

Lyon 1

Université Claude Bernard – Lyon 1

INSTITUT DE SCIENCE FINANCIERE ET D'ASSURANCES



Mémoire présenté
devant l'Institut de Science Financière et d'Assurances
pour l'obtention du diplôme d'Actuaire de l'Université de Lyon
le 5 décembre 2012

Par : Clémence MICHAUD

Titre : La mesure du risque de provisionnement à horizon un an en assurance non-vie -
Prise en compte du jugement d'expert et de données exogènes

Confidentialité : ☒ NON ☐ OUI (Durée : ☐ 1 an ☐ 2 ans)

Membres du jury de l'Institut des Actuares

M. Frédéric PLANCHET

M. Pierre THEROND

Entreprise :

FIXAGE

Membres du jury I.S.F.A.

Mme Flavia BARSOTTI

M. Alexis BIENVENÛE

M. Areski COUSIN

Mme Diana DOROBANTU

Mme Anne EYRAUD-LOISEL

M. Nicolas LEBOISNE

M. Stéphane LOISEL

Mlle Esterina MASIELLO

Mme Véronique MAUME-DESCHAMPS

M. Frédéric PLANCHET

Mme Béatrice REY-FOURNIER

M. Pierre RIBEREAU

M. Christian-Yann ROBERT

M. Didier RULLIERE

M. Pierre THEROND

Directeur de mémoire en entreprise :

M. Emmanuel TASSIN

Invité :

**Autorisation de mise en ligne sur
un site de diffusion de documents
actuariels (après expiration de
l'éventuel délai de confidentialité)**

Signature du responsable entreprise

Secrétariat

Mme Marie-Claude MOUCHON

Signature du candidat

Bibliothèque :

Mme Patricia BARTOLO

Résumé

Mots-clés :

Solvabilité II, provisionnement non-vie, incertitude à un an, Chain Ladder, Bornhuetter-Ferguson, facteur de queue de développement, stochastique, bootstrap, Merz Wüthrich.

Les compagnies d'assurance doivent désormais répondre à de nouvelles exigences réglementaires introduites par la directive Solvabilité II. En assurance non-vie, les provisions pour sinistres à payer sont souvent les plus importantes en montant et nécessitent, par conséquent, une évaluation très précise. Celles-ci correspondent aux sinistres survenus à la date d'inventaire mais qui n'ont pas encore été réglés. Jusqu'à maintenant, on s'intéressait à la volatilité à l'ultime des provisions ainsi constituées, c'est-à-dire à l'incertitude à long terme liée au déroulement complet de la charge sinistre. Il s'agit dorénavant de mesurer le risque de provisionnement à un an. On entend par là la variabilité de la prédiction de la charge ultime à la date $t = I$ par rapport à sa valeur un an plus tard, en $t = I + 1$. Ce dernier point constituera le principal objet d'étude de ce mémoire.

D'autre part, le provisionnement est souvent affaire de choix ; lorsque l'actuaire retient une méthode d'évaluation, cela a nécessairement une conséquence importante sur le niveau des provisions à constituer. C'est ainsi que le savoir et l'expérience de l'actuaire, autrement dit le jugement d'expert, interviennent dans le processus d'évaluation des provisions techniques. La méthode Bornhuetter-Ferguson permet de surcroît d'introduire ce jugement d'expert dans le calcul même des provisions. Elle possède également un avantage majeur, à savoir la possibilité de prendre en compte des données exogènes qui mesurent l'exposition au risque, comme les primes. C'est pour l'ensemble de ces raisons que cette méthode sera particulièrement présente dans ce mémoire. Le jugement d'expert permet également à l'actuaire d'utiliser ses connaissances dans certaines branches dites à développement long : en effet, il pourra juger d'une éventuelle insuffisance des provisions calculées et insérer ainsi un facteur de queue de développement. Désormais, la question à se poser est : comment intégrer la prise en compte de données exogènes et du jugement d'expert dans la mesure du risque de volatilité à horizon un an ?

Abstract

Key words :

Solvency II, non-life reserving, uncertainty on a one-year period, Chain Ladder, Bornhuetter-Ferguson, tail factor, stochastique, bootstrap, Merz Wüthrich.

The recent Solvency II Directive introduced many regulatory requirements that insurers have now to face. For a non-life insurer and in terms of volume, the most important reserves to be constituted are claims reserves. That is why those reserves need a very accurate estimation. They correspond to claims that occurred before the date of inventory but have not been paid yet. So far, insurers were interested in the ultimate volatility of the claims reserves, that is to say in the uncertainty on the long run linked to the complete development of the claims payments. The objective is now to measure the reserving risk on a one-year horizon. This means the variability of the ultimate payment prediction at $t = I$ compared to its value one year later, at $t = I + 1$. This last point will be the main subject of this study.

Besides, reserving often stems from the choices the actuary has made ; indeed, when the latter selects an estimation method, this will necessarily lead to an important consequence on the level of the reserves to constitute. That is how the actuary's knowledge and experience, that is to say his expert judgement, are involved in the estimation process of the technical reserves. What is more, the Bornhuetter-Ferguson method enables to introduce this expert judgement in the reserves' calculation. It has also a major advantage, namely the possibility to take into account exogenous data that assess the risk exposure, such as premiums. For all these reasons, this method will be specially studied in this paper. The expert judgement also enables the actuary to use his knowledge as far as long-tail development lines of business are concerned: indeed, he can notice a possible insufficiency of the estimated reserves and, as a consequence, insert a tail factor. From now on, the question to be asked is: how to deal with both exogenous data and expert judgement in the assessment of the one-year horizon volatility risk?

Table des matières

Résumé	1
Abstract	2
Introduction	4
I Le provisionnement en assurance non-vie	6
1 La vie d'un sinistre	6
1.1 Le cycle de production en assurance non-vie	6
1.2 La liquidation du sinistre	7
2 La provision pour sinistres à payer dans le référentiel Solvabilité I	9
3 La nécessité de mesurer le risque de provisionnement à horizon un an dans le référentiel Solvabilité II	10
3.1 L'évaluation des provisions techniques	11
3.2 L'exigence de capital réglementaire sous Solvabilité II	12
3.3 Le risque de provisionnement à horizon un an	13
II Les conséquences des choix de l'actuaire sur la mesure de l'incertitude à l'ultime des provisions	15
4 L'approche Chain Ladder et le jugement d'expert	16
4.1 L'estimation des provisions	16
4.1.1 La méthode Chain-Ladder	16
4.1.2 La détermination d'un éventuel facteur de queue de développement	17
4.2 La mesure de l'incertitude à l'ultime	18
4.2.1 La mesure de la variance des provisions par la méthode de Mack	18
4.2.2 La mesure de l'incertitude à partir d'une distribution complète des provisions par la méthode bootstrap	22
5 La prise en compte de données exogènes par l'approche Bornhuetter-Ferguson	28
5.1 L'estimation des provisions	28
5.1.1 La méthode Bornhuetter-Ferguson	28
5.1.2 La méthode Benktander	29
5.2 La mesure de l'incertitude à l'ultime des provisions	29
5.2.1 La méthode de Mack (2008)	29
5.2.2 La méthode d'Alai, Merz et Wüthrich	32
III L'influence du jugement d'expert et de données exogènes sur la vision à un an du risque de provisionnement	35

6	L'incertitude à un an par une formule fermée	36
6.1	La formule de Merz et Wüthrich	36
6.2	L'insertion d'un facteur de queue de développement	45
7	La mesure de l'incertitude à partir d'une distribution complète du <i>Claims Development Result</i>	47
7.1	La méthode bootstrap	47
7.1.1	La méthode bootstrap de re-provisionnement	47
7.1.2	La méthode bootstrap améliorée	49
7.1.3	La méthode récursive	50
7.2	L'insertion d'un facteur de queue de développement	55
7.2.1	La méthode de re-provisionnement améliorée	55
7.2.2	La méthode récursive	55
7.3	La prise en compte de données exogènes	56
7.3.1	La méthode de re-provisionnement améliorée suivant l'ap- proche Bornhuetter-Ferguson	56
7.3.2	La méthode de re-provisionnement améliorée suivant l'ap- proche Benktander	57
8	La détermination du capital réglementaire	60
8.1	La détermination du SCR à partir d'une formule fermée	60
8.2	La détermination du SCR par une approche simulatoire	60
IV	La détermination de l'incertitude sur les provisions pour des données en assurance non-vie	62
9	Présentations des données	62
9.1	Les triangles de liquidation	62
9.2	Les données exogènes aux triangles	65
10	La mesure de l'incertitude à l'ultime de la provision pour sinistres à payer	65
10.1	L'approche Chain Ladder de la provision estimée	66
10.1.1	Les méthodes analytiques Chain Ladder et Mack	66
10.1.2	Les méthodes par simulations bootstrap	70
10.2	L'approche Bornhuetter-Ferguson du risque de provisionnement	76
10.2.1	L'estimation de la PSAP par les méthodes déterministes choisies	77
10.2.2	La mesure de l'incertitude à l'ultime de la PSAP par les mé- thodes stochastiques	77
11	La vision à un an du risque de provisionnement	88
11.1	La méthode de référence Merz-Wüthrich	88
11.2	La méthode bootstrap de re-provisionnement	91
11.3	La méthode bootstrap récursive	94
12	La détermination du capital réglementaire	98
	Conclusion	101
	Remerciements	103

Annexes	106
A. La méthode Delta	106
B. La méthode d'Alai, Merz et Wüthrich	108
C. La méthode bootstrap de De Felice et Moriconi	110
D. L'équivalence des résultats de la méthode récursive avec ceux de la méthode analytique de Merz et Wüthrich	112
E. La formule fermée pour le calcul du capital réglementaire en assurance non-vie	117
F. Les tests de normalité utilisés	119

Introduction

L'assurance est une activité économique singulière dont la spécificité est l'inversion du cycle de production. En effet, l'organisme assureur touche une prime de la part de l'assuré et s'engage, en contrepartie, à lui verser une prestation en cas de réalisation d'un risque aléatoire prévu au contrat. L'assuré verse donc dans un premier temps une prime à l'assureur et, si un sinistre survient pendant la période de couverture, l'assureur règle le sinistre. Pour honorer ses engagements envers l'assuré en cas de réalisation du sinistre, l'assureur se doit alors de constituer des provisions au passif de son bilan. Il existe différents types de provisions selon la branche d'activité considérée et le risque couvert. Or, dans ce mémoire, on se place dans le cadre de l'assurance non-vie. Les provisions pour sinistres à payer sont ainsi souvent les plus importantes en montant et nécessitent, par conséquent, une évaluation très précise ; elles correspondent aux sinistres survenus à la date d'inventaire mais qui n'ont pas encore été réglés.

Ces provisions sont des outils importants de solvabilité, de gestion du capital ou encore de comptabilité. Il était donc nécessaire d'instaurer un certain cadre réglementaire. La Directive Solvabilité II (2009) est la récente réforme du monde de l'assurance. Des règles de solvabilité communes aux pays de l'Union européenne, une meilleure prise en compte du profil de risque de l'assureur, une protection accrue des consommateurs sont autant d'objectifs de ce nouveau système qui s'apprête à modifier en profondeur le marché assurantiel. C'est d'ailleurs la raison pour laquelle ce sujet est à la fois technique et politique. Comme le rappelle B. Spitz¹, Président de la Fédération Française des Sociétés d'Assurances (FFSA), une forte adaptation est en jeu pour les assureurs, que ce soit au niveau de la gestion des risques et des besoins en capitaux que de la compétitivité mondiale. L'objectif est ainsi de renforcer la protection des assurés et d'éviter de trop importantes hausses de tarifs tout en finançant l'économie et la croissance. Cela passe notamment par une juste appréciation des risques, une simplification des propositions et des mesures de transition appropriées.

Jusqu'à maintenant, les assureurs s'intéressaient uniquement à la variabilité à long terme des provisions constituées à la date d'inventaire. En effet, de manière assez intuitive, on voulait connaître l'incertitude liée au déroulement complet de la charge sinistre. Désormais, la Directive Solvabilité II instaure une nouvelle vision du risque lié au provisionnement en effectuant un profond changement de perspective : dorénavant, les assureurs doivent également mesurer ce risque à horizon un an. Il s'agit alors d'étudier la variabilité de la prédiction de la charge ultime à la date $t = I$ par rapport à sa valeur un an plus tard, en $t = I + 1$. Ce dernier point constituera le principal objet d'étude de ce mémoire.

Les méthodes de provisionnement évoluent à la fois pour répondre à cette nouvelle exigence réglementaire et en vue de satisfaire également à certains critères. En effet, de nombreuses méthodes ont vu le jour pour s'adapter à ce que J.M. Nessi² nomme « critères de performance » : données fiables ou nombreuses, branche peu volatile, etc. Les méthodes progressent pour faire face à la non-réalisation d'un ou

1. Dans l'article *Solvabilité II, cinq conditions* [33].

2. Dans la formation SEPIA *Méthodes de provisionnement en assurance non-vie* [27].

plusieurs de ces critères. Chain Ladder est incontestablement la méthode la plus utilisée en assurance non-vie ; elle est devenue une référence pour la plupart des assureurs qui voient en elle une méthode simple et relativement efficace par rapport à ces critères de performance. D'ailleurs, les modèles de provisionnement relativement innovateurs sont souvent appliqués à la méthode Chain Ladder dans un premier temps. On citera notamment Mack pour la mesure de la volatilité des provisions à l'ultime, et Merz et Wüthrich pour la détermination de l'incertitude des provisions à horizon un an.

Toutefois, on est en droit de se demander si la méthode Chain Ladder est la plus appropriée dans la plupart des cas. Le provisionnement est souvent affaire de choix et, selon Hans Bühlmann³, « le provisionnement n'est pas un problème de modélisation complexe et sophistiqué, mais plutôt un exercice de choix de modèle ». En effet, lorsque l'actuaire retient une méthode d'évaluation, cela a nécessairement une conséquence importante sur le niveau des provisions à constituer. C'est ainsi que le savoir et l'expérience de l'actuaire, autrement dit le jugement d'expert, interviennent dans le processus d'évaluation des provisions techniques. La méthode Bornhuetter-Ferguson permet de surcroît d'introduire ce jugement d'expert dans le calcul même des provisions, en choisissant notamment un ratio sinistres à primes attendu. Elle possède également un avantage majeur, à savoir la possibilité de prendre en compte des données exogènes qui mesurent l'exposition au risque, comme les primes. C'est pour l'ensemble de ces raisons que cette méthode sera particulièrement présente dans ce mémoire. Le jugement d'expert permet également à l'actuaire d'utiliser ses connaissances dans certaines branches dites à développement long : en effet, il pourra juger d'une éventuelle insuffisance des provisions calculées et insérer ainsi un facteur de queue de développement, ce qui, en pratique, a une importance capitale dans le montant des provisions à constituer. Dès lors, la question à se poser est : comment intégrer la prise en compte de données exogènes et du jugement d'expert dans la mesure du risque de provisionnement à horizon un an ?

On introduit tout d'abord le provisionnement en assurance non-vie puis les principales exigences de Solvabilité II qui s'y rapportent, notamment le risque de provisionnement à horizon un an. Dans un deuxième temps, on s'intéresse aux conséquences des choix de l'actuaire sur la mesure de l'incertitude à l'ultime des provisions. En effet, on cherche à mesurer l'impact qu'aura la décision de l'actuaire sur la volatilité à l'ultime des provisions, selon que l'on se place dans le cadre des méthodes Chain Ladder ou Bornhuetter-Ferguson. Puis on cherche à déterminer quelle est l'influence du jugement d'expert et de la prise en compte de données exogènes sur la mesure du risque de provisionnement à un an. Enfin, on s'appuie sur des données en assurance non-vie pour illustrer, comprendre et comparer les notions introduites tout au long de ce mémoire.

3. Dans la postface du tome II de Charpentier et Denuit [10].

Première partie

Le provisionnement en assurance non-vie

On introduit dans cette partie la notion de provisionnement en assurance non-vie, et notamment la dynamique de liquidation des sinistres. En effet, l'activité d'assurance est marquée par un cycle de production inversé qui nécessite ainsi la constitution de provisions au passif de la compagnie. Si on se limitait jusqu'à présent à l'étude de la volatilité des provisions à l'ultime, c'est-à-dire jusqu'au déroulement complet de la charge de sinistre, il s'agit désormais d'évaluer le risque de provisionnement à un an.

1 La vie d'un sinistre

On présente dans ce qui suit la dynamique de la vie d'un sinistre, ainsi que la représentation classique de la liquidation des sinistres.

1.1 Le cycle de production en assurance non-vie

L'activité d'assurance est marquée par une importante spécificité, à savoir l'inversion du cycle de production. Un contrat d'assurance est une convention établie entre deux parties : l'assureur promet à une personne appelée l'assuré, moyennant une prime, de lui verser une prestation garantie ou une indemnité d'assurance en cas de réalisation d'un risque aléatoire prévu au contrat ou de réalisation du sinistre. L'assuré verse d'abord une prime à l'assureur et, si un sinistre survient pendant la période de couverture, l'assureur règle le sinistre. Toutefois, le processus d'indemnisation n'est pas instantané et, au contraire, s'étale dans le temps. Le schéma ci-dessous illustre les étapes de la vie d'un sinistre.

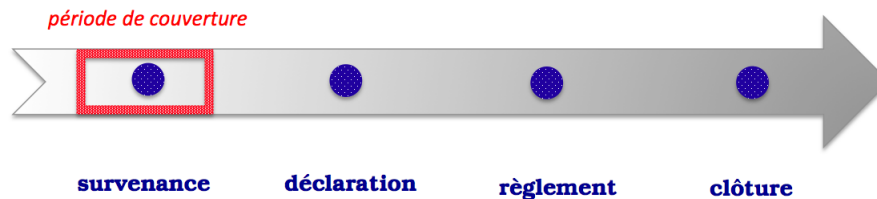


FIGURE 1 – La dynamique de la vie d'un sinistre.

Si un sinistre survient pendant la période de couverture, l'assuré le déclare à l'assureur qui effectue des remboursements. Le sinistre est clôturé une fois qu'il n'y a plus de règlements. Le montant de ces paiements est estimé par la compagnie d'assurance à partir de différents critères. Mais cela ne concerne que les sinistres survenus et effectivement déclarés à la compagnie. La difficulté réside alors dans

l'estimation des règlements au titre des sinistres survenus mais non encore déclarés à la compagnie à la date de l'inventaire, comme l'exige la réglementation. On parle alors de sinistres tardifs, ou sinistres *Incurred But Not Reported (IBNR)*. Les sinistres survenus après la date d'inventaire, qu'ils soient connus ou non, doivent également être sujets à une évaluation.

Ainsi, l'organisme assureur doit constituer des provisions afin d'honorer ses engagements envers les assurés et les bénéficiaires des contrats. Ces provisions permettent donc de tenir compte de l'inversion du cycle de production et de régler les prestations futures. Pour évaluer le montant des provisions, on présente souvent les données sous forme triangulaire.

1.2 La liquidation du sinistre

Les données sont très souvent présentées sous la forme d'un triangle qui reflète la dynamique des sinistres. On introduit les notations usuelles liées à ce triangle dit de liquidation, ou triangle *run-off*, et qui seront utilisées dans ce mémoire.

- L'exercice de rattachement de la sinistralité est désigné par la lettre i , $i \in \{1, \dots, I\}$, et il correspond, pour la plupart des contrats, à l'année de survenance du sinistre. La prime acquise est la part de la prime destinée à couvrir les risques afférents à l'exercice considéré. Dans d'autres cas, on peut parler d'année de souscription ou de déclaration.
- Les années de développement, ou de déroulé des sinistres sont désignées par la lettre j , $j \in \{1, \dots, J\}$.
- $I = J$ car on se place dans le cadre d'un triangle de liquidation.
- $X_{i,j}$ désigne le paiement décumulé ou incrémental réglé l'année de développement j pour un sinistre survenu l'année i ,
- $C_{i,j}$ désigne les règlements cumulés en j années de développement pour un sinistre survenu l'année i .

On note $\mathcal{D}_I$ l'information contenue dans ce triangle. $\mathcal{D}_{I+1}$ correspond à l'information comprise dans le triangle auquel une diagonale supplémentaire a été ajoutée. Cette nouvelle diagonale représente les règlements effectués pendant l'année comptable $(I, I + 1]$ au titre des années de survenance 2 à I .

Le schéma suivant illustre ces notations et on propose un exemple.

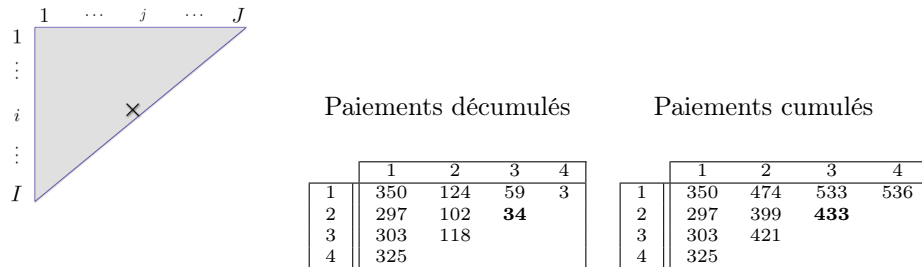


FIGURE 2 – Les notations d'un triangle de liquidation théorique et un exemple introductif.

La croix figurant sur le triangle peut désigner le montant incrémental $X_{i,j}$ ou le montant cumulé $C_{i,j}$ des règlements incrémentaux de la même année de survenance i et après j années de déroulé, selon que l'on se place dans le cadre d'un triangle décumulé ou cumulé. Si on prend l'exemple ci-dessus, on observe qu'un montant de 34 a été réglé au titre des sinistres survenus l'année 2 la troisième année de développement. Parallèlement, on peut dire qu'un montant de 433 a été réglé au titre de ces mêmes sinistres au cours des trois années de développement. Les paiements décumulés de la première colonne sont ceux effectués la même année que la survenance des sinistres. D'autre part, on appelle charge ultime de l'année de survenance i le montant des règlements cumulés au bout d'un certain nombre d'années de développement jugé suffisant pour la clôture du sinistre. Dans ce triangle, la charge ultime de l'année 1 est de 536. En revanche, à la fin de l'année 4, on ne connaît pas encore les charges ultimes afférentes aux années de survenance 2 à 4. Les méthodes de provisionnement permettent alors d'estimer ces montants et de pouvoir en déduire la provision correspondante. La provision pour l'année de survenance i correspond à la différence entre ce que l'on s'attend à payer au total pour cette année de survenance, à savoir la charge ultime estimée, et ce que l'on a déjà réglé, c'est-à-dire le montant cumulé figurant sur la dernière diagonale connue $C_{i,J-i+1}$.

Il existe différentes approches dans les méthodes de provisionnement. Les méthodes liquidatives reposent sur des facteurs de développement f_j et des cadences de règlement p_j définis de la manière suivante :

$$f_j = \frac{\sum_{i=1}^{I-j} C_{i,j+1}}{\sum_{i=1}^{I-j} C_{i,j}} \text{ et } p_j = \frac{\sum_{i=1}^{I-j} C_{i,j}}{\sum_{i=1}^{I-j} C_{i,J}} \text{ pour } j \in \{1, \dots, J-1\}.$$

Il existe une correspondance entre ces deux types de paramètres. En effet,

$$f_j = \frac{p_{j+1}}{p_j} \text{ et } p_j = \frac{1}{f_j \times f_{j+1} \times \dots \times f_{J-1}}.$$

D'autres méthodes s'appuient quant à elles sur les ratios sinistres à primes, ou *loss ratio*. Raisonner en termes de coût moyen constitue également un autre axe de méthodes.

Les triangles de liquidation ne sont pas toujours homogènes. En effet, il se peut qu'une diagonale diffère des autres, ce qui peut traduire un changement dans le processus de gestion des sinistres. Si un changement survient dans les lignes du triangles, le profil de risque de la compagnie peut avoir été modifié à partir d'une certaine année de survenance. Des sinistres dits graves peuvent également être observés à travers des règlements importants singuliers dans le triangle. Ainsi, un triangle ne ressemble pas à un autre et il s'agit toujours d'étudier les données avant d'appliquer une méthode en particulier.

D'autre part, certaines branches d'assurance ont des spécificités. En particulier, les sinistres des branches à développement long, c'est-à-dire les branches où les règlements s'effectuent pendant un grand nombre d'années, ne sont pas nécessairement clos à la fin des J années de développement présentées dans le triangle de liquidation. C'est pourquoi l'actuaire peut trouver intéressant d'insérer un facteur de queue de développement, à savoir un facteur de développement supplémentaire qui permettrait d'assurer la liquidation totale des sinistres.

On se place par la suite dans le cadre de triangles de paiements de sinistres et à la date $31/12/I$.

2 La provision pour sinistres à payer dans le référentiel Solvabilité I

Le cycle de production en assurance étant inversé, les organismes assureurs doivent répondre à des contraintes spécifiques dans le cadre du système de réglementation actuel Solvabilité I (2002). Cela concerne notamment la constitution de provisions techniques ; en effet, l'assureur doit constituer de telles provisions afin d'être en mesure d'indemniser l'assuré si un sinistre survient pendant la période de couverture future. L'article R331-1 du Code des assurances stipule que ces provisions techniques doivent être « *suffisantes pour le règlement intégral de leurs engagements vis-à-vis des assurés, des entreprises réassurées et bénéficiaires de contrats* ». Il existe différents types de provisions selon la branche d'assurance considérée et le risque couvert. Dans ce mémoire, on se place dans le cadre de l'assurance non-vie ; les provisions pour sinistres à payer sont alors très souvent les plus importantes en montant.

D'après l'article R331-6, la provision pour sinistres à payer (PSAP) est la « *valeur estimative des dépenses en principal et en frais, tant internes qu'externes, nécessaires au règlement de tous les sinistres survenus et non payés, y compris les capitaux constitutifs des rentes non encore mises à la charge de l'entreprise* ». Le Code des assurances stipule également les modalités de calcul de cette provision dans l'article R331-15 :

« *La provision pour sinistres à payer est calculée exercice par exercice. Sans préjudice de l'application des règles spécifiques à certaines branches prévues à la présente section, l'évaluation des sinistres connus est effectuée dossier par dossier, le coût d'un dossier comprenant toutes les charges externes individualisables ; elle est augmentée d'une estimation du coût des sinistres survenus mais non déclarés [ou sinistres Incurred But Not Reported (IBNR)]. La provision pour sinistres à payer doit toujours être calculée pour son montant brut, sans tenir compte des recours à exercer ; les recours à recevoir font l'objet d'une évaluation distincte. Par dérogation aux dispositions du deuxième alinéa du présent article, l'entreprise peut, avec l'accord de l'Autorité de contrôle prudentiel, utiliser des méthodes statistiques pour l'estimation des sinistres survenus au cours des deux derniers exercices.* »

La PSAP est ainsi constituée d'une provision Dossier/Dossier et d'une provision pour les IBNR, comme l'illustre le schéma de la page suivante. Cette dernière est évaluée par des méthodes statistiques, et on peut notamment utiliser le triangle de liquidation introduit précédemment pour estimer la charge de sinistres ultime. Chaque année, des observations supplémentaires viennent compléter le triangle de liquidation : une diagonale constituée du montant des règlements effectués dans l'année qui s'est écoulée s'ajoute au triangle. La charge ultime est alors ré-estimée pour tenir compte de cette nouvelle information disponible. Si l'estimation de la provision est revue à la hausse, un mali est constaté. En revanche, si cette estimation est revue à la baisse, on parle de boni pour l'organisme assureur.

Pour l'assureur, le risque lié aux provisions est alors déterminé par l'étude de la volatilité de la charge de sinistres ultime estimée. C'est en effet une vision assez intuitive du risque de provisionnement du point de vue de l'assureur. Mais cette perception du risque est désormais modifiée par la Directive Solvabilité II destinée à pallier les limites du référentiel actuel : une harmonisation de la réglementation au niveau européen et une plus grande importance du profil de risque de l'assureur

sont notamment recherchées.

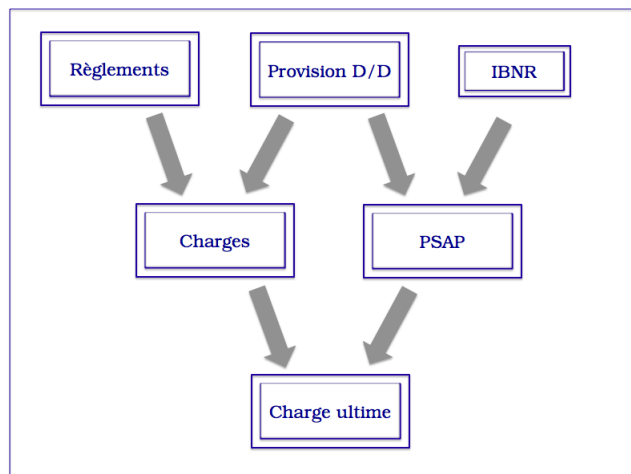


FIGURE 3 – La décomposition de la charge ultime.

3 La nécessité de mesurer le risque de provisionnement à horizon un an dans le référentiel Solvabilité II

La récente Directive Solvabilité II (2009) marque un véritable changement dans le monde de l'assurance au niveau européen. L'un des objectifs de cette réforme est une harmonisation des règles de solvabilité dans le but de mieux protéger les assurés. Les organismes assureurs doivent donc s'adapter et mieux gérer les risques couverts, tout en promouvant leur compétitivité.

Solvabilité II s'inscrit dans le cadre du processus Lamfalussy que l'Union européenne a l'habitude d'utiliser pour les textes réglementaires dans le secteur financier. Ce processus distingue trois niveaux de mesures, des grands principes jusqu'aux mesures opérationnelles et pratiques. Le texte adopté par le Parlement et le Conseil européens fait partie du premier niveau de mesures et définit les grands objectifs de la réforme. Le deuxième niveau correspond aux mesures techniques d'application et, enfin, le troisième niveau est l'ensemble des recommandations pratiques de l'*European Insurance and Occupational Pensions Authority* (EIOPA)⁴. À noter également l'existence d'études quantitatives d'impact, ou *Quantitative Impact Study* (QIS) destinées à mesurer l'impact de ces nouvelles perspectives réglementaires. Plusieurs études se sont succédées et la dernière en date, le QIS 5, a été publiée en 2011. Ces études permettent de faire le lien entre les différents acteurs du marché et sont susceptibles de modifier certaines des règles proposées par l'EIOPA. Par la suite, on se base sur les hypothèses du QIS 5.

4. Anciennement le *Committee of European Insurance and Occupational Pensions Supervisors* (CEIOPS).

Les idées nouvelles de cette réforme sont en particulier une vision économique du bilan avec la notion de *fair value* ou encore le fait que la vision économique soit désormais plus importante que la forme juridique. Solvabilité II se compose de trois piliers : tout d’abord, le premier pilier correspond aux exigences quantitatives auxquelles les organismes d’assurance et de réassurance doivent répondre. Puis, le second pilier concerne les contraintes qualitatives, la gestion et la gouvernance des risques, notamment dans le cadre de l’ORSA. Enfin, le troisième pilier vise une communication financière plus importante des compagnies au public et aux autorités de contrôle, dans un souci de transparence.

Dans ce qui suit, on se concentre sur le premier pilier qui concerne à la fois les provisions techniques et les fonds propres.

3.1 L’évaluation des provisions techniques

L’organisme assureur doit constituer un capital au passif de son bilan afin d’honorer ses engagements vis-à-vis des assurés. Les provisions techniques font partie de ce capital et doivent être rattachées à des classes de risques homogènes, ou *Lines of Business* (LoB). Elles sont constituées d’une provision Best Estimate et d’une marge pour risque.

□ La provision Best Estimate

L’article 77 de la directive du parlement européen définit la provision Best Estimate, ou la meilleure estimation, comme « *la moyenne pondérée par leur probabilité des flux de trésorerie futurs, compte tenu de la valeur temporelle de l’argent (valeur actuelle attendue des flux de trésorerie futurs), estimés sur la base de la courbe des taux sans risque pertinente* ».

La provision Best Estimate brute de réassurance peut s’écrire de la manière suivante :

$$BE = \sum_{i \geq I+1} \frac{CF_i}{(1 + r_i)^i},$$

avec CF_i les cash flows de l’année comptable i et r_i le taux sans risque de maturité i obtenu grâce à la courbe des taux sans risque à la date d’évaluation. Les cash flows à considérer ici doivent être associés aux contrats actuels.

En assurance non-vie, la provision best estimate est composée d’une provision pour primes nécessaire pour couvrir les sinistres à venir sur des contrats déjà existants, et d’une provision pour sinistres afin de régler les sinistres déjà survenus. Dans la suite du mémoire, on se concentrera sur la provision pour sinistres.

□ La marge pour risque

Dans le même article de la directive du parlement européen, la marge pour risque est définie comme « *le coût que représente la mobilisation d’un montant de fonds propres éligible égal au Solvency Capital Requirement nécessaire pour faire face aux engagements d’assurance et de réassurance sur toute la durée de vie de ceux-ci* ». Elle représente ainsi le coût de l’immobilisation du capital ⁵.

5. Pour plus de détails sur le calcul de la marge pour risque dans le cadre du risque de provisionnement, on pourra se référer à l’article de C.Y. ROBERT [32]

3.2 L'exigence de capital réglementaire sous Solvabilité II

Dans le premier pilier de Solvabilité II, les compagnies d'assurance doivent répondre à des exigences quantitatives liées aux fonds propres, en sus des provisions.

❑ Le *Minimum Capital Requirement* (MCR)

Le capital minimum requis est le montant minimum de fonds propres que doit détenir la compagnie d'assurance. Si ces fonds propres tombent en-dessous de ce seuil, l'autorité de contrôle se réserve le droit de retirer l'agrément de l'assureur.

❑ Le *Solvency Capital Requirement* (SCR)

Le capital de solvabilité requis est défini comme le montant de fonds propres nécessaire à une compagnie pour limiter la probabilité de ruine à 0,5 % sur un horizon d'un an. Ce capital permet d'affronter une sinistralité exceptionnelle. Il peut également être défini comme la Value-at-Risk ⁶ à 99,5 % des fonds propres à horizon un an.

On définit la *Net Asset Value* comme la situation nette de la compagnie, soit les fonds propres dont l'entreprise dispose après avoir couvert l'ensemble de ses engagements. A l'instant t , on peut l'écrire de la manière suivante :

$$NAV_t = \text{Valeur de marché de l'actif} - \text{Valeur économique du passif},$$

l'actif correspondant principalement aux placements et le passif aux engagements vis-à-vis des assurés. Ainsi, il est nécessaire que la *NAV* reste positive pour que l'entreprise soit solvable. Le SCR peut finalement être défini comme :

$$SCR = VaR_{99,5\%}(NAV_{t=1} \text{ actualisé} - NAV_{t=0}).$$

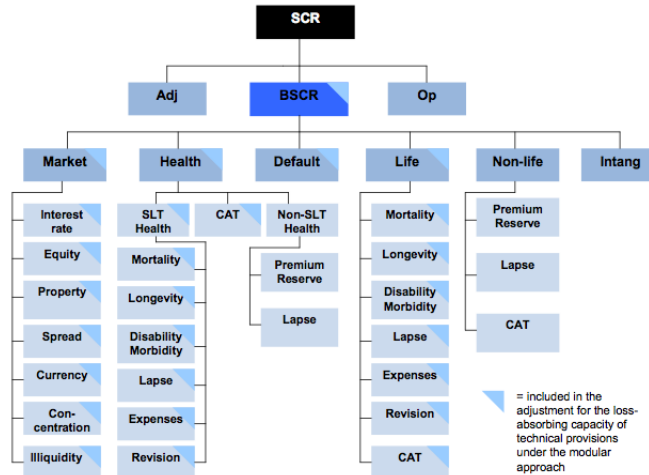


FIGURE 4 – L'aspect modulaire du SCR.

Un tel montant peut être déterminé par la formule standard commune à toutes les compagnies d'assurance et fournie par les spécifications techniques de Solvabilité

6. La Value-at-Risk (VaR) est une mesure de risque très utilisée et représente le quantile de niveau $\alpha \in [0, 1]$ du risque X , c'est-à-dire $VaR_\alpha(X) = \inf\{x \in \mathcal{R} | F(x) \geq \alpha\}$.

II. Comme le témoigne le schéma ci-dessus, le SCR est composé de différents modules et sous-modules de risques. Il s'agit ainsi d'agréger ces nombreux modules et de tenir compte des corrélations existantes. Toutefois, il n'est pas obligatoire d'avoir recours à cette formule et les assureurs peuvent adopter une méthode alternative pour déterminer le SCR. En effet, avec l'accord préalable de l'autorité de contrôle, un modèle interne spécifique à la compagnie et à son profil de risque peut ainsi être adopté.

Dans la suite de ce mémoire, on se concentre sur le module du risque de souscription non-vie, lui-même constitué des sous-modules suivants :

- le risque de primes et de provisions,
- le risque de rachat,
- le risque de catastrophe.

Le premier sous-module est ainsi scindé en deux parties : le risque que les primes ne suffisent pas à couvrir les sinistres futurs liés aux contrats en cours, et le risque de provisionnement qui est l'objet principal de ce mémoire.

3.3 Le risque de provisionnement à horizon un an

On s'intéresse désormais uniquement au $SCR_{\text{provisions}}$ destiné à couvrir une insuffisance des provisions constituées pour les sinistres survenus. On présente dans cette partie la formule standard établie pour l'ensemble des branches d'activités, ou *Lines of Business* (LoB) dans les spécifications techniques du QIS 5. On suppose par la suite que la marge pour risque est nulle.

Pour déterminer le risque de provisionnement, la provision Best Estimate et la volatilité de première année de déviation qui lui est associée sont nécessaires. La volatilité à un an est en effet une nouvelle notion introduite par la Directive Solvabilité II et est une vision à court terme du risque. Le QIS 5 fournit une estimation σ_{LoB} de cette volatilité par branche d'activité, mais les compagnies d'assurance peuvent choisir d'évaluer les volatilités qui leur sont propres par des méthodes dont l'étude fait l'objet de ce mémoire.

Le SCR pour une branche d'activité est ainsi donné par :

$$SCR_{LoB} = \rho(\sigma_{LoB}) \times V_{LoB},$$

avec

- $\rho(\sigma_{LoB})$ une fonction de la volatilité,
- et V_{LoB} une mesure de volume correspondant au Best Estimate.

La fonction de volatilité $\rho(\sigma_{LoB})$ est définie en supposant une distribution des risques selon une loi log-normale⁷ :

$$\rho(\sigma_{LoB}) = \frac{e^{VaR_{99,5\%}(\Phi(0,1))\sqrt{\log(\sigma_{LoB}^2+1)}}}{\sqrt{\sigma_{LoB}^2+1}} - 1,$$

7. On pourra se reporter à l'annexe pour la démonstration.

où $\Phi(0, 1)$ désigne la fonction de répartition de la loi normale centrée réduite.

Pour l'ensemble des branches d'activité, le $SCR_{\text{provisions}}$ est alors donné par :

$$SCR_{\text{provisions}} = \rho(\sigma) \times V,$$

avec

- σ la volatilité agrégée des branches et définie par :

$$\sigma = \sqrt{\frac{1}{V^2} \sum_{i,j} CorrLoB_{i,j} \sigma_i \sigma_j V_i V_j},$$

où $CorrLoB_{i,j}$ correspond aux coefficients de la matrice de corrélation des branches d'activité fournie par l'EIOPA dans le QIS 5,

- et V une mesure de volume correspondant au Best Estimate agrégé, c'est-à-dire à la somme des Best Estimate de toutes les branches d'activité.

Deuxième partie

Les conséquences des choix de l'actuaire sur la mesure de l'incertitude à l'ultime des provisions

L'actuaire dispose de nombreux modèles pour constituer une provision de sinistres. Les méthodes déterministes permettent d'avoir une estimation de cette provision. L'objectif est alors de savoir à quel point la provision connue a posteriori diffère de l'estimation qui avait été faite auparavant ; les méthodes stochastiques permettent de répondre à cette problématique en évaluant la variabilité de la charge sinistre prévue par le modèle. La propriété fondamentale de ces méthodes est la modélisation des règlements incrémentaux par des variables aléatoires. Des estimations et des intervalles de confiance peuvent alors être obtenus et la sinistralité d'exercices futurs simulée. Les méthodes par simulations apportent de surcroît une distribution complète de la provision pour sinistres, ce qui permet la détermination de quantiles. Le schéma ci-dessous illustre de manière simpliste ces différents types de méthodes de provisionnement.

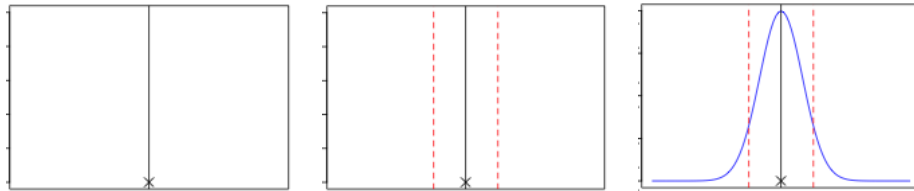


FIGURE 5 – Evolution du calcul de la provision pour sinistres et du risque associé.

Le premier graphique représente l'estimation de la provision par une méthode déterministe. Les méthodes stochastiques déterminent un intervalle de confiance de la charge sinistre représenté en pointillés sur le deuxième graphique. Une distribution de la provision obtenue par les méthodes simulatoires est présentée sur le dernier graphique.

Dans de nombreuses compagnies d'assurance, la méthode déterministe de référence est la méthode Chain Ladder. La popularité de cette méthode a conduit Mack à développer le premier modèle stochastique de provisionnement ; la méthode simulatoire bootstrap associée à Chain Ladder est également très répandue et robuste. Toutefois, cette approche Chain Ladder n'est pas toujours la plus adaptée pour certaines données. C'est pourquoi l'insertion d'un facteur de queue de développement par l'actuaire est étudiée dans le cas de branches à développement long ; il s'agit dès lors de déterminer les conséquences de ce choix à la fois sur le montant et sur la volatilité de l'estimation de la provision.

L'actuaire peut également choisir de privilégier des méthodes qui tiennent compte de données exogènes au triangle de liquidation des sinistres, comme les méthodes déterministes Bornhuetter-Ferguson et Benktander. On peut alors se demander dans quelle mesure l'estimation de la provision et l'incertitude à l'ultime associée à cette estimation sont impactées par la prise en compte d'informations supplémentaires.

Dans cette partie, les méthodes de référence de l'approche Chain Ladder, à savoir les méthodes de Mack et du bootstrap, sont tout d'abord présentées. L'impact du jugement d'expert est également étudié à travers l'insertion d'un facteur de queue de développement dans chacune de ces méthodes. L'influence des données exogènes au triangle de liquidation sur le montant et la volatilité des provisions est approfondie dans un second temps. Les méthodes déterministes Bornhuetter-Ferguson et Benktander sont introduites pour l'estimation des provisions. Les méthodes stochastiques de Mack (2008) et Alai, Merz et Wüthrich (2009) sont ensuite développées pour déterminer l'incertitude à l'ultime des provisions.

Notations :

On rappelle les principales notations introduites précédemment et nécessaires à la compréhension de cette partie :

- les années de survenance des sinistres sont désignées par la lettre i , $i \in \{1, \dots, I\}$,
- les années de développement des sinistres sont désignées par la lettre j , $j \in \{1, \dots, J\}$,
- $I = J$ car on se place dans le cadre d'un triangle de liquidation,
- $X_{i,j}$ désigne le paiement décumulé ou incrémental réglé l'année de développement j pour un sinistre survenu l'année i ,
- $C_{i,j}$ désigne les règlements cumulés en j années de développement pour un sinistre survenu l'année i .

