

EURIA - Euro Institute of Actuarial Studies
Bureau d'études

Sous la direction de
Aude Goichon (KPMG)
Franck Vermet (UBO)
Françoise Pène (UBO)

Estimation de l'erreur de prédiction dans
le cas de l'utilisation d'une combinaison
de méthodes pour le calcul de provisions
en assurance IARD

A. Le Tesson
A. Lenain
S. Samba
J. Ung



Brest, le 26 mai 2014

Remerciements

Nous tenons à remercier toutes les personnes qui ont contribué directement ou indirectement à la rédaction de ce rapport, que ce soit par leurs questions ou par leur soutien.

Nous remercions tout particulièrement Mme Goichon pour nous avoir proposé ce sujet de bureau d'études particulièrement intéressant et qui malgré ses fortes contraintes professionnelles, a toujours su nous guider par ses réponses, sa documentation judicieuse, ainsi que pour ses remarques et critiques qui ont découlé de la lecture de nos travaux. Le travail fourni tout au long de l'année a été très enrichissant pour chacun d'entre nous.

Nous voulons également remercier nos référents EURIA, M. Vermet et Mme Pène pour leur suivi tout au long de l'année, leur intérêt pour nos travaux, leur disponibilité, leurs éclaircissements sur les différentes théories mathématiques abordées, mais aussi pour leurs conseils et astuces concernant la rédaction sous L^AT_EX.

M. Franke et M. Kamega, pour avoir assisté à nos présentations et qui de par leurs multiples interrogations ou réponses à nos questions, ont sans nul doute contribué à ce rapport.

Aussi, nous tenons à remercier la promotion EURIA, pour avoir assisté à nos présentations et émis des remarques pertinentes.

Enfin, nous souhaitons témoigner notre gratitude envers Mme Bellanger-Gruet pour son professionnalisme, sa bonne humeur ainsi que pour l'impression de nos documents.

Table des matières

I	Introduction	7
0.1	Notions générales	11
0.2	Données utilisées	12
II	Méthodes déterministes	13
1	Méthode Chain-Ladder	14
1.1	Présentation du principe général	14
1.2	Exemple d'application	15
1.3	Validation du modèle	17
2	Méthode London-Chain	21
2.1	Modèle théorique	21
2.2	Exemple d'application	23
3	Méthode de Bornhuetter-Ferguson	24
3.1	Présentation du principe général	24
3.2	Exemple d'application	26
4	Combinaison des deux méthodes	28
4.1	Comparaison de ces deux méthodes	28
4.2	Raisonnement théorique	29
4.3	Application	30
III	Méthodes stochastiques	31
5	Méthode de Mack	33
5.1	Hypothèses du modèle de Mack	33
5.2	Mesure de l'incertitude de l'estimateur	36
5.3	Exemple d'application	36
6	Méthode du Bootstrap - Variante Chain-Ladder	38

7	Méthode du Bootstrap - Combinaison Chain-Ladder/Bornhuetter-Ferguson	41
8	Méthode Bornhuetter-Ferguson Stochastique	43
8.1	Introduction des notations	43
8.2	Calcul de l'erreur de prédiction	44
8.3	Application numérique	47
IV	Etude des résultats	51
9	Résumé des résultats obtenus et commentaires	52
10	Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement dans le cadre de la méthode du Bootstrap CL/BF	54
V	Conclusion	59
VI	Annexes	61
A	Données utilisées	62
B	Codes R	65
B.1	Méthodes déterministes	66
B.1.1	Méthode Chain-Ladder	66
B.1.2	Méthode London-Chain	68
B.1.3	Méthode de Bornhuetter-Ferguson	69
B.1.4	Analyse graphique des loss ratios S/P	71
B.2	Méthodes Stochastiques	72
B.2.1	Méthode du Bootstrap/Chain-Ladder	72
B.2.2	Méthode du Bootstrap - Combinaison Chain-Ladder/Bornhuetter-Ferguson	76
B.2.3	Méthode Bornhuetter-Ferguson Stochastique	80

Première partie

Introduction

Une assurance est généralement décrite comme un service fournissant une prestation à la survenance d'un sinistre. La prestation, généralement de nature financière, est destinée à des assurés de différentes catégories juridiques : il peut s'agir d'une entreprise, d'une association ou d'un particulier. Toutefois, en échange de cette "protection", chaque assuré s'engage à verser auprès de l'assureur, et selon des modalités bien définies, une cotisation ou prime qui lui permettra d'effectuer son provisionnement.

La problématique du provisionnement est liée à la nature même de l'activité d'assurance à savoir, *l'inversion du cycle de production*. Les assureurs ayant pris l'engagement d'indemniser tous les sinistres survenus pendant la période de couverture, il convient de constituer des provisions pour indemniser les victimes d'un sinistre, même si celui-ci n'est déclaré, puis clôturé que des années plus tard. Nous distinguons trois catégories de sinistres :

- *Les sinistres clos* : A la date d'inventaire, le sinistre est déclaré clos si nous estimons que plus aucune indemnité ne sera versée à l'assuré. Il n'y a donc aucun engagement résiduel ;
- *Les sinistres ouverts* : Ce sont les sinistres qui ont été déclarés à l'assureur avant la date d'inventaire mais dont nous ne connaissons pas avec certitude les montants restants à régler à l'assuré ;
- *Les sinistres tardifs* : Ces sinistres sont les plus difficiles à estimer pour l'assureur car, bien qu'ils soient survenus durant la période de couverture, l'assureur ne connaît ni la quantité, ni la gravité de ces sinistres.

En particulier, pour les deux dernières catégories de sinistres, l'assureur devra donc estimer au mieux leur montant final dans le but d'assurer la solvabilité de sa compagnie et de manière plus globale, la pérennité du marché de l'assurance.

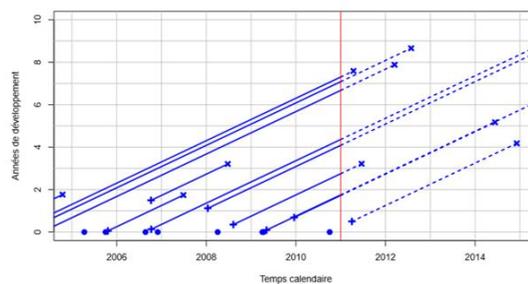


FIGURE 1 – Diagramme de Lexis

Le but de ce bureau d'études sera alors d'estimer *la provision pour sinistre à payer* (PSAP), constituant la majeure partie des provisions techniques en assurance non vie puis, d'en estimer l'erreur de prédiction. Pour permettre l'éva-

luation actuarielle de la PSAP, les modèles que nous utiliserons s'appuient sur l'étude des triangles de liquidation, et plus particulièrement dans le cadre de notre étude, sur les triangles de paiements. Ces triangles se décomposent en deux parties :

- La partie supérieure du triangle où figurent les montants des paiements déjà versés ;
- La partie inférieure dont nous devons estimer les montants encore inconnus à la date de calcul.

De plus, l'information présente dans ces triangles peut s'appréhender de trois façons distinctes : par année de survenance, par année de déroulé et par année calendaire.

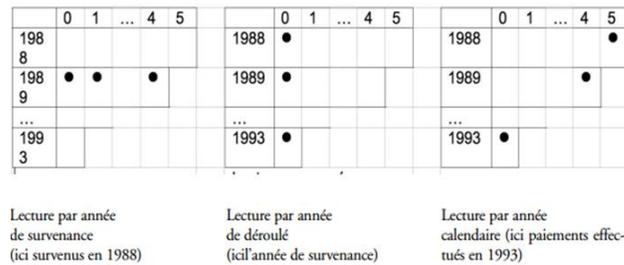


FIGURE 2 – Les différentes lectures d'un triangle de paiements

Pour obtenir le montant des provisions, nous devons estimer les montants de la partie inférieure du triangle de liquidation à partir de l'information disponible dans la partie supérieure de celui-ci.

Pour ce faire, nous verrons dans un premier temps des méthodes dites déterministes (Chain-Ladder, London-Chain, Bornhuetter-Ferguson) car ne possédant pas de cadre probabiliste. Comme le calcul de l'ensemble des provisions à partir d'une méthode n'est pas toujours optimal, les compagnies utilisent en pratique diverses méthodes de provisionnement. Nous nous focaliserons dans ce bureau d'études sur la combinaison des méthodes Chain-Ladder et Bornhuetter-Ferguson.

En revanche, les méthodes déterministes ne nous permettent pas de calculer l'erreur de prédiction. C'est pourquoi, dans un second temps, nous utiliserons des méthodes stochastiques inspirées de ces méthodes déterministes que sont la méthode de Mack, la méthode du Bootstrap et la méthode de Bornhuetter-Ferguson Stochastique.

Néanmoins, ces méthodes permettent uniquement le calcul de l'erreur de prédiction des provisions dans le cadre de l'utilisation d'une méthode bien précise :

à titre d'exemple, la méthode de Mack permet seulement d'estimer l'incertitude de la provision déterminée par la méthode Chain-Ladder. Ainsi, nous étudierons dans une dernière partie comment combiner plusieurs méthodes stochastiques pour estimer le montant des provisions et l'erreur de prédiction qui lui est associée.

0.1 Notions générales

Avant de débiter l'étude des différentes méthodes de provisionnement, introduisons quelques notions essentielles à la bonne compréhension des principes qui vont suivre ainsi que diverses notations que nous conserverons tout au long de ce rapport.

- en ligne, l'année de survenance, $i = 0, \dots, n$;
- en colonne, l'année de développement, $j = 0, \dots, n$;
- $Y_{i,j}$, les incréments de paiements pour l'année de développement j pour les sinistres survenus l'année i ;
- $C_{i,j}$, les paiements cumulés des j premières années de développement de l'année de survenance i .

Ainsi, à partir de ces notations, nous pouvons d'ores et déjà établir le triangle des incréments de paiements et le triangle des paiements cumulés associé.

Accident	Années de développement								
	0	1	...	k	...	n-i	...	n-1	n
0	$Y_{0,0}$	$Y_{0,1}$...	$Y_{0,k}$...	$Y_{0,n-i}$...	$Y_{0,n-1}$	$Y_{0,n}$
1	$Y_{1,0}$	$Y_{1,1}$...	$Y_{1,k}$...	$Y_{1,n-i}$...	$Y_{1,n-1}$	
						
i	$Y_{i,0}$	$Y_{i,1}$...	$Y_{i,k}$...	$Y_{i,n-i}$			
			...						
n-k	$Y_{n-k,0}$	$Y_{n-k,1}$...	$Y_{n-k,k}$					
n-1	$Y_{n-1,0}$	$Y_{n-1,1}$							
n	$Y_{n,0}$								

TABLE 1 – Triangle des incréments de paiements, $Y = (Y_{i,j})$.

Accident	Années de développement								
Années	0	1	...	k	...	n-i	...	n-1	n
0	$C_{0,0}$	$C_{0,1}$...	$C_{0,k}$...	$C_{0,n-i}$...	$C_{0,n-1}$	$C_{0,n}$
1	$C_{1,0}$	$C_{1,1}$...	$C_{1,k}$...	$C_{1,n-i}$...	$C_{1,n-1}$	
						
i	$C_{i,0}$	$C_{i,1}$...	$C_{i,k}$...	$C_{i,n-i}$			
			...						
n-k	$C_{n-k,0}$	$C_{n-k,1}$...	$C_{n-k,k}$					
n-1	$C_{n-1,0}$	$C_{n-1,1}$							
n	$C_{n,0}$								

TABLE 2 – Triangle des paiements cumulés, $C = (C_{i,j})$.

Nous pouvons, par ailleurs, établir la relation suivante :

$\forall i, k \in (0, 1, \dots, n), :$

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

0.2 Données utilisées

En ce qui concerne les données que nous utiliserons par la suite, nous avons décidé d'appliquer les différentes méthodes sur des données de marché d'un assureur automobile, extraites de l'article *Dependent Loss Reserving Using Copulas*, et dont nous avons joint le détail en annexe. Ces données présentent l'avantage de fournir les primes annuelles dont nous aurons besoin plus loin.

Deuxième partie

Méthodes déterministes

Chapitre 1

Méthode Chain-Ladder

1.1 Présentation du principe général

La méthode Chain-Ladder est une méthode standard, couramment utilisée sur les marchés depuis les années 30 du fait de sa simplicité.

Cette méthode est applicable à toute sorte de triangles (de dépenses, de charges, de nombre de sinistres). Cependant, au cours de notre étude, nous nous contenterons de l'appliquer uniquement au triangle des paiements cumulés $C = (C_{i,j})$.

Remarquons également que la méthode Chain-Ladder nécessite l'hypothèse forte suivante :

Pour $j = 0, 1, \dots, n-1$, les ratios des facteurs adjacents $\frac{C_{i,j+1}}{C_{i,j}}$ sont indépendants de l'année d'origine i .
Ainsi, $\forall j \in (0, 1, \dots, n-1)$:

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

Plus simplement, nous supposons que les années de survenance sont indépendantes.

En plus de ces deux hypothèses, la méthode Chain-Ladder repose sur le calcul des différents facteurs de développement (λ_j), également appelés *link ratios*.

Cette méthode de provisionnement suggère d'approcher ces *link ratios* par un estimateur naturel $\hat{\lambda}_j$ calculé à partir de la formule suivante :

$$\forall j = 0, \dots, n-1, \hat{\lambda}_j = \frac{\sum_{i=0}^{n-j-1} C_{i,j+1}}{\sum_{i=0}^{n-j-1} C_{i,j}}$$

Ainsi, à partir de ces conditions et de l'estimation des facteurs de développe-

ment, nous pouvons alors compléter le triangle inférieur des paiements cumulés :

$$\forall i, j \in (1, \dots, n),$$

$$C_{i,j} = C_{i,n-i} \hat{\lambda}_{n-i} \dots \hat{\lambda}_{j-1} = C_{i,n-i} \prod_{k=n-i}^{j-1} \lambda_k$$

En particulier, nous calculons alors les charges ultimes (S_i) pour chaque année d'accident i :

$$\forall i \in (0, 1, \dots, n),$$

$$S_i = C_{i,n} = C_{i,n-i} \prod_{k=n-i}^{n-1} \lambda_k$$

Enfin, nous déterminons respectivement les provisions par année d'accident (P_i) et la provision totale P par :

$$\forall i \in (1, \dots, n),$$

$$R_i = S_i - C_{i,n-i}$$

$$R = \sum_{i=1}^n R_i$$

1.2 Exemple d'application

Dans le but d'illustrer cette méthode, nous l'appliquons aux données présentées dans l'introduction et disponibles en annexe.

j	0	1	2	3	4	5	6	7	8
λ_j	1,794	1,195	1,090	1,045	1,020	1,010	1,004	1,002	1,001

TABLE 1.1 – Facteurs de développement en fonction des années de développement.

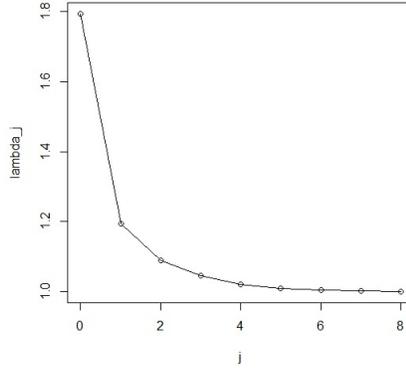


FIGURE 1.1 – Courbe des facteurs de développement ($\hat{\lambda}_j$).

