MODELISATION STOCHASTIQUE DES PROVISIONS TECHNIQUES EN ASSURANCE NON-VIE

KERDALI Abida[1]

LATRECHE Abdelouahab[2]

Résumé :
Nous présentons dans ce travail quelques méthodes et modèles de réservation récursives pour le calcul stochastique des provisions techniques d'une compagnie d'assurance. Pour cela, nous avons considéré plusieurs types de modèles pour l'estimation des valeurs du triangle de développement ; nous nous sommes inspirés d'une modélisation autorégressive d'ordre 1, 2, 3 selon le modèle choisi. Dans une première étape, nous avons proposé une méthode directe, c'est-à-dire que le calcul n'est pas récursif, pour déterminer un estimateur des paramètres du modèle. Ensuite, dans une seconde étape, nous l'avons amélioré pour lui permettre de calculer récursivement les paramètres du modèle proposé évitant ainsi l'inversion d'une matrice de grande taille, seulement il faut prendre quelques précautions pour que ce nouvel estimateur ne s'éloigne pas trop de la vraie valeur des paramètres à estimer puisque les méthodes récurrentes sont très sensibles aux valeurs initiales. Pour terminer, nous avons comparé les résultats obtenus avec celles des modèles classiques de provisionnement.

Mots clés : modélisation stochastiques, provisions techniques, modèle de réservation, assurance non-vie, estimation récursive.

1. 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 une cotisation ou une prime qui lui permettra son approvisionnement. En assurance, le cycle de production est dit inversé, cette inversion du cycle rend impossible ; la détermination exacte de la richesse d'une société d'assurance à un instant donné puisqu'elle ne connait pas avec exactitude ses engagements. 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.

Pour calculer la charge ultime, la méthode la plus utilisée est la méthode de Chaine-Ladder, ce modèle permet de calculer une cadence de développement moyenne à partir des données historiques. Un très grand nombre de modélisations sont proposées dans la littérature pour déterminer le montant de provisions ; ces méthodes se divisent en deux catégories. (a) les méthodes déterministes et les méthodes stochastiques : La première classe ne fait aucune hypothèse probabiliste permettant de mesurer l'incertitude associée à la prédiction du montant de la réserve, elles reposent sur l'hypothèse de stabilité du délai s'écoulant entre la survenance d'un sinistre et le règlement en absence de l'inflation, de changement de structure portefeuille des garanties de contrat et plus généralement la gestion des sinistres. Parmi ces méthodes on trouve : Méthode chaine-Ladder, Méthode de London-Chain, Méthode de De Vylder, Méthode de Bornuhetter-Ferguson, etc. (b) la seconde classe a été introduite justement afin d'obtenir une estimation de cette erreur de prévision. Le recours aux modèles stochastiques est justifié par le besoin de mesurer l'incertitude présentée dans les triangles et les résultats issus des méthodes déterministes. Ces méthodes ont toutes été développées dans le but de palier à un problème spécifique (manque de données, historique insuffisant, développement long, environnement incertain, etc. ; parmi ces méthodes on trouve : Modèles de Mack, Méthode de Boostrap, Modèle linéaire de Christophide, etc. Toutes ces méthodes classiques ont chacune leurs avantages et inconvénients.

Nous considérons dans ce travail un certain nombre de méthodes et modèles de réservation récursives pour le calcul stochastique des provisions techniques. Les algorithmes récurrents ont occupé une place considérable due au développement du traitement numérique et à l'augmentation constante de la puissance des calculateurs permettant ainsi l'implémentation en temps réel d'algorithmes de plus en plus sophistiqués. L'intérêt de la forme récursive des algorithmes récursifs est dû à la possibilité du traitement en ligne des donnés en temps réel grâce à une mémorisation finie dans un vecteur de taille fixé. En effet l'estimation courante est remise à jour de manière à ce que le nombre d'opérations et l'espace mémoire ne croient pas avec le nombre d'observations et ne dépendent que du modèle choisi

Pour bien comprendre quels sont les effets du modèle de réservation récursive, il est très intéressant de comparer les résultats obtenus avec les modèles classiques de provisionnement et pour cela nous avons considéré plusieurs types de modèles possibles pour l'estimation des valeurs du triangle de développement ; nous nous sommes inspirés d'une modélisation autorégressive d'ordre 1, 2, 3 selon le modèle choisi, nous proposons au début une méthode permettant de déterminer un estimateur des paramètres du modèle proposé d'une manière directe, c'est-à-dire que le calcul n'est pas récursif, hors l'inconvénient de cette méthode est qu'elle nécessite l'inversion d'une matrice de grande taille ce qui est très couteux en temps de traitement du programme si le nombre de paramètre à estimer est grand, pour résoudre ce problème, nous avons proposé une seconde méthode qui améliore beaucoup la première permettant de calculer récursivement les paramètres du modèle proposé évitant l'inversion d'une matrice de grande taille, seulement il faut prendre quelques précautions pour que cet estimateur ne s'éloigne pas trop de la vraie valeur des paramètres à estimer puisque les méthodes récurrentes sont très sensibles aux valeurs initiales.

Ce papier qui débute par une introduction des modèles considérés, nous présentons dans la section 2, les différentes méthodes d'estimation récursive des paramètres des modèles proposés. Les sections 3,4 sont consacrées respectivement aux choix des valeurs initiales et aux facteurs d'oubli. La section 5 est réservée aux expérimentations numériques et des comparaisons des résultats de chaque modèle avec les modèles classiques déterministes et stochastiques. Une conclusion termine ce travail.

2. Modèles proposés

Nous supposons dans cette modélisation que la valeur du triangle de développement $D_{i,j}$ (prestation, des sinistres de l'année i, réalisée après j années) s'écrit comme une combinaison (linéaire, quadratique, etc.) des valeurs $D_{k,l}$, $k \leq i$ et $l \leq j$ et ($k \neq i$ ou $l \neq j$), voir la Figure 1, plus une certaine erreur considérée comme une variable aléatoire.