4 L'approche Chain Ladder et le jugement d'expert

On cherche dans un premier temps à déterminer le montant des provisions à constituer par la méthode Chain Ladder ; puis on étudie l'impact de l'insertion d'un facteur de queue de développement sur ce montant. Une méthode permettant d'obtenir les incertitudes à l'ultime correspondantes est ensuite développée.

4.1 L'estimation des provisions

On présente tout d'abord la méthode de provisionnement Chain Ladder et la détermination d'un éventuel facteur de queue de développement.

4.1.1 La méthode Chain-Ladder

La méthode Chain-Ladder est la méthode déterministe de référence pour l'estimation des provisions. Elle est très facile à appliquer et à comprendre, ce qui la rend très populaire. L'idée de cette méthode est de supposer que la liquidation future est similaire à la liquidation passée. Les hypothèses du modèle sont les suivantes :

- les règlements incrémentaux sont indépendants par année de survenance,
- les années de développement sont les variables explicatives du comportement des sinistres futurs.

Le déroulement des règlements est régi par des facteurs de développement strictement positifs f_j , $j \in \{1, \dots, J-1\}$, qui ne dépendent que de l'année de développement j . Ces facteurs sont tels que:

$$\mathbb{E}[C_{i,j}|C_{i,1}, \dots, C_{i,j-1}] = f_{j-1} \times C_{i,j-1} \text{ pour } i \text{ et } j \in \{2, \dots, J\}.$$

Les facteurs de développement f_j sont estimés par :

$$\hat{f}_j = \frac{\sum_{i=1}^{I-j} C_{i,j+1}}{\sum_{i=1}^{I-j} C_{i,j}} \text{ pour } j \in \{1, \dots, J-1\}.$$

Une fois que ces facteurs ont été estimés, les règlements cumulés du triangle inférieur sont déterminés par les estimations $\hat{C}_{i,j}$ suivantes pour $i \in \{2, \dots, I\}$ et $j \in \{2, \dots, J\}$:

$$\hat{C}_{i,j} = (\hat{f}_{J+1-i} \times \dots \times \hat{f}_{j-1}) \times C_{i,J+1-i}.$$

• L'estimation de la provision

L'estimation de la provision $\hat{R}_i$ par année de survénance i pour $i \in \{1, \dots, I\}$ est définie par la relation suivante :

$$\hat{R}_i = \hat{C}_{i,J} - \hat{C}_{i,J+1-i}.$$

La provision $\hat{R}$ à constituer pour l'ensemble des années de survénance est égale à la somme des provisions pour chaque année de survénance :

$$\hat{R} = \sum_{i=1}^I \hat{R}_i.$$

4.1.2 La détermination d'un éventuel facteur de queue de développement

Pour les branches à développement long, il est souvent nécessaire d'approximer la queue de développement lorsque les règlements des années de survénance les plus anciennes sont encore en développement ; un facteur de queue de développement est alors estimé pour prolonger le développement des sinistres et ainsi éviter que les charges ultimes soient sous-estimées.

Le facteur de queue de développement peut être estimé par une régression log-linéaire⁸ sur une fonction des facteurs de développement f_j pour $j \in \{1, \dots, J-1\}$:

$$\log(f_j - 1) = a + b \times j.$$

Les facteurs de développement $\hat{f}_j$ pour $j \geq J$ sont obtenus par extrapolation de la fonction précédente :

$$\hat{f}_j = 1 + \exp(\hat{a} + \hat{b} \times j).$$

Le facteur de queue de développement $f_{ult} > 1$ est estimé par :

$$\hat{f}_{ult} = \prod_{j \geq J} \hat{f}_j.$$

8. Il existe bien d'autres méthodes pour estimer un tel facteur ; pour plus de détails à ce sujet, on peut se référer par exemple à l'article [6] de J. BOOR qui développe un large panel de méthodes.

Il est nécessaire de vérifier la concordance de ce facteur avec l'expérience acquise par l'actuaire sur le développement futur des sinistres.

La charge ultime pour l'année de survenance i est alors estimée par :

$$\hat{C}_{i,ult} = \hat{C}_{i,J} \times \hat{f}_{ult}.$$

Le modèle Chain Ladder est simple à appliquer et à comprendre, ce qui explique son succès dans la plupart des compagnies d'assurance. Il possède cependant plusieurs limites. Les hypothèses sont très contraignantes : les facteurs de développement sont supposés constants pour toutes les années d'origine, c'est-à-dire que les règlements cumulés $C_{i,j}$ et $C_{i,j+1}$ doivent être proportionnels. De plus, la provision dépend fortement du dernier montant de charge connu ; si le triangle est peu stable et irrégulier, ce modèle n'est pas adapté.

4.2 La mesure de l'incertitude à l'ultime

Il s'agit désormais de quantifier l'incertitude sur les provisions ainsi estimées, c'est-à-dire l'erreur de prédiction du modèle de provisionnement utilisé. En effet, estimer la provision pour sinistres à payer est un processus de prédiction : on se base sur des observations passées pour prédire les règlements futurs. La méthode analytique de Mack puis la méthode par simulations bootstrap sont développées dans cette partie. On mesure la volatilité créée par l'aléa tout au long du développement des sinistres en l'absence de jugement d'expert. L'erreur de prédiction du modèle est ensuite estimée lorsqu'un facteur de queue de développement est pris en considération dans chacune de ces deux méthodes.

4.2.1 La mesure de la variance des provisions par la méthode de Mack

Mack est l'un des premiers à développer une formule analytique simple de l'incertitude sur les estimations des provisions par Chain Ladder dans l'article *Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates* (1993). Cette méthode fournit une estimation des deux premiers moments de la distribution du montant de la provision pour sinistres à payer. Une hypothèse sur la loi de distribution de la provision totale $\hat{R}$ permet d'estimer les quantiles de cette distribution ; les lois normale ou log-normale permettent d'obtenir des intervalles de confiance pour la provision. Cette méthode conserve exactement l'approche Chain Ladder pour la provision moyenne et permet d'estimer numériquement l'écart type de cette provision à partir des facteurs de développement.

- **Le modèle de Mack repose sur trois hypothèses importantes :**

Hypothèse 1 :

Les années de survenance sont indépendantes : pour $i \neq j$, $\{C_{i,1}, \dots, C_{i,J}\}$ est indépendant de $\{C_{j,1}, \dots, C_{j,J}\}$.

Cette hypothèse peut être validée graphiquement en utilisant les résidus, c'est-à-dire les différences entre les règlements connus a posteriori et leur estimation. Ces résidus ne doivent pas être corrélés, ni afficher une tendance nette.

Hypothèse 2 :

Seul le montant des paiements $C_{i,j}$ réglé en j années de développement pour les sinistres survenus l'année i , $i \in \{1, \dots, I\}$, sert à déterminer le montant $C_{i,j+1}$ réglé en $j+1$ années de développement pour les sinistres survenus la même année. De plus, le facteur f_j est indépendant de l'année de survenance i . La propriété de linéarité de Chain Ladder est :

$$\mathbb{E}[C_{i,j}|C_{i,1}, \dots, C_{i,j-1}] = f_{j-1} \times C_{i,j-1} \text{ pour } i \text{ et } j \in \{2, \dots, J\}.$$

Cette hypothèse peut être validée graphiquement : les $(C_{i,j}, C_{i,j+1})$ doivent être sensiblement alignés sur une droite de pente f_j passant par l'origine. Les facteurs de développement f_j sont estimés de la même manière que pour la méthode Chain Ladder.

Ces deux premières hypothèses sont similaires à celles de la méthode Chain Ladder ; Mack ajoute une troisième hypothèse qui concerne la variance des règlements cumulés.

Hypothèse 3 :

Une dernière hypothèse est nécessaire : il existe des paramètres σ_j^2 strictements positifs tels que :

$$\text{Var}(C_{i,j}|C_{i,1}, \dots, C_{i,j-1}) = \sigma_{j-1}^2 \times C_{i,j-1},$$

pour $i \in \{2, \dots, I\}, j \in \{2, \dots, J\}$. La variance du montant $C_{i,j}$ ne dépend que du montant de l'année de développement précédente. Un estimateur sans biais du paramètre σ_j^2 est donné par :

$$\hat{\sigma}_j^2 = \frac{1}{I-j} \sum_{i=1}^{I-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2, \text{ pour } j \in \{1, \dots, J-2\}.$$

Si le dernier facteur de développement est égal à un et que les règlements sont supposés constants, alors $\hat{\sigma}_{I-1}^2 = 0$. Sinon, on peut écrire l'égalité suivante :

$$\hat{\sigma}_{I-1}^2 = \min\left(\frac{\hat{\sigma}_{I-2}^4}{\hat{\sigma}_{I-3}^2}, \min(\hat{\sigma}_{I-3}^2, \hat{\sigma}_{I-2}^2)\right).$$

• L'erreur de prédiction des provisions

Soit $\mathcal{D}_I$ l'information disponible au moment de l'estimation des provisions : $\mathcal{D}_I = \{C_{i,j} | i+j \leq I+1\}$. L'erreur de prédiction $MSEP(\hat{C}_{i,J})$, ou *Mean Square Error of Prediction*, de l'estimation de la charge ultime $\hat{C}_{i,J}$ est définie par l'espérance conditionnelle suivante :

$$\begin{aligned} MSEP(\hat{C}_{i,J}) &= \mathbb{E}[(\hat{C}_{i,J} - C_{i,J})^2 | \mathcal{D}_I], \\ &= \underbrace{\text{Var}(C_{i,J} | \mathcal{D}_I)}_{\text{Erreur de processus}} + \underbrace{(\mathbb{E}[C_{i,J} | \mathcal{D}_I] - \hat{C}_{i,J})^2}_{\text{Erreur d'estimation}}. \end{aligned}$$

L'erreur de prédiction est composée de deux termes. L'erreur de processus intrinsèque au modèle provient du mouvement stochastique du processus $C_{i,J}$ et l'erreur d'estimation porte sur les facteurs de développement f_j .

Soit $\hat{R}_i$ l'estimation de la provision pour l'année de survenance i , $i \in \{1, \dots, I\}$. On remarque que :

$$MSEP(\hat{R}_i) = \mathbb{E}[(\hat{R}_i - R_i)^2 | \mathcal{D}_I] = \mathbb{E}[(\hat{C}_{i,J} - C_{i,J})^2 | \mathcal{D}_I] = MSEP(\hat{C}_{i,J}).$$

Sous les hypothèses du modèle, l'erreur de prédiction de la provision pour l'année de survénance i , $i \in \{1, \dots, I\}$, est :

$$\widehat{MSEP}(\hat{R}_i) = \hat{C}_{i,J}^2 \sum_{j=I+1-i}^{I-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2} \left(\frac{1}{\hat{C}_{i,j}} + \frac{1}{\sum_{l=1}^{I-j} C_{l,j}} \right).$$

L'erreur de prédiction correspond à la volatilité créée par l'aléa tout au long du développement des sinistres.

L'erreur de prédiction de la provision constituée pour l'ensemble des années de survénance est donnée par :

$$\widehat{MSEP}(\hat{R}) = \sum_{i=2}^I \left[\widehat{MSEP}(\hat{R}_i) + C_{i,J} \left(\sum_{l=i+1}^I C_{l,J} \right) \sum_{j=I+1-i}^{I-1} \frac{2\hat{\sigma}_j^2/\hat{f}_j^2}{\sum_{m=1}^{I-j} C_{m,j}} \right].$$

Pour obtenir l'incertitude à l'ultime de la provision constituée pour l'ensemble des années de survénance, la formule de Mack est légèrement modifiée pour tenir compte des covariances entre ces provisions.

• L'insertion d'un facteur de queue de développement

En 1999, Mack propose dans son article *The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor* de prendre en considération un facteur de queue de développement dans la détermination de l'erreur de prédiction des provisions. Comme on se place dans un cadre stochastique, il s'agit désormais d'estimer la volatilité associée à ce facteur de queue de développement.

Mack étend les résultats précédents à des facteurs de développement Chain Ladder pondérés et au cas de triangles à déroulement incomplet. Les facteurs de développement sont alors :

$$\hat{f}_j = \frac{\sum_{i=1}^{I-j} w_{i,j} C_{i,j}^\alpha F_{i,j}}{\sum_{i=1}^{I-j} w_{i,j} C_{i,j}^\alpha} \text{ pour } j \in \{1, \dots, J-1\},$$

avec

$$\left\{ \begin{array}{ll} F_{i,j} & \text{les facteurs de développement individuels } F_{i,j} = C_{i,j+1}/C_{i,j}, \\ & \text{pour } i \in \{1, \dots, I\} \text{ et } j \in \{1, \dots, J-1\}, \\ \alpha \in \{0, 1, 2\}, & \begin{array}{l} \text{si } \alpha = 0, \text{ on obtient la moyenne des facteurs de développement individuels,} \\ \text{si } \alpha = 1, \text{ on obtient les facteurs de développement Chain Ladder,} \\ \text{si } \alpha = 2, \text{ on a effectué une régression de } C_{i,j+1} \text{ contre } C_{i,j} \text{ avec un intercept nul,} \end{array} \\ w_{i,j} \in [0, 1] & \text{des pondérations arbitraires.} \end{array} \right.$$

Avec ces notations, les hypothèses du modèle sont :

Hypothèse 1 :

Les années de survénance sont indépendantes : pour $i \neq j$, $\{C_{i,1}, \dots, C_{i,J}\}$ est indépendant de $\{C_{j,1}, \dots, C_{j,J}\}$.

Hypothèse 2 :

$$\mathbb{E}[F_{i,j} | C_{i,1}, \dots, C_{i,j}] = f_j, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \in \{1, \dots, J-1\}.$$

Hypothèse 3 :

$$\mathbb{V}ar(F_{i,j}|C_{i,1}, \dots, C_{i,j}) = \frac{\sigma_j^2}{w_{i,j}C_{i,j}^\alpha}, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \in \{1, \dots, J-1\}.$$

Un estimateur non biaisé du paramètre σ_j^2 est :

$$\hat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=1}^{I-j} w_{i,j} C_{i,j}^\alpha (F_{i,j} - \hat{f}_j)^2 \text{ pour } j \in \{1, \dots, J-2\}.$$

L'erreur de prédiction de la provision pour l'année de survénance i , $i \in \{1, \dots, I\}$, est :

$$\widehat{MSEP}(\hat{R}_i) = \hat{C}_{i,J}^2 \sum_{j=J+1-i}^{J-1} \frac{\widehat{MSE}(F_{i,j}) + \widehat{MSE}(\hat{f}_j)}{\hat{f}_j^2}$$

avec

$$\begin{cases} \widehat{MSE}(F_{i,j}) & \text{un estimateur de } \mathbb{V}ar(F_{i,j}|C_{i,1}, \dots, C_{i,j}), \\ \widehat{MSE}(\hat{f}_j) & \text{un estimateur de } \mathbb{V}ar(\hat{f}_j) = \frac{\sigma_j^2}{\sum_{i=1}^{I-j} w_{i,j} C_{i,j}^\alpha}. \end{cases}$$

La formule de récursivité suivante est obtenue :

$$\widehat{MSEP}(\hat{C}_{i,j+1}) = \hat{C}_{i,j}^2 \left(\widehat{MSE}(F_{i,j}) + \widehat{MSE}(\hat{f}_j) \right) + \widehat{MSEP}(\hat{C}_{i,j}) \hat{f}_j^2,$$

avec $\widehat{MSEP}(\hat{C}_{i,J+1-i}) = 0$.

Pour les triangles à développement long, l'actuaire insère le facteur de développement f_{ult} pris en compte dans la formule de récursivité suivante :

$$\widehat{MSEP}(\hat{C}_{i,ult}) = \hat{C}_{i,J}^2 \left(\widehat{MSE}(F_{i,ult}) + \widehat{MSE}(\hat{f}_{ult}) \right) + \widehat{MSEP}(\hat{C}_{i,J}) \hat{f}_{ult}^2.$$

L'actuaire est chargé d'estimer le facteur de queue de développement $\hat{f}_{ult}$ et l'erreur d'estimation associée $\widehat{MSE}(\hat{f}_{ult})$. Il s'agit de déterminer à quel point l'estimation $\hat{f}_{ult}$ sera éloignée de la réalisation f_{ult} . Mack propose de déterminer cette erreur d'estimation à partir des erreurs $\widehat{MSE}(\hat{f}_j)$ historiques : un critère de choix pourrait être de trouver un $j < J$ tel que :

$$\hat{f}_{j-1} > \hat{f}_{ult} > \hat{f}_j,$$

et de fixer l'erreur $\widehat{MSE}(\hat{f}_{ult})$ telle que :

$$\widehat{MSE}(\hat{f}_{j-1}) > \widehat{MSE}(\hat{f}_{ult}) > \widehat{MSE}(\hat{f}_j).$$

L'erreur d'estimation $\widehat{MSE}(F_{i,ult})$ associée au facteur individuel de queue de développement doit également être connue. Mack propose là encore de trouver un $j < J$ tel que :

$$\hat{f}_{j-1} > \hat{f}_{ult} > \hat{f}_j,$$

et de fixer l'erreur $\widehat{MSE}(F_{i,ult})$ pour un i fixé, telle que :

$$\widehat{MSE}(F_{i,j-1}) > \widehat{MSE}(F_{i,ult}) > \widehat{MSE}(F_{i,j}).$$

La méthode de Mack conserve exactement l'approche Chain Ladder pour le calcul de la provision moyenne et elle permet d'obtenir un intervalle de confiance pour cette provision. Mack étend également sa méthode de telle manière qu'on puisse prendre en considération un facteur de queue de développement. Cependant, les hypothèses de ce modèle sont souvent considérées comme trop restrictive. La principale limite du méthode de Mack est de déterminer uniquement les moments d'ordres un et deux de la distribution des provisions. C'est pourquoi on s'intéresse désormais à une méthode palliant cette limite.

4.2.2 La mesure de l'incertitude à partir d'une distribution complète des provisions par la méthode bootstrap

Le bootstrap est une méthode simulatoire introduite en 1979 par Efron et dont le principal atout est de fournir une distribution complète d'un estimateur. Beaucoup d'informations, comme les quantiles de la distribution, peuvent être obtenues par cette méthode à partir d'un unique échantillon.

Dans le cadre de l'assurance non-vie, le bootstrap est de plus en plus utilisé et permet d'obtenir une distribution des provisions lorsqu'on fait intervenir les modèles linéaires généralisés (MLG). Le bootstrap repose sur la technique de rééchantillonnage avec remise des résidus issus de l'application du modèle statistique sous-jacent. Il suppose que les incréments $X_{i,j}$ du triangle de liquidation de départ sont indépendants et identiquement distribués, ce qui n'est pas toujours vérifié lorsqu'on se place dans un cadre MLG. Ainsi, il est préférable d'utiliser les résidus de Pearson plutôt que d'appliquer le bootstrap directement sur les incréments. En effet, ces résidus sont calculés entre les valeurs observées et les valeurs prédites par le modèle et sont supposés être à la fois indépendants et identiquement distribués.

Renshaw et Verrall ont montré dans l'article *A Stochastic Model Underlying the Chain-Ladder Technique* (1998) que le modèle de Poisson surdispersé avec fonction de lien logarithmique détaillé ci-dessous reproduit exactement les estimations Chain Ladder des provisions pour un triangle à incréments positifs. On se place dans le cadre de ce modèle GLM où la distribution du terme d'erreur est de type Poisson surdispersé. La méthode bootstrap tient alors compte de cette propriété et est simplifiée.

- **Le modèle de Poisson surdispersé**

La variable aléatoire Y appartient à la famille exponentielle si sa densité peut s'écrire sous la forme

$$f_Y(y, \theta, \phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right\}.$$

Le modèle de Poisson surdispersé appartient à cette famille avec:

$$\begin{cases} a(\phi) = \phi \\ b(\theta) = e^\theta \\ c(y, \phi) = -\ln y \end{cases}$$

avec ϕ le paramètre de dispersion, θ le paramètre de moyenne et $b''(\theta)$ la fonction de variance.

Ce modèle a été introduit dans un contexte bayésien par Renshaw et Verrall ; il diffère du modèle de Poisson dans la mesure où la variance est supérieure à l'espé-

rance. Un avantage certain de ce modèle est le fait d'obtenir les mêmes cadences de règlements que le modèle Chain-Ladder qui sert souvent de référence.

Les hypothèses de ce modèle sont les suivantes :

- les règlements incrémentaux $X_{i,j}$ sont indépendants et distribués selon le modèle de Poisson surdispersé (ODP) et il existe des paramètres positifs $\gamma_1, \dots, \gamma_I, m_1, \dots, m_I$ et $\phi > 0$ tels que:

$$\mathbb{E}[X_{i,j}] = \mu_{i,j} = m_i \gamma_j$$

$$\text{Var}(X_{i,j}) = \phi \mu_{i,j}$$

$$\text{et } \sum_{j=1}^J \gamma_j = 1,$$

avec ϕ le paramètre de dispersion et γ_k les cadences de règlements non cumulées.

Si $X_{i,j}$ suit une loi de Poisson surdispersé, on peut écrire :

$$\frac{X_{i,j}}{\phi} \sim \mathcal{Poisson}\left(\frac{\mu_{i,j}}{\phi}\right).$$

Les caractéristiques de ce modèle MLG sont :

- i. la loi de probabilité de la variable aléatoire $X_{i,j}$ est de type exponentiel,
- ii. la fonction lien logarithmique g strictement monotone et dérivable établit un lien entre les composantes aléatoires $X_{i,j}$ et le prédicteur linéaire $\eta_{i,j}$:

$$g(\mathbb{E}[X_{i,j}]) = \eta_{i,j} = \log(\mu_{i,j}) = \mu + \alpha_i + \beta_j \text{ pour tout } 0 \leq i \leq I \text{ et } 0 \leq j \leq J,$$

$$\text{tel que : } \mu_{i,j} = e^{\mu + \alpha_i + \beta_j}, \text{ avec } \alpha_1 = \beta_1 = 0.$$

• La procédure bootstrap

La procédure bootstrap se décompose en plusieurs étapes détaillées ci-dessous.

1. Utiliser le modèle sur le triangle d'incrémentaux supérieur pour obtenir les valeurs prévues $\hat{\mu}_{i,j}$. Comme le modèle ODP reproduit celui de Chain Ladder, cela revient à appliquer cette dernière méthode. Déterminer alors les facteurs de développement $\hat{f}_j, j \in \{1, \dots, J-1\}$ à partir du triangle de charges cumulées, puis estimer la provision $\hat{R}$ par la méthode Chain Ladder décrite précédemment.
2. Estimer le triangle ajusté : calculer les valeurs prédites par le modèle $\hat{C}_{i,j}$ pour la partie supérieure du triangle à partir de la diagonale que l'on laisse inchangée.

$$\hat{C}_{i,j} = \frac{\hat{C}_{i,J}}{\prod_{k=j}^{J-1} \hat{f}_k}, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \leq J - i + 1.$$

3. Dédire du triangle précédent les règlements incrémentaux ajustés $\hat{X}_{i,j}$ pour la partie supérieure du triangle :

$$\hat{X}_{i,j} = \hat{C}_{i,j} - \hat{C}_{i,j-1}, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \leq J - i + 1.$$

4. Calculer les résidus de Pearson qui correspondent aux résidus bruts standardisés définis par :

$$r_{i,j}^P = \frac{X_{i,j} - \hat{X}_{i,j}}{\sqrt{\hat{X}_{i,j}}}.$$

Ces résidus $r_{i,j}^P$ sont ensuite ajustés pour corriger le biais dans la comparaison des estimations analytiques et bootstrap de l'erreur de prédiction. On considère pour cela les degrés de liberté des données. Les résidus de Pearson ajustés sont définis par :

$$r_{i,j}^{P_{ajus}} = \sqrt{\frac{n}{n-p}} r_{i,j}^P,$$

avec

$$\begin{cases} n = \frac{I(I+1)}{2} \text{ le nombre de données,} \\ p = 2I - 1 \text{ le nombre de degrés de liberté.} \end{cases}$$

Les résidus étant en moyenne proches de zéro, l'ajustement par le nombre de paramètres augmente la variance. On estime alors le paramètre d'échelle $\hat{\phi}$:

$$\hat{\phi} = \sum_{i+j \leq I+1} (r_{i,j}^{P_{ajus}})^2.$$

5. Répéter N fois les étapes suivantes :

- Effectuer un rééchantillonnage aléatoire avec remise des résidus de Pearson ajustés. On obtient un nouveau triangle de résidus $\tilde{r}_{i,j}^{(k)}$ pour la k^{ieme} boucle bootstrap.
- Reconstituer le pseudo-triangle des règlements incrémentaux par la relation suivante :

$$\tilde{X}_{i,j}^{(k)} = \hat{X}_{i,j} + \tilde{r}_{i,j}^{(k)} \sqrt{\hat{X}_{i,j}}, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \leq J - i + 1.$$

- Réestimer le modèle pour obtenir les nouvelles valeurs prévues $\tilde{\mu}_{i,j}^{(k)}$. On simplifie cette étape en calculant les facteurs de développement Chain Ladder du nouveau pseudo-triangle des règlements cumulés et en utilisant cette méthodologie pour obtenir les valeurs du triangle inférieur. L'erreur d'estimation est ainsi prise en compte dans cette étape.
- Effectuer un tirage aléatoire sur le triangle inférieur : simuler un règlement selon une distribution de moyenne $\tilde{\mu}_{i,j}^{(k)}$ et de variance $\hat{\phi} \tilde{\mu}_{i,j}^{(k)}$. Cette étape prend alors en compte l'erreur de processus.

Pour simuler les paiements, on peut choisir une loi de Poisson surdispersée : on simule une réalisation d'une loi de Poisson d'espérance $\tilde{\mu}_{i,j}^{(k)}/\hat{\phi}$ que l'on multiplie par $\hat{\phi}$. Les modélisations sont alors des multiples de $\hat{\phi}$. Pour faire face à cet inconvénient, il peut être intéressant d'utiliser une loi Gamma telle que son espérance soit $\tilde{\mu}_{i,j}^{(k)}$ et sa variance $\hat{\phi} \tilde{\mu}_{i,j}^{(k)}$.

- Estimer la provision $\hat{R}^{(k)}$ par Chain Ladder et stocker le résultat.

6. Obtenir une distribution empirique de l'estimation de la provision ; déterminer l'erreur de prédiction associée à cette estimation.

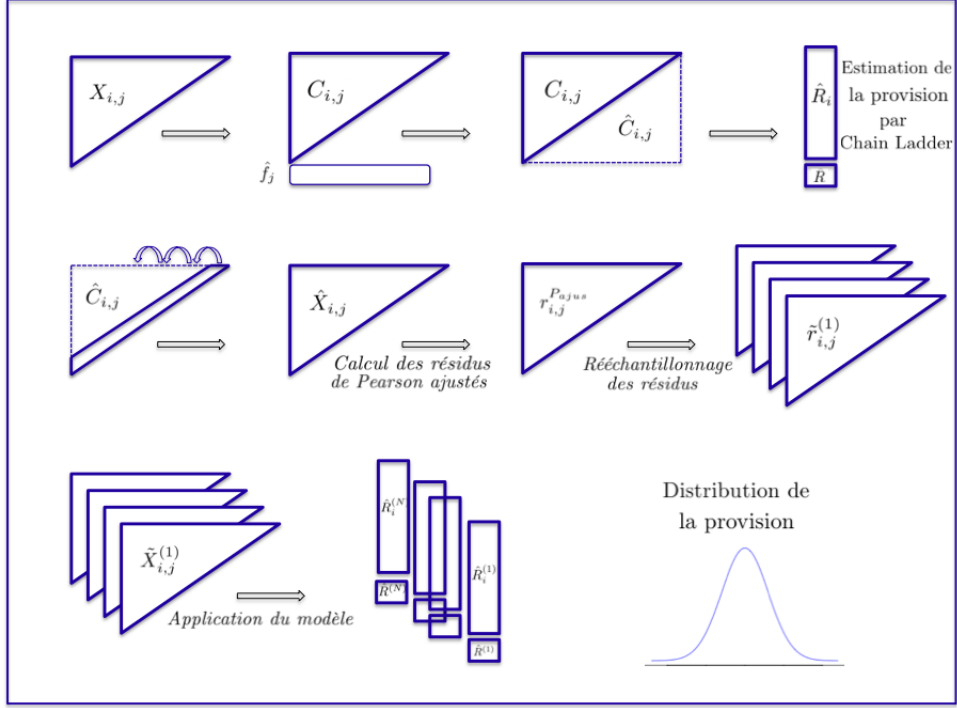


FIGURE 6 – Illustration de la procédure bootstrap.

• L'erreur de prédiction

Deux approches différentes permettent d'estimer l'erreur de prédiction des provisions par cette méthode.

L'approche simulatoire consiste à modifier la modélisation MLG de la k^{ieme} boucle bootstrap comme cela a été développé précédemment dans la procédure. L'erreur de prédiction associée à cette estimation de la provision est alors calculée en calculant l'écart-type des provisions stockées.

On suppose les incréments indépendants. Dans l'approche semi-simulatoire, l'erreur de processus des provisions est considérée comme la variance de la somme des incréments du triangle inférieur :

$$\mathbb{V}ar(\hat{R}) = \mathbb{V}ar\left(\sum_{i+j>I} X_{i,j}\right) = \sum_{i+j>I} \mathbb{V}ar(X_{i,j}).$$

L'erreur de processus intrinsèque au modèle est alors estimée par :

$$\widehat{\mathbb{V}ar}(\hat{R}) = \sum_{i+j>I} \hat{\phi}\tilde{\mu}_{i,j}^{(k)}.$$

L'erreur d'estimation des paramètres est définie par:

$$\mathbb{E} = \left[(\hat{R} - \mathbb{E}[R])^2 \right], \text{ et est estimée par } \frac{n}{n-p} \sum_{k=1}^N \left(\hat{R}^{(k)} - \frac{1}{N} \sum_{k=1}^N \hat{R}^{(k)} \right)^2.$$

L'erreur de prédiction qui correspond à l'incertitude à l'ultime des provisions est

ainsi estimée par :

$$\widehat{MSEP}(\hat{R}) = \underbrace{\sum_{i+j>I} \hat{\phi} \tilde{\mu}_{i,j}^{(k)}}_{\text{Erreur de processus}} + \underbrace{\frac{n}{n-p} \sum_{k=1}^N \left(\hat{R}^{(k)} - \frac{1}{N} \sum_{k=1}^N \hat{R}^{(k)} \right)^2}_{\text{Erreur d'estimation}}.$$

• L'insertion d'un facteur de queue de développement

Un facteur de queue de développement peut être pris en compte dans l'estimation des provisions dans la méthode Chain Ladder. L'incertitude à l'ultime de cette estimation a pu également être déterminée par la méthode analytique de Mack (1999). On s'intéresse désormais à l'incertitude à l'ultime des provisions en présence d'un tel facteur de queue de développement par la méthode simulateur bootstrap. On procède alors en deux étapes :

1. Déterminer un facteur de queue de développement.

Il s'agit d'estimer un facteur de queue de développement $\hat{f}_{ult}^{(k)}$, $k \in \{1, \dots, N\}$, à chacun des N passages dans la boucle bootstrap. Dans le cadre MLG Poisson surdispersé, la logique est la même que pour Chain Ladder. Pour chacun des pseudo-triangles obtenus par rééchantillonnage, le facteur de queue de développement $\hat{f}_{ult}^{(k)}$ est déterminé par extrapolation.

2. Simulation du facteur de queue de développement.

Dans chaque itération bootstrap k , $k \in \{1, \dots, N\}$, le facteur de queue de développement est déterminé par simulation d'une loi d'espérance $\hat{f}_{ult}^{(k)}$ et de variance $\mathbb{V}ar(\hat{f}_{ult}^{(k)})$ obtenue par la méthode Delta.

On suppose que le facteur de queue de développement $\hat{f}_{ult}^{(k)}$ a été déterminé par régression :

$$\hat{f}_{ult}^{(k)} = f(\hat{a}, \hat{b}) = \prod_{j \geq J} \hat{f}_j,$$

où $\hat{a}$ et $\hat{b}$ sont les estimateurs par maximum de vraisemblance.

$$\hat{Z} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \sim \mathcal{N}(\mu, \Sigma),$$

où la matrice de variance-covariance est $\Sigma = \begin{pmatrix} \mathbb{V}ar(\hat{a}) & \text{Cov}(\hat{a}, \hat{b}) \\ \text{Cov}(\hat{a}, \hat{b}) & \mathbb{V}ar(\hat{b}) \end{pmatrix}$.

La méthode Delta⁹ permet alors d'écrire :

$$\hat{f}_{ult}^{(k)} \sim \mathcal{N}(f(\mu), {}^t \nabla f \Sigma \nabla f), \text{ avec } \nabla f = \begin{pmatrix} \partial f / \partial a \\ \partial f / \partial b \end{pmatrix} \text{ le gradient de } f.$$

La variance du facteur de queue de développement $\hat{f}_{ult}^{(k)}$ est alors :

$$\begin{aligned} \mathbb{V}ar(\hat{f}_{ult}^{(k)}) &= \left(\frac{\partial f}{\partial b}(\hat{f}_{ult}^{(k)}) \right)^2 \mathbb{V}ar(\hat{b}) + \left(\frac{\partial f}{\partial a}(\hat{f}_{ult}^{(k)}) \right)^2 \mathbb{V}ar(\hat{a}) \\ &+ 2 \frac{\partial f}{\partial a}(\hat{f}_{ult}^{(k)}) \frac{\partial f}{\partial b}(\hat{f}_{ult}^{(k)}) \text{Cov}(\hat{a}, \hat{b}). \end{aligned}$$

9. On peut se reporter à l'annexe pour un raisonnement plus détaillé.

$\hat{f}_{ult}^{(k)}$ est l'estimateur par maximum de vraisemblance de

$$f_{ult} = \prod_{j \geq J} (1 + \exp(a \times j + b)).$$

Dans un cadre asymptotique, $\hat{f}_{ult}^{(k)}$ est gaussien et le facteur de queue de développement peut alors être modélisé par une loi normale d'espérance $\hat{f}_{ult}^{(k)}$ et de variance $\mathbb{V}ar(\hat{f}_{ult}^{(k)})$. La loi log-normale peut être utilisée si l'approche se veut plus prudente.

La méthode du bootstrap fournit une distribution des règlements futurs. Cela permet d'avoir de nombreuses informations à partir du seul triangle de liquidation. Si ce triangle de départ est modifié, le bootstrap reste pertinent, ce qui fait la robustesse de ce modèle. Un intervalle de confiance de l'estimation de la provision peut être obtenu. Il est également possible d'insérer un facteur de queue de développement dans la procédure bootstrap, ce qui permet à l'actuaire de garder la robustesse du modèle bootstrap pour son estimation des provisions tout en apportant son expérience.

En revanche, cette méthode ne permet pas de fournir une information fiable sur un quantile très élevé de la distribution des provisions : l'article *bootstrap et méthodes de triangles* de Planchet [29] rappelle que le rééchantillonnage du bootstrap s'effectue sur peu de données issues du coeur de la distribution. Les données du triangle de départ ne reflètent pas les observations concernant les quantiles, c'est-à-dire les extrêmes de la distribution ; cela pose problème si on est dans une approche modèle interne. La méthode bootstrap est également restrictive dans le choix de la famille de loi qui est souvent à support positif ; les incréments doivent donc être positifs. Si les données ne sont pas assez homogènes, cette méthode n'est pas très adaptée puisque l'erreur de prédiction sera surestimée. Cette erreur est souvent supérieure à celle estimée par la méthode de Mack, ce qui est un inconvénient supplémentaire.

La mesure de l'incertitude à l'ultime des provisions dépend du choix de l'actuaire : dans l'approche Chain Ladder, celui-ci doit en effet effectuer un arbitrage entre la robustesse du modèle bootstrap et une volatilité moindre dans le cas du modèle de Mack. Il a également la possibilité d'apporter son jugement d'expert : déterminer et insérer un facteur de queue de développement dans la détermination de la provision reflète mieux la réalité, mais la volatilité de cette provision estimée augmente. L'actuaire peut dès lors choisir d'utiliser d'autres méthodes de provisionnement et faire valoir son expérience d'une manière différente ou tenir compte de données supplémentaires.

5 La prise en compte de données exogènes par l'approche Bornhuetter-Ferguson

On dispose parfois d'informations complémentaires au triangle de liquidation que l'on souhaiterait utiliser pour déterminer la provision la plus adéquate possible. Certaines méthodes alternatives à Chain Ladder permettent d'insérer un indicateur d'exposition au risque qui peut être :

- un montant de primes,
- un nombre de sinistres,
- un nombre de contrats.

Les méthodes Bornhuetter-Ferguson et Benktander utilisent de telles données exogènes au triangle de liquidation et permettent de réduire l'instabilité de la méthode de référence Chain Ladder.

5.1 L'estimation des provisions

5.1.1 La méthode Bornhuetter-Ferguson

Bornhuetter et Ferguson proposent en 1972 une méthode alternative à Chain Ladder dans *The actuary and IBNR* pour déterminer une estimation des provisions. Une propriété fondamentale de cette méthode est l'indépendance entre la provision à constituer et le dernier montant connu des règlements $C_{i,I+1-i}$, $i \in \{1, \dots, I\}$. Les auteurs introduisent une estimation a priori des montants de charge ultime $\hat{U}_i$, $i \in \{1, \dots, I\}$, obtenue généralement par le calcul suivant :

$$\hat{U}_i = \Phi_i \times P_i,$$

avec Φ_i le ratio sinistres à primes (ou loss ratio) attendu pour l'année de survénance i , et P_i le montant de primes acquises relatif à cette même année.

L'estimateur $\hat{C}_{i,J}^{BF}$ de la charge ultime par cette méthode pour l'année de survénance i , $i \in \{1, \dots, I\}$, est :

$$\hat{C}_{i,J}^{BF} = C_{i,I-i} + (1 - \hat{z}_{I+1-i})\hat{U}_i,$$

avec $\hat{z}_j$ la cadence de règlement cumulée après j années de développement. Cette cadence est le plus souvent issue de la méthode Chain Ladder.

L'estimation de la charge ultime correspond à une pondération entre une estimation Chain Ladder et la charge de sinistres ultime supposée connue.

On déduit l'estimateur $\hat{R}_i^{BF}$ de la provision à constituer pour cette même année de survénance :

$$\hat{R}_i^{BF} = (1 - \hat{z}_{I+1-i})\hat{U}_i,$$

et l'estimateur $\hat{R}$ de la provision globale est :

$$\hat{R}^{BF} = \sum_{i=1}^I \hat{R}_i^{BF}.$$

L'actuaire apporte son expérience dans l'estimation de la charge ultime a priori, ce qui rend l'application de cette méthode spécifique à chaque actuaire. La provision estimée paraît plus adéquate que la méthode Chain Ladder car elle est propre

à la branche étudiée et indépendante de la dernière charge connue. Ainsi, elle est souvent supérieure à la charge sinistre estimée par Chain Ladder mais permet d'en réduire l'instabilité.

5.1.2 La méthode Benktander

Benktander propose de modifier la méthode Bornhuetter-Ferguson dans *An approach to credibility in calculating IBNR for casualty excess reinsurance* (1976) afin de mieux tenir compte des données du triangle de liquidation. Cette méthode s'inscrit dans le prolongement de la méthode Bornhuetter-Ferguson et il s'agit là encore d'utiliser des estimateurs a priori.

L'estimateur de la charge ultime $\hat{C}_{i,J}^{BE}$ par cette méthode pour l'année de survenance $i, i \in \{1, \dots, I\}$, fait intervenir l'estimateur $\hat{C}_{i,J}^{BF}$ de la méthode Bornhuetter-Ferguson :

$$\hat{C}_{i,J}^{BE} = C_{i,I+1-i} + (1 - \hat{z}_{I+1-i})\hat{C}_{i,J}^{BF}.$$

La provision $\hat{R}_i^{BE}$ est alors :

$$\begin{aligned}\hat{R}_i^{BE} &= (1 - \hat{z}_{I+1-i})\hat{C}_{i,J}^{BF} \\ &= (1 - \hat{z}_{I+1-i}) \times (C_{i,I+1-i} + \hat{R}_i^{BF}) \\ &= \frac{\hat{z}_{I+1-i}(1 - \hat{z}_{I+1-i})C_{i,I+1-i}}{\hat{z}_{I+1-i}} + (1 - \hat{z}_{I+1-i})\hat{R}_i^{BF} \\ &= \hat{z}_{I+1-i}\hat{R}_i^{CL} + (1 - \hat{z}_{I+1-i})\hat{R}_i^{BF}\end{aligned}$$

avec $\hat{R}_i^{CL}$ la provision Chain Ladder. La provision à constituer par la méthode Benktander est donc un barycentre des provisions Chain Ladder et Bornhuetter-Ferguson. La méthode Benktander est simple à appliquer et donne une estimation plus précise que ces deux méthodes très utilisées en pratique. Mack [23] insiste sur la précision de cette méthode en 2000 en s'appuyant sur la théorie de la crédibilité. Dans la suite de ce mémoire, on s'intéresse surtout à la méthode Bornhuetter-Ferguson dans la mesure où la problématique reste la prise en compte de données exogènes et, de ce point de vue, cette méthode marque un véritable changement.

5.2 La mesure de l'incertitude à l'ultime des provisions

5.2.1 La méthode de Mack (2008)

Dans *The Prediction Error of Bornhuetter-Ferguson* (2008), Mack propose un modèle stochastique pour estimer la variance de l'estimateur des provisions $\hat{R}_i^{BF}$ obtenu par la méthode Bornhuetter-Ferguson et défini par :

$$\hat{R}_i^{BF} = (1 - \hat{z}_{I+1-i})\hat{U}_i,$$

avec $\hat{U}_i$ un estimateur a priori de la charge ultime pour l'année de survenance i et $\hat{z}_j$ la cadence de règlement cumulée après j années de développement.

Mack est alors le premier à proposer une formule fermée pour l'erreur de prédiction des provisions Bornhuetter-Ferguson. La méthode Bornhuetter-Ferguson de ces provisions considère le développement des sinistres jusqu'à l'ultime ; l'estimation des paramètres dans le modèle de Mack doit donc prendre en compte la queue de développement. C'est pourquoi les années de développement j appartiennent à

l'intervalle $\llbracket 1, I + 1 \rrbracket$ dans ce modèle.

Selon Mack, l'actuaire est chargé d'opérer une sélection des facteurs de développement qui diffèrent par conséquent des facteurs Chain Ladder couramment utilisés. Un estimateur a priori de la charge ultime $\hat{U}_i$ et l'incertitude associée à cet estimateur sont également demandés à l'actuaire. Ainsi, l'erreur de prédiction des provisions Bornhuetter-Ferguson dépend en grande partie du jugement de l'actuaire.

Mack modélise les règlements incrémentaux $X_{i,j}$, $i \in \{1, \dots, I\}$, $j \in \{1, \dots, I+1\}$ de la manière suivante :

- les incréments $X_{i,j}$ sont indépendants,
- il existe des paramètres x_i , y_j tels que :

$$\mathbb{E}[X_{i,j}] = x_i y_j \text{ et } \sum_{j=1}^{I+1} y_j = 1,$$

avec $x_i = \mathbb{E}[\hat{U}_i]$ et y_j la cadence de règlement non cumulée.

- il existe des constantes s_j^2 telles que $\mathbb{V}ar(X_{i,j}) = x_i s_j^2$.