Observations :

A partir du graphique précédent, nous observons une forte décroissance des facteurs de développement au cours du temps. En effet, pour un sinistre déclaré, la majorité des assurés sera en règle générale dédommagée les premières années. Cette dynamique est également observable dans le triangle des incréments de paiements (fournis en annexe) où nous constatons une décroissance des prestations pour une année d'accident i donnée. Ainsi, bien que la croissance des provisions cumulées, matérialisée par les facteurs de développement (supérieurs à 1), soit continue, celle-ci est moins importante que pour les premières années.

Années de développement										
Années	0	1	2	3	4	5	6	7	8	9
1988	1 376	2 588	3 123	3 437	3 605	3 685	3 725	3 740	3 750	3 755
1989	1 576	3 013	3 666	4 009	4 197	4 274	4 309	4 326	4 339	4 344
1990	1 763	3 304	3 982	4 347	4 524	4 602	4 649	4 675	4 688	4 693
1991	1 780	3 278	3 940	4 261	4 424	4 508	4 562	4 582	4 596	4 601
1992	1 843	3 417	4 030	4 329	4 506	4 613	4 660	4 681	4 695	4 700
1993	1 962	3 483	4 065	4 413	4 650	4 744	4 792	4 814	4 828	4 833
1994	2 033	3 464	4 097	4 530	4 733	4 828	4 877	4 899	4 913	4 919
1995	2 072	3 531	4 258	4 641	4 849	4 946	4 997	5 020	5 034	5 040
1996	2 211	3 728	4 455	4 856	5 074	5 176	5 229	5 252	5 268	5 273
1997	2 207	3 958	4 730	5 156	5 387	5 495	5 551	5 576	5 593	5 599

TABLE 1.2 – Triangle des paiements cumulés (en milliers) $C = (C_{i,j})$ et leurs projections futures $\hat{C} = (\hat{C}_{i,j})$.

De plus, l'objectif même de l'assureur étant de déterminer le montant des prestations qu'il devra verser dans les années à venir, nous en déduisons le tableau des provisions suivant :

Années	Provisions CL
1988	-
1989	4 727
1990	18 653
1991	38 931
1992	87 089
1993	182 757
1994	389 111
1995	781 855
1996	1 545 034
1997	3 391 735
Total	6 439 892

TABLE 1.3 – Provisions Chain-Ladder.

Remarque : Nous avons également programmé cette méthode sur R, le code étant joint en annexe. Ce programme nécessite le triangle des paiements cumulés et calcule :

1. Le D-triangle que nous développerons plus loin ;
2. Les facteurs de développements λ_j correspondants à chaque année de développement ;
3. Le triangle des paiements cumulés C avec leurs projections futures \hat{C} ;
4. La provision totale.

1.3 Validation du modèle

Bien que nous ayons déterminé la provision totale, vérifions à présent que les hypothèses établies en début de chapitre sont vérifiées et notamment, l'hypothèse de la constance des cadences de paiements.

Ainsi, en fixant j , si nous montrons que les $(n-j)$ couples $(C_{i,j}, C_{i,j+1})$, pour tout $i \in (0, \dots, n-j-1)$, sont "sensiblement" alignés suivant une droite passant par l'origine, cela supposera intuitivement l'existence d'un facteur de développement λ_j , constant par année de développement, tel que :

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

Nous nous contenterons ici de vérifier l'hypothèse pour les deux premières années de développement, c'est-à-dire, pour $j=0,1$. La représentation graphique

correspondante prendra la dénomination de C-C plots. Nous obtenons alors, sous R, les graphiques ci-dessous :

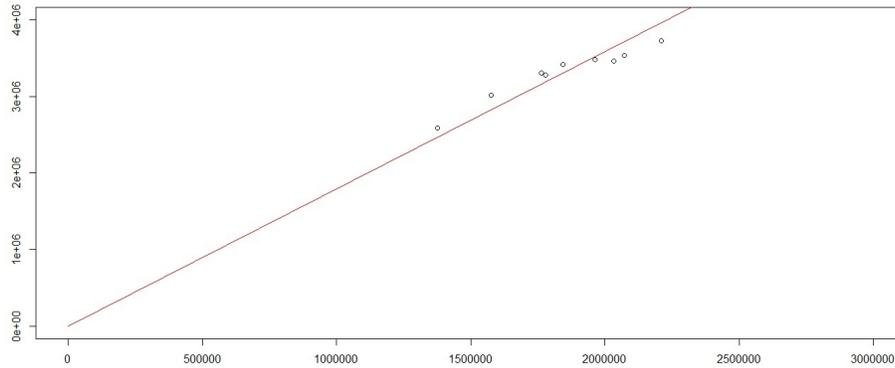


FIGURE 1.2 – C-C plot pour $j=0$.

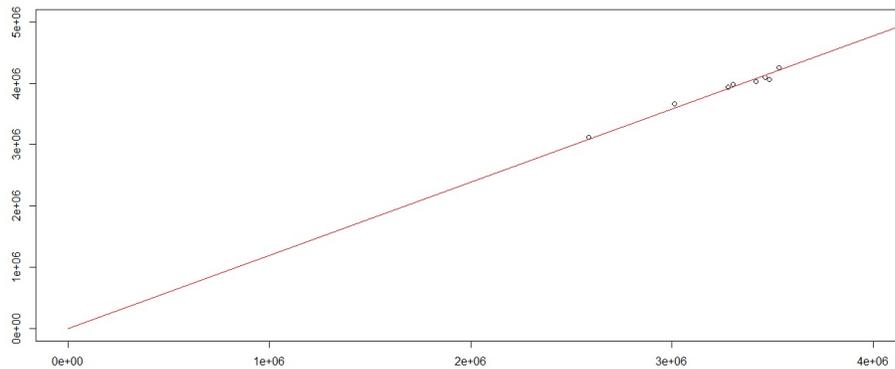


FIGURE 1.3 – C-C plot pour $j=1$.

Cependant, la validation essentielle repose sur l'étude du triangle de développement, dit *D-triangle*, composé des facteurs individuels suivants :

$$f_{i,j} = \frac{C_{i,j+1}}{C_{i,j}} \text{ pour } i + j \leq n - 1$$

Années	0	1	2	3	4	5	6	7	8
1988	1.8800	1.2071	1.1005	1.0489	1.0222	1.0106	1.0040	1.00209	1.0011
1989	1.9117	1.2165	1.0935	1.0471	1.0183	1.0082	1.0040	1.0029	
1990	1.8735	1.2055	1.0915	1.0407	1.0173	1.0103	1.0054		
1991	1.8420	1.2018	1.0816	1.0382	1.0191	1.0119			
1992	1.8537	1.1794	1.0743	1.0408	1.0236				
1993	1.7747	1.1671	1.0855	1.0540					
1994	1.7035	1.1829	1.1055						
1995	1.7039	1.2059							
1996	1.6864								

TABLE 1.4 – D-Triangle ($f_{i,j}$).

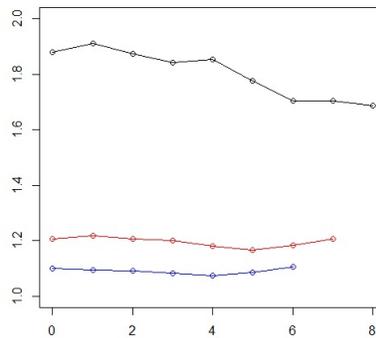


FIGURE 1.4 – Colonne du D-Triangle pour $j=0,1,2$.

Nous en déduisons alors quelques statistiques représentées dans le tableau qui suit :

j	0	1	2	3	4	5	6	7
Moyenne	1.8033	1.1958	1.0903	1.0450	1.0201	1.0102	1.0045	1.0029
Ecart-Type	0.1169	0.0287	0.0160	0.0091	0.0035	0.0021	0.0010	0.0000
$\frac{\text{Ecart-type}}{\text{Moyenne}} \cdot 100$	6.5%	2.4%	1.5%	0.9%	0.3%	0.2%	0.1%	0.0%

TABLE 1.5 – Observation via des données statistiques du D-Triangle.

Observations : Comme nous l'avons vu précédemment, la méthode Chain-Ladder s'appuie sur l'hypothèse forte de l'indépendance des années de surveillance i . Ainsi, pour vérifier cette condition, nous nous intéressons aux données relatives au D-triangle.

Dans notre cas, nous nous intéressons à l'étude de la moyenne et de l'écart-type des colonnes du D-triangle. Ainsi, pour $j=0, \dots, 7$, nous observons une variation maximale de 6.5% par rapport à la valeur moyenne (en valeur absolue). Nous supposons alors l'hypothèse vérifiée dans notre étude.

Chapitre 2

Méthode London-Chain

2.1 Modèle théorique

La méthode dite *London-Chain* a été introduite par Benjamin et Eagles pour le calcul des réserves au Lloyd's en 1986. L'hypothèse sous-jacente du modèle est que la dynamique des $C_{i,j}$, pour j fixé, est donnée par le modèle affine suivant :

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

avec, λ_j et β_j , des paramètres à estimer.

Ainsi, pour j toujours fixé, les $(n-j)$ couples $(C_{i,j+1}, C_{i,j})$ doivent être "sensiblement" alignés sur une même droite affine.

De plus, pour les estimations de λ_j et β_j , elles sont déterminées par la méthode des moindres carrés en minimisant :

$$\delta_j = \sum_{i=0}^{n-j-1} (C_{i,j+1} - \beta_j - \lambda_j C_{i,j})^2 \quad (1)$$

Ainsi, en annulant simultanément les dérivées partielles $\frac{\partial \delta_j}{\partial \lambda_j}$ et $\frac{\partial \delta_j}{\partial \beta_j}$, nous obtenons les estimateurs de λ_j et β_j dont les expressions respectives sont les suivantes :

$$\begin{cases} \lambda_j^{LC} = \frac{\frac{1}{n-j} (\sum_{i=0}^{n-j-1} C_{i,j} C_{i,j+1}) - \bar{C}_j \bar{C}_{j+1}}{\frac{1}{n-j} (\sum_{i=0}^{n-j-1} C_{i,j}^2) - \bar{C}_j^2} \\ \beta_j^{LC} = \bar{C}_{j+1} - \lambda_j^{LC} \bar{C}_j \end{cases}$$

où, \bar{C}_j et \bar{C}_{j+1} sont respectivement les moyennes des colonnes j et $(j+1)$:

$$\begin{cases} \bar{C}_j = \frac{1}{n-j} \sum_{i=0}^{n-j-1} C_{i,j} \\ \bar{C}_{j+1} = \frac{1}{n-j} \sum_{i=0}^{n-j-1} C_{i,j+1} \end{cases}$$

Démonstration :

En dérivant **(1)** respectivement par rapport à λ_j et β_j , et en en supposant les dérivées partielles nulles, nous obtenons :

$$\begin{cases} \frac{\partial \delta_j}{\partial \lambda_j} = -2 \sum_{i=0}^{n-j-1} C_{i,j} (C_{i,j+1} - \beta_j^{LC} - \lambda_j^{LC} C_{i,j}) = 0 & \mathbf{(2)} \\ \frac{\partial \delta_j}{\partial \beta_j} = -2 \sum_{i=0}^{n-j-1} (C_{i,j+1} - \beta_j^{LC} - \lambda_j^{LC} C_{i,j}) = 0 & \mathbf{(3)} \end{cases}$$

De **(3)**, nous en déduisons :

$$\beta_j^{LC} \frac{1}{n-j} \sum_{i=0}^{n-j-1} (C_{i,j+1} - \lambda_j^{LC} C_{i,j})$$

puis,

$$\boxed{\beta_j^{LC} = \bar{C}_{j+1} - \lambda_j^{LC} \bar{C}_j} \quad \mathbf{(4)}$$

Par substitution du résultat **(4)** dans **(2)**,

$$\sum_{i=0}^{n-j-1} C_{i,j} [C_{i,j+1} - \bar{C}_{j+1} + \lambda_j^{LC} (\bar{C}_j - C_{i,j})] = 0$$

Ceci étant équivalent à :

$$\lambda_j^{LC} = \frac{\sum_{i=0}^{n-j-1} C_{i,j} C_{i,j+1} - C_{i,j} \bar{C}_{j+1}}{\sum_{i=0}^{n-j-1} C_{i,j}^2 - C_{i,j} \bar{C}_j}$$

En divisant le numérateur et le dénominateur par $(n-j)$, nous obtenons finalement que :

$$\boxed{\lambda_j^{LC} = \frac{\frac{1}{n-j} (\sum_{i=0}^{n-j-1} C_{i,j} C_{i,j+1}) - \bar{C}_j \bar{C}_{j+1}}{\frac{1}{n-j} (\sum_{i=0}^{n-j-1} C_{i,j}^2) - \bar{C}_{i,j}}}$$

■

Remarque :

En particulier, pour $j=n-1$, nous posons $\beta_{n-1}^{LC} = 0$.

Ainsi, $\lambda_{n-1}^{LC} = \frac{\bar{C}_{j+1}}{\bar{C}_j}$

2.2 Exemple d'application

j	1	2	3	4	5	6	7	8	9
λ_j^{LC}	1.2524	1.1052	1.0516	1.0295	1.0120	1.0127	1.0103	1.0028	1.0011
β_j^{LC}	999 219	292 504	147 610	63 468	33 397	- 7 568	- 24 270	402	0

TABLE 2.1 – Estimation des paramètres λ_j^{LC} et β_j^{LC} par année de développement.

Années	Années de développement									
	0	1	2	3	4	5	6	7	8	9
1988	1 376	2 588	3 123	3 437	3 605	3 685	3 725	3 740	3 750	3 755
1989	1 576	3 013	3 666	4 009	4 197	4 274	4 310	4 326	4 339	4 344
1990	1 763	3 304	3 982	4 347	4 524	4 602	4 649	4 675	4 688	4 693
1991	1 780	3 278	3 940	4 261	4 424	4 508	4 562	4 584	4 597	4 602
1992	1 843	3 417	4 030	4 329	4 506	4 612	4 660	4 684	4 698	4 703
1993	1 962	3 483	4 065	4 412	4 650	4 740	4 790	4 815	4 829	4 834
1994	2 033	3 464	4 097	4 530	4 727	4 818	4 868	4 894	4 908	4 913
1995	2 072	3 531	4 258	4 625	4 825	4 917	4 969	4 995	5 010	5 015
1996	2 211	3 728	4 413	4 788	4 993	5 08	5 141	5 169	5 184	5 190
1997	2 207	3 763	4 452	4 829	5 035	5 130	5 184	5 213	5 228	5 233

TABLE 2.2 – Tableau des paiements cumulés.

Années	Provisions LC
1988	-
1989	4 727
1990	18 589
1991	40 801
1992	90 135
1993	183 571
1994	383 555
1995	757 458
1996	1 461 701
1997	3 026 472
Total	5 967 009

TABLE 2.3 – Provisions London-Chain.

Chapitre 3

Méthode de Bornhuetter-Ferguson

3.1 Présentation du principe général

Une autre méthode classique de provisionnement est la méthode de *Bornhuetter-Ferguson*, développée en 1972. Sa particularité est d'estimer l'ultime et les provisions en y intégrant une information exogène au triangle de paiements. En effet, les provisions sous Bornhuetter-Ferguson se déterminent à partir d'une estimation des ultimes appelés *ultimes a priori* qu'on notera U_i .