$\ddots$	$\vdots$	$\vdots$	$\vdots$
$\cdots$	$D_{i-2,j-2}$	$D_{i-2,j-1}$	$D_{i-2,j}$
$\cdots$	$D_{i-1,j-2}$	$D_{i-1,j-1}$	$D_{i-1,j}$
$\cdots$	$D_{i,j-2}$	$D_{i,j-1}$	$D_{i,j}$

Figure 1 : Les différentes valeurs possibles qui précèdent $D_{i,j}$.

Nous considérons plusieurs types de modèles possibles pour l'estimation des valeurs du triangle de développement, à savoir :

- M100s : $D^s_{i,j} = \alpha_0 + \alpha_1 D^s_{i,j-1} + u_{i,j}$ pour $i = 0,\ldots,n-1$, $j = 1,\ldots,n-i$ et $s = 1, 2, 3,\ldots$
- M001s : $D^s_{i,j} = \alpha_0 + \alpha_1 D^s_{i-1,j} + u_{i,j}$ pour $i = 1,\ldots,n$, $j = 0,\ldots,n-i$ et $s = 1, 2, 3,\ldots$
- M010s : $D^s_{i,j} = \alpha_0 + \alpha_1 D^s_{i-1,j-1} + u_{i,j}$
- M110s : $D^s_{i,j} = \alpha_0 + \alpha_1 D^s_{i,j-1} + \alpha_2 D^s_{i-1,j-1} + u_{i,j}$
- M011s : $D^s_{i,j} = \alpha_0 + \alpha_1 D^s_{i-1,j-1} + \alpha_2 D^s_{i-1,j} + u_{i,j}$
- M101s : $D^s_{i,j} = \alpha_0 + \alpha_1 D^s_{i,j-1} + \alpha_2 D^s_{i-1,j} + u_{i,j}$

- M111s : $D_{i,j}^{s} = \alpha_0 + \alpha_1 D_{i,j-1}^{s} + \alpha_2 D_{i-1,j-1}^{s} + \alpha_3 D_{i-1,j}^{s} + u_{i,j}$
  pour $i = 1,\ldots,n-1$, $j = 1,\ldots,n-i$ et $s = 1, 2, 3,\ldots$
- M200s : $D_{i,j}^{s} = \alpha_0 + \alpha_1 D_{i,j-1}^{s} + \alpha_2 D_{i,j-2}^{s} + u_{i,j}$
  pour $i = 0,\ldots,n-1$, $j = 2,\ldots,n-i$ et $s = 1, 2, 3,\ldots$
- M002s : $D_{i,j}^{s} = \alpha_0 + \alpha_1 D_{i-1,j}^{s} + \alpha_2 D_{i-2,j}^{s} + u_{i,j}$
  pour $i = 2,\ldots,n$, $j = 0,\ldots,n-i$ et $s = 1, 2, 3,\ldots$
- M202s :

$$D_{i,j}^{s} = \alpha_0 + \alpha_1 D_{i,j-1}^{s} + \alpha_2 D_{i,j-2}^{s} + \alpha_3 D_{i-1,j}^{s} + \alpha_4 D_{i-2,j}^{s} + u_{i,j}$$

  pour $i = 2,\ldots,n$, $j = 2,\ldots,n-i$ et $s = 1, 2, 3,\ldots$
- Etc.

où $u_{i,j}$ est l'erreur d'estimation considérée comme une variable aléatoire.

3. Estimation récursive des réserves

Toute méthode récurrente part d'une fonction objectif à optimiser et se termine par une méthode ou un algorithme récursif pour la mise en œuvre effective sur les données. Théoriquement, plusieurs critères à optimiser peuvent être considérés (approche Bayesienne, maximum de vraisemblance, etc.). Mais souvent le critère retenu est la minimisation de l'erreur quadratique moyenne (Mean Square Error ou MSE). Cette approche permet de construire une suite d'estimateurs du paramètre à estimer qui permet d'approcher de plus en plus la vraie valeur de ce paramètre.

Tous les modèles proposés ci-dessous peuvent se mettre sous la forme matricielle suivante : $D_{i,j}^{s} = \varphi_{i,j}^{T}\theta + u_{i,j}$. Pour le modèle M111s, par exemple nous avons :

$$\varphi_{i,j} = \begin{pmatrix} 1 \\ D_{i,j-1}^{s} \\ D_{i-1,j-1}^{s} \\ D_{i-1,j}^{s} \end{pmatrix} \text{ Et } \theta = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$ qui sont deux vecteurs colonne de taille 4.

Nous avons une représentation autorégressif d'ordres 1, 2 ou 3 selon le modèle choisi. La meilleure estimation de $D_{i,j}^{s}$ au sens des moindres carrés et que l'on note $\hat{D}_{i,j}^{s}(\theta)$ est par conséquent donnée par $\hat{D}_{i,j}^{s}(\theta) = \varphi_{i,j}^{t}\theta$. L'erreur d'estimation

$\hat{u}_{i,j}(\theta) = u_{i,j}$ est donc égal à $D_{i,j}^{s} - \hat{D}_{i,j}^{s}(\theta)$. Nous avons la relation $\hat{u}_{i,j}(\theta^{*}) = u_{i,j}$ où θ^{*} est la vraie valeur du paramètre.

Le critère d'optimisation est donc : $\min\limits_{\theta} f(\theta)$

où $f(\theta) = \frac{1}{N}\sum_{i}\sum_{j}\hat{u}_{i,j}(\theta)^2 = \frac{1}{N}\sum_{i}\sum_{j}(D_{i,j}^{s} - \varphi_{i,j}^{T}\theta)^2$

$\hat{u}_{i,j}(\theta)$ n'est définie que pour certaines valeurs de i et de j. Par exemple pour le modèle M111s, nous avons $i = 1,\ldots,n-1$ et $j = 1,\ldots,n-i$ et $N = n(n-1)/2$ étant le nombre d'éléments.