Sachant les ultimes, les estimateurs suivants sont sans biais :

$$\hat{y}_j = \frac{\sum_{i=1}^{I+1-j} X_{i,j}}{\sum_{i=1}^{I+1-j} x_i} \text{ pour } j \in \{1, \dots, I\},$$

$$\hat{s}_j^2 = \frac{1}{I-j} \sum_{i=1}^{I+1-j} \frac{(X_{i,j} - x_i \hat{y}_j)^2}{x_i} \text{ pour } j \in \{1, \dots, I-1\}.$$

Tout comme le modèle Chain-Ladder de Mack, le modèle Bornhuetter-Ferguson stochastique comporte beaucoup de paramètres ; un inconvénient majeur de ce modèle est d'avoir des cadences incrémentales dont la somme est souvent différente de l'unité. Cependant, l'actuaire peut effectuer une régression sur les logarithmes des facteurs de développement pour réduire le nombre de paramètres et équilibrer les estimations. Après l'ajustement et l'extrapolation des paramètres $\hat{y}_1^*, \dots, \hat{y}_{J+1}^*$ et $\hat{s}_1^*, \dots, \hat{s}_{J+1}^*$, la nouvelle provision Bornhuetter-Ferguson est estimée par :

$$\hat{R}_i^* = (\hat{y}_{J+2-i}^* + \dots, \hat{y}_{J+1}^*) \times \hat{U}_i = (1 - \hat{z}_{J+1-i}^*) \times \hat{U}_i \text{ avec } \hat{z}_k^* = \sum_{j=1}^k \hat{y}_j^*.$$

• **L'erreur de prédiction de la provision Bornhuetter-Ferguson $\hat{R}_i^*$**

L'erreur de prédiction de la provision Bornhuetter-Ferguson $\hat{R}_i^*$ est définie par l'espérance conditionnelle suivante :

$$MSEP(\hat{R}_i^*) = \mathbb{E}[(\hat{R}_i^* - R_i)^2 | X_{i,1}, \dots, X_{i,J+1-i}].$$

L'hypothèse d'indépendance des $X_{i,j}$ permet d'écrire les égalités suivantes :

$$\begin{aligned} MSEP(\hat{R}_i^*) &= \mathbb{E}[(\hat{R}_i^* - R_i)^2] \\ &= \mathbb{V}ar(\hat{R}_i^* - R_i) + (\mathbb{E}[\hat{R}_i^*] - \mathbb{E}[R_i])^2 \\ &= \mathbb{V}ar(\hat{R}_i^*) + \mathbb{V}ar(R_i). \end{aligned}$$

L'erreur de prédiction est ainsi la somme de l'erreur d'estimation $\mathbb{V}ar(\hat{R}_i^*)$ liée aux paramètres et de l'erreur de processus $\mathbb{V}ar(R_i)$. L'erreur de processus est estimée simplement par :

$$\widehat{\mathbb{V}ar}(R_i) = (\hat{s}_{J+2-i}^{2*} + \dots + \hat{s}_{J+1}^{2*}) \times \hat{U}_i.$$

Pour obtenir l'erreur d'estimation, il faut auparavant déterminer l'incertitude sur les ultimes a priori $\hat{U}_i$ et sur les cadences de règlements cumulées estimées $\hat{z}_j^*$. L'erreur d'estimation des ultimes a priori s'écrit :

$$\widehat{MSE}(\hat{U}_i) = \frac{P_i}{I-1} \sum_{j=1}^J P_j \left(\frac{\hat{U}_j}{P_j} - \frac{\sum_{k=1}^I \hat{U}_k}{\sum_{k=1}^I P_k} \right)^2, \text{ pour } i \in \{1, \dots, I\},$$

sous l'hypothèse de non-corrélation des $\hat{U}_i$.

L'erreur d'estimation des cadences de règlements cumulées $\hat{z}_j^*$ est :

$$\begin{cases} \widehat{MSE}(\hat{z}_j^*) = \min \left((\widehat{MSE}(\hat{y}_1^*) + \dots + \widehat{MSE}(\hat{y}_j^*), \widehat{MSE}(\hat{y}_{j+1}^*) + \dots + \widehat{MSE}(\hat{y}_{J+1}^*)) \right), \\ \widehat{MSE}(\hat{z}_{J+1}^*) = 0. \end{cases}$$

L'incertitude sur les cadences de règlements cumulées $\hat{z}_j^*$ est donc estimée à partir de l'erreur d'estimation des cadences de règlements incrémentales $\hat{y}_j^*$:

$$\widehat{MSE}(\hat{y}_j^*) = \frac{\hat{s}_j^{2*}}{\sum_{i=1}^{I+1-j} \hat{U}_i}, \text{ pour } j \in \{1, \dots, J\}.$$

L'incertitude sur le dernier facteur de développement $\hat{y}_{J+1}^*$ est déterminée de manière exogène. Il est fréquent de poser :

$$\widehat{MSE}(\hat{y}_{J+1}^*) = \frac{1}{2} \hat{y}_{J+1}^*.$$

On peut finalement écrire l'erreur de prédiction de la provision Bornhuetter-Ferguson $\hat{R}_i^*$ pour l'année de survenance i , $i \in \{1, \dots, I\}$, de la manière suivante :

$$\begin{aligned} \widehat{MSEP}(\hat{R}_i^*) &= \underbrace{(\hat{s}_{J+2-i}^{2*} + \dots + \hat{s}_{J+1}^{2*}) \hat{U}_i}_{\text{Erreur de processus}} \\ &+ \underbrace{(\hat{U}_i^2 + \widehat{MSE}(\hat{U}_i)) \times \widehat{MSE}(\hat{z}_{J+1-i}^*) + \widehat{MSE}(\hat{U}_i) \times (1 - \hat{z}_{J+1-i}^*)^2}_{\text{Erreur d'estimation}}. \end{aligned}$$

L'erreur de prédiction de la provision globale $\hat{R}^*$ est la somme des erreurs de processus et d'estimation. Elle est estimée par :

$$\begin{aligned} \widehat{MSEP}(\hat{R}^*) &= \widehat{\mathbb{V}ar}(R) + \widehat{\mathbb{V}ar}(R^*) \\ &= \underbrace{\sum_{i=1}^I (\hat{s}_{I+2-i}^{2*} + \dots + \hat{s}_{I+2-i}^{2*}) \times \hat{U}_i}_{\text{Erreur de processus } \widehat{\mathbb{V}ar}(R)} \\ &+ \underbrace{\sum_{i=1}^I (\hat{U}_i^2 + \widehat{MSE}(\hat{U}_i)) \times \widehat{MSE}(\hat{z}_{J+1-i}^*) + \widehat{MSE}(\hat{U}_i) \times (1 - \hat{z}_{J+1-i}^*)^2 + 2 \sum_{i < j} \widehat{\text{Cov}}(\hat{R}_i^*, \hat{R}_j^*)}_{\text{Erreur d'estimation } \widehat{\mathbb{V}ar}(R^*)}. \end{aligned}$$

La covariance est définie par :

$$\begin{aligned}\widehat{\text{Cov}}(\hat{R}_i^*, \hat{R}_j^*) &= \hat{\rho}_{i,j}^U \sqrt{\widehat{MSE}(\hat{U}_i)} \sqrt{\widehat{MSE}(\hat{U}_j)} (1 - \hat{z}_{J+1-i}^*)(1 - \hat{z}_{J+1-j}^*) \\ &\quad + \hat{\rho}_{i,j}^z \sqrt{\widehat{MSE}(\hat{z}_{J+1-i}^*)} \sqrt{\widehat{MSE}(\hat{z}_{J+1-j}^*)} \hat{U}_i \hat{U}_j,\end{aligned}$$

où les estimations des coefficients de corrélation sont les suivantes :

$$\hat{\rho}_{i,j}^z = \sqrt{\frac{\hat{z}_{J+1-j}^*(1 - \hat{z}_{J+1-i}^*)}{\hat{z}_{J+1-i}^*(1 - \hat{z}_{J+1-j}^*)}}, \text{ pour } i < j \text{ et } \hat{z}_1^* \leq \dots \leq \hat{z}_{J+1}^*,$$

et

$$\begin{cases} \hat{\rho}_{i,j}^U = \frac{1}{\sqrt{I}} & \text{pour un coefficient de corrélation constant,} \\ \text{ou } \hat{\rho}_{i,j}^U = \frac{1}{1+|i-j|} & \text{pour un coefficient de corrélation décroissant.} \end{cases}$$

Un avantage certain de ce modèle est la liberté laissée à l'actuaire d'apporter ses connaissances spécifiques à la branche considérée : l'estimation de la charge ultime a priori et des paramètres est à la charge de l'actuaire. Les cadences de règlements diffèrent des cadences Chain Ladder, ce qui montre l'indépendance de cette méthode et le poids des choix de l'actuaire. Cette méthode peut néanmoins nécessiter du temps lors de l'estimation des paramètres.

5.2.2 La méthode d'Alai, Merz et Wüthrich

L'article *Mean Square Error of Prediction in the Bornhuetter-Ferguson Claims Reserving Method* (2009) de D.H. Alai, M. Merz et M.V. Wüthrich permet de déterminer l'erreur de prédiction à l'ultime des provisions dans le cadre de la méthode Bornhuetter-Ferguson. Pour cela, les auteurs utilisent les modèles linéaires généralisés et le modèle de Poisson surdispersé (ODP) qui permet de répliquer les provisions Bornhuetter-Ferguson. En 2010, les auteurs simplifient l'expression de l'erreur de prédiction grâce aux propriétés asymptotiques du maximum de vraisemblance dans l'article *Prediction Uncertainty in the Bornhuetter-Ferguson Claims Reserving Method: Revisited*.

• Le modèle de Poisson surdispersé

Ce modèle a déjà été évoqué précédemment dans la méthode du bootstrap. On ne rappelle ici que les hypothèses du modèle :

- les règlements incrémentaux $X_{i,j}$ sont indépendants et distribués selon le modèle de Poisson surdispersé (ODP) et il existe des paramètres positifs $\gamma_1, \dots, \gamma_I, \mu_1, \dots, \mu_I$ et $\phi > 0$ tels que:

$$\begin{aligned}\mathbb{E}[X_{i,j}] &= m_{i,j} = \mu_i \gamma_j \\ \text{Var}(X_{i,j}) &= \phi m_{i,j} \\ \text{et } \sum_{j=1}^I \gamma_j &= 1,\end{aligned}$$

avec ϕ le paramètre de dispersion et γ_k les cadences de règlements non cumulées,

- les estimateurs non biaisés $\hat{\nu}_k$, $k \in \{1, \dots, I\}$, des charges ultimes a priori $\mu_k = \mathbb{E}[C_{k,I}]$ sont des variables aléatoires indépendantes,
- les règlements $X_{i,j}$ et les estimateurs des charges ultimes $\hat{\nu}_k$ sont indépendants pour tout i, j, k .

Remarques :

- Pour l'estimation de l'erreur de prédiction, une estimation de l'incertitude des charges ultimes a priori $\hat{\nu}_k$ est nécessaire. On fait l'hypothèse qu'une variance a priori $\widehat{\text{Var}}(\hat{\nu}_k)$ est donnée de manière exogène.
- Dans ce modèle, les règlements incrémentaux doivent être positifs. Le paramètre de dispersion ϕ ne dépend ni de l'année de survenance i ni de l'année de développement j . Il peut être estimé par la méthode du maximum de vraisemblance ou bien en utilisant les résidus de Pearson :

$$\hat{\phi} = \frac{1}{d} \sum_{i+j \leq I+1} \frac{(X_{i,j} - \hat{m}_{i,j})^2}{\hat{m}_{i,j}},$$

avec $d = \frac{I(I+1)}{2} - 2I + 1$ le degré de liberté du modèle et $\hat{m}_{i,j} = \hat{\mu}_i \hat{\gamma}_j$.

La fonction lien de ce modèle ODP est la fonction logarithmique. Avec les notations introduites précédemment pour les modèles linéaires généralisés, le prédicteur linéaire est :

$$\eta_{i,j} = g(m_{i,j}) = \log(m_{i,j}) = m + \alpha_i + \beta_j, \text{ avec } \alpha_1 = \beta_1 = 0.$$

• **L'estimateur de la charge ultime par Bornhuetter-Ferguson**

L'estimateur de la charge ultime espérée par Bornhuetter-Ferguson repose sur les données $\mathcal{D}_I = \{X_{i,j}, i + j \leq I + 1\}$ disponibles à l'instant I et sur les données externes ou le jugement d'expert. L'estimateur s'écrit:

$$\hat{C}_{i,I}^{BF} = C_{i,I-i+1} + \hat{\nu}_i \sum_{j>I-i+1} \hat{\gamma}_j,$$

avec $\hat{\gamma}_j$ l'estimateur par maximum de vraisemblance de γ_j qui correspond à la cadence Chain-Ladder non cumulée, et $\hat{\nu}_i$ un estimateur a priori de $\mathbb{E}[C_{i,I}]$ déterminé de manière exogène.

• **L'erreur de prédiction de l'estimateur $\hat{C}_{i,j}^{BF}$**

L'erreur de prédiction *conditionnelle* de l'estimateur de la charge ultime par Bornhuetter-Ferguson $\hat{C}_{i,j}^{BF}$ pour chaque année de survenance i , $i \in \{1, \dots, I\}$ est définie par :

$$MSEP_{C_{i,I}|\mathcal{D}_I}(\hat{C}_{i,I}^{BF}) = \mathbb{E}[(\hat{C}_{i,I}^{BF} - C_{i,I})^2 | \mathcal{D}_I].$$

Cette espérance peut se réécrire de la manière suivante¹⁰ :

$$MSEP_{C_{i,I}|\mathcal{D}_I}(\hat{C}_{i,I}^{BF}) = \underbrace{\sum_{j>I-i+1} \text{Var}(X_{i,j})}_{\text{Erreur de processus}} + \underbrace{\left(\sum_{j>I-i+1} \hat{\gamma}_j \right)^2 \text{Var}(\hat{\mu}_i) + \mu_i^2 \left(\sum_{j>I-i+1} \hat{\gamma}_j - \sum_{j>I-i+1} \gamma_j \right)^2}_{\text{Erreur d'estimation}}.$$

L'**erreur de processus** provient du mouvement stochastique de $X_{i,j}$; elle correspond à l'incertitude liée aux hypothèses du modèle. L'**erreur d'estimation** comporte deux termes : il s'agit d'une part de l'incertitude sur les estimations des ultimes a priori $\hat{\mu}_i$ et, d'autre part, de l'incertitude sur les estimations des cadences de règlements $\hat{\gamma}_j$ obtenues par maximum de vraisemblance .

10. Les étapes du raisonnement sont disponibles en annexe.

L'erreur de prédiction pour chaque année de survénance $i \in \{1, \dots, I\}$ est :

$$\widehat{MSEP}_{C_{i,I}|\mathcal{D}_I}(\hat{C}_{i,I}^{BF}) = \sum_{j>I-i+1} \hat{\phi}\hat{\nu}_i\hat{\gamma}_j + \left(\sum_{j>I-i+1} \hat{\gamma}_j \right)^2 \widehat{\mathbb{V}ar}(\hat{\nu}_i) + \hat{\nu}_i^2 \sum_{\substack{j>I-i+1 \\ l>I-i+1}} g_{j,l}$$

avec $\mathcal{G} = (g_{j,l})_{j,l=1,\dots,I}$ une matrice établie à partir de la matrice d'information de Fisher. On pourra se reporter à l'annexe pour plus de détails sur cette matrice.

L'erreur de prédiction *conditionnelle* de l'estimateur de la charge ultime par Bornhuetter-Ferguson $\hat{C}_{i,j}^{BF}$ pour l'ensemble des années de survénance est définie par :

$$MSEP_{\sum_{i=1}^I C_{i,I}|\mathcal{D}_I} \left(\sum_{i=1}^I \hat{C}_{i,I}^{BF} \right) = \mathbb{E} \left[\left(\sum_{i=1}^I \hat{C}_{i,I}^{BF} - \sum_{i=1}^I C_{i,I} \right)^2 \middle| \mathcal{D}_I \right],$$

que l'on peut réécrire ainsi :

$$MSEP_{\sum_{i=1}^I C_{i,I}|\mathcal{D}_I} \left(\sum_{i=1}^I \hat{C}_{i,I}^{BF} \right) = \sum_{i=1}^I MSEP_{C_{i,I}|\mathcal{D}_I}(\hat{C}_{i,I}^{BF}) + 2 \sum_{i<k} \mu_i \mu_k \sum_{\substack{j>I-i+1 \\ l>I-k+1}} (\hat{\gamma}_j - \gamma_j)(\hat{\gamma}_l - \gamma_l).$$

L'erreur de prédiction pour l'ensemble des années de survénance est :

$$\widehat{MSEP}_{\sum_{i=1}^I C_{i,I}|\mathcal{D}_I} \left(\sum_{i=1}^I \hat{C}_{i,I}^{BF} \right) = \sum_{i=1}^I \widehat{MSEP}_{C_{i,I}|\mathcal{D}_I}(\hat{C}_{i,I}^{BF}) + 2 \sum_{i<k} \hat{\nu}_i \hat{\nu}_k \sum_{\substack{j>I-i+1 \\ l>I-k+1}} g_{j,l}.$$

Cette méthode a l'avantage de répliquer les provisions estimées par Bornhuetter-Ferguson et permet d'obtenir la variance des provisions à partir d'une formule fermée. Contrairement à la méthode de Mack (2008), elle sépare les erreurs d'estimation de la charge ultime a priori et des cadences de règlements. Les cadences sont d'ailleurs issues du modèle Chain Ladder, à la différence du modèle du Mack (2008). Ainsi, l'actuaire peut choisir entre deux approches dans l'estimation des paramètres.

L'approche Bornhuetter-Ferguson de la mesure du risque à l'ultime fait intervenir des données exogènes au triangle de liquidation. La provision estimée est donc plus proche de la provision connue a posteriori que la provision Chain Ladder, et l'incertitude des provisions moindre.

L'actuaire peut opter pour l'approche Chain Ladder et apporter son jugement d'expert ; grâce au facteur de queue de développement, il augmente l'incertitude sur les provisions. Il décide alors quel modèle doit être privilégié en fonction de la robustesse des résultats ou d'une incertitude moindre sur les provisions. L'actuaire peut aussi faire valoir son expérience à travers l'approche Bornhuetter-Ferguson et utiliser des données exogènes au triangle de liquidation. La prise en compte d'un indicateur d'exposition au risque réduit ainsi l'incertitude à l'ultime des provisions. Les deux méthodes stochastiques présentées diffèrent principalement dans l'estimation des paramètres et il revient à l'actuaire de choisir la démarche Bornhuetter-Ferguson qu'il juge la plus appropriée.

Troisième partie

L'influence du jugement d'expert et de données exogènes sur la vision à un an du risque de provisionnement

Une des principales exigences réglementaires introduites par Solvabilité II est la nécessité de mesurer la volatilité de première année de déviation des provisions. Le point de vue est ainsi entièrement modifié. Jusqu'à présent, on s'attachait à quantifier l'incertitude à long terme liée au déroulement complet de la charge sinistre ; désormais, l'horizon temporel à se fixer est la situation comptable dans un an. On se place alors dans une vision à court terme.

Chaque année, des observations supplémentaires viennent compléter le triangle de liquidation : une diagonale constituée du montant des règlements effectués dans l'année qui s'est écoulée s'ajoute au triangle. La charge ultime est alors ré-estimée pour tenir compte de cette nouvelle information disponible. Si l'estimation de la provision est revue à la hausse, un mali est constaté. En revanche, si cette estimation est revue à la baisse, on parle de boni pour l'organisme assureur. L'objectif est alors de quantifier l'incertitude associée à ce montant de boni/mali. En d'autres termes, le risque de provisionnement à un an est une estimation de la variance de la variation calendaire de la charge de sinistres ultime estimée. Cette problématique est illustrée par le schéma suivant :

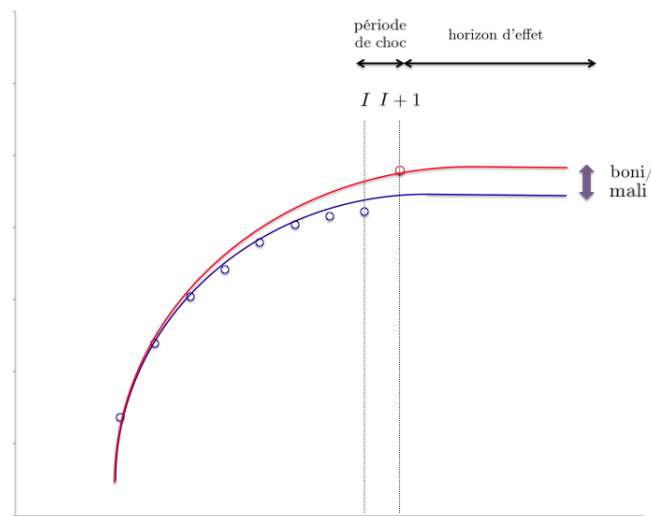


FIGURE 7 – Réestimation de la charge de sinistre ultime et détermination du boni/mali.

On observe en bleu l'estimation du modèle à la date $t = I$ qui prédit la charge de sinistre ultime $\hat{R}_I$ après quelques années de développement. Un scénario de règlements pour l'année comptable $(I, I + 1]$ est généré et est ici illustré par le point rouge. Le modèle est alors réestimé à la date $t = I + 1$ (courbe rouge) en tenant

compte de ce scenario ; on obtient la charge de sinistre ultime réestimée $\hat{R}_{I+1}$ qui est en l'occurrence supérieure à l'estimation précédente. Un mali est ainsi constaté dans cet exemple.

Depuis quelques années, plusieurs méthodes permettant d'atteindre l'horizon temporel d'un an ont vu le jour. Le modèle introduit par Merz et Wüthrich en 2008 est devenu une référence. Ceux-ci ont développé une formule fermée pour déterminer la variance des bonis/malis en s'appuyant sur les travaux de Mack [21]. En 2011, Angoua, Boisseau, Boumezoued et Devineau ont même proposé une formule fermée permettant d'insérer un facteur de queue de développement. En parallèle, des méthodes simulatoires reposant sur la technique du bootstrap faisaient leur apparition, permettant ainsi d'insérer un tel facteur de queue et, de surcroît, de prendre en compte des données exogènes au triangle de liquidation. La problématique qui s'ensuit est de déterminer le capital réglementaire exigé par Solvabilité II grâce à ces méthodes.

6 L'incertitude à un an par une formule fermée

6.1 La formule de Merz et Wüthrich

Dans l'article *Modelling the Claims Development Result for Solvency Purposes*, Merz et Wüthrich proposent une méthode pour obtenir une estimation de la variance de la variation calendaire de la charge de sinistres ultime estimée. Les auteurs désignent d'ailleurs cette variation calendaire comme le **Claims Development Result (CDR)** pour l'année comptable $(I, I + 1]$: il correspond donc à la différence entre les provisions constituées à la date $t = I$ et celles ajustées grâce aux règlements observés pendant l'année comptable $(I, I + 1]$ à la date $t = I + 1$ pour les années de survénance antérieures à I . Le CDR d'une compagnie d'assurance impacte directement son compte de résultat. C'est pourquoi cette vision à court terme du risque de provisionnement est très importante du point de vue de la solidité financière de la compagnie.

Pour étudier le CDR, on peut se placer sous deux angles de vue différents :

- Le point de vue prospectif :

On se place à la date $t = I$ et on fait l'hypothèse que le CDR pour l'année comptable $(I, I + 1]$ sera nul. Il s'agit d'analyser l'incertitude de cette prédiction en répondant à la question suivante : « à quel point la réalisation du CDR va-t-elle dévier de zéro? »

- Le point de vue rétrospectif :

On se place à la date $t = I + 1$ et on obtient une réalisation du CDR. Il faut alors déterminer si cette observation est dans un intervalle raisonnable autour de zéro.

$\mathcal{D}_I$ et $\mathcal{D}_{I+1}$ désignent l'information disponible aux dates $t = I$ et $t = I + 1$ respectivement :

$$\begin{cases} \mathcal{D}_I &= \{C_{i,j}, i + j \leq I \text{ et } i \leq I\}, \\ \mathcal{D}_{I+1} &= \{C_{i,j}, i + j \leq I + 1 \text{ et } i \leq I\} = \mathcal{D}_I \cup \{C_{i,I-i+1}, i \leq I\}. \end{cases}$$

On observe une diagonale supplémentaire des paiements dans l'ensemble $\mathcal{D}_{I+1}$. En termes mathématiques, la tribu $\mathcal{D}_I$ est étendue à la tribu $\mathcal{D}_{I+1}$.

On note R_i^I et R_i^{I+1} les provisions estimées aux dates $t = I$ et $t = I + 1$ et définies respectivement par :

$$\begin{cases} R_i^I &= C_{i,J} - C_{i,I-i+1}, \\ R_i^{I+1} &= C_{i,J} - C_{i,I-i+2}. \end{cases}$$

Le modèle sous-jacent que les auteurs utilisent est celui proposé par Mack [21] pour l'approche Chain Ladder. Les hypothèses de ce dernier modèle doivent donc être vérifiées. Toutefois, les hypothèses du modèle Merz-Wüthrich sont un peu plus restrictives :

Hypothèse 1 :

Les règlements cumulés $C_{i,j}$ sont indépendants pour $i \in \{1, \dots, I\}$.

Hypothèse 2 :

Les règlements cumulés $C_{i,j}$, $j \in \{1, \dots, J\}$, forment une chaîne de Markov et il existe des constantes strictement positives f_j et σ_j telles que pour $j \in \{2, \dots, J\}$ et $i \in \{1, \dots, I\}$:

$$\begin{aligned} \mathbb{E}[C_{i,j}|C_{i,j-1}] &= f_{j-1}C_{i,j-1}, \\ \text{et } \text{Var}(C_{i,j}|C_{i,j-1}) &= \sigma_{j-1}^2 C_{i,j-1}. \end{aligned}$$

Cette hypothèse est en effet plus forte que celle de Mack [21] dans la mesure où ce dernier supposait seulement que le premier et le second moment de $C_{i,j+1}$ ne dépendaient que de $C_{i,j}$ et non de $C_{i,l}$, $l < j$. Les hypothèses Chain Ladder sont également vérifiées et on a :

$$\begin{aligned} \mathbb{E}[C_{i,J}|\mathcal{D}_I] &= C_{i,I-i+1} \prod_{j=I-i+1}^{J-1} f_j, \\ \text{et } \mathbb{E}[C_{i,J}|\mathcal{D}_{I+1}] &= C_{i,I-i+2} \prod_{j=I-i+2}^{J-1} f_j. \end{aligned}$$

Les facteurs de développements Chain Ladder sont estimés par :

$$\hat{f}_j^I = \frac{\sum_{i=1}^{I-j} C_{i,j+1}}{\sum_{i=1}^{I-j} C_{i,j}} \text{ et } \hat{f}_j^{I+1} = \frac{\sum_{i=1}^{I-j+1} C_{i,j+1}}{\sum_{i=1}^{I-j+1} C_{i,j}}.$$

Mack montre dans [21] que ces estimateurs sont sans biais et non corrélés.

Un estimateur non biaisé de $\mathbb{E}[C_{i,J}|\mathcal{D}_I]$, $j \geq I - i + 1$, conditionnellement à $C_{i,I-i}$, est :

$$\hat{C}_{i,J}^I = C_{i,I-i+1} \times \hat{f}_{I-i+1}^I \times \dots \times \hat{f}_{j-1}^I.$$

Un estimateur non biaisé de $\mathbb{E}[C_{i,J}|\mathcal{D}_{I+1}]$, $j \geq I - i + 2$, conditionnellement à $C_{i,I-i+2}$, est :

$$\hat{C}_{i,J}^{I+1} = C_{i,I-i+2} \times \hat{f}_{I-i+2}^{I+1} \times \dots \times \hat{f}_{j-1}^{I+1}.$$

Un estimateur non biaisé de $\mathbb{E}[R_i^I|\mathcal{D}_I]$ sachant $C_{i,I-i+1}$ est :

$$\hat{R}_i^{\mathcal{D}_I} = \hat{C}_{i,J}^I - C_{i,I-i+1}.$$

Un estimateur non biaisé de $\mathbb{E}[R_i^{I+1}|\mathcal{D}_{I+1}]$ sachant $C_{i,I-i+2}$ est :

$$\hat{R}_i^{\mathcal{D}_{I+1}} = \hat{C}_{i,J}^{I+1} - C_{i,I-i+2}.$$

• **La définition mathématique du CDR**

On définit le CDR *réel* pour l'année de survenance i et l'année comptable $(I, I+1]$ par :

$$\begin{aligned} CDR_i(I+1) &= \mathbb{E}[R_i^I|\mathcal{D}_I] - (X_{i,I-i+2} + \mathbb{E}[R_i^{I+1}|\mathcal{D}_{I+1}]) \\ &= \mathbb{E}[C_{i,J}|\mathcal{D}_I] - \mathbb{E}[C_{i,J}|\mathcal{D}_{I+1}], \end{aligned}$$

où les $X_{i,I-i+1} = C_{i,I-i+1} - C_{i,I-i}$ désignent les règlements incrémentaux.

Le CDR réel agrégé s'écrit simplement $\sum_{i=2}^I CDR_i(I+1)$.

Comme les $(C_{i,j})_{j \geq 1}$ ont les propriétés d'une martingale, on peut écrire :

$$\mathbb{E}[CDR_i(I+1)|\mathcal{D}_I] = 0. \quad (1)$$

Pour les facteurs de développement Chain Ladder connus f_j , le CDR attendu (vu à la date $t = I$) est nul. Ainsi, on ne doit pas anticiper de boni ni de mali.

Mais les facteurs de développement Chain Ladder sont en général inconnus, et, par conséquent, le CDR réel n'est pas observable. C'est pourquoi on s'intéresse dès lors au CDR *observable* ; il permet d'estimer le CDR réel et est défini de la manière suivante :

$$\begin{aligned} \widehat{CDR}_i(I+1) &= \hat{R}_i^{\mathcal{D}_I} - (X_{i,I-i+2} + \hat{R}_i^{\mathcal{D}_{I+1}}) \\ &= \hat{C}_{i,J}^I - \hat{C}_{i,J}^{I+1}, \end{aligned}$$

pour l'année de survenance i et l'année comptable $(I, I+1]$.

$\hat{R}_i^{\mathcal{D}_I}$ est un estimateur non biaisé de $\mathbb{E}[R_i^I|\mathcal{D}_I]$ sachant $C_{i,I-i+1}$. Il est tel que :

$$\hat{R}_i^{\mathcal{D}_I} = \hat{C}_{i,J}^I - C_{i,I-i+1}, \text{ pour } i \in \{2, \dots, I\}.$$

$\hat{R}_i^{\mathcal{D}_{I+1}}$ est un estimateur non biaisé de $\mathbb{E}[R_i^{I+1}|\mathcal{D}_{I+1}]$ sachant $C_{i,I-i+2}$. Il est tel que :

$$\hat{R}_i^{\mathcal{D}_{I+1}} = \hat{C}_{i,J}^{I+1} - C_{i,I-i+2}, \text{ pour } i \in \{2, \dots, I\}.$$

Le CDR observable agrégé s'écrit $\sum_{i=2}^I \widehat{CDR}_i(I+1)$.

Il s'agit désormais de quantifier la qualité de l'approximation du CDR réel par le CDR observable. On se place ainsi dans l'approche rétrospective. Le CDR observable apparaît dans le compte de résultat de l'organisme assureur au 31/12/ I et est prédit par la valeur nulle. Dans l'approche prospective, il faut donc juger la qualité de cette prédiction.

• **L'erreur de prédiction conditionnelle du CDR**

Le CDR observable $\widehat{CDR}_i(I+1)$ est un prédicteur $\mathcal{D}_I$ -mesurable du CDR réel $CDR_i(I+1)$. L'erreur de prédiction conditionnelle du CDR dans l'approche prospective mesure l'incertitude sur la valeur nulle prédite pour le CDR observable à la fin de l'année comptable. Cette erreur est définie par :

$$MSEP_{\widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) = \mathbb{E}[(\widehat{CDR}_i(I+1) - 0)^2 | \mathcal{D}_I].$$

L'erreur de prédiction conditionnelle du CDR dans l'approche rétrospective analyse la distance entre le CDR réel et le CDR observable. Elle est définie par :

$$MSEP_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)) = \mathbb{E}[(CDR_i(I+1) - \widehat{CDR}_i(I+1))^2 | \mathcal{D}_I]. \quad (2)$$

Comme $CDR_i(I+1)$ est $\mathcal{D}_I$ -mesurable et de moyenne nulle conditionnellement à $\mathcal{D}_I$, on a également la décomposition suivante :

$$MSEP(\mathbb{E}[\widehat{CDR}_i(I+1)|\mathcal{D}_I]) = \underbrace{\mathbb{V}ar(CDR_i(I+1)|\mathcal{D}_I)}_{\text{Erreur de processus du CDR}} + \underbrace{\hat{\mathbb{E}}\left[\mathbb{E}[\widehat{CDR}_i(I+1)|\mathcal{D}_I]^2 | \mathcal{D}_I\right]}_{\text{Erreur d'estimation du CDR}}.$$

Pour déterminer les erreurs de prédiction conditionnelles ci-dessus, on définit un estimateur non biaisé pour la variance des paramètres σ_j^2 :

$$\hat{\sigma}_j^2 = \frac{1}{I-j} \sum_{i=1}^{I-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j^I \right)^2.$$

Sous les hypothèses du modèle et avec l'information $\mathcal{D}_I$ disponible, l'erreur de processus du CDR pour l'année de survenance i et l'année comptable $(I, I+1]$ est :

$$\mathbb{V}ar(CDR_i(I+1)|\mathcal{D}_I) = \mathbb{E}[C_{i,J}|\mathcal{D}_I]^2 \times \frac{(\sigma_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}},$$

et un estimateur est donné par :

$$\widehat{\mathbb{V}ar}(CDR_i(I+1)|\mathcal{D}_I) = \mathbb{E}[C_{i,J}|\mathcal{D}_I]^2 \times \frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}}.$$

Pour déterminer l'erreur d'estimation du CDR, on définit l'estimateur suivant :

$$\hat{\Delta}_{i,J}^I = \frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{S_{I-i+1}^I} + \sum_{j=I-i+2}^{J-1} \left(\frac{C_{I-j+1,j}}{S_j^I} \right)^2 \frac{(\hat{\sigma}_j^I)^2 / (\hat{f}_j^I)^2}{S_j^I},$$

avec

$$S_j^I = \sum_{i=1}^{I-j} C_{i,j} \text{ et } S_j^{I+1} = \sum_{i=1}^{I-j+1} C_{i,j}.$$

Sous les hypothèses du modèle et avec l'information $\mathcal{D}_I$ disponible, l'erreur d'estimation du CDR pour l'année de survenance i et l'année comptable $(I, I+1]$ est :

$$\hat{\mathbb{E}}\left[\mathbb{E}[\widehat{CDR}_i(I+1)|\mathcal{D}_I]^2 | \mathcal{D}_I\right] = (\hat{C}_{i,J}^I)^2 \hat{\Delta}_{i,J}^I.$$

□ L'erreur de prédiction conditionnelle du CDR réel

D'après les égalités (1) et (2), on peut écrire :

$$MSEP_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)) = \phi_{i,J}^I + (\mathbb{E}[\widehat{CDR}_i(I+1)|\mathcal{D}_I])^2,$$

avec

$$\begin{aligned} \phi_{i,J}^I &= \mathbb{V}ar(CDR_i(I+1) - \widehat{CDR}_i(I+1)|\mathcal{D}_I) \\ &= \underbrace{\mathbb{V}ar(CDR_i(I+1)|\mathcal{D}_I)}_{(A)} + \underbrace{\mathbb{V}ar(\widehat{CDR}_i(I+1)|\mathcal{D}_I)}_{(B)} \\ &\quad - \underbrace{2 \times \text{Cov}(\widehat{CDR}_i(I+1), CDR_i(I+1)|\mathcal{D}_I)}_{(C)}. \end{aligned}$$

$\phi_{i,J}^I$ correspond à l'incertitude liée au processus stochastique et comporte trois termes :

- (A) est l'erreur de modèle conditionnelle du CDR réel,
- (B) est l'erreur de modèle conditionnelle du CDR observable,
- (C) traduit la relation qui existe entre le CDR observable et le CDR réel.

Sous les hypothèses du modèle, l'estimateur de l'erreur de prédiction conditionnelle de l'estimateur du CDR réel à la date $t = I+1$ pour l'année comptable $(I, I+1]$ et pour l'année de survénance i est :

$$\widehat{MSEP}_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)) = (\hat{C}_{i,J}^I)^2(\hat{\Phi}_{i,J}^I + \hat{\Delta}_{i,J}^I),$$

avec $\hat{\Phi}_{1,J}^I = 0$ et pour $i > 1$,

$$\hat{\Phi}_{i,J}^I = \left[1 + \frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}} \right] \left(\prod_{l=I-i+2}^{J-1} \left(1 + \frac{(\hat{\sigma}_l^I)^2 / (\hat{f}_l^I)^2}{(S_l^{I+1})^2} C_{I-l+1,l} \right) - 1 \right).$$

Les mêmes observations sont utilisées pour estimer les facteurs de développement Chain Ladder et sont appliquées à des années de survénance différentes ; c'est pourquoi l'agrégation autour des années de survénance doit tenir compte des corrélations entre celles-ci.

Sous les hypothèses du modèle, l'estimateur de l'erreur de prédiction conditionnelle de l'estimateur du CDR réel agrégé à la date $t = I+1$ pour l'année comptable $(I, I+1]$ est :

$$\begin{aligned} \widehat{MSEP}_{\sum_{i=1}^I CDR_i(I+1)|\mathcal{D}_I} \left(\sum_{i=1}^I \widehat{CDR}_i(I+1) \right) &= \sum_{i=1}^I \widehat{MSEP}(\widehat{CDR}_i(I+1)) \\ &\quad + 2 \times \sum_{k>i>0} \hat{C}_{i,J}^I \hat{C}_{k,J}^I (\hat{\Psi}_{i,J}^I + \hat{\Lambda}_{i,J}^I), \end{aligned}$$

avec $\hat{\Psi}_{i,1}^I = 0$ pour $i > 1$ et pour $k > i > 0$,

$$\begin{aligned}\hat{\Psi}_{i,J}^I &= \left(1 + \frac{(\hat{\sigma}_{I-i+1}^I)^2/(\hat{f}_{I-i+1}^I)^2}{S_{I-i+1}^{I+1}}\right) \left(1 + \frac{(\hat{\sigma}_{I-i+1}^I)^2/(\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}}\right)^{-1} \times \hat{\Phi}_{i,J}^I, \\ \hat{\Lambda}_{i,J}^I &= \frac{C_{i,I-i+1}}{S_{I-i+1}^{I+1}} \frac{(\hat{\sigma}_{I-i+1}^I)^2/(\hat{f}_{I-i+1}^I)^2}{S_{I-i+1}^I} + \sum_{j=I-i+2}^{J-1} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}}\right)^2 \frac{(\hat{\sigma}_j^I)^2/(\hat{f}_j^I)^2}{S_j^I}.\end{aligned}$$

□ L'erreur de prédiction conditionnelle du CDR observable

D'après les articles [26], [25] et [20], la variance du CDR observable est composée du biais assimilé à l'erreur d'estimation et d'un terme de variance de processus.

Avec l'information $\mathcal{D}_I$ disponible, l'erreur d'estimation du CDR observable pour l'ensemble du portefeuille et pour l'année comptable $(I, I+1]$ est estimée par :

$$\begin{aligned}(\text{u-bias})^2 &= \hat{\mathbb{E}}_{\mathcal{D}_I} \left[\left(\mathbb{E} \left[\sum_{i=1}^I \widehat{CDR}_i(I+1) \middle| \mathcal{D}_I \right] \right)^2 \right] \\ &= \sum_{i=2}^I (\hat{C}_{i,J}^I)^2 \hat{\Delta}_{i,J}^I + 2 \times \sum_{i>k>1} \hat{C}_{i,J}^I \hat{C}_{k,J}^I \hat{\Lambda}_{k,J}^I\end{aligned}$$

La variance de processus du CDR observable pour l'année comptable $(I, I+1]$ est estimée par :

$$\widehat{\mathbb{V}ar} \left(\sum_{i=1}^I \widehat{CDR}_i(I+1) \middle| \mathcal{D}_I \right) = \sum_{i=2}^I (\hat{C}_{i,J}^I)^2 \hat{\Gamma}_{i,J}^I + 2 \sum_{i>k>1} \hat{C}_{i,J}^I \hat{C}_{k,J}^I \hat{\Upsilon}_{i,k}^I,$$

avec

$$\begin{aligned}\hat{\Gamma}_{i,J}^I &= \widehat{\mathbb{V}ar}(\widehat{CDR}_i(I+1) | \mathcal{D}_I) \\ &= \left(1 + \frac{(\hat{\sigma}_{I-i+1}^I)^2/(\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}}\right) \times \prod_{l=I-i+2}^{J-1} \left(1 + \frac{(\hat{\sigma}_l^I)^2/(\hat{f}_l^I)^2}{(S_l^{I+1})^2} C_{I-l+1,l}\right) - 1,\end{aligned}$$

et pour $k > i > 0$,

$$\begin{aligned}\hat{\Upsilon}_{i,k}^I &= \widehat{\text{Cov}}(\widehat{CDR}_i(I+1), \widehat{CDR}_k(I+1) | \mathcal{D}_I) \\ &= \left(1 + \frac{(\hat{\sigma}_{I-k+1}^I)^2/(\hat{f}_{I-k+1}^I)^2}{S_{I-k+1}^{I+1}}\right) \times \prod_{l=I-k+2}^{J-1} \left(1 + \frac{(\hat{\sigma}_l^I)^2/(\hat{f}_l^I)^2}{(S_l^{I+1})^2} C_{I-l+1,l}\right) - 1.\end{aligned}$$

Finalement, l'estimateur de l'erreur de prédiction conditionnelle du CDR obser-

vale agrégé à la date $t = I + 1$ pour l'année comptable $(I, I + 1]$ est :

$$\begin{aligned}
\widehat{MSEP}_{\sum_{i=1}^I \widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) &= \mathbb{E} \left[\left(\sum_{i=1}^I \widehat{CDR}_i(I+1) - 0 \right)^2 \right] \\
&= (\text{u-bias})^2 + \widehat{\text{Var}} \left(\sum_{i=1}^I \widehat{CDR}_i(I+1) | \mathcal{D}_I \right) \\
&= \sum_{i=1}^I (\hat{C}_{i,J}^I)^2 (\hat{\Gamma}_{i,J}^I + \hat{\Delta}_{i,J}^I) + 2 \sum_{k>i>1} \hat{C}_{i,J}^I \hat{C}_{k,J}^I (\hat{\Upsilon}_{i,k}^I + \hat{\Lambda}_{i,J}^I).
\end{aligned}$$

• **L'approximation de l'erreur de prédiction du CDR**

Merz et Wüthrich simplifient les formules précédentes dans leur article [26] en approchant les produits par des sommes. Ceux-ci s'appuient sur le lemme suivant :

Si les a_j sont des scalaires positifs tels que $1 \gg a_j$, alors :

$$\begin{aligned}
\prod_{j=1}^J (1 + a_j) - 1 &\approx \sum_{i=1}^J a_j, \\
\text{ie } \frac{(\hat{\sigma}_j)^2}{(\hat{f}_j^I)^2} &<< C_{I-j+1,j}.
\end{aligned}$$

On peut alors simplifier les estimateurs introduits précédemment de la manière suivante :

$$\begin{aligned}
\hat{\Phi}_{i,J}^I &\approx \sum_{j=I-i+2}^{J-1} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}} \right)^2 \frac{(\hat{\sigma}_j^I)^2 / (\hat{f}_j^I)^2}{C_{I-j+1,j}}, \\
\hat{\Psi}_{i,J}^I &\approx \frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}}, \\
\hat{\Gamma}_{i,J}^I &\approx \hat{\Phi}_{i,J}^I + \hat{\Psi}_{i,J}^I \geq \hat{\Phi}_{i,J}^I, \\
\hat{\Upsilon}_{i,J}^I &\approx \hat{\Phi}_{i,J}^I + \frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{S_{I-i+1}^{I+1}} \geq \hat{\Phi}_{i,J}^I.
\end{aligned}$$

Par exemple, l'estimateur $\hat{\Phi}_{i,J}^I$ se simplifie de la manière suivante :

$$\begin{aligned}
\hat{\Phi}_{i,J}^I &= \left[1 + \frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}} \right] \left(\prod_{l=I-i+2}^{J-1} \left(1 + \frac{(\hat{\sigma}_l^I)^2 / (\hat{f}_l^I)^2}{(S_l^{I+1})^2} C_{I-l+1,l} \right) - 1 \right) \\
&\approx \left[1 + \frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}} \right] \sum_{j=I-i+2}^{J-1} \frac{(\hat{\sigma}_j^I)^2 / (\hat{f}_j^I)^2}{C_{I-j+1,j}} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}} \right)^2 \\
&\approx \sum_{j=I-i+2}^{J-1} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}} \right)^2 \frac{(\hat{\sigma}_j^I)^2 / (\hat{f}_j^I)^2}{C_{I-j+1,j}}.
\end{aligned}$$

En se référant au lemme précédent, cette approximation est valable dans la mesure où, en pratique, on a souvent :

$$1 \gg \frac{(\hat{\sigma}_j^I)^2/(\hat{f}_j^I)^2}{C_{I-j+1,j}}.$$

On se place à la date $t = I + 1$ et on considère l'année comptable $(I, I + 1]$. Les estimateurs de l'erreur de prédiction conditionnelle des estimateurs du CDR réel et du CDR observable pour l'année de survenance i sont respectivement :

$$\begin{aligned}\widehat{MSEP}_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)) &= (\hat{C}_{i,J}^I)^2(\hat{\Phi}_{i,J}^I + \hat{\Delta}_{i,J}^I), \\ \widehat{MSEP}_{\widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) &= (\hat{C}_{i,J}^I)^2(\hat{\Gamma}_{i,J}^I + \hat{\Delta}_{i,J}^I).\end{aligned}$$

On remarque que :

$$\begin{aligned}\widehat{MSEP}_{\widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) &= \widehat{MSEP}_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)) + \widehat{\mathbb{V}ar}(CDR_i(I+1)|\mathcal{D}_I) \\ &\geq \widehat{MSEP}_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)),\end{aligned}$$

car $\widehat{\mathbb{V}ar}(CDR_i(I+1)|\mathcal{D}_I)$ est estimé par $(\hat{C}_{i,J}^I)^2\hat{\Phi}_{i,J}^I \geq 0$.