Pour déterminer l'ultime de Bornhuetter-Ferguson, nous appliquons la relation suivante :

$$\forall i \in (0, \dots, n-1), C_{i,n} = C_{i,n-i+1} + (1 - Z_{n-i+1})U_i \quad (1)$$

avec,

$$\begin{cases} \forall i \in (0, \dots, n-1), Z_i = \frac{1}{\prod_{k=i}^n \lambda_k} \\ U_i, \text{ un estimateur a priori de la charge ultime} \end{cases}$$

Il existe différentes méthodes pour déterminer l'ultime a priori. Voici la méthode que nous avons décidé de retenir :

1. Nous appliquons la méthode Chain-Ladder standard au triangle des paiements cumulés $C = (C_{i,j})$ dans le but d'obtenir les *charges ultimes Chain-Ladder* $U_{CL} = (U_{CL_i})$ tel que :

$$\forall i \in (0, \dots, n-1), U_{CL_i} = C_{i,n}$$

2. Comme nous l'avions énoncé dans l'introduction, la méthode de Bornhuetter-Ferguson implique l'utilisation d'une donnée exogène. En effet, la méthode requiert un avis d'expert (ou *expert judgement*) dans le but de déterminer le loss ratio (ou ratio S/P) correspondant au quotient entre le coût des sinistres et les primes reçues par l'assureur. C'est le travail le plus délicat à effectuer du fait qu'il n'existe pas de méthode bien définie pour déterminer ces ratios.

Dans notre étude, nous nous appuyerons sur un graphique représentatif de l'évolution des ratios. En effet, à partir des informations déjà connues (triangles des paiements cumulés et primes), nous pouvons déterminer, pour chaque année d'accident et chaque année de survenance, les loss ratios correspondants. Formellement,

$$\forall i, j \in (1, \dots, n-1), \quad (S/P)_{i,j} = \frac{C_{i,j}}{P_i}$$

Remarque :

- Idéalement, en plus de cette observation graphique, nous devrions nous appuyer sur des données supplémentaires pour pouvoir expliquer cette approche graphique. A titre d'exemple, dans le cadre des données automobiles comme c'est le cas dans notre étude, nous devrions nous appuyer en plus sur les chiffres de la sécurité routière, l'évolution du montant des primes des contrats d'assurance, le montant de chaque sinistre, etc. Cependant, nous nous contenterons uniquement de l'analyse graphique.
 - Le programme R permettant d'obtenir le graphique est joint en annexe.
3. A présent, dans le but d'obtenir l'ultime a priori $U = (U_i)$ associé aux années de survenance i , nous appliquons le calcul suivant :

$$\forall i \in (0, \dots, n-1), U_i = (S/P)_{retenu} * P_i$$

4. Puis, en remarquant le raisonnement suivant :

$$\begin{aligned} Z_{n-i+1} &= \frac{1}{\prod_{k=n-i+1}^n \lambda_k} \\ &= \frac{C_{i,n-i+1}}{C_{i,n-i+1} * \prod_{k=n-i+1}^n \lambda_k} \\ &= \frac{ChargeActuelle}{UltimeChain-Ladder} \end{aligned}$$

Nous en déduisons l'ultime de Bornhuetter-Ferguson.

5. Enfin, en appliquant (1), nous obtenons les charges ultimes $(C_{i,n})$ puis nous en déduisons les provisions pour chaque année d'accident i et la provision finale.

3.2 Exemple d'application

Dans le but de rendre l'exemple le plus exhaustif possible, nous choisissons de reprendre chacune des étapes présentées dans la partie théorique.

Etape 1 :

Années	U_{CL}
1988	3 754 555
1989	4 343 687
1990	4 693 275
1991	4 600 603
1992	4 699 623
1993	4 833 181
1994	4 918 780
1995	5 039 555
1996	5 273 289
1997	5 598 621

TABLE 3.1 – Calcul des ultimes Chain-Ladder $U_{CL} = (U_{CL_i})$.

Etape 2 :

Nous obtenons le graphique ci-dessous dont la représentation est l'évolution des loss ratios.

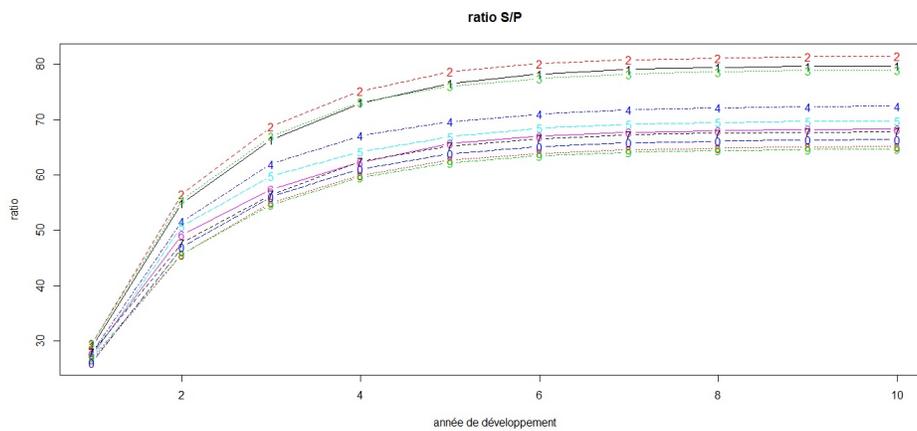


FIGURE 3.1 – Evolution des loss ratios en fonction des années de développement.

Nous observons une tendance dans l'évolution des ratios. En effet, ils croissent rapidement durant les premières années, puis cette croissance s'essouffle et les ratios stagnent sur les dernières années. Pour les trois premières années, les loss ratios tendent vers une limite avoisinant 80% puis, pour les années suivantes, cette limite est plus faible (environ 70%). En prenant compte de ces observations, un ratio S/P retenu de 72% semble être un bon compromis pour l'utilisation de la méthode de Bornhuetter-Ferguson.

Etape 3 :

Années	Ultime a Priori
1988	3 769 066
1989	4 268 420
1990	4 758 003
1991	5 083 358
1992	5 390 538
1993	5 663 555
1994	5 803 866
1995	6 191 503
1996	6 523 252
1997	6 748 734

TABLE 3.2 – Calcul des ultimes a priori $U = (U_i)$.

Etape 4 :

Années	Ultime BF	Provisions BF
1988	3 754 555	-
1989	4 343 605	4 645
1990	4 693 532	18 910
1991	4 604 689	43 017
1992	4 712 426	99 892
1993	4 864 580	214 156
1994	4 988 797	459 128
1995	5 218 273	960 573
1996	5 639 519	1 911 264
1997	6 295 378	4 088 492
Total		7 800 077

TABLE 3.3 – Calcul de l'ultime BF $C_{.,n} = (C_{i,n})$ et des provisions BF.

Chapitre 4

Combinaison des deux méthodes

Maintenant que nous avons étudié séparément les méthodes Chain-Ladder et Bornhuetter-Ferguson, intéressons-nous à la combinaison de ces deux méthodes.

4.1 Comparaison de ces deux méthodes

La méthode de Bornhuetter-Ferguson ayant été développée plus récemment que la méthode Chain-Ladder, nous pourrions supposer a priori qu'elle est plus performante. Néanmoins, ces méthodes ont chacune leurs avantages et inconvénients.

– **Avantages Chain-Ladder :**

La méthode de provisionnement Chain-Ladder demeure la méthode la plus connue et la plus appliquée de toutes les méthodes déterministes utilisant le développement d'un triangle. Elle doit son succès notamment à sa facilité de mise en œuvre et peut être employée sur différents types de données (triangle de paiements cumulés, charges de sinistres, etc.).

– **Inconvénients Chain-Ladder :**

En revanche, celle-ci présente quelques inconvénients. En effet, nous ne pouvons pas appliquer la méthode Chain-Ladder sur n'importe quel triangle de paiements, il faut que celui-ci vérifie certaines hypothèses. Par exemple, il existe des triangles dits instables où le nombre de sinistres varie brutalement d'une année à une autre pour une année d'accident i fixée.

De plus, le calcul des provisions est fortement lié au dernier montant de sinistre connu. Dès lors, il y a un risque de biais sur le montant des provisions calculées.

Enfin, pour les années récentes, l'incertitude est très importante : le coefficient multiplicatif de la dernière année est le produit de $(n-1)$ estimations de

coefficients de proportionnalité. Ainsi, l'incertitude est d'autant plus grande que les branches sont longues.

– **Avantages Bornhuetter-Ferguson :**

Un des avantages de la méthode Bornhuetter-Ferguson est qu'en plus d'utiliser le triangle de paiements, elle incorpore une information exogène appelée avis d'expert pour estimer la charge ultime des sinistres d'une année d'incidence donnée. En effet, elle se base sur une combinaison des méthodes du Loss Ratio (méthode selon laquelle nous supposons connaître le rapport du coût des sinistres sur la prime ultime) et de Chain-Ladder.

Un autre avantage de ne pas dépendre du dernier montant observé est qu'il n'est pas nécessaire d'avoir un triangle contenant beaucoup d'années pour obtenir de bonnes provisions.

– **Inconvénients Bornhuetter-Ferguson :**

La détermination du Loss ratio est fortement subjective et dépend beaucoup de l'expert. Si ce Loss ratio est mal estimé, nous courrons ainsi le risque d'estimer une charge ultime des sinistres fortement biaisée étant donné qu'elle dépend fortement de ce ratio. La méthode ne garantit donc pas une stabilité du résultat année après année.

Conclusion :

Compte-tenu de ces remarques, il paraît logique que certaines compagnies d'assurance utilisent une combinaison de ces deux méthodes de provisionnement pour utiliser leur avantages respectifs et minimiser leurs défauts.

Ainsi, pour un triangle avec un nombre d'années de paiements non négligeable, nous retiendrons que la meilleure approche est de garder les premières provisions fournies par la méthode Chain-Ladder puis de conserver les dernières provisions déduites de l'application de la méthode Bornhuetter-Ferguson. En effet, nous avons vu que la méthode Chain-Ladder est efficace pour estimer les provisions des survenances anciennes mais donne des estimations approximatives pour les années récentes en raison du peu d'informations disponibles. A contrario, les provisions récentes sous Bornhuetter-Ferguson seront davantage calculées à partir de l'information exogène au triangle. Les provisions alors calculées seront meilleures que les provisions approximatives fournies par Chain-Ladder, du fait de la prise en compte des ultimes a priori.

4.2 Raisonnement théorique

Comme nous venons de le voir, chacune des deux méthodes est à appliquer dans des situations distinctes, les durées d'application étant déterminées par l'utilisateur. Par exemple, ce dernier pourra décider d'appliquer la méthode Chain-Ladder standard sur les m premières années puis la méthode Bornhuetter-Ferguson, sur les $(n-m+1)$ dernières années.

Il semble tout à fait légitime de se demander, à présent, comment nous devrions procéder. N'oublions pas que le problème de l'assureur est de déterminer ses provisions.

Ainsi, pour résoudre ce problème à partir de la combinaison des deux méthodes, il devra appliquer chacune d'entre elles en totalité puis sélectionner les provisions qui l'intéressent comme le présente le tableau suivant :

Années	Provisions CL	Provisions BF	Provisions retenues
0	-	-	-
1	P_{CL_1}	P_{BF_1}	P_{CL_1}
m-1	$P_{CL_{m-1}}$	$P_{BF_{m-1}}$	$P_{CL_{m-1}}$
m	P_{CL_m}	P_{BF_m}	P_{CL_m}
m+1	$P_{CL_{m+1}}$	$P_{BF_{m+1}}$	$P_{BF_{m+1}}$
n	P_{CL_n}	P_{BF_n}	P_{BF_n}

TABLE 4.1 – Tableau des provisions retenues.

4.3 Application

Dans notre exemple, nous choisissons de raisonner de la façon suivante :

- Nous appliquons la méthode Chain-Ladder standard sur la période de survivance 1988-1994 ;
- Puis, la méthode Bornhuetter-Ferguson aux années 1995-1997.

Années	Provisions CL	Provisions BF	Provisions retenues
1988	-	-	-
1989	4 727	4 645	4 727
1990	18 653	18 910	18 653
1991	38 931	43 017	38 931
1992	87 089	99 892	87 089
1993	182 757	214 156	182 757
1994	389 111	459 128	389 111
1995	781 855	960 573	960 573
1996	1 545 034	1 911 264	1 911 264
1997	3 391 735	4 088 492	4 088 492
Total	6 439 892	7 800 077	7 681 597

TABLE 4.2 – Tableau récapitulant les provisions respectives Chain-Ladder et Bornhuetter-Ferguson ainsi que les provisions retenues et la provision totale.

Troisième partie

Méthodes stochastiques

Objectifs :

- Estimer les provisions ;
- Connaitre les propriétés stochastiques de ces provisions, notamment l'erreur de prédiction dans le but d'évaluer le degré de certitude ou d'incertitude des prévisions.

Remarque : En ce qui concerne les raisonnements qui vont suivre, nous reprenons les notations de T. Mack introduites dans l'article *Distribution-free calculation of the standard error of chain ladder reserve estimates* (1993) afin de faciliter la lecture de ce document. **Ainsi, nous considérons des années d'origine et délais de règlement des triangles variant de 1 à n** (contre 0 à n précédemment).

Chapitre 5

Méthode de Mack

5.1 Hypothèses du modèle de Mack

La méthode de Mack est la première méthode faisant intervenir la notion d'incertitude dans la méthode déterministe Chain-Ladder. En effet, elle permet de mesurer l'incertitude associée à la prédiction du montant des provisions que doit faire l'assureur.

Cette méthode s'appuie sur trois hypothèses :

(H1) : Les années de survenance des sinistres sont indépendantes les unes des autres, c'est-à-dire, $C_{i,j}$ et $C_{k,j}$ sont indépendants si $i \neq k$.

(H2) :
 $\forall i = 1, \dots, n, \forall j = 1, \dots, n - 1,$

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

Cette seconde hypothèse suppose alors que le passage d'une année de développement à l'autre est décrit en termes d'espérance.

Rappelons que λ_j correspond au facteur de développement associé à l'année de développement j dont l'estimateur $\hat{\lambda}_j$ est déterminé par la relation suivante :

$$\forall j = 1, \dots, n - 1, \hat{\lambda}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}$$

Il est important de noter que les facteurs $\hat{\lambda}_j$ sont des estimateurs de λ_j sans biais.

Démonstration :

Soit, pour $j=1, \dots, n$, $T_j = \{C_{i,h} : h \leq j, i + j \leq n + 1\}$, la partie du triangle supérieur limitée au délai de règlement j , nous avons :

$$\mathbb{E}[C_{i,j+1}|T_j] = \mathbb{E}[C_{i,j+1}|C_{i,1}, \dots, C_{i,j}] = \lambda_j C_{i,j}$$

puis,

$$\mathbb{E}\left[\frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}} \middle| T_j\right] = \frac{\sum_{i=1}^{n-j} \mathbb{E}[C_{i,j+1}|T_j]}{\sum_{i=1}^{n-j} C_{i,j}} = \frac{\sum_{i=1}^{n-j} \lambda_j C_{i,j}}{\sum_{i=1}^{n-j} C_{i,j}} = \lambda_j$$

puis, par conditionnement :

$$\mathbb{E}[\hat{\lambda}_j] = \mathbb{E}[\mathbb{E}[\hat{\lambda}_j|T_j]] = \mathbb{E}[\lambda_j] = \lambda_j$$

Cette dernière égalité démontrant alors l'assertion selon laquelle $\hat{\lambda}_j$ est un estimateur sans biais de λ_j .

