Mémoire présenté le : 6 septembre 2017

pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA et l'admission à l'Institut des Actuaires

Par : Vincent BONANNO

Titre : Modèle markovien de provisionnement des sinistres en RC automobile

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

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus.

Membres présents du jury de l'IA — *Signature*

Mme Catherine PIGEON

M. Frédéric SCHWACH

Membres présents du jury de l'ISFA

Mme Véronique MAUME-DESCHAMPS

M. Didier RULLIERE

Entreprise

Nom : MACSF

Signature :

Directeur de mémoire en entreprise

Nom : Jérôme DANNEQUIN

Signature :

Invité

Nom :

Signature :

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

Signature du responsable entreprise

Signature du candidat

Secrétariat :
Mme Christine DRIGUZZI

Bibliothèque :
Mme Patricia BARTOLO

Résumé

Mots clés : *Provisionnement, Chain Ladder, Chaîne de Markov, Matrice de transition, Classification, Segmentation, Théorie des valeurs extrêmes, Régression logistique, Loi log-normale*

L'hétérogénéité des sinistres corporels peut s'avérer être une source d'incertitude non négligeable en assurance automobile. C'est pourquoi ce mémoire propose un modèle adapté dont l'intérêt est notamment de provisionner plus précisément les sinistres les plus graves. Une méthode markovienne permettant d'y parvenir ayant déjà été développée a servi de base afin de construire ce nouveau modèle. Une segmentation des sinistres selon leur tranche de coût a dans un premier temps été définie selon un raisonnement basique, mais cette classification a été revue par la suite à l'aide d'une étude plus approfondie. Cela conduit logiquement à l'utilisation de la théorie des chaînes de Markov afin d'estimer, en fonction du stock de sinistres à provisionner à la clôture 2015, le nombre de sinistres dans chaque classe à la fin de leur liquidation. En poursuivant l'étude, il nous a paru pertinent de faire dépendre les classes non plus seulement de la charge des sinistres mais également de deux autres variables qualitatives : le nombre de pièces de gestion qui ont été reçues par la compagnie et l'état dans lequel les gestionnaires ont placé le sinistre (en cours ou terminé). Cette nouvelle classification a permis aux probabilités de transition de ne plus dépendre de la maturité des sinistres et donc d'évaluer les classes ultime de notre flotte de sinistres à un horizon qui n'est plus borné dans le temps. L'évaluation de la charge ultime a enfin été réalisée de deux manières : en utilisant simplement les coûts moyens des sinistres dans chaque classe ou en effectuant des simulations reposant sur l'hypothèse que les charges des sinistres suivent une loi log-normale. Dans l'application finale, la charge ultime a été estimée à la fois brute et nette de réassurance, ce qui permet de combler une limite des méthodes classiques.

Abstract

Key words : *Claims reserve, Chain Ladder, Markov chain, Transition Matrix, Classification, Segmentation, Extreme value theory, Logistic regression, Log-normal distribution*

One of the sources of uncertainty in motor vehicle insurance is the heterogeneity of corporal claims, therefore this work aims to provide a claims reserving model whose goal is to improve the precision of the reserves of the most serious claims. John B. Mahon already developed a Markov model taking into account that disparity which is our starting point to build our new model. First, intervals of size of loss are defined in order to classify claims depending on how serious they are. That classification has been carried out by two ways : on the one hand a not extremely rigorous but much simpler one, on the other hand by conducting a more precise study. The Transition Matrix approach allows us to predict how many claims are supposed to be in each state when they are all closed. It has occurred to be relevant that the size of loss may not be the only variable for the states to depend on. New states are defined as the combination of the size of loss, the number of documents received by the insurance company and whether claims are open or closed. That new way to classify claims allows the model not to depend of the claims maturities, so the transition probabilities of the Markov process are no longer time-dependent. It is therefore possible for the ultimate states to be evaluated without being bounded by any number of maturity years. The ultimate loss development has yet to be estimated. That can be made either by using the average loss of the claims in each state, or by using a simulation approach based on the assumption that the losses are log-normally distributed. It is to be pointed out that the final model predicts the reserve either net or gross of reinsurance, which means that the model approach the amount of losses paid by the company in a better way that non-life reserving methods usually do.

Remerciements

Je tiens tout d'abord à remercier les équipes de la MACSF de m'avoir acceuilli au sein de la Direction Technique et de m'avoir permis d'y réaliser mon mémoire.

Des remerciements particuliers s'adressent à mon tuteur de mémoire M. Jérôme Dannequin, responsable du service Actuariat IARD, pour son accompagnement et pour les précieux conseils qu'il a pu me fournir afin de mener à bien cette étude.

Je souhaite également remercier M. Pierre Thérond pour avoir encadré mon mémoire en tant que tuteur académique.

Je remercie enfin ma famille et mes amis pour leur soutien.

Table des matières

Introduction **9**

1 Méthodes classiques de provisionnement non-vie **11**
1.1 Triangles de liquidation 11
1.2 Méthodes déterministes 13
1.2.1 Méthode Chain Ladder 13
1.2.2 Vérification des hypothèses du modèle 14
1.3 Méthodes stochastiques 16
1.3.1 Erreur de prédiction 16
1.3.2 Modèle de Mack 16
1.3.3 Vérification des hypothèses 17
1.4 Application 18
1.4.1 Vérification des hypothèses 18
1.4.2 Résultats obtenus 20

2 Chaînes de Markov **23**
2.1 Premières définitions 23
2.2 Matrice de transition 24
2.3 Propriétés des chaînes de Markov 25
2.3.1 États 25
2.3.2 Périodicité 25
2.3.3 Récurrence et transience 25
2.3.4 Probabilité stationnaire 26

3 Méthode des matrices de transition **27**
3.1 Présentation des données 27
3.1.1 Retraitement des montants en euros constants 27
3.1.2 Description des sinistres 28
3.2 Présentation de la méthode 30
3.2.1 Classes 31
3.2.2 Hypothèses du modèle 32
3.3 Construction des matrices 32
3.3.1 Matrices de transition en nombres de sinistres 33
3.3.2 Matrices de transition stochastiques 33
3.3.3 Matrices de transition à l'ultime 34
3.4 Projection des sinistres 35
3.4.1 Objectif de la méthode 35
3.4.2 Sinistres tardifs 35

3.4.3 Projection du nombre de sinistres dans chaque classe . . . 36
3.4.4 Évaluation de la charge ultime . . . 36

4 Première amélioration de la méthode 40
4.1 Préliminaires . . . 40
4.2 Présentation du modèle . . . 41
4.2.1 Hypothèses . . . 42
4.2.2 Lissage des moments . . . 43
4.2.3 Estimation de la tendance . . . 47
4.2.4 Construction des nouvelles matrices . . . 47
4.2.5 Alternative à l'hypothèse de log-normalité . . . 48
4.3 Application et défauts de la méthode . . . 51
4.3.1 Résultats . . . 51
4.3.2 Défauts de la méthode . . . 53

5 Deuxième amélioration de la méthode 54
5.1 Nouvelles variables . . . 55
5.1.1 Position du sinistre . . . 55
5.1.2 Variable d'état . . . 55
5.2 Changement des seuils de charge . . . 57
5.2.1 Premiers seuils arbitraires . . . 57
5.2.2 Première approche de détermination de seuils : méthode du boni mali . . . 57
5.2.3 Méthodes basées sur la théorie des valeurs extrêmes . . . 58
5.2.4 Application des méthodes . . . 62
5.2.5 Conclusion : nouvelle segmentation des charges . . . 68
5.3 Nouvelles classes . . . 68
5.4 Validation des classes . . . 69
5.4.1 Sinistres clos . . . 69
5.4.2 Régression logistique . . . 69
5.5 Chaîne de Markov . . . 72
5.5.1 Construction de la matrice . . . 72
5.5.2 Propriétés de la chaîne de Markov . . . 73
5.5.3 Première projection : approche déterministe . . . 74
5.5.4 Deuxième projection : approche stochastique . . . 76

6 Évaluation de la charge ultime 78
6.1 Approche déterministe basée sur les coûts moyens . . . 78
6.1.1 Description de la méthode . . . 78
6.1.2 Résulats obtenus . . . 79
6.2 Approche stochastique . . . 80
6.2.1 Présentation de l'approche . . . 81
6.2.2 Choix de la loi simulée . . . 81
6.2.3 Simulations et résultats obtenus . . . 84

Conclusion 86

Bibliographie 88

A Applications des méthodes de provisionnement classiques **89**
A.1 Triangles 89
A.1.1 Triangle des règlements 89
A.1.2 Triangle des charges 90
A.2 Validation des hypothèses 91
A.2.1 Hypothèse H2 91
A.2.2 Hypothèse H3 93

B Résultats de la première proposition d'amélioration **95**
B.1 Résultats de l'algorithme de Levenberg-Marquardt 95
B.1.1 Résultats obtenus pour les fonctions (f_k) 96
B.1.2 Résultats obtenus pour les fonctions (g_k) 96
B.2 Graphes du skewness et du kurtosis 97
B.2.1 Skewness 97
B.2.2 Kurtosis 98

C Régression logistique **99**

D Matrices **102**

E Convergences des simulations **106**

Introduction

L'évaluation des engagements qu'une compagnie d'assurance a envers ses assurés est une des problématiques majeures du secteur. Il existe de nombreuses méthodes de provisionnement connues mais celles qui sont les plus répandues ne sont pas forcément les plus efficaces et les assureurs ont toujours besoin de s'interroger sur la pertinence des modèles et leurs améliorations possibles. Nous nous intéresserons dans ce mémoire à une étude concernant l'évaluation des provisions pour sinistres à payer de la MACSF sur la garantie "Responsabilité Civile Corporelle" pour les sinistres automobiles.

La garantie RC corporelle en assurance auto a la particularité de présenter une certaine hétérogénéité dans ses sinistres. On trouve d'une part une grande majorité de sinistres dits attritionnels ou non graves qui ont une liquidation assez rapide. Mais il existe des sinistres bien plus importants en montant malgré le fait qu'ils soient beaucoup plus rares en fréquence. Il peut donc être judicieux de prendre en compte cette hétérogénéité dans le calcul des provisions afin d'éviter le sous-provisionnement des sinistres graves, ce que la MACSF a décidé de mettre en œuvre.

Plutôt que de procéder à l'application de méthodes de provisionnement classiques en segmentant les sinistres en plusieurs tranches de coût, ce qui correspond à une approche assez répandue sur le marché, John B. Mahon a eu l'idée de prendre en compte cette disparité de sinistres en développant une méthode basée sur des matrices de transition. Son approche est décrite dans son article MAHON [2005]. Il s'agit d'une méthode markovienne qui se base également sur une segmentation des sinistres en fonction de certains seuils mais Mahon utilise cette segmentation pour estimer l'intervalle de coût dans lequel chaque sinistre se situera à la fin de sa liquidation, cette estimation étant effectuée à l'aide de matrices de transition construites d'après les données empiriques. Par ailleurs, la méthode proposée par Mahon se base des matrices différentes à chaque pas de maturité de la cohorte de sinistres. Nous complèterons ces estimations par une évaluation de la charge ultime.

Comme nous pourrons le constater, la méthode de Mahon présente quelques imperfections et l'objectif de ce mémoire sera d'en proposer une amélioration et de développer deux autres approches basées sur cette première méthode des matrices de transition.

L'une d'entre elles n'a pas pu fournir de résultats satisfaisants étant données les hypothèses beaucoup trop fortes sur lesquelles elle repose. Son but était de

proposer une nouvelle manière de construire les matrices. Pour cela, l'idée était de palier à l'incertitude générée lors de la construction des matrices par le faible nombre de sinistres présents dans les classes les plus coûteuses. Les classes les plus peuplées en sinistres (i.e. les plus faibles en coût) et donc les plus fiables auraient permis de modéliser mathématiquement les probabilités de transition relatives aux sinistres les plus graves. Les hypothèses et la façon dont nous aurions aimé que cette méthode fonctionne seront présentées dans le chapitre dédié.

L'autre alternative est quant à elle plus intéressante. Ce modèle a pour objectif de se libérer d'une hypothèse très contraignante de la méthode de Mahon : la durée maximale de la liquidation des sinistres. Il faut pour cela revoir la façon dont est défini le modèle. En effet, la chaîne de Markov construite par Mahon dépend de la maturité des sinistres, ce qui l'oblige à considérer que les sinistres les plus vieux de l'historique ont forcément atteint la fin de leur vie. Des probabilités de transition ne dépendant pas du temps nous permettraient de ne pas avoir à faire cette hypothèse. Comme nous le verrons dans le chapitre consacré à cette méthode, la classification des sinistres doit alors nécessairement être plus fine que précédemment si nous voulons y parvenir. Le coût ou la charge des sinistres est dans ce cas une information certes utile, mais insuffisante. Ce nouveau modèle sera accompagné d'une proposition d'évaluation de la charge ultime un peu plus approfondie que celle abordée à la suite de la méthode de Mahon.

Chapitre 1

Méthodes classiques de provisionnement non-vie

Dans cette première partie, nous allons aborder les méthodes de provisionnement non-vie les plus répandues. Dans la mesure où il s'agit des méthodes usuelles, elles ne prennent pas en compte l'hétérogénéité des sinistres.

Il existe deux catégories de méthodes qui permettent de calculer la valeur des engagements d'assurance de l'entreprise. Nous trouvons d'une part des méthodes dites déterministes qui se contentent de fournir un résultat qui sera l'estimation du montant à provisionner. Mais ces méthodes se révèlent insuffisantes car elles ne prennent pas en compte l'erreur d'estimation qui est commise. Nous avons donc besoin également des méthodes stochastiques qui permettent quant à elles de mesurer cette erreur.

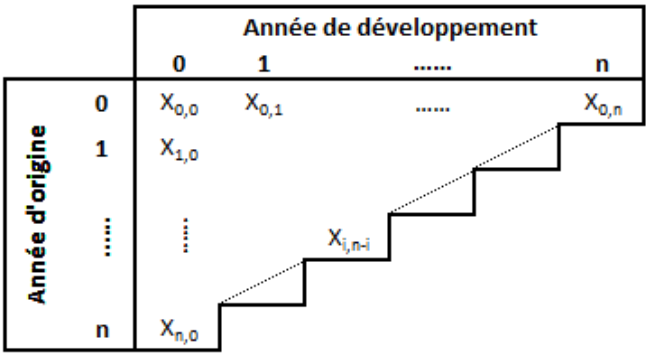
1.1 Triangles de liquidation

Avant de décrire les méthodes, des définitions qui serviront de préliminaires sur la façon dont nous procédons pour appliquer ces modèles seront présentées ici. Les montants de sinistres payés sont représentés sous forme de triangles de liquidation. En effet, dans la mesure où les sinistres sont rarement entièrement réglés dès l'année où ils ont lieu, notre analyse se doit de prendre en compte deux dates : l'année de survenance du sinistre et l'année de développement ou de liquidation (i.e. l'année au cours de laquelle le sinistre a été réglé).

On note I le nombre d'années de survenance connues et J le nombre d'années de développement connues, et on suppose que $I = J = n$. On note $X_{i,j}$ le montant réglé au cours de l'année $i + j$ pour les sinistres ayant eu lieu durant l'année i, et on note $C_{i,j}$ le montant cumulé correspondant, autrement dit :

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

Les informations sont donc connues pour $i + j \leqslant n$. Ces deux dimensions nous conduisent à représenter nos données sous la forme d'un triangle de liquidation :



Une troisième dimension est définie à partir des deux premières : il s'agit de l'année comptable (qui correspond donc aux diagonales, i.e. à un $i + j$ donné).

Il est également possible de construire un triangle à partir d'autres données que les règlements pour appliquer nos méthodes de provisionnement, comme par exemple la charge ou le nombre de sinistres (ouverts ou réglés).

L'objectif de la plupart des méthodes est de calculer une estimation des montants futurs, autrement dit de compléter de la partie inférieure du triangle. Cela permet notamment d'obtenir la charge ultime par année de survenance i égale à $C_{i,n}$.

La provision par année d'origine i se déduit naturellement de cette charge ultime :

$$R_i = C_{i,n} - C_{i,n-i} = \sum_{k=n-i+1}^{n} X_{i,j}$$

.

Et la provision totale sera égale à la somme des provisions par année de survenance :

$$R = \sum_{i=0}^{n} R_i$$

.

On définit également les facteurs de développement individuels et les cadences de règlements par :

$$f_{i,j} = \frac{C_{i,j+1}}{C_{i,j}} \quad \text{et} \quad p_{i,j} = \frac{C_{i,j}}{C_{i,n}}$$

On remarque alors que $f_{i,j} = \frac{p_{i,j+1}}{p_{i,j}}$. Ces outils seront utilisés dans les méthodes décrites dans les parties suivantes.

1.2 Méthodes déterministes

1.2.1 Méthode Chain Ladder

Présentation de la méthode

Le modèle Chain Ladder est la méthode de provisionnement la plus répandue dans le milieu de l'assurance non-vie. Il s'agit aussi de la méthode la plus simple à mettre en place et repose sur deux hypothèses :

H1 Les paiements cumulés $C_{i,j}$ ne dépendent pas de l'année de survenance i

H2 Il existe des facteurs de développement $f_0, \ldots, f_{J-1}$ tels que pour tout $0 \leqslant i \leqslant n$ et $0 \leqslant j \leqslant n$, on ait :

$$\mathrm{E}[C_{i,j}|C_{i,0}, \ldots, C_{i,j-1}] = \mathrm{E}[C_{i,j}|C_{i,j-1}] = f_{j-1} \times C_{i,j-1}$$

Nous détaillerons plus loin des tests permettant de vérifier si ces hypothèses sont respectées.

Dans la deuxième hypothèse, on remarque que les facteurs de développement f_j ne dépendent pas de l'année d'origine i. Cela signifie que les facteurs de développement devraient idéalement satisfaire la relation $f_j = f_{0,j} = ... = f_{n-1,j}$.
C'est pourquoi les facteurs de développement f_j sont estimés par la moyenne des facteurs de développement individuels $f_{i,j}$:

$$\widehat{f_j} = \frac{\sum_{i=0}^{n-j-1} C_{i,j+1}}{\sum_{i=0}^{n-j-1} C_{i,j}}$$

On peut montrer que les estimateurs $\widehat{f_j}$ sont sans biais et non corrélés.

Les hypothèses du modèle nous permettent alors d'estimer les paiements cumulés futurs par :

$$\widehat{C}_{i,j+1} = \mathrm{E}[C_{i,j+1}|C_{i,j}] = \widehat{f_j} \times C_{i,j}$$

Et on en déduit le montant de la charge ultime pour l'année de survenance i :

$$\widehat{C}_{i,n} = C_{i,n-i} \times \prod_{k=n-i}^{n-1} \widehat{f_k}$$

On en déduit également les estimations du montant de la provision par année de survenance et de la provision totale :

$$\widehat{R}_i = \widehat{C}_{i,n} - C_{i,n-i}$$

$$\widehat{R} = \sum_{i=0}^{n} \widehat{R}_i$$

Adaptation de la méthode

La méthode peut éventuellement être adaptée dans le cas où certains facteurs de développement individuels ne nous paraissent pas pertinents, par exemple si ils causent un effet calendaire dans le triangle. En effet, ne pas prendre en compte certaines données peut parfois permettre d'obtenir des résultats plus cohérents. Il suffit pour cela d'adapter le calcul de l'estimation des facteurs de développement :

$$\widehat{f_j} = \frac{\sum_{i \in A} C_{i,j+1}}{\sum_{i \in A} C_{i,j}} \quad \text{où} \quad A \subset \{0, ..., n-j-1\}$$

Il s'agit en réalité d'un cas particulier de la méthode Chain Ladder pondérée qui consiste à affecter un poids de pondération à chaque facteur de développement individuel (ici les poids valent 0 ou 1).

1.2.2 Vérification des hypothèses du modèle

Bien que le modèle Chain Ladder soit très utilisé, ses hypothèses ne sont pas toujours vérifiées. Nous rappelons ci-dessous les hypothèses H1 et H2 sur lesquels est basée la méthode :

H1 Les paiements cumulés $C_{i,j}$ ne dépendent pas de l'année de survenance i

H2 Il existe des facteurs de développement $f_0, \ldots, f_{J-1}$ tels que pour tout $0 \leqslant i \leqslant n$ et $0 \leqslant j \leqslant n$, on ait :

$$\mathrm{E}[C_{i,j}|C_{i,0}, \ldots, C_{i,j-1}] = \mathrm{E}[C_{i,j}|C_{i,j-1}] = f_{j-1} \times C_{i,j-1}$$

Nous allons maintenant expliquer comment nous pouvons déterminer si un triangle est apte à se voire appliquer la méthode Chain Ladder (cf BOUTHIER, HIPPOLYTE, [2015]).

Vérification de l'hypothèse H1

Cette première hypothèse concerne l'indépendance des années de survenance, autrement dit l'absence d'effet calendaire dans les diagonales du triangle. Il existe un test mis en place par Thomas Mack permettant de nous assurer que le triangle considéré vérifie bien cette hypothèse.

La méthode consiste à étudier les facteurs de développement individuels et à étudier les tendances haussières ou baissières sur chaque diagonale.

Appelons D_j les éléments d'une diagonale j $(1 \leqslant j \leqslant n) : D_j = \{C_{j,i}, ..., C_{i,j}\}$. On note n_j le nombre d'éléments de D_j.

À chaque facteur de développement individuel est associé :
- la lettre G si $f_{i,j}$ est supérieur à la médiane des facteurs de développement individuels de la colonne j
- la lettre P si $f_{i,j}$ est inférieur à la médiane des facteurs de développement individuels de la colonne j
- la lettre N si $f_{i,j}$ est égal à la médiane des facteurs de développement individuels de la colonne j

Cette règle nous permet de construire un nouveau triangle constitué des lettres G, P et N. La méthode repose sur le fait que, en théorie, les nombres de P et de G sur une même diagonale devraient être à peu près identiques. Le cas contraire mettrait en évidence la présence d'un effet calendaire.

Nous appelons respectivement G_j et P_j le nombre de G et de P sur la colonne j. Nous avons également besoin d'introduire une variable aléatoire $Z_j = \min(G_j, P_j)$ qui devrait normalement être proche de la moyenne de G_j et de P_j : $\frac{G_j+P_j}{2}$.

Si on suppose qu'il y a bien autant de G que de P sur chaque diagonale, la variable aléatoire représentant la probabilité qu'un élément de la diagonale soit associé à un P suit une loi de Bernoulli de paramètre $\frac{1}{2}$. Sous cette hypothèse, G_j et P_j suivent chacun une loi Binomiale de paramètres $(n_j, \frac{1}{2})$.

Nous utilisons les résultats suivants montrés par Mack :

$$\mathrm{E}[Z_j] = \frac{n_j}{2} - \binom{n_j - 1}{m_j} \times \frac{n_j}{2^{n_j}}$$

$$\mathrm{Var}(Z_j) = \frac{n_j(n_j-1)}{4} - \binom{n_j - 1}{m_j} \times \frac{n_j(n_j-1)}{2^{n_j}} + \mathrm{E}[Z_j] - \mathrm{E}[Z_j]^2$$

où $m_j = \lfloor \frac{n_j-1}{2} \rfloor$.

Nous définissons maintenant la variable $Z = \sum_{j=2}^{n-1} Z_j$. L'hypothèse d'indépendance des années de survenance nous assure que les $(Z_j)_{1 \leqslant j \leqslant n}$ sont indépendants, ce qui nous permet d'obtenir les résultats suivants : $\mathrm{E}[Z] = \sum_{j=2}^{n-1} \mathrm{E}[Z_j]$ et $\mathrm{Var}(Z) = \sum_{j=2}^{n-1} \mathrm{Var}(Z_j)$. Mack utilise ensuite le fait que la somme de lois binomiales converge aléatoirement vers une loi normale pour considérer que Z suit une distribution de loi normale.

On en déduit que la condition pour que l'hypothèse soit vérifiée au seuil 5% correspond à l'intervalle de confiance : $[\mathrm{E}[Z] - 1.96\sqrt{\mathrm{Var}(Z)}, \mathrm{E}[Z] + 1.96\sqrt{\mathrm{Var}(Z)}]$. Nous rejetterons donc l'hypothèse si Z n'appartient pas à cet intervalle.

Vérification de l'hypothèse H2

Cette hypothèse signifie que les facteurs de développements individuels $f_{i,j}$ sont indépendants de l'année d'origine i.

Nous pouvons vérifier cette hypothèse en construisant sur un graphe les points correspondants aux couples $(C_{i,j}, C_{i,j+1})$ pour une année de développement j donnée. Si l'hypothèse est effectivement vérifiée, les points ainsi tracés sont censés se trouver à peu près alignés sur une droite. Nous pouvons donc effectuer la régression linéaire de ces points et considérer qu'un coefficient de détermination R^2 proche de 1 correspond à la véracité de l'hypothèse.

1.3 Méthodes stochastiques

Nous allons maintenant nous intéresser à des méthodes permettant de quantifier le risque de provisionnement, autrement dit l'erreur commise par notre estimation. Nous utilisons pour cela l'erreur quadratique moyenne de prédiction (ou mean square error of prediction ou MSEP).

1.3.1 Erreur de prédiction

Dans un cadre plus général, l'erreur quadratique moyenne d'un estimateur $\widehat{X}$ correspondant à une variable aléatoire X est définie par :

$$\mathrm{E}[(\widehat{X} - X)^2]$$

Ici, nous nous intéresserons par exemple à l'estimateur $\widehat{R}$ du montant de la provision, donc :

$$\mathrm{MSEP}(\widehat{R}) = \mathrm{E}[(\widehat{R} - R)^2]$$

Nous pourrions également étudier l'erreur associée à l'estimation de la charge ultime mais étant donné que $\widehat{R} - R = \widehat{C}_{i,n} - C_{i,n}$, ces deux erreurs de prédiction sont donc identiques.

On définit aussi l'erreur standard de prédiction par : $\mathrm{sep}(\widehat{R}) = \sqrt{(\mathrm{MSEP}(\widehat{R}))}$.

L'objectif des méthodes de provisionnement stochastiques est d'estimer cette erreur de prédiction.

1.3.2 Modèle de Mack

Le modèle de Mack correspond à l'adaptation stochastique du modèle Chain Ladder. Il s'agit donc d'un modèle très utilisé. En effet, le but de ce modèle proposé par Mack en 1993 est de nous permettre d'évaluer l'erreur de prédiction commise lors de l'évaluation des engagements par la méthode Chain Ladder.

Il est dont cohérent que les deux hypothèses H1 et H2 du modèle Chain Ladder soient encore exigées ici. Ces hypothèses sont rappelées ci-dessous.

H1 Les paiements cumulés $C_{i,j}$ ne dépendent pas de l'année de survenance i

H2 Il existe des facteurs de développement $f_0, \ldots, f_{J-1}$ tels que pour tout $0 \leqslant i \leqslant n$ et $0 \leqslant j \leqslant n$, on ait :

$$\mathrm{E}[C_{i,j}|C_{i,0}, \ldots, C_{i,j-1}] = \mathrm{E}[C_{i,j}|C_{i,j-1}] = f_{j-1} \times C_{i,j-1}$$

Il nous faut de plus ici une hypothèse supplémentaire concernant la variance des estimations :

H3 $\mathrm{Var}(C_{i,j}|C_{i,0}, \ldots, C_{i,j-1}) = \mathrm{Var}(C_{i,j}|C_{i,j-1}) = \sigma_{j-1}^2 \times C_{i,j-1}$

L'estimation des facteurs de développement f_j est identique à celle utilisée dans la partie déterministe :

$$\widehat{f_j} = \frac{\sum_{i=0}^{n-j-1} C_{i,j+1}}{\sum_{i=0}^{n-j-1} C_{i,j}}$$

De plus, les paramètres de variance σ_j^2 sont estimés par :

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

La formule précédente n'étant valable que pour $j < n-1$, on définit l'estimateur du dernier paramètre σ_{n-1}^2 par : $\widehat{\sigma}_{n-1}^2 = \min \left(\frac{\widehat{\sigma}_{n-2}^4}{\widehat{\sigma}_{n-3}^2}, \max(\widehat{\sigma}_{n-2}^2, \widehat{\sigma}_{n-3}^2) \right)$.

Nous pouvons maintenant définir l'estimateur de l'erreur quadratique moyenne de prédiction $\text{MSEP}(\widehat{R}_i)$ fournie par Mack pour chaque année d'origine i :

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

L'erreur de prédiction est composée d'une erreur de modèle (premier terme) et d'une erreur de prédiction (deuxième terme).

Ces estimateurs ne sont pas indépendants (en effet, ils dépendent des mêmes facteurs $\widehat{f_j}$). L'erreur quadratique moyenne de prédiction agrégée ne peut donc pas être obtenue en se contentant de sommer toutes les estimations des $\text{MSEP}(\widehat{R}_i)$. L'erreur de prédiction totale $\text{MSEP}(\widehat{R})$ est estimée par la quantité suivante :

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

1.3.3 Vérification des hypothèses

Le modèle de Mack est basé sur des hypothèses qui ne sont pas toujours vérifiées, ce qui peut poser problème. Les hypothèses sont les suivantes :

H1 Les paiements cumulés $C_{i,j}$ ne dépendent pas de l'année de survenance i

H2 Il existe des facteurs de développement $f_0, \ldots, f_{J-1}$ tels que pour tout $0 \leqslant i \leqslant n$ et $0 \leqslant j \leqslant n$, on ait :

$$\text{E}[C_{i,j}|C_{i,0}, \ldots, C_{i,j-1}] = \text{E}[C_{i,j}|C_{i,j-1}] = f_{j-1} \times C_{i,j-1}$$

H3 $\text{Var}(C_{i,j}|C_{i,0}, \ldots, C_{i,j-1}) = \text{Var}(C_{i,j}|C_{i,j-1}) = \sigma_{j-1}^2 \times C_{i,j-1}$

Les méthodes permettant de vérifier les hypothèses H1 et H2 ont déjà été traitées dans la partie concernant le modèle Chain Ladder. En effet, il s'agit des

mêmes hypothèses. Il ne nous reste donc plus qu'à expliquer comment s'assurer de la véracité de l'hypothèse H3 (cf BOUTHIER, HIPPOLYTE, [2015]). Cette hypothèse signifie que la volatilité des paiements ne dépend que de l'année de développement et pas de l'année de survenance.

Cette dernière hypothèse peut se vérifier en étudiant graphiquement les résidus définis par :

$$R_{i,j} = \frac{C_{i,j+1} - C_{i,j} f_j}{\sqrt{C_{i,j}}}$$

La véracité de l'hypothèse se traduit alors par le fait que les résidus soient parfaitement aléatoires. Ce type de jugement possède forcément une certaine subjectivité.

1.4 Application

Dans cette section, nous présenterons les résultats obtenus avec les méthodes de provisionnement décrites précédemment, c'est-à-dire le modèle Chain Ladder et le modèle de Mack, ainsi que l'application des méthodes permettant de vérifier les hypothèses sous-jacentes.

L'application détaillée ci-dessous concerne dans un premier temps le triangle des règlements cumulés (bruts de recours). Ce triangle est présenté en annexe A.1.1.

1.4.1 Vérification des hypothèses

Hypothèse H1

Nous allons ici vérifier l'hypothèse selon laquelle les paiements ne dépendent pas de l'année de survenance. Pour rappel, la méthode décrite précédemment permettant de valider cette hypothèse définit une variable aléatoire notée Z et en propose une estimation théorique. Nous rappelons que l'intervalle de confiance à 95% permettant de vérifiant si la valeur de Z obtenue empiriquement est :

$$[\mathrm{E}[Z] - 1.96\sqrt{\mathrm{Var}(Z)}, \mathrm{E}[Z] + 1.96\sqrt{\mathrm{Var}(Z)}]$$

Le tableau ci-dessous récapitule les valeurs de Z observées ainsi que les estimations théoriques de son espérance et de sa variance. Il s'agit de vérifier que Z soit compris entre les bornes inférieures et supérieures de l'intervalle de confiance à 95%. En plus de l'estimation globale de Z, les résultats sont également présentés par année de liquidation.

Année de liquidation	Z	E[Z]	Var(Z)	Borne inf	Borne sup
2	1	0,50	0,25	−0,48	1,48
3	0	0,50	0,25	−0,48	1,48
4	2	1,25	0,44	−0,05	2,55
5	2	1,56	0,37	0,37	2,76
6	3	2,06	0,62	0,52	3,61
7	2	2,06	0,62	0,52	3,61
8	2	2,91	0,80	1,15	4,66
9	2	2,41	0,55	0,95	3,86
10	5	3,77	0,99	1,82	5,72
11	3	3,27	0,74	1,59	4,95
Total	22	20,29	5,63	15,64	24,94

Au vu des résultats ci-dessus, nous pouvons considérer que le triangle des règlements cumulés vérifie l'hypothèse H1, c'est-à-dire l'indépendance des années de survenance.

Hypothèse H2

La deuxième hypothèse impose que les facteurs de développements individuels dépendent uniquement de l'année de liquidation (et non de l'exercice de survenance).

Comme expliqué précédemment, la vérification de cette hypothèse consiste à observer pour chaque année de développement j les graphes représentant les points $(C_{i,j}, C_{i,j+1})$. Ces graphes sont représentés en annexe A.2.1 pour $j = 1, \ldots, 8$. En effet, les maturités supérieures seraient vérifiées à partir d'un nombre de données trop faibles pour être fiables et ont un enjeu relativement faible sur l'estimation de la provision globale puisque les facteurs de développement commencent à être très proches de 1.

