

Comparative Study of Technical Provisioning Methods in Health Insurance: An Analytical Approach Using the Chain Ladder Method, the Mack Method and Generalized Linear Models

Ayyoub Saoudi¹, Ghita Hajraoui², Jamal Zahi³

^{1,2,3}Université Hassan I^{er}, Faculté d'Economie et de Gestion, LM2CE, Settat, Maroc

E-mail: a.saoudi@uhp.ac.ma, ghita.hajraoui@gmail.com, zahi71@hotmail.com

Article history

Received Mar 21, 2024
Revised Dec 08, 2024
Accepted Jan 11, 2025
Published Feb 18, 2025

ABSTRACT

This paper presents an in-depth analysis of technical reserving in health insurance, using three different methods: the Chain Ladder method (deterministic), the Mack stochastic method (stochastic) and Generalized Linear Models (GLM). The aim of this study is to compare these methods in terms of forecast accuracy, robustness and ability to consider the uncertainty inherent in health insurance claims. The scientific contribution of this research lies in the proposal of methods for improving technical reserving, thereby improving the financial management of health insurance companies.

Keywords : Health insurance, Chain Ladder, Financial management, Uncertainty, Mack, Generalized Linear Models, Technical reserving.

I. INTRODUCTION

Les compagnies d'assurance se voient contraintes d'estimer correctement leurs réserves en utilisant des méthodes de provisionnement, qui ont évolué des méthodes déterministes aux méthodes d'apprentissage automatique grâce au progrès technologique et à la disponibilité des données massives. En particulier, au Maroc, il est crucial de mettre en place des méthodes d'apprentissage automatique pour optimiser le provisionnement technique de l'assurance maladie. Cependant, la question centrale est de savoir dans quelle mesure les méthodes d'apprentissage automatique sont plus performantes que les méthodes déterministes.

Les méthodes d'apprentissage automatique offrent plusieurs avantages pour l'estimation des réserves en assurance, notamment l'analyse de données permettant de prédire les résultats en identifiant les tendances et les modèles, la réduction de l'incertitude en utilisant des modèles sophistiqués prenant en compte les facteurs variables, et l'amélioration de la prise de décision en fournissant des informations à jour et en temps réel sur les risques et les tendances du marché. Il est important de noter que ces méthodes ne sont pas largement adoptées dans le secteur de l'assurance et que leur implémentation nécessite une évaluation des avantages potentiels par rapport aux coûts et aux défis.

II. MÉTHODOLOGIE

Le présent article met l'accent sur la différence entre les méthodes classiques dites déterministes, qui se basent sur la sinistralité passée sans prendre en considération les aléas liés au marché des assurances, et les méthodes d'apprentissage automatique conçues pour remédier à ce désavantage et pour assurer l'exactitude des provisionnements. Pour ce faire, nous avons entamé notre travail par l'application de la méthode déterministe Chain Ladder puis la méthode de Mack afin d'en prédire l'erreur. Par la suite, nous avons opté pour l'application du modèle linéaire généralisé (GLM) en tant que méthode d'apprentissage automatique à des fins de comparaison entre les deux approches précitées.

Les données ont été fournies par une compagnie d'assurance maladie désirant garder l'anonymat. Elles sont relevées de l'exercice de l'année 2019. Le traitement a été effectué moyennant le langage R.

III. TRIANGLE DE LIQUIDATION

En vue de créer un tableau de provisionnement pour les sinistres à payer (PSAP) en fonction de l'année d'occurrence des sinistres, avec une répartition mensuelle, les données relatives aux prestations maladie de l'exercice 2019 ont été extraites, comprenant les montants des prestations pour chaque mois de survenance, de janvier à décembre. Ce processus a été

essentiel pour élaborer le triangle de liquidation comme représenté sur le Tableau 1 :

TABLE I. TRIANGLE DE LIQUIDATION DE L'EXERCICE 2019.