■

(H3) :

$\forall i = 1, \dots, n, \forall j = 1, \dots, n - 1$, il existe un paramètre σ_j tel que :

$$\text{Var}(C_{i,j+1}|C_{i,1}, \dots, C_{i,j}) = \sigma_j^2 C_{i,j}$$

$\hat{C}_{i,n}$ est un estimateur de $C_{i,n}$ avec :

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

Dans son raisonnement, Mack propose également un estimateur sans biais de σ_j^2 que nous noterons ici S_j^2 :

$$\begin{cases} \forall 1 \leq j \leq n - 2, S_j^2 = \frac{1}{n-j-1} \sum_{i=1}^{n-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \lambda_j\right)^2 \\ S_{n-1}^2 = \min\left\{\frac{S_{n-2}^4}{S_{n-3}^2}, S_{n-2}^2, S_{n-3}^2\right\} \end{cases}$$

Démonstration :

$\forall 1 \leq j \leq n - 2,$

$$\begin{aligned}
\mathbb{E}[S_j^2|T_j] &= \mathbb{E}\left[\frac{1}{n-j-1} \sum_{i=1}^{n-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \lambda_j\right)^2 | T_j\right] \\
&= \frac{1}{n-j-1} \left(\sum_{i=1}^{n-j} \frac{\mathbb{E}[C_{i,j+1}^2 | T_j]}{C_{i,j}} - 2\mathbb{E}\left[\sum_{i=1}^{n-j} C_{i,j+1} \hat{\lambda}_j | T_j\right] + \mathbb{E}\left[\sum_{i=1}^{n-j} C_{i,j} \hat{\lambda}_j^2 | T_j\right] \right) \\
&= \frac{1}{n-j-1} \left(\sum_{i=1}^{n-j} \frac{\mathbb{E}[C_{i,j+1}^2 | T_j]}{C_{i,j}} - 2\mathbb{E}\left[\sum_{i=1}^{n-j} C_{i,j} \hat{\lambda}_j^2 | T_j\right] + \mathbb{E}\left[\sum_{i=1}^{n-j} C_{i,j} \hat{\lambda}_j^2 | T_j\right] \right) \\
&= \frac{1}{n-j-1} \left(\sum_{i=1}^{n-j} \frac{\mathbb{E}[C_{i,j+1}^2 | T_j]}{C_{i,j}} - \mathbb{E}\left[\sum_{i=1}^{n-j} C_{i,j} \hat{\lambda}_j^2 | T_j\right] \right)
\end{aligned}$$

Or,

$$\begin{aligned}
\text{Var}[\hat{\lambda}_j | T_j] &= \text{Var}\left[\frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}} | T_j\right] \\
&= \frac{\sum_{i=1}^{n-j} \text{Var}[C_{i,j+1} | T_j]}{(\sum_{i=1}^{n-j} C_{i,j})^2} \\
&= \frac{\sum_{i=1}^{n-j} C_{i,j} \sigma_j^2}{(\sum_{i=1}^{n-j} C_{i,j})^2} \\
&= \frac{\sigma_j^2}{\sum_{i=1}^{n-j} C_{i,j}}
\end{aligned}$$

Puis,

$$\begin{aligned}
\mathbb{E}[S_j^2|T_j] &= \frac{1}{n-j-1} \left(\sum_{i=1}^{n-j} \frac{C_{i,j} \sigma_j^2 + C_{i,j}^2 \lambda_j^2}{C_{i,j}} - \sum_{i=1}^{n-j} C_{i,j} \left(\frac{\sigma_j^2}{\sum_{i=1}^{n-j} C_{i,j}} + \lambda_j^2\right) \right) \\
&= \sigma_j^2
\end{aligned}$$

Enfin, en composant par l'espérance, nous obtenons alors :

$$\mathbb{E}[S_j^2] = \mathbb{E}[\mathbb{E}[S_j^2|T_j]] = \sigma_j^2$$

■

Remarque : Le modèle de Mack considère implicitement que les paiements cumulés suivent une distribution normale avec une moyenne et une variance qui sont décrites dans les hypothèses **(H2)** et **(H3)**.

5.2 Mesure de l'incertitude de l'estimateur

Sous ces hypothèses, Mack obtient une formule fermée pour la variance de la charge ultime. Nous calculons l'erreur quadratique moyenne (MSEP - Mean Square of Error of Prediction) en conditionnant par rapport aux données passées :

$$MSEP(\hat{C}_{i,n}) = \mathbb{E}[(\hat{C}_{i,n} - C_{i,n})^2 | C_{i,j} : i + j \leq n + 1]$$

Nous en déduisons :

$$MSEP(\hat{C}_{i,n}) = \text{Var}[C_{i,n} | C_{i,j} : i + j \leq n + 1] + (\mathbb{E}[C_{i,n} | C_{i,j} : i + j \leq n + 1] - \hat{C}_{i,n})^2$$

De plus, en posant $\hat{R}_i = \hat{C}_{i,n} - C_{i,n-i+1}$, la provision étudiée, $\hat{R}_i - R_i = \hat{C}_{i,n} - C_{i,n}$ conduit à $MSEP(\hat{R}_i) = MSEP(\hat{C}_{i,n})$ et à l'erreur standard $SEP(\hat{R}_i) = \sqrt{MSEP(\hat{R}_i)}$

Puis, avec $\hat{R} = \sum_{i=2}^n \hat{R}_i$, des estimateurs de $MSEP(\hat{R}_i)$ et $MSEP(\hat{R})$ sont respectivement :

$$\left\{ \begin{array}{l} M\hat{S}EP(\hat{R}_i) = \hat{C}_{i,n}^2 \sum_{j=n-i+1}^{n-1} \frac{S_j^2}{\hat{\lambda}_j^2} \left(\frac{1}{\hat{C}_{i,j}} + \frac{1}{\sum_{k=1}^{n-j} C_{k,j}} \right), \text{ pour } i = 2, \dots, n \\ M\hat{S}EP(\hat{R}) = \sum_{i=2}^n \{ M\hat{S}EP(\hat{R}_i) + \hat{C}_{i,n} (\sum_{k=i+1}^n \hat{C}_{k,n}) (\sum_{j=n-i+1}^{n-1} \frac{2S_j^2}{\hat{\lambda}_j^2 \sum_{l=1}^{n-j} C_{l,j}}) \} \end{array} \right.$$

5.3 Exemple d'application

j	1	2	3	4	5	6	7	8	9
λ_j	1,794	1,195	1,090	1,045	1,020	1,010	1,004	1,003	1,001
S_j^2	14 202	966	455	156	31	10	3	0.0004	6.10^{-8}

TABLE 5.1 – Estimation des paramètres $(\lambda_j)_{(j=1,\dots,n-1)}$ et $(\sigma_j)_{(j=1,\dots,n-1)}$.

i	1989	1990	1991	1992	1993	1994	1995	1996	1997
\hat{R}_i	4727	18653	38931	87089	182757	389111	781855	1545034	3391735
$M\hat{S}EP(\hat{R}_i)$	0.5535	3150	$1.91 \cdot 10^7$	$7.97 \cdot 10^7$	$2.64 \cdot 10^8$	$1.18 \cdot 10^9$	$3.86 \cdot 10^9$	$9.83 \cdot 10^9$	$8.16 \cdot 10^{10}$
$S\hat{E}P(\hat{R}_i)$	0.7440	56	4376	8927	16262	34287	62107	99142	285577
$\frac{S\hat{E}P(\hat{R}_i)}{\hat{R}_i}$	0,02%	0.3%	11,2%	10,2%	8,9%	8,9%	7,9%	6,4%	8,4%

TABLE 5.2 – Mesures de risque conditionnel de prédiction.

avec,

$$\begin{cases} \hat{R} = 6\,439\,892 \\ M\hat{S}EP(\hat{R}) = 1,04 \cdot 10^{11} \end{cases}$$

Nous en déduisons alors l'erreur de prédiction suivante :

$$\text{Erreur de Prédiction} = \sqrt{M\hat{S}EP\hat{R}} = 322\,527$$

Chapitre 6

Méthode du Bootstrap - Variante Chain-Ladder

La méthode du Bootstrap est une méthode relativement récente consistant à fabriquer de l'information et à fournir des réponses là où les autres méthodes ne sont pas applicables (manque d'information, calculs impossibles...)

Le principe général de la méthode du Bootstrap est le rééchantillonnage par placement. Dans notre étude, nous choisissons d'appliquer la méthode du Bootstrap dans le cadre des hypothèses de Mack.

L'erreur de prédiction peut s'interpréter comme la combinaison de deux erreurs sous-jacentes que sont respectivement l'erreur de processus et l'erreur de simulation.

– *L'erreur d'estimation* est introduite lors du ré-échantillonnage par remplacement des résidus de Pearson.

Plus généralement, les résidus peuvent se traduire mathématiquement de la façon suivante :

$$r_{i,j} = \frac{f_{i,j} - \mathbb{E}[f_{i,j}]}{\sqrt{\text{Var}[f_{i,j}]}}$$

De plus, des hypothèses **(H2)** et **(H3)** de Mack et de l'égalité $f_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$ où $(f_{i,j})$ sont les coordonnées du D-triangle, nous en déduisons pour tout $i=1, \dots, n$:

$$\begin{cases} \mathbb{E}[f_{i,j} | C_{i,1}, \dots, C_{i,j}] = \lambda_j \\ \text{Var}[f_{i,j} | C_{i,1}, \dots, C_{i,j}] = \frac{\sigma_j^2}{C_{i,j}} \end{cases}$$

Par substitution, nous obtenons alors les résidus de Pearson propres au modèle de Mack :

$$r_{i,j} = r_{PS}(f_{i,j}, \lambda_j, w_{i,j}, \sigma_j) = \frac{\sqrt{w_{i,j}}(f_{i,j} - \lambda_j)}{\sigma_j}, \text{ avec } w_{i,j} = C_{i,j-1}$$

- *L'erreur de processus*, quant à elle, est introduite lors d'un tirage aléatoire des éléments du triangle inférieur projeté ($C_{i,j}^*$) obtenu après inversion des résidus rééchantillonnés. Dans notre étude, nous effectuons ce tirage aléatoire par le biais d'une loi normale dont l'espérance et la variance sont déduites des hypothèses **(H2)** et **(H3)** de Mack :

$$\tilde{C}_{i,j}|C_{i,j}^* = N(\tilde{\lambda}_j C_{i,j}^*, \sigma_j^2 C_{i,j}^*)$$

avec $\tilde{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} w_{i,j} f_{i,j}^B}{\sum_{i=1}^{n-j+1} w_{i,j}}$ dont les nouveaux facteurs de développement ($f_{i,j}^B$) sont issus de l'inversion des résidus de Pearson remplacés.

Pour plus de clarté, résumons à présent les différentes étapes :

1. Calcul du D-triangle $(f_{i,j})$, des (λ_j) et des (σ_j^2) ;
2. Calcul du triangle des résidus supérieurs :

$$r_{i,j} = \frac{\sqrt{w_{i,j}}(f_{i,j} - \lambda_j)}{\sigma_j}$$

Boucle (avec N, le nombre de simulations) :

3. Rééchantillonnage par remplacement des résidus que nous notons alors :

$$r_{i,j}^B = \frac{\sqrt{w_{i,j}}(f_{i,j}^B - \lambda_j)}{\sigma_j}$$

4. Calcul des link ratios simulés :

$$\tilde{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} w_{i,j} f_{i,j}^B}{\sum_{i=1}^{n-j+1} w_{i,j}}, \text{ avec } f_{i,j}^B = r_{i,j}^B \frac{\sigma_j}{\sqrt{w_{i,j}}} + \lambda_j$$

5. Calcul du triangle inférieur des paiements cumulés $(C_{i,j}^*)$ à partir de la méthode Chain-Ladder ;
6. Simulation par une loi normale des données $(\tilde{C}_{i,j})$ tel que :

$$\tilde{C}_{i,j+1} | C_{i,j}^* = N(\tilde{\lambda}_j C_{i,j}^*, \sigma_j^2 C_{i,j}^*)$$

Enfin, nous obtenons N estimations de la provision totale \hat{R} dont l'écart-type correspond à l'erreur de processus.

Chapitre 7

Méthode du Bootstrap - Combinaison Chain- Ladder/Bornhuetter-Ferguson

Dans le but d'apporter une première réponse au problème posé, la méthode du Bootstrap appliquée à la combinaison Chain-Ladder/Bornhuetter-Ferguson semble être appropriée.

Nous aurons ici besoin de revenir sur les acquis de la combinaison de ces méthodes étudiée dans la première partie. En effet, chaque méthode doit être appliquée dans un contexte bien défini : la méthode Chain-Ladder s'applique parfaitement aux sinistres anciens et la méthode Bornhuetter-Ferguson davantage adaptée aux années récentes.

Ainsi, une fois l'étape 6 (du raisonnement du chapitre précédent) atteinte, calculons à partir des données simulées $(\tilde{C}_{i,j})$:

- Les provisions mathématiques Chain-Ladder aux k années les plus anciennes ;
- Les provisions mathématiques Bornhuetter-Ferguson aux $(n-k)$ années les plus récentes.

Remarque : Dans le second cas, il nous suffira d'appliquer la méthode Bornhuetter-Ferguson déterministe pour parvenir au calcul de la provision.

Une fois ces provisions Chain-Ladder/Bornhuetter-Ferguson et les N simulations établies, l'erreur de prédiction sera finalement déduite de l'écart type de la provision totale.

Ci-dessous, un schéma du raisonnement :

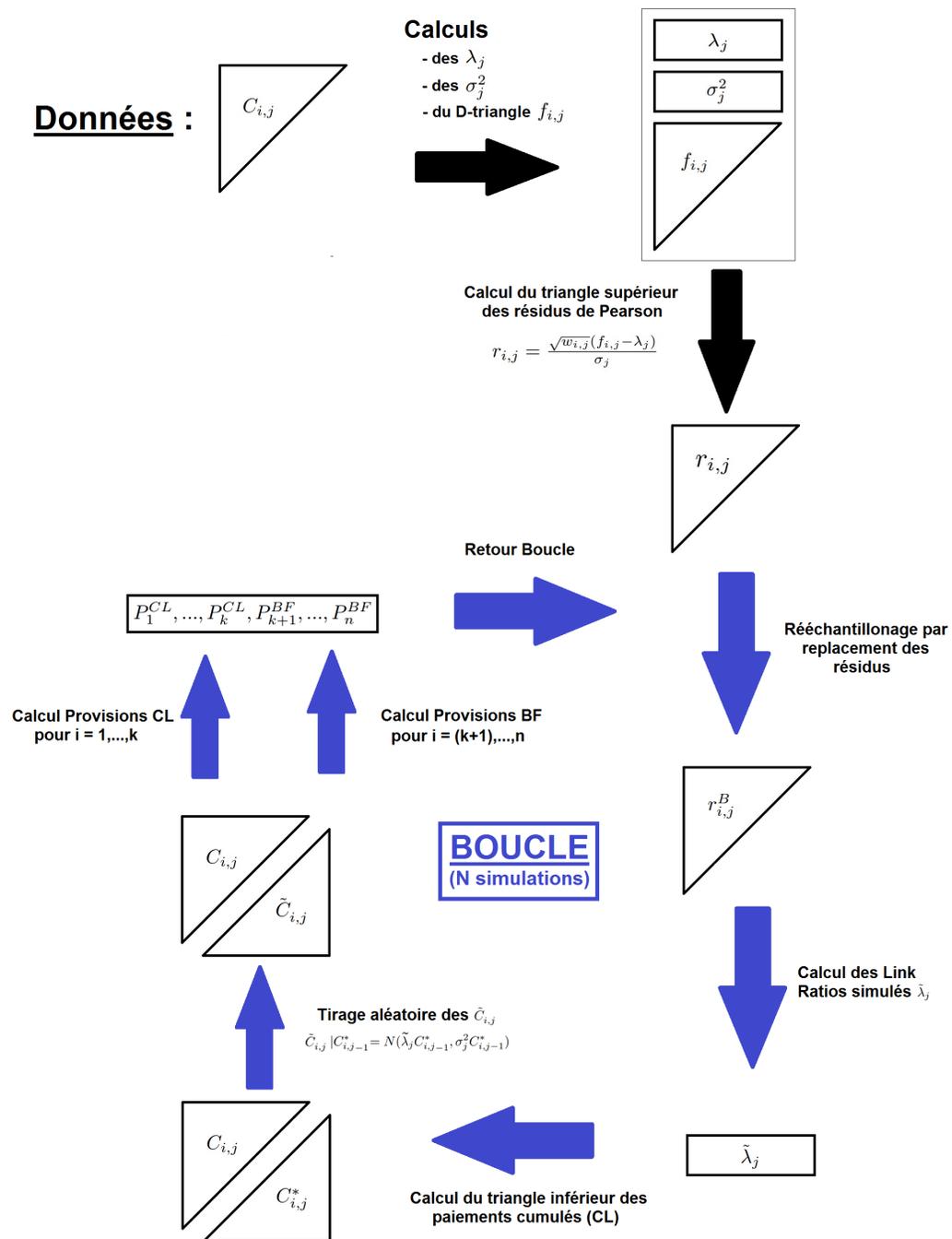


FIGURE 7.1 – Raisonnement théorique de la méthode du Bootstrap combinée CL/BF.

