

Analyse et application d’une méthode de provisionnement sans triangle en assurance santé

Jude KLEVINE^{a,1}

^aUniversité de Strasbourg

^bFrance Mutuelle

^cInstitut des Actuaires

1. Introduction

Les méthodes déterministes, souvent critiquées pour leur perte d’information, ont longtemps été utilisées par France Mutuelle pour le calcul de ses provisions, donnant de bons résultats.

Cependant, avec les récentes innovations technologiques, les délais de remboursement ont été accélérés, diminuant en moyenne de 10 jours entre les années 2021 et 2023 comme le montre la figure 1. Cette réduction des délais de paiement entraîne des biais dans les méthodes d’estimation des provisions, car celles-ci supposent des cadences de paiement relativement stables.

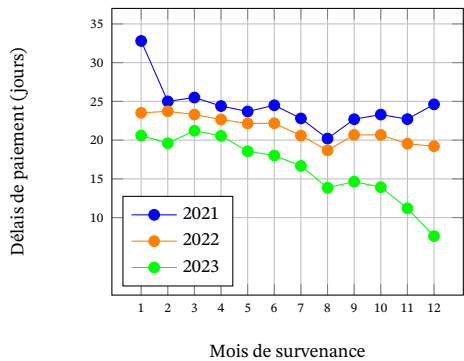


Figure 1. Délais de paiement moyen par période de survénance

Les provisions sont calculées en agrégeant les données dans un triangle 36×36 , où 36 représente les mois sur 3 ans. Les provisions calculées en utilisant les données de 2021 à 2023 pour estimer celles de 2024, sans prendre en compte l’effet d’accélération, sont donc biaisées. L’objectif est de proposer une méthode de calcul sans biais, prenant en compte l’accélération des cadences de paiement.

2. Méthodes conventionnelles de provisionnement agrégé

Le calcul des provisions se faisait en utilisant des triangles agrégés, comme illustré dans la figure 2, avec $Y_{i,j}$ et $C_{i,j}$ représentant respectivement les paiements non cumulés et cumulés à la date j pour des survénances à la date i , où $C_{i,j} = \sum_{k=0}^j Y_{i,k}$.

i, j	0	...	j	...	n
0	$Y_{0,0} C_{0,0}$...	$Y_{0,j} C_{0,j}$...	$Y_{0,n} C_{0,n}$
\vdots	\vdots		\vdots		
i	$Y_{i,0} C_{i,0}$...	$Y_{i,j} C_{i,j}$		
\vdots	\vdots				
n	$Y_{n,0} C_{n,0}$				

Figure 2. Triangle de paiement cumulé et non cumulé

France Mutuelle utilisait principalement des méthodes déterministes, complétées par des méthodes stochastiques pour quantifier les erreurs. Cependant, ces deux dernières années, ces approches, autrefois fiables (Taylor et De Vylder), ont surestimé les provisions par rapport aux valeurs observées, comme le montre la figure 5.

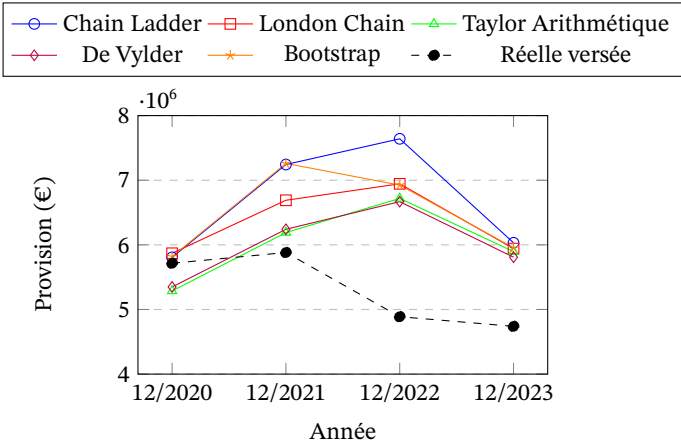


Figure 3. Évolution des provisions IBNR par méthode

Les méthodes déterministes reposent sur une hypothèse forte de stabilité des paiements, ce qui n’est pas le cas ici. Avec des délais de paiement de plus en plus courts, les cadences de paiement des années 2021, 2022 et 2023 diffèrent, entraînant le non-respect de plusieurs hypothèses liées à ces méthodes. Il est donc nécessaire de mettre en place une autre approche pour le calcul des provisions.