Comme nous pouvons le constater, les points semblent bien alignés à partir de la troisième année de développement. Cependant, ce n'est pas tout à fait le cas pour les deux premières années de liquidation. Cette observation se confirme par les coefficients de détermination R^2 des régressions linéaires qui sont présentés ci-dessous :

Année de liquidation	R^2
1	0,234
2	0,383
3	0,867
4	0,836
5	0,711
6	0,899
7	0,752
8	0,999

Les estimations des deux premiers facteurs de développement ne sont donc pas tout à fait fiables, ce qui implique que les estimations des deux années de survenance les plus récentes sont à considérer avec prudence.

Hypothèse H3

Il s'agit de l'hypothèse spécifique au modèle de Mack puisqu'elle concerne la variance des quantités considérées (en l'occurrence les règlements). Nous avons tracé pour chaque année de liquidation les graphes représentant les points $(R_{i,j}, C_{i,j})$ où $R_{i,j}$ est le résidu défini précédemment :

$$R_{i,j} = \frac{C_{i,j+1} - C_{i,j} f_j}{\sqrt{C_{i,j}}}$$

Comme pour l'hypothèse H2, les graphes sont représentés en annexe A.2.2 jusqu'au huitième pas de maturité.

Comme nous l'avions indiqué, cette hypothèse est vérifiée si les résidus sont parfaitement aléatoires. Les observations graphiques ne semblent pas révéler de dépendance entre les résidus $R_{i,j}$ et les paiements cumulés $C_{i,j}$), c'est pourquoi nous faisons le choix de considérer l'hypothèse comme valable même si ce choix est à considérer avec prudence.

1.4.2 Résultats obtenus

Les résultats par années de survenance obtenus avec le modèle Chain Ladder et le modèle de Mack sont présentés dans le tableau ci-dessous :

Année de survenance	Montant réglé au 31/12/2015 (k€)	Charge ultime (k€)	Provision (k€)	Erreur de prédiction standard (k€)
2004	13 537	13 537	0	0
2005	18 662	18 782	119	0
2006	18 338	18 856	518	20
2007	17 539	18 304	765	37
2008	20 141	21 655	1 513	91
2009	14 630	17 564	2 934	602
2010	15 795	20 295	4 500	745
2011	11 541	16 065	4 525	852
2012	10 941	18 382	7 441	1 027
2013	10 459	21 058	10 599	1 177
2014	7 117	23 097	15 979	1 407
2015	2 435	37 996	35 562	2 780
Total	161 136	245 590	84 455	4 854

L'erreur de prédiction du modèle de Mack permet de plus d'évaluer les quantiles de la distribution du montant de provision à effectuer sous-hypothèse de log-normalité et leurs estimations sont présentées dans le tableau ci-dessous :

Probabilité	Quantile (k€)	Ratio quantile / espérance
75%	85 305	101,0%
85%	87 097	103,1%
90%	88 332	104,6%
95%	90 194	106,8%
99%	93 795	111,1%
99,5%	95 148	112,7%

L'évaluation des flux futurs obtenus avec la méthode Chain Ladder nous permet de plus de calculer la duration de l'écoulement, qui est donc estimée à 3,35 ans. Cela traduit une liquidation lente des sinistres.

Liquidation de la charge

Comme nous l'indiquions, ces méthodes peuvent également être appliquées au triangle des charges (présenté en annexe A.1.2). Nous procédons ici de la même manière que pour les règlements, en présentant les vérifications des hypothèses d'une part et les résultats d'autre part.

Hypothèse H1 Les résultats du test statistique servant vérifier l'hypothèse H1 sont présentés ci-dessous :

Année de liquidation	Z	$\mathrm{E}[Z]$	$\mathrm{Var}(Z)$	Borne inf	Borne sup
2	1	0,50	0,25	−0,48	1,48
3	0	0,75	0,19	−0,10	1,60
4	1	0,75	0,19	−0,10	1,60
5	2	1,56	0,37	0,37	2,76
6	2	2,06	0,62	0,52	3,61
7	1	1,56	0,37	0,37	2,76
8	1	2,41	0,55	0,95	3,86
9	2	3,27	0,74	1,59	4,95
10	5	3,77	0,99	1,82	5,72
11	3	3,27	0,74	1,59	4,95
Total	18	19,90	5,00	15,52	24,28

D'après les résultats ci-dessus, il est possible de considéré que l'hypothèse H1 est vérifiée puisque les valeurs de Z obtenues se situent dans l'intervalle de confiance indiqué précédemment : $[\mathrm{E}[Z] - 1.96\sqrt{\mathrm{Var}(Z)}, \mathrm{E}[Z] + 1.96\sqrt{\mathrm{Var}(Z)}]$.

Hypothèses H2 et H3 Les graphiques à observer afin de contrôler ces hypothèses sont présentés en annexe.

Tout d'abord pour l'hypothèse H2, les points $(C_{i,j}, C_{i,j+1})$ semblent être alignés à partir de la deuxième année de liquidation, ce qui signifie que la provision estimée pour l'année de survenance la plus récente (i.e. 2015) sera moins fiable.

En ce qui concerne l'hypothèse H3, la conclusion est similaire à celle que nous avions faite pour le triangle des règlements, autrement dit nous considérons que les résidus observés dans les graphiques des points $(R_{i,j}, C_{i,j})$ semblent effectivement aléatoires.

Résultats Les provisions obtenues par année de survenance en liquidant ce triangle avec le modèle Chain Ladder ainsi que les volatilités associées sont présentées ci-dessous :

Année de survenance	Charge ultime (k€)	Provision (k€)	Erreur de prédiction standard (k€)
2004	15 783	2 246	0
2005	19 996	1 334	0
2006	19 867	1 530	279
2007	34 952	17 413	1 138
2008	37 056	16 914	2 850
2009	18 356	3 726	3 458
2010	30 355	14 559	4 809
2011	19 563	8 022	4 460
2012	20 204	9 263	6 079
2013	35 298	24 839	10 300
2014	31 601	24 484	14 084
2015	32 120	29 685	25 657
Total	315 151	154 015	38 726

Ces montants sont sensiblement plus élevés que ceux obtenus avec le triangle des règlements, ce qui s'explique notamment par les montants parfois très importants de provisions dossier présents dans les charges à l'inventaire du 31/12/2015, en particulier sur les exercices 2007, 2008 et 2010.

Chapitre 2

Chaînes de Markov

Ce deuxième chapitre présentera certaines notions sur les chaînes de Markov qui s'avéreront utiles pour la suite dans la mesure où nous présenterons dans ce mémoire un modèle se basant sur ce type d'objet mathématique. Les chaînes de Markov ont été développées par Andreï Markov et permettent de modéliser un système constitué de plusieurs états et de prévoir l'évolution du système dans le futur. Dans l'application que nous ferons de ce modèle, les différents états décriront différentes classes de sinistres automobiles.

2.1 Premières définitions

Une chaîne de Markov se définit tout d'abord à l'aide d'un nombre d'états qui constituent un ensemble E. Il faut préciser que E peut-être fini ou infini mais il doit nécessairement être dénombrable.

Définition d'une chaîne de Markov Une suite de variables aléatoires $(X_n)_{n\geqslant 0}$ sur E est une chaîne de Markov si :

$$\forall i_0, \ldots, i_n, j, \mathbb{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n)$$

On retrouve alors le fait que l'estimation de X_{n+1} ne dépend que de l'état du système dans le présent, c'est-à-dire de X_n : on appelle ce résultat la propriété de Markov.

Par abus de langage, on désigne la plupart du temps par "chaîne de Markov" une chaîne de Markov dont la probabilité de transition ne dépend pas de n :

$$\forall i, j \in E, \mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_1 = j | X_0 = i)$$

Une chaîne de Markov vérifiant cette propriété porte le nom de chaîne de Markov homogène. Pour définir une chaîne de Markov homogène, nous avons également besoin d'une loi de probabilité $p_{i,j}$ sur E^2. Il s'agit de probabilités de transition définies par :

$$\forall i, j \in E, \mathbb{P}(X_1 = j | X_0 = i) = p_{i,j}$$

Il est important de préciser que $\forall i \in E, \sum_{j \in E} p_{i,j} = 1$.

Si de plus une distribution initiale sur E (i.e. une distribution de X_0) est connue, on peut en déduire la formule suivante :

$$\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_0 = i_0) \times p_{i_0,i_1} \times ... \times p_{i_{n-1},i_n}$$

On retrouve le fait que $\sum_{j \in E} p_{i,j} = 1$ comme indiqué plus haut.

La différence entre une chaîne de Markov homogène et non-homogène aura son importance dans la suite du mémoire. Néanmoins, afin d'alléger le vocabulaire, toutes les chaînes de Markov considérées jusqu'à la fin ce chapitre seront des chaînes de Markov homogènes même si ce ne sera pas précisé.

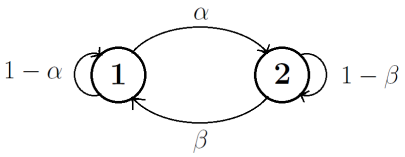
2.2 Matrice de transition

Cette partie présente quelques rappels théoriques sur la façon de modéliser mathématiquement une chaîne de Markov. En effet, une chaîne de Markov peut être représentée visuellement par un graphe probabiliste. Il s'agit d'une modélisation assez intuitive mais qui ne permet pas de manipuler la chaîne de Markov mathématiquement.

Si le nombre d'états est fini, il est également possible de modéliser une chaîne de Markov par une matrice de transition. Il s'agit d'une notion qui sera utilisée dans les différents modèles décrits dans les chapitres suivants du mémoire.

En notant N le nombre d'état, il s'agit d'une matrice carrée de taille N dont les coefficients de chaque ligne i et de chaque colonne j sont égaux à $p_{i,j}$. Chaque ligne correspond à un état initial et les colonnes correspondent aux états d'arrivée. On notera cette matrice P.

Voici un exemple simple de graphe probabiliste et de matrice de transition modélisant la même chaîne de Markov à deux états avec $0 \leqslant \alpha \leqslant 1$ et $0 \leqslant \beta \leqslant 1$:



$$\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Donc la somme des coefficients de chaque ligne de la matrice est égale à 1 puisque $\forall i \in E, \sum_{j \in E} p_{i,j} = 1$. Une matrice vérifiant cette condition (avec de plus $0 \leqslant p_{i,j} \leqslant 1 \forall i, j$) est appelée une matrice stochastique.

Une chaîne de Markov $(X_n)_{n \geqslant 0}$ est entièrement caractérisée par sa matrice de transition et par la distribution de X_0. En effet, on déduit des résultats précédents que si on note μ_n le vecteur de la distribution de X_n, on a :

$$\mu_n = \mu_0 \times P^n$$

Donc si la suite $(P^n)_{n\geqslant 0}$ est convergente et qu'on note L sa limite, on en déduit que la suite (X_n) converge en loi vers $\mu_0 \times L$.

2.3 Propriétés des chaînes de Markov

2.3.1 États

Définitions

- Un état i d'une chaîne de Markov est dit absorbant si $p_{i,i} = 1$ et $p_{i,j} = 0$ pour $i \neq j$.
- L'état j est accessible à partir de l'état i s'il existe $n \in \mathbb{N}$ tel que $\mathbb{P}(X_n = j|X_0 = i) > 0$.
- Deux états i et j sont communicants si l'état j est accessible à partir de l'état i et si l'état i est accessible à partir de l'état j.

On remarque que tout état communique forcément au moins avec lui-même (pour le vérifier, il suffit de prendre $n = 0$ dans la définition).

Si tous les états d'une chaîne de Markov communiquent entre eux, la chaîne est dite irréductible.

2.3.2 Périodicité

Définition Soit une chaîne de Markov $(X_n)_{n\geqslant 0}$ sur E et soit $i \in E$.
On définit la période de i comme le PGCD de tous les entiers $n \geqslant 1$ tels que $\mathbb{P}(X_n = j|X_0 = j) > 0$. On note alors cette période $d(i)$.
S'il n'existe pas de tels entiers, alors $d(i) = +\infty$.
Si $d(i) = 1$, alors l'état i est dit apériodique.

Propriété Si i est un état périodique (de période finie) qui communique avec l'état j, alors j est un état périodique et $d(i) = d(j)$.

2.3.3 Récurrence et transience

Définition Soit une chaîne de Markov $(X_n)_{n\geqslant 0}$ sur E.
On note $\tau_i = \inf\{n \geqslant 1, X_n = i\}$ le temps d'atteinte de l'état i.

- L'état i est transient ou transitoire si $\mathbb{P}(\tau_i < \infty|X_0 = i) < 1$.
- L'état i est récurrent si $\mathbb{P}(\tau_i < \infty|X_0 = i) = 1$.
- L'état i est récurrent positif si $\mathrm{E}[\tau_i|X_0 = i] < \infty$.
- L'état i est récurrent nul si $\mathbb{P}(\tau_i|X_0 = i) = 1$ et $\mathrm{E}[\tau_i|X_0 = i] = \infty$.

Propriété

- L'état i est transient si et seulement si $\sum_{n\geqslant 1} \mathbb{P}(X_n = i|X_0 = i) < \infty$
- L'état i est récurrent si et seulement si il existe presque sûrement une infinité d'indice n tels que $X_n = i$ sachant que $X_0 = i$

Si l'état j est transient, on a donc nécessairement :

$$\forall i \in E, \lim_{n\mapsto+\infty} \mathbb{P}(X_n = j|X_0 = i) = 0$$

Propriété Si les états i et j communiquent, alors ils sont de même nature (transients, récurrents nuls, récurrents positifs).

Corollaire Si la chaîne de Markov est irréductible, alors tous les états sont de même nature (transients, récurrents nuls, récurrents positifs). On dit que la chaîne est transiente, récurrente ou périodique.

2.3.4 Probabilité stationnaire

Définition Une probabilité stationnaire d'une chaîne de Markov est une mesure de probabilité π sur E telle que $\pi = \pi P$ où P est la matrice de transition de la chaîne de Markov.
Il existe une propriété stipulant que $\forall i, j \in E, \pi(i)p_{i,j} = \pi(j)p_{j,i}$.
On a de plus : $\sum_{i \in E} \pi(i)p_{i,j} = \pi(j)$.

La probabilité stationnaire représente la fraction du temps passé dans chaque état de la chaîne de Markov. Si la distribution initiale de la chaîne de Markov $(X_n)_{n \geqslant 0}$ est aussi une loi de probabilité stationnaire, alors les $(X_n)_{n \geqslant 0}$ sont identiquement distribués.

Propriété Si π est une probabilité stationnaire et si l'état i est transitoire, alors $\pi(i) = 0$. De plus, si une chaîne de Markov ne possède que des états transitoires, alors il n'existe pas de probabilité stationnaire.

Il existe de nombreuses propriétés basées sur les probabilités stationnaires. Nous ne nous intéresserons qu'à celles concernant les chaînes de Markov dont le nombre d'états est fini dans la mesure où il s'agit du seul cas qui sera mis en application dans la suite du mémoire.

Dans toute la suite, $(X_n)_{n \geqslant 0}$ désignera une chaîne de Markov dont le nombre d'états est fini.

Propriété

- $(X_n)_{n \geqslant 0}$ admet au moins une probabilité stationnaire.
- Si $(X_n)_{n \geqslant 0}$ est irréductible, on a de plus l'unicité de la probabilité stationnaire. Elle est défini $\forall i \in E$ par :

$$\pi(i) = \frac{1}{\mathrm{E}[\tau_i | X_0 = i]}$$

- Si $(X_n)_{n \geqslant 0}$ est irréductible et apériodique, alors $X_n \xrightarrow{\mathcal{L}} \pi$ indépendamment de la distribution initiale, autrement dit la suite $(P^n)_{n \geqslant 0}$ converge vers la matrice dont toutes les lignes sont égales à π.

Théorème ergodique On considère que $(X_n)_{n \geqslant 0}$ est irréductible.
Soit π sa distribution stationnaire et $f : E \mapsto \mathbb{R}$.

$$\lim_{n \mapsto +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \sum_{i \in E} \frac{f(i)}{\mathrm{E}[\tau_i | X_0 = i]} = \int f \mathrm{d}\pi \quad \text{p.s.}$$

Chapitre 3

Méthode des matrices de transition

3.1 Présentation des données

Avant de présenter le modèle, nous allons donner dans cette partie certaines indications importantes sur les données à partir desquelles l'étude sera effectuée. Tout d'abord, il faut préciser que la garantie Responsabilité Civile Corporelle couvre uniquement les dommages corporels infligés aux différentes victimes. La difficulté d'évaluer le montant total qui sera réglé pour chaque sinistre dès leur survenance est d'autant plus importante que leur liquidation est longue : pour rappel, nous avions calculé une duration de 3,35 ans avec la méthode Chain Ladder dans le premier chapitre. En effet, l'état des victimes étant susceptible d'évoluer au cours du temps, il peut arriver qu'un sinistre mette plusieurs années avant d'être entièrement réglé. La règlementation relative aux remboursements sur cette garantie n'a pas évolué depuis la date où débutent nos données, ce qui n'impose donc pas une éventuelle problématique sur des retraitements à considérer.

3.1.1 Retraitement des montants en euros constants

Dans la mesure où nous allons être amenés à comparer des montants payés durant des années différentes voire même très éloignées, tous les montants de charge sont passés en euros constants afin que la comparaison soit valable.

La nomenclature servant de référence pour l'indemnisation des préjudices corporels est la Nomenclature Dinthillac. Elle répertorie les préjudices selon différents postes dont les principaux au sens économique sont les postes ATP (Assistance par tierce personne) et PGPF (Perte de gains professionnels futurs. Les règlements sur ces deux postes étant des indemnisations de salaire, nous faisons le choix de déterminer les indices de revalorisation selon la valeur du SMIC horaire durant chaque année comptable par rapport à sa valeur en 2015. Les taux à appliquer aux montants en euros courants pour les convertir en euros constants sont les suivants :

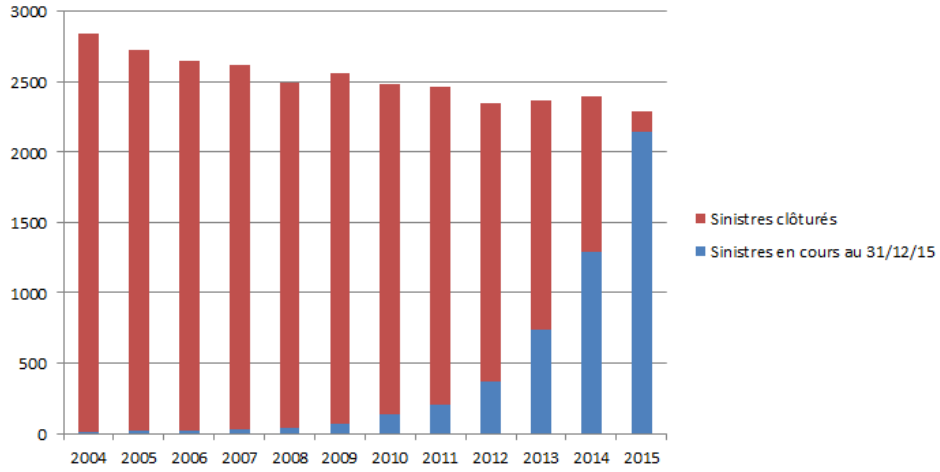
Année	SMIC horaire brut (€)	Indice de revalorisation
2004	7,61	1,263
2005	8,03	1,197
2006	8,27	1,162
2007	8,44	1,139
2008	8,63	1,114
2009	8,82	1,090
2010	8,86	1,085
2011	9,00	1,068
2012	9,22	1,042
2013	9,43	1,019
2014	9,53	1,008
2015	9,61	1,000

3.1.2 Description des sinistres

En ce qui concerne les sinistres contenus dans la base de données, ils sont au nombre de 30 254 dont 4 886 d'entre eux sont en cours à fin décembre 2015. Un détail des nombres de sinistres par année de survenance est indiqué dans le tableau ci-dessous :

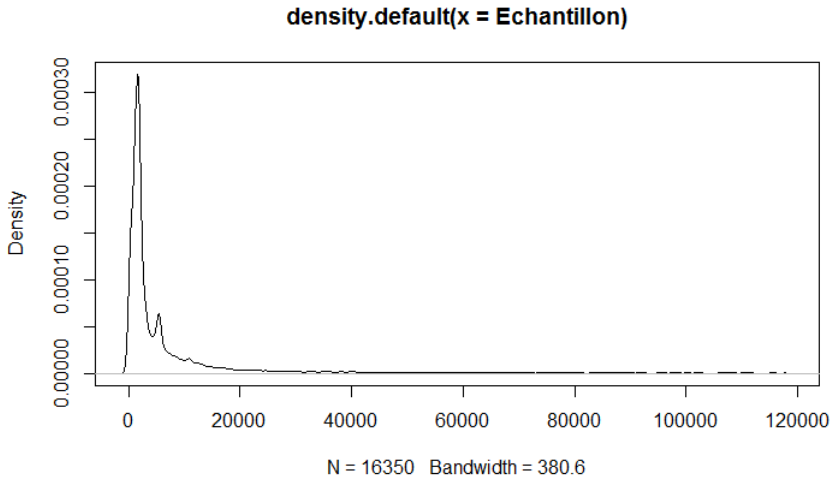
Année de survenance	Nombre de sinistres déclarés	Nombre de sinistres en cours au 31/12/2015)	Ratio sinistres en cours / total
2004	2 840	10	0,4%
2005	2 728	16	0,6%
2006	2 649	24	0,9%
2007	2 616	30	1,1%
2008	2 497	36	1,4%
2009	2 566	73	2,8%
2010	2 488	132	5,3%
2011	2 461	204	8,3%
2012	2 350	347	15,9%
2013	2 371	738	31,1%
2014	2 397	1 287	53,7%
2015	2 291	2 148	93,8%
Total	30 254	4 886	16,8%

Ces informations sont également représentées de manière graphique ci-dessous :



Comme nous pouvions nous y attendre, la proportion de sinistres en cours est beaucoup plus importante pour les survenances récentes.

De plus, sur notre stock de 30 254 sinistres, 13 604 soit 45% d'entre eux sont de charge nulle (i.e. sans suite) au 31/12/2015. La charge maximale (en euros constants) observée pour un sinistre à la date de l'étude sur l'ensemble de nos données est de 13 846 k€. Nous présentons ci-dessous le graphe représentant la densité empirique de notre échantillon réalisé avec la fonction *density* du logiciel R (à partir de tous les sinistres hors sans suite dont les 300 plus grandes valeurs ont été retirées afin de rendre le graphe plus lisible) :



Les quatre premiers moments empiriques que nous observons sur notre échantillon privé des sinistres sans suite sont les suivants :

- Espérance = 18,9 k€
- Écart-type = 214,6 k€

- Skewness = 37,0
- Kurtosis = 1 744,8

Enfin, nous présentons ci-dessous les principaux quantiles de notre échantillon (toujours hors sinistres sans suite) :

Probabilité	Quantile (k€)
10%	0,5
20%	1,2
30%	1,5
40%	1,6
50%	2,1
60%	3,0
70%	5,4
80%	8,2
90%	18,2
95%	40,8
99%	191,9

Tous ces éléments confirment que la garantie RC corporelle est effectivement très hétérogène et nous confortent dans notre choix d'avoir recours à un modèle adapté.

3.2 Présentation de la méthode

Dans cette partie, nous allons introduire une méthode basée sur des matrices de transition qui consiste, en fonction d'un certain nombre de classes correspondant à des intervalles sur la charge des sinistres, à estimer les probabilités de passer d'une classe à une autre lorsque la maturité du sinistre augmente. Cette méthode introduite dans l'article de MAHON [2005] permet d'obtenir une distribution du nombre de sinistres dans chaque tranche de charge à la fin de leur développement. La différence majeure avec les méthodes de provsionnement décrites dans le chapitre est que celles-ci se basaient sur les montants totaux réglés pour la flotte de sinitres, alors que celle-ci apporte une vision par sinistre. Elle présente l'avantage de prendre en compte l'hétérogénéité du coût des sinistres. En effet, l'étude des sinistres graves en RC automobile nécessitent un soin particulier, les méthodes de provisionnement classiques comme la méthode Chain Ladder ayant souvent tendance à ne pas les traiter convenablement.

L'article de Mahon avait uniquement pour objectif d'être appliqué à des problématiques de réassurance, il n'avait donc pas besoin de calculer la charge ultime des sinistres mais il est possible d'en proposer tout de même une évaluation à l'aide des coûts moyens des sinistres dans chaque classe.

3.2.1 Classes

Afin de définir les différentes classes qui serviront à construire les matrices de transition, il faut choisir les paliers qui serviront de bornes inférieures et supérieures aux intervalles de charge. Nous avons tout d'abord besoin d'un palier à 0 € permettant d'avoir une classe pour les sinistres sans suite. Ensuite, Mahon fixait son premier palier positif à 200 k€ mais ce n'est pas pertinent compte tenu de nos données donc nous faisons le choix de le fixer à 20 k€. De plus il peut être judicieux d'avoir recours à une échelle logarithmique. En effet, les sinistres ayant les plus gros montants sont beaucoup moins présents que les sinistres peu coûteux, donc le fait que l'amplitude des classes augmente soit croissante semble être une bonne idée. Cela facilite également la modélisation par une loi log-normale. Les montants des paliers sont donc multipliés par 2 pour le passage d'une classe à la suivante.

Les paliers supérieurs à 20 k€ sont donc choisis parmi les montants de charge suivants :

- 40 k€
- 80 k€
- 160 k€
- 320 k€
- 640 k€
- 1 280 k€
- 2 560 k€
- 5 120 k€
- 10 240 k€
- etc...

En se basant sur nos données, on se rend compte que le nombre de sinistres ayant une charge supérieure à 2 560 k€ à un moment donné de leur liquidation est extrêmement faible, ce montant est donc défini comme étant notre dernier palier.

Cela nous donne un total de 10 classes, plus une onzième classe appelée "IBNR" (pour Incurred But Not Reported) qui concernera les sinistres tardifs. Ces classes sont récapitulées ci-dessous :

Classe	Borne inf (k€)	Borne sup (k€)
IBNR		
0	$-\infty$	0
1	0	20
2	20	40
3	40	80
4	80	160
5	160	320
6	320	640
7	640	1 280
8	1 280	2 560
9	2 560	$+\infty$

Mahon préférait de plus effectuer une séparation entre les sinistres ouverts et les sinistres clôturés. L'intérêt serait que les classes correspondants aux sinistres terminés soient absorbantes, autrement dit le bloc de matrice de transition qui correspond à ces classes serait égal à la matrice identité. Cependant, le modèle dépend de la maturité des sinistres donc l'absorbance est déjà prise en compte dans le modèle puisque les sinistres finissent tous par être clôturés au bout d'un certain temps. Il ne paraît dans un premier temps pas nécessaire de segmenter les classes selon la position ("en cours" / "terminé") des sinistres. Cela permet également de limiter le nombre de classes.

3.2.2 Hypothèses du modèle

Propriété de Markov

Étant donné que l'objectif est de construire un modèle markovien, il est logique d'y retrouver la propriété fondamentale des chaînes de Markov décrite dans la partie 2.1 que nous rappelons ci-dessous :

$$\forall i_0, \ldots, i_n, j, \mathbb{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n)$$

La suite de variables aléatoires étant considérée ici est bien la classe de nos sinistres. Cela signifie que la classe de gravité dans laquelle se situe un sinistre durant l'année de liquidation $j+1$ ne dépend que de sa classe durant l'année de liquidation j.

Durée maximale de liquidation

Dans la mesure où le modèle est construit d'après les données empiriques dont nous disposons, nous sommes contraints de supposer que les plus vieux sinistres de notre historique ont atteint la fin de leur vie et ne susciterons plus de règlements. Dans notre cas, il s'agit des sinistres survenus en 2004. La dernière clôture considérée étant celle de décembre 2015, nous faisons l'hypothèse que chaque sinistre met au maximum 11 ans pour être complètement liquidé.

Cette hypothèse est discutable. En effet, la durée de liquidation maximale est indexée par la survie des victimes, autrement dit la durée de vie humaine. C'est pourquoi la proposition d'amélioration de la méthode abordée dans le chapitre 5 aura pour objectif de se libérer de cette hypothèse.

3.3 Construction des matrices

Une matrice de transition est construite pour chaque passage d'une année de liquidation à la suivante. Ce point est très important. Certes la probabilité de transition dépend de la classe actuelle du sinistre et pas de ses classes antérieures, ce qui rend bien le modèle markovien. Néanmoins, comme cette probabilité évolue au cours du temps, nous sommes dans le cas d'une chaîne de Markov non-homogène.

Une chaîne de Markov homogène serait donc modélisable par une unique matrice de transition. Ici nous sommes forcés à en construire plusieurs.

Étant donné que les données traitées concernent les sinistres survenus de 2004 à 2015 (le dernier inventaire considéré étant celui du 31 décembre 2015), nous serons en mesure d'obtenir un total de 11 matrices. En effet, comme nous l'expliquions dans la sous-partie 3.2.2, nous sommes obligés de considérer que les sinistres survenus en 2004 sont complètement liquidés en 2015, donc l'ultime correspond à un délai de 11 ans, ce qui implique que la dernière matrice construite à partir des observations empiriques sera celle concernant le passage de la maturité 10 ans à l'ultime. On suppose alors que les matrices de transition associées aux maturités supérieures sont toutes égales à la matrice identité I_{11}.

On note $M_{k,k+1}$ la matrice permettant de passer de la maturité k à la maturité $k+1$ pour $0 \leqslant k \leqslant n$ (avec $n = 10$).

3.3.1 Matrices de transition en nombres de sinistres

Nous présentons ci-dessous la matrice construite à partir des données entre la maturité 0 et la maturité 1 an.

	IBNR	0	1	2	3	4	5	6	7	8	9	Total
IBNR	137	472	1 220	39	23	6	4	-	-	-	-	1 901
0	-	4 139	331	13	6	5	2	1	1	-	-	4 498
1	-	6 626	13 251	328	125	28	18	7	1	1	1	20 386
2	-	53	231	304	97	27	10	3	-	-	-	725
3	-	23	32	58	174	56	25	10	6	3	2	389
4	-	3	3	2	15	73	21	6	-	-	-	123
5	-	-	-	1	1	11	29	8	4	1	1	56
6	-	-	-	-	-	-	2	15	1	-	-	18
7	-	-	-	-	-	-	-	-	4	-	-	4
8	-	-	-	-	-	-	-	-	2	8	3	13
9	-	-	-	-	-	-	-	-	-	-	4	4
Total	137	11 316	15 068	745	441	206	111	50	19	13	11	28 117

Cette matrice a été construite en comptant le nombre de sinistres passant d'un état à un autre entre la maturité 0 et la maturité 1 an.
Par exemple, sur les 20 386 sinistres qui appartenaient à la classe 1 à l'inventaire de leur année de survenance, 328 d'entre eux sont passés dans la classe 2 l'année suivante.

Nous nous sommes bien entendu limités aux années pour lesquelles les données sont disponibles. En l'occurrence, il n'y a que les sinistres de 2015 dont on ne connaisse pas l'état 1 an après leur survenance (i.e. à l'inventaire du 31 décembre 2016) donc la matrice $M_{0,1}$ est construite à partir des données de sinistres survenus entre 2004 et 2014.

3.3.2 Matrices de transition stochastiques

Mais la matrice ci-dessus n'est pas celle que nous voulons obtenir. Il faut maintenant transformer cette matrice en une matrice de probabilités pour construire une chaîne de Markov. Il suffit de diviser chaque "case" par le total de la ligne à laquelle elle appartient. Par exemple, la probabilité qu'un sinistre dans la classe 1 reste dans la classe 1 est égale à 13 251/20 386. Ainsi les probabilités totales de chaque ligne seront égales à 1.

Il y a tout de même un cas ambigu : s'il n'y a aucun sinistre dans la classe de départ, nous ne pouvons pas appliquer cette méthode. Cela ne concerne aucune classe dans l'exemple indiqué mais il est possible que ce soit le cas pour d'autres maturités. On suppose donc arbitrairement que ces classes sont stationnaires, autrement dit si un sinistre avait été dans une de ces classes dans l'état initial, il y serait forcément resté l'année suivante.

La matrice $M_{0,1}$ peut donc désormais être construite. Elle est représentée ci-dessous :

	IBNR	0	1	2	3	4	5	6	7	8	9	Total
IBNR	0,07	0,25	0,64	0,02	0,01	0,00	0,00	-	-	-	-	1,00
0	-	0,92	0,07	0,00	0,00	0,00	0,00	0,00	0,00	-	-	1,00
1	-	0,33	0,65	0,02	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,00
2	-	0,07	0,32	0,42	0,13	0,04	0,01	0,00	-	-	-	1,00
3	-	0,06	0,08	0,15	0,45	0,14	0,06	0,03	0,02	0,01	0,01	1,00
4	-	0,02	0,02	0,02	0,12	0,59	0,17	0,05	-	-	-	1,00
5	-	-	-	0,02	0,02	0,20	0,52	0,14	0,07	0,02	0,02	1,00
6	-	-	-	-	-	-	0,11	0,83	0,06	-	-	1,00
7	-	-	-	-	-	-	-	-	1,00	-	-	1,00
8	-	-	-	-	-	-	-	-	0,15	0,62	0,23	1,00
9	-	-	-	-	-	-	-	-	-	-	1,00	1,00

La matrice se lit de la même manière que la précédente mais avec des probabilités. Par exemple, un sinistre qui était dans la classe 1 a une probabilité égale à environ 0,02 de passer dans la classe 2 entre les maturité 0 et 1 an. La matrice confirme aussi que les totaux de toutes les lignes sont effectivement égaux à 1, ce qui est une propriété de base des chaînes de Markov.

Les matrices $M_{1,2}, M_{2,3}, .., M_{n-1,n}$ (avec $M_{n-1,n} = M_{10,11}$) se construisent de la même manière que la matrice $M_{0,1}$. La seule différence est que nous faisons le choix de prendre en compte plus de données pour les dernières matrices. La matrice $M_{8,9}$ prendra aussi en compte les sinistres survenus en 2003, la matrice $M_{9,10}$ contiendra ceux survenus en 2002 et la matrice $M_{10,11}$ aura en plus ceux de 2001.