Chapitre 8

Méthode Bornhuetter-Ferguson Stochastique

Il existe différents modèles de la méthode de Bornhuetter-Ferguson Stochastique pour déterminer l'erreur de prédiction. Parmi elles, figure une méthode proposée par Mack, introduite dans le document *The Prediction Error of Bornhuetter-Ferguson* (2008), et une méthode développée par Alai, Merz et Wüthrich présentée dans l'article *Mean Square Error of Prediction in the Bornhuetter-Ferguson Claims Reserving Methods* (2009). Nous choisirons ici de développer la première méthode.

8.1 Introduction des notations

Dans son modèle, Mack reprend la formule des réserves introduite dans la méthode Bornhuetter-Ferguson déterministe étudiée préalablement que nous rappelons ici :

$$\forall i \in (0, \dots, n-1), \hat{R}_i^{BF} = (1 - Z_{n-i+1})U_i$$

avec,

$$\begin{cases} \forall i \in (0, \dots, n-1), Z_i = \frac{1}{\prod_{k=i}^n \lambda_k}, \text{ la cadence des règlements cumulés} \\ U_i, \text{ un estimateur a priori de la charge ultime} \end{cases}$$

Les provisions calculées à partir de cette méthode considèrent le développement des sinistres jusqu'à l'ultime ; l'estimation des paramètres dans le modèle de Mack doit donc prendre en compte la queue de développement. C'est pourquoi, les années de développement j appartiennent à l'intervalle $[1, n+1]$ dans notre modèle.

Comme nous l'avons étudié dans le cadre déterministe, l'avantage de la méthode Bornhuetter-Ferguson résulte de l'intégration de données exogènes. En effet, la méthode requiert l'intégration d'une estimation de la charge ultime a priori (U_i) et de l'incertitude associée à cet estimateur. Ainsi, l'erreur de prédiction obtenue dépendra alors en partie de l'intégration de ce jugement d'expert.

Dans sa méthode, Mack introduit le modèle suivant en notant $S_{i,k}$, $i = 1, \dots, n$, $k = 1, \dots, n+1$, les incréments de paiements :

- Les incréments $S_{i,k}$ sont indépendants ;
- x_i, y_k sont des paramètres inconnus tels que : $\mathbb{E}[S_{i,k}] = x_i y_k$
avec ,

$$\begin{cases} \sum_{k=1}^{n+1} y_k = 1 \text{ avec } k = i, \dots, n \text{ et où } y_i \text{ s'interprète comme la cadence de règlements non cumulés} \\ x_i = \mathbb{E}[U_i] \end{cases}$$

- De même, Mack introduit le vecteur s_k^2 de la manière suivante :

$$\text{Var}[S_{i,k}] = x_i s_k^2 ;$$

Pour la suite du raisonnement, nous estimerons les x_i par \hat{U}_i , ce qui nous permettra de définir les estimateurs suivants :

$$\begin{cases} \hat{y}_k = \frac{\sum_{i=1}^{n+1-k} S_{i,k}}{\sum_{i=1}^{n+1-k} x_i}, 1 \leq k \leq n; \\ \hat{s}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} \frac{(S_{i,k} - x_i \hat{y}_k)^2}{x_i}, 1 \leq k \leq n-1; \end{cases}$$

Par la suite, l'actuaire devra déterminer \hat{y}_i^* et \hat{s}_i^{2*} pour $1 \leq n+1$, à partir des méthodes de régression et de lissage des données. En ce qui concerne les estimations des cadences des règlements non cumulés, nous vérifierons en plus $\sum_{i=1}^{n+1} \hat{y}_i^* = 1$.

L'ajustement et l'extrapolation de ces paramètres nous permettent alors d'obtenir une estimation de la provision Bornhuetter-Ferguson :

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

8.2 Calcul de l'erreur de prédiction

Dans ce cadre, nous définissons la MSEP (ou *Mean Squared Error of Prediction*) de la provision Bornhuetter-Ferguson \hat{R}_i^* , pour chaque année de survenance i , par l'espérance conditionnelle :

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

A partir de l'hypothèse d'indépendance des $S_{i,k}$, nous obtenons :

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

Pour déterminer l'erreur de prédiction, il suffit donc de faire la somme de l'erreur d'estimation $\text{Var}[\hat{R}_i^*]$ et de l'erreur de processus $\text{Var}[R_i^*]$. L'erreur de processus est estimée simplement par :

$$\hat{\text{V}}\text{ar}[R_i] = \hat{U}_i(\hat{s}_{n+2-i}^{2*} + \dots + \hat{s}_{n+1}^{2*})$$

En revanche, déterminer l'erreur d'estimation est un peu plus compliqué. En effet, pour commencer nous devons d'abord calculer l'incertitude des ultimes a priori \hat{U}_i :

$$\hat{M}SE(\hat{U}_i) = \frac{P_i}{n-1} \sum_{j=1}^n P_j \left(\frac{\hat{U}_j}{P_j} - \frac{\sum_{k=1}^n \hat{U}_k}{\sum_{k=1}^n P_k} \right)^2$$

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

Ensuite, il nous faudra estimer l'erreur sur les cadences de règlements incrémentaux \hat{y}_k :

$$\hat{M}SE(\hat{y}_k^*) = \frac{\hat{s}_k^{2*}}{\sum_{j+1}^{n+1-k} \hat{U}_j}$$

En ce qui concerne le dernier facteur \hat{y}_{n+1} , l'incertitude sera déterminée de manière exogène ; il est fréquent de poser :

$$\hat{M}SE(\hat{y}_{n+1}^*) = \frac{1}{2} \hat{y}_{n+1}^*$$

Il en découlera alors l'erreur d'estimation des cadences de règlements cumulés \hat{z}_k :

$$\begin{cases} \hat{M}SE(\hat{z}_k^*) = \min(\hat{M}SE(\hat{y}_1^*) + \dots + \hat{M}SE(\hat{y}_k^*), \hat{M}SE(\hat{y}_{k+1}^*) + \dots + \hat{M}SE(\hat{y}_{n+1}^*)), \\ \hat{M}SE(\hat{z}_{n+1}^*) = 0. \end{cases}$$

Après ces calculs intermédiaires, l'erreur de prédiction de la provision Bornhuetter-Ferguson \hat{R}_i pour l'année de survenance i , $i = 1, \dots, n$, s'écrira de la manière suivante :

$$\begin{aligned}
M\hat{S}EP(\hat{R}_i^*) &= \underbrace{\hat{U}_i(\hat{s}_{n+2-i}^{2*} + \dots + \hat{s}_{n+1}^{2*})}_{\text{Erreur de processus}} \\
&\quad + \underbrace{(\hat{U}_i^2 + M\hat{S}E(\hat{U}_i))M\hat{S}E(\hat{z}_{n+1-i}) + M\hat{S}E(\hat{U}_i)(1 - \hat{z}_{n+1-i}^*)}_{\text{Erreur d'estimation}}
\end{aligned}$$

L'erreur de prédiction de la provision totale \hat{R}^* sera un peu plus complexe que la simple somme des \hat{R}_i^* , du fait de leur corrélation, et sera calculée ainsi :

$$\begin{aligned}
M\hat{S}EP(\hat{R}^*) &= \hat{V}\text{ar}[R] + \hat{V}\text{ar}[R^*] \\
&= \underbrace{\sum_{i=1}^n \hat{U}_i(\hat{s}_{n+2-i}^{2*} + \dots + \hat{s}_{n+1}^{2*})}_{\text{Erreur de processus } \hat{V}\text{ar}[R]} \\
&\quad + \underbrace{\sum_{i=1}^n (\hat{U}_i^2 + M\hat{S}E(\hat{U}_i))M\hat{S}E(\hat{z}_{n+1-i}) + M\hat{S}E(\hat{U}_i)(1 - \hat{z}_{n+1-i}^*) + 2 \sum_{i < j} \hat{C}\text{ov}(\hat{R}_i^*, \hat{R}_j^*)}_{\text{Erreur d'estimation } \hat{V}\text{ar}[R^*]}
\end{aligned}$$

La covariance étant donnée par la formule suivante :

$$\begin{aligned}
\hat{C}\text{ov}(\hat{R}_i^*, \hat{R}_j^*) &= \hat{\rho}_{i,j}^U \sqrt{M\hat{S}E(\hat{U}_i)} \sqrt{M\hat{S}E(\hat{U}_j)(1 - \hat{z}_{n+1-i}^*)(1 - \hat{z}_{n+1-j}^*)} \\
&\quad + \hat{\rho}_{i,j}^z \sqrt{M\hat{S}E(\hat{z}_{n+1-i}^*)} \sqrt{M\hat{S}E(\hat{z}_{n+1-j}^*)} \hat{U}_i \hat{U}_j
\end{aligned}$$

et les coefficients de corrélation estimés comme il suit :

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

et

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

Pour déterminer l'incertitude liée aux provisions de Bornhuetter-Ferguson, il nous faudra alors prendre la racine carrée du $M\hat{S}EP(\hat{R}^*)$. L'avantage de ce modèle réside dans la grande liberté laissée à l'actuaire pour déterminer les paramètres. Il pourra ainsi les ajuster selon la branche considérée.

8.3 Application numérique

La première étape est donc de déterminer \hat{y}_{11}^* à partir des \hat{y}_j^* calculés grâce à la formule vue précédemment. Nous avons donc ajusté une régression log-linéaire pour expliquer les cadences de règlement incrémentales \hat{y}_j^* en fonction des années de développement, et obtenu des résultats très satisfaisants, avec un R_{ajust}^2 de 99%. De cette manière, nous avons obtenu par extrapolation une valeur pour \hat{y}_{11}^* .

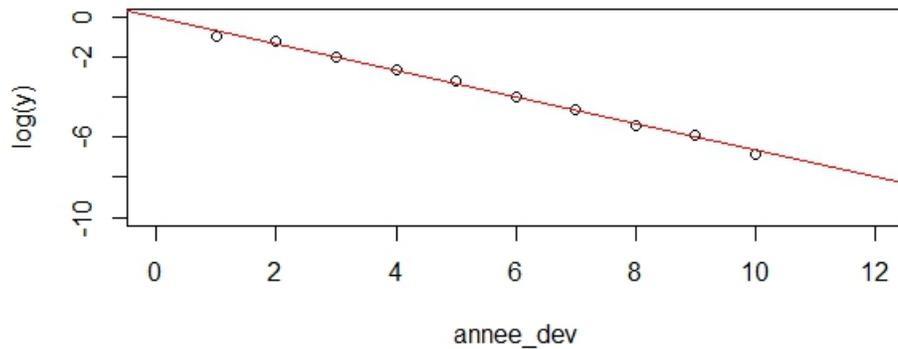


FIGURE 8.1 – Représentation des logarithmes des cadences \hat{y}_j^* en fonction des années de développement.

De la même manière, mais en utilisant la racine carrée plutôt que le logarithme pour cause de meilleurs résultats, nous avons ajusté une régression sur les paramètres \hat{s}_j^{2*} en fonction des valeurs absolues $|\hat{y}_j^*|$, pour un R_{ajust}^2 de 89%, et ainsi extrapoler \hat{s}_{10}^{2*} et \hat{s}_{11}^{2*} .

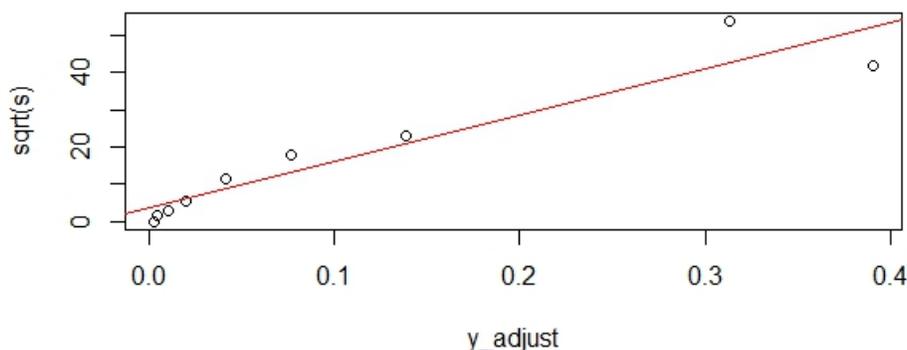


FIGURE 8.2 – Représentation de la racine des \hat{s}_j^{2*} en fonction des valeurs absolues $|\hat{y}_j^*|$.

Pour les paramètres \hat{y}_j^* et \hat{s}_j^{2*} , nous obtenons alors les valeurs suivantes :

j	1	2	3	4	5	6	7	8	9	10	11
\hat{y}_j^*	0.3906	0.3130	0.1388	0.0768	0.0418	0.0195	0.0102	0.0045	0.0029	0.0011	0.0007
\hat{s}_j^{2*}	2747.5	1827.4	442.6	177.3	80.1	38.1	25.1	18.6	16.9	15.1	14.7

TABLE 8.1 – Paramètres estimés dans la méthode de Mack

Nous avons également dû porter un jugement d'expert sur l'incertitude de la cadence de règlement incrémentale \hat{y}_{11}^* . En observant que les cadences \hat{y}_{10}^* et \hat{y}_{11}^* étaient très proches, nous avons établi l'incertitude sur la dernière cadence de règlement de sorte qu'elle soit égale à $M\tilde{S}E(\hat{y}_{10}^*)$.

Impact du coefficient de corrélation sur l'erreur de prédiction de la provision

Dans la partie théorique vue précédemment, nous avons fait l'hypothèse que les \hat{U}_i n'étaient pas corrélés pour le calcul de l'erreur associée. Or dans la pratique, il est très rare que cette hypothèse soit vérifiée et les \hat{U}_i sont très souvent positivement corrélés. Nous avons alors distingué deux cas, un premier où le coefficient de corrélation est constant et un second où il est décroissant.

