la prime collective Ξ_p doit être calculée en utilisant la distribution a priori. Dans le cas où la prime collective Ξ_p est inconnue, nous pouvons surmonter ce problème en résolvant le problème de minimisation 2 sous la classe $L_E(\hat{\xi}^p)$ et en établissant les estimateurs de crédibilité de quantiles homogène de Ξ_{i,n_i+1}^p pour $i = 1, \dots, K$.

3.3 La dérivation de la prime de crédibilité des quantiles homogène

Théorème 3.2 Selon les hypothèses ci-dessus, l'estimateur de crédibilité homogène de Ξ_{i,n_i+1} pour $i = 1, 2, \dots, k$, peut être exprimé comme

$$\widehat{\Xi_{i,n_i+1}}^{\text{hom}} = Z_{i1}\hat{\xi}_i^p + (1 - Z_{i1})\overline{\xi}^{\overline{p}^d} \quad (3.21)$$

Où $Z_{i1} = (\eta + (1 - \rho)\Psi_p)d_i$.

Preuve a partir (3.8), il apparaît que

$$PROJ\left(\Xi_{i,n_i+1}^p \mid L_E\left(\hat{\xi}^p\right)\right) = PROJ\left(PROJ\left(\Xi_{i,n_i+1}^p \mid L\left(\hat{\xi}^p, 1\right)\right) \mid L_E\left(\hat{\xi}^p\right)\right)$$

Théorème (3.1) indique que

$$PROJ\left(\Xi_{i,n_i+1}^p \mid L\left(\hat{\xi}^p, 1\right)\right) = Z_{i1}\hat{\xi}_i^p + Z_{i2}\overline{\xi}^{\overline{p}^d} + (1 - Z_{i1} - Z_{i2})\Xi_p$$

Puisque $\hat{\xi}_i^p, \overline{\xi}^{\overline{p}^d} \in L_E\left(\hat{\xi}^p\right)$, on peut écrire

$$\widehat{\Xi_{i,n_i+1}}^{\text{hom}} = Z_{i1}\hat{\xi}_i^p + Z_{i2}\overline{\xi}^{\overline{p}^d} + (1 - Z_{i1} - Z_{i2})PROJ\left(\Xi_p \mid L_E\left(\hat{\xi}^p\right)\right) \quad (3.22)$$

3.3. La dérivation de la prime de crédibilité des quantiles homogène

De plus, $Cov(Y, X) = 0$, pour $Y = \Xi_p(\theta)$ et $X = \hat{\xi}^p$ dans (3.8), Alors il apparaît que

$$PROJ\left(\Xi_p \mid L_E\left(\hat{\xi}^p\right)\right) = \frac{\Xi_p E\left(\hat{\xi}^p\right) \sum_{\hat{\xi}^p}^{-1} \hat{\xi}^p}{E\left(\hat{\xi}^p\right) \sum_{\hat{\xi}^p}^{-1} E\left(\hat{\xi}^p\right)} \quad (3.23)$$

En remplaçant les équations (3.10) et (3.13) dans (3.23), on obtient

$$\begin{aligned} & PROJ\left(\Xi_p \mid L_E\left(\hat{\xi}^p\right)\right) = \\ & \frac{\Xi_p^2 \mathbf{1}'_N \left(\text{Diag}\left(\wedge_1^{-1}, \dots, \wedge_k^{-1}\right) - \frac{\rho \Psi_p}{1 + \rho D \Psi_p} \begin{pmatrix} \wedge_1^{-1} \mathbf{1}_{n_1} \\ \vdots \\ \wedge_k^{-1} \mathbf{1}_{n_k} \end{pmatrix} \left(\mathbf{1}'_{n_1} \wedge_1^{-1}, \dots, \mathbf{1}'_{n_k} \wedge_k^{-1}\right) \right)}{\Xi_p \mathbf{1}'_N \left(\text{Diag}\left(\wedge_1^{-1}, \dots, \wedge_k^{-1}\right) - \frac{\rho \Psi_p}{1 + \rho D \Psi_p} \begin{pmatrix} \wedge_1^{-1} \mathbf{1}_{n_1} \\ \vdots \\ \wedge_k^{-1} \mathbf{1}_{n_k} \end{pmatrix} \left(\mathbf{1}'_{n_1} \wedge_1^{-1}, \dots, \mathbf{1}'_{n_k} \wedge_k^{-1}\right) \right)} \Xi_p \mathbf{1}_N \\ & = \overline{\hat{\xi}^p}^d \end{aligned}$$

Ensuite, l'estimateur de crédibilité homogène $\widehat{\Xi}_{i, n_i+1}^p \text{ hom}$ devient comme mentionné dans (3.21) :

$$\begin{aligned} \widehat{\Xi}_{i, n_i+1}^p \text{ hom} &= Z_{i1} \hat{\xi}_i^p + Z_{i2} \overline{\hat{\xi}^p}^d + (1 - Z_{i1} - Z_{i2}) \overline{\hat{\xi}^p}^d \\ &= Z_{i1} \hat{\xi}_i^p + (Z_{i2} + 1 - Z_{i1} - Z_{i2}) \overline{\hat{\xi}^p}^d \\ &= Z_{i1} \hat{\xi}_i^p + (1 - Z_{i1}) \overline{\hat{\xi}^p}^d \end{aligned}$$

Ce qui donne le résultat.

Les deux théorèmes ci-dessus fournissent les estimateurs des primes de crédibilité non homogènes et homogènes pour chaque contrat, mais les paramètres structuels Ξ_p , σ_p^2 , Ψ_p, η et ρ dans les applications réelles sont généralement inconnus en formule

3.3. La dérivation de la prime de crédibilité des quantiles homogène

et cela nous amène à les estimer sur la base des quantiles $\hat{\xi}^p = (\hat{\xi}_1^{p'}, \hat{\xi}_2^{p'}, \dots, \hat{\xi}_k^{p'})$.

Concernant l'estimateur de Ξ_p , il ne représente pas un sujet de grande controverse car il est déjà disponible dans la formule pour $\widehat{\Xi}_{i, n_i+1}^p$ ^{hom} et bien établie dans la littérature actuarielle (voir **Pitselis (2013)**).

Nous étudierons donc uniquement l'estimation des trois paramètres structurels de σ_p^2 , Ψ_p et η .

Les estimateurs de σ_p^2 , Ψ_p et η sont donnés dans la proposition suivante, Nous supposons également que ces estimateurs sont sans biais.