Les estimateurs de l'erreur de prédiction conditionnelle des estimateurs du CDR réel et du CDR observable agrégés sont respectivement :

$$\begin{aligned}\widehat{MSEP}_{\sum_{i=1}^I CDR_i(I+1)|\mathcal{D}_I}\left(\sum_{i=1}^I \widehat{CDR}_i(I+1)\right) &= \sum_{i=1}^I (\hat{C}_{i,J}^I)^2(\hat{\Phi}_{i,J}^I + \hat{\Delta}_{i,J}^I) \\ &\quad + 2 \sum_{k>i>0} \hat{C}_{i,J}^I \hat{C}_{k,J}^I (\hat{\Psi}_{k,J}^I + \hat{\Lambda}_{k,J}^I),\end{aligned}$$

$$\widehat{MSEP}_{\sum_{i=1}^I \widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) = \sum_{i=1}^I (\hat{C}_{i,J}^I)^2(\hat{\Gamma}_{i,J}^I + \hat{\Delta}_{i,J}^I) + 2 \sum_{k>i>0} \hat{C}_{i,J}^I \hat{C}_{k,J}^I (\hat{\Upsilon}_{k,J}^I + \hat{\Lambda}_{k,J}^I).$$

On remarque également que :

$$\begin{aligned}\widehat{MSEP}_{\sum_{i=1}^I \widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) &= \widehat{MSEP}_{\sum_{i=1}^I CDR_i(I+1)|\mathcal{D}_I}\left(\sum_{i=1}^I \widehat{CDR}_i(I+1)\right) \\ &\quad + \sum_{i=1}^I \widehat{\mathbb{V}ar}(CDR_i(I+1)|\mathcal{D}_I) + 2 \sum_{k>i>0} \hat{C}_{i,J}^I \hat{C}_{k,J}^I (\hat{\Upsilon}_{k,J}^I + \hat{\Lambda}_{k,J}^I) \\ &\geq \widehat{MSEP}_{\sum_{i=1}^I CDR_i(I+1)|\mathcal{D}_I}\left(\sum_{i=1}^I \widehat{CDR}_i(I+1)\right).\end{aligned}$$

Ainsi, l'estimateur de l'erreur conditionnelle des estimateurs du CDR observable est supérieur à celui du CDR réel ; on observe cette inégalité à la fois dans le cas agrégé et par exercice de survenance.

• **La comparaison avec la méthode de Mack [21]**

L'erreur de prédiction par la méthode de Mack se présente sous la forme suivante pour l'année de survenance i :

$$\widehat{MSEP}(\hat{R}_i) = \hat{C}_{i,J}^2 \sum_{j=I-i+1}^{J-1} \left(\underbrace{\frac{(\hat{\sigma}_j^I)^2 / (\hat{f}_j^I)^2}{C_{i,j}}}_{\text{Erreur de processus}} + \underbrace{\frac{(\hat{\sigma}_j^I)^2 / (\hat{f}_j^I)^2}{S_j}}_{\text{Erreur d'estimation}} \right)$$

L'erreur de prédiction du CDR réel par la méthode de Merz et Wüthrich s'écrit quant à elle :

$$\begin{aligned} \widehat{MSEP}_{\sum_{i=1}^I \widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) &\approx (\hat{C}_{i,J})^2 \left[\underbrace{\frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{C_{i,I-i+1}} + \frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{S_{I-i+1}^I}}_A \right. \\ &\quad \left. + \underbrace{\sum_{j=I-i+2}^{J-1} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}} \right)^2 \frac{(\hat{\sigma}_j^I)^2 / (\hat{f}_j^I)^2}{S_j^I}}_B \right]. \end{aligned}$$

Le terme (A) correspond à l'erreur de processus et d'estimation sur la première diagonale ($j = I - i + 1$) et le terme (B) à l'erreur d'estimation uniquement sur les diagonales suivantes ($j \geq I - i + 1$). De plus, le terme d'erreur d'estimation est négligeable car $\frac{C_{I-j+1,j}}{S_j^{I+1}} \leq 1$.

Dans le cas agrégé, Merz et Wüthrich estime l'erreur de prédiction par :

$$\begin{aligned} \widehat{MSEP}_{\sum_{i=1}^I \widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) &= \sum_{i=1}^I \widehat{MSEP}_{\widehat{CDR}_i(I+1)|\mathcal{D}_I}(0) \\ &+ 2 \sum_{k>i>0} \hat{C}_{i,J}^I \hat{C}_{k,J}^I \left[\frac{(\hat{\sigma}_{I-i+1}^I)^2 / (\hat{f}_{I-i+1}^I)^2}{S_{I-i+1}^I} + \sum_{j=I-i+2}^{J-1} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}} \right)^2 \frac{(\hat{\sigma}_j^I)^2 / (\hat{f}_j^I)^2}{S_j^I} \right]. \end{aligned}$$

L'erreur de prédiction du CDR réel agrégé considère l'erreur d'estimation pour la diagonale suivante ($j = I - i$) mais écrase toutes les erreurs d'estimation des diagonales d'après ($j \geq I - i$) par le facteur $\frac{C_{I-j,j}}{S_j^{I+1}} \leq 1$. Ces comparaisons avec les formules de Mack ne sont possibles qu'après approximation.

La méthode de Merz et Wüthrich est une méthode non simulatoire, au même titre que celle de Mack [21]. Elle est devenue en quelques années seulement une référence dans la détermination du risque de provisionnement à horizon un an ; elle fournit en effet une expression fermée de la volatilité à un an de la variation calendaire de la charge de sinistres ultime estimée. L'importante utilisation de la méthode Chain Ladder induit une préférence pour ce modèle Merz Wüthrich qui est relativement simple à implémenter et non paramétrique. Il est possible dans certains triangles que l'incertitude à un an soit supérieure à l'incertitude à long terme.

En revanche, cette méthode ne permet pas a priori d'appliquer un facteur de queue de développement, ce qui peut véritablement s'éloigner de la réalité de la

sinistralité de certaines branches. D'autre part, il est impossible d'utiliser des modèles sous-jacents de provisionnement autres que le modèle de Mack dans l'approche Chain Ladder, ce qui est un inconvénient majeur. Cette méthode ne fournit pas une distribution du CDR, mais uniquement une estimation des deux premiers moments de $\widehat{CDR}(I+1)$, c'est-à-dire :

$$\mathbb{E}(\widehat{CDR}(I+1)|\mathcal{D}_I) \text{ et } \widehat{\mathbb{V}ar}(\widehat{CDR}(I+1)|\mathcal{D}_I).$$

C'est pourquoi on s'intéresse désormais à une autre approche, à savoir la technique du bootstrap, afin de trouver une alternative aux problèmes évoqués.

6.2 L'insertion d'un facteur de queue de développement

Une formule fermée tenant compte d'un facteur de queue de développement pour l'incertitude à un an du risque de provisionnement a été développée récemment dans l'article *One-year reserve risk including a tail factor: closed formula and bootstrap approaches* (2011). Cet article permet donc de remédier à l'une des limites du modèle de Merz et Wüthrich.

La variance du CDR est là encore la somme de l'erreur d'estimation et de l'erreur de processus :

$$MSEP_{CDR_i(I+1)|\mathcal{D}_I}(0) = \underbrace{MSEP_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1))}_{\text{Erreur d'estimation}} + \underbrace{\widehat{\mathbb{V}ar}(CDR_i(I+1))}_{\text{Erreur de processus}}.$$

L'erreur d'estimation tenant compte du facteur de développement $\hat{f}_{ult}$ et de sa volatilité associée $\hat{\sigma}_{ult}$ pour l'année de survénance i est donnée par :

$$\begin{aligned} & MSEP_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)) \\ &= \hat{C}_{i,ult}^2 \left[\left(1 + \frac{\hat{\sigma}_{ult}^2}{\hat{f}_{ult}^2} \right) \left(1 + \frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{I-i+1}^2 S_{I-i+1}^I} + \sum_{j=I-i+2}^{I-1} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}} \right)^2 \frac{\hat{\sigma}_j^2}{\hat{f}_j^2 S_j^I} \right) - 1 \right], \end{aligned}$$

avec

$$S_j^I = \sum_{i=1}^{I-j} C_{i,j} \text{ et } S_j^{I+1} = \sum_{i=1}^{I-j+1} C_{i,j}.$$

Et on a pour la première année de survénance :

$$\mathbb{V}ar(\widehat{CDR}_1) = C_{1,J}^2 \hat{\sigma}_{ult}^2.$$

Pour l'ensemble des années de survénance, l'erreur d'estimation totale est :

$$\begin{aligned} MSEP_{\sum_i CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)) &= \sum_{i=1}^I MSEP_{CDR_i(I+1)|\mathcal{D}_I}(\widehat{CDR}_i(I+1)) \\ &\quad + 2 \sum_i \sum_j \text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j), \end{aligned}$$

avec, pour i et $j \geq 2$,

$$\begin{aligned} & \text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j) \\ & \approx \hat{C}_{i,ult} \hat{C}_{j,ult} \left[\left(1 + \frac{\hat{\sigma}_{ult}^2}{\hat{f}_{ult}^2} \right) \left(1 + \frac{\hat{\sigma}_{I-i+1}^2 C_{i,I-i+1}}{\hat{f}_{I-i+1}^2 S_{I-i+1}^I S_{I-i+1}^{I+1}} + \sum_{m=I-i+2}^{I-1} \left(\frac{C_{I-m+1,m}}{S_m^{I+1}} \right)^2 \frac{\hat{\sigma}_m^2}{\hat{f}_m^2 S_m^I} \right) - 1 \right], \\ & \text{et } \text{Cov}(\widehat{CDR}_i, \widehat{CDR}_1) = \hat{C}_{i,ult} \hat{C}_{1,ult} \frac{\hat{\sigma}_{ult}^2}{\hat{f}_{ult}^2}. \end{aligned}$$

En ce qui concerne l'erreur de processus, on peut écrire pour l'année de survénance i :

$$\widehat{\mathbb{V}ar}(CDR_i(I+1)) = \hat{C}_{i,ult}^2 \left[\left(1 + \frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{I-i+1}^2 C_{i,I-i+1}} \right) \prod_{j=I-i+2}^{I-1} \left(1 + \frac{\hat{\sigma}_j^2 C_{I-j+1,j}}{\hat{f}_j^2 (S_j^{I+1})^2} \right) - 1 \right].$$

Et on a également : $\widehat{\mathbb{V}ar}(CDR_1(I+1)) = 0$.

Pour l'ensemble des années de survénance, on a :

$$\begin{aligned} \widehat{\mathbb{V}ar}\left(\sum_{i=1}^I CDR_i(I+1)\right) &= \sum_{i=1}^I \widehat{\mathbb{V}ar}(CDR_i(I+1)) \\ &+ 2 \sum_i \sum_j \text{Cov}(CDR_i, CDR_j), \end{aligned}$$

avec, pour i et $j \geq 2$,

$$\begin{aligned} & \text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j) \\ & \approx \hat{C}_{i,ult} \hat{C}_{j,ult} \left[\left(1 + \frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{I-i+1}^2 S_{I-i+1}^{I+1}} \right) \prod_{m=I-i+2}^{I-1} \left(1 + \frac{\hat{\sigma}_m^2 C_{I-m+1,m}}{\hat{f}_m^2 (S_m^{I+1})^2} \right) - 1 \right], \\ & \text{et pour } i \geq 2, \text{ Cov}(\underbrace{CDR_i CDR_1}_0) = 0. \end{aligned}$$

7 La mesure de l'incertitude à partir d'une distribution complète du *Claims Development Result*

La méthode Merz Wüthrich présentée précédemment possède le défaut majeur de ne fournir que les moments d'ordre un et deux. C'est pourquoi on s'intéresse désormais à la technique du bootstrap qui permet notamment d'obtenir une distribution complète d'un estimateur.

On se place dans un premier temps dans le cadre de l'approche Chain Ladder et trois méthodes différentes du bootstrap sont développées. Certaines de ces méthodes sont ensuite modifiées afin de pouvoir insérer un facteur de queue de développement. Enfin, une méthode bootstrap appliquée aux méthodes Bornhuetter-Ferguson et Benktander est explicitée.

7.1 La méthode bootstrap

On se place dans le cadre Chain Ladder. La méthode du bootstrap à l'ultime est modifiée pour tenir compte de la Directive Solvabilité II et mesurer le risque de provisionnement à horizon un an. Ainsi, on présente dans cette partie trois méthodes qui reposent sur la technique du bootstrap et permettent de déterminer l'erreur de prédiction à un an. La première méthode, dite de re-provisionnement, a été développée dès 2008 par Ohlsson & Lauzenings [28] et Diers [13] ; c'est une adaptation directe de la méthode bootstrap à l'ultime. D'autre part, il est rappelé dans [4] comment Boisseau a révisé cette méthode. Celle-ci sera appelée « méthode améliorée » par la suite. Enfin, Angoua, Boisseau, Boumezoued et Devineau [4] proposent une méthode bootstrap dite récursive en s'appuyant notamment sur les travaux de De Felice et Moriconi [12]. Cette méthode permet de répliquer les résultats de Merz et Wüthrich [26].

Comme dans le bootstrap à l'ultime, on se place dans le modèle MLG de Poisson surdispersé (ODP) ; les règlements incrémentaux $X_{i,j}$ sont modélisés par la loi ODP de moyenne $\mu_{i,j}$ et de variance $\phi\mu_{i,j}$ tels que :

$$\mathbb{E}[X_{i,j}] = \mu_{i,j} = \exp(\mu + \alpha_i + \beta_j).$$

7.1.1 La méthode bootstrap de re-provisionnement

Diers et Ohlsson & Lauzenings introduisent la notion de méthode bootstrap de re-provisionnement qui est une adaptation directe de la méthode bootstrap à l'ultime dans le cadre Chain Ladder. Elle est une version MLG de la méthode Merz-Wüthrich ; elle consiste à simuler N fois la diagonale future et à réestimer le montant des provisions par Chain Ladder pour chacune de ces simulations. Le but est d'obtenir une distribution du Claims Development Result (CDR) défini précédemment.

La procédure bootstrap est composée des trois étapes principales suivantes :

- Estimer les provisions à la date $t = I$ par une méthode de provisionnement, en l'occurrence Chain Ladder.

- Simuler les règlements futurs pour l'année calendaire $(I, I + 1]$, c'est-à-dire ceux de la diagonale suivante.
- Estimer les provisions à la date $t = I + 1$ par une méthode de provisionnement sur le triangle augmenté de cette diagonale.

Lacoume explicite ces étapes dans son mémoire [17]. Les premières étapes sont semblables à celles de la procédure à l'ultime et seule la boucle bootstrap est différente :

1. Les quatre premières étapes sont identiques à celles présentées dans la méthode classique du bootstrap. On obtient une estimation des provisions $\hat{R}$ à la date $t = I$ par la méthode Chain Ladder.

2. Répéter N fois les étapes suivantes :

- Effectuer un rééchantillonnage aléatoire avec remise des résidus de Pearson ajustés. On obtient un nouveau triangle de résidus $\tilde{r}_{i,j}^{(k)}$ pour la k^{ieme} boucle bootstrap.
- Reconstituer le pseudo-triangle des règlements incrémentaux par la relation suivante :

$$\tilde{X}_{i,j}^{(k)} = \hat{X}_{i,j} + \tilde{r}_{i,j}^{(k)} \sqrt{\hat{X}_{i,j}}, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \leq J - i + 1.$$

- Calculer les facteurs de développement Chain Ladder à partir du nouveau pseudo-triangle des règlements cumulés.
 - Estimer la diagonale suivante grâce à ces facteurs de développement. Estimer les règlements incrémentaux de la diagonale dont les valeurs déterminent la moyenne utilisée pour la simulation de l'étape suivante.
 - Simuler les règlements de la diagonale par une loi gamma de moyenne $\tilde{\mu}_{i,j}$ et de variance $\hat{\phi}\tilde{\mu}_{i,j}$. On tient compte de l'erreur de processus dans cette étape.
 - Estimer les provisions par la méthode Chain Ladder et stocker le résultat. L'erreur d'estimation est prise en compte dans cette étape.
3. Obtenir les distributions de l'estimation de la provision et du CDR et déterminer les erreurs de prédiction qui leur sont associées.

On illustre cette procédure par le schéma ci-après.

Cette méthode bootstrap de re-provisionnement fournit une distribution du CDR, de sorte qu'elle permet de répondre aux exigences de Solvabilité II. De plus, cette méthode repose sur des simulations et on peut l'étendre à horizon k années. L'insertion d'un facteur de queue est possible dans cette procédure, tout comme le fait d'utiliser un modèle alternatif à Chain Ladder, comme Bornhuetter-Ferguson, ce qui fait également son intérêt.

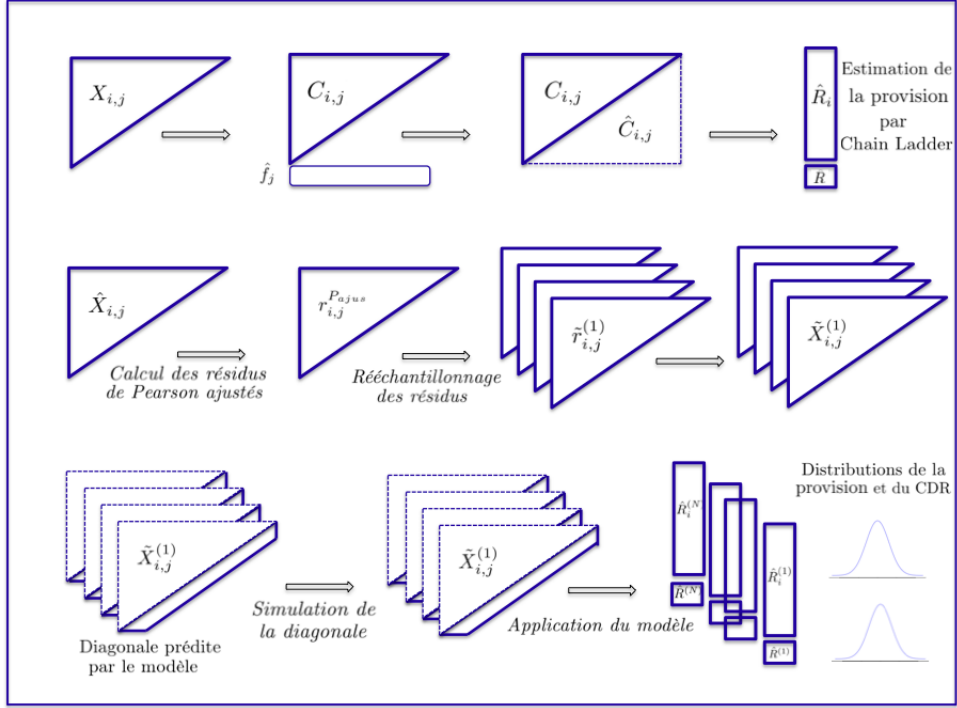


FIGURE 8 – Illustration de la procédure bootstrap de re-provisionnement à un an.

En revanche, les auteurs de l'article [4] soulèvent deux limites majeures de cette méthode déjà évoquées par Boisseau en 2010.

Tout d'abord, la prédiction de la diagonale par le modèle (avant la simulation) est faite à partir des paramètres MLG estimés sur le triangle supérieur. C'est pourquoi, lorsque les paramètres MLG sont ensuite réestimés par maximum de vraisemblance à partir du trapèze, les variables aléatoires du triangle supérieur et de la diagonale ne sont pas indépendantes. Ainsi, le cadre de l'estimation par maximum de vraisemblance n'est pas vérifié, l'utilisation d'un modèle MLG supposant l'indépendance des variables aléatoires.

D'autre part, cette méthode bootstrap estime les paramètres MLG à deux reprises : ceux-ci sont estimés une première fois pour obtenir les règlements incrémentaux de la diagonale, puis pour déterminer les règlements futurs attendus du triangle inférieur. Cette approche tend à augmenter de manière significative l'erreur d'estimation et, par conséquent, l'erreur de prédiction, par rapport aux résultats que l'on obtient par la méthode Merz-Wüthrich précédente.

7.1.2 La méthode bootstrap améliorée

Les limites de la méthode bootstrap de re-provisionnement précédente ont donc amené Boisseau à modifier la procédure dans l'article *One-year reserve risk including a tail factor: closed formula and bootstrap approaches*. L'idée principale de cette méthode bootstrap améliorée est un rééchantillonnage avec remise des résidus à partir du triangle supérieur afin d'obtenir un trapèze de résidus. La diagonale correspondant à l'année $(I, I + 1]$ n'est donc plus simulée par une loi gamma dans

cette méthode.

La procédure bootstrap est la suivante :

1. Les quatre premières étapes sont identiques à celles présentées dans la méthode bootstrap de re-provisionnement. On obtient une estimation des provisions $\hat{R}$ à la date $t = I$ par la méthode Chain Ladder.

2. Répéter N fois les étapes suivantes :

- Effectuer un rééchantillonnage aléatoire avec remise des $\frac{I(I+1)}{2}$ résidus de Pearson ajustés et obtenir un nouveau trapèze de $\frac{I(I+3)}{2} - 1$ résidus $\tilde{r}_{i,j}^{(k)}$ pour la k^{ieme} boucle bootstrap, et non plus seulement un triangle. La première diagonale est donc comprise dans ce trapèze.
- Reconstituer le pseudo-trapèze des règlements incrémentaux par la relation suivante :

$$\tilde{X}_{i,j}^{(k)} = \hat{X}_{i,j} + \tilde{r}_{i,j}^{(k)} \sqrt{\hat{X}_{i,j}}, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \leq J - i + 2.$$

On obtient alors une réalisation des règlements de la diagonale en utilisant directement les résidus et les règlements obtenus précédemment par Chain Ladder. L'erreur de processus est ainsi prise en compte à partir des résidus rééchantillonnés ; il en est de même pour l'erreur d'estimation. Il y a bien indépendance entre les variables aléatoires du trapèze supérieur.

- Calculer les facteurs de développement Chain Ladder à partir du nouveau pseudo-trapèze des règlements cumulés.
 - Estimer les diagonales suivantes grâce à ces facteurs de développement et déduire les règlements incrémentaux de ces diagonales.
 - Estimer les provisions par la méthode Chain Ladder et stocker le résultat.
3. Obtenir les distributions de l'estimation de la provision et du CDR et déterminer les erreurs de prédiction qui leur sont associées.

Cette méthode du bootstrap de re-provisionnement améliorée est illustrée par le schéma suivant.

7.1.3 La méthode récursive

De Felice et Moriconi proposent dans leur article *Process Error and Estimation Error of Year-End Reserve Estimation in the Distribution Free Chain-Ladder Model* (2006) une méthode bootstrap qui repose sur la version récursive du modèle Chain Ladder. Angoua et al. [4] s'appuient en 2011 sur cet article pour développer une méthode bootstrap que l'on nomme récursive. Elle permet d'obtenir une distribution du montant des provisions et présente l'avantage de répliquer les résultats de l'erreur de prédiction du CDR observable de Merz-Wüthrich [26]. Le rééchantillonnage de cette méthode bootstrap s'effectue sur les résidus des facteurs de développement individuels. Les étapes de cette procédure bootstrap sont détaillées ci-après.

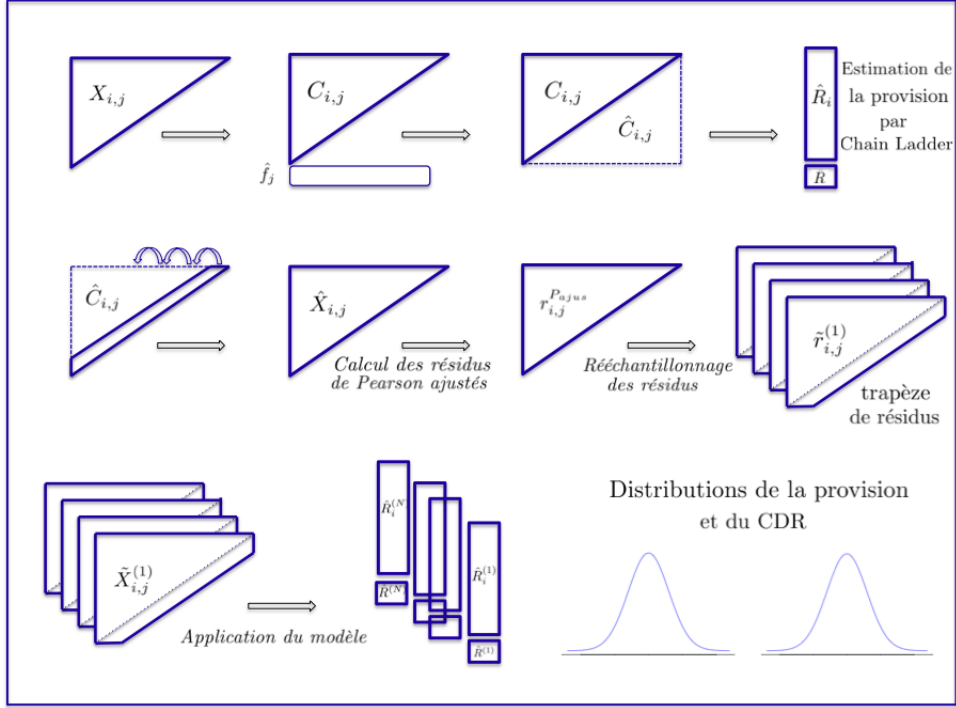


FIGURE 9 – Illustration de la procédure bootstrap de re-provisionnement améliorée à un an.

• **La procédure bootstrap**

1. Estimer les facteurs de développement $\hat{f}_j$, $j \in \{1, \dots, J-1\}$, et les variances de ces estimateurs $\hat{\sigma}_j$.
2. Déterminer le triangle supérieur des facteurs de développement individuels définis par :

$$F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}, \text{ pour } 2 \leq i+j \leq I-1.$$

3. Estimer la provision $\hat{R}$ par la méthode Chain Ladder.
4. Calculer les résidus des facteurs de développement individuels tels que :

$$r_{i,j} = \frac{\sqrt{C_{i,j}}(F_{i,j} - \hat{f}_j)}{\hat{\sigma}_j}, \text{ pour } 2 \leq i+j \leq I-1.$$

Pour corriger le biais dans la comparaison des estimations analytiques et bootstrap de l'erreur de prédiction, les résidus sont ensuite ajustés de la façon suivante :

$$r^{ajus}_{i,j} = \sqrt{\frac{I-j}{I-j-1}} \frac{\sqrt{C_{i,j}}(F_{i,j} - \hat{f}_j)}{\hat{\sigma}_j} \text{ pour } 2 \leq i+j \leq I.$$

5. Répéter N fois les étapes suivantes :

- Effectuer un rééchantillonnage aléatoire avec remise des résidus ajustés. On obtient un nouveau triangle de résidus $\tilde{r}_{i,j}^{(k)}$ pour la k^{ieme} boucle bootstrap.
- Reconstituer le pseudo-triangle des facteurs de développement individuels à la date $t = I$ par la relation suivante :

$$\tilde{F}_{i,j}^{k,I} = \tilde{r}_{i,j}^{(k)} \sqrt{\frac{\hat{\sigma}_j^2}{C_{i,j}}} + \hat{f}_j, \text{ pour } 2 \leq i + j \leq I - 1.$$

- Calculer les facteurs de développement Chain Ladder $\tilde{f}_j^{k,I}$ à la date $t = I$ à partir des nouveaux facteurs de développement individuels :

$$\tilde{f}_j^{k,I} = \frac{\sum_{i=1}^{I-j} C_{i,j} \tilde{F}_{i,j}^{k,I}}{\sum_{i=1}^{I-j} C_{i,j}}, \text{ pour } j \in \{1, \dots, I - 1\}.$$

- Simuler les règlements de l'année $(I, I + 1]$ correspondant à la diagonale suivante du triangle de liquidation initial. Pour cela, simuler les règlements cumulés $\tilde{C}_{i,J-i+2}^k$ par une loi normale de moyenne $C_{i,I-i+1} \tilde{f}_{I-i+1}^{k,I}$ et de variance $C_{i,I-i+1} \hat{\sigma}_{I-i+1}^2$. On déduit alors les règlements futurs pour l'année comptable $(I, I + 1]$:

$$\tilde{P}^{k,I+1} = \sum_{i=1}^I (\tilde{C}_{i,J-i+2}^k - C_{i,J-i+1}).$$

- Déterminer les facteurs de développement individuels $\tilde{F}_{i,j}^{k,I+1}$ en considérant la nouvelle diagonale simulée. On en déduit les nouveaux facteurs de développement $\tilde{f}_j^{k,I+1}$ à la fin de l'année $(I, I + 1]$ grâce à la relation suivante :

$$\tilde{f}_j^{k,I+1} = \frac{\sum_{i=1}^{I-j} (C_{i,j} F_{i,j}) + C_{I-j+1,j} F_{I-j+1,j}^{k,I+1}}{\sum_{i=1}^{I-j+1} C_{i,j}}, \text{ pour } j \in \{1, \dots, I - 1\}.$$

- Estimer la provision $\hat{R}^{(k)}$ par la méthode Chain Ladder.
- Calculer le CDR^k par la relation suivante :

$$CDR^k = \hat{R} - \tilde{P}^{k,I+1} - \hat{R}^{(k)}.$$

- Stocker les résultats.

6. Obtenir les distributions empiriques de l'estimation de la provision et du CDR.

• **L'équivalence avec les résultats analytiques de Merz et Wüthrich [26]**

Les auteurs ont montré que la procédure du bootstrap récursif permet d'obtenir la même erreur de prédiction que celle du CDR observable proposée par Merz et Wüthrich [26].

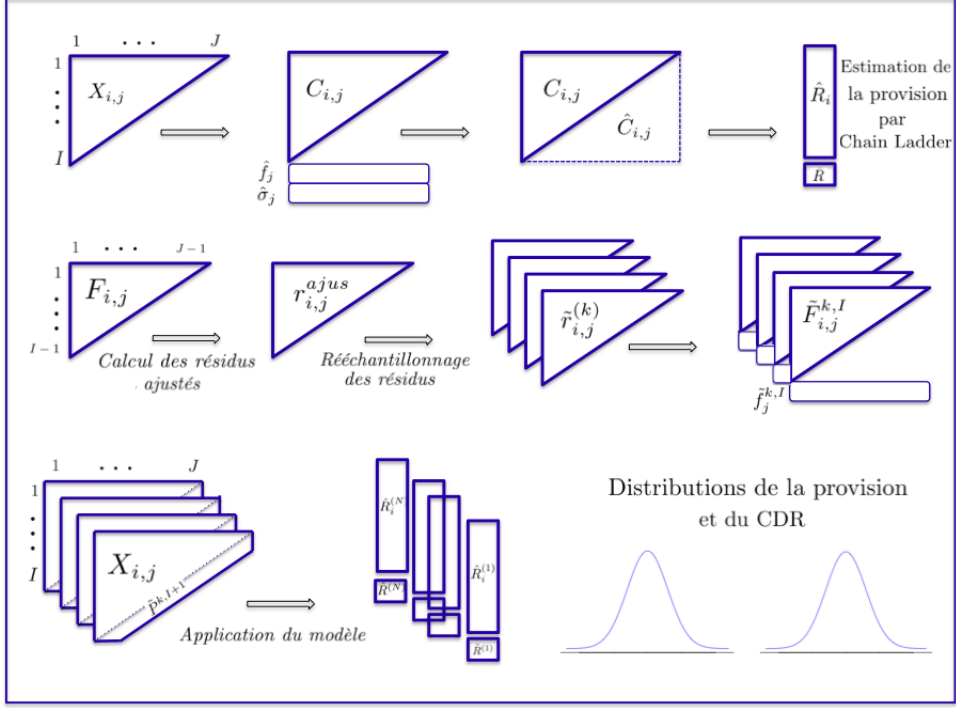


FIGURE 10 – Illustration de la procédure bootstrap récursive à un an.

Erreur d'estimation

Soit $\widehat{CDR}_i$ le CDR ne tenant compte que de l'erreur d'estimation. Pour $i \in \{2, \dots, I\}$, $C_{i,I-i+2}^k = \tilde{f}_{i,I-i+1}^{(k,I)} C_{i,I-i+1}$. La variance du CDR est approximée par :

$$\text{Var}(\widehat{CDR}_i) \approx \hat{C}_{i,I}^2 \left(\frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{i,I-i+1}^2 S_{I-i+1}^I} + \sum_{j=I-i+2}^{I-1} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}} \right)^2 \frac{\hat{\sigma}_j^2}{\hat{f}_j^2 S_j^I} \right).$$

Cette formule correspond à l'estimateur de l'erreur d'estimation proposé par Merz et Wüthrich pour l'année de survénance i .

Pour l'ensemble des années de survénance, l'erreur d'estimation s'écrit :

$$\text{Var} \left(\sum_{i=1}^I \widehat{CDR}_i \right) = \sum_{i=1}^I \text{Var}(\widehat{CDR}_i) + 2 \sum_{1 \leq i < j \leq I} \text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j),$$

avec

$$\text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j) = \mathbb{E}[\widehat{CDR}_i \widehat{CDR}_j] - \mathbb{E}[\widehat{CDR}_i] \mathbb{E}[\widehat{CDR}_j].$$

Or $\text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j) = \mathbb{E}[\widehat{CDR}_i \widehat{CDR}_j]$ et la covariance est approximativement égale à :

$$\mathbb{E}[\widehat{CDR}_i \widehat{CDR}_j] \approx C_{i,I} C_{j,I} \left(\frac{C_{i,I-i+1} \hat{\sigma}_{I-i+1}^2}{S_{I-i+1}^{I+1} S_{I-i+1}^I \hat{f}_{I-i+1}^2} + \sum_{m=I-i+2}^{I-1} \left(\frac{C_{I-m+1,m}}{S_m^{I+1}} \right)^2 \frac{\hat{\sigma}_m^2}{S_m^I \hat{f}_m^2} \right).$$

On retrouve la même covariance que celle proposée par Merz et Wüthrich et le même estimateur $(u - \text{bias})^2$ de l'erreur d'estimation agrégée.

Erreur de processus

Soit CDR_i le CDR ne prenant en compte que l'erreur de processus. L'erreur de processus est égale à :

$$\begin{aligned}\mathbb{V}ar(CDR_i) &= (C_{i,I-i+1}\hat{\sigma}_{I-i+1}^2 + C_{i,I-i+1}^2\hat{f}_{I-i+1}^2) \prod_{j=I-i+2}^{I-1} \left(\hat{f}_j^2 + \frac{\hat{\sigma}_j^2 C_{I-j+1,j}}{(S_j^{I+1})^2} \right) - \hat{C}_{i,I}^2 \\ &= \hat{C}_{i,I}^2 \left[\left(1 + \frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{I-i+1}^2 C_{i,I-i+1}} \right) \prod_{j=I-i+2}^{I-1} \left(1 + \frac{\hat{\sigma}_j^2 C_{I-j+1,j}}{\hat{f}_j^2 (S_j^{I+1})^2} \right) - 1 \right].\end{aligned}$$

On retrouve bien l'estimateur de l'erreur de processus proposé par Merz et Wüthrich pour l'année de survenance i .

Pour l'ensemble des années de survenance, l'erreur de processus s'écrit :

$$\mathbb{V}ar\left(\sum_{i=1}^I CDR_i\right) = \sum_{i=1}^I \mathbb{V}ar(CDR_i) + 2 \sum_{1 \leq i < j \leq I} \text{Cov}(CDR_i, CDR_j),$$

avec

$$\text{Cov}(CDR_i, CDR_j) = \mathbb{E}[CDR_i CDR_j] - \mathbb{E}[CDR_i] \mathbb{E}[CDR_j].$$

On peut écrire :

$$\begin{aligned}\mathbb{E}[CDR_i] &= \mathbb{E}\left[C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - C_{i,I-i+2}^{(k)} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)}\right] \\ &= C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - \underbrace{\mathbb{E}[C_{i,I-i+2}^{(k)}]}_{\hat{f}_{I-i+1} C_{i,I-i+1}} \prod_{j=I-i+2}^{I-1} \underbrace{\mathbb{E}[\tilde{f}_j^{(k,I+1)}]}_{\hat{f}_j} \\ &= 0.\end{aligned}$$

D'où $\text{Cov}(CDR_i, CDR_j) = \mathbb{E}[CDR_i CDR_j]$.

L'erreur de processus s'écrit finalement ¹¹ :

$$\mathbb{E}[CDR_i CDR_j] = \hat{C}_{i,I} \hat{C}_{j,I} \left[\left(1 + \frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{I-i+1}^2 S_{I-i+1}^{I+1}} \right) \prod_{m=I-i+1}^{I-1} \left(1 + \frac{\hat{\sigma}_m^2 C_{I-m+1,m}}{\hat{f}_m^2 (S_m^{I+1})^2} \right) - 1 \right].$$

Ce résultat correspond bien à celui de Merz et Wüthrich et permet d'obtenir l'erreur de processus agrégée.

Erreur de prédiction

L'erreur de prédiction s'obtient en sommant les erreurs d'estimation et de processus obtenues précédemment :

$$MSEP(CDR_i) = \mathbb{V}ar(CDR_i) + \mathbb{V}ar(\widehat{CDR_i}),$$

$$MSEP(CDR) = \mathbb{V}ar(CDR) + \mathbb{V}ar(\widehat{CDR}).$$

11. On pourra se reporter à l'article [4] pour un développement plus détaillé.

Contrairement aux méthodes bootstrap développées précédemment, cette méthode récursive permet de répliquer les résultats analytiques de Merz et Wüthrich [26], ce qui est un avantage certain. Mais cette méthode permet de surcroît d'obtenir une distribution complète du CDR et non pas uniquement les deux premiers moments comme le proposent Merz et Wüthrich. De plus, il est possible d'insérer un facteur de queue de développement dans cette méthode et ce point sera étudié par la suite.

7.2 L'insertion d'un facteur de queue de développement

Il est possible d'insérer un facteur de queue de développement dans les méthodes bootstrap introduites précédemment. On s'intéresse dans ce qui vient à la méthode de re-provisionnement dite « améliorée » et à la méthode récursive.

7.2.1 La méthode de re-provisionnement améliorée

On peut insérer un facteur de queue de développement $\hat{f}_{ult}$ dans la méthode bootstrap de re-provisionnement améliorée. Celui-ci est à nouveau déterminé par une régression sur les facteurs de développement et une volatilité lui est associée par la méthode Delta.

Il s'agit d'estimer un facteur de queue de développement $\hat{f}_{ult}^{(k)}$, $k \in \{1, \dots, N\}$, à chacun des N passages dans la boucle bootstrap. Dans le cadre MLG Poisson sur-dispersé, la logique est la même que pour Chain Ladder. Pour chacun des pseudo-trapèzes obtenus par rééchantillonnage, le facteur de queue de développement $\hat{f}_{ult}^{(k)}$ est déterminé par extrapolation.

7.2.2 La méthode récursive

On s'appuie sur l'article *One-year reserve risk including a tail factor: closed formula and bootstrap approaches* [4] pour insérer un facteur de développement dans la méthode récursive du bootstrap. Il est noté $\hat{f}_{ult}$ et la volatilité associée est $\hat{\sigma}_{ult}^2$.

1. Dans le calcul de la provision à la date $t = I$.

Le facteur de queue de développement est déterminé par extrapolation comme on a pu le faire précédemment. Il est ensuite multiplié à la charge ultime estimée par Chain Ladder, de sorte que :

$$\hat{C}_{i,ult} = \hat{C}_{i,J} \times \hat{f}_{ult}.$$

La provision à constituer pour l'ensemble des années de survivance à la date $t = I$ est ainsi :

$$\hat{R} = \sum_{i=1}^I \hat{C}_{i,ult} - C_{i,J-i+1}.$$

2. Dans la boucle bootstrap.

Une fois que les facteurs de développement $\tilde{f}_j^{(k,I+1)}$, $k \in \{1, \dots, N\}$, ont été réestimés à la date $t = I + 1$, le facteur de queue de développement $\hat{f}_{ult}^k$ est

simulé par une loi normale ou log-normale d'espérance $\hat{f}_{ult}$ et de variance $\hat{\sigma}_{ult}^2$ déterminées à la date $t = I$.

$\hat{f}_{ult}^{(k)}$ est l'estimateur par maximum de vraisemblance de $f_{ult} = \prod_{j \geq J} (1 + \exp(a \times j + b))$. Dans un cadre asymptotique, $\hat{f}_{ult}^{(k)}$ est gaussien. C'est pourquoi le facteur de queue de développement peut être modélisé par une loi normale d'espérance $\hat{f}_{ult}^{(k)}$ et de variance $\mathbb{V}ar(\hat{f}_{ult}^{(k)})$. La loi log-normale peut être utilisée si l'approche se veut plus prudente.

Finalement, la provision à constituer pour l'ensemble des années de survivance à la date $t = I + 1$ est ainsi :

$$\hat{R}^{(k)} = \sum_{i=1}^I \tilde{C}_{i,ult} - \tilde{C}_{i,J-i+2}.$$

7.3 La prise en compte de données exogènes

On peut à nouveau se demander s'il est possible de tenir compte de données exogènes au triangle de liquidation utilisé dans les méthodes précédentes. On se propose dans cette partie d'adapter la méthode du bootstrap de re-provisionnement dans l'approche Bornhuetter-Ferguson.