Cela nous permet de toujours baser nos matrices sur un minimum de quatre années de survenance pour avoir des données plus crédibles, même si nous ne cherchons pas à évaluer la charge ultime des sinistres survenus avant 2004. Cependant, cette prolongation de l'historique peut ne toujours pas être suffisante.

3.3.3 Matrices de transition à l'ultime

Nous pouvons désormais déduire de ces matrices intermédiaires la matrice permettant de passer d'une maturité quelconque k (pour $0 \leqslant k \leqslant 10$) à l'ultime. Cette matrice est notée $M_{k,\infty}$. On remarque que $M_{10,\infty} = M_{n-1,n} = M_{10,11}$ (en effet, il s'agit de la matrice correspondant à la dernière maturité).

Les propriétés des chaînes de Markov nous permettent de savoir que la matrice "finale" peut être obtenue en effectuant un produit matriciel :

$$M_{k,\infty} = M_{k,k+1} \times M_{k+1,\infty} \quad \text{pour} \quad k \leqslant 9$$

En raisonnant par récurrence, on en déduit que pour $0 \leqslant k \leqslant 10$, on a :

$$M_{k,\infty} = \prod_{j=k}^{n-1} M_{j,j+1}$$

En particulier, $M_{0,\infty} = M_{0,1} \times M_{1,2} \times ... \times M_{n-1,n}$.

Cette matrice $M_{0,\infty}$ est représentée ci-dessous :

	IBNR	0	1	2	3	4	5	6	7	8	9	Total
IBNR	-	0,38	0,58	0,02	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,00
0	-	0,91	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00
1	-	0,44	0,54	0,02	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,00
2	-	0,17	0,46	0,20	0,10	0,05	0,02	0,01	0,00	0,00	0,00	1,00
3	-	0,11	0,23	0,18	0,22	0,13	0,06	0,03	0,01	0,02	0,00	1,00
4	-	0,06	0,11	0,11	0,22	0,29	0,14	0,04	0,01	0,01	0,00	1,00
5	-	0,02	0,06	0,06	0,13	0,28	0,21	0,12	0,03	0,07	0,01	1,00
6	-	0,01	0,02	0,02	0,06	0,19	0,23	0,29	0,07	0,10	0,01	1,00
7	-	0,00	0,00	0,00	0,01	0,06	0,12	0,33	0,14	0,30	0,03	1,00
8	-	0,00	0,00	0,00	0,00	0,01	0,03	0,13	0,12	0,64	0,07	1,00
9	-	0,00	0,00	0,00	0,00	0,00	0,01	0,07	0,16	0,60	0,15	1,00

Cette nouvelle matrice se lit de la même manière que les matrices illustrées précédemment : un sinistre dans la classe 1 lors de son premier inventaire a une probabilité de 0,54 environ de terminer sa liquidation en étant toujours dans la classe 1 (il peut cependant très bien avoir de classe entre temps).

Maintenant que nous avons construit les matrices de transition, il reste à les utiliser pour évaluer la charge ultime et la provision.

3.4 Projection des sinistres

3.4.1 Objectif de la méthode

Cette partie détaillera comment appliquer la méthode. Tout d'abord, nous définissons $V_{i,j} \in \mathbb{R}_+^{11}$ comme étant le vecteur contenant le nombre de sinistres appartenant à chacune de nos onze classes (IBNR, 0, 1, ..., 9) pour l'année d'origine i et la maturité j. On remarque que pour une année i donnée, la somme des éléments de $V_{i,j}$ est constante quelque que soit la maturité j.

L'objectif est de proposer une estimation des vecteurs $V_{i,j}$ pour $i + j > n$, où n est la dernière année comptable pour laquelle les données sont connues, en l'occurrence 2015. Comme le passage d'une maturité à la suivante ne dépend pas des maturités précédentes, nous n'avons besoin que des observations correspondant à l'inventaire 2015 qui sera notre état initial, c'est-à-dire les vecteurs $V_{i,n-i}$.

3.4.2 Sinistres tardifs

Les sinistres tardifs doivent faire l'objet d'une étude particulière. En effet, le nombre de tardifs à l'inventaire 2015 est par définition inconnu.
L'évaluation du nombre de sinistres dans la classe IBNR s'effectue en procédant de la manière suivante :

– on construit le triangle cumulé des nombres de sinistres en développement $N_{i,j}$ pour $i + j \leqslant n$
– on le complète à l'aide de la méthode Chain Ladder pour estimer le nombre ultime de sinistres en développement par année de survenance $\widehat{N}_{i,n}$
– le nombre de sinistres tardifs à l'inventaire 2015 pour l'année d'origine i est alors égal à $\widehat{N}_{i,n}$–$N_{i,n-i}$

Résultat Cette estimation donne un total de 183 sinistres tardifs estimés au 31 décembre 2015 dont voici le détail par année de survenance :

Exercice	Nombre de sinistres tardifs au 31/12/2015
2004	0
2005	0
2006	0
2007	0
2008	0
2009	0
2010	0
2011	1
2012	2
2013	3
2014	12
2015	166

3.4.3 Projection du nombre de sinistres dans chaque classe

Nous sommes donc en mesure de définir tous les vecteurs $V_{i,n-i}$. Il va maintenant falloir prévoir leur évolution pour chaque délai futur jusqu'à l'ultime. Toujours grâce aux propriétés de chaînes de Markov, on sait qu'on peut calculer les estimations des vecteurs futurs avec un produit matriciel :

$$\forall i, j, \widehat{V}_{i,j+1} = \mathrm{E}[V_{i,j+1}|V_{i,j}] = V_{i,j} \times M_{j,j+1}$$

On définit également $V_{i,\infty} = V_{i,n} = V_{i,11}$. La formule ci-dessus est adaptée de la manière suivante :

$$\widehat{V}_{i,\infty} = \mathrm{E}[V_{i,\infty}|V_{i,n-i}] = V_{i,n-i} \times M_{n-i,\infty} = V_{i,n-i} \times \prod_{j=n-i}^{n-1} M_{j,j+1}$$

Nous pouvons donc déterminer des estimations des nombres de sinistres dans chaque tranche pour toutes les années de survenance et toutes les maturités. Si nous avons besoin de connaître l'évolution des sinistres pendant toute la durée de leur liquidation, il va falloir utiliser les calculs permettant de connaître les états intermédiaires.

3.4.4 Évaluation de la charge ultime

Il nous reste maintenant à évaluer les charge futurs en fonction des nombre de sinistres dans chaque classe et à chaque maturité. Dans son article, Mahon se contentait de donner une méthode pour déterminer le nombre de sinistres

présents dans chaque classe mais ne s'intéressait pas à la charge ultime.

Nous utilisons une approche simple se basant sur les coûts moyens. Cela consiste à calculer les coûts moyens des sinistres situés dans chaque classe et à le multiplier par le nombre de sinistres présents dans la tranche. Le calcul du coût moyen est effectué sur l'ensemble des sinistres tous exercices confondus (ouverts et clos), le coût considéré étant celui au 31 décembre 2015.

Ces coûts moyens sont présentés dans le tableau ci-dessous :

Classe	Borne inf (k€)	Borne sup (k€)	Coût moyen (k€)
IBNR	0	0	0
0	$-\infty$	0	0
1	0	20	2,7
2	20	40	27,9
3	40	80	56,7
4	80	160	110,0
5	160	320	217,7
6	320	640	439,0
7	640	1 280	881,6
8	1 280	2 560	1 659,2
9	2 560	$+\infty$	3 714,2

Ce vecteur de coûts moyens est noté $CM \in \mathbb{R}^{11}$. On l'utilise directement pour estimer les paiements cumulés futurs à partir des nombres de sinistres dans chaque classe :

$$\widehat{C}_{i,j} = \widehat{V}'_{i,j} \times CM \quad \text{pour} \quad i+j > n$$

Et nous obtenons donc le triangle des charges complété.

Retraitement des résultats en euros courants

Toutes les données ayant été traitées en euros 2015, il reste désormais à les convertir à nouveau en euros historiques. Les indices d'inflation ne peuvent s'appliquer simplement au triangle des charges complété dans la mesure les charges contiennent elles-même de la provision. Nous ne connaissons que le montant total des règlements futurs pour chaque exercice sans avoir le détail de l'écoulement. Or, les cadences de règlements sont primordiales pour connaître l'inflation à intégrer.

Nous supposons pour cela que la provision évaluée avec cette méthode suit la même cadence d'écoulement que les paiements futurs évalués avec la méthode Chain Ladder dans le chapitre 1. Cela nous permet d'obtenir cette fois-ci le triangle complété des règlements, auquel nous pouvons intégrer l'inflation des années future. Une projection du SMIC sur les années futures nous a permis d'obtenir les indices suivants (ces indices s'appliquent donc aux montants en euros 2015) :

Année de liquidation	Indice de revalorisation
2016	1,004
2017	1,011
2018	1,020
2019	1,032
2020	1,045
2021	1,061
2022	1,078
2023	1,097
2024	1,116
2025	1,137
2026	1,159

Résultats obtenus

À l'aide de la méthode décrite plus haut, les charges ultimes et les provisions obtenues pour chaque année de survenance sont les suivantes :

Année de survenance	Charge ultime (k€)	Provision (k€)
2004	15 783	2 246
2005	18 245	−417
2006	18 412	74
2007	21 574	4 034
2008	21 375	1 233
2009	17 724	3 094
2010	23 211	7 415
2011	20 000	8 459
2012	18 652	7 711
2013	23 457	12 998
2014	20 655	13 538
2015	24 753	22 318
Total	243 840	82 704

Dans le tableau suivant, la provision obtenue avec cette méthode des matrices de transition est comparée à celle obtenue avec le modèle Chain Ladder :

Année de survenance	Provision *Matrices de transition* (k€)	Provision *Chain Ladder* (k€)	Écart
2004	2 246	0	−100,0%
2005	−417	119	−128,6%
2006	74	518	594,4%
2007	4 034	765	−81,0%
2008	1 233	1 513	22,7%
2009	3 094	2 934	−5,2%
2010	7 415	4 500	−39,3%
2011	8 459	4 525	−46,5%
2012	7 711	7 441	−3,5%
2013	12 998	10 599	−18,5%
2014	13 538	15 979	18,0%
2015	22 318	35 562	59,3%
Total	82 704	84 455	2,1%

L'écart observé entre les deux méthodes sur l'exercice 2004 est uniquement du à la différence entre les règlements cumulés et la charge au 31/12/2015 : cela signifie que les charges contiennent encore de la provision pour sinistre à payer.

Sur les autres années de survenance, les écarts peuvent être dus en partie à l'absence de distinction entre les sinistres ouverts et clos, tous les sinistres sont donc projetés, ce qui diminue la précision de l'estimation. Cette distinction sera une des améliorations qu'apportera la méthode présentée dans le chapitre 5.

Chapitre 4

Première amélioration de la méthode

4.1 Préliminaires

Un défaut de la méthode décrite par Mahon est que les classes correspondant aux sinistres les plus coûteux sont déterminées à partir d'un nombre de données empiriques que l'on ne considère pas comme suffisant. Nous allons donc tenter de proposer un modèle permettant de construire les matrices $(M_{j,j+1})_{j\geqslant 0}$ autrement.

Cette nouvelle approche serait basée sur l'idée que les matrices de transition "intermédiaires" convergent vers la matrice identité lorsque la maturité croit, ce qui est logique puisque tous les sinistres finissent forcément par se clôturer au bout d'un certain temps.

$$\lim_{j\to+\infty} M_{j,j+1} = Id$$

De plus, les sinistres attritionnels ont généralement une gestion plus simple et plus rapide. Nous faisons donc l'hypothèse supplémentaire que les premières lignes des matrices, c'est-à-dire celles qui correspondent aux classes initiales les moins élevées convergent plus rapidement que celles correspondant aux "grandes" classes.

L'objectif est donc ici de trouver une tendance qui permettrait de déduire la convergence des classes les plus grandes à partir de celle des classes les plus petites. Cela nous permettrait de construire de nouvelles matrices dont les lignes correspondant aux classes initiales les plus élevées seraient construites non pas empiriquement mais théoriquement (comme nous le verrons, le modèle pourra tout de même être appliqué aux autres lignes).

Il va donc falloir définir $k_0 \in \mathbb{N}$ délimitant les classes en deux catégories : celles dont les observations empiriques sont considérées comme suffisamment nombreuses pour être fiables (i.e. celles contenant les sinistres les moins graves) et celles pour lequel l'étude sera nécessaire.

Il est important dans cette méthode d'augmenter le nombre de classes afin d'avoir plus d'évaluations à partir desquelles déterminer la tendance. Pour cela, on fait le choix de conserver l'échelle logarithmique : des paliers supplémentaires sont obtenus en multipliant chacun des paliers que nous avions déjà dans l'autre méthode par $\sqrt{2}$.

Les nouvelles classes sont présentées dans le tableau ci-dessous :

Classe	Borne inf (en k€)	Borne sup (en k€)
0	$-\infty$	0
1	0	20
2	20	28,3
3	28,3	40
4	40	56,6
5	56,6	80
6	80	113,1
7	113,1	160
8	160	226,3
9	226,3	320
10	320	452,5
11	452,5	640
12	640	905,1
13	905,1	1 280
14	1 280	1 810,2
15	1 810,2	2 560
16	2 560	$+\infty$

Il peut de plus s'avérer utile de passer à un pas de maturité semestriel. Cela nous permettrait de modéliser les tendances à partir de plus de points donc avec plus de précision.

De plus, la classe IBNR est bien entendu conservée. L'étude de cette classe ne diffère pas fondamentalement par rapport à la méthode décrite précédemment dans la sous-partie 3.4.2. Un seul changement est opéré : le Chain Ladder effectué n'est plus annuel mais semestriel pour être cohérent avec la modification décrite ci-dessus. Le nombre de sinistres tardifs estimés au 31 décembre 2015 est désormais égal à 186.

4.2 Présentation du modèle

On appelle $T_{j+1,k}$ la variable aléatoire correspondant à la classe dans lequel appartiendra un sinistre durant le semestre de développement $j+1$ sachant qu'il se situe dans la classe k durant le semestre de développement j.

Cette notation nous permet de déduire que la matrice $M_{j,j+1}$ est égale à :

$$\begin{pmatrix} \mathbb{P}(T_{j+1,0}=0) & \mathbb{P}(T_{j+1,0}=1) & \cdots & \mathbb{P}(T_{j+1,0}=k') & \cdots & \mathbb{P}(T_{j+1,0}=16) \\ \mathbb{P}(T_{j+1,1}=0) & \mathbb{P}(T_{j+1,1}=1) & \cdots & \mathbb{P}(T_{j+1,1}=k') & \cdots & \mathbb{P}(T_{j+1,1}=16) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbb{P}(T_{j+1,k}=0) & \mathbb{P}(T_{j+1,1}=k) & \cdots & \mathbb{P}(T_{j+1,k}=k') & \cdots & \mathbb{P}(T_{j+1,k}=16) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbb{P}(T_{j+1,16}=0) & \mathbb{P}(T_{j+1,16}=k) & \cdots & \mathbb{P}(T_{j+1,16}=k') & \cdots & \mathbb{P}(T_{j+1,16}=16) \end{pmatrix}$$

4.2.1 Hypothèses

Hypothèses de la méthode de Mahon

Les deux hypothèses sur lesquelles la méthode de Mahon reposait, qui étaient décrites dans la sous-partie 3.2.2, sont toujours valables ici. Pour rappel, il s'agissait de supposer :

- d'une part que la classe d'un sinistre durant le semestre de développement $j+1$ ne dépend que de sa classe durant le semestre de développement j (ce qui rend la notation $T_{j+1,k}$ valable)
- d'autre part que les sinistres ont une durée de liquidation qui ne peut excéder 11 ans en raison l'historique de données considéré

Hypothèse supplémentaire

Nous devons faire une hypothèse sur la loi de $T_{j+1,k}$. Une première approche consiste à faire une hypothèse de log-normalité. Une alternative à cette approche sera développée plus loin.

Comme $T_{j+1,k}$ est une variable discrète, cette hypothèse est légèrement adaptée de la manière suivante :

$$T_{j+1,k} = \begin{cases} \max(0, \min(16, \lfloor T'_{j+1,k} \rfloor)) & \text{si} \quad T'_{j+1,k} - \lfloor T'_{j+1,k} \rfloor < 0.5 \\ \max(0, \min(16, \lfloor T'_{j+1,k} \rfloor + 1)) & \text{sinon} \end{cases}$$

où $T'_{j+1,k}$ suit une loi log-normale dont on cherchera à estimer les paramètres.

Deux autres hypothèses sur les moments de $T_{j+1,k}$ sont aussi nécessaires :

$$\lim_{j \to +\infty} \mathrm{E}[T_{j+1,k}] = k \qquad \text{et} \qquad \lim_{j \to +\infty} \mathrm{Var}[T_{j+1,k}] = 0$$

On considère que $\mathrm{E}[T_{j+1,k}] = \mathrm{E}[T'_{j+1,k}]$ et $\mathrm{Var}[T_{j+1,k}] = \mathrm{Var}[T'_{j+1,k}]$.

Choix de k_0

On choisit $k_0 = 7$, c'est-à-dire le plus petit nombre de classes tel que le modèle soit construit à partir de 99,5% des sinistres. Le modèle prendra alors en compte les données correspondant à huit classes (0, 1, ..., 7).
Le montant de charge délimitant le seuil à partir duquel les observations sont considérées comme trop peu nombreuses est égal à 160k€.

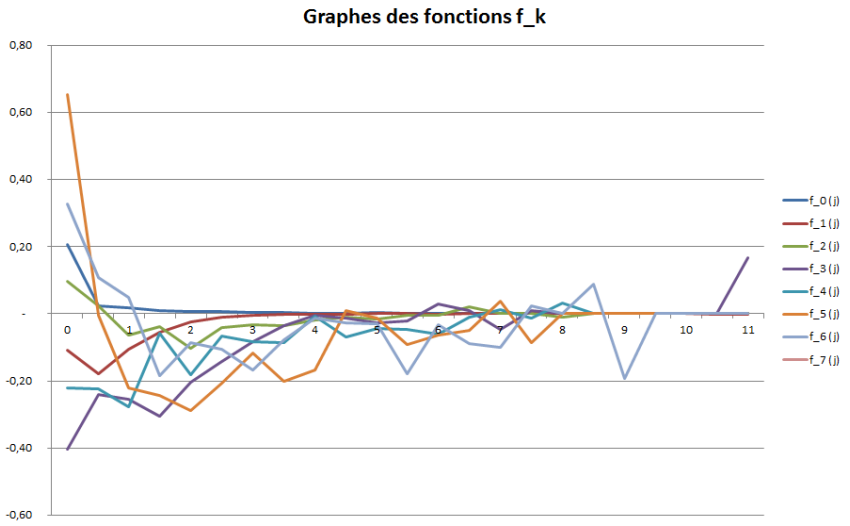
4.2.2 Lissage des moments

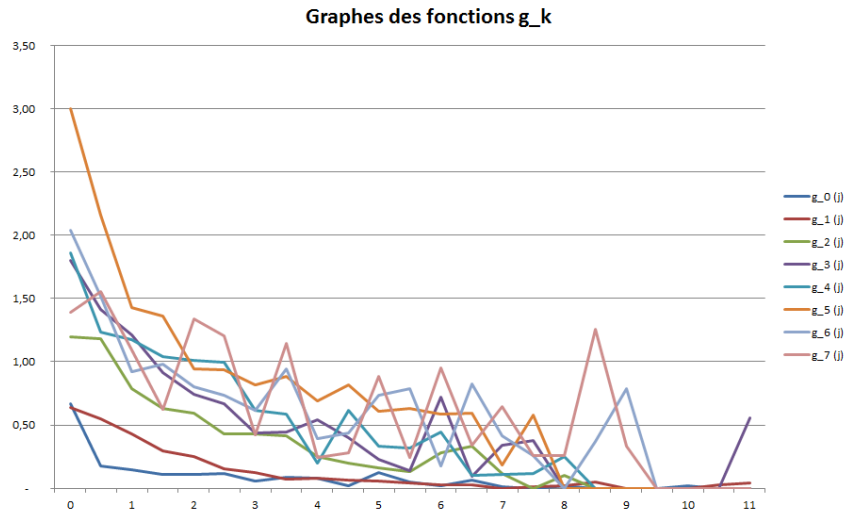
Pour chaque classe $k \leqslant k_0$ donnée, le but est d'obtenir une estimation paramétrique des moments de $T_{j+1,k}$.

Il va falloir pour cela déterminer un lissage des deux fonctions définies par :

- $f_k : j \longmapsto \mathrm{E}[T_{j+1,k}] - k$
- $g_k : j \longmapsto \sqrt{\mathrm{Var}(T_{j+1,k})}$

Nous présentons ci-dessous deux graphiques représentant les fonctions (f_k) et (g_k) en fonction de la maturité pour $k = 0, ..., k_0$.





Il s'agit ainsi de fonctions convergeant vers 0, c'est pourquoi les plusieurs lissages que nous allons tester à partir de différentes fonctionnelles seront faits de telle sorte que toutes les fonctionnelles testées convergent vers 0.

Les différentes fonctionnelles testées sont :

- la famille exponentielle : $h(x) = a\exp(-bx)$
- la famille Weibull : $h(x) = \frac{1}{1-\exp(-a(x+1)^b)} - 1$
- la famille Gamma : $h(x) = a(x+1)^b \times \exp(-(x+1)^c)$
- la famille puissance inverse : $h(x) = a \times (x+1)^{-b}$
- la famille puissance : $h(x) = a^{(b^x)} - 1$

Le fait que ces fonctionnelles convergent vers 0 se vérifie facilement.

Il va maintenant falloir déterminer quelles fonctionnelles sont éligibles pour chaque lissage. Nous basons pour cela sur les variations que sont censées avoir les deux fonctions considérées.

En se basant sur les données empiriques correspondant aux quelques classes les plus faibles, on constate que :

- $\forall k, f_k$ semble présenter une partie décroissante (en étant négative puis positive) puis une partie croissante. Cela signifie qu'au tout début de sa vie, un sinistre est plus susceptible de passer dans une classe plus élevée, ensuite il aura plutôt tendance à passer dans une classe plus faible, enfin il se stabilisera.
- $\forall k, g_k$ semble quant à elle être strictement décroissante. Ce point rejoint l'intuition que l'on pouvait avoir : plus le sinistre est jeune, plus il est susceptible de changer de classe.

Comme l'objectif est que l'estimation théorique de f_k et g_k que nous allons proposer pour $k \leqslant k_0$ (k_0 a été déterminé plus haut) puisse se prolonger afin d'avoir une expression théorique de f_k et g_k pour $k > k_0$. Il est donc nécessaire que tous les lissages de f_k (et respectivement tous les lissages de g_k) soient effectués à partir de la même famille de fonctionnelles.

Mais avant de choisir quelles fonctionnelles serviront pour faire le lissage, nous devons d'abord estimer les paramètres de chaque fonctionnelles pour chaque classe $k \leqslant k_0$.

Estimation des paramètres

Les paramètres de ces fonctionnelles s'estiment à l'aide de la méthode des moindres carrés. Cette méthode consiste à résoudre les problèmes d'optimisation suivants en fonction des paramètres que l'on cherche à estimer :

$$\min \sum_{j=0}^{n} [f_k(j) - \widehat{f}_k(j)]^2 \quad \text{et} \quad \min \sum_{j=0}^{n} [g_k(j) - \widehat{g}_k(j)]^2$$

Afin de simplifier les calculs, on peut également considérer ces problèmes de minimisation à la place (qui sont en réalité identiques aux deux précédents) :

$$\min \sum_{i=1}^{n} [(f_k(j) - 1) - (\widehat{f}_k(j) - 1)]^2 \quad \text{et} \quad \min \sum_{i=1}^{n} [(g_k(j) - 1) - (\widehat{g}_k(j) - 1)]^2$$

La minimisation à l'aide de cette méthode s'effectue généralement en appliquant la fonction logarithme à $f_k(j)$ et à $\widehat{f}_k(j)$ (respectivement à $g_k(j)$ et à $\widehat{g}_k(j)$) qui conserve le minimum étant donné qu'il s'agit d'une fonction bijective croissante. Cela permet de dériver plus facilement la nouvelle expression à minimiser par rapport aux paramètres à estimer, il ne reste plus qu'à annuler les dérivées pour obtenir les paramètres optimaux. Le problème auquel nous sommes ici confrontés est que le logarithme n'est pas toujours défini.

On préfère donc ici l'algorithme de Levenberg-Marquardt : il s'agit d'un algorithme itératif qui est une bonne alternative pour résoudre les problèmes d'optimisation.

L'objectif est, étant donnée une fonction F deux fois dérivable sur $\mathbb{R}^n$, de déterminer $\min\limits_{X \in \mathbb{R}^n} F(X)^2$.

L'algorithme de Levenberg-Marquardt calcule un minimum local, il est donc possible que le résultat diffère selon le point initial choisi.

Données :
- $F : \mathbb{R}^n \to \mathbb{R}$ deux fois dérivables
- $x_0 \in \mathbb{R}^n$: le point initial (ou graine)
- $\epsilon, \alpha, \beta > 0$

On définit le paramètre d'amortissement initial par $\lambda_0 = F(x_0)^2$.

À chaque itération k, on effectue les instructions suivantes tant que la condition $\mid x_{k+1} - x_k \mid > \epsilon$ est vérifiée :

1 On calcule le pas d_k solution de :

$$\lambda Id + \nabla F(x_k) \nabla F(x_k)^{\mathsf{T}} \times d_k = -\nabla F(x_k) F(x_k)$$

où ∇F désigne le gradient de F (car F est à valeur dans $\mathbb{R}$, sinon il aurait fallu prendre à la place la matrice jacobienne) et λ est le paramètre d'amortissement.

2 Si $F(x_k + d_k) \leqslant \gamma F(x_k)$, alors $x_{k+1} = x_k + d_k$.
Sinon : soit m_k le plus petit entier naturel tel que :

$$\phi(x_k + \beta^{m_k} d_k)^2 - \phi(x_k^2) \leqslant \alpha \beta^{m_k} \nabla(x_k)^{\mathsf{T}} d_k$$

et on pose $x_{k+1} = x_k + \beta^{m_k} d_k$

3 On définit $\lambda_{k+1} = F(x_{k+1})^2$.

Lorsque l'algorithme s'arrête, c'est-à-dire lorsque $\mid x_{k+1} - x_k \mid \leqslant \epsilon$, l'estimation du minimum est égale à λ_{k+1} et ce minimum est atteint en x_{k+1}.

Résultats Les paramètres de chaque fonctionnelle estimés avec l'algorithme de Levenberg-Marquardt pour les fonctions (f_k) et (g_k) sont présentés respectivement en annexes B.1.1 et B.1.2.

Choix de la fonctionnelle

Après avoir estimé les paramètres des fonctionnelles de lissage, il va désormais falloir déterminer lesquelles nous ferons le choix de retenir. Pour cela, le critère des moindres carrés est la solution la plus évidente mais comme nous l'avons expliqué plus haut, il est important de rappeler que les $(f_k)_{k=0,\dots,k_0}$ et les $(g_k)_{k=0,\dots,k_0}$ doivent être respectivement lissés à partir des mêmes fonctionnelles. Nous avons donc besoin de définir des critères en plus.

Étant donnée une fonctionnelle donnée que l'on nomme $(\widehat{f_k})_{k=0,\dots,k_0}$ (respectivement $(\widehat{g}_k)_{k=0,\dots,k_0}$), s'il existe $k_1 \in [\![0, k_0]\!]$ tel que l'algorithme de Levenberg-Marquardt ne converge pas pour f_{k_1} (respectivement pour g_{k_1}), alors la fonctionnelle $(\widehat{f_k})_{k=0,\dots,k_0}$ (respectivement $(\widehat{g}_k)_{k=0,\dots,k_0}$) ne pourra pas être retenue pour modéliser les $(f_k)_{k=0,\dots,k_0}$ (respectivement pour les $(g_k)_{k=0,\dots,k_0}$).

Donc pour qu'une fonctionnelle soit retenue, il faut donc non seulement que l'algorithme converge $\forall k \in [\![0, k_0]\!]$ mais aussi que la somme des critères des moindres carrés soit la plus faible. Il s'agit donc de résoudre les problèmes de minimisation suivants :

$$\min_{(\widehat{f_k})} \sum_{k=0}^{k_0} \sum_{j=0}^{n} [\widehat{f_k}(j) - f_k(j)]^2 \quad \text{et} \quad \min_{(\widehat{g_k})} \sum_{k=0}^{k_0} \sum_{j=0}^{n} [\widehat{g}_k(j) - g_k(j)]^2$$

Résultats Voici les résultats que donne le critère des moindres carrés sur les fonctionnelles testées pour les fonctions (f_k) et (g_k) :

Famille	(f_k)	(g_k)
Exponentielle	1,275 4	0,872 6
Weibull	$15{,}8 \times 10^3$	1,070 4
Puissance inverse	0,984 3	1,359 5
Puissance	$40{,}0 \times 10^4$	0,855 7
Gamma	$80{,}0 \times 10^4$	

D'après les résultats ci-dessus, les familles de fonctionnelles à retenir respectivement pour (f_k) et (g_k) seraient idéalement la famille Puissance inverse et la famille Puissance. Or, comme nous le verrons, ces fonctionnelles doivent également permettre d'identifier une tendance sur les paramètres dont elles dépendent, ce qui ne sera pas forcément le cas. Nous pouvons nous autoriser à sélectionner une famille de fonctionnelles plus adaptée à notre besoin si cela est nécessaire.

4.2.3 Estimation de la tendance

Une fois que les lissages des "petites" classes sont bien définis (i.e. choix de la fonctionnelle et estimation des paramètres), nous pouvons passer à l'étape suivante qui consiste à en déduire une tendance permettant de proposer une estimation des "grandes" classes. Pour cela, nous avons besoin d'effectuer un nouveau lissage de fonctions qui concernera cette fois-ci les paramètres du premier lissage.

On note :

$$\begin{cases} \widehat{f}_k(j) = \widehat{f}(j, \ldots, a_k^1, a_k^2, \ldots, a_k^p) \\ \widehat{g}_k(j) = \widehat{g}(j, \ldots, b_k^1, b_k^2, \ldots, b_k^q) \end{cases}$$

Remarque : dans les familles de fonctionnelles que nous avons testées dans notre étude, p et q ne sont jamais supérieurs à 3.

On connait donc pour l'instant les $(a_k^i)_{1 \leqslant i \leqslant p}$ et les $(b_k^i)_{1 \leqslant i \leqslant q}$ pour $k \leqslant k_0$.

Le fait d'obtenir des estimations $(\widehat{a}_k^i)_{1 \leqslant i \leqslant p}$ et $(\widehat{b}_k^i)_{1 \leqslant i \leqslant q} \forall k$ nous permettra d'en déduire respectivement les fonctions $\widehat{f}_k$ et $\widehat{g}_k$ pour $k > k_0$:

$$\begin{cases} \widehat{f}_k(j) = f(j, \widehat{a}_k^1, \widehat{a}_k^2, \ldots, \widehat{a}_k^p) \\ \widehat{g}_k(j) = f(j, \widehat{b}_k^1, \widehat{b}_k^2, \ldots, \widehat{b}_k^q) \end{cases} \qquad \text{pour} \quad k > k_0$$

Ces estimations peuvent être effectuées de différentes manières :
- en choisissant la moyenne
- à l'aide d'une régression linéaire
- en faisant un lissage par l'algorithme de Levenberg-Marquardt comme précédemment

Alternative Une autre option se propose alors : le modèle théorique pourrait très bien s'appliquer également aux données relatives aux "petites" classes (autrement dit celles ayant servis à construire le modèle). En effet, on peut considérer que les expressions théoriques de $\widehat{a}_k^1, \ldots, \widehat{a}_k^p$ et $\widehat{b}_k^1, \widehat{b}_k^2, \ldots, \widehat{b}_k^q$ sont valables pour les valeurs de $k \leqslant k_0$.

Nous obtenons alors des expressions de $\widehat{a}_k^1, \ldots, \widehat{a}_k^p$ et $\widehat{b}_k^1, \widehat{b}_k^2, \ldots, \widehat{b}_k^q$ quelque soit la classe initiale k. Les $(\widehat{f}_k(j))$ et les $(\widehat{g}_k(j))$ sont donc également connues pour toutes les valeurs de k.

Cela nous permet de connaître des estimations de $\mathrm{E}[T_{j+1,k}]$ et de $\mathrm{Var}[T_{j+1,k}]$ pour toutes les valeurs de j et k. En effet, on rappelle que $f_k(j) = \mathrm{E}[T_{j+1,k}] - k$ et $g_k(j) = \mathrm{Var}[T_{j+1,k}]$.

Remarque : Il s'agit précisément de cette étape de la méthode qui a posé problème et a empêché de poursuivre l'étude. Nous reviendrons sur cet élément dans la partie 4.3.

4.2.4 Construction des nouvelles matrices

Les hypothèses concernant la distribution de $T_{j+1,k}$ sont rappelées ci-dessous :

Hypothèse 1

$$T_{j+1,k} = \begin{cases} \max(0, \min(16, \lfloor T'_{j+1,k} \rfloor)) & \text{si} \quad T'_{j+1,k} - \lfloor T'_{j+1,k} \rfloor < 0.5 \\ \max(0, \min(16, \lfloor T'_{j+1,k} \rfloor + 1)) & \text{sinon} \end{cases}$$

où $T'_{j+1,k}$ suit une loi log-normale.

Hypothèse 2 $\mathrm{E}[T_{j+1,k}] = \mathrm{E}[T'_{j+1,k}]$ et $\mathrm{Var}[T_{j+1,k}] = \mathrm{Var}[T'_{j+1,k}]$.