Proposition 3.1 Les estimateurs de σ_p^2 , η et Ψ_p sont donnés par

$$(1) \hat{\sigma}_p^2 = E \left[\hat{V}_p(\theta_i) \right] = \frac{1}{K} \sum_{i=1}^K Var \left(\hat{\xi}_i^p \mid \theta_i \right) = \frac{1}{K} \sum_{i=1}^K \frac{\hat{\omega}_p(\theta_i)}{n_i} \quad (3.24)$$

$$(2) \hat{\eta} = \hat{\sigma}_p^2 - \frac{1}{K} \sum_{i=1}^K \frac{1}{n_i - 1} \sum_{i=1}^{n_i} (\hat{\xi}_i^p - \bar{\hat{\xi}}^p)^2 \quad (3.25)$$

$$(3) \hat{\Psi}_p = Var(\Xi_p(\theta_i)) = \frac{1}{1-\rho} \left(\frac{1}{K-1} \sum_{i=1}^K (\hat{\xi}_i^p - \bar{\hat{\xi}}^p)^2 - \frac{1}{K} \sum_{i=1}^K \frac{\hat{\sigma}_p^2 + (n_i - 1)\hat{\eta}}{n_i} \right) \quad (3.26)$$

Preuve

1) Soit $S_i^2 = \frac{1}{n_i-1} \sum_{i=1}^{n_i} (\hat{\xi}_i^p - \bar{\hat{\xi}}^p)^2$, où $\bar{\hat{\xi}}^p = \frac{1}{K} \sum_{i=1}^K \hat{\xi}_i^p$

On a

$$E(S_i^2) = E \left[\frac{1}{n_i - 1} \sum_{i=1}^{n_i} (\hat{\xi}_i^p - \bar{\hat{\xi}}^p)^2 \right] = \frac{1}{n_i - 1} \sum_{i=1}^{n_i} E \left(\hat{\xi}_i^p - \bar{\hat{\xi}}^p \right)^2$$

$$\begin{aligned}
 E \left[\left(\hat{\xi}_i^p - \bar{\xi}^p \right)^2 \mid \theta_i \right] &= \text{Var} \left[\left(\hat{\xi}_i^p - \bar{\xi}^p \right) \mid \theta_i \right] \\
 &= \text{Var}(\hat{\xi}_i^p \mid \theta_i) + \text{Var}(\bar{\xi}^p \mid \theta_i) - 2\text{Cov} \left(\hat{\xi}_i^p, \bar{\xi}^p \mid \theta_i \right) \\
 &= \nu_p(\theta_i) + \frac{1}{n_i} \nu_p(\theta_i) - 2 \frac{1}{n_i} \nu_p(\theta_i) = \frac{n_i - 1}{n_i} V_p(\theta_i)
 \end{aligned}$$

On sait que

$$\begin{aligned}
 E \left(\left(\hat{\xi}_i^p - \bar{\xi}^p \right)^2 \right) &= E \left(E \left(\hat{\xi}_i^p - \bar{\xi}^p \mid \theta_i \right) \right) = E \left(\frac{n_i - 1}{n_i} V_p(\theta_i) \right) \\
 &= \frac{n_i - 1}{n_i} E(V_p(\theta_i)) = \frac{n_i - 1}{n_i} \sigma_p^2(\theta_i)
 \end{aligned}$$

Et

$$\begin{aligned}
 E(S_i^2) &= \frac{1}{n_i - 1} \sum_{i=1}^{n_i} \frac{n_i - 1}{n_i} \sigma_p^2(\theta_i) = \sigma_p^2 \\
 E(\hat{\sigma}_p^2) &= \frac{1}{k} \sum_{i=1}^k E(S_i^2) = \sigma_p^2, \\
 E(\hat{\sigma}_p^2) &= \frac{1}{k} \sum_{i=1}^k E(V_p(\theta_i)) = E \left(\frac{1}{k} \sum_{i=1}^k V_p(\theta_i) \right)
 \end{aligned}$$

On obtient donc immédiatement que

$$\hat{\sigma}_p^2 = \frac{1}{k} \sum_{i=1}^k V_p(\theta_i)$$

En plus, pour $\text{Var}(\hat{\xi}_i^p \mid \theta_i) = V_p(\theta_i) = \frac{\hat{\omega}_p(\theta_i)}{n_i}$, l'estimateur de σ_p^2 peut s'écrire comme suit

$$\hat{\sigma}_p^2 = \frac{1}{k} \sum_{i=1}^k \text{Var}(\hat{\xi}_i^p \mid \theta_i) = \frac{1}{K} \sum_{i=1}^K \frac{\hat{\omega}_p(\theta_i)}{n_i}$$

3.3. La dérivation de la prime de crédibilité des quantiles homogène

2) Pour $\hat{\eta}$ on a

$$\sum_{i=1}^{n_i} (\hat{\xi}_i^p - \bar{\hat{\xi}}^p)^2 = \sum_{i=1}^{n_i} (\hat{\xi}_i^p - \Xi_p(\theta_i))^2 - n_i (\bar{\hat{\xi}}^p - \Xi_p(\theta_i))^2$$

Et à partir de la loi de l'espérance itérée, on obtient

$$E \left(\frac{1}{n_j - 1} \sum_{i=1}^{n_i} (\hat{\xi}_i^p - \bar{\hat{\xi}}^p)^2 \right) = E \left(E \left(\frac{1}{n_j - 1} \sum_{i=1}^{n_i} (\hat{\xi}_i^p - \bar{\hat{\xi}}^p)^2 \mid \theta \right) \right) = \sigma^2 - \eta$$

Ensuite

$$\hat{\eta} = \hat{\sigma}_p^2 - \frac{1}{K} \sum_{i=1}^K \frac{1}{n_j - 1} \sum_{i=1}^{n_i} (\hat{\xi}_i^p - \bar{\hat{\xi}}^p)^2$$