7.3.1 La méthode de re-provisionnement améliorée suivant l'approche Bornhuetter-Ferguson

Dans cette partie, on adapte la méthode du bootstrap de re-provisionnement pour pouvoir appliquer la méthode Bornhuetter-Ferguson et non plus Chain Ladder. La procédure est alors la suivante :

1. Estimer les facteurs de développement Chain Ladder $\hat{f}_j$, $j \in \{1, \dots, J - 1\}$.
2. Déterminer la charge de sinistre ultime $\hat{C}_{i,J}^{CL,I}$, $i \in \{1, \dots, I\}$, à la date $t = I$ par la méthode Chain Ladder.
3. Estimer la charge de sinistre ultime a priori $\hat{U}_i$ obtenue généralement par la relation suivante :

$$\hat{U}_i = \Phi_i \times P_i,$$

avec

$$\begin{cases} \Phi_i & \text{le ratio sinistres à primes attendu pour l'année de survivance } i, \\ P_i & \text{le montant de primes acquises relatif à cette même année.} \end{cases}$$

4. Déterminer la charge de sinistre ultime estimée $\hat{C}_{i,J}^{BF,I}$ par la méthode Bornhuetter-Ferguson et en déduire la provision $\hat{R}$ à constituer à la date $t = I$.

5. Répéter N fois les étapes suivantes :

- Effectuer un rééchantillonnage aléatoire avec remise des $\frac{I(I+1)}{2}$ résidus de Pearson ajustés et obtenir un nouveau trapèze de $\frac{I(I+3)}{2} - 1$ résidus $\tilde{r}_{i,j}^{(k)}$ pour la k^{ieme} boucle bootstrap.
- Reconstituer le pseudo-trapèze des règlements incrémentaux par la relation suivante :

$$\tilde{X}_{i,j}^{(k)} = \hat{X}_{i,j} + \tilde{r}_{i,j}^{(k)} \sqrt{\hat{X}_{i,j}}, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \leq J - i + 1.$$

On obtient alors une réalisation des règlements de la diagonale en prenant en compte l'erreur de processus à partir des résidus rééchantillonnés ; on prend ainsi en compte l'erreur d'estimation. Il y a bien indépendance entre les variables aléatoires du trapèze supérieur.

- Calculer les facteurs de développement Chain Ladder à partir du nouveau pseudo-trapèze des règlements cumulés.
 - Estimer les diagonales suivantes grâce à ces facteurs de développement et déduire la charge de sinistre ultime estimée Chain Ladder $\hat{C}_{i,J}^{CL,I+1}$.
 - Déterminer la charge de sinistre ultime estimée $\hat{C}_{i,J}^{BF,I+1}$ par la méthode Bornhuetter-Ferguson, la charge ultime a priori restant inchangée. En déduire la provision $\hat{R}^{(k)}$ à constituer à la date $t = I + 1$.
 - Stocker le résultat.
6. Obtenir les distributions de l'estimation de la provision et du CDR et déterminer les erreurs de prédiction qui leur sont associées.

Cette méthode du bootstrap de re-provisionnement améliorée dans l'approche Bornhuetter-Ferguson est illustrée par le schéma suivant.

7.3.2 La méthode de re-provisionnement améliorée suivant l'approche Benktander

On se propose dans cette partie d'adapter la procédure du bootstrap de re-provisionnement à la méthode Benktander ; on se place donc dans le prolongement de la méthode Bornhuetter-Ferguson. On développe ci-dessous la procédure bootstrap tout en s'appuyant sur la méthode de re-provisionnement améliorée suivant l'approche Bornhuetter-Ferguson.

1. Les trois premières étapes sont identiques à celles de la procédure bootstrap Bornhuetter-Ferguson précédente.
2. Estimer la charge de sinistre ultime $\hat{C}_{i,J}^{BF,I}$ par la méthode Bornhuetter-Ferguson à la date $t = I$.

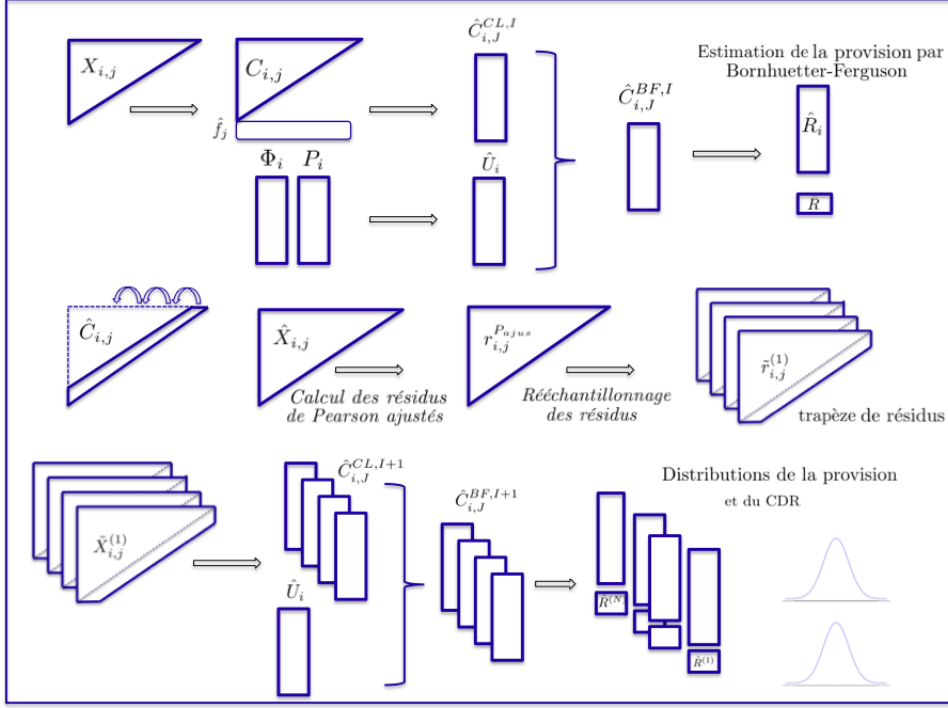


FIGURE 11 – Illustration de la procédure bootstrap de re-provisionnement par l’approche Bornhuetter-Ferguson à un an.

3. Estimer la charge de sinistre ultime $\hat{C}_{i,J}^{BE,I}$ par la méthode Benktander et en déduire la provision $\hat{R}$ à constituer à la date $t = I$.
4. Répéter N fois les étapes suivantes :

- Effectuer un rééchantillonnage aléatoire avec remise des $\frac{I(I+1)}{2}$ résidus de Pearson ajustés et obtenir un nouveau trapèze de $\frac{I(I+3)}{2} - 1$ résidus $\tilde{r}_{i,j}^{(k)}$ pour la k^{ieme} boucle bootstrap.
- Reconstituer le pseudo-trapèze des règlements incrémentaux par la relation suivante :

$$\tilde{X}_{i,j}^{(k)} = \hat{X}_{i,j} + \tilde{r}_{i,j}^{(k)} \sqrt{\hat{X}_{i,j}}, \text{ pour } i \in \{1, \dots, I\} \text{ et } j \leq J - i + 1.$$

On obtient alors une réalisation des règlements de la diagonale en prenant en compte l’erreur de processus à partir des résidus rééchantillonnés ; on prend ainsi en compte l’erreur d’estimation. Il y a bien indépendance entre les variables aléatoires du trapèze supérieur.

- Calculer les facteurs de développement Chain Ladder à partir du nouveau pseudo-trapèze des règlements cumulés.
- Estimer les diagonales suivantes grâce à ces facteurs de développement et déduire la charge de sinistre ultime estimée Chain Ladder $\hat{C}_{i,J}^{CL,I+1}$.

- Déterminer la charge de sinistre ultime estimée $\hat{C}_{i,J}^{BF,I+1}$ par la méthode Bornhuetter-Ferguson, la charge ultime a priori restant inchangée.
 - Estimer la charge de sinistre ultime $\hat{C}_{i,J}^{BE,I+1}$ par la méthode Benktander et en déduire la provision $\hat{R}^{(k)}$ à constituer à la date $t = I + 1$.
 - Stocker le résultat.
5. Obtenir les distributions de l'estimation de la provision et du CDR et déterminer les erreurs de prédiction qui leur sont associées.

Le schéma suivant illustre cette procédure.

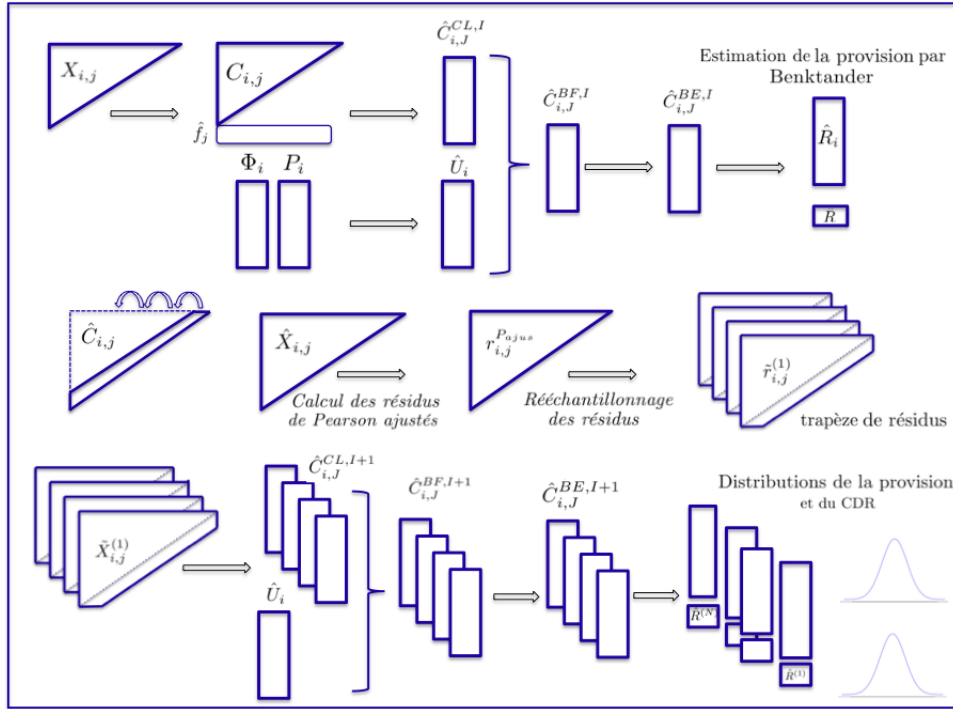


FIGURE 12 – Illustration de la procédure bootstrap de re-provisionnement par l'approche Benktander à un an.

Deux méthodes de bootstrap à un an ont été développées dans le cadre de Bornhuetter-Ferguson et de la méthode Benktander, permettant ainsi de déterminer les distributions de la provision et du CDR. Ces méthodes permettent de répondre aux exigences de Solvabilité II tout en utilisant des informations exogènes au triangle de liquidation utilisé dans toutes les méthodes.

L'incertitude à un an que l'on a calculée grâce à l'une des méthodes développées dans cette partie permet de déduire l'exigence de capital réglementaire pour le risque non-vie.

8 La détermination du capital réglementaire

On présente dans cette partie deux méthodes de détermination de l'exigence de capital réglementaire, ou *Solvency Capital Requirement* (SCR), pour le risque non-vie selon la méthode utilisée pour mesurer l'incertitude à horizon un an.

8.1 La détermination du SCR à partir d'une formule fermée

La première méthode pour déterminer le SCR repose sur la mesure de l'incertitude à horizon un an par une formule fermée, en l'occurrence, dans ce mémoire, celle proposée par Merz et Wüthrich. Le capital réglementaire est alors également calculé à partir d'une formule fermée ; pour une certaine branche d'activité, ou *Line of Business* (LoB), on a la relation suivante :

$$SCR_{LoB} = \rho(\sigma_{LoB}) \times V_{LoB},$$

avec

- $\rho(\sigma_{LoB})$ une fonction de la volatilité et σ_{LoB} une mesure de volatilité pour le risque de provisionnement à un an estimée par :

$$\hat{\sigma}_{LoB} = \frac{\sqrt{MSEP_{\text{Merz Wüthrich}}}}{\text{Best Estimate}},$$

- et V_{LoB} une mesure de volume correspondant au Best Estimate.

La fonction de volatilité $\rho(\sigma_{LoB})$ est définie en supposant une distribution des risques selon une loi log-normale :

$$\rho(\sigma_{LoB}) = \frac{e^{VaR_{99,5\%}(\Phi(0,1))\sqrt{\log(\sigma_{LoB}^2+1)}}}{\sqrt{\sigma_{LoB}^2+1}} - 1,$$

où $\Phi(0,1)$ désigne la fonction de répartition de la loi normale centrée réduite.

8.2 La détermination du SCR par une approche simulateur

La seconde méthode proposée pour déterminer le SCR à partir de simulations consiste à obtenir l'opposé du quantile à 0,5 % de la distribution du CDR. En effet, cette méthode permet d'estimer le montant de provisions requis afin de ramener la probabilité d'une insuffisance des provisions à horizon un an à 0,5 %.

Ce capital réglementaire est donc obtenu par :

$$SCR = -VaR_{0,5\%}(CDR).$$

Deux approches sont possibles pour mesurer l'incertitude à un an du risque de provisionnement. La formule fermée de Merz et Wüthrich qui a tout d'abord été développée est sur le point de devenir une référence dans la détermination de cette incertitude ; elle est en effet relativement simple à implémenter et elle s'inscrit dans l'approche Chain Ladder du provisionnement qui est utilisée dans la plupart des compagnies d'assurance à l'heure actuelle. Plusieurs méthodes bootstrap ont ensuite été explicitées dans le but de trouver une alternative simulateur au modèle de Merz et Wüthrich. La méthode de re-provisionnement paraît la plus intuitive mais la méthode dite récursive a le principal atout de pouvoir répliquer les résultats de Merz et Wüthrich tout en bénéficiant des avantages de la procédure bootstrap.

Enfin, les auteurs de l'article [4] ont proposé une formule fermée de la volatilité à un an de la variation calendaire de la charge de sinistre ultime estimée, la particularité étant de tenir compte d'un facteur de queue de développement. D'un point de vue simulateur, des modifications ont été apportées aux méthodes bootstrap évoquées ci-dessus : il s'agissait, d'une part, d'insérer un facteur de queue de développement et, d'autre part, de prendre en compte des données exogènes au triangle par l'approche Bornhuetter-Ferguson.

Dans la partie suivante, on se propose d'appliquer les méthodes de provisionnement à l'ultime et à un an qui ont été développées dans ce mémoire à des données en assurance non-vie, le but étant d'obtenir le SCR.

Quatrième partie

La détermination de l'incertitude sur les provisions pour des données en assurance non-vie

L'objectif de cette partie est d'obtenir à la fois la provision pour sinistres à payer sur la branche d'assurance non-vie considérée pour deux compagnies d'assurance, l'incertitude à l'ultime de cette provision et l'incertitude à horizon un an. Les données dont on dispose sont tout d'abord présentées ; on s'intéresse ensuite à la volatilité à l'ultime des provisions grâce aux modèles stochastiques qui ont été introduits dans la deuxième partie de ce mémoire. Enfin, l'incertitude à un an du risque de provisionnement est déterminée afin d'obtenir le SCR non-vie exigé par la Directive Solvabilité II.

Les logiciels utilisés pour cette étude sont R et Excel.

9 Présentations des données

On dispose de triangles de liquidation relatifs à la branche « Accidents du travail » obtenus à partir des états réglementaires de deux compagnies d'assurance non-vie européennes que l'on nomme A et B. Ces triangles concernent les années de survenance 2001 à 2010 et sont bruts de réassurance. Cette branche peut être considérée comme à développement mi-long.

9.1 Les triangles de liquidation

• Compagnie A

On présente tout d'abord les triangles des règlements incrémentaux et cumulés de la compagnie A pour cette branche.

	1	2	3	4	5	6	7	8	9	10
2001	3 803	9 442	4 555	3 984	3 666	3 686	1 359	942	1 283	925
2002	4 610	14 777	8 823	8 056	9 646	5 624	3 282	2 959	2 717	
2003	7 684	21 888	15 792	26 465	14 152	10 355	6 177	6 500		
2004	4 684	12 626	22 189	15 887	10 879	6 621	6 707			
2005	3 952	18 565	15 033	11 770	7 601	6 376				
2006	5 926	19 565	18 352	10 904	11 114					
2007	5 671	21 166	16 645	11 085						
2008	6 367	26 589	15 897							
2009	6 293	17 959								
2010	6 439									

TABLE 1 – Triangle des règlements incrémentaux.

	1	2	3	4	5	6	7	8	9	10
2001	3 803	13 245	17 800	21 784	25 450	29 136	30 495	31 437	32 720	33 645
2002	4 610	19 387	28 210	36 266	45 912	51 536	54 818	57 777	60 494	
2003	7 684	29 572	45 364	71 829	85 981	96 336	102 513	109 013		
2004	4 684	17 310	39 499	55 386	66 265	72 886	79 593			
2005	3 952	22 517	37 550	49 320	56 921	63 297				
2006	5 926	25 491	43 843	54 747	65 861					
2007	5 671	26 837	43 482	54 567						
2008	6 367	32 956	48 853							
2009	6 293	24 252								
2010	6 439									

TABLE 2 – Triangle des règlements cumulés.

On remarque que pour les années de survenance 2003 et ultérieures, les montants des règlements ont augmenté par rapport à ceux de 2001 et 2002. Cela s'explique par une mutualisation de deux portefeuilles. Pour les années de survenance 2006 à 2010, on constate une certaine homogénéité des règlements.

Le tableau suivant présente les facteurs de développements $\hat{f}_j$ associées à ce dernier triangle, ainsi que les cadences de règlement cumulées correspondantes $\hat{p}_j$ pour $j = 1, \dots, 9$.

j	1	2	3	4	5	6	7	8	9
$\hat{f}_j$	4,32	1,63	1,34	1,20	1,12	1,07	1,06	1,04	1,03
$\hat{p}_j$	6,5%	28,2%	45,9%	61,7%	73,8%	82,4%	88,2%	93,1%	97,3%

TABLE 3 – Facteurs de développement et cadences de règlements

Au terme de dix années de développement, seuls 97,3 % du montant des sinistres ont été réglés. Ainsi, l'insertion d'un facteur de queue de développement paraît nécessaire pour obtenir la liquidation totale des sinistres. On détermine un tel facteur de queue de développement $\hat{f}_{ult}$ par régression linéaire sur les $\log(\hat{f}_j - 1)$.

Il s'agit tout d'abord de déterminer $\hat{a}$ et $\hat{b}$ tels que $\log(\hat{f}_j - 1) = \hat{a} + \hat{b} \times j$.

La volatilité estimée $\hat{\sigma}_{ult}$ associée à ce facteur de queue de développement est obtenue par la méthode Delta¹². Les résultats sont présentés ci-dessous.

$\hat{f}_{ult}$	1,028699
$\hat{\sigma}_{ult}$	$1,604428 e^{-06}$
R^2	91,7 %
$\hat{a}$	0,82416
$\hat{b}$	-0,52771

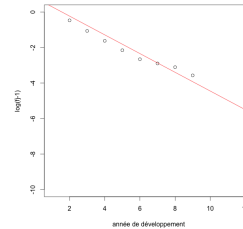


FIGURE 13 – Résultats et droite de régression $y = \hat{a} + \hat{b} \times j$ obtenus.

12. cf Annexe.

Ainsi, prendre en considération $\hat{f}_{ult}$ dans le calcul des provisions équivaut à augmenter la charge ultime de 2,87 %.

D'autre part, le coefficient de détermination R^2 permet de juger la qualité de l'ajustement du modèle de régression. Si on considère une variable réponse Y à expliquer à l'aide de p variables explicatives et pour v observations indépendantes de $Y_1, \dots, Y_v$ de Y , le coefficient de détermination est défini par :

$$R^2 = \frac{\sum_{i=1}^v (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^v (Y_i - \bar{Y})^2}.$$

Le R^2 est ici supérieur à 90 %, ce qui signifie que l'ajustement est très satisfaisant.

• Compagnie B

On présente ci-dessous les triangles des règlements incrémentaux et cumulés de la compagnie B pour cette même branche.

	1	2	3	4	5	6	7	8	9	10
2001	6 438	22 694	14 556	12 624	4 126	3 806	3 489	2 663	1 262	558
2002	6 112	15 951	15 876	7 602	5 540	3 180	2 172	571	335	
2003	4 526	11 940	11 887	13 162	4 963	1 854	2 740	713		
2004	3 225	12 336	15 398	9 303	5 241	3 117	503			
2005	3 424	17 083	14 844	6 452	2 815	2 747				
2006	8 714	21 268	14 050	6 772	3 187					
2007	8 855	23 735	17 580	7 832						
2008	7 697	22 656	17 658							
2009	6 688	18 621								
2010	6 414									

TABLE 4 – Triangle des règlements incrémentaux.

	1	2	3	4	5	6	7	8	9	10
2001	6 438	29 132	43 688	56 312	60 438	64 244	67 733	70 396	71 658	72 216
2002	6 112	22 063	37 939	45 541	51 081	54 261	56 433	57 004	57 339	
2003	4 526	16 466	28 353	41 515	46 478	48 332	51 072	51 785		
2004	3 225	15 561	30 959	40 262	45 503	48 620	49 123			
2005	3 424	20 507	35 351	41 803	44 618	47 365				
2006	8 714	29 982	44 032	50 804	53 991					
2007	8 855	32 590	50 170	58 002						
2008	7 697	30 353	48 011							
2009	6 688	25 309								
2010	6 414									

TABLE 5 – Triangle des règlements cumulés.

Le tableau suivant présente les facteurs de développements $\hat{f}_j$ associées à ce dernier triangle, ainsi que les cadences de règlement cumulées correspondantes $\hat{p}_j$ pour $j = 1, \dots, 9$.

j	1	2	3	4	5	6	7	8	9
$\hat{f}_j$	3,99	1,62	1,24	1,09	1,06	1,04	1,02	1,01	1,01
$\hat{p}_j$	10,0%	39,7%	64,3%	79,4%	86,9%	92,0%	95,8%	98,0%	99,2%

TABLE 6 – Facteurs de développement et cadences de règlements

Les cadences de règlement montrent que les sinistres semblent liquidés au terme de dix années de développement ; c’est pourquoi un facteur de queue de développement ne paraît pas nécessaire dans le cas de la compagnie B.

9.2 Les données exogènes aux triangles

Les primes et les ratios sinistres à primes (S/P) sont présentés ci-dessous pour chacune des compagnies et par année de survenance. Il s’agit des primes acquises, c’est-à-dire les primes correspondant à la période qui s’écoule entre la date d’échéance de la prime payable d’avance et la date d’inventaire.

Compagnie A			Compagnie B		
Année	Primes	S/P	Année	Primes	S/P
2001	45 867	83%	2001	65 719	115%
2002	50 643	108%	2002	66 465	93%
2003	122 016	92%	2003	90 782	64%
2004	126 111	95%	2004	79 032	74%
2005	134 468	84%	2005	92 485	73%
2006	142 176	89%	2006	109 691	73%
2007	147 336	87%	2007	127 652	79%
2008	158 835	86%	2008	146 614	82%
2009	159 138	75%	2009	154 447	78%
2010	162 982	62%	2010	172 349	73%

TABLE 7 – Primes et ratios sinistres à primes pour les deux compagnies et par année de survenance.

On remarque une rupture dans le volume des primes de la compagnie A entre les années de survenance 2002 et 2003 ; celles-ci ont en effet doublé de volume d’une année sur l’autre en raison de l’acquisition d’un nouveau portefeuille. Les ratios sinistres à primes de la compagnie A sur cette branche sont bons puisqu’ils sont majoritairement inférieurs à 90%. Toutefois, ceux de la compagnie B sont meilleurs dans la mesure où ils sont très souvent inférieurs à 80 %.

10 La mesure de l’incertitude à l’ultime de la provision pour sinistres à payer

On cherche dans cette partie à déterminer la provision pour sinistres à payer (PSAP) des deux compagnies pour cette branche. Il s’agit ensuite de mesurer la volatilité de cette estimation des provisions par des méthodes stochastiques analytiques et simulateurs.

10.1 L'approche Chain Ladder de la provision estimée

On s'intéresse tout d'abord aux méthodes analytiques Chain Ladder et Mack, puis on estimera la provision et l'incertitude à l'ultime par les méthodes simulatoires du bootstrap.

10.1.1 Les méthodes analytiques Chain Ladder et Mack

Hypothèses du modèle

- ✓ Les graphiques ci-après représentent les facteurs de développement par année de survenance pour les deux compagnies A et B. On cherche ainsi à montrer l'indépendance des années de survenance.

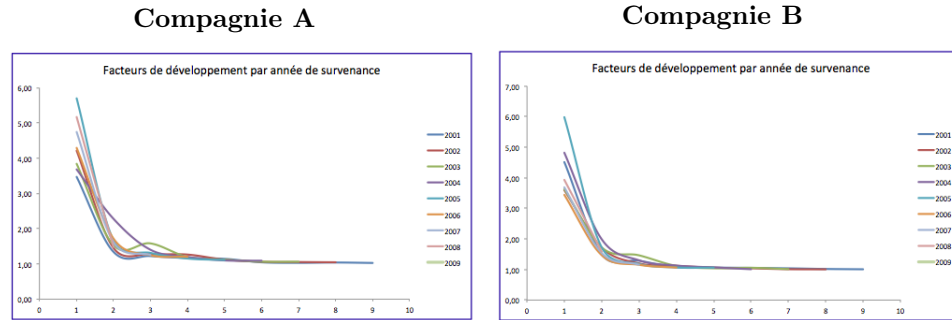


FIGURE 14 – Facteurs de développement par année de survenance pour les compagnies A et B sur la branche Accidents du travail.

Les facteurs de développement sont sensiblement superposés, une nuance devant toutefois être apportée concernant la première année de développement; en effet, pour les deux compagnies, les facteurs de développement relatifs à cette année de développement sont assez dispersés; l'hypothèse d'égalité des facteurs ne paraît alors pas réaliste. A noter également une légère différence pour les années de survenance les plus anciennes; cela peut être lié au changement dans le processus de gestion de la compagnie. On suppose que les facteurs sont quasiment identiques quelle que soit l'année de survenance et, par conséquent, que l'hypothèse permettant l'application de Chain Ladder est vérifiée.

- ✓ Les graphiques suivants représentent l'alignement des couples $(C_{i,j}, C_{i,j+1})$ pour un j fixé et $i \in \{1, \dots, I - j\}$ pour les compagnies A et B. On souhaite montrer l'existence des facteurs de développement.

Pour que l'hypothèse Chain Ladder soit vérifiée, il faut que les points soient alignés sur une droite passant par l'origine. Pour la première année de développement, en principe, on ne semble pas pouvoir valider cette propriété. En revanche, à partir de l'année suivante, on peut considérer que l'hypothèse est relativement respectée.

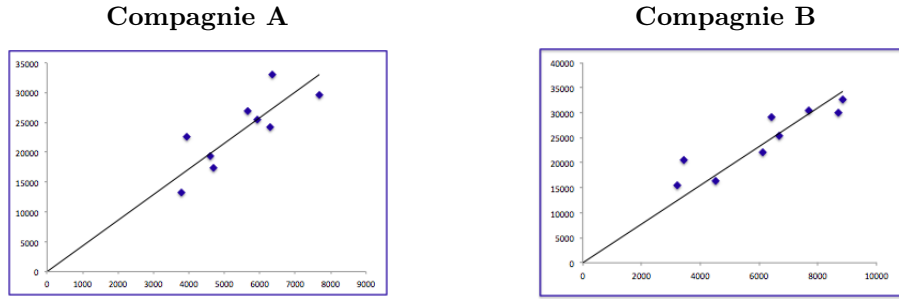


FIGURE 15 – Alignement des couples $(C_{i,1}, C_{i,2})_{i=1,\dots,I-1}$.

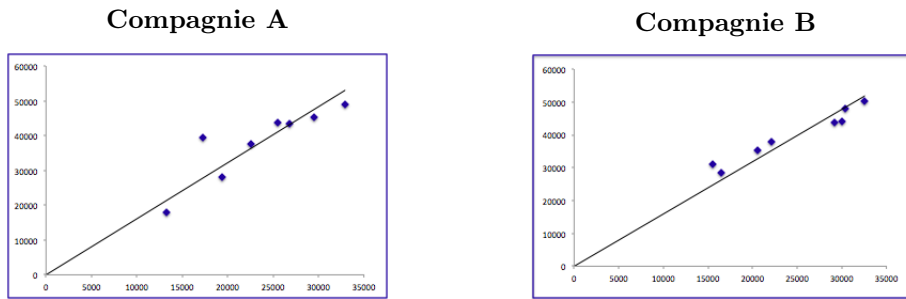


FIGURE 16 – Alignement des couples $(C_{i,2}, C_{i,3})_{i=1,\dots,I-2}$.

✓ Enfin, on présente les graphes des résidus égaux à $\frac{C_{i,j+1} - \hat{f}_j C_{i,j}}{\sqrt{C_{i,j}}}$ pour un j fixé et $i = 1 \dots, I - j$ en fonction des années de survenance. Les résidus ne doivent présenter aucune structure non aléatoire.

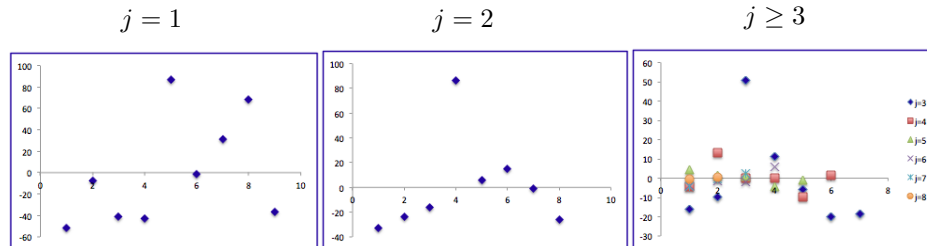


FIGURE 17 – Représentations des résidus en fonction des années de survenance pour $j = 1$, $j = 2$ et $j \geq 3$ pour la compagnie A.

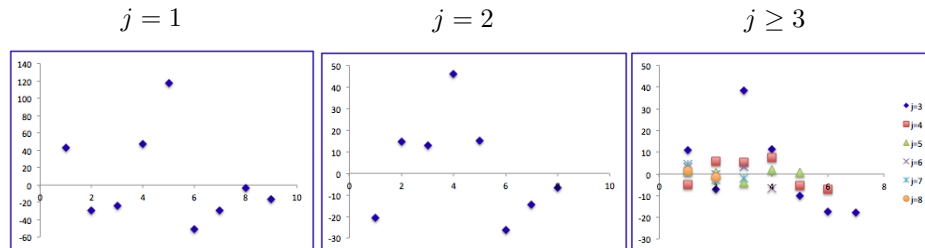


FIGURE 18 – Représentations des résidus en fonction des années de survenance pour $j = 1$, $j = 2$ et $j \geq 3$ pour la compagnie B.

Aucune tendance évidente ne semblent apparaître dans ces graphiques. On peut néanmoins remarquer un point atypique dans chacune des compagnies pour les trois premières années de développement. Pour remédier à ce problème, certains facteurs de développement individuels pourraient être modifiés afin que la condition aléatoire soit véritablement respectée.

La détermination de la PSAP et de l'incertitude à l'ultime associée

Les triangles de liquidation peuvent ainsi être complétés par la méthode Chain Ladder. La provision pour sinistres à payer (PSAP) et l'incertitude à l'ultime $\hat{\sigma}$ associée sont alors déterminées par exercice de survenance et au global par les méthodes Chain Ladder et Mack. On rappelle que la PSAP déterminée par la méthode de Mack est identique à la PSAP Chain Ladder. La méthode de Mack permet ensuite d'en mesurer l'incertitude à l'ultime. On présente dans les tableaux suivants les estimations obtenues dans le cas standard, puis en tenant compte du facteur de queue de développement $\hat{f}_{ult}$ dans le cas de la compagnie A.

• Compagnie A

	1	2	3	4	5	6	7	8	9	10
2001	3 803	13 245	17 800	21 784	25 450	29 136	30 495	31 437	32 720	33 645
2002	4 610	19 387	28 210	36 266	45 912	51 536	54 818	57 777	60 494	62 204
2003	7 684	29 572	45 364	71 829	85 981	96 336	102 513	109 013	113 901	117 121
2004	4 684	17 310	39 499	55 386	66 265	72 886	79 593	84 001	87 767	90 248
2005	3 952	22 517	37 550	49 320	56 921	63 297	67 736	71 487	74 692	76 804
2006	5 926	25 491	43 843	54 747	65 861	73 529	78 686	83 043	86 766	89 219
2007	5 671	26 837	43 482	54 567	65 328	72 934	78 049	82 371	86 064	88 497
2008	6 367	32 956	48 853	65 692	78 646	87 803	93 961	99 164	103 610	106 539
2009	6 293	24 252	39 437	53 030	63 488	70 880	75 851	80 051	83 641	86 005
2010	6 439	27 807	45 219	60 805	72 796	81 271	86 971	91 787	95 902	98 613

TABLE 8 – Triangle des règlements cumulés complété par Chain Ladder.

	PSAP estimée	Erreur de prédiction	Normal IC 95 %		Log-normal IC 95 %	
2001	-	-	-	-	-	-
2002	1 710	92	1 529	1 891	1 698	1 718
2003	8 108	476	7 175	9 041	8 039	8 150
2004	10 655	1 337	8 035	13 275	10 247	10 908
2005	13 507	1 837	9 907	17 107	12 902	13 883
2006	23 358	2 313	18 825	27 891	22 795	23 702
2007	33 930	3 469	27 130	40 730	33 059	34 464
2008	57 686	10 661	36 790	78 582	53 040	60 667
2009	61 753	16 725	28 972	94 534	51 737	68 671
2010	92 174	23 755	45 615	138 733	78 485	101 508
Total	302 882	34 220	235 811	369 954	293 427	308 699

TABLE 9 – PSAP, incertitude à l'ultime et intervalles de confiance par Mack sans facteur de queue de développement.

	PSAP estimée	Erreur de prédiction	Normal IC 95 %		Log-normal IC 95 %	
2001	966	143	686	1 246	915	998
2002	3 495	233	3 038	3 952	3 457	3 518
2003	11 469	594	10 305	12 633	11 392	11 515
2004	13 245	1 403	10 495	15 995	12 881	13 469
2005	15 711	1 906	11 975	19 447	15 148	16 059
2006	25 919	2 395	21 225	30 613	25 374	26 252
2007	36 740	3 579	29 455	43 485	35 607	36 998
2008	60 744	10 971	39 241	82 247	56 060	63 740
2009	64 221	17 207	30 495	97 947	54 001	71 259
2010	95 004	24 438	47 106	142 903	80 943	104 588
Total	327 244	35 251	258 152	396 336	317 940	332 956

TABLE 10 – PSAP, incertitude à l’ultime et intervalles de confiance par Mack avec le facteur de queue de développement $\hat{f}_{ult}$.

L’erreur de prédiction correspond à l’incertitude à l’ultime des provisions constituées, c’est-à-dire jusqu’à la liquidation totale des sinistres. Logiquement, insérer un facteur de queue de développement augmente l’estimation de la provision puisque la charge ultime calculée par Chain Ladder a été accrue de 2,87 %. L’incertitude est elle aussi en hausse avec un tel facteur.

L’erreur de prédiction permet de déterminer des intervalles de confiance, et ceux à 95 % ont été présentés dans ces tableaux, selon que l’on se place sous l’hypothèse de normalité ou de log-normalité de la distribution des provisions. On constate que l’intervalle de confiance sous l’hypothèse de log-normalité est bien plus précis que le premier qui, d’ailleurs, fait une forte hypothèse sur la symétrie de la distribution des provisions. L’intervalle de confiance log-normal semble donc à privilégier ici.

• Compagnie B

	1	2	3	4	5	6	7	8	9	10
2001	6 438	29 132	43 688	56 312	60 438	64 244	67 733	70 396	71 658	72 216
2002	6 112	22 063	37 939	45 541	51 081	54 261	56 433	57 004	57 339	57 785
2003	4 526	16 466	28 353	41 515	46 478	48 332	51 072	51 785	52 434	52 842
2004	3 225	15 561	30 959	40 262	45 503	48 620	49 123	50 229	50 859	51 255
2005	3 424	20 507	35 351	41 803	44 618	47 365	49 322	50 433	51 066	51 463
2006	8 714	29 982	44 032	50 804	53 991	57 191	59 554	60 895	61 659	62 139
2007	8 855	32 590	50 170	58 002	63 434	67 194	69 971	71 546	72 443	73 007
2008	7 697	30 353	48 011	59 326	64 882	68 727	71 567	73 179	74 097	74 674
2009	6 688	25 309	40 991	50 651	55 395	58 678	61 103	62 479	63 262	63 755
2010	6 414	25 569	41 412	51 172	55 965	59 281	61 731	63 121	63 913	64 410

TABLE 11 – Triangle des règlements cumulés complété par Chain Ladder.

L’intervalle de confiance sous l’hypothèse de normalité des provisions conduit à des bornes inférieures négatives. Il est donc préférable de se reporter à l’intervalle de confiance sous l’hypothèse de log-normalité des provisions.

	PSAP estimée	Erreur de prédiction	Normal IC 95 %		Log-normal IC 95 %	
2001	-	-	-	-	-	-
2002	446	370	-279	1 171	120	978
2003	1 057	677	-270	2 384	448	1 770
2004	2 132	1 214	-247	4 511	1 057	3 249
2005	4 098	1 711	744	7 452	2 742	5 215
2006	8 148	2 034	4 161	12 135	7 005	8 921
2007	15 005	3 020	9 086	20 924	13 587	15 926
2008	26 663	6 828	13 280	40 046	22 748	29 328
2009	38 446	8 906	20 990	55 902	33 736	41 582
2010	57 996	14 516	29 545	86 447	49 823	63 530
Total	153 993	20 981	112 869	195 117	147 070	158 299

TABLE 12 – PSAP, incertitude à l’ultime et intervalles de confiance par Mack sans facteur de queue de développement.

10.1.2 Les méthodes par simulations bootstrap

La provision pour sinistres à payer peut être déterminée à partir de méthodes par simulations bootstrap qui permettent également d’obtenir l’incertitude à l’ultime. Les procédures bootstrap utilisées sont les suivantes :

- la méthode bootstrap Poisson (1) obtenue avec la fonction *BootChainLadder* du package *ChainLadder* ou avec le code fourni par Lacoume [17] sous le logiciel R,
- la méthode bootstrap Poisson (2) obtenue à partir du code R fourni par Lacoume [17] et modifiée pour permettre l’insertion du facteur de queue de développement $\hat{f}_{ult}$.

50 000 simulations ont été effectuées.

Hypothèses du modèle

Pour pouvoir appliquer le rééchantillonnage des résidus de Pearson ajustés égaux à $r_{i,j}^{P_{ajus}} = \sqrt{\frac{n}{n-p}} \frac{X_{i,j} - \hat{X}_{i,j}}{\sqrt{\hat{X}_{i,j}}}$, il faut, en premier lieu, s’assurer que ceux-ci sont indépendants et identiquement distribués.

- ✓ Pour cela, on se propose de représenter dans un premier temps les résidus ajustés en fonction des résidus empiriques d’une loi normale centrée réduite grâce à un QQ-plot¹³.

13. On pourra se référer à l’annexe pour plus de précision.

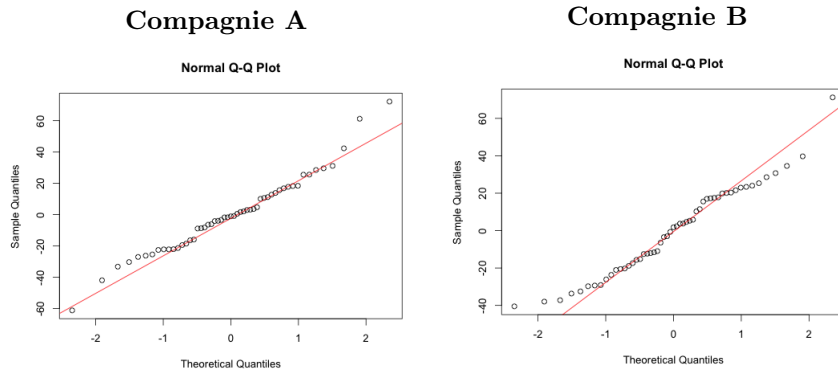


FIGURE 19 – QQ-plots des résidus pour les compagnies A et B.

Plus les résidus sont proches de la droite dite de Henry, plus l'hypothèse que les résidus suivent une loi normale est vérifiée. On constate que les résidus de la compagnie A sont relativement alignés le long de cette droite, ceux de la compagnie B en étant assez proches. On peut penser que l'hypothèse est vérifiée en respectant une certaine prudence. D'ailleurs, il convient de rester vigilant en ce qui concerne les queues de distribution. Pour confirmer cette interprétation, on effectue le test statistique de normalité de Shapiro-Wilk¹⁴.

	statistique du test	p-value
Compagnie A	0,9742	0,3035
Compagnie B	0,9674	0,1562

TABLE 13 – Résultats du test de Shapiro-Wilk.

On observe des p-values supérieures à 5 % pour les deux compagnies ; on ne rejette donc pas l'hypothèse de normalité des résidus.

- ✓ Les graphiques ci-dessous représentent les résidus par années de survénance, par années de développement et par rapport aux valeurs prévues par le modèle.

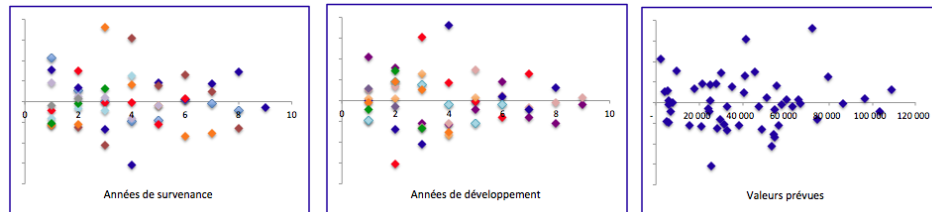


FIGURE 20 – Représentations des résidus en fonction des années de survénance, des années de développement et des valeurs prévues pour la compagnie A.

14. On trouvera en annexe une description de ce test.

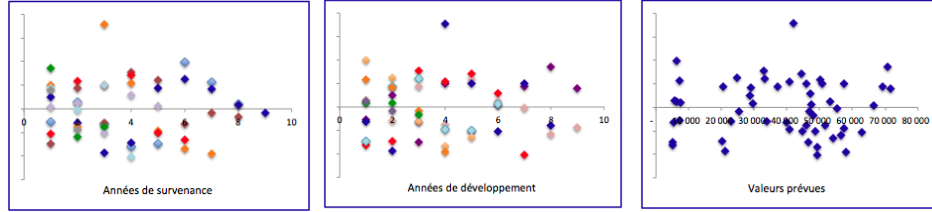


FIGURE 21 – Représentations des résidus en fonction des années de survénance, des années de développement et des valeurs prévues pour la compagnie B.

Aucun de ces graphiques ne fait apparaître de tendance non aléatoire ; les résidus semblent répartis de manière non structurée. On peut donc dire que la propriété d’espérance nulle et d’homoscédasticité des résidus est respectée.

- ✓ On présente désormais deux ACF représentant la fonction d’autocorrélation empirique des résidus. On souhaite ainsi montrer l’indépendance des résidus.