Nous avons donc commencé par étudier le cas d'un coefficient de corrélation constant. Nous avons alors :

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

L'erreur d'estimation des ultimes a priori s'écrit alors :

$$M\hat{S}E(\hat{U}_i) = \frac{P_i}{n-\sqrt{n}} \sum_{j=1}^n P_j \left(\frac{\hat{U}_j}{P_j} - \frac{\sum_{k=1}^n \hat{U}_k}{\sum_{k=1}^n P_k} \right)^2, \text{ pour } i = 1, \dots, n$$

Dès lors, nous avons obtenu les résultats suivants :

	PSAP estimée	Erreur de prédiction
1988	2 571	10 584
1989	7 763	16 783
1990	22 109	21 104
1991	42 616	23 636
1992	91 619	27 804
1993	189 319	34 887
1994	401 522	50 490
1995	816 642	83 565
1996	1 650 152	148 122
1997	3 742 933	299 068
TOTAL	6 967 247	425 483

TABLE 8.2 – PSAP et incertitude à l'ultime estimées.

Nous nous sommes ensuite intéressés au cas d'un coefficient de corrélation décroissant. Nous avons alors :

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

L'erreur d'estimation des ultimes a priori s'écrit alors :

$$M\hat{S}E(\hat{U}_i) = \frac{P_i}{n-\sqrt{2n}} \sum_{j=1}^n P_j \left(\frac{\hat{U}_j}{P_j} - \frac{\sum_{k=1}^n \hat{U}_k}{\sum_{k=1}^n P_k} \right)^2, \text{ pour } i = 1, \dots, n$$

De ces formules, nous avons obtenu les résultats récapitulés dans le tableau suivant :

	PSAP estimée	Erreur de prédiction
1988	2 571	10 589
1989	7 763	16 793
1990	22 109	21 129
1991	42 616	23 706
1992	91 619	28 058
1993	189 319	35 713
1994	401 522	52 953
1995	816 642	89 533
1996	1 650 152	160 586
1997	3 742 933	325 986
TOTAL	6 967 247	482 428

TABLE 8.3 – PSAP et incertitude à l’ultime estimées.

Remarque : Sous des hypothèses portant sur la distribution des provisions, il est possible de construire des intervalles de prédiction pour R . Nous pouvons cependant noter le caractère arbitraire d’un tel choix, à l’opposé du caractère non paramétrique de cette approche.

Prenons l’exemple d’une distribution Normale $N(\mu, \sigma^2)$ avec $\mu = \hat{R}, \sigma = \sqrt{M\hat{S}EP(\hat{R})}$, un intervalle de prédiction pour R calculé avec les quantiles à 40% et à 80% est :

$$[\hat{R} + q_{40}\sqrt{M\hat{S}EP(\hat{R})}; \hat{R} + q_{80}\sqrt{M\hat{S}EP(\hat{R})}]$$

où q_{40} et q_{80} sont les quantiles de la loi normale standard.

Pour le cas du coefficient de corrélation de décroissant pour les \hat{U}_i , cela nous conduit à un intervalle de prédiction : [6 845 025;7 373 269], et pour un coefficient de corrélation constant on obtient l’intervalle suivant : [6 859 452;7 325 343]

Quatrième partie
Etude des résultats

Chapitre 9

Résumé des résultats obtenus et commentaires

Méthode de Prov.	Méthodes Déterministes			Méthodes Stochastiques			
	CL Standard	LC	BF Dét.	Bootstrap CL	Bootstrap CL/BF	Mack	BF Stoch.
Prov. Totale	6 440	5 967	7 506	6 614	7 455	6 440	6 967
Erreur de Prédiction				166	90	323	425

TABLE 9.1 – Résumé des différents résultats obtenus précédemment (en milliers).

Observations : Tout d’abord, nous pouvons noter l’écart existant sur l’estimation de la provision (de l’ordre d’une dizaine de pour cent entre la plus petite et la plus grande valeur) selon la méthode utilisée. En effet, les méthodes reposant sur Chain-Ladder semblent légèrement sous-estimer la provision et à l’inverse, celles reposant sur Bornhuetter-Ferguson ont une tendance à la surestimation ; le pilotage de l’actuaire via son ratio S/P a priori doit être prudent. Notons également que la méthode Bornhuetter-Ferguson stochastique donne un résultat intermédiaire. De plus, nous remarquons que la combinaison de ces deux méthodes via le Bootstrap nous apporte un résultat de l’ordre de grandeur de celui de la méthode de Bornhuetter-Ferguson déterministe, ceci s’expliquant par l’importance des provisions pour les années de survénances les plus récentes.

En ce qui concerne l’erreur de prédiction, les résultats que nous avons obtenus, qui sont perfectibles pour certaines méthodes (prise en compte de l’erreur relative au choix du ratio S/P, de l’année de changement), sont meilleurs dans le cadre de l’utilisation de la méthode du Bootstrap. En effet, le Bootstrap variante Chain-Ladder et la méthode de Mack, reposant toutes deux sur Chain-Ladder,

apportent des erreurs de prédiction variant du simple au double (166 026 pour le Bootstrap contre 322 527 pour Mack). Quant à la méthode de Bornhuetter-Fergusson stochastique, nous obtenons un résultat encore plus élevé (de 425 483).

Les MSEP n'étant pas additives, nous n'avons pas pu combiner les méthodes de Mack et de Bornhuetter-Ferguson stochastique. Ainsi, nous n'avons également pas pu comparer cette approche au résultat obtenu à partir la combinaison des deux méthode via le Bootstrap qui est relativement faible (89 750). Ceci reste un axe de recherche suite à ce bureau d'étude.

Chapitre 10

Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement dans le cadre de la méthode du Bootstrap CL/BF

Comme indiqué dans l'intitulé de ce chapitre, cette partie est exclusivement dédiée à l'étude du résultat obtenu à partir de la méthode du Bootstrap combinant Chain-Ladder et Bornhuetter-Ferguson.

Rappelons avant tout l'objet de notre étude dont le but final est de déterminer une provision et l'erreur de prédiction qui lui est associée. Dans le raisonnement que nous avons suivi jusque là, nous avons apporté une réponse partielle au problème. En effet, dans notre réflexion, nous avons avant tout appliqué l'erreur de prédiction à la méthode Chain-Ladder. Bien que les choix du loss ratio S/P et de l'année de changement soient laissés à l'appréciation de l'actuaire, il semble être légitime de se poser la question de leurs impacts respectifs sur la valeur de la provision totale ainsi que sur son erreur de prédiction.

Dans le but d'étudier un éventuel impact du choix du loss ratios et de l'année de changement, nous avons alors décidé, dans le cadre de notre bureau d'études, de faire varier ces paramètres et d'en étudier leurs impacts. Dans notre cas, nous avons observé des loss ratios variant de 65% à 85% et des années de changement de 1 à 9.

Ci-dessous figurent les deux graphiques 3D où sont respectivement représentées la valeur de la provision totale et l'erreur de prédiction en fonction de l'évolution du ratio S/P et de l'année de changement.

Remarque : Le détail des données simulées a été joint en fin de ce chapitre.

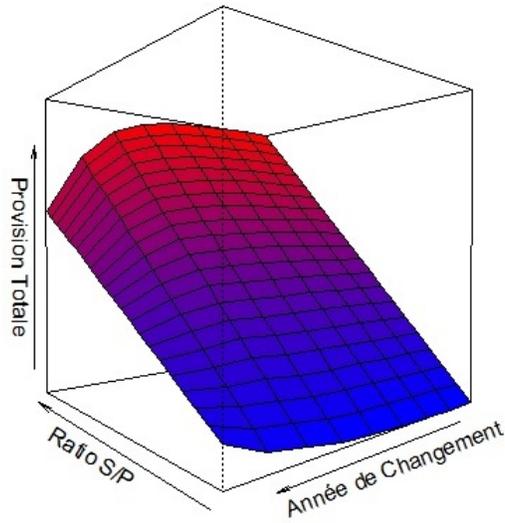


FIGURE 10.1 – Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement sur la provision totale.

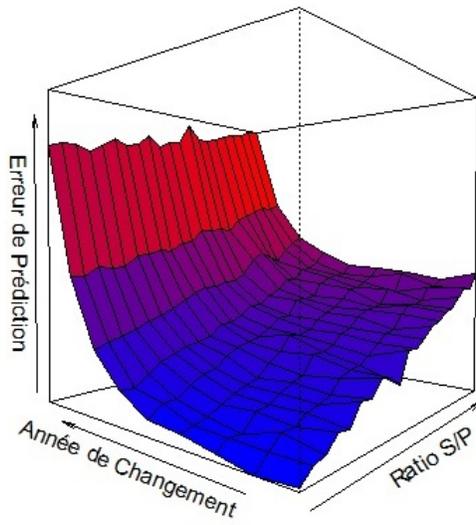


FIGURE 10.2 – Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement sur l'erreur de prédiction.

Observations :

- Etude du premier graphique : nous remarquons que les choix des deux paramètres influent sur la provision totale. Il est également intéressant de remarquer que le choix du ratio S/P joue un rôle prépondérant sur la valeur de la provision finale. A contrario, en dehors des années de changement 8 et 9, la provision totale reste constante lorsque ce deuxième paramètre varie (entre 1 et 7) pour un ratio S/P fixé ;
- Etude du deuxième graphique : de même que précédemment, nous constatons que les choix de ces deux paramètres ont une influence non négligeable sur la valeur de la provision totale. Nous observons également que le ratio S/P semble être légèrement plus influent que l'année de changement bien que ce deuxième paramètre reste tout aussi important.

Comme nous venons de l'observer, les choix respectifs du ratio S/P et de l'année de changement ont un effet non-négligeable sur la provision et l'erreur de prédiction associée. Ainsi, il est important de souligner que notre calcul de l'erreur de prédiction de la méthode du Bootstrap combinée ne prend pas en considération l'erreur liée aux choix de ces deux paramètres.

	Année de Changement CL => BF								
	1	2	3	4	5	6	7	8	9
65%	6 345 891	6 345 694	6 348 481	6 352 518	6 358 925	6 371 332	6 408 751	6 445 166	6 538 164
66%	6 443 817	6 444 006	6 446 712	6 448 457	6 451 511	6 465 907	6 496 282	6 520 451	6 588 004
67%	6 542 375	6 541 676	6 543 291	6 546 832	6 548 323	6 557 966	6 582 797	6 596 785	6 641 080
68%	6 638 381	6 639 528	6 641 180	6 642 323	6 643 626	6 650 492	6 668 843	6 667 959	6 691 125
69%	6 737 578	6 736 684	6 737 879	6 738 977	6 741 543	6 745 044	6 757 090	6 744 510	6 743 315
70%	6 833 917	6 835 090	6 834 598	6 836 968	6 835 501	6 837 405	6 844 461	6 820 132	6 796 235
71%	6 931 736	6 932 966	6 932 604	6 933 663	6 932 271	6 929 871	6 930 841	6 896 452	6 845 325
72%	7 028 640	7 030 069	7 030 345	7 029 786	7 025 629	7 022 564	7 016 928	6 970 007	6 898 244
73%	7 127 306	7 128 152	7 128 995	7 127 573	7 122 253	7 116 283	7 106 934	7 045 033	6 947 495
74%	7 223 681	7 224 602	7 224 371	7 224 006	7 218 316	7 208 199	7 193 086	7 120 370	7 001 940
75%	7 323 123	7 322 243	7 321 633	7 320 555	7 313 862	7 301 317	7 280 149	7 195 721	7 051 014
76%	7 419 639	7 420 364	7 418 769	7 416 145	7 409 298	7 394 012	7 366 858	7 271 579	7 101 112
77%	7 518 219	7 518 800	7 515 763	7 512 062	7 505 380	7 487 325	7 453 247	7 346 632	7 151 756
78%	7 615 331	7 615 210	7 613 494	7 612 092	7 600 472	7 579 316	7 540 005	7 421 387	7 202 652
79%	7 712 830	7 713 022	7 711 682	7 707 792	7 696 425	7 673 156	7 627 261	7 496 169	7 254 833
80%	7 809 710	7 810 573	7 809 699	7 804 218	7 792 538	7 767 454	7 714 958	7 570 910	7 307 467
81%	7 908 548	7 907 059	7 905 878	7 901 911	7 887 187	7 858 304	7 803 976	7 645 623	7 357 020
82%	8 005 400	8 004 219	8 003 449	7 998 701	7 983 508	7 951 285	7 889 148	7 719 957	7 407 511
83%	8 102 748	8 104 909	8 101 430	8 093 542	8 078 896	8 045 146	7 975 932	7 795 111	7 458 877
84%	8 201 430	8 200 938	8 198 536	8 191 967	8 172 007	8 138 531	8 063 579	7 872 711	7 509 346
85%	8 298 547	8 297 811	8 296 516	8 288 806	8 270 200	8 229 983	8 149 034	7 947 952	7 559 398

TABLE 10.1 – Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement sur la provision totale dans le cadre de la méthode du Bootstrap CL/BF.

	Année de Changement CL => BF								
	1	2	3	4	5	6	7	8	9
65%	68 728	68 210	69 213	70 195	70 642	73 813	79 396	91 157	114 382
66%	70 428	70 300	70 443	70 684	72 680	73 712	80 362	91 080	115 278
67%	70 869	72 097	70 385	71 209	72 667	75 453	80 796	92 602	115 379
68%	72 940	72 160	72 849	72 443	74 151	77 008	81 744	92 203	114 600
69%	72 648	73 343	73 144	73 653	75 597	78 041	82 761	92 547	115 938
70%	75 566	73 976	74 152	74 333	75 897	78 691	83 658	94 497	117 632
71%	75 464	74 769	75 413	74 867	77 474	79 100	84 575	94 779	116 543
72%	77 855	76 601	75 935	76 809	78 481	79 820	85 989	95 331	117 124
73%	77 275	77 322	78 800	78 455	79 267	81 598	87 054	95 149	119 298
74%	75 547	77 463	78 808	80 880	79 833	81 550	87 849	96 884	117 316
75%	80 414	78 739	79 529	80 472	81 083	82 749	88 054	97 067	118 688
76%	80 177	80 172	81 031	81 007	82 605	84 275	88 722	99 402	118 847
77%	81 482	81 464	81 674	82 479	82 646	84 284	88 920	99 475	122 968
78%	83 008	81 914	83 337	83 591	85 356	85 719	90 085	99 876	120 301
79%	83 711	84 184	84 370	84 386	85 558	87 170	92 900	99 583	120 288
80%	84 875	85 480	84 678	85 696	85 858	88 799	94 359	101 840	121 076
81%	85 330	85 618	84 834	87 454	88 348	88 767	93 397	102 771	121 910
82%	86 956	85 942	87 828	87 663	89 227	90 905	92 559	103 847	122 917
83%	88 653	88 081	86 953	88 426	89 772	91 002	94 855	104 912	123 092
84%	88 257	88 233	90 567	89 603	89 869	92 063	96 452	104 646	124 332
85%	90 955	89 435	90 692	91 081	91 679	94 247	97 728	105 758	124 565

TABLE 10.2 – Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement sur l'erreur de prédiction dans le cadre de la méthode du Bootstrap CL/BF.