Ces hypothèses nous permettent de calculer les paramètres de la loi log-normale à l'aide de la méthode des moments :

$$\mu_{j+1,k} = \ln(\mathrm{E}[T_{j+1,k}]) - \frac{1}{2} \ln \left(1 + \frac{\mathrm{Var}(T_{j+1,k})}{\mathrm{E}[T_{j+1,k}]^2} \right)$$

$$\sigma_{j+1,k} = \sqrt{\ln \left(1 + \frac{\mathrm{Var}(T_{j+1,k})}{\mathrm{E}[T_{j+1,k}]^2} \right)}$$

On peut maintenant proposer des évaluations de $T_{j+1,k}$ pour $k > k_0$. Nous faisons le choix de procéder en effectuant des simulations. En effet, les hypothèses précédentes nous permettent de déduire facilement une simulation de $T_{j+1,k}$ à partir d'une simulation d'une loi log-normale. Si on choisit un nombre de simulations égal à N, on note $T^i_{j+1,k}$ la i-ième simulation de $T_{j+1,k}$.

On pose alors :

$$\mathbb{P}(\widehat{T}_{j+1,k} = k') = \frac{1}{N} \sum_{i=1}^{N} 1_{\{T^i_{j+1,k} = k'\}}$$

Ce qui nous permet de construire les nouvelles matrices $(M_{j,j+1})_{j \geqslant 0}$. Comme nous l'avions précisé plus haut, deux alternatives sont possibles :

- construire les lignes de la matrice correspondant aux classes $k \leqslant k_0$ à partir des données empiriques (de manière identique à la première approche) et construire les lignes supérieures à k_0 avec les résultats des simulations fournies par le modèle
- ou alors se fier exclusivement au modèle et donc effectuer les simulations sur toutes les classes, celles inférieures ou égales à k_0 n'auront alors servi qu'à paramétrer le modèle

à l'exception de la classe "IBNR" qui est dans tous les cas construite de la façon décrite précédemment.

Une fois que les matrices sont définies, l'évaluation du nombre de sinistres dans chaque classe à l'ultime s'effectue en procédant de manière identique à la partie précédente.

4.2.5 Alternative à l'hypothèse de log-normalité

L'hypothèse de log-normalité n'étant pas toujours vérifiée, nous allons donc décrire une approche pouvant servir d'alternative.

Cette approche est basée sur l'approximation décrite dans l'article de CORNISH et FISHER [1938]. L'approximation de Cornish-Fisher est une estimation

des quantiles d'une variable aléatoire donnée en fonction de ses quatre premiers moments. Dans le cas précédent, nous ne prenions en compte que l'espérance et la variance, ici nous avons également besoin de l'asymétrie (skewness) et de l'aplatissement (kurtosis).

En effet, il parait plausible que ces moments soient pris en compte. Les sinistres étant dans une petite classe vont plus facilement pouvoir passer dans une classe plus grande alors que ce serait plutôt le contraire pour les ceux appartenant à une grande classe, ce qui pourrait traduire une information sur l'asymétrie des distributions étudiées. De plus, on pourrait supposer que l'aplatissement décroit lorsque la classe augmente.

Approximation de Cornish-Fisher

Soit X une variable aléatoire. On pose :
- $\mu = \mathrm{E}[X]$
- $\sigma^2 = \mathrm{Var}(X)$
- $s = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}^3\right)\right]$ (skewness)
- $k = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}^4\right)\right]$ (kurtosis)

On rappelle que le quantile d'ordre α de X est noté $\mathrm{VaR}_\alpha(X)$ (pour "Value at Risk). L'approximation de Cornish-Fisher des quantiles de X est donnée par l'expression suivante.

$$\widehat{\mathrm{VaR}}_\alpha^{CF}(X) = \mu + \sigma \times \left[\begin{array}{l} \Phi^{-1}(\alpha) + \frac{(\Phi^{-1}(\alpha)^2-1)s}{6} \\ + \frac{(\Phi^{-1}(\alpha)^3-3\Phi^{-1}(\alpha))(k-3)}{24} - \frac{(2\Phi^{-1}(\alpha)^3-5\Phi^{-1}(\alpha))s^2}{36} \end{array} \right]$$

où $\Phi^{-1}(\alpha)$ désigne le quantile d'ordre α de la loi normale centrée réduite.

Réarrangement Il existe des cas où l'approximation de Cornish-Fisher peut poser problème. En effet, l'approximation des quantiles fournie par cette approche n'est pas nécessairement croissante, ce qui est théoriquement impossible. Pour une variable aléatoire X donnée, le résultat suivant doit être toujours vérifié :

$$\forall \alpha, \beta \in]0,1[, \alpha \leqslant \beta \Rightarrow \mathrm{VaR}_\alpha(X) \leqslant \mathrm{VaR}_\beta(X)$$

Il existe un moyen de corriger cela détaillé dans CHERNOZHUKOV et *al.* [2007].

On note $f :]0,1[\to \mathbb{R}$ la fonction estimée par l'approximation de Cornish-Fisher. La méthode consiste à réarranger les valeurs de $f(x)$ de telle sorte qu'elles soient ordonnées dans l'ordre croissant. Pour cela, une autre fonction est introduite :

$$F_f(y) = \int_0^1 1_{\{f(u) \leqslant y\}} \mathrm{d}u$$

Il s'agit de la fonction de répartition de $f(X)$ où X suit une loi uniforme sur $[0,1]$. Cette fonction nous sert à définir la fonction donnant le réarrangement de l'approximation de Cornish-Fisher :

$$\begin{array}{rccl} f^\star : &]0,1[& \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \inf\{y \in \mathbb{R} / F_f(y) \geqslant x\} \end{array}$$

Dans le cas où les estimations des quantiles ne sont pas croissantes, ce sera donc cette expression qui sera prise en compte.

Description de l'approche

Lissage du skewness et du kurtosis Cette approche nécessite d'effectuer un lissage sur le skewness et sur le kurtosis en procédant de manière similaire à ce nous avions fait pour l'espérance et pour la variance. Les graphes représentant le skewness et le kurtosis de notre distribution pour les classes $k \leqslant 7$ sont présentés en annexe B.2.

L'algorithme de Levenberg-Marquardt est encore une fois la méthode utilisée afin d'y parvenir.

Comme nous avions obtenu des estimations de $\mathrm{E}[T_{j+1,k}]$ et de $\mathrm{Var}(T_{j+1,k})$, les nouveaux lissages nous permettent de pouvoir également proposer des estimations de $\mathrm{E}\left[\left(\frac{T_{j+1,k}-\mathrm{E}[T_{j+1,k}]}{\sqrt{\mathrm{Var}(T_{j+1,k})}}\right)^3\right]$ et de $\mathrm{E}\left[\left(\frac{T_{j+1,k}-\mathrm{E}[T_{j+1,k}]}{\sqrt{\mathrm{Var}(T_{j+1,k})}}\right)^4\right]$ pour chaque valeur de j et de k.

Estimation des quantiles Nous pouvons désormais appliquer la formule ci-dessus pour obtenir les quantiles de $T_{j+1,k}$, qui seront directement utilisés pour évaluer les probabilités de transition afin de construire les nouvelles matrices.

Comme $T_{j+1,k}$ est une variable entière prenant ses valeurs dans $[\![0, 16]\!]$, la formule de Cornish-Fisher est adaptée de la manière suivante :

$$\widehat{\mathrm{VaR}}_\alpha^{CF}(T_{j+1,k}) = \begin{cases} \max(0, \min(16, \lfloor \mathrm{VaR}_\alpha^{CF}(T_{j+1,k}) \rfloor)) \\ \qquad\qquad \text{si} \quad \mathrm{VaR}_\alpha^{CF}(T_{j+1,k}) - \lfloor \mathrm{VaR}_\alpha^{CF}(T_{j+1,k}) \rfloor < 0.5 \\ \max(0, \min(16, \lfloor \mathrm{VaR}_\alpha^{CF}(T_{j+1,k}) \rfloor)) \qquad \text{sinon} \end{cases}$$

Construction de la matrice Quant à la construction de la matrice, la démarche est légèrement différente que lorsque nous avions fait l'hypothèse de log-normalité dans la mesure où il n'est pas nécessaire d'effectuer des simulations.

Soit $N > 1$ un entier naturel. On considère les quantiles d'ordre $\frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}$ (les quantiles d'ordre 0 et 1 n'étant pas défini pour la loi normale), soit un total de $N-1$ pris en compte. On ne choisit donc plus le nombre de simulations comme précédemment mais le nombre de quantiles.

Les nouvelles probabilités de passage (qui servent à construire les nouvelles matrices de transition) sont données par la formule suivante :

$$\mathbb{P}(T_{j+1,k} = k') = \frac{1}{N-1} \sum_{i=1}^{N-1} 1_{\{\widehat{\mathrm{VaR}}_{\frac{i}{N}}^{CF}(T_{j+1,k})=k'\}}$$

Limites

Cette méthode présente quelques problèmes. En effet, l'approximation de Cornish-Fisher n'est définie que pour des petites valeurs de s et de $(k-3)$, c'est-à-dire pour des variables aléatoires ayant des distributions proches de la loi normale (en effet, une loi normale a un skewness nul et un kurtosis égal à 3).

Comme c'est le cas dans notre étude (nous reviendrons sur ce point plus tard), des valeurs trop élevées pour le skewness et le kurtosis augmentent considérablement la probabilité de se retrouver dans les extrêmes et cela risque de nous poser problème.

4.3 Application et défauts de la méthode

4.3.1 Résultats

Comme nous l'expliquions dans l'introduction du mémoire, cette méthode ne fourni les résultats que nous espérions. La principale difficulté rencontrée lors de l'application de cette méthode est la difficulté à trouver un lissage convenable, en particulier pour les fonctions (f_k).

Lissage des fonctions (g_k)

Nous nous intéressons tout d'abord aux fonctions (g_k) qui, pour rappel, sont définies par :

$$g_k(j) = \sqrt{\mathrm{Var}(T_{j+1,k})}$$

Nous indiquions dans la section 4.2.2 que la famille de fonctionnelles fournissant le meilleur lissage selon le critère des moindres carrés était la famille Puissance. Nous rappelons ci-dessous les paramètres de lissage estimés par l'algorithme sur cette famille de fonctionnelles (ces résultats viennent du tableau présenté en annexe B.1.2) :

k	$\widehat{a}_k$	$\widehat{b}_k$
0	1,63	0,27
1	1,68	0,63
2	2,27	0,74
3	2,74	0,77
4	2,77	0,78
5	3,71	0,78
6	2,70	0,83
7	2,46	0,87

Or, il n'est pas du tout aisé de définir une tendance permettant d'estimer les paramètres $(\widehat{a}_k)$ et $(\widehat{b}_k)$ pour $k > 7$. C'est pourquoi nous sommes forcés de choisir une autre famille.

La famille Exponentielle semble être une alternative convenable. En effet, son critère des moindres carrés est de 0,873 (contre 0,856 pour la famille Puissance). Les paramètres de lissage obtenus avec l'algorithme sont les suivants :

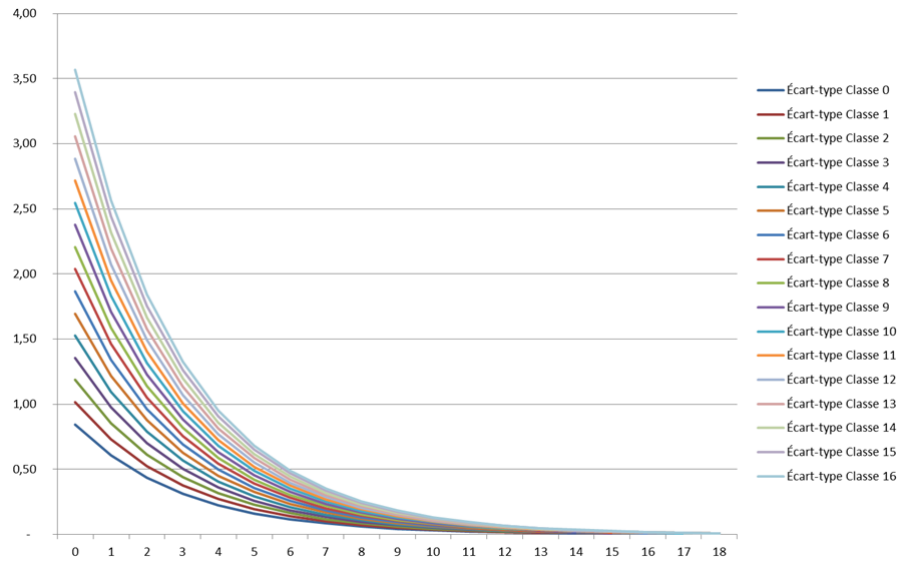
k	$\widehat{a}_k$	$\widehat{b}_k$
0	0,63	1,58
1	0,67	0,53
2	1,23	0,37
3	1,68	0,34
4	1,72	0,32
5	2,56	0,34
6	1,62	0,24
7	1,42	0,18

Les paramètres $(a_k)_{k\geqslant k_0}$ ci-dessus s'adaptent convenablement à une régression linéaire ($R^2 = 0{,}76$), ce que nous avons donc utilisé pour estimer les $(\widehat{a}_k)\forall k$. Les paramètres $(\widehat{b}_k)$ ont été quant à eux estimés par la moyenne des $(b_k)_{k\geqslant k_0}$ à cause de leur faible volatilité (écart-type de 0,11 si on exclut $\widehat{b}_0$) et sont donc constants.

Cela nous permet de modéliser, pour $k = 0, ..., 16$ et pour toute maturité j :

$$\widehat{g}_k(j) = \widehat{a}_k \exp(-\widehat{b}_k j) \quad \text{où} \quad \begin{cases} \widehat{a}_k = 0{,}762 + k \times 0{,}194 \\ \\ \widehat{b}_k = 0{,}330 \end{cases}$$

Les graphes des fonctions $j \longmapsto \widehat{g}_k(j)$ estimées par le modèle pour l'ensemble des classes k sont représentés ci-dessous :



Lissage des fonctions (f_k)

Un problème se pose pour le lissage de l'espérance. Les fonctionnelles testées donnent globalement des lissages satisfaisants selon le critère des moindres carrés, en particulier la famille exponentielle et la famille puissance inverse. Mais contrairement aux lissages obtenus pour les fonctions (g_k), les paramètres des fonctionnelles ne sont pas du tout exploitables. Leur évolution a une forme très volatile qui ne permet pas d'estimer ceux correspondant aux classes supérieures,

ni par un lissage, ni par une régression linéaire. Il est donc difficile de conclure ici. En effet, si nous voulions essayer d'estimer les paramètres posant problème, l'incertitude de ces estimations serait beaucoup trop élevée.

Skewness et kurtosis

Enfin, même s'ils ne sont nécessaires que pour la deuxième approche, le skewness et le kurtosis sont encore plus complexes à modéliser à cause de leur évolution très volatile. Et comme nous l'avons précisé précédemment, le fait qu'ils prennent des valeurs particulièrement élevées rend l'approximation de Cornish-Fisher beaucoup moins efficace. En effet, elle donnerait des probabilités beaucoup trop élevées de se retrouver dans la classe 0 et dans la classe 16, quelque soit la classe de départ.

4.3.2 Défauts de la méthode

Au vu des résultats décrits précédemment, cette méthode n'a pas du tout été satisfaisante. En effet, elle reposait sur des hypothèses beaucoup trop optimistes et générait de très grandes incertitudes.

Nous faisons le choix d'abandonner cette méthode. La partie suivante développera une autre méthode qui se basera également sur l'article de Mahon avec pour objectif de l'améliorer mais cette autre approche procédera de manière très différente à ce que nous avons tenté ici.

Chapitre 5

Deuxième amélioration de la méthode

Pour rappel, la méthode des matrices de transition décrite par Mahon consistait à construire une chaîne de Markov non-homogène. Cela signifie que les probabilités de transition représentant l'évolution de la classe dans laquelle se situera un sinistre dans le futur dépendent non seulement de sa classe actuelle mais également sa maturité. C'est pourquoi nous avions besoin de construire une matrice de transition par couple d'unités de temps (années ou semestres de liquidation) représentant une transition.

Cela ne nous permet pas d'appliquer toutes les propriétés dont nous disposons sur les chaînes de Markov. Par exemple, la première méthode nous forçait à supposer que les sinistres arrivaient tous à la fin de leur liquidation durant l'intervalle de temps dont nous disposons, hypothèse que nous n'avons pas besoin de faire ici.

Notre objectif pour cette nouvelle approche est de construire un modèle basé sur une chaîne de Markov homogène qui ne prendra pas en compte les maturités de liquidation.

Cela signifie que la matrice de transition qui va être construite ne dépendra pas des maturités des sinistres. Nous faisons le choix de rester sur un pas de maturité semestriel comme dans la méthode décrite dans la partie précédente même si cela ne paraît pas aussi important que dans l'autre approche.

Hypothèse Comme nous venons de le préciser, l'hypothèse relative à la durée maximale de liquidation des sinistres décrite dans la sous-partie 3.2.2 n'est plus nécessaire ici. Néanmoins, nous allons un nouvelle fois avoir affaire à un modèle markovien, ce qui signifie que la propriété fondamentale des chaînes de Markov (également rappelée dans la sous-partie 3.2.2) est toujours aussi primordiale. Pour rappel, cela signifie que la classe dans laquelle se situe un sinistre durant le semestre de liquidation $j+1$ ne dépend que de sa classe durant le semestre de liquidation j. Le seul élément de cette propriété qui diffèrera sera la façon dont les classes seront définies mais nous verrons cela plus loin.

Définition Dans toute la suite, le terme *transition* désignera un couple de classes : celle dans laquelle un sinistre donné se situe durant le semestre de liquidation j et celle dans laquelle ce même sinistre ce situe durant le semestre de liquidation $j+1$.
Dans notre modèle, nous ne nous intéressons pas aux sinistres mais aux transitions. De plus, toutes les transitions sont considérées de la même manière, c'est-à-dire indépendamment de j.

5.1 Nouvelles variables

Il est évident qu'on ne peut pas se contenter de garder les mêmes classes en ne modifiant que la façon de compter les sinistres. Il va falloir changer de procédure. Les classes ne seront plus constituées d'une unique variable (l'intervalle dans lequel se situe la charge), il va falloir en ajouter.

5.1.1 Position du sinistre

Tout d'abord, bien que le choix d'introduire des classes absorbantes n'ait pas été retenu précédemment, cela paraît indispensable ici. En effet, dans l'étude précédente, la durée de vie des sinistres était bornée par la durée sur laquelle s'étendaient les données. Mais ici nous avons besoin de distinguer clairement les sinistres selon qu'ils soient susceptibles d'évoluer ou non.

Pour cela, nous utilisons une variable remplie par les gestionnaires de sinistres. Il s'agit de la position du sinistre : elle est notée E si le sinistres est en cours, T si le sinistre est terminé.

Il faut noter qu'il existe quelques imperfections dans les données : certains sinistres notés comme clôturés peuvent parfois continuer à être modifiés. Afin que l'étude soit crédible, nous calculerons une fois que toutes les classes seront complètement la probabilité pour qu'un sinistre clos change quand même de classe en espérant que ces anomalies ne représentent qu'une part négligeable de notre flotte de sinistres.

5.1.2 Variable d'état

Nous faisons également le choix d'introduire une variable supplémentaire. L'idée est que cette variable donne une indication non seulement sur l'âge du sinistre mais également sur la significativité qu'elle joue sur les transitions de classe (sinon on ferait directement dépendre les classes de la maturité).

Choix de la variable

Plusieurs variables ont été envisagées :
- le nombre de victimes
- le nombre de victimes vivantes ou décédées
- les types de victimes (piéton, cycliste, conducteur)
- les filières de gravité des victimes : il s'agit d'un entier compris entre 1 et 4 que les gestionnaires attribuent aux victimes en fonction des données à

leur disposition, la filière 4 étant la plus grave (cet indicateur est spécifique à la MACSF)
- les taux d'IPP (incapacité partielle et permanente) des victimes
- le nombre de pièces reçues et traitées

Parmi ces variables, nous privilégions celles qui évoluent au cours du temps dans la mesure où sont plus susceptibles d'avoir un impact sur l'évolution du sinistre au cours de sa liquidation. Les variables parmi celles-ci-dessus qui sont variables dans le temps sont :
- le nombre de victimes vivantes ou décédées
- les filières de gravité des victimes
- les taux d'IPP des victimes
- le nombre de pièces reçues et traitées

Intuitivement, le nombre de pièces est une variable que nous considérons comme pouvant satisfaire nos attentes. En effet, on peut penser que le fait de recevoir régulièrement des pièces pour un sinistre peut indiquer qu'il est susceptible de continuer à être réglé donc de changer de classe. Il est tout de même important de valider ce choix à l'aide d'une régression logistique qui sera décrite plus loin.

Choix des classes correspondantes

La question que nous nous posons pour l'instant concerne la segmentation qu'il va falloir effectuer afin de définir des classes dépendant du nombre de pièces reçues et traitées pour un sinistre donné.

Nous utilisons pour cela un algorithme de classification basé sur les centroïdes que nous décrivons ci-dessous :

On note $X_1, \ldots, X_n$ nos observations.

1. On définit un nombre k de classes voulues et des centroïdes initiaux $C_1, \ldots, C_k \in \mathbb{R}$ distincts.
2. Chaque observation X_i est associée au centroïde ayant la distance la plus faible au sens de la valeur absolue, on note alors $c(X_i)$ le centroïde qui minimise $|X_i – C_j|$ pour $j = 1, ..., k$.
3. De nouveaux centroïdes sont définis par la moyenne des observations associées à chaque centroïde de départ : $C'_j = \frac{\sum_{i=1}^{n} X_i 1_{\{c(X_i)=C_j\}}}{\sum_{i=1}^{n} 1_{\{c(X_i)=C_j\}}}$.
4. Si $\forall j = 1, \ldots, k, C'_j = C_j$, alors l'algorithme s'arrête et les classes sont définies par les X_i associés aux mêmes centroïdes. Sinon, on recommence à partir de l'étape 2 en considérant que les $C'_1, \ldots, C'_k$ sont les nouveaux centroïdes initiaux.

Cette méthode s'effectue à l'aide de la procédure Fastclus de SAS. Concernant le nombre de classes souhaité, nous le fixons arbitrairement à trois. Néanmoins, si l'algorithme donne une classification telle qu'une ou plusieurs classe contiennent

trop peu de données (généralement il s'agit des plus élevées), on s'autorise à augmenter le nombre de classes sur lequel s'effectue l'algorithme, puis à regrouper les classes négligeables dans la classe la plus proche.

Résulats obtenus

La méthode ci-dessus donne des résultats nous permettant d'effectuer la segmentation suivante :

A les sinistres pour lesquels la mutuelle a reçu (et traité) entre 0 et 6 pièces

B les sinistres pour lesquels la mutuelle a reçu (et traité) entre 7 et 38 pièces

C les sinistres pour lesquels la mutuelle a reçu (et traité) plus de 39 pièces

5.2 Changement des seuils de charge

Le fait de faire dépendre les classes de plus d'une variable a pour conséquence d'augmenter considérablement le nombre de classes. Il serait donc préférable de diminuer le nombre de paliers concernant les charges des sinistres. Nous allons donc essayer de choisir des seuils cohérents permettant d'alléger le modèle.

5.2.1 Premiers seuils arbitraires

Tout d'abord, le seuil à 0 € pour décrire les sinistres sans suite est conservé. Le plus haut palier est quant à lui fixé à 1 000 k€. Ce choix semble pertinent étant donné qu'il s'agit du montant à partir du lequel la MACSF passe ses sinistres en réassurance. Cela nous conduira, lors de l'évaluation de la charge ultime, à estimer beaucoup plus simplement la charge nette de réassurance. Un sinistre dont le coût est supérieur à 1 000 k€ sera donc réglé 1 000 k€.

On pourrait être tenté de se limiter à trois classes de charge pour segmenter les sinistres : les sans suite, ceux compris entre 0 € et 1 000 k€ et ceux passés en réassurance donc supérieurs à 1 000 k€.

Mais un problème est susceptible d'être rencontré. Le nombre de sinistres passant de la troisième à la deuxième classe est beaucoup plus élevé que ceux passant de la deuxième à la troisième classe. On risque donc d'avoir un nombre insuffisant de sinistres se retrouvant dans la troisième classe à l'ultime.

Un ou plusieurs paliers intermédiaires sont donc tout de même nécessaires. Afin de trouver ces paliers, nous allons appliquer méthodes de segmentation étant souvent utilisées en provisionnement non-vie et notamment sur la garantie RC corporelle pour segmenter les sinistres (la plupart du temps pour séparer les sinistres attritionnels des sinistres graves). Ces méthodes seront l'objet des sous-parties suivantes.

5.2.2 Première approche de détermination de seuils : méthode du boni mali

On rappelle que le boni mali entre deux dates t_1 et t_2 est défini par la différence entre l'estimation de la charge ultime en t_2 et l'estimation de la charge

ultime en t_1.

Cette première méthode est basée sur l'hypothèse selon laquelle les sinistres les plus graves ont généralement tendance à générer du mali et que les sinistres attritionnels génèrent quant à eux plutôt du boni. En effet, la plupart du temps les sinistres graves ont une liquidation plus lente que les autres, ainsi il peut paraître intuitif d'imaginer que ces sinistres sont souvent sous-provisionnés.

La méthode consiste donc à classer les sinistres pour une survenance donnée (à condition qu'elle soit suffisamment ancienne pour que les sinistres soient bien écoulés) du plus grave au moins grave (la gravité étant définie en fonction de la charge et non des règlements). Les boni mali des sinistres sont ensuite étudiés en fonction de ce classement : le boni mali considéré est celui correspondant à l'inventaire du 31 décembre 2015 et l'inventaire de la survenance.

En traçant le graphe correspondant, on remarque que la courbe est globalement située en dessous de 0 pour les premiers sinistres et passe au dessus de 0 à partir d'un certain moment : cela confirme que les sinistres graves ont tendance à générer du mali alors que les moins graves génèrent du boni. La courbe des boni mali étant assez volatile, on peut tracer la moyenne mobile (en prenant par exemple les 10 valeurs précédentes et les 10 valeurs suivantes) afin d'obtenir un graphe plus lisible.

Le seuil que nous choisirons sera donc le montant du sinistre à partir duquel on considère que les boni mali seront globalement positifs. On peut réitérer la méthode pour d'autres survenances afin de vérifier le premier résultat obtenu : on devrait normalement arriver au même résultat à chaque fois.

En pratique, la méthode des boni mali ne s'avère pas donner des résultats très crédibles. Les seuils obtenus en appliquant la méthode sur différentes années de survenance sont toujours beaucoup trop faibles. Le postulat sur lequel se base la méthode (les sinistres graves génèrent du mali et les non-graves du boni) n'est peut-être pas pertinent.

5.2.3 Méthodes basées sur la théorie des valeurs extrêmes

Contrairement à la méthode précédente qui était spécifique au provisionnement, les méthodes statistiques présentées ci-dessous sont plus générales dans la mesure où elles peuvent s'appliquer à d'autres problématiques.

Rappels sur la théorie des valeurs extrêmes

La théorie des valeurs extrêmes propose certains outils statistiques permettant des fournir des résultats sur la loi au-delà d'un certain seuil, aussi appelée la distribution des excès (i.e. la variable aléatoire $X - u|X > u$). En effet, on se base sur le postulat que nos observations suivent une loi à queue épaisse, c'est-à-dire que les sinistres graves (donc situés sur la queue de distribution) ont une distribution plus importante que si les observations suivaient une loi exponentielle. Par exemple, la loi log-normale, la loi de Pareto et la loi de Pareto généralisée (Generalized Pareto Distribution ou GPD) sont des lois à queues

épaisses. La distribution GPD est définie par :

$$G_{\xi,\sigma}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ \\ 1 - \exp(-\frac{x}{\sigma}) & \text{si } \xi = 0 \end{cases}$$

pour $x \geqslant 0$ si $\xi \geqslant 0$ et $0 \leqslant x \leqslant -\frac{\sigma}{\xi}$ si $\xi < 0$.

On va donc s'intéresser à l'indice de queue ξ de la distribution. Cet indice mesure l'épaisseur de la queue : plus sa valeur absolue est élevée, plus les extrêmes sont "lourds" dans la distribution initiale. Cet indice est le même que dans la distribution GEV (Generalized Extreme Values i.e. distribution généralisée des valeurs extrêmes) dont nous rappelons la définition ci-dessous :

$$H_{\xi,\mu,\sigma}(x) = \begin{cases} \exp\left\{-\left(1 + \xi \times \frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right\} & \text{si } \xi \times \frac{x-\mu}{\sigma} > 0 \text{ et } \xi \neq 0 \\ \\ \exp\left\{-\exp(-\frac{x-\mu}{\sigma})\right\} & \text{si } \xi = 0 \end{cases}$$

La distribution GEV peut être égale à trois distributions différentes selon la valeur de ξ :

- si $\xi > 0$, il s'agit d'une loi de Fréchet de paramètre $\alpha = \frac{1}{\xi}$
- si $\xi < 0$, il s'agit d'une loi de Weibull de paramètre $\alpha = -\frac{1}{\xi}$
- si $\xi = 0$, il s'agit d'une loi de Gumbel

Théorème de Gnedenko Si $X_1, ..., X_n$ est une suite de variables aléatoires indépendantes et identiquement distribuées (comme c'est le cas dans notre échantillon), et qu'il existe deux suites $a_n > 0$ et $b_n \in \mathbb{R}$, et une fonction de répartition H non dégénérée telles que :

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leqslant x\right) = H(x) \quad \text{où} \quad M_n = \max(X_1, ..., X_n)$$

alors H suit une distribution GEV (on dit alors que la loi des (X_i) appartient au domaine d'attraction de H).

Le théorème de Pickands fait quant à lui le lien entre la distribution GEV et la distribution de excès.

Théorème de Pickands Soit F une fonction de répartition appartenant au domaine d'attraction d'une loi GEV.
On note $\omega_F = \sup\{x/F(x) < 1\}$ et F_u la fonction de répartition conditionnelle des excès définie par $F_u(x) = \mathbb{P}(X - u \leqslant x | X > u)$.
Le théorème nous assure que $\lim_{u\to\omega_F} F_u(x) = G_{\xi,\sigma}(x)$.

QQ-plot

Afin de s'assurer que la distribution à laquelle nous ans affaire est effectivement à queue épaisse, nous pouvons observer le QQ-plot de la distribution. On

rappelle qu'un QQ-plot est un graphe dans lequel sont tracés les quantiles empiriques basés sur les observations en fonction des quantiles de la loi attendue. Ainsi, si les données suivent effectivement la loi avec laquelle le QQ-plot a été tracé, le graphe obtenu sera linéaire.

Pour tracer le QQ-plot dans le cadre de la théorie des valeurs extrêmes, la loi de référence est la loi exponentielle.

Il faut donc tracer les points $(F^{-1}(\frac{i}{n+1}), X_{(i)})$ où les $X_{(i)}$ sont les montants de chaque sinistre classés tels que $X_{(1)} \leqslant ... \leqslant X_{(n)}$ et F^{-1} est la fonction quantile de la loi exponentielle, c'est-à-dire $F^{-1}(\frac{i}{n+1}) = -\ln(1 - \frac{i}{n+1})$.

Comme on veut que les données suivent une loi à queue épaisse, on s'attend à ce que le QQ-plot ait une forme concave.

Fonction d'espérance résiduelle

La fonction d'espérance résiduelle, aussi appelée fonction moyenne des excès (ou FME) est définie (en fonction du seuil u) par :

$$\mathrm{FME}(u) = \mathrm{E}[X - u | X > u]$$

Cette fonction est estimée empiriquement par la moyenne des excès des sinistres dépassant le seuil u :

$$\widehat{\mathrm{FME}}(u) = \frac{\sum_{i=1}^{n}(X_i - u)}{\sum_{i=1}^{n} 1_{X_i > u}}$$

Il existe une propriété assurant que si X suit une distribution GPD de paramètres (ξ, σ), alors $(X - u | X > u)$ suit une loi GPD de paramètres $(\xi, \sigma + \xi u)$, ce qui permet d'en déduire que pour $\xi < 1$, il existe une constant $\beta(u)$ telle que pour tout $v \geqslant u$ on ait :

$$\mathrm{E}[X - v | X > v] = \mathrm{FME}(v) = \frac{\beta + \xi(v - u)}{1 - \xi}$$

Nous pouvons ainsi en déduire que le seuil u est convenable si la partie du graphe de la FME située après u est de forme affine de manière suffisamment stable. Donc le seuil de segmentation des sinistres choisi est la plus petite valeur de u satisfaisant cette condition. On considère alors qu'un sinistre dont le montant est situé dans la partie affine fait partie de la queue de distribution.

Cette méthode permet également de proposer une estimation de ξ puisque la pente de la partie affine est censée être égale à $\frac{\xi}{1-\xi}$.

Estimateurs de l'indice de queue ξ

Nous allons maintenant présenter quelques estimateurs (dépendant du seuil choisi) basés sur la théorie des valeurs extrêmes permettant d'approximer le paramètre de queue de la distribution de Pareto ou GPD. Leurs calculs nécessitent de classer les observations par ordre décroissant (contrairement au calcul du QQ-plot où les sinistres étaient ordonnés dans l'ordre croissant) : les observations

sont notées $X_{(1)}, ..., X_{(n)}$ et sont telles que $X_{(1)} \geqslant ... \geqslant X_{(n)}$. Les estimateurs sont alors calculés en fonction du nombre de sinistres graves que l'on considère (plus il y a de sinistres graves, plus le seuil est faible).