3) Pour $\hat{\Psi}_p$, nous devons d'abord calculer ces quantités

$$Var \left(\hat{\xi}_i^p \right) = \frac{\sigma_p^2 + (n_i - 1)\eta}{n_i} + \Psi_p$$

$$Var \left(\bar{\hat{\xi}}^p \right) = \frac{1}{K^2} \sum_{i=1}^K \frac{\sigma_p^2 + (n_i - 1)\eta}{n_i} + \rho\Psi_p + \frac{(1 - \rho)\Psi_p}{K}$$

$$Cov \left(\hat{\xi}_i^p, \bar{\hat{\xi}}^p \right) = \rho\Psi_p + \frac{(1 - \rho)\Psi_p}{K} + \frac{\sigma_p^2 + (n_i - 1)\eta}{Kn_i}$$

aussi,

$$\begin{aligned} \frac{1}{K-1} \sum_{i=1}^K E \left(\hat{\xi}_i^p - \bar{\hat{\xi}}^p \right)^2 &= \frac{1}{K-1} \sum_{i=1}^K \left[Var \left(\hat{\xi}_i^p \right) + Var \left(\bar{\hat{\xi}}^p \right) - 2Cov \left(\hat{\xi}_i^p, \bar{\hat{\xi}}^p \right) \right] \\ &= (1 - \rho)\Psi_p + \frac{1}{K} \sum_{i=1}^K \frac{\sigma_p^2 + (n_i - 1)\eta}{n_i} \end{aligned}$$

Le paramètre de structure $\hat{\Psi}_p$ est sans biais car

$$E \left[\hat{\Psi}_p \right] = \frac{1}{1 - \rho} E \left[\frac{1}{K-1} \sum_{i=1}^K \left(\hat{\xi}_i^p - \bar{\hat{\xi}}^p \right)^2 - \frac{1}{K} \sum_{i=1}^K \frac{\hat{\sigma}_p^2 + (n_i - 1)\hat{\eta}}{n_i} \right] = \hat{\Psi}_p$$

3.4 Application numérique

Après avoir démontré comment obtenir les deux estimateurs de $\widehat{\Xi}_{i,n_i+1}^{p, non\ hom}$ et $\widehat{\Xi}_{i,n_i+1}^{p, hom}$ par les théorèmes (3.1) et (3.2), dans cette section, nous voulons faire une comparaison entre les résultats des estimateurs des quantiles de crédibilité avec une corrélation identique entre les risques et les estimateurs de crédibilité des quantiles dans **Pitselis(2013)** et **Wang et al (2020)**. Pour cette raison, nous utilisons les mêmes données que **Wang et al (2020)**. nous générons n =1000 groupes de trajectoire de sinistres sur 10 années pour 10 individus différents et nous présentons les données d'expérience pour le 1^{er} groupe dans le tableau (3.1). Ensuite, les statistiques descriptives des primes de crédibilité de quantile avec p=0.5 pour 10 individus dans ce groupe sont données dans le tableau (3.2).

Groupe	Contrat	$\mathbf{X}_{1.1}$	$\mathbf{X}_{1.2}$	$\mathbf{X}_{1.3}$	$\mathbf{X}_{1.10}$
1	1	268.71	182.69	394.92	283.68
	2	296.85	347.79	370.99	374.21
	3	367.35	213.87	215.27	387.67
	4	337.74	443.31	332.42	371.85
	5	309.75	338.14	221.49	244.74
	6	314.40	317.58	374.51	265.73
	7	265.50	409.18	394.40	340.09
	8	181.41	215.33	370.74	332.11
	9	322.73	388.63	306.62	384.12
	10	329.49	326.14	272.04	323.35

Tableau 3.1. Données de portefeuille.

3.4. Application numérique

Contrat	Moyenne	Mediane	Minimum	Maximum	Ecart-type
1	313.23	309.06	182.69	396.30	67.05
2	318.74	332.38	132.85	416.51	81.35
3	292.43	317.26	184.44	387.66	75.83
4	350.54	338.39	247.74	443.31	66.23
5	277.10	278.14	194.31	339.05	49.66
6	340.68	339.77	265.73	401.97	39.56
7	280.31	302.79	84.68	451.73	129.96
8	290.28	271.71	181.41	458.11	81.69
9	305.30	319.45	106.62	388.63	84.48
10	295.46	306.24	210.88	348.69	45.80

Tableau 3.2. statistiques descriptives .

Le Tableau (3.3) montre une comparaison entre les résultats de notre modèle dans les cas homogène et non homogène et ceux du modèle de Pitselis (2013) où l'hypothèse de dépendance n'est pas considérée, et le modèle de Wang et al avec effets communs (2020) où ils ont supposé que Θ_i varie parmi 10 individus mais que la variable aléatoire d'effet commun Λ reste la même. Le quantile que nous utilisons est la médiane (l'ordre $p = 0,5$). Pour montrer comment le coefficient de corrélation ρ (qui indique la dépendance entre les risques) affecte les estimateurs de nos primes, nous considérons qu'il prend deux valeurs connues : $\rho = 0,4$ avec $\eta = 0,32$ et $\rho = 0,65$ avec $\eta = 0,12$. Ensuite, par des calculs simples, nous pouvons obtenir les estimateurs de crédibilité pour la période suivante (11ème année).

Maintenant, par des calculs simple, nous pouvons obtenir les estimateurs de crédibilité pour la période suivante (11^{ème} Année) qui sont repris dans le tableau suivant :

3.4. Application numérique

Contrat	1	2	3	9	10
$\widehat{\xi}_i^{0.5}$	309.06	332.26	317.26	319.45	306.24
$\overline{\xi}^{0.5}$	311.52	311.52	311.52	311.52	311.52
Modèle de crédibilité des quantiles avec effet commun (2020) pour $\Lambda = 132.12$						
Θ_i	98.80	175.68	144.48	69.69	111.81
Z_1	0.43	0.43	0.43	0.43	0.43
Z_2	0.18	0.18	0.18	0.18	0.18
Z_3	0.39	0.39	0.39	0.39	0.39
$\widehat{\Xi}_{i,n_i+1}^p{}^C$	305.96	315.99	309.43	310.43	304.75
$\Xi_P = 300$	$\sigma_{0,P}^2 = 836.01$	$\sigma_{1,P}^2 = 1126.77$	$\sigma_{2,P}^2 = 89.10$			
Modèle de crédibilité des quantiles sans dépendance (Pitselis 2013)						
Z	0.75	0.75	0.75	0.75	0.75
$\widehat{\Xi}_{i,n_i+1}^p{}^{Pitselis}$	309.67	327.22	315.84	317.49	307.54
Modèle de crédibilité des quantiles avec une corrélation identique pour $\rho = 0.4$ et $\eta = 0.32$						
Z_1	0.52	0.52	0.52	0.52	0.52
Z_2	0.12	0.12	0.12	0.12	0.12
Z_3	0.36	0.36	0.36	0.36	0.36
$\widehat{\Xi}_{i,n_i+1}^p{}^{non\ hom}$	310.23	322.45	314.53	315.68	308.75
$\widehat{\Xi}_{i,n_i+1}^p{}^{hom}$	310.23	322.45	314.53	315.68	308.75
Modèle de crédibilité des quantiles avec une corrélation identique pour $\rho = 0.65$ et $\eta = 0.12$						
Z_1	0.27	0.27	0.27	0.27	0.27
Z_2	0.24	0.24	0.24	0.24	0.27
Z_3	0.49	0.49	0.49	0.49	0.49
$\widehat{\Xi}_{i,n_i+1}^p{}^{non\ hom}$	310.84	317.20	313.08	313.68	310.08
$\widehat{\Xi}_{i,n_i+1}^p{}^{hom}$	310.84	317.20	313.08	313.68	310.08