Cinquième partie

Conclusion

En assurance IARD, l'un des objectifs de l'actuaire est d'effectuer des provisionnements en vue des paiements des différents sinistres. Les actuaires ont développé de nombreuses techniques dites déterministes qui permettent d'obtenir un montant des provisions qui couvrira les sinistres à payer, puis des méthodes stochastiques dérivées des méthodes déterministes qui, en plus de fournir des provisions, permettent de calculer l'incertitude autour de celles-ci. Néanmoins chaque méthode stochastique existante permet uniquement de déterminer une erreur de prédiction pour une méthode déterministe bien précise. Or, pour obtenir des résultats plus fiables les compagnies utilisent souvent une combinaison de plusieurs méthodes de provisionnement. En effet, comme chaque méthode possède des avantages et des inconvénients, il paraît naturel d'établir des combinaisons pour obtenir des provisions optimales. L'objectif de notre bureau d'études était donc de déterminer l'erreur de prédiction pour la combinaison des méthodes très connues : Chain-Ladder et Bornhuetter-Ferguson.

Dans un souci d'exhaustivité, nous avons dans notre première partie étudié les méthodes Chain-Ladder, Bornhuetter-Ferguson déterministes ainsi que la méthode London-Chain qui est une extension de la méthode Chain-Ladder. Puis nous avons étudié la méthode de Mack qui permet de calculer l'erreur de prédiction de la méthode Chain-Ladder, la méthode Bornhuetter-Ferguson stochastique qui permet de calculer l'erreur de prédiction relative à la méthode Bornhuetter-Ferguson déterministe ainsi que la méthode de Bootstrap qui nous a permis de trouver une réponse à notre problème. Dans notre dernière partie, nous détaillons et comparons tous les résultats obtenus précédemment. L'erreur de prédiction que nous avons trouvée n'est pas totalement satisfaisante car bien qu'elle répercute entièrement l'erreur liée à la méthode Chain-Ladder elle ne définit que partiellement l'erreur de prédiction de la méthode Bornhuetter-Ferguson. En effet, l'erreur calculée n'explique pas l'erreur liée au choix du S/P ni à l'année de changement de méthode.

En guise d'ouverture, on peut s'interroger sur une manière d'incorporer l'erreur liée au jugement d'expert (choix du S/P et de l'année de changement). Des études annexes et surtout des échanges avec les autres services de la compagnie (commerciaux, économistes, management, etc.) permettront aux actuaires d'appréhender au mieux ces paramètres exogènes sans pour autant en estimer l'erreur. Aussi, il serait intéressant d'élaborer une démarche qui répondrait au problème en combinant les méthodes de Mack et Bornhuetter-Ferguson stochastique.

Sixième partie

Annexes

Annexe A

Données utilisées

En ce qui concerne les données que nous utiliserons par la suite, nous avons choisi d'appliquer les différentes méthodes sur des données de marché d'un assureur automobile.

Années	Primes	Années de développement									
		0	1	2	3	4	5	6	7	8	9
1988	4 711 333	1 376 384	1 211 168	535 883	313 790	168 142	79 972	39 235	15 030	10 865	4 086
1989	5 335 525	1 576 278	1 437 150	652 445	342 694	188 799	76 956	35 042	17 089	12 507	
1990	5 947 504	1 763 277	1 540 231	678 959	364 199	177 108	78 169	47 391	25 288		
1991	6 354 197	1 779 698	1 498 531	661 401	321 434	162 578	84 581	53 449			
1992	6 738 172	1 843 224	1 573 604	613 095	299 473	176 842	106 296				
1993	7 079 444	1 962 385	1 520 298	581 932	347 434	238 375					
1994	7 254 832	2 033 371	1 430 541	633 500	432 257						
1995	7 739 379	2 072 061	1 458 541	727 098							
1996	8 154 065	2 210 754	1 517 501								
1997	8 435 918	2 206 886									

TABLE A.1 – Triangle des incréments de paiements.

Années	Primes	Années de développement									
		0	1	2	3	4	5	6	7	8	9
1988	4 711 333	1 376 384	2 587 552	3 123 435	3 437 225	3 605 367	3 685 339	3 724 574	3 739 604	3 750 469	3 754 555
1989	5 335 525	1 576 278	3 013 428	3 665 873	4 008 567	4 197 366	4 274 322	4 309 364	4 326 453	4 338 960	
1990	5 947 504	1 763 277	3 303 508	3 982 467	4 346 666	4 523 774	4 601 943	4 649 334	4 674 622		
1991	6 354 197	1 779 698	3 278 229	3 939 630	4 261 064	4 423 642	4 508 223	4 561 672			
1992	6 738 172	1 843 224	3 416 828	4 029 923	4 329 396	4 506 238	4 612 534				
1993	7 079 444	1 962 385	3 482 683	4 064 615	4 412 049	4 650 424					
1994	7 254 832	2 033 371	3 463 912	4 097 412	4 529 669						
1995	7 739 379	2 072 061	3 530 602	4 257 700							
1996	8 154 065	2 210 754	3 728 255								
1997	8 435 918	2 206 886									

TABLE A.2 – Triangle des paiements cumulés.

Annexe B

Codes R

Les programmes suivants sont directement applicables à partir des données programmées ci-dessous.

```
#####  
### Triangle supérieur des paiements cumulés ###  
#####  
  
a1988=c(1376384,2587552,3123435,3437225,3605367,3685339,3724574,3739604,3750469,3754555)  
a1989=c(1576278,3013428,3665873,4008567,4197366,4274322,4309364,4326453,4338960,NA)  
a1990=c(1763277,3303508,3982467,4346666,4523774,4601943,4649334,4674622,NA,NA)  
a1991=c(1779698,3278229,3939630,4261064,4423642,4508223,4561672,NA,NA,NA)  
a1992=c(1843224,3416828,4029923,4329396,4506238,4612534,NA,NA,NA,NA)  
a1993=c(1962385,3482683,4064615,4412049,4650424,NA,NA,NA,NA,NA)  
a1994=c(2033371,3463912,4097412,4529669,NA,NA,NA,NA,NA,NA)  
a1995=c(2072061,3530602,4257700,NA,NA,NA,NA,NA,NA,NA)  
a1996=c(2210754,3728255,NA,NA,NA,NA,NA,NA,NA,NA)  
a1997=c(2206886,NA,NA,NA,NA,NA,NA,NA,NA,NA)  
  
primes=c(4711333,5335525,5947504,6354197,6738172,7079444,7254832,7739379,8154065,8435918)  
  
triangle=rbind(a1988,a1989,a1990,a1991,a1992,a1993,a1994,a1995,a1996,a1997)
```

B.1 Méthodes déterministes

B.1.1 Méthode Chain-Ladder

```
CL=function(triangle){

#####
### CALCUL DU D-TRIANGLE ###
#####

n=dim(triangle)[1]
Dtriangle=matrix(rep(0,n*n),nrow=n)
Dtriangle[,1]=1

for(j in 1:(n-1)){
  for(i in 1:(n-j)){
    Dtriangle[i,j+1]=triangle[i,j+1]/triangle[i,j]
  }
}

#####
### CALCUL FACTEURS DE DEVELOPPEMENT ###
#####

lambda=rep(1,n)

for(k in 2:n){
  lambda[k]=sum(triangle[1:(n+1-k),k])/sum(triangle[1:(n+1-k),k-1])
}

#####
### CALCUL TRIANGLE INFERIEUR ###
#####

M_cumulé=triangle

for(j in 1:(n-1)){
  for(i in (n+1-j):n){
    M_cumulé[i,j+1]=M_cumulé[i,j]*lambda[j+1]
  }
}
```

```

#####
### CALCUL PROVISIONS CL ###
#####

prov_CL=0

for(i in 2:n){
  prov_CL[i]=M_cumulé[i,n]-M_cumulé[i,(n+1-i)]
}

return(list(Dtriangle=Dtriangle,lambda=lambda,M_cumulé=M_cumulé,prov_CL=sum(prov_CL)))
}

```

B.1.2 Méthode London-Chain

#NECESSITE L'IMPLEMENTATION DU PROGRAMME CL

```
LC=function(triangle){

  n=dim(triangle)[1]

  lambda=0
  beta=0

  for(j in 0:(n-2)){
    Cj_bar0=mean(triangle[1:(n-j-1),j+1])
    Cj_bar1=mean(triangle[1:(n-j-1),j+2])

    num=mean(triangle[1:(n-j-1),j+1]*triangle[1:(n-j-1),j+2])-Cj_bar0*Cj_bar1
    denom=mean(triangle[1:(n-j-1),j+1]^2)-Cj_bar0^2

    lambda[j+1]=num/denom
    beta[j+1]=Cj_bar1-lambda[j+1]*Cj_bar0
  }

  lambda[n-1]=Cj_bar1/Cj_bar0
  beta[n-1]=0

  M_cumulé=triangle

  for(j in 1:(n-1)){
    for(i in (n+1-j):n){
      M_cumulé[i,j+1]=M_cumulé[i,j]*lambda[j]+beta[j]
    }
  }

  prov_LC=0

  for(i in 2:n){
    prov_LC[i]=M_cumulé[i,n]-M_cumulé[i,(n+1-i)]
  }

  return(list(lambda_LC=lambda,beta_LC=beta,M_cumulé=M_cumulé,prov_LC=sum(prov_LC)))
}
```

B.1.3 Méthode de Bornhuetter-Ferguson

#NECESSITE L'IMPLEMENTATION DU PROGRAMME CL

```
BF=function(triangle,primes,ratio){

  n=dim(triangle)[1]
  M=CL(triangle)[[3]]

  #####
  ### CALCUL de la charge actuelle ###
  #####

  charge_actuelle=rep(1,n)
  for (i in 1:n){
    charge_actuelle[i]=triangle[i,(n-i+1)]
  }

  #####
  ### CALCUL de l'ultime à priori ###
  #####

  ultime_priori=primes*ratio

  #####
  ### CALCUL part de la charge ###
  #####

  part=rep(1,n)
  for (i in 1:n){
    part[i]=charge_actuelle[i]/M[i,n]
  }

  #####
  ### CALCUL de la provision de BF ###
  #####

  prov_BF=rep(1,n)
  for (i in 1:n){
    prov_BF[i]=(1-part[i])*ultime_priori[i]
  }
}
```

```
prov=sum(prov_BF)

#####
### CALCUL de l'ultime de BF ###
#####

ultime=rep(1,n)
ultime=prov_BF+charge_actuelle

return(list(provision=prov,part=part,ultime_priori=ultime_priori,ultime=ultime,
vecteur_prov=prov_BF))

}
```

B.1.4 Analyse graphique des loss ratios S/P

```
#####  
##### Fonction Ratio #####  
#####  
  
Ratio=function(triangle,primes){  
  
    n=dim(triangle)[1]  
    ratio_ultime_CL=matrix(1,10,10)  
  
    for (j in 1:10){  
        for (i in 1:10){  
            ratio_ultime_CL[i,j]=(CL(triangle)[[3]][i,j])/primes[i]*100  
        }  
    }  
  
    seq=seq(n)  
    return(matplot(seq,aperm(ratio_ultime_CL),type = "b",  
        xlab = "année de développement", ylab = "loss ratio",  
        main = "évolution du loss ratio"))  
  
}
```

B.2 Méthodes Stochastiques

B.2.1 Méthode du Bootstrap/Chain-Ladder

```
Bootstrap_CL=function(triangle){

  N=10^4
  n=nrow(triangle)

  M_cumulé=triangle # Triangle des paiements cumulés initial

  #####
  ### CALCUL DU D-TRIANGLE ###
  #####

  Dtriangle=matrix(0,n,n)
  Dtriangle[,1]=1

  for(j in 1:(n-1)){
    for(i in 1:(n-j)){
      Dtriangle[i,j+1]=M_cumulé[i,j+1]/M_cumulé[i,j]
    }
  }

  #####
  ### CALCUL LINK RATIOS : lambda ###
  #####

  lambda=0

  for(k in 1:(n-1)){
    lambda[k]=sum(triangle[1:(n-k),k+1])/sum(triangle[1:(n-k),k])
  }

  #####
  ### CALCUL SIGMA^2 ###
  #####

  S2=0
  for(j in 1:(n-2)){
    S2[j]=(1/(n-j-1))*sum(triangle[1:(n-j),j]*((triangle[1:(n-j),j+1]
      /triangle[1:(n-j),j])-lambda[j])^2)
```

```

}
S2[n-1]=min((S2[n-2]^2)/(S2[n-3]),S2[n-2],S2[n-3])

#####
### CALCUL TRIANGLE DES RESIDUS : res ###
#####

res=matrix(0,n,n)
for(j in 2:n){
  for(i in 1:(n-j+1)){
    res[i,j]=sqrt(M_cumulé[i,j-1]/S2[j-1])*(Dtriangle[i,j]-lambda[j-1])
  }
}

#####
### BOUCLE ###
#####

A=matrix(0,n,N)

for(k in 1:N){

#####
### CALCUL TRIANGLE DES RESIDUS "REAJUSTES/REPLACES" : res_adj ###
#####

V=as.vector(res)
V=V[V!=0] # Suppression des résidus valant 0 (par construction)

res_adj=matrix(0,n,n)

for(j in 2:n){
  for(i in 1:(n-j+1)){
    res_adj[i,j]=sample(V,1)
  }
}

#####
### CALCUL Dtriangle simulé : Dtriangle_sim : ###
#####

Dtriangle_sim=Dtriangle

```

```

for(j in 2:n){
  for(i in 1:(n-j+1)){
    Dtriangle_sim[i,j]=res_adj[i,j]*sqrt(S2[j-1]/M_cumulé[i,j-1])+lambda[j-1]
  }
}

#####
### CALCUL LAMBDA SIMULES : lambda_sim ###
#####

lambda_sim=1

for(j in 2:n){
  lambda_sim[j]=sum(M_cumulé[1:(n-j+1),j-1]*Dtriangle_sim[1:(n-j+1),j])
  /sum(M_cumulé[1:(n-j+1),j-1])
}

lambda_sim=lambda_sim[-1]

#####
### CALCUL TRIANGLE DES PAIEMENTS CUMULES (C-L "inversé") : N_cumulé ###
#####

M_cumulé_sim=M_cumulé

for(j in 2:n){
  M_cumulé_sim[,j]=lambda_sim[j-1]*M_cumulé_sim[,j-1]
}

####

T_cumulé_sim=M_cumulé_sim

for(j in 1:(n-1)){
  for(i in (n+1-j):n){
    T_cumulé_sim[i,j+1]=rnorm(1,lambda_sim[j]*M_cumulé_sim[i,j],sqrt(S2[j]
    *M_cumulé_sim[i,j]))
  }
}

#####
### CALCUL PROVISIONS : prov ###

```

```
#####  
  
temp=0  
  
for(i in 2:n){  
  temp[i]=T_cumulé_sim[i,n]-T_cumulé_sim[i,(n+1-i)]  
}  
  
A[,k]=temp  
  
}  
  
prov=apply(A,2,sum)  
  
return(list(prov_tot=mean(prov),prediction_error=sd(prov)))  
  
}
```