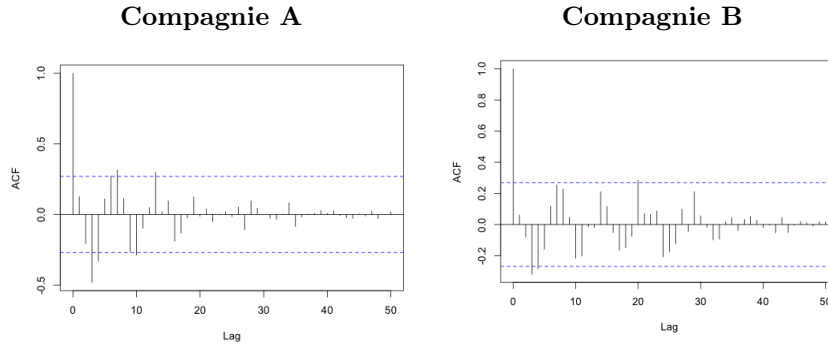


FIGURE 22 – ACF des résidus pour les compagnies A et B.

Concernant le premier graphique, on peut constater des covariances non nulles entre certains résidus, mais leur nombre reste faible. La plupart des covariances tracées sur ces graphiques sont contenues dans l’intervalle en pointillé bleu, soit l’intervalle de confiance à 95 %. Par définition de la fonction d’autocorrélation, la première valeur est toujours égale à l’unité. On ne rejète pas l’hypothèse d’indépendance des résidus, bien qu’il faille rester prudent quant aux résultats obtenus par la suite.

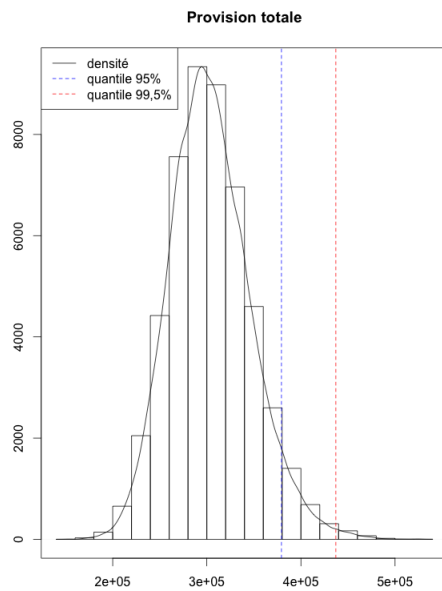
La détermination de la PSAP et de l’incertitude à l’ultime associée

Pour chacune des compagnies A et B, on présente ci-après l’estimation de la provision pour sinistres à payer et l’erreur de prédiction qui lui est associée.

- **Compagnie A**

	bootstrap Poisson (1)		bootstrap Poisson (2) avec $\hat{f}_{ult}$	
	PSAP estimée	Erreur de prédiction	PSAP estimée	Erreur de prédiction
2001	-	-	923	439
2002	1 740	1 732	3 462	2 383
2003	8 196	3 962	11 439	5 288
2004	10 738	3 798	13 242	4 778
2005	13 581	3 944	15 724	4 733
2006	23 467	5 305	25 934	6 148
2007	34 104	6 615	36 503	7 389
2008	57 903	9 763	60 868	10 664
2009	62 053	12 210	64 383	12 836
2010	93 422	30 785	96 086	31 789
Total	305 205	44 234	328 565	51 685

TABLE 14 – PSAP et incertitude à l’ultime estimées par bootstrap.



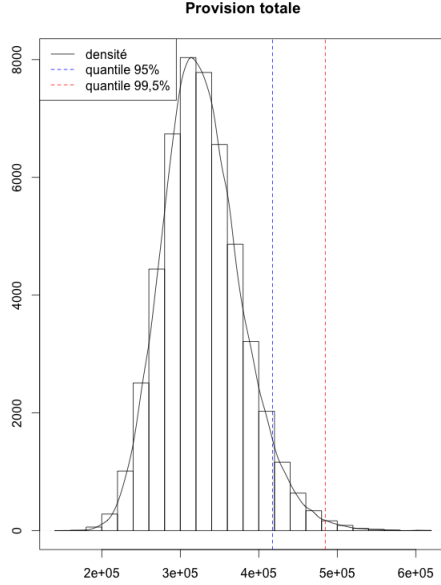
On présente ci-contre la densité empirique obtenue par la méthode bootstrap (1) pour l'estimation de la provision. La moyenne de cette distribution est égale à la PSAP estimée et l'erreur de prédiction correspond à l'incertitude autour de cette moyenne.

FIGURE 23 – Histogramme, densité et quantiles de la provision totale dans le cas du bootstrap Poisson (1).

On peut également observer les quantiles à 95 % (bleu) et 99,5 % (rouge) égaux à 382 160 et 440 578 respectivement. Le quantile d'ordre 95 % est le montant tel que 95 % des observations au maximum sont inférieures à celui-ci et 5 % sont supérieures.

On obtient avec cette méthode bootstrap des PSAP estimées équivalentes à celles obtenues par la méthode Chain Ladder. On s'attendait à ce résultat dans la mesure où le modèle sous-jacent utilisé reproduit les PSAP Chain Ladder. En revanche, l'incertitude à l'ultime estimée par bootstrap est bien plus importante que

celle obtenue avec le modèle de Mack ; on a donc une variabilité des montants plus grande dans la méthode par simulation.



On présente ci-contre la densité empirique obtenue par la méthode bootstrap (2) pour l'estimation de la provision. On constate que l'insertion d'un facteur de queue de développement a augmenté la moyenne de la distribution.

FIGURE 24 – Histogramme, densité et quantiles de la provision totale dans le cas du bootstrap Poisson (2), c'est-à-dire avec le facteur de queue $\hat{f}_{ult}$.

Les quantiles à 95 % (bleu) et 99,5 % (rouge) sont cette fois-ci naturellement supérieurs aux précédents et sont égaux à 419 518 et 487 367 respectivement. On présente dans le tableau suivant le skewness et le kurtosis des deux distributions ainsi obtenues :

	bootstrap (1)	bootstrap (2)
skewness	0,379	0,489
kurtosis	0,376	0,651

TABLE 15 – Coefficients de skewness et de kurtosis des deux distributions obtenues, c'est-à-dire avec ou sans facteur de queue de développement.

Le coefficient d'asymétrie, ou skewness, S est un indicateur pour étudier l'asymétrie d'une distribution et le coefficient d'aplatissement, ou kurtosis, K mesure le degré d'aplatissement de cette distribution. Ils sont définis par :

$$S = \frac{\mathbb{E}[(X - \bar{X})^3]}{(\sqrt{\mathbb{E}[(X - \bar{X})^2]})^3} \text{ et } K = \frac{\mathbb{E}[(X - \bar{X})^4]}{(\mathbb{E}[(X - \bar{X})^2])^2} - 3$$

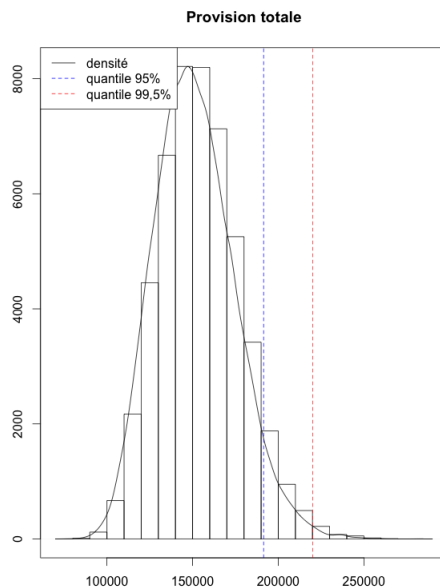
On observe pour ces distributions des coefficients positifs, ce qui signifie que celles-ci sont asymétriques. En effet, le skewness indique que la plupart des valeurs sont concentrées à gauche de la moyenne et il y a des valeurs extrêmes à droite. Le kurtosis témoigne également de valeurs concentrées autour de la moyenne et de queues relativement fines. L'insertion d'un facteur de queue de développement ne semble pas modifier fondamentalement la forme de la distribution.

- **Compagnie B**

Contrairement à la compagnie A, la méthode bootstrap Poisson avec prise en compte d'un facteur de queue de développement n'est pas retenue pour la compagnie B.

	bootstrap Poisson (1)	
	PSAP estimée	Erreur de prédiction
2001	-	-
2002	446	680
2003	1 061	947
2004	2 133	1 290
2005	4 102	1 788
2006	8 154	2 544
2007	15 011	3 608
2008	26 682	5 210
2009	38 532	7 709
2010	58 593	18 945
Total	154 715	23 721

TABLE 16 – PSAP et incertitude à l'ultime estimées par bootstrap.



On présente ci-contre la densité obtenue par cette méthode bootstrap pour l'estimation de la provision. On observe également les quantiles à 95 % (bleu) et 99,5 % (rouge) égaux à 195 257 et 224 702 respectivement.

FIGURE 25 – Histogramme, densité et quantiles de la provision totale dans le cas du bootstrap Poisson (1).

L'approche Chain Ladder

Pour comparer les résultats de ces différentes méthodes, on utilisera l'indice de comparaison suivant :

$$I = \frac{\text{incertitude à l'ultime de la PSAP}}{\text{PSAP estimée}}.$$

• Compagnie A

	Méthode	PSAP	Incertainitude	Indice I
Sans f_{ult}	Mack	302 882	34 220	11,30%
	bootstrap Poisson (1)	305 205	44 234	14,49%
Avec f_{ult}	Mack	327 244	35 251	10,77 %
	bootstrap Poisson (2)	328 565	51 685	15,73 %

TABLE 17 – Synthèse de l'approche Chain Ladder

• Compagnie B

	Méthode	PSAP	Incertainitude	Indice I
Sans f_{ult}	Mack	153 993	20 981	13,63%
	bootstrap Poisson (1)	154 715	23 721	15,33%

TABLE 18 – Synthèse de l'approche Chain Ladder

10.2 L'approche Bornhuetter-Ferguson du risque de provisionnement

Pour cette approche, on souhaite utiliser un indicateur d'exposition au risque, à savoir le montant de primes acquises par année de survenance i , $i \in \{1, \dots, I\}$. Les données exogènes qui interviennent dans les méthodes suivantes sont donc les primes P_i et les ratios sinistres à primes attendus et fixés à l'avance $\frac{S}{P}_i$; ces données permettent de déterminer les estimations a priori des charges ultimes $\hat{U}_i$ présentées dans le tableau ci-dessous :

Compagnie A

Année de survenance	Charge ultime a priori $\hat{U}_i$
2001	38 212
2002	54 753
2003	112 398
2004	119 550
2005	113 085
2006	126 433
2007	127 490
2008	135 823
2009	119 790
2010	100 594

Compagnie B

Année de survenance	Charge ultime a priori $\hat{U}_i$
2001	75 389
2002	61 547
2003	58 299
2004	58 680
2005	67 877
2006	80 420
2007	101 406
2008	120 710
2009	120 992
2010	125 893

TABLE 19 – Charges ultimes a priori par année de survenance.

10.2.1 L'estimation de la PSAP par les méthodes déterministes choisies

Les tableaux suivants présentent les PSAP obtenues par les méthodes Bornhuetter-Ferguson et Benktander pour les deux compagnies.

Compagnie A			Compagnie B		
	Bornhuetter-Ferguson	Benktander		Bornhuetter-Ferguson	Benktander
2001	-	-	2001	-	-
2002	1 505	1 710	2002	476	446
2003	7 781	8 106	2003	1 167	1 057
2004	14 114	10 703	2004	2 441	2 133
2005	19 887	13 704	2005	5 405	4 106
2006	33 101	24 026	2006	10 545	8 189
2007	48 880	36 128	2007	20 842	15 252
2008	73 542	62 335	2008	43 100	28 758
2009	86 011	74 259	2009	72 961	50 997
2010	94 026	93 792	2010	113 357	102 880
Total	378 847	324 763	Total	270 294	213 820

TABLE 20 – PSAP estimées par les méthodes Bornhuetter-Ferguson et Benktander.

Pour rappel, la PSAP estimée par la méthode Chain-Ladder s'élève à 302 882 et à 153 993 pour les compagnies A et B respectivement. On constate que les PSAP estimées par les méthodes Bornhuetter-Ferguson et Benktander sont nettement supérieures aux PSAP Chain Ladder, ce qui montre la prudence de ces deux méthodes. En effet, les ultimes a priori apportent une information supplémentaire que ces méthodes prennent en considération dans l'estimation de la PSAP.

10.2.2 La mesure de l'incertitude à l'ultime de la PSAP par les méthodes stochastiques

Les méthodes de Mack (2008) et d'Alai, Merz, Wüthrich (2009) permettent d'obtenir l'incertitude associée à la PSAP dans le cadre de Bornhuetter-Ferguson. Ces méthodes utilisent deux approches différentes pour l'estimation des cadences de règlements : Mack estime que l'information présente dans les estimations a priori des charges ultimes doit impacter ces cadences, tandis que Alai, Merz et Wüthrich considèrent que les cadences obtenues avec Chain Ladder sont les plus adéquates.

La méthode de Mack

Estimation des paramètres

Selon Mack [24], l'information a priori sur les ultimes doit impacter les cadences de règlement ; c'est pourquoi celles-ci diffèrent des cadences Chain Ladder. Dans l'approche de Mack, il est possible d'ajuster les cadences de règlement non cumulées en effectuant une régression log-linéaire. Le nombre de paramètres est alors réduit et les estimations équilibrées. Le tableau suivant présente les cadences de règlement incrémentales $\hat{y}_j$ estimées, puis les cadences retenues après lissage $\tilde{y}_j$ et normalisation $\hat{y}_j^*$. Les paramètres $\hat{s}_j^2$ également introduits dans ce modèle sont obtenus d'une

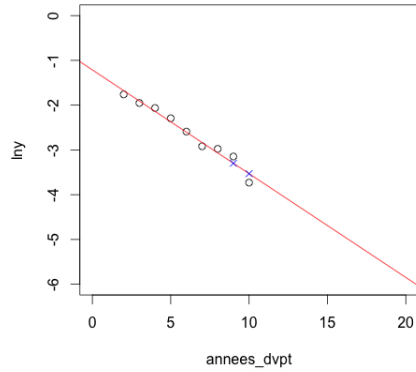
manière identique. Les paramètres affichés en gras indiquent une modification.

• **Compagnie A**

j	1	2	3	4	5	6	7	8	9	10	10+
$\hat{y}_j$	5,3 %	17,2 %	14,2 %	12,7 %	10,1 %	7,5 %	5,4 %	5,1%	4,3 %	2,4 %	-
$\hat{y}_j$	5,3 %	17,2 %	14,2 %	12,7 %	10,1 %	7,5 %	5,4 %	5,1%	3,7 %	2,9 %	2,8 %
$\hat{y}_j^*$	6,1 %	19,8 %	16,3 %	14,7 %	11,6 %	8,6 %	6,2 %	5,8 %	4,3 %	3,3 %	3,2 %
$\hat{s}_j^2$	28,2	188,6	54,2	307	108,6	44,4	5,2	16,1	9,2	-	-
$\hat{s}_j^*$	28,2	188,6	54,2	307	108,6	44,4	13,6	16,1	9,2	4,6	4,4

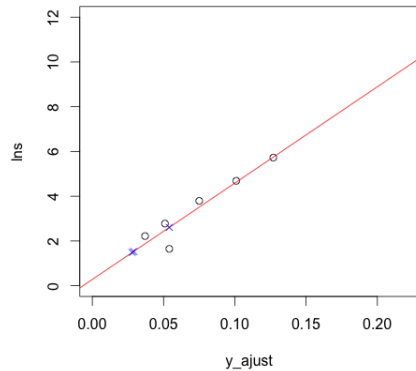
TABLE 21 – Paramètres estimés dans la méthode de Mack.

Une régression log-linéaire a été effectuée sur les cadences de règlement incrémentales $\hat{y}_j$ à partir de $\hat{y}_2$, c'est-à-dire pour la partie décroissante des cadences. Sur le schéma suivant, les logarithmes des cadences $\hat{y}_j$ sont représentés en fonction des années de développement.



Le coefficient de détermination R^2 est supérieur à 97 %, ce qui traduit une excellente qualité d'ajustement. On a ainsi ajusté les cadences $\hat{y}_9$ et $\hat{y}_{10}$ et extrapolé ces cadences pour obtenir un facteur de queue de développement cohérent avec la valeur estimée au début de cette étude.

En ce qui concerne les paramètres $\hat{s}_j^*$, une régression log-linéaire a été une fois encore effectuée sur les $\hat{s}_j^2$. Le schéma suivant présente les logarithmes des paramètres $\hat{s}_j^2$ en fonction des valeurs absolues $|\hat{y}_j^*|$ pour les valeurs décroissantes, c'est-à-dire pour les années de développement $j \geq 4$.



Le coefficient de détermination R^2 est de 90 %, ce qui est satisfaisant dans la mesure où on dispose de peu de données. Comme précédemment, on ajuste les paramètres à partir de cette droite de régression, et on extrapole les paramètres $\hat{s}_{10}^*$ et $\hat{s}_{10+}^*$.

Un autre point où l'actuaire doit faire intervenir son jugement d'expert est l'incertitude sur la cadence de règlement incrémentale $\hat{y}_{10+}^*$.

j	1	2	3	4	5	6	7	8	9	10	10+
$\hat{y}_j^*$	6,1 %	19,8 %	16,3 %	14,7 %	11,6 %	8,6 %	6,2 %	5,8 %	4,3 %	3,3 %	3,2 %
$\widehat{MSE}(\hat{y}_j^*)$	0,52 %	1,41 %	0,81 %	2,11 %	1,39 %	1,01 %	0,65 %	0,89 %	0,99 %	1,10 %	1,10 %

TABLE 22 – Cadences de règlement incrémentales et incertitudes associées.

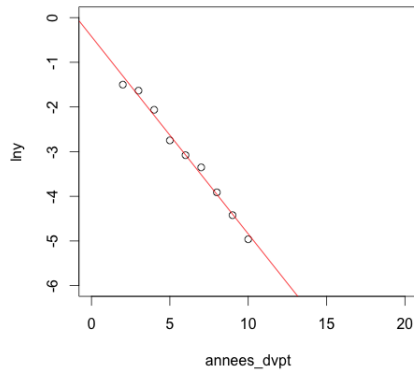
L'incertitude sur la dernière cadence de règlement incrémentale a été choisie de sorte qu'elle soit équivalente à $\widehat{MSE}(\hat{y}_{10}^*)$. En effet, les deux cadences $\hat{y}_{10}^*$ et $\hat{y}_{10+}^*$ sont très proches l'une de l'autre.

• Compagnie B

Comme dans les méthodes de provisionnement précédentes, on ne considère pas de facteur de queue de développement dans le cas de la compagnie B. On présente les résultats obtenus de façon similaire à ceux de la compagnie A.

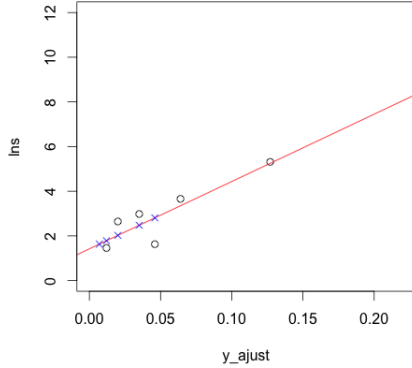
j	1	2	3	4	5	6	7	8	9	10
$\hat{y}_j$	7,1 %	22,3 %	19,5 %	12,7 %	6,4 %	4,6 %	3,5 %	2 %	1,2 %	0,7 %
$\hat{y}_j^*$	7,1 %	22,3 %	19,5 %	12,7 %	6,4 %	4,6 %	3,5 %	2 %	1,2 %	0,7 %
$\hat{s}_j^2$	37,5	187,9	131,5	203,1	38,8	5,1	19,7	14,1	4,3	-
$\hat{s}_j^*$	37,5	187,9	131,5	203,1	38,8	16,6	11,9	7,6	5,9	5,1

TABLE 23 – Paramètres estimés dans la méthode de Mack.



Le coefficient de détermination R^2 est de 99 %, ce qui traduit une excellente qualité d'ajustement. Aucune cadence n'est ajustée et, dans la mesure où la somme de ces cadences $\hat{y}_j^*$ est égale à 100 %, une normalisation n'est pas nécessaire.

En revanche, les paramètres $\hat{s}_j^*$ ont quant à eux nécessité un ajustement. On considère les logarithmes de ces paramètres en fonction des $|\hat{y}_j^*|$ pour les années de développement $j \geq 4$, à savoir la partie décroissante des $\hat{s}_j^2$.



Le coefficient de détermination R^2 est de 78 % ; il s'agit donc d'être prudent quant à l'ajustement des paramètres. $\hat{s}_6^*$ à $\hat{s}_9^*$ ont été ajustés, et $\hat{s}_{10}^*$ obtenu par extrapolation.

Erreur de prédiction de la provision en fonction du coefficient de corrélation

Comme c'est souvent le cas en pratique, les ultimes a priori $\hat{U}_i$ sont positivement corrélés, de sorte qu'on ne peut utiliser l'égalité proposée dans la deuxième partie de ce mémoire pour le calcul de l'erreur de ces $\hat{U}_i$. Dans le calcul de l'erreur d'estimation de la provision, un terme de covariance apparaît. Ce dernier contient les estimations des coefficients de corrélations $\hat{\rho}_{i,j}^z$ et $\hat{\rho}_{i,j}^U$, $\hat{\rho}_{i,j}^z$ étant défini par :

$$\hat{\rho}_{i,j}^z = \sqrt{\frac{\hat{z}_{J+1-j}^*(1 - \hat{z}_{J+1-i}^*)}{\hat{z}_{J+1-i}^*(1 - \hat{z}_{J+1-j}^*)}}, \text{ pour } i < j \text{ et } \hat{z}_1^* \leq \dots \leq \hat{z}_{J+1}^*.$$

□ On se place tout d'abord dans le cas d'un coefficient de corrélation $\hat{\rho}_{i,j}^U$ constant. Quelle que soit la compagnie (A ou B), on a alors :

$$\hat{\rho}_{i,j}^U = \frac{1}{\sqrt{I}} = \frac{1}{\sqrt{10}} \approx 31,62\%.$$

Dans ce cadre, l'erreur d'estimation des ultimes a priori $\hat{U}_i$ s'écrit :

$$\widehat{MSE}(\hat{U}_i) = \frac{P_i}{I - \sqrt{I}} \sum_{j=1}^J P_j \left(\frac{\hat{U}_j}{P_j} - \frac{\sum_{k=1}^I \hat{U}_k}{\sum_{k=1}^I P_k} \right)^2, \text{ pour } i \in \{1, \dots, I\}.$$

L'actuaire peut alors vérifier la vraisemblance de ces erreurs de la manière suivante : sous l'hypothèse d'une loi normale, alors l'intervalle

$$[\hat{U}_i - 2\widehat{MSE}(\hat{U}_i); \hat{U}_i + 2\widehat{MSE}(\hat{U}_i)]$$

doit contenir la valeur $\mathbb{E}[\hat{U}_i]$ avec une probabilité de 95 %. On présente dans le tableau ci-dessous cette valeur ainsi que l'intervalle défini ci-dessus.

Le tableau suivant donne l'estimation de la provision pour sinistres à payer ainsi que l'erreur de prédiction qui lui est associée pour les deux compagnies A et B.

	Compagnie A			Compagnie B		
	Borne inférieure	$\mathbb{E}[U_i]$	Borne supérieure	Borne inférieure	$\mathbb{E}[U_i]$	Borne supérieure
2001	18 152	34 878	58 272	52 863	72 216	97 915
2002	33 674	64 090	75 831	38 893	57 908	84 201
2003	79 680	121 188	145 115	31 824	53 173	84 774
2004	86 288	99 520	152 812	33 977	52 002	83 383
2005	78 738	89 175	147 431	41 154	53 669	94 600
2006	91 115	105 658	161 750	51 318	66 051	109 522
2007	91 538	109 548	163 443	70 011	81 358	132 801
2008	98 493	127 367	173 152	87 064	94 900	154 356
2009	82 425	113 057	157 154	86 459	101 807	155 525
2010	62 781	100 903	138 407	89 414	121 098	162 372

TABLE 24 – Intervalle de confiance à 95 % pour l'espérance de $\hat{U}_i$.

	Compagnie A		Compagnie B	
	PSAP estimée	Erreur de prédiction	PSAP estimée	Erreur de prédiction
2001	1 233	679	-	-
2002	3 596	1 312	569	635
2003	12 175	3 092	1 388	937
2004	19 927	4 228	2 879	1 290
2005	25 878	5 184	6 304	1 999
2006	39 797	7 191	12 060	3 044
2007	54 981	9 894	23 356	4 844
2008	78 514	13 816	46 889	9 134
2009	88 805	16 201	76 498	13 188
2010	94 464	19 774	114 684	18 798
Total	419 370	55 403	284 627	40 225

TABLE 25 – PSAP et incertitude à l'ultime estimées.

□ On se place désormais dans le cas d'un coefficient de corrélation $\hat{\rho}_{i,j}^U$ décroissant défini par :

$$\hat{\rho}_{i,j}^U = \frac{1}{1 + |i - j|}.$$

Dans ce cadre, l'erreur d'estimation des ultimes a priori $\hat{U}_i$ s'écrit :

$$\widehat{MSE}(\hat{U}_i) = \frac{P_i}{I - \sqrt{2I}} \sum_{j=1}^J P_j \left(\frac{\hat{U}_j}{P_j} - \frac{\sum_{k=1}^I \hat{U}_k}{\sum_{k=1}^I P_k} \right)^2, \text{ pour } i \in \{1, \dots, I\}.$$

On présente à nouveau le tableau permettant à l'actuaire de vérifier que l'erreur sur les ultimes a priori $\hat{U}_i$ est plausible.

Le tableau suivant donne l'estimation de la provision pour sinistres à payer ainsi que l'erreur de prédiction qui lui est associée pour les deux compagnies A et B.

	Compagnie A			Compagnie B		
	Borne inférieure	$\mathbb{E}[U_i]$	Borne supérieure	Borne inférieure	$\mathbb{E}[U_i]$	Borne supérieure
2001	15 902	34 878	60 522	50 336	72 216	100 442
2002	31 310	64 090	78 195	36 352	57 908	86 742
2003	76 009	121 188	148 786	28 853	53 173	87 745
2004	82 556	99 520	156 544	31 206	52 002	86 154
2005	74 885	89 175	151 284	38 157	53 669	97 597
2006	87 153	105 658	165 712	48 053	66 051	112 787
2007	87 504	109 548	167 476	66 489	81 358	136 323
2008	94 306	127 367	177 339	83 290	94 900	158 130
2009	78 233	113 057	161 346	82 585	101 807	159 399
2010	58 539	100 903	142 649	85 321	121 098	166 465

TABLE 26 – Intervalle de confiance à 95 % pour l'espérance de $\hat{U}_i$.

	Compagnie A		Compagnie B	
	PSAP estimée	Erreur de prédiction	PSAP estimée	Erreur de prédiction
2001	1 233	699	-	-
2002	3 596	1 357	569	624
2003	12 175	3 214	1 388	879
2004	19 927	4 441	2 879	1 135
2005	25 878	5 529	6 304	1 563
2006	39 797	7 686	12 060	2 118
2007	54 981	10 592	23 356	3 218
2008	78 514	14 781	46 889	6 373
2009	88 805	17 549	76 498	7 394
2010	94 464	21 580	114 684	8 791
Total	419 370	57 153	284 627	34 610

TABLE 27 – PSAP et incertitudes à l'ultime estimées.

La méthode d'Alai, Merz et Wüthrich

La principale différence du modèle d'Alai, Merz et Wüthrich [3] par rapport à celui de Mack [24] concerne l'estimation des cadences de développement. Celles-ci sont estimées par la méthode Chain Ladder dans ce modèle. Les deux tableaux suivants les rappellent :

• Compagnie A

j	1	2	3	4	5	6	7	8	9	10
$\hat{y}_j$	6,5 %	21,7 %	17,7 %	15,8 %	12,2 %	8,6 %	5,8 %	4,9 %	4,2 %	2,7 %

• Compagnie B

j	1	2	3	4	5	6	7	8	9	10
$\hat{y}_j$	10,0 %	29,7 %	24,6 %	15,2 %	7,4 %	5,1 %	3,8 %	2,2 %	1,2 %	0,8 %

TABLE 28 – Cadences de règlement incrémentales.

Pour l'estimation de l'erreur de prédiction, une estimation de l'incertitude des charges ultimes a priori $\hat{\nu}_k$ est nécessaire. La variance a priori $\widehat{\text{Var}}(\hat{\nu}_k)$ est supposée donnée de manière exogène. On fait ici l'hypothèse que l'incertitude des charges a priori $\hat{\nu}_k$ est obtenue par un coefficient de variation de 5 %,

$$\text{ie } \widehat{\text{Var}}(\hat{\nu}_k) = \hat{\nu}_k^2 (0,05)^2.$$

On cherche désormais le paramètre de dispersion ϕ qui ne dépend ni de l'année de survenance i ni de l'année de développement j . Il peut être estimé par la méthode du maximum de vraisemblance ou bien en utilisant les résidus de Pearson comme c'est le cas ici. Avec les notations du modèle, les résidus de Pearson sont définis par :

$$\hat{R}_{i,j}^P = \frac{X_{i,j} - \hat{m}_{i,j}}{\sqrt{\hat{m}_{i,j}}}, \text{ avec } \hat{m}_{i,j} = \hat{\mu}_i \hat{\gamma}_j,$$

et on en déduit l'estimation du paramètre de dispersion :

$$\hat{\phi} = \frac{1}{d} \sum_{i+j \leq I+1} \frac{(X_{i,j} - \hat{m}_{i,j})^2}{\hat{m}_{i,j}} = \frac{\sum_{2 \leq i+j \leq I+1} (\hat{R}_{i,j}^P)^2}{N - p},$$

avec

$$\begin{cases} d &= \frac{I(I+1)}{2} - 2I + 1 \text{ le degré de liberté du modèle,} \\ N &= |\mathcal{D}_I| = 55 \text{ le nombre d'observations } X_{i,j} \text{ à la date } t = I, \\ p &= I + J - 1 = 19 \text{ le nombre de paramètres estimés.} \end{cases}$$

On estime ainsi le paramètre de dispersion par :

	$\hat{\phi}$
Compagnie A	557
Compagnie B	528

TABLE 29 – Estimation du paramètre ϕ pour les deux companies.

On présente alors la provision pour sinistres à payer et l'erreur de prédiction associée pour les deux compagnies dans le tableau sur la page suivante.

On constate comme prévu que ce modèle ODP appliqué à Bornhuetter-Ferguson réplique bien la provision estimée par Bornhuetter-Ferguson avec les cadences de règlement Chain Ladder.

L'approche Bornhuetter-Ferguson

On compare tout d'abord dans les tableaux suivants les cadences de règlement cumulées utilisées dans le modèle de Mack [24] $\hat{z}_j = \sum_{k=1}^j \hat{y}_k^*$ et dans le modèle d'Alai, Merz et Wüthrich (AMW) $\hat{\beta}_j = \sum_{k=1}^j \hat{\gamma}_k$ pour $j \in \{1, \dots, J\}$.

La méthode de Mack conduit à tenir compte d'un facteur de queue de développement, ce qui permet d'être cohérent avec l'étude préalable du triangle de liquidation

	Compagnie A		Compagnie B	
	PSAP estimée	Erreur de prédiction	PSAP estimée	Erreur de prédiction
2001	-	-	-	-
2002	1 505	1 474	476	681
2003	7 781	3 543	1 167	987
2004	14 114	4 276	2 441	1 363
2005	19 887	4 593	5 405	1 998
2006	33 101	5 661	10 545	2 785
2007	48 880	6 548	20 842	3 962
2008	73 542	7 921	43 100	5 744
2009	86 011	8 377	72 961	7 508
2010	94 026	8 662	113 357	9 666
Total	378 847	25 310	270 294	16 550

TABLE 30 – Estimation de la PSAP et de l’incertitude à l’ultime pour les deux compagnies.

	Compagnie A					Compagnie B			
	Mack		AMW			Mack		AMW	
	$\hat{z}_j$	$\widehat{MSE}(\hat{z}_j)$	$\hat{\beta}_j$	$\widehat{MSE}(\hat{\beta}_j)$		$\hat{z}_j$	$\widehat{MSE}(\hat{z}_j)$	$\hat{\beta}_j$	$\widehat{MSE}(\hat{\beta}_j)$
1	6,1 %	0,5 %	6,5 %	0,7 %	1	8,9 %	0,7 %	10,0 %	0,9 %
2	25,9 %	1,5 %	28,2 %	1,6 %	2	36,8 %	1,6 %	39,7 %	1,8 %
3	42,2 %	1,7 %	45,9 %	2,1 %	3	61,2 %	2,0 %	64,3 %	2,0 %
4	56,9 %	2,7 %	61,7 %	2,4 %	4	77,0 %	1,3 %	79,5 %	1,9 %
5	68,5 %	2,4 %	73,8 %	2,6 %	5	85,0 %	1,0 %	86,9 %	1,7 %
6	77,1 %	2,1 %	82,4 %	2,7 %	6	90,7 %	0,9 %	92,0 %	1,5 %
7	83,3 %	2,0 %	88,2 %	2,6 %	7	95,1 %	0,8 %	95,8 %	1,3 %
8	89,2 %	1,8 %	93,1 %	2,5 %	8	97,6 %	0,6 %	98,0 %	1,0 %
9	93,4 %	1,6 %	97,3 %	2,1 %	9	99,1 %	0,5 %	99,2 %	0,7 %
10	96,8 %	1,1 %	100 %	0 %	10	100 %	0 %	100 %	0 %
10+	100 %	0 %	-	-					

TABLE 31 – Paramètres et incertitudes associées pour les méthodes de Mack (2008) et Alai, Merz et Wüthrich.

de la compagnie A. En revanche, la méthode AMW utilise les cadences Chain Ladder mais ne permet pas l’insertion d’un tel facteur, ce qui peut être un inconvénient. On constate également une diminution plus rapide de l’incertitude dans la méthode de Mack.

On peut également souligner le rôle comparable des paramètres $\hat{s}_j^*$ introduits précédemment dans le modèle de Mack et du paramètre $\hat{\phi}$ dans le modèle AMW. Ces deux types de paramètres diffèrent dans la mesure où les $\hat{s}_j^*$ dépendent des années de développement j . C’est pourquoi la principale différence entre les deux modèles réside dans l’erreur de processus. La variance du règlement incrémental s’écrit :

$$\begin{cases} \text{Var}(X_{i,j}) = \phi m_{i,j} & \text{dans le modèle de Mack,} \\ \text{Var}(X_{i,j}) = s_j^2 m_{i,j} & \text{dans le modèle AMW.} \end{cases}$$

La volatilité des paramètres s_j^2 impacte alors fortement l’erreur de processus. Dans le tableau suivant, on observe que la divergence majeure entre les deux modèles réside dans cette erreur, tandis que les erreurs d’estimation sont moins éloignées l’une de l’autre¹⁵.

15. Pour rappel, dans le modèle AMW, l’erreur d’estimation comporte à la fois l’erreur d’estimation des paramètres et l’erreur des ultimes a priori.

	Compagnie A		Compagnie B	
	Erreur de processus	Erreur d'estimation	Erreur de processus	Erreur d'estimation
Mack ^(*)	40 573	37 726	30 533	26 188
AMW	14 533	20 721	11 950	11 450

(*) On utilise le cas $\hat{\rho}_{i,j}^U$ constant.

TABLE 32 – Comparaison des erreurs de processus et d'estimation des méthodes de Mack (2008) et Alai, Merz et Wüthrich.

Même si les erreurs d'estimation peuvent paraître relativement éloignées dans le cas des compagnies A et B, pour des montants de provisions estimées bien importants elles restent bien moins espacées que les erreurs de processus¹⁶.

• Compagnie A

	Méthode	PSAP	Incertitude	Indice I
Sans f_{ult}	Bornhuetter-Ferguson	378 847	-	-
	Benktander	324 763	-	-
	Alai, Merz, Wüthrich	378 847	25 310	6,7 %
Avec f_{ult}	Mack avec $\hat{\rho}_{i,j}^U$ constant	419 370	55 403	13,2 %
	Mack avec $\hat{\rho}_{i,j}^U$ décroissant	419 370	57 153	13,6 %

TABLE 33 – Synthèse de l'approche Bornhuetter-Ferguson.

• Compagnie B

	Méthode	PSAP	Incertitude	Indice I
Sans f_{ult}	Bornhuetter-Ferguson	270 294	-	-
	Benktander	213 820	-	-
	Mack avec $\hat{\rho}_{i,j}^U$ constant	284 627	40 225	14,1 %
	Mack avec $\hat{\rho}_{i,j}^U$ décroissant	284 627	34 610	12,2 %
	Alai, Merz, Wüthrich	270 294	16 550	6,1 %

TABLE 34 – Synthèse de l'approche Bornhuetter-Ferguson.

La première remarque que l'on peut faire après la lecture de ces deux tableaux concerne la méthode d'Alai, Merz et Wüthrich; en effet, celle-ci réplique bien la provision Bornhuetter-Ferguson, ce qui est un avantage pour ce modèle stochastique. En revanche, la provision estimée dans le modèle de Mack est plus élevée. On rappelle que cette divergence est due à l'utilisation de cadences de règlement différentes. D'autre part, quel que soit le coefficient de corrélation des ultimes $\hat{\rho}_{i,j}^U$ dans le modèle de Mack, la PSAP estimée reste la même.

16. On peut se reporter à l'article [3] pour de tels résultats.

En ce qui concerne l'incertitude à l'ultime des provisions, on constate une nette différence entre le modèle de Mack et le modèle d'Alai Merz et Wüthrich. ce dernier modèle permet d'avoir une incertitude beaucoup plus réduite que le modèle de Mack. Néanmoins, le modèle de Mack se place dans un cadre où les données exogènes sont encore plus présentes que dans le cadre de Bornhuetter-Ferguson ; en effet, il permet à l'actuaire de faire intervenir l'information des ultimes a priori dans l'estimation des cadences de règlement. C'est pourquoi l'actuaire peut choisir l'approche stochastique qu'il souhaite privilégier, à l'intérieur même du cadre de la méthode Bornhuetter-Ferguson.

La vision à l'ultime

• Compagnie A

		Méthode	PSAP	Incertitude	Indice I
Approche Chain Ladder	Sans f_{ult}	Mack	302 882	34 220	11,30%
		bootstrap Poisson (1)	305 205	44 234	14,49%
	Avec f_{ult}	Mack	327 244	35 251	10,77 %
		bootstrap Poisson (2)	328 565	51 685	15,73 %
Approche Bornhuetter Ferguson	Sans f_{ult}	Bornhuetter-Ferguson	378 847	-	-
		Benktander	324 763	-	-
		Alai, Merz, Wüthrich	378 847	25 310	6,7 %
	Avec f_{ult}	Mack avec $\hat{\rho}_{i,j}^U$ constant	419 370	55 403	13,2 %
		Mack avec $\hat{\rho}_{i,j}^U$ décroissant	419 370	57 153	13,6 %

TABLE 35 – Synthèse de la vision à l'ultime pour la compagnie A.

• Compagnie B

		Méthode	PSAP	Incertitude	Indice I
Approche Chain-Ladder	Sans f_{ult}	Mack	153 993	20 981	13,63%
		bootstrap Poisson (1)	154 715	23 721	15,33%
Bornhuetter-Ferguson		270 294	-	-	
Benktander		213 820	-	-	
Mack avec $\hat{\rho}_{i,j}^U$ constant		284 627	40 225	14,1 %	
Mack avec $\hat{\rho}_{i,j}^U$ décroissant		284 627	34 610	12,2 %	
Alai, Merz, Wüthrich		270 294	16 550	6,1 %	
Approche Bornhuetter Ferguson					

TABLE 36 – Synthèse de la vision à l'ultime pour la compagnie B.

Ce que l'on constate en premier lieu est la grande diversité des estimations des provisions et de l'incertitude à l'ultime. On remarque toutefois que la PSAP estimée dans l'approche Bornhuetter-Ferguson est bien plus importante que celle estimée dans l'approche Chain Ladder. L'utilisation de données exogènes reflète donc une plus grande prudence et est sans doute plus juste sur la réalité de la branche. A

noter toutefois que la méthode Benktander se place finalement entre Chain Ladder et Bornhuetter-Ferguson pour l'estimation de la provision. D'ailleurs, lorsque l'actuaire de la compagnie A insère un facteur de queue de développement dans la méthode Chain Ladder, la provision estimée est très proche de la provision Benktander qui tient compte d'informations complémentaires. Elle peut donc être très appropriée pour l'actuaire si ce dernier souhaite utiliser les informations exogènes au triangle dont il dispose, tout en restant proche de la provision Chain Ladder qui sert bien souvent de référence.

Dans le cadre Chain Ladder, la méthode bootstrap Poisson fournit une incertitude à l'ultime plus importante qu'avec la méthode Mack (1993), que ce soit avec ou sans facteur de queue de développement. Elle permet en revanche d'obtenir une distribution de la provision et, par conséquent, beaucoup d'informations. À l'inverse, la méthode de Mack (2008) dans le cadre Bornhuetter-Ferguson estime une incertitude à l'ultime bien plus élevée que celle du modèle d'Alai, Merz et Wüthrich qui se rapproche finalement plus de l'incertitude de Mack Chain Ladder (1993).

Ainsi, selon que l'actuaire souhaite privilégier le jugement d'expert, à travers l'utilisation d'un facteur de queue de développement par exemple, ou bien valoriser des informations dont il dispose, son choix pourra se porter sur l'un des modèles évoqués. Constituer une provision très prudente et appropriée à la réalité ou bien utiliser une méthode simulateur pour les nombreux avantages qu'elle comporte sont autant de possibilités offertes à l'actuaire. Celui-ci doit finalement effectuer un arbitrage entre ces méthodes selon les critères qu'il veut mettre en avant. Dans le cas des compagnies A et B, on pourrait être tenté de choisir le modèle d'Alai Merz et Wüthrich dans la mesure où les hypothèses dans l'approche Chain Ladder suggèraient une certaine prudence et l'erreur de prédiction associée à l'estimation de la provision dans ce modèle paraît raisonnable.

Dans la partie suivante, on cherche à répondre aux exigences de la Directive Solvabilité II en matière de provisionnement. On souhaite donc calculer l'incertitude à un an de la variation calendaire de la charge de sinistres ultime estimée afin de calculer le capital réglementaire.

11 La vision à un an du risque de provisionnement

Si un organisme assureur a besoin de connaître l'incertitude à l'ultime des provisions qu'il a constituées pour des raisons assez intuitives, la Directive Solvabilité II lui impose désormais de déterminer son capital réglementaire, ou Solvency Capital Requirement (SCR). Et, pour cela, il est nécessaire de connaître la volatilité de première année de déviation des provisions. On cherche ainsi, dans cette partie, à estimer l'incertitude à horizon un an des triangles des compagnies A et B étudiés précédemment et en déduire le SCR.

11.1 La méthode de référence Merz-Wüthrich

On veut tout d'abord obtenir l'incertitude à horizon un an par la méthode Merz Wüthrich qui peut désormais être considérée comme une référence. Les auteurs désignent la variation calendaire comme le **Claims Development Result (CDR)** pour l'année comptable $(I, I + 1]$: il correspond à la différence entre les provisions constituées à la date $t = I$ et celles ajustées grâce aux règlements observés pendant l'année comptable $(I, I + 1]$ à la date $t = I + 1$ pour les années de survenance antérieures à I . Le CDR *réel* pour l'année de survenance i et l'année comptable $(I, I + 1]$ est défini par :

$$CDR_i(I + 1) = \mathbb{E}[R_i^I | \mathcal{D}_I] - (X_{i, I-i+2} + \mathbb{E}[R_i^{I+1} | \mathcal{D}_{I+1}]),$$

où les $X_{i, I-i+1} = C_{i, I-i+1} - C_{i, I-i}$ désignent les règlements incrémentaux.