Tableau 3.3. Comparaison des primes de crédibilité des quantiles

3.4. Application numérique

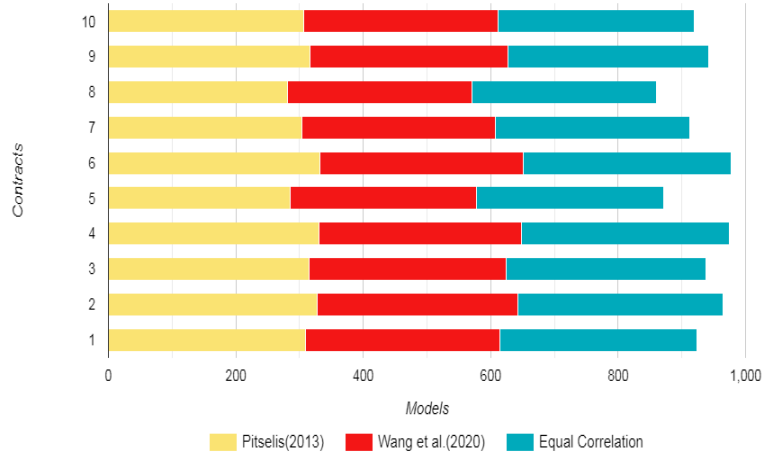


Figure 3.1. Comparaison des trois primes avec $\rho = 0.4$

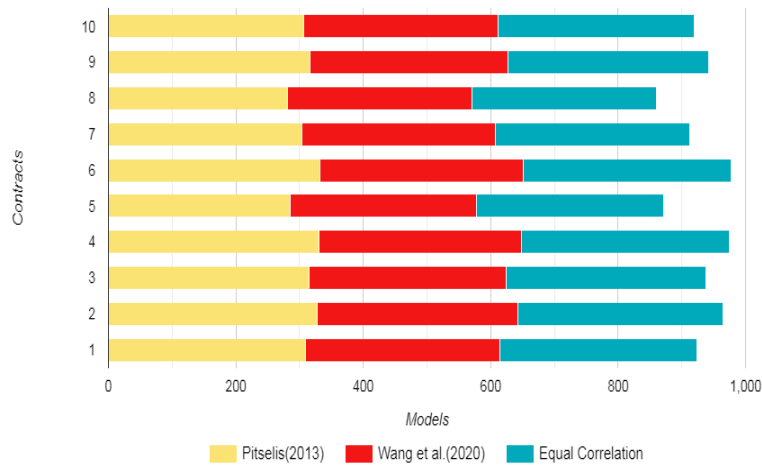


Figure 3.2 Comparaison des trois primes avec $\rho = 0.65$

D'après le tableau (3.3) et comme le montrent les figures (3.1) et (3.2) :

(i) Notre modèle accorde plus de poids à l'expérience individuelle par rapport aux primes de crédibilité de quantile dans les modèles de Wang et al (2020) et de Pitselis (2013), ce qui explique le fait que les valeurs du modèle de Pitselis sont plus proches de la perte moyenne du portefeuille (la prime collective est de 311,52).

(ii) À mesure que la valeur de ρ augmente, les primes des trois modèles se rapprochent. Cette approximation entre les trois primes est justifiée par la convergence du paramètre vers 1, cette dernière valeur signifiant une indépendance totale entre les risques, ce qui nous ramène au modèle de Pitselis.

3.5 Perspectives

Selon les résultats de la comparaison, la prime de crédibilité nouvellement définie devrait être privilégiée dans la pratique car l'illustration numérique sur des données réelles montre la faisabilité du modèle. Le fait que cette hypothèse soit plus pragmatique sur le marché de l'assurance lui donne de plus en plus l'opportunité d'être étudiée et exploitée par les actuaires. Cette hypothèse simplifie l'analyse du portefeuille d'assurance en supposant un niveau uniforme de corrélation entre les risques. Elle peut être utilisée dans certains modèles mathématiques ou calculs au sein de la théorie de la crédibilité en assurance pour rendre les calculs plus gérables ou pour fournir un scénario de référence pour l'analyse. Cependant, il est essentiel de reconnaître qu'en pratique, la corrélation entre les risques dans un portefeuille d'assurance ne peut pas être identique entre toutes les paires de risques. Les corrélations réelles peuvent changer selon divers facteurs comme : la nature des risques, la localisation géographique, les conditions économiques...etc. Bien que l'hypothèse de structure de corrélation identique fournisse une simplification pratique, il est important pour les professionnels de l'assurance et les actuaires d'évaluer de manière critique si elle reflète avec précision les schémas de corrélation dans le portefeuille d'assurance spécifique analysé. Dans de nombreux cas, des modèles plus sophistiqués peuvent être nécessaires pour capturer efficacement la véritable structure de corrélation.

Conclusion générale

Dans cette thèse, nous avons examiné l'impact de la théorie de la crédibilité sur la tarification des assurances, en mettant en lumière son évolution, ses applications actuelles et ses avancées récentes. Tout d'abord, nous avons repensé les fondements historiques de la théorie de la crédibilité, en nous concentrant sur les modèles de Bühlmann et de Bühlmann-Straub, qui ont été un tournant dans l'évaluation des primes d'assurance à partir de l'expérience passée des sinistres. Ensuite, notre analyse s'est étendue en étudiant la relation entre les risques d'assurance, un élément souvent négligé dans les modèles de crédibilité classiques. En incorporant des mesures de dépendance, les copules et l'effet commun, on peut mieux comprendre les liens complexes entre les sinistres, ce qui permet une tarification plus précise et adaptée au profil de risque des portefeuilles d'assurance. Finalement, dans le troisième chapitre, nous avons introduit un nouveau modèle de crédibilité, qui est le modèle de crédibilité des quantiles avec une corrélation identique entre les risques individuels. Ce modèle novateur répond aux contraintes des méthodes traditionnelles en intégrant de manière plus robuste et souple la variabilité et la corrélation entre les risques, tout en préservant les bénéfices de la théorie de la crédibilité. Ce modèle offre de nouvelles opportunités pour améliorer les méthodes de tarification, en particulier dans des situations où les répartitions des sinistres sont asymétriques ou comportent des dépendances complexes.