Considérons le théorème suivant qui nous donne un estimateur des paramètres du modèle. Cet estimateur est calculé de manière directe (c'est-à-dire que le calcul n'est pas récursif).

Théorème 1. Supposons que la matrice $\sum_{i}\sum_{j}\varphi_{i,j}\varphi_{i,j}^{T}$ soit inversible, nous avons :

$$\hat{\theta}_N = \left(\sum_{i}\sum_{j}\varphi_{i,j}\varphi_{i,j}^{T}\right)^{-1}\sum_{i}\sum_{j}D_{i,j}^{s}\varphi_{i,j}$$

Preuve. En calculant et en annulant les dérivées premières $\frac{\partial f}{\partial \hat{\alpha}_i}$ pour $i = 0,1,\ldots,q$ (q=1, 2 ou 3, selon le modèle choisi), nous obtenons pour le modèle M111s :

$$\frac{\partial f(\hat{\alpha}_0,\ldots,\hat{\alpha}_3)}{\partial \hat{\alpha}_0} = -\frac{2}{N}\sum_{i}\sum_{j}(D_{i,j}^{s} - \varphi_{i,j}^{T}\hat{\theta}_N) = 0$$

$$\frac{\partial f(\hat{\alpha}_0,\ldots,\hat{\alpha}_3)}{\partial \hat{\alpha}_1} = -\frac{2}{N}\sum_{i}\sum_{j}(D_{i,j}^{s} - \varphi_{i,j}^{T}\hat{\theta}_N)D_{i,j-1}^{s} = 0$$

$$\frac{\partial f(\hat{\alpha}_0,\ldots,\hat{\alpha}_3)}{\partial \hat{\alpha}_2} = -\frac{2}{N}\sum_{i}\sum_{j}(D_{i,j}^{s} - \varphi_{i,j}^{T}\hat{\theta}_N)D_{i-1,j-1}^{s} = 0$$

$$\frac{\partial f(\hat{\alpha}_0,\ldots,\hat{\alpha}_3)}{\partial \hat{\alpha}_3} = -\frac{2}{N}\sum_{i}\sum_{j}(D_{i,j}^{s} - \varphi_{i,j}^{T}\hat{\theta}_N)D_{i-1,j}^{s} = 0$$

Donc $\sum_i \sum_j (D^s_{i,j} - \varphi^T_{i,j}\hat{\theta}_N)\varphi_{i,j} = \underline{0}$

Ainsi $\sum_i \sum_j D^s_{i,j}\varphi_{i,j} - \sum_i \sum_j \varphi^T_{i,j}\hat{\theta}_N\varphi_{i,j} = \underline{0}$

D'où $\sum_i \sum_j D^s_{i,j}\varphi_{i,j} = \sum_i \sum_j \varphi^T_{i,j}\hat{\theta}_N\varphi_{i,j}$

où $\underline{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$: un vecteur à q+1 éléments nuls.

Finalement, $\sum_i \sum_j D^s_{i,j}\varphi_{i,j} = \sum_i \sum_j \left(\varphi_{i,j}\varphi^T_{i,j}\right)\hat{\theta}_N$ et

$$\hat{\theta}_N = \left(\sum_i \sum_j \varphi_{i,j}\varphi^T_{i,j}\right)^{-1} \sum_i \sum_j D^s_{i,j}\varphi_{i,j}$$

sous l'hypothèse que la matrice $\sum_i \sum_j \varphi_{i,j}\varphi^T_{i,j}$ (de taille $(q+1)\times(q+1)$) soit inversible. □

Considérons maintenant le théorème suivant permettant de calculer récursivement un estimateur des paramètres d'un modèle. C'est-à-dire que le résultat de la proposition 1 peut être obtenu de manière récursive en calculant $\hat{\theta}_N$ à partir de $\hat{\theta}_{N-1}$ et d'une nouvelle observation. Pour cela, rangeons les paires de valeurs (i, j) dans une liste indicée de *1* à *N*. (i_m, j_m) désigne la paire d'indice *m*. Nous pouvons donc écrire $\hat{\theta}_N = \left(\sum_{m=1}^{N} \varphi_{i_m,j_m}\varphi^T_{i_m,j_m}\right)^{-1} \sum_{m=1}^{N} D^s_{i_m,j_m}\varphi_{i_m,j_m}$.

Théorème 2. Pour $k > 1$, nous avons :

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \frac{D^s_{i_k,j_k} - \varphi^T_{i_k,j_k}\hat{\theta}_{k-1}}{k}\overline{R}_k^{-1}\varphi_{i_k,j_k}$$

avec $\overline{R}_k = \overline{R}_{k-1} + \frac{1}{k}(\varphi_{i_k,j_k}\varphi_{i_k,j_k}^T - \overline{R}_{k-1})$

Preuve. Du théorème 1, posons :

$\hat{\theta}_k = R_k^{-1}\sum_{m=1}^{k} D_{i_m,j_m}^s \varphi_{i_m,j_m}$ avec $R_k = \sum_{m=1}^{k}\varphi_{i_m,j_m}\varphi_{i_m,j_m}^T$

(1)

Nous avons :

$$\hat{\theta}_k = R_k^{-1}\sum_{m=1}^{k} D_{i_m,j_m}^s \varphi_{i_m,j_m}$$

$$= R_k^{-1}\left(\sum_{m=1}^{k-1} D_{i_m,j_m}^s \varphi_{i_m,j_m} + D_{i_k,j_k}^s \varphi_{i_k,j_k}\right)$$

$$= R_k^{-1}\left(R_{k-1}\hat{\theta}_{k-1} + D_{i_k,j_k}^s \varphi_{i_k,j_k}\right)$$ (car d'après (1), nous avons $\sum_{m=1}^{k} D_{i_m,j_m}^s \varphi_{i_m,j_m} = R_k\hat{\theta}_k$)

$$= R_k^{-1}\left(R_k\hat{\theta}_{k-1} - \varphi_{i_k,j_k}\varphi_{i_k,j_k}^T\hat{\theta}_{k-1} + D_{i_k,j_k}^s \varphi_{i_k,j_k}\right)$$