3. Études statistiques

Il est crucial de bien comprendre les données avant toute étude. Les différences entre variables d’étude (sexe, type de bénéficiaire, etc.) sont minimales, bien que certaines catégories déclarent moins de sinistres. L’analyse montre une baisse progressive des délais moyens de paiement et une hausse du coût des sinistres due à l’inflation.

Plus de 90% des délais de paiement sont inférieurs à 3 mois, et plus de 90% des montants sont inférieurs à 200%. Les montants restent globalement faibles, à l’exception de rares cas extrêmes qui seront étudiés dans le chapitre suivant.

4. Études des délais de paiement

4.1. Seuils des délais et montants de paiement

L’étude des valeurs extrêmes a permis de mettre en évidence des montants et des délais de paiement inhabituels, c’est-à-dire relativement rares.

Table 1. Seuils des délais et montants de paiement

Variab	Seuil	% des fréquences	% des montants
Délais de paiement	200 jours	1,36%	2,02%
Montants	10 000	0,025%	8,34%

Note : La table présente les seuils retenus pour les délais et montants de paiement.

Une fois cette mise en évidence réalisée, la base de données est séparée en deux catégories :

- **Sinistres graves** : paiements au-delà du seuil.
- **Sinistres attritionnels** : paiements en dessous du seuil.

Étant donné la part négligeable des sinistres avec des délais supérieurs à 200 jours, les calculs seront effectués en tenant compte uniquement du seuil des montants.

4.2. Analyse des délais de paiement

L'analyse des délais de paiement repose sur l'application de modèles de survie, notamment **Kaplan-Meier** et **Nelson-Aalen**. Le **test de log-rank** met en évidence des différences significatives entre les courbes de survie des années 2021, 2022 et 2023, révélant une hétérogénéité temporelle.

L'estimation par les **modèles de Cox** permet de quantifier l'impact des différentes variables explicatives sur les délais de paiement. Les variables considérées incluent :

Table 2. Liste des variables étudiées pour l'analyse des délais de paiement.

Variable	Description	Modalités / Type
CB	Catégorie de bénéficiaire	{A, E, C}
SB	Sexe du bénéficiaire	{F, M}
TP	Type de paiement	{I, V, N}
AB	Âge du bénéficiaire	Numérique
TC	Type de contrat	{CO, IN}
PS	Période de survie	[1, 36] mois
RB	Région du bénéficiaire	14 régions
CG	Catégories de codes d'actes	9 groupes

Le modèle de Cox simple a été testé, mais le non-respect de l'hypothèse de proportionnalité a conduit à un modèle de Cox stratifié sur quatre variables, affectant les performances. Un modèle de Cox temporel a ensuite été évalué, sans amélioration notable.

Pour dépasser ces limites, des approches de machine learning ont été explorées. L'analyse a identifié le **type de paiement** et la **période de survie** comme variables clés. Parmi les modèles testés, **DeepSurv** a offert les meilleures performances pour estimer les courbes de survie, confirmant l'impact des variables sur les délais de paiement.

5. Formalisation des modèles de calculs

Les délais de paiement diminuent, notamment ceux < 10 jours (68 % des données, 58 % des prestations), faussant les provisions et rendant les méthodes déterministes inadéquates.