En conclusion, cette thèse démontre que la théorie de la crédibilité, même si elle repose sur des principes traditionnels, continue d'évoluer et de s'ajuster aux défis contemporains engendrés par la complexité grandissante des risques assurantiels. Dans les compagnies d'assurance, les avancées récentes, telles que le modèle de crédibilité de quantile démontrent que cette théorie demeure un outil indispensable pour les actuaires, offrant ainsi une gestion plus précise et plus efficace des risques. Ces progrès laissent aussi entrevoir des perspectives prometteuses pour la recherche à venir, en particulier en ce qui concerne la modélisation de la dépendance et l'intégration des techniques d'apprentissage automatique dans les méthodes de crédibilité.

Annexes

A) Dans chapitre 2 (Propriétés asymptotiques)

P2) Lorsque $T \rightarrow \infty$, nous avons $\omega_1 \rightarrow 1$, $\omega_{i \neq 1} \rightarrow 0$ et $\omega_{\theta, \lambda} \rightarrow 0$ car

$$\begin{aligned} \omega_1 &= \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 T + \sigma_x^4} \\ &= \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 \frac{1}{T}}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 \frac{1}{T} + \sigma_x^4 \frac{1}{T^2}} \\ &\rightarrow \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2} = 1 \end{aligned}$$

D'où on peut conclure que $\omega_1 \rightarrow 1$, lorsque $T \rightarrow \infty$.

$$\begin{aligned} \omega_{i \neq 1} &= \frac{(I - 1) \sigma_\lambda^2 \sigma_x^2 \frac{1}{T}}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 \frac{1}{T} + \sigma_x^4 \frac{1}{T^2}} \\ &\rightarrow \frac{0}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2} = 0 \end{aligned}$$

Alors $\omega_{i \neq 1} \rightarrow 0$ Lorsque $T \rightarrow \infty$.

$$\begin{aligned} \omega_{\theta, \lambda} &= \frac{\sigma_\theta^2 \sigma_x^2 \frac{1}{T} + \sigma_x^4 \frac{1}{T^2}}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 \frac{1}{T} + \sigma_x^4 \frac{1}{T^2}} \\ &\rightarrow \frac{0}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2} = 0 \end{aligned}$$

D'où on peut en déduire que $\omega_{\theta, \lambda} \rightarrow 0$ lorsque $T \rightarrow \infty$.

P3) Lorsque $I \rightarrow \infty$, nous avons $\omega_1 \rightarrow \frac{\sigma_\theta^2 T}{\sigma_x^2 + \sigma_\theta^2 T}$, $\omega_{i \neq 1} \rightarrow \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2 T}$ et $\omega_{\theta, \lambda} \rightarrow 0$ car

$$\begin{aligned}
\omega_1 &= \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 T + \sigma_x^4} \\
&= \frac{\sigma_\theta^2 T^2 \sigma_\lambda^2 I + \sigma_\theta^4 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T}{\sigma_\theta^2 T^2 \sigma_\lambda^2 I + \sigma_\theta^4 T^2 + \sigma_x^2 \sigma_\lambda^2 I T + 2\sigma_\theta^2 \sigma_x^2 T + \sigma_x^4} \\
&= \frac{\sigma_\theta^2 T^2 \sigma_\lambda^2 + \sigma_\theta^4 T^2 \left(\frac{1}{I}\right) + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T \left(\frac{1}{I}\right)}{\sigma_\theta^2 T^2 \sigma_\lambda^2 + \sigma_\theta^4 T^2 \left(\frac{1}{I}\right) + \sigma_x^2 \sigma_\lambda^2 T + 2\sigma_\theta^2 \sigma_x^2 T \left(\frac{1}{I}\right) + \sigma_x^4 \left(\frac{1}{I}\right)} \\
&\rightarrow \frac{T^2 \sigma_\theta^2 \sigma_\lambda^2}{\sigma_\lambda^2 \sigma_\theta^2 T + \sigma_x^2 \sigma_\lambda^2 T^2} = \frac{\sigma_\theta^2 T}{\sigma_x^2 + \sigma_\theta^2 T}
\end{aligned}$$

D'où on peut conclure que $\omega_1 \rightarrow \frac{\sigma_\theta^2 T}{\sigma_x^2 + \sigma_\theta^2 T}$ lorsque $I \rightarrow \infty$.

$$\begin{aligned}
\omega_{i \neq 1} &= \frac{T(I-1) \sigma_\lambda^2 \sigma_x^2}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)} \\
&= \frac{\sigma_\lambda^2 \sigma_x^2 T I - T \sigma_\lambda^2 \sigma_x^2}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 T + \sigma_x^4} \\
&= \frac{\sigma_\lambda^2 \sigma_x^2 T - T \sigma_\lambda^2 \sigma_x^2 \left(\frac{1}{I}\right)}{\sigma_\lambda^2 \sigma_\theta^2 T^2 + \sigma_\theta^4 T^2 \left(\frac{1}{I}\right) + \sigma_\lambda^2 \sigma_x^2 T + 2\sigma_\theta^2 \sigma_x^2 T \left(\frac{1}{I}\right) + \sigma_x^4 \left(\frac{1}{I}\right)} \\
&\rightarrow \frac{\sigma_\lambda^2 \sigma_x^2 T}{\sigma_\lambda^2 \sigma_x^2 T + \sigma_\lambda^2 \sigma_\theta^2 T^2} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2 T}
\end{aligned}$$

Alors on en déduit que $\omega_{i \neq 1} \rightarrow \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2 T}$ lorsque $I \rightarrow \infty$

$$\begin{aligned}
\omega_{\theta, \lambda} &= \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2)}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_x^2 + \sigma_\theta^2 T)} \\
&= \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2) \left(\frac{1}{I}\right)}{\sigma_\lambda^2 \sigma_\theta^2 T^2 + \sigma_\theta^4 T^2 \left(\frac{1}{I}\right) + \sigma_\lambda^2 \sigma_x^2 T + 2\sigma_\theta^2 \sigma_x^2 T \left(\frac{1}{I}\right) + \sigma_x^4 \left(\frac{1}{I}\right)} \\
&\rightarrow \frac{0}{\sigma_\lambda^2 \sigma_\theta^2 T^2 + \sigma_\lambda^2 \sigma_x^2 T} = 0
\end{aligned}$$

Alors comme conclusion $\omega_{\theta, \lambda} \rightarrow 0$ lorsque $I \rightarrow \infty$.