(car d'après (1), nous avons $R_k = R_{k-1} + \varphi_{i_k,j_k}\varphi_{i_k,j_k}^T$ donc $R_{k-1} = R_k - \varphi_{i_k,j_k}\varphi_{i_k,j_k}^T$)

$$= R_k^{-1}\left(R_k\hat{\theta}_{k-1} + \varphi_{i_k,j_k}(-\varphi_{i_k,j_k}^T\hat{\theta}_{k-1} + D_{i_k,j_k}^s)\right)$$

$$= \hat{\theta}_{k-1} + R_k^{-1}\varphi_{i_k,j_k}(D_{i_k,j_k}^s - \varphi_{i_k,j_k}^T\hat{\theta}_{k-1})$$

(2)

Aussi :

$$\overline{R}_k = \frac{1}{k}R_k$$

$$= \frac{1}{k}(R_{k-1} + \varphi_{i_k,j_k}\varphi_{i_k,j_k}^T)$$ (car d'après (1), nous avons $R_k = R_{k-1} + \varphi_{i_k,j_k}\varphi_{i_k,j_k}^T$)

$$= \frac{1}{k}((k-1)\overline{R}_{k-1} + \varphi_{i_k,j_k}\varphi_{i_k,j_k}^T)$$

$$= \frac{k-1}{k}\bar{R}_{k-1} + \frac{1}{k}\varphi_{i_k,j_k}\varphi_{i_k,j_k}^T$$

$$= \bar{R}_{k-1} + \frac{1}{k}(\varphi_{i_k,j_k}\varphi_{i_k,j_k}^T - \bar{R}_{k-1})$$

Finalement : $\hat{\theta}_k = \hat{\theta}_{k-1} + \dfrac{D_{i_k,j_k}^s - \varphi_{i_k,j_k}^T\hat{\theta}_{k-1}}{k}\bar{R}_k^{-1}\varphi_{i_k,j_k}$.

□

Les équations de ce théorème possèdent théoriquement l'aspect d'un algorithme récursif, mais cet algorithme n'est pas encore exploitable du point de vue numérique. L'inconvénient de cette méthode est qu'elle nécessite l'inversion d'une matrice de taille $(q+1)\times(q+1)$, ce qui est très coûteux en temps de traitement du programme si le nombre de paramètres à estimer est grand. Pour remédier à cela, nous proposons le résultat suivant :

Théorème 3. Pour $k > 1$, nous avons : $\hat{\theta}_k = \hat{\theta}_{k-1} + (D_{i_k,j_k}^s - \varphi_{i_k,j_k}^T\hat{\theta}_{k-1})L_k$

avec $L_k = \dfrac{1}{1+\varphi_{i_k,j_k}^T P_{k-1}\varphi_{i_k,j_k}} P_{k-1}\varphi_{i_k,j_k}$ et $P_k = P_{k-1} - L_k\varphi_{i_k,j_k}^T P_{k-1}$.

Preuve. Nous avons :

$$R_k^{-1} = (R_{k-1} + \varphi_{i_k,j_k}\varphi_{i_k,j_k}^T)^{-1}$$

$$= R_{k-1}^{-1} - \frac{R_{k-1}^{-1}\varphi_{i_k,j_k}\varphi_{i_k,j_k}^T R_{k-1}^{-1}}{1+\varphi_{i_k,j_k}^T R_{k-1}^{-1}\varphi_{i_k,j_k}} \qquad \text{(car } (A+bb^T)^{-1} = A^{-1} - \frac{A^{-1}bb^T A^{-1}}{1+b^T A^{-1} b}$$

) (3)

Donc $R_k^{-1}\varphi_{i_k,j_k} = R_{k-1}^{-1}\varphi_{i_k,j_k} - \dfrac{R_{k-1}^{-1}\varphi_{i_k,j_k}\varphi_{i_k,j_k}^T R_{k-1}^{-1}\varphi_{i_k,j_k}}{1+\varphi_{i_k,j_k}^T R_{k-1}^{-1}\varphi_{i_k,j_k}}$

$$= \frac{R_{k-1}^{-1}\varphi_{i_k,j_k}}{1+\varphi_{i_k,j_k}^T R_{k-1}^{-1}\varphi_{i_k,j_k}}$$

En posant $R_k^{-1} = P_k$ et $L_k = \dfrac{1}{1+\varphi_{i_k,j_k}^T P_{k-1}\varphi_{i_k,j_k}} P_{k-1}\varphi_{i_k,j_k}$, on obtient :

$P_k = P_{k-1} - L_k \varphi^T_{i_k, j_k} P_{k-1}$ d'après (3) et $\hat{\theta}_k = \hat{\theta}_{k-1} + (D^s_{i_k, j_k} - \varphi^T_{i_k, j_k} \hat{\theta}_{k-1}) L_k$ d'après (2). □

Cette méthode est plus intéressante puisqu'elle ne nécessite pas l'inversion d'une matrice de taille $(q+1)\times(q+1)$, ce qui minimise la propagation des erreurs d'arrondi dans les calculs et accélère et diminue le temps d'exécution.

3. Choix des valeurs initiales

La simplicité des méthodes récurrentes étudiées ne doit pas faire oublier qu'en pratique leur implémentation numérique n'est pas une tâche facile puisque ces méthodes récurrentes sont très sensibles aux valeurs initiales. Pour empêcher que l'estimateur ne soit très sensible aux valeurs de départ (les premières valeurs des données), il faut bien choisir ces valeurs initiales. Il faut donc prendre quelques précautions pour que l'estimateur ne s'éloigne pas de la vraie valeur du paramètre à estimer.