La méthode est la même pour chaque estimateur. Si la distribution des excès converge bien vers une distribution GPD, l'estimateur devrait logiquement être relativement stable à partir d'un certain nombre de sinistres considérés : il est en effet naturel que l'estimation soit très variable si le nombre de sinistres est trop faible. En effet, si on ne prend pas toute la queue de distribution, il n'est pas étonnant que l'indice de queue ne puisse pas être correctement estimé.

On note k le nombre de sinistres considérés comme graves et u le seuil associé à ce nombre de sinistres.

Estimateur de Hill

$$\widehat{\xi}^{Hill}(u) = \frac{1}{k} \sum_{i<k} \ln \left(\frac{X_{(i)}}{X_{(k)}} \right)$$

Il s'agit de l'estimateur le plus répandu. Il produit généralement des estimations efficaces, par contre il a le défaut de n'être valable que pour les distributions de Fréchet, c'est-à-dire si $\xi > 0$. De plus, on peut montrer que l'estimateur de Hill est identique à l'estimateur du maximum de vraisemblance pour la loi de Pareto.

Il existe un test permettant de vérifier si les observations suivent bien une distribution de Fréchet. Comme la fonction de répartition de la loi de Fréchet est égale à $F_\alpha(x) = \exp(-x^{-\alpha})$, on en déduit que $\ln(\ln(\frac{1}{F_\alpha(x)})) = -\alpha \ln(x)$. Donc si on trace les points $\left(\ln(x), \ln(\ln(\frac{1}{\widehat{F}(x)}) \right)$ où $\widehat{F}$ est notre fonction de répartition empirique, on devrait obtenir des points alignés.

Pour tester s'il s'agit d'une distribution de Gumbel, nous pourrions appliquer le même raisonnement et en déduire que cette fois-ci ce sont les points $\left(x, \ln(\ln(\frac{1}{\widehat{F}(x)}) \right)$ qui seraient censés être alignés.

Estimateur de Pickands

$$\widehat{\xi}^{Pickands}(u) = \frac{1}{\ln 2} \times \ln \frac{X_{(k)} - X_{(2k)}}{X_{(2k)} - X_{(4k)}}$$

Cet estimateur est quant à lui censé être le plus volatile puisqu'il ne dépend que de trois paramètres, mais il a l'avantage par rapport à l'estimateur de Hill de pouvoir être utilisé même si $\xi = 0$.

Estimateur DEdh ou estimateur des moments

$$\widehat{\xi}^{Moments}(u) = \widehat{\xi}^{Hill(1)}(u) + 1 - \frac{1}{2}\left(1 - \frac{(\widehat{\xi}^{Hill(1)}(u))^2}{\widehat{\xi}^{Hill(2)}(u)}\right)^{-1}$$

où $\widehat{\xi}^{Hill(j)}(u) = \frac{1}{k}\sum_{i<k}(\ln(\frac{X_{(i)}}{X_{(k)}}))^j$ pour $j = 1, 2$.

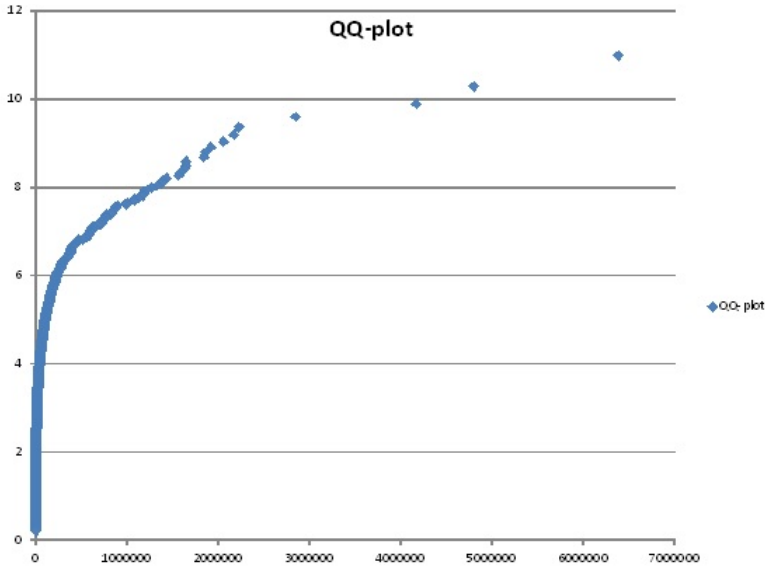
Tout comme l'estimateur de Hill, cet estimateur n'est valable que pour les distributions de Fréchet.

5.2.4 Application des méthodes

Comme nous l'avions précisé, la méthode des boni mali n'est pas retenue. Nous ne considérons donc que les résultats des méthodes présentées dans la sous-partie précédente.

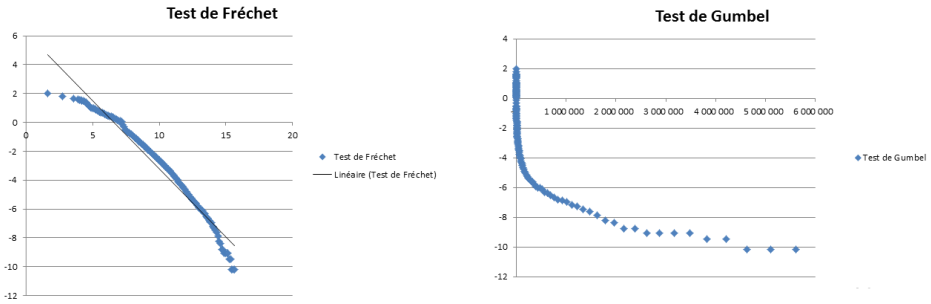
Tests de la distribution

QQ-plot Nous commençons par tracer le QQ-plot exponentiel à partir de nos observations pour vérifier que la distribution des excès est à queue épaisse, sinon les estimations ne pourront pas être valables :



On remarque que le QQ-plot a une forme concave. Cela semble confirmer l'hypothèse selon laquelle nous avons bien une distribution à queue épaisse, comme la loi log-normale, la loi de Pareto ou la distribution GPD.

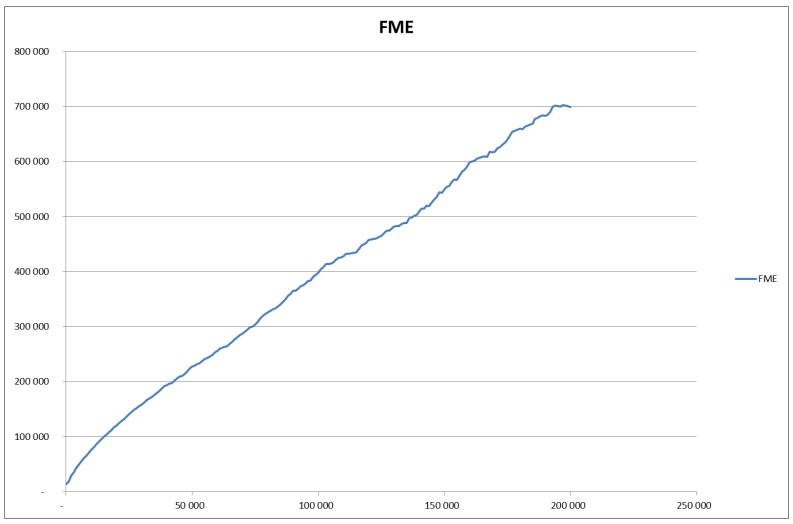
Test de la distribution de Fréchet Il est également préférable d'effectuer dès maintenant le test permettant de vérifier si notre échantillon suit bien une distribution de Fréchet. Afin de valider le résultat de ce test, nous pouvons aussi appliquer l'autre test pour s'assurer qu'il ne s'agit pas d'une distribution de Gumbel. Les résultats de ces tests sont présentés ci-dessous :



Les points du test de la distribution de Fréchet sont assez alignés donc l'hypothèse devrait être vérifiée. Cela semble se confirmer par le fait que le test de la distribution de Gumbel ne soient pas du tout alignés. On peut donc considérer que l'estimateur de Hill, et par extension l'estimateur des moments (qui seront calculés plus loin), sont valables pour nos données.

Fonction d'espérance résiduelle

Le graphe représentant le fonction moyenne des excès en fonction du seuil considéré est tracé ci-dessous.

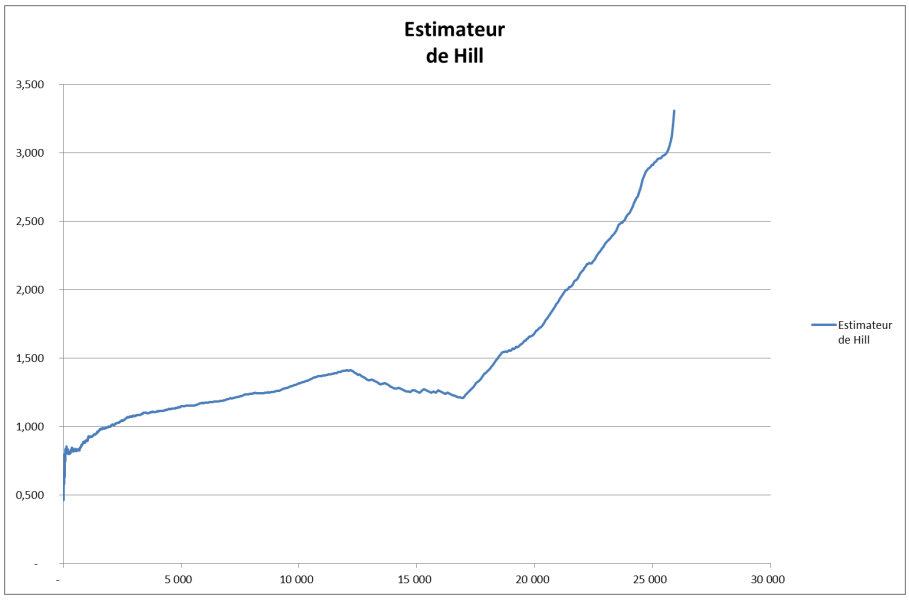


Comme nous pouvons le constater, ce graphe n'est pas exploitable. En effet, le principe de cette méthode est de déterminer le plus petit montant possible tel que le graphe de la FME ait une forme affine. Or, notre application nous donne un graphe qui semble à peu près linéaire dès le début (le coefficient de détermination R^2 de la régression linéaire calculé à partir de l'ensemble des données est très proche de 1). Il est donc difficile de conclure à l'aide de cette méthode.

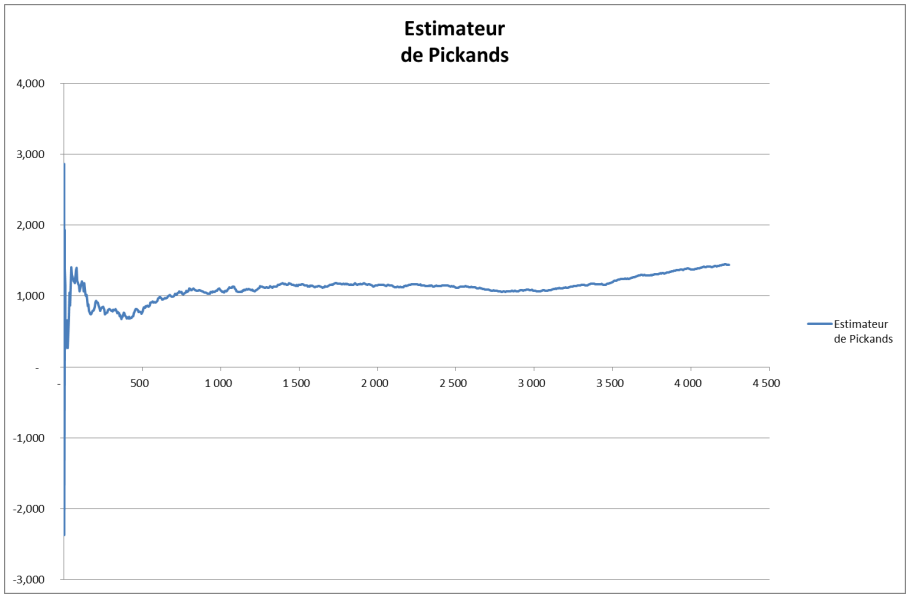
Estimation de l'indice de queue

Les graphes des estimateurs de Hill, de Pickands et DEdh en fonction du nombre de sinistres considérés comme graves sont représentés ci-après.

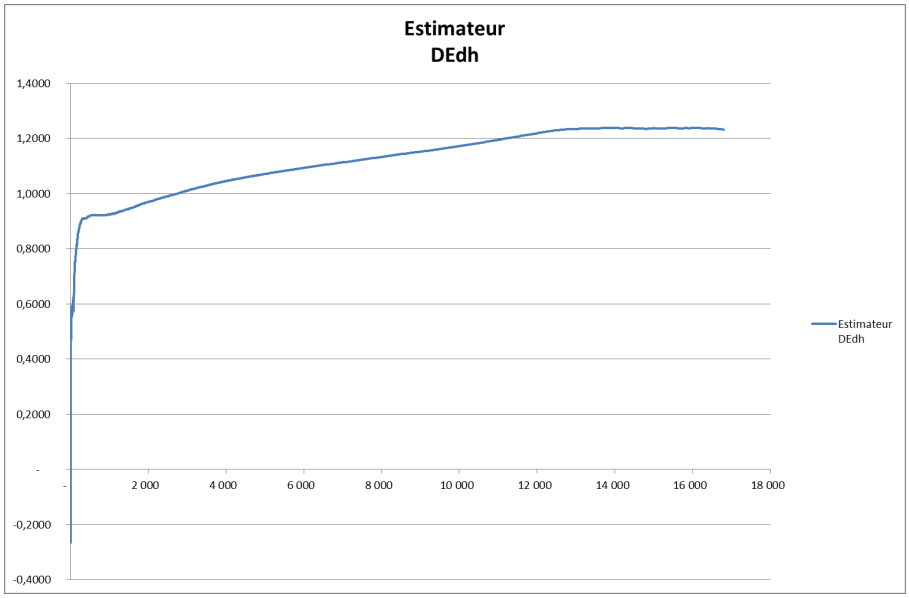
Estimateur de Hill



Estimateur de Pickands



Estimateur des moments (ou estimateur DEdh)



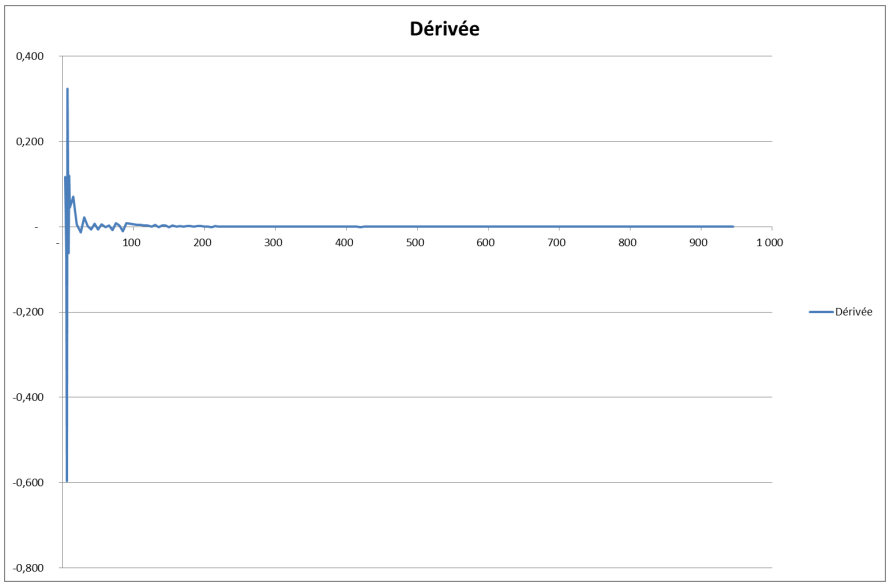
Choix des seuils

L'estimateur de Hill fournissant des résultats très étonnants, et l'estimateur de Pickands étant censé être le plus volatil des trois, nous choisissons de retenir l'estimateur des moments. Les résultats ne permettent pas de mettre en évidence graphiquement un seuil de manière évidente donc nous allons procéder autrement : nous allons dans un premier temps calculer la dérivée théorique de l'estimateur puis chercher différents stades à partir desquels cette dérivée s'approche de 0.

La dérivée théorique est calculée de la manière suivante. Si on note $(u_n)_{n\geqslant 0}$ les valeurs telles que les estimateurs $\widehat{\xi}^{Moments}(u_n)$ sont calculés (on note que $u_0 = 2$ car l'estimateur des moments n'est pas défini pour $u \leqslant 1$. La dérivée est déterminée par :

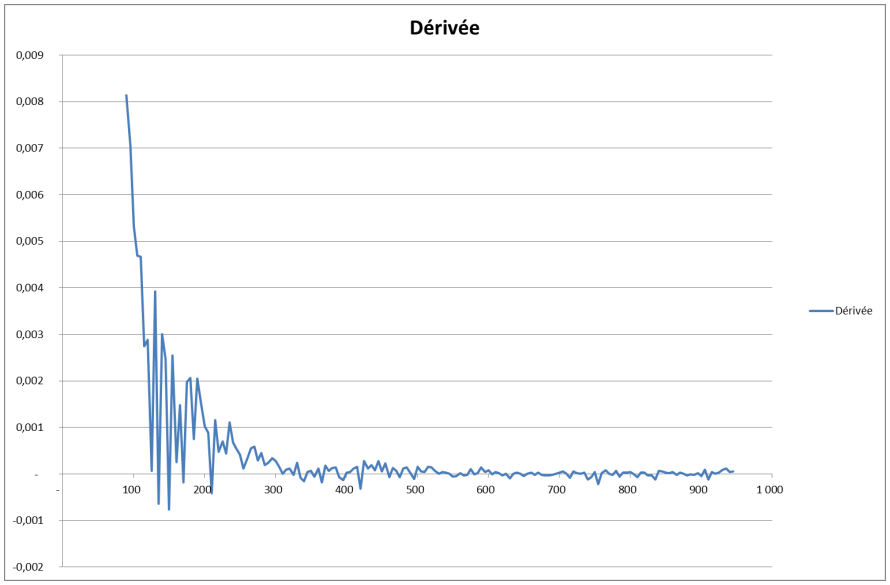
$$\frac{\mathrm{d}}{\mathrm{d}u}\widehat{\xi}^{Moments}(u_n) = \frac{\widehat{\xi}^{Moments}(u_n) - \widehat{\xi}^{Moments}(u_{n-1})}{u_n - u_{n-1}}$$

Le graphe de cette dérivée est représenté ci-dessous :



On observe que la dérivée se rapproche fortement de 0 à partir du point correspondant à 95 sinistres graves considérés, c'est-à-dire un seuil égal à une charge de 427 k€ que l'on arrondi à 400 k€.

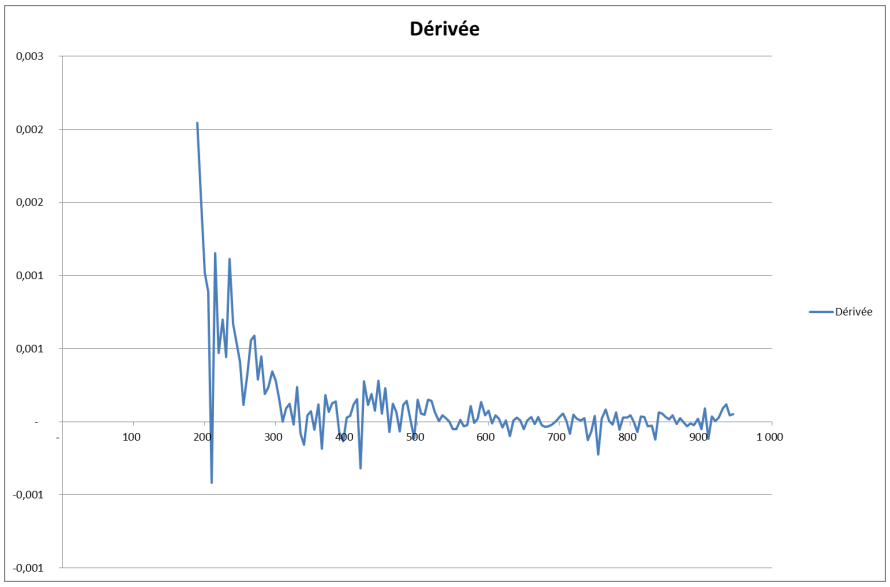
Cependant il s'agit d'une observation graphique, donc peu précise et il peut donc paraître judicieux de tracer le graphe de la dérivée privé des points correspondants aux 95 sinistres les plus graves. Cela pourrait permettre de déterminer un nouveau seuil à l'aide de la nouvelle échelle. Ce graphe est représenté ci-dessous :



Un nouveau palier s'observe effectivement à partir de 195 sinistres graves,

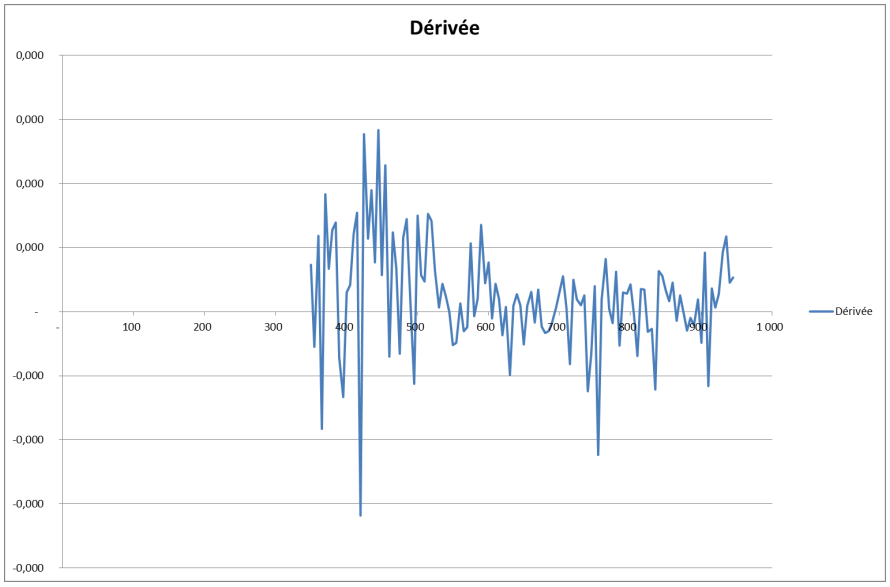
c'est-à-dire 210 k€ de charge que l'on arrondi à 200 k€.

Nous réitérons maintenant le processus en traçant le graphe de la dérivée privée des 195 sinistres les plus graves.



Cette fois-ci le nouveau palier semble être aux alentours des 350 sinistres, autrement dit une charge de 124 k€ que l'on arrondi à 100 k€.

Nous poursuivons la démarche en traçant cette fois-ci le graphe de la dérivée privée de 350 sinistres.



Cette fois, ci, nous n'observons pas réellement de tendance ni de convergence vers 0 permettant d'en déduire un palier supplémentaire. La méthode nous a

donc fourni trois seuils permettant de délimiter les charges des sinistres : 400k€, 200k€ et 100k€.

5.2.5 Conclusion : nouvelle segmentation des charges

Cela nous donne un total de six nouvelles classes de charges intégrées dans le modèle qui sont donc indiquées dans le tableau suivant :

Classe	Borne inf (k€)	Borne sup (k€)
0	$-\infty$	0
1	0	100
2	100	200
3	200	400
4	400	1 000
5	1 000	$+\infty$

Cette segmentation peut sembler questionnable dans la mesure où, comme nous l'avons vu dans la section 3.1.2, plus de 95% des sinistres hors sans suite ont une charge inférieure à 100 k€. Cependant, les sinistres de cette tranche sont peu volatils, la charge ultime de ces sinistre peut donc être estimée de manière suffisamment fiable. D'autre part, bien que ces sinistres représentent une part importante du nombre de sinistres total, leur part dans la charge globale est beaucoup moins significative. L'estimation de la provision a donc un enjeu plus important sur les tranches correspondant aux sinistres supérieurs à 100 k€, c'est pourquoi le choix de cette segmentation nous semble tout de même pertinent.

5.3 Nouvelles classes

Les parties précédentes nous on permis de définir des nouvelles classes dépendant de plusieurs variables. Pour rappel, ces variables sont :

- la charge du sinistre
- la position du sinistre
- le nombre de pièces reçues et traitées pour le sinistre

Classe	Charge (en k€)
0 ••	$]-\infty, 0]$
1 ••	$]0, 100]$
2 ••	$]100, 200]$
3 ••	$]200, 400]$
4 ••	$]400, 1\,000]$
5 ••	$]1\,000, +\infty[$

Classe	Nombre de pièces
•A•	$[\![0, 6]\!]$
•B•	$[\![7, 38]\!]$
•C•	$[\![39, +\infty[\![$

Classe	Position
•• E	En cours
•• T	Terminé

Exemple : La classe 3BE contient les sinistres en cours, dont la charge est comprise entre 200k€ et 400k€, et pour lesquels la MACSF a reçu et traité entre 7 et 38 pièces.

Cela nous donne un total de $2 \times 3 \times 6 = 36$ nouvelles classes, ce qui fait un total de 37 classes avec la classe IBNR qui est conservée, comme dans les deux approches ayant déjà été abordées (pour rappel, le nombre de tardifs estimés au 31 décembre 2015 est égal à 186).

5.4 Validation des classes

Maintenant que des classes ont été définies, la question qui se pose est de s'assurer que ces classes soient pertinentes. Le but de cette partie est d'une part de vérifier que le nombre de sinistres clos changeant de classe entre deux semestres de développement soit négligeable, et d'autre part de valider le choix du nombre de pièces comme "variable d'état" ou de proposer une autre variable s'il en existe une mieux adaptée.

5.4.1 Sinistres clos

Comme nous l'avions indiqué, la base de données à partir de laquelle le modèle est construit possède quelques imperfections. En effet, il existe des sinistres notés comme des sinistres clos mais qui continuent à recevoir des règlements ou des pièces de gestion. Ces sinistres sont donc susceptibles de changer de classe, ce qui pose problème dans la mesure où l'intérêt de faire dépendre la classe de la position était d'avoir des classes absorbantes.

Nous allons donc compter le nombre de fois que les sinistres de notre base dont la classe de départ est une classe correspondant à un sinistre terminé (i.e les classes $\bullet\bullet$T) changent de classe. Sur l'ensemble des semestres de survenance ou de liquidation, cela arrive seulement 220 fois sur un total de plus de 271 000 transitions considérées. Les transitions erronées représentent donc seulement 0,08% de la base complète.

Ce rapide calcul nous permet de nous rendre compte que la proportion de sinistres changeant de classe après avoir appartenu à une classe $\bullet\bullet$T est très faible. Lorsque la matrice de transition de la chaîne de Markov sera construite, les classes $\bullet\bullet$T seront "forcées" pour qu'elles soient bien des classes absorbantes comme nous l'avions voulu.

5.4.2 Régression logistique

Nous voulons maintenant vérifier que le choix du nombre de pièces comme variable supplémentaire est pertinent. La régression logistique est un moyen efficace d'y parvenir.

Description du modèle

Il s'agit d'un modèle linéaire généralisé permettant de mettre en évidence la significativité de plusieurs variables explicatives sur une variable à expliquer. La variable à expliquer est une variable qualitative à K modalités où $K \geqslant 2$, que l'on note $Y \in \{y_1, \ldots, y_K\}$. Si $K = 2$, on dit que la régression logistique est binaire, dans le cas contraire le modèle est dit multinomial (l'étude dont il question ici se place dans le cas multinomial).

Les variables explicatives peuvent quant à elles être aussi bien qualitatives que quantitatives et sont notées $X = (X_1, \ldots, X_p)^\mathsf{T}$.

On défini $\pi_k = \mathbb{P}(Y = y_k|X)$. On a naturellement la relation $\sum_{k=1}^{K} \pi_k = 1$. Une des modalités de Y doit de plus être définie comme la modalité de référence. Nous choisissons ici de manière arbitraire la K-ième modalité.

Le modèle de régression logistique s'écrit de la manière suivante :

$$\forall k \in \{1, \ldots, K-1\}, \ln\left(\frac{\pi_k}{\pi_K}\right) = \beta_{0,k} + \beta_{1,k}X_1 + \ldots + \beta_{p,k}X_p$$

Cette hypothèse permet d'en déduire les expressions de $\pi_1, ..., \pi_K$ suivantes :

$$\begin{cases} \pi_k &= \dfrac{\exp\left(\beta_{0,k} + \beta_{1,k}X_1 + \ldots + \beta_{p,k}X_p\right)}{1 + \exp\left(\beta_{0,k} + \beta_{1,k}X_1 + \ldots + \beta_{p,k}X_p\right)} \quad \text{pour} \quad 1 \leqslant k < K \\ \\ \pi_K &= 1 - \sum_{k=1}^{K-1} \pi_k \end{cases}$$

Les paramètres du modèle $(\beta_{i,k})_{(i,k)\in[\![1,p]\!]\times[\![1,K-1]\!]}$ s'interprètent comme une mesure d'association de X par rapport à Y. En d'autres termes, plus $\|\beta_{i,k}\|$ est élevé, plus la variable X_i du vecteur X influe sur $\mathbb{P}(Y = y_k)$. Ces paramètres peuvent être estimées par la méthode du maximum de vraisemblance.

Application

Dans notre étude, la variable à expliquer est le couple de classes représentant la transition d'un sinistre entre deux semestres de liquidation. Il s'agit donc bien d'une variable qualitative qui pourrait prendre un maximum de 37^2 modalités mais certaines d'entre elles ne sont pas possibles en pratique.

Objectif L'idée est d'inclure la classe de départ parmi les variables explicatives. Si cette variable explicative s'avère être la plus significative du modèle, on pourra considérer que le choix de classes défini dans la sous-partie précédente est validé.

Il s'agit encore une fois d'une variable qualitative, cette fois-ci pouvant prendre jusqu'à 37 modalités. Une de ces modalités devant être considérée comme modalité de référence, il y a 36 paramètres à estimer pour cette variable mais seulement un d'entre eux sera non nul. Cela s'explique naturellement à l'aide d'un exemple. Si on considère la transition entre la classe 0AE et la classe 0AT, la seule classe de départ pouvant être significative pour expliquer cette transition est évidemment la classe 0AE, c'est pourquoi les paramètres associés à toutes les autres modalités sont nécessairement nuls.

Autres variables explicatives Il faut également définir précisément les autres variables explicatives qui seront présentes dans le modèle. Elles correspondent à celles que nous avions déjà sélectionnées dans la sous-partie 5.1.2 :

- le nombre de victimes vivantes
- le nombre de victimes décédées
- le nombre de victimes dont la filière de gravité est égale à 4
  (pour rappel, il s'agit d'un indicateur spécifique à la MACSF attribué par les gestionnaires de sinistres qui correspond à un entier compris entre 1 et 4)
- le nombre de victimes dont la filière de gravité est comprise entre 1 et 3
  (la somme des deux variables précédentes n'est pas forcément égales au nombre total de victimes puisqu'il existe des cas où la filière de gravité n'est pas renseignée dans les données)
- le nombre de victimes dont le taux d'IPP est égal à 0 %
- le nombre de victimes dont le taux d'IPP est supérieur ou égal à 40 %
  (le seuil de 40% a été déterminé avec la méthode de segmentation décrite dans la sous-partie 5.1.2)
- le maximum des taux d'IPP de toutes les victimes du sinistre

Toutes les variables énumérées ci-dessus sont des variables quantitatives. Un seul paramètre par variable est donc à estimer.

L'estimation des paramètres du modèle par la méthode du maximum de vraisemblance des paramètres du modèle a été effectuée à l'aide de la procédure Logistic du logiciel SAS.

Résultats

Les résultats provenant de la procédure Logistic de SAS correspondent à une liste des estimations $(\widehat{\beta}_{i,k})_{(i,k)\in[\![1,p]\!]\times[\![1,K-1]\!]}$ des paramètres du modèle. Nous pouvons donc construire à partir de cette sortie SAS la liste des équations du modèle de régression logistique. Ces équations sont présentées en annexe C.

Remarque : Certains retraitements ont été effectués à partir des résultats obtenus avec la procédure de SAS pour construire les équations du modèle. En effet, le logiciel estime non nuls certains paramètres étant justement censés être nuls. Cependant leur erreur type associée étant très élevée par rapport aux autres estimations, nous ne considérons pas ces résultats comme pertinents.

Il reste maintenant à vérifier que, pour la k-ième des équations du modèle, le paramètre correspondant à la classe précédente est bien le plus élevé en valeur absolu (mis en part le paramètre $\beta_{0,k}$ bien entendu puisqu'il n'est associé à aucune variable). On constate que cela ne produit pas dans la totalité des cas. Mais si on pondère chacune des équations par le nombre de transitions considérées, la condition est vérifiée dans 98% des cas, ce qui est un résultat très convenable.

Conclusion D'après les résultats de la régression logistique, la façon dont les classes ont été définies est pertinente dans la mesure où pour expliquer qu'un sinistre passe d'une classe à une autre entre deux semestres de liquidation,

aucune autre variable n'est plus significative que la classe dans laquelle il se trouvait au départ.

5.5 Chaîne de Markov

5.5.1 Construction de la matrice

Une fois que toutes les classes sont bien définies, et surtout qu'elles peuvent être considérée comme pertinente, la construction de la matrice s'effectue de manière relativement similaire à précédemment, à l'exception qu'aucune distinction n'est faite entre les maturités des sinistres considérés dans le modèle.

Comme expliqué au début de la partie, le comptage s'effectue en prenant en compte toutes les transitions d'une classe à la suivante lorsque le pas de maturité augmente d'une unité mais sans se préoccuper de l'âge des sinistres. Ainsi, un sinistre est compté une fois par transition, ce qui implique que le nombre de fois qu'un sinistre est compté est plus élevé pour les plus vieux sinistres que les plus récents, mais cela ne constitue pas un problème.