P4) Lorsque $\sigma_x^2 \rightarrow \infty$, nous avons $\omega_1 \rightarrow 0$, $\omega_{i \neq 1} \rightarrow 0$ et $\omega_{\theta, \lambda} \rightarrow 1$ car

$$\begin{aligned}
 \omega_1 &= \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 T + \sigma_x^4} \\
 &= \frac{\left(\frac{1}{\sigma_x^4}\right) (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T \left(\frac{1}{\sigma_x^2}\right)}{\left(\frac{1}{\sigma_x^4}\right) (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \left(\frac{1}{\sigma_x^2}\right) T + 1} \\
 &\rightarrow \frac{0}{0+1} = 0
 \end{aligned}$$

Alors on en déduit que $\omega_1 \rightarrow 0$ lorsque $\sigma_x^2 \rightarrow \infty$.

$$\begin{aligned}
 \omega_{i \neq 1} &= \frac{T(I-1) \sigma_\lambda^2 \sigma_x^2}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)} \\
 &= \frac{\sigma_\lambda^2 \left(\frac{1}{\sigma_x^2}\right) T I - T \sigma_\lambda^2 \left(\frac{1}{\sigma_x^2}\right)}{\left(\frac{1}{\sigma_x^4}\right) (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \left(\frac{1}{\sigma_x^2}\right) T + 1} \\
 &\rightarrow \frac{0}{0+1} = 0
 \end{aligned}$$

D'où on peut conclure que $\omega_{i \neq 1} \rightarrow 0$ lorsque $\sigma_x^2 \rightarrow \infty$.

$$\begin{aligned}
 \omega_{\theta, \lambda} &= \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2)}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_x^2 + \sigma_\theta^2 T)} \\
 &= \frac{\sigma_x^2 \sigma_\theta^2 T + \sigma_x^4}{\sigma_\lambda^2 \sigma_\theta^2 T^2 I + \sigma_\theta^4 T^2 + \sigma_\lambda^2 \sigma_x^2 T I + 2\sigma_\theta^2 \sigma_x^2 T + \sigma_x^4} \\
 &= \frac{\left(\frac{1}{\sigma_x^2}\right) \sigma_\theta^2 T + 1}{\left(\frac{1}{\sigma_x^4}\right) \sigma_\lambda^2 \sigma_\theta^2 T^2 I + \left(\frac{1}{\sigma_x^4}\right) \sigma_\theta^4 T^2 + \sigma_\lambda^2 \left(\frac{1}{\sigma_x^2}\right) T I + 2\sigma_\theta^2 \left(\frac{1}{\sigma_x^2}\right) T + 1} \\
 &\rightarrow \frac{1}{1} = 1
 \end{aligned}$$

Alors, lorsque $\sigma_x^2 \rightarrow \infty$, nous avons $\omega_{\theta, \lambda} \rightarrow 1$.

P5) Lorsque $\sigma_\theta^2 \rightarrow \infty$, nous avons $\omega_1 \rightarrow 1$, $\omega_{i \neq 1} \rightarrow 0$ et $\omega_{\theta, \lambda} \rightarrow 0$ car

$$\begin{aligned}
 \omega_1 &= \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 T + \sigma_x^4} \\
 &= \frac{\sigma_\lambda^2 \sigma_\theta^2 T^2 I + \sigma_\theta^4 T^2 + \sigma_\lambda^2 \sigma_x^2 T + \sigma_\theta^2 \sigma_x^2 T}{\sigma_\lambda^2 \sigma_\theta^2 T^2 I + \sigma_\theta^4 T^2 + \sigma_\lambda^2 \sigma_x^2 T I + 2\sigma_\theta^2 \sigma_x^2 T + \sigma_x^4} \\
 &\longrightarrow \frac{T^2}{T^2} = 1
 \end{aligned}$$

D'où on peut conclure que, lorsque $\sigma_\theta^2 \rightarrow \infty$, alors $\omega_1 \rightarrow 1$.

$$\begin{aligned}
 \omega_{i \neq 1} &= \frac{(\sigma_\lambda^2 \sigma_x^2 T I - \sigma_\lambda^2 \sigma_x^2 T) \left(\frac{1}{\sigma_\theta^4} \right)}{\left(\frac{1}{\sigma_\theta^2} \right) \sigma_\lambda^2 T^2 I + T^2 + \sigma_\lambda^2 \sigma_x^2 T I \left(\frac{1}{\sigma_\theta^4} \right) + 2 \left(\frac{1}{\sigma_\theta^2} \right) \sigma_x^2 T + \left(\frac{1}{\sigma_\theta^4} \right) \sigma_x^4} \\
 &\longrightarrow \frac{0}{T^2} = 0
 \end{aligned}$$

D'où on peut conclure que, lorsque $\sigma_\theta^2 \rightarrow \infty$, alors $\omega_{i \neq 1} \rightarrow 0$.

$$\begin{aligned}
 \omega_{\theta, \lambda} &= \frac{\sigma_\theta^2 \sigma_x^2 T + \sigma_x^4}{\sigma_\theta^2 \sigma_\lambda^2 T^2 I + \sigma_\theta^4 T^2 + \sigma_x^2 \sigma_\lambda^2 T I + 2\sigma_\theta^2 \sigma_x^2 T + \sigma_x^4} \\
 &= \frac{\sigma_x^2 \left(\frac{1}{\sigma_\theta^2} \right) T + \sigma_x^4 \left(\frac{1}{\sigma_\theta^4} \right)}{\sigma_\lambda^2 \left(\frac{1}{\sigma_\theta^2} \right) T^2 I + T^2 + \sigma_x^2 \sigma_\lambda^2 T I \left(\frac{1}{\sigma_\theta^4} \right) + 2 \left(\frac{1}{\sigma_\theta^2} \right) \sigma_x^2 T + \left(\frac{1}{\sigma_\theta^4} \right) \sigma_x^4} \\
 &\longrightarrow \frac{0}{T^2} = 0
 \end{aligned}$$

Donc on peut en déduire que, lorsque $\sigma_\theta^2 \rightarrow \infty$, alors $\omega_{\theta, \lambda} \rightarrow 0$.