3.1 Choix de $\hat{\theta}_1$

Nous avons observé que la bonne valeur pour $\hat{\theta}_1$ est $\underline{0}$ **(**$\underline{0}$ est un vecteur ayant tous ces éléments égaux à *0* et est de dimension (q+*1***)**

3.2 Choix de P_1

Nous avons observé aussi que la valeur initiale de P_1 influence très faiblement la valeur finale de $\hat{\theta}$. Nous avons choisi de prendre comme valeur initiale Id **.**

4. Facteur d'oubli

Comme nous l'avons déjà mentionné les méthodes récurrentes sont très sensibles aux valeurs initiales. Pour remédier à cet inconvénient, nous proposons d'incorporer un facteur d'oubli. C'est-à-dire une suite de valeurs λ_k qui permet d'atténuer les variations des paramètres à estimer.

Nous avons proposé la suite de valeurs $\lambda_k = \lambda_{k-1} - \delta$ et nous avons choisi les conditions : $\lambda_1 = 1$ et $\delta = 0.0001$. Le théorème 3 devient alors :

Théorème 4 Pour $k > 1$, nous avons : $\hat{\theta}_k = \hat{\theta}_{k-1} + (D^s_{i_k,j_k} - \varphi^T_{i_k,j_k}\hat{\theta}_{k-1})L_k$

avec $L_k = \dfrac{\lambda_k}{\lambda_k + \varphi^T_{i_k,j_k} P_{k-1}\varphi_{i_k,j_k}} P_{k-1}\varphi_{i_k,j_k}$, $\quad P_k = P_{k-1} - L_k \varphi^T_{i_k,j_k} P_{k-1}$ et $\lambda_k = \lambda_{k-1} - \delta$.

et comme valeurs initiales : $\hat{\theta}_1 = \underline{0}$, $P_1 = Id$, $\lambda_1 = 1$ et $\delta = 0.0001$.

5. Expérimentation numériques

A. Données générées aléatoirement : Nous avons comparé notre modèle sur les données tirées de la référence [BEN], ces données proviennent des 5000 simulations générées par le programme proposé dans le papier. Voici le triangle moyen utilisé dans la référence :

Année de survenance	Année de développement					
	1	2	3	4	5	6
1997	4 740	210 938	1 850 617	2 608 283	3 495 492	4 607 766
1998	181 773	1 314 395	3 609 713	4 881 558	5 880 198	
1999	112 302	1 137 729	2 286 295	2 962 406		
2000	283 555	977 041	2 131 768			
2001	474 694	1 078 972				
2002	304 668					

Table 1. Triangle des paiements cumulés.

Le principal résultat des différents modèles étudiés est l'obtention de la moyenne des réserves prévues. Voici la table2 comparative obtenu :

Année	1998	1999	2000	2001	2002	Somme
Chain-Ladder	1 871 092	1 925 887	2 614 473	5 440 206	7 913 296	19 764 953
London-Chain	1 871 092	1 904 664	2 590 249	5 491 631	8 390 847	20 248 482
Analyse	1 411 248	1 635 890	2 332 903	5 053 456	5 236 463	15 669 959
Bornhuetter-Ferguson	1 689 737	1 969 897	2 754 257	4 172 463	4 814 633	15 400 987
Mack	1 871 092	1 925 887	2 614 473	5 440 206	7 913 296	19 764 953
Bootstrap	1 878 463	1 929 746	2 597 262	5 444 237	7 506 241	19 355 950
Avis Experts CL	1 180 000	1 510 000	2 216 000	4 992 000	5 176 000	15 070 000
Avis Experts BF	1 869 000	1 933 000	2 626 000	4 475 000	4 920 000	15 890 000
Corrélations	1 871 092	2 006 246	2 861 360	4 840 828	7 524 391	19 157 234
Modèle individuel	1 178 648	1 593 484	3 022 185	5 078 058	4 847 637	15 720 012

Table 2. Résultats de l'expérimentation (source [BEN]).

La principale remarque de l'auteur est qu'il y a deux catégories de modèles. D'une part, il y a les modèles présentant des résultats similaires à la méthode Chain-Ladder classique. Et d'autre part, il y a les modèles résultant d'une analyse plus fine de la sinistralité.

Les sommes des réserves obtenues par nos modèles sont données dans la table 3 suivante.

Variante	s=			
	1	2	3	4
100	16123919	14989696	15790368	13895734
001	20686079	18129396	15836711	8854923
010	21531075	20117780	17361606	14127364
101	15071934	11780139	40270713	47233775
110	13628232	66973353	78052432	430398063
011	21697982	34786495	193396183	102047750
111	13033545	53527820	62472167	68651186
200	15033918	14976496	27617976	147518226
002	15697119	17126019	16982701	16974243
202	16581981	22849124	16646444	11846550

Table 3. Résultats 1

<u>Commentaire</u> : Nous remarquons que les résultats colorés en vert sont proches des résultats de la méthode de Chain-Ladder (première catégorie). Tandis que les résultats colorés en bleu sont proches des valeurs résultant d'une analyse plus fine de la sinistralité (deuxième catégorie). Les valeurs en rouge et en jaune sont à rejeter car elles sont très éloignées des valeurs obtenues par les méthodes classiques.

<u>B. Données du marché I</u> : Nous avons comparé aussi notre modèle sur les données tirées de la référence [TLS]. En ce qui concerne les données utilisées, elles proviennent du marché d'un assureur automobile. Voici le triangle utilisé dans la référence :

An.	Années de développement									
	0	1	2	3	4	5	6	7	8	9
88	1376384	2587552	3123435	3437225	3605367	3685339	3724574	3739604	3750469	3754555
89	1576278	3013428	3665873	4008567	4197366	4274322	4309364	4326453	4338960	
90	1763277	3303508	3982467	4346666	4523774	4601943	4649334	4674622		
91	1779698	3278229	3939630	4261064	4423642	4508223	4561672			
92	1843224	3416828	4029923	4329396	4506238	4612534				
93	1962385	3482683	4064615	4412049	4650424					
94	2033371	3463912	4097412	4529669						
95	2072061	3530602	4257700							
96	2210754	3728255								
97	2206886									

Table 4. Triangle des paiements cumulés (source [TLS]).