Rappel : Le pas de maturité considéré est toujours semestriel.

La matrice construite de cette manière est représentée en annexe D (on rappelle que les lignes correspondent aux classes de départ et les colonnes aux classes d'arrivée).

Comme précédemment, il reste à passer cette matrice de nombre de transitions sous la forme d'une matrice stochastique, c'est-à-dire une matrice de probabilités. On conserve la condition que nous avions mis en place concernant les classes ne contenant pas de sinistres dans nos observations (comme c'est le cas ici de la classe 5AT) : ces classes sont supposées absorbantes.

De plus, comme expliqué dans la sous-partie 5.4.1, on remarque dans la matrice de nombres qu'il y a certaines transitions ayant pour classe de départ une classe • • T et pour classe d'arrivée une autre classe • • T (i.e le bloc inférieur droit de la matrice), ce qui n'est pas normal puisque la lettre T correpond aux sinistres terminés, donc ne pouvant plus changer de classe. Par exemple, on observe 9 transitions de la classe 0BT vers la classe 1BT. Nous avions considéré ces anomalies comme négligeables, ces classes sont donc "forcées" de façon à être des classes absorbantes.

La matrice stochastique qui est l'unique matrice de transition associée à la chaîne de Markov de notre modèle est également représentée en annexe D.
Cette matrice est notée M.

La matrice M peut être décrite selon ses quatre "blocs" :

1. Le bloc supérieur gauche contient les probabilités que les sinistres en cours restent en cours au prochain semestre
2. Le bloc supérieur droit contient les probabilités que les sinistres en cours se soient clôturés au prochain semestre

3. Le bloc inférieur gauche est nul : il correspondrait aux sinistres terminés qui redeviendraient "en cours"
4. Le bloc inférieur droit est égal à la matrice identité I_{18} : ce sont les sinistres terminés qui restent forcément dans leur classe, permis par le retraitement indiqué plus haut.

5.5.2 Propriétés de la chaîne de Markov

Tout d'abord, on peut remarquer que la présence d'états absorbants empêche nécessairement la chaîne de Markov d'être irréductible. En effet, on rappelle que tous les états d'une chaîne de Markov irréductible communiquent entre eux, ce qui n'est pas possible dans notre cas dans la mesure où un état absorbant ne communique qu'avec lui-même.

D'après les propriétés décrites dans la sous-partie 2.3.4 du mémoire, ce résultat nous montre que bien que notre chaîne de Markov admet une probabilité stationnaire (puisque le nombre d'états est fini), mais cette probabilité n'a pas de raison d'être unique.

Mais la propriété la plus importante est celle qui indique qu'on ne peut pas en déduire que la suite $(M^n)_{n\geqslant 0}$ converge vers la matrice dont toutes les lignes sont égales à la probabilité stationnaire.

Le fait que cette propriété ne soit pas vérifiée n'est pas surprenant dans la mesure où cela signifierait que la distribution du nombre de sinistres dans chaque classe à l'ultime ne dépendrait pas de la distribution initiale, ce qui semblerait peu vraisemblable.

Nous allons donc nous intéresser à la limite de la suite $(M^n)_{n\geqslant 0}$ si elle existe. Cette limite est importante puisqu'elle nous permet de déterminer la distribution du nombre de sinistres dans chaque classe à l'ultime.
On note cette matrice M^∞.

Calcul de M^∞

Nous abordons ici à la suite de matrices $(M^n)_{n\geqslant 0}$. Cette suite est la matrice qui décrit l'état de tous les sinistres après n semestres de liquidation, c'est-à-dire que toutes les probabilités de se retrouver dans une des classes $\bullet\bullet$ E sont de moins en moins importantes. Si on s'intéresse à la fin de leur liquidation (autrement à la limite de la suite), ces probabilités sont même nulles.

Cela nous permet de construire la matrice M^∞ à partir de ses quatre blocs :

1. Le bloc supérieur gauche est composé uniquement de zéros (comme expliqué précédemment)
2. Le bloc inférieur gauche est toujours nul puisque c'était le cas dès le début.
3. Le bloc inférieur droit est bien entendu toujours égal à la matrice identité I_{18}.
4. Soit $\epsilon > 0$. On cherche le plus petit entier naturel N tel que tous les coefficients du bloc supérieur gauche de M^N soient inférieurs à ϵ. Le bloc supérieur droit de M^∞ est construit à partir du bloc supérieur droit de M^N. Un retraitement est nécessaire : chaque ligne est multipliée par le

facteur adéquat de telle sorte que la somme de chaque ligne soit toujours égale à 1.

Résultat obtenu En choisissant $\epsilon = 0{,}001$, on obtient $N = 274$. Cela signifie qu'il faudrait en théorie un maximum de 137 années pour qu'un sinistre se clôture. La matrice M^{∞} obtenue est présentée en annexe D.

5.5.3 Première projection : approche déterministe

Description de la démarche

L'application déterministe du modèle consiste à projeter le nombre le sinistre dans chaque classe à l'inventaire 2015 en fonction de la distribution limite de la chaîne de Markov. Il s'agit d'appliquer un résultat déduit des propriétés dont nous disposons sur les chaînes de Markov.

Nous allons adapter la notation $V_{i,j}$ définie dans la partie 3.4 lorque la première méthode de Mahon était détaillée. Pour rappel, $V_{i,j}$ désignait le vecteur contenant les nombres de sinistres dans chaque classe pour l'année d'origine i et l'année de développement j, qui était pertinent puisque la maturité des sinistres avait son importance dans le modèle.

Le vecteur utilisé ici n'a besoin que d'être unidimensionnel. Nous définissons $V_n \in \mathbb{R}^{37}$ comme étant le vecteur contenant le nombre de sinistres dans chaque classe à l'inventaire du semestre n, celui correspondant à l'inventaire du 31 décembre 2015 étant noté V_0.

Les deux résultats suivants sont ceux qui importent dans cette application :

- Le nombre de sinistres de sinistres dans chaque classe dans n semestres est estimé par :
$$\widehat{V}_n = V_0 \times M^n$$
- Et le nombre de sinistres de sinistres dans chaque classe à l'ultime est estimé par :
$$\widehat{V}^{\infty} = \lim_{n \mapsto +\infty} \widehat{V}_n = V_0 \times M^{\infty}$$

Les sinistres présents dans les classes absorbantes peuvent toutefois ne pas être inclus dans les vecteurs $(V_n)_{n \geqslant 0}$ dans la mesure où leur classe ne changera pas et n'a donc pas besoin de faire l'objet d'une estimation.

Une fois que la projection à l'ultime est effectuée, les classes de sinistres sont regroupées en six classes qui ne dépendent que de la charge des sinistres. En effet, nous rappelons que les autres variables ont été introduites uniquement dans le but de calibrer le modèle mais il ne s'agit pas de données que l'on souhaite évaluer.

Vecteur V_0

Les éléments du vecteur V_0, autrement dit les nombres de sinistres appartenant à chaque classe au 31 décembre 2015 sont répertoriés dans les tableaux ci-dessous :

Classe	Effectif
IBNR	186
0AE	87
1AE	2 126
2AE	2
3AE	0
4AE	2
5AE	0

Classe	Effectif
0BE	37
1BE	2 194
2BE	48
3BE	18
4BE	10
5BE	6

Classe	Effectif
0CE	2
1CE	230
2CE	40
3CE	25
4CE	21
5CE	38

Classe	Effectif
0AT	10 544
1AT	7 435
2AT	20
3AT	5
4AT	3
5AT	0

Classe	Effectif
0BT	2 889
1BT	4 036
2BT	53
3BT	14
4BT	4
5BT	2

Classe	Effectif
0CT	45
1CT	273
2CT	32
3CT	8
4CT	4
5CT	1

Rappel : Le nombre de sinistres présents dans la classe IBNR a été estimé en effectuant un Chain Ladder semestriel sur le triangle des nombres de sinistres en gestion.

Cela nous donne un total de 30 440 sinistres. Parmi ceux-là, 25 368 sont déjà terminés donc leur classe actuelle ne changera pas. Il reste donc à évaluer la classe à l'ultime de 5 072 sinistres, c'est-à-dire ceux qui sont listés dans les trois tableaux de la première ligne.

Estimation du vecteur V^{∞}

On rappelle que le vecteur V^{∞} est estimé à l'aide de la relation suivante :

$$\widehat{V}^{\infty} = V_0 \times M^{\infty}$$

La matrice M^{∞} ayant été calculée lors de la sous-partie 5.5.2, il est donc possible d'évaluer une espérance du nombre de sinistres appartenant à chaque classe à la fin de leur liquidation.

Les éléments du vecteur $\widehat{V}^{\infty}$ obtenus avec la formule ci-dessus sont listés dans les tableaux ci-dessous. Ces résultats ont été arrondis à l'unité dans la mesure où un nombre de sinistres est censé être un entier naturel.

Classe	Effectif
0AT	11 578
1AT	8 264
2AT	22
3AT	6
4AT	4
5AT	0

Classe	Effectif
0BT	3 788
1BT	5 661
2BT	82
3BT	21
4BT	6
5BT	3

Classe	Effectif
0CT	85
1CT	706
2CT	118
3CT	48
4CT	27
5CT	21

Le modèle ayant été conçu pour que les classes $\bullet\bullet$E ne contiennent aucun sinistres à l'ultime, elles n'ont pas été répertoriées ici. On retrouve bien un total de 30 440 sinistres, soit autant qu'au 31 décembre 2015, ce qui est logique

puisque parmi les 30 440 de départ, on trouvait déjà des sinistres tardifs "fictifs".

On pourrait aussi bien n'effectuer l'estimation que sur les 5 072 sinistres qui étaient en cours au 31 décembre 2015, c'est-à-dire ceux qui étaient soit dans la classe IBNR, soit dans une des classes • • E. Ces sinistres étaient représentés dans les trois tableaux de la ligne supérieure lorsque nous avions répertorié les éléments de V_0.

Les tableaux suivants contiennent les effectifs dans chaque classe de ces 5 072 sinistres à l'ultime.

Classe	Effectif
0AT	1 034
1AT	829
2AT	2
3AT	1
4AT	1
5AT	0

Classe	Effectif
0BT	899
1BT	1 625
2BT	29
3BT	7
4BT	2
5BT	1

Classe	Effectif
0CT	40
1CT	433
2CT	86
3CT	40
4CT	23
5CT	20

Le total de 5 072 sinistres est bien retrouvé. Les résultats ont cette fois encore été arrondis à l'unité. Ce sont ces sinistres qui sont vraiment ceux que le modèle permet vraiment d'évaluer.

Regroupement des classes en fonction de la charge

Les deux variables que nous avions introduites afin d'avoir une classification plus fine, à savoir le nombre de pièces et la position du sinistre, ne servaient qu'à calibrer le modèle. Elles ne font pas partie des données que nous souhaitons évaluer. Il est donc intéressant, une fois les résultats obtenus, de répartir très simplement les sinistres du vecteur $\widehat{V}^{\infty}$ dans leur état final en six classes de 0 à 5 prenant en compte uniquement la segmentation que nous avons effectué sur la charge.

Les deux tableaux ci-dessous présentent les résultats obtenus après ce regroupement, d'une part en considérant tous les 30 440 sinistres, d'autre part en ne considérant que les 5 072 qui ont réellement fait l'objet d'une estimation :

Classe	Effectif global	En cours au 31/12/2015	Ratio
0 • T	15 451	1 973	12,8 %
1 • T	14 661	2 887	19,7 %
2 • T	222	117	52,7 %
3 • T	75	48	64,2 %
4 • T	37	26	70,2 %
5 • T	24	21	87,7 %

Le fait que le ratio indiqué dans la dernière colonne soit croissant semble confirmer l'idée que plus un sinistre est grave, plus sa liquidation est lente.

5.5.4 Deuxième projection : approche stochastique

Cette sous-partie va détailler une approche alternative basée sur des simulations pour projeter les classes des 5 072 sinistres en cours. Cela permet non

seulement d'obtenir une estimation du vecteur V^{∞}, mais aussi d'estimer sa distribution grâce à laquelle nous pourrons par exemple calculer des quantiles.

Explication de l'approche

L'algorithme appliqué prend en entrée une matrice de transition, en l'occurrence la matrice M, et un vecteur de classes de sinistres qui est dans notre cas de taille 5 072 et que l'on note :

$$\mu_0 = (c_1^0, ..., c_{5072}^0)^{\mathsf{T}}$$

À chaque itération, on simule le vecteur μ_k en fonction du vecteur μ_{k-1}. Pour $j = 1, ..., 5072$, la classe c_j^k est déterminée à partir de la classe c_j^{k-1} en simulant une loi de probabilité discrète donnée par la matrice M. La ligne de la matrice donnant les probabilités de cette loi est bien entendu celle correspondant à la classe de départ c_j^{k-1}.

L'algorithme s'arrête lorsque les classes contenues dans le vecteur μ_k ne sont que des classes $\bullet\bullet$ T.

Si on note S le nombre de simulations effectuées, les vecteurs obtenus pour chaque simulation sont notés $\mu_{\infty}^{(1)}, ..., \mu_{\infty}^{(S)}$.

De chaque vecteur $\mu_{\infty}^{(s)}$, on peut déduire simplement un vecteur $V_{(s)}^{\infty}$. Comme précédemment, les sinistres sont ensuite sont regroupés selon les six classes ne dépendant que de la charge.

Résultats des simulations

Le nombre de simulations choisi est $S = 10\,000$. Ces simulations ont été effectuées à l'aide du logiciel SAS. Nous obtenons donc 10 000 possibilités de répartitions des sinistres par classe.

Le fait de n'avoir plus un seul vecteur $\widehat{V}^{\infty}$ nous permettra dans le chapitre suivant non seulement d'estimation une charge ultime correspondant à ce vecteur mais aussi une charge ultime correspondant aux 10 000 simulations qui ont été effectuées et donc d'en déduire une distribution empirique.

Remarque Il peut être judicieux de suivre l'évolution de la répartition de sinistres dans chaque classe en fonction du nombre de simulations considérées. Pour cela, nous proposons d'observer, pour $k = 0, ..., 5$, le ratio suivant :

$$p(k, s) = \%\text{ de sinistres dans les classes } k\bullet\bullet \text{ pour les simulations 1 à } s$$

Nous nous attendons à ce que ce ratio converge à partir d'un certain stade, ce qui semble être le cas au vu des graphiques des $p(k, s)$ présentés en annexe E. En effet, selon les classes, la convergence semble se produire entre la 1000-ième et la 4000-ième simulation, ce qui traduit une volatilité faible des simulations.

Chapitre 6

Évaluation de la charge ultime

Le chapitre précédent nous a permis de proposer une estimation de la classe dans laquelle sera chaque sinistre à l'ultime. Mais cela ne nous suffit pas, il faut maintenant évaluer la charge ultime correspondante et par extension le montant de la provision à constituer. La charge évaluée sera nette de réassurance, la raison nous permettant ce choix est expliquée plus loin.

6.1 Approche déterministe basée sur les coûts moyens

Cette approche a déjà rapidement expliquée dans la sous-partie 3.4.4 lorsque nous avions fait notre première projection de classes avec la méthode de Mahon. Pour rappel, Mahon ne proposait pas de moyen d'évaluer la charge ultime dans son article.

6.1.1 Description de la méthode

Il s'agit d'une méthode très simple qui consiste à calculer le coût moyen des sinistres présents dans les classes :

- 0 • •
- 1 • •
- 2 • •
- 3 • •
- 4 • •
- 5 • •

En effet, on rappelle que les variables ayant été ajoutées dans le modèle, c'est-à-dire le nombre de pièces reçues par la MACSF et la position des sinistres, n'ont plus d'importance une fois que les classes des sinistres ont été projetées à l'ultime.

Ces coûts moyens sont ensuite multipliés par les effectifs de sinistres appartenant à ces classes qui ont été évalués dans le chapitre précédent pour obtenir

une estimation de la charge ultime.

Deux de ces classes font l'objet d'un traitement légèrement différent que nous explicitons ci-dessous.

Classe $0\bullet\bullet$ Il s'agit de la classe des sinistres sans suite, ce qui signifie que le coût moyen n'a pas besoin d'être calculé puisqu'il est égal à 0 €.

Classe $5\bullet\bullet$ Il s'agit quant à elle de la classe contenant les sinistres dont la charge est supérieure ou égale à 1 000 000 €. Ce seuil correspond au montant à partir duquel les sinistres sont passés en réassurance par la MACSF. La MACSF règlera donc forcément 1 000 000 € pour chaque sinistres de la classe $5\bullet\bullet$, il n'y a donc cette fois-ci encore pas besoin de calculer le coût moyen. En plus de fournir des résultats plus intéressants puisque plus proches du montant que la mutuelle sera effectivement amenée à payer, le fait de s'intéresser à la charge nette de réassurance permet également de rendre l'étude beaucoup plus simple.

L'évaluation de la charge ne s'effectue bien entendu que sur les sinistres qui étaient encore en cours au 31 décembre 2015. On rappelle qu'ils sont au nombre de 5 072, dont 186 d'entre eux sont des sinistres fictifs représentant une estimation du nombre de sinistres tardifs.

La charge ultime se calcule de la manière suivante :

$$\hat{C} = \sum_{k=0}^{5} \text{Coût moyen } (k\bullet\bullet) \times \text{Effectif } (k\bullet\bullet)$$

6.1.2 Résulats obtenus

À partir de la projection déterministe

Les coûts moyens considérés pour chacune des classes ainsi que les nombres de sinistres estimés par la première approche déterministe dans la sous-partie 5.5.3 sont présentés dans le tableau ci-dessous :

Classe	Coût moyen	Effectif
$0\bullet\bullet$	0 €	1 973
$1\bullet\bullet$	6 378 €	2 887
$2\bullet\bullet$	138 268 €	117
$3\bullet\bullet$	266 661 €	48
$4\bullet\bullet$	619 794 €	26
$5\bullet\bullet$	1 000 000 €	21

Ces coûts moyens sont calculés sur tous les sinistres, ouverts et clos, quelque que soit l'exercice de survenance.

Cela nous donne une charge ultime estimée à 84 505 k€.

Dans les méthodes précédentes, la provision était obtenue en effectuant la différence entre la charge ultime et le total des règlements au 31 décembre 2015.

Ici, la charge ultime évaluée n'est pas la charge de la totalité de notre flotte de sinistres mais seulement ceux qui étaient encore en cours à fin 2015. Ce sont bien uniquement les paiements de ces sinistres qu'il faut soustraire, soit 55 138 k€.

La provision nette de réassurance estimée est donc égale à **29 367 k€**. Ce montant n'est pas comparable aux précédentes évaluations dans la mesure où celles-ci étaient brutes de réassurance, et donc nécessairement beaucoup plus élevées.

À partir de la projection stochastique

Nous allons maintenant appliquer cette méthode à l'approche stochastique de projection des sinistres que nous avions décrite dans la sous-partie 5.5.4.

Pour rappel, nous avions effectué 10 000 simulations, ce qui signifie que nous pouvons appliqué les coûts moyens présentés plus haut aux 10 000 vecteurs $V_1^\infty, ..., V_{10000}^\infty$, et ainsi obtenir 10 000 estimations différentes pour la charge ultime que l'on note $\hat{C}_1, ...\hat{C}_{10000}$.

On s'intéresse à quelques résultats statistiques tel que la moyenne où certains quantiles de la distribution de la charge ultime estimée de cette manière. Ces résultats sont présentés ci-dessous :

- $\bar{C} = 84\,897\,942$ €
- $\text{VaR}_{90\%}(C) = 90\,793\,429$ €
- $\text{VaR}_{95\%}(C) = 92\,374\,371$ €
- $\text{VaR}_{99\%}(C) = 95\,919\,077$ €
- $\text{VaR}_{99,5\%}(C) = 97\,204\,029$ €

Le quantile d'ordre α est déterminé comme étant la $\big(10\,000 \times (1-\alpha)\big)$-ième plus grande valeur des $(\hat{C}_s)_{s=1,...,10000}$.

La moyenne des estimations effectuées à partir des 10 000 simulations est légèrement supérieure à l'estimation que nous avions faite à partir de la projection déterministe. Un nombre plus élevé de simulations aurait permis d'obtenir un résultat plus proche, cependant l'écart est très minime (0,5%).

Nous pouvons de plus constater que les quantiles obtenus sont relativement peu élevés : par exemple le quantile n'est supérieur à la moyenne que de 14%. Cela s'explique par le fait que la variance des simulations soit très faible, ce que nous mettions en évidence dans la remarque située à la fin de la section 5.5.4.

Pour plus de prudence, nous pouvons par exemple choisir de conserver le quantile à 95%, soit une estimation de la charge ultime égale à 92 374 k€ et donc une provision de **32 237 k€**.

6.2 Approche stochastique

L'approche décrite dans la partie précédente a un défaut important. Celui de ne pouvoir proposer comme charge ultime qu'une combinaison linéaire

du vecteur des coûts moyens. Cela signifie que la charge ultime peut toujours s'exprimer de la manière suivante :

$$\hat{C} = 0 \times n_0 + 6\,378 \times n_1 + 138\,268 \times n_2 + 266\,661 \times n_3 + 619\,794 \times n_4 + 1\,000\,000 \times n_5$$

où $\sum_{i=0}^{5} n_i = 5\,072$

Il peut être judicieux d'avoir recours à des simulations afin de se libérer de cette condition contraignante, d'autant plus que cette alternative permettra également de déterminer la volatilité des estimations.

Remarque : Dans l'approche précédente, seule la charge nette de réassurance était évaluée car la volatilité de la dernière classe était trop élevée pour estimer un coût moyen. Ici, nous nous intéresserons à la fois à la charge brute et nette de réassurance, ce qui nous permettra également d'estimer le coût du traité de réassurance.

6.2.1 Présentation de l'approche

Cette méthode d'évaluation de la charge ultime se base sur les résultats des projections des sinistres effectuées dans la section 5.5.4. Pour rappel, nous avions obtenus 10 000 scénarios modélisant les classes de tous les sinistres à l'ultime. Comme nous le verrons, cette approche permet d'estimer la provision brute et nette de réassurance, ainsi que la volatilité du modèle.

Afin d'évaluer la charge ultime de ces sinistres, nous procédons de la manière suivante :

1. Tout d'abord, nous simulons 500 000 sinistres d'une distribution donnée dont le choix sera abordé plus loin. À chaque sinistre simulé est associée une classe de 0 à 5 correspond aux seuils définis par la classification de notre modèle.
2. Les sinistres dont nous devons évaluer la charge ultime sont tous les sinistres des 10 000 scénarios de répartitions des sinistres par classe que nous avions simulées à la section 5.5.4. La charge ultime s'obtient en tirant aléatoirement un des 500 000 sinistres que nous venons de simuler en se limitant à ceux dont la classe est identique à celle du sinistre dont nous voulons évaluer la charge.
3. À partir de nos 10 000 scénarios, nous obtenons donc 10 000 estimations différentes de la charge ultime.

6.2.2 Choix de la loi simulée

Nous cherchons maintenant à déterminer quelle loi de probabilité reproduit le mieux notre échantillon. La description des données effectuée dans la section 3.1.2 du mémoire nous permet de sélectionner quelques distributions possibles. En effet, la loi voulue doit satisfaire au moins les quelques critères suivants :

- elle doit être définie au minimum sur $\mathbb{R}_+^*$ (puisque les sinistres sans suite sont traités à part) et ne doit pas être bornée
- elle doit présenter une queue de distribution

– son coefficient d'asymétrie doit être positif (pour rappel, le skweness empirique de la distribution est de 37,0)

Les lois de probabilité possibles envisagées pour l'instant sont donc les suivantes :
– la loi log-normale
– la loi Gamma
– la loi de Weibull
– loi de Pareto

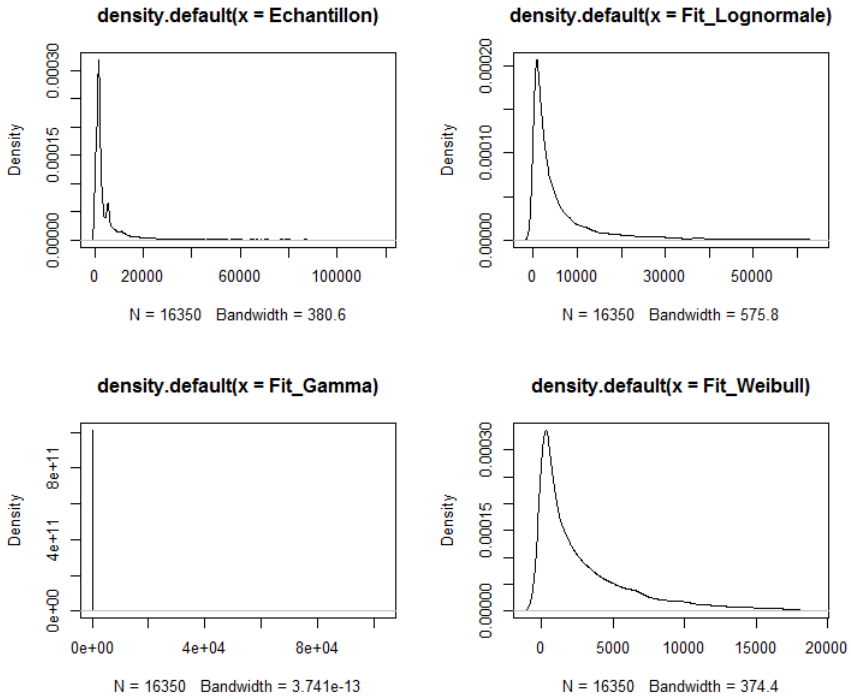
Nous procédons donc à l'adéquation des paramètres de ces trois lois avec notre échantillon. Cette adéquation a été réalisé avec la méthode du maximum de vraisemblance pour les lois log-normale, Weibull et Pareto. L'adéquation avec la loi Gamma a quant à elle été effectué via la méthode des moments pour la loi Gamma dans la mesure où les estimateurs du maximum de vraisemblance de cette distribution n'ont pas de forme paramétrique(cf MINKA [2002]).

Les densités de probabilité qui seront comparées à notre densité empirique sont donc respectivement les suivantes :

- $f_{\text{LN}}(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$ où $\mu = 7{,}929$ et $\sigma = 1{,}488$
- $f_{\text{Gamma}}(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} \exp(-\lambda x)$ où $k = 7{,}729 \times 10^{-3}$ et $\lambda = 4{,}097 \times 10^{-7}$
- $f_{\text{Weibull}}(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \exp\left(\left(-\frac{x}{\lambda}\right)^k\right)$ où $k = 0{,}545$ et $\lambda = 6\,075{,}87$

Pour la loi de Pareto, dont la densité est $f_{\text{Pareto}}(x) = \alpha \frac{\beta^\alpha}{x^{\alpha+1}}$ pour $x \geqslant \beta$, nous estimons $\alpha = 0{,}47$ et $\beta = 5{,}79$. Nous choisissons donc d'écarter cette distribution dans la mesure où nous aurions voulu que α soit supérieur à 1 (sinon, la distribution n'admet pas de moments finis).

Les graphiques correspondant aux densités retenues ont été tracés avec la fonction *density* du logiciel R (les 300 plus grandes valeurs ont été retirées afin de rendre les graphes plus lisibles). Ces densités, ainsi que celle de notre échantillon (qui était déjà présentée à la section 3.1.2 du mémoire), sont représentées ci-dessous :



L'observation graphique nous permet de constater que la loi Gamma ne semble pas du tout correspondre à la distribution empirique. La loi log-normale et la loi de Weibull peuvent quant à elles être envisageables. Afin de déterminer quelle loi sera effectivement retenue, nous avons recours à la distance entre la distribution empirique et chaque distribution estimée correspondant à la statistique du test de Kolmogorov-Smirnov.

Soient $x_1, ..., x_N > 0$ nos observations empiriques, $\hat{F}$ la fonction de répartition empirique et F_0 la fonction de répartition d'une distribution donnée. Pour rappel, le test de Kolmogorov-Smirnov est un test d'adéquation qui permet de vérifier, pour un échantillon (de fonction de répartition empirique $\hat{F}$) et une loi continue F_0 donnée, l'hypothèse suivante :

$$H_0 : \hat{F} = F_0$$

L'hypothèse alternative est donc :

$$H_1 : \hat{F} \neq F_0$$

La statistique du test de Kolmogorov-Smirnov est définie de la manière suivante :

$$D = \max_{1=1,...,N} |F(x_i) - \hat{F}(x_i)|$$

Plus cette quantité est faible, plus la distribution associée à la fonction F est une estimation satisfaisante de la distribution empirique.

Le test de Kolmogorov-Smirnov associe à cette statistique une p-value correspond au degré de confiance de l'adéquation des deux lois. Ainsi, l'hypothèse d'adéquation est validée au seuil de confiance α si la p-value estimée est supérieure ou égale à α.

Application Les valeurs de la statistique du test de Kolmogorov-Smirnov obtenues pour les trois lois testées sont les suivantes :

Distribution	D
Log-normale	0,122
Gamma	0,962
Weibull	0,256

Ces valeurs sont très élevées et le logiciel R estime des p-value inférieures à 10^{-15}, ce qui est insatisfaisant dans la mesure où l'hypothèse H_0 est rejetée pour les trois distributions.

Le fait que la distance de Kolmogorov-Smirnov soit si élevée peut s'expliquer par le fait que la distribution des excès de notre échantillon soit à queue épaisse (comme l'indiquaient les résultats présentés dans la section 5.2.4 du mémoire). Cela peut donc générer des écarts importants sur la distribution des extrêmes Nous choisissons tout de même de retenir la loi log-normale pour effectuer nos simulations.

6.2.3 Simulations et résultats obtenus

Nous simulons donc 500 000 sinistres suivant une loi log-normale de paramètres estimés par la méthode du maximum de vraisemblance :

- $\mu = 7{,}929$
- $\sigma = 1{,}488$

Comme nous l'indiquions, nous attribuons à chacun des 500 000 sinistres la tranche de coût (1, 2, 3, 4 ou 5) associée telle que définie par notre classification. Puis ce stock de sinistres sert à sélectionner aléatoirement (uniformément et avec remise) un sinistre correspondant à une classe donnée, et ce pour chaque sinistres de nos 10 000 scénarios (à l'exception des sinistres sans suite dont la charge est directement connue).

En ce qui concerne les sinistres de la classe 5 • •, deux alternatives s'offrent à nous :

- Il est dans un premier temps possible de procéder comme pour les autres classes, ce qui permet d'obtenir la charge brute de réassurance
- Mais nous pouvons également fixer tous ces sinistres à 1 000 k€, ce qui correspond à leur charge nette de réassurance

Ces deux alternatives permettent donc également d'estimer le montant qui pris en charge par la réassurance. Cependant, comme indiqué plus haut, la distribution des extrêmes de nos simulations est à considérer avec prudence.

Nous obtenons donc une estimations de la charge par scénario (correspondant aux simulations de la section 5.5.4), soit 10 000 estimations, à partir desquels nous pouvons déduire une distribution de la provision.

Remarque La charge ultime déterminée est une fois encore en euros constants et la méthode décrite ne donne pas d'informations sur l'écoulement des règlements futurs afin de convertir les montants en euros courants. Nous faisons donc l'hypothèse que la cadence de liquidation des montants provisionnés est identique à

celle obtenue en liquidant les règlements avec la méthode Chain Ladder. Cette hypothèse n'est toutefois pas entièrement satisfaisante dans la mesure où le modèle était construit dans le but d'étendre la projection des sinistres à un horizon plus lointain que celui imposé par Chain Ladder, ce qui crée un biais. Les taux d'inflation appliqués sont ceux indiqués dans la section 3.4.4.

Résultats obtenus

Cette approche nous permet donc d'obtenir une espérance et un écart-type de la provision brute et nette de réassurance ainsi que du montant pris en charge par la réassurance que nous indiquons dans le tableau ci-dessous :

	PSAP brute de de réassurance	PSAP nette de de réassurance	Montant réglé par la réassurance
Espérance	105 147 k€	57 250 k€	47 897 k€
Écart-type	30 265 k€	4 761 k€	28 689 k€

Les principaux quantiles (sous hypothèse de log-normalité) sont également présentés ci-dessous :

Probabilité	Quantile - *PSAP brute de réassurance* (k€)	Quantile - *PSAP nette de réassurance* (k€)
75%	122 224	60 340
80%	128 125	61 183
85%	135 364	62 180
90%	145 056	63 459
95%	160 713	65 402
99%	194 783	69 209
99,5%	208 986	70 657

Conclusion

Ce mémoire a donc rempli son objectif qui était de proposer un modèle permettant d'évaluer une charge de sinistres qui tienne compte des fortes disparités de coût que l'on constate régulièrement en RC corporelle.

La première méthode markovienne qui a été abordée dans notre étude, c'est-à-dire la méthode de Mahon, a fourni un résultat satisfaisant dans une certaine mesure. En effet, la provision obtenue a été inférieure de 2% par rapport à la provision évaluée par le modèle Chain Ladder. Cependant et comme nous l'avons vu, cette méthode n'est pas exempte de reproches, c'est pourquoi des possibilités d'amélioration ont été tentées.

La première d'entre elles a probablement été trop ambitieuse. Les conditions à remplir pour que la modélisation que nous espérions puisse être effectuée semblent irréalisables. L'idée de prendre comme point de départ la méthode de Mahon en construisant les matrices de transitions à l'aide d'un modèle mathématique au lieu d'être contraint de faire uniquement confiance aux données empiriques parfois trop volatile paraissait pourtant pertinente. Il parait difficile de trouver des pistes d'amélioration permettant de poursuivre cette tentative.

Le modèle suivant a été celui ayant finalement été retenu. Comme nous l'avons vu, le fait d'avoir des probabilités de transitions constantes dans le temps permet de libérer de certaines contraintes, notamment de durée de liquidation. La nouvelle classification des sinistres a été effectuée en se fiant à des notions mathématiques a priori plus solides que le rapide raisonnement à partir duquel les classes avaient été définies dans la première méthode.