P6) Lorsque $\sigma_\lambda^2 \rightarrow \infty$, nous avons $\omega_1 \rightarrow \frac{\sigma_\theta^2 I T + \sigma_x^2}{\sigma_\theta^2 I T + \sigma_x^2 I}$, $\omega_{i \neq 1} \rightarrow \frac{\sigma_x^2 (I-1)}{\sigma_\theta^2 I T + \sigma_x^2 I}$ et $\omega_{\theta, \lambda} \rightarrow 0$

car

$$\begin{aligned}
 \omega_1 &= \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 T + \sigma_x^4} \\
 &= \frac{\sigma_\lambda^2 \sigma_\theta^2 T^2 I + \sigma_\theta^4 T^2 + \sigma_\lambda^2 \sigma_x^2 T + \sigma_\theta^2 \sigma_x^2 T}{\sigma_\lambda^2 \sigma_\theta^2 T^2 I + \sigma_\theta^4 T^2 + \sigma_\lambda^2 \sigma_x^2 T I + 2\sigma_\theta^2 \sigma_x^2 T + \sigma_x^4} \\
 &= \frac{\sigma_\theta^2 T^2 I + \left(\frac{1}{\sigma_\lambda^2}\right) \sigma_\theta^4 T^2 + \sigma_x^2 T + \sigma_\theta^2 \sigma_x^2 T \left(\frac{1}{\sigma_\lambda^2}\right)}{\sigma_\theta^2 T^2 I + \sigma_\theta^4 T^2 \left(\frac{1}{\sigma_\lambda^2}\right) + \sigma_x^2 T I + 2\sigma_\theta^2 \sigma_x^2 T \left(\frac{1}{\sigma_\lambda^2}\right) + \sigma_x^4 \left(\frac{1}{\sigma_\lambda^2}\right)} \\
 &\rightarrow \frac{(\sigma_\theta^2 I T + \sigma_x^2) T}{(\sigma_\theta^2 I T + \sigma_x^2 I) T} = \frac{\sigma_\theta^2 I T + \sigma_x^2}{\sigma_\theta^2 I T + \sigma_x^2 I}
 \end{aligned}$$

Donc on peut en déduire que, lorsque $\sigma_\lambda^2 \rightarrow \infty$, alors $\omega_1 \rightarrow \frac{\sigma_\theta^2 I T + \sigma_x^2}{\sigma_\theta^2 I T + \sigma_x^2 I}$.

$$\begin{aligned}
 \omega_{i \neq 1} &= \frac{T(I-1) \sigma_\lambda^2 \sigma_x^2}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)} \\
 &= \frac{T I \sigma_x^2 - T \sigma_x^2}{\sigma_\theta^2 T^2 I + \left(\frac{1}{\sigma_\lambda^2}\right) \sigma_\theta^4 T^2 + T I \sigma_x^2 + 2\sigma_\theta^2 \sigma_x^2 T \left(\frac{1}{\sigma_\lambda^2}\right) + \sigma_x^4 \left(\frac{1}{\sigma_\lambda^2}\right)} \\
 &\rightarrow \frac{(I-1) T \sigma_x^2}{(\sigma_\theta^2 T I + \sigma_x^2 I) T} = \frac{(I-1) \sigma_x^2}{\sigma_\theta^2 T I + \sigma_x^2 I}
 \end{aligned}$$

Donc on peut en déduire que, lorsque $\sigma_\lambda^2 \rightarrow \infty$, alors $\omega_{i \neq 1} \rightarrow \frac{\sigma_x^2 (I-1)}{\sigma_\theta^2 T I + \sigma_x^2 I}$.

$$\begin{aligned}
 \omega_{\theta, \lambda} &= \frac{\sigma_\theta^2 \sigma_x^2 T \left(\frac{1}{\sigma_\lambda^2}\right) + \sigma_x^4 \left(\frac{1}{\sigma_\lambda^2}\right)}{\sigma_\theta^2 T^2 I + \sigma_\theta^4 T^2 \left(\frac{1}{\sigma_\lambda^2}\right) + \sigma_x^2 T I + 2\sigma_\theta^2 \sigma_x^2 T \left(\frac{1}{\sigma_\lambda^2}\right) + \sigma_x^4 \left(\frac{1}{\sigma_\lambda^2}\right)} \\
 &\rightarrow \frac{0}{\sigma_\theta^2 T^2 I + \sigma_x^2 T I} = 0
 \end{aligned}$$

D'où on peut conclure que, lorsque $\sigma_\lambda^2 \rightarrow \infty$, alors $\omega_{\theta, \lambda} \rightarrow 0$.

P6) i) (évident)

P6) ii) (évident)

P6) iii) Comme $\sigma_\lambda^2 \rightarrow \infty$, de plus si $\sigma_x^2 \rightarrow \infty$, alors $\omega_1 \rightarrow \frac{1}{I}$, $\omega_{i \neq 1} \rightarrow \frac{I-1}{I}$, $\omega_{\theta, \lambda} \rightarrow 0$ et $\mu_{1, T+1} \rightarrow \bar{\bar{X}}$ car : on sait que la prime de crédibilité

$$\begin{aligned} \mu_{1, T+1} &= \omega_1 \bar{X}_1 + \omega_{i \neq 1} \bar{\bar{X}}_{i \neq 1} + \omega_{\theta, \lambda} (\mu_\theta + \mu_\lambda) \\ &= \frac{1}{I} \left(\frac{1}{T} \sum_{t=1}^T X_{1,t} \right) + \frac{I-1}{I} \left(\frac{1}{(I-1)T} \sum_{i=2}^I \sum_{t=1}^T X_{i,t} \right) \\ &= \frac{1}{TI} \sum_{t=1}^T X_{1,t} + \frac{1}{TI} \sum_{i=2}^I \sum_{t=1}^T X_{i,t} \\ &= \frac{1}{TI} \sum_{i=1}^I \sum_{t=1}^T X_{i,t} \\ &= \bar{\bar{X}}, \text{ qui est une moyenne globale.} \end{aligned}$$

BIBLIOGRAPHIE

- [1] Arafa, G., Metiri, F., Sadoun, A., & Remita, M. R. (2024). On Quantile Credibility Estimators Under An Equal Correlation Structure Over Risks. *Statistics, Optimization & Information Computing*, 12(5), 1325-1341.
- [2] Bailey, A. L. (1950). Credibility Procedures : Laplace's generalization of Bayes' Rule and the combination of collateral knowledge with observed data.
- [3] Blomqvist, N. (1950). On a measure of dependence between two random variables. *The Annals of Mathematical Statistics*, 593-600.
- [4] Bolancé, C., Guillén, M., & Pinquet, J. (2003). Time-varying credibility for frequency risk models : estimation and tests for autoregressive specifications on the random effects. *Insurance : Mathematics and Economics*, 33(2), 273-282.
- [5] Bozikas, A., & Pitselis, G. (2020). Incorporating crossed classification credibility into the Lee–Carter model for multi-population mortality data. *Insurance : Mathematics and Economics*, 93, 353-368.
- [6] Bühlmann, H. (1967). Experience rating and credibility. *ASTIN Bulletin : The Journal of the IAA*, 4(3), 199-207.
- [7] Bühlmann, H., & Gisler, A. (2005). A course in credibility theory and its applications. Berlin, Heidelberg : Springer Berlin Heidelberg.