La Table 5 suivant Résume les différents résultats obtenus (en milliers).

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

Table 5. Résultats de l'expérimentation (source [TLS]).

La remarque donnée par les auteurs est « 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. Aussi, la méthode Bornhuetter-Ferguson stochastique donne un résultat intermédiaire.»

Les sommes des réserves obtenues par nos modèles sont données dans la table 6 suivante:

variante	s=			
	1	2	3	4
100	4824893	4016880	4417567	4970764
001	12323824	10623558	9477589	9084955
010	7722468	6674067	6818381	7050535
101	10868616	8166849	7602934	7288919
110	7098752	5833152	6004289	6336115
011	10751956	8449032	7509507	7176740
111	7238057	8474592	15566029	9315366
200	5998650	4635315	5085452	5859897
002	11602320	8906021	13827450	10606529
202	7737582	10605922	1031999512	1130273098

Table 6. Résultats 2

Commentaire : Nous remarquons que les résultats colorés en vert sont proches des résultats de la méthode de Chain-Ladder. Tandis que les résultats colorés en bleu sont proches des valeurs résultant de la méthode de Bornhuetter-Ferguson. Les valeurs en rouge et en jaune sont à rejeter car elles sont très éloignées des valeurs obtenues par les méthodes classiques.

C. Données du marché II : Nous avons comparé aussi notre modèle sur les données tirées de la référence [BOU], elles proviennent de la branche responsabilité civil de la compagnie d'assurance CAAT. Voici le triangle utilisé dans la référence :

An. de sur.	Année de développement										
	1	2	3	4	5	6	7	8	9	10	11
2002	457839860	887023764	1068425322	1208097790	1295340211	1360222837	1407217827	1443384360	1466384360	1486384360	1502384360
2003	671537689	1180029254	1419536816	1586757331	1704352943	1776283795	1818498536	1849142900	1870531264	1888531264	
2004	858562086	1510227839	1797074483	1975167597	2099864885	2184211615	2232436717	2260819349	2287628478		
2005	1123371387	1889996522	2179293626	2371404723	2506864257	2573387600	2623990455	2659105095			
2006	1251342558	2169168341	2513777315	2739206303	2861175638	2939507991	2997812607				
2007	1697774013	2978991755	3430276860	3681628578	3820166424	3920648147					
2008	1920386534	3471396131	3916400458	4185137257	4368920546						
2009	1807079871	3185369533	3690539487	3985513372							
2010	1447258303	2695370642	3145964854								
2011	2010995472	3704295507									
2012	2145515517										

Table 7. Triangle des paiements cumulés (source [BOU]).

La table 8 suivante résumé les différents résultats obtenus en (DA).

Méthode	Somme des provisions
Chain Ladder	7859220425
London Chain	6161517267
Modèle de Mach	7859220425
Modèle log-normale	7866600302

Table 8. Résultats de l'expérimentation (source [BOU]).

La principale remarque donnée par l'auteur est que la méthode London Chain sous-estime la somme des provisions.

Les sommes des réserves obtenues par nos modèles sont données dans la table 9 suivante.

<u>Commentaire</u> : Nous remarquons que les résultats colorés en vert sont proches des résultats de la méthode de Chain-Ladder. Tandis que les résultats colorés en bleu sont une surestimation. Les résultats colorés en jaune sont une sous-estimation. Les valeurs en rouge sont à rejeter car elles sont très éloignées des valeurs obtenues par les méthodes classiques.

variante	s=			
	1	2	3	4
100	12550653826	12769656677	13663135630	13589476655
001	107796804517	119405394035	121933117858	128922189273
010	26741916296	30985966766	34251779552	36408483908
101	17190474727	15833456678	16793504178	17016046808
110	13246241115	12827584618	13398135075	14117182391
011	74293082733	74732581523	64629988125	52414161140
111	10467426958	11420769161	11018626393	11645658993
200	7762054712	7183894130	6278868535	4673920532
002	648053335931	1405327925106	2403516655010	1493385702149
202	10373975121	10948874344	5399013142	5488613836

Table 9. Résultats 3

D. Données du marché III : Montants des sinistres payés par la compagnie d'assurance des hydrocarbures CASH SPA pour les sinistres survenus entre 2004 et 2011 [SBM].

	1	2	3	4	5	6	7	8
1	264588118	364602784	374576654	375330264	380932678	381329660	381329660	3813397888
2	10574209	462845145	604801128	639971700	640642943	640642943	640642943	
3	12156497	149143027	155986202	220259084	256665794	294053390		
4	57295731	342702810	485672884	497304294	510037912			
5	253891795	403699937	843721476	14134189				
6	14276932	70298208	778784302					
7	17073563	80502839						
8	174345152							

Table 10. Triangle des paiements cumulés (source [SBM]).

La synthèse des méthodes est donnée à la table 11.

Méthode	provision
Chain-ladder	1595970615
London chaine	1481814182
Modèle de Mack	1595970615
Modèle log-normal	1246987855

Table 11. Résultats de l'expérimentation (source [SBM]).

Les résultats de nos modèles sont donnés à la table 12.

variante	s=			
	1	2	3	4
100	2413963336	8337737475	15590417169	27271768436
001	30744947541	37663492938	37663589296	34361076318
010	18575484972	19538542063	19946726646	18179569655
101	20618235931	95927693269	100436045780	80002864840
110	5637999539	27147938214	92430378355	81680949194
011	23940323131	33943993062	19691276943	22498146998
111	91407878792	84477755316	19436655720	21724720695
200	3442874921	21505910570	27667153019	35977286218
002	28966751465	46136409764	105887210608	57138643873
202	176849546533	128988268248	106252842635	94667776202

Table 12. Résultat 4

Commentaires : Les valeurs en bleu sont une surestimation des valeurs classiques tandis que les valeurs en rouge sont très éloignées des valeurs classiques.