Comme les $(C_{i, j})_{j \geq 1}$ ont les propriétés d'une martingale, on peut écrire :

$$\mathbb{E}[CDR_i(I + 1) | \mathcal{D}_I] = 0.$$

Pour les facteurs de développement Chain Ladder connus f_j , le CDR attendu (vu à la date $t = I$) est nul. On ne doit donc pas anticiper de boni ni de mali. Mais ces facteurs de développement sont en général inconnus, et le CDR réel n'est pas observable. C'est pourquoi on s'intéresse au CDR *observable* ; il permet d'estimer le CDR réel et est défini de la manière suivante :

$$\widehat{CDR}_i(I + 1) = \hat{R}_i^{\mathcal{D}_I} - (X_{i, I-i+2} + \hat{R}_i^{\mathcal{D}_{I+1}}),$$

pour l'année de survenance i et l'année comptable $(I, I + 1]$.

Hypothèses du modèle

- ✓ Dans le modèle de Merz et Wüthrich, les règlements du triangle de liquidation sont supposées vérifier la relation suivante :

$$X_{i, j} = f_{j-1} \times X_{i, j-1} + \sigma_{j-1} \times \sqrt{X_{i, j-1}} \times \epsilon_{i, j},$$

où ϵ sont des variables aléatoires indépendantes conditionnellement aux $X_{1,1}, \dots, X_{I,1}$ et $\mathbb{E}[\epsilon_{i, j} | X_{1,1}, \dots, X_{I,1}] = 0$. Afin de vérifier cette hypothèse, on représente sur les graphes suivants les résidus égaux à $\frac{X_{i, j+1} - \hat{f}_j X_{i, j}}{\hat{\sigma}_j \sqrt{X_{i, j}}}$ par années de survenance, années de développement et années calendaires.

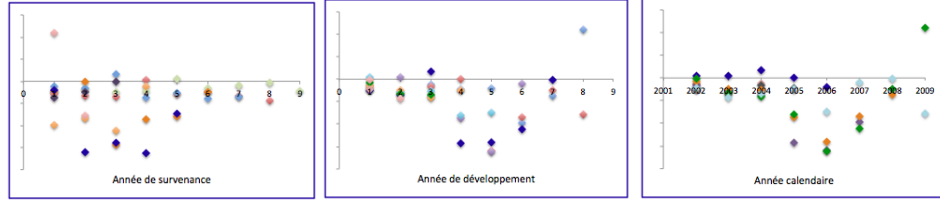


FIGURE 26 – Représentations des résidus en fonction des années de survénance, des années de développement et des années calendaires pour la compagnie A.

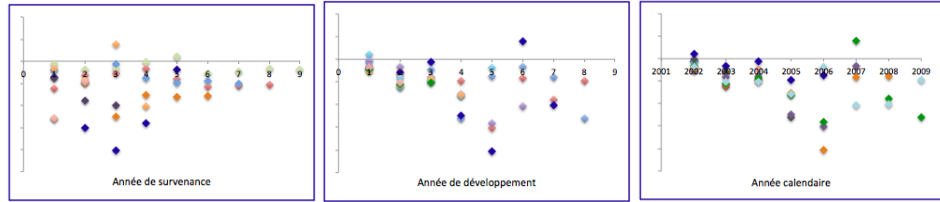


FIGURE 27 – Représentations des résidus en fonction des années de survénance, des années de développement et des années calendaires pour la compagnie B.

Ces graphiques ne font pas apparaître une tendance aléatoire ; les résidus semblent répartis dans la partie inférieure du graphe. Les hypothèses du modèle ne paraissent donc pas entièrement satisfaites.

- ✓ On présente ensuite deux ACF représentant la fonction d'autocorrélation empirique des résidus. On souhaite ainsi montrer l'indépendance des résidus.

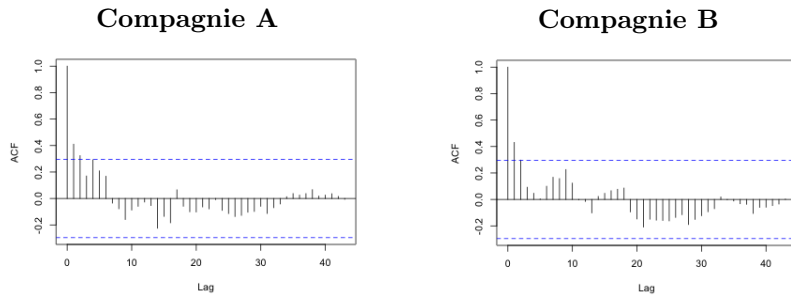


FIGURE 28 – ACF des résidus pour les compagnies A et B.

Sur les deux graphiques, on constate des covariances non nulles entre certains résidus aux premiers lags. La plupart des autres covariances tracées sur ces graphiques sont contenues dans l'intervalle en pointillé bleu, soit l'intervalle de confiance à 95 %. Par définition de la fonction d'autocorrélation, la première valeur est toujours égale à l'unité. On ne rejette pas l'hypothèse d'indépendance des résidus, bien qu'il faille rester prudent quant aux résultats obtenus par la suite.

- ✓ Dans les deux tableaux suivants, on rappelle les facteurs de développement $\hat{f}_j$ Chain Ladder des deux triangles A et B, puis on présente les estimations des paramètres σ_j du modèle de Mack [21] ainsi qu'un ratio justifiant l'utilisation de formules approchées.

• Compagnie A

j	1	2	3	4	5	6	7	8	9
$\hat{f}_j$	4,32	1,63	1,34	1,20	1,12	1,07	1,06	1,04	1,03
$\hat{\sigma}_j$	50,94	38,29	25,38	7,53	3,27	4,34	3,54	0,89	0,22
$\frac{\hat{\sigma}_j^2/\hat{f}_j^2}{C_{I-j+1,j}}$	$2,2e^{-2}$	$2,3e^{-2}$	$7,3e^{-3}$	$7,2e^{-4}$	$1,3e^{-4}$	$2,6e^{-4}$	$1,4e^{-4}$	$6,6e^{-6}$	$7,7e^{-7}$

• Compagnie B

j	1	2	3	4	5	6	7	8	9
$\hat{f}_j$	3,99	1,62	1,24	1,09	1,06	1,04	1,02	1,01	1,01
$\hat{\sigma}_j$	53,62	24,20	20,39	6,61	2,39	4,80	3,97	2,14	1,16
$\frac{\hat{\sigma}_j^2/\hat{f}_j^2}{C_{I-j+1,j}}$	$2,8e^{-2}$	$8,8e^{-3}$	$5,7e^{-3}$	$6,3e^{-4}$	$9,4e^{-5}$	$4,5e^{-4}$	$3,1e^{-4}$	$8,6e^{-5}$	$2,3e^{-5}$

TABLE 37 – Vérification de l’hypothèse pour l’approximation pour les compagnies A et B.

On constate bien que tous les ratios $\frac{\hat{\sigma}_j^2/\hat{f}_j^2}{C_{I-j+1,j}}$ sont nettement inférieurs à l’unité ; c’est pourquoi on peut utiliser les formules approchées de l’article de Merz et Wüthrich pour le calcul des erreurs.

L’incertitude à horizon un an

On présente dans le tableau suivant le CDR *observable* par année de survenance et au global pour les compagnies A et B. Il a pu être calculé en simulant la diagonale correspondant à l’année calendaire $(I, I + 1]$ par la technique du bootstrap. L’erreur de prédiction conditionnelle du CDR *observable* par rapport à 0 obtenue par l’approximation proposée par les auteurs est également donnée dans ce tableau.

	$\widehat{CDR}_i(I + 1)$		$\widehat{MSEP}_{\widehat{CDR}_i(I+1) \mathcal{D}_I}(0)^{1/2}$	
	Compagnie A	Compagnie B	Compagnie A	Compagnie B
2001	-	-	-	-
2002	-802	-282	92	370
2003	-6 660	-5 973	463	601
2004	-2 327	-3 479	1 296	1 039
2005	251	-3 233	1 436	1 243
2006	-4 293	207	1 253	809
2007	4 849	2 028	2 659	2 090
2008	-412	4 663	9 965	6 140
2009	3 659	-1 272	13 896	6 437
2010	3 996	-8 345	15 616	11 489
Total	-1 738	-15 686	26 255	16 477

TABLE 38 – CDR moyen et incertitude à un an par année de survenance et au global pour les deux compagnies.

On est dans le cadre de la vision prospective ; cette erreur de prédiction mesure l'incertitude sur la prédiction du CDR *observable* par la valeur nulle. On raisonne ainsi en termes de solvabilité et la marge de solvabilité est déduite de cette valeur. L'erreur de prédiction du CDR *observable* autour de 0 est de 26 255 et 16 477 pour les compagnies A et B respectivement. Ainsi, observer des $\widehat{CDR}_i(I+1)$ moyens de -1 738 et -15 686 n'est pas une situation improbable.

On ne présente pas ici d'application de la formule fermée proposée par Angoua et al. [4] pour la prise en compte d'un facteur de queue de développement. En effet, la méthode récursive que ceux-ci proposent a été préférée ici dans la mesure où on obtient en théorie les mêmes résultats tout en bénéficiant des atouts des méthodes simulatoires, comme l'obtention de la distribution de l'estimateur.

11.2 La méthode bootstrap de re-provisionnement

La méthode bootstrap de re-provisionnement a été appliquée aux triangles de liquidation A et B, ce qui a permis d'obtenir une distribution des CDR. On s'intéresse tout d'abord à la méthode dite directe, puis à la version améliorée. Certaines caractéristiques de cette distribution seront présentées uniquement pour la version améliorée. D'ailleurs, un facteur de queue de développement a été inséré pour la compagnie A dans la version améliorée.

La méthode directe de re-provisionnement

Ce tableau reprend les $\widehat{CDR}_i(I+1)$ moyens obtenus par années de survenance et les erreurs $\sigma(\widehat{CDR}_i(I+1))$ qui leur sont associées pour 50 000 simulations.

	Compagnie A		Compagnie B	
	$\widehat{CDR}_i(I+1)$	$\sigma(\widehat{CDR}_i(I+1))$	$\widehat{CDR}_i(I+1)$	$\sigma(\widehat{CDR}_i(I+1))$
2001	-	-	-	-
2002	-23	6 047	-7	5 552
2003	-52	8 851	10	5 333
2004	-70	7 900	-30	5 301
2005	-50	7 481	5	5 377
2006	-62	8 632	3	6 172
2007	-136	9 588	51	7 001
2008	-195	12 655	0	7 922
2009	-332	14 776	-83	9 562
2010	-1 177	33 118	-428	20 269
Total	-2 097	50 150	-480	29 423

TABLE 39 – CDR moyen et incertitude à un an par année de survenance et au global pour les deux compagnies.

La méthode améliorée de re-provisionnement

On présente ci-après les résultats obtenus par la méthode améliorée pour la compagnie A, avec et sans insertion de facteur de queue de développement, puis pour la compagnie B.

• **Compagnie A**

	sans facteur de queue		avec le facteur de queue $\hat{f}_{ult}$	
	$\widehat{CDR}_i(I+1)$	$\sigma(\widehat{CDR}_i(I+1))$	$\widehat{CDR}_i(I+1)$	$\sigma(\widehat{CDR}_i(I+1))$
2001	-	-	39	440
2002	-2	983	78	1 288
2003	-62	3 144	36	4 379
2004	-69	3 068	-3	4 105
2005	-76	3 059	-8	3 901
2006	-89	4 068	-38	5 024
2007	-113	5 012	-21	5 868
2008	-163	7 290	-74	8 222
2009	-148	8 679	-92	9 356
2010	-248	14 392	-229	15 137
Total	-970	29 416	-311	38 314

TABLE 40 – CDR moyen et incertitude à un an par année de survenance et au global pour les deux compagnies.

Certaines caractéristiques des distributions obtenues sont présentées ci-dessous.
sans facteur de queue *avec le facteur de queue $\hat{f}_{ult}$*

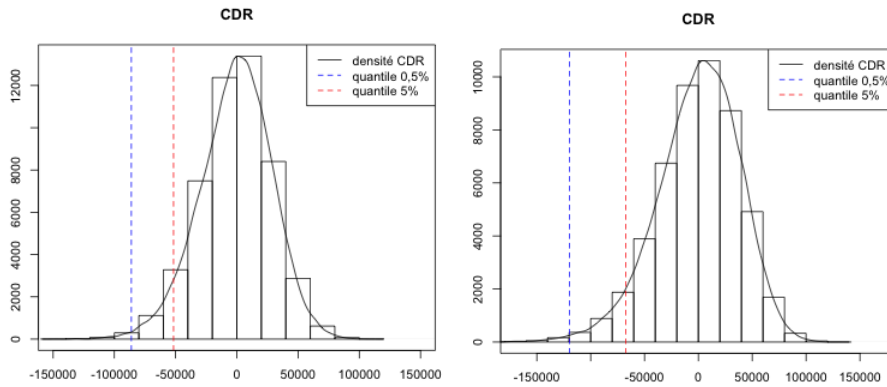


FIGURE 29 – Histogrammes, densités et quantiles à 0,5 % et 5 % du CDR pour la compagnie A, sans facteur de queue de développement et avec l'insertion d'un facteur $\hat{f}_{ult}$.

	sans facteur de queue	avec facteur de queue
- quantile 5 %	51 684	67 456
- quantile 0,5 %	86 161	119 719
skewness	-0,297	-0,505
kurtosis	0,300	0,563

TABLE 41 – Caractéristiques des distributions avec et sans facteur de queue de développement.

Le tableau ci-dessus présente tout d'abord les opposés des quantiles à 5 % et à 0,5 %, ce dernier étant nécessaire pour l'obtention du SCR par la suite. On constate que l'insertion d'un facteur de queue de développement augmente de manière significative le montant de ce quantile et, par conséquent, le SCR. On peut également comparer les indicateurs d'asymétrie (skewness) et d'aplatissement (kurtosis) : on observe des coefficients négatifs pour le skewness, c'est-à-dire qu'on constate une asymétrie. La plupart des valeurs sont concentrées à droite de la moyenne. Le coefficient de kurtosis est en revanche positif ; les observations sont ainsi concentrées autour de la moyenne avec des queues. L'insertion du facteur de queue de développement accentue ces caractéristiques.

• Compagnie B

	$\overline{CDR}_i(I+1)$	$\sigma(\overline{CDR}_i(I+1))$
2001	-	-
2002	3	486
2003	4	712
2004	-5	954
2005	-2	1 287
2006	12	1 802
2007	-1	2 571
2008	-15	3 806
2009	-32	5 276
2010	-60	8 939
Total	-96	13 705

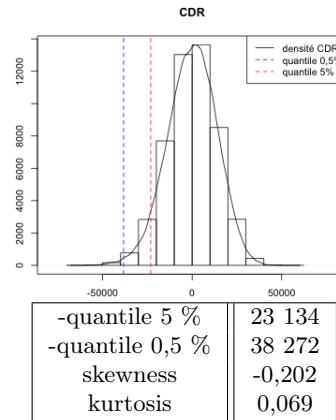


FIGURE 30 – Histogrammes, densités et quantiles à 0,5 % et 5 % du CDR pour la compagnie A, sans facteur de queue de développement et avec l'insertion d'un facteur $\hat{f}_{ult}$.

Pour les deux méthodes de re-provisionnement et pour les deux compagnies, on constate des CDR moyens négatifs. Cela correspond à des malis alors qu'on est supposé constater un CDR moyen nul. On peut remédier à ce problème en simulant un plus grand nombre de CDR ; par la loi des grands nombres, le CDR moyen doit en effet tendre vers zéro.

D'autre part, les erreurs associées aux CDR dans le cas de la méthode directe sont bien plus importantes que celles obtenues par la méthode améliorée. En effet, Angoua et al. [4] ont montré que la méthode directe tendait à augmenter significativement l'erreur d'estimation et, par là même, l'erreur de prédiction du CDR. Les erreurs d'estimation de la méthode directe sont respectivement de 42 092 et 21 911 au global pour les compagnies A et B. On remarque alors que ces erreurs sont supérieures aux erreurs de prédiction obtenues par la méthode améliorée et, par conséquent, supérieures aux erreurs d'estimation de cette méthode. De surcroît, on obtient des incertitudes à un an supérieures aux incertitudes à l'ultime avec cette première méthode de re-provisionnement.

Ces méthodes de re-provisionnement permettent également l'utilisation d'autres modèles déterministes ; elles ont ainsi été adaptées aux méthodes Bornhuetter-Ferguson et Benktander. Ces dernières méthodes bootstrap permettent ainsi de répondre aux exigences de Solvabilité II tout en tenant compte de données exogènes,

en l'occurrence ici les primes et les ratios sinistres à primes des deux compagnies. Le tableau ci-dessous rappelle succinctement les résultats obtenus.

	Bornhuetter-Ferguson		Benktander	
	Compagnie A	Compagnie B	Compagnie A	Compagnie B
$\widehat{CDR}(I+1)$	713	-3 994	-3 174	-6 766
$\sigma(\widehat{CDR}(I+1))$	24 526	17 125	28 865	18 977
- quantile 5 %	40 022	31 945	51 575	38 409
- quantile 0,5 %	63 179	47 958	83 643	57 205
skewness	-0,002	0,021	-0,171	-0,089
kurtosis	0,061	0,003	0,123	-0,014

TABLE 42 – Résultats obtenus pour la méthode bootstrap de re-provisionnement améliorée dans une approche Bornhuetter-Ferguson et Benktander.

11.3 La méthode bootstrap récursive

On applique, dans cette partie, la méthode récursive issue de l'article [4] aux triangles de liquidation A et B. On considère tout d'abord le cas d'un développement complet du triangle supposé répliquer les résultats analytiques de Merz et Wüthrich. Puis un facteur de queue de développement est inséré par la suite dans le cas de la compagnie A.

Le tableau ci-dessous donne les incertitudes du CDR par année de survenance et au global obtenues en appliquant la méthode récursive aux triangles A et B pour 50 000 simulations. On compare également ces résultats à ceux de Merz et Wüthrich déjà présentés en calculant les écarts relatifs qui les séparent.

	Compagnie A			Compagnie B		
	récursive	MW	écart	récursive	MW	écart
2001	-	-	-	-	-	-
2002	16	92	82,61 %	354	370	4,32 %
2003	531	463	14,69 %	617	601	2,66 %
2004	1 347	1 296	3,94 %	1 056	1 039	1,64 %
2005	1 468	1 436	2,23 %	1 260	1 243	1,37 %
2006	1 283	1 253	2,39 %	818	809	1,11 %
2007	2 693	2 659	1,28 %	2 122	2 090	1,53 %
2008	10 043	9 965	0,78 %	6 176	6 140	0,59 %
2009	14 006	13 896	0,79 %	6 481	6 437	0,68 %
2010	15 721	15 616	0,67 %	11 607	11 489	1,03 %
Total	26 385	26 255	0,50 %	16 730	16 477	1,54 %

TABLE 43 – Comparaison entre les résultats obtenus avec la méthode bootstrap récursive et avec la méthode de Merz et Wüthrich.

On constate que les écarts relatifs entre les deux méthodes sont très faibles pour la plupart des résultats, des nuances devant être apportées concernant les écarts des premières années de développement. Les montants relativement bas et un nombre

insuffisant de simulations peuvent en être la cause. Toutefois, il semble bien que l'estimateur proposé par Merz et Wüthrich dans leur article soit répliqué par cette méthode récursive.

On s'intéresse en particulier à la compagnie A pour laquelle un facteur de queue de développement est désormais inséré. Celui-ci est simulé dans chacune des itérations bootstrap par une loi normale de moyenne $\hat{f}_{ult} \approx 1,028699$ et d'écart-type $\hat{\sigma}_{ult} \approx 1,604428e^{-06}$. Le tableau suivant présente l'incertitude à un an obtenue par cette méthode, par année de survenance et au global. Le CDR moyen est par ailleurs égal à 1 257.

	$\sigma(\widehat{CDR}(I+1))$
2001	1
2002	16
2003	548
2004	1 392
2005	1 500
2006	1 319
2007	2 768
2008	10 309
2009	14 254
2010	16 069
Total	29 961

TABLE 44 – Incertitude à un an obtenue avec la méthode récursive et insertion d'un facteur de queue de développement.

On constate que l'insertion d'un facteur de queue de développement dans cette méthode augmente l'incertitude du CDR. C'est, d'une part, une conséquence directe de la volatilité du facteur de queue. D'autre part, l'erreur de processus des règlements cumulés de la diagonale se propage dans le triangle à mesure qu'on le complète. Ainsi, le fait d'ajouter des années de développement supplémentaires grâce à ce facteur de queue augmente la variance du CDR.

L'incertitude à un an

Le tableau suivant synthétise les erreurs de prédictions à horizon un an obtenues pour chacune des méthodes et pour les deux compagnies A et B.

Le premier constat est une grande disparité des incertitudes à un an. La méthode bootstrap de re-provisionnement dite directe fournit l'erreur de prédiction la plus importante ; elle est en effet presque deux fois plus grande que celles déterminées par les autres méthodes. On a précédemment souligné le fait que cette méthode ne vérifiait pas l'hypothèse d'indépendance des variables aléatoires et que, dans chacune des boucles bootstrap, deux estimations du modèle avaient lieu. L'erreur d'estimation des provisions en $t = I + 1$ est donc plus importante, de même que l'erreur de prédiction des règlements effectuées entre les dates $t = I$ et $t = I + 1$. En revanche, la méthode améliorée permet d'obtenir une erreur plus réduite.

		Compagnie A	Compagnie B
Merz	Wüthrich	26 255	16 477
bootstrap de re-provisionnement	direct	50 150	29 423
	amélioré	29 416	13 705
	amélioré avec $\hat{f}_{ult}^{(1)}$	38 814	-
	amélioré - BF ⁽²⁾	24 526	17 125
	amélioré - BE ⁽³⁾	28 865	18 977
bootstrap récursif	récursif	26 385	16 730
	récursif avec $\hat{f}_{ult}^{(1)}$	29 961	-

⁽¹⁾ : on désigne par là l'insertion d'un facteur de queue de développement.

⁽²⁾ : il s'agit de la méthode améliorée avec Bornhuetter-Ferguson.

⁽³⁾ : il s'agit de la méthode améliorée avec Benktander.

TABLE 45 – Synthèse de la vision à un an.

Il semble également que, globalement, les incertitudes à un an obtenues par les méthodes bootstrap sont légèrement supérieures à celles obtenues par une formule fermée, à savoir la méthode Merz et Wüthrich. En effet, dans ce type de méthodes simulatoires, on observe une plus grande variabilité des règlements. De plus, concernant les hypothèses du modèle, s'il existe des résidus atypiques dans les triangles, alors cela aura une forte influence sur la détermination de l'erreur de prédiction. C'est d'ailleurs le cas dans nos triangles dans la mesure où on a pu constater des facteurs de développement atypiques. En revanche, le bootstrap apparaît comme une méthode qui simule bien la réalité de la sinistralité, sauf si on observe des tendances et que l'on aimerait en tenir compte. Elle est à privilégier surtout lors de l'étude d'un triangle stable.

Toutefois, on remarque que la méthode bootstrap récursive réplique bien les résultats de Merz et Wüthrich, ce qui constitue un avantage certain. Comme les hypothèses du modèle Merz Wüthrich sont assez fortes, cela peut être une excellente alternative. De surcroît, la méthode récursive a été adaptée pour tenir compte d'un facteur de queue de développement. On constate que la présence d'un tel facteur n'augmente pas de manière excessive l'incertitude à un an. En effet, on peut considérer que la valeur relativement faible de ce facteur n'impacte pas considérablement l'erreur de prédiction. Avec la méthode récursive, ce facteur est simulé dans chaque boucle bootstrap par une loi normale, tandis qu'il est extrapolé N fois dans la méthode de re-provisionnement. Comme la volatilité du facteur est très importante dans le calcul de l'erreur, cela explique la différence entre les résultats obtenus pour ces deux méthodes bootstrap.

Enfin, en ce qui concerne l'approche Bornhuetter-Ferguson et Benktander, l'erreur de prédiction est du même ordre de grandeur que celles obtenues dans le cadre de Chain Ladder. C'est pourquoi tenir compte d'informations complémentaires aux triangles peut être judicieux. L'actuaire peut apporter ses connaissances de la branche à travers divers indicateurs sans augmenter excessivement l'incertitude à un an.

La comparaison des incertitudes à un an et à l'ultime

Tout d'abord, on compare les incertitudes obtenues par une méthode analytique, à savoir les erreurs de prédiction de Mack [21] et Merz Wüthrich [26].

	Compagnie A	Compagnie B
$MSEP$ Mack	34 220	20 981
$MSEP$ MW	26 255	16 477
$\frac{MSEPMW}{MSEPMack}$	77 %	79%

TABLE 46 – Comparaison des erreurs à l'ultime et à un an.

Le tableau précédent donne les erreurs de prédiction à l'ultime de Mack et à horizon un an de Merz et Wüthrich. On constate que l'incertitude à un an est inférieure. Cela peut paraître cohérent puisque l'incertitude à l'ultime résulte d'une projection à long terme. On pouvait s'attendre à un tel résultat dans la mesure où les deux formules de l'erreur ont un terme en commun mais les autres termes sont inférieurs dans la formule de Merz et Wüthrich car ce sont des diviseurs des termes de la formule de Mack. Le terme en commun correspond à la diagonale suivante du triangle, tandis que les termes suivants se réfèrent aux diagonales suivantes.

L'incertitude à un an contribue dans notre cas à plus de 70 % de l'incertitude à l'ultime. Ainsi, l'exercice comptable suivant, c'est-à-dire la diagonale suivante, impacte beaucoup l'incertitude à l'ultime des provisions. L'étude de l'ASAM-ACME [1] a montré qu'en général l'incertitude à un an était inférieure à l'incertitude ultime d'un tiers. Pour les exercices de survénance les plus éloignés, ce ratio est souvent plus élevé car le sinistre est presque totalement liquidé. Cependant, sur certains triangles très instables, on peut effectuer des ajustements pour ne pas que l'incertitude à un an soit supérieure à la variance de la charge ultime. D'ailleurs, on constate que l'incertitude à un an obtenue par la méthode bootstrap de re-provisionnement directe est nettement supérieure à l'incertitude à l'ultime. On a pu évoquer les raisons d'une telle différence précédemment et la méthode améliorée a résolu cet écueil.

En revanche, si l'on se place dans le cadre des méthodes Bornhuetter-Ferguson et Benktander, les erreurs de prédiction à un an sont supérieures à l'incertitude à l'ultime. Toutefois, pour les années de survénances les plus récentes, on obtient un risque à un an plus faible car on utilise plus l'information donnée a priori. Ces méthodes ne dépendent pas du passé et en l'occurrence de la dernière charge connue.

Finalement, selon le triangle de liquidation et des données dont on dispose, on peut être amené à arbitrer entre une incertitude à horizon un an inférieure à une vision à l'ultime, une incertitude mieux adaptée à la prise en considération de données exogènes ou encore une méthode simulateur ou non. La méthode récursive avec la présence d'un facteur de queue de développement, lorsque cela a un sens, peut apparaître comme une solution répondant à de nombreux critères de satisfaction. Si on s'est intéressé à la mesure de l'incertitude à un an jusqu'ici, le but est de déterminer le capital réglementaire que l'entreprise doit constituer à partir de cette valeur.

12 La détermination du capital réglementaire

Pour déterminer le risque de provisionnement, il faut utiliser le Best Estimate, et non pas la provision globale déterminée par les méthodes précédentes. Pour cela, les paiements futurs ont été actualisés grâce à la courbe des taux swap Euro 2010 fournie par l'EIOPA. La branche considérée, à savoir « Accidents du travail », est de l'assurance non-vie ; c'est pourquoi ces taux ont été complétés par une prime d'illiquidité égale à 50 % des primes d'illiquidité de l'EIOPA. Cette prime permet d'intégrer le fait que les contrats d'assurance sont moins liquides qu'un actif sans risque. On fait l'hypothèse que les règlements ont lieu au 31 décembre de l'année n et le risque de prime n'est pas étudié.

D'autre part, lorsqu'on détermine le risque de provisionnement à l'aide des méthodes Merz Wüthrich et bootstrap présentées précédemment, il s'agit de déterminer également le Best Estimate à la date $t = I + 1$. Dans ce but, les taux forward un an plus tard $f_{1,d}$, pour $d = 1, \dots, 10$ ont été calculés à partir des taux swaps précédents par la relation suivante¹⁷ :

$$f_{1,d} = \left(\frac{(1 + r_{1+d})^{1+d}}{1 + r_1} \right)^{\frac{1}{d}} - 1,$$

avec r_i le taux swap de maturité i années. Les taux utilisés sont présentés dans le tableau suivant.

Maturité	taux swap r_i	taux forward $f_{1,d}$
1	2,299 %	-
2	2,525 %	1,773 %
3	2,859 %	2,188 %
4	3,172 %	2,531 %
5	3,479 %	2,858 %
6	3,672 %	3,032 %
7	3,931 %	3,298 %
8	4,088 %	3,438 %
9	4,216 %	3,548 %
10	4,327 %	3,642 %
11	4,396 %	3,691 %

TABLE 47 – Taux utilisés pour l'actualisation du Best Estimate.

On souhaite tout d'abord déterminer le risque de provisionnement des deux compagnies sur la branche « Accidents du travail » par la formule standard du QIS 5. Cette branche correspond à la *Line of Business* (LoB) « Workers' compensation » dont le paramètre de volatilité σ_{LoB} est fourni par le QIS 5 et vaut 10 %.

En revanche, si on veut utiliser l'erreur de prédiction à un an $MSEP_{MerzWüthrich}^{1/2}$ obtenue par la méthode Merz Wüthrich, on estime le paramètre de volatilité σ_{LoB} par le ratio suivant :

$$\frac{MSEP_{MerzWüthrich}^{1/2}}{\text{Best Estimate}}.$$

17. cf Finance de marché [30].

On obtient les volatilités $\hat{\sigma}_A = 9,5\%$ et $\hat{\sigma}_B = 11,5\%$ pour les compagnies A et B. On constate que ces valeurs sont très proches de celle utilisée dans la formule standard. Toutefois, la valeur fournie par le QIS 5 semble être sous-estimée si on se place du point de vue de la compagnie B.

Pour ces deux méthodes, le $SCR_{\text{provisions}}$ pour cette *LoB* est ensuite obtenu à l'aide de la relation suivante :

$$SCR_{\text{provisions}} = \rho(\sigma_{LoB}) \times \text{Best Estimate},$$

la fonction ρ ayant été définie précédemment.

Les Best Estimate obtenus à la date $t = I$ sont respectivement de 276 904 et 143 685 pour les compagnies A et B.

Enfin, si on utilise les méthodes par simulations bootstrap pour obtenir le risque de provisionnement, il suffit de déterminer l'opposé du quantile à 0,5 % du CDR désormais défini par :

$$CDR = BE_I - \frac{BE_{I+1} + P_{(I,I+1]}}{1 + r_1},$$

où BE_I et BE_{I+1} désignent les Best Estimate calculés aux dates $t = I$ et $t = I + 1$ respectivement, et $P_{(I,I+1]}$ les paiements effectués durant l'année comptable $(I, I + 1]$.

Le tableau ci-après présente les résultats obtenus pour le risque de provisionnement et pour chacune des méthodes évoquées.

		Compagnie A	Compagnie B
formule standard		79 496	41 250
Merz Wüthrich		74 980	48 038
bootstrap de re-provisionnement	direct	145 865	81 138
	amélioré	81 702	37 057
	amélioré avec $\hat{f}_{ult}^{(1)}$	109 570	-
	amélioré - BF ⁽²⁾	50 714	28 351
	amélioré - BE ⁽³⁾	64 766	36 489
bootstrap récursif	récursif	78 416	51 429
	récursif avec $\hat{f}_{ult}^{(1)}$	81 033	-

(1) : on désigne par là l'insertion d'un facteur de queue de développement.

(2) : il s'agit de la méthode améliorée avec Bornhuetter-Ferguson.

(3) : il s'agit de la méthode améliorée avec Benktander.

TABLE 48 – Résultats obtenus pour le risque de provisionnement selon les différentes méthodes et la compagnie considérée.

On observe une très grande différence entre les SCR obtenus pour les deux compagnies. On s'attendait à de tels résultats dans la mesure où la détermination de la volatilité à un an offrait déjà de telles différences. On remarque que les SCR obtenus se situent dans un certain intervalle autour du SCR calculé à l'aide de la

formule standard fournie par le QIS 5. Certains SCR sont ainsi bien inférieurs au SCR standard ; c'est le cas notamment du SCR obtenu grâce à la méthode bootstrap de re-provisionnement améliorée appliquée aux méthodes Bornhuetter-Ferguson et Benktander. Ainsi, tenir compte d'informations supplémentaires semble permettre de mieux appréhender le risque et de réduire le SCR. La méthode de Merz-Wüthrich, qui apparaît déjà comme la méthode de référence, permet d'obtenir un SCR proche de la formule standard, tout comme la méthode récursive répliquant les résultats de Merz et Wüthrich.

En revanche, le SCR obtenu par la méthode de re-provisionnement directe est presque égal au double du SCR donné par la formule standard, ce qui constitue un inconvénient majeur de cette méthode. En effet, si une compagnie utilise cette méthode pour calculer l'incertitude à un an des provisions, cela signifie qu'elle devra affecter au SCR un montant deux fois plus important qu'avec les autres méthodes. En ce qui concerne la présence d'un facteur de queue de développement, les SCR obtenus sont supérieurs aux précédents, mais la méthode récursive permet d'obtenir un résultat bien plus faible que la méthode de re-provisionnement améliorée ; c'est pourquoi la première méthode semble être à privilégier.

Finalement, la méthode de provisionnement à un an que l'organisme assureur doit utiliser en priorité dépend nécessairement du montant du SCR qu'il juge le plus approprié. Il s'agit à la fois de respecter le principe de prudence à l'origine même de la définition du SCR, et d'effectuer un arbitrage entre les méthodes selon les critères à privilégier, comme un SCR proche de la formule standard, l'insertion d'un facteur de queue ou encore l'influence de données exogènes au triangle de liquidation.

Conclusion

Les méthodes de provisionnement évoluent beaucoup et offrent aujourd'hui une panoplie de possibilités à l'actuaire pour exercer son jugement d'expert. Celui-ci choisit une méthode en fonction de son expérience et des informations dont il dispose et qu'il souhaite valoriser. On a ainsi pu constater lors de l'application sur les données en assurance non-vie qu'un arbitrage doit être fait entre les méthodes ; les provisions pour sinistres à payer et les incertitudes à l'ultime que l'on obtient selon les modèles sont en effet très éparses. Il s'agit donc pour l'actuaire de sélectionner les critères à privilégier, comme le fait de constituer une provision prudente et appropriée à la réalité grâce à l'insertion d'un facteur de queue de développement, la prise en compte de données exogènes, ou encore le choix des ratios sinistres à primes attendus. Il peut également choisir d'opter pour une méthode par simulations pour les nombreux avantages qu'elle comporte. Dans l'application de ce mémoire, la méthode qui paraissait la plus appropriée était celle d'Alai, Merz et Wüthrich dans l'approche Bornhuetter-Ferguson ; en effet, l'erreur de prédiction obtenue semblait raisonnable et adéquate, sans compter que les hypothèses dans l'approche Chain Ladder invitaient à la prudence.

Si le dessein de la détermination de l'incertitude à l'ultime des provisions est assez intuitif pour les assureurs, la notion d'incertitude à un an l'est beaucoup moins. Il ne s'agit pas de la volatilité des règlements effectués pendant la prochaine année comptable, mais de l'incertitude liée au boni/mali de l'assureur, soit la révision de l'estimateur des provisions. Une attention toute particulière doit donc être portée à cette nouvelle notion. Plusieurs méthodes ont ainsi été développées dans ce mémoire pour déterminer une telle incertitude. Celle de Merz et Wüthrich est désormais une référence en la matière, mais diverses méthodes bootstrap ont également montré leurs atouts. Des données exogènes sont facilement exploitables dans ces modèles, de même qu'un facteur de queue de développement est aisément inséré. La méthode dite récursive permet en particulier d'allier facteur de queue de développement, obtention de la distribution du CDR et réplcation des résultats de référence Merz et Wüthrich. A noter également que cette méthode bootstrap est applicable à différents modèles sous-jacents.

L'incertitude à horizon un an est ainsi une notion assez complexe. Initialement, la Directive Solvabilité II était supposée donner comme objectif aux assureurs de mieux maîtriser leurs risques dans le but de respecter absolument leurs engagements vis-à-vis des assurés. Mais Charpentier et al. [11] mettent en garde avec justesse sur le fait que cet objectif risque d'être supplanté par l'envie de minimiser l'incertitude sur les boni/mali, les intéressés étant alors les directeurs financiers et non plus les assurés eux-mêmes.

Il existe certaines difficultés techniques dans l'application de la Directive Solvabilité II, mais cette réforme a surtout une dimension politique. En effet, elle est supposée financer l'économie en réduisant notamment l'exposition aux risques de volatilité à court terme des assureurs. A en croire J.H. Lorenzi [19], l'impact macroéconomique de Solvabilité II aurait été sous-estimé jusqu'à présent alors que la réforme se place au cœur de la question de la croissance.

Enfin, on pourra noter que les modèles bayésiens n'ont pas été étudiés dans ce

mémoire dans la mesure où l'objectif n'était pas de proposer une liste exhaustive des méthodes existantes. En revanche, ces modèles sont pertinents et mériteraient un mémoire à eux seuls. L'incertitude à horizon un an que l'on a déterminée dans ce mémoire a ainsi permis de calculer le capital économique exigé par la Directive à des fins de solvabilité. Mais les calculs ont été effectués bruts de réassurance. Il serait alors intéressant d'étudier l'impact de la réassurance sur la détermination de ce capital économique.

Remerciements

Je tiens à remercier en premier lieu M. Michel PIERMAY pour m'avoir acceptée dans son entreprise et permis de me consacrer à ce mémoire durant mon année d'apprentissage. Je lui témoigne également ma gratitude pour l'attention qu'il a pu porter à mon travail.

Je remercie également mon tuteur en entreprise M. Emmanuel TASSIN pour sa disponibilité et la confiance qu'il m'a accordée durant les missions.

Je tiens également à remercier l'ensemble des consultants de FIXAGE et tout particulièrement M. Hadi BOUOKBA pour son écoute, son attention et pour tout l'intérêt qu'il a porté à ce mémoire.

Enfin, je remercie mes professeurs de l'Institut de Sciences Financière et d'Assurance (ISFA) et en particulier mon tuteur universitaire M. Christian ROBERT pour son attention et son amabilité.

Références

- [1] AISAM/ACME Study on non-life long tail liabilities - Reserve risk and risk margin assessment under Solvency II (2007), *aisam.org*.
- [2] ALAI, D.H., MERZ, M., WÜTHRICH, M. V. (2009). Mean square Error of Prediction in the Bornhuetter-Ferguson Claims Reserving Method, *Annals of Actuarial Science* 4, no.1, 7-31.
- [3] ALAI, D.H., MERZ, M., WÜTHRICH, M. V. (2010). Prediction Uncertainty in the Bornhuetter-Ferguson Claims Reserving Method: Revisited, *Annals of Actuarial Science* 5, no.1, 7-17.
- [4] ANGOUA, Y., BOISSEAU, J.P., BOUMEZOUED, A., DEVINEAU, L. (2011). One-year reserve risk including a tail factor: closed formula and bootstrap approaches, *Laboratoire de Sciences Actuarielle et Financière*.
- [5] BENKTANDER, G. (1976). An approach to credibility in calculating IBNR for casualty excess reinsurance, *The Actuarial Review*, April, 7.
- [6] BOOR, J. (2006). Estimating Tail Development Factors: What to do When the Triangle Runs Out, *Casualty Actuarial Society Forum*, 345-390.
- [7] BORNHUETTER, R.L., FERGUSON, R.E. (1972). The actuary and IBNR, *Proc Casualty Actuarial Society*, Vol. LIX, 181-195.
- [8] BOUOKBA, H. (2010). Les méthodes statistiques de provisionnement utilisées en assurance dommage, *Intervention IMA, FIXAGE Actuariat*.
- [9] BOUTHIER, D., HIPPOLYTE, H. (2012). Mesure du risque de provisionnement à un an en assurance IARD, *Mémoire d'actuariat, ENSAE*.
- [10] CHARPENTIER, A., DENUIT, M. (2005). *Mathématiques de l'Assurance non-vie, Tome II : tarification et provisionnement*, Economica.
- [11] CHARPENTIER, A., DEVINEAU, L., NESSI, J.M. (2010). Mesurer le risque lors du calcul des provisions pour sinistres à payer, *Revue Risques*, no. 83.
- [12] DE FELICE, M., MORICONI, F. (2006). Process Error and Estimation Error of Year-End Reserve Estimation in the Distribution Free Chain-Ladder Model, *Alef Working Paper, Version B, Rome*.
- [13] DIERS, D. (2009). Stochastic re-reserving in multi-year internal models - An approach based on simulations, *ASTIN Colloquium in Helsinki*.
- [14] ENGLAND, P.D., VERRALL, R.J. (2002). Stochastic Claims Reserving in General Insurance, *Institute of Actuaries and Faculty of Actuaries*.
- [15] FIXAGE (2010). La mise en place d'un modèle interne en dommages conforme aux Consultation Papers, *Formation FIXAGE*.
- [16] JAZIRI, S. (2011). Méthodes de provisionnement non-vie et risque de réserve à un an, *Mémoire d'actuariat, ISFA*.
- [17] LACOUME, A. (2008). Mesure du risque de réserve sur un horizon de un an, *Mémoire d'actuariat, ISFA*.
- [18] LECOEUR, E., NESSI, J.M., NISIPASU, E., PARTRAT, C., REIZ, O. (2007). *Provisionnement technique en Assurance non-vie*, Economica.
- [19] LORENZI, J.H. (2011). *Article paru dans la Revue Risques*, no.85.
- [20] LYSENKO, N., MERZ, M., WÜTHRICH, M. V. (2007). Uncertainty of the Claims Development Result in the Chain Ladder Method, *Scandinavian Actuarial Journal*.
- [21] MACK, T. (1993). Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates, *Astin Bulletin*, Vol 23, no. 2, 213-225.

- [22] MACK, T. (1999). The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor, *Astin Bulletin*, Vol 29, no. 2, 361-366.
- [23] MACK, T. (2000). Credible Claims Reserves: the Benktander Method, *Astin Bulletin*, Vol 30, no. 2.
- [24] MACK, T. (2008). The Prediction Error of Bornhuetter-Ferguson, *Casualty Actuarial Society E-Forum*, Fall 2008, 222-240.
- [25] MERZ, M., WÜTHRICH, M. V. (2007). Prediction Error of the Expected Claims Development Result in the Chain Ladder Method, *North American Actuarial J.*
- [26] MERZ, M., WÜTHRICH, M. V. (2008). Modelling the Claims Development Result for Solvency Purposes, *Casualty Actuarial Society Forum*, 542-568.
- [27] NESSI, J.M. (2008) Méthodes de provisionnement en assurance non-vie, *formation SEPIA*.
- [28] OHLSSON, E., LAUZENINGKS, J. (2008). The one-year non-life insurance risk, *ASTIN Colloquium in Manchester*.
- [29] PLANCHET, F. (2011). bootstrap et méthodes de triangles, actudactulaires.typepad.com/laboratoire/2011/07/bootstrap-et-méthodes-de-triangles.html.
- [30] PONCET, P., PORTRAIT, R. (2009). *Finance de marché - Instruments de base, produits dérivés, portefeuilles et risques*, 2^e édition, Dalloz.
- [31] RENSHAW, A.E., VERRALL, R.J. (1998). A Stochastic Model Underlying the Chain-Ladder Technique, *B.A.J.* 4, IV, 903-923.
- [32] ROBERT, C.Y. (2011). Market Value Margin calculations under the Cost of Capital approach within a Bayesian chain ladder framework.
- [33] SPITZ, B. (2011). Solvabilité II, cinq conditions, *Revue Risques*, no. 85.
- [34] VOGT, A. (2010). Modélisation stochastique en assurance non vie : calcul de l'exigence de capital réglementaire sous Solvency II, *Mémoire d'actuariat, UFR-FSEG*.

Annexes

A. La méthode Delta

La méthode Delta permet d'obtenir une expression analytique asymptotique de la variance d'un estimateur. Si on dispose d'un estimateur asymptotiquement gaussien, toute transformation bijective de cet estimateur reste asymptotiquement gaussienne.

Théorème :

Soit $(X_n)_{n \geq 0}$ une suite de vecteurs aléatoires de $\mathbb{R}^p$ asymptotiquement normaux de moyenne μ et de matrice de variance-covariance $d_n^2 \Sigma^2$ avec d_n qui tend vers 0 lorsque n tend vers l'infini, et Σ^2 semi-définie positive.

On note $h : \begin{cases} \mathbb{R}^p \rightarrow \mathbb{R}^q \text{ une application différentiable dans un voisinage de } \mu, \\ x \rightarrow (h_1(x), \dots, h_q(x)) \end{cases}$

et $(H_{i,j})_{\substack{1 \leq i \leq q \\ 1 \leq j \leq p}}$ la matrice telle que $H_{i,j} = \left. \frac{\partial h_i}{\partial x_j} \right|_{x=\mu}$.

Si tous les éléments diagonaux de la matrice $H \Sigma^2 H'$ sont non nuls, alors :

$$\frac{1}{d_n} (h(X_n) - h(\mu)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, H \Sigma^2 H').$$

Estimation de la variance du facteur de queue de développement $\hat{f}_{ult}$ ¹⁸

On souhaite estimer la variance du facteur de queue de développement. On suppose que celui-ci a été estimé par régression :

$$\hat{f}_{ult} = f(\hat{a}, \hat{b}) = \prod_{j \geq J} \hat{f}_j,$$

où $\hat{a}$ et $\hat{b}$ sont des estimateurs obtenus par maximum de vraisemblance.

$$\text{Soient } Y = \begin{pmatrix} \log(\hat{f}_o - 1) \\ \vdots \\ \log(\hat{f}_{I-1} - 1) \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ \vdots & \vdots \\ I-1 & 1 \end{pmatrix}, \text{ et } Z = \begin{pmatrix} a \\ b \end{pmatrix}.$$

L'estimateur par maximum de vraisemblance de Z est $\hat{Z} = (X'X)^{-1}X'Y$.

La variance du facteur de queue de développement extrapolé s'écrit :

$$\mathbb{V}ar(\hat{f}_{ult}) = \mathbb{V}ar\left(\prod_{j=1}^{I_{ult}-1} \hat{f}_j\right) = \mathbb{V}ar\left(\prod_{j=1}^{I_{ult}-1} (1 + \exp(\hat{a} \times j + \hat{b}))\right) = \mathbb{V}ar(H_{I_{ult}-1}(\hat{Z})),$$

$$\text{avec } H_{I_{ult}-1}(\hat{Z}) = \prod_{j=1}^{I_{ult}-1} (1 + \exp(\hat{a} \times j + \hat{b})).$$

La méthode Delta permet d'approximer la variance de $\hat{f}_{ult}$ par :

$$\widehat{\mathbb{V}ar}(\hat{f}_{ult}) \approx (\nabla H_{I_{ult}-1}(\hat{Z}))' \widehat{\Sigma}(\hat{Z}) \nabla H_{I_{ult}-1}(\hat{Z}),$$

18. On se réfère à l'article *One-year reserve risk including a tail factor: closed formula and bootstrap approaches* d'Angoua, Boisseau, Boumezoued et Devineau pour cette démonstration.

où $\widehat{\Sigma(\hat{Z})}$ est un estimateur de la matrice de variance-covariance de $\hat{Z}$. Comme l'estimateur du maximum de vraisemblance $\hat{Z}$ est efficace, $\widehat{\Sigma(\hat{Z})}$ est l'inverse de l'estimation de la matrice d'information de Fisher $\hat{I} : \widehat{\Sigma(\hat{Z})} = \hat{I}^{-1}$. Cette matrice $\hat{I}^{-1}$ est telle que :

$$\hat{I}^{-1} = \hat{\sigma}^2(X'X)^{-1}, \text{ avec } \hat{\sigma}^2 = \frac{1}{I} \sum_{j=1}^{J-1} (\hat{f}_j - \hat{a} \times j - \hat{b})^2$$

l'estimateur biaisé de la variance des résidus.

Le facteur de queue de développement extrapolé s'écrit :

$$H_{I_{ult}-1}(a, b) = \prod_{j=I}^{I_{ult}-1} (1 + \exp(a \times j + b)).$$

Pour simplifier les formules, une expression récursive de $\nabla H_{I_{ult}-1}(a, b)$ est proposée :

Dérivée partielle par rapport à a :

Soit $j \geq 2$. $H_{I+j}(a, b) = (1 + e^{a(I+j)+b})H_{I+j-1}(a, b)$. On peut alors écrire :

$$\frac{\partial H_{I+j}}{\partial a}(a, b) = (I+j)e^{a(I+j)+b}H_{I+j-1}(a, b) + (1 + e^{a(I+j)+b})\frac{\partial H_{I+j-1}}{\partial a}(a, b),$$

avec $\frac{\partial H_I}{\partial a}(a, b) = Ie^{aI+b}$.

Dérivée partielle par rapport à b :

Soit $j \geq 2$. On peut écrire :

$$\frac{\partial H_{I+j}}{\partial b}(a, b) = (I+j)e^{a(I+j)+b}H_{I+j-1}(a, b) + (1 + e^{a(I+j)+b})\frac{\partial H_{I+j-1}}{\partial b}(a, b),$$

avec $\frac{\partial H_I}{\partial b}(a, b) = e^{aI+b}$.

Ces deux derniers résultats permettent d'avoir l'expression récursive du gradient de la fonction $H_{I_{ult}-1}$:

$$\nabla H_{I_{ult}-1}(a, b) = \begin{pmatrix} \frac{\partial H_{I_{ult}-1}}{\partial a}(a, b) \\ \frac{\partial H_{I_{ult}-1}}{\partial b}(a, b) \end{pmatrix}.$$

On peut alors approximer la variance du facteur de queue de développement $\hat{f}_{ult}$ par la méthode Delta :

$$\widehat{\text{Var}}(\hat{f}_{ult}) \approx (\nabla H_{I_{ult}-1}(\hat{Z}))' \widehat{\Sigma(\hat{Z})} \nabla H_{I_{ult}-1}(\hat{Z}).$$

On peut utiliser la fonction *deltamethod* sous le logiciel R pour utiliser la méthode Delta.

B. La méthode d'Alai, Merz et Wüthrich

• L'erreur de prédiction conditionnelle

L'erreur de prédiction conditionnelle de l'estimateur $\hat{C}_{i,j}^{BF}$ pour chaque année de survenance $i \in \{1, \dots, I\}$ est définie par :

$$\begin{aligned}
 MSEP_{C_{i,I}|\mathcal{D}_I}(\hat{C}_{i,I}^{BF}) &= \mathbb{E}[(\hat{C}_{i,I}^{BF} - C_{i,I})^2 | \mathcal{D}_I] \\
 &= \mathbb{E} \left[\left(C_{i,I-i} - C_{i,I} + \hat{\mu}_i \sum_{j>I-i+1} \hat{\gamma}_j \right)^2 \middle| \mathcal{D}_I \right] \\
 &= \mathbb{E} \left[\left(\sum_{j>I-i+1} X_{i,j} + \hat{\mu}_i \sum_{j>I-i+1} \hat{\gamma}_j \right)^2 \middle| \mathcal{D}_I \right] \\
 &= \sum_{j>I-i+1} \text{Var}(X_{i,j}) + \mathbb{E} \left[\left(\sum_{j>I-i+1} \mathbb{E}[X_{i,j}] - \hat{\mu}_i \left(\sum_{j>I-i+1} \hat{\gamma}_j \right) \right)^2 \middle| \mathcal{D}_I \right] \\
 &\quad + \underbrace{2\mathbb{E} \left[\left(\sum_{j>I-i+1} (X_{i,j} - \mathbb{E}[X_{i,j}]) \right) \left(\sum_{j>I-i+1} \mathbb{E}[X_{i,j}] - \hat{\mu}_i \left(\sum_{j>I-i+1} \hat{\gamma}_j \right) \right) \middle| \mathcal{D}_I \right]}_{\text{terme A}}.
 \end{aligned}$$

On sait que:

- $\hat{\mu}_i$ est indépendant de $X_{k,j}$ pour tout k, j ,
- $\sum_{j>I-i+1} \hat{\gamma}_j$ est $\mathcal{D}_I$ -mesurable,
- $\mathbb{E}[\hat{\mu}_i] = \mu_i$.

On en déduit que le terme A disparaît et l'égalité suivante est obtenue :

$$MSEP_{C_{i,I}|\mathcal{D}_I}(\hat{C}_{i,I}^{BF}) = \sum_{j>I-i+1} \text{Var}(X_{i,j}) + \left(\sum_{j>I-i+1} \hat{\gamma}_j \right)^2 \text{Var}(\hat{\mu}_i) + \mu_i^2 \left(\sum_{j>I-i+1} \hat{\gamma}_j - \sum_{j>I-i+1} \gamma_j \right)^2.$$

• L'erreur de prédiction

Les deux premiers termes de l'erreur de prédiction conditionnelle $MSEP_{C_{i,I}|\mathcal{D}_I}(\hat{C}_{i,I}^{BF})$ peuvent être estimés en remplaçant les inconnues par leurs estimations. En revanche, le terme $\mu_i^2 \left(\sum_{j>I-i+1} \hat{\gamma}_j - \sum_{j>I-i+1} \gamma_j \right)^2$ est estimé par l'espérance inconditionnelle :

$$\mathbb{E} \left[\left(\sum_{j>I-i+1} (\hat{\gamma}_j - \gamma_j) \right)^2 \right] = \sum_{\substack{j>I-i+1 \\ l>I-i+1}} \mathbb{E} \left[(\hat{\gamma}_j - \gamma_j)(\hat{\gamma}_l - \gamma_l) \right].$$

En ne tenant pas compte du fait que les estimations obtenues par maximum de vraisemblance peuvent être biaisées, l'approximation suivante est possible :

$$\sum_{\substack{j>I-i+1 \\ l>I-i+1}} \mathbb{E} \left[(\hat{\gamma}_j - \gamma_j)(\hat{\gamma}_l - \gamma_l) \right] \approx \sum_{\substack{j>I-i+1 \\ l>I-i+1}} \text{Cov}(\hat{\gamma}_j, \hat{\gamma}_l).$$

Pour estimer la covariance, la propriété asymptotique du maximum de vraisemblance suivante est utilisée :

$$\sqrt{n}(\hat{\zeta} - \zeta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, H(\zeta, \phi)^{-1}) \text{ lorsque } n \text{ tend vers } \infty,$$

avec $H(\zeta, \phi) = (h_{r,s}(\zeta, \phi))_{r,s=1,\dots,m}$ la matrice d'information de Fisher telle que :

$$h_{r,s} = -\mathbb{E}_{\zeta} \left[\frac{\partial^2}{\partial \zeta_r \partial \zeta_s} l_{\mathcal{D}_I}(\zeta, \phi) \right],$$

avec

$$\begin{cases} \zeta = (\zeta_1, \dots, \zeta_{2I-1}) = (\mu_1, \dots, \mu_I, \gamma_1, \dots, \gamma_{I-1}), \\ \hat{\zeta} \text{ l'estimateur par maximum de vraisemblance,} \\ l_{\mathcal{D}_I}(\zeta, \phi) = l_{\mathcal{D}_I}(\mu_i, \gamma_j, \phi) \text{ la log-vraisemblance des données.} \end{cases}$$

$$\begin{aligned} l_{\mathcal{D}_I}(\mu_i, \gamma_j, \phi) &= \sum_{\substack{i+j \leq I+1 \\ j < i}} \left(\frac{1}{\phi} (X_{i,j} \log(\mu_i \gamma_j) - \mu_i \gamma_j) + \log c(X_{i,j}, \phi) \right) \\ &\quad + \left(\frac{1}{\phi} \left(X_{1,I} \log \left[\mu_1 \left(1 - \sum_{n=1}^{J-1} \gamma_n \right) \right] - \mu_1 \left(1 - \sum_{n=1}^{J-1} \gamma_n \right) \right) + \log c(X_{1,I}, \phi) \right), \end{aligned}$$

avec $c(\cdot, \phi)$ une fonction qui ne dépend pas des paramètres.

La matrice d'information de Fisher H est définie ainsi :

$$\begin{aligned} h_{i,i} &= \frac{\mu_i^{-1}}{\phi} \sum_{j=1}^{J-i+1} \gamma_j, & i \in \{1, \dots, I\}, \\ h_{I+j, I+j} &= \frac{\gamma_j^{-1}}{\phi} \sum_{i=1}^{I-j+1} \mu_i + \frac{\mu_1}{\phi(1 - \sum_{n=1}^{I-1} \gamma_n)}, & j \in \{1, \dots, I-1\}, \\ h_{I+j, I+l} &= \frac{\mu_1}{\phi(1 - \sum_{n=1}^{I-1} \gamma_n)}, & j, l \in \{1, \dots, I-1\}, j \neq l, \\ h_{i, I+j} &= \frac{1}{\phi}, & i \in \{2, \dots, I\}, j \in \{1, \dots, I-i+1\}, \\ h_{I+j, i+1} &= \frac{1}{\phi}, & j \in \{1, \dots, I-1\}, i \in \{1, \dots, I-j\}. \end{aligned}$$

Les autres éléments de la matrice de taille $(2I-1) \times (2I-1)$ sont nuls.

La matrice inverse d'information de Fisher est ensuite simplifiée en retirant l'information sur les μ_i qui n'est pas nécessaire. La nouvelle matrice de taille $I \times I$ $\mathcal{G} = (g_{j,l})_{j,l=1,\dots,I}$ est définie par :

$$\begin{aligned} g_{j,l} &= \widehat{\text{Cov}}(\hat{\gamma}_j, \hat{\gamma}_l) = H(\hat{\zeta}, \hat{\phi})_{I+2+j, I+2+l}^{-1}, & j, l \in \{1, \dots, I-1\}, \\ g_{j,I} &= g_{I,j} = \widehat{\text{Cov}}(\hat{\gamma}_j, \hat{\gamma}_I) = - \sum_{m=1}^{I-1} H(\hat{\zeta}, \hat{\phi})_{I+2+j, I+2+m}^{-1}, & j \in \{1, \dots, I-1\}, \\ g_{I,I} &= \widehat{\text{Var}}(\hat{\gamma}_I) = \sum_{\substack{1 \leq m \leq I-1 \\ 1 \leq n \leq I-1}} H(\hat{\zeta}, \hat{\phi})_{I+2+m, I+2+n}^{-1}. \end{aligned}$$

C. La méthode bootstrap de De Felice et Moriconi

De Felice et Moriconi proposent dans leur article *Process Error and Estimation Error of Year-End Reserve Estimation in the Distribution Free Chain Ladder* (2006) une méthode bootstrap qui repose sur la version récursive du modèle Chain Ladder. On se propose de détailler cette procédure bootstrap :

1. Estimer les facteurs de développement $\hat{f}_j$, $j \in \{1, \dots, J-1\}$, et les variances de ces estimateurs $\hat{\sigma}_j$.
2. Déterminer le triangle supérieur des facteurs de développement individuels définis par :

$$F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}, \text{ pour } 2 \leq i+j \leq I-1.$$

3. Calculer les résidus en fonction des facteurs de développement individuels tels que :

$$\epsilon_{i,j} = \frac{\sqrt{C_{i,j-1}}(F_{i,j-1} - \hat{f}_{j-1})}{\hat{\sigma}_{j-1}}.$$

Ces résidus sont ensuite ajustés :

$$\bar{\epsilon}_{i,j} = \frac{\epsilon_{i,j} - \mu_\epsilon}{\sigma_\epsilon}$$

avec

$$\begin{aligned} n &= \frac{(I-1)I}{2} \text{ le nombre de résidus,} \\ \mu_\epsilon &= \frac{1}{n} \sum_{i=1}^{I-1} \sum_{j=2}^{I-i+1} \epsilon_{i,j}, \\ \sigma_\epsilon^2 &= \frac{1}{n} \sum_{i=1}^{I-1} \sum_{j=2}^{I-i+1} (\epsilon_{i,j} - \mu_\epsilon)^2. \end{aligned}$$

4. Répéter N fois les étapes suivantes :

- Effectuer un rééchantillonnage aléatoire avec remise des résidus ajustés. On obtient un nouveau triangle de résidus $\tilde{\epsilon}_{i,j}^{(k)}$ pour la k^{ieme} boucle bootstrap.
- Reconstituer le pseudo-triangle des facteurs de développement individuels à la date $t = I$ par la relation suivante :

$$\tilde{F}_{i,j-1}^{k,I} = \tilde{\epsilon}_{i,j}^{(k)} \sqrt{\frac{\hat{\sigma}_{j-1}^2}{C_{i,j-1}}} + \hat{f}_{j-1}, \text{ pour } 2 \leq i+j \leq I-1.$$

- Calculer les facteurs de développement $\tilde{f}_{j-1}^{k,I}$ à la date $t = I$ à partir des nouveaux facteurs de développement individuels :

$$\tilde{f}_{j-1}^{k,I} = \frac{\sum_{i=1}^{I-j+1} C_{i,j-1} \tilde{F}_{i,j-1}^{k,I}}{\sum_{i=1}^{I-j} C_{i,j-1}}, \text{ pour } j \in \{2, \dots, J\}.$$

- Déterminer les règlements de l'année $(I, I+1]$ correspondant à la diagonale suivante du triangle de liquidation initial de manière récursive à partir des paramètres $\tilde{f}_j^{k,I}$ et $\hat{\sigma}_j$:

$$\tilde{C}_{i,J-i+2} = \tilde{f}_{J-i+1}^{k,I} C_{i,J-i+1} + \hat{\sigma}_{J-i+1} \sqrt{C_{i,J-i+1}} \epsilon_{i,J-i+2}, \text{ pour } i \in \{2, \dots, I\}.$$

- Déterminer ensuite le triangle inférieur des règlements cumulés :

$$\tilde{C}_{i,j} = \tilde{f}_{j-1}^{k,I} C_{i,j-1} + \hat{\sigma}_{j-1} \sqrt{\tilde{C}_{i,j-1}} \epsilon_{i,j}, \text{ pour } i \in \{3, \dots, I\} \text{ et } j \in \{J-i+3, \dots, J\}.$$

Certaines hypothèses techniques doivent être envisagées pour les résidus $\epsilon_{i,j}$. Il est nécessaire de faire un grand nombre de simulations afin d'assurer une certaine fiabilité du résultat.

- Déterminer les facteurs de développement $\tilde{f}_j^{k,I+1}$ à la fin de l'année $(I, I+1]$ grâce à la relation suivante :

$$\tilde{f}_{j-1}^{k,I+1} = \frac{\sum_{i=1}^{I-j+1} C_{i,j} + \tilde{C}_{i,I-j+2}}{\sum_{i=1}^{I-j+2} C_{i,j-1}}, \text{ pour } j \in \{2, \dots, J\}.$$

- Déterminer la charge ultime $\tilde{C}_{i,J}^{(k,I+1)}$ en appliquant Chain Ladder au triangle élargi à la nouvelle diagonale bootstrap :

$$\tilde{C}_{i,J}^{(k,I+1)} = \tilde{C}_{i,J-i+2} \tilde{f}_{J-i+2}^{k,I+1} \dots \tilde{f}_{J-1}^{k,I+1}, \text{ pour } i \in \{3, \dots, I\}.$$

- Estimer la provision $\hat{R}^{(k)}$ par la méthode Chain Ladder.
- Calculer le CDR^k par la relation suivante :

$$CDR^k = \hat{R} - \tilde{P}^{k,I+1} - \hat{R}^{(k)}.$$

5. Obtenir les distributions empiriques de l'estimation de la provision et du CDR.

Une distribution empirique des provisions est obtenue en combinant toutes les itérations, l'erreur de prédiction est estimée comme l'écart-type des provisions simulées. L'erreur de processus est obtenue en simulant seulement les résidus $\epsilon_{i,j}$ sans bootstrap, et l'erreur d'estimation en considérant uniquement la variabilité du bootstrap sans ajouter les termes d'erreur $\epsilon_{i,j}$.

D. L'équivalence des résultats de la méthode récursive avec ceux de la méthode analytique de Merz et Wüthrich

Les auteurs ont montré que la procédure du bootstrap récursif permet d'obtenir la même erreur de prédiction que celle du CDR observable proposée par Merz et Wüthrich [26].

Erreur d'estimation

Soit $\widehat{CDR}_i$ le CDR ne tenant compte que de l'erreur d'estimation. Pour $i \in \{2, \dots, I\}$, $C_{i,I-i+2}^k = \tilde{f}_{I-i+1}^{(k,I)} C_{i,I-i+1}$. La variance du CDR s'écrit :

$$\mathbb{V}ar(\widehat{CDR}_i) = \mathbb{V}ar\left(C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - \tilde{f}_{I-i+1}^{(k,I)} C_{i,I-i+1} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)}\right).$$

Pour $j \in \{1, \dots, J-1\}$, on a les égalités suivantes :

$$\begin{cases} \mathbb{E}[\tilde{f}_j^{(k,I)}] = \hat{f}_j, \\ \mathbb{E}[(\tilde{f}_j^{(k,I)})^2] = \mathbb{V}ar(\tilde{f}_j^{(k,I)}) + \hat{f}_j^2 = \frac{\hat{\sigma}_j^2}{S_j^I} + \hat{f}_j^2, \\ \tilde{f}_j^{(k,I+1)} = \frac{S_j^I}{S_j^{I+1}} \hat{f}_j + \tilde{f}_j^{(k,I)} \frac{C_{I-j+1,j}}{S_j^{I+1}}, \\ \mathbb{V}ar(\tilde{f}_j^{(k,I+1)}) = \left(\frac{C_{I-j+1,j}}{S_j^{I+1}}\right)^2 \mathbb{V}ar(\tilde{f}_j^{(k,I)}) = \left(\frac{C_{I-j+1,j}}{S_j^{I+1}}\right)^2 \frac{\hat{\sigma}_j^2}{S_j^I}, \\ \mathbb{E}[(\tilde{f}_j^{(k,I+1)})^2] = \left(\frac{C_{I-j+1,j}}{S_j^{I+1}}\right)^2 \frac{\hat{\sigma}_j^2}{S_j^I} + \hat{f}_j^2. \end{cases}$$

En utilisant l'indépendance des pseudo-facteurs de développement pour différentes années de développement, on obtient :

$$\begin{aligned} & \mathbb{E}\left[C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - \tilde{f}_{I-i+1}^{(k,I)} C_{i,I-i+1} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)}\right] \\ &= C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - \underbrace{\mathbb{E}[\tilde{f}_{I-i+1}^{(k,I)}]}_{\hat{f}_{I-i+1}} C_{i,I-i+1} \prod_{j=I-i+2}^{I-1} \underbrace{\mathbb{E}[\tilde{f}_j^{(k,I+1)}]}_{\hat{f}_j} \\ &= 0. \end{aligned}$$

Par conséquent, la variance du CDR s'écrit :

$$\begin{aligned} \mathbb{V}ar(\widehat{CDR}_i) &= \mathbb{E}\left[\left(C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - \tilde{f}_{I-i+1}^{(k,I)} C_{i,I-i+1} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)}\right)^2\right] \\ &= C_{i,I-i+1}^2 \underbrace{\left(\prod_{j=I-i+1}^{I-1} \hat{f}_j^2 + \mathbb{E}[(\tilde{f}_{I-i+1}^{(k,I)})^2] \prod_{j=I-i+2}^{I-1} \mathbb{E}[(\tilde{f}_j^{(k,I+1)})^2]\right)}_{A1} \\ &\quad - 2 \underbrace{\mathbb{E}\left[\tilde{f}_{I-i+1}^{(k,I)} \hat{f}_{I-i+1} \prod_{j=I-i+2}^{I-1} \hat{f}_j \tilde{f}_j^{(k,I+1)}\right]}_{B1}. \end{aligned}$$

Pour le calcul de $A1$, on utilise le fait que $\left(\frac{C_{I-j+1,j}}{S_j^{I+1}}\right)^2 \frac{\hat{\sigma}_j^2}{\hat{f}_j^2 S_j^I} \leq \frac{\hat{\sigma}_j^2}{S_j^I} \approx 0$ et l'approximation linéaire suivante :

$$\prod_j (1 + x_j) \approx 1 + \sum_j x_j.$$

Pour le calcul de $B1$, on utilise à nouveau l'indépendance des pseudo-facteurs de développement, de sorte que :

$$B1 = \prod_{j=I-i+1}^{I-1} \hat{f}_j^2.$$

Finalement, on obtient :

$$\text{Var}(\widehat{CDR}_i) \approx \hat{C}_{i,I}^2 \left(\frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{I-i+1}^2 S_{I-i+1}^I} + \sum_{j=I-i+2}^{I-1} \left(\frac{C_{I-j+1,j}}{S_j^{I+1}} \right)^2 \frac{\hat{\sigma}_j^2}{\hat{f}_j^2 S_j^I} \right).$$

Cette formule correspond à l'estimateur de l'erreur d'estimation proposé par Merz et Wüthrich pour l'année de survenance i .

Pour l'ensemble des années de survenance, l'erreur d'estimation s'écrit :

$$\text{Var}\left(\sum_{i=1}^I \widehat{CDR}_i\right) = \sum_{i=1}^I \text{Var}(\widehat{CDR}_i) + 2 \sum_{1 \leq i < j \leq I} \text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j),$$

avec

$$\text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j) = \mathbb{E}[\widehat{CDR}_i \widehat{CDR}_j] - \mathbb{E}[\widehat{CDR}_i] \mathbb{E}[\widehat{CDR}_j].$$

Or $\text{Cov}(\widehat{CDR}_i, \widehat{CDR}_j) = \mathbb{E}[\widehat{CDR}_i \widehat{CDR}_j]$. On suppose $i < j$. On a alors :

$$\begin{aligned} & \mathbb{E}[\widehat{CDR}_i \widehat{CDR}_j] \\ &= \mathbb{E}\left[\left(C_{i,I-i+1} \prod_{m=I-i+1}^{I-1} \hat{f}_m - \tilde{f}_{I-i+1}^{(k,I)} C_{i,I-i+1} \prod_{m=I-i+2}^{I-1} \tilde{f}_m^{(k,I+1)}\right) \right. \\ & \quad \left. \left(C_{j,I-j+1} \prod_{m=I-j+1}^{I-1} \hat{f}_m - \tilde{f}_{I-j+1}^{(k,I)} C_{j,I-j+1} \prod_{m=I-j+2}^{I-1} \tilde{f}_m^{(k,I+1)}\right)\right] \\ &= C_{i,I-i+1} C_{j,I-j+1} \mathbb{E}\left[\left(\prod_{m=I-i+1}^{I-1} \hat{f}_m\right) \left(\prod_{m=I-j+1}^{I-1} \hat{f}_m\right) \right. \\ & \quad \left. + \underbrace{\tilde{f}_{I-i+1}^{(k,I)} \tilde{f}_{I-j+1}^{(k,I)} \left(\prod_{m=I-i+2}^{I-1} \tilde{f}_m^{(k,I+1)}\right) \left(\prod_{m=I-j+2}^{I-1} \tilde{f}_m^{(k,I+1)}\right)}_A \right. \\ & \quad \left. - \underbrace{\tilde{f}_{I-i+1}^{(k,I)} \left(\prod_{m=I-i+2}^{I-1} \tilde{f}_m^{(k,I+1)}\right) \left(\prod_{m=I-j+1}^{I-1} \hat{f}_m\right)}_B \right. \\ & \quad \left. - \underbrace{\tilde{f}_{I-j+1}^{(k,I)} \left(\prod_{m=I-j+2}^{I-1} \tilde{f}_m^{(k,I+1)}\right) \left(\prod_{m=I-i+1}^{I-1} \hat{f}_m\right)}_C \right]. \end{aligned}$$

En utilisant l'indépendance des pseudo-facteurs de développement vus à la date $t = I + 1$, on obtient :

$$\mathbb{E}[B] = \mathbb{E}[C] = \left(\prod_{m=I-i+1}^{I-1} \hat{f}_m \right) \left(\prod_{m=I-j+1}^{I-1} \hat{f}_m \right).$$

Le calcul de $\mathbb{E}[A]$ est plus complexe¹⁹ et on obtient :

$$\mathbb{E}[A] \approx \left(\prod_{m=I-i+1}^{I-1} \hat{f}_m \right) \left(\prod_{m=I-j+1}^{I-1} \hat{f}_m \right) \left(1 + \frac{C_{i,I-i+1} \hat{\sigma}_{I-i+1}^2}{S_{I-i+1}^{I+1} S_{I-i+1}^I \hat{f}_{I-i+1}^2} + \sum_{m=I-i+2}^{I-1} \left(\frac{C_{I-m+1,m}}{S_m^{I+1}} \right)^2 \frac{\hat{\sigma}_m^2}{S_m^I \hat{f}_m^2} \right).$$

Finalement, la covariance est approximativement égale à :

$$\mathbb{E}[\widehat{CDR_i} \widehat{CDR_j}] \approx C_{i,I} C_{j,I} \left(\frac{C_{i,I-i+1} \hat{\sigma}_{I-i+1}^2}{S_{I-i+1}^{I+1} S_{I-i+1}^I \hat{f}_{I-i+1}^2} + \sum_{m=I-i+2}^{I-1} \left(\frac{C_{I-m+1,m}}{S_m^{I+1}} \right)^2 \frac{\hat{\sigma}_m^2}{S_m^I \hat{f}_m^2} \right).$$

On retrouve la même covariance que celle proposée par Merz et Wüthrich et le même estimateur $(u - bias)^2$ de l'erreur d'estimation agrégée.

Erreur de processus

Soit CDR_i le CDR ne prenant en compte que l'erreur de processus. Dans ce cas, on a :

$$\begin{aligned} \tilde{f}_j^{(k,I)} &= \hat{f}_j, \\ \tilde{f}_j^{(k,I+1)} &= \frac{S_j^I}{S_j^{I+1}} \hat{f}_j + \frac{C_{I-j+1,j+1}^{(k)}}{S_j^{I+1}}, \end{aligned}$$

avec

$$C_{I-j+1,j+1}^{(k)} = \hat{f}_j C_{I-j+1,j} + \hat{\sigma}_j \sqrt{C_{I-j+1,j} \epsilon_{I-j+1,j+1}^{(k)}} \text{ où } \epsilon_{I-j+1,j+1}^{(k)} \sim \mathcal{N}(0, 1).$$

Pour $i \in \{1, \dots, I-1\}$ et $j \in \{1, \dots, J-1\}$, on a les égalités suivantes :

$$\left\{ \begin{array}{l} \mathbb{V}ar(C_{i,I-i+2}^{(k)}) = C_{i,I-i+1} \hat{\sigma}_{I-i+1}^2, \\ \mathbb{E}[(C_{i,I-i+2}^{(k)})^2] = \mathbb{V}ar(C_{i,I-i+2}^{(k)}) + \mathbb{E}[(C_{i,I-i+2}^{(k)})]^2 = C_{i,I-i+1} \hat{\sigma}_{I-i+1}^2 + C_{i,I-i+1}^2 \hat{f}_{I-i+1}^2, \\ \mathbb{V}ar(\tilde{f}_j^{(k,I+1)}) = \frac{\mathbb{V}ar(C_{I-j+1,j+1}^{(k)})}{(S_j^{I+1})^2} = \frac{C_{I-j+1,j} \hat{\sigma}_j^2}{(S_j^{I+1})^2}, \\ \mathbb{E}[(\tilde{f}_j^{(k,I+1)})^2] = \mathbb{V}ar(\tilde{f}_j^{(k,I+1)}) + \mathbb{E}[\tilde{f}_j^{(k,I+1)}]^2 = \frac{C_{I-j+1,j} \hat{\sigma}_j^2}{(S_j^{I+1})^2} + \hat{f}_j^2. \end{array} \right.$$

La variance du CDR s'écrit :

$$\begin{aligned} \mathbb{V}ar(CDR_i) &= \mathbb{V}ar \left(C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - C_{i,I-i+2}^{(k)} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)} \right) \\ &= \mathbb{V}ar \left(C_{i,I-i+2} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)} \right). \end{aligned}$$

19. On pourra se référer à l'article [4] pour un développement détaillé de cette étape.

En utilisant l'indépendance des $\tilde{f}_j^{(k,I+1)}$ pour les différentes années de développement, on a :

$$\begin{aligned}\mathbb{E}\left[\left(C_{i,I-i+2}^{(k)} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)}\right)^2\right] &= \mathbb{E}[(C_{i,I-i+2}^{(k)})^2] \prod_{j=I-i+2}^{I-1} \mathbb{E}[(\tilde{f}_j^{(k,I+1)})^2] \\ &= (C_{i,I-i+1} \hat{\sigma}_{I-i+1}^2 + C_{i,I-i+1}^2 \hat{f}_{I-i+1}^2) \prod_{j=I-i+2}^{I-1} \left(\hat{f}_j^2 + \frac{\hat{\sigma}_j^2 C_{I-j+1,j}}{(S_j^{I+1})^2}\right).\end{aligned}$$

Et on a d'autre part :

$$\mathbb{E}\left[C_{i,I-i+2}^{(k)} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)}\right] = C_{i,I-i+1} \hat{f}_{I-i+1} \prod_{j=I-i+2}^{I-1} \hat{f}_j = \hat{C}_{i,I}.$$

On obtient finalement l'erreur de processus :

$$\begin{aligned}\mathbb{V}ar(CDR_i) &= (C_{i,I-i+1} \hat{\sigma}_{I-i+1}^2 + C_{i,I-i+1}^2 \hat{f}_{I-i+1}^2) \prod_{j=I-i+2}^{I-1} \left(\hat{f}_j^2 + \frac{\hat{\sigma}_j^2 C_{I-j+1,j}}{(S_j^{I+1})^2}\right) - \hat{C}_{i,I}^2 \\ &= \hat{C}_{i,I}^2 \left[\left(1 + \frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{I-i+1}^2 C_{i,I-i+1}}\right) \prod_{j=I-i+2}^{I-1} \left(1 + \frac{\hat{\sigma}_j^2 C_{I-j+1,j}}{\hat{f}_j^2 (S_j^{I+1})^2}\right) - 1\right].\end{aligned}$$

On retrouve bien l'estimateur de l'erreur de processus proposé par Merz et Wüthrich pour l'année de survenance i .

Pour l'ensemble des années de survenance, l'erreur de processus s'écrit :

$$\mathbb{V}ar\left(\sum_{i=1}^I CDR_i\right) = \sum_{i=1}^I \mathbb{V}ar(CDR_i) + 2 \sum_{1 \leq i < j \leq I} \text{Cov}(CDR_i, CDR_j),$$

avec

$$\text{Cov}(CDR_i, CDR_j) = \mathbb{E}[CDR_i CDR_j] - \mathbb{E}[CDR_i] \mathbb{E}[CDR_j].$$

On peut écrire :

$$\begin{aligned}\mathbb{E}[CDR_i] &= \mathbb{E}\left[C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - C_{i,I-i+2}^{(k)} \prod_{j=I-i+2}^{I-1} \tilde{f}_j^{(k,I+1)}\right] \\ &= C_{i,I-i+1} \prod_{j=I-i+1}^{I-1} \hat{f}_j - \underbrace{\mathbb{E}[C_{i,I-i+2}^{(k)}]}_{\hat{f}_{I-i+1} C_{i,I-i+1}} \prod_{j=I-i+2}^{I-1} \underbrace{\mathbb{E}[\tilde{f}_j^{(k,I+1)}]}_{\hat{f}_j} \\ &= 0.\end{aligned}$$

D'où $\text{Cov}(CDR_i, CDR_j) = \mathbb{E}[CDR_i CDR_j]$.

L'erreur de processus s'écrit finalement²⁰ :

$$\mathbb{E}[CDR_i CDR_j] = \hat{C}_{i,I} \hat{C}_{j,I} \left[\left(1 + \frac{\hat{\sigma}_{I-i+1}^2}{\hat{f}_{I-i+1}^2 S_{I-i+1}^{I+1}}\right) \prod_{m=I-i+1}^{I-1} \left(1 + \frac{\hat{\sigma}_m^2 C_{I-m+1,m}}{\hat{f}_m^2 (S_m^{I+1})^2}\right) - 1\right].$$

20. On pourra se reporter à l'article [4] pour un développement plus détaillé.

Ce résultat correspond bien à celui de Merz et Wüthrich et permet d'obtenir l'erreur de processus agrégée.

Erreur de prédiction

L'erreur de prédiction s'obtient en sommant les erreurs d'estimation et de processus obtenues précédemment :

$$MSEP(CDR_i) = \mathbb{V}ar(CDR_i) + \mathbb{V}ar(\widehat{CDR_i}),$$

$$MSEP(CDR) = \mathbb{V}ar(CDR) + \mathbb{V}ar(\widehat{CDR}).$$

E. La formule fermée pour le calcul du capital réglementaire en assurance non-vie

La fonction de volatilité $\rho(\sigma)$ est définie par :

$$\rho(\sigma) = \frac{e^{VaR_{99,5\%}(\Phi(0,1))\sqrt{\log(\sigma^2+1)}}}{\sqrt{\sigma^2+1}} - 1,$$

où $\Phi(0,1)$ désigne la fonction de répartition de la loi normale centrée réduite.

Démonstration

Soit X une variable aléatoire correspondant à la variation de la quantité suivante :

$$\frac{P^{I+1} + BE^I}{Primes^I}.$$

Elle représente donc le ratio combiné suivant : la variation des provisions sur le passé auxquelles s'ajoute la somme des coûts de l'année à venir, le tout rapporté aux primes de l'année.

X est à valeurs dans $]0, \infty[$ et l'hypothèse du QIS 5 est de considérer que X suit une loi log-normale d'espérance μ_X et de variance σ_X^2 .

Par définition, on a : $\mathbb{P}(X < VaR_{99,5\%}(X)) = 99,5\%$. On appelle aussi $VaR_{99,5\%}(\Phi(0,1))$ le quantile à 99,5 % de la loi normale centrée réduite.

Or la fonction logarithme est croissante. On peut donc écrire :

$$\mathbb{P}(\log X < \log(VaR_{99,5\%}(X))) = 99,5\%.$$

Soit $Y = \log(X)$ une variable aléatoire suivant une loi normale $\mathcal{N}(\mu_Y, \sigma_Y^2)$ telle que :

$$\begin{cases} \mu_Y &= \log(\mu_X) - \frac{1}{2} \left(\frac{1+\sigma_X^2}{\mu_X^2} \right), \\ \sigma_Y^2 &= \log \left(\frac{1+\sigma_X^2}{\mu_X^2} \right). \end{cases}$$

Par définition de la Value-at-Risk, on obtient :

$$\mathbb{P} \left[\frac{Y - \mu_Y}{\sigma_Y} < \frac{\log(VaR_{99,5\%}(X)) - \mu_Y}{\sigma_Y} \right].$$

Comme $\frac{Y - \mu_Y}{\sigma_Y}$ suit une loi normale centrée réduite, on peut en déduire :

$$\begin{aligned} \frac{\log(VaR_{99,5\%}(X)) - \mu_Y}{\sigma_Y} &= VaR_{99,5\%}(\Phi(0,1)) \\ \Leftrightarrow \log(VaR_{99,5\%}(X)) &= VaR_{99,5\%}(\Phi(0,1)) \times \sigma_Y + \mu_Y \\ \Leftrightarrow VaR_{99,5\%}(X) &= \exp(VaR_{99,5\%}(\Phi(0,1)) \times \sigma_Y + \mu_Y) \\ \Leftrightarrow VaR_{99,5\%}(X) &= \exp(VaR_{99,5\%}(\Phi(0,1)) \times \sigma_Y) \times \exp(\mu_Y) \end{aligned}$$

On peut alors réécrire cette égalité de la manière suivante :

$$VaR_{99,5\%}(X) = \exp\left\{VaR_{99,5\%}(\Phi(0,1))\sqrt{\log\left(\frac{1+\sigma_X^2}{\mu_X^2}\right)}\right\} \times \exp\left\{\log(\mu_X) - \frac{1}{2}\log\left(\frac{1+\sigma_X^2}{\mu_X^2}\right)\right\}.$$

Si on suppose que le profit moyen comptable est nul, alors $\mu_X = 1$ et on a la formule suivante :

$$VaR_{99,5\%}(X) = \frac{e^{VaR_{99,5\%}(\Phi(0,1))\sqrt{\log(\sigma_X^2+1)}}}{\sqrt{\sigma_X^2+1}}.$$

La fonction $\rho(\sigma)$ représente donc l'écart entre la Value-at-Risk à 99,5 % de X et sa moyenne supposée unitaire.

F. Les tests de normalité utilisés

Le QQ-plot

Un QQ-plot est un graphique quantile-quantile qui permet de comparer des quantiles empiriques et des quantiles théoriques d'un modèle. Un bon alignement des points sur la droite dite de Henry traduit un bon ajustement du modèle.

La loi théorique très fréquemment utilisée est la loi normale centrée réduite. Si les points sont alignés, l'hypothèse de normalité est acceptée. On s'intéresse dans ce qui suit à la loi normale.

Soit X une variable aléatoire suivant une loi normale centrée réduite $\mathcal{N}(0, 1)$. On note Φ sa fonction de répartition telle que $\Phi(x) = \mathbb{P}(X \leq x)$. La fonction inverse Φ^{-1} est la fonction quantile. On souhaite alors vérifier si les points $(\Phi^{-1}(\frac{i}{n+1}), x_{i:n})$ sont alignés, $x_{i:n}$ étant la i^e valeur de l'échantillon, pour $i = 1 \dots, n$. Si ces points sont alignés, l'échantillon suit une loi normale.

Le test de Shapiro-Wilk

Ce test a été proposé par Shapiro et Wilk en 1965 et est relativement puissant en ce qui concerne les petits effectifs $n \leq 50$. L'étude présentée dans ce mémoire correspond bien puisqu'on utilise 45 résidus pour ce test. La statistique du test est définie par:

$$W = \frac{\left(\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} a_i (x_{(n-i+1)} - x_{(i)})\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

avec

$$\begin{cases} x_{(i)} & \text{est la série des données triées,} \\ \lfloor \frac{n}{2} \rfloor & \text{est la partie entière de } \frac{n}{2} \\ a_i & \text{est une constante générée à partir de l'espérance et de la matrice de} \\ & \text{variance-covariance des quantiles d'un échantillon de taille } n \text{ suivant la loi normale.} \end{cases}$$

W est peut s'interpréter comme le coefficient de détermination entre la série des quantiles générées à partir de la loi normale et des quantiles empiriques obtenues à partir des données. Plus W est élevé, plus l'hypothèse de normalité est vérifiée. La région de rejet est telle que $W < W_R$, W_R étant la valeur seuil pour un risque α et un effectif n donnée dans la table de Shapiro-Wilk.

Sous le logiciel R, on peut utiliser la fonction *shapiro.test* de la librairie *nortest*. On peut ainsi également vérifier l'hypothèse de normalité en s'intéressant à la p-value du test. En effet, si cette p-value est supérieure à 5 %, alors on ne rejette pas l'hypothèse de normalité pour un seuil $\alpha = 0,05$.