M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
43272	46904	14572	4401,	1539,	2223,	4272,	1887,		1106,	5801,	962,6
,86	,89	,28	17	08	14	96	4	1312	6	05	8
11206	89619	10243	721,3	4130,	217,8	1192,	730,8	3416,		1524,	
,46	,7	,95	5	47		9	6	85	330		76
12083	45262	17852	730,4	3856,	2093,	15651	2867,	1071,		958,6	
,78	,74	,36		67	31	,04	61	06			
17728	47633	6478,	1639,	11949	15384	4143,	2276,	1651,			
,69	,7	2	81	,9	,8	92	54	72			
16799	49377	3981,	349,6	3309,	1969,		1488	2152,			
,17	,43	89		38	44		36				
12174	84121	3666,	7679,		4226,	1938,					
,06	,74	27	18		492,8	82	79				
12012	10951	6812,	346,2	907,0	1605,						
,17	0,9	48	1	4	41						
37575	57058	8929,	7892,	6786,							
,13	,16	15	98	6							
27461	86746	4146,	2594,								
,97	,66	51	64								
17686	53560	3471,									
,08	,12	41									
11204	35680										
,53	,86										
6367,											
98											

L'élaboration du triangle de liquidation est une étape essentielle dans la construction du triangle cumulé. Ce dernier sert de base de calcul pour toutes les méthodes de provisionnement. Il est obtenu en additionnant, pour chaque ligne, le montant progressif des mois précédents au montant du mois en cours. Cela permet de calculer le montant cumulatif pour chaque mois de développement, conduisant ainsi au résultat final synthétisé dans le Tableau 2 :

TABLE II. TRIANGLE CUMULÉ DE LIQUIDATION.

M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
43272	90177	10475	10915	11069	11291	11718	11907	12038	12149	12729	12825
		0	1	0	3	6	3	5	2	3	6
11206	10082	11107	11179	11592	11613	11733	11806	12148	12181	12333	
	6	0	1	1	9	2	3	0	0	5	
12083	57346	75198	75929	79785	81879	97530		10039	10146	10242	
								7	8	7	
17728	65362	71840	73480	85430	10081	10495	10723	10888			
					5	9	5	7			
16799	66176	70158	70508	73817	75786	77274	79427				
12174	96295	99902	10764	10813	11236	11429					
			1	4	0	9					
12012	12152	12833	12868	12958	13119						
	3	5	1	8	4						
37575	94633	10356	11145	11824							
		2	5	2							
27461	11420	11835	12094								
	8	5	9								
17686	71246	74717									
11204	46885										
6367											

IV. RÉSULTATS ET DISCUSSION

A. Chain ladder

E. Astesan a développé en 1938, la technique des triangles de liquidation, aboutissant ainsi à la méthode très répandue parmi les actuaires, dite de Chain-Ladder. [1]

Cette méthode compte parmi les plus utilisées par les actuaires, vue sa simplicité de compréhension et de mise en place. La méthode de Chain Ladder est basée sur les facteurs de développement. Elle suppose que ces facteurs sont constants pour l'ensemble des années.

Le facteur de développement individuel est donné par :

$$\lambda_j =, \text{ tel que } j= 0, \dots, n-1 \quad (1)$$

Cette méthode se base sur l'hypothèse d'indépendance des facteurs de développement λ_j et les années de survenance.

La Figure 1 ci-après illustre les données du Tableau 2 à l'aide de la bibliothèque R Chain Ladder du paquet CRAN (notez que dans ces graphiques, les indices de l'année de développement j sont décalés par défaut de 1).

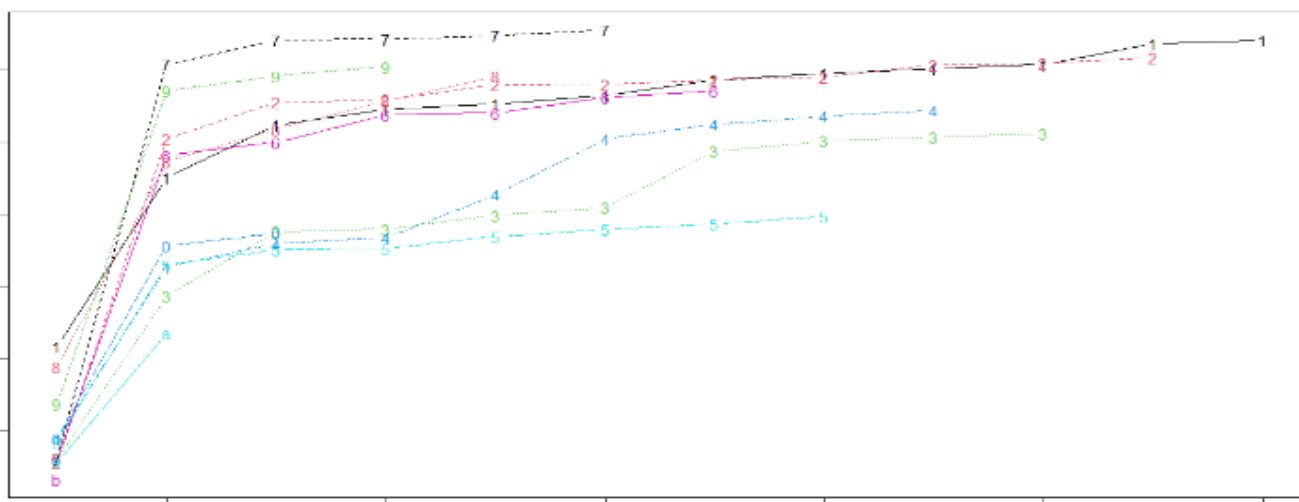


Figure 1: Développement des sinistres par mois de survenance.

TABLE III. COEFFICIENTS DE PASSAGE.

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
4,2184	1,0913	1,0298	1,0418	1,0394	1,0478
λ_7	λ_8	λ_9	λ_{10}	λ_{11}	
1.0192	1.0167	1.006979	1.0301	1.0076	

Les facteurs incrémentaux calculés permettent d'estimer le triangle inférieur et ainsi d'obtenir un tableau de liquidation complet et cohérent. Les résultats d'estimation sont donnés par le tableau 4 ci-dessous :

TABLE IV. TRIANGLE DE LIQUIDATION COMPLETE PAR LA METHODE CHAIN LADDER.

M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
43272	90177	104750	109151	110690	112913	117186	119073	120385	121492	127293	128256
11206	100826	111070	111791	115921	116139	117332	118063	121480	121810	123335	124268
12083	57346	75198	75929	79785	81879	97530	100397	101468	102427	105511	106309
17728	65362	71840	73480	85430	100815	104959	107235	108887	109646	112948	113802
16799	66176	70158	70508	73817	75786	77274	79427	80757	81321	83770	84403
12174	96295	99962	107641	108134	112360	114299	116502	118454	119281	122872	123802
12012	121523	128335	128681	129588	131194	137467	140117	142465	143459	147779	148897
37575	94633	103562	111455	118242	122902	128779	131262	133461	134392	138439	139486
27461	114208	118355	120949	126005	130971	137235	139880	142224	143216	147529	148645
17686	71246	74717	76946	80163	83322	87307	88990	90481	91113	93856	94566
11204	46885	51166	52693	54896	57059	59788	60940	61961	62394	64273	64759
6367	26858,66	29311	30185	31447	32687	34250	34910	35495	35743	36819	37098

L'estimation de la partie inférieure du triangle se fait par le biais des facteurs de développements calculés auparavant, il s'agit de multiplier les montants des diagonales par le facteur de passage correspondant, ci-dessous un exemple de calcul : La première diagonale correspondant au mois de développement de janvier est multipliée par le premier facteur de passage λ_1 .

Le triangle de liquidation complété permet d'estimer les montants des provisions mensuelles. Plaçons-nous sur la dernière colonne, cette colonne constitue la charge ultime mensuelle pour chaque mois de survenance. Pour avoir le montant de la PSAP mensuelle on en retranche le montant des règlements sur la diagonale correspondante.

TABLE V. PSAP ESTIMEE PAR LA METHODE CHAIN LADDER.

Mois	PSAP (DHS)	Charges ultimes (DHS)
Janvier	0	128 256
Février	933,06	124 268,06
Mars	3 882,37	106 309,37
Avril	4 915,91	113 802,91
Mai	4 976,75	84 403,75
Juin	9 503,21	123 802,21
Juillet	17 703,51	148 897,51
Août	21 244,74	139 486,74
Septembre	27 696,23	148 645,23
Octobre	19 849,56	94 566,56
Novembre	17 874,27	64 759,27
Décembre	30 731,16	37 098,16
Total	159 310,8	1 314 295,77

Le montant global de la provision pour sinistres à payer est constitué par le cumul des PSAP mensuelles, indiqué dans la dernière ligne du Tableau 4, ce montant avoisine les 159 310.8 Dhs.

B. Méthode stochastique de Mack

Modèle de Mack ou distribution free Chain-Ladder, introduit en 1993 par Mack [2] rend aléatoire le modèle Chain Ladder en y ajoutant des indicateurs de risque de prédiction, il est non-paramétrique dans la mesure où aucune hypothèse de distribution n'est faite sur les composantes du triangle.

La méthode de Mack est la première méthode qui fait intervenir la dimension stochastique dans le modèle déterministe de Chain Ladder. Ceci rend possible la mesure de l'incertitude liée à la prédiction du montant des provisions.

TABLE VI. PSAP ESTIMEE PAR LA METHODE MACK.

	Totaux
Latest:	1 154 985.00
Dev:	0.88
Ultimate:	1 314 295.78
PSAP:	159 310.78
Mack.S.E	53 967.39
CV(PSAP):	0.34

Après avoir calculé le provisionnement moyennant la méthode de Mack (résultats synthétisés dans le Tableau 5) qui a pour objectif d'estimer l'erreur liée au provisionnement obtenu par la méthode Chain Ladder, nous avons constaté que le modèle est fiable à hauteur de 66%. Chose qui nous a amenés à trouver une méthode plus performante et plus précise grâce aux méthodes d'apprentissage automatique.

C. Modèle linéaire généralisé (GLM)

Les modèles linéaires généralisés (GLM) sont devenus couramment utilisés dans divers domaines statistiques depuis leur introduction par J. Nelder et R. Wedderburn en 1972 [2] comme une extension du modèle linéaire normal. À l'origine utilisés dans les années 1980 pour la tarification des risques, ils ont depuis été largement adoptés. Leur application dans les années 1990 pour la détermination stochastique des provisions pour sinistres a été particulièrement fructueuse ainsi que Zhou et Garrido en 2009 [3] qui ont appliqué le GLM pour calculer le provisionnement individuel. La mise en œuvre pratique de ces modèles repose sur des procédures spécialisées intégrées aux logiciels statistiques standards. » [4].

Les techniques de calcul des provisions pour sinistres à payer basées sur les principes des modèles linéaires généralisés ont été largement développées et adoptées. Les facteurs explicatifs à tenir en compte sont de trois types, à savoir les facteurs lignes, les facteurs colonnes et les facteurs diagonaux, comme illustré sur le Tableau 6.

TABLE VII. COEFFICIENTS DU MODELE GLM.

	Coefficients				
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	10.010708	0.002364	4233.760	< 2e-16	***
Lig2	1.439460	0.002376	605.937	< 2e-16	***
Lig3	1.526925	0.002378	642.232	< 2e-16	***
Lig4	1.558583	0.002398	649.932	< 2e-16	***
Lig5	1.603694	0.002434	658.827	< 2e-16	***
Lig6	1.642557	0.002475	663.532	< 2e-16	***
Lig7	1.697668	0.002533	670.197	< 2e-16	***
Lig8	1.723582	0.002608	660.976	< 2e-16	***
Lig9	1.748288	0.002676	653.385	< 2e-16	***
Lig10	1.734769	0.002816	616.051	< 2e-16	***
Lig11	1.731024	0.003026	572.036	< 2e-16	***
Lig12	1.751075	0.003659	478.576	< 2e-16	***
Col2	-0.006318	0.001306	-4.838	1.31e-06	***
Col3	-0.292723	0.001475	-198.415	< 2e-16	***
Col4	-0.228371	0.001519	-150.333	< 2e-16	***
Col5	-0.407569	0.001698	-239.971	< 2e-16	***
Col6	-0.031679	0.001609	-19.686	< 2e-16	***
Col7	0.168689	0.001629	103.585	< 2e-16	***
Col8	0.067325	0.001828	36.834	< 2e-16	***
Col9	0.160502	0.001981	81.037	< 2e-16	***
Col10	-0.289896	0.002754	-105.281	< 2e-16	***
Col11	-0.693171	0.004372	-158.533	< 2e-16	***
Col12	-1.251825	0.012753	-98.156	< 2e-16	***
Signif. Codes :	0	****	0.001	***	0.01
				**	0.05
				*	0.1
				.	1

Sous le logiciel R le triangle de liquidation est automatiquement calculé selon la requête donnée par les

annexes, le résultat est illustré par le **Erreur ! Source du renvoi introuvable.7** :

TABLE VIII. TRIANGLE DE LIQUIDATION COMPLETE PAR LA METHODE DE GLM.

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]
[1,]	22263.6	93917.2	102501.6	105798.5	110680.42	115066.52	121585.90	124777.88	127898.97	126181.65	125709.87	128256.00
[2,]	22123.4	93325.7	101856.03	105132.1	109983.34	114341.81	120820.13	123992.01	127093.45	125386.94	124918.13	127448.23
[3,]	16613.7	70083.7	76489.7	78949.9	82592.94	85865.97	90730.92	93112.87	95441.92	94160.41	93808.35	95708.34
[4,]	17718.0	74742	81573.7	84197.4	88082.65	91573.23	96761.54	99301.81	101785.66	100418.97	100043.51	102069.80
[5,]	14811.2	62479.9	68190.8	70384.1	73631.88	76549.81	80886.93	83010.44	85086.80	83944.32	83630.46	85324.32
[6,]	21569.4	90988.6	99305.4	102499.4	107229.16	111478.49	117794.58	120887.03	123910.80	122247.03	121789.96	124256.69
[7,]	26354.6	111174.6	121336.4	125239.1	131018.13	136210.18	143927.51	147706.02	151400.62	149367.74	148809.27	151823.25
[8,]	23814.1	100457.8	109640.1	113166.5	118388.49	123080.05	130053.46	133467.74	136806.19	134969.28	134464.64	137188.09
[9,]	26139.7	110268.2	120347.2	124217.99	129949.91	135099.63	142754.04	146501.74	150166.22	148149.91	147595.99	150585.40
[10,]	16660.8	70282.1	76706.16	79173.3	82826.71	86109.01	90987.73	93376.42	95712.06	94426.92	94073.87	95979.24
[11,]	11131.5	46957.5	51249.6	52898	55338.90	57531.89	60791.51	62387.46	63947.97	63089.33	62853.44	64126.48
[12,]	6367	26858.7	29313.7	30256.5	31652.66	32907.01	34771.44	35684.29	36576.87	36085.74	35950.82	36678.97

L'estimation de la partie inférieure du rectangle de liquidation nous a permis d'avoir les résultats qui suivent :

TABLE IX. ESTIMATION DE LA PSAP ET LES CHARGES ULTIMES PAR LA METHODE GLM.

Mois	Charge Ultime	PSAP
Janvier	128 256	0
Février	127 448,2	2 530,1
Mars	95 708,34	1 547,93
Avril	102 069,8	284,14
Mai	85 324,32	2 313,88
Juin	124 256,7	6 462,11
Juillet	151 823,3	15 613,07
Août	137 188,1	18 799,6
Septembre	150 585,4	26 367,41
Octobre	95 979,24	19 273,08
Novembre	64 126,48	17 169,02
Décembre	36 678,97	30 311,97
Total	1 299 445	140 672,3

Le Tableau 8 présente les charges ultimes de chaque exercice aussi que le montant de la PSAP à constituer pour chaque exercice.

D. Discussion

Après avoir utilisé la méthode classique de Chain Ladder, la plus couramment utilisée dans le contexte marocain, pour calculer les réserves techniques, nous avons obtenu un montant de 159 310,8 Dhs. Cependant, cette méthode n'a pas fourni la valeur de l'erreur due à ce calcul. Par conséquent, nous avons eu recours à la méthode de Mack, qui a indiqué un taux d'erreur (CV(PSAP)) de 34%, un pourcentage élevé par rapport à l'importance de la précision requise lors du calcul du provisionnement. Afin de remédier à ce problème, nous avons fait appel à la méthode GLM, qui a estimé le montant de la PSAP à 140 672,3 Dhs avec une précision de 95%. Cette précision a permis de résoudre le problème de sur-provisionnement, qui impacte la solvabilité de la compagnie d'assurance. Ces résultats mettent en évidence les risques liés à l'utilisation des méthodes classiques pour le provisionnement, qui ne prennent pas en considération les aléas liés au marché des assurances. Par conséquent, il est nécessaire d'avoir recours aux méthodes d'apprentissage automatique pour garantir la capacité de la compagnie d'assurance à faire face à ses engagements futurs envers ses assurés, sans dépasser le montant requis dédié aux éventuelles indemnités.

V. CONCLUSION

Notre étude comparative entre la méthode déterministe Chain Ladder et la méthode d'apprentissage automatique basée sur le modèle linéaire généralisé met en évidence l'avantage de cette dernière. En utilisant les données de l'exercice 2019 provenant d'une compagnie d'assurance, nous avons pu constater que le modèle GLM offre une estimation plus précise des réserves, surpassant ainsi la méthode traditionnelle Chain Ladder. Cette performance supérieure du modèle GLM peut être attribuée à sa capacité à prendre en compte les subtilités complexes du marché des assurances, garantissant ainsi une approche plus adaptative et précise pour l'optimisation du provisionnement technique.

L'intégration des méthodes d'apprentissage automatique, notamment le GLM, dans le processus de provisionnement technique en assurance au Maroc, se révèle, donc, non seulement primordiale mais également bénéfique. La transition vers des approches plus sophistiquées et flexibles est essentielle pour répondre efficacement aux défis actuels du secteur, offrant ainsi aux compagnies d'assurance une meilleure maîtrise de leur solvabilité et une estimation plus fiable de leurs réserves.

REFERENCES

- [1] Astesan, E. (1938). *Les réserves techniques des sociétés d'assurances contre les accidents d'automobiles*. Librairie générale de droit et de jurisprudence.
- [2] Mack, T. (1993). Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin*, 23, 213-225.
- [3] Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 135(3), 370-384.
- [4] Zhou, J., & Garrido, J. (2009). A loss reserving method based on generalized linear models. *Society of Actuaries*.
- [5] Partrat, P. (2007). *Provision technique en assurance non-vie, perspectives actuarielles modernes*. Paris: Economica.
- [6] Denuit, M., & Charpentier, A. (2005). *Mathématiques de l'Assurance Non-Vie. Tome II: Tarification et Provisionnement*. Economica.
- [7] Thomas, M. (1993). Measuring the variability of chain ladder reserve estimates. *Casualty Actuarial Society E-Forum, Spring, Vol. 1*, 101-182.
- [8] England, P., & Verrall, R. (2001). A Flexible Framework for Stochastic Claims. *Casualty Actuarial Society*, 1-38.
- [9] Bențe, C. C. (2017). Actuarial Estimation Of Technical Reserves In Insurance Companies. Basic Chain Ladder Method. *The Annals of the University of Oradea*, 227.
- [10] Belkacem-Nacer, M., & Latreche, A. (2015). Estimation de l'incertitude du montant de provisions pour sinistre à payer: Modèle de Mack "cas de la SAA". *Les Cahiers du CREAD*, 112, 85-112.