F. Données du marché IV : Application des méthodes sur la branche Automobile [SAM]. Le triangle de liquidation de la branche automobile est donné dans la table 13.

	1	2	3	4	5	6	7	8
1	9730946	19745612	22311606	26402789	26770053	27140469	27140469	27183444
2	10574209	23130898	27376094	27936252	29222933	29223933	29300815	
3	32371776	60365499	66971001	75091880	75129013	75512562		
4	52840089	85804490	99326476	99652851	102036342			
5	45385992	92106154	92698656	96075035				
6	51902315	54366260	65141380					
7	16209479	53657410						
8	69362667							

Table 13. Triangle des paiements cumulés (source [SAM]).

Les résultats trouvés par les auteurs sont donnés par la table 13.

Méthode	Somme des provisions
Chain Ladder	99172973
London Chain	89597802
Modèle de Mack	99172973
Modèle christophide	78031201

Table 14. Résultats de l'expérimentation (source [SAM]).

Les résultats trouvés par nos modèles sont donnés à la table 15.

variante	s=			
	1	2	3	4
100	178221755	151948482	141891904	150208613
001	370613135	460305329	554506526	588915408
010	126381160 9	357162879	349471651	337769454
101	258284253	178944948	179589719	148087007
110	231378606	167306957	168807143	144493589
011	383051973	410801401	3282475593	356332376
111	203398303	165837281	135527294	362143433 0
200	89275439	83282100	108529107	142394991
002	205247254 3	210748721 8	4249419145	473313187 3
202	277361407	152245721 6	701636305	618632003

Table 15. Résultat 5

Commentaire : Les valeurs en vert sont proches de la valeur de Chain- Ladder. Les valeurs en jaune sont proches des valeurs de London Chain. Les valeurs en bleu sont une surestimation des valeurs de Chain ladder. Les valeurs en rouge sont très éloignées des valeurs classiques.

G. Données du marché V : Application des méthodes sur la branche Transport [SAM], Le triangle de liquidation de la branche transport est donné comme suit

	1	2	3	4	5	6	7	8
1	1238660	12987361 5	15648480 0	16372381 3	20079782 3	20094247 3	20094247 3	20094247 3
2	1766565	25757642	27201637	35012602	52653413	52656499	52663689	

	0	0	5	0	5	3	5	
3	42145068	129899850	204109599	206187747	209521384	221172957		
4	50914418	78907376	397448527	397475424	398802878			
5	4742318	36975511	37047183	39401988				
6	30138926	30459912	148729616					
7	309246	108187704						
8	6741089							

Table 16. Triangle des paiements cumulés (source [SAM]).

Les résultats obtenus par les méthodes de provisionnement sont donnés par la table 17.

Méthode	Somme des provisions
Chain Ladder	289991907
London Chain	245996911
Modèle de Mach	289991907
Bootstrap	274177575
Modèle christophide	66940964

Table 17. Résultats de l'expérimentation (source [SAM]).

Les résultats trouvés par nos modèles sont donnés à la table 18.

variante	s=			
	1	2	3	4
100	1147205043	1195614110	1321126105	1407081368
001	3022879775	3102257306	5407276007	5031545155
010	2394596626	2531734447	4122738734	3691939125
101	296454739	880548872	1028714199	949246321494
110	1089264440	807382709	786753301	2750292525
011	6841056483	7818854270	8185770649	4407267378
111	905527407	1373997449	441287539912	4480397045625
200	1228101424	1518668042	1072908475	1291135689
002	6957258714	8646518324	9151924691	5167212169
202	403463745491	2237074725274	5777902387577	13050145453041

Table 15. Résultat 6

Commentaire : Les valeurs en vert sont proches de la valeur de Chain Ladder, les valeurs en bleu sont une sur estimation, les valeurs en rouge sont erronées.

CONCLUSION

Nous nous sommes intéressés dans ce travail à l'estimation récursive des paramètres d'un modèle stochastique de provisionnement. Ainsi, nous avons considéré le critère de la minimisation de l'erreur quadratique (Mean Square error, MSE). A cet effet nous avons développé des méthodes d'estimation directe et récursive pour le calcul stochastique des réserves pour une compagnie d'assurance. Pour cela, nous avons proposé de les confronter aux différentes méthodes actuarielles classiques.

Les principales remarques sur les expérimentations réalisées sont :

- Les modèles M100s et M200s avec s=1 sont proches des modèles classiques.
- L'efficacité des autres modèles diffère d'une expérimentation à une autre.
- La valeur minimale de tous les modèles est une sous-estimation de la somme des réserves par rapport à la valeur de Chain-Ladder.
- La valeur maximale de tous les modèles est une valeur erronée car elle est très éloignée des valeurs des méthodes classiques.
- La valeur moyenne de tous les modèles proposés peut être une valeur à considérer et peut être confrontée aux autres modèles classiques.

Le choix d'une méthode de provisionnement s'avère très délicat, nous ne pouvons pas dire laquelle est préférable à l'autre et mettre en œuvre une solution pour choisir la plus appropriée, ce choix dépend de son adéquation aux donnée, comme le notait Hans Bulhlman dans le postface du tome 2 de Denuit et Charpentier (2005) '*'le provisionnement n'est pas un problème de modélisation complexe et sophistiquée mais plutôt un exercice de choix de modèle''*. Il n'existe donc pas une méthode universelle qui soit applicable à toutes les compagnies d'assurance ; chaque compagnie à sa propre Politique de gestion des sinistres. En pratique, face à l'ensemble des estimations possibles, il peut s'avérer difficile de trouver le niveau de provisionnement adéquat. Néanmoins, quelques astuces peuvent être utiles, à savoir le biais de l'estimation comme critère et ceci en comparant les valeurs obtenues par un modèle avec les valeurs dont on dispose. Un assureur prudent devrait appliquer diverses méthodes ou imaginer de nouvelles méthodes adaptées à sa gestion et déduire par comparaison des diverses évaluation un montant pour la provision pour sinistres à payer.