5.1. Problématiques

Posons pour une date j de l'année y , avec $N = 365$ jours et Y_i^y les paiements pour une survie en y avec un délai i , la cadence de paiement de la manière suivante :

$$\text{cadence}_j^y = \frac{\sum_{i=0}^j Y_i^y}{\sum_{i=0}^N Y_i^y}$$

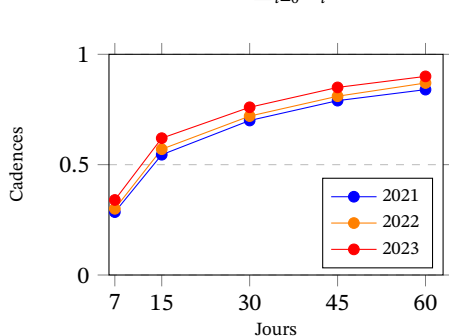


Figure 4. Cadences de paiements pas année de survie

D'après le graphique 4, les cadences varient sur les 3 ans d'étude, avec 2023 s'écartant le plus dès le 7^e jour. Supposer une cadence

similaire reviendrait à aligner les courbes des 3 années, ce qui n'est pas le cas ici.

5.2. Approche fréquence-coût de Parodi

Avec des cadences de paiement variables, il est crucial d'en tenir compte pour le calcul des provisions. La méthode de Parodi permet d'intégrer ces spécificités en modélisant les IBNR comme suit :

- Modélisation des IBNeR avec des modèles de régression
- Modélisation des IBNyR avec des modèles fréquence-coût

Ici, les IBNeR ne sont pas considérés, tous les sinistres étant clos. Le modèle fréquence-coût utilise la distribution des délais de paiement pour estimer la fréquence des sinistres survenus mais non déclarés :

$$\hat{\mu}_t = \frac{\int_0^t v(\tau) d\tau}{\int_0^t v(\tau) F(t - \tau) d\tau} r_t$$

Avec :

- r_t : sinistres connus sur $[0, t]$
- v : densité des fréquences
- F : distribution des délais de paiement

L'objectif est d'estimer $\hat{\mu}_t$, la fréquence des sinistres attendus déclarés avant t . Les coûts sont ramenés au coût moyen indexé sur l'inflation annuelle.

5.3. Approche simplifiée : Parodi 90%

Étant donné que 90 % des déclarations sont réglées en trois mois (voir figure 4), les provisions 2024 couvriront principalement les sinistres des trois derniers mois de 2023.

Une simplification de la méthode de Parodi consiste donc à estimer les fréquences et coûts sur cette période, puis à calculer les IBNR comme suit :

$$\text{IBNR} = \frac{\text{IBNR}_{90\%}}{0.9}$$

6. Calcul des provisions

6.1. Modélisation des provisions

Les IBNR sont modélisés selon la méthode de Parodi, avec un ajustement log-normal pour la distribution F des délais de paiement. La densité v des fréquences est estimée en intégrant la saisonnalité des déclarations.

Les provisions sont calculées une fois toutes les estimations réalisées.

6.2. Calcul des IBNR

Les deux provisions issues des méthodes de Parodi sont les plus proches des prestations réellement payées en fin d'année 2024.

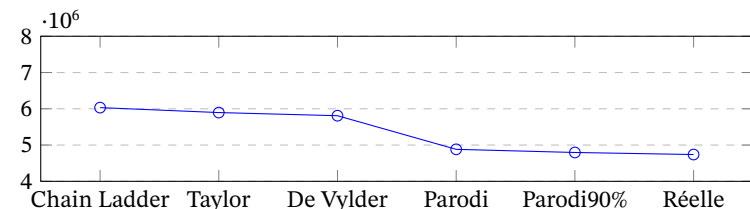


Figure 5. Provisions par méthodes de calcul

La méthode Parodi90% fournit de meilleures provisions, car en excluant les neuf premiers mois, elle réduit l'effet de saisonnalité, améliorant ainsi les estimations.

7. Conclusion

Très bon travail dans l'ensemble.