

Techniques stochastiques de provisionnement sur données agrégées en IARD

Parcours IMSA
Master 2 MAS, 2023-2024

Xavier Milhaud

Maitre de Conférences, Aix-Marseille Université
Actuaire certifié

- 1 Projection de triangles de liquidation : introduction
- 2 Approches déterministes du provisionnement
- 3 Approches stochastiques de provisionnement
- 4 Distributions prédictives et mesures de risque
- 5 Cas particuliers
- 6 Calcul de SCR provisionnement à un an
- 7 Extensions de ces modèles

Organisation

Le cours se