De plus, bien que ce modèle soit bâti, comme la méthode de Mahon, exclusivement sur des données empiriques, il ne possède pas le défaut que nous essayions de corriger avec la méthode précédente. En effet, le fait de ne pas séparer les transitions de classes en fonction de la maturité des sinistres implique qu'il y a toujours beaucoup plus de données qui sont considérées lors du calcul des probabilités de passage, quel que soit le niveau de gravité considéré.

L'approche simulatoire permettant d'évaluer la charge ultime semble a priori également plus convenable que l'approche plus simple consistant à ne retenir que les coûts moyens afin d'éviter de se limiter aux combinaisons linéaires de plusieurs valeurs. Néanmoins la difficulté rencontrée pour déterminer une distribution convenant à notre distribution est à prendre en compte et peut avoir

impliqué des imprécisions dans les estimations, en particulier sur les extrêmes. En revanche, le fait de pouvoir évaluer la provision nette de réassurance est un avantage non négligeable dans la mesure où cela permet de combler une limite de Chain Ladder (qui s'adapte très bien à la réassurance proportionnelle mais ne convient pas aux traités en excédent de sinistres, comme c'est le cas pour la MACSF). Cependant, on peut regretter le fait que les autres méthodes testées ne soient pas adaptées à la provision nette de réassurance dans la mesure où cela nous empêche de comparer ce résultat avec d'autres méthodes.

Bibliographie

- MAHON J.B. [2005] - *"Transition Matrix Theory and Individual Claim Loss Development"* - Casualty Actuarial Society Forum
- PAYRE B., SAVARRE S. [2012] - *"Charge ultime nette de réassurance en RC corporelle : 2 modèles stochastiques pour les flottes automobiles" (Mémoire d'actuariat)*
- GIRODOT T. [2012] - *"Méthodes alternatives pour le calcul du Best Estimate des provisions des sinistres graves en assurance automobile" (Mémoire d'actuariat)*
- CORNISH E.A., FISHER R.A. [1937] - *"Moments and cumulants in the specification of distributions"*
- CHERNOZHUKOV V., FERNANDEZ-VAL I., GALICHON A. [2007] - *"Rearranging Edgewoth-Cornish-Fisher expansions"*
- BENLAGHA N., GRUN-RÉHOMME M., VASECHKO O. [2009] - *Les sinistres graves en assurance automobile : Une nouvelle approche par la théorie des valeurs extrêmes* - Revue MODULAD
- JAZIRI S. [2011] - *"Méthode de provisionnement non-vie et risque de réserve à un an" (Mémoire d'actuariat)*
- BOUTHIER D., HIPPOLYTE H. [2012] - *"Mesure du risque de provisionnement à un an en assurance IARD" (Mémoire d'actuariat)*
- PARTRAT C. [2007] - *"Provisionnement technique en assurance non-vie"*
- MINKA T. [2002] - *"Estimating a Gamma distribution"*

Annexe A

Applications des méthodes de provisionnement classiques

A.1 Triangles

Cette annexe présente le triangle de règlements cumulés brut de recours issu de nos données que l'on utilise pour appliquer la méthode Chain Ladder dans le chapitre 1. Le triangles des charges est également présenté.

A.1.1 Triangle des règlements

	0	1	2	3	4	5
2004	975k€	5 410k€	7 904k€	9 006k€	10 725k€	11 152k€
2005	1 035k€	5 940k€	9 626k€	12 039k€	15 832k€	16 517k€
2006	1 140k€	5 351k€	8 988k€	11 319k€	13 609k€	14 461k€
2007	1 439k€	5 716k€	9 550k€	11 668k€	13 273k€	13 652k€
2008	1 282k€	6 006k€	9 289k€	11 294k€	13 887k€	14 894k€
2009	1 324k€	5 988k€	8 909k€	10 200k€	12 636k€	13 920k€
2010	934k€	5 574k€	9 265k€	10 636k€	12 699k€	15 795k€
2011	1 313k€	5 464k€	8 550k€	10 173k€	11 541k€	
2012	1 152k€	5 418k€	9 098k€	10 941k€		
2013	1 324k€	5 986k€	10 459k€			
2014	1 385k€	7 117k€				
2015	2 435k€					

	6	7	8	9	10	11
2004	12 555k€	12 695k€	13 041k€	13 214k€	13 451k€	13 537k€
2005	16 876k€	17 251k€	18 030k€	18 217k€	18 662k€	
2006	15 192k€	17 461k€	17 966k€	18 338k€		
2007	14 488k€	17 217k€	17 539k€			
2008	16 800k€	20 141k€				
2009	14 630k€					
2010						
2011						
2012						
2013						
2014						
2015						

A.1.2 Triangle des charges

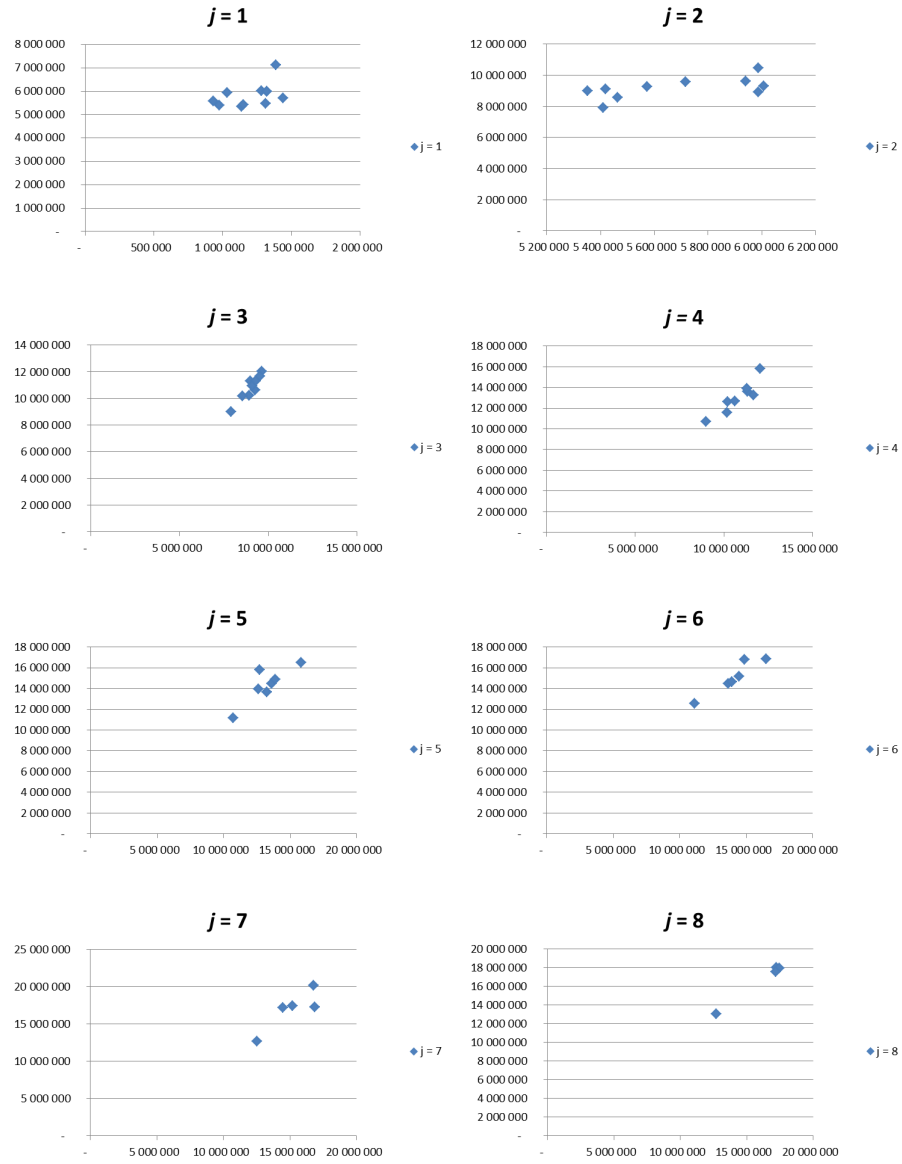
	0	1	2	3	4	5
2004	14 871k€	16 804k€	15 814k€	15 419k€	16 166k€	16 217k€
2005	16 427k€	22 944k€	22 277k€	23 140k€	22 989k€	23 801k€
2006	21 784k€	24 240k€	27 129k€	27 597k€	28 396k€	27 956k€
2007	20 168k€	27 131k€	26 230k€	25 355k€	35 783k€	35 737k€
2008	20 434k€	27 383k€	30 056k€	30 171k€	36 002k€	35 095k€
2009	18 053k€	21 778k€	21 203k€	21 736k€	22 084k€	21 201k€
2010	13 560k€	24 611k€	31 944k€	34 652k€	37 579k€	33 264k€
2011	14 988k€	20 149k€	19 897k€	19 909k€	22 074k€	
2012	11 671k€	18 064k€	22 143k€	20 415k€		
2013	19 363k€	32 360k€	35 389k€			
2014	15 433k€	29 594k€				
2015	21 194k€					

	6	7	8	9	10	11
2004	16 177k€	16 319k€	16 705k€	16 466k€	15 823k€	15 783k€
2005	23 836k€	23 796k€	19 543k€	19 879k€	20 047k€	
2006	28 008k€	20 427k€	20 249k€	20 182k€		
2007	35 402k€	34 028k€	35 486k€			
2008	35 180k€	38 680k€				
2009	19 930k€					
2010						
2011						
2012						
2013						
2014						
2015						

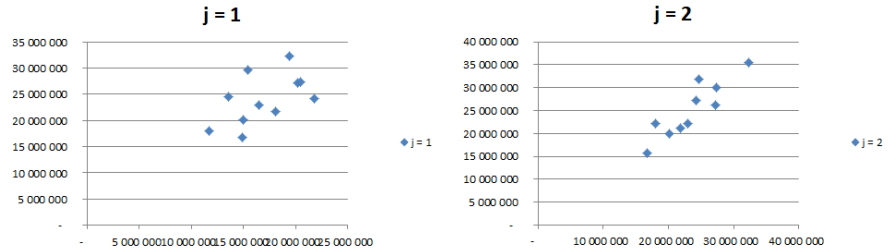
A.2 Validation des hypothèses

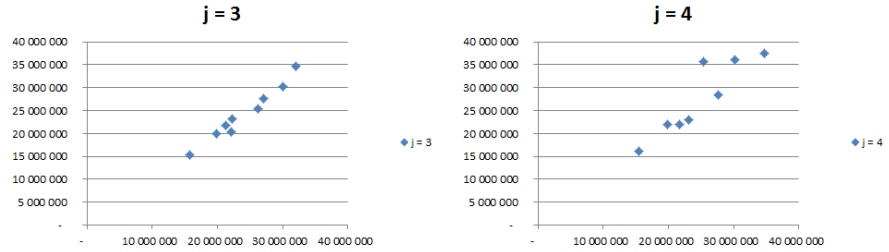
A.2.1 Hypothèse H2

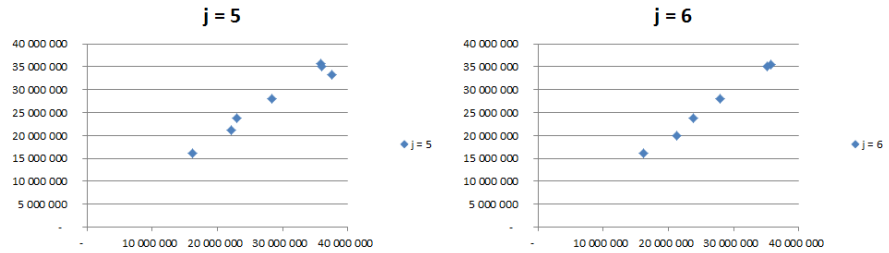
Pour le triangle des règlements

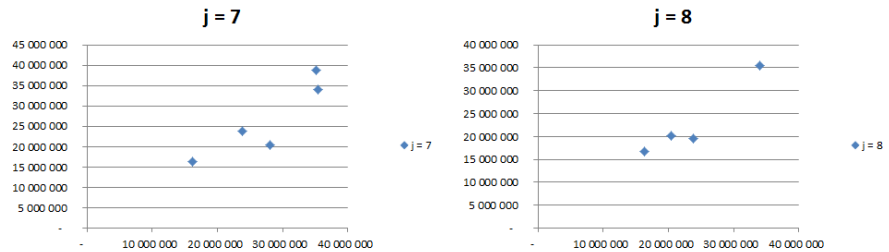


Pour le triangle des charges



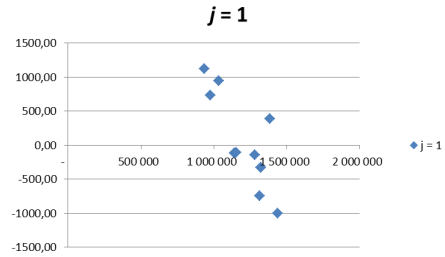


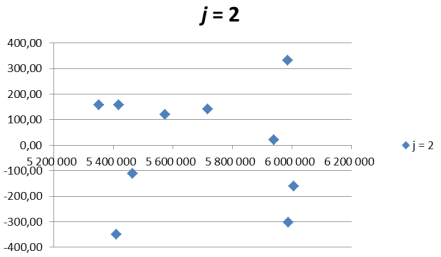


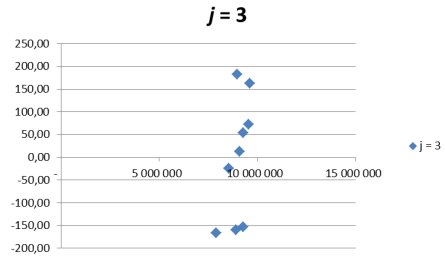


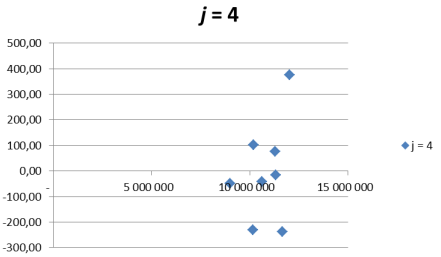
A.2.2 Hypothèse H3

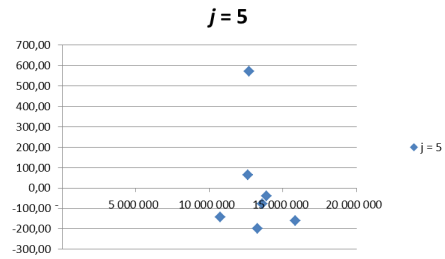
Pour le triangle des règlements

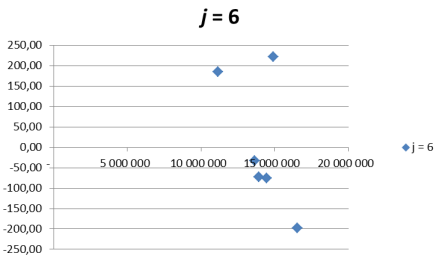


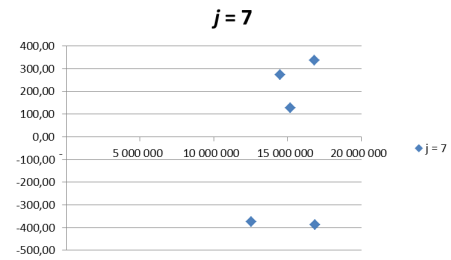


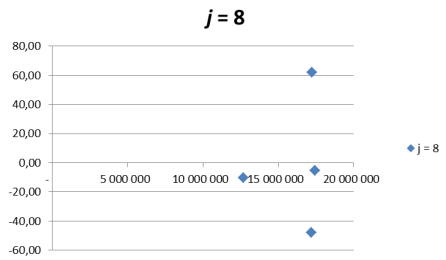




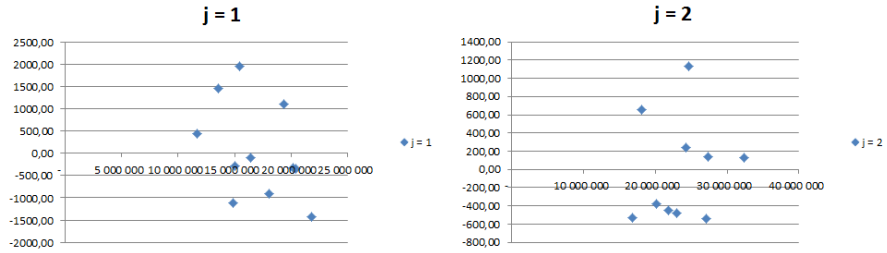


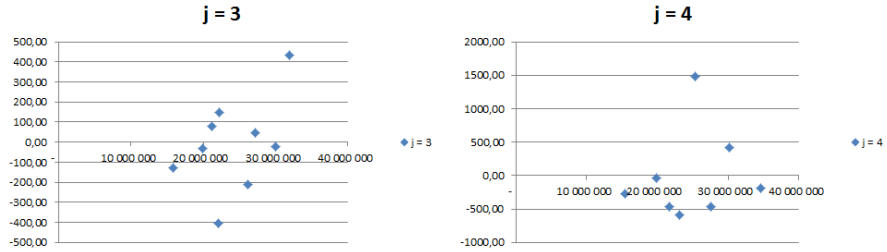


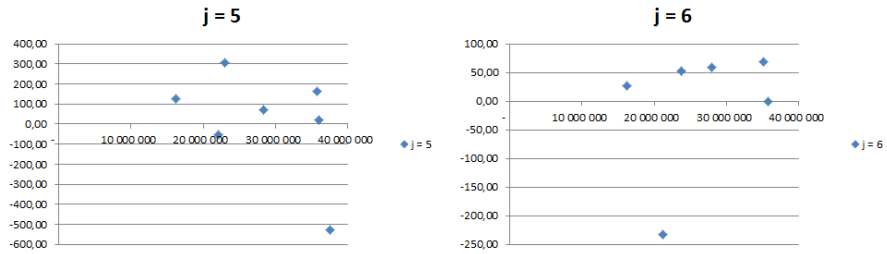


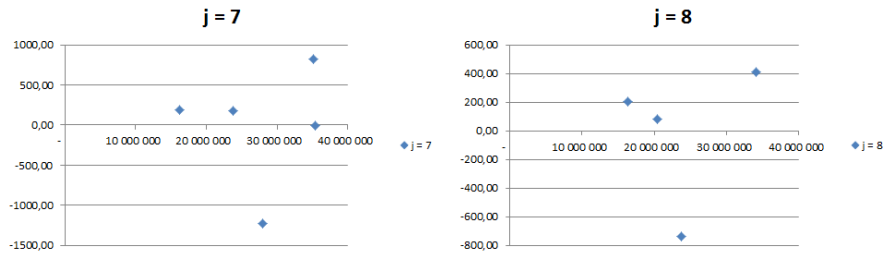


Pour le triangle des charges









Annexe B

Résultats de la première proposition d'amélioration

B.1 Résultats de l'algorithme de Levenberg-Marquardt

Ici, nous présentons les différents paramètres de lissage estimés par l'algorithme de Levenberg-Marquardt que nous abordons dans la section 4.2.2. L'algorithme a été appliqué aux fonction (f_k) et (g_k), pour $k \leqslant 7$ et plusieurs familles de fonctionnelles différentes ont été testées.

B.1.1 Résultats obtenus pour les fonctions (f_k)

k	Exponentielle	Weibull	Puissance inverse	Puissance	Gamma
0	$\widehat{a}_0 = 0{,}21$ $\widehat{b}_0 = 3{,}96$	$\widehat{a}_0 = 1{,}76$ $\widehat{b}_0 = 1{,}75$	$\widehat{a}_0 = 0{,}21$ $\widehat{b}_0 = 4{,}68$	$\widehat{a}_0 = 1{,}21$ $\widehat{b}_0 = 0{,}02$	$\widehat{a}_0 = 0{,}56$ $\widehat{b}_0 = -4{,}68$ $\widehat{c}_0 = 0{,}00$
1	$\widehat{a}_1 = -0{,}15$ $\widehat{b}_1 = 0{,}66$	$\widehat{a}_1 = 27{,}51$ $\widehat{b}_1 = 0{,}75$	$\widehat{a}_1 = -0{,}16$ $\widehat{b}_1 = 1{,}33$	$\widehat{a}_1 = 0{,}84$ $\widehat{b}_1 = 0{,}50$	$\widehat{a}_1 = -0{,}31$ $\widehat{b}_1 = 3{,}50$ $\widehat{c}_1 = 1{,}75$
2	$\widehat{a}_2 = -0{,}02$ $\widehat{b}_2 = 0{,}08$	$\widehat{a}_2 = 2{,}43$ $\widehat{b}_2 = 1{,}56$	$\widehat{a}_2 = -0{,}01$ $\widehat{b}_2 = -0{,}02$	$\widehat{a}_2 = 0$ $\widehat{b}_2 = 0$	$\widehat{a}_2 = -0{,}00$ $\widehat{b}_2 = 8{,}10$ $\widehat{c}_2 = 1{,}52$
3	$\widehat{a}_3 = -0{,}40$ $\widehat{b}_3 = 0{,}49$	$\widehat{a}_3 = 36{,}4 \times 10^9$ $\widehat{b}_3 = -9{,}50$	$\widehat{a}_3 = -0{,}45$ $\widehat{b}_3 = 1{,}25$	$\widehat{a}_3 = 0{,}60$ $\widehat{b}_3 = 0{,}58$	$\widehat{a}_3 = -0{,}97$ $\widehat{b}_3 = 1{,}02$ $\widehat{c}_3 = 0{,}96$
4	$\widehat{a}_4 = -0{,}26$ $\widehat{b}_4 = 0{,}38$	$\widehat{a}_4 = 72{,}12$ $\widehat{b}_4 = -0{,}94$	$\widehat{a}_4 = -0{,}29$ $\widehat{b}_4 = 1{,}02$	$\widehat{a}_4 = 0{,}74$ $\widehat{b}_4 = 0{,}66$	$\widehat{a}_4 = -0{,}64$ $\widehat{b}_4 = 0{,}92$ $\widehat{c}_4 = 0{,}86$
5	$\widehat{a}_5 = -0{,}05$ $\widehat{b}_5 = 0{,}03$	$\widehat{a}_5 = 0{,}93$ $\widehat{b}_5 = 8{,}68$	$\widehat{a}_5 = 0{,}65$ $\widehat{b}_5 = 42{,}93$	$\widehat{a}_5 = 0{,}95$ $\widehat{b}_5 = 0{,}97$	$\widehat{a}_5 = 0$ $\widehat{b}_5 = 0$ $\widehat{c}_5 = 0$
6	$\widehat{a}_6 = 0{,}33$ $\widehat{b}_6 = 2{,}84$	$\widehat{a}_6 = 1{,}39$ $\widehat{b}_6 = 1{,}54$	$\widehat{a}_6 = 0{,}33$ $\widehat{b}_6 = 4{,}55$	$\widehat{a}_6 = 1{,}33$ $\widehat{b}_6 = 0{,}07$	$\widehat{a}_6 = 0$ $\widehat{b}_6 = 0$ $\widehat{c}_6 = 0$
7	$\widehat{a}_7 = -0{,}12$ $\widehat{b}_7 = 0{,}10$	$\widehat{a}_7 = 20{,}85$ $\widehat{b}_7 = 2{,}86$	$\widehat{a}_7 = -0{,}13$ $\widehat{b}_7 = 0{,}38$	$\widehat{a}_7 = 0{,}88$ $\widehat{b}_7 = 0{,}90$	$\widehat{a}_7 = -0{,}09$ $\widehat{b}_7 = 2{,}47$ $\widehat{c}_7 = 0{,}83$

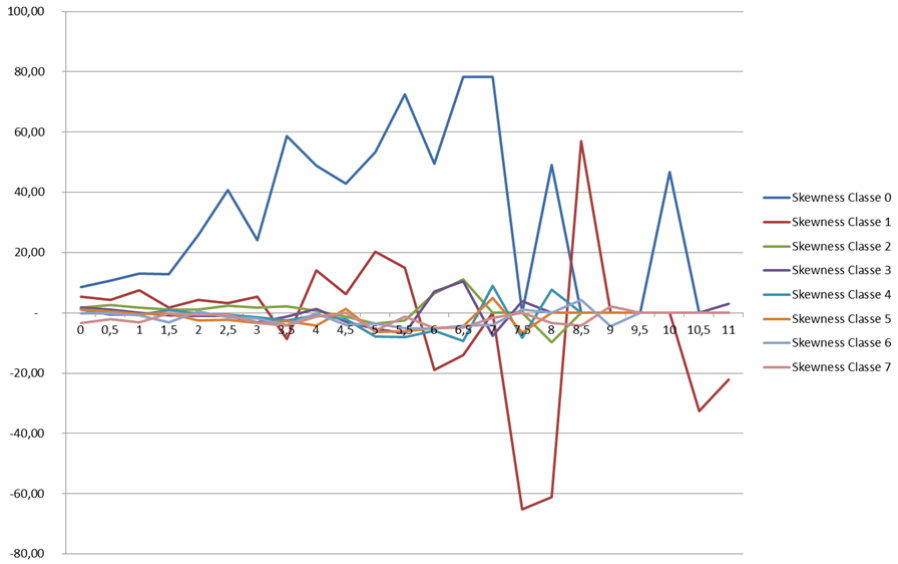
B.1.2 Résultats obtenus pour les fonctions (g_k)

k	Exponentielle	Weibull	Puissance inverse	Puissance
0	$\widehat{a}_0 = 0{,}63$ $\widehat{b}_0 = 1{,}58$	$\widehat{a}_0 = 1{,}00$ $\widehat{b}_0 = 0{,}85$	$\widehat{a}_0 = 0{,}62$ $\widehat{b}_0 = 1{,}84$	$\widehat{a}_0 = 1{,}63$ $\widehat{b}_0 = 0{,}27$
1	$\widehat{a}_1 = 0{,}67$ $\widehat{b}_1 = 0{,}53$	$\widehat{a}_1 = 0{,}87$ $\widehat{b}_1 = 0{,}64$	$\widehat{a}_1 = 0{,}73$ $\widehat{b}_1 = 1{,}22$	$\widehat{a}_1 = 1{,}68$ $\widehat{b}_1 = 0{,}63$
2	$\widehat{a}_2 = 1{,}23$ $\widehat{b}_2 = 0{,}37$	$\widehat{a}_2 = 0{,}54$ $\widehat{b}_2 = 0{,}64$	$\widehat{a}_2 = 1{,}41$ $\widehat{b}_2 = 1{,}01$	$\widehat{a}_2 = 2{,}27$ $\widehat{b}_2 = 0{,}74$
3	$\widehat{a}_3 = 1{,}68$ $\widehat{b}_3 = 0{,}34$	$\widehat{a}_3 = 0{,}42$ $\widehat{b}_3 = 0{,}66$	$\widehat{a}_3 = 1{,}96$ $\widehat{b}_3 = 0{,}96$	$\widehat{a}_3 = 2{,}74$ $\widehat{b}_3 = 0{,}77$
4	$\widehat{a}_4 = 1{,}72$ $\widehat{b}_4 = 0{,}32$	$\widehat{a}_4 = 0{,}41$ $\widehat{b}_4 = 0{,}67$	$\widehat{a}_4 = 2{,}01$ $\widehat{b}_4 = 0{,}96$	$\widehat{a}_4 = 2{,}77$ $\widehat{b}_4 = 0{,}78$
5	$\widehat{a}_5 = 2{,}06$ $\widehat{b}_5 = 0{,}34$	$\widehat{a}_5 = 0{,}29$ $\widehat{b}_5 = 0{,}76$	$\widehat{a}_5 = 3{,}11$ $\widehat{b}_5 = 1{,}02$	$\widehat{a}_5 = 3{,}71$ $\widehat{b}_5 = 0{,}78$
6	$\widehat{a}_6 = 1{,}62$ $\widehat{b}_6 = 0{,}24$	$\widehat{a}_6 = 0{,}41$ $\widehat{b}_6 = 0{,}57$	$\widehat{a}_6 = 2{,}04$ $\widehat{b}_6 = 0{,}82$	$\widehat{a}_6 = 2{,}70$ $\widehat{b}_6 = 0{,}83$
7	$\widehat{a}_7 = 1{,}92$ $\widehat{b}_7 = 0{,}18$	$\widehat{a}_7 = 0{,}47$ $\widehat{b}_7 = 0{,}44$	$\widehat{a}_7 = 1{,}67$ $\widehat{b}_7 = 0{,}63$	$\widehat{a}_7 = 2{,}46$ $\widehat{b}_7 = 0{,}87$

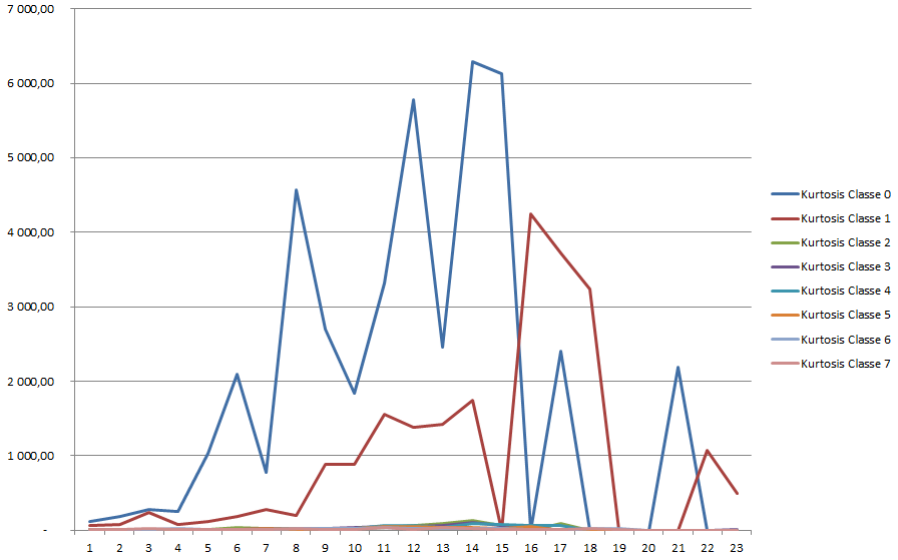
B.2 Graphes du skewness et du kurtosis

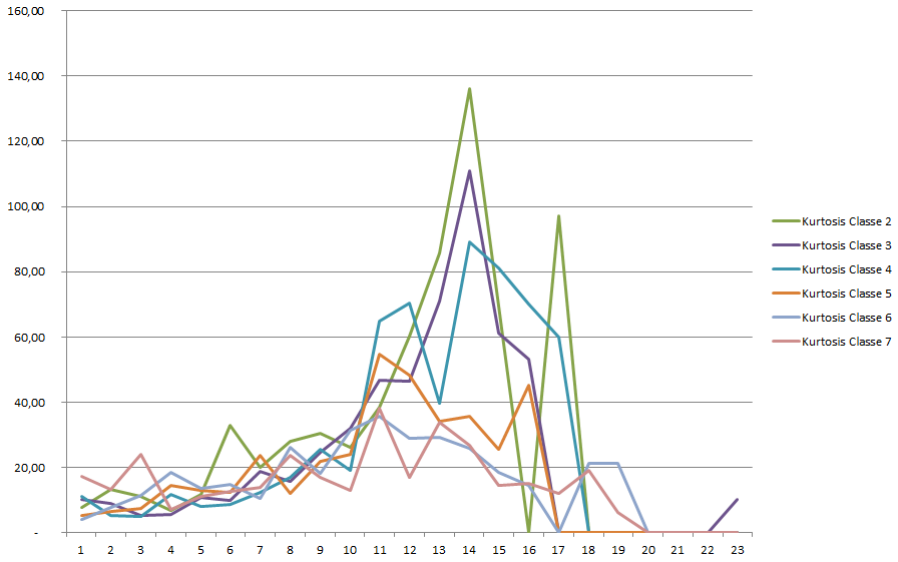
Dans cette annexe, nous présentons les graphes des skewness et kurtosis empiriques de la variables $T_{j+1,k}$ en fonction de la maturité j utilisés pour l'approximation de Cornish-Fisher, dans le chapitre 4.

B.2.1 Skewness



B.2.2 Kurtosis





Annexe C

Régression logistique

Cette annexe présente les résultats de la régression logistique effectuée dans la sous-partie 5.4.2. Le tableau ci-dessous ne correspond pas directement aux résultats sortant de la procédure SAS. Les paramètres estimés par le logiciel ont d'abord mis sous la forme des équations du modèle.