-
- [8] Charpentier, A., & Segers, J. (2007). Lower tail dependence for Archimedean copulas : characterizations and pitfalls. *Insurance : Mathematics and Economics*, 40(3), 525-532.
- [9] Clauss, P., & des Risques, F. G. (2011). Théorie des copules. Notes de cours, ENSAI.
- [10] Ebrahimzadeh, M., Ibrahim, N. A., Jemain, A. A., & Kilicman, A. (2013). Claim dependence induced by common effects in hierarchical credibility models. *Communications in Statistics-Theory and Methods*, 42(18), 3373-3400.
- [11] Embrechts, P., Lindskog, F., & McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. *Handbook of heavy tailed distributions in finance*, 8(1), 329-384.
- [12] Frees, E. W., & Valdez, E. A. (1998). Understanding relationships using copulas. *North American actuarial journal*, 2(1), 1-25.
- [13] Frees, E. W., Young, V. R., & Luo, Y. (1999). A longitudinal data analysis interpretation of credibility models. *Insurance : Mathematics and Economics*, 24(3), 229-247.
- [14] Frees, E. W., Young, V. R., & Luo, Y. (2001). Case studies using panel data models. *North American Actuarial Journal*, 5(4), 24-42.
- [15] Gómez-Déniz, E., Perez-Sanchez, J. M., & Vázquez-Polo, F. J. (2006). On the use of posterior regret Γ -minimax actions to obtain credibility premiums. *Insurance : Mathematics and Economics*, 39(1), 115-121.
- [16] Goovaerts, M. J., Kaas, R., Van Heerwaarden, A. E., & Bauwelinckx, T. (1990). *Effective actuarial methods*. AmsterdamNorth-Holland.
- [17] Heilmann, W. R. (1989). Decision theoretic foundations of credibility theory. *Insurance : Mathematics and Economics*, 8(1), 77-95.
- [18] Herbert, A. D., & Nagaraja, H. N. (2003). *Order statistics*. Series in probability and statistics (3rd ed.). Hoboken : Wiley.
- [19] Herzog, T. N. (1996). *Introduction to credibility theory* (Winsted, Connecticut, Actex Publications Inc.

-
- [20] Joe, H. (2014). Dependence modeling with copulas. CRC press.
- [21] Kahn, P. M. (Ed.). (1975). Credibility : Theory and Applications : Proceedings of the Berkeley Actuarial Research Conference on Credibility, September 19-21, 1974, the University of California, Berkley [sic]. Academic Press.
- [22] Landsman, Z., & Makov, U. E. (1999). Credibility evaluation for the exponential dispersion family. *Insurance : Mathematics and Economics*, 24(1-2), 23-29.
- [23] Leong, Y. K., & Valdez, E. A. (2005). Claims prediction with dependence using copula models. *Insurance : Mathematics and Economics*.
- [24] Mowbray, A. H. (1914). How extensive a payroll exposure is necessary to give a dependable pure premium ? *Proceedings of the Casualty Actuarial Society*, 1, 24-30.
- [25] Nelsen, R. B. (2006). An introduction to copulas. New York, NY : Springer New York.
- [26] Oyakapelia, O., Ikamarib, C., & Karanjaha, A. (2024). Premium Estimation of General Insurance Contracts Using Buhlmann Credibility Model. *Asian Journal of Probability and Statistics*, 26(5), 64-71.
- [27] Parzen, E. (1979). Nonparametric statistical data modeling. *Journal of the American statistical association*, 74(365), 105-121.
- [28] Pitselis, G. (2013). Quantile credibility models. *Insurance : Mathematics and Economics*, 52(3), 477-489.
- [29] Pitselis, G. (2016). Credible risk measures with applications in actuarial sciences and finance. *Insurance : Mathematics and Economics*, 70, 373-386.
- [30] Pitselis, G. (2017). Risk measures in a quantile regression credibility framework with Fama/French data applications. *Insurance : Mathematics and Economics*, 74, 122-134.
- [31] Pitselis, G. (2020). Multi-stage nested classification credibility quantile regression model. *Insurance : Mathematics and Economics*, 92, 162-176.
- [32] Promislow, S. D., & Young, V. R. (2000). Equity and exact credibility. *ASTIN Bulletin : The Journal of the IAA*, 30(1), 3-11.

-
- [33] Purcaru, O., & Denuit, M. (2002). On the dependence induced by frequency credibility models. *Belgian Actuarial Bulletin*, 2(1), 73-79.
- [34] Purcaru, O., & Denuit, M. (2003). Dependence in dynamic claim frequency credibility models. *ASTIN Bulletin : The Journal of the IAA*, 33(1), 23-40.
- [35] Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. In *Annales de l'ISUP* (Vol. 8, No. 3, pp. 229-231).
- [36] Wang, W., Wen, L., Yang, Z., & Yuan, Q. (2020). Quantile credibility models with common effects. *Risks*, 8(4), 100.
- [37] Wen, L., & Deng, W. (2011). The credibility models with equal correlation risks. *Journal of Systems Science and Complexity*, 24(3), 532-539.
- [38] Wen, L., Wu, X., & Zhou, X. (2009). The credibility premiums for models with dependence induced by common effects. *Insurance : Mathematics and Economics*, 44(1), 19-25.
- [39] Whitney, A. W. (1918). The theory of experience rating. *Proceedings of the Casualty Actuarial Society*, 4, 274-292.
- [40] Yeo, K. L., & Valdez, E. A. (2006). Claim dependence with common effects in credibility models. *Insurance : Mathematics and Economics*, 38(3), 609-629.
- [41] Young, V. R. (2000). Credibility using semiparametric models and a loss function with a constancy penalty. *Insurance : Mathematics and Economics*, 26(2-3), 151-156.
- [42] Zhang, Q., & Chen, P. (2018). Credibility estimators with dependence structure over risks and time under balanced loss function. *Statistica Neerlandica*, 72(2), 157-173.