Provisionnement non vie

11 mai 2021

M1 Actuariat, année 2020 - 2021

Durée : 1h30

La calculatrice est autorisée ainsi qu'une feuille **manuscrite** *seulement recto.*

Toutes les réponses doivent être soigneusement justifiées.

Exercice Nous considérons le triangle de paiements de sinistres incrémentaux, effectués par une compagnie d'assurance pour son portefeuille Corps maritimes au titre des exercices de 1984 à 1991 :

	1	2	3	4	5	6	7	8
1	1381	4399	4229	435	465	205	110	67
2	859	6940	2619	1531	517	572	287	
3	6482	6463	3995	1420	547	723		
4	2899	16428	5521	2424	477			
5	3964	15872	8178	3214				
6	6809	24484	27928					
7	11155	38229						
8	10641							

Le triangle des paiements cumulés est également disponible :

	1	2	3	4	5	6	7	8
1	1381	5780	10009	10444	10909	11114	11224	11291
2	859	7799	10418	11949	12466	13038	13325	
3	6482	12945	16940	18360	18907	19630		
4	2899	19327	24848	27272	27749			
5	3964	19836	28014	31228				
6	6809	31293	59221					
7	11155	49384						
8	10641							

Dans un premier temps, nous explorons la possibilité d'utiliser quelques méthodes déterministes afin d'effectuer un calcul de provision.

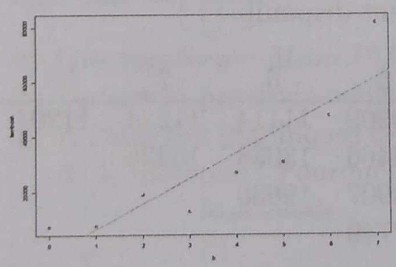
a) Avant d'appliquer la méthode de Chain-Ladder, nous avons calculé le d-triangle ainsi que des statistiques élémentaires par colonnes :

	1	2	3	4	5	6	7
1	4,1854	1,7317	1,0435	1,0445	1,0188	1,0099	1,0060
2	9,0792	1,3358	1,1470	1,0433	1,0459	1,0220	
3	1,9971	1,3086	1,0838	1,0298	1,0382		
4	6,6668	1,2857	1,0976	1,0175			
5	5,0040	1,4123	1,1147				
6	4,5958	1,8925					
7	4,4271						
$\hat{\mu}$	5,1365	1,4944	1,0973	1,0338	1,0343	1,016	
$\hat{\sigma}$	2,0515	0,2326	0,0342	0,01103	0,0114	0,0061	
$\hat{CV}$	0,399	0,156	0,031	0,011	0,011	0,006	

Au vu de cela, l'application de cette méthode vous semble-t-elle envisageable ? Oui ? Non ? Pourquoi ?

b) Y aurait-il un variant de la méthode Chain Ladder qui permettrait d'améliorer les résultats que l'on aurait obtenu par Chain Ladder ? Sans l'implémenter, expliquer la démarche de la méthode proposée et la raison pour laquelle, selon vous, elle devrait donner de meilleurs résultats.

c) Nous envisageons maintenant de calculer le montant de provision globale par la méthode de séparation arithmétique de Verbeek-Taylor. Pour cela, nous décidons d'extrapoler de façon linéaire les facteurs d'inflation $(\lambda_h)_{h \geq n+1}$ selon le modèle $\lambda_h = ah + b$. Les paramètres a et b ont été estimés par régression linéaire sur les valeurs antérieures de λ_h et nous avons obtenu : $\hat{b} = -2890$ et $\hat{a} = 9142$. Dans le graphique ci-dessous, nous représentons les données avec la droite de régression estimée (en rouge).