Équations du modèle (1/2)

	Intercept	Classe départ	nb_vivant	nb_deces	nb_fil4	nb_fil123	nb_taux_ipp1	nb_taux_ipp3	ipp_max
0AE-0AT	-0,07	0,00	-0,03	0,00	0,01	-0,05	-0,03	0,18	-0,04
0AE-0BE	-3,48	0,00	-0,03	0,18	0,09	-0,04	-0,03	0,17	-0,04
0AE-0BT	-3,67	0,00	-0,03	-0,02	0,00	-0,05	-0,02	0,18	-0,04
0AE-1AE	-1,46	0,00	-0,02	0,05	0,09	0,04	-0,02	0,16	-0,04
0AE-1AT	-6,78	0,00	-0,04	-0,02	0,00	0,03	-0,04	0,18	-0,04
0AE-1BE	-4,03	0,00	0,00	0,68	0,14	0,06	0,01	0,14	-0,04
0AE-1BT	-7,47	0,00	-0,03	-0,02	0,00	-0,06	-0,04	0,19	-0,03
0AE-2AE	-6,81	0,00	0,01	-0,39	2,69	0,02	0,03	-0,25	-0,06
0AE-3AE	-7,48	0,00	-0,04	-0,38	2,70	-0,07	-0,04	-0,23	-0,05
0AE-3BE	-7,50	0,00	0,01	-0,03	-0,02	0,12	-0,06	0,16	0,03
0AE-4AE	-7,50	0,00	0,01	-0,39	2,68	0,02	-0,06	-0,25	0,02
0AE-5AE	-7,52	0,00	0,06	-0,41	2,68	0,01	0,10	-0,26	-0,07
0AT-0AT	4,37	0,34	-0,01	-0,10	-0,12	-0,04	0,02	-0,15	-0,02
0AT-0BT	-2,49	0,23	0,01	0,14	0,00	-0,02	0,01	-0,13	-0,01
0AT-1AT	-5,81	0,23	-0,02	-0,10	-0,10	-0,02	-0,02	-0,10	0,00
0BE-0BE	-1,81	12,08	0,01	-0,05	-0,07	-0,05	0,03	-0,08	0,00
0BE-0BT	-2,04	12,08	0,01	0,12	-0,04	-0,05	0,03	-0,09	0,00
0BE-0CE	-7,45	12,08	0,16	-0,13	-0,13	-0,01	0,15	-0,12	0,05
0BE-1BE	-4,24	12,07	0,02	-0,11	0,01	0,03	0,03	-0,10	0,00
0BE-1BT	-6,73	12,08	0,04	-0,11	-0,10	0,06	0,08	-0,09	-0,02
0BE-2BE	-6,71	12,08	-0,04	5,60	1,08	-0,02	0,01	-0,33	-0,01
0BT-0BT	2,55	1,91	0,01	0,10	-0,05	-0,01	0,03	-0,13	-0,01
0BT-0CT	-4,61	1,80	0,03	1,11	0,95	-0,01	0,01	-0,30	0,03
0BT-1BT	-5,24	1,79	0,04	-0,13	-0,12	0,06	0,06	-0,12	-0,01
0CE-0CE	-7,41	13,88	-0,07	11,30	2,26	-0,07	0,01	-0,59	-0,01
0CE-0CT	-6,74	13,88	0,10	-0,11	-0,12	0,00	0,07	-0,10	0,06
0CT-0CT	-2,21	13,18	0,05	1,27	0,88	0,03	0,02	0,92	0,04
0CT-1CT	-7,42	13,17	0,05	-0,10	-0,11	0,01	0,00	-0,10	0,07
1AE-0AE	-2,75	0,70	0,01	-0,12	0,01	-0,02	0,03	-0,13	-0,02
1AE-0AT	1,62	0,72	-0,01	-0,08	-0,10	-0,03	0,02	-0,11	-0,02
1AE-0BE	-6,04	0,70	-0,02	2,74	0,47	0,01	0,01	-0,24	0,00
1AE-0BT	-1,48	0,70	0,01	-0,02	-0,09	-0,03	0,03	-0,11	-0,02
1AE-1AE	3,02	0,76	0,00	-0,05	-0,01	0,01	-0,01	-0,13	0,00
1AE-1AT	1,49	0,71	0,00	-0,07	-0,05	0,02	-0,02	-0,11	0,01
1AE-1BE	0,77	0,70	0,01	0,11	0,02	0,02	0,01	-0,12	-0,01
1AE-1BT	-2,30	0,70	0,00	0,03	-0,07	0,03	-0,01	-0,12	0,01
1AE-1CE	-5,68	0,66	0,09	-0,14	-0,13	0,16	0,15	-0,14	-0,05
1AE-2AE	-3,08	0,69	0,01	1,30	1,70	0,01	-0,02	-0,42	0,02
1AE-2AT	-7,43	0,71	-0,08	11,29	2,24	-0,08	0,00	-0,62	-0,03
1AE-2BE	-3,56	0,70	-0,02	2,34	1,65	-0,02	-0,02	0,73	0,01
1AE-2CE	-7,52	0,63	0,11	11,22	2,19	0,26	0,27	-0,69	-0,09
1AE-3AE	-4,49	0,69	0,00	1,45	2,10	-0,03	-0,02	-0,48	0,02
1AE-3BE	-4,66	0,70	-0,03	5,40	2,38	-0,04	0,00	2,87	-0,01
1AE-4AE	-5,49	0,70	0,00	1,20	2,47	-0,03	0,00	7,33	-0,01
1AE-4BE	-6,04	0,70	-0,01	-0,37	1,91	-0,04	0,00	-0,42	-0,01
1AE-5AE	-5,83	0,69	0,01	-0,40	2,04	-0,01	0,02	-0,44	-0,02
1AE-5BE	-5,36	0,69	0,00	2,36	3,11	-0,05	0,02	6,23	-0,03
1AT-1AT	3,97	0,50	0,00	-0,07	-0,04	0,02	-0,02	-0,13	0,01
1AT-1BT	-4,10	0,38	0,02	-0,12	-0,02	0,04	-0,01	-0,13	0,02
1BE-0BE	-7,44	1,77	0,03	-0,13	-0,14	0,10	-0,02	-0,12	0,05
1BE-0BT	0,14	1,79	0,01	0,10	-0,07	0,00	0,03	-0,13	-0,01
1BE-0CT	-5,24	1,77	0,03	-0,17	0,16	0,06	0,00	-0,17	0,03
1BE-1BE	2,21	1,85	0,01	0,10	0,08	0,03	-0,01	-0,10	0,02
1BE-1BT	0,81	1,80	0,01	0,01	0,00	0,04	-0,01	-0,11	0,02
1BE-1CE	-1,43	1,76	0,06	0,35	0,47	0,10	0,01	0,03	0,03
1BE-1CT	-4,18	1,77	0,04	-0,15	0,07	0,09	-0,03	-0,16	0,04
1BE-2BE	-2,75	1,77	0,01	1,13	1,87	-0,02	-0,03	-0,44	0,03
1BE-2BT	-7,42	1,77	-0,08	11,27	2,22	-0,08	0,00	-0,62	-0,03
1BE-2CE	-4,87	1,77	0,00	-0,39	1,94	-0,02	-0,06	-0,43	0,05
1BE-3BE	-3,91	1,77	0,01	1,31	2,03	-0,02	-0,02	1,15	0,02
1BE-3CE	-5,65	1,76	0,03	1,47	2,50	0,02	-0,02	-0,55	0,01
1BE-4BE	-4,66	1,78	0,00	-0,47	2,11	-0,03	-0,03	16,74	-0,01
1BE-4CE	-6,75	1,75	0,00	5,21	3,74	-0,05	0,03	-0,79	0,00
1BE-5BE	-5,83	1,76	0,01	1,86	2,50	-0,02	0,04	-0,55	-0,02
1BE-5CE	-6,74	1,78	0,00	-0,53	2,36	-0,04	-0,05	27,01	0,01
1BT-1BT	2,72	1,62	0,01	0,09	0,04	0,04	-0,01	-0,13	0,02
1BT-1CT	-4,45	1,51	0,03	0,89	1,04	0,04	-0,03	-0,31	0,04
1CE-0CT	-4,60	10,40	0,03	1,78	1,05	0,07	0,00	2,91	0,04
1CE-1CE	-0,64	10,43	0,08	0,36	0,57	0,11	0,01	0,61	0,04
1CE-1CT	-2,10	10,41	0,07	0,45	0,60	0,11	0,00	-0,25	0,04
1CE-2CE	-4,73	10,40	0,06	-0,40	2,02	0,06	-0,03	-0,45	0,05
1CE-3CE	-7,41	10,40	-0,01	-0,46	2,58	-0,08	-0,08	-0,51	0,07
1CE-4CE	-7,43	10,39	0,04	-0,82	5,27	-0,09	-0,09	-0,94	0,05
1CT-1CT	-0,53	10,76	0,06	0,53	0,82	0,10	-0,01	-0,26	0,04
2AE-0AT	-6,33	12,52	0,00	3,32	3,34	-0,05	0,03	-0,67	0,00
2AE-1AE	-4,71	12,53	0,01	0,29	2,72	-0,04	-0,05	-0,53	0,06
2AE-1AT	-5,35	12,52	0,01	6,77	3,01	-0,01	0,00	-0,68	0,02
2AE-1BE	-6,72	12,53	-0,01	-0,46	2,57	-0,07	-0,07	-0,49	0,07
2AE-2AE	-1,91	12,54	0,01	1,84	2,60	-0,02	-0,03	-0,09	0,05
2AE-2AT	-4,78	12,53	0,00	2,88	2,46	-0,02	-0,03	-0,53	0,05
2AE-2BE	-3,71	12,53	0,01	1,21	2,76	-0,03	-0,03	0,79	0,05
2AE-3AE	-4,39	12,54	0,04	0,62	2,63	0,02	0,00	2,08	0,04
2AE-3BE	-6,03	12,54	0,02	-0,38	1,89	0,00	-0,06	-0,40	0,07
2AE-4AE	-6,72	12,52	0,02	-0,65	3,91	-0,08	-0,08	-0,70	0,06
2AE-4BE	-6,32	12,51	-0,04	7,24	3,23	-0,07	-0,03	-0,70	0,03
2AE-5AE	-7,54	12,57	0,35	-0,54	2,49	0,53	0,16	-0,61	-0,03
2AE-5BE	-7,48	12,48	0,13	-1,60	10,63	-0,12	0,23	-1,77	-0,08
2AT-2AT	-1,90	12,93	0,01	2,40	2,49	-0,02	-0,03	-0,52	0,05

Équations du modèle (2/2)

	Intercept	Classe départ	nb_vivant	nb_deces	nb_fil4	nb_fil123	nb_taux_ipp1	nb_taux_ipp3	ipp_max
2BE-0BT	-6,79	11,32	0,07	16,54	6,01	-0,07	0,17	-1,31	0,01
2BE-1BE	-3,84	11,44	0,01	0,45	2,71	-0,05	-0,06	0,99	0,06
2BE-1BT	-3,86	11,44	-0,01	2,17	2,60	-0,05	-0,05	-0,54	0,05
2BE-1CE	-5,82	11,47	0,05	-0,43	1,98	0,05	-0,03	-0,42	0,05
2BE-1CT	-7,49	11,31	0,14	10,49	7,53	-0,12	0,13	-1,47	0,01
2BE-2BE	-1,23	11,45	0,01	2,06	2,88	-0,03	-0,04	-0,36	0,05
2BE-2BT	-3,66	11,43	-0,01	2,75	2,68	-0,05	-0,04	-0,56	0,04
2BE-2CE	-3,52	11,43	0,01	2,72	2,96	-0,02	-0,03	-0,61	0,05
2BE-2CT	-7,43	11,41	-0,03	11,23	2,17	0,01	0,07	-0,61	-0,03
2BE-3BE	-3,93	11,44	0,02	0,86	2,97	-0,03	-0,04	-0,59	0,05
2BE-3CE	-6,04	11,45	0,04	-0,50	2,52	0,01	-0,03	-0,50	0,05
2BE-4BE	-6,33	11,44	0,02	-0,65	3,29	-0,05	-0,08	17,73	0,06
2BE-4CE	-6,77	11,44	0,08	5,34	2,33	0,14	0,18	-0,58	-0,05
2BE-5CE	-6,85	11,43	0,39	5,11	3,63	0,18	0,34	-0,86	-0,07
2BT-2BT	-1,53	12,53	0,00	2,55	2,62	-0,04	-0,04	-0,24	0,05
2BT-2CT	-7,45	12,54	0,10	-0,50	2,55	0,10	-0,02	-0,52	0,04
2CE-0CT	-7,46	12,82	-0,04	22,64	4,54	0,00	0,05	-1,14	0,03
2CE-1CE	-4,80	12,86	0,07	-0,48	2,56	0,05	-0,03	-0,51	0,05
2CE-1CT	-4,73	12,84	0,03	4,03	3,86	-0,01	-0,02	-0,77	0,05
2CE-2CE	-2,09	12,86	0,05	2,00	2,82	0,02	-0,03	-0,58	0,05
2CE-2CT	-4,60	12,85	0,05	0,22	2,54	0,02	-0,02	-0,51	0,05
2CE-3CE	-4,86	12,85	0,04	1,34	2,51	0,02	-0,04	-0,52	0,05
2CT-2CT	-2,79	13,48	0,03	1,35	2,66	0,01	-0,03	-0,54	0,05
3AE-1AT	-6,34	13,13	0,05	-0,54	2,30	0,04	-0,01	36,17	0,00
3AE-1BE	-6,39	13,15	0,17	-0,55	2,41	0,27	0,15	17,77	-0,04
3AE-2AE	-5,10	13,12	-0,01	-0,46	2,54	-0,07	-0,07	4,99	0,05
3AE-2AT	-5,81	13,12	0,01	-0,46	2,58	-0,04	-0,06	-0,53	0,07
3AE-2BE	-6,71	13,11	-0,04	5,37	2,21	-0,07	-0,03	26,97	-0,05
3AE-3AE	-2,48	13,13	0,02	1,61	2,48	0,00	-0,02	4,14	0,04
3AE-3AT	-5,80	13,12	-0,02	4,24	2,44	-0,05	-0,03	-0,57	0,03
3AE-3BE	-4,29	13,12	0,01	3,57	2,77	-0,02	-0,03	4,16	0,03
3AE-4AE	-6,05	13,12	0,04	2,38	3,16	0,01	-0,03	-0,67	0,05
3AT-3AT	-3,59	13,70	-0,03	5,93	2,41	-0,04	-0,01	-0,55	0,02
3BE-1BE	-7,43	12,14	0,03	-1,06	4,44	-0,08	-0,08	109,20	-0,11
3BE-1BT	-6,03	12,29	-0,01	2,45	2,47	-0,07	-0,01	-0,61	0,03
3BE-1CE	-6,78	12,19	0,05	10,46	7,37	-0,05	0,10	25,99	-0,02
3BE-2BE	-4,16	12,29	0,00	2,61	2,98	-0,05	-0,04	-0,70	0,05
3BE-2BT	-5,22	12,29	0,01	2,16	2,13	-0,01	-0,05	5,55	0,04
3BE-2CE	-6,74	12,29	0,04	-0,67	3,89	-0,04	0,00	-0,81	0,05
3BE-2CT	-7,41	12,27	-0,07	11,27	2,20	-0,07	-0,08	-0,69	0,07
3BE-3BE	-1,65	12,29	0,01	2,19	2,81	-0,03	-0,04	6,34	0,03
3BE-3BT	-4,85	12,28	0,01	1,26	2,88	-0,04	-0,06	3,57	0,04
3BE-3CE	-3,97	12,28	0,03	3,79	3,32	-0,01	-0,01	0,94	0,04
3BE-4BE	-4,86	12,27	0,03	0,30	3,02	-0,03	-0,05	16,25	0,03
3BE-4CE	-6,32	12,30	0,03	-0,48	2,55	-0,01	-0,08	-0,59	0,06
3BE-5BE	-6,37	12,20	0,06	6,65	6,52	-0,04	0,06	35,34	0,00
3BT-3BT	-3,09	13,59	0,02	0,45	2,70	-0,03	-0,06	4,53	0,05
3CE-1CE	-7,41	13,11	-0,01	-0,46	2,58	-0,07	-0,07	-0,49	0,07
3CE-1CT	-7,50	13,16	0,25	-0,96	4,81	0,27	-0,05	54,09	0,00
3CE-2CE	-4,95	13,12	0,06	-0,54	2,98	0,00	-0,03	4,01	0,05
3CE-2CT	-5,22	13,12	0,01	0,80	2,83	-0,04	-0,05	-0,56	0,06
3CE-3CE	-2,34	13,12	0,04	1,98	3,14	0,00	-0,04	0,05	0,05
3CE-3CT	-5,51	13,11	0,02	11,01	4,13	0,02	0,04	-0,92	0,03
3CE-4CE	-6,05	13,12	0,05	2,29	3,82	0,01	-0,05	-0,75	0,05
3CT-3CT	-4,51	13,83	0,04	7,43	3,33	0,06	0,04	-0,74	0,04
4AE-3AE	-5,48	13,57	0,04	1,19	2,42	0,03	-0,02	15,13	0,01
4AE-4AE	-3,33	13,55	0,01	2,00	2,57	-0,02	-0,02	28,74	-0,01
4AE-4AT	-6,32	13,58	-0,01	3,35	3,37	-0,08	-0,05	-0,75	0,06
4AE-4BE	-5,49	13,55	0,05	2,61	3,73	0,00	-0,01	38,48	-0,02
4AE-5AE	-7,40	13,52	-0,01	-0,55	2,20	-0,07	-0,07	54,51	-0,09
4AT-4AT	-3,78	13,74	-0,01	3,15	3,33	-0,08	-0,05	-0,65	0,06
4BE-2BT	-7,41	13,10	-0,01	-0,45	2,58	-0,07	-0,07	-0,67	0,07
4BE-3BE	-6,32	12,99	0,02	-0,65	3,20	-0,05	-0,05	35,90	0,01
4BE-3BT	-7,41	13,10	-0,01	-0,45	2,58	-0,07	-0,07	-0,67	0,07
4BE-4BE	-2,36	13,03	0,01	1,29	2,78	-0,03	-0,04	24,34	0,00
4BE-4BT	-6,73	13,02	0,04	-0,51	2,37	0,01	-0,04	26,85	0,02
4BE-4CE	-4,17	13,03	0,04	1,61	3,35	-0,01	-0,04	24,55	0,03
4BE-5BE	-5,62	12,99	0,02	-0,53	2,31	-0,04	-0,03	36,04	-0,03
4BE-5CE	-5,80	13,00	0,01	-0,52	2,34	-0,03	-0,04	32,37	-0,03
4BT-4BT	-4,46	13,81	0,00	-0,54	2,22	-0,04	-0,05	51,66	-0,07
4CE-2CE	-6,73	13,31	-0,05	11,11	3,57	-0,08	0,00	-0,94	0,02
4CE-2CT	-6,32	13,31	-0,01	3,46	2,46	-0,04	-0,02	-0,67	0,04
4CE-3CE	-6,36	13,29	0,11	-0,63	3,32	0,10	-0,02	17,56	0,04
4CE-4CE	-2,63	13,27	0,07	1,62	3,95	0,02	-0,04	27,50	0,04
4CE-4CT	-6,04	13,31	0,06	-0,55	3,24	-0,01	-0,05	-0,75	0,06
4CE-5CE	-5,34	13,26	0,01	-0,51	2,38	-0,04	-0,07	26,91	0,03
4CT-4CT	-5,47	13,86	0,01	-0,44	2,60	-0,05	-0,08	-0,49	0,07
5AE-4AE	-6,75	13,58	0,03	5,35	2,20	0,01	0,07	26,85	-0,03
5AE-5AE	-3,43	13,53	0,03	0,32	2,19	0,00	-0,04	50,37	-0,04
5AE-5BE	-5,03	13,53	0,04	-0,59	2,43	0,01	-0,05	54,41	-0,03
5BE-4BE	-7,46	12,91	0,02	11,17	1,79	0,10	-0,03	54,12	0,04
5BE-4BT	-6,71	12,88	-0,01	-0,55	2,19	-0,07	-0,07	54,26	-0,09
5BE-5BE	-2,43	12,92	0,03	0,37	2,42	0,00	-0,04	49,38	-0,03
5BE-5BT	-6,73	13,04	0,01	-0,50	2,37	-0,03	-0,08	26,71	0,06
5BE-5CE	-4,18	12,92	0,04	0,67	3,09	-0,01	-0,02	49,85	-0,03
5BT-5BT	-5,13	13,87	0,03	-0,54	2,26	0,00	-0,09	43,55	0,06
5CE-3CT	-7,41	13,32	-0,01	-0,44	2,59	-0,07	-0,07	-0,76	0,07
5CE-4CE	-6,32	13,16	0,00	3,28	3,10	-0,04	-0,05	35,78	-0,05
5CE-5CE	-2,51	13,11	0,05	0,78	3,01	-0,01	-0,03	51,36	-0,02
5CE-5CT	-7,40	13,08	-0,01	-0,55	2,19	-0,07	-0,07	54,32	-0,09
IBNR-0AE	-2,05	0,00	0,00	-0,10	-0,07	-0,05	0,02	-0,09	-0,01
IBNR-0AT	-2,69	0,00	-0,01	0,01	-0,07	-0,05	0,01	-0,09	-0,02
IBNR-0BE	-3,10	0,00	0,01	0,05	-0,03	-0,05	0,03	-0,10	0,00
IBNR-0BT	-3,52	0,00	0,00	0,60	-0,01	-0,03	0,03	-0,12	-0,01
IBNR-0CE	-6,33	0,00	0,02	3,68	0,67	-0,02	0,08	-0,28	-0,02
IBNR-1AE	0,25	0,00	0,00	-0,07	-0,08	0,00	0,02	-0,10	-0,01
IBNR-1AT	-4,47	0,00	-0,01	-0,10	-0,10	-0,03	0,00	-0,09	-0,01
IBNR-1BE	-0,96	0,00	0,01	-0,01	-0,02	0,02	0,04	-0,11	-0,01
IBNR-1BT	-4,93	0,00	0,00	-0,10	-0,10	-0,01	0,01	-0,09	0,00
IBNR-1CE	-6,34	0,00	0,04	-0,11	-0,12	0,10	-0,02	-0,11	0,06
IBNR-2AE	-7,42	0,00	-0,01	-0,46	2,60	-0,08	0,00	-0,50	-0,02
IBNR-2BE	-5,35	0,00	0,02	-0,38	1,92	0,01	0,06	-0,41	-0,03
IBNR-2CE	-7,44	0,00	0,03	-0,47	2,59	0,01	0,07	-0,52	-0,03
IBNR-3AE	-7,41	0,00	-0,01	-0,10	-0,10	0,02	0,00	-0,09	-0,01
IBNR-4AE	-7,41	0,00	-0,08	-0,10	-0,09	0,03	0,01	-0,08	-0,01
IBNR-4BE	-7,42	0,00	-0,01	-0,46	2,60	-0,08	0,00	-0,50	-0,02
IBNR-IBNR	-0,62	0,00	-0,01	0,12	0,00	-0,02	0,01	-0,11	-0,01

Annexe D

Matrices

Cette annexe présente les différentes matrices construites dans le chapitre concernant la deuxième amélioration de la méthode de Mahon :

- la matrice des nombres de transitions
- la matrice de transition M de la chaîne de Markov
- la matrice M^{∞}

Nombres de transitions

	8NR	0AE	1AE	2AE	3AE	4AE	5AE	0BE	1BE	2BE	3BE	4BE	5BE	0CE	1CE	2CE	3CE	4CE	5CE
8NR	852	212	2 122	1	1	1	-	75	643	8	-	-	-	3	3	1	-	-	-
0AE	-	1 679	413	2	1	1	1	54	32	-	1	-	-	-	6	-	-	-	-
1AE	-	107	34 257	78	19	7	5	4	3 626	48	16	4	8	-	6	1	-	-	-
2AE	-	-	15	247	21	2	1	-	2	41	4	3	1	-	-	-	-	-	-
3AE	-	-	-	10	141	4	-	-	3	2	23	-	-	-	-	-	-	-	-
4AE	-	-	-	-	7	60	-	-	-	-	-	7	-	-	-	-	-	-	-
5AE	-	-	-	-	-	2	54	-	-	-	-	-	11	-	-	-	-	-	-
0BE	-	-	-	-	-	-	-	269	-	-	-	-	-	1	-	-	-	-	-
1BE	-	-	-	-	-	-	-	1	15 422	2	34	16	5	-	412	13	6	2	2
2BE	-	-	-	-	-	-	-	-	36	108	33	3	-	-	5	50	4	2	2
3BE	-	-	-	-	-	-	-	-	1	26	322	13	3	-	2	2	32	3	-
4BE	-	-	-	-	-	-	-	-	-	-	3	138	6	-	-	-	-	26	26
5BE	-	-	-	-	-	-	-	-	-	-	-	1	148	1	-	-	-	-	-
0CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	897	15	1	1	-
1CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	14	210	13	-	-
2CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	12	162	4	-
3CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	3	124	8
4CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	137
5CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	0AT	1AT	2AT	3AT	4AT	5AT	0BT	1BT	2BT	3BT	4BT	5BT	0CT	1CT	2CT	3CT	4CT	5CT
8NR	112	19	-	-	-	-	49	12	-	-	-	-	-	-	-	-	-	-
0AE	1 645	2	-	-	-	-	45	1	-	-	-	-	-	-	-	-	-	-
1AE	8 485	7 422	1	-	-	-	380	168	-	-	-	-	-	-	-	-	-	-
2AE	3	8	14	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3AE	-	3	5	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-
4AE	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5AE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0BE	-	-	-	-	-	-	215	-	-	-	-	-	-	-	-	-	-	-
1BE	-	-	-	-	-	-	1 941	2	-	-	-	-	9	26	-	-	-	-
2BE	-	-	-	-	-	-	2	3 793	1	-	-	-	-	1	1	-	-	-
3BE	-	-	-	-	-	-	-	35	43	13	-	-	-	-	-	-	-	-
4BE	-	-	-	-	-	-	-	4	9	1	2	-	-	-	-	-	-	-
5BE	-	-	-	-	-	-	-	-	1	-	-	2	-	-	-	-	-	-
0CE	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-
1CE	-	-	-	-	-	-	-	-	-	-	-	-	17	209	-	-	-	-
2CE	-	-	-	-	-	-	-	-	-	-	-	-	1	15	17	-	-	-
3CE	-	-	-	-	-	-	-	-	-	-	-	-	-	1	9	7	-	-
4CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	4	-
5CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	1
0AT	132 462	-	-	-	-	-	139	-	-	-	-	-	-	-	-	-	-	-
1AT	-	89 033	-	-	-	-	-	28	-	-	-	-	-	-	-	-	-	-
2AT	-	-	250	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3AT	-	-	-	46	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4AT	-	-	-	-	38	-	-	-	-	-	-	-	-	-	-	-	-	-
5AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0BT	-	-	-	-	-	-	21 489	9	-	-	-	-	-	-	-	-	-	-
1BT	-	-	-	-	-	-	-	25 701	-	-	-	-	-	-	-	-	-	-
2BT	-	-	-	-	-	-	-	362	-	-	-	-	-	-	-	-	-	-
3BT	-	-	-	-	-	-	-	-	76	-	-	-	-	-	-	-	-	-
4BT	-	-	-	-	-	-	-	-	-	19	-	-	-	-	-	-	-	-
5BT	-	-	-	-	-	-	-	-	-	-	10	-	-	-	-	-	-	-
0CT	-	-	-	-	-	-	-	17	-	-	-	185	1	-	-	-	-	-
1CT	-	-	-	-	-	-	-	20	-	-	-	998	-	-	-	-	-	-
2CT	-	-	-	-	-	-	-	-	1	-	-	-	103	-	-	-	-	-
3CT	-	-	-	-	-	-	-	-	-	-	-	-	-	19	-	-	-	-
4CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7	-	-	-
5CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Matrice de transition M

	0BNR	0AE	1AE	2AE	3AE	4AE	5AE	0BE	1BE	2BE	3BE	4BE	5BE	0CE	1CE	2CE	3CE	4CE	5CE	0AT	1AT	2AT	3AT	4AT	5AT	0BT	1BT	2BT	3BT	4BT	5BT	0CT	1CT	2CT	3CT	4CT	5CT
0BNR	0,21	0,05	0,51	0,00	0,00	0,00	0,00	0,02	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
0AE	0,05	0,43	0,11	0,00	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,42	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1AE	0,00	0,63	0,04	0,00	0,00	0,00	0,00	0,00	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,16	0,14	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2AE	0,04	0,69	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,02	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3AE	0,05	0,68	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,03	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4AE	0,03	0,77	0,01	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,03	0,03	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5AE	0,09	0,72	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
0BE	0,00	0,52	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1BE	0,00	0,71	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2BE	0,00	0,69	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3BE	0,01	0,75	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4BE	0,01	0,78	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5BE	0,01	0,83	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
0CE	0,33	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1CE	0,79	0,05	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2CE	0,01	0,78	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3CE	0,05	0,83	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4CE	0,00	0,86	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5CE	0,02	0,96	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
0AT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1AT	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,42	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2AT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,16	0,14	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3AT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,02	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4AT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5AT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
0BT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,42	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1BT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2BT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,17	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3BT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4BT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5BT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
0CT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1CT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2CT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3CT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4CT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5CT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

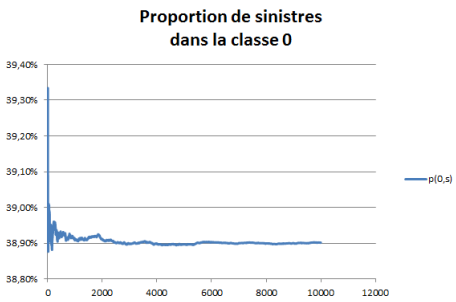
Matrice M^{∞}

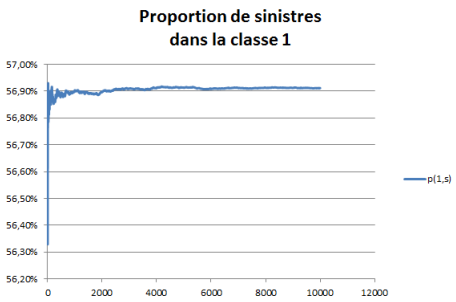
	IBNR	0AE	1AE	2AE	3AE	4AE	5AE	0BE	1BE	2BE	3BE	4BE	5BE	0CE	1CE	2CE	3CE	4CE	5CE	0AT	1AT	2AT	3AT	4AT	5AT	0BT	1BT	2BT	3BT	4BT	5BT	0CT	1CT	2CT	3CT	4CT	5CT
IBNR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0AE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,35	0,25	0,00	0,00	0,00	-	0,15	0,20	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00
1AE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,83	0,07	0,00	0,00	0,00	-	0,06	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2AE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,42	0,37	0,00	0,00	0,00	-	0,07	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3AE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,08	0,13	0,14	0,02	0,01	-	0,04	0,15	0,10	0,02	0,01	0,01	0,01	0,01	0,00	0,00	0,02	0,02
4AE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,02	0,08	0,12	0,10	0,01	-	0,01	0,12	0,09	0,06	0,01	0,01	0,01	0,10	0,13	0,06	0,03	0,02
5AE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,01	0,03	0,05	0,04	0,17	-	0,01	0,05	0,05	0,04	0,03	0,01	0,01	0,08	0,16	0,10	0,10	0,07
0BE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,00	0,00	0,01	0,01	0,09	-	0,00	0,01	0,01	0,01	0,06	0,00	0,00	0,06	0,17	0,23	0,14	0,10
1BE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,91	0,60	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00
2BE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,05	0,38	0,22	0,02	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00
3BE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,05	0,11	0,14	0,13	0,01	0,01	0,01	0,19	0,01	0,04	0,02	0,01
4BE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,00	0,01	0,03	0,03	0,00	0,00	0,01	0,10	0,23	0,17	0,05	0,03
5BE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,00	0,00	0,03	0,03	0,06	0,01	0,01	0,10	0,26	0,17	0,05	0,03
0CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,00	0,00	0,00	0,07	0,00	0,00	0,06	0,18	0,25	0,14	0,13
1CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	0,00	0,00	0,00	0,00	0,00	0,00
2CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5CE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,00	0,00	0,00	0,00	0,00	0,00	0,00
0AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5AT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-	-	-	-	-
0BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-	-	-	-
1BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-	-	-
2BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-	-
3BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-	-
4BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-	-
5BT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-	-
0CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-	-
1CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-	-
2CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-	-
3CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-	-
4CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-	-
5CT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,00	-

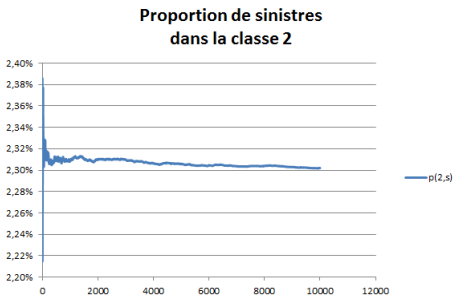
Annexe E

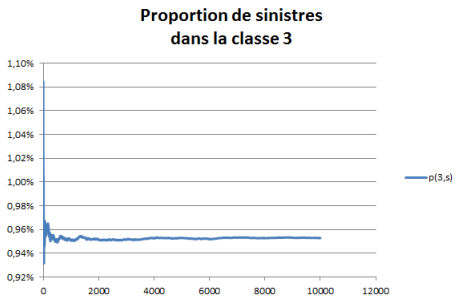
Convergences des simulations

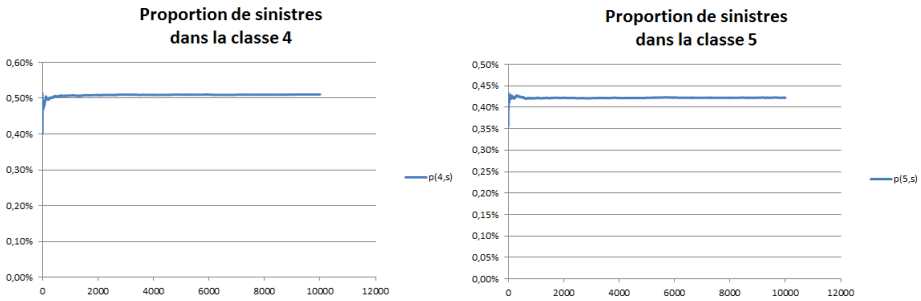
Cette annexe présente la proportion du nombre dans chaque classe en fonction du nombre de simulations effectuées dans la section 5.5.4.










Proportion de sinistres
dans la classe 4
0,60%
0,50%
0,40%
0,30%
0,20%
0,10%
0,00%
0
2000
4000
6000
8000
10000
12000
p(4,s)
Proportion de sinistres
dans la classe 5
0,50%
0,45%
0,40%
0,35%
0,30%
0,25%
0,20%
0,15%
0,10%
0,05%
0,00%
0
2000
4000
6000
8000
10000
12000
p(5,s)