B.2.2 Méthode du Bootstrap - Combinaison Chain-Ladder/Bornhuetter-Ferguson

```
#NECESSITE L'IMPLEMENTATION DU PROGRAMME CL et BF

Bootstrap_CLBF=fonction(triangle,année_changement,primes,ratio){

  N=10^4
  n=nrow(triangle)

  M_cumulé=triangle # Triangle des paiements cumulés initial

  #####
  ### CALCUL DU D-TRIANGLE ###
  #####

  Dtriangle=matrix(0,n,n)
  Dtriangle[,1]=1

  for(j in 1:(n-1)){
    for(i in 1:(n-j)){
      Dtriangle[i,j+1]=M_cumulé[i,j+1]/M_cumulé[i,j]
    }
  }

  #####
  ### CALCUL LINK RATIOS : lambda ###
  #####

  lambda=0

  for(k in 1:(n-1)){
    lambda[k]=sum(triangle[1:(n-k),k+1])/sum(triangle[1:(n-k),k])
  }

  #####
  ### CALCUL SIGMA^2 ###
  #####

  S2=0
  for(j in 1:(n-2)){
    S2[j]=(1/(n-j-1))*sum(triangle[1:(n-j),j]*((triangle[1:(n-j),j+1]
```

```

        /triangle[1:(n-j),j])-lambda[j])^2)
    }
S2[n-1]=min((S2[n-2]^2)/(S2[n-3]),S2[n-2],S2[n-3])

#####
### CALCUL TRIANGLE DES RESIDUS : res ###
#####

res=matrix(0,n,n)
for(j in 2:n){
  for(i in 1:(n-j+1)){
    res[i,j]=sqrt(M_cumulé[i,j-1]/S2[j-1])*(Dtriangle[i,j]-lambda[j-1])
  }
}

#####
### BOUCLE ###
#####

A=matrix(0,n,N)

for(k in 1:N){

#####
### CALCUL TRIANGLE DES RESIDUS "REAJUSTES/REPLACES" : res_adj ###
#####

V=as.vector(res)
V=V[V!=0] # Suppression des résidus valant 0 (par construction)

res_adj=matrix(0,n,n)

  for(j in 2:n){
for(i in 1:(n-j+1)){
    res_adj[i,j]=sample(V,1)
  }
}

#####
### CALCUL Dtriangle simulé : Dtriangle_sim : ###
#####

```

```

Dtriangle_sim=Dtriangle

for(j in 2:n){
  for(i in 1:(n-j+1)){
    Dtriangle_sim[i,j]=res_adj[i,j]*sqrt(S2[j-1]/M_cumulé[i,j-1])+lambda[j-1]
  }
}

#####
### CALCUL LAMBDA SIMULES : lambda_sim ###
#####

lambda_sim=1

for(j in 2:n){
  lambda_sim[j]=sum(M_cumulé[1:(n-j+1),j-1]*Dtriangle_sim[1:(n-j+1),j])
  /sum(M_cumulé[1:(n-j+1),j-1])
}

lambda_sim=lambda_sim[-1]

#####
### CALCUL TRIANGLE DES PAIEMENTS CUMULES (C-L "inversé") : N_cumulé ###
#####

M_cumulé_sim=M_cumulé

for(j in 2:n){
  M_cumulé_sim[,j]=lambda_sim[j-1]*M_cumulé_sim[,j-1]
}

###

T_cumulé_sim=M_cumulé_sim

for(j in 1:(n-1)){
  for(i in (n+1-j):n){
    T_cumulé_sim[i,j+1]=rnorm(1,lambda_sim[j]*M_cumulé_sim[i,j],sqrt(S2[j]
    *M_cumulé_sim[i,j]))
  }
}

#####

```

```

### CALCUL PROVISIONS CHAIN LADDER : ###
#####

temp=rep(0,n)

for(i in 2:année_changement){
  temp[i]=T_cumulé_sim[i,n]-T_cumulé_sim[i,(n+1-i)]
}

A[,k]=temp

#####
### CALCUL PROVISIONS BORNHUETTER-FERGUSON : ###
#####

prov_BF=BF(T_cumulé_sim,primes,ratio)[[5]]

A[(année_changement+1):n,k]=prov_BF[(année_changement+1):n]
}

prov=apply(A,2,sum)

return(list(prov_tot=mean(prov),prediction_error=sd(prov)))
}

```

B.2.3 Méthode Bornhuetter-Ferguson Stochastique

#NECESSITE L'IMPLEMENTATION DU PROGRAMME BF

```
BF_Sto=function(triangle,primes,ratio){

  n=dim(triangle)[1]
  annee_dev=c(1:10)
  U=BF(triangle,primes,ratio)[[4]]

  #####
  ### Calcul de y ###
  #####

  triangle_decumule=triangle

  for(j in 1:(n-1)){
    triangle_decumule[,j+1]=triangle[,j+1]-triangle[,j]
  }

  y=rep(1,n)
  for (i in 1:n){
    y[i]=sum(triangle_decumule[,i],na.rm=TRUE)/sum(U[1:(n+1-i)])
  }

  fit=lm(log(y)~annee_dev)
  #summary(fit)
  #plot(annee_dev,log(y),xlim=c(0,12),ylim=c((-10),0))
  #abline(fit,col='red')

  #Calcul de y10+
  y=c(y,(exp(fit$coefficients[2]*11+fit$coefficients[1])))

  #normalisation
  y=y/sum(y)

  #####
  ### Calcul des sj**2 ###
  #####

  s=rep(1,(n-1))
  for (k in 1:(n-1)){
    s[k]=1/(n-k)*sum(((triangle_decumule[1:(n+1-k),k]-U[1:(n+1-k)])*y[k])^2)/U[1:(n+1-k)]
  }
}
```

```

}

y_abs=abs(y)
fit2=lm(sqrt(s)~y_abs[-c(10,11)])
#summary(fit2)
#plot(y_abs[-c(10,11)],sqrt(s),xlab="y_adjust")
#abline(fit2,col='red')

#Calcul de s10 et s10+
s=c(s,((fit2$coefficients[2]*y[n]+fit2$coefficients[1]))^2,
((fit2$coefficients[2]*y[n+1]+fit2$coefficients[1]))^2)

#lissage
s_adj=1
s_adj=(fit2$coefficients[2]*y+fit2$coefficients[1])^2
s=s_adj

#####
### Calcul de la provision BF ###
#####

R=1
for (i in 1:10){
  R[i]=sum(y[(n+2-i):(n+1)])*U[i]
}

#####
### Calcul de l'erreur de Processus de R[i] ###
#####

err_proc=rep(1,n)
for (i in 1:n){
  err_proc[i]=sum(s[(n+2-i):(n+1)])*U[i]
}

#####
### Calcul de l'erreur de y ###
#####

mse_y=rep(1,n)
for(i in 1:n){
  mse_y[i]=s[i]/sum(U[1:(n+1-i)])
}

```

```

#estimation de mse_y[11] dans la pratique
mse_y=c(mse_y,mse_y[n])

#####
### calcul de l'erreur de z ###
#####

mse_z=rep(0,(n+1))
for (i in 1:n){
  mse_z[i]=min(sum(mse_y[1:i]),sum(mse_y[(i+1):(n+1)]))
}

#####
### Calcul de l'erreur dans le cas d'un coefficient de corrélation décroissant ###
#####

#####
### Calcul de l'erreur de U ###
#####

q=sum(U)/sum(primes)
mse_U=1
for(i in 1:n){
  mse_U[i]=primes[i]/(n-sqrt(2*n))*sum(primes[1:n]*(U[1:n]/primes[1:n]-q)**2)
}

#####
### Calcul de l'erreur d'estimation ###
#####

err_est=rep(1,n)

z=y
for (i in 1:(n)){
  z[i+1]=z[i]+y[i+1]
}

for (i in 1:n){
  err_est[i]=(U[i]**2+mse_U[i])*mse_z[n+1-i]+mse_U[i]*((1-z[n+1-i])**2)
}

```

```

#####
### Calcul du MSEP de R[i] ###
#####

msep_decroissant=1
msep_decroissant=err_proc+err_est

#####
### Calcul de la covariance Ri Rj ###
#####

coeff_z=matrix(NA,n,n)
for (j in 2:n){
  for (i in 1:(j-1)){
    coeff_z[i,j]=sqrt(z[n+1-j]*(1-z[n+1-i])/z[n+1-i]/(1-z[n+1-j]))
  }
}

coeff_U_decroissant=matrix(1,n,n)
for (i in 1:n){
  for (j in 1:n){
    coeff_U_decroissant[i,j]=1/(1+abs(i-j))
  }
}

Covariance=matrix(NA,n,n)
for (j in 1:n){
  for (i in 1:(j-1)){
    Covariance[i,j]=coeff_U_decroissant[i,j]*sqrt(mse_U[i])
    *sqrt(mse_U[j])*(1-z[n+1-i])*(1-z[n+1-j])
    +coeff_z[i,j]*sqrt(mse_z[n+1-i])*sqrt(mse_z[n+1-j])*U[i]*U[j]
  }
}

#####
### Calcul du MSEP de R ###
#####

MSEP_decroissant=sum(err_proc)+sum(err_est)+2*sum(Covariance,na.rm=TRUE)

#####
### Calcul de l'erreur dans le cas d'un coefficient de corrélation constant ###

```

```

#####

#####
### Calcul de l'erreur de U ###
#####

mse_U=1
for(i in 1:n){
  mse_U[i]=primes[i]/(n-sqrt(n))*sum(primes[1:n]*(U[1:n]/primes[1:n]-q)**2)
}

#####
### Calcul de l'erreur d'estimation ###
#####

err_est=rep(1,n)

z=y
for (i in 1:(n)){
  z[i+1]=z[i]+y[i+1]
}

for (i in 1:n){
  err_est[i]=(U[i]**2+mse_U[i])*mse_z[n+1-i]+mse_U[i]*((1-z[n+1-i])**2)
}

#####
### Calcul du MSEP de R[i] ###
#####

msep_constant=1
msep_constant=err_proc+err_est

#####
### Calcul de la covariance Ri Rj ###
#####

coeff_z=matrix(NA,n,n)
for (j in 2:n){
  for (i in 1:(j-1)){
    coeff_z[i,j]=sqrt(z[n+1-j]*(1-z[n+1-i])/z[n+1-i]/(1-z[n+1-j]))
  }
}

```

```

}

coeff_U_constant=1/sqrt(n)

Covariance=matrix(NA,n,n)
for (j in 1:n){
  for (i in 1:(j-1)){
    Covariance[i,j]=coeff_U_constant*sqrt(mse_U[i])*sqrt(mse_U[j])
      *(1-z[n+1-i])*(1-z[n+1-j])+coeff_z[i,j]*sqrt(mse_z[n+1-i])
      *sqrt(mse_z[n+1-j])*U[i]*U[j]
  }
}

#####
### Calcul du MSEP de R ###
#####

MSEP_constant=sum(err_proc)+sum(err_est)+2*sum(Covariance,na.rm=TRUE)

return(list(R,sum(R),sqrt(MSEP_decroissant),sqrt(msep_decroissant),
sqrt(MSEP_constant),sqrt(msep_constant)))

}

```

Table des figures

1	Diagramme de Lexis	8
2	Les différentes lectures d'un triangle de paiements	9
1.1	Courbe des facteurs de développement ($\hat{\lambda}_j$).	16
1.2	C-C plot pour $j=0$	18
1.3	C-C plot pour $j=1$	18
1.4	Colonne du D-Triangle pour $j=0,1,2$	19
3.1	Evolution des loss ratios en fonction des années de développement.	26
7.1	Raisonnement théorique de la méthode du Bootstrap combinée CL/BF.	42
8.1	Représentation des logarithmes des cadences \hat{y}_j^* en fonction des années de développement.	47
8.2	Représentation de la racine des \hat{s}_j^{2*} en fonction des valeurs abso- lues $ \hat{y}_j^* $	48
10.1	Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement sur la provision totale.	55
10.2	Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement sur l'erreur de prédiction.	55

Liste des tableaux

1	Triangle des incréments de paiements, $Y = (Y_{i,j})$	11
2	Triangle des paiements cumulés, $C = (C_{i,j})$	12
1.1	Facteurs de développement en fonction des années de développe- ment.	15
1.2	Triangle des paiements cumulés (en milliers) $C = (C_{i,j})$ et leurs projections futures $\hat{C} = (\hat{C}_{i,j})$	16
1.3	Provisions Chain-Ladder.	17
1.4	D-Triangle $(f_{i,j})$	19
1.5	Observation via des données statistiques du D-Triangle.	20
2.1	Estimation des paramètres λ_j^{LC} et β_j^{LC} par année de développe- ment.	23
2.2	Tableau des paiements cumulés.	23
2.3	Provisions London-Chain.	23
3.1	Calcul des ultimes Chain-Ladder $U_{CL} = (U_{CLi})$	26
3.2	Calcul des ultimes a priori $U = (U_i)$	27
3.3	Calcul de l'ultime BF $C_{.,n} = (C_{i,n})$ et des provisions BF.	27
4.1	Tableau des provisions retenues.	30
4.2	Tableau récapitulant les provisions respectives Chain-Ladder et Bornhuetter-Ferguson ainsi que les provisions retenues et la pro- vision totale.	30
5.1	Estimation des paramètres $(\lambda_j)_{(j=1,\dots,n-1)}$ et $(\sigma_j)_{(j=1,\dots,n-1)}$	36
5.2	Mesures de risque conditionnel de prédiction.	37
8.1	Paramètres estimés dans la méthode de Mack	48
8.2	PSAP et incertitude à l'ultime estimées.	49
8.3	PSAP et incertitude à l'ultime estimées.	50
9.1	Résumé des différents résultats obtenus précédemment (en milliers). 52	

10.1	Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement sur la provision totale dans le cadre de la méthode du Bootstrap CL/BF.	57
10.2	Etude de l'influence des choix respectifs du ratio S/P et de l'année de changement sur l'erreur de prédiction dans le cadre de la méthode du Bootstrap CL/BF.	58
A.1	Triangle des incréments de paiements.	63
A.2	Triangle des paiements cumulés.	64

Bibliographie

- [1] Baron T., *Mesure de l'Incertitude sur les Provisions en Assurance Non Vie Approche Bornhuetter-Ferguson*.
- [2] Charpentier A., *Risque et Assurance de Arthur Charpentier*, Chapitre 1.
- [3] Denuit M., Charpentier A., *Mathématiques de l'Assurance Non-Vie - Tome II : Tarification et Provisionnement*, Economica, 2005.
- [4] England P.D., *Addendum to "Analytic and bootstrap estimates of prediction errors in claims reserving"*, Report No. Actuarial Research Paper No. 138, London, UK : Faculty of Actuarial Science & Insurance, City University London, 2001.
- [5] England P.D., Verral R.J., *Predictive distributions of outstanding liabilities in general insurance*.
- [6] Goichon A., *Provisionnement en Assurance IARD dans le cas de segments corrélés*, 2011.
- [7] Jaziri. S., *Méthodes de provisionnement non- vie et risque de réserve à un an*, 2011.
- [8] Mack T., *Distribution-free calculation of the standard error of chain ladder reserve estimates*, Astin Bulletin 23, 1993.
- [9] Mack T., *The prediction error of Bornhuetter/Ferguson stochastic*, Astin Bulletin, 2008.
- [10] Michaud C., *La mesure du risque de provisionnement à horizon un an en assurance non-vie - Prise en compte du jugement d'expert et de données exogènes*, 2012.
- [11] Schmidt K., *Bornhuetter-Ferguson as a general principle of loss reserving*, Astin Colloque.
- [12] Partrat C., Lecoœur E., Nessi J.M., Nisipau E. ; Reiz O., *Provisionnement technique en Assurance Non-Vie, Perspectives actuarielles modernes*, Economica, 2007.
- [13] Peng S., Edward W.F., *Dependent Loss Reserving Copulas*, 2010