Face à l'ouverture du marché de l'assurance en Algérie, les institutions financières nationales ; se sont confronter à d'autres organismes assureurs étrangers évoluant en conformité avec les méthodes de provisionnement les plus récentes, élaborées en réponse aux besoins des réalités économiques et afin d'assurer leur survie dans un environnement prévu plus hostile du point de vue concurrentiel, nos entreprises devront se familiariser obligatoirement avec les nouvelles techniques et mécanisme financières imposées par la nouvelle donnée économique, les méthodes actuarielles développées dans ce document ; particulièrement, le modèle récursif constitue une très bonne approche de la variabilité des réserves, nous vivons dans un monde incertain, complexe et volatile où les changements économiques ; réglementaires, démographique et sociaux modifient notre paysage, les assureurs doivent faire preuve d'adaptation et d'anticipation au quotidien pour pouvoir imaginer des réponses à toutes les inquiétudes et risque de la vie tels que, la dépendance, le prix du pétrole, le terrorisme. Toutes ces craintes doivent pouvoir être atténuées grâce à des garanties de prévoyance dont la maitrise passe par l'exploitation des nouvelles techniques de calcul rendues possible grâce au progrès de l'informatique.

Comme perspective à notre travail, nous pensons à :

- Améliorer les méthodes proposées.
- Faire plus d'expérimentations numériques.
- Faire des études plus poussées sur l'estimation des valeurs initiales.
- Etudier plus à fond les facteurs d'oubli.

REFERENCES

[BEN] Beneteau (G). *Modèle de provisionnement sur données détaillées en assurance non-vie*. Mémoire d'actuariat 2004, ENSAE, France, 2004.

[BEE] Benjamin (S) and Eagles (LM). *Reserves in Lloyd's and the London Market.* Journal of the institute of actuaries 113.2, pages 197-256, 1986.

[BOL] Bollerslev (T). *Generalized autoregressive conditional heteroscedasticity*. journal of econometrics 31, 327-367, 1986.

[BEN] Bollerslev (T), Engle (RF) and Nelson (DB). *Financial market efficiency tests*. in M. H. Pesaran et Mickens Eds. Handbook of Applied Econometrics, vol.4, North Holland, Amsterdam, 1994.

[BOU] Boudaoud (W). *Détermination d'un niveau de provisions techniques requis pour une compagnie d'assurance non-vie par des modèles stochastiques*, Mémoire de Magister, ENSSEA, Alger, 2014.

[CHR] Christophide (S). *Regression models based on log-incremental payments claims reserving,* manual 2. Institute of Actuaries, 1990.

[DPV] Denuit (M), Purcaru (O) and Van Keilegom (I). *bivariate Archimedean copula modelling for Loss-ALAE data in nonlife insurance*. Working paper 04-03, institute des sciences Actuarielles, université catholique de Louvain, Belgium, 2004.

[DEC] Denuit, M. and Charpentier A. (2005) *Mathématique de l'assurance non-vie : Tarification et provisionnement*. Tome 1. Economica.

[DEN] Denuit (M) and Charpentier (A). *Mathématique de l'assurance non-vie : Tarification et provisionnement*. Tome 2. Economica, 2005.

[DIV] Dionne (G) and Vanesse (C). *Automobile insurance ratemaking in the presence of assymmetrcal information*. Journal of Applied Econometrics 7, 149-165, 1992.

[DOR] Dodge (Y) and Rousson (V). *Analyse de Regression Appliquée*. Dunod, Paris.

[EFB] Efron (B). *Bootstrap methods: Another look at th jackknife*. The annals of statistics, vol. 7, no1 p. 1-26, 1979.

[GOU] Gourierou (C). *Modèle ARCH et Applications Financières*. Ed. Economica, 1992.

[GOR] Gourierou (C). *Statistique de l'assurance*. Ed. Economica, Paris, 1999.

[JPB] Droesbeke (J) Fichet (B) et Tassi (P). *Modélisation ARCH : Théorie statistique et applications dans le domaine de la finance*. Editions de l'Université de Bruxelles, Bruxelles, Association pour la statistique et ses utilisations, 1984.

[KRA] Krauth (G). *Provisionnement et corrélation entre branche*. Mémoire d'économiste de CNAM spécialité Actuariat. CNAM, Paris, 2007.

[KER] Kerdali (A). *Estimation Récursive des modèles conditionnellement hétéros-cédastiques Périodiques*. Mémoire de Magister, USTHB, Alger, Algérie, 2005.

[MEH] Mehdi (A). *Calcul des Provisions techniques dans une compagnie d'assurance non-vie*. PGS En Actuariat, USTHB, Alger, Algérie, 2014.

[PLM] Partrat (C), Lecoeur (E), Nessi (JM) et Nisipasu (E). *Provisionnement technique en assurance non-vie, perspective actuarielles moderne*. Edition Economica, 2007.

[PAB] Partrat (C) et Besson (JL). *Assurance non-vie ; Modélisation, Simulation*. Edition Economica, 2005.

[SAM] Serir (A) et Ali (M). *Calcul stochastique et déterministe des provisions technique dans une compagnie d'assurance*. Master en maths financière. USTHB, Alger, 2012.

[KDS] Schmidt (KD). *Modèles et méthodes de réservation*. Petit cours donné à l'université de Strasbourg, 2003.

[TLS] Le Tesson (A), Lenain (A), Samba (S) et Ung (J). *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*. Europe Institute of Actuarial Studies EURIA, 2014.

[YOU] Young (P). *Recursive estimation and time series analysis: An introduction*. Springer Verlag, 1984.