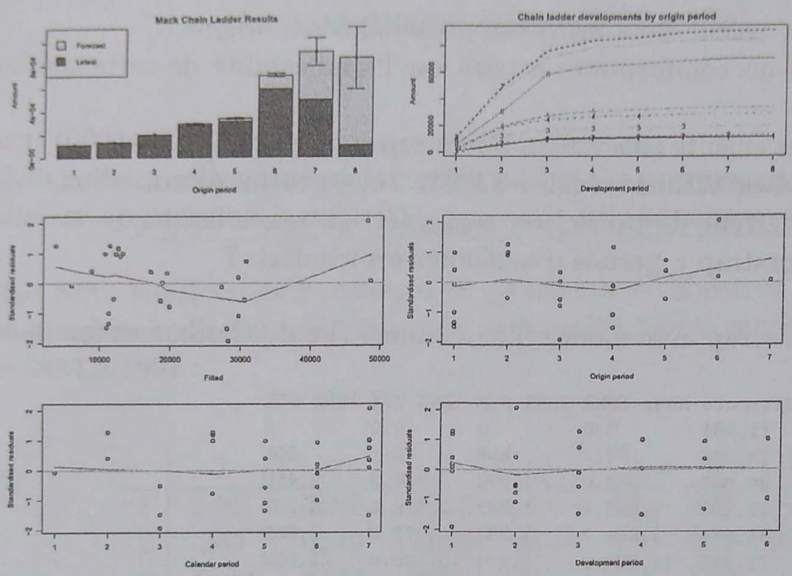
Je ne vous demande pas d'implémenter la méthode mais juste de répondre à cette question : à votre avis, aurait-il été possible d'améliorer les résultats de cette méthode ? Si oui, comment ?

Dans un deuxième temps, nous estimons plusieurs modèles stochastiques à l'aide du logiciel R. Les résultats sont présentés ci-dessous :

1) modèle de Mack (*N.B. les colonnes* f_j *et* σ_j^2 *ont été rajoutées*)

	Latest	Dev.To.Date	Ultimate	IBNR	Mack.S.E	CV(IBNR)	f_j	sigma^2_j
1	11,291	1.000	11,291	0.0	0.0	NaN	4.362693	1.214391e+04
2	13 325				32.3			1.338376e+03
3	19,630	0.978	20,072	441.8	183.6	0.415	1.100012	1.640612e+01
4	27,749	0.944	29,380	1,631.1	410.7	0.252	1.029489	2.852305e+00
5	31,228	0.917	34,039	2,810.6	596.0	0.212	1.035476	2.265839e+00
6	59,221	0.834	71,007	11,785.9	1,728.8	0.147	1.016438	8.806148e-01
7	49,384	0.541	91,248	41,864.3	12,155.7	0.290	1.005969	3.572896e-02

```
8 10,641        0.124    85,778 75,137.0 26,779.8      0.356   1.000000
                Totals
Latest:    222,469.00
Dev:             0.62
Ultimate:  356,219.13
IBNR:      133,750.13
Mack.S.E    31,223.77
CV(IBNR):        0.23
```



i) Compléter la ligne 2 du tableau en détaillant le calcul de chacune des quantités manquantes et en donnant une interprétation de ces mêmes quantités.

ii) Commenter globalement les résultats du modèle de Mack, à l'aide de la table et de chacun des graphiques ci-dessus (10 lignes maximum).

2) modélisation Log Poisson surdispersée (loi de Poisson surdispersée et fonction lien ln). Nous avons obtenu le tableau ci-dessous : $\mathbb{E}[\hat{R}_3] =$

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.2447     0.3083  23.498  < 2e-16 ***
anF2          0.1716     0.3627   0.473  0.64105
anF3          0.5753     0.3358   1.713  0.10142
anF4          0.9563     0.3186   3.002  0.00679 **
anF5          1.1035     0.3140   3.514  0.00206 **
anF6          1.8388     0.2954   6.224 3.57e-06 ***
anF7          2.0896     0.3048   6.857 8.89e-07 ***
anF8          2.0278     0.4128   4.912 7.37e-05 ***
devF2         1.2127     0.1761   6.888 8.30e-07 ***
devF3         0.8588     0.2048   4.193  0.00041 ***
devF4        -0.3969     0.3450  -1.150  0.26287
devF5        -1.5229     0.6584  -2.313  0.03095 *
devF6        -1.3090     0.7588  -1.725  0.09920 .
devF7        -2.0434     1.4406  -1.418  0.17073
devF8        -3.0400     3.4725  -0.875  0.39123
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
(Dispersion parameter for quasipoisson family taken to be 801.5309)
```

ce qui nous a permis de calculer les quantités suivantes :

```
     R_i  sep(R_i)  CV(R_i)
1      0         0      NaN
2     80       354     4.43
3    442       818     1.85
4   1631      1575     1.38
```

5	2811	2052	0.73
6	11786	4747	0.40
7	41864	9873	0.24
8	75137	24332	0.32

R	133750
sep(R)	30460.9
CV(R)	0.228

a) Détailler le calcul de la provision pour l'année d'origine 3.

b) Avez-vous un commentaire à faire sur l'applicabilité de cette méthode sur ce triangle ?

quasi-poisson donne comme Mack

c) Nous avons ensuite effectué un Bootstrap sur ce modèle (B=5000) et nous avons obtenu un best estimate égale à 136042.9 et un risque d'estimation de 564676602. Calculer l'erreur de prédiction $sep_{boot}(\hat{R})$ et le coefficient de variation. Est-ce que le Bootstrap a permis d'améliorer les résultats ?

3) méthode Bootstrap avec modèle Log Gamma (loi de Gamma et fonction lien ln) :

	Latest	Mean Ultimate	Mean IBNR	IBNR.S.E	IBNR 75%	IBNR 95%
1	11,291	11,291	0.0	0	0.0	0
2	13,325	13,398	73.3	496	43.1	805
3	19,630	20,083	453.0	1,023	730.8	2,410
4	27,749	29,409	1,660.4	1,842	2,538.5	5,237
5	31,228	34,077	2,849.2	2,337	4,078.0	7,352
6	59,221	71,155	11,934.3	5,191	15,003.6	21,552
7	49,384	91,801	42,417.0	10,544	48,617.0	60,893
8	10,641	87,902	77,260.6	26,554	89,424.4	128,422

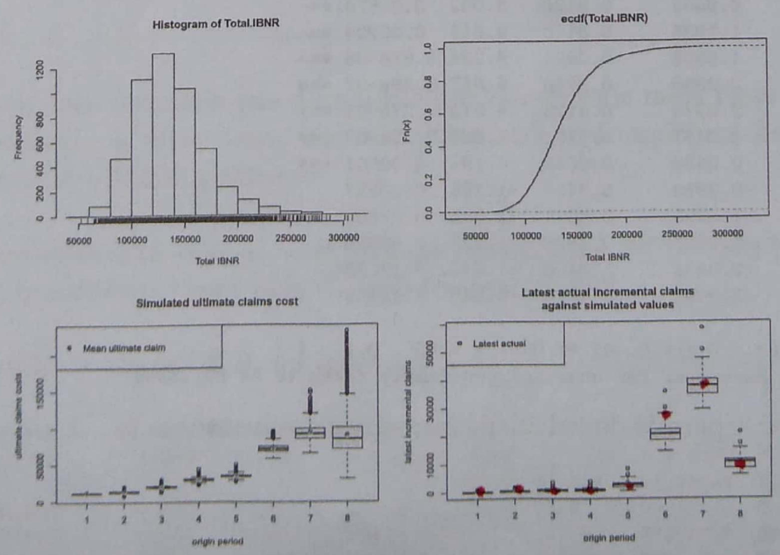
	Totals
Latest:	222,469
Mean Ultimate:	359,117
Mean IBNR:	136,648
IBNR.S.E	33,281
Total IBNR 75%:	154,235
Total IBNR 95%:	198,805

moyenne pos...

a) Que représente *Mean IBNR* (3ème colonne) ?

b) Donner la provision suffisante dans 95% des cas.

c) À l'aide des graphiques ci-dessous, que pouvez-vous conclure sur la performance de la méthode ? Pourquoi ?



4) Au finale, au vu des résultats obtenus, quelle méthode stochastique auriez-vous adopté ? Pourquoi ?

celle avec la plus petite deviance dispersion pourrait être au plus juste

Bootstrap n'est pas bon pour les last actual incrément

Exercice :

a) Compte tenu des résultats du d-triangle, l'application de la méthode de Chain-Ladder ne semble pas envisageable car pour accepter cette méthode avec le d-triangle on attend à ce que les facteurs de développem^ts soient constants par colonne.

Or ici, ça n'est clairement pas le cas.

b) Nous pourrions opter pour la méthode Chain Ladder Pondérées.

Cette méthode consiste à accorder ± d'importance à certaines années par des poids que l'on attribue aux facteurs de dvpt individuel f_{ij} lors du calcul des f_j.

$$f_j = \frac{\sum_{i=0}^{m-j-1} f_{ij} w_{ij}}{\sum_{i=0}^{m-j-1} w_{ij}} \qquad j \in [\![0, m-1]\!].$$

c) Oui, on peut améliorer les résultats de la méthode en utilisant un ajustement exponentiel de $\hat{\lambda}_k$ plutôt qu'un linéaire.

1) i) Ligne 2 :

Latest : 13325 ← dernier montant connu pour les sinistres survenus l'année 2.

Dev. to. date : $\frac{\text{Latest}}{\text{Ultimate}} = 0,9941$ ← Cadence de règlement pour l'année 2.

Ultimate : 13 404,54 ← Charge ultime de l'année 2 estimée par le modèle de Mack.

IBNR : Ultimate - Latest = 79,54 ← Provis° estimée pour l'année 2.

$CV(IBNR) = \frac{\text{Mack.S.E}}{\text{IBNR}} = 0.4061$ ← Coeff de variat°.

$$f_j = \frac{\sum_{i=0}^{m-j-1} C_{i,j+1}}{\sum_{i=0}^{m-j-1} C_{ij}} \Rightarrow f_2 = 1,5403939$$ ← facteur de dvpt pour l'année 2

ii) OK

2) a) $\mathbb{E}[\hat{R}_3] = e^{\hat{\mu}+\hat{\alpha}_3}\left(e^{\hat{\beta}_7} + e^{\hat{\beta}_8}\right)$

$$= e^{7.2447+0.5753}\left(e^{-2.0434} + e^{-3.0400}\right)$$

$$\approx 442$$

b) Les p.value de anF2, anF3, devF4, devF6, devF7 et devF8 > 5%
⟹ elles ne sont pas significatives

⟹ Attention à ce que ces variables n'influent pas trop.

c) $Sep_{boot}(\hat{R}) = \sqrt{MSE_{boot} + Var_{boot}}$

$$= \sqrt{564676602 + 136042,9}$$

$$= 23\ 765,79$$

$$CV = \frac{Sep_{boot}}{R} = \frac{23\ 765,79}{136042,9} = 0,1747$$

↳ meilleure estimation car Sep et CV sont plus faible par bootstrap.

3) a) Mean IBNR correspond à la provision moyenne

b) La provision suffisante dans 95% des cas est : 198 805

c) Méthode relativement performante mais attention à l'année 6 pour les Latest actual incremental Claims. moyenne en dehors du boxplot.
De +, pour les années le + récentes il y a une forte variance.

4) Pour Mack :

* CV = 0,23
* $\hat{R}$ = 133750,13
* OK mais attent° aux années les + anciennes.

Pour Log-Poisson surdispersé :

* CV = 0,228
* R = 133750
* Des Var non signif
* CV des années les + anciennes très élevé

Pour Bootstrap Log-Gamma :

* CV = 0,24
* R = 136 648
* CV des années les + anciennes élevés

Pour Bootstrap Log-Poisson surdispersé :

* CV = 0,1747
* R = 136 043

→ On choisit Bootstrap Log-Poisson car le CV est le plus faible ⇒ proche de la réalité et le R permet de rester prudent p/r aux autres méthodes.