



**Mémoire présenté devant le jury de l'ENSAE en vue de l'obtention du
Diplôme d'Actuaire ENSAE
et de l'admission à l'Institut des Actuaires**

le 4 Juillet 2013

Par : Phi-Hung LE et Gaël VIROT

Titre : Méthode d'évaluation stochastique des provisions techniques d'une entreprise d'assurance dommages par une approche ligne à ligne

Confidentialité : non.

**Membres présents du jury de l'Institut
des Actuaires**

Signature :

Signature :

Membres présents du jury de l'ENSAE

Xavier MILHAUD

Manasa PATNAM

Entreprise

AXA GIE

Directeurs de mémoire en entreprise

Ngoc An DINH et Charles-Erwin ELIACHAR

Signatures :

Invité

Signature :

**Autorisation de publication et de mise en ligne sur un site de diffusion
de documents actuariels**

Signatures des responsables entreprise

Signatures des candidats

Bibliothèque :

Secrétariat :

Résumé

Le sujet de ce mémoire est de proposer un modèle de provisionnement non-vie « ligne à ligne », c'est-à-dire en utilisant les données individuelles des tous les sinistres déjà observés, et non des données agrégées. Notre travail s'est articulé en deux étapes : d'abord, l'établissement d'un modèle théorique s'inscrivant dans le cadre développé par Chau et Dinh (2012) [7]. Nous nous sommes concentrés sur l'identification d'hypothèses permettant une modélisation simplifiée, mais qui ne sont pas restrictives, afin de permettre une utilisation sur des cas réels. A partir des hypothèses faites, un estimateur consistant de la moyenne des réserves est donné. Cet estimateur est comparé au comportement asymptotique de l'estimateur Chain-Ladder, après agrégation des données. Nous prouvons une équivalence asymptotique entre les deux estimateurs. Des simulations permettent de vérifier, qu'à distance finie, l'estimateur ligne à ligne et l'estimateur Chain-Ladder sont proches. Nous proposons également des procédures bootstrap pour estimer la distribution des réserves, dans une optique de mesure du risque de provisionnement. Ensuite, le modèle est testé sur une base de données de sinistres automobiles. L'étude met en lumière l'importance d'une bonne segmentation de la base entre sinistres attritionnels et sinistres graves. Nous présentons une méthode de « segmentation dynamique », qui permet de prendre en compte qu'un sinistre d'abord évalué comme attritionnel peut par la suite devenir grave, et vice-versa. Nous effectuons également une estimation des réserves liées aux sinistres tardifs.

Abstract

The purpose of this study is to build an individual claim loss reserving model, which consists in estimating future payments directly upon claim-by-claim data. First, based on the framework developed in [7], a stochastic model is built, which leads to significant simplification in regards to what is proposed in the literature, while not making too restrictive assumptions to be considered enough realistic in a variety of practical cases. Under these assumptions, a consistent estimator of the reserve is derived, which is compared to the Chain-Ladder estimator after data aggregation. We find an asymptotic equivalence between the Chain-Ladder estimator and the estimator based on individual claims. A simulation study was conducted in order to verify that the claim-by-claim estimator and the Chain-Ladder estimator are near at finite distance. We also propose bootstrap procedures to estimate the distribution of reserve, which could be used to quantify reserve risk. Then, the model is tested on an automobile claims database. It reveals the importance of defining a good segmentation between attritional claims and large claims. We expose a "dynamic segmentation" method allowing to take into account a transition from attritional to large claims status during claims life cycle. Late claims have also been estimated.

Remerciements

Nous aimerions tout d'abord vivement remercier nos encadrants Ngoc An DINH et Charles-Erwin ELIACHAR, pour nous avoir fait bénéficier de leur expérience, de leurs conseils précieux, et de leur supervision très active tout au long de l'année. Ils nous ont énormément appris, tant au niveau de la théorie que de la pratique de l'actuariat.

Nous remercions également Xavier MILHAUD et le corps enseignant de l'ENSAE pour la qualité des cours qui nous ont été dispensés.

Table des matières

Introduction	1
1 Le provisionnement en assurance non-vie : aspects économiques, réglementaires et pratiques	5
1.1 L'activité d'assurance non-vie	6
1.1.1 Le contrat d'assurance	6
1.1.2 Le périmètre de l'assurance non-vie	7
1.1.3 L'inversion du cycle de production et ses conséquences	8
1.2 Réglementation	11
1.2.1 Normes comptables françaises	14
1.2.2 Directive Solvabilité II	18
1.2.3 Normes IFRS 4	25
1.2.4 Bilan	28
1.3 Marché de l'assurance non-vie	30
1.3.1 Classification par branches techniques	30
1.3.2 Les classifications officielles	32
1.3.3 Niveau de granularité	32
1.3.4 Classification par contrats et garanties	33
1.3.5 Poids des branches	34
1.4 La gestion des sinistres en pratique	36
1.4.1 La survenance du sinistre	38
1.4.2 La déclaration à l'assureur	41
1.4.3 L'ouverture du dossier sinistre	42
1.4.4 L'expertise	44
1.4.5 Indemnités	45
1.4.6 Clôture, recours de l'assureur, réouverture	48
1.4.7 La gestion des provisions de sinistres	50
2 Méthodes de provisionnement classiques	55
2.1 Construction du triangle de développement	55
2.2 La méthode de référence : Chain-Ladder	58
2.2.1 Hypothèses et estimation	58
2.2.2 Validation de Chain Ladder	59
2.2.3 Limites de la méthode Chain-Ladder	60
2.2.4 Autres modèles déterministes	60
2.3 Chain Ladder Stochastique : Mack (1993)	62
2.3.1 Hypothèses	62

2.3.2	Estimation	62
2.3.3	Validation du modèle	63
2.3.4	Erreurs de prédiction	63
2.4	Modélisation GLM	64
2.4.1	Modèles GLM	64
2.4.2	Estimation du modèle	65
2.4.3	Qualité de l'ajustement	65
2.4.4	Modèles usuels	66
2.5	Bootstrap	67
2.5.1	Principe	67
2.5.2	Application à la modélisation GLM	67
2.6	Facteurs de queue	68
3	Établissement d'un cadre pour une évaluation des provisions par une approche ligne à ligne	69
3.1	Revue succincte de la littérature sur les modèles individuels	69
3.2	Démarche adoptée et présentation du cadre	71
3.3	Définitions de base	73
3.3.1	Processus d'états	74
3.3.2	Processus de paiements	76
3.3.3	Processus de charges	77
3.4	Ensembles d'information	78
3.4.1	Sous-ensembles de sinistres d'intérêt	78
3.4.2	Information disponible à la date t	81
3.5	Mesures de provision	83
3.5.1	Mesures de provision cohérentes	83
3.5.2	Mesures de provision retenues	84
4	Modèle et estimation	87
4.1	Modélisation et estimation des processus	87
4.1.1	Processus d'états	87
4.1.2	Processus de règlements	90
4.2	Estimation des provisions	98
4.2.1	Flux de trésorerie à un an	98
4.2.2	Réserve à l'ultime	100
4.3	Comparaison avec la méthode agrégée Chain-Ladder	103
4.3.1	Flux de trésorerie à un an	103
4.3.2	Réserve ultime	106
4.4	Distribution des réserves	108
4.4.1	Bootstrap « total »	108
4.4.2	Bootstrap en utilisant la durée de vie estimée	108
4.4.3	Bootstrap et distribution non-paramétrique des facteurs de développement	109
5	Exercice de simulations	111
5.1	Générateur de bases de données	111
5.2	Générateur de scénario central	113
5.2.1	Paramétrage retenu	113
5.2.2	Étude d'une réalisation du PGD central	114

5.2.3	Sensibilité au nombre de sinistres simulés	117
5.3	Augmentation de la volatilité de la première charge	118
5.4	Augmentation de la volatilité des facteurs de développement	119
5.5	Branche longue	119
5.6	Levée de l'hypothèse d'équi-distribution du premier paiement	120
6	Applications à des données réelles	121
6.1	Description de la base	121
6.2	Profondeur de la base	122
6.3	Construction de la base d'étude	122
6.4	Traitements effectués	122
6.4.1	Durée d'observation	122
6.4.2	Cas de réouvertures	122
6.4.3	Cas de valeurs nulles ou négatives	123
6.5	Démarche adoptée	124
6.6	Statistiques descriptives	125
6.6.1	Charges	126
6.6.2	Durée de vie	128
6.6.3	Facteurs de développement	128
6.7	Vérification des hypothèses	129
6.7.1	Indépendance entre les facteurs de développement et la durée de vie . . .	129
6.7.2	Indépendance entre les facteurs de développement et le passé	130
6.7.3	Indépendance des facteurs de développement	132
6.8	Estimation sans segmentation	134
6.8.1	Description de la démarche	134
6.8.2	Résultats	134
6.9	Segmentation <i>a priori</i>	136
6.9.1	Démarche	136
6.9.2	Choix du seuil	136
6.9.3	Résultats	136
6.9.4	Indépendance des facteurs avec le passé après segmentation	139
6.9.5	Sensibilité par rapport au seuil	142
6.9.6	Segmentation <i>a priori</i> sur la dernière charge observée	144
6.10	Segmentation dynamique	146
6.10.1	Modèle	146
6.10.2	Procédure bootstrap	147
6.10.3	Le choix du seuil	148
6.10.4	Résultats	148
6.10.5	Sensibilité par rapport au seuil	150
6.10.6	Comparaison entre la segmentation <i>a priori</i> et segmentation dynamique .	152
6.11	Analyses complémentaires	154
6.11.1	Temps de calcul	154
6.11.2	Comparaison de la distribution avec les modèles stochastiques agrégés . .	155
6.12	Base retraitée	157
6.12.1	Retraitements effectués	157
6.12.2	Modélisation des tardifs	157
6.12.3	Nombre de tardifs	158

TABLE DES MATIÈRES

6.12.4 Résultats	159
6.12.5 Limites	159
6.12.6 Pistes d'amélioration	161
Conclusion	163
Bibliographie	165
Annexe A Divers	169
A.1 Exemple de bilans	170
A.2 Classification officielle des branches	172
A.3 Estimation non-paramétrique	175

Introduction

En assurance non-vie, tous les engagements d'un contrat ne sont pas quantitativement fixés à sa signature. L'assureur s'engage à verser, en cas de survenance d'un sinistre, une indemnité dont le montant n'est pas connu à l'avance.

L'assureur doit être en mesure d'évaluer correctement ses dettes probables, appelées **provisions techniques** (ou réserves), c'est-à-dire être en mesure d'estimer à tout moment le montant des engagements qu'il a pris vis-à-vis de ses assurés.

Un mauvais provisionnement (*reserving*) de la part de l'assureur peut avoir des conséquences graves. L'insuffisance des provisions pour sinistres a été la cause principale de faillite des compagnies d'assurance aux Etats-Unis, pour la période 1969 - 2002, d'après une étude d'A.M. Best. Par exemple, sur la branche RC et accidents en Assurance Dommages, sur la période 1969 - 2002 aux Etats-Unis, 37% des cas de faillite étaient dus à une insuffisance des provisions pour sinistres¹.

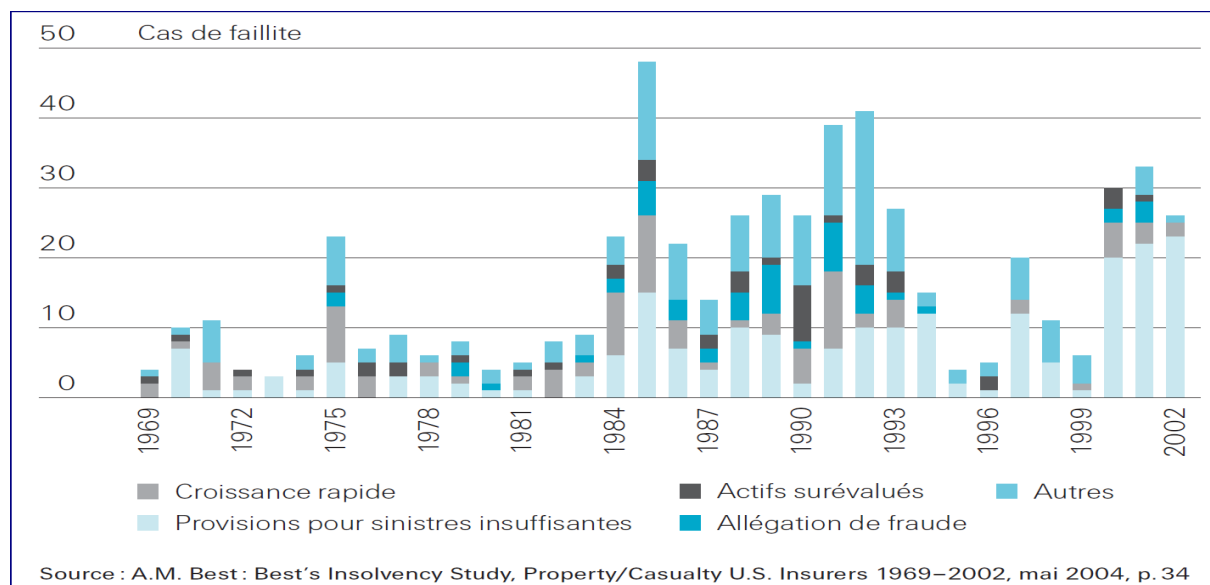


FIGURE 1 – Causes des faillites enregistrées des compagnies d'assurance américaines dommages, responsabilité civile et accidents, période 1969-2002.

1. N.B. : l'insuffisance de réserve peut avoir comme cause première une sous-tarification du produit d'assurance vendu.

INTRODUCTION

De manière symétrique, un sur-provisionnement revient à immobiliser du capital supplémentaire qui pourrait être investi dans des valeurs à de meilleurs taux de retours sur investissement¹ que celui généré par les actifs de la compagnie. Par ailleurs, un sur-provisionnement peut aussi entraîner un redressement fiscal, puisqu'il entraîne un résultat non déclaré.

Les méthodes traditionnelles de provisionnement reposent sur des techniques de triangulation. Les paiements effectués par l'assureur liés aux indemnités de sinistres sont agrégés par année de survenance et par année de développement. Seul l'étalement des montants agrégés au cours du temps est analysé : les données spécifiques à chaque sinistre (les données individuelles) ne sont pas prises en compte.

England et Verrall (2002) [13] notent que les « techniques traditionnelles [de provisionnement] ont été développées avant l'avènement des ordinateurs personnels², en ayant recours à des méthodes qui pouvaient être utilisées avec un crayon et un papier ».

Le recours à une modélisation individuelle, **par une approche ligne à ligne** sur l'ensemble des sinistres, pourrait apporter de nombreux avantages. Il permettrait notamment :

- de valider ou non l'utilisation de méthodes agrégées en fonction de la perte d'information ou des caractéristiques individuelles des sinistres. Notamment, étudier les données individuelles permet de segmenter de manière dynamique les portefeuilles pour obtenir des ensembles plus homogènes.
- de potentiellement dépasser certaines limites des méthodes agrégées. Parmi ces limites, on peut citer les problèmes de sur-paramétrisation des modèles (due à un faible nombre d'observations après agrégation des données). Les méthodes agrégées sont aussi généralement peu adaptées en présence de données négatives dans les triangles, cas relativement courants avec des données de charges (paiements effectués et réserves dossier par dossier constituées par le gestionnaire de sinistres) .
- d'obtenir des estimations de la distribution des engagements futurs probables, permettant de mesurer la capacité de l'assureur à faire face à ceux-ci dans les cas les plus extrêmes. Les méthodes agrégées fournissent une distribution prédictive des réserves au prix soit de formuler une hypothèse sur la loi des provisions (ou des paiements), ou soit en faisant appel à des techniques de ré-échantillonnage sur un échantillon initial contenant une information trop succincte. Les distributions prédictives de réserve obtenues avec un modèle ligne à ligne permettraient d'exploiter l'ensemble de l'information disponible.
- d'appliquer des limites de contrat (montant maximum des indemnités), et des structures de réassurance non-proportionnelle. L'approche ligne à ligne permettrait ainsi de calculer les réserves cédées au réassureur (et les réserves nettes), par exemple dans le cas de traités de type « excédent de sinistres » (*excess of loss*³), et d'étudier la volatilité des réserves brutes et des réserves nettes.

1. p. ex. un investissement dans le capital d'une compagnie d'assurance d'un pays émergent.

2. L'utilisation de la méthode Chain-Ladder remonterait aux années 30 (Denuit et Charpentier, 2005 [6]) .

3. c'est-à-dire lorsque le réassureur prend en charge les indemnités liés à un sinistre qu'à partir d'un certain seuil de sinistre (la priorité) et avec une limite d'intervention (la garantie).

Le mémoire s'articule selon le plan suivant :

- dans la première partie, nous rappellerons l'environnement réglementaire. Nous détaillons aussi les différentes étapes de la vie d'un sinistre, de sa survenance à sa clôture, afin de mieux se familiariser avec le type d'information qui sera disponible dans les données individuelles de sinistres.
- dans la seconde partie, nous exposerons brièvement les principales méthodes agrégées de provisionnement, en mettant l'accent sur la méthode déterministe de référence : Chain-Ladder, et les principales méthodes stochastiques.

Il est conseillé au lecteur familier avec les notions relatives au secteur de l'assurance non-vie et avec les enjeux du provisionnement de se rendre directement en partie 3.

- dans la troisième partie, nous présenterons le cadre que nous avons retenu pour étudier les sinistres par une approche ligne à ligne.
- dans la quatrième partie, nous exposerons le modèle stochastique permettant d'évaluer les provisions par une approche ligne à ligne. Nous comparerons le comportement asymptotique de nos estimateurs avec ceux de la méthode agrégée Chain-Ladder.
- dans la cinquième partie, nous réaliserons une étude de simulations pour étudier l'écart à distance finie entre les estimateurs de la méthode ligne à ligne et ceux de la méthode Chain-Ladder.
- dans la sixième partie, nous effectuerons une application sur des données réelles issues d'une branche d'assurance automobile particulier. Nous présenterons les méthodes de segmentation étudiées qui permettent d'obtenir des sous-ensembles de sinistres où les hypothèses du modèle ne sont pas invalidées.

Partie 1

Le provisionnement en assurance non-vie : aspects économiques, réglementaires et pratiques

Dans une démarche de modélisation, il est essentiel de connaître les spécificités des risques et leur importance. Cette première partie est donc l'occasion de rappeler les principales dimensions de l'activité d'assurance non-vie, notamment pour mettre en exergue le rôle central des provisions. L'assurance non-vie couvre des catégories de risque très diverses. Chaque catégorie possède ses propres caractéristiques, et représente un poids plus ou moins considérable dans l'activité de chaque assureur.

Nous décrivons aussi l'environnement réglementaire et comptable dans lequel les provisions s'inscrivent. Si l'évaluation des engagements auxquels fait face l'assureur est primordiale pour la gestion stratégique des risques, les provisions constituent également des éléments réglementés. Il est donc important de déterminer comment un modèle de provisionnement interagit avec les attentes réglementaires.

Enfin, il nous paraît également important d'évoquer les principales étapes de la gestion des sinistres et de la constitution des provisions. La connaissance du processus de gestion des sinistres est importante pour bien appréhender les données que nous serons amenées à manipuler, d'autant plus que nous utilisons des données de charges (règlements effectués et provisions dossier-dossier constituées par les gestionnaires de sinistres). Elle permet également d'aborder les spécificités (légales, ou bien propres aux branches ou aux compagnies) qu'une modélisation doit tenter de prendre en compte, ainsi que de dégager les éventuelles limites opérationnelles.

1.1 L'activité d'assurance non-vie

1.1.1 Le contrat d'assurance

Une opération d'assurance est une opération synallagmatique : une partie - l'assureur - s'engage à payer une somme ou une prestation à une autre partie - l'assuré - en cas de survenance d'un sinistre prédéfini à l'avance. En échange, l'assureur reçoit une prime (ou cotisation) payée par l'assuré à la date définie par le contrat ¹ (voir figure 1.1).

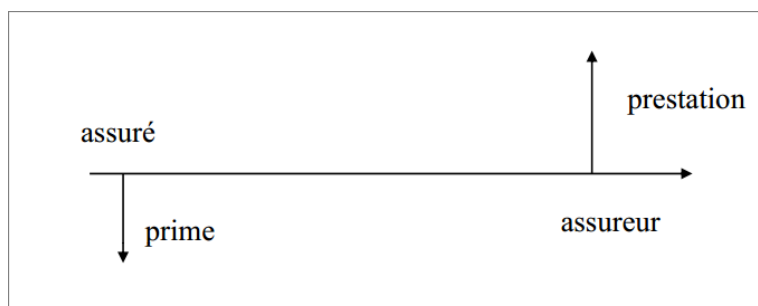


FIGURE 1.1 – Schéma d'un contrat d'assurance (Source : [30])

Dans un contrat d'assurance, il convient de distinguer :

- l'assuré, qui court le risque,
- le souscripteur, qui signe le contrat,
- le bénéficiaire, qui perçoit la prestation.

Par exemple, dans le cas d'un accident de voiture, le souscripteur peut être le propriétaire du véhicule, l'assuré le conducteur occasionnel du véhicule qui a causé l'accident, couvert par la garantie, et le bénéficiaire le tiers accidenté.

Dans le cas où le souscripteur est une personne physique, et l'assuré une ou plusieurs personnes, le contrat est individuel. Le contrat peut également être collectif s'il est souscrit par une personne morale et compte de nombreux assurés (par exemple, l'assurance d'une flotte de véhicules d'entreprises).

Un risque assurable est un événement futur, aléatoire (quant à sa probabilité ou sa date de réalisation), légalement assurable² et mesurable³. Sa réalisation ne doit pas dépendre de la volonté exclusive d'une partie. La nature du (des) risque(s) garanti(s) par le contrat (ou encore l'objet de la garantie) permet de délimiter l'activité d'assurance non-vie de l'assurance-vie, la première étant le champ d'application de ce mémoire.

1. Généralement à l'initiation du contrat, ou périodiquement (p. ex. en assurance emprunteur dans le cas d'un crédit immobilier), voire en fin de contrat (p. ex. pour l'assurance grêle pour les exploitants agricoles)

2. Les amendes pénales et administratives sont par exemple non assurables.

3. Les risques nucléaire et de guerre sont généralement exclus des garanties en raison de leur gravité.

1.1.2 Le périmètre de l'assurance non-vie

On distingue trois catégories d'assurance différentes, en fonction de leur finalité : les assurances de choses (ou assurances de biens), les assurances de responsabilité et les assurances de personnes¹.

Les **assurances de biens et de responsabilité** (ou assurances de dommages²) ont pour objet la protection du patrimoine de l'assuré. Les assurances de biens et de responsabilité se subdivisent en deux catégories :

- L'**assurance de biens** couvre des biens matériels appartenant à l'assuré contre les accidents, incendies, vols et autres dommages involontaires. En général, l'assuré, le souscripteur et le bénéficiaire ne forment qu'une seule et même personne. Des pertes immatérielles, dites « pertes indirectes » (telles que les pertes d'exploitation, ou les pertes de loyers) peuvent aussi être couvertes, car elles sont souvent consécutives à des pertes matérielles directement causées par le sinistre.
- Les **assurances de responsabilité** prennent en charge les conséquences financières des dommages (matériels ou corporels) dont l'assuré est responsable vis-à-vis d'un (ou de) tiers : l'assureur indemnise les victimes à la place de l'assuré. Le bénéficiaire est ici systématiquement un tiers.

Les **assurances de personnes** ont pour objet de garantir la personne humaine. L'assureur s'engage à verser une rente ou un capital en cas de réalisation d'un risque touchant la personne même de l'assuré (maladie, accident, décès, survie). On peut distinguer :

- les assurances sur la vie, dites « assurances-vie » : le risque porte sur le décès ou la survie de l'assuré ;
- les contrats de capitalisation : il s'agit d'un contrat d'épargne garantissant un capital déterminé à l'échéance du contrat.
- les assurances de dommages corporels regroupent les assurances des risques d'atteinte à l'intégrité physique en cas de maladie ou d'accident corporel.

L'assurance non-vie comprend les assurances de biens et de responsabilité et les assurances de dommages corporels. Si, en assurance-vie, l'incertitude peut porter uniquement sur la date de réalisation d'un événement (cas d'un capital-décès par exemple), l'incertitude en assurance non-vie porte la plupart du temps à la fois sur la réalisation du risque (l'assureur ignore si un sinistre va affecter le bénéficiaire du contrat au cours de la période de couverture), et le cas échéant, sur la date de réalisation de ce risque et les montants qui seront à verser. L'incertitude sur les prestations que l'assureur devra verser est un des aspects de l'inversion du cycle de production en assurance.

1. Source : FFSA.

2. ou IARD, pour Incendie, Accidents et Risque Divers, ou encore *Property and Casualty* en anglais.

1.1.3 L'inversion du cycle de production et ses conséquences

Dans une entreprise traditionnelle, le prix de revient du produit est connu et payé avant le prix de vente. En assurance, **le cycle de production est inversé** : les primes (le prix de vente) sont connues (et également perçues dans la grande majorité des cas) avant que les règlements des indemnités de sinistres (le prix de revient) interviennent, le cas échéant.

En encaissant directement le prix de vente (la cotisation), l'assureur dispose immédiatement de fonds, mais il s'est engagé à verser une indemnité en cas de survenance au cours de la période de validité du contrat. Du fait des conséquences différées des sinistres¹, des délais d'obtention de l'information, des expertises, des recours, etc. (voir partie 1.4), l'assureur décaissera son prix de revient (le règlement effectif) dans le futur, souvent au cours d'un exercice comptable différent de celui auquel la cotisation a été émise. Cependant, la sinistralité d'un contrat est (dans la grande majorité des branches d'assurance) rattachée à l'exercice comptable d'émission de la prime du contrat, afin de dégager le résultat annuel de l'assureur. De plus, dans l'intervalle de temps entre l'encaissement du prix de vente et le décaissement du prix de revient, il est important de mesurer la capacité de l'assureur à honorer les engagements pris vis-à-vis des assurés.

Ainsi, l'inversion du cycle de production de l'activité d'assurance entraîne un traitement comptable singulier par rapport à la comptabilité générale : une compagnie d'assurances ne doit pas seulement comptabiliser ses dettes certaines, mais également ses **dettes probables** à l'égard de ses assurés et des bénéficiaires de contrats, ce qu'on appelle les **provisions techniques** :

« Les provisions techniques sont les provisions destinées à permettre le règlement intégral des engagements pris envers les assurés et bénéficiaires de contrats. Elles sont liées à la technique même de l'assurance, et imposées par la réglementation. » (Simonet, 1998 [47]).

De manière plus détaillée, les provisions techniques correspondent (voir figure 1.2) :

- aux charges à prévoir pour faire face aux prestations non encore versées mais prévisibles des contrats en cours. Par exemple, dans le cas d'un sinistre automobile déjà déclaré à l'assureur, aux sommes qui seront probablement versées aux bénéficiaires. Ces prestations concernent à la fois les sinistres déclarés à l'assureur (sinistres « ouverts ») et les sinistres déjà survenus mais pas encore notifiés à l'assureur (sinistres « tardifs ») ;
- à une anticipation des prestations futures auxquelles l'assureur devra faire face lorsqu'un engagement prendra effet : par exemple, une partie des primes d'assurance payée au cours d'une période comptable peut être destinée à couvrir des risques de la période comptable suivante.

1. Les dommages peuvent se produire longtemps après la survenance du fait dommageable, par exemple dans le cas des risques liés à l'amiante.

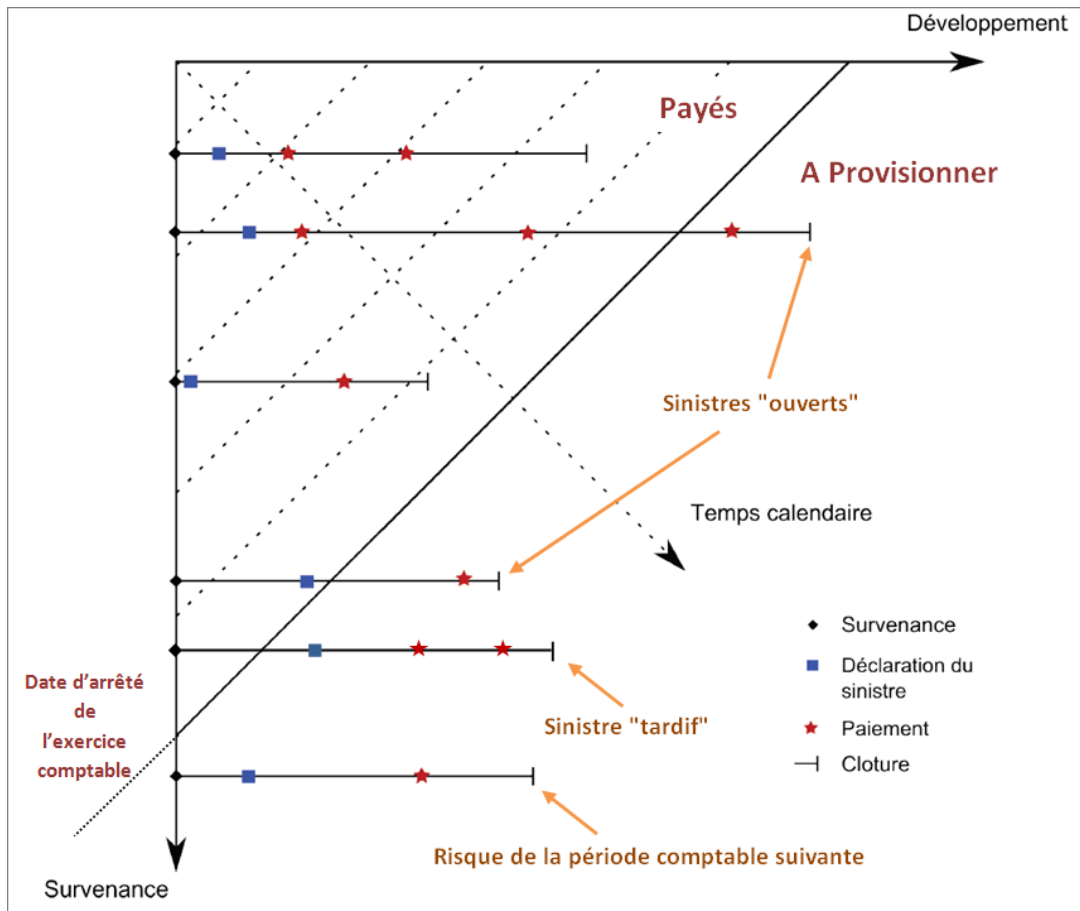


FIGURE 1.2 – Diagramme de Lexis permettant de localiser les engagements futurs de l'assureur par rapport au temps calendaire (exercice comptable)

Le bilan d'une compagnie d'assurance reflète l'inversion du cycle de production et se lit de droite à gauche. De manière schématique, le bilan traduit comment les engagements de l'assureur contractés à l'égard des assurés qu'il devra honorer dans le futur (les provisions techniques, inscrites au passif du bilan) sont couverts par les placements (l'emploi des primes perçues dans le passé), inscrits à l'actif du bilan (voir figure 1.3).

ACTIF	PASSIF
Placements	Fonds propres
	Provisions techniques

FIGURE 1.3 – Bilan simplifié d'une compagnie d'assurance

Les fonds propres, constitués des capitaux initiaux et des réserves accumulées (résultats des

exercices précédents non distribués aux actionnaires¹ et résultat courant) traduisent la situation nette de l'entreprise : une situation nette positive signifie que les biens et les créances de la compagnie sont plus importants que ses engagements et ses dettes. Dans le cas d'une situation nette négative, la compagnie est dans une situation d'insolvabilité : elle doit faire appel aux actionnaires² ou à des fonds de garantie pour remplir ses engagements et ne pas faire banqueroute.

En assurance, les capitaux propres ne servent donc pas à financer l'activité (les investissements nécessaires à la production d'un bien ou d'un service), mais protègent les assurés propriétaires du passif : la situation nette doit être suffisante pour pouvoir faire face aux différents risques les plus imprévus. C'est l'une des raisons pour lesquelles les États³, au travers d'autorités de contrôle⁴, exigent que les compagnies d'assurance détiennent un montant de fonds propres minimal (capital réglementaire ou capital immobilisé) pour exercer l'activité d'assurance.

La manière de comptabiliser les différents éléments du bilan d'une compagnie d'assurance varie selon les référentiels réglementaires. Dans la partie suivante, nous allons aborder notamment les exigences en termes d'évaluation des provisions techniques. Nous appréhenderons également le risque de provisionnement dans le spectre réglementaire. En première définition, il s'agit du risque que les provisions techniques ne soient pas suffisantes pour honorer les engagements de l'assureur (les fonds propres doivent alors être mis à contribution pour pallier l'insuffisance des provisions techniques). Le but est de pouvoir lier par la suite la modélisation des provisions techniques avec les exigences réglementaires.

1. c.-à-d. mis en réserve.

2. ou à ses sociétaires dans le cas d'une mutuelle.

3. voire les entités supra-nationales comme l'Union Européenne.

4. En France, l'ACP : l'Autorité de Contrôle Prudential.

1.2 Réglementation

Les entreprises d'assurance sont soumises ou seront soumises prochainement à plusieurs normes comptables et financières, françaises et internationales.

L'appréhension de la solvabilité d'une compagnie d'assurance, sa capacité à faire face à ses engagements futurs, reposent sur les trois blocs principaux du bilan (Tosetti *et. al.*, 2011 [51]). Les contrats de la société considérée vont générer des résultats ou des pertes futurs qui ne figurent pas dans les capitaux propres à la date d'évaluation. Ainsi :

1. Une compagnie d'assurance doit évaluer de manière correcte et suffisante le montant de ses engagements vis-à-vis des assurés (les provisions techniques) : des provisions insuffisantes pouvant générer des pertes futures ;
2. En représentation de ses engagements, la compagnie doit posséder un montant suffisant d'actifs sûrs, liquides et rentables (règle d'évaluation, de dispersion et de répartition prudentes) : la compagnie doit pouvoir être capable de faire face à des dépréciations futures de ses actifs pesant sur ses résultats ;
3. La compagnie doit posséder plus d'actifs réels que de dettes et engagements, c'est-à-dire des ressources supplémentaires lui permettant de rester solvable dans l'éventualité de résultats négatifs futurs (défaut d'un réassureur, catastrophes naturelles, etc.).

Les différents cadres réglementaires et de valorisation des compagnies d'assurance, répondant à des objectifs différents, induisent des manières différentes de comptabilisation et de représentation (règles de valorisation des provisions techniques et des actifs, calcul d'un capital minimal de solvabilité, etc.). Ils influencent la lecture du bilan d'une compagnie et l'image de sa situation nette. Il s'agit :

- des comptes sociaux en normes comptables locales. Ils servent notamment au calcul des impôts vis-à-vis de l'État. Nous nous concentrerons sur les **normes comptables françaises** (*French GAAP*) ;
- des comptes consolidés en normes **IFRS**, pour le *reporting* financier, notamment au niveau international ;
- de la norme **Solvabilité II** (*Solvency II*), qui servira notamment au contrôle prudentiel du capital réglementaire, harmonisé au niveau européen.

Pour l'instant, les entreprises d'assurance françaises sont soumises aux normes comptables françaises (code des assurances). Le régime européen Solvabilité I, issu de directives européennes des années 70¹, a été transposé dans le droit français : une marge de solvabilité minimale est exigée pour pouvoir exercer l'activité d'assurance.

De plus, depuis 2005, toutes les sociétés cotées en Europe ou faisant appel à l'épargne publique sont tenues de présenter leurs comptes consolidés selon les normes IFRS. C'est par exemple le cas d'AXA. Élaborées par l'IASB (*International Accounting Standards Board*), ces normes ont pour objectif d'harmoniser, d'améliorer et de clarifier les principes comptables internationaux.

La Directive Solvabilité II, adoptée en 2009 par le conseil de l'Europe et le parlement européen, pour remplacer la directive Solvabilité I, doit entrer en vigueur à l'horizon 2016 (selon

1. Il a connu une révision en 2002.

la dernière date de mise en application communiquée). La directive introduit un nouveau mode de calcul du capital réglementaire, selon une approche fondée sur les risques. Les règles de construction du bilan Solvabilité II diffèrent également des normes comptables françaises en vigueur actuellement, et seront communes à tous les pays de l'Union Européenne.

Les différents référentiels diffèrent sensiblement dans la manière d'appréhender les trois blocs du bilan simplifié (figure 1.3 *supra*). Dans le cadre de notre mémoire, nous sommes plus particulièrement intéressés par la valorisation des provisions, et les aspects de la solvabilité qui sont liés au risque de provisionnement ¹.

Avant de détailler ces deux points, nous pouvons esquisser les principaux points de divergence entre les différents référentiels :

a) Engagements

On peut citer deux divergences d'approche majeures dans la valorisation des engagements de l'assureur :

- **Évaluation prudente vs *Best Estimate*** Les normes comptables françaises font référence à une notion de suffisance/prudence de la valeur des engagements. Dans la norme Solvabilité II, c'est la notion de *Best Estimate* qui est retenue. Il s'agit de la valeur probable moyenne des engagements, en ayant recours aux hypothèses les plus réalistes, compte-tenu de l'information dont dispose l'assureur.
- **Non actualisation vs Actualisation** En assurance non-vie, les provisions techniques (en comptabilité française) sont en général non escomptées. À l'inverse, le *Best Estimate* fait l'objet d'une actualisation. Concernant les normes IFRS, l'escompte est pour l'instant optionnel.

b) Valorisation des actifs

En normes comptables françaises, « les engagements réglementés doivent être, à toutes époques, représentés par des actifs équivalents ». Ces actifs sont évalués au coût historique.

En normes IFRS et Solvabilité II, c'est l'évaluation *Marked to market* (ou *fair-value*) qui prévaut. Les actifs doivent être valorisés à la valeur à laquelle ils pourraient être échangés entre deux parties informées et consentantes sur un marché concurrentiel. Lorsque la valeur de marché de l'actif n'est pas disponible, il est possible d'utiliser un modèle mathématique (*mark-to-model*), ou un *proxy* (p. ex. prix de transaction d'éléments similaires). Si l'utilisation de la valeur économique n'est pas appropriée, les actifs sont à défaut évalués selon la valeur comptable nationale.

1. Le risque de provisionnement fait référence à l'incertitude concernant les résultats de la souscription de l'assureur : la valeur de ses engagements peut être affectée par des fluctuations affectant la date de survenance, la fréquence et la gravité des événements assurés ainsi que la date et le montant des règlements de sinistres [40].

c) Solvabilité

La marge de solvabilité française repose pour l'instant sur la taille de l'activité générée par la compagnie d'assurance (chiffre d'affaire, montant des sinistres).

L'exigence réglementaire de fonds propres sous Solvabilité II intègre les risques inhérents à l'activité d'assurance. L'exigence de fonds propres sous Solvabilité II est fondée sur l'exposition aux risques de la compagnie, en incorporant tous les risques liés à son activité, c.-à-d. principalement : le risque de souscription, le risque de marché, le risque de crédit, le risque opérationnel, le risque de liquidité.

Nous allons maintenant détailler chacun des cadres réglementaires.

1.2.1 Normes comptables françaises

1.2.1.1 Provisions techniques

L'article R331-1 du code des assurances, définit les engagements réglementés dont les entreprises d'assurance doivent à toute époque, être en mesure de justifier l'évaluation. Notamment :

Les provisions techniques **suffisantes** pour le règlement intégral de leurs engagements vis-à-vis des assurés, des entreprises réassurées et bénéficiaires de contrats (sans déduction des réassurances cédées à des entreprises agréées ou non).

Le qualificatif technique permet de faire la distinction avec les autres provisions telles que la provision pour risques et charges, provisions pour dépréciation, etc.

Les assureurs doivent provisionner des montants suffisants pour honorer leurs engagements, c'est-à-dire des montants supérieurs aux montants probables (c.-à-d. les montants moyens attendus (Tosetti *et al.*, 2011)).

En assurance non-vie, les provisions techniques sont essentiellement constituées de l'estimation des sinistres et des frais de gestion associés que l'assureur est engagé à payer : **la provision pour sinistres à payer**.

En général, les provisions techniques non-vie ne sont pas escomptés (sauf cas particulier de la provision mathématique des rentes et la provision pour risques croissants (PRC)¹, voir Tosetti *et al.* (2011), p. 81 [51]).

a) La provision pour sinistres à payer

L'article R331-6 du code des assurances définit la provision pour sinistres à payer (PSAP) :

La provision pour sinistres à payer représente 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 ».

Les articles R331-15 et sseq. définissent les modalités de calcul de la PSAP. On distingue :

- Les sinistres « ouverts »², qui correspondent aux sinistres déclarés pour lesquels le règlement des prestations n'est pas terminé à la date d'inventaire. L'assureur doit procéder à une évaluation « dossier par dossier »³, « le coût d'un dossier comprenant toutes les charges externes individualisables ».
- Les sinistres « tardifs »⁴ qui correspondent aux sinistres survenus mais non encore déclarés à l'assureur à la date de l'exercice comptable.

1. Il s'agit de la provision pouvant être exigée pour les opérations d'assurance contre les risques de maladie et d'invalidité et égale à la différence des valeurs actuelles des engagements respectivement pris par l'assureur et par les assurés.

2. RBNS pour *Reported But Not Settled*

3. Pour chaque dossier de sinistre déclaré, l'assureur estime le montant probable du décaissement correspondant.

4. ou IBNyR pour *Incurred But Not yet Reported*

La PSAP doit être déterminée :

- **brute de recours** : les recours sont à évaluer séparément sous la forme de prévisions de recours à encaisser¹ ;
- **brute de réassurance** : les provisions pour sinistres cédés apparaissent à l'actif du bilan ;
- en incluant les **charges externes individualisables** (ou ALAE (*Allocated Loss Adjustment Expense*)) Il s'agit des frais d'experts, honoraires d'avocats, etc.

La ventilation de la PSAP selon les différentes branches d'assurance non-vie est donnée dans l'annexe A344-10 état C10.

La PSAP pour les sinistres ouverts doit être systématiquement évaluée par une méthode dossier-dossier.

Un recours à des méthodes statistiques est autorisé mais ne concerne que les sinistres de moins de deux ans :

- sinistres matériels de la branche automobile,
- avec autorisation de la commission de contrôle pour les sinistres d'autres branches.

Les sinistres autres que corporels de la branche automobile sont évalués en utilisant concurremment les trois méthodes suivantes (Article R331-26) :

- Évaluation dossier par dossier (optionnelle) ;
- Évaluation par référence au coût moyen des sinistres des exercices antérieurs ;
- Évaluation basée sur les cadences de règlement observées dans l'entreprise au cours des exercices antérieurs.

L'évaluation la plus élevée résultant de ces deux (ou trois) méthodes est retenue.

b) Autres provisions techniques

Provisions pour sinistres

Les autres provisions pour sinistres que l'on peut mentionner sont :

- La provision pour sinistre non encore manifesté (PSNEM) (R331-17 du code des assurances) dans le cadre de l'assurance construction et de la garantie décennale (garanties décennales de responsabilité civile et de dommage aux ouvrages), qui fixe les modalités d'estimation des sinistres tardifs dans cette branche. Il s'agit : « d'une estimation, calculée selon les méthodes fixées par arrêté du ministre de l'économie, du coût des sinistres non encore manifestés et qui devraient se manifester d'ici à l'expiration de la période de prescription décennale. » Les modalités de calcul sont fixés par l'article A331-21 du code des assurances.²

1. Ces prévisions constituent une estimation de la somme restant à recevoir des actions exercées par la compagnie d'assurance en vue d'obtenir le remboursement par le responsable d'un préjudice d'une indemnité ou partie d'indemnité versée au titre d'un sinistre.

2. Dans le détail, on a :

$$PSNEM = \max(s_n S_n, p_n P_n),$$

où :

- n est l'ancienneté du traité,
- S_n montant total des sinistres nets de recours, d'ancienneté n , qui se sont déjà manifesté à la date d'inventaire,

- La provision pour égalisation : elle a pour but de prendre en compte les charges exceptionnelles liées aux événements naturels, risques atomiques, terrorisme, etc. Elle peut aussi permettre de provisionner des pertes en fin d'exercice dans l'assurance crédit, ou encore une fluctuation anticipée de sinistralité en dommages corporels. En France, les conditions de comptabilisation de ces provisions sont définies dans l'article 39 du Code Général des impôts.

Provisions pour primes

Les sinistres non survenus et garantis font l'objet de la provision pour primes non acquises (PPNA) (UPR : *unearned premium reserve*) et de la provision pour risque en cours (PREC) (ou PDR pour *Premium deficiency reserve* ou encore URR pour *Unexpired Risk Reserve*) (Article R331-6, alinéa 2 et 2 bis) [41].

- La PPNA représente la proportion des primes émises qui correspond aux risques non encore courus (calculé au *prorata temporis*).
- La PREC, qui est additionnée à la PPNA couvre les risques futurs (dérive de sinistralité anticipée sur les risques non encore courus). Elle permet de couvrir la charge attendue si celle-ci est supérieure à la prime reçue placée en PPNA.

1.2.1.2 Marge de solvabilité

La Situation Nette Comptable (SNC) de la compagnie, c'est-à-dire les fonds propres, est calculée comme la différence entre l'actif réel (biens et créances) et le passif réel (dettes et engagements).

En fait, la réglementation française retient plutôt la notion de marge de solvabilité (Article R334-3 du code des assurances). Cette marge est calculée comme la somme de la SNC retraitée¹, des plus-values latentes (c'est-à-dire de l'excédent de la valeur de réalisation de l'ensemble des placements sur leur valeur comptable), de la dette subordonnée. La marge de solvabilité a pour but d'harmoniser la comptabilité d'Europe continentale, qui repose sur le principe du coût historique, et la comptabilité anglo-saxonne, qui retient le principe de Juste Valeur.

A la marge de solvabilité est comparée l'exigence minimale de marge de solvabilité (Article R334-5), qui vient de la transposition du régime Solvabilité 1 transposé en droit français dans les années 70. Il s'agit du montant des fonds propres que les autorités de régulation assurantielles imposent aux compagnies pour couvrir les risques qu'elles gèrent ou collectent. Ce montant a pour but d'assurer la solvabilité de la compagnie, sur une période de temps et à un seuil de confiance prédéfinis. Il permet de faire face à des risques systémiques ou des pertes extrêmes non couverts par les provisions, qui reflètent, elles, les charges certaines ou prévisibles à rattacher à

– P_n : montant total des primes émises nettes des frais d'acquisition,
– p_n et s_n : coefficients réglementés appliqués aux montants des primes et sinistres respectivement pour les traités d'âge n .

1. La SNC retraitée est constituée du capital versé au fonds d'établissement, de la moitié du capital non versé (dans le cas des sociétés anonymes) et des réserves, auxquels sont retranchés les actifs incorporels, les pertes, et la part des frais d'acquisition non admise en représentation des engagements réglementés.

l'exercice comptable.

Le minimum réglementaire de marge est basé sur le plus élevé des montants suivants :

- un premier montant basé sur les primes : 16% à 18% des primes (en fonction de tranches définies par la réglementation)
- un montant basé sur les sinistres : 23% à 26% des sinistres (en fonction des tranches)

Ces montants sont multipliés par un ratio réducteur visant à tenir compte de la réassurance (ratio sinistres payés dans le passé en net de cessions en réassurance sur les sinistres payés en brut de cession).



FIGURE 1.4 – Le calcul de la marge de solvabilité en Assurance non-vie (Source : [42])

1.2.1.3 Poids des provisions

Les provisions techniques représentent environ trois-quarts du passif au bilan des compagnies d'assurance non-vie. La PSAP constitue à elle seule généralement plus de 50% du passif (dans les compagnies d'assurance matures).

Le niveau des engagements peut cependant varier grandement selon les assureurs. A titre d'exemple, nous nous sommes intéressés au bilan 2011 de deux grands assureurs dommages français (Voir annexe A.1). Pour le premier, les provisions techniques brutes (hors réassurance) représentaient 70% du passif, la PSAP à elle seule représentant 64% du passif. Pour le second, les provisions techniques brutes ne représentaient que 50% du passif, les provisions pour sinistres (PSAP et autres provisions de sinistres) comptant pour 41% du passif.

1.2.2 Directive Solvabilité II

Sous la réglementation actuellement en vigueur (Solvabilité I), la solvabilité future ne dépend que de la situation actuelle, et il y a une absence de distinction entre les risques. La réforme réglementaire européenne de l'assurance, Solvabilité II, introduit un nouveau calcul du capital réglementaire selon une approche fondée sur les risques. L'exigence réglementaire de fonds propres sous Solvabilité II incorpore les risques effectifs auxquels les compagnies d'assurance sont exposées : risques liés aux engagements relatifs aux contrats d'assurance, mais également risques pris sur les marchés financiers au travers de l'allocation d'actifs, risque de défaut des contreparties, ...

Nous commençons d'abord par décrire les exigences de fonds propres sous Solvabilité II, avant d'aborder la prise en compte des provisions.

1.2.2.1 Seuil réglementaire de fonds propre

Deux niveaux de seuils réglementaires pour les fonds propres sont définis par la directive Solvabilité II : le SCR (*Solvency Capital Requirement*) et le MCR (*Minimum Capital Requirement*).

Le SCR doit représenter la valeur des fonds propres nécessaires pour éviter la faillite à un horizon d'un an avec un intervalle de confiance à 99,5% (exigence de capital cible correspondant à une probabilité de ruine de 1 sur 200 sur un horizon d'un an). Il s'agit donc d'une *Value at Risk* (VaR^1) à 99,5% des fonds propres.

Le MCR représente le niveau minimum de fonds propres en dessous duquel l'intervention de l'autorité de contrôle est automatique. La formule de calcul est assez similaire au mode de calcul du besoin minimum en marge de solvabilité dans Solvabilité I. Il est calculé à partir d'une fonction linéaire sur les primes émises et les provisions techniques en *Best Estimate*, calibrée de manière à reproduire une VaR à 85% sur des fonds propres, sur une période d'un an. Il doit situer dans un corridor de 25-45 % du SCR et ne pas se situer en-dessous d'un plancher absolu (Art. 130 de la Directive [40]).

Si une compagnie n'est pas en mesure de démontrer que son niveau de fonds propres est suffisant pour couvrir le SCR, elle devra soumettre dans les 2 mois à l'Autorité de contrôle un plan de redressement, précisant comment et quand elle pourra à nouveau respecter les exigences en capital réglementaire (le niveau de fonds propres éligibles couvrant le capital de solvabilité doit être rétabli dans un délai maximal de 6 mois) (Art. 138 de la Directive [40]). Si la compagnie franchit à la baisse le seuil du *MCR*, l'autorité de contrôle peut retirer l'agrément en l'absence d'un plan de rétablissement convaincant soumis dans un délai de 1 mois, avec obligation de rétablissement dans les 3 mois (Art. 139 de la Directive [40]).

a) Calcul du SCR : les principes de la formule standard

Les compagnies d'assurance peuvent calculer le SCR à l'aide d'une formule standard, d'un modèle interne, ou d'un modèle interne partiel.

1. La VaR au seuil α d'une v.a. X est le quantile de niveau α de la loi de X , c.-à-d. : $VaR_\alpha(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq \alpha\}$

La formule standard consiste en une approche modulaire : le risque global supporté par l'assureur est découpé en modules de risques (non-vie, opérationnel, crédit, ...), eux-mêmes découpés en sous-modules représentant les différents sous-risques qui les composent. Dans chacun des sous-modules, un capital de solvabilité élémentaire est déterminé (le SCR du sous-risque considéré). Une fois les SCR de chaque sous-module calculés, ils sont agrégés à l'aide d'une matrice de corrélation afin d'obtenir le SCR du module, lequel est également agrégé avec les autres modules pour obtenir le SCR global.

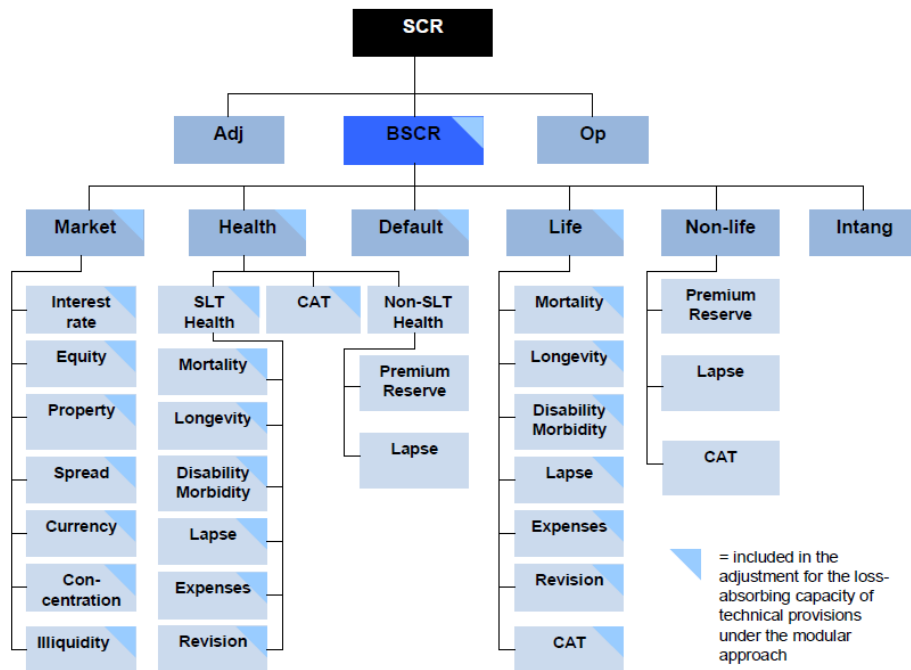


FIGURE 1.5 – Décomposition du SCR en formule standard

Le SCR est calculé comme suit :

$$SCR = BSCR + SCR \text{ Opérationnel} - Ajustement,$$

avec :

- *BSCR* le capital de solvabilité requis de base ;
- *SCR Opérationnel*, le capital de solvabilité au titre du risque opérationnel¹ ;
- *Ajustement* permet de tenir compte des pertes futures qui seront compensées par la réduction des impôts différés à payer et par la réduction de la participation aux bénéfices futurs à distribuer aux assurés.

Le BSCR est constitué des modules de risques, représentant les grandes familles de risques auxquelles les sociétés d'assurance sont confrontées. Dans notre cas, il s'agit des risques sur l'activité non-vie, les risques de marché, les risques de défauts des contreparties et le risque sur

1. Il s'agit du risque de pertes directes ou indirectes résultant de processus internes, de systèmes inadaptés ou défectueux, d'erreurs commises par des personnes,...

1.2. RÉGLEMENTATION

les intangibles (actifs incorporels). L'ensemble de ces risques est ensuite agrégé selon la formule suivante¹ :

$$BSCR = \sqrt{\sum_{ij} Corr_{ij} \times SCR_i \times SCR_j} + SCR_{intangible},$$

où $Corr_{ij}$ est le coefficient de la matrice de corrélation entre les modules de risque i et j (Voir Figure 1.6).

j \ i	Marché	Défaut	Non vie
Marché	1,00		
Défaut	0,25	1,00	
Non vie	0,25	0,50	1,00

FIGURE 1.6 – Corrélations entre les modules de risques du BSCR

Le SCR des sous-modules de risque se calcule généralement selon une des deux approches suivantes :

- Approche facteur : formule fermée permettant un calcul direct du SCR (p. ex. Risque catastrophe),
- Scénario de choc : calcul de la diminution des fonds propres économiques (NAV pour *Net Asset Value*²) suite à un choc instantané portant sur un facteur de risque R (Voir figure 1.7)³. Dans un scénario de choc, seule la variation du *Best Estimate* est considérée, la marge pour risque, qui sera définie ci-dessous, ne doit pas varier.

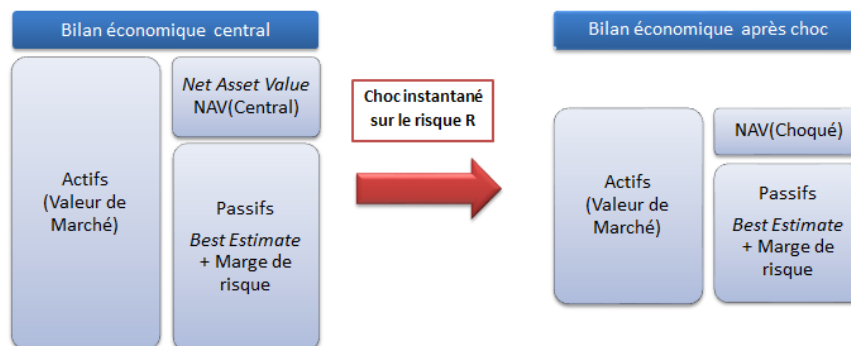


FIGURE 1.7 –

Calcul du capital de solvabilité au titre du risque R par une approche de type Formule Standard avec scénario de choc.

1. Le SCR Intangible représente le capital requis pour couvrir le risque sur les actifs incorporels.
 2. *Net Asset Value* désigne la valeur de marché des actifs diminuée du *Best Estimate* des engagements au passif.
 3. Si le scénario engendre une augmentation des fonds propres économiques, le SCR pour le risque concerné est nul.

Le SCR élémentaire du sous-module de risque est alors égal à :

$$\Delta NAV = NAV(Central) - NAV(Choqué).$$

1.2.2.2 Les fonds propres

Sous Solvabilité II, les fonds propres économiques sont composés à la fois des fonds propres de base et des fonds propres auxiliaires sous réserve de critères d'éligibilité.

- Les fonds propres de base éligibles sont constitués de l'excédent des actifs par rapport aux passifs (les capitaux propres réglementaires) et les passifs subordonnés¹.
- Les fonds propres auxiliaires éligibles sont des éléments de fonds propres hors-bilan qui peuvent être appelés pour absorber des pertes (fraction non versée du capital social, lettres de crédit, garanties,...). Lorsqu'ils sont payés ou appelés, ils sont assimilés aux fonds propres de base.

Par ailleurs, les fonds propres sont classés en trois catégories ou *Tiers*. Ce classement est fait en fonction de leur qualité évaluée au regard de leur disponibilité², de leur degré de subordination et de leur durée ou permanence. Selon leurs caractéristiques, les fonds propres de base peuvent être classés en fonds propres de *Tier* 1, 2 ou 3. Les fonds propres auxiliaires peuvent être classés en fonds propres de *Tier* 2 ou 3.

Des limites quantitatives sont imposées pour déterminer le montant des fonds propres éligibles en couverture respectivement du SCR et du MCR. Le SCR devait ainsi être couvert au moins à moitié par des éléments de *Tier* 1 et à 15% maximum par des fonds propres classés en *Tier* 3. Le MCR doit être contenu dans les éléments des Tiers 1 et 2, l'ensemble des éléments du *Tier* 1 devant dépasser 80% du MCR.

1.2.2.3 Provisions techniques

Le passif d'une compagnie doit être décomposé en distinguant :

- les passifs répliquables (*hedgeable*)
- les passifs non répliquables (*non-hedgeable*)

Un passif est répliquable si les flux qu'il engendre peuvent être parfaitement répliqués à l'aide d'instruments financiers disponibles sur un marché suffisamment actif, liquide et transparent. Il est alors évalué à l'aide de ce portefeuille répliquant valorisé en *Fair Value (Marked-to-Market)*.

Les provisions techniques constituent généralement des passifs non répliquables³ : leur évaluation doit correspondre à la somme du *Best Estimate* des flux futurs actualisés et d'une marge pour le risque (*risk margin*).

1. Un élément des passifs est subordonné si, en cas de liquidation, il est disponible pour absorber des pertes et son remboursement est refusé à son détenteur jusqu'à ce que tous les autres engagements d'assurance et de réassurance vis-à-vis des assurés soient honorés.

2. Capacité à être appelé sur demande pour absorber des pertes que ce soit dans le cadre de l'exploitation continue ou en cas de liquidation.

3. sauf cas particulier en assurance-vie des produits libellés en unités de compte sans garantie plancher.

a) Best Estimate

L'article 77 de la directive Solvabilité II ([40]) définit la composante *Best Estimate* des provisions techniques 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), estimée sur la base de la courbe des taux sans risque pertinents. »

La prise en compte de l'incertitude des flux de trésorerie futurs par le *Best Estimate* ne doit pas entraîner l'inclusion d'une marge additionnelle : le *Best Estimate* doit être la moyenne non biaisée de tous les scénarii possibles¹.

Le *Best Estimate* doit refléter toutes les caractéristiques du contrat d'assurance : comportement de l'assuré, existence d'options et de garanties financières, politique d'investissement et de gestions, prestations futures discrétionnaires, etc.². Le calcul *Best Estimate* des PSAP se fait généralement par des méthodes de provisionnement et non à partir des provisions dossier-dossier.

Quatre étapes sont requises dans l'élaboration du *Best Estimate* ([48]) :

1. Collecte et analyse des données (qualités, critère de sélection, avis d'expert)
2. Détermination des hypothèses nécessaires à l'évaluation (incertitude des CF)
3. Sélection du modèle
4. Validation et documentation (pertinence du modèle, communication)

Les spécifications techniques de l'EIOPA³ ([12]) indiquent que le *Best Estimate* doit être calculé brut de réassurance. Les provisions pour primes et les provisions de sinistres doivent être évaluées séparément. L'assureur doit évaluer tous les flux futurs de trésorerie susceptibles d'être encourus pour honorer les engagements faits aux assurés. Ainsi, toutes les dépenses liés au contrat doivent aussi être prises en compte (frais administratifs, de gestion des sinistres, etc.). Dans les flux de trésorerie entrants, les recours doivent être pris en considération.

1. En pratique, selon la nature des risques considérés, il n'est pas nécessaire d'incorporer explicitement tous les scénarii possibles dans l'évaluation des provisions, ou de déterminer les probabilités explicites dans tous les cas.

2. Les causes d'incertitude que la compagnie d'assurance doit prendre en compte sont les suivantes ([48] / TP 2.4) :

- Fluctuations sur la gravité et la fréquence des sinistres ;
- Montant et date des versement des prestations ;
- Variations de la valeur d'un indice de marché utilisé ;
- Facteurs spécifiques (juridiques, sociaux, économiques, démographiques, technologiques, ...)
- Incertitude sur le comportement des assurés ;
- Montant des frais ;
- dépendance des flux aux conditions économiques antérieures ;
- Interdépendance entre plusieurs causes d'incertitudes.

3. *European Insurance and Occupational Pensions Authority*, ex-CEIOPS), qui est un organe consultatif indépendant européen regroupant les autorités de contrôle des états membres. Il conseille la Commission Européenne sur les aspects techniques de Solvabilité II, et participe à l'élaboration des mesures d'application détaillées de la directive.

Les approches pour calculer le *Best Estimate* peuvent être diverses. La compagnie doit utiliser des techniques actuarielles et statistiques qui reflètent de manière appropriée les risques qui affectent les flux de trésorerie. Plusieurs approches sont possibles :

- Une approche stochastique : par simulation Monte-Carlo, Bootstrap, approche bayésienne,
- ...
- Une séries de projections déterministes avec des probabilités attribuées,
- Une évaluation déterministe (un seul scénario) : Chain-Ladder, coût moyen, analyse de sensibilités...

La compagnie doit être en mesure de justifier le choix de la technique retenue, et de prouver que les hypothèses sous-jacentes sont réalistes.

b) Marge pour risque

La marge pour risque (*Risk Margin* ou *Market Value Margin*) constitue une marge de prudence dans l'estimation des provisions techniques. Elle est définie comme le montant destiné à ([40]) :

« garantir que la valeur des provisions techniques est équivalente au montant que les entreprises d'assurance et de réassurance demanderaient pour reprendre et honorer les engagements d'assurance et de réassurance. »

La marge pour le risque est donc calculée dans une logique de transfert. Elle représente le coût financier de l'incertitude des flux futurs dans le cadre d'un transfert de risque.

La logique de transfert est la suivante :

- Dans un premier temps, transfert du portefeuille complet d'engagements d'assurance (pour lequel est calculé la marge pour risque) à une autre entreprise de référence qui n'avait pas d'engagements d'assurance ni de fonds propres avant ce transfert.
- Après le transfert, cette entreprise de référence constitue un montant de fonds propres égal au SCR nécessaire pour supporter les engagements d'assurance sur toute leur durée.

Pour le calcul de cette marge, une méthode coût du capital a été retenue. Plus précisément, il s'agit du : « coût que représente la mobilisation d'un montant de fonds propres éligible au *Solvency Capital Requirement* nécessaire pour faire face aux engagements d'assurance et de réassurance sur toute la durée de ceux-ci ». Formellement, le calcul de la marge pour risque se fait de la manière suivante ([12]) :

$$CoCM = CoC \times \sum_{t \geq 0} \frac{\mathbb{E}[SCR_t]}{(1 + r_{t+1})^{t+1}},$$

où :

- $CoCM$ est la marge pour risque,
- SCR_t est le SCR de l'année t ,
- CoC est le coût du capital fixé à 6% par l'EIOPA à la date d'écriture de ce mémoire,
- r_t correspond au taux d'intérêt sans risque de maturité t . La courbe de taux sans risques doit être choisie de manière cohérente avec l'unité monétaire utilisée pour établir les comptes financiers de la compagnie.

Plusieurs simplifications sont acceptées [12], par exemple :

- Approximation des risques dans les modules de calcul du SCR ;
- Approximation du SCR total pour les années futures ;
- Calcul de la marge pour le risque comme un pourcentage du *Best Estimate* (si l'activité de la compagnie est limitée à une seule branche).

1.2.2.4 Le risque de réserve en Formule standard

Le module risques non-vie en Formule Standard se décompose en 3 composantes, une des composantes étant le risque de prime et de réserve. Le risque de prime correspond au risque que les montants des pertes et les frais liés aux sinistres qui surviendront dans un an soient plus élevés que ce qui avait été anticipé lors de l'élaboration des tarifs. Le risque de réserve résulte de fluctuations dans le déroulé temporel et les montants des règlements de sinistres. Il couvre deux sources d'incertitudes : l'erreur de modèle (variabilité des montants réels des provisions) et l'erreur d'estimation (écart entre l'estimateur statistique et la variable estimée). Dans l'étude quantitative d'impact 5 ([5]), l'hypothèse d'une distribution log-normale du risque de prime et de réserve pour chaque branche était faite pour calculer la *VaR* à 99,5%.

1.2.3 Normes IFRS 4

Les Normes IFRS (*International Financial Reporting Standards*) sont élaborées par l'IASB (*International Accounting Standards Board*, Anciennement l'IASC¹).

Elles ont été introduites afin d'encourager l'harmonisation comptable dans le monde. Le but est d'élaborer un corps unique de normes comptables de haute qualité qui doit accroître la confiance des investisseurs. L'information comptable doit être transparente, cohérente entre pays, facilement comparables, établie selon des approches peu subjectives.

Le 11 Septembre 2002 est paru au JO.C.E. (Journal Officiel des Communautés Européennes), le règlement CE N°1606/2002 dit IFRS 2005 qui impose aux sociétés européennes cotées qui publient des comptes consolidés ou qui émettent de la dette au niveau international, l'application des IAS/IFRS pour les exercices ouverts à partir du 1^{er} Janvier 2005. Ainsi, depuis 2005, ces normes s'appliquent à toutes les sociétés cotées ou faisant appel public à l'épargne.

La norme IFRS 4 relative aux contrats d'assurance ne traite que des contrats d'assurance (ou de réassurance) présentant un risque d'assurance². Elle est structurée en deux phases :

- La phase 1, dite transitoire, durant laquelle comme son nom l'indique des mesures transitoires ont été adoptées. L'IFRS 4 a été publié le 31 mars 2004 et est entré en vigueur au 1er Janvier 2005, avec l'application effective de la Directive 1606/2002.
- La phase 2 est la phase de publication de la norme définitive, actuellement en cours. Elle doit résoudre les aspects conceptuels et la mise en œuvre de normes comptables relatives à l'évaluation des passifs d'assurance. Le 30 Juillet 2010, l'IASB a publié un exposé sondage (*Exposure Draft* ED/2010/8 [26]). Une phase de consultation auprès de l'industrie a été ouverte, et l'IASB a publié ses réponses en Février 2013.

1.2.3.1 Phase 1

La norme IFRS 4 phase I définit notamment la notion de contrat d'assurance. Les contrats d'investissement n'ayant pas de clause de participation aux bénéfices discrétionnaires sont traités comme des instruments financiers en application de la norme IAS 39 En matière comptable, elle propose des modifications limitées et permet l'application d'une grande variété de pratiques comptables.

Au niveau des provisions techniques, IFRS 4 Phase I autorise les entreprises à s'appuyer sur les normes comptables locales. Il est ainsi possible de conserver provisoirement des principes comptables locaux : par exemple, la vision prudente et le non-escompte des provisions en normes comptables françaises. Les provisions pour catastrophes et les provisions d'égalisation sont cependant supprimées.

1. *International Accounting Standards Committee.*, fondé en 1973

2. ainsi qu'aux contrats d'investissement ayant une clause de participation aux bénéfices discrétionnaires, par dérogation à IAS 39/IFRS 9.

1.2.3.2 Phase 2

Comme Solvabilité II, les normes IFRS 4 Phase II autorise la valorisation de contrats d'assurance répliquable par la méthode des portefeuilles répliquants.

L'évaluation des contrats d'assurance en IFRS 4 Phase II repose sur trois blocs :

- un *Best Estimate* ;
- une marge pour risque ;
- une marge résiduelle.

a) Best Estimate

L'estimation des flux de trésorerie doivent refléter les perspectives de l'entité : le *Best Estimate* doit être calculé comme la valeur actuelle probable des flux de trésorerie auxquels serait confronté l'assureur pour honorer lui-même le contrat (*current fulfillment value*). Il ne s'agit donc pas de la valeur actuelle probable pour transférer le contrat à un autre assureur (*current exit value*). Cela conduit à retenir des hypothèses et des données propres à la compagnie. Pour les variables de marché cependant, l'évaluation doit être consistante avec les prix de marché observé.

L'évaluation doit incorporer de manière non biaisée, toute l'information disponible à la date de calcul au niveau des montants, dates et incertitudes liés à tous les flux de trésorerie permettant à l'assureur d'honorer ses engagements.

La mesure doit être brute de réassurance et seuls les flux liés aux contrats doivent être pris en compte (traitement séparé des frais d'acquisition).

L'actualisation doit refléter les caractéristiques des cash-flows au passif (duration, monnaie, liquidité). Elle ne doit pas intégrer des risques qui ne sont pas présents dans le passif (si les flux de trésorerie ne dépendent pas de la performance d'actifs spécifiques, la courbe des taux sans risque doit être retenue).

b) Marge pour risque

La marge pour risque doit mesurer l'incertitude associée aux flux de trésorerie futurs, c'est-à-dire représenter le montant qui fournit une forte probabilité que l'assureur puisse remplir ses obligations. Elle doit fournir des informations au sujet du niveau de risque associé au contrat et la manière dont la compagnie gère ses risques.

L'exposé-sondage de Juillet 2010 définit la marge pour le risque comme : « le montant maximum qu'un assureur serait prêt à payer rationnellement pour être libéré du risque que les flux de trésorerie finaux soit supérieurs à ceux attendus. ».

Trois méthodes sont proposées pour estimer la marge pour risque :

- une méthode reposant sur les intervalles de confiance (approche *Value-at-Risk* VaR) ;
- une méthode *Conditional Tail Expectation* (ou $TVaR^1$) ;
- une méthode Coût du capital.

Suite aux commentaires formulés par les acteurs du marché sur l'exposé-sondage, l'IASB a décidé en Février 2013 de ne pas limiter aux trois méthodes précédentes les techniques pour

1. Pour une variable aléatoire X continue et intégrale, la $TVaR$ est définie par : $TVaR_\alpha = \mathbb{E}[X|X \geq VaR_\alpha(X)]$

mesurer l'ajustement pour le risque [27].

La marge pour risque doit être estimée au niveau des portefeuilles de contrats. La prise en compte d'un bénéfice de diversification entre les portefeuilles n'est pas autorisée. L'évaluation ne tient compte que des risques liés au passif d'assurance. Pour chaque portefeuille, l'assureur doit décider quelle mesure il souhaite appliquer. Le choix du seuil de confiance (pour la VaR, ou la TVaR, etc.), ainsi que le niveau d'agrégation, doit être justifié par la compagnie.

c) Marge résiduelle

La marge résiduelle (*Residual margin*) permet de lisser le gain à la souscription du contrat sur toute la durée de couverture du contrat.

Elle existe lorsque la valeur actualisée des flux entrants est supérieure à la somme de la valeur actualisée des flux sortants et de la marge pour le risque. La marge résiduelle doit être déterminée au niveau d'un portefeuille (et au sein d'un portefeuille, par contrats ayant une date de souscription et une période de couverture similaire). Elle doit être reconnue au début de la période de couverture des contrats. Le rythme de relâchement est déterminé à la souscription puis bloqué. Le relâchement se fait au *prorata temporis* ou, si cela reflète mieux l'exposition relative de la couverture, sur la base des prestations attendues.

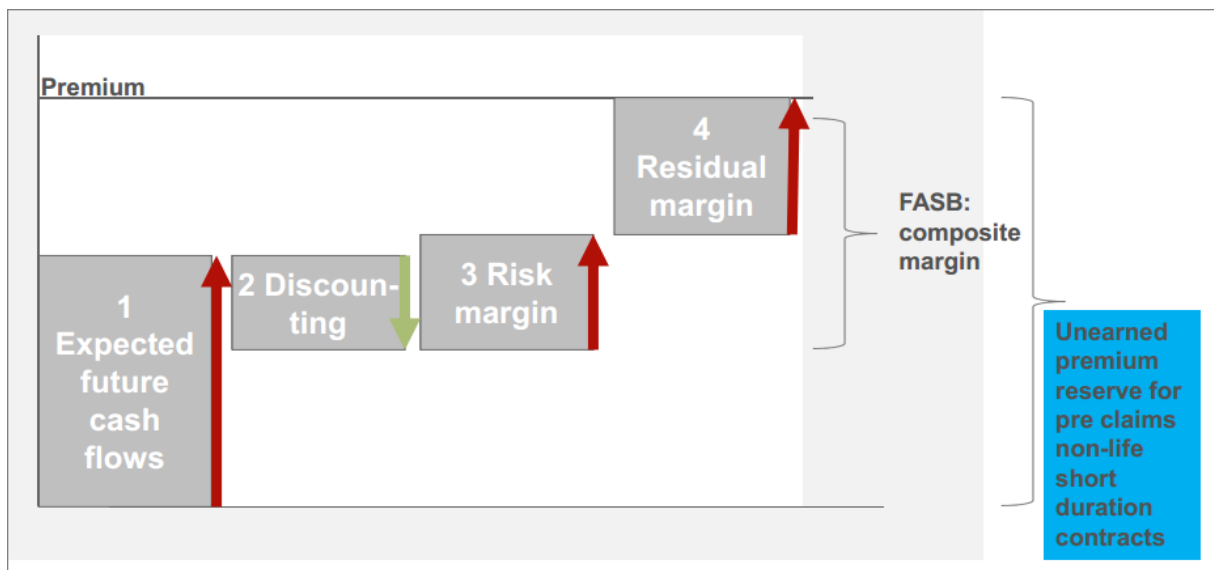


FIGURE 1.8 – Provisions sous IFRS4 - Phase II. Source : [42])

1.2.4 Bilan

La réglementation française retient une approche prudente, sans qu'il y ait une mesure quantitative de la marge de prudence qui est retenue par l'entreprise. Par exemple, prenons le cas où l'entreprise, dans ses processus internes, choisit de constituer des provisions dossier par dossier prudentes telles que les provisions estimées soient suffisantes pour régler tous les sinistres dans 9 cas sur 10 (VaR à 90%). En sommant les provisions dossier par dossier pour constituer la provision au niveau du portefeuille (pour les sinistres déclarés), il n'est pas possible d'explicitement la marge de prudence sans modélisation supplémentaire (La VaR n'est en général ni additive ni sous-additive).

Au contraire, IFRS 4 et Solvabilité II décomposent les provisions en plusieurs blocs, permettant de séparer le *Best Estimate* (espérance) de la marge de risque au-delà du *Best Estimate*, comme le montre la figure 1.9. La prudence implicite des normes comptables françaises est ici extériorisée. Notons que la mesure de la marge d'ajustement au risque dans les normes IFRS est propre à la compagnie (*entity specific measure*), tandis qu'avec Solvabilité II, la méthode de coût d'immobilisation du capital se veut plus consistant avec le marché.

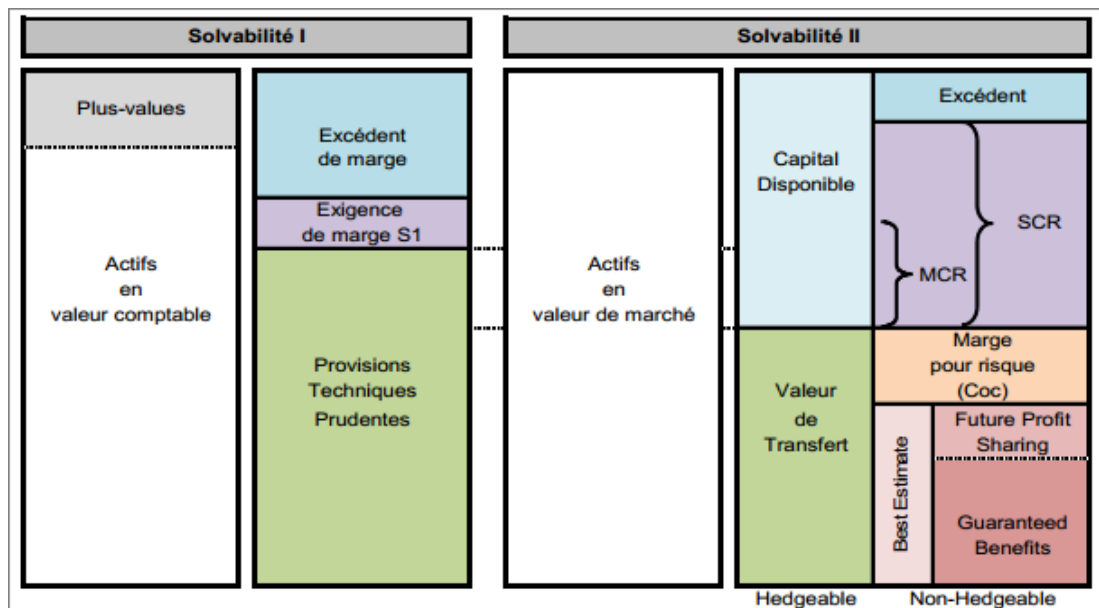


FIGURE 1.9 – Comparaison des bilans sous Solvabilité I (French GAAP) et Solvabilité II (Source : [42])

- Dans ce contexte, les méthodes statistiques déterministes ou stochastiques peuvent servir :
- Dans la réglementation française, à déterminer la provision pour les sinistres tardifs, ainsi qu'à analyser la pertinence des provisions dossier-dossier.
 - Dans les référentiels Solvabilité II et IFRS 4 Phase II, à fonder une approche *Best Estimate*. De plus, elles peuvent permettre de quantifier l'incertitude autour du *Best Estimate*, pour le calcul de la marge de risque en IFRS 4 (VaR à 75% par exemple) ou le SCR lié au risque de réserve en modèle interne (partiel).

L'utilisation de mesure de risque sur le provisionnement pose la question de l'agrégation. Si en vision *Best Estimate*, l'espérance est un opérateur linéaire, la mesure de l'incertitude autour de la moyenne peut faire apparaître des bénéfices de diversification entre les portefeuilles en les agréant. Les risques peuvent être de nature très différentes selon les branches, et les mesures d'incertitude pertinentes à retenir doivent tenir compte de la nature des risques au sein de chaque portefeuille. Il est donc important de connaître les caractéristiques propres à chaque branche pour faire du provisionnement et être en mesure de provisionner correctement. La prochaine partie présente brièvement les différents branches techniques en assurance non-vie.

1.3 Marché de l'assurance non-vie

1.3.1 Classification par branches techniques

a) Assurance dommages corporels

Cette catégorie recouvre les frais de soins, les rentes invalidités ou d'incapacité, les accidents du travail (*workers' compensation*).

Les indemnités du travail sont des assurances procurant des indemnités de remplacement de salaire et des prestations couvrant les frais de santé en cas de blessures ou d'accidents au cours de leur activité professionnelle ou en cas de maladie professionnelle (liée à l'amiante par exemple). En France, les salariés sont couverts par l'Assurance Maladie en cas d'accidents du travail. Des complémentaires santé sont souscriptibles.

b) Assurance automobile

L'assurance automobile couvre les dommages causés « à » (dommages matériels) ou « avec » (responsabilité, dommages corporels) un véhicule automobile.

Une segmentation supplémentaire consiste à séparer les contrats monovéhicules (pour les particuliers) et les contrats assurant plusieurs véhicules (flotte d'entreprise).

c) Assurance dommages aux biens

L'assurance dommages aux biens peut-être divisée en trois catégories :

- Dommage aux biens des particuliers :
- Dommages aux biens des professionnels, qui protège le patrimoine des entreprises (bâti-ments, machine, matériel informatique, marchandises). On y trouve notamment l'assurance bris de machine, dont l'objet est de couvrir les dommages à un outil artisanal ou industriel (grippage, surtension, ...).
- Dommages aux biens agricoles : perte d'exploitation, mortalité du bétail, grêle.

d) Assurance responsabilité civile générale (RC)

Plusieurs catégories peuvent être distinguées :

- La responsabilité civile « vie privée des particuliers » : cette assurance couvre l'assuré et sa famille vivant sous son toit pour les dommages qu'ils causent à autrui.
- La responsabilité civile de l'entreprise : elle vise à réparer les dommages corporels, matériels ou immatériels causés à un des tiers ou des clients lors de leur exploitation ou de l'exécution de tâches annexes à leur activité.

e) Assurance construction

En France, l'assurance construction peut se décomposer en deux catégories :

- Les assurance de dommages à l'ouvrage : il s'agit d'une assurance de choses à charge du propriétaire de l'ouvrage (maître d'ouvrage¹). La qualité d'assuré reste acquise aux propriétaires successifs du bien. Elles garantissent les dommages à l'ouvrage en cours de

1. Personne physique ou morale propriétaire du terrain pour le compte de laquelle est réalisée l'opération de construction

construction (éléments constitutifs ou éléments d'équipement affectant la solidité de l'ouvrage).

- Les assurances responsabilité décennale : c'est une assurance de responsabilité incombant aux constructeurs au sens large (entreprises de construction, maîtres d'œuvre, bureaux de contrôle, etc.).

Ces deux assurances sont obligatoires.

f) Assurance transport

Au sein de cette branche, on peut distinguer :

- Les assurances maritimes, qui sont destinées à couvrir les corps de navire et autre construction *off-shore*.
- Les assurances aviation, qui couvrent les risques liés au transport aéronautique
- Les assurances spatiales, qui sont destinées à la couverture des satellites et des lanceurs.
- Les assurances de marchandises transportées (dans le domaine terrestre, maritime, fluvial et aérien). Elles garantissent les marchandises contre les avaries que ces dernières peuvent subir au cours de leur transport : accidents (naufrage des navires), vol, casse (chute de colis pendant les opérations de chargement/déchargement), voire la grève ...

g) Assurance crédit

Il s'agit d'assurance portant sur le risque commercial, c'est-à-dire des opérations reposant sur des marchandises ou des prestations de service entre industriels ou commerçants : elle couvre contre le défaut de paiements d'une contrepartie (insolvabilité, le risque politique dans les contrats à l'exportation dans les contrats de crédit-export, crédits à la consommation, crédits agricoles, prêts immobiliers, ...).

h) Assistance

Il s'agit d'une assurance permettant d'être indemnisé des frais annexes¹ en cas d'accident, maladie ou autres situations d'urgence lors de voyages à l'étranger (p. ex. organisation du rapatriement, du transport et de l'admission en service hospitalier, remboursement des frais de location d'un autre véhicule en cas de vol du véhicule pendant des vacances à l'étranger, etc.).

i) Protection juridique

L'assuré bénéficie de la prise en charge des frais de procédure de justice l'opposant à un tiers (frais d'expertise, honoraires d'avocats, coût des actes d'huissier, etc.).

L'assurance prend en charge les frais de justice de l'assuré dans deux types de situations (et peut éventuellement le représenter) :

- soit l'assuré demande réparation auprès d'un tiers pour un préjudice qu'il estime avoir subi,
- soit l'assuré doit se défendre, car un tiers estime avoir subi un dommage dont il est responsable.

1. c.-à-d. Elle n'a pas en général vocation à réparer un préjudice.

j) Catastrophes Naturelles

En France, les catastrophes naturelles sont généralement traitées dans une branche spécifique, du fait de la spécificité de ces sinistres et des primes collectées pour la garantie catastrophes naturelles (sur-prime réglementaire sur le contrat d'assurance dommages).

Depuis la loi du 13 juillet 1982, toutes les personnes possédant une assurance dommages s'acquittent d'une prime additionnelle obligatoire au titre de la garantie Catastrophes Naturelles. Elles sont automatiquement couvertes en cas de dommages causés par un événement extrême déclaré catastrophe naturelle. Pour les contrats garantissant les risques aux biens de particuliers, cette prime s'élève à ce jour à 12% de l'ensemble des cotisations afférentes au contrat. Pour la branche Automobile, le taux annuel de la prime additionnelle est fixée à 0,5% des primes ou cotisations afférentes aux garanties dommages.

1.3.2 Les classifications officielles

Législation française

Le Code des assurances comporte une classification par branches de toutes les opérations d'assurance pouvant être pratiquées. Elle est donnée dans l'article R321-1 du Code des assurances (voir annexe A.2).

Solvabilité II

La directive Solvabilité II décompose les branches de la façon suivante (classification par risque) :

- Dépenses médicales ;
- Protection du revenu ;
- Indemnités de travail ;
- Responsabilité automobile, tiers ;
- Automobile, autres classes ;
- Marine, aviation et transport ;
- Incendie et autres dommages de biens ;
- Responsabilité civile, tiers ;
- Crédit et assurance caution ;
- Juridique ;
- Assistance ;
- Assurance non-vie divers.

1.3.3 Niveau de granularité

Des segmentations plus fines que les lignes d'activité sont généralement nécessaires afin d'obtenir des estimations précises des provisions. Il s'agit de prendre en compte les similitudes dans la cadence du développement des sinistres, dans le comportement des assurés, dans les propriétés des produits (garanties), les actions futures en terme de gestion de sinistres, etc. Par exemple, dans le cas de l'assurance marine, on sépare généralement les bateaux de plaisance des bateaux de transport des marchandises. Une segmentation par réseau de distribution (vente directe par téléphone ou internet, intermédiaires, etc.) peut aussi être nécessaire si les profils de risque des

contrats sont très différents, ou simplement si le suivi de la charge est effectué par réseau de distribution.

Un arbitrage avec la quantité de données disponibles est à faire. Si la taille du portefeuille est trop petite, une analyse statistique sera difficile à mettre en place. Pour certains risques particuliers, l'analyse des engagements repose sur des notions économiques, juridiques et sociales, et le recours à des opinions d'expert plutôt que le recours aux statistiques. On peut par exemple citer le cas de l'amiante. Pour les risques émergents, comme le risque de pollution, les nouveaux risques médicaux, l'historique est très fiable voire inexistant.

1.3.4 Classification par contrats et garanties

Dans les faits, les contrats délivrés par les compagnies d'assurance révèle généralement de plusieurs branches et plusieurs finalités.

Un contrat multirisque automobile réunit des garanties relevant de plusieurs branches :

- la branche « corps de véhicules terrestres à moteur » : tous les dommages subis par le véhicule assuré à la suite :
 - d'une collision ou d'un choc (garantie dommages tous accidents),
 - d'un incendie ou d'une explosion (garantie incendie),
 - d'un vol (garantie vol),
 - etc.
- la branche « responsabilité civile véhicules terrestres à moteur », tous les dommages causés aux tiers à la suite :
 - d'un accident,
 - d'un incendie,
 - etc.

Un contrat multi-risque habitation relève également de plusieurs garanties. A titre d'exemple, le poids des différentes garanties des contrats MRH en France est donné dans la figure 1.10.

RÉPARTITION DES GARANTIES AU SEIN DES CONTRATS MULTIRISQUES HABITATION EN 2011	
SOURCE : FFSA-GEMA	
1. INCENDIE	18,0%
2. VOL	16,5%
3. DÉGÂTS DES EAUX	15,5%
4. RESPONSABILITÉ CIVILE	12,5%
5. TEMPÊTE, GRÊLE, NEIGE	9,5%
6. CATASTROPHES NATURELLES	8,0%
7. BRIS DE GLACE	5,5%
8. DIVERS	14,5%

FIGURE 1.10 – MRH Source : FFSA Rapport annuel 2011

L'état oblige à assurer certains risques :

- RC civile automobile,
- RC locative,
- RC professionnelles dans de nombreuses professions (médicales, juridiques, financières, etc.)

1.3. MARCHÉ DE L'ASSURANCE NON-VIE

- RC décennale en assurance construction,
- RC chasse,
- etc.

L'inclusion de certaines garanties est obligatoire si une garantie principale est souscrite :

- En assurance automobile, la garantie catastrophes naturelles et la garantie catastrophes technologiques si au moins une garantie dommage accidentel a été souscrite.

1.3.5 Poids des branches

Dans une démarche de modélisation, il semble important de connaître quelles sont les branches les plus importantes et auxquelles le modèle doit s'adresser en premier lieu. Pour cela, on peut par exemple au chiffre d'affaires, c'est-à-dire aux primes reçues. La figure 1.11 donne les cotisations sur le marché français en 2011. Les primes liés aux branches automobiles et dommages aux biens représentent près de trois quarts des primes non-vie.

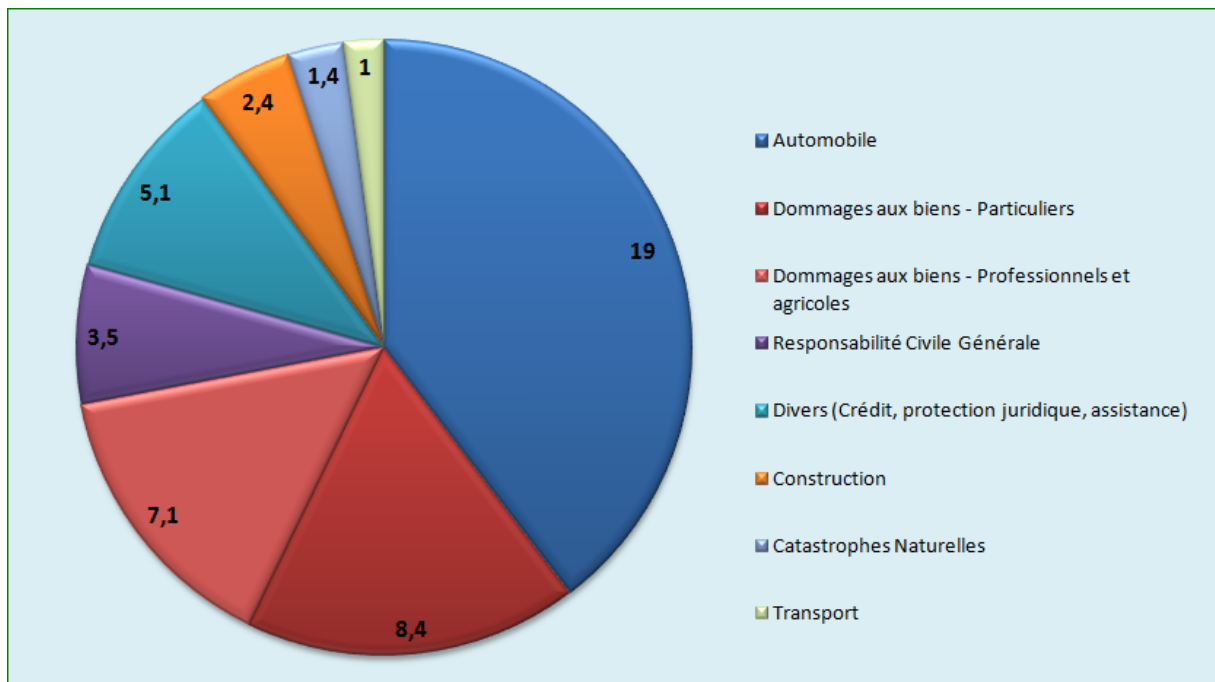


FIGURE 1.11 – Les cotisations sur le marché français en 2011 (affaires directes, en milliards d'euros, Source : FFSA [15])

On peut également s'intéresser aux prestations versées et aux dotations aux provisions (c.-à-d. la charge). Le graphique 1.12 présente ces données pour 2011 sur le marché français. Le poids de chaque branche est sensiblement équivalent à celui dans le chiffre d'affaire du marché non-vie français.

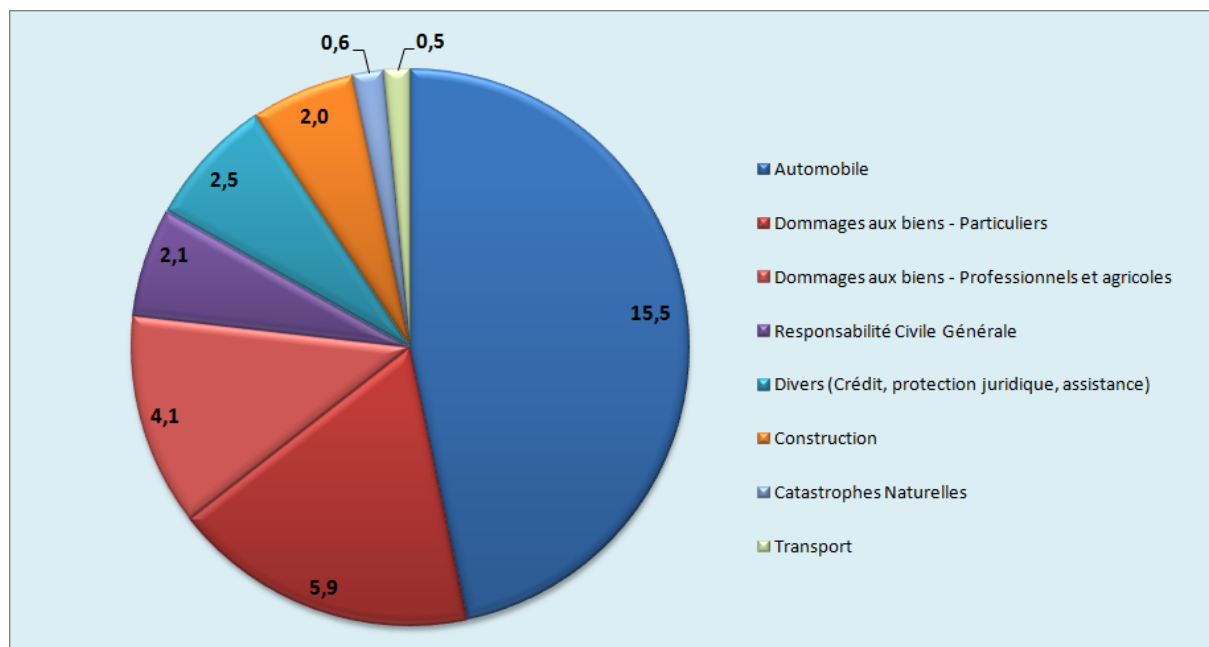


FIGURE 1.12 – La charge des sinistres par branche sur le marché français en 2011 (affaires directes, en milliards d'euros, Source : FFSA [15])

N.B. : Les affaires directes correspondent aux primes ou cotisations relatives à des contrats distribués directement ou par le biais d'un intermédiaire pour lequel l'assureur reste seul responsable du paiement des prestations vis à vis de l'assuré (c.-à-d. hors réassurance et rétrocession).

Un modèle ligne à ligne qui ferait des hypothèses assez réalistes au niveau de la modélisation des sinistres des branches automobiles et dommages permettrait donc de couvrir le *reserving* de 70% de la charge. Cependant, au niveau des provisions à l'ultime, les branches longues comme la Responsabilité Civile peuvent représenter davantage du poids.

1.4 La gestion des sinistres en pratique

Dans cette partie, nous présentons les principales étapes du processus de gestion, d'indemnisation et de provisionnement des sinistres en assurance non-vie. Nous mettons également en lumière quelques spécificités propres à certaines branches.

Comprendre les étapes-clefs dans la gestion de sinistre permet de mieux postuler un modèle statistique et de réfléchir à l'acceptabilité des hypothèses. Le déroulement d'un sinistre dépend fortement du type de risque considéré : pour certains types de sinistre, les prestations payées par l'assureur peuvent s'étendre sur des périodes de longue durée (« branches longues »). Il est en particulier possible que la proportion de sinistres survenus durant un exercice comptable mais non encore déclarés à la fin de celui-ci (les sinistres « tardifs ») soit très importante. Avoir à l'esprit le déroulement usuel pour les grands types de risque, ainsi que les particularités de certaines branches permet de mieux délimiter le champ d'application des modèles et leurs limites.

D'abord, donnons la définition d'un sinistre : c'est la réalisation de l'évènement prévu par la police d'assurance et entraînant la mise en jeu de la garantie. Il se décompose en un fait dommageable¹, à l'origine du dommage, et un préjudice résultant du dommage. Plus précisément : « le fait dommageable est celui qui constitue la cause génératrice du dommage. Un ensemble de faits dommageables ayant la même cause technique est assimilé à un fait dommageable unique. »

Par exemple, considérons une période de gel provoque la rupture d'une canalisation. Ici, le fait dommageable est le gel. Le sinistre survient lorsque la canalisation cède, et le préjudice résultant du dommage sont les dégâts causés par le liquide qui s'est écoulé de la canalisation. Un autre exemple est le cas de l'assurance construction : un vice de construction portant sur la solidité de l'ouvrage (fait dommageable) peut entraîner la survenance d'un dommage des années plus tard (p. ex. chute d'un balcon, effondrement de la charpente soutenant la toiture, etc.). Ainsi, la survenance du sinistre (du dommage) peut être plus ou moins rapprochée dans le temps du fait dommageable.

Ensuite, les principales étapes dans la vie d'un sinistre sont les suivantes (voir Figure 1.13) :

- la notification à l'assureur et l'ouverture du dossier sinistre par le gestionnaire de sinistre,
- la mise en œuvre de différentes expertises,
- le paiement d'indemnités à la victime,
- la clôture du dossier sinistre,
- d'éventuels recours menés par l'assureur contre le tiers responsable et/ou l'assureur de ce dernier.

Pour la description de ces différentes étapes, nous nous sommes notamment référés à Noël (2010) [36].

1. ou fait générateur (au choix)

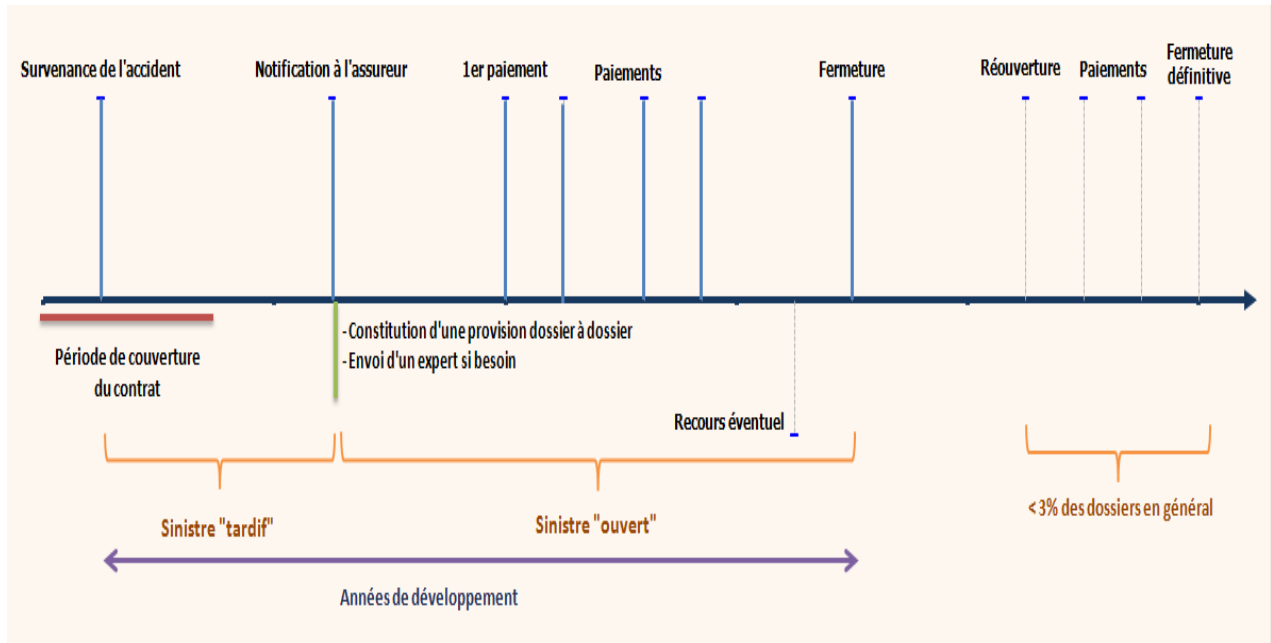


FIGURE 1.13 – Cycle de vie d'un sinistre

Dans les systèmes d'information des compagnies d'assurance, d'où sont extraites les données individuelles des sinistres, il est généralement possible d'extraire :

- les dates de survenance et de déclaration du sinistre,
- le montant et la date des paiements effectués, ainsi que des provisions dossier-dossier constituées par les gestionnaires de sinistres à différentes dates de vision,
- certaines données qualitatives : numéro d'identification du sinistre, caractère litigieux du sinistre, recours, date des expertises, agence, entité légale, etc.

Dans la suite, nous allons détailler chacune de ces étapes. Nous terminerons en abordant la gestion des provisions techniques au sein des compagnies.

1.4.1 La survenance du sinistre

La détermination de la date de survenance du sinistre est primordiale, puisque pour la majorité des contrats d'assurance non-vie (automobile, habitation, santé, etc.), l'exercice de rattachement de la sinistralité est l'exercice de survenance. Les primes acquises au titre d'un exercice sont mises en face des sinistres survenus au cours de la même période. L'établissement de la date de survenance permet à l'assureur de déterminer si la garantie était toujours en vigueur au moment des faits.

Cependant, la notion de date de survenance peut être plus ou moins difficile à définir. Par exemple :

- Dans le cas d'un sinistre automobile corporel ou matériel, la date de survenance est évidente car il s'agit de la date de l'accident (inscrite sur le constat amiable ou le procès-verbal établi par les forces de l'ordre).
- Dans le cas d'un dégât des eaux, la date de survenance est beaucoup moins évidentes : est-ce la date à laquelle la tuyauterie a commencé à fuir ou celle à laquelle les dégâts sont apparus à l'étage inférieur ? [6]

Le cas de l'assurance responsabilité civile (RC) est plus complexe. Il existe deux modes de fonctionnement pour l'application de la garantie dans le temps. On distingue ainsi :

- La garantie RC en « Faits dommageable » (FD) : elle peut être déclenchée si le FD survient durant la période de couverture de la garantie, quelque soit la date de réclamation (en tenant cependant compte des règles de prescription en vigueur).
- la garantie RC en « base réclamation » : la première réclamation adressée à l'assuré ou son assureur (par le tiers lésé) doivent être intervenues entre la date d'effet et la date d'expiration du contrat. C'est ici la date de première réclamation qui est le critère déterminant pour déterminer l'exercice de rattachement de la sinistralité.

Ainsi, si on prend l'exemple d'une réparation mal faite sur une installation (fait dommageable) qui génère une explosion faisant plusieurs victimes (survenance du dommage) :

- Si la garantie RC du professionnel qui a effectué la réparation est en FD, le sinistre est à rattacher aux primes acquises lors de l'exercice où cette réparation a été effectuée,
- Si la garantie est en « Base réclamation », le sinistre est à rattacher aux primes reçus lors de l'exercice de la date de première réclamation faite par les victimes de l'explosion.

Des règles fixées dans la loi¹ (clause de reprise du passé inconnu, existence d'une subsé-
quente, règles de priorité d'application entre les contrats) permettent de s'assurer la continuité de la couverture RC pour les activités passées de l'assuré et d'éviter les litiges entre assureurs successifs.

Dans certaines branches, un seul type de garantie est possible :

- Dans le cas de la RC vie privée des particuliers, seule la garantie FD est possible :
« Toutefois, lorsqu'elle couvre la responsabilité des personnes physiques en dehors de leur activité professionnelle, la garantie est déclenchée par le fait dommageable. » (Article L 124-5 du Code des Assurances)
- Dans certaines branches RC professionnelles. Par exemple, en responsabilité médicale, seule la garantie « base réclamation » est possible (Loi n° 2002-1577 du 30 décembre 2002).

1. Article 80 de la loi de Sécurité Financière du 1er août 2003, entrée en application le 2 novembre 2003

Liens avec les données ligne à ligne

Selon le type de garanties, le délai entre la date de survenance du dommage (ou du « fait dommageable ») et la première notification à l'assureur peut être très éloigné dans le temps. Dans les grandes lignes, on observe généralement :

- Peu de sinistres tardifs pour les branches dommages aux biens et santé. Les sinistres sont ordinairement connus et réglés en moins de 24 mois.
- Une proportion importante de sinistres tardifs pour les branches RC. Notamment, les assurances de responsabilité décennales en assurance construction connaissent des déroulés très longs : la garantie étant acquise pour 10 ans à compter de la date de réception des travaux peut conduire à des déclarations intervenant jusqu'à treize ans après la date d'ouverture du chantier.

En RC, le poids des sinistres tardifs devrait être moins important pour une garantie en base réclamation que pour une garantie en FD, puisque les provisions doivent servir à couvrir les sinistres pour lesquels une réclamation a lieu pendant la période de validité des contrats. Cependant, l'existence d'une subséquence implique des développements qui peuvent être longs pour certains sinistres.

Le rapport entre le stock de provisions et les flux de primes collectées au cours d'une année peut donner un premier indicateur de l'importance des tardifs (ainsi que des durées de développement) au sein d'une branche. Sur la figure 1.14, on observe ainsi, pour le marché français en 2011, un rapport provisions à primes très élevé pour les branches RC générale et Construction, tandis qu'il est faible pour les branches Dommages aux Biens.

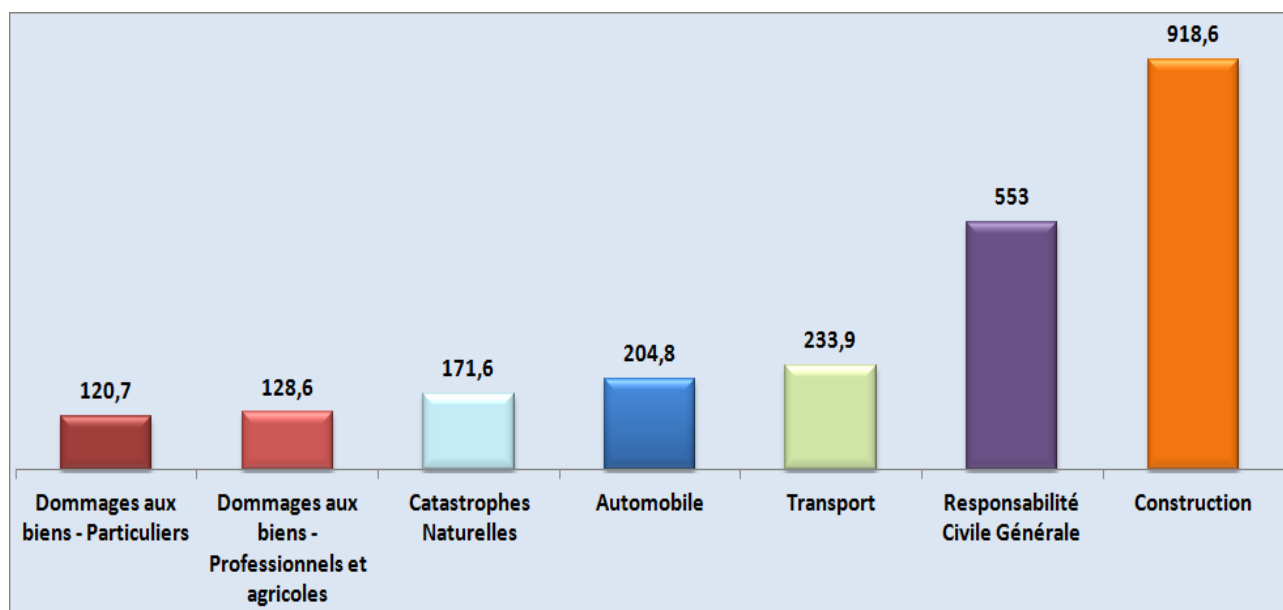


FIGURE 1.14 – Total des provisions en pourcentage des primes nettes, 2011, France (Source : FFSA [15])

Un autre moyen d'évaluer l'importance des tardifs est de s'intéresser aux cadences de règlements au sein d'une branche (voir figure 1.15). Les cadences donnent, pour une année de survenance donnée, les pourcentages de la charge ultime qui est réglé durant la première année de développement (c.-à-d. l'année de survenance, n), l'année de développement suivante ($n + 1$), etc.

On constate ainsi que pour les branches RC générale et Dommages Corporels en assurance automobile, moins de 15% des sinistres sont réglés après 1 an, ce qui implique des délais de déclaration longs et un déroulé des sinistres plus longs que dans d'autres branches (p. ex. Multirisque Habitation).

Règlements en	n	$n + 1$	$n + 2$	$n + 3$	$n + 4$
Multirisque Habitation	55%	90%	94%	95%	96%
Automobile	55%	79%	84%	89%	90%
<i>dont corporel</i>	13%	38%	50%	65%	72%
Responsabilité Civile	10%	25%	35%	40%	45%

FIGURE 1.15 – Cadences de règlements pour différentes branches (Source : Charpentier et Denuit, 2005 [6])

1.4.2 La déclaration à l'assureur

La première formalité dans le processus de gestion des sinistre est la déclaration du sinistre. Le client (ou un autre assureur, ou le garage conventionné, etc.) informe l'assureur de la survenance d'un sinistre qui est susceptible de mettre en jeu la garantie souscrite.

La déclaration se fait généralement sur papier libre, sur des imprimés spécifiques (bris de glace), ou bien par courriel, ou encore par téléphone.

La déclaration se doit d'être fidèle dans la description des circonstances et des conséquences. Dans le cas contraire, l'assuré s'expose à des sanctions contractuelles allant jusqu'à la déchéance (c.-à-d. perte du droit à l'indemnité, définies par les articles L113-2, L113-11 du code des assurances).

Délai de déclaration

Les délais de déclaration sont donnés par l'article L 113-2 du Code des Assurances.

L'assuré doit donner avis à l'assureur, dès qu'il en a eu connaissance, et au plus tard dans le délai fixé par le contrat, de tout sinistre de nature à entraîner la garantie de l'assureur. Ce délai contractuel ne peut être inférieur à cinq jours ouvrés (ramené à quatre jours ouvrés en cas de grêle, deux jours ouvrés en cas de vol et à 24 heures en cas de mortalité de bétail). Cependant, la perte de la garantie n'est pas automatiquement encourue en cas de non respect du délai de déclaration. Dans les faits, l'assureur doit prouver que ce retard lui a porté préjudice¹.

1. Loi Evin du 31 Décembre 1989 : " Lorsqu'elle est prévue par une clause du contrat, la déchéance pour déclaration tardive [...] ne peut être opposée à l'assuré que si l'assureur établit que le retard dans la déclaration lui a causé un préjudice."

1.4.3 L'ouverture du dossier sinistre

La déclaration entraîne l'instruction du dossier par l'assureur. Il recueille auprès de l'assuré un certain nombre d'éléments qui lui permettent de vérifier l'acquisition de la garantie, notamment :

- que le contrat est en règle (que la cotisation est acquittée, le contrat ne doit pas faire l'objet d'une suspension ou d'une résiliation pour non-paiement - L113-3) ;
- que le risque sinistré concerne bien une ou plusieurs des garanties souscrites ;
- que les circonstances autorisent l'application de la garantie.

Sur la base des informations collectées, le gestionnaire procède à une évaluation du sinistre. Il peut avoir recours à un forfait d'ouverture (souvent un coût moyen ultime estimé, voir encadré sur la page suivante), ou si le dossier est déjà bien étayé, à une évaluation plus spécifique pour constituer la provision dossier à dossier.

Cette estimation doit être complétée d'une estimation des frais allouables aux sinistres de manière individuelle par rapport au sinistre considéré, les ALAE (*Allocated Loss Adjustment Expense*). Il peut s'agir des honoraires d'avocats, des frais d'experts, des coûts administratifs, des frais de commissaire priseur, etc. Dans le cas de dommages causés par un tiers, la plupart de ces dépenses sont recouvrables par l'assureur s'il y a recours contre le tiers responsable. Dans la pratique, les ALAE et les montants d'indemnisation ne sont pas nécessairement distinguables dans les bases de données des assureurs.

Liens avec les données ligne à ligne

Il est possible d'observer de grandes différences dans les provisions d'ouverture dossier-dossier et l'estimation des ALAE selon les branches et les pratiques de l'assureur.

Par exemple, certains assureurs choisissent d'utiliser un unique coût moyen ultime estimé à l'ouverture du sinistre. (voir Savarre et Payre, 2012 [46]). D'autres assureurs effectuent des provisions d'ouverture plus différenciées, avant même le recours à l'expert ou à des évaluations plus détaillées, en intégrant des informations spécifiques à chaque cas (estimation au coût réel).

Les données individuelles sont donc très sensibles aux méthodes de gestion des sinistres retenues par la compagnie. Dans le cas de sinistres co-assurés, lorsque la compagnie est « pure » suiveuse (*follower*), les données collectées dépendent aussi des processus internes de la compagnie apériteuse (*leader*).

L'acquisition d'informations de plus en plus précises conduit à une provision dossier-dossier de plus en plus affinée pour chaque sinistre.

Méthodes des coûts moyens estimés à l'ultime

Ces méthodes consistent à évaluer la charge moyenne à l'ultime des sinistres d'un exercice en partant des coûts moyens (à l'ultime) estimés lors des exercices antérieurs. Il est nécessaire de prendre en compte l'inflation de la branche considérée : soit l'inflation du marché ^a, soit l'inflation observée dans le passé par la société. Pour estimer les coûts moyens à l'ultime des années de survenance précédentes, il faut disposer :

- d'un bon dénombrement des sinistres,
- de la charge totale à l'ultime pour les années de survenance passées considérées (soit en supposant que les sinistres sont totalement développés pour ces années-là, soit en ayant recours à des méthodes d'extrapolation).

Il est possible d'avoir recours à la triangulation par construction du triangle des coûts moyens à partir des triangles du nombre de sinistre et de charges des sinistres. Puis, le triangle des coûts moyens est projeté à l'aide de la méthode de triangulation retenue (p. ex. *Chain-Ladder*)

^a. Indice de réparation des véhicules publié par l'INSEE, indice de la Fédération Nationale du Bâtiment (FNB) pour les garanties Incendies, indices de l'Assemblée plénière des sociétés d'assurances dommage (ASPAD), etc.

1.4.4 L'expertise

Le gestionnaire de sinistre décide également s'il y a lieu de procéder à une expertise.

Les deux principales missions de l'expert sont de :

1. Vérifier la conformité du risque (coïncidence avec la réalité physique) et les éléments sur lesquels la tarification du contrat a été établie. Par exemple, dans le cas d'une garantie Habitation, l'expert peut être amené à vérifier le nombre de pièces principales indiquées, la surface pour une habitation louée, la nature des matériaux employés pour la construction, etc. La non-conformité du risque aux termes du contrat entraîne des sanctions (nullité L113-8 ou règle proportionnelle L113-9).

Liens avec les données ligne à ligne

Les cas d'irrégularités dans la déclaration faite par l'assuré peuvent amener à observer deux types de cas particuliers dans les données :

- Le cas des sinistres clos sans suite (*nil claims*) : l'assureur ouvre un dossier pour le sinistre, puis le clôturé sans qu'il y ait de paiements à l'assuré (déchéance suite à une fausse déclaration du sinistre, ou nullité du contrat dans le cas d'une fausse déclaration à caractère intentionnel lors de l'établissement du contrat).
- Dans le cas d'une application de règles proportionnelles, la réduction du forfait d'ouverture ou de l'évaluation du coût du sinistre faite par l'assureur.

2. Vérifier l'évaluation faite par l'assuré du sinistre (cas des sinistres dommages), ou bien établir une évaluation du sinistre (cas des expertises médicales dans le cas des accidents avec dommages corporels) (voir section suivante Indemnités).

1.4.5 Indemnités

1.4.5.1 Évaluation des indemnités

a) Cas des dommages matériels

Il appartient à l'assuré de chiffrer son préjudice, soit :

- en fournissant des devis et/ou des factures à l'assureur,
- en faisant appel à un « expert-assuré ». L'expert-assuré est un représentant de l'assuré, qui discute sur un pied d'égalité avec l'expert-compagnie.

L'assureur décide ensuite :

- de régler le dossier de gré à gré, sans avoir recours à un expert, sur la base de la facture ou du devis fourni par l'assuré. Les petits sinistres font ainsi généralement l'objet d'un règlement sur dossier.
- de vérifier l'évaluation de la demande faite par l'assuré.

L'assureur peut avoir recours à deux types de personnes pour effectuer cette vérification :

- un expert indépendant (profession libérale) ;
- un salarié ou un mandataire (agent ou courtier). Ce dernier est lié à la compagnie, mais censé réaliser son travail d'évaluation de manière équitable.

Plusieurs modes de résolution sont alors possibles :

- une expertise simple, dans le cas où il n'y a pas de tiers en cause. L'expert se rend sur place et évalue le montant des préjudices.
- une expertise contradictoire, qui met en présence l'expert-assuré et l'expert-compagnie. Les deux experts se mettent d'accord sur le montant du préjudice. Il est également possible qu'un expert représentant le tiers responsable du dommage (locataire, voisins, automobiliste, etc.) soit présent.

b) Cas des dommages corporels

L'assureur peut faire appel à un expert médical chargé d'évaluer le dommage : blessé léger, traumatisme crânien, troubles locomoteurs, etc.

Il est également possible de reconstituer l'accident et les vitesses des véhicules à partir de la déformation des pièces (études CESVI).

Les dommages corporels font également l'objet d'une étape de consolidation (au « moment où les lésions se fixent et prennent un caractère permanent »). La date de consolidation est fixée par le médecin expert ou judiciaire. C'est à ce moment qu'une offre d'indemnisation des préjudices permanents doit être faite, et que les préjudices temporaires sont soldés.

1.4.5.2 Litiges

En France, en cas de litiges, l'assuré ou la victime peut mener les actions suivantes :

- Saisir les instances internes de la compagnie d'assurance (service clientèle ou la direction de la qualité),

- Saisir le médiateur des assurances (médiateur de la Fédération française des sociétés d'assurances (FFSA) ou médiateur du Groupement des entreprises mutuelles d'assurance (GEMA)).
- Saisir l'autorité de contrôle prudentiel (ACP), s'il estime que les pratiques de la compagnie d'assurance relèvent d'anomalies et de dysfonctionnements.
- Enfin, saisir la justice. La juridiction à saisir dépend du montant d'indemnisation sur lequel porte le litige.

Liens avec les données ligne à ligne

Après expertise, la provision dossier-dossier n'est plus constituée selon le forfait d'ouverture mais avec une estimation au coût réel.

Sur certaines branches (par exemple bris de glace), ou pour certains sinistres dont le coût ne dépasse pas un certain seuil, une expertise n'est pas menée et seul le forfait d'ouverture est utilisé. La provision dossier-dossier n'évolue pas, jusqu'à la clôture du sinistre où le paiement réel apparaît.

Le délai entre la déclaration et l'estimation au coût réel dépend en partie des processus internes de chaque compagnie.

Les cas litigieux conduisent nécessairement à un développement plus long des sinistres. Les procédures peuvent dépendre des spécificités propres à chaque pays, branches, et compagnies, et rendent le traitement statistique plus difficile.

1.4.5.3 Le règlement de l'indemnité

Une fois l'indemnité fixée (parfois par la justice), l'assureur doit payer la victime dans le délai contractuel figurant dans la police (généralement 30 jours). Cependant, le règlement n'est généralement pas unique :

- Tout d'abord, il est possible que l'assureur ait déjà versé un ou plusieurs acomptes. C'est notamment le cas lorsque l'assuré doit faire face à des mesures d'urgence (hébergement dans un autre lieu, etc.). Dans les dommages corporels, l'assureur a déjà pris en charge tous les frais de santé immédiats de la victime, et le règlement de l'indemnité finale concerne les préjudices permanents.
- Ensuite, le règlement peut être fractionné. Par exemple, dans le cas d'une assurance en valeur à neuf : une évaluation définitive des dommages (en valeur à neuf) est d'abord effectuée. L'assureur verse dans un premier temps la valeur d'usage. Ensuite, la différence entre la valeur à neuf et la valeur d'usage est réglée par l'assureur sur présentation des factures (sans excéder le montant maximum de l'évaluation). En cas de contestation sur les factures, l'assureur peut avoir à tout moment recours à un expert.
- Dans le cas des dommages corporels, lorsqu'il y a transaction entre les parties, l'assureur peut s'orienter soit vers le versement d'une rente, soit le versement d'un capital.

Liens avec les données ligne à ligne

L'assuré dispose d'un délai de deux ans pour faire valoir son droit à l'indemnisation contractuelle, sinon il s'expose à une prescription qui lui fait perdre le droit à toute indemnité. Cependant, des causes d'interruption existe : assignation en justice, référé, opérations d'expertise, versement d'acompte. Dans les faits, le déroulement de sinistres même mineurs (p. ex. petit dégât des eaux) peut potentiellement s'étaler sur de nombreuses années. Dans le cas des méthodes agrégées type Chain-Ladder, des méthodes d'extrapolation sont généralement utilisées pour traiter la queue du triangle.

Le délai entre l'ouverture et le premier versement peut être très long, s'il y a un désaccord entre les parties prenantes et qu'un recours à la justice a lieu. Le cas des recours judiciaires est potentiellement à traiter à part au cours de l'étude d'une base de sinistres ligne à ligne (cas généralement minoritaires en nombre et spécifiques rendant difficile la postulation de propriétés statistiques et l'estimation).

1.4.6 Clôture, recours de l'assureur, réouverture

Clôture

La clôture intervient lorsque les règlements de la victime et des tiers-payeurs sont effectués.

Recours de l'assureur

L'assureur dispose, après avoir réglé son client, du droit à récupérer auprès des tiers responsables les sommes versées.

« L'assureur qui a payé l'indemnité est subrogé, jusqu'à concurrence de cette indemnité, dans les droits et actions de l'assuré contre les tiers, qui par leur fait ont causé le dommage ayant donné lieu à la responsabilité de l'assureur » (Article L 121-12 du code civil).

Dans les cas où l'assureur ne peut pas mener une action contre le tiers responsable¹, une action contre l'assureur de ce dernier reste généralement possible.

Dans d'autres cas, il existe des accords de convention entre les assureurs. C'est notamment le cas de la RC automobile matériel en France. La plupart des assureurs français ont signé la convention IRSA (Indemnisation Règlement des Sinistres Automobiles). L'assureur de la victime de l'accident indemnise directement cette dernière en lieu de l'assureur du responsable de l'accident, à travers la convention IDA (Indemnisation Directe de l'Assuré). L'assuré peut cependant demandé d'être soumis au droit commun (procédure classique d'indemnisation). Ensuite, l'assureur de la victime se retourne contre l'assureur de la partie adverse, selon les modalités établies par la convention IRSA. Ainsi, si le montant des dommages est inférieur à 6500 euros HT, le recours est forfaitaire, dans la limite de 1236 € (en 2013). Le recours exercé est proportionnel au niveau de la responsabilité de chaque partie dans l'accident (Barème IRSA). Source : <http://www.index-assurance.fr/pratique/sinistre/convention-irsa>

Réouverture

La réouverture concerne généralement peu de cas. Les cas de réouverture peuvent être plus fréquents pour les dommages corporels. Il s'agit notamment de réouvertures liées à des aggravations médicales, qui peuvent survenir très tardivement après la date de survenance du sinistre. En effet, la loi n°2008-561 du 17 juin 2008 (publié au JO le 18 juin 2008) a modifié les délais de prescription en la matière :

« L'action en responsabilité née à raison d'un événement ayant entraîné un dommage corporel, engagée par la victime directe ou indirecte des préjudices qui en résultent, se prescrit par dix ans à compter de la date de consolidation du dommage initial ou aggravé. »

Ainsi, une victime d'accident de la circulation qui subit un traumatisme crânien en 1995, qui est soignée et paraît guérie, mais qui présente en 2008 des troubles anormaux que les experts

1. Ces cas se présentent notamment dans les relations entre locataires et propriétaires :

- Clause de renonciation prévue dans le contrat d'assurance (l'assuré a renoncé à d'éventuels recours contre le locataire responsable)
- Renonciation à recours du propriétaire envers le locataire prévue dans le bail.

Par ailleurs, l'assureur ne peut pas avoir recours contre toute personne vivants au foyer (sauf en cas de malveillance de leur part).

rattachent à l'accident initial, disposera alors d'un nouveau délai de 10 ans pour agir en Justice pour réparer le préjudice aggravé ([10]).

D'autres types de réouverture sont possibles, notamment :

- économiques : mise à jour des paramètres économiques et règlementaires (p. ex. coût horaire d'une tierce personne) ;
- situationnelles (décès d'un parent en charge de la victime, ...).

Enfin, lorsque un tiers victime exerce un recours contre l'assuré, l'assuré peut mener une action en responsabilité contre son assureur pour être dédommagé : le délai de la prescription de deux ans ne court alors que du jour où ce tiers a exercé une action en justice contre l'assuré ou a été indemnisé par ce dernier. Cela peut engendrer des cas de réouverture de dossier : par exemple, si le gestionnaire de sinistres n'a pas réussi à établir tous les tiers victimes d'un accident susceptibles de se retourner contre l'assuré responsable.

Liens avec les données ligne à ligne

Les recours faisant l'objet d'une évaluation distincte (prévisions de recours restant à encaisser à la clôture de l'exercice inventorié), il conviendrait de les traiter à part s'il est possible de les distinguer dans les données sinistres individuelles. En effet, le travail à partir des charges de sinistres ou les provisions dossier-dossier brutes de recours permet de s'affranchir des spécificités propres à certains pays en terme de recours. Par exemple, en Malaisie, il existe des tribunaux islamiques pour les musulmans, et un système juridique laïc pour les non-musulmans. De plus, au niveau de la modélisation statistique, il peut s'avérer plus difficile de modéliser des paiements incrémentaux d'abord positifs (indemnisation de la victime) puis négatifs généralement en fin de cycle de vie du sinistre (recours de l'assureur). La décision d'effectuer des recours dépend aussi des *process* internes de l'entreprise (seuil minimum d'indemnités à partir duquel l'assureur est prêt à mener un recours, etc.)

Pour les petits sinistres RC automobile dommages matériels, dans certaines compagnies, les gestionnaires de sinistres inscrivent un forfait d'ouverture égal au forfait IRSA (qui est établi à partir d'un calcul de coût moyen) (Kelle, 2007 [28]).

Les cas de réouverture sont généralement trop peu nombreux pour que le recours à des méthodes statistiques soit possible.

1.4.7 La gestion des provisions de sinistres

Comment évoluent les provisions de sinistres au cours de la vie des sinistres ?

On peut décomposer la charge ultime vue à la date d'aujourd'hui pour une année de survenance en quatre éléments (voir figure 1.16) :

- Un premier bloc est constitué par la charge dossier-dossier (*Incurred*), qui se décompose en :
 - Les montants déjà réglés par l'assureur (*payments*),
 - Les provisions dossier-dossier (*Case reserves*),
- Ensuite, on distingue les provisions IBNR, composées de :
 - La provision IBNyR (*Incurred But Not yet Reported*) pour les sinistres tardifs,
 - La provision « IBNeR » (*Incurred But Not enough Reported*) qui complète les provisions dossier-dossier pour les sinistres survenus (dans certains cas, de manière négative, lorsque ces dernières sont trop prudentes).

Lorsque l'assureur n'a pas définitivement réglé tous les sinistres de l'année de survenance considérée, il s'agit d'une charge ultime estimée.

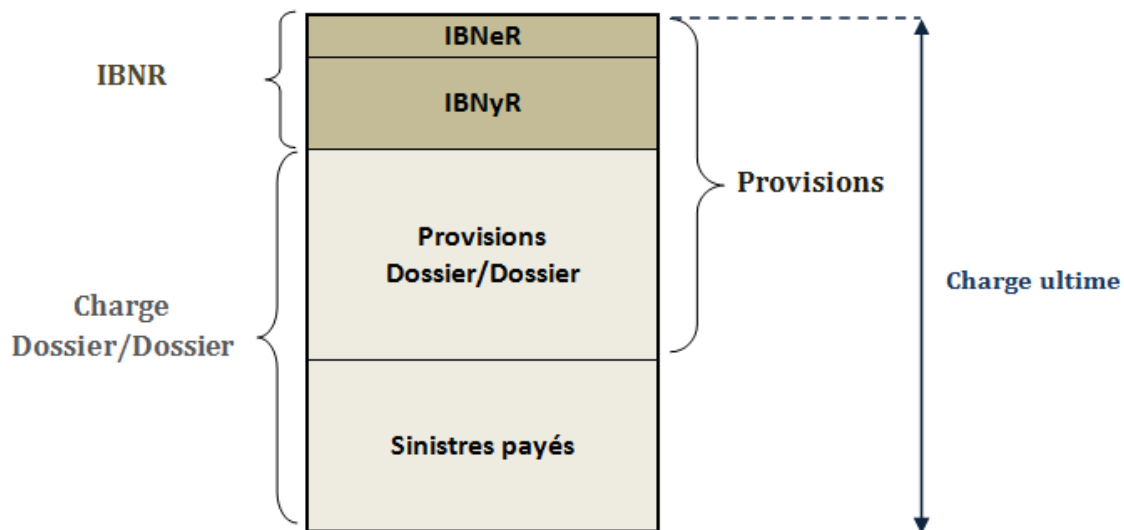


FIGURE 1.16 – Répartition de la charge ultime (estimée) des sinistres (dans le cas d'IBNeR positifs)

1.4.7.1 La charge dossier-dossier

La charge dossier-dossier pour les sinistres survenus (*incurred claims*) est constituée des montants qui ont déjà été réglés et de la provision dossier-dossier. La provision dossier-dossier est constituée par le gestionnaire de sinistres pour chaque sinistre déclaré à l'assureur. En fonction

des *process* internes à la compagnie d'assurance, cette provision est réévaluée plus ou moins fréquemment.

- A l'ouverture du sinistre, la charge dossier-dossier n'est constituée que d'une provision pour sinistre à payer,
- La provision dossier-dossier est diminuée des montants réglés au fur à mesure de la vie du sinistre,
- A la clôture du sinistre, la provision a été entièrement liquidée.

Les boni / mali de liquidation sont les variations de charge ultime (estimée) des sinistres (pour un même exercice de survenance) entre deux dates de vision (voir figure 1.17). Ces boni/mali peuvent survenir suite :

- à la ré-estimation des provisions dossier-dossier lors du passage du forfait d'ouverture à l'estimation au coût réel, de la revue des provisions ou de l'arrivée de nouvelles informations sur les sinistres,
- à des écarts constatés entre les règlements d'indemnités effectués et les provisions qui avaient été passées.

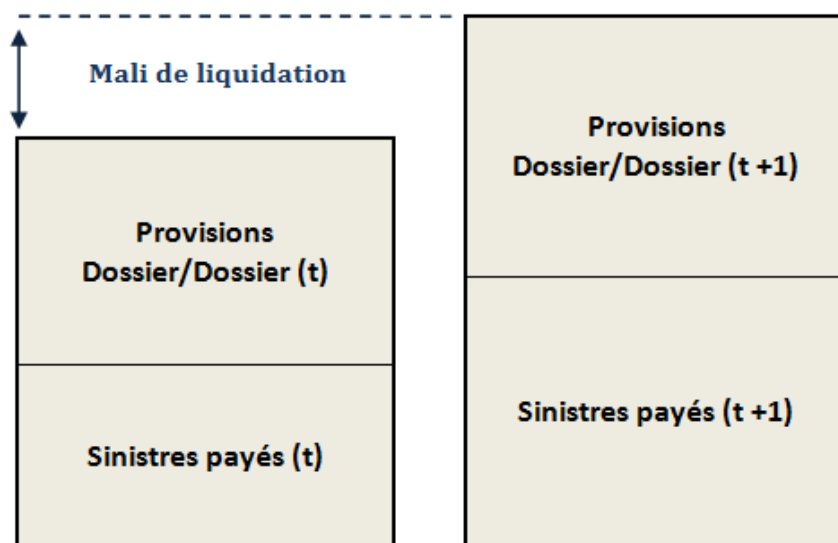


FIGURE 1.17 – Mali de liquidation pour la charge dossier-dossier de sinistres (de même année de survenance) estimée à deux dates différentes (t et t+1)

Selon les compagnies, la façon de constituer la charge dossier-dossier peut être différente :

- Les gestionnaires de sinistres peuvent chercher à approcher au plus près les charges ultimes des sinistres (vision *Best Estimate*) ; les provisions estimées sont alors suffisantes pour régler les sinistres en espérance, la marge de prudence nécessaire en réglementation française pour la PSAP étant ajoutée ensuite.
- Dans d'autres compagnies, les gestionnaires de sinistres peuvent avoir des estimations plus prudentes lors de la constitution des provisions dossier-dossier.

Par exemple : en cas de litige, lorsque l'issue d'un procès est incertaine, les gestionnaires de sinistre peuvent chercher à provisionner un montant suffisant dans 50% des issues (vision *Best Estimate*), suffisant 75% des issues (vision prudente), en se basant sur les observations passées ; ou encore faire appel à un jugement d'expert.

1.4.7.2 Les provisions IBNeR

La provision « IBNeR » (*Incurring But Not enough Reported*) correspond aux provisions pour sinistres déclarés mais non suffisamment (ou trop) provisionnés. Elle vient compléter les provisions dossier-dossier pour les sinistres ouverts. Elle est calculée à un niveau agrégée, et reflète des changements plus globaux : modifications des taux techniques pour les rentes qui impactent les provisions estimées pour le capital de la rente (cf. infra), changement dans le processus de gestion des sinistres, etc. La provision IBNeR peut être également négative si les provisions dossier-dossier sont trop prudentes. En raison des délais de gestion, il est aussi possible qu'une provision dossier-dossier n'est pas été constituée pour des sinistres déclarés à l'assureur. Dans la pratique, il peut cependant être difficile de distinguer la provision IBNeR des boni/mali de liquidation.

1.4.7.3 Les provisions pour sinistres tardifs (IBNyR)

Les réserves IBNyR permettent de prendre en compte les sinistres tardifs. Une des manières les plus simples de calculer les réserves IBNyR est la suivante :

- Un taux de sinistres tardifs sur la base des observations passées (cadence/ méthode Chain-Ladder sur le nombre de sinistres).
- Le nombre estimé de sinistres tardifs est ensuite multiplié par le coût moyen ultime des sinistres de l'année de survenance à laquelle ils sont rattachés (ou le coût moyen des sinistres déclarés tardivement sur les survenances précédentes).

1.4.7.4 Cas des rentes

Dans le cas des dommages corporels, le cas de versement d'une rente est particulier (voir Figure 1.18) :

- Avant l'attribution de la rente, des provisions sont constituées qui évaluent à la fois rente potentielle que l'assureur versera et toutes les autres indemnités annexes (frais de santé avant l'attribution de la rente, etc.)
- A l'attribution de la rente (c.-à-d. à la date du jugement ou de la transaction), un Capital Constitutif de Rente (CCR) est payé. Il correspond à la somme d'argent nécessaire à la date d'attribution de la rente pour payer la rente jusqu'à l'extinction de celle-ci.
- Dès le lendemain de l'attribution de la rente, une Provision Mathématique (PM) est constituée. Elle doit servir à payer les arrérages restant à payer d'une rente. Seule la PM est encadrée légalement, c'est-à-dire que le taux de capitalisation et la table de mortalité (le cas échéant) à utiliser sont donnés par la loi.

Les rentes prennent généralement en compte l'inflation (par exemple, indexation sur le taux d'inflation des pensions versées par la Sécurité Sociale), et peuvent être viagères. L'assureur doit alors se protéger du risque de longévité et du risque d'inflation. Le versement en rente ou en capital est une décision discrétionnaire du juge ou déterminé librement dans le cadre de la transaction [41].

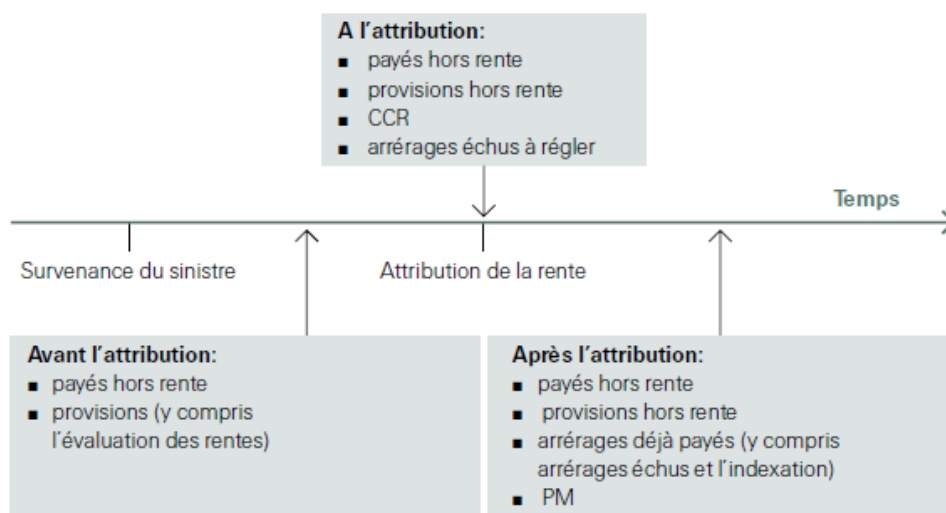


FIGURE 1.18 – Paiements et provisions dans le cas d'une rente (Source : Swiss Re [4])

Bilan

Dans cette partie, nous avons rappelé le rôle central de l'estimation des provisions dans les compagnies d'assurances dommage. Nous avons évoqué les enjeux réglementaires liés aux provisions, afin de cerner quelles contraintes les modèles de provisionnement devaient respecter (actualisation des flux futurs, mesure du risque de réserve, etc.). Enfin, nous avons décrit les étapes de la vie d'un sinistre, afin d'avoir une idée précise des informations qui seront disponibles dans les données ligne à ligne et des phénomènes que l'on pourrait observer (utilisation d'un forfait d'ouverture, estimation au coût réel, recours, réouverture, etc.).

Nous allons maintenant exposer les méthodes de provisionnement classique agrégées, qui serviront de référence pour le modèle ligne à ligne. La méthode Chain-Ladder servira de point de comparaison à l'estimation de la moyenne des réserves. Les méthodes stochastiques permettront de comparer les distributions prédictives des provisions obtenues dans le cadre ligne à ligne et dans le cadre agrégé.

Partie 2

Méthodes de provisionnement classiques

Dans cette partie, nous présentons les principales méthodes agrégées, dites classiques, notamment la méthode Chain-Ladder, qui constitue une méthode standard pour évaluer les provisions de sinistres à payer. Nous présenterons les limites des méthodes classiques et dirons quels peuvent être les avantages du recours à une modélisation individuelle, en utilisant les données ligne à ligne de la vie des sinistres.

2.1 Construction du triangle de développement

Les méthodes agrégées sont basées sur les triangles de développement (*run-off triangles*) reflétant la dynamique des sinistres. Les données étudiées (notamment dans le cadre de Chain-Ladder) peuvent être de différentes natures : règlements, charges, nombre de sinistres, ALAE, etc. Dans la suite, nous exposons les méthodes pour les triangles de règlements. Les données peuvent également être rapportées à des périodes différentes : années, semestres, etc. Par la suite, on va considérer des triangles de développement annuels. On se place aussi dans le cadre de l'étude d'un portefeuille pré-défini.

Le triangle de développement est un tableau à double entrée : les lignes correspondent aux années de survenance des sinistres, les colonnes aux années de développement des sinistres.

Nous introduisons les notations suivantes :

- L'exercice de rattachement de la sinistralité (c.-à-d. l'année de survenance) désignée par la lettre i , $i = 0, \dots, n$;
- Les années de développement (ou de déroulé de sinistres), désignée par la lettre j , $j = 0, \dots, n$;
- $X_{i,j}$ désigne le paiement incrémental réglé durant l'année de développement j pour les sinistres survenus durant l'année i ;
- $C_{i,j}$ désigne les règlements cumulés réglés de l'année 0 à l'année de développement j , pour les sinistres survenus durant l'année i : $C_{i,j} = \sum_{k=0}^j X_{i,k}$

On suppose que n désigne l'année actuelle, c'est-à-dire l'année d'inventaire où l'étude de provisionnement est menée. Pour les sinistres survenus en année i , on observe le développement jusqu'à l'année $j = n - i$. On parle ainsi de triangle, car au-delà de la dernière diagonale, les

2.1. CONSTRUCTION DU TRIANGLE DE DÉVELOPPEMENT

montants ne sont pas connus (voir figure 2.1). Il s'agit de les prévoir pour déterminer les réserves à effectuer.

$i \backslash j$	0	1	...	j	...	$n-1$	n
0	$C_{0,0}$	$C_{0,1}$...	$C_{0,j}$...	$C_{0,n-1}$	$C_{0,n}$
1	$C_{1,0}$	$C_{1,1}$...	$C_{1,j}$...	$C_{1,n-1}$	
⋮	⋱		
i	$C_{i,0}$	$C_{i,1}$...	$C_{i,j}$			
⋮	⋱				
$n-1$	$C_{n-1,0}$	$C_{n-1,1}$					
n	$C_{n,0}$						

TABLE 2.1 – Triangle des paiements cumulés

On suppose généralement que les sinistres sont ou seront tous clos après $n+1$ années de développement, quelle que soit la date de survenance. En particulier, pour les sinistres survenus en année $i=0$, les sinistres sont tous clos à la date actuelle n . Si ce n'est pas, il faut utiliser des méthodes d'extrapolation pour modéliser les années de développement non connus (voir partie 2.6).

Il est également possible de travailler avec des trapèzes lorsque l'on a des historiques de données plus profonds. Lors de revues de provisions, les gestionnaires de risques peuvent également s'intéresser seulement aux dernières diagonales connues du triangle, lorsqu'ils ne se focalisent que sur les évolutions récentes, ou que l'information antérieure n'est pas disponible (voir figure 2.1).

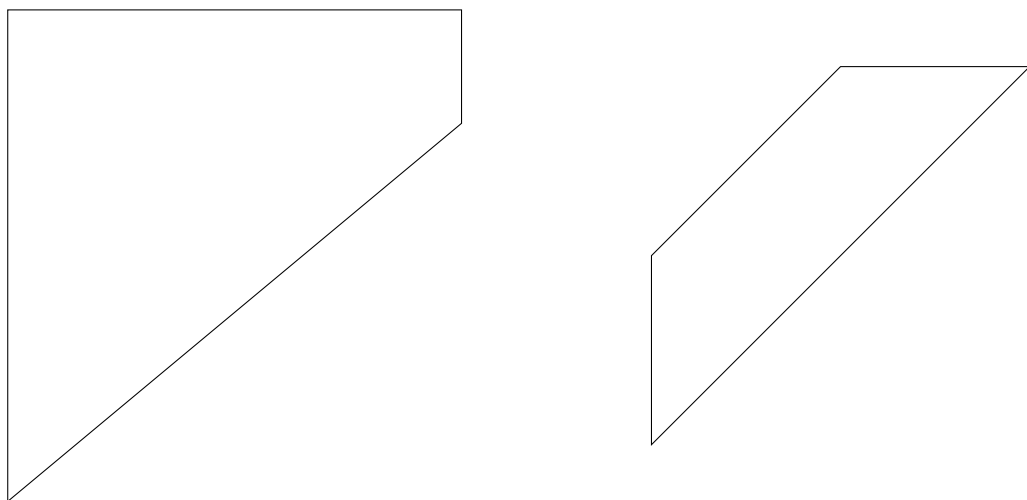


FIGURE 2.1 – Autres formes d'agrégation des données possibles

Nous travaillons désormais avec le triangle des règlements cumulés. L'information disponible

apportée par une ligne i du triangle est :

$$\mathcal{H}_{i,n} = \{(C_{i,j}) | 0 \leq j \leq n - i\}$$

L'information disponible est notée \mathcal{H}_n :

$$\mathcal{H}_n = \{(C_{i,j}) | 0 \leq i + j \leq n\}$$

Les diagonales correspondent aux années calendaires. On peut voir apparaître certains changements dans le processus de gestion des sinistres (révision de la réserve de tous les dossiers quelle que soit l'année de survenance), ou les effets de l'inflation.

Après avoir estimé la partie basse du triangle de liquidation par les diverses méthodes décrites ci-dessous, ($\{\hat{C}_{i,j} | i + j > n, i \leq n, j \leq n\}$), il est possible de déterminer :

- la réserve à un an : en faisant la différence entre la somme de la première sous-diagonale et la somme de la diagonale n .
- la réserve à l'ultime, en faisant la différence de la somme de la dernière colonne et de la somme de la diagonale n .

Au lieu de modéliser directement les paiements cumulés futurs, certaines méthodes (dont Chain-Ladder) s'intéressent aux facteurs individuels de développement (ou *link-ratios*). Ces derniers sont définis par :

$$f_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}, 0 \leq i \leq n - 1, 0 \leq j \leq n - 1 \quad (2.1)$$

Remarque

Notons qu'au lieu de calculer des facteurs de développement, on peut également s'intéresser à des cadences de développement sur les règlements cumulés ($\gamma_{i,j}$) ou incrémentaux ($\varphi_{i,j}$) :

$$C_{i,j} = \gamma_{i,j} C_{i,n} \text{ ou } X_{i,j} = \varphi_{i,j} C_{i,n}$$

Les relations entre γ , φ et f sont les suivantes :

$$\gamma_{i,j} = \prod_{k=j}^{n-1} \frac{1}{\lambda_{i,k}} \text{ et } \varphi_{i,j} = \begin{cases} \gamma_{i,j} & \text{si } j = 0 \\ \gamma_{i,j} - \gamma_{i,j-1} & \text{si } j > 0 \end{cases}$$

2.2 La méthode de référence : Chain-Ladder

La méthode Chain-Ladder constitue certainement la méthode de provisionnement la plus simple et la plus utilisée par les compagnies d'assurance.

2.2.1 Hypothèses et estimation

La méthode Chain-Ladder est fondée sur l'hypothèse de proportionnalité entre les années de règlements. Elle s'applique aux triangles des règlements cumulés.

Formellement, l'hypothèse est la suivante :

Hypothèse 2.2.1 *Pour $j = 1, \dots, n$, les facteurs de développement sont indépendants de l'année de survenance du sinistre i .*

Sous cette hypothèse, on a :

$$\frac{C_{0,j+1}}{C_{0,j}} = \frac{C_{1,j+1}}{C_{1,j}} = \dots = \frac{C_{n-j-1,j+1}}{C_{n-j-1,j}}.$$

On estime alors le facteur de développement Chain-Ladder lié à l'année de développement j par :

$$\hat{f}_j = \frac{\sum_{i=0}^{n-j-1} C_{i,j+1}}{\sum_{i=0}^{n-j-1} C_{i,j}}, j = 0, \dots, n-1. \quad (2.2)$$

Ainsi, sous l'hypothèse de Chain-Ladder 2.2.1, au lieu de faire une moyenne simple des facteurs de développement individuels, on estime \hat{f}_j par une moyenne pondérée :

$$\hat{f}_j = \sum_{i=0}^{n-j-1} \frac{C_{i,j}}{\sum_{i=0}^{n-j-1} C_{i,j}} \frac{C_{i,j+1}}{C_{i,j}}, j = 0, \dots, n-1. \quad (2.3)$$

L'exposition au risque est ainsi prise en compte, c'est-à-dire que l'on donne plus d'importance aux années où les paiements cumulés sont les plus importants. La croissance du portefeuille est ainsi implicitement prise en compte.

Une fois que les facteurs sont estimés, l'estimation des paiements cumulés futurs est donné par :

$$\hat{C}_{i,j} = \hat{f}_{j-1} \times \dots \times \hat{f}_{n-i} \times C_{i,n-i} \quad (2.4)$$

Le triangle est ainsi complétée (voir figure 2.2).

La réserve à un an pour les sinistres survenus durant l'année i vaut¹ :

$$\hat{R}_i^1 = \hat{C}_{i,n-i+1} - C_{i,n-i}, i = 1, \dots, n$$

La réserve à un an pour l'ensemble des sinistres vaut :

1. Avec le triangle des paiements cumulés, la réserve à un an est à entendre comme les paiements incrémentaux qui seront à effectuer à horizon d'un an.

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

La réserve à l'ultime pour les sinistres survenus durant l'année i vaut :

$$\hat{R}_i^u = \hat{C}_{i,n} - C_{i,n-i}, i = 1, \dots, n$$

Enfin, la réserve à l'ultime pour l'ensemble des sinistres vaut :

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

$i \backslash j$	0	1	...	j	...	$n-1$	n	R_i^u
0	$C_{0,0}$	$C_{0,1}$...	$C_{0,j}$...	$C_{0,n-1}$	$C_{0,n}$	0
1	$C_{1,0}$	$C_{1,1}$...	$C_{1,j}$...	$C_{1,n-1}$	$\hat{C}_{1,n}$	$\hat{C}_{1,n} - C_{1,n-1}$
⋮
i	$C_{i,0}$	$C_{i,1}$...	$C_{i,j}$	$\hat{C}_{i,n}$	$\hat{C}_{i,n} - C_{i,j}$
⋮
$n-1$	$C_{n-1,0}$	$C_{n-1,1}$...	$\hat{C}_{n-1,j}$	$\hat{C}_{n-1,n}$	$\hat{C}_{n-1,n} - C_{n-1,1}$
n	$C_{n,0}$	$\hat{C}_{n,1}$	$\hat{C}_{n,n}$	$\hat{C}_{n,n} - C_{n,0}$

TABLE 2.2 – Triangle complété de paiements cumulés

2.2.2 Validation de Chain Ladder

La validation de la méthode Chain-Ladder est purement empirique, puisqu'il s'agit d'une méthode déterministe.

2.2.2.1 Triangle des facteurs de développement (d-triangle)

Il s'agit du triangle constitué des facteurs individuels de développement. On calcule alors la moyenne, l'écart-type, et le coefficient de variation (défini comme le rapport de l'écart-type à la moyenne) de ces facteurs par année de développement. Un faible coefficient de variation permet de valider empiriquement l'hypothèse de Chain-Ladder. L'examen de ce triangle permet de repérer les facteurs atypiques.

2.2.2.2 C-C plots

Il s'agit d'une méthode graphique [41]. Sous l'hypothèse Chain-Ladder 2.2.1, pour une année de développement donné $j = 0, 1, \dots$ les couples $(C_{i,j}, C_{i,j+1})_{i=0, \dots, n-j-1}$ doivent être alignés par une droite passant par l'origine.

Il est également possible d'effectuer une régression linéaire, pour vérifier que l'ordonnée à l'origine estimée n'est pas significative (c.-à-d. on rejette l'hypothèse qu'elle est statistiquement différente de 0), et que les écarts entre les points et la droite d'ajustement (résidus) sont faibles.

2.2.2.3 Triangle rétrospectif

Il est également possible d'appliquer de manière rétrospective les facteurs de développement aux triangles, et de comparer les écarts entre les données historiques observées et les données estimées avec les facteurs de développement Chain-Ladder.

2.2.3 Limites de la méthode Chain-Ladder

Pour que l'hypothèse d'indépendance des facteurs de développement avec les années de survenance soit valide, il faut que les conditions suivantes sur les données soit réunies (Partrat *et al.*, 2007 [41]) :

- Le passé doit être régulier. Par exemple, il ne doit pas y avoir de changements dans la politique de gestion des dossiers, notamment une accélération des remboursements au fil des ans. Une évolution dans la composition du portefeuille (changement dans la fréquence et/ou des coûts moyens, changements de souscripteurs ou des modalités de souscription, ...) rendra l'estimation des facteurs plus difficile. Dans ce cas-là, une subdivision du triangle peut être envisagé.
- Le présent et le futur doivent être structurellement peu différents du passé. Un changement futur de jurisprudence n'est par exemple pas pris en compte. De même, au niveau de l'inflation, l'inflation passée est implicite dans le calcul des facteurs de développement. Il faut donc supposer que l'inflation moyenne passée (pondérée) sera répétée dans le futur.
- La branche doit être peu volatile. Notamment, si pour une année ou plusieurs années données, des sinistres graves se produisent, ils doivent généralement être traités à part. S'il y a une valeur aberrante (*outlier*) dans la dernière diagonale connue, cet valeur sera projetée jusqu'à la fin.
- Les données doivent être nombreuses et fiables : il faut un portefeuille de taille suffisante.

Un autre désavantage, commun aux méthode agrégées, est qu'il n'est pas possible d'appliquer de la réassurance non proportionnelle pour calculer les réserves cédées.

Il y a également très peu de données disponible pour les derniers facteurs de développement f_{n-1}, f_{n-2} , etc. De plus, l'estimation de ces derniers facteurs est basé sur des sinistres survenus dans les années de survenance les plus anciennes, dont le développement peut ne plus être représentatif par rapport au développement des sinistres survenus récemment.

Le calcul des réserves estimées par Chain-Ladder est fortement sensible aux premières valeurs du triangle. Ainsi, une baisse de 1% de la cadence du premier paiement $\frac{C_{i,0}}{C_{i,n}}$ peut entraîner une baisse de 10% de la charge finale estimée (Habib et Riban, 2012 [22]).

2.2.4 Autres modèles déterministes

Plusieurs variantes déterministes de Chain-Ladder ont été introduites :

- La méthode **London-Chain**. Elle a été introduite par Benjamin et Eagles (1986). Au lieu de considérer une relation affine entre les $C_{i,j}$ et les $C_{i,j+1}$, on considère une relation linéaire, autorisant le cas où l'ordonnée à l'origine n'est pas nulle. On suppose ainsi que : $\forall j \in \{0, \dots, n-1\}, \exists f_j, \alpha_j$ tels que :

$$\forall i \in \{0, \dots, 1, \dots, n-j-1\}, C_{i,j+1} = f_j C_{i,j} + \alpha_j$$

Le modèle est estimé par moindres carrés ordinaires. Un inconvénient de ce modèle est le grand nombre de paramètre.

- La méthode **London-Pivot** (Straub, 1988) est à mi-chemin entre Chain-Ladder et London-Chain. La relation supposée est cette fois-ci : $\forall j \in \{0, \dots, n-1\}, \exists f_j, \alpha$ tels que :

$$\forall i \in \{0, \dots, 1, \dots, n-j-1\} = C_{i,j+1} = f_j(C_{i,j} + \alpha)$$

Cette méthode peut être mise en œuvre lorsque les droites des C-C plots, pour $j = 0, 1, \dots$, sont concourantes en un point $(a, -a)$.

- La méthode de **De Vylder**. Il s'agit d'une méthode basée sur les paiements incrémentaux. L'hypothèse faite est que le règlement incrémental de l'année de survenance i et de développement j est un pourcentage du règlement total versé par la compagnie pour les années de développement j [22].
- On peut également citer la méthode Bornhuetter-Ferguson, qui incorpore des hypothèses exogènes sur la charge finale.

Les méthodes déterministes sont simples à implémenter, mais elles ne donnent qu'une estimation de la charge ultime et sont sensibles à des petites variations dans les données utilisées. Pour obtenir de l'information sur la volatilité des provisions, il faut recourir à des modèles stochastiques. Avec les modèles stochastiques, il est possible de valider les hypothèses du modèle par une analyse des résidus. La détection des irrégularités potentielles dans le triangle est ainsi facilitée.

2.3 Chain Ladder Stochastique : Mack (1993)

Le modèle de Mack (1993) ([31]) est un modèle non-paramétrique (*distribution free*) qui permet d'estimer une marge d'erreur sur le montant des provisions.

La méthode de Mack constitue le pendant de la méthode Chain-Ladder au niveau stochastique, dans le sens où, sous les hypothèses du modèle, l'estimateur Chain-Ladder de $C_{i,n}$ est un estimateur sans biais de $\mathbb{E}[C_{i,n}|\mathcal{H}_n]$ (c.-à-d. de l'espérance de la charge ultime sachant l'information disponible aujourd'hui). Dans le cadre de notre mémoire, ce modèle est intéressant car nous chercherons également à faire le parallèle entre notre estimateur dans le modèle ligne à ligne et l'estimateur Chain-Ladder. De plus, le modèle peut être vu comme Markovien, et c'est également ce type de modèle récursif que nous retiendrons dans notre approche ligne à ligne.

2.3.1 Hypothèses

Le modèle de Mack repose sur les trois hypothèses suivantes :

Hypothèse 2.3.1 *Les paiements cumulés $(C_{i,j})_{j \geq 0}$ sont indépendants année de survenance par année de survenance i , c.-à-d. :*

$\forall i \neq i', (C_{i,j})_{j \geq 0}$ et $(C_{i',j})_{j \geq 0}$ sont indépendants.

Hypothèse 2.3.2 *Il existe des paramètres f_0, \dots, f_{n-1} tels que :*

$$\mathbb{E}[C_{i,j+1}|C_{i,j}, \dots, C_{i,0}] = \mathbb{E}[C_{i,j+1}|C_{i,j}] = f_j \times C_{i,j}$$

Hypothèse 2.3.3 *Il existe des paramètres $\sigma_0, \dots, \sigma_{n-1}$ tels que :*

$$\text{Var}[C_{i,j+1}|C_{i,j}, \dots, C_{i,0}] = \text{Var}[C_{i,j+1}|C_{i,j}] = \sigma_j^2 \times C_{i,j} \text{ pour } j = 0, \dots, n-1$$

2.3.2 Estimation

Sous les hypothèses 2.3.1 et 2.3.2, on peut montrer, en utilisant la loi des espérances itérées, que :

$$\mathbb{E}[C_{i,n}|\mathcal{H}_n] = f_{j-1} \times f_j - 2 \times \dots \times f_{n-i} C_{i,n-i}$$

On montre ensuite que les estimateurs Chain-Ladder \hat{f}_j (2.2) sont des estimateurs non biaisés et non corrélés de f_j . Ainsi, l'estimateur Chain-Ladder de la charge ultime pour l'année i (2.4) est un estimateur non biaisé de $\mathbb{E}[C_{i,n}|\mathcal{H}_n]$ ([23]).

Un estimateur de la variance est donné par :

– Pour $j = 0, 1, 2, \dots, n-2$:

$$\hat{\sigma}_j^2 = \frac{1}{n-j-1} \sum_{i=0}^{n-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)$$

– Si le dernier facteur de développement est égal à un et que les règlements sont supposés constants (tous les sinistres sont clos à cette date en n), on a $\hat{\sigma}_{n-1}^2 = 0$. Dans le cas contraire, on estime σ_{n-1}^2 par :

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

2.3.3 Validation du modèle

L'hypothèse 2.3.1 peut être validée de la même manière qu'avec Chain-Ladder, en étudiant la volatilité des facteurs de développement individuels $f_{i,j}$.

De même, l'hypothèse 2.3.2 peut être validé graphiquement en utilisant des C-C plots. Il est également possible d'effectuer un test de non corrélation des facteurs de développement successifs, comme le test de Spearman (Hubert et Guay (2009) [25]).

Pour tester l'hypothèse 2.3.3, on peut tracer le graphique des résidus normalisés en fonction de $C_{i,j}$ ([17]), où les résidus normalisés sont définies par :

$$r_{i,j} = \frac{C_{i,j} - \hat{f}_j C_{i,j}}{\sqrt{C_{i,j}}}$$

Les résidus ne doivent pas présenter de tendance.

2.3.4 Erreurs de prédiction

Le modèle de Mack permet d'obtenir une estimateur de la variance des provisions :

$$\widehat{\text{Var}}[R_i^u | \mathcal{H}_n] = \hat{C}_{i,n}^2 \sum_{j=n-i}^{n-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j} \frac{1}{\sum_{i=0}^{n-j} C_{i,j}}$$

L'erreur de prédiction MSEP (*Mean Squared Error of Prediction*) de l'estimation de la charge ultime \hat{R}_i^u pour les sinistres survenus durant l'année i est donnée par :

$$\text{MSEP}(\hat{R}_i^u) = \mathbb{E} \left[(\hat{R}_i^u - R_i^u)^2 | \mathcal{H}_n \right]$$

La MSEP quantifie la différence entre l'estimation de la provision \hat{R}_i^u et la vraie valeur (aléatoire) R_i^u , la différence étant mesurée en terme d'écarts quadratiques espérés. On montre que :

$$\text{MSEP}(\hat{R}_i^u) = \widehat{\text{Var}}[R_i^u | \mathcal{H}_n] + (\mathbb{E}[(R_i^u | \mathcal{H}_n)] - \hat{R}_i^u)^2$$

Une estimation de la MSEP pour les réserves à l'ultime correspondant aux sinistres survenus en année i est alors :

$$\widehat{\text{MSEP}}(\hat{R}_i^u) = \hat{C}_{i,n}^2 \sum_{j=n-i}^{n-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j} \left(\frac{1}{\hat{C}_{i,j}} + \frac{1}{\sum_{i=0}^{n-j} C_{i,j}} \right) \quad (2.5)$$

De même, Mack dérive un estimateur de la MSEP pour la réserve ultime totale.

Comme Mack considère un modèle non-paramétrique, avec seulement des hypothèses sur les deux premiers moments, il n'est pas possible d'obtenir la distribution complète des réserves. Nous disposons seulement d'estimateurs des deux premiers moments. Il n'est donc pas possible de calculer avec ce modèle une VaR à 99,5% par exemple, dans une perspective de solvabilité. Cependant, en pratique, une hypothèse de distribution sur les réserves est généralement faites (p. ex. normales ou log-normales) afin d'obtenir des quantiles des provisions à l'ultime. Les modèles GLM, présentés dans la section suivante, sont eux des modèles paramétriques.

2.4 Modélisation GLM

Les modèles linéaires généralisés sont utilisés pour évaluer et quantifier la relation entre une variable réponse (expliquée) et des variables explicatives. Ils peuvent être appréhendés comme une généralisation du modèle linéaire normal.

2.4.1 Modèles GLM

2.4.1.1 Composante aléatoire

On suppose que les paiements incrémentaux sont indépendants et suivent une loi appartenant à la famille exponentielle.

La fonction de densité d'une variable aléatoire $X_{i,j}$ est dit de type exponentielle si elle peut s'écrire sous la forme :

$$f(x_{i,j}, \theta_{i,j}, \phi) = \exp \left\{ \frac{\omega_{i,j}}{\phi} [\theta_{i,j} x_{i,j} - b(\theta_{i,j})] + c(x_{i,j}, \omega_{i,j}, \phi) \right\}, \quad (2.6)$$

où :

- $\theta_{i,j}$ est un paramètre réel ;
- $\omega_{i,j}$ est une pondération donnée pour la variable ; aléatoire $Y_{i,j}$ (p. ex. durée d'exposition au risque). Par la suite, on suppose qu'il n'y a pas de pondération ($\omega_{i,j} = 1$) ;
- $\phi > 0$ est un paramètre de dispersion indépendant de i et j ;
- $b(\cdot)$ est une fonction suffisamment régulière (de classe C^2) ;
- $c(\cdot)$ est une fonction ne dépendant pas de $\theta_{i,j}$.

La loi normale appartient à la famille exponentielle, mais également de nombreuses autres lois usuelles : poisson, gamma, binomiale...

2.4.1.2 Fonction lien

C'est la fonction qui fait le lien entre la moyenne des règlements incrémentaux $\mathbb{E}[X_{i,j}] = \mu_{i,j}$ et les variables explicatives. On suppose qu'une transformation de la moyenne est linéairement reliée aux variables explicatives, au travers d'une fonction lien $g(\cdot)$:

$$\eta_{i,j} = g(\mu_{i,j}) = Z'_{i,j} \gamma,$$

où $\eta_{i,j}$ est appelé la composante systématique, $Z_{i,j}$ est la matrice des variables explicatives et γ le vecteur des paramètres.

Dans le cadre du provisionnement, on retient généralement comme composante systématique $\eta_{i,j} = \mu + \alpha_i + \beta_j$, avec :

- Le paramètre μ qui capture l'effet des années calendaire (inflation) ;
- Les paramètres (α_i) , qui capturent l'effet des années de survenance ;
- Les paramètres (β_j) , qui capturent l'effet des années de développement.

Pour que le modèle soit identifiable, il faut poser par exemple $\alpha_0 = \beta_0 = 0$.

Si la fonction lien retenue est la fonction log, on parle de lien multiplicatif, puisque on a alors :

$$\mu_{i,j} = \exp(\mu) \exp(\alpha_i) \exp(\beta_j).$$

Un modèle GLM introduit généralement de l'hétéroscédasticité. On a en effet :

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

c'est-à-dire que la variance est fonction de la moyenne. Dans le cas d'une distribution normale, on a $v(\mu_{i,j}) = 1$ (homoscédasticité).

2.4.2 Estimation du modèle

Le modèle est estimé par maximum de vraisemblance. La log-vraisemblance s'écrit :

$$l(\beta, \phi, y) = \frac{1}{\phi} \sum_{i=1}^n (y_i \theta_i - b(\theta_i)) + \sum_{i=1}^n c(y_i, \phi)$$

En dérivant la log-vraisemblance par rapport au vecteur des paramètres γ , les équations de vraisemblance s'écrivent ([23]) :

$$\sum_{0 \leq i,j \leq n} \frac{x_{i,j} - b'(\gamma_{i,j})}{b''(\gamma_{i,j})} \frac{\partial \mu_{i,j}}{\partial \eta_{i,j}} b_{i,j}^{(k)} = 0, \text{ pour } k = 1, \dots, 2n + 1,$$

où $b_{i,j}^{(k)}$ est la dérivée partielle de $\eta_{i,j}$ par rapport au $k^{\text{ème}}$ élément du vecteur γ :

$$\gamma = \begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_n \\ \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},$$

Généralement, le système n'est pas résoluble analytiquement : on utilise alors des procédures numériques, comme l'algorithme de Newton-Raphson.

Une fois le modèle estimé, on peut estimer la réserve pour les sinistres survenus en années i ainsi :

$$\hat{\mathbb{E}}[R_i] = \sum_{j=n-i+1}^n \hat{\mathbb{E}}[X_{i,j}] = \sum_{j=n-i+1}^n \hat{\mu}_{i,j}$$

2.4.3 Qualité de l'ajustement

La qualité de l'ajustement est évaluée à l'aide de la statistique du χ^2 de Pearson.

Les résidus de Pearson sont définis comme :

$$r_{i,j}^P = \frac{x_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\text{Var}(\mu_{i,j})}}$$

La statistique du χ^2 de Pearson est calculée comme :

$$\chi^2 = \frac{1}{\phi} \sum_{0 \leq i+j \leq n} (r_{i,j}^P)^2$$

χ^2 est asymptotiquement distribué selon une loi du chi-2, de paramètre :

$$p = \frac{(n+1)(n+2)}{2} - (2n+1)$$

Un estimateur de ϕ est :

$$\hat{\phi} = \frac{\phi\chi^2}{p} = \frac{1}{p} \sum_{0 \leq i+j \leq n} (r_{i,j}^P)^2$$

Les résidus de Pearson doivent asymptotiquement suivre une loi $\mathcal{N}(0, \phi)$.

2.4.4 Modèles usuels

- $X_{i,j}$ suit une loi de Poisson (ou Poisson surdispersé) avec une fonction lien multiplicative. Les valeurs prédites par un modèle GLM de Poisson (simple ou surdispersé¹) à lien multiplicatif (log) coïncident avec celles prédites par la méthode Chain-Ladder standard. La démonstration a été faite par Renshaw et Verrall (1998) ([43]).
- $X_{i,j}$ suit une loi normale avec une fonction lien multiplicative : Méthode de De Vylder.
- $X_{i,j}$ modélisé par une loi Gamma avec une fonction lien multiplicative : Méthode de Mack.

En pratique, le modèle de Poisson surdispersé est très utilisé. Cependant, il suppose implicitement que les paiements incrémentaux sont tous positifs. Cela pose problème lorsque les paiements sont étudiés nets de recours (Sous Solvabilité II, le *Best Estimate* doit être obtenue net de recours). Le problème est aussi présent lorsque l'on travaille avec des charges. Il est très possible d'avoir des charges incrémentales négatives : si les gestionnaires de sinistre ont tendance à être trop conservateurs à l'ouverture du sinistre dans la constitution des provisions, il est possible d'observer une baisse des charges cumulées d'une année de développement à une autre.

1. On dit que Y suit une loi de Poisson dispersé de paramètre $(\lambda, \phi) \in \mathbb{R}_+^2$ si et seulement si $\frac{Y}{\phi} \sim \mathcal{P}(\frac{\lambda}{\phi})$.

2.5 Bootstrap

2.5.1 Principe

La méthode Bootstrap a été introduite par Efron (1978). L'idée générale est un ré-échantillonnage des données. Supposons que nous observons un échantillon i.i.d. (Z_1, \dots, Z_n) de données suivant une loi \mathcal{F} inconnue, et que nous cherchons à estimer un paramètre d'intérêt θ . Supposons que nous avons obtenu un estimateur de θ : $\hat{\theta}_n = g(z_1, \dots, z_n)$ (par exemple θ est la moyenne de F et $\hat{\theta}_n$ la moyenne empirique). Nous aimerions avoir d'autres échantillons de réalisations de la loi \mathcal{F} , pour calculer de nouveaux estimateurs $\tilde{\theta}_n$ et obtenir une distribution de l'estimateur pour mesurer l'erreur d'estimation. Mais nous ne connaissons pas la loi \mathcal{F} .

Cependant, l'estimateur usuel de la fonction de répartition de F , la fonction de répartition empirique \hat{F}_n estimé à partir de l'échantillon (Z_1, \dots, Z_n) est un estimateur "naturel" et convergent de la fonction de répartition F . L'idée est donc de créer B nouveaux échantillons indépendants de taille n , en tirant avec remise dans l'échantillon initial (Z_1, \dots, Z_n) . Chaque échantillon $b = 1, \dots, B$ donne un nouvel estimateur $\hat{\theta}_n^b$. Nous pouvons alors obtenir une distribution empirique pour $\hat{\theta}_n$.

2.5.2 Application à la modélisation GLM

La procédure bootstrap se décompose en plusieurs étapes :

1. Estimer le modèle par GLM.
2. Calculer les résidus de Pearson.
3. Procéder à B ré-échantillonnage avec remise des résidus de Pearson. On obtient un nouveau triangle des résidus $r_{i,j}^b, b = 1, \dots, B$. On génère alors B échantillons Bootstrap de paiements incrémentaux :

$$x_{i,j}^b = \hat{\mu}_{i,j} + r_{i,j}^b \sqrt{\text{Var}(\hat{\mu}_{i,j})}.$$

On a alors B triangles de pseudo-données.

4. Estimer le modèle GLM pour les B pseudo-triangles et calculer les réserves associées. On obtient alors une distribution empirique de la réserve ultime.

Dans le cadre d'un modèle GLM, on peut également utiliser un Bootstrap paramétrique, puisque l'on a supposé une loi paramétrique pour les données. Dans l'étape 5 précédente, on estime le modèle GLM pour les B pseudo-triangles. On en tire des estimateurs $\hat{\mu}_{i,j}^b$ et $\text{Var}(\hat{\mu}_{i,j}^b)$. Au lieu d'estimer les $X_{i,j}$ du triangle inférieur en utilisant $\hat{\mu}_{i,j}^b$, on simule des $X_{i,j}$ de moyenne $\hat{\mu}_{i,j}^b$ et de variance $\text{Var}(\hat{\mu}_{i,j}^b)$, selon la loi que l'on a supposé dans le modèle GLM (p. ex. une loi de Poisson surdispersée).

2.6 Facteurs de queue

Pour les branches longues, le triangle n'est parfois pas complet : le développement complet de la sinistralité n'est pas disponible. Pour ne pas sous-estimer la charge ultime en raison de cette troncature à droite des données, il est alors nécessaire d'estimer une queue du développement du triangle.

Pour les facteurs de développement de Chain-Ladder, on peut par exemple chercher à ajuster une fonction $y = f(j)$ régulière et vérifiant $f(j) \geq 1$ pour tout j . Compain (2010) [9] propose par exemple : $f(j) = 1 + a(1 + j)^{-b}$. Les coefficients a et b sont obtenus par minimisation des carrés des écarts aux facteurs empiriques. Il est également possible d'extrapoler directement les règlements décumulés [39].

Dans le cadre de son modèle, Mack a proposé en 1999 [32] de prendre en considération un facteur de queue pour déterminer l'erreur de prédiction des provisions.

Dans la procédure bootstrap avec GLM, Michaud (2012) [34] propose d'estimer un facteur de queue de développement pour chaque pseudo-triangle obtenu par rééchantillonnage.

Bilan

Dans cette partie, nous avons présenté les modèles classiques de provisionnement agrégé. Une revue plus détaillée des modèles peut être trouvée dans England et Verrall (2002) [13]. Les modèles agrégés possèdent plusieurs inconvénients, notamment :

- Le faible nombre d'observation dans un triangle agrégé $((n + 1)(n + 2)/2)$ rend sensible les différentes méthodes d'estimation aux *outliers*. Ce faible nombre entraîne aussi une sur-paramétrisation des méthodes (par exemple, dans les modèles GLM, $2n + 1$ paramètres : dans un triangle 6x6, cela donne 21 observations et 11 paramètres à estimer). L'étude séparée des sinistres ouverts mais non clos et des IBNyR est également difficile à mener.
- La plupart des modèles développés sont généralement mieux adaptés aux données de paiements qu'aux données de charges. Certains modèles, comme le modèle GLM Poisson-surdispersé, n'autorisent pas à avoir des charges incrémentales négatives dans le triangle.

La modélisation individuelle des sinistres peut permettre de dépasser ces différentes limites. Dans la section 1.4 ci-dessus, nous avons évoqué les différents événements qui peuvent intervenir dans la vie d'un sinistre. Pour modéliser les sinistres ligne à ligne, nous avons besoin d'établir un cadre qui définit quels éléments vont être retenus dans la modélisation, et ceux qui vont être écartés. C'est ce que nous faisons dans la partie suivante.

Partie 3

Établissement d'un cadre pour une évaluation des provisions par une approche ligne à ligne

Nous faisons d'abord une brève revue de la littérature, puis nous expliquons notre démarche avant de développer le cadre que nous avons retenu.

3.1 Revue succincte de la littérature sur les modèles individuels

L'idée de modéliser individuellement les sinistres n'est pas nouvelle.

Les modèles de provisionnement individuel proposés dans la littérature s'inscrivent le plus souvent dans le cadre des Processus de Poisson marqués (*point process*) proposé par Norberg (1993, 1999) [37], [38]. Il s'agit de modèles en temps continu représentant la survenance des sinistres à l'aide de processus de Poisson. À la survenance d'un sinistre dans le temps est associée un vecteur de variables aléatoires (la marque) modélisant tous les éléments du cycle de vie d'un sinistre.

Ces modèles sont très flexibles et permettent de prendre en compte de manière exhaustive les caractéristiques d'un sinistre :

- la date de survenance
- le délai de notification à l'assureur,
- les dates des différents paiements et leur montants,
- la date de clôture.

Des informations sur les caractéristiques du preneur d'assurance, des effets saisonniers, le niveau d'exposition au risque de l'assureur, etc., peuvent également être prises en compte.

Les méthodes d'estimation associées sont également très variées : Haarstrup et Arjas (1996) [21] utilisent des méthodes bayésiennes non-paramétriques. Larsen (2007) [29] utilise la distribution Pareto Généralisée pour modéliser la série de paiements incrémentaux. D'autres auteurs ont recours aux modèles linéaires généralisés (GLM) (voir Taylor *et al.* (2008) [50] par exemple). Zhao *et al.* (2009) [53] et Zhao et Zhou (2010) [52] utilisent des techniques de l'analyse de survie et des méthodes de copules.

L'inconvénient du cadre des processus de Poisson marqués est sa complexité et le nombre de processus (et paramètres associés) à estimer. Antonio et Plat (2012) [1] utilisent par exemple trois types d'événements pour modéliser la vie future des sinistres ouverts : classement direct du dossier sans paiement, paiement sans classement, paiement avec classement. Pour prédire la vie future, il faut alors simuler le temps jusqu'au prochain événement, simuler le type d'événement, simuler le paiement associé, et recommencer jusqu'à la clôture. L'erreur de modèle dans le choix d'une distribution pour chaque processus et l'erreur d'estimation (incertitude sur les paramètres estimés) peuvent conduire à des difficultés dans le calibrage et un manque de robustesse. De plus, les études de cas pratiques restent pour l'instant assez limitées.

Un autre inconvénient est que les processus de Poisson marqués sont définis en temps continu. D'un point de vue pratique, cela demande une base de données extrêmement détaillée, qui contient la date précise des événements de la vie d'un sinistre (date de survenance, date de paiements, etc.).

D'autres études se penchent sur une méthode d'estimation particulière des réserves au niveau individuel. Elles ne développent non pas un cadre général mais un cadre analytique propre au modèle statistique retenu : Dittmer (2006) [11] propose d'utiliser la méthode des k -voisins les plus proches pour estimer les provisions liées aux sinistres "tardifs" exceptionnels (*large claims*). Mahon (2005) [33] applique la théorie des matrices de transition pour modéliser le développement années après années des pertes liées à des sinistres individuels exceptionnels. Roselund (2012) [45] met en place une méthode *bootstrap* sur l'historique des sinistres individuels pour estimer les réserves.

3.2 Démarche adoptée et présentation du cadre

Notre démarche consiste à établir un premier cadre général simple pour une modélisation individuelle des sinistres, quitte à ce qu'il ne puisse pas prendre en compte tous les éléments de la vie des sinistres et tous les types de sinistre (notamment les tardifs). Nous cherchons des hypothèses qui permettent une modélisation aisée, mais qui doivent aussi rester réalistes pour que les modèles développés à partir de ce cadre soient utilisables dans des cas pratiques. Nous reprenons le cadre proposé par Chau et Dinh (2012) [7], discutons de son champ d'application et l'étendons à la modélisation des charges de sinistre.

Les intérêts de retenir un cadre simple sont multiples :

- D'abord, un cadre simple permet d'avoir un nombre restreint de processus et de paramètres à estimer le cas échéant. Cela permet de mieux analyser les erreurs de prédiction commises
- Ensuite, ce cadre permet d'étudier facilement les hypothèses sous lesquelles on peut établir un lien entre la modélisation individuelle et les modèles agrégés traditionnelles (démarche micro-fondée). Ce lien est entendu au sens où **l'estimation des provisions au niveau individuel et l'estimation des provisions après agrégation sont asymptotiquement équivalentes**. Nous nous concentrons sur la méthode Chain-Ladder, qui constitue un véritable *benchmark*. L'intérêt de créer un modèle qui reproduit (asymptotiquement) une estimation des réserves proches de Chain-Ladder peut sembler « futile, bien qu'un éclairage considérable peut être obtenu en commençant par ces modèles » (England et Verrall, 2002) [13]. En effet, l'établissement d'un lien entre la méthode Chain-Ladder et la méthode ligne à ligne permettrait de cerner les cas où la première méthode présente une perte d'information importante ou enfreint des caractéristiques saillantes des sinistres au niveau micro.

Le cadre proposé par Chau et Dinh (2012) [7] caractérise chaque sinistre d'une base de données "ligne à ligne" par trois éléments :

- une date de survenance ;
- un processus d'états du sinistre (ouvert/clos) ;
- un processus de règlements.

Les méthodes agrégées, comme mentionnées ci-dessus (partie 2.1), peuvent être utilisées sur un triangle de paiements, ou bien sur un triangle de charges. La charge du sinistre à la date t est la somme des paiements effectués jusqu'à cette date t et de la provision dossier-dossier à cette date estimée par le gestionnaire de sinistres. Un désavantage potentiel de considérer la charge est qu'elle est dépendante de la politique de gestion de sinistre de la compagnie (choix du forfait d'ouverture, rapidité de la collecte d'information, etc.). L'avantage du travail avec les charges est que les processus sont généralement plus stables, et strictement positifs¹, ce qui permet d'étudier des ratios de charge (p. ex. charge ultime sur charge actuelle, comme le font Taylor *et. al.* (2008) [50]). Pour cette raison, nous allons étendre le modèle de Chau et Dinh en incluant un processus de charges.

1. Hors cas des *nil claims*, et en considérant que la provision dossier-dossier est systématiquement constituée à la date de notification du sinistre à l'assureur. L'utilisation d'un forfait d'ouverture peut être estimée comme quasi-systématique dans les protocoles de gestion des sinistres au sein des compagnies d'assurance.

Les trois processus (état, paiements, charge), qui sont discrets, peuvent potentiellement être étudiés selon différents découpages temporels : annuels, semestriels, trimestriels, ... (de la même manière que les montants dans Chain-Ladder peuvent être rapportés à des périodes de longueurs différentes). Nous retenons par la suite un découpage par année. Un découpage temporel plus fin peut se révéler coûteux en terme de traitement des données et de temps d'estimation.

Il serait intéressant d'étudier la stabilité de la méthode à un changement de date d'inventaire (fin d'année, milieu d'année, ...). Cependant, c'est généralement à la date d'arrêté comptable que les *case estimates* des gestionnaires de sinistre sont les plus complets. Se placer à cette date permet de diminuer le nombre potentiel de dossiers de sinistres qui ont été notifiés à l'assureur mais qui ne disposent pas encore d'une estimation spécifique (sinistres provisionnés en IBNeR).

Nous allons maintenant présenter les notions de base qui permettent de définir un cadre simple pour la modélisation ligne à ligne.

3.3 Définitions de base

Le caractère aléatoire du provisionnement des sinistres est décrit par l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ sur lequel les variables aléatoires ci-dessous seront définies, sauf indication contraire.

La définition mathématique d'un sinistre est :

Définition 3.3.1 *Un sinistre S_i est un quadruplet (n_i, X_i, I_i, D_i) , où :*

1. $n_i \in \mathbb{N}$ est l'**année d'origine** du sinistre ;
2. X_i le **processus des règlements incrémentaux**, défini comme une séquence de variables aléatoires $(X_{i,j})_{j \geq 0}$, où $X_{i,j}$ est la v.a. représentant le paiement associé au sinistre S_i en année de déroulement j ;
3. I_i le **processus de charges**, défini comme une séquence de variables aléatoires $(I_{i,j})_{j \geq 0}$, où $I_{i,j}$ est la v.a. représentant la charge associée à au sinistre S_i en année de déroulement j ;
4. D_i est le **processus d'états**, défini comme une suite de variables aléatoires de Bernoulli $(D_{i,j})_{j \geq 0}$, avec la convention suivante :

$$D_{i,j} = \begin{cases} 1 & \text{si le sinistre } S_i \text{ est en cours en année de développement } j, \\ 0 & \text{si le sinistre } S_i \text{ est clos en année de développement } j. \end{cases}$$

Nous noterons par la suite \mathcal{S} l'ensemble des sinistres étudiés.

Remarques

L'année d'origine n_i marque le début de la vie d'un sinistre. Par exemple, si un ensemble de sinistres RC en base réclamation est étudié, l'année de première réclamation pourrait être retenue comme année d'origine. Par la suite, nous retiendrons comme année d'origine l'**année de survenance** du sinistre.

Concernant les processus de paiements incrémentaux et de charges, dans la pratique, il est possible qu'il y ait plusieurs paiements ou modifications de la réserve dossier-dossier au cours d'une année. On considère alors qu'on se place en fin d'année de déroulement j (date d'inventaire), et l'on considère la somme des paiements incrémentaux ayant eu lieu au cours de l'année. Pour déterminer la charge, on utilise la dernière estimation dossier-dossier constituée avant la date d'inventaire.

Le processus d'états nous informe si le sinistre est ouvert ou clos. Il existe une ambiguïté concernant le choix de la date de clôture. Il peut s'agir de l'année du dernier paiement final, ou bien de l'année où le gestionnaire de sinistres ferme administrativement le dossier (classement du dossier). En effet, les dossiers peuvent être laissés ouverts après le dernier paiement, lorsque le gestionnaire s'attend encore à de nouveaux règlements dans le futur, qui n'ont finalement pas lieu. Cette date de classement est liée à la politique de gestion des dossiers de la compagnie et à des considérations de prudence.

Néanmoins, retenir la date du dernier paiement ou bien la date de fermeture administrative ne devrait pas avoir *a priori* une très grande incidence dans l'estimation des réserves. La méthode proposée permet de prendre en compte la stabilité de la charge et des paiements cumulés d'une année à l'autre. Nous retiendrons donc la date de classement du dossier, qui est généralement celle qui est la plus facile à obtenir dans les données ligne à ligne.

3.3.1 Processus d'états

Nous allons faire une première hypothèse sur le processus d'états :

Hypothèse 3.3.1 *Pour tout $S_i \in \mathcal{S}$, le processus d'états $(D_{i,j})_{j \geq 0}$ vérifie les deux conditions suivantes :*

$$\mathbb{P}[D_{i,0} = 1] = 1, \tag{3.1}$$

$$\forall j \geq 1, \forall k \geq j, \mathbb{P}[D_{i,k} = 0 | D_{i,j} = 0] = 1. \tag{3.2}$$

Remarques

- La condition (3.1) revient à exclure de l'analyse les sinistres « tardifs », qui ne sont pas encore ouverts durant l'année de survenance. Comme vu dans la partie 1.4.1, dans certaines branches, notamment les branches Responsabilité Civile (automobile, médicale, vie privée, etc.), le nombre de sinistres tardifs peut être très important.
- La condition (3.2) revient à ne pas considérer dans l'analyse les cas où les sinistres sont réouverts après leur clôture par le gestionnaire de sinistre.

La non-modélisation d'un délai de notification pour les sinistres tardifs et leur traitement à part ne rend pas possible l'établissement d'une provision à un an pour les tardifs (IBNyR), mais seulement à l'ultime. De même, devoir traiter les cas de réouverture séparément peut constituer une limite au cadre développé. Pour prendre en compte un délai de notification et/ou les cas de réouverture, nous aurions pu considérer un processus à plusieurs états discrets (survenu mais non notifié, ouvert, clos, à nouveau ouvert, etc.).

Une solution alternative serait de considérer que n_i représente l'année de notification du sinistre, et non l'année de survenance. Il serait alors possible d'estimer la réserve pour les tardifs observés dans la base (en faisant p. ex. l'hypothèse supplémentaire que le développement des tardifs est le même que celui des non-tardifs). Pour les tardifs non encore observés, un recours à des méthodes agrégées serait possible (triangle des nombres de sinistre et estimation à un coût moyen à l'ultime).

La connaissance de l'état initial (\mathbb{P} -p.s.) du processus D_i et le caractère absorbant de l'état 0 (la trajectoire de $D_{i,j}$ ne quitte plus cet état une fois qu'il est atteint) entraînent une propriété particulière du processus d'état (voir partie 4.2.2.1). Cette propriété facilitera les estimations des provisions à l'ultime. De plus, la notion de durée de vie du sinistre, que nous allons introduire ci-dessous (c.f. 3.3), prend un sens très intuitif et univoque avec les conditions imposées par 3.3.1.

Cette notion de durée de vie permet de « résumer » le processus D_i à l'année de développement à laquelle il atteint l'état absorbant 0.

A partir du processus D_i , nous pouvons définir la variable aléatoire suivante représentant l'année de clôture du sinistre S_i (alternativement, sa durée de vie) :

$$\Gamma_i = \inf \{j \geq 0 | D_{i,j+1} = 0\}. \quad (3.3)$$

Le déroulé d'un sinistre ne dure pas indéfiniment. Nous faisons donc l'hypothèse suivante :

Hypothèse 3.3.2

$$\Gamma^{max} = \sup_{S_i \in \mathcal{S}} \Gamma_i < \infty$$

La durée de vie de l'ensemble des sinistres est ainsi finie, et on a :

$$\forall S_i \in \mathcal{S}, \mathbb{P}[D_{i,\Gamma^{max}+1} = 0] = 1.$$

Remarques

L'hypothèse 3.3.2 répond à deux préoccupations :

- d'un point de vue théorique, faciliter l'étude de la convergence, en s'assurant que l'état 1 n'est jamais absorbant et que l'état 0 est atteint à distance finie.
- d'un point de vue pratique, dans une base de données ligne à ligne, nous utilisons les sinistres observés déjà clos pour estimer la durée de vie maximale. Statistiquement, il ne sera pas possible de mesurer une probabilité d'avoir une durée de vie supérieure à la durée de vie maximale observée dans la base. En termes opérationnels, il faudra donc disposer d'une base avec un historique suffisamment profond, notamment pour les branches longues.

Si jamais l'historique n'est pas assez profond, et que les sinistres les plus longs sont censurés à droite (c.-à-d. à la dernière date d'observation, le dossier est toujours ouvert), il pourrait être intéressant de réfléchir à des méthodes d'extrapolation, par analogie aux méthodes d'extrapolation pour les méthodes agrégées (voir partie 2.6). Avec un phénomène de censure à droite pour le sinistre observé le plus long et non clos, nous pourrions seulement estimer la probabilité suivante : $\mathbb{P}[\Gamma_i > k]$, où k est la profondeur maximale de la base. Une méthode d'extrapolation consisterait à fixer un âge maximal Γ^{max} (donc telle que $\mathbb{P}[\Gamma_i > \Gamma^{max}] = 0$ et à utiliser des méthodes d'interpolation pour établir une estimation de : $\mathbb{P}[\Gamma_i = k + 1], \dots, \mathbb{P}[\Gamma_i = \Gamma^{max}]$. Il serait également possible d'utiliser une loi paramétrique pour modéliser la queue de la distribution. Il faudrait aussi dans ce cas-là établir une méthode d'extrapolation des paiements futurs, ce qui est moins évident, et nécessite une approche "facteur de queue".

3.3.2 Processus de paiements

Le processus de règlements cumulés liés au sinistre S_i s'écrit :

$$C_{i,j} = \sum_{k=0}^j X_{i,k}.$$

Le montant réellement payé par l'assureur pour le sinistre S_i jusqu'à l'année de développement j est :

$$Y_{i,j} = \sum_{k=0}^j D_{i,k} X_{i,k},$$

c'est-à-dire que pour les années de développement $j > \Gamma_i$, $X_{i,j}$ est un processus « fictif » inobservable. De même, la charge I_i n'est plus observée pour les années de développement $j > \Gamma_i$.

On a donc en utilisant la durée de vie Γ_i :

$$Y_{i,j} = \sum_{k=0}^{\min\{j, \Gamma_i\}} X_{i,k}.$$

Les relations entre $Y_{i,j}$, $C_{i,j}$ et $X_{i,j}$ peuvent s'exprimer de la manière suivante :

$$Y_{i,j} = C_{i,j} \mathbb{1}\{D_{i,j} = 1\} + \sum_{k=0}^{j-1} C_{i,k} \mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\}. \quad (3.4)$$

Par la suite, nous ferons l'hypothèse suivante :

Hypothèse 3.3.3 *Le processus d'états D_i et le processus de règlements incrémentaux X_i (resp. le processus de charges I_i) sont indépendants.*

Remarque : validité de l'hypothèse

La décomposition (3.4), sous l'hypothèse 3.3.3, permet d'étudier séparément les processus d'états et les processus de règlements incrémentaux (et cumulés C_i). Cette décomposition est particulièrement utile pour estimer les valeurs futures de $Y_{i,j}$, en utilisant les résultats de convergence sur les fonctions de variables aléatoires indépendantes. De plus, pour l'espérance, nous pouvons utiliser la propriété que l'espérance d'un produit de v.a. indépendantes et le produit des espérances.

L'hypothèse 3.3.3 revient à considérer que la durée de vie d'un sinistre est indépendante des montants de la prestation versée. Cette hypothèse peut être un peu forte, dans le sens où il est possible qu'on observe en moyenne des déroulés plus longs pour les sinistres mettant en jeu de fortes sommes. D'un autre côté, le processus de paiements relève plus de l'aléa lié au sinistre même, tandis que la durée de vie peut être plus liée aux *process* internes de la compagnie (rapidité à obtenir de l'information sur les sinistres, politique en termes de classement des dossiers,

...). Il conviendra donc de vérifier si l'on observe une corrélation entre la durée de vie de sinistres et par exemple le premier paiement, ou la charge initiale, dans les données.

Une solution alternative proposée dans la littérature est de ne considérer qu'un processus de paiements cumulés, et de le traiter comme une variable censurée (Rose, 2009 [44]) : si le sinistre est toujours ouvert à la date d'inventaire, la charge finale du sinistre n'est pas connue (censure à droite), seulement les paiements effectués aujourd'hui. Sinon, il n'y a pas de censure, la charge finale est connue. L'estimation de la charge finale pour les sinistres ouverts est ensuite faite à l'aide de modèles de survie paramétriques. Cette solution a cependant l'inconvénient de ne pas modéliser les durées de liquidation des sinistres, et donc de ne pouvoir établir qu'une provision à l'ultime. En modélisant la durée de vie, nous pouvons estimer par exemple les flux de trésorerie à un an, ou bien la variation à un an de charge dossier-dossier.

3.3.3 Processus de charges

Le processus de charges I_i pour le sinistre S_i en année de développement j est la somme du règlement cumulé en année j , $Y_{i,j}$, et de la provision dossier-dossier effectuée par le gestionnaire à la date j .

La relation entre la charge à l'ultime et les paiements cumulés à l'ultime (c.-à-d. à l'année de clôture du sinistre) est :

$$\boxed{\forall i, \forall j \geq \Gamma_i, I_{i,j} = Y_{i,j}.} \quad (3.5)$$

3.4 Ensembles d'information

Pour pouvoir étudier les propriétés de convergence des estimateurs des processus que nous allons définir par la suite, nous avons besoin de définir les différents sous-ensembles de sinistres auxquels nous allons nous intéresser.

3.4.1 Sous-ensembles de sinistres d'intérêt

Soit une année calendaire $t \in \mathbb{N}$ fixée. Dans la suite du rapport, cette année calendaire représentera la dernière année observée. Pour chaque $n \in \{0, 1, \dots, \Gamma^{max}\}$, on définit le sous-ensemble de sinistres suivants :

$$\mathcal{A}_n(t) = \{S_i \in \mathcal{S} | t - n_i = n\}, \quad (3.6)$$

l'ensemble des sinistres qui sont en année de développement $j = n$ à la date t .

On considéra également le sous-ensemble de $\mathcal{A}_n(t)$ suivant :

$$\mathcal{A}_n^*(t) = \{S_i \in \mathcal{S} | t - n_i = n, D_{i,n} = 1\}, \quad (3.7)$$

l'ensemble des sinistres qui sont en année de développement $j = n$ et non clôturés à la date t .

On s'intéressera également à l'ensemble suivant :

$$\mathcal{B}_n(t) = \{S_i \in \mathcal{S} | t - n_i \geq n\} = \cup_{k \geq n} \mathcal{A}_k(t), \quad (3.8)$$

l'ensemble des sinistres qui sont au moins en année de développement $j = n$ à la date t ,

et le sous-ensemble de $\mathcal{B}_n(t)$:

$$\mathcal{B}_n^*(t) = \{S_i \in \mathcal{S} | t - n_i \geq n, D_{i,n} = 1\}, \quad (3.9)$$

l'ensemble des sinistres qui sont au moins en année de développement $j = n$ à la date t et qui sont toujours ouverts à cette année de développement n .

Nous allons illustrer ces ensembles sur un exemple. Sur la figure 3.1, nous avons représenté une base de données contenant les paiements observés $Y_{i,j}$. Les lignes contiennent les différents sinistres i . La date d'inventaire t est fixée à 2012. Pour un sinistre i , la première valeur observée $Y_{i,0}$ nous donne l'année d'origine n_i . la dernière colonne indique le statut du sinistre à la date d'inventaire.

Année	$t - n_1$							t	$D_{i,t-n_i}$
	2005	2006	2007	2008	2009	2010	2011	2012	
1	$Y_{1,0}$	$Y_{1,1}$	$Y_{1,2}$	$Y_{1,3}$	$Y_{1,4}$	$Y_{1,5}$	$Y_{1,6}$	$Y_{1,7}$	0
2		$Y_{2,0}$	$Y_{2,1}$	$Y_{2,2}$	$Y_{2,3}$	$Y_{2,4}$	$Y_{2,5}$	$Y_{2,6}$	1
3		$Y_{3,0}$	$Y_{3,1}$	$Y_{3,2}$	$Y_{3,3}$	$Y_{3,4}$	$Y_{3,5}$	$Y_{3,6}$	0
4			$Y_{4,0}$	$Y_{4,1}$	$Y_{4,2}$	$Y_{4,3}$	$Y_{4,4}$	$Y_{4,5}$	0
5				$Y_{5,0}$	$Y_{5,1}$	$Y_{5,2}$	$Y_{5,3}$	$Y_{5,4}$	0
6					$Y_{6,0}$	$Y_{6,1}$	$Y_{6,2}$	$Y_{6,3}$	1
7					$Y_{7,0}$	$Y_{7,1}$	$Y_{7,2}$	$Y_{7,3}$	0
8					$Y_{8,0}$	$Y_{8,1}$	$Y_{8,2}$	$Y_{8,3}$	1
9					$Y_{9,0}$	$Y_{9,1}$	$Y_{9,2}$	$Y_{9,3}$	0
10						$Y_{10,0}$	$Y_{10,1}$	$Y_{10,2}$	1
11						$Y_{11,0}$	$Y_{11,1}$	$Y_{11,2}$	0
12							$Y_{12,0}$	$Y_{12,1}$	1

FIGURE 3.1 – Exemple de données ligne à ligne

Sur les figures 3.2 et 3.3, nous avons surligné respectivement les ensembles $\mathcal{A}_3(t)$ et $\mathcal{A}_3^*(t)$. Lorsque l'on agrège les données par année de survenance pour construire le triangle de développement, on considère les ensembles $\mathcal{A}_n(t)$: même si le sinistre est clos en année de développement n , le paiement cumulé associé à ce sinistre sera pris en compte dans le triangle. Dans les données ligne à ligne, on s'intéresse aux ensembles $\mathcal{A}_n^*(t)$, c'est-à-dire que l'on cherche à estimer les paiements futurs pour les sinistres qui sont encore ouverts en année de développement $t - n_i$.

Année	$t - n_1$							t	$D_{i,t-n_i}$
	2005	2006	2007	2008	2009	2010	2011	2012	
1	$Y_{1,0}$	$Y_{1,1}$	$Y_{1,2}$	$Y_{1,3}$	$Y_{1,4}$	$Y_{1,5}$	$Y_{1,6}$	$Y_{1,7}$	0
2		$Y_{2,0}$	$Y_{2,1}$	$Y_{2,2}$	$Y_{2,3}$	$Y_{2,4}$	$Y_{2,5}$	$Y_{2,6}$	1
3		$Y_{3,0}$	$Y_{3,1}$	$Y_{3,2}$	$Y_{3,3}$	$Y_{3,4}$	$Y_{3,5}$	$Y_{3,6}$	0
4			$Y_{4,0}$	$Y_{4,1}$	$Y_{4,2}$	$Y_{4,3}$	$Y_{4,4}$	$Y_{4,5}$	0
5				$Y_{5,0}$	$Y_{5,1}$	$Y_{5,2}$	$Y_{5,3}$	$Y_{5,4}$	0
6					$Y_{6,0}$	$Y_{6,1}$	$Y_{6,2}$	$Y_{6,3}$	1
7					$Y_{7,0}$	$Y_{7,1}$	$Y_{7,2}$	$Y_{7,3}$	0
8					$Y_{8,0}$	$Y_{8,1}$	$Y_{8,2}$	$Y_{8,3}$	1
9					$Y_{9,0}$	$Y_{9,1}$	$Y_{9,2}$	$Y_{9,3}$	0
10						$Y_{10,0}$	$Y_{10,1}$	$Y_{10,2}$	1
11						$Y_{11,0}$	$Y_{11,1}$	$Y_{11,2}$	0
12							$Y_{12,0}$	$Y_{12,1}$	1

FIGURE 3.2 – L'ensemble $\mathcal{A}_3(t)$

		$t - n_1$							t	
		t								
Année	2005	2006	2007	2008	2009	2010	2011	2012	$D_{i,t-n_i}$	
}	1	$Y_{1,0}$	$Y_{1,1}$	$Y_{1,2}$	$Y_{1,3}$	$Y_{1,4}$	$Y_{1,5}$	$Y_{1,6}$	$Y_{1,7}$	0
	2		$Y_{2,0}$	$Y_{2,1}$	$Y_{2,2}$	$Y_{2,3}$	$Y_{2,4}$	$Y_{2,5}$	$Y_{2,6}$	1
	3		$Y_{3,0}$	$Y_{3,1}$	$Y_{3,2}$	$Y_{3,3}$	$Y_{3,4}$	$Y_{3,5}$	$Y_{3,6}$	0
	4			$Y_{4,0}$	$Y_{4,1}$	$Y_{4,2}$	$Y_{4,3}$	$Y_{4,4}$	$Y_{4,5}$	0
	5				$Y_{5,0}$	$Y_{5,1}$	$Y_{5,2}$	$Y_{5,3}$	$Y_{5,4}$	0
	6					$Y_{6,0}$	$Y_{6,1}$	$Y_{6,2}$	$Y_{6,3}$	1
	7					$Y_{7,0}$	$Y_{7,1}$	$Y_{7,2}$	$Y_{7,3}$	0
	8					$Y_{8,0}$	$Y_{8,1}$	$Y_{8,2}$	$Y_{8,3}$	1
	9					$Y_{9,0}$	$Y_{9,1}$	$Y_{9,2}$	$Y_{9,3}$	0
	10						$Y_{10,0}$	$Y_{10,1}$	$Y_{10,2}$	1
	11						$Y_{11,0}$	$Y_{11,1}$	$Y_{11,2}$	0
	12							$Y_{12,0}$	$Y_{12,1}$	1

FIGURE 3.3 – L'ensemble $\mathcal{A}_3^*(t)$

En termes de triangle agrégé des données :

- $\sum_{S_i \in \mathcal{A}_{t-k}(t)} Y_{i,j}$ correspond à la cellule (k, j) du triangle des paiements cumulés (voir figure 3.4).
- Pour $m \geq j$, $\sum_{S_i \in \mathcal{B}_m(t)} Y_{i,j}$ correspond à la somme des éléments de la colonne numéro j , de l'année de survenance 0 jusqu'à l'année de survenance m du triangle agrégé des paiements cumulés (voir figure 3.5). Cette somme entre ainsi en jeu dans le calcul des facteurs Chain-Ladder au niveau agrégé.

k, j	0	1	...	j	...	$t - 1$	t
0	\ddots
1	\ddots	
\vdots	\ddots		
k	...	$\sum_{S_i \in \mathcal{A}_{t-k}(t)} Y_{i,1}$...	$\sum_{S_i \in \mathcal{A}_{t-k}(t)} Y_{i,j}$...		
\vdots	\ddots				
$t - 1$...	\ddots					
t	\ddots						

FIGURE 3.4 – Les ensembles $\sum_{S_i \in \mathcal{A}_{t-k}(t)} Y_{i,1}$ et $\sum_{S_i \in \mathcal{A}_{t-k}(t)} Y_{i,j}$ dans le triangle des données agrégés

k,j	0	1	...	j	...	$t-1$	t
0
1
⋮
k
⋮
$t-1$
t

FIGURE 3.5 – L'ensemble $\sum_{S_i \in \mathcal{B}_1(t)} Y_{i,0}$ dans le triangle agrégé des données

3.4.2 Information disponible à la date t

On se place à l'année t , date à laquelle l'étude de provisionnement est réalisée (année de l'inventaire).

Nous faisons l'hypothèse suivante :

Hypothèse 3.4.1 *A la date t , l'assureur sait si un sinistre non clos sera encore ouvert à la date $t + 1$.*

Remarque

Le raisonnement sous-jacent à l'hypothèse 3.4.1 est le suivant : en se plaçant en date d'inventaire (t), le gestionnaire de sinistres sait s'il s'attend encore à des paiements ou des évolutions du dossier l'an prochain, ou si le règlement final a été réglé au cours de l'année. Dans le premier cas, le gestionnaire laissera la date de clôture indéterminée. Dans le second cas, on considère que le sinistre était ouvert en année t , date du dernier paiement, et qu'il est fermé en année $t + 1$. Ainsi, les sinistres survenus et clos administrativement la même année auront pour processus d'état : $D_{i,0} = 1$, $D_{i,1} = 0$. Si la date calendaire t correspond à l'année de développement 0 du sinistre, le statut en année de développement 1 est connu.

En revanche, l'assureur observe les processus $C_{i,j}$ seulement pour les sinistres encore ouverts à la date t .

Finalement, à la date t , l'assureur observe :

- Les réalisations des $Y_{i,j}$ pour tout $S_i \in \mathcal{B}_j(t)$;
- Les réalisations des $D_{i,j}$ pour tout $S_i \in \mathcal{B}_{j-1}(t)$;

- Les réalisations des $C_{i,j}$ et des $I_{i,j}$ sont observées seulement pour les $i \in \mathcal{B}_j^*(t)$.

L'information disponible à la date t s'écrit donc :

$$\begin{aligned} \mathcal{F}_t = \sigma(\{ & Y_{i,j} | S_i \in \mathcal{B}_j(t), j \in \{0, 1, \dots, \Gamma^{max}\} \} \\ & \cup \{ D_{i,j} | S_i \in \mathcal{B}_{j-1}(t), j \in \{1, \dots, \Gamma^{max}\} \} \\ & \cup \{ I_{i,j} | S_i \in \mathcal{B}_j^*(t), j \in \{0, 1, \dots, \Gamma^{max}\} \} \}. \end{aligned}$$

Dans toute la suite, on considérera toujours la même année d'inventaire t et nous noterons quelquefois $\mathcal{B}_n, \mathcal{A}_n, \mathcal{F}$ au lieu de $\mathcal{B}_n(t), \mathcal{A}_n(t), \mathcal{F}_t$ pour alléger les notations. Nous noterons aussi $i \in \mathcal{A}$ pour indiquer qu'un sinistre S_i appartient à l'ensemble \mathcal{A} .

Maintenant que nous avons défini les processus fondamentaux et la notion d'information disponible, nous pouvons développer la notion de mesures de provision dans le cadre d'un modèle ligne à ligne.

3.5 Mesures de provision

Dans le cadre développé par Chau et Dinh (2012) ([7]), une provision est une fonction définie sur l'ensemble des sinistres \mathcal{S} à valeurs dans \mathbb{R}^+ .

$R[\mathbf{A}|\mathcal{F}_t] = R_t[\mathbf{A}]$ est appelée une mesure de provision (*Stochastic measures of payments*) de l'ensemble $\mathbf{A} \subset \mathcal{S}$ sachant l'information \mathcal{F}_t .

La notion de mesures de provision ainsi introduite est très large et recouvre un ensemble de fonctions très variées (par exemple, une fonction associant toujours le même montant à un sinistre). Il semble donc important de donner des propriétés désirables de ces mesures de provision en développant une approche axiomatique. Chau et Dinh [7] introduisent la notion de mesures de provision cohérentes, par analogie avec les mesures de risque cohérentes¹.

3.5.1 Mesures de provision cohérentes

La définition suivante, introduite par Chau et Dinh [7], a pour but de se restreindre à l'étude d'une sous-classe de mesures de provision qui possèdent des propriétés souhaitables.

Définition 3.5.1 Une mesure de provision R est une mesure de provision cohérente si elle vérifie les trois propriétés suivantes :

1. **Positivité** : $\forall S_i \in \mathcal{S}$,

$$R_t[S_i] \geq 0, \quad (3.10)$$

$$R_t[S_i] = 0 \text{ si } D_{i,t+1} = 0. \quad (3.11)$$

2. **Additivité** : $\forall \mathbf{A}, \mathbf{B} \subseteq \mathcal{S}$,

$$\mathbf{A} \cap \mathbf{B} = \emptyset \Rightarrow R_t[\mathbf{A} \cup \mathbf{B}] = R_t[\mathbf{A}] + R_t[\mathbf{B}]. \quad (3.12)$$

3. **Invariance par changement de numéraire** : Soit une constante $c > 0$. On considère le sous-ensemble de sinistres $c\mathbf{A} = \{(n_i, cX_i, D_i) | S_i \in \mathbf{A}\}$

$$R_t[c\mathbf{A}|c\mathcal{F}_t] = cR_t[\mathbf{A}]. \quad (3.13)$$

Remarques

- La condition de positivité de la mesure de provision cohérente peut être discutée. Les provisions peuvent en effet être négatives dans certains cas, p. ex. lorsqu'il y a un taux de recours très importants. De plus, dans le cadre d'une provision IBNeR, il est tout à fait possible d'obtenir une provision négative, lorsque les gestionnaires ont tendance à être trop prudents dans leur estimation de charge dossier-dossier par rapport à la charge finale. Cependant, si l'on s'intéresse systématiquement à la PSAP brute de recours, cette condition ne pose pas problème.

1. Pour davantage de d'informations sur la notion de mesures de risque cohérentes, on peut se référer à [19].

- La condition d’additivité nous permet de calculer la provision pour un ensemble de sinistres comme la somme des provisions pour chacun des sinistres individuels. De la même manière que l’axiome de sous-additivité des mesures de risques cohérentes est critiquable (Une fusion de deux portefeuilles d’assurance crée-t-elle nécessairement moins de risques?), la condition d’additivité des mesures de provision est-elle souhaitable? Par exemple, supposons que nous nous plaçons dans le cadre d’un provisionnement prudent, et que la mesure de provision retenue soit la *VaR* à 75% de la charge ultime des sinistres. La propriété d’additivité n’est alors pas vérifiée. La condition d’additivité revient à placer la notion de mesures de provision cohérentes dans le cadre du *Best Estimate*, où l’on souhaite obtenir la moyenne non biaisée de la charge finale, l’espérance étant un opérateur linéaire.
- La condition d’invariance par changement de numéraire permet de se placer indifféremment sous l’unité monétaire souhaitée¹

3.5.2 Mesures de provision retenues

Dans la suite, nous allons nous intéresser à la mesure de provision suivante :

Provision ultime

$\mathbb{E}R_t^u[i]$ (resp. $\mathbb{E}R_t^u[A]$) est appelée provision à l’ultime du sinistre S_i (resp. du sous-ensemble de sinistres $A \subseteq \mathcal{S}$) si :

$$\begin{aligned}\mathbb{E}R_t^u[i] &= \mathbb{E}[Y_{i,\Gamma^{max}}|\mathcal{F}_t] - Y_{i,t-n_i} \\ &= \mathbb{E}[I_{i,\Gamma^{max}}|\mathcal{F}_t] - Y_{i,t-n_i} \text{ d'après la propriété (3.5),} \\ \mathbb{E}R_t^u[\mathbf{A}] &= \sum_{i \in \mathbf{A}} R_t^u[i],\end{aligned}$$

où $\mathbb{E}[Y_{i,\Gamma^{max}}|\mathcal{F}_t]$ et $\mathbb{E}[I_{i,\Gamma^{max}}|\mathcal{F}_t]$ sont respectivement la moyenne des paiements cumulés à l’ultime et la moyenne la charge ultime du sinistre S_i , sachant toute l’information disponible aujourd’hui \mathcal{F}_t .

Pour estimer cette mesure de provision, il faut exhiber des estimateurs $\hat{\mathbb{E}}[Y_{i,\Gamma^{max}}]$ et $\hat{\mathbb{E}}[I_{i,\Gamma^{max}}]$ \mathcal{F}_t -mesurable sans biais de $\mathbb{E}[Y_{i,\Gamma^{max}}|\mathcal{F}_t]$ et $\mathbb{E}[I_{i,\Gamma^{max}}|\mathcal{F}_t]$, ce qui sera l’objet de la partie suivante.

Cette mesure de provision correspond à une vision long-terme du provisionnement, dans une approche *Best Estimate*. La provision ultime correspond à l’espérance des flux futurs de paiements qui le sinistre i génère jusqu’à sa clôture.

Pour entrer dans le cadre IFRS 4 Phase II, une actualisation des flux pourrait être envisagée : en effet, nous verrons que dans le chapitre suivant que l’estimateur de $\mathbb{E}[Y_{i,\Gamma^{max}}|\mathcal{F}_t]$ est construit à partir de tous les flux attendus chaque année jusqu’à la clôture. Nous pourrions donc facilement actualiser en utilisant une courbe des taux d’intérêt *forward* (déterministe). .

1. Des problématiques liées aux évolutions des taux de change dans le temps pourraient être considérées. Cependant, les compagnies mettent en place une couverture de change à l’actif du bilan, associée aux flux de trésorerie futurs au passif. Le risque de change est donc traité séparément.

Une autre mesure à laquelle nous pourrions nous intéresser, qui n'est pas à proprement parler une provision, est la suivante :

Flux de trésorerie à un an

$\mathbb{E}CF_t^1[i]$ (resp. $\mathbb{E}CF_t^1[A]$) est appelée flux de trésorerie attendu à un an du sinistre S_i (resp. du sous-ensemble de sinistres $A \subseteq \mathcal{S}$) si :

$$\begin{aligned}\mathbb{E}CF_t^1[i] &= \mathbb{E}[Y_{i,t-n_i+1}|\mathcal{F}_t] - Y_{i,t-n_i}, \\ \mathbb{E}CF_t^1[A] &= \sum_{i \in \mathbf{A}} \mathbb{E}CF_t^1[i],\end{aligned}$$

où $\mathbb{E}[Y_{i,t-n_i+1}|\mathcal{F}_t]$ est la moyenne des paiements cumulés du sinistre S_i en année $t+1$, sachant toute l'information disponible aujourd'hui \mathcal{F}_t .

Cette mesure a trois intérêts principaux :

- Pour une compagnie d'assurance, il peut être intéressant de connaître quels seront les paiements à faire à un an, dans une optique d'investissement et de gestion actif-passif.
- D'un point de vue théorique, l'estimation des flux de trésorerie à un an constitue la pierre angulaire pour l'estimation de la réserve à l'ultime. En effet, dans les méthodes agrégées, avec un triangle des paiements cumulés, nous pouvons calculer la réserve ultime comme la somme des différences deux à deux des diagonales projetées. Ici, l'estimation de la charge ultime sera aussi envisagée comme les flux futurs estimés de paiements incrémentaux chaque année, en prenant en compte le statut clos ou non clos du sinistre.
- D'un point de vue pratique, cette mesure ne fait pas intervenir une estimation de la variable de durée de vie. Elle peut donc aider à discerner quelle estimation (processus de paiements ou processus d'états) est la moins robuste.

Enfin, une troisième mesure pourrait être analysée, qui serait la variation de charge à un an :

Évolution de la charge à un an

$\mathbb{E}R_t^1[i]$ (resp. $\mathbb{E}R_t^1[A]$) est appelée variation de charge à un an du sinistre S_i (resp. du sous-ensemble de sinistres $A \subseteq \mathcal{S}$) si :

$$\begin{aligned}\mathbb{E}R_t^1[i] &= \mathbb{E}[I_{i,t-n_i+1}|\mathcal{F}_t] - I_{i,t-n_i}, \\ \mathbb{E}R_t^1[A] &= \sum_{i \in \mathbf{A}} \mathbb{E}R_t^1[i],\end{aligned}$$

où $\mathbb{E}[I_{i,t-n_i+1}|\mathcal{F}_t]$ est la moyenne de la charge du sinistre S_i en année $t+1$, sachant toute l'information disponible aujourd'hui \mathcal{F}_t correspond à la réserve dossier-dossier faite en t pour le sinistre i .

Remarques En fait, dans le cadre de Solvabilité II, nous serions plutôt intéressés, non pas par un estimateur de la moyenne, mais par un estimateur de la volatilité de cette réserve.

En effet, si les gestionnaires de sinistres établissent leur charge dans un cadre *Best Estimate*, on devrait avoir : $\mathbb{E}[I_{i,t-n_i+1}|\mathcal{F}_t] = I_{i,t-n_i}$. On cherche plutôt à mesurer de combien la réalisation de $I_{i,t-n_i+1} - I_{i,t-n_i}$ aura dévié de 0 en $t + 1$.

Dans le calcul du SCR dans le cadre d'un modèle interne (éventuellement partiel), nous chercherions à déterminer le quantile à 99,5% de la variation de *Best Estimate* à un an. Nous proposerons des pistes pour étudier le risque de réserve à un an dans le cadre de notre modèle dans la partie ???. Il peut se révéler ensuite difficile de réunir cette volatilité avec la volatilité estimée sur les IBNR, afin de déterminer la volatilité totale du portefeuille.

Bilan

Dans cette partie, nous avons établi un cadre simple pour évaluer les provisions par une approche ligne à ligne. Ce cadre est restreint aux sinistres notifiés la même année que celle de leur survenance, mais des extensions du cadre peuvent être facilement envisagées. Nous avons fait le lien entre les ensembles de sinistres dans une base de données ligne à ligne, et ceux du triangle agrégé, pour illustrer le passage de la dimension individuelle à la dimension agrégée. Nous avons également défini la notion d'information disponible : comme dans le cas agrégé, les estimations des flux futurs de sinistres sont faites sachant les réalisations des variables aléatoires jusqu'à aujourd'hui. Enfin, nous avons présenté les mesures de provision auxquelles nous allons nous intéresser.

Nous allons maintenant proposer un modèle particulier s'inscrivant de ce cadre. Plus précisément, nous allons nous restreindre à une classe de processus de paiements (et de charges). Nous pourrions ensuite déduire des estimateurs des processus, qui permettraient dans un second temps d'estimer les mesures de provision d'intérêt. Nous allons enfin comparer le comportement asymptotique de l'estimateur des provisions ligne à ligne avec celui de Chain-Ladder.

Partie 4

Modèle et estimation

Pour estimer les mesures de provision d'intérêt, nous avons besoin d'hypothèses supplémentaires sur les processus facilitant les estimations. Nous pouvons considérer séparément l'estimation des processus d'états et les processus de paiements (comme mentionné dans la partie 3.3.3). Nous commençons par exposer la modélisation et l'estimation des processus d'états.

4.1 Modélisation et estimation des processus

4.1.1 Processus d'états

4.1.1.1 Hypothèse

Estimer le processus d'états $(D_{i,j})_{j \geq 0}$ revient à estimer la fonction de survie de la variable d'âge Γ_i , puisqu'on a :

$$\forall j > 0, \mathbb{P}(D_{i,j} = 1) = \mathbb{P}(\Gamma_i \geq j).$$

Pour pouvoir étudier statistiquement la durée de vie des sinistres, nous reprenons l'hypothèse de Chau et Dinh (2012) [7] :

Hypothèse 4.1.1 *Pour tout $i \in \mathcal{S}$, les variables d'âge ultime Γ_i sont indépendamment et identiquement distribuées.*

Remarque

Pour que cette hypothèse soit valide, il est nécessaire d'avoir une segmentation des bases de sinistres permettant d'obtenir une bonne homogénéité. Par exemple, dans le cas d'une garantie dommages tempête, il est possible d'observer une forte dissymétrie et des effets de regroupement (*clusters*) entre les régions affectées par des sinistres attritionnels (sinistres à haute fréquence et coût faible, réglés probablement rapidement), et des régions qui ont subi une tempête exceptionnelle (sinistres à faible fréquence et coût élevé, pouvant entraîner un déroulé plus long, dû aux délais de gestion allongés par la surcharge).

Il peut être aussi nécessaire de faire une segmentation temporelle. Par exemple, si un changement dans le protocole de gestion des sinistres, visant à traiter les dossiers plus rapidement et accélérer les remboursements, a eu lieu, l'hypothèse 4.1.1 ne sera plus valide. La mise en place

de conventions de recours entre assureurs (comme la convention IRSA) peut également modifier la durée de vie des sinistres, entre ceux traités avant et ceux traités après la mise en place.

4.1.1.2 Estimation

Pour estimer la durée de vie, nous avons donc recours aux techniques de l'analyse de survie (modèles de durée).

Comme certaines données sont nécessairement censurées à droite (les sinistres qui ne sont pas encore clos et dont nous devons prédire la vie future), nous ne pouvons pas utiliser les estimateurs usuels : par exemple, l'estimateur empirique usuel de la fonction de répartition est biaisé et non consistant¹ dans ce cas.

On peut utiliser l'estimateur non-paramétrique de Kaplan-Meier, qui estime la fonction de survie. Pour un sinistre S_i , la fonction de survie de Γ_i s'écrit :

$$S_{\Gamma_i}(j) = \mathbb{P}[\Gamma_i > j] = 1 - F_{\Gamma_i}(j).$$

La méthode d'estimation de Kaplan-Meier repose sur une idée simple : être encore ouvert à l'année j , c'est être ouvert juste avant (en année $j - 1$), et ne pas sortir de l'état "ouvert" entre $j - 1$ et j (clôture ou censure).

Plus précisément, on note :

- d_j , le nombre de transitions dans l'intervalle $]j - 1, j]$, c'est-à-dire le nombre de dossiers qui étaient ouverts en $j - 1$ et sont clos en j ;
- c_j , le nombre d'observations censurées dans l'intervalle $]j - 1, j]$, c'est-à-dire dont l'on connaissait le statut en $j - 1$ (ouvert/clos), mais que l'on n'observe plus dans leur année de développement j ;
- r_j , le nombre d'observations "à risque", c'est-à-dire les sinistres qui sont toujours ouverts en année de développement $j - 1$, et qui peuvent potentiellement connaître une censure ou une fermeture sur l'intervalle $]j - 1, j]$ (Il s'agit de $Card(A_{j-1}^*(t))$, où $Card(\cdot)$ désigne le cardinal d'un ensemble).

L'estimateur de Kaplan-Meier de la fonction de survie est :

$$\hat{S}_{\Gamma}(k) = \prod_{j|j \leq k} \frac{r_j - d_j}{r_j}.$$

Propriétés

A moins qu'aucune observation ne soit censurée², l'estimateur de Kaplan-Meier de la fonction de survie possède un biais positif (soit un biais négatif pour l'estimateur de la fonction de répartition $\hat{F}_{\Gamma}(k) = 1 - \hat{S}_{\Gamma}(k)$) (Gill, 1980 [20]). Plus précisément, l'estimateur est sans biais si la plus longue durée de vie $\Gamma(n)$ observée dans la base n'est pas censurée, ou si c'est le cas, il reste

1. Convergeant en probabilité.

2. Dans ce cas-là, l'estimateur de Kaplan-Meier coïncide avec l'estimateur empirique usuel de la fonction de survie/de répartition.

sans biais pour $j < \Gamma(n)$ (Fleming et Harrington, 2005 [16]). Il serait envisageable de réduire ce biais en utilisant la technique du Jackknife appliquée au cas de l'estimateur de Kaplan-Meier (proposé par Stute et Wang, 1994 [49]).

Asymptotiquement, l'estimateur est consistant : plus précisément on obtient une convergence p.s. uniforme, sous l'hypothèse que la fonction de répartition de la variable de survie et celle de la variable de censure n'ont aucune discontinuité commune¹.

L'estimateur est également asymptotiquement normal :

$$\sqrt{n}[\hat{S}_\Gamma(k) - S_\Gamma(k)] \xrightarrow[n \rightarrow \infty]{} \mathcal{N}(0, V(k)),$$

où $V(k)$ est la variance asymptotique :

$$V(k) = S_\Gamma^2(k) \int_0^k \frac{f(t)dt}{S_\Gamma^2(t)G(t)},$$

avec $G(t)$ la fonction de survie de la variable de censure.

Un estimateur de la variance asymptotique ([24]) est :

$$\hat{V}(k) = -\hat{S}_\Gamma(k) \prod_{j|j \leq k} \frac{d_j}{r_j(r_j - d_j)}.$$

On peut montrer que l'estimateur de Kaplan-Meier est un estimateur non-paramétrique du maximum de vraisemblance ([18], ch.2).

Bilan

L'estimateur de Kaplan-Meier est asymptotiquement sans biais, mais biaisé à distance finie. Cependant, ce biais apparaît en queue de distribution de la durée de vie. Avec un échantillon suffisamment large, l'erreur commise en vision *Best Estimate* devrait donc être faible.

Une solution alternative pourrait être d'utiliser des modèles paramétriques, en postulant une loi paramétrique pour la densité de Γ . Il serait également possible d'introduire des variables explicatives en explicitant un lien entre le paramètre de la loi et des variables individuelles liées au sinistre (année de déclaration, charge initiale, etc.). Cela a été testé par Beneteau (2002) [3]. Cependant, le risque de faire une erreur de modèle reste important.

Maintenant que nous avons défini un estimateur de la durée de vie des sinistres, nous allons passer à la modélisation des processus de paiements.

1. Les énoncés complets des différents résultats de convergence de l'estimateur de Kaplan-Meier peuvent être trouvés dans [14].

4.1.2 Processus de règlements

4.1.2.1 Hypothèses

La distribution des $(C_{i,j})_{j \geq 0}$ a potentiellement une structure très complexe. Plusieurs choix permettant une modélisation simplifiée pourraient être retenus pour modéliser les processus, notamment les processus de paiements (cumulés) : modélisation paramétrique, modèles bayésiens, etc.

Nous avons retenu une hypothèse de dynamique markovienne pour ces processus, comme cela est parfois fait dans les modèles agrégés, notamment le modèle de Mack (1993) (voir partie 2.3). Cette hypothèse, couplée à celle d'invariance par changement de numéraire, permet de se ramener à l'étude de **facteurs de développement individuel** $F_{i,j}$, définis par : $C_{i,j+1} = F_{i,j}C_{i,j}$.

L'idée d'étudier les facteurs de développement individuels plutôt que la suite des paiements a déjà été introduite par Murphy et McLennan (2006) [35]. Le modèle proposé ici permet d'obtenir l'indépendance des facteurs de développement avec le passé, que les auteurs avaient supposé dans leur étude pratique.

L'obtention de propriétés intéressantes sur les facteurs de développement individuels, notamment leur indépendance avec le passé, facilitera également la comparaison avec Chain-Ladder après agrégation des données.

Nous allons donc nous restreindre à une certaine classe de processus : les Chaînes de Markov invariantes par changement de numéraire. Les hypothèses suivantes sont faites :

Hypothèse 4.1.2

– *Indépendance* :

$$\text{Les sinistres } (S_i)_{i=1, \dots, \text{Card}(\mathcal{S})} \text{ sont indépendants.} \quad (4.1)$$

– *Dynamique markovienne des paiements cumulés* :

$$\forall i \in \mathcal{S}, \forall j \geq 0, \forall x \in \mathbb{R}, \mathbb{P}(C_{i,j+1} \leq x | C_{i,j}, C_{i,j-1}, \dots, C_{i,0}) = \mathbb{P}(C_{i,j+1} \leq x | C_{i,j}). \quad (4.2)$$

– *Invariance par changement de numéraire* :

$$\begin{aligned} \forall i \in \mathcal{S}, \forall j \geq 0, \forall \alpha > 0, \forall x \in \mathbb{R}, \mathbb{P}(C_{i,j+1} \leq \alpha x | C_{i,j} = \alpha c_j, \dots, C_{i,0} = \alpha c_0) \\ = \mathbb{P}(C_{i,j+1} \leq x | C_{i,j} = c_j, \dots, C_{i,0} = c_0) \end{aligned} \quad (4.3)$$

– *Même dynamique conditionnelle pour tous les sinistres* :

$$\forall x \in \mathbb{R}, \mathbb{P}(C_{i,j+1} \leq x | C_{i,j}, \dots, C_{i,0}) \text{ ne dépend pas de l'indice } i. \quad (4.4)$$

Remarques

Nous avons formulé les hypothèses 4.1.2 pour le processus de paiements cumulés. Nous faisons également **les mêmes hypothèses pour le processus de charges**. L'idée est de pouvoir étudier les paiements cumulés et les charges, comme dans la méthode Chain-Ladder. Nous présentons les résultats sur les paiements cumulés dans la suite, par analogie à Chain-Ladder, où les résultats sont généralement présentés sur le triangle des règlements cumulés.

- L'hypothèse d'indépendance des sinistres (4.1) est vérifiée si le portefeuille considéré est suffisamment diversifié. S'il y a des effets de concentration trop importants (p. ex. le cas d'une garantie incendie avec de nombreux assurés vivant dans un même immeuble), l'hypothèse peut ne pas être valide.
- L'hypothèse (4.2) peut ne pas être vérifiée en pratique. Par exemple, on peut observer une série de paiements incrémentaux ($X_{i,0} = 10, X_{i,1} = 10, X_{i,2} = 10$) et une autre série ($X'_{i,0} = 0, X'_{i,1} = 0, X'_{i,2} = 30$), qui conduisent au même montant cumulé C_2 par des chemins différents. Mais la prévision des paiements l'année prochaine par un gestionnaire de sinistres ne serait vraisemblablement la même. C'est également vraie pour les charges. De plus, l'ensemble d'information que constitue $C_{i,j}$ peut être assez sensible au découpage temporel choisi (inventaire fait en milieu d'année ou en fin d'année), ce qui peut rendre la dernière information insuffisante pour estimer le prochain paiement ou charge.
- L'hypothèse (4.3) est assez naturelle (l'unité monétaire utilisée ne doit pas changer la mesure des provisions), même si l'inflation est liée à la monnaie (cela reste un effet de second ordre).
- L'hypothèse d'équi-distribution conditionnelle (4.4) implique que toutes les autres caractéristiques d'un sinistre S_i (p. ex. l'année de survenance) n'influencent pas la distribution conditionnelle des paiements cumulés.

En utilisant :

- l'hypothèse d'indépendance entre les sinistres (4.1),
- et l'hypothèse, pour un sinistre donné, d'indépendance entre les processus de paiements et d'états (3.3.3),

l'hypothèse (4.2) ci-dessus s'écrit (de manière équivalente) :

$$\mathbb{P}(C_{i,t-n_i+1} \leq x | \mathcal{F}_t) = \mathbb{P}(C_{i,t-n_i+1} \leq x | C_{i,t-n_i}),$$

pour toute année t .

En utilisant un théorème de théorie du contrôle¹, la relation précédente se réécrit :

1. Il s'agit de la représentation d'une chaîne de Markov (inhomogène) définie sur un espace non dénombrable sous forme de suite récurrente aléatoire. L'énoncé du théorème et de la preuve peuvent être trouvées dans Benaïm et El Karoui (2005), p. 90 et sseq. [2].

$$C_{i,j+1} = \phi_{i,j}(C_{i,j}, U_{i,j}),$$

où $\phi_{i,j}$ une fonction sur \mathbb{R}^{\neq} et $U_{i,j}$ suit une loi uniforme, indépendante de \mathcal{F}_t avec $j = t - n_i$.

D'après la remarque ci-dessus, l'hypothèse (4.3) est équivalente à :

$$\forall \alpha > 0, \forall y \in \mathbb{R}, \forall x \in \mathbb{R}, \mathbb{P}(\phi_{i,j}(C_{i,j}, U_{i,j}) \leq \alpha x | C_{i,j} = \alpha y) = \mathbb{P}(\phi_{i,j}(C_{i,j}, U_{i,j}) \leq x | C_{i,j} = y)$$

ou de façon équivalente :

$$\forall \alpha > 0, \forall y \in \mathbb{R}, \forall x \in \mathbb{R}, \mathbb{P}(\phi_{i,j}(\alpha y, U_{i,j}) \leq \alpha x) = \mathbb{P}(\phi_{i,j}(y, U_{i,j}) \leq x).$$

Cela implique (sous certaines contraintes de régularité sur $\phi_{i,j}$) :

$$\forall \alpha > 0, \forall y \in \mathbb{R}, \forall u \in [0, 1], \phi_{i,j}(\alpha y, u) = \alpha \phi_{i,j}(y, u),$$

c'est-à-dire que $\phi_{i,j}$ est homogène de degré 1 dans son premier argument.

En prenant : $y = 1, u = U_{i,j}, \alpha = C_{i,j}$ et :

$$\phi_{i,j}(C_{i,j}, U_{i,j}) = \phi_{i,j}(1, U_{i,j})C_{i,j},$$

on obtient finalement :

$$C_{i,j+1} = C_{i,j}F_{i,j} \tag{4.5}$$

avec $F_{i,j} = \phi_{i,j}(1, U_{i,j})$ indépendant de \mathcal{F}_t , avec t tel que $j = t - n_i$. En particulier, $F_{i,j}$ est indépendant de $C_{i,j}$. Les $F_{i,j}$ sont appelés les **facteurs de développement individuels** (stochastiques), par analogie avec Chain-Ladder.

L'hypothèse (4.4) implique que la loi des $F_{i,j}$ est la même pour tout i (elle ne dépend que de l'indice j).

Résumé des hypothèses faites jusqu'à présent

1. Les sinistres sont indépendants (3.1). Les processus de paiements (resp. de charges) sont indépendants des processus d'états (3.3.3).
2. Les processus d'états sont i.i.d. (4.1.1).
3. Les variables $F_{i,j}$ définies par (4.5) sont indépendantes du passé (c'est-à-dire de \mathcal{F}_u , avec u tel que $j = u - n_i$). Pour un indice j donné, les $F_{i,j}$ suivent la même loi pour tout i . Nous noterons F_j la variable aléatoire représentant la loi des $F_{i,j}$.

4.1.2.2 Estimation

Avec la restriction à la classe de processus NIPMD, le problème de prédiction des paiements cumulés futurs se ramène à un problème de prédiction des facteurs de développement individuels futurs. Nous avons besoin maintenant de pouvoir estimer $\mathbb{E}[F_j]$ afin d'estimer les réserves.

Pour cela, nous voulons utiliser la même idée au niveau individuel que l'estimateur Chain-Ladder au niveau agrégé. En effet, les facteurs de développement pour les « petits » sinistres, que cela soit au niveau des paiements cumulés ou des charges, risquent d'être très volatils et de contenir des valeurs « aberrantes » (*outliers*). Au contraire, pour les sinistres ayant des grandes charges, les facteurs seront certainement plus stables. La moyenne empirique, estimateur usuel de la moyenne, risque d'être trop sensible aux valeurs extrêmes, surtout s'il y a un grand nombre de « petits » sinistres dans la base. L'idée est donc d'utiliser une moyenne pondérée par l'exposition, comme dans Chain-Ladder, c'est-à-dire en affectant un poids plus important aux plus gros sinistres.

Pour que la moyenne empirique pondérée soit convergente, nous avons besoin d'utiliser une généralisation de la loi des grands nombres, dans le cas de variables indépendantes mais non identiquement distribuées. Nous avons besoin de conditions techniques supplémentaires pour pouvoir appliquer ce théorème.

Hypothèse 4.1.3 : Conditions techniques

Il existe $\delta > 0$ tel que :

1. $\forall j \geq 0, \mathbb{E}[(F_j)^{1+\delta}] < \infty$;
2. $\sup_{i \in \mathcal{S}} \left\{ \mathbb{E}[(C_{i,0})^{1+\delta}] \right\} < \infty$;

De plus, on suppose que :

3. $\exists c > 0, \forall S_i \in \mathcal{S}, C_{i,0} > c$ \mathbb{P} -p.s. ;
4. $\forall j \geq 0, \exists c' > 0, \forall S_i \in \mathcal{S}, F_{i,j} \geq c'$ \mathbb{P} -p.s.

Remarques

- La condition 1 est une condition classique qui stipule qu'il existe un moment d'ordre strictement supérieur à 1 fini.
- La condition 2 permet de s'assurer qu'aucun facteur de développement théorique ne domine tous les autres.
- La condition 3 implique que le premier paiement cumulé doit être nécessairement strictement positif. C'est une condition très forte si l'on considère les paiements cumulés (il n'y a pas forcément d'indemnisation lors de la première année de notification). En considérant les charges en lieu des paiements, cette condition est beaucoup moins forte : à la notification, le gestionnaire de sinistres constitue généralement immédiatement une provision dossier-dossier (forfait d'ouverture). Il peut cependant exister des délais de gestion qui font

qu'un sinistre déclaré n'est pas nécessairement immédiatement provisionné.

- La condition 4 revient à exclure les cas où des paiements cumulés futurs (ou les charges) seraient nuls, voire négatifs (c.-à-d. à travailler brut de recours et ne traitant pas les *nil claims*).

Notons que les conditions 2 et 3 impliquent : $\sup_{i \in \mathcal{S}} \mathbb{E}[C_{i,0}] < \infty$ et les conditions 1 et 4 impliquent : $\mathbb{E}[F_j] < \infty$ (les moments d'ordre inférieurs à $1 + \delta$ existent).

Ces conditions sont toutes vérifiées si l'on suppose que les premiers paiements $(C_{i,0})$ pour tout $S_i \in \mathcal{S}$ suivent la même loi à valeurs réelles strictement positives (par exemple une loi log-normale), tronquée en-dessous d'un seuil minimum c , et que les variables $F_{i,j}, C_{i,0}$ disposent tous un moment d'ordre strictement supérieur à 1 finie (par exemple, si elles sont de variance finie).

On évite de supposer une propriété d'équi-distribution sur les premiers paiements afin de ne pas exclure les cas où la base n'est pas suffisamment homogène pour que cette propriété soit vérifiée.

Loi des grands nombres pour des variables i.n.i.d.

Soit $(A_i)_{i \geq 1}$ une suite de variables aléatoires indépendantes non identiquement distribuées (i.n.i.d.), de moyennes finies $\mathbb{E}[A_i] = \mu_i$ pour tout $i \geq 1$. S'il existe $\delta > 0$ tel que :

$$\sum_{i=1}^{\infty} \frac{\mathbb{E} \left[|A_i - \mu_i|^{1+\delta} \right]}{i^{1+\delta}} < \infty, \quad (4.6)$$

alors :

$$\boxed{\left(\frac{\sum_{t=1}^n A_t}{n} - \frac{\sum_{i=1}^n \mu_i}{n} \right) \xrightarrow[n \rightarrow +\infty]{p.s.} 0.}$$

L'énoncé et la preuve peuvent être trouvés dans Chung, 2001 (p. 132) [8]. La condition (4.6) qui restreint les moments de (A_i) est appelée condition de Markov. Le cas $\delta = 1$ est la loi des grands nombres de Kolmogorov.

Proposition

La condition de Markov (4.6) est vérifiée si : $\exists \delta > 0, M_\delta = \sup_{i > 0} \mathbb{E} \left[|A_i|^{1+\delta} \right] < \infty$.

Preuve de la proposition

Par l'inégalité de Hölder, on a :

$$|\mu_i| = |\mathbb{E}[A_i]| \leq \mathbb{E}[|A_i|] \leq \left(\mathbb{E}[|A_i|^{1+\delta}] \right)^{\frac{1}{1+\delta}},$$

d'où :

$$|\mu_i|^{1+\delta} \leq M_\delta. \quad (4.7)$$

De plus, l'inégalité c_r , avec $r = 1 + \delta$ ici donne :

$$\mathbb{E}[|A_i - \mu_i|^{1+\delta}] = 2^\delta \left(\mathbb{E}[|A_i|^{1+\delta}] + |\mu_i|^{1+\delta} \right) \quad (4.8)$$

Les équations (4.7) et (4.8), ainsi que l'hypothèse de la proposition donnent :

$$\mathbb{E}[|A_i - \mu_i|^{1+\delta}] < 2^\delta M_\delta,$$

et la série $2^\delta M_\delta \sum_{i=1}^{\infty} \frac{1}{i^{1+\delta}}$ est convergente.

Lemme 1

Soit $\delta > 0$ et $\{A_i, B_i | i \in \mathbb{N}^*\}$ un ensemble de variables aléatoires avec A_i et B_i indépendantes, tel qu'il existe $\delta > 0$ tel que :

1. les variables (B_i) sont i.i.d, d'espérance b et admet un moment d'ordre $1 + \delta$ fini ;
2. les variables (A_i) sont indépendantes, d'espérance $\mathbb{E}[A_i] = a_i$, et satisfont :
 - (a) $M_\delta = \sup_{i>0} \mathbb{E}[|A_i|^{1+\delta}] < \infty$.
 - (b) $\exists a > 0, \forall i > 0, A_i > a$ \mathbb{P} -p.s.

Alors :

$$\boxed{\frac{\sum_{i=1}^n A_i B_i}{\sum_{i=1}^n A_i} \xrightarrow[n \rightarrow +\infty]{p.s.} b.}$$

Théorème Sous les conditions techniques 4.1.3), $\frac{\sum_{i \in \mathcal{B}_{j+1}^*} C_{i,j+1}}{\sum_{i \in \mathcal{B}_{j+1}^*} C_{i,j}}$ est un estimateur consistant de $\mathbb{E}[F_j]$, i.e. :

$$\frac{\sum_{i \in \mathcal{B}_{j+1}^*} C_{i,j+1}}{\sum_{i \in \mathcal{B}_{j+1}^*} C_{i,j}} \xrightarrow[\text{Card}(\mathcal{B}_{j+1}^*) \rightarrow \infty]{} \mathbb{E}[F_j].$$

Preuve du théorème

Il suffit d'appliquer le lemme 1 à $B_i = \frac{C_{i,j+1}}{C_{i,j}} = F_{i,j}$ et $A_i = C_{i,j}$ pour $i \in \mathcal{B}_{j+1}^* \subseteq \mathcal{S}$. Les variables $F_{i,j}$ et $C_{i,j}$ vérifient bien les hypothèses du lemme 1, d'après les conditions techniques supposées 4.1.3).

Démonstration du lemme 1

En posant $C_i = A_i B_i$, on a $\mathbb{E}[|C_i|^{1+\delta}] = \mathbb{E}[|A_i|^{1+\delta}] \mathbb{E}[|B_i|^{1+\delta}]$, d'où :

$$\sup_{i>0} \mathbb{E}[|C_i|^{1+\delta}] < \infty.$$

L'hypothèse de la proposition est vérifiée, on a donc :

$$\frac{1}{n} \sum_{i=1}^n C_i - \frac{1}{n} \sum_{i=1}^n a_i b \xrightarrow[n \rightarrow +\infty]{} 0.$$

En appliquant la loi des grands nombres à la suite A_i , on a aussi :

$$\frac{1}{n} \sum_{i=1}^n A_i - \frac{1}{n} \sum_{i=1}^n a_i \xrightarrow[n \rightarrow +\infty]{} 0$$

La démonstration du lemme 1 se termine en utilisant le lemme ci-dessous, en notant que $\frac{1}{n} \sum_{i=1}^n A_i > a > 0$.

Lemme 2

Soit G_n, H_n deux suites de variable aléatoires et g_n, h_n deux suites réelles telles que :

1. $\exists a > 0, \forall n > 1, H_n > a$ \mathbb{P} -p.s. ;
2. $(G_n - g_n) \xrightarrow[n \rightarrow +\infty]{} 0$ et $(H_n - h_n) \xrightarrow[n \rightarrow +\infty]{} 0$ \mathbb{P} -p.s. ;
3. $\lim_{n \rightarrow \infty} \frac{g_n}{h_n} = \alpha$.

Alors :

$$\frac{G_n}{H_n} \xrightarrow[n \rightarrow +\infty]{p.s.} \alpha. \tag{4.9}$$

Démonstration du lemme 2

$$\begin{aligned} \left| \frac{G_n}{H_n} - \alpha \right| &\leq \left| \frac{G_n - g_n}{H_n} \right| + \left| \frac{g_n}{H_n} - \frac{g_n}{h_n} \right| + \left| \frac{g_n}{h_n} - \alpha \right| \\ &\leq \left| \frac{G_n - g_n}{a} \right| + \left| \frac{g_n (H_n - h_n)}{h_n a} \right| + \left| \frac{g_n}{h_n} - \alpha \right| \\ &\xrightarrow[n \rightarrow +\infty]{} 0 \text{ } \mathbb{P}\text{-p.s.} \end{aligned}$$

Estimateurs de la moyenne des facteurs de développement

Nous pouvons maintenant donner deux estimateurs consistants de $\mathbb{E}[F_j]$.

Notons d'abord que : $F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$, et qu'à la date t , on connaît $(C_{i,j+1}, C_{i,j}, F_{i,j})$ uniquement pour les sinistres $i \in \mathcal{B}_{j+1}^*(t)$.

Si $\mathcal{B}_{j+1}^*(t)$ est un ensemble non vide, les deux estimateurs décrits ci-dessous sont asymptotiquement consistants :

- L'estimateur usuel de la moyenne, la moyenne empirique :

$$\hat{f}_j = \frac{1}{\text{Card}(\mathcal{B}_{j+1}^*)} \sum_{i \in \mathcal{B}_{j+1}^*} \frac{C_{i,j+1}}{C_{i,j}}$$

- L'estimateur de la moyenne calculé par analogie avec Chain-Ladder :

$$\widehat{\mathbb{E}F}_j(t) = \frac{\sum_{i \in \mathcal{B}_{j+1}^*} C_{i,j+1}}{\sum_{i \in \mathcal{B}_{j+1}^*} C_{i,j}}$$

Ces deux estimateurs convergent \mathbb{P} -p.s. vers l'espérance de F_j , le premier par la loi des grands nombres classique, le second par le lemme 1 précédent.

Maintenant que nous avons obtenu des estimateurs consistants pour les processus de base, nous pouvons passer à l'estimation des flux de trésorerie à un an et des réserves à l'ultime.

4.2 Estimation des provisions

4.2.1 Flux de trésorerie à un an

Si on se place à la date t , la somme des paiements incrémentaux à la date $t + 1$ peut s'écrire :

$$CF^1 = CF^1(t) = \sum_{i \in \mathcal{S}} X_{i,t-n_i+1} D_{i,t-n_i+1}.$$

En séparant les sinistres selon les différentes années de développement (à la date t), on écrit¹ : $CF^1 = \sum_{n=0}^{\infty} CF_n^1$, avec :

$$CF_n^1 = \sum_{i \in \mathcal{A}_n} (Y_{i,n+1} - Y_{i,n}) \quad (4.10)$$

On souhaite à partir de l'information disponible \mathcal{F}_t , estimer la distribution de CF^1 . En particulier, dans cette partie, on va estimer son espérance conditionnelle $\mathbb{E}[CF^1 | \mathcal{F}_t]$, c'est-à-dire les flux de paiements à un an défini dans la partie 3.5.2.

Remarquons que l'on peut écrire :

$$\frac{Y_{i,n+1}}{Y_{i,n}} = \begin{cases} \frac{C_{i,n+1}}{C_{i,n}} = F_{i,n} & \text{si le sinistre } i \text{ est en cours en année de déroulement } n+1 \text{ } (i \in \mathcal{A}_{n+1}^*), \\ 1 & \text{si le sinistre } i \text{ est clos avant l'année de déroulement } n+1. \end{cases}$$

L'équation (4.10) se réécrit donc :

$$CF_n^1 = \sum_{i \in \mathcal{A}_n} Y_{i,n} D_{i,n+1} (F_{i,n} - 1)$$

En passant à l'espérance, et comme par hypothèse (voir 3.4.1), $D_{i,n+1}$ est \mathcal{F}_t -mesurable pour tout $i \in \mathcal{A}_{n+1}^*$, on a :

$$\mathbb{E}[CF_n^1 | \mathcal{F}_t] = \sum_{i \in \mathcal{A}_n} Y_{i,n} D_{i,n+1} (\mathbb{E}[F_{i,n}] - 1)$$

Un estimateur de $\mathbb{E}[CF_n^1 | \mathcal{F}_t]$ est donné par :

$$\widehat{ECF}_n^1(t) = \sum_{i \in \mathcal{A}_n(t)} Y_{i,n} D_{i,n+1} (\widehat{EF}_n(t) - 1),$$

qui est particulier un estimateur \mathcal{F}_t -mesurable.

Nous avons de plus le théorème suivant :

Théorème

$$\boxed{\frac{\widehat{ECF}_n^1 - \mathbb{E}[CF_n^1]}{\text{Card}(\mathcal{A}_n)} \xrightarrow{p.s.} 0 \text{ quand } \min(\text{Card}(\mathcal{A}_n), \text{Card}(\mathcal{B}_{n+1}^*)) \rightarrow \infty.}$$

1. On somme jusqu'à l'infini pour alléger les notations, avec l'hypothèse 3.3.2, les flux de trésorerie à un an pour les sinistres à un stade de développement $n \geq \Gamma^{max}$ en année t sont nuls

Preuve

Pour $i \in \mathcal{A}_n$, nous avons :

$$Y_{i,n}D_{i,n+1} = C_{i,n}D_{i,n+1},$$

c'est-à-dire que le processus de paiements $C_{i,n}$ est observé.

De plus, par indépendance entre le processus de paiements et le processus d'états (3.3.3), on a $\mathbb{E}[C_{i,n}D_{i,n+1}] = \mathbb{E}[C_{i,n}]p_{n+1}$, (avec $p_{n+1} = \mathbb{P}[D_{i,n+1} = 1] = \mathbb{E}[D_{i,n+1}]$). Par la loi des grands nombres, nous obtenons alors :

$$\frac{\sum_{i \in \mathcal{A}_n} C_{i,n}D_{i,n+1}}{\text{Card}(\mathcal{A}_n)} - \frac{\mathbb{E}[\sum_{i \in \mathcal{A}_n} C_{i,n}]p_{n+1}}{\text{Card}(\mathcal{A}_{n+1}^*)} \xrightarrow[\mathcal{A}_{n+1}^* \rightarrow \infty]{p.s.} 0,$$

De plus, nous avons aussi $\widehat{\mathbb{E}F}_n \xrightarrow{p.s.} \mathbb{E}[F_n]$ lorsque $\text{Card}(\mathcal{B}_n^*) \rightarrow \infty$.

Il reste à montrer que :

$$\mathbb{E} \left[\sum_{i \in \mathcal{A}_n} C_{i,n} \right] p_{n+1} (\mathbb{E}[F_j] - 1) = \mathbb{E}[CF_n^1].$$

Or :

$$\begin{aligned} \mathbb{E} \left[\sum_{i \in \mathcal{A}_n} C_{i,n} \right] p_{n+1} (\mathbb{E}[F_j] - 1) &= \mathbb{E}[C_{i,n+1} - C_{i,n}]p_{n+1} \\ &= \mathbb{E}[(C_{i,n+1} - C_{i,n})D_{i,n+1}] \\ &= \mathbb{E}[Y_{i,n+1} - Y_{i,n}]. \end{aligned}$$

Remarques

La convergence dans notre modèle s'opère de manière très différente de celle des méthodes agrégées. Par exemple, dans un modèle GLM sur le triangle agrégé, la consistance des estimateurs s'obtient lorsque le nombre de cellules du triangle tend vers l'infini, c'est-à-dire lorsque la durée d'observation tend vers l'infini. Dans le modèle ligne à ligne, c'est le nombre de sinistres dans une cellule qui doit tendre vers l'infini. C'est pourquoi nous nous intéressons à des convergences "en moyenne", c'est-à-dire en normalisant par le nombre de sinistres.

Le calcul de $\widehat{\mathbb{E}F}_n(t)$ requiert que $\mathcal{A}_n^*(t)$ soit non-vide. Pour la convergence de $\widehat{\mathbb{E}CF}_n^1$, nous avons besoin de :

- soit $\mathcal{A}_{n+1}^*(t)$ non-vide, afin d'estimer la moyenne des facteurs de développement ($\widehat{\mathbb{E}F}_n$). Cela est équivalent à ce que $t - \min_{i \in \mathcal{S}} n_i \geq n + 1$. D'où la nécessité de supposer que $\Gamma^{max} \leq t - \min_{i \in \mathcal{S}} n_i$, c'est-à-dire qu'il n'est pas possible pour un sinistre d'avoir une durée de vie supérieure à la fenêtre maximale d'observation de la vie des sinistres.
- soit \mathcal{A}_{n+1}^* vide lorsque $\text{Card}(\mathcal{A}_{n+1}^*) \rightarrow \infty$, et la durée d'observation est telle que $t - \min_{i \in \mathcal{S}} n_i \geq n + 1$. Cela entraîne $p_{n+1} = 0$, et les flux de trésorerie à un an CF_n^1 sont nuls.

4.2.2 Réserve à l'ultime

Nous allons maintenant chercher à estimer les réserves ultimes dans le cadre de notre modèle ligne à ligne. L'estimation de la réserve à l'ultime est plus compliquée, puisqu'elle fait intervenir les dates de clôture futures. Nous avons donc d'abord besoin d'exhiber une propriété particulière du processus d'états.

4.2.2.1 Propriété du processus d'états

La définition retenue pour les processus d'états permet d'obtenir une propriété qui est utile lors de l'étude de la convergence de l'estimateur de la réserve à l'ultime. Notamment, si nous savons qu'un sinistre est ouvert aujourd'hui, nous pouvons exprimer facilement la probabilité qu'un sinistre soit clos à une date future.

En effet, soit D une suite de variables de Bernoulli satisfaisant :

$$\mathbb{P}[D_0 = 1] = 1 \tag{4.11}$$

$$\forall j \geq 1, \mathbb{P}[D_{j+1} = 0 | D_j = 0] = 1, \tag{4.12}$$

On pose : $\mathbb{P}[D_j = 1] = p_j \in [0, 1]$.

Alors, on peut montrer la propriété suivante sur cette variable :

Propriété

Pour tout $0 \leq a < b$,

$$\mathbb{P}[D_{a+1} = D_{a+2} = \dots = D_b = 1, D_{b+1} = 0 | D_a = 1] = \frac{p_b - p_{b+1}}{p_a} \tag{4.13}$$

Preuve

On rappelle que la durée de vie Γ est telle que $\mathbb{P}[D_j = 1] = \mathbb{P}[\Gamma \geq j]$. On a donc :

$$\begin{aligned} \mathbb{P}[D_{a+1} = D_{a+2} = \dots = D_b = 1, D_{b+1} = 0 | D_a = 1] &= \mathbb{P}[\Gamma = b | \Gamma \geq a] \\ &= \frac{\mathbb{P}[\Gamma \geq b] - \mathbb{P}[\Gamma \geq b+1]}{\mathbb{P}[\Gamma \geq a]} \\ &= \frac{p_b - p_{b+1}}{p_a} \end{aligned}$$

4.2.2.2 Estimation de la réserve à l'ultime

À la date t , la somme des paiements futurs sur les sinistres $i \in \mathcal{S}$ (déjà déclarés) s'écrit :

$$CF^u = CF^u(t) = \sum_{i \in \mathcal{S}} \sum_{s > t} X_{i,s-n_i} D_{i,s-n_i}$$

En séparant les sinistres en différentes années de développement (à la date t), on peut écrire $CF^u = \sum_{n=0}^{\infty} CF_n^u$, avec :

$$CF_n^u = \sum_{i \in \mathcal{A}_n(t)} (Y_{i,u} - Y_{i,n}),$$

où $Y_{i,u} = \lim_{j \rightarrow \infty} Y_{i,j}$. Notons que pour tout j suffisamment grand, (notamment $j > \Gamma^{max}$), $D_{i,j} = 0$ et donc $Y_{i,j} = Y_{i,j-1}$, d'après l'hypothèse 3.3.2.

On a d'après l'équation (3.4) de décomposition entre les processus de paiements et ceux d'états :

$$Y_{i,j} = C_{i,j} \mathbb{1}\{D_{i,j} = 1\} + \sum_{k=0}^{j-1} C_{i,k} \mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\}$$

Nous pouvons séparer $\mathcal{A}_n(t)$ entre les sinistres qui sont ouverts l'année prochaine et ceux qui seront fermés l'année prochaine, comme $D_{i,n+1}$ est \mathcal{F}_t -mesurable. Ainsi :

– pour le sous-ensemble de sinistres $\{S_i | D_{i,n+1} = 1\}$, on a :

$$Y_{i,u} = \sum_{k=n+1}^{\infty} C_{i,k} \mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\},$$

et

$$Y_{i,n} = C_{i,n}.$$

– pour le sous-ensemble de sinistres $\{S_i | D_{i,n+1} = 0\}$, $Y_{i,n+1} - Y_{i,n} = 0$.

Il s'en suit :

$$\begin{aligned} Y_{i,u} - Y_{i,n} &= D_{i,n+1} \sum_{k=n+1}^{\infty} (C_{i,k} - C_{i,n}) \mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\} \\ &= D_{i,n+1} C_{i,n} \sum_{k=n+1}^{\infty} \left(\prod_{l=n}^{k-1} F_i^l - 1 \right) \mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\} \end{aligned}$$

Et ainsi en passant à l'espérance :

$$\begin{aligned} \mathbb{E}R_t^u[\mathcal{A}_n] &= \mathbb{E}[Y_i^u - Y_{i,n} | \mathcal{F}] = D_{i,n+1} C_{i,n} \sum_{k=n+1}^{\infty} \left(\prod_{l=n}^{k-1} \mathbb{E}[F_i^l] - 1 \right) \mathbb{E}[\mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\} | D_{i,n+1} = 1] \\ &= D_{i,n+1} C_{i,n} \sum_{k=n+1}^{\infty} \left(\prod_{l=n}^{k-1} \mathbb{E}[F_i^l] - 1 \right) \frac{p_k - p_{k+1}}{p_{n+1}}, \end{aligned}$$

où $p_j = P(D_{i,j} = 1)$, en appliquant la propriété (4.13).

$\mathbb{E}R_t^u[\mathcal{A}_{n+1}]$ peut alors être estimé par :

$$\widehat{\mathbb{E}R}_n^u = \sum_{i \in \mathcal{A}_n(t)} D_{i,n+1} C_{i,n} \sum_{k=n+1}^{\infty} \left(\prod_{l=n}^{k-1} \widehat{\mathbb{E}F}_l(t) - 1 \right) \frac{\hat{p}_k - \hat{p}_{k+1}}{\hat{p}_{n+1}}, \quad (4.14)$$

avec \hat{p}_j des estimateurs consistants et \mathcal{F}_t -mesurables de p_j .

Dans la section 4.3.2 ci-dessous, nous montrerons que l'estimateur donné par 4.14 est consistant, sous l'hypothèse supplémentaire d'équi-distribution du premier paiement.

4.3 Comparaison avec la méthode agrégée Chain-Ladder

Nous allons maintenant étudier le comportement asymptotique de l'estimateur de Chain-Ladder dans le cadre de notre modèle individuel, après agrégation des données. Les résultats de convergence sont donnés p.s..

4.3.1 Flux de trésorerie à un an

Dans la méthode de Chain-Ladder classique, on ne distingue pas les sinistres clos et non clos. A la date t , on dispose du triangle de paiements cumulés (L_s^j) avec $\min_{i \in \mathcal{S}} n_i \leq s \leq t$ et $j \leq t - n$:

$$L_s^j = \sum_{i \in \mathcal{S}, n_i = s} Y_{i,j},$$

la somme des paiements en année de développement j des sinistres de l'année d'origine s .

L'estimation de la réserve à un an par Chain-Ladder s'écrit dans le cadre de notre modèle :

$$CLR^1 = \sum_s L_s^{t-s} G_{t-s} = \sum_j L_{t-j}^j (G_j - 1),$$

avec G_j le facteur de développement, défini par :

$$G_j = \frac{\sum_{s \leq t-j-1} L_s^{j+1}}{\sum_{s \leq t-j-1} L_s^j} = \frac{\sum_{i \in \mathcal{B}_{j+1}} Y_{i,j+1}}{\sum_{i \in \mathcal{B}_{j+1}} Y_{i,j}}.$$

Il faut donc comparer $CL_j = L_{t-j}^j (G_j - 1)$ et $\widehat{\mathbb{E}CF}_j^1$, où :

$$\begin{aligned} \widehat{\mathbb{E}CF}_j^1 &= \left(\sum_{i \in \mathcal{A}_j} Y_{i,j} D_{i,j+1} \right) \left(\frac{\sum_{i \in \mathcal{B}_{j+1}^*} C_{i,j+1}}{\sum_{i \in \mathcal{B}_{j+1}^*} C_{i,j}} - 1 \right), \\ CL_j &= L_{t-j}^j (G_j - 1) = \left(\sum_{i \in \mathcal{A}_j} Y_{i,j} \right) \left(\frac{\sum_{i \in \mathcal{B}_{j+1}} Y_{i,j+1}}{\sum_{i \in \mathcal{B}_{j+1}} Y_{i,j}} - 1 \right). \end{aligned}$$

On va montrer qu'en général, G_j ne converge pas vers $\mathbb{E} \left[\frac{Y_{i,j+1}}{Y_{i,j}} \right]$. Néanmoins, on obtient quand même la convergence de CL_j rapporté au nombre de sinistre (moyenne de l'estimateur) vers la réserve théorique rapporté au nombre de sinistres (moyenne de la réserve théorique).

Plus précisément, on va montrer le théorème suivant :

Théorème

$$\boxed{\frac{CL_j - CF_j^1}{\text{Card}(\mathcal{A}_j)} \rightarrow 0 \text{ lorsque } \text{Card}(\mathcal{A}_j) \rightarrow \infty.}$$

Remarques

On a :

$$\frac{CF_j^1 - \mathbb{E}CF_j^1}{\text{Card}(\mathcal{A}_j)} \longrightarrow 0 \text{ lorsque } \text{Card}(\mathcal{A}_j) \longrightarrow \infty.$$

Les théorèmes (4.2.1) et (4.3.1) donnent bien une équivalence asymptotiques entre l'estimateur Chain-Ladder et l'estimateur ligne à ligne des flux de trésorerie à un an.

Pour démontrer le théorème, nous avons besoin du lemme suivant :

Lemme 3

Soit $A \subseteq \mathcal{S}$ un ensemble (non nul) de sinistres, on pose :

- $s_A = \sum_{i \in A} \mathbb{E}[C_{i,0}]$;
- $fc_j = \prod_{k=0}^{j-1} \mathbb{E}[F_k]$;
- $h_j = p_j fc_j + \sum_{k=0}^{j-1} (p_k - p_{k+1}) fc_k$.

Alors, pour tout j on a :

$$\left(\frac{\sum_{i \in A} C_{i,j}}{\text{Card}(A)} - \frac{s_A fc_j}{\text{Card}(A)} \right) \xrightarrow{\text{Card}(A) \rightarrow +\infty} 0,$$

et :

$$\left(\frac{\sum_{i \in A} Y_{i,j}}{\text{Card}(A)} - \frac{s_A h_j}{\text{Card}(A)} \right) \xrightarrow{\text{Card}(A) \rightarrow +\infty} 0.$$

Démonstration du lemme 3

En appliquant la loi de grandes nombres de Markov, pour la première limite, il suffit de montrer que :

$$\mathbb{E} \left[\sum_{i \in A} C_{i,j} \right] = s_A fc_j$$

On a en effet : $\mathbb{E}[C_{i,j}] = \mathbb{E}[C_{i,0} \prod_{k=0}^{j-1} F_k] = \mathbb{E}[C_{i,0}] \prod_{k=0}^{j-1} \mathbb{E}[F_k] = \mathbb{E}[C_{i,0}] fc_j$.

On écrit :

$$Y_{i,j} = C_{i,j} \mathbb{1}\{D_{i,j} = 1\} + \sum_{k=0}^{j-1} C_{i,k} \mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\}.$$

En utilisant la même démarche, on peut montrer que (en notant que que les variables D et C sont indépendantes et que $\mathbb{E}[D_{i,j}] = p_j$) :

$$\left(\frac{\sum_{i \in A} C_{i,j} D_{i,j}}{\text{Card}(A)} - \frac{s_A fc_j p_j}{\text{Card}(A)} \right) \xrightarrow{\text{Card}(A) \rightarrow \infty} 0.$$

On a :

$$\left(\frac{\sum_{i \in A} C_{i,k} \mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\}}{\text{Card}(A)} - \frac{s_A f c_j (p_k - p_{k+1})}{\text{Card}(A)} \right) \rightarrow 0 \text{ quand } \text{Card}(A) \rightarrow \infty,$$

car $\mathbb{E}[\mathbb{1}\{D_{i,k+1} = 0, D_{i,k} = 1\}] = p_k - p_{k+1}$.

Finalement :

$$\left(\frac{\sum_{i \in A} Y_{i,j}}{\text{Card}(A)} - \frac{s_A h_j}{\text{Card}(A)} \right) \rightarrow 0 \text{ quand } \text{Card}(A) \rightarrow \infty.$$

Démonstration du lemme terminée

Démonstration du théorème

En appliquant le lemme 3 à G_j , on a :

$$G_j = \frac{\sum_{i \in \mathcal{B}_{j+1}} Y_{i,j+1}}{\sum_{i \in \mathcal{B}_{j+1}} Y_{i,j}} \rightarrow \frac{h_{j+1}}{h_j} \text{ quand } \text{Card}(\mathcal{B}_{j+1}) \rightarrow \infty.$$

Or :

$$\frac{\sum_{i \in \mathcal{A}_j} Y_{i,j} - s_{\mathcal{A}_j} h_j}{\text{Card}(\mathcal{A}_j)} \rightarrow 0 \text{ quand } \text{Card}(\mathcal{A}_j) \rightarrow \infty.$$

On obtient donc :

$$\frac{(\sum_{i \in \mathcal{A}_j} Y_{i,j}) G_j - s_{\mathcal{A}_j} h_{j+1}}{\text{Card}(\mathcal{A}_j)} \rightarrow 0 \text{ quand } \min(\text{Card}(\mathcal{A}_j), \text{Card}(\mathcal{B}_{j+1})) \rightarrow \infty.$$

Notons que :

$$\frac{\sum_{i \in \mathcal{A}_j} Y_{i,j+1} - s_{\mathcal{A}_j} h_{j+1}}{\text{Card}(\mathcal{A}_j)} \rightarrow 0 \text{ quand } \text{Card}(\mathcal{A}_j) \rightarrow \infty,$$

$$\frac{(\sum_{i \in \mathcal{A}_j} Y_{i,j}) G_j - \sum_{i \in \mathcal{A}_j} Y_{i,j+1}}{\text{Card}(\mathcal{A}_j)} \rightarrow 0 \text{ quand } \min(\text{Card}(\mathcal{A}_j), \text{Card}(\mathcal{B}_{j+1})) \rightarrow \infty.$$

Finalement :

$$\frac{(\sum_{i \in \mathcal{A}_j} Y_{i,j})(G_j - 1) - \sum_{i \in \mathcal{A}_j} (Y_{i,j+1} - Y_{i,j})}{\text{Card}(\mathcal{A}_j)} \rightarrow 0 \text{ quand } \min(\text{Card}(\mathcal{A}_j), \text{Card}(\mathcal{B}_{j+1})) \rightarrow \infty.$$

En d'autres termes :

$$\frac{CL_j - CF_j^1}{\text{Card}(\mathcal{A}_j)} \rightarrow 0 \text{ quand } \min(\text{Card}(\mathcal{A}_j), \text{Card}(\mathcal{B}_{j+1})) \rightarrow \infty.$$

Démonstration du théorème terminée

4.3.2 Réserve ultime

Il faut comparer $CL_j^u = L_{t-j}^j(\prod_{k=j}^{\infty} G_k - 1)$ et $\widehat{\mathbb{E}R}_j^u$. Notons que pour tout $k > j_{max} = t - \min_{i' \in \mathcal{S}} n_{i'}$, nous avons $G_{i,k} = 1$ (hypothèse 3.3.2), le produit des G_k est fini.

Nous allons montrer que, comme dans le cas des flux de paiement à un an, ces deux quantités sont asymptotiquement égales. Plus précisément, nous avons :

Théorème

$$\frac{CL_j^u - R_j^u}{Card(\mathcal{A}_j)} \longrightarrow 0 \text{ quand } \min(Card(\mathcal{A}_j), Card(\mathcal{B}_{j+1})) \longrightarrow \infty$$

Démonstration

Pour simplifier les formulations, nous faisons ici l'hypothèse d'équi-distribution sur les premiers paiements, ce qui implique que tous les processus de paiements sont i.i.d. Notons qu'avec la même démarche que dans le cas des flux de paiement à un an, nous pouvons démontrer le résultat sans cette hypothèse.

D'après la loi de grands nombres :

$$G_k = \frac{\sum_{i \in \mathcal{B}_{k+1}} Y_{i,k+1}}{\sum_{i \in \mathcal{B}_{k+1}} Y_{i,k}} \longrightarrow \frac{\mathbb{E}[Y_{i,k+1}]}{\mathbb{E}[Y_{i,k}]} \text{ quand } Card(\mathcal{B}_{k+1}) \longrightarrow \infty,$$

et :

$$\frac{L_{t-j}^j}{Card(\mathcal{A}_j)} = \frac{\sum_{i \in \mathcal{S}, n_i = t_j} Y_{i,j}}{Card(\mathcal{A}_j)} \longrightarrow \mathbb{E}[Y_{i,j}] \text{ quand } Card(\mathcal{A}_j) \longrightarrow \infty.$$

Nous obtenons donc :

$$\frac{CL_j^u}{Card(\mathcal{A}_j)} \longrightarrow \mathbb{E}[Y_{i,j}] \left(\prod_{k=j}^{\infty} \frac{\mathbb{E}[Y_{i,k+1}]}{\mathbb{E}[Y_{i,k}]} - 1 \right) = \mathbb{E}[Y_{i,\infty}] - \mathbb{E}[Y_{i,j}]$$

La méthode Chain-Ladder mesure bien ce que nous voulons. De façon similaire :

$$\widehat{\mathbb{E}F}_l = \frac{\sum_{i \in \mathcal{B}_{l+1}^*} C_{i,l+1}}{\sum_{i \in \mathcal{B}_{l+1}^*} C_i^l} \longrightarrow \mathbb{E} \left[\frac{C_{i,l+1}}{C_{i,j}} \right] \text{ quand } Card(\mathcal{B}_{l+1}^*) \longrightarrow \infty,$$

et :

$$\frac{\sum_{i \in \mathcal{A}_j(t)} D_{i,j+1} C_{i,j}}{Card(\mathcal{A}_j)} \longrightarrow \mathbb{E}[C_{i,j}] \mathbb{E}[D_{i,j+1}] \text{ quand } Card(\mathcal{A}_j) \longrightarrow \infty$$

Sous l'hypothèse que les estimateur \hat{p}_k sont consistants : $\hat{p}_k \longrightarrow \mathbb{E}[D_{i,k}]$, nous avons :

$$\frac{\widehat{\mathbb{E}R}_j^u}{Card(\mathcal{A}_j)} = \frac{\sum_{i \in \mathcal{A}_j(t)} D_{i,j+1} C_{i,j}}{Card(\mathcal{A}_j)} \sum_{k=j+1}^{\infty} \left(\prod_{l=j}^{k-1} \widehat{\mathbb{E}F}_l(t) - 1 \right) \frac{\hat{p}_k - \hat{p}_{k+1}}{\hat{p}_{j+1}}$$

tend quand $\text{Card}(\mathcal{A}_j) \rightarrow \infty$ vers :

$$\begin{aligned} \sum_{k=j+1}^{\infty} \mathbb{E}[C_{i,j}] p_{j+1} \left(\prod_{l=j}^{k-1} \mathbb{E}[F_i^l] - 1 \right) \frac{\hat{p}_k - \hat{p}_{k+1}}{\hat{p}_{j+1}} &\rightarrow \sum_{k=j+1}^{\infty} (\mathbb{E}[C_{i,k}] - \mathbb{E}[C_{i,j}]) (p_k - p_{k+1}) \\ &= \sum_{k=j+1}^{\infty} (\mathbb{E}[C_{i,k}] - \mathbb{E}[C_{i,j}]) \mathbb{P}(D_k = 1, D_{k+1} = 0) \\ &= \mathbb{E}[Y_{i,\infty}] - \mathbb{E}[Y_{i,j}] \end{aligned}$$

La démonstration est terminée.

Nous avons obtenu des **estimateurs ponctuels** de la moyenne de la réserve à un an et à l'ultime, et nous les avons comparé à l'estimateur ponctuel Chain-Ladder. Les estimateurs ponctuels sont asymptotiquement des **Best Estimate** des réserves. Cependant, nous nous intéressons maintenant à une **distribution prédictive** des réserves, pour quantifier le risque de déviation des provisions par rapport à la moyenne. Par exemple, cette distribution prédictive peut servir à calculer une VaR ou une TVaR, dans le cadre de la marge de risque IFRS 4 Phase II. Nous proposons des méthodes pour estimer la distribution des flux de paiements à un an et à l'ultime.

4.4 Distribution des réserves

Pour estimer la distribution des réserves, nous pourrions par exemple paramétriser la loi des paiements, ou celle des facteurs de développement individuels. Cependant, utiliser une procédure Bootstrap permet d'estimer la distribution de la réserve, sans avoir à faire d'hypothèse supplémentaire. C'est la démarche utilisée par Murphy et McLennan (2006) [35], Hubert et Guay (2009) [25], et Rosleund (2012) [45]. Nous avons envisagé plusieurs procédures :

- un bootstrap « total », qui consiste à estimer les flux futurs des sinistres non-clos à l'aide d'un ré-échantillonnage « pur » ;
- un bootstrap utilisant l'estimation de la durée de vie ;
- un bootstrap utilisant une estimation non paramétrique de la distribution des facteurs de développement.

Nous donnons ci-dessous les étapes pour la mise en œuvre en pratique dans ces différentes procédures.

4.4.1 Bootstrap « total »

Cette procédure est celle proposée par Murphy et McLennan [35] et utilisée par Hubert et Guay [25]. En date t , on considère l'ensemble des sinistres de \mathcal{S} à provisionner, c'est-à-dire $\mathcal{S}^P = \{S_i \in \mathcal{S} | D_{i,t+1} = 1\}$. Considérons un sinistre S_k de l'ensemble \mathcal{S}^P , en année de développement $j = t - n_k$.

Une estimation Bootstrap de la réserve ultime est construite ainsi :

- On tire un sinistre $S_{i_1} \in \mathcal{B}_{j+1}(t)$.
- Si le sinistre i_1 tiré est tel que $D_{i_1,j+2} = 0$, on applique le facteur de développement tiré $f_{i_1,j}^b$ à $Y_{k,j}$ et on arrête la procédure.
- Sinon, on tire un sinistre $S_{i_2} \in \mathcal{B}_{j+2}(t)$. Si le sinistre S_{i_2} est tel que $D_{i_2,j+3} = 0$, on applique les facteurs de développement tiré $f_{i_1,j}^b \times f_{i_2,j+1}^b$ à $Y_{k,j}$ et on arrête la procédure.
- Sinon, on continue la procédure jusqu'à son arrêt.

Pour chaque sinistre de l'ensemble \mathcal{S}^P , on tire B pseudo-observations en utilisant cette procédure. On obtient alors B pseudo-échantillons de la charge ultime finale pour les sinistres déclarés mais non clos, qui permettent d'estimer la distribution de la charge finale. N.B. : si $Card(\mathcal{S})$ est suffisamment petit, on peut envisager de tirer tous les chemins possibles pour les sinistres de l'ensemble \mathcal{S}^P .

4.4.2 Bootstrap en utilisant la durée de vie estimée

Dans la procédure précédente, il y a à la fois un tirage des facteurs de développement et implicitement de la durée de vie. Comme nous avons fait l'hypothèse dans le cadre proposé que la durée de vie est indépendante des facteurs de développement, nous pouvons séparer les deux tirages.

Notons $\mathcal{E}_j = \left\{ \frac{C_{i,j+1}}{C_{i,j}} \mid i \in \mathcal{B}_j^*(t) \right\}$, l'ensemble des réalisations de facteur de développement F_j (entre les années de développement j et $j+1$), connues à la date de réalisation de l'étude t ($0 \leq j \leq \Gamma^{max}$).

La procédure est la suivante. On considère à nouveau l'ensemble \mathcal{S}^P et on réalise B simulations, indicées $b = 1, \dots, B$, pour chaque sinistre k de cet ensemble.

Pour chaque simulation b :

1. Tirage d'une durée de vie résiduelle Γ_k^b pour le sinistre k à partir de la loi estimée de $\Gamma \mid \Gamma \geq t - n_k + 1$.
2. Tirage aléatoire des facteurs de développement $f_{i,j}^b, f_{i,j+1}^b, \dots, f_{i,\Gamma_k^b}^b$ dans l'échantillon correspondant $\mathcal{E}_j, \mathcal{E}_{j+1}, \dots, \mathcal{E}_{\Gamma_k^b}$.
3. Calcul de la charge ultime $CU_k^b = c_{k,j} \prod_{k=j}^{\Gamma_k^b} f_{i,j}^b$.

4.4.3 Bootstrap et distribution non-paramétrique des facteurs de développement

Il est également possible de chercher à estimer non paramétriquement la distribution des facteurs F_j à partir des réalisations correspondantes \mathcal{E}_j (voir annexe A.3).

Ainsi, dans l'étape 2 de la procédure précédente, nous ne tirerions pas un facteur appartenant aux ensembles \mathcal{E}_j , mais nous tirerions directement (par simulation) une valeur dans la distribution estimée (non paramétriquement) des facteurs F_j . Cependant, les méthodes d'estimation non-paramétrique conduisent à des estimateurs biaisés de la densité des F_j . De plus, cette procédure est plus coûteuse en terme de temps de calcul, ce qui peut conduire à l'écartier.

Bilan

Dans la comparaison entre l'estimation des réserves (à un an ou à l'ultime) par Chain-Ladder (notée CL ci-après) et l'estimation de la moyenne des réserves par la méthode ligne à ligne (notée ML ci-après), nous avons obtenu une convergence des moyennes. Il est donc intéressant d'étudier la vitesse de convergence entre l'estimation des réserves totales par Chain-Ladder et par la méthode ligne à ligne, en fonction du nombre de données disponibles. Pour cela, nous proposons de mener une étude simulatoire.

Cette étude sera aussi l'occasion de comparer l'estimation de la distribution des réserves par bootstrap et la distribution théorique retenue.

Partie 5

Exercice de simulations

Dans cette partie, nous allons créer des bases de données simulées, afin d'étudier, à distance finie, la convergence entre l'estimateur Chain-Ladder (CL) et l'estimateur ligne à ligne (ML) de la moyenne des réserves. Nous allons également tester les différentes procédures bootstrap proposées dans notre approche ligne à ligne, et comparer la moyenne de la distribution bootstrap avec ML et CL, et la distribution théorique. Nous avons réalisé cette étude à l'aide du logiciel R ¹.

Dans la première partie, nous rappellerons quels sont les différents éléments fondamentaux constituant une base de données sinistres, dans le cadre développé dans ce mémoire. Nous rappellerons également les hypothèses que ces éléments doivent respecter pour que ML soit asymptotiquement équivalent à CL. Ensuite, nous générons des bases de données selon plusieurs scénarii. Nous discutons de la distance entre CL et ML selon les cas de figures retenus, et comparons la distribution prédictive des réserves estimées par bootstrap, et la distribution théorique estimée par simulations Monte-Carlo.

5.1 Générateur de bases de données

Dans le modèle développé dans ce mémoire, un sinistre i peut se modéliser en un vecteur $(n_i, I_{i,j}, \Gamma_i)$, avec n_i l'année de survenance, $I_{i,j}$ le processus de charges, et Γ_i la v.a. représentant la durée de vie du sinistre. Nous ne considérons que les sinistres non tardifs. Nous avons fait les hypothèses suivantes :

Résumé des hypothèses sur les sinistres, la durée de vie et les facteurs de développement

1. Les sinistres sont indépendants. Pour un sinistre donné, le processus de charges est indépendant de la durée de vie ;
2. Les durées de vie Γ_i sont i.i.d. ;
3. Les variables $F_{i,j}$ définies par $I_{i,j+1} = F_{i,j}I_{i,j}$ sont indépendantes du passé (c'est-à-dire de \mathcal{F}_u , avec u la date telle que $j = u - n_i$). Pour une année de développement j donné, les $F_{i,j}$ suivent la même loi pour tout i .

Notons que nous avons retenu la définition d'un sinistre avec le processus de charges et non

1. <http://cran.r-project.org/>

celui des paiements. Ce premier processus est en effet mieux à même de vérifier les conditions techniques sous lesquelles ML et CL sont asymptotiquement équivalents, notamment d'être toujours supérieurs à un seuil donné. Nous rappelons ses conditions ci-dessous.

Conditions techniques

Il existe $\delta > 0$ tel que :

1. $\forall j \geq 0, \mathbb{E}[(F_j)^{1+\delta}] < \infty$;

2. $\sup_{i \in \mathcal{S}} \left\{ \mathbb{E}[(I_{i,0})^{1+\delta}] \right\} < \infty$.

De plus, on suppose que :

3. $\exists c > 0, \forall S_i \in \mathcal{S}, I_{i,0} > c$ \mathbb{P} -p.s. ;

4. $\forall j \geq 0, \exists c' > 0, \forall S_i \in \mathcal{S}, F_{i,j} \geq c'$ \mathbb{P} -p.s. ;

5. Les premières charges $I_{i,0}$ sont équidistribuées (hypothèse supplémentaire faite pour la convergence des réserves à l'ultime).

Nous avons donc besoin des éléments suivants pour simuler une base de sinistres \mathcal{S} permettant de comparer la convergence entre CL et ML :

- Une première date d'observation, qui sera égale à la plus ancienne date de survenance des sinistres simulés : $\min_{i \in \mathcal{S}} n_i$.
- Une date d'observation finale : t . La profondeur de la base (c.-à-d. la durée d'observation) sera donc : $T = t - \min_{i \in \mathcal{S}} n_i$.
- Un nombre de sinistres survenus pour chaque année de survenance $n_i = 0, \dots, t$. Ce nombre peut être fixé à l'avance, ou bien choisi de manière aléatoire (p. ex. en utilisant une loi de Poisson). Il est également possible de modéliser une croissance du portefeuille en termes de nombre d'assurés.
- Une distribution i.i.d pour la durée de vie d'un sinistre (p. ex. une loi de Poisson ou une loi géométrique). L'hypothèse (3.3.2) implique que la durée de vie maximale Γ^{max} doit être égale à T .
- Une distribution i.i.d. de la première charge $I_{i,0}$ à valeurs strictement positives (p. ex. une loi log-normale, ou une loi gamma).
- Une distribution i.i.d. des facteurs de développement $(F_{i,j})_{0 \leq j \leq \Gamma_i - n_i}$. Elle peut être différente par année de développement j . Par souci de réalisme, nous proposons d'utiliser une loi composée d'une masse de Dirac en 1 et d'une loi continue (à valeurs strictement positives), pour refléter le cas très fréquent où il n'y a pas d'évolution des charges d'une année à l'autre.

Nous proposons maintenant un premier paramétrage, qui constituera le scénario de référence.

5.2 Générateur de scénario central

5.2.1 Paramétrage retenu

Nous avons retenu le paramétrage suivant pour le générateur de scénario central :

- La durée d’observation est de $T = 7$ ans.
- Le nombre de sinistres par année de survenance est de $N_s = 3000$ (soit 21000 sinistres au total).
- La durée de vie d’un sinistre suit une loi de Poisson de paramètre 3 tronquée à 7 (puis translatée de 1). La durée de vie moyenne est de 4 ans.
- Les premières charges suivent des distributions log-normales de paramètre $(0, 1.2)$: $I_{i,0} \sim \mathcal{LN}(0, 1.2)$, et tronquées à 0.01. La médiane du premier paiement est 1 (représentant 1000€ p.ex.), la moyenne 2.05, le quantile à 90% est 4.65 et celui à 99.5% est 22. Il y a donc très peu de chance d’obtenir un sinistre exceptionnel.
- Les facteur de développement suivent (pour toutes les années de développement) une loi composée à partir :
 - Une loi de Bernoulli de paramètre 0.5, qui représente la probabilité qu’il y ait une révision de la charge : $B \sim \mathcal{Ber}(0.5)$.
 - Une loi log-normale X qui représente la variation de charge en cas d’évolution : $X \sim (\mathcal{LN}(0.1, 0.25))$, tronquée à 0.01. On autorise donc des facteurs de développement inférieurs à 1 (c.-à-d. des recours). En moyenne, il y aura une révision de 14% de la charge à la hausse.
 - La loi des facteurs de développement est alors : $F = B \times X + (1 - B) \times 1$.

Maintenant que nous avons défini le processus générateur de données (PGD) central, nous allons étudier une réalisation particulière de ce processus.

5.2.2 Étude d'une réalisation du PGD central

Le tableau 5.1 donne un exemple de triangle de charges agrégées obtenu avec une réalisation du PGD centrale. Les charges augmentent de moins en moins avec les années de développement, en raison des sinistres qui sont clos.

i, j	0	1	2	3	4	5	6
0	6205	6567	6880	7189	7337	7450	7493
1	6262	6617	6971	7275	7432	7499	
2	6351	6799	7231	7607	7772		
3	5880	6248	6593	6848			
4	6220	6629	7032				
5	6107	6481					
6	6542						

TABLE 5.1 – Exemple de triangle de charges dans le scénario de référence.

5.2.2.1 Comparaison entre ML et CL sur la réali

Les facteurs de développement et les réserves calculées avec la méthode Chain-Ladder sont les suivants (Réserve 1 an est ici la variation de charge à un an) :

f_0	f_1	f_2	f_3	f_4	f_5	Réserve 1 an	Réserve ultime
1.062	1.056	1.045	1.0213	1.012	1.006	1 373	3 454

Avec le modèle ligne à ligne, on obtient des estimations des réserves très proches : il y a 2,5% d'écart relatif sur la réserve à un an et sur la réserve ultime. On peut s'intéresser à l'estimation de la moyenne des facteurs individuels $F_{i,j}$, définie par :

$$\tilde{f}_j = \frac{\sum_{i \in \mathcal{B}_{j+1}^*} I_{i,j+1}}{\sum_{i \in \mathcal{B}_{j+1}^*} I_{i,j}},$$

et les comparer aux facteurs Chain-Ladder. Les facteurs \tilde{f}_j sont plus élevés que les facteurs f_j , puisque les premiers seront à appliquer aux sinistres ouverts en $j + 1$, tandis que les seconds prennent en compte tous les sinistres. Les résultats sont présentés dans le tableau suivant :

\hat{f}_0	\hat{f}_1	\hat{f}_2	\hat{f}_3	\hat{f}_4	\hat{f}_5	Réserve 1 an	Réserve ultime
1.066	1.070	1.076	1.059	1.062	1.0490	1 338	3 371

La moyenne théorique des facteurs individuels $F_{i,j}$ est de 1.07. Les estimateurs \tilde{f} sont proches. \tilde{f}_5 donne une estimation plus éloignée de la moyenne théorique, mais il y a moins de sinistres toujours ouverts au bout en $j = 6$ pour estimer la moyenne. Nous avons aussi calculé la moyenne empirique \hat{f} des $F_{i,j}$, par année de développement j , définie par :

$$\hat{f}_j = \frac{1}{\text{Card}(\mathcal{B}_{j+1}^*)} \sum_{i \in \mathcal{B}_{j+1}^*} \frac{I_{i,j+1}}{I_{i,j}}.$$

La moyenne empirique \hat{f}_j est proche de la moyenne empirique pondérée \tilde{f}_j et la moyenne théorique du PGD central de 1.07 :

\tilde{f}_0	\tilde{f}_1	\tilde{f}_2	\tilde{f}_3	\tilde{f}_4	\tilde{f}_5
1.069	1.072	1.071	1.073	1.066	1.062

Bilan

Sur cette réalisation particulière du PGD central, avec 21 000 sinistres simulés, nous obtenons bien une proximité à distance finie entre l'estimateur CL et l'estimateur ML. Nous allons maintenant nous intéresser à l'estimation de la distribution prédictive des réserves par bootstrap. Elle pourra être comparée à la distribution théorique des réserves définie par le PGD central et sachant la réalisation observée du triangle supérieur (Table 5.1).

5.2.2.2 Comparaison entre la distribution prédictive par bootstrap et la distribution théorique

Nous procédons maintenant à une estimation de la distribution des réerves par bootstrap, à partir de la réalisation du PGD central. Nous avons implémenté dans cette partie :

- le bootstrap qui utilise la durée de vie résiduelle estimée (BST) (section 4.4.2),
- le bootstrap utilisant l'estimation non-paramétrique des facteurs de développement (section 4.4.3) (BST NP).

Nombre de tirages

Nous avons réalisé 10 000 tirages par sinistre. Le temps de calcul de la version non-paramétrique est sensiblement plus long. Nous effectuons également 10 000 tirages pour estimer la distribution théorique des réserves par Monte-Carlo (en utilisant le paramétrage retenu dans le PGD).

Résultat : réserve un an

Les résultats pour la réserve un an sont les suivants :

	Distribution	Moyenne	Quantile 90%	Quantile 95%	Quantile 99.5%
Réserve 1 an	Théorique	1 393	1 520	1 565	1 687
	BST	1 395	1 533	1 578	1 694
	BST NP	1 393	1 525	1 577	1 703

La distribution théorique et les distributions bootstrap (BST) et (BST NP) sont très proches. Cela est également mis en évidence sur la figure 5.1. Un test de Kolmogorov-Smirnov ne rejette pas l'hypothèse nulle d'équi-distribution des différentes lois.

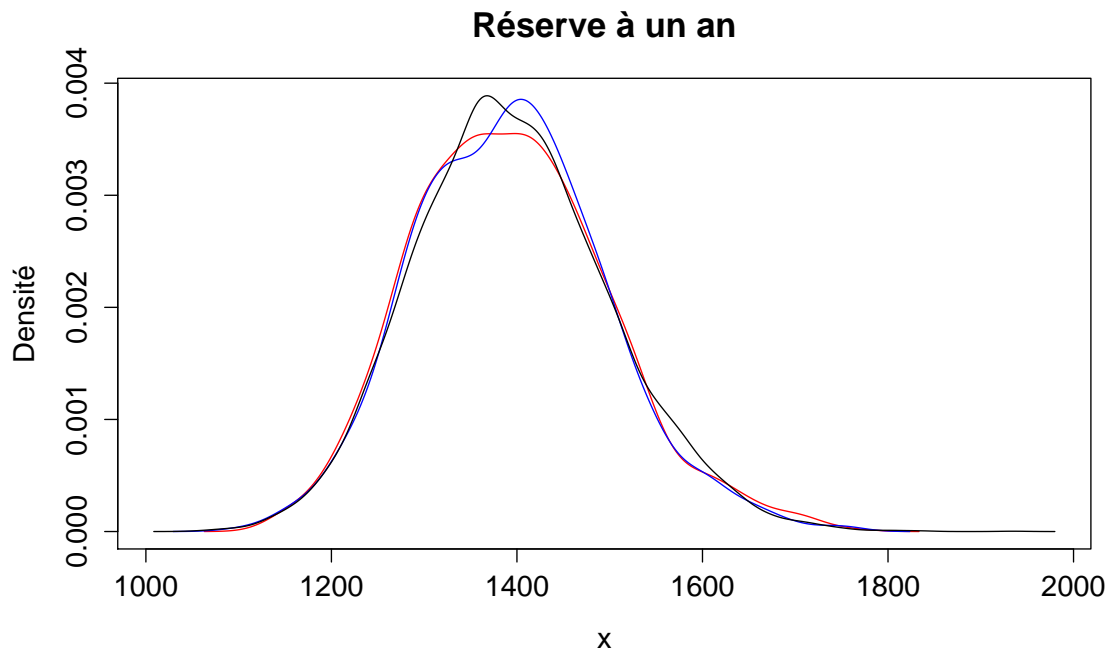


FIGURE 5.1 – En bleu : distributions de la réserve à un an : théorique (bleue), par BST (noir) et par BST NP (rouge)

Résultat : réserve ultime

Pour la réserve ultime, au vu des temps de calcul de la méthode BST NP, nous n'avons retenu que la méthode BST. Les résultats sont les suivants :

	Distribution	Moyenne	Quantile 90%	Quantile 95%	Quantile 99.5%
Réserve ultime	Théorique	3 574	3 834	3 926	4 216
	BST	3 585	3 853	3 945	4 255

Ici aussi, un test de Kolmogorov-Smirnov ne rejette pas l'hypothèse nulle d'équi-distribution entre les deux lois.

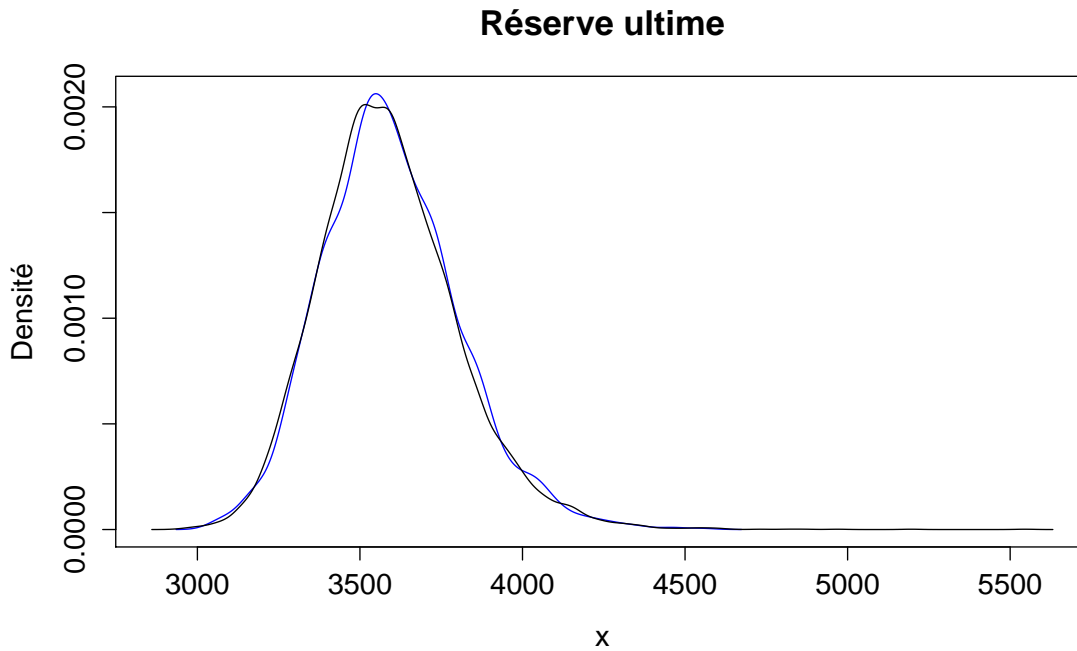


FIGURE 5.2 – En bleu : distributions de la réserve à l’ultime : théorique (bleue) et par BST (noir)

5.2.3 Sensibilité au nombre de sinistres simulés

Nous faisons maintenant varier le nombre de sinistres N_s simulés par année de survenance (N.B. : 100 sinistres par année de survenance correspond à une base totale de 700 sinistres). Nous comparons ensuite l’écart relatif fait entre CL et ML. Pour chaque N_s fixé, nous réalisons 3 000 simulations du scénario (afin de limiter les temps de calcul). Nous calculons ensuite l’écart relatif moyen sur les 3 000 simulations. Les résultats sont donnés dans la table 5.2.

Nbre total	N_s	$\frac{ CL^1 - ML^1 }{CL^1}$ (%)	$\frac{ CL^u - ML^u }{CL^u}$ (%)
700	100	7.85	10.91
3 500	500	3.13	4.22
7 000	1 000	2.21	3.19
35 000	5 000	1.03	1.46
70 000	10 000	0.70	0.99

TABLE 5.2 – Ecart relatif moyen entre CL et ML, à un an et à l’ultime en fonction du nombre de sinistres (3 000 simulations)

Ils font apparaître une vitesse de convergence en \sqrt{n} : lorsque le nombre de sinistres est multiplié par 100, l’écart relatif est divisé par 10.

Dans cette sous-section, nous avons montré que les estimateurs ML et CL étaient proches, avec le PGD central défini. L'estimation de la distribution prédictive des réserves par bootstrap était également proche de la distribution théorique. Nous allons maintenant créer de nouveaux PGD faisant à chaque fois varier un paramètre du PGD central, afin de voir comment la distance entre les estimateurs ML et CL évolue.

5.3 Augmentation de la volatilité de la première charge

On suppose ici que les premiers paiements suivent des distributions log-normales : $\exp(\mathcal{N}(0, 2))$. Cela autorise la génération de sinistres graves, comme le montre le tableau suivant donnant des quantiles de la loi log-normale de paramètres $(0, 2)$:

α	α -quantile
50%	1
75%	3,85
90%	12,97
95%	26,83
99%	104,9
99,5%	172,7

Hormis cette modification de la loi des premiers paiements, les autres paramétrages du PGD central sont retenus. Les résultats en fonction du nombre de sinistres par an simulés sont les suivants :

Nbre total	N_s	$\frac{ CL^l - ML^l }{CL^l}$ (%)	$\frac{ CL^u - ML^u }{CL^u}$ (%)
700	100	20.83	31.67
3 500	500	9.51	12.16
7 000	1 000	6.82	9.33
35 000	5 000	3.10	4.20
70 000	10 000	2.05	2.98

Nous constatons qu'il faut un nombre plus considérable de sinistres dans la base pour obtenir un écart relatif moyen acceptable entre ML et CL. Nous allons maintenant définir un nouveau PGD, qui cette fois-ci modifie la volatilité des facteurs de développement par rapport au PGD central.

5.4 Augmentation de la volatilité des facteurs de développement

Nous testons maintenant une augmentation de la volatilité des facteurs de développements. La composante X dans la loi de F suit maintenant $X \sim (\mathcal{LN}(0.1, 0.5))$.

Nbre total	N_s	$\frac{ CL^l - ML^l }{CL^l}$ (%)	$\frac{ CL^u - ML^u }{CL^u}$ (%)
700	100	10.60	14.48
3 500	500	4.25	5.73
7 000	1 000	3.14	4.14
35 000	5 000	1.31	1.81
70 000	10 000	0.93	1.30

Nous constatons aussi que les écarts relatifs moyens sont plus élevés que dans le scénario de référence, mais restent de même ordre.

5.5 Branche longue

Par rapport au paramétrage du PGD central, nous considérons maintenant une durée de vie $T = 14$. Le paramètre la loi de Poisson pour la durée de vie est changé à 10. Les résultats sont les suivants :

Nbre total	N_s	$\frac{ CL^l - ML^l }{CL^l}$ (%)	$\frac{ CL^u - ML^u }{CL^u}$ (%)
700	50	5.60	7.22
1400	100	4.01	5.35
3 500	250	2.43	3.48
7 000	500	1.63	2.34
14 000	1 000	1.17	1.74
35 000	2 500	0.74	1.14
70 000	5 000	0.53	0.79
140 000	10 000	0.37	0.57

Nous constatons que changer seulement la durée d'observation (et la durée de vie des sinistres en conséquence) par rapport au scénario de référence ne change pas de manière significative les écarts relatifs pour un nombre de sinistres par année de survenance N_s donné.

5.6 Levée de l'hypothèse d'équi-distribution du premier paiement

Dans cette partie, nous levons l'hypothèse d'équi-distribution du premier paiement. Nous voulons tester si c'est cette hypothèse, nous observons une convergence entre ML à l'ultime et CL à l'ultime. Pour cela, par rapport au scénario de référence, nous ajoutons seulement un facteur d'inflation à la première charge en fonction de l'année de survenance : $I_{i,0} \sim \mathcal{LN}(0, 1.2) \times (1 + 0.05)^{n_i}$, où n_i sont les années de survenance de 0 à $T - 1$. Les résultats sont les suivants :

Nbre total	N_s	$\frac{ CL^l - ML^l }{CL^l}$ (%)	$\frac{ CL^u - ML^u }{CL^u}$ (%)
700	100	19.43	27.45
3500	500	8.73	11.37
7 000	1 000	6.29	8.21
35 000	5 000	2.90	4.09
70 000	10 000	2.21	3.16

L'écart relatif moyen entre CL et ML diminue bien avec le nombre de sinistre simulés. Cela est aussi vrai pour la réserve ultime, ce qui semble indiquer que l'équivalence asymptotique entre CL et ML à l'ultime pourrait s'obtenir sans hypothèse d'équi-distributon.

Bilan

Avec l'étude de simulation, nous avons montré que les estimateurs CL et ML sont proches à distance finie. Cependant, lorsque l'on augmente la volatilité des processus générateurs de données, il est nécessaire d'avoir un nombre de sinistres importants dans la base pour obtenir un écart relatif faible.

De nombreuses études complémentaires pourraient être menées :

- étude d'autres paramétrages ;
- étude de l'écart entre CL et ML au niveau de chaque cellule du triangle agrégé ;
- lever de certaines hypothèses du modèle pour discerner s'il y a convergence entre CL et ML sous des hypothèses moins fortes.

Nous allons passer maintenant à l'application du modèle sur une base de données réelle.

Partie 6

Applications à des données réelles

Nous utilisons la base de données automobiles particuliers d'une compagnie d'assurances. Les sinistres étudiés sont des dommages corporels (*Bodily Injury*) et des dommages matériels (*Own Damage*).

6.1 Description de la base

La base comprend pour chaque sinistre :

- Un numéro de sinistre,
- La date de survenance de l'événement générant le sinistre,
- La date de notification du sinistre à l'assureur,
- Les paiements cumulés ainsi que les dates de vision (sur une base trimestrielle). Ils sont donnés bruts et nets de réassurance.
- les provisions dossier-dossier constitués, ainsi que les dates de vision (sur une base trimestrielle).
- La date de clôture du sinistre le cas échéant.

Certaines informations ne sont pas disponibles :

- La déclinaison des montants par nature : principal (indemnisation pure), recours, paiements d'honoraires (ALAE), ... En particulier, les flux de paiements dûs aux recours ne sont pas isolés.
- Les informations sur les caractéristiques du preneur d'assurance et certains facteurs de risques (véhicules, ...). En particulier, le forfait d'ouverture automatique pour les petits sinistres est différencié en fonction du type de risque et est calculé et réévalué par des actuaires en fonction des coûts moyens finaux estimés. Les sinistres dont la charge attendue est au-delà d'un seuil prédéfini doivent être systématiquement évalués au coût réel. Une segmentation par type de risque aurait donc pu être intéressante pour évaluer la pertinence des forfaits d'ouverture au coût moyen.
- Le caractère litigieux ou non du sinistre n'est également pas disponible.

6.2 Profondeur de la base

Les données contiennent des sinistres survenus entre 1984 et le 1^{er} trimestre 2013. Il est à noter qu'il y a eu un changement dans la gestion des sinistres. En 2010, les protocoles de gestion des sinistres corporels ont été revus, et beaucoup de dossiers ont été réévalués.

6.3 Construction de la base d'étude

La base initiale comprend une **ligne par date de vision par sinistre**, avec exclusion des dates de vision auxquelles aucun changement de charge n'a été constaté (c.-à-d. un changement de paiements ou un changement dans la provision dossier-dossier). Nous avons procédé à une agrégation des données afin de passer à une **ligne par sinistre** résumant sa vie au cours des années de développement :

- une première base contient l'évolution de la charge des sinistres, de sa notification à l'assureur à sa clôture, ainsi que son statut (clos ou ouvert fin 2012).
- une seconde base contient les paiements cumulés.

Lors de l'agrégation des données, des informations ont été perdues : si plusieurs opérations avaient lieu la même année, seule la dernière opération a été retenue. Lorsqu'un sinistre entraînait à la fois un dommage corporel et un dommage matériel, nous avons scindé ce sinistre en deux sinistres différents, un par garantie en jeu.

6.4 Traitements effectués

6.4.1 Durée d'observation

la première date de vision des gestionnaires de sinistres dans la base est le 1^{er} trimestre 2002. Les sinistres survenus avant cette date ont donc été supprimés, car nous ne disposons pas de l'historique des opérations entre la survenance et le 1^{er} trimestre 2002. De plus, nous avons supprimé les informations relatives à l'année 2013 : nous nous intéressons à l'évolution de la charge/des paiements cumulés au cours d'années entières, et nous n'avons qu'un seul trimestre d'observations pour 2013. Or, nous menons l'étude de provisionnement en nous plaçant fin 2012.

6.4.2 Cas de réouvertures

Les cas de réouvertures peuvent biaiser l'estimation de la durée de vie des sinistres, puisqu'il y aura de l'hétérogénéité entre les sinistres non réouverts et les sinistres réouverts, par exemple au niveau de la durée de vie. Nous avons donc cherché à séparer tous les événements de la vie des sinistres qui relevaient de la réouverture. Segmenter la base entre sinistres réouverts et un sinistre non réouverts pouvait constituer une autre solution : cependant, pour projeter un sinistres non clos, il aurait alors fallu estimer la probabilité qu'il soit réouvert.

Les cas de réouvertures ont été traités de la manière suivante : lorsqu'il y avait plusieurs dates de clôture pour un même sinistre, nous avons toujours considéré la première date de clôture comme la date de clôture officielle. Les opérations après cette date officielle sont considérées

comme des cas de réouverture. Nous avons été amené à distinguer :

- Les cas de réouvertures "administratives" : il s'agissait de cas où le sinistre a été réétudié après sa date de clôture officielle par le gestionnaire de sinistre, sans qu'il y ait de nouveaux paiements ou de reconstitution d'une provision dossier-dossier.
- Les cas de réouvertures "réelles", ou de nouveaux montants ont été payés et/ou provisionnés par le gestionnaire de sinistres. Dans ce cas-là, nous avons gardé le paiement final (si la réouverture avait été close) ou la dernière charge estimée (si le sinistre réouvert n'avait pas été clos), afin de pouvoir déterminer le poids de ces réouvertures par rapport à la charge finale.

6.4.3 Cas de valeurs nulles ou négatives

Nous avons aussi observé des sinistres avec des charges finales négatives. Ces charges négatives peuvent être dues à plusieurs cas :

- D'abord, à des cas de recours. En effet, dans le droit du pays de l'entité locale considérée, le droit à la subrogation s'exerce jusqu'à hauteur des indemnités versées à l'assuré, comme en France, et si l'assureur recouvre des sommes supérieures aux indemnités versés à l'assuré, il doit les retourner à l'assuré. Cependant, la composante intérêts des dommages-intérêts, fixée par les parties ou par la justice, est généralement alloué à l'assureur¹. Cela peut donc conduire l'assureur à avoir une charge finale négative sur un sinistre.
- A des accords avec des garages conventionnés. Lorsque les garages conventionnés vendent les épaves des véhicules endommagés suite à un accident, ou les mettent à la casse, une part du profit revient à la compagnie. Les gestionnaires de sinistres ont pris l'habitude d'enregistrer ces rentrées d'argent comme un sinistre dans la base. Les cas d'accords ont été identifiés et supprimés.
- A des cas de fraude à l'assurance (cas similaire à un recours).

Des cas de valeurs nulles ou problématiques sont aussi apparus. Nous avons repéré plusieurs cas :

- Les *nil claims* : on peut distinguer les *nil claims* "purs" (dont la charge est toujours estimée à zéro) et des *nil claims* qui ont connu une provision dossier-dossier dans un premier temps, avant d'être reconnu comme *nil claims* et finir avec une charge finale nulle.
- Les délais de gestion : une provision dossier-dossier n'a pas été tout de suite constituée à la déclaration du sinistre (par manque d'information sur le sinistre, ou par manque de temps au moment de la date de clôture des comptes)². La (ou les premières) charge(s) est

1. Dans la jurisprudence anglo-saxonne : *Brown v. Hiatts*, 82 U.S. (15 Wall.) 177, 185 (1872) : "Interest is the compensation allowed by law, or fixed by the parties, for the use or forbearance of money, or as damages for its detention ..."

2. Cela concerne généralement les sinistres anciens, le forfait d'ouverture automatique devant être maintenant systématiquement utilisé.

alors nulle(s). Ces cas-là peuvent être considérés comme des IBNeR.

- Les cas de recours complet, qui donnent une charge finale nulle. Parfois, le recours est aussi indiquée par un code sur la réserve dossier-dossier (exemple : réserve dossier-dossier fixée à 3 = attente de recours), qui est laissé lorsque le recours a lieu.
- Les problèmes dûs au traitement informatique des données (extraction de la base), qui font apparaître des données de charges en 10^{-13} ou 10^{-14} . Ces cas là ont été ramenés à des charges nulles.

6.5 Démarche adoptée

D'un côté, les cas de charges finales nulles ou négatives ne respectent pas les hypothèses avec laquelle on peut comparer l'estimateur Chain-Ladder au niveau agrégé et l'estimateur de la réserve en ligne à ligne proposé (voir 4.1.3). De l'autre côté, dans l'estimation de la distribution des réserves par Bootstrap, rien n'empêcherait *a priori* de tirer des facteurs de développement nuls pour prendre en compte ces cas de figure dans l'estimation des réserves. Une approche en deux temps a été retenue :

1. **Approche théorique avec la base non traitée.** Dans un premier temps, nous avons cherché à comparer les estimateurs CL et ML. Pour se placer dans le cadre où les hypothèses sont respectées, nous avons exclu de la base les sinistres dont l'une des charges observées est inférieure à un certain seuil (10€). **Seuls les sinistres non tardifs sont étudiés, et les réserves sont calculées pour les non tardifs qui ne sont pas encore clos.** Cette approche est théorique, au sens où en faisant cette exclusion, nous nous attendons à sur-estimer la charge finale : en effet, lorsque nous projetons la charge ultime d'un sinistre, nous ne prenons pas en compte la possibilité de connaître un recours ou de finir sans suite (*nil claims*).
2. **Approche pratique avec la base retraitée.** Dans un second temps, nous avons retraité des valeurs problématiques, négatives et certaines valeurs nulles. Tous les sinistres sont pris en compte. Les estimateurs ML et CL ne peuvent plus être comparées, mais nous pouvons estimer une distribution des sinistres par bootstrap. Les sinistres tardifs ont également été ré-intégrés. Nous verrons que les méthodes proposées donnent des résultats moins satisfaisants. Nous proposons ensuite des pistes d'amélioration.

Nous allons maintenant présenter les résultats sur la base non retraitée. Nous montrerons que les hypothèses du modèle sont mieux respectées si nous segmentons entre sinistres attritionnels et sinistres graves. Deux méthodes de segmentation seront proposées.

6.6 Statistiques descriptives

Le tableau suivant donne des information sur l'ensemble des sinistres qui étaient présents dans la base en fonction de leur année de survenance.

Année Surv.	Non-tardifs	(>10)	NA	Négative	Zéro	Tardifs	Non-clos >10	%
2002	4624	4242	182	0	200	246	3306	0.71
2003	4538	4113	113	6	306	396	3098	0.68
2004	3354	2998	113	0	243	78	2281	0.68
2005	3491	2999	163	8	321	94	2547	0.73
2006	4389	3744	307	2	336	110	3147	0.72
2007	5546	4799	467	4	276	242	4423	0.80
2008	6111	5145	678	6	282	259	4755	0.78
2009	6796	5745	780	4	267	325	5437	0.80
2010	6253	5602	461	3	187	248	5418	0.87
2011	6023	5352	505	8	158	250	5243	0.87
2012	5568	5125	271	8	164	196	5125	0.92
Total	56 693	49 864	4 040	49	2740	2444	44780	0,79

TABLE 6.1 – Nombre de sinistres dans la base

Pour les sinistres non tardifs :

- **> 10** correspond au nombre de sinistres dont la première charge est supérieure à 10 €.
- **0** correspond au nombre de sinistres dont la première charge est comprise entre 0 et 10.
- **Positive (resp. Négative)** correspond au nombre de sinistres dont la première charge est positive (resp. négative).
- **NA** correspond au nombre de sinistres dont la première charge n'est pas renseigné.
- **Étudiés** correspond au nombre de sinistres après application du seuil de 10€ sur toutes les charges observées. Le pourcentage correspond à la proportion de sinistres étudiés par rapport au nombre de sinistres non-tardifs.

Nous excluons donc environ 20% des sinistres non-tardifs de la base. En terme de charges finales ou dernières charges observées, les sinistres exclus représentent les proportions suivantes :

Charge ultime	OD	BI	Total
Sinistres clos observés	145 284 542	85 035 731	230 320 274
Sinistres clos observés exclus	229 789	133 458	363 247
% exclus	0,16%	0,2%	0,2%
Dernière charge observée	OD	BI	Total
Sinistres observés	182 574 872	160 287 262	342 862 134
Sinistres observés exclus	209 947	150 390	360 337
% exclus	0,11%	0,09%	0,11%

Nous voyons donc qu'ils représentent une toute petite part de la charge finale. C'est attendu car il s'agit essentiellement de sans-suites et de recours. Cependant, nous nous attendons à ce que cette exclusion conduise à une sur-estimation de la charge finale : nous ne modélisons pas la probabilité pour un sinistre non-clos de connaître un recours complet ou de finir sans suite.

Selon la réglementation, les branches dommages corporels et dommages tardifs doivent être traitée séparément. Nous allons nous concentrer sur la branche des **dommages matériels**, qui

6.6. STATISTIQUES DESCRIPTIVES

représentent 80% des sinistres et où les tardifs sont beaucoup moins nombreux. Les données étudiées sont les suivantes :

Année Surv.	Non-tardifs	(>10)	NA	Négative	Zéro	Tardifs	Non-clos >10	%
2002	4337	4001	144	0	192	227	3125	0.72
2003	4163	3816	59	6	282	365	2925	0.70
2004	2964	2723	22	0	219	49	2107	0.71
2005	3010	2672	15	8	315	52	2267	0.75
2006	3811	3398	79	2	332	64	2847	0.75
2007	4585	4196	127	3	259	133	3899	0.85
2008	4861	4424	173	6	258	117	4130	0.85
2009	4975	4560	169	4	242	121	4366	0.88
2010	4110	3829	125	3	153	65	3700	0.90
2011	3811	3519	167	5	120	74	3433	0.90
2012	3552	3334	92	6	120	65	3334	0.94
Total	44 179	40 472	1 172	43	2 492	1 332	36 133	0.82

6.6.1 Charges

La figure 6.1 présente la distribution de la première charge pour les sinistres étudiés. On constate le fort pic aux alentours de 2800, qui correspond à un forfait d'ouverture automatique très utilisé par les gestionnaires. On constate également pour la première année de développement, des pics à 10000, 11000, etc, qui correspondent à une première évaluation manuelle faite par le gestionnaire de sinistres pour les sinistres supérieurs à 10k€.

Sur la figure suivante 6.2, nous pouvons voir également que le forfait d'ouverture de 2802 est réévalué assez rapidement avec les années de développement.

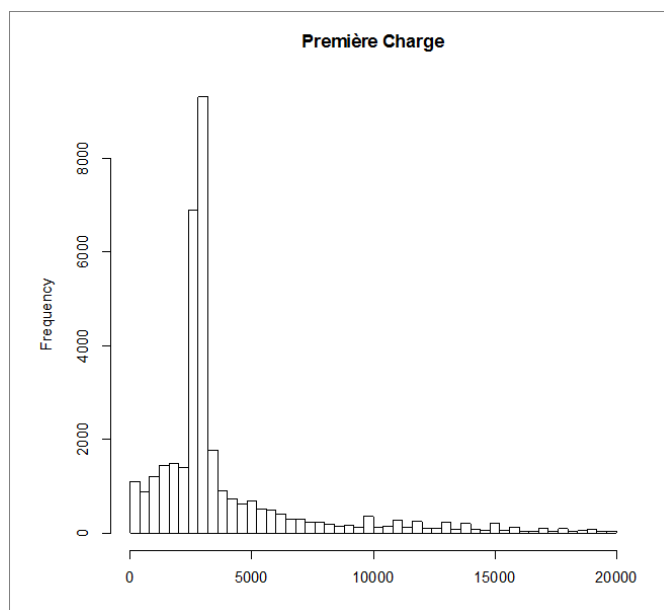


FIGURE 6.1 – Distribution de charges de la première année de développement (tronquée à 20000 - Il reste encore environ 4% de sinistres dépassant ce seuil).

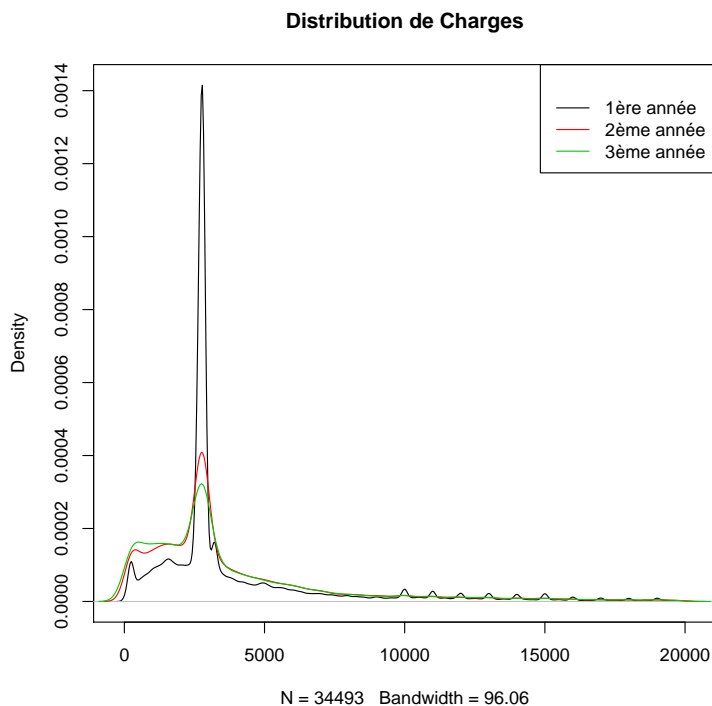


FIGURE 6.2 – Distributions de charges dans les trois premières années de développement (tronquée à 20000 - Il reste encore environ 4% de sinistres dépassant ce seuil).

6.6.2 Durée de vie

L'estimation de la durée de vie d'un sinistre est donnée par la figure 6.3, nous remarquons que plus de 95% de sinistres durent moins de 5 ans. La proportion de sinistres censurée ayant plus de 10 ans d'âge est minimale. Nous aurons donc un biais de l'estimateur de Kaplan-Meier à distance finie, qui ne devrait pas être trop important. Par simplification, nous supposons que ces sinistres sont clos au bout de 11 ans. En étudiant attentivement les données, nous remarquons qu'il s'agit majoritairement de sinistres dont la charge évaluée n'évolue pas depuis des années. Un traitement des queues, par analogie aux méthodes agrégées, pourrait être envisagé et intégré au modèle.

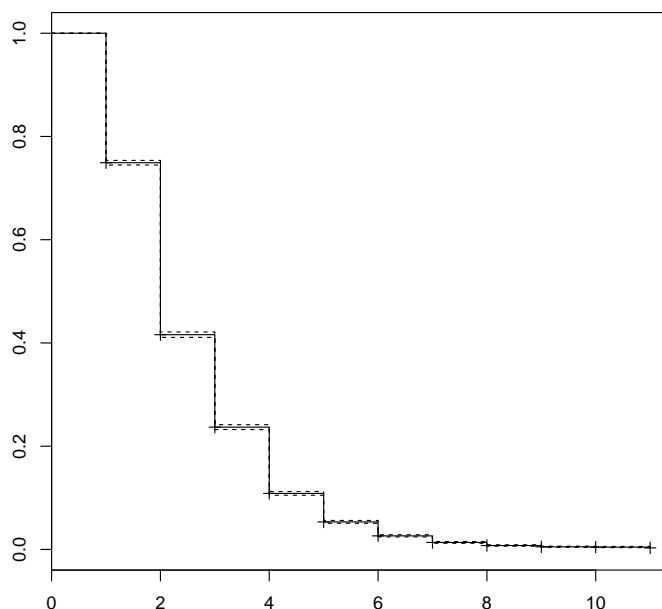


FIGURE 6.3 – Probabilité de survie estimée par Kaplan-Meier (intervalles de confiance à 95% en pointillés)

6.6.3 Facteurs de développement

La table 6.2 présente le nombre de facteurs de développement F_j (c.-à-d. le nombre de sinistres qui étaient ouverts en $j + 1$, ligne $\text{Card}(F)$), et leur répartition. Nous pouvons voir qu'une grande proportion de facteurs de développement est égale à 1 chaque année, c'est-à-dire reflétant une stabilité de la charge. Les facteurs de développement inférieurs à un peuvent concerner des cas de révision à la baisse de la provision dossier-dossier et des cas de recours.

j	0	1	2	3	4	5
Card(F)	24 892	12 484	6 317	2 517	1 049	397
%(F>1)	19.3	8.7	8.6	8.7	11.9	7.9
%(F<1)	32.7	25.2	27.2	29.3	29.7	27.0
%(F=1)	48.0	66.1	64.1	62.0	58.5	65.2
Fmax	82.82	28	100.71	20.23	105	6.45

TABLE 6.2 – Description des facteurs de développement

6.7 Vérification des hypothèses

Dans cette partie nous allons tester sur les données les hypothèses du modèle. Nous rappelons les trois hypothèses centrales faites dans le modèle :

Résumé des hypothèses

1. Les sinistres sont indépendants. Pour un sinistre, le processus de charges est indépendant de la durée de vie ;
2. Les durées de vie Γ_i sont i.i.d. ;
3. Les variables $F_{i,j}$ définies par $I_{i,j+1} = F_{i,j}I_{i,j}$ sont indépendantes du passé (c'est-à-dire de \mathcal{F}_u , avec u la date telle que $j = u - n_i$). Pour une année de développement j donné, les $F_{i,j}$ suivent la même loi pour tout i .

6.7.1 Indépendance entre les facteurs de développement et la durée de vie

La table 6.3 met en évidence une faible corrélation empirique entre les facteurs de développement et la durée de vie.

j	0	1	2	3	4	5
Card(F)	24 892	12 484	6 317	2 517	1 049	397
Corr(F, Γ)	0.008	0.041	0.004	0.037	-0.011	0.083

TABLE 6.3 – Facteurs de développement et corrélation empirique avec la durée de vie

La figure 6.4 présente le nuage de points des durées de vie et des premières charges, pour tous les sinistres étudiés entre 2002 et 2012 qui sont clos avant fin 2012. Sur les sinistres d'âge inférieur à 8 ans et de charges inférieurs à 50 000, l'hypothèse d'indépendance semble bien vérifiée. Il reste la queue de la distribution bi-variée empirique, où l'on observe des âges supérieurs à 9 ans seulement pour les petits sinistres, et des durées de vie faibles pour les très gros sinistres (effet diagonale sur le graphique). L'indépendance entre la durée de vie et le processus de charges ne semble pas fondamentalement remise en cause, du fait qu'on dispose de beaucoup de moins d'observations pour les sinistres graves.

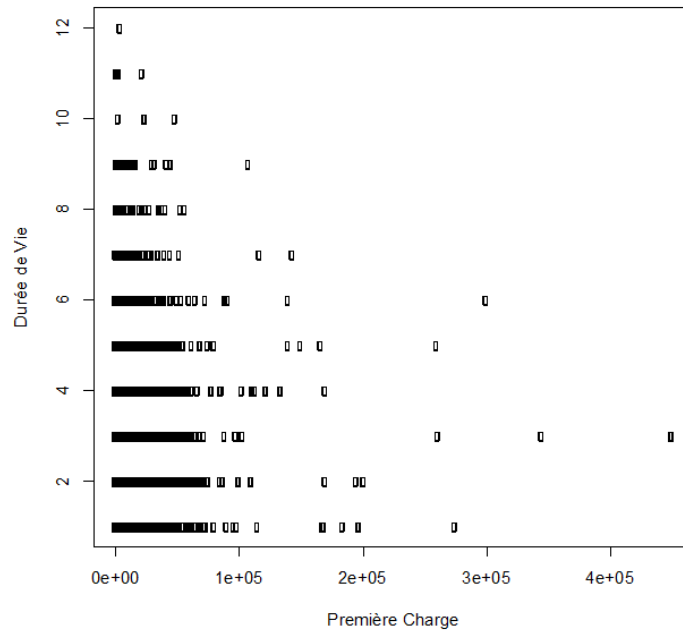


FIGURE 6.4 – Corrélation entre l'âge du sinistre et la première charge

6.7.2 Indépendance entre les facteurs de développement et le passé

A première vue, la corrélation linéaire entre les facteurs de développement et les charges (entre F_i^j et C_i^j) est très faible, ce qui peut nous laisser penser que l'hypothèse d'indépendance n'est pas invalidée (voir table 6.4).

j	0	1	2	3	4	5
Card(F)	24 892	12 484	6 317	2 517	1 049	397
Corr(C,F)	-0.038	-0.024	-0.020	-0.025	-0.020	0.014

TABLE 6.4 – Corrélation empirique des Facteurs de développement $F_{i,j}$ avec la charge $C_{i,j}$

Cependant, une étude de la distribution empirique jointe, pour j donné, de $F_{i,j}$ et $C_{i,j}$ permet de mettre en lumière des phénomènes de dépendance (voir figure 6.5). En effet, on observe une très forte volatilité des facteurs de développement pour les charges petites, et une distribution très resserrée autour de un pour les grandes charges. Les valeurs négatives des corrélations nous suggèrent la même chose : plus la charge est déjà élevée, moins il est probable que les facteurs de développement suivants soient élevés.

Ce résultat suggère de procéder à une segmentation de la base, entre sinistres attritionnels et sinistres graves, afin d'avoir une meilleure homogénéité des facteurs de développement en fonction des charges, ce que nous ferons dans les parties 6.9 et 6.10.

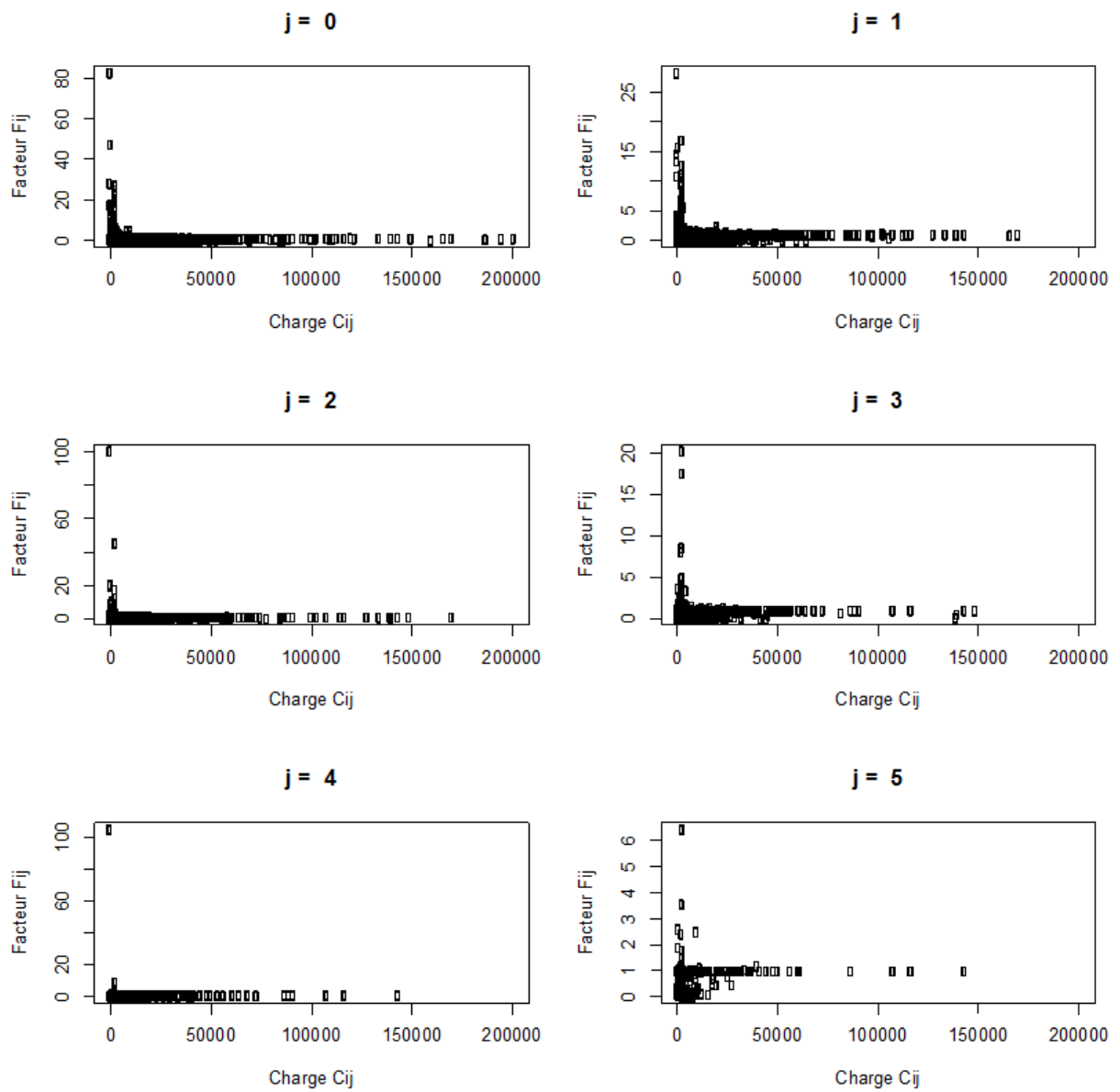


FIGURE 6.5 – Distribution croisées des charges et des facteurs

6.7.3 Indépendance des facteurs de développement

Nous réalisons des tests ANOVA sur la moyenne des facteurs de développement. L'hypothèse nulle est : il y a au moins deux catégories pour lesquelles les moyennes de la variable sont différentes. Nous considérons les échantillons F_0 en fonction de l'année de survenance. Les résultats sont les suivants.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(d\$org)	9	56.38	6.26	5.44	1.69e-7
Residuals	25320	29131.20	1.15		

Nous rejettons l'hypothèse nulle (la p-value $Pr(> F)$ est inférieure à 0,05%). Cependant, l'ANOVA est valide dans un cadre gaussien ou quasi-gaussien, tandis que nous avons plutôt des distributions proches de loi log-normales pour les facteurs de développement. Sur la figure 6.6, nous avons tracé la fonction de répartition des F_0 par année de survenance (d'abord \mathbb{R}^+ , puis en tronquant les valeurs inférieures à 1). On constate que la proportion de facteurs égale à un an peut varier assez sensiblement selon les années. Sur la queue de distribution à droite, il y a également une fonction de répartition empirique qui est nettement plus basse que les autres (il s'agit de celle des premiers facteurs de l'année de survenance 2005). L'hypothèse d'indépendance n'est donc que partiellement vérifiée, mais les différences entre les distributions par année de survenance ne semblent pas non plus trop marquées.

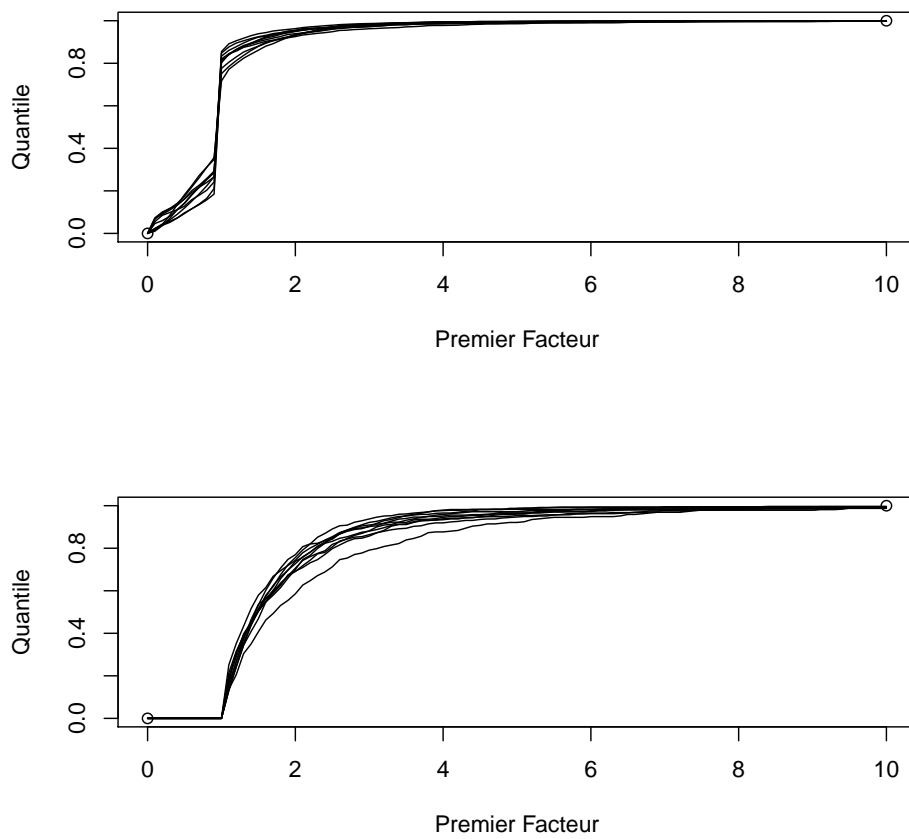


FIGURE 6.6 – Fonction de répartition empirique du premier facteur de développement selon les années de survenance

6.8 Estimation sans segmentation

6.8.1 Description de la démarche

Nous allons considérer la base des dommages corporels, qui contient environ 36000 sinistres. Nous nous intéressons aux charges et non au paiement, puisque ces premières respectent mieux les hypothèses faites dans le modèle. Nous ne faisons pas de segmentation, pour vérifier que les procédures bootstrap ne sont pas adaptées dans ce cas-là.

Nous allons nous intéresser aux réserves :

- à horizon un an : la différence entre la charge totale dans un an et la charge totale actuelle (variation de charge à un an).
- à horizon ultime : la différence entre la charge ultime (qui sera également égal au paiement cumulé ultime) et le paiement cumulé actuel (réserve ultime).

Pour estimons chaque réserve avec les méthodes suivantes :

- Chain Ladder (CL) sur le triangle agrégé des charges ;
- Ligne à Ligne (ML) sur les données non-agrégées de charges
- Méthode Bootstrap) sur les données non-agrégées de charges. Nous utilisons la méthode bootstrap avec estimation de la durée de vie présentée dans la partie 4.4.2 (BST) et la méthode bootstrap « total » (4.4.1) (BST TOT). Notons que les deux méthodes sont identiques pour la réserve à horizon un an.

6.8.2 Résultats

Les résultats des estimations sont présentés sur les tableaux 6.5 et 6.6. L'écart relatif entre la méthode Chain-Ladder et les autres méthodes est donné.

On constate que les résultats de Chain-Ladder et de la méthode ligne à ligne sont proches. La convergence entre les deux semble bien vérifiée. La variation de charge à un an est négative, ce qui signifie que les gestionnaires de sinistres ont tendance à être trop conservateurs lors des premières évaluations. Nous constatons aussi que la charge ultime est encore plus basse que la charge à un, cela souligne l'importance des recours. Le *Best Estimate* reste cependant positif.

Nous constatons en revanche une surévaluation très importante de la moyenne de la distribution des provisions obtenues par bootstrap. Nous pouvons remarquer sur la figure 6.7 que les distributions estimées par bootstrap ont des queues de distribution très longues. Sans segmentation, la méthode peut associer des facteurs de développement importants à des sinistres qui ont déjà de très grande charges. Une telle association est trop irréaliste et biaise l'estimation de la provision. L'hypothèse d'indépendance entre les facteurs de développement et le passé est donc cruciale pour que la méthode soit valide. Nous procédons donc maintenant à la segmentation de la base.

Méthode	Charge à un an	$\Delta CL / CL$	Δ Charge un an	$\Delta CL / CL$	Écart-type empirique
CL	35 589 006	-	-2 679 317	-	-
ML	35 824 580	0.007	-2 443 742	0.088	-
BST	38 427 808	0.080	159 485	1.060	2 193 989

TABLE 6.5 – Réserve à horizon un an - résultats des différentes méthodes (Données 2002-2012 Own Damage). 20 000 tirages sont réalisés pour chaque sinistre ouvert à la date d'étude.

Méthode	Charge Ultime	$\Delta CL / CL$	Réserve	$\Delta CL / CL$	Écart-type empirique
CL	31 913 148	-	5 995 328	-	-
ML	33 181 616	0.040	7 263 796	0.212	-
BST	37 590 086	0.178	11 672 266	0.947	3 816 895
BST TOT	36 882 313	0.156	10 964 493	0.829	3 639 875

TABLE 6.6 – Réserve à horizon ultime - résultats des différentes méthodes (Données 2002-2012 Own Damage). 20 000 tirages sont réalisés pour chaque sinistre ouvert à la date d'étude.

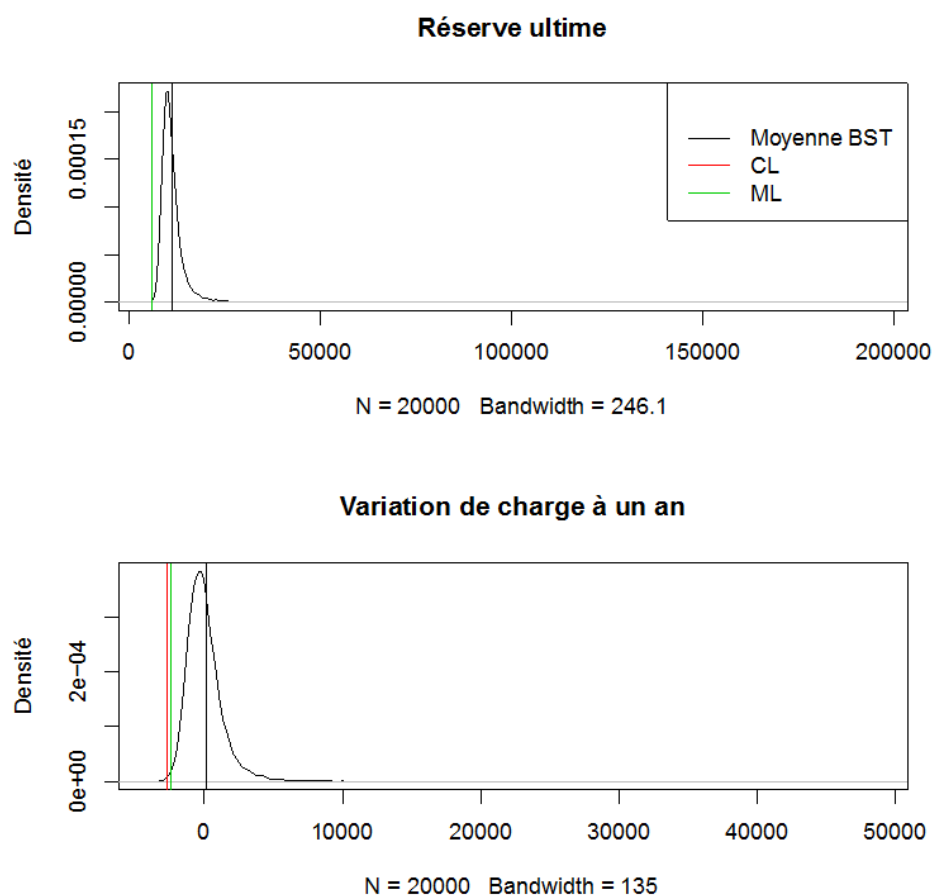


FIGURE 6.7 – Distribution de la variation de charge à un an et de la réserve ultime.

6.9 Segmentation *a priori*

6.9.1 Démarche

La première idée intuitive pour segmenter la base est de séparer les sinistres graves des sinistres attritionnels en fonction des charges observées. Cette méthode, appelée **Bootstrap avec Segmentation *a priori*** (BST APR), est implémentée de la manière suivante :

1. Division de l'ensemble des sinistres en deux sous-ensembles (ou potentiellement davantage) selon un critère fixé *a priori* sur les charges observées des sinistres. Le seuil pour déterminer à quelle classe appartient un sinistre peut être appliqué à différentes charges, notamment :
 - La première charge observée des sinistres (première colonne du triangle agrégé) ;
 - la dernière charge observée des sinistres (diagonale du triangle agrégé)Nous retenons la première solution, qui nous apparaît la plus cohérente. Avec la seconde, nous appliquons le seuil à la fois sur des sinistres qui sont clos et des sinistres qui sont au tout début de leur cycle de vie.
2. Application de la procédure bootstrap séparément sur chaque sous-ensemble de sinistres.
3. Réunion des deux ensembles pour déterminer la distribution de la réserve au niveau du portefeuille.

6.9.2 Choix du seuil

Pour fixer le seuil *a priori*, nous pouvons regarder, sur l'ensemble des sinistres clos, quelle est la proportion de sinistres qui a une charge finale au-delà de 10000, 15000, 20000 sachant que sa charge initiale est inférieure au niveau $I_{i,0}$ (10000, 15000, 20000 sont respectivement les quantiles 90%, 95% et 97% de la charge finale). Le résultat est présenté sur la figure 6.8 semble indiquer qu'un seuil fixé à 3000 est raisonnable. On constate que, tant qu'on considère l'ensemble des sinistres qui ont une charge inférieure à 3000, il y a une très faible chance qu'un des sinistres de cet ensemble termine avec une charge très élevée.

6.9.3 Résultats

Nous avons fixé à 3000 le seuil de segmentation *a priori* après avoir examiné la distribution de la première charge, c'est-à-dire au-dessus du forfait d'ouverture très couramment utilisé pour les sinistres attritionnels (2800).

Une segmentation par première charge semble résoudre de manière satisfaisante le problème de la sur-évaluation des réserves par la méthode bootstrap sans segmentation. Les résultats sont donnés sur la figure 6.9 ainsi que les tables 6.7 et 6.8. Avec segmentation, la moyenne de la distribution des réserves estimée par bootstrap est plus proche de Chain Ladder (pour différents choix du Seuil). L'écart-type empirique est également beaucoup moins important : la queue de la distribution est moins épaisse.

Cependant, les quantiles empiriques restent assez élevés : en effet, les échantillons bootstrap de l'ensemble des « sinistres attritionnels » contiennent des grands facteurs (voir la figure 6.5) : un sinistre dans cet ensemble commençant sa vie avec une petite charge peut accumuler des grands facteurs de développement et terminer sa vie avec une charge très importante. Dans la pratique, des sinistres d'abord évalués comme attritionnels peuvent être ré-évalués plus tard

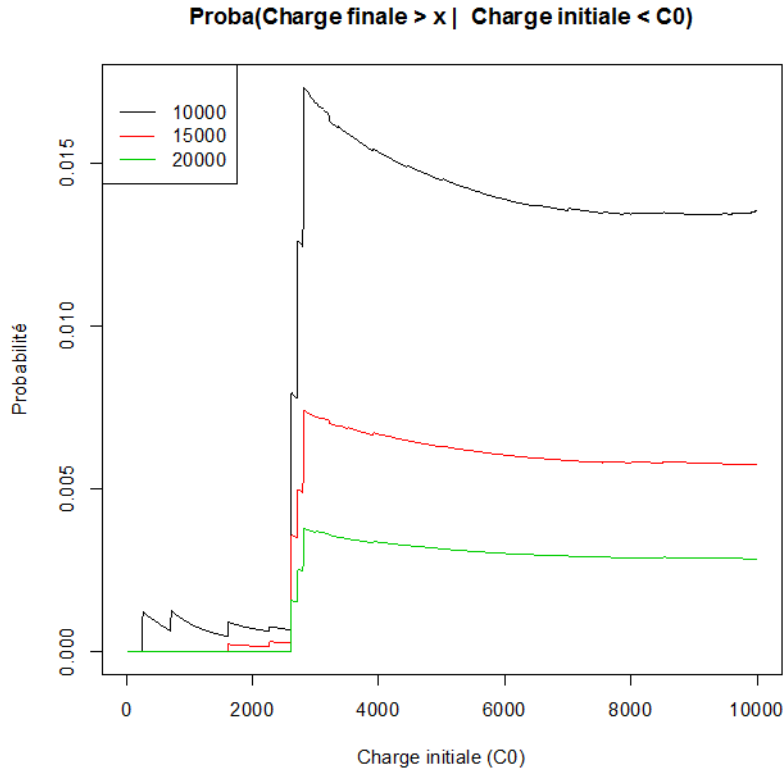


FIGURE 6.8 – Probabilité empirique d’avoir une charge finale supérieure à 10 000, 15 000 et 20 000 € en fonction de la charge initiale (calculé sur les sinistres clos)

comme graves par les gestionnaires de sinistres, ou entraîner des paiements beaucoup plus importants que prévus. Cependant, une fois qu’ils deviennent graves, il est irréaliste de continuer à tirer des facteurs de développement élevé, multipliant la charge par 20 ou plus. Ce problème est traité dans la section suivante, en introduisant une segmentation dynamique.

Méthode	Charge à un an	$\Delta CL / CL$	Δ Charge un an	$\Delta CL / CL$	Écart-type empirique
CL	31 913 148	-	5 995 328	-	-
BST APR	32 630 622	0.022	6 712 802	0.120	1 066 832

TABLE 6.7 – **Segmentation a priori** - Réserve à horizon un an. . Le seuil de segmentation a été fixé à 3 000. 20 000 tirages sont réalisés pour chaque sinistre ouvert.

Méthode	Charge ultime	$\Delta CL / CL$	Réserve	$\Delta CL / CL$	Écart-type empirique
CL	32 377 535	-	5 996 216	-	-
BST APR	32 013 047	0.011	5 631 727	0.061	1 276 860

TABLE 6.8 – **Segmentation a priori** - Réserve à horizon ultime. Le seuil de segmentation a été fixé à 3 000. 20 000 tirages sont réalisés pour chaque sinistre ouvert.

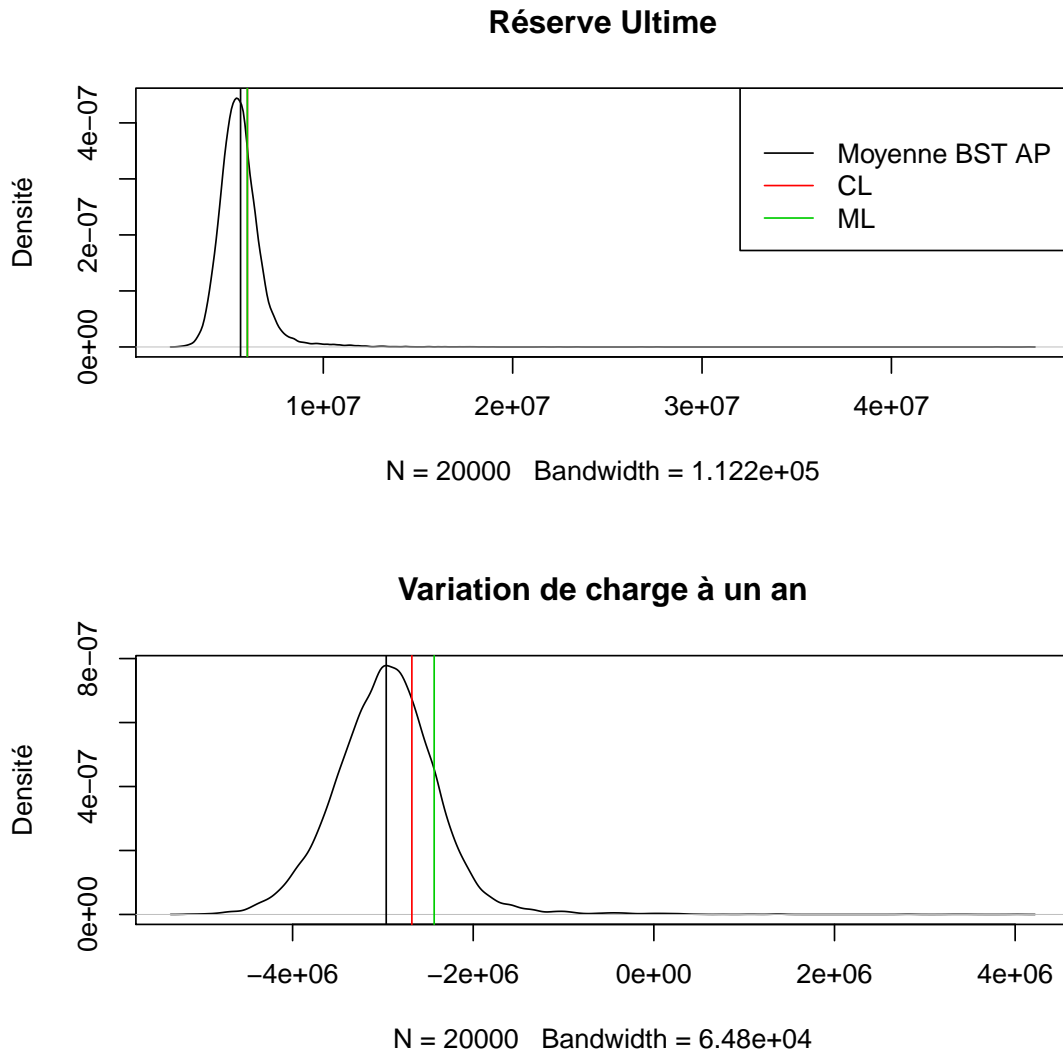


FIGURE 6.9 – **Segmentation *a priori*** - Distribution de la variation de charge à un an et de la réserve ultime.

6.9.4 Indépendance des facteurs avec le passé après segmentation

Nous allons étudier si la corrélation empirique entre les facteurs de développement et le passé est plus faible après segmentation. Sur la table 6.9, nous pouvons voir qu'il n'y a pas de réels améliorations pour les sinistres graves (par rapport à la table 6.4). Pour les sinistres attritionnels, on constate une corrélation plus faible. Pour les deux ensembles de sinistres corrélation est plus forte pour les dernières années de développement.

Sinistres	j	0	1	2	3	4	5
Graves	Card(F)	7 337	3 972	2 117	905	387	157
	Corr(C,F)	-0.032	-0.036	-0.035	-0.042	-0.060	0.130
Attritionnels	Card(F)	17 993	8 738	4 335	1 663	699	262
	Corr(C,F)	-0.008	-0.006	-0.021	-0.005	-0.034	-0.029

TABLE 6.9 – : Corrélation empirique entre les facteurs de développement $F_{i,j}$ et la charge $C_{i,j}$

Les figures 6.10 et 6.11 illustrent le phénomène sous-jacent : les petits sinistres peuvent devenir gros au cours de leur vie (et avoir alors des facteurs de développement stables proches de 1). De même, des sinistres d'abord évalués comme graves peuvent être réévalués comme attritionnels. La segmentation *a priori* sur la première charge ne permet pas de modéliser ce phénomène. La méthode de segmentation de la partie suivante est pour but de le modéliser. Avant cela, nous allons nous intéresser à la sensibilité de la méthode de segmentation *a priori* à un changement dans le seuil.

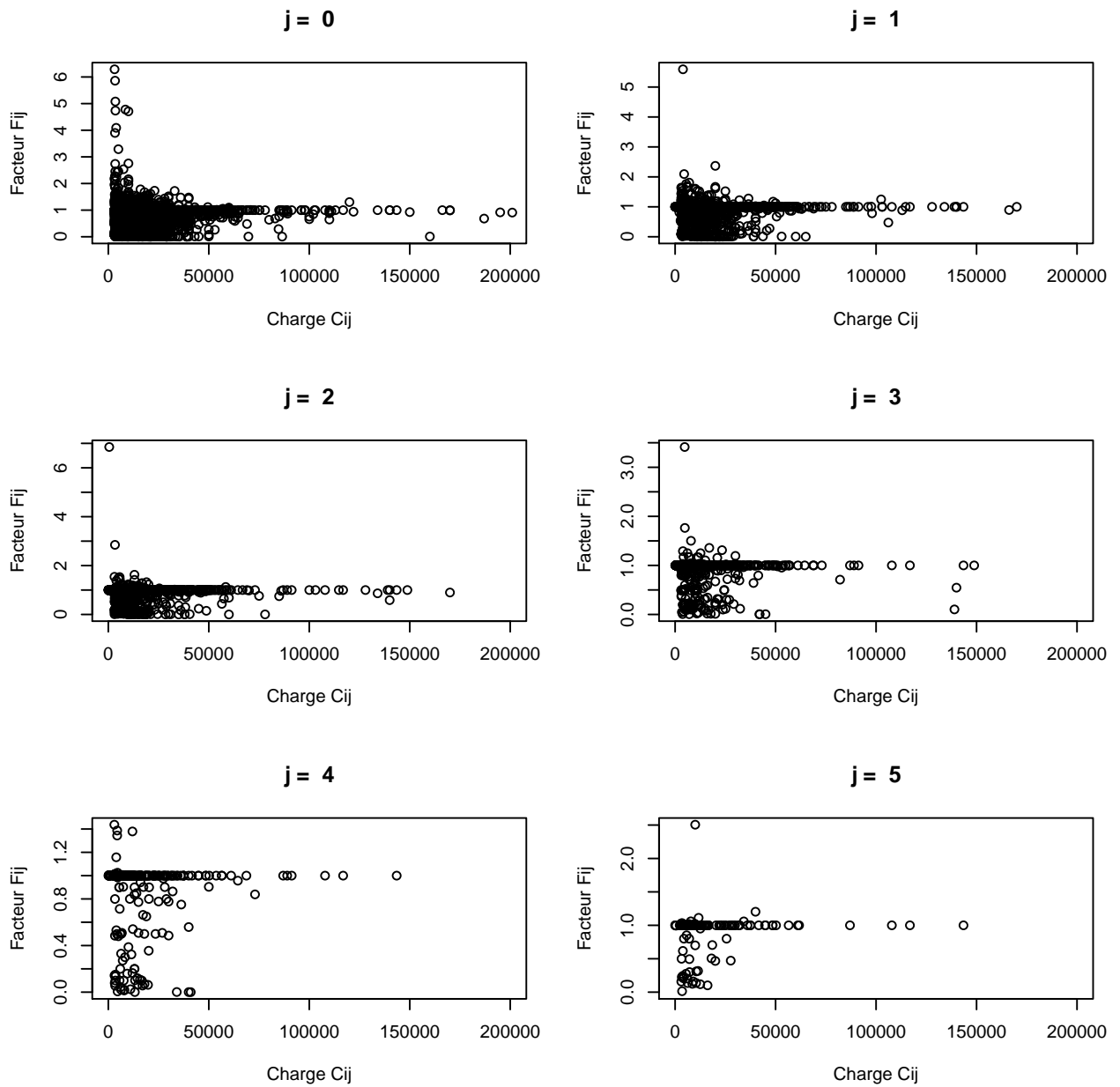


FIGURE 6.10 – **Segmentation *a priori*** - Charges vs Facteurs pour l'ensemble des sinistres graves

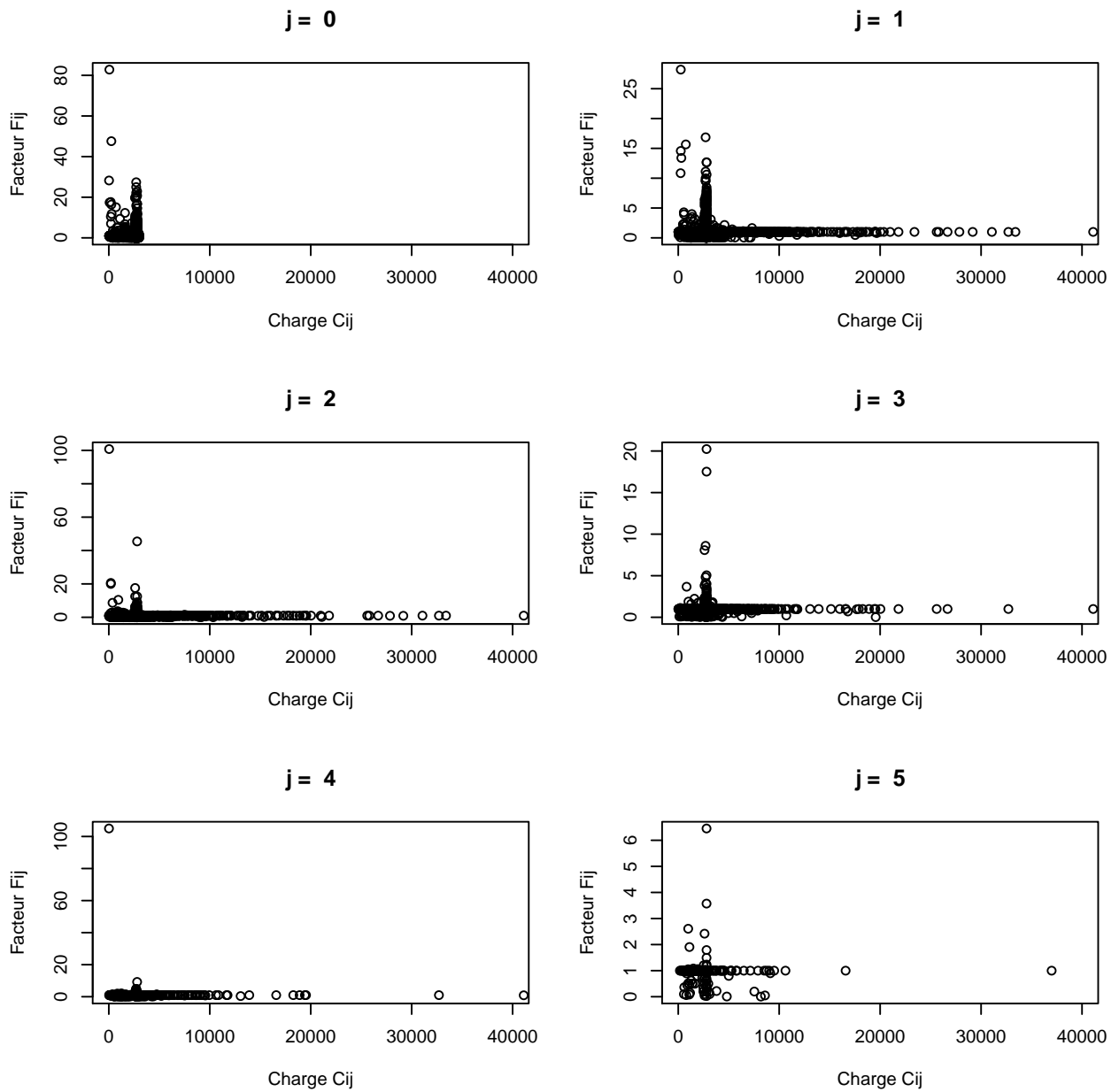


FIGURE 6.11 – **Segmentation *a priori*** - Charges vs Facteurs pour l'ensemble des sinistres attritionnels

6.9.5 Sensibilité par rapport au seuil

Nous allons intéresser à la robustesse de la méthode de segmentation *a priori* : l'estimation des provisions est-elle sensible à un changement dans le choix du seuil fixé ?

Les figures 6.13 et 6.12 nous montrent que les résultats restent assez stables à un changement de seuil. Le quantile extrême 99.5 % reste très fluctuant : en effet, nous n'avons réalisé que 3000 simulations pour les différents seuils pour limiter les temps de calcul. Cela rend l'estimation empirique de ce quantile peu robuste, d'autant plus que la queue de distribution est épaisse.

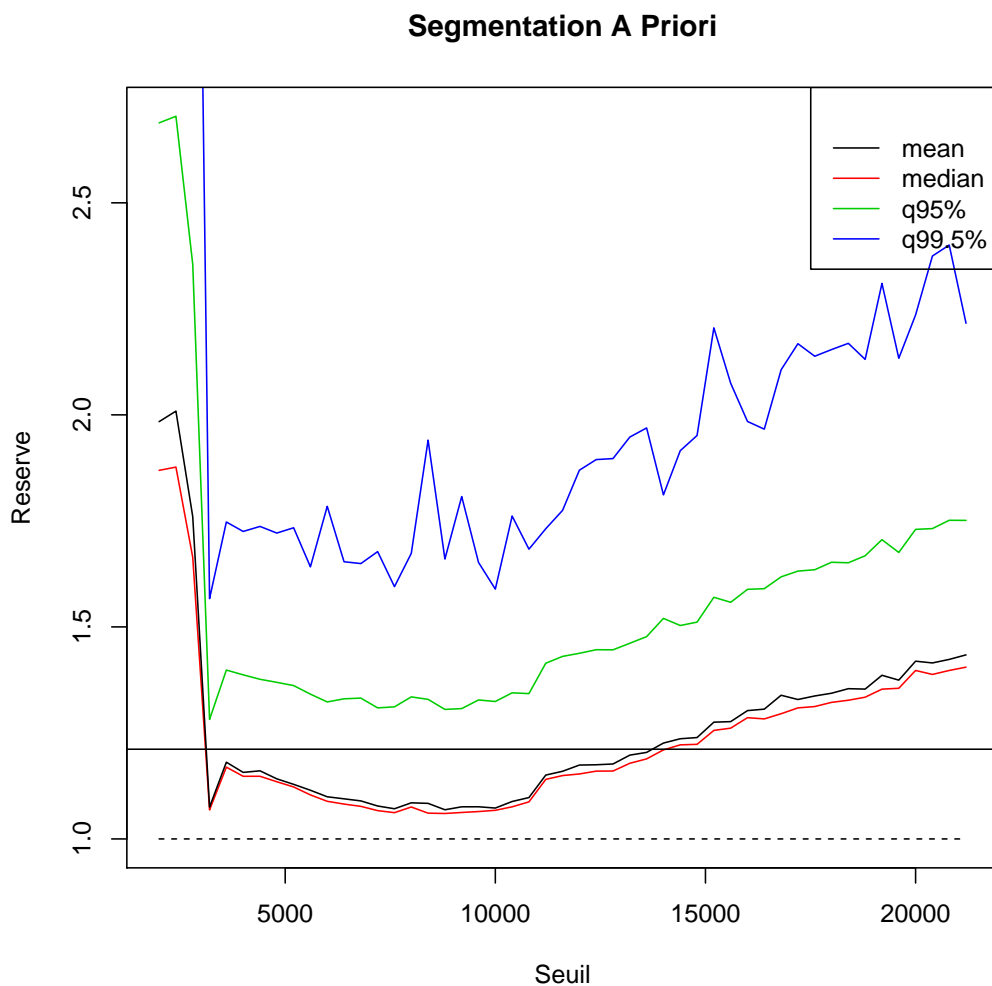


FIGURE 6.12 – Moyenne et quantiles de la réserve ultime en fonction du seuil de segmentation. Le résultat de Chain-Ladder est pris comme référence (ligne pointillé). 3 000 tirages pour chaque sinistre.

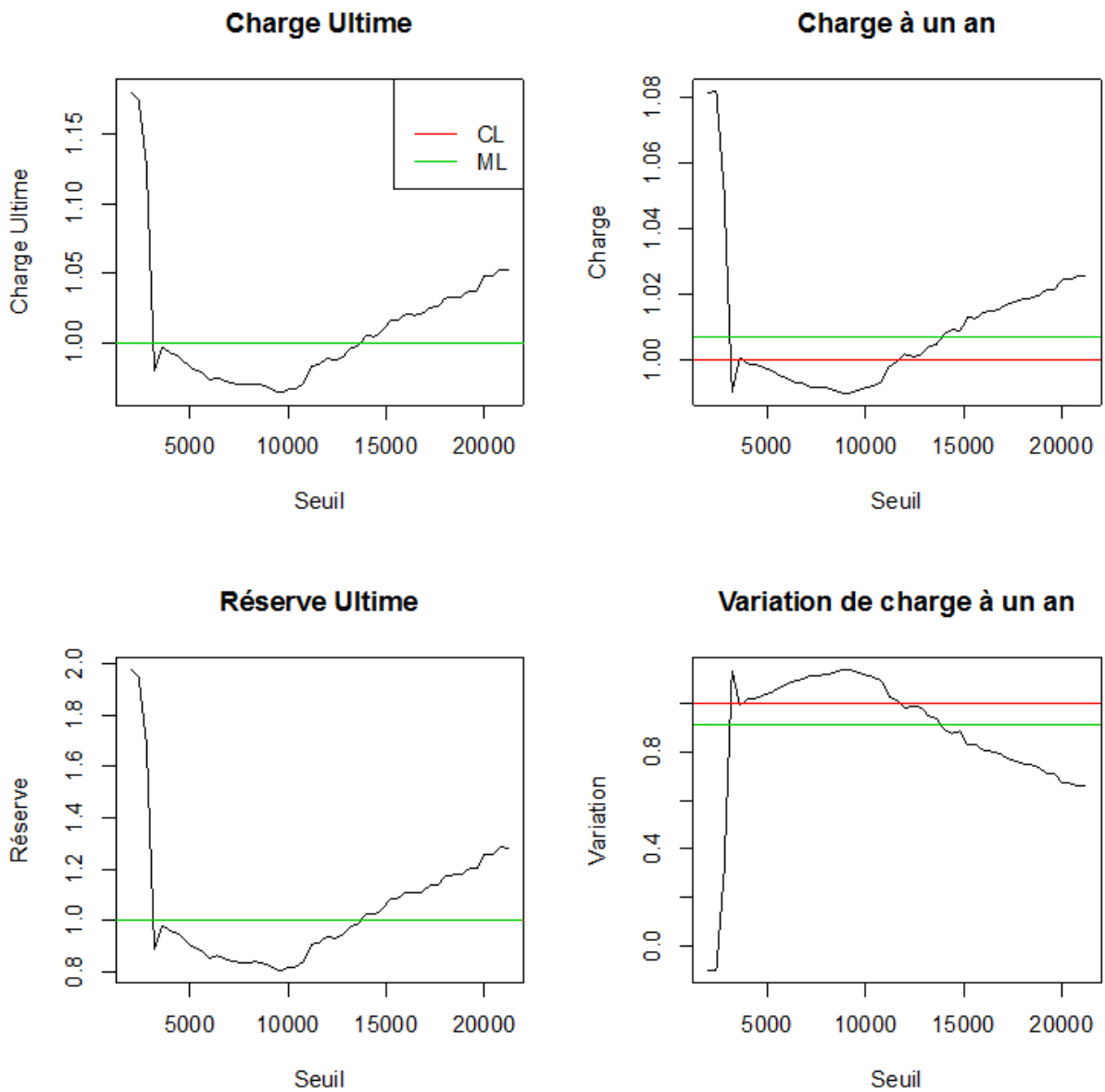


FIGURE 6.13 – Moyenne estimée par la méthode BST AP de la charge à un an et à l’ultime, de la variation de charge à un an et de la réserve ultime en fonction du seuil. Le résultat de Chain-Ladder est pris comme référence. A gauche, la réserve ligne à ligne et Chain-Ladder se superposent. 3 000 tirages pour chaque sinistre.

6.9.6 Segmentation *a priori* sur la dernière charge observée

Nous avons aussi utilisé la dernière charge connue comme critère de segmentation, en la comparant à un seuil fixé (qui peut dépendre éventuellement de l'âge du sinistre à la date de la dernière observation). Mais cette segmentation n'apporte pas de réelle amélioration par rapport au bootstrap sans segmentation. La moyenne des réserves ultimes estimées par cette méthode, pour un seuil de 3000, est 57 801 186 avec un écart-type empirique de 6 850 984.

Sur l'ensemble des sinistres graves, on a toujours une probabilité non négligeable d'associer un facteur important à une charge déjà élevée. En effet, les échantillons de cet ensemble contiennent des facteurs très importants, correspondant aux premières années de développement)où certains sinistres n'étaient pas encore évalués comme grave. Ces facteurs sont ensuite attribués à des sinistres qui sont évalués comme grave dès le début de leur vie.

En outre, il est difficile de définir un modèle sous-jacent dans ce cas, et donner des hypothèses qui assurent que la segmentation par la dernière charge soit une méthode de séparation consistante pour l'évaluation des réserves : l'année de développement de la dernière charge observée n'est pas la même pour tous les sinistres. De plus, pour les sinistres clos, il s'agit de la vraie charge ultime, tandis qu'il s'agit de la dernière charge estimée pour les sinistres non clos.

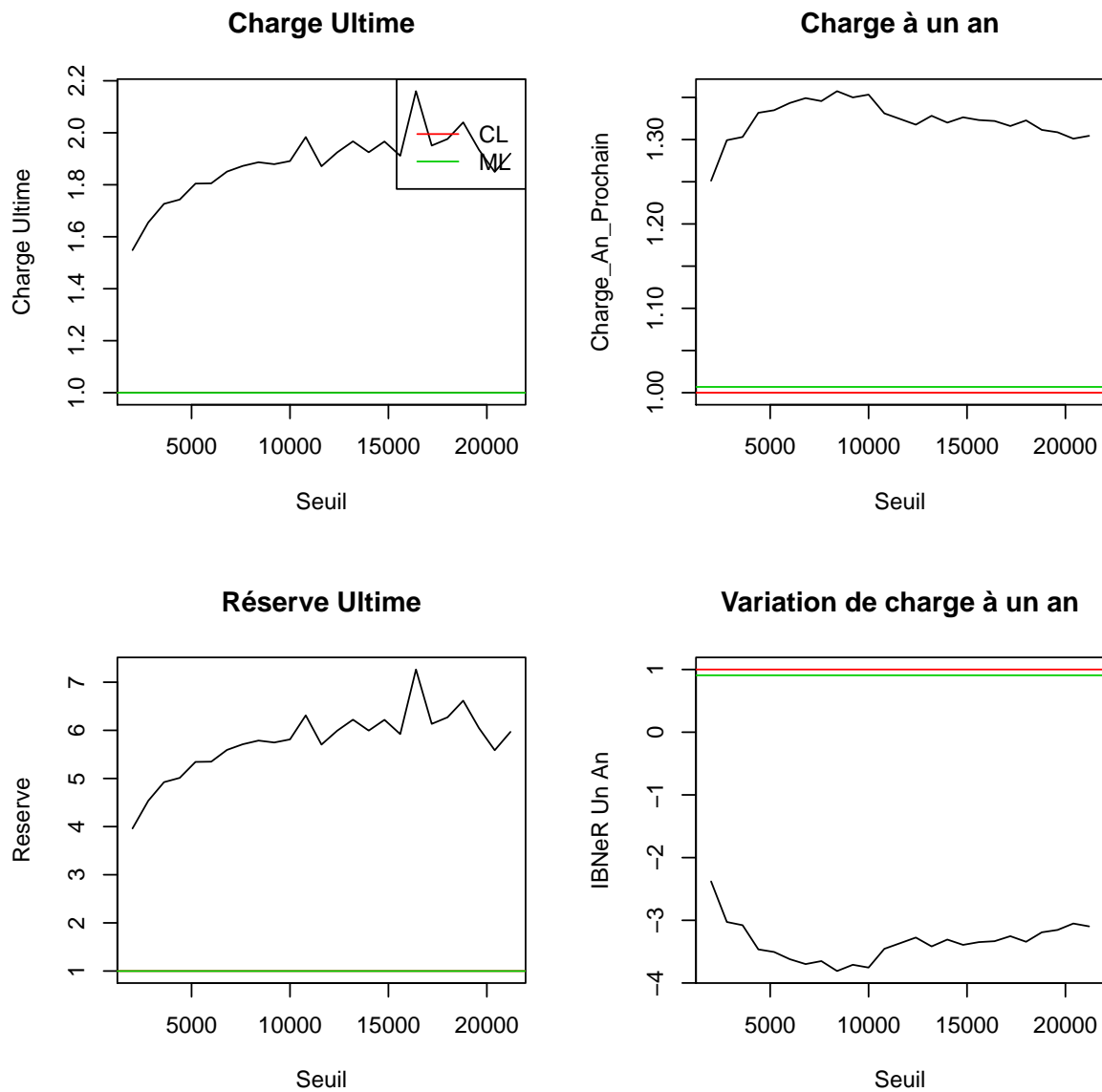


FIGURE 6.14 – Méthode de segmentation *a priori* (sur la dernière charge). Charge à un an et à l'ultime, variation de charge à un an et réserve ultime en fonction du seuil. Le résultat de Chain-Ladder est pris comme référence.

6.10 Segmentation dynamique

Nous proposons une méthode de « segmentation dynamique », qui consiste à ne plus segmenter les sinistres mais à segmenter les facteurs selon les charges qui précèdent. L'idée consiste pour une année de développement donnée, l'échantillon des facteurs en deux parties : les facteurs précédés par des petites charges et ceux précédés par des grandes charges. Ces deux catégories de charges sont définies en fixant un seuil exogène.

Au niveau de la procédure bootstrap, selon la charge atteinte par le sinistre en année de développement j , nous tirons dans l'ensemble de facteurs de développement correspondant. En faisant des échantillons de facteurs de développement séparés pour les grandes charges et pour les petites charges, nous n'autorisons plus la procédure bootstrap à tirer des facteurs de développement élevés pour des charges déjà très importantes. Nous modélisons ainsi le fait qu'un sinistre qui a atteint le stage de développement où il est évalué comme grave par les questionnaires de sinistre, connaîtra des variations de charge moins importantes que les sinistres qui restent attritionnels.

6.10.1 Modèle

Nous reprenons le cadre général défini dans le chapitre 3. Nous retenons les mêmes hypothèses faites sur l'indépendance mutuelle des sinistres et l'indépendance, pour un sinistre, entre le processus d'états et de charges. Pour simplifier la modélisation, nous supposons cette fois-ci l'équi-distribution de la première charge (premier paiement).

L'hypothèse faite sur la dynamique du processus de charges (4.1.2.1) est remplacée par la suivante :

Hypothèse : Il existe une constante K (qui peut dépendre éventuellement de j) et des variables aléatoire $H_{i,j}$ et $B_{i,j}$ tels que :

1.

$$\frac{C_{i,j+1}}{C_{i,j}} = \mathbf{1}_{C_{i,j} > K} H_{i,j} + \mathbf{1}_{C_{i,j} \leq K} B_{i,j},$$

pour tout sinistre i et pour toute année de développement j .

2. Les variables $H_{i,j}$ et $B_{i,j}$ sont mutuellement indépendantes et sont indépendantes du passé.

3. Les lois de $H_{i,j}$ et de $B_{i,j}$ ne dépendent pas de i .

Cette hypothèse implique que les facteurs $F_{i,j}$ ne sont plus indépendants du passé. Leur loi dépend de dernier paiement $C_{i,j}$.

Notons que sous $\{C_{i,j} > K\}$, on a : $\frac{C_{i,j+1}}{C_{i,j}} = H_{i,j}$ est indépendant de $C_{i,j}$. De façon similaire, sous $\{C_{i,j} \leq K\}$, on a : $\frac{C_{i,j+1}}{C_{i,j}} = B_{i,j}$ est indépendant de $C_{i,j}$.

La décomposition de $F_{i,j}$ en deux variables $H_{i,j}$ et $B_{i,j}$ nous permet de modéliser d'une manière simple la dépendance entre les facteurs et les charges.

6.10.2 Procédure bootstrap

Nous détaillons dans cette partie la mise en œuvre de la méthode bootstrap dans le cadre de la segmentation dynamique, d'abord pour la réserve à un an, puis pour la réserve ultime.

6.10.2.1 Réserve à un an

La procédure est la suivante :

1. Échantillonnage

Nous fixons tout d'abord un seuil K afin de séparer les grandes charges et des petites charges.

Notons $\mathcal{EH}_j = \left\{ \frac{C_{i,j+1}}{C_{i,j}} \mid i \in \mathcal{B}_j^*(t), C_{i,j}^j > K \right\}$, l'ensemble des réalisations de facteur de développement F_j (entre les années de développement j et $j+1$, $0 \leq j \leq \Gamma^{max}$), connues à la date d'inventaire t , et **qui arrivent après une grande charge**.

Notons aussi $\mathcal{EB}_j = \left\{ \frac{C_{i,j+1}}{C_{i,j}} \mid i \in \mathcal{B}_j^*(t), C_{i,j} \leq K \right\}$, l'ensemble **correspondant aux charges attritionnels**.

2. **Bootstrap** On tire M échantillons bootstrapés, indicé $m = 1$ to M , de la manière suivante : Pour chaque sinistre non-clos S_i qui est en année de développement j (c.-à-d. $S_i \in \mathcal{A}_j^*(t)$, avec $j = t - n_i$), on tire (aléatoirement) des facteurs de développement $f_{i,j}^m$ dans l'échantillon correspondant \mathcal{EH}_j si $C_{i,j} > K$ ou dans \mathcal{EB}_j si $C_{i,j} \leq K$. On calcule $cf_i^m = C_{i,j}(f_{i,j}^m - 1)$, le flux de trésorerie à un an bootstrapé pour le sinistre S_i .

On effectue ensuite la somme des cf_i^m pour tous les sinistre non-clos i , ce qui donne une réserve à un an r^m .

3. **Estimation.** On estime la distribution empirique de la réserve à un an (par une méthode de noyau) à partir de r^1, r^2, \dots, r^M .

6.10.2.2 Réserve ultime

Pour chaque boucle du bootstrap $m = 1, \dots, M$:

1. Pour chaque sinistre non clos i qui est en année de développement j (c.-à-d. $i \in \mathcal{A}_j^*(t)$ et $j = t - n_i$), tirage aléatoire d'une durée de vie résiduelle pour ce sinistre ($\tilde{\Gamma}_i^m$), sachant qu'il est encore ouvert à la date d'étude. Initialisation de $CU_i^m = C_{i,j}$. Pour $k = j$ à $(\tilde{\Gamma}_i^m - 1)$, on procède ainsi :
 - Tirage d'un facteur de développement $f_{i,k}^m$ dans l'échantillon \mathcal{EH}_k si $CU_i^m > K$, ou dans \mathcal{EB}_k si $CU_i^m \leq K$
 - Calculer la nouvelle charge $CU_i^m = CU_i^m f_{i,k}^m$, et continuer ainsi jusqu'à $k = \tilde{\Gamma}_i^m - 1$
2. Faire la somme des CU_i^m pour tous les sinistre non-clos i , pour obtenir une estimation de la réserve ultime $r^{u,m}$.
3. Estimation de la distribution empirique de la réserve ultime (par une méthode de noyau) à partir de l'échantillon bootstrapé $r^{u,1}, r^{u,1}, \dots, r^{u,M}$.

6.10.3 Le choix du seuil

Dans cette méthode, le choix de K est crucial : un seuil trop petit ou trop grand laisse une probabilité trop grande que des grandes charges soient associées à des facteurs élevés. Nous proposons ici deux approches pour fixer le seuil K :

1. **Choix *a priori*** : Le Seuil K est tel que : conditionnellement à un des deux événements $\{C_{i,j} > K\}$ et $\{C_{i,j} \leq K\}$, nous devons avoir $\frac{C_{i,j+1}}{C_{i,j}}$ indépendant de $C_{i,j}$. Plusieurs méthodes peuvent être envisagés :
 - **Analyse graphique.** Sur les distributions croisées des charges et des facteurs (figure ??), nous cherchons le niveau de charge qui segmente au mieux les facteurs de développement. Notons que nous pouvons choisir un seuil par année de développement.
 - **Minimisation la corrélation empirique.** Un seuil K satisfaisant les hypothèses du modèle doit annuler la quantité suivante (condition nécessaire) :

$$L = [Corr(\frac{C_{i,j+1}}{C_{i,j}}, C_{i,j} | C_{i,j} > K)]^2 + [Corr(\frac{C_{i,j+1}}{C_{i,j}}, C_{i,j} | C_{i,j} \leq K)]^2$$

Cela nous suggère de choisir le seuil \hat{K} qui minimise la version empirique de cette quantité.

- **Seuil construit à partir d'un seuil sur les facteurs de développement** Pour chaque j , nous prenons K_j comme la charge maximale qui précède un facteur dépassant un seuil fixé sur le facteur. Plus précisément, nous définissons d'abord un seuil $f_j > 0$ sur les facteurs de développement (5, par exemple). Nous choisissons ensuite K_j de la manière suivante :

$$K_j = \max_i C_{i,j} , \text{ tel que : } \frac{C_{i,j+1}}{C_{i,j}} \geq f_j$$

2. **Validation croisée** : Nous effectuons des estimations avec de différents niveaux de K et retenons le résultat qui semble le meilleur à partir de certains critères définis en avance. Par exemple, nous pouvons demander que la valeur moyenne obtenue par bootstrap avec segmentation soit proche de résultat de Chain-Ladder ou de la méthode ligne à ligne.

6.10.4 Résultats

Les résultats sont donnés sur la figure 6.15 ainsi que les tables 6.10 et 6.11. Nous avons effectué le choix des seuils par la méthode de minimisation de corrélation empirique. Nous constatons une amélioration significative en comparant la distribution obtenue à celle de bootstrap sans segmentation. La queue de distribution est beaucoup moins épaisse.

Méthode	Charge à un an	$\Delta CL / CL$	Δ Charge un an	$\Delta CL / CL$	Écart-type empirique
CL	35 589 006	-	-2 679 317	-	-
BST SEG	35 218 603	0.010	-3 049 720	0.138	558 462

TABLE 6.10 – **Segmentation dynamique** - Réserve à horizon un an. Le seuil de segmentation a été fixé à 3000. 20 000 tirages sont réalisés pour chaque sinistre ouvert.

Méthode	Charge ultime	$\Delta CL / CL$	Réserve	$\Delta CL / CL$	Écart-type empirique
CL	32 377 535	-	5 996 216	-	-
BST SEG	32 706 011	0.025	6 788 191	0.132	763 715

TABLE 6.11 – **Segmentation dynamique** - Réserve à horizon ultime. Le seuil de segmentation a été fixé à 3000. 20 000 tirages sont réalisés pour chaque sinistre ouvert.

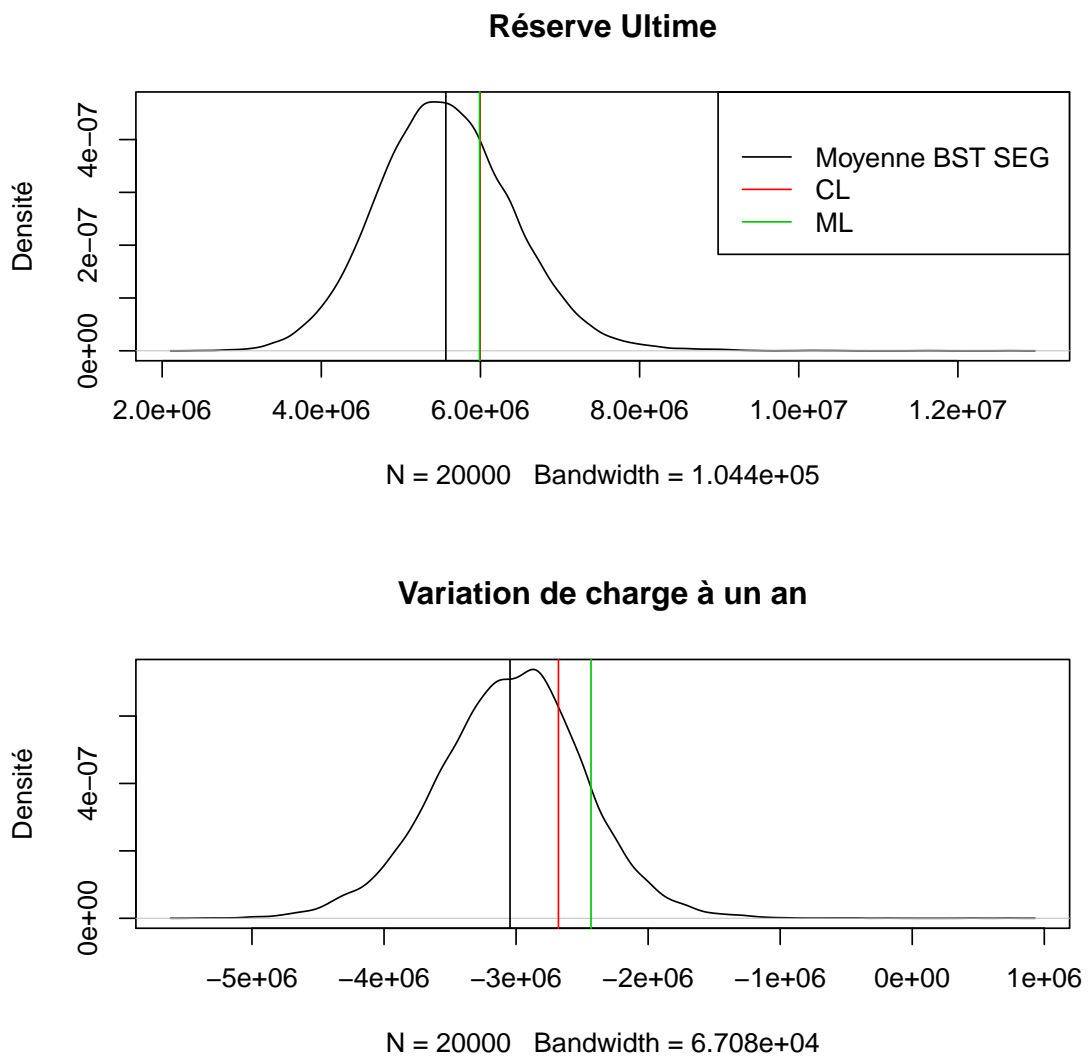


FIGURE 6.15 – **Segmentation dynamique** - Distribution de la variation de charge à un an et de la réserve ultime.

6.10.5 Sensibilité par rapport au seuil

Le résultat d'estimation pour différents Seuils est présenté sur les figures 6.16. Nous pouvons remarquer que le résultat reste plutôt stable quand le seuil change. La figure 6.17 donne des estimations de quantile correspondant à différentes valeurs du seuil K .

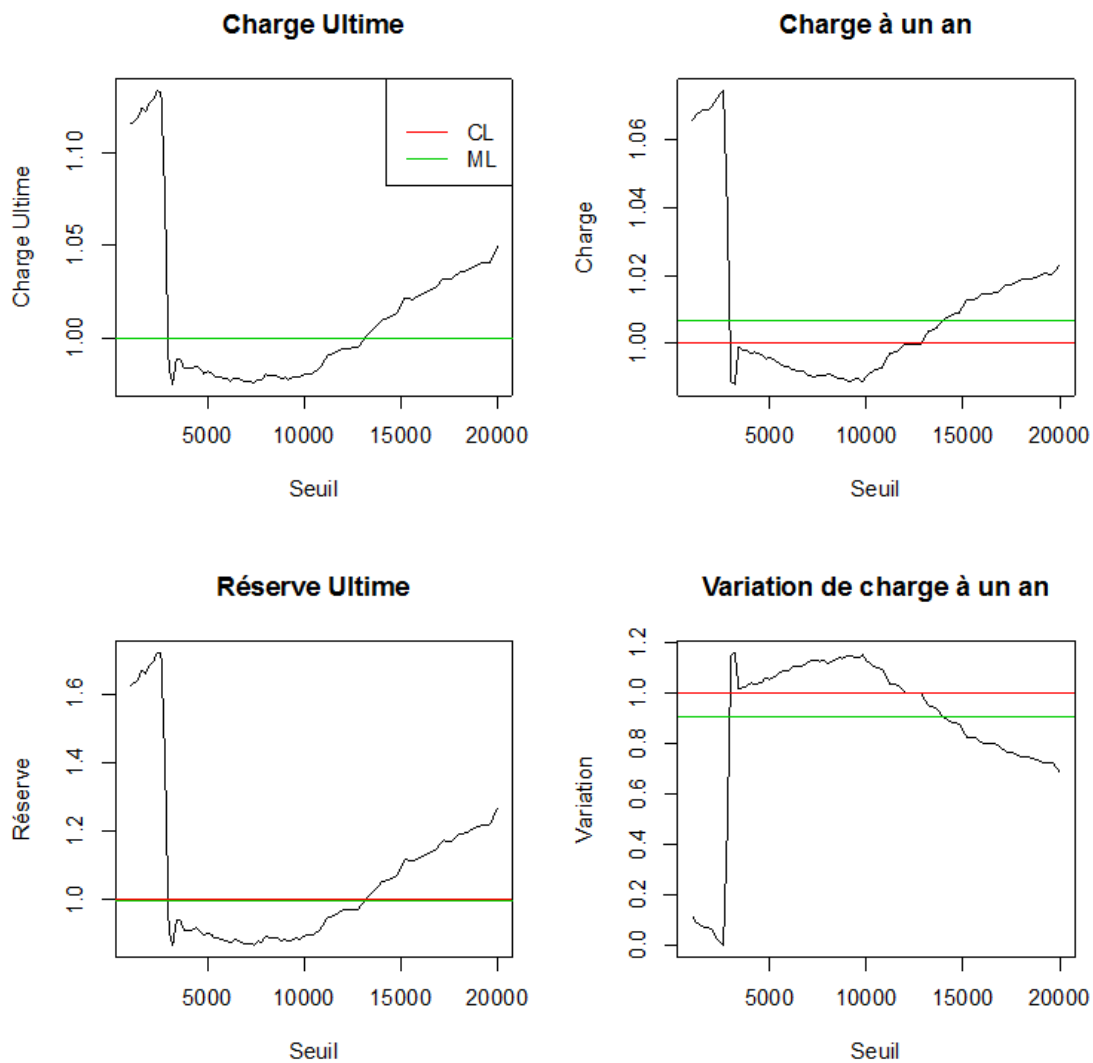


FIGURE 6.16 – **Segmentation dynamique.** Charge à un an et à l'ultime, variation de charge à un an et réserve ultime en fonction du seuil K . Le résultat de Chain-Ladder est pris comme référence.

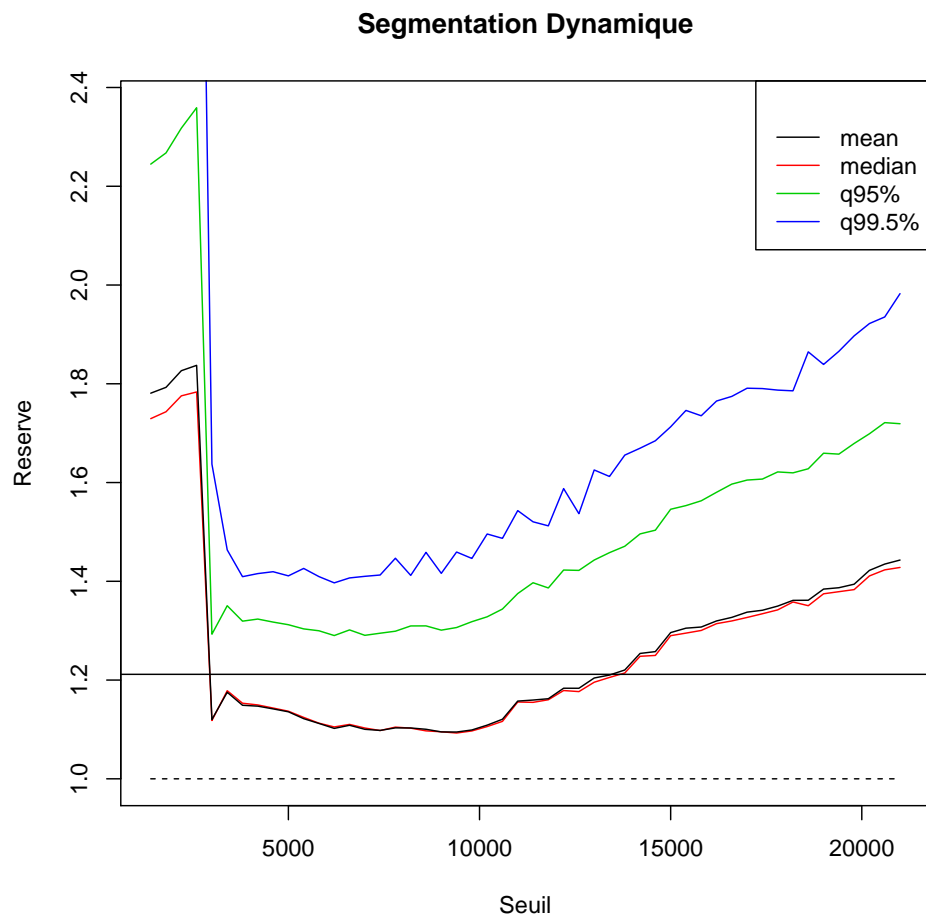


FIGURE 6.17 – **Segmentation dynamique** - Moyenne et quantiles de la réserve ultime en fonction du seuil de segmentation. Le résultat de Chain-Ladder est pris comme référence. 3 000 tirages pour chaque sinistre.

6.10.6 Comparaison entre la segmentation *a priori* et segmentation dynamique

Nous comparons les quantiles de la distribution obtenue par la méthode sans segmentation, la méthode de segmentation *a priori* (avec le seuil de 3000), la méthode de segmentation dynamique (avec la méthode de minimisation de la corrélation empirique). Nous voyons que les queues de distribution pour les deux méthodes de segmentation sont très proches, mais commencent à diverger au-delà du quantile à 95%.

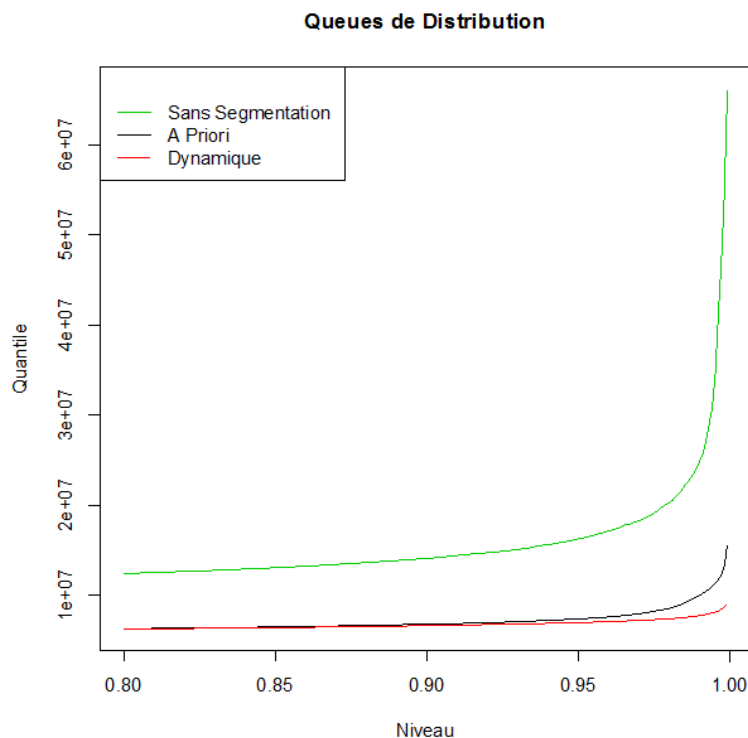


FIGURE 6.18 – Queue de distributions empiriques pour les 3 approches : sans segmentation, avec segmentation *a priori* et avec segmentation dynamique

La méthode de segmentation *a priori* produit donc des résultats qui paraissent raisonnables. On a très peu de grands facteurs pour l'ensemble des petits sinistres, la probabilité que la procédure bootstrap attribue plusieurs fois de suite des grands facteurs à un sinistre et conduisent à des charges finales irréalistes est très faible. S'il y a très peu de changements de catégories, comme sembler le montrer le graphique 6.8, et qu'il y a également peu de facteurs extrêmes pour l'ensemble des petits sinistres, cette méthode pourrait être suffisante, étant plus facile à implémenter que la segmentation dynamique.

Nous comparons aussi en fonction du seuil choisi. Nous obtenons une courbe de quantile à 99.5% plus régulière pour la méthode de segmentation dynamique, qui semble plus robuste dans la queue de distribution.

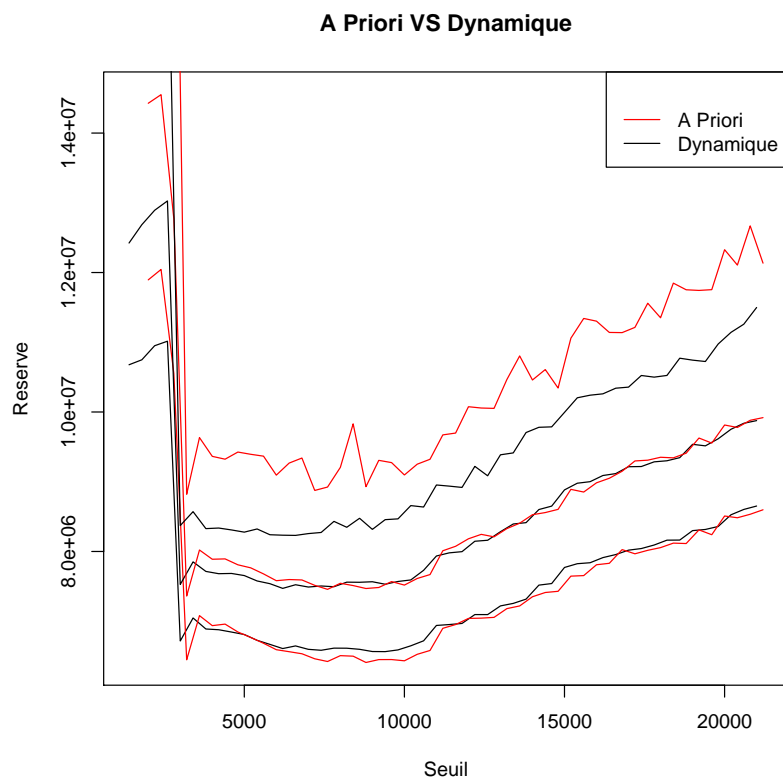


FIGURE 6.19 – Réserve Ultime - de bas en haut : moyenne, quantile de 90% et quantile de 99%. 3000 simulations pour chaque Bootstrap.

6.11 Analyses complémentaires

6.11.1 Temps de calcul

Pour discuter de l'implémentabilité de la méthode, nous nous sommes intéressés au temps de calcul. Le graphique suivant 6.20 donne le temps de calcul des différentes méthodes bootstrap utilisées (approche naïve, totale, segmentation *a priori*, segmentation *dynamique*) en fonction du nombre de tirages faits pour simuler la vie future d'un sinistre ouvert à la date d'étude. L'étude a été réalisée pour les 36 000 sinistres de la base et la profondeur de 10 ans (2002-2012). Le temps de calcul est linéaire avec le nombre de tirages.

La méthode bootstrap total prend davantage de temps, puisque pour chaque sinistre tiré, il faut vérifier s'il est clos ou ouvert pour passer à l'étape suivante. Dans les autres approches, la durée de vie (résiduelle) est tirée d'emblée.

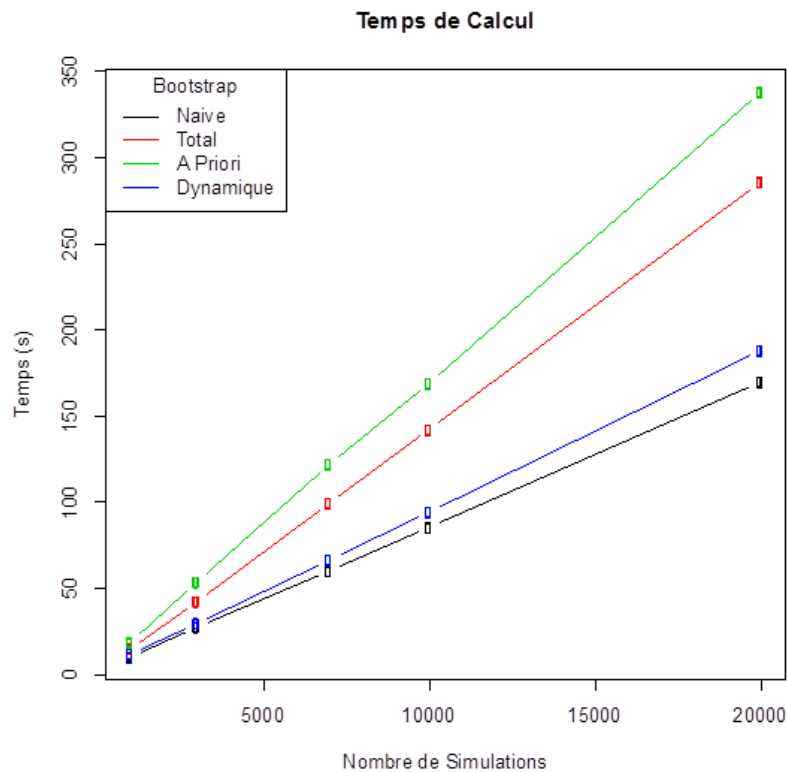


FIGURE 6.20 – Temps de calcul en fonction du nombre de tirages (secondes)

6.11.2 Comparaison de la distribution avec les modèles stochastiques agrégés

Nous comparons la distribution de la réserve ultime obtenue avec notre méthode ligne à ligne avec les distributions obtenues dans le cas des modèles stochastiques agrégés. Nous utilisons deux méthodes :

- La méthode de Mack, avec une hypothèse de normalité sur la distribution de la réserve à un an. Nous avons choisi une distribution gaussienne plutôt que log-normale pour modéliser une queue de distribution à gauche des réserves négatives.
- La méthode Bootstrap par Chain-Ladder avec modélisation GLM¹ (CL BST). Nous avons utilisé une loi Gamma pour la loi des charges incrémentales.

Nous avons utilisé le package Chain-Ladder de R² pour les méthodes stochastiques agrégées. Sur la figure 6.21, on constate que les méthodes ligne à ligne ont moins de variabilité.

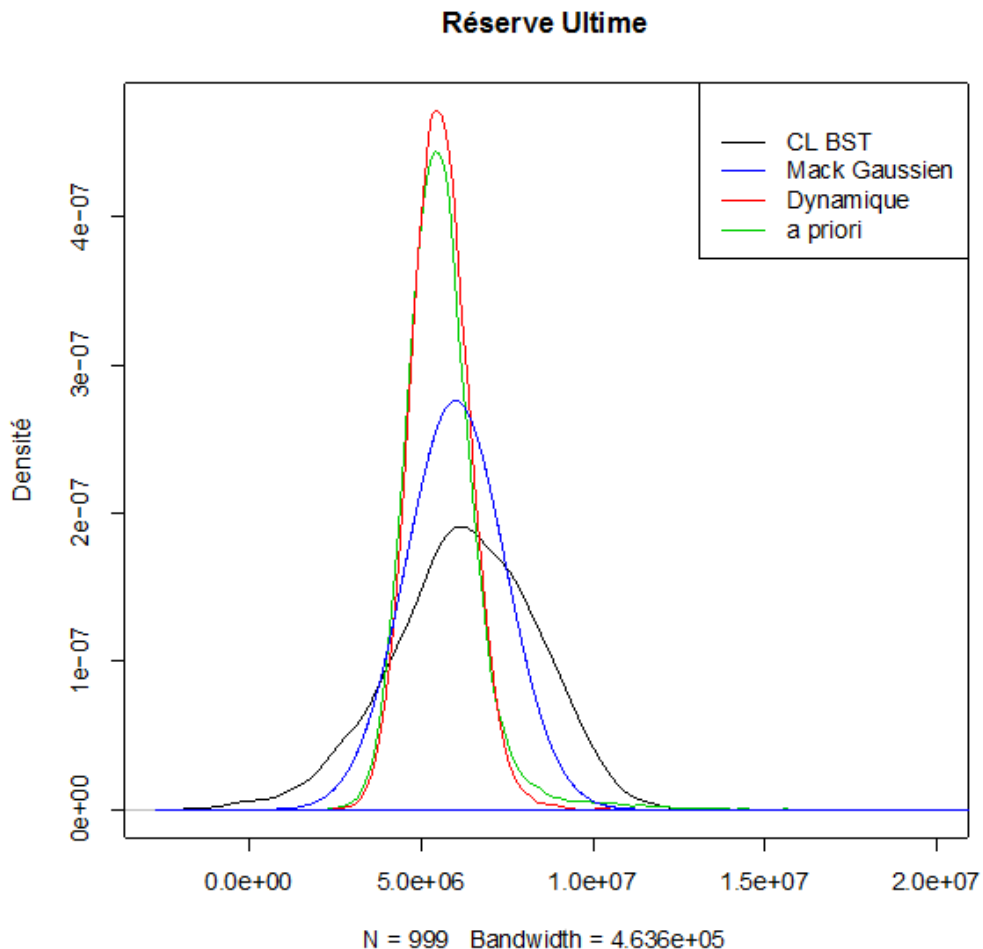


FIGURE 6.21 – Comparaison des distributions

1. Le triangle est d'abord projeté par Chain-Ladder pour calculer les résidus de Pearson
 2. <http://cran.r-project.org/web/packages/ChainLadder/>

Si l'on s'intéresse à la moyenne et à la queue de distribution à droite, on a les résultats suivants.

	Moyenne	0.5	0.9	0.95	0.99	0.995
<i>A priori</i>	5 631 728	5 502 619	6 812 350	7 404 465	10 057 173	11 333 072
Dynamique	5 565 267	5 535 480	6 662 595	6 998 702	7 775 511	8 149 482
Mack Gaussien	5 996 216	5 996 216	7 848 024	8 372 985	9 357 726	9 718 220
CL BST	11 234 752	10 427 718	14 128 719	16 288 674	25 004 351	34 482 837

TABLE 6.12 – Quantiles des distributions prédictives des provisions selon les différentes méthodes

La moyenne avec la méthode de Mack est bien entendu la même que Chain-Ladder. La méthode CL BST semble sur-estimer fortement les réserves. Nous voyons également que la queue de distribution avec la méthode de segmentation *a priori* est au-dessus de celle de Mack à partir du quantile à 99%, tandis que celle de la méthode de segmentation dynamique reste en-dessous. Nous sommes néanmoins dans des ordres de grandeurs comparables.

Bilan d'étape

Dans cette première sous-partie, nous avons utilisé la base de données non retraitée, en excluant les charges inférieures à 10. Nous nous sommes restreint à l'estimation des flux futurs liés aux non-tardifs. L'idée était de pouvoir étudier le modèle théorique sur une base de données réelles. Les estimateurs CL et ML étaient proches. Grâce à des méthodes de segmentation, nous avons trouvé des distributions des provisions qui sont cohérentes avec les méthodes stochastiques agrégées.

Cependant, cette approche de validation du modèle théorique souffre deux défauts majeurs :

- En excluant les cas de charges nulles, nous biaisons l'estimation des réserves : nous ne modélisons en effet pas pour un sinistre non clos la possibilité de connaître un recours ou de finir sans suite.
- Nous n'avons projeté que les sinistres non-tardifs. Pour que le modèle soit utilisable en pratique, nous devons estimer la provision totale.

Nous allons essayer de contourner ces deux défauts dans la partie suivante.

6.12 Base retraitée

Dans cette partie, nous allons réintégrer les cas exclus dans la base non traitée, c'est-à-dire les charges nulles et les charges négatives. Dans les procédures bootstrap, les facteurs de développement tirés pourront être égal à 0, pour modéliser les cas de recours complet et de sinistres sans suite. D'abord, nous présentons comme nous avons retraité la base pour prendre en compte tous les sinistres.

6.12.1 Retraitements effectués

6.12.1.1 Charges négatives

Les cas de facteur de développement négatif risquent de poser problème pour la prévision des réserves. En effet, si un sinistre ayant une charge de 100€ passe l'année suivante à une charge de -10€, cela conduit à un facteur de développement de $-1/10$. Lors de l'application du bootstrap, en tirant ce facteur de développement et en l'appliquant à des sinistres de plus gros montants, nous pouvons sous-estimer fortement les provisions. Notamment, pour les recours, on peut imaginer qu'il y a un effet de seuil en raison de la subrogation de l'assureur, et il n'est pas très réaliste qu'une charge passe de 100000€ à -10000€ (même avec les intérêts perçus). Nous les avons transformé les charges négatives et les paiements négatifs à zéro, quitte à ensuite « translater » les réserves des montants négatifs retirés par la suite.

6.12.1.2 Charges nulles

On peut distinguer trois cas :

- Traitement des nil claims « purs » (c.-à-d. = 0 tout au long de leur vie). On garde juste la première charge à zéro, afin que les autres zéros soient exclus de la base où on tire les facteurs de développement pour le bootstrap (pour ne pas avoir de problème de facteurs égal à 0/0). La date de clôture a été modifiée en conséquent.
- Traitement des sinistres d'abord estimé comme *nil claims* puis ayant des charges positives par la suite. On considère que ce sont des IBNR. Nous avons considéré que leur année de notification était celle où la première charge était positive.
- Traitement des zéros en « cours de vie » du sinistre. Seul le premier zéro a été considéré, toujours pour ne pas avoir deux charges nulles à la suite. S'il y avait à nouveau des charges positives ensuite, elles ont été traitées comme un cas de réouverture.

Notons une première limite de ce retraitement : si l'on observe une première charge nulle en 2012 pour un sinistre à provisionner, on ne sait pas s'il s'agit d'un tardif ou d'un sans suite. Nous allons maintenant exposer la modélisation des tardifs que nous avons retenu :

6.12.2 Modélisation des tardifs

Pour modéliser les tardifs, nous allons procéder de la manière suivante :

- **Sinistres tardifs observés non clos** : nous supposons que le **développement des tardifs est le même que le développement des non-tardifs**. Nous faisons donc les mêmes hypothèses théoriques. Le seul changement à faire dans le modèle théorique est de remplacer l'année de survenance par l'année de notification du sinistre. Nous pouvons ensuite utiliser les procédures bootstrap comme dans la partie précédente.
- **Sinistres tardifs non observés (TNO)**
 1. Nous estimons le **nombre de tardifs** à venir en utilisant la méthode Chain-Ladder sur le nombre de sinistres cumulés. Murphy et McLennan [35] proposent également possible des lois de poissons pour simuler la fréquence de sinistres tardifs, en prenant en compte des données exogènes (exposition - c.-à-d. nombre de polices dans le portefeuille, montant de la prime, etc.).
 2. Estimation des paiements futurs générés par ce tardif. Une première méthode simple est de considérer un coût moyen à l'ultime (CM) (par exemple, par année de survenance), à partir des sinistres observés non clos simulés. Une méthode plus sophistiquée est de tirer une vie pour les sinistres tardifs parmi les sinistres observés clos et les observés non-clos simulés jusqu'à clôture. Il pourrait s'agir : d'un bootstrap « complet » (tirage d'un sinistre entier), ou d'un bootstrap décomposé en plusieurs étapes (tirage d'une charge initiale, d'une durée de vie, de facteurs de développement).

Dans ce mémoire, en première approche, nous avons retenu une méthode coût moyen à l'ultime pour les sinistres tardifs non observés. Remarquons que dans la base de données dommages matériels, il y a peu de sinistres tardifs observés, et donc une estimation faible des sinistres tardifs non observés, comme le montre la sous-section suivante.

6.12.3 Nombre de tardifs

Le triangle complété par Chain-Ladder du nombre cumulé de sinistres est le suivant (nous retrouvons en ligne les années de survenance et en colonne les années de développement) :

	0	1	2	3	4	5	6	7	8	9	10
2002	3 806	431	11	12	3	3	2	0	0	0	0
2003	3 490	71	16	8	1	1	0	2	0	0	0
2004	2 686	53	4	0	0	0	0	0	0	0	0
2005	2 644	61	0	1	0	0	0	0	0	0	0
2006	3 403	180	1	2	0	0	0	0.54	0	0	0
2007	4 118	230	1	0	0	0	0.52	0.65	0	0	0
2008	4 305	266	2	0	1	0.86	0.54	0.69	0	0	0
2009	4 444	211	3	2	0.90	0.88	0.55	0.70	0	0	0
2010	3 762	177	3	3.24	0.76	0.74	0.47	0.59	0	0	0
2011	3 450	221	4.38	3.02	0.71	0.69	0.44	0.55	0	0	0
2012	3 276	172.47	4.12	2.84	0.67	0.65	0.41	0.52	0	0	0

Finalement, il y a 39 384 non-tardifs (première colonne du triangle). On observe 1980 tardifs. L'estimation Chain-Ladder conduit à estimer 204 sinistres tardifs non observés.

6.12.4 Résultats

Les résultats des différentes méthodes appliquées à la base retraitée (tirage de zéro autorisé) sont les suivants :

	Réserve TNO (CM)	Réserve ultime totale	dCL
CL	-	3 996 552	0.00
ML	1 260 812	6 282 916	0.57
Moyenne BST	1 311 021	7 528 348	0.88
Moyenne BST APR	1 241 924	5 814 381	0.45
Moyenne BST SEG	1 260 206	6 267 862	0.57

De manière assez contre-intuitive, en intégrant la possibilité de recours complet de sans suite, nous voyons que les différentes méthodes proposées donnent toutes des réserves ultimes bien plus élevées que Chain-Ladder. En fait, ce premier retraitement proposé souffre de plusieurs limites.

6.12.5 Limites

La procédure d'estimation souffre de plusieurs limitations.

D'abord, l'hypothèse faites sur les tardifs observés ne semblent pas vérifiée. La table 6.13 montre que les tardifs observés n'ont pas le même développement que les non-tardifs : nous observons des statistiques différentes sur les charges et les durées de vie selon le délai de notification. Les durées de vie moyennes des tardifs observés sont par exemple plus faibles.

Délai de notification	0	1	2	≥ 3
Nbre	347 34	1 698	38	35
Moyenne 1ère Charge	4 778	4 923	5 352	2 545
Moyenne dernière Charge	3 958	4 416	5 167	2 272
Moyenne âge	2.37	1.90	1.82	1.34
Moyenne 1 ^{er} facteur	0.99	2.28	1.12	0.80

TABLE 6.13 – Statistiques descriptives des non-tardifs (délai de notification = 0) et des tardifs

De plus, les méthodes de segmentation sont moins performantes et discriminantes lorsque l'on autorise des facteurs de développement négatifs. En effet, à la fois les sinistres graves ou les sinistres attritionnels peuvent être affectés par des recours complets. La forme en L des distributions empiriques de charges et de facteurs ressort beaucoup moins pour les années de développement supérieures à 1, comme le montre la figure 6.22.

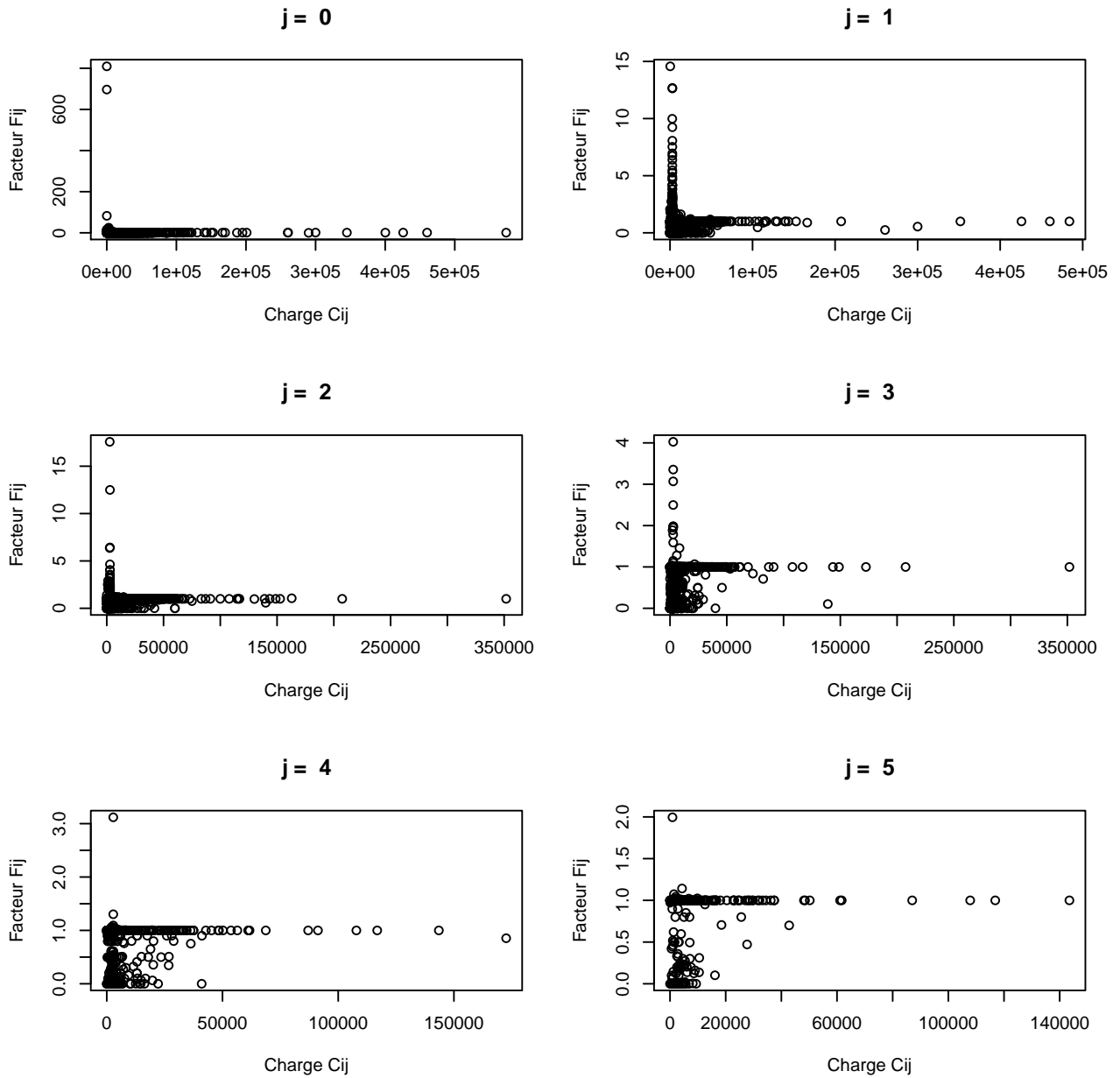


FIGURE 6.22 – Charges vs Facteurs pour l'ensemble des sinistres

Enfin, la corrélation entre les facteurs et les durées de vie est beaucoup plus forte. En particulier, dans la procédure bootstrap, la vie d'un sinistre s'arrête s'il tire un facteur 0. Cette dépendance remet en cause l'hypothèse d'indépendance entre le processus d'âge et le processus de paiements.

6.12.6 Pistes d'amélioration

La piste que nous avons suivie pour prendre en compte toutes les données et estimer la réserve globale n'est pas complètement satisfaisante. Les hypothèses du modèle sont trop invalidées lorsqu'on autorise le tirage de facteurs de développement nuls. Pour le futur, nous proposons une autre piste qui pourrait s'avérer plus satisfaisante : il s'agirait de continuer à travailler avec des facteurs de développement strictement positif pour la procédure bootstrap. Le cas des *nil claims* et des recours complets seraient traités à part de la procédure bootstrap, par des méthodes idoines (p. ex. on estimerait la probabilité pour un sinistre non clos d'avoir une charge finale nulle, sur la base des sinistres clos observés). Pour les sinistres tardifs non observés, il serait possible de tirer une première charge initiale et de leur simuler ensuite une vie future par bootstrap. Les différents blocs de ce modèle sont détaillés sur la figure 6.23. Le modèle théorique permettant de bien représenter les liens entre les sous-modules et permettant d'établir les hypothèses sous-jacentes reste à construire.

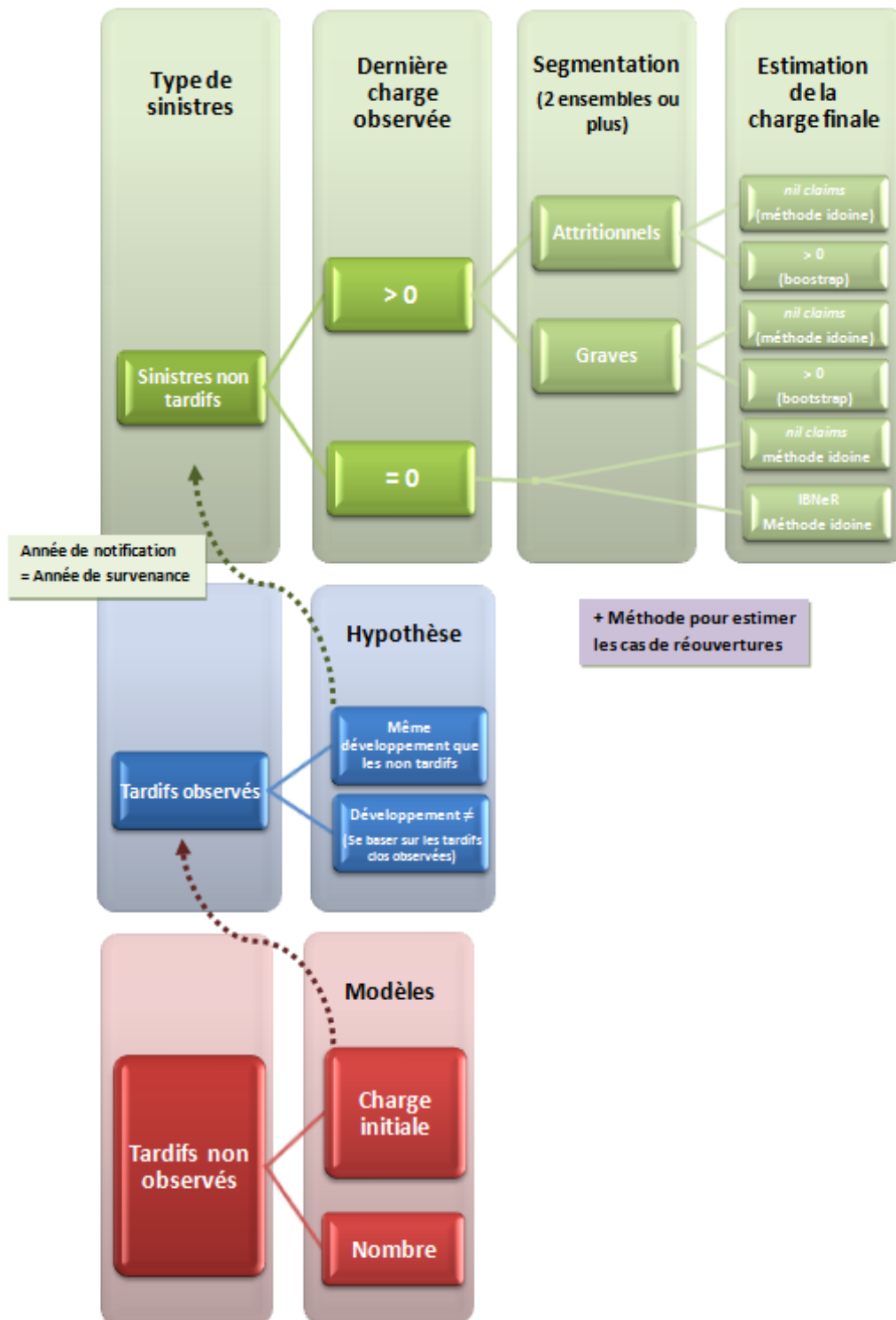


FIGURE 6.23 – Vers une modélisation de tous les cas de figures possibles dans les données ligne à ligne...

Conclusion

Dans ce mémoire, nous avons proposé un modèle ligne à ligne pour évaluer les provisions techniques reposant sur les facteurs individuels de développement. L'idée d'utiliser ces facteurs individuels a déjà été proposée dans l'article de Murphy et McLennan (2006) [35]. Nos avancées par rapport à cet article sont :

- Le remplacement de l'étude des facteurs individuels dans un cadre théorique, celui développé par Chau et Dinh [7]. Nous avons dans ce cadre établi un modèle et formulé des hypothèses qui permettent d'obtenir les propriétés des facteurs individuels de développement supposées par Murphy et McLennan (notamment indépendance entre eux et indépendance avec le passé).
- Nous avons aussi établi une comparaison entre le modèle individuel et le modèle Chain-Ladder. Nous avons énoncé des hypothèses sous lesquelles les estimateurs de la méthode ligne à ligne et les estimateurs Chain-Ladder sont asymptotiquement équivalents. A distance finie, l'étude de simulations montre que l'on peut obtenir des estimateurs proches avec un nombre raisonnable de sinistres.
- Enfin, nous avons fait des apports concernant la segmentation des sinistres. Dans l'étude de Murphy et McLennan, les sinistres ne pouvaient pas changer de catégorie.¹ dans les deux sens au cours de leur vie. La méthode de segmentation dynamique proposée permet de dépasser cette limite.

Au niveau de l'application à des données réelles, le bilan est le suivant :

- Lorsque nous n'avons utilisé qu'une partie des données pour se replacer dans le cadre du modèle théorique, les résultats se sont révélés convenables : l'estimation de la moyenne des réserves par la méthode ligne à ligne et par Chain-Ladder étaient proches. Les distributions prédictives des réserves obtenues par bootstrap, au niveau ligne à ligne sont comparables avec celles obtenues avec les méthodes agrégées.
- Lorsque nous avons réalisé une estimation globale des réserves, pour une utilisation *in vivo* du modèle, nous avons mis en évidence l'existence de certaines limitations. Le cas de charges nulles est mal supporté par le modèle. Cependant, des pistes pour modéliser de manière séparée les charges nulles, ou encore retraiter d'une manière plus habile les données sont certainement envisageables. Par exemple, le problème pourrait être décomposé de la manière suivante
 - Pour les charges nulles correspondant à des **recours complets**, le travail était effectué brut de recours. Ici, nous n'avons pas accès à la nature des montants, et ne pouvons pas ainsi identifier les recours. Si les recours sont traités à part, il faudra cependant alors modéliser les recours (le *Best Estimate* dans Solvabilité II doit être calculé net de recours).

1. Les auteurs considéraient deux catégories de sinistres graves.

CONCLUSION

- Pour les charges nulles dans la base dues à des réserves constituées puis déconstituées (**sans-suite**), une probabilité qu'un sinistre ouvert aujourd'hui termine soit un sans suite était estimée.
- Pour les cas de **premières charges nulles**, une probabilité d'être un tardif (au sens large) ou un sans-suite était estimée.

Un cadre plus complet pourrait être construit pour modéliser les interactions entre chaque sous-ensemble du problème.

Le travail souligne le besoin pour les compagnies de mettre en place des protocoles de collecte des données de plus en plus complets, si elles veulent effectuer d'autres études de ce type dans le futur. La politique de classement des dossiers doit aussi être auditée au préalable : ne pas clore un dossier pour lequel l'assuré ne s'est pas manifesté dans le délai réglementaire (de deux ans en France) pour faire valoir son droit à l'indemnisation contractuelle n'a pas d'incidence sur l'estimation des provisions dossier-dossier, mais a des conséquences dans notre modélisation (estimation de la durée de vie des sinistres, estimation des facteurs de développement).

La mise au point d'un modèle ligne à ligne parfaitement opérationnel pourra permettre de nombreuses applications au centre des problématiques actuarielles actuelles, par exemple :

- une application à la réassurance non proportionnelle, permettant d'étudier la variabilité des réserves brutes et nettes de réassurance. Une segmentation *a priori* de la base à partir des priorités des différents traités de réassurance pourrait être utilisée.
- Une application au calcul du risque de réserve à un an, dans le cadre Solvabilité II. La procédure bootstrap permet en effet de simuler des trajectoires de changements de charge à un an. Un *Best Estimate* « dans un an » sachant la nouvelle information (les réalisations de charge dans un an) peut alors être calculé et comparé au *Best Estimate* actuel.

Bibliographie

- [1] Katrien ANTONIO et Richard PLAT : Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, 2012. Accepted, to be published.
- [2] Michel BENAÏM et Nicole EL KAROUI : *Promenade aléatoire - Chaînes de Markov et simulations ; martingales et stratégies*. Les Éditions de l'École Polytechnique, 2005.
- [3] Guillaume BENETEAU : *Modèle de provisionnement sur données détaillées en Assurance Non-Vie*. Mémoire IA, ENSAE, 2004.
- [4] Anne BONTOUX : *Les rentes en assurance et en réassurance non-vie : ce qu'il faut savoir*. Swiss Re, 2007. Technical publishing Casulaty.
- [5] CEIOPS : *QIS5 Technical Specifications*, Juillet 2010.
- [6] Arthur CHARPENTIER et Michel DENUIT : *Mathématiques de l'assurance non-vie : Tome 2, Tarification et provisionnement*. Economie et statistiques avancées. Economica, 2005.
- [7] Gilles CHAU et Ngoc An DINH : *Mesures de Provision Cohérentes et Méthodes Lignes à Lignes pour des Risques Non-Vie*. Mémoire IA, ENSAE, 2012.
- [8] Kai Lai CHUNG : *A Course in Probability Theory*. Academic Press, 2000.
- [9] Hélène COMPAIN : *Analyse du risque de provisionnement non-vie dans le cadre de la réforme Solvabilité II*. Mémoire IA, Paris Dauphine, 2010.
- [10] Ministère de la JUSTICE : Loi relative à la prescription en matière civile. <http://www.textes.justice.gouv.fr/lois-et-ordonnances-10180/loi-relative-a-la-prescription-en-matiere-civile-15569.html>.
- [11] Jens M. DITTMER : Nearest-Neighbour Methods for Reserving with respect to Individual Losses. *Blätter der DGVMF*, 27(4):647–664, 2006.
- [12] EIOPA : *Technical Specifications for the Solvency II valuation and Solvency Capital Requirements calculations (Part I)*, Octobre 2012.
- [13] P.D. ENGLAND et R.J. VERRALL : Stochastic Claims Reserving In General Insurance. *In Institute of Actuaries*, 2002.
- [14] Jean-David FERMANIAN : *Modèles de durée*. Cours 3A, ENSAE, 2010 - 2011. http://www.crest.fr/ckfinder/userfiles/files/Pageperso/fermania/JDF_duree3.pdf.
- [15] FFSA-GEMA : Les assurances de biens et de responsabilité : Données clés 2011, 2012.
- [16] Thomas R. FLEMING et David P. HARRINGTON : *Counting processes and survival analysis* . Wiley Series in Probability and Statistics. Wiley, 2005.
- [17] Morgane FORT : *Méthodes de provisionnement en assurance non-vie et extrapolation des triangles*. Mémoire IA, EURIA, 2010.
- [18] Denis FOUGÈRE : *Duration Model*. Cours 3A, ENSAE, 2012 - 2013. <http://www.crest.fr/ses.php?user=2944>.

- [19] Christian FRANCO : *Mesures de risque*. Cours 3A, ENSAE, 2012 - 2013. <http://perso.univ-lille3.fr/~cfranco/Christian-Franco/mesures-de-risque.htm>.
- [20] Richard D. GILL : Censoring and stochastic integrals. *Statistica Neerlandica*, 34(2):124, 1980.
- [21] Svend HAARSTRUP et Elja ARJAS : Claims Reserving in Continuous Time : A Nonparametric Bayesian Approach. *Astin Bulletin*, 26(2):139–164, 1999.
- [22] Ilan HABIB et Stéphane RIBAN : *Quelle méthode de provisionnement pour des engagements non-vie dans Solvabilité 2 ?* Mémoire IA, ENSAE, 2012.
- [23] Donatien HAINAUT : *Econométrie de l'assurance*. Cours 3A, ENSAE, 2012 - 2013.
- [24] Guillaume HORNY : *Econométrie des modèles de durée*. Master 2 MOSEF, Paris 1, 2009 - 2010. <http://horny.economics.free.fr/PI/chap2.pdf>.
- [25] Catherine HUBERT et Philippe GUAY : *Mesure d'incertitude sur le provisionnement IARD : modélisation des branches longues et traitement des valeurs extrêmes*. Mémoire IA, CEA, 2009.
- [26] IASB : *Exposure Draft ED/2010/8 Insurance Contracts*, Juillet 2010. http://www.ifrs.org/Current-Projects/IASB-Projects/Insurance-Contracts/Exposure-draft-2010/Documents/ED_Insurance_Contracts_Standard_WEB.pdf.
- [27] IASB : *Effect of board redeliberations on ED Insurance Contracts*, Février 2013. <http://www.ifrs.org/Current-Projects/IASB-Projects/Insurance-Contracts/Documents/Status/Effect-Board-decisions-on-ED-February-2013.pdf>.
- [28] Magali KELLE : Provisionnement pour Sinistres à Payer : Analyses et Modélisations sur Données Détaillées. *Bulletin français d'actuariat*, 7(13):122–142, 2007.
- [29] Christian Roholte LARSEN : An individual claims reserving model. *Astin Bulletin*, 37(1): 113–132, 2007.
- [30] Frank LE VALLOIS : *Réglementation et assurance*. Cours 3A, ENSAE, 2012 - 2013.
- [31] Thomas MACK : Distribution-free calculation of the standard error of chain ladder reserve estimate. *Astin Bulletin*, 23(2), 1993.
- [32] Thomas MACK : The standard error of chain ladder reserve estimate. *Astin Bulletin*, 29(2), 1999.
- [33] John B. MAHON : Transition Matrix Theory And Individual Claim Loss Development. *CAS Forum Spring 2005*, pages 115–170, 2005.
- [34] Clémence MICHAUD : *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*. Mémoire IA, ISFA, 2012.
- [35] Karl MURPHY et Andrew MCLENNAN : A Method For Projecting Individual Large Claims. *In Casualty Actuary Society Forum*, 2006.
- [36] Francis NOËL : *La Gestion des Sinistres IRD*. Les Pédagogiques. SEFI, 2ème édition, 2010.
- [37] Ragnar NORBERG : Prediction of outstanding liabilities in non-life insurance. *Astin Bulletin*, 23(1), 1993.
- [38] Ragnar NORBERG : Prediction of outstanding liabilities : II Model variations and extensions. *Astin Bulletin*, 29(1):5–25, 1999.
- [39] Laurence PACQUIER : *Solvabilité II et IARD : Provisionnement, exigence de solvabilité, modèle interne*. Mémoire IA, ISUP, 2012.

-
- [40] PARLEMENT EUROPÉEN ET CONSEIL DE L'EUROPE : *Directive 2009/138/CE sur l'accès aux activités de l'assurance et de la réassurance et de leur exercice (Solvabilité II)*, Novembre 2009.
- [41] Christian PARTRAT, Eric LECŒUR, Jean-Marie NESSI, Ecaterina NISIPASU et Olivier REIZ : *Provisionnement Technique en Assurance Non Vie : Perspectives actuarielles modernes*. Assurance Audit Actuariat. Economica, 2007.
- [42] Nicolas PATRIGOT : *Nouvelles normes comptables et réglementation financière*. Cours 3A, ENSAE, 2012 - 2013.
- [43] A.E RENSHAW et R.J. VERRALL : A stochastic model underlying the chain-ladder technique. *British Actuarial Journal*, 4(4), 1998.
- [44] Noémie ROSE : *Provisionnement en assurance non-vie : utilisation de modèles paramétriques censurés*. Mémoire IA, ISUP, 2009.
- [45] Stig ROSENLUND : Bootstrapping Individual Claim Histories. *Astin Bulletin*, 42(1):291–324, 2012.
- [46] Sabrina SAVARRE et Benoît PAYRE : *Charge ultime nette de réassurance en RC corporelle : 2 modèles stochastiques pour les flottes automobiles*. Mémoire IA, CEA, 2012.
- [47] Guy SIMONET : *La comptabilité des entreprises d'assurances*. Les Fondamentaux de l'Assurance. L'argus de l'assurance, 5ème édition, 1998.
- [48] STÉPHANIE RUAULT AND MARC JUILLARD AND FRÉDÉRIC PLANCHET : *IFRS/Solvabilité 2 - Calculs des best estimate et marge pour risque : mise en situation*, Novembre 2010.
- [49] Winfried STUTE et Jane-Ling WANG : The jackknife estimate of Kaplan-Meier integral. *Biometrika*, 81(3):602–606, 1994.
- [50] Greg TAYLOR, Grainne MCGUIRE et James SULLIVAN : Individual Claim Loss Reserving Conditioned by Case Estimates. *Annals of Actuarial Science*, 3(1-2):215–256, 2008.
- [51] Alain TOSETTI, Thomas BÉHAR, Michel FROMENTEAU et Stéphane MÉNART : *Assurance : Comptabilité - Réglementation - Actuariat*. Assurance Audit Actuariat. Economica, 2011.
- [52] Xiao Bing ZHAO et Xian ZHOU : Applying copula models to individual claim loss reserving methods. *Insurance : Mathematics and Economics*, 46(2), 2010.
- [53] Xiao Bing ZHAO, Xian ZHOU et Jing Long WANG : Semiparametric model for prediction of individual claim loss reserving. *Insurance : Mathematics and Economics*, 45(1), 2009.

BIBLIOGRAPHIE

Annexe A

Divers

A.1 Exemple de bilans

Bilan de la MAIF

Passif	
	2011
Capitaux propres	1 286 383
fonds d'établissement	278 094
– <i>fonds d'établissement</i>	260 272
– <i>fonds d'établissement - droits d'adhésion</i>	17 822
autres réserves	965 525
– <i>réserve de sécurité</i>	405 869
– <i>réserve pour fonds de développement des filiales et activités nouvelles</i>	428 696
– <i>réserve de capitalisation</i>	130 534
– <i>réserve achat d'œuvres originales d'artistes vivants</i>	426
report à nouveau	– 6 013
résultat de l'exercice	48 777
Provisions techniques	2 998 627
provisions pour primes non acquises (non vie)	3 127
provisions pour sinistres (non vie)	2 471 160
provisions d'égalisation	138 101
autres provisions techniques (non vie)	386 239
Provisions	55 550
Dettes pour dépôts en espèces reçus des cessionnaires	3 987
Autres dettes	1 595 917
dettes nées d'opérations d'assurance directe	805 165
– <i>dettes envers les sociétaires (contrats de mensualisation, cotisations perçues d'avance)</i>	789 134
– <i>autres dettes nées d'opérations d'assurance directe</i>	16 031
dettes nées d'opérations de réassurance	2 642
dettes envers les établissements de crédit	33 441
autres dettes	754 669
– <i>autres emprunts, dépôts et cautionnements reçus</i>	401
– <i>personnel</i>	68 721
– <i>État, organismes de sécurité sociale, collectivités publiques</i>	59 659
– <i>créanciers divers</i>	625 888
Compte de régularisation - Passif	18 050
TOTAL DU PASSIF	5 958 514

FIGURE A.1 – Bilan publié par la MAIF en 2011

Bilan de la MAAF

» PASSIF	
EN MILLIERS D'EUROS	MONTANT NET 2011
▶ Capitaux propres	1 775 141
- Capital social	160 000
- Primes liées au capital social	249 978
- Autres réserves	1 194 083
- Report à nouveau	
RÉSULTAT DE L'EXERCICE	171 080
▶ Provisions techniques brutes	5 506 852
- Provisions pour primes non acquises	57 810
- Provisions pour sinistres	4 943 052
- Provisions pour égalisation	101 648
- Autres provisions techniques	404 342
▶ Provisions pour risques et charges	20 887
▶ Dettes pour dépôts en espèces reçus des cessionnaires	14 263
▶ Autres dettes	599 922
- Dettes nées d'opérations d'assurance directe	228 230
- Dettes nées d'opérations de réassurance	7 783
- Dettes envers des établissements de crédit	109 710
- Autres dettes :	
Autres emprunts, dépôts et cautionnements reçus	190
Personnel	49 544
État, organismes de Sécurité sociale et collectivités publiques	50 437
Créanciers divers	154 028
▶ Comptes de régularisation - passif	28 201
▶ TOTAL DU PASSIF	7 945 266

FIGURE A.2 – Bilan publié par la MAAF Assurances S.A. (contrôlé par la mutuelle MAAF Assurances) en 2011

A.2. CLASSIFICATION OFFICIELLE DES BRANCHES

NATURE	PROVISIONS TECHNIQUES	PART DES CESSIONNAIRES DANS LES PROVISIONS TECHNIQUES
▶ PROVISIONS POUR SINISTRES À PAYER	5 105 820	240 751
▶ PROVISIONS MATHÉMATIQUES	399 445	
▶ PROVISIONS POUR ÉGALISATION	101 648	
▶ PROVISIONS POUR RISQUES EN COURS	4 897	
▶ PROVISIONS DE RECOURS À ENCAISSER	-293 993	
▶ PROVISIONS TECHNIQUES ACCEPTATION DE SINISTRES	131 180	
▶ PROVISIONS SINISTRES CO-ASSURANCES	45	
▶ TOTAL	5 449 042	240 751
▶ PROVISIONS TECHNIQUES ACCEPTATION DE PRIMES	19	
▶ PROVISIONS POUR PRIMES NON ACQUISES	57 791	2 486
▶ TOTAL PROVISIONS TECHNIQUES BRUTES	5 506 852	243 237

FIGURE A.3 – Détails des provisions dans le bilan de la MAAF Assurances S.A. en 2011

A.2 Classification officielle des branches

L'agrément administratif prévu par l'article L. 321-1 est accordé par l'Autorité de contrôle prudentiel. Pour l'octroi de cet agrément, les opérations d'assurance sont classées en branches et sous-branches de la manière suivante :

1. Accidents (y compris les accidents de travail et les maladies professionnelles) : a) Prestations forfaitaires ;

- b) Prestations indemnitaires ;
- c) Combinaisons ;
- d) Personnes transportées.

2. Maladie :

- a) Prestations forfaitaires ;
- b) Prestations indemnitaires ;
- c) Combinaisons.

3. Corps de véhicules terrestres (autres que ferroviaires) :

Tout dommage subi par :

- a) Véhicules terrestres à moteur ;
- b) Véhicules terrestres non automoteurs.

4. Corps de véhicules ferroviaires :

Tout dommage subi par les véhicules ferroviaires.

5. Corps de véhicules aériens :

Tout dommage subi par les véhicules aériens.

6. Corps de véhicules maritimes, lacustres et fluviaux :

Tout dommage subi par :

- a) Véhicules fluviaux ;
- b) Véhicules lacustres ;
- c) Véhicules maritimes.

7. Marchandises transportées (y compris les marchandises, bagages et tous autres biens) :

Tout dommage subi par les marchandises transportées ou bagages, quel que soit le moyen de transport.

8. Incendie et éléments naturels :

Tout dommage subi par les biens (autres que les biens compris dans les branches 3, 4, 5, 6 et 7) lorsqu'il est causé par :

- a) Incendie ;
- b) Explosion ;
- c) Tempête ;
- d) Éléments naturels autres que la tempête ;
- e) Énergie nucléaire ;
- f) Affaissement de terrain.

9. Autres dommages aux biens :

Tout dommage subi par les biens (autres que les biens compris dans les branches 3, 4, 5, 6 et 7) lorsque ce dommage est causé par la grêle ou la gelée, ainsi que par tout événement, tel le vol, autre que ceux compris dans la branche 8.

10. Responsabilité civile véhicules terrestres automoteurs :

Toute responsabilité résultant de l'emploi de véhicules terrestres automoteurs (y compris la responsabilité du transporteur).

11. Responsabilité civile véhicules aériens :

Toute responsabilité résultant de l'emploi de véhicules aériens (y compris la responsabilité du transporteur).

12. Responsabilité civile véhicules maritimes, lacustres et fluviaux :

Toute responsabilité résultant de l'emploi de véhicules fluviaux, lacustres et maritimes (y compris la responsabilité du transporteur).

13. Responsabilité civile générale :

Toute responsabilité autre que celles mentionnées sous les numéros 10, 11 et 12.

14. Crédit :

- a) Insolvabilité générale ;
- b) Crédit à l'exportation ;
- c) Vente à tempérament ;
- d) Crédit hypothécaire ;
- e) Crédit agricole.

15. Caution :

- a) Caution directe ;

b) Caution indirecte.

16. Pertes pécuniaires diverses :

- a) Risques d'emploi ;
- b) Insuffisance de recettes (générale) ;
- c) Mauvais temps ;
- d) Pertes de bénéfices ;
- e) Persistance de frais généraux ;
- f) Dépenses commerciales imprévues ;
- g) Perte de la valeur vénale ;
- h) Pertes de loyers ou de revenus ;
- i) Pertes commerciales indirectes autres que celles mentionnées précédemment ;
- j) Pertes pécuniaires non commerciales ;
- k) Autres pertes pécuniaires.

17. Protection juridique.

18. Assistance :

Assistance aux personnes en difficulté, notamment au cours de déplacements.

20. Vie-Décès :

Toute opération comportant des engagements dont l'exécution dépend de la durée de la vie humaine autre que les activités visées aux branches 22, 23 et 26.

21. Nuptialité-Natalité :

Toute opération ayant pour objet le versement d'un capital en cas de mariage ou de naissance d'enfants.

22. Assurances liées à des fonds d'investissement :

Toutes opérations comportant des engagements dont l'exécution dépend de la durée de la vie humaine et liées à un fonds d'investissement.

Les branches mentionnées aux 20, 21 et 22 comportent la pratique d'assurances complémentaires au risque principal, notamment celles ayant pour objet des garanties en cas de décès accidentel ou d'invalidité.

23. Opérations tontinières :

Toutes opérations comportant la constitution d'associations réunissant des adhérents en vue de capitaliser en commun leurs cotisations et de répartir l'avoir ainsi constitué soit entre les survivants, soit entre les ayants droit des décédés.

24. Capitalisation :

Toute opération d'appel à l'épargne en vue de la capitalisation et comportant, en échange de versements uniques ou périodiques, directs ou indirects, des engagements déterminés quant à leur durée et à leur montant.

25. Gestion de fonds collectifs :

Toute opération consistant à gérer les placements et notamment les actifs représentatifs des réserves d'entreprises autres que celles mentionnées à l'article L. 310-1 et qui fournissent des prestations en cas de vie, en cas de décès ou en cas de cessation ou de réduction d'activités.

A.3 Estimation non-paramétrique

Nous rappelons rapidement les principes de la méthode d'estimation non paramétrique par Kernel (ou noyau ou méthode de Parzen-Rozenblatt). Pour un échantillon d'observations $x_i, i = 1, \dots, N$ donné, la fonction de densité f estimée au point x_0 à l'aide de cette méthode est :

$$\hat{f}(x_0) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x_i - x_0}{h}\right),$$

où N est le nombre total d'observations, h un paramètre de largeur de bande, et $K(\cdot)$ la fonction Kernel.

L'estimateur de $\hat{f}(x)$ est généralement biaisé. Le biais au point x_0 vaut :

$$b(x_0) = \mathbb{E}[\hat{f}(x_0)] - f(x_0) = \frac{1}{2}h^2 f'' \int z^2 K(z) dz$$

Le paramètre de largeur de bande h permet de faire un arbitrage entre le biais (h petit) et le lissage de la densité estimée (h grand). Le critère généralement retenu et implémenté dans les logiciels pour choisir un h optimal est la minimisation la MISE, l'erreur quadratique moyenne intégrée :

$$MISE(h) = \int MSE[\hat{f}(x)] dx,$$

où la MSE désigne l'erreur quadratique moyenne (c.-à-d. la somme de la variance de $\hat{f}(x)$ et du carré du biais de $\hat{f}(x)$).

Il existe plusieurs choix de fonctions Kernel (voir tableau A.4).

Kernel	Kernel Function $K(z)$
Uniform (or box or rectangular)	$\frac{1}{2} \times \mathbf{1}(z < 1)$
Triangular (or triangle)	$(1 - z) \times \mathbf{1}(z < 1)$
Epanechnikov (or quadratic)	$\frac{3}{4}(1 - z^2) \times \mathbf{1}(z < 1)$
Quartic (or biweight)	$\frac{15}{16}(1 - z^2)^2 \times \mathbf{1}(z < 1)$
Triweight	$\frac{35}{32}(1 - z^2)^3 \times \mathbf{1}(z < 1)$
Tricubic	$\frac{70}{81}(1 - z ^3)^3 \times \mathbf{1}(z < 1)$
Gaussian (or normal)	$(2\pi)^{-1/2} \exp(-z^2/2)$
Fourth-order Gaussian	$\frac{1}{2}(3 - z)^2(2\pi)^{-1/2} \exp(-z^2/2)$
Fourth-order quartic	$\frac{15}{32}(3 - 10z^2 + 7z^4) \times \mathbf{1}(z < 1)$

FIGURE A.4 – Fonctions Kernel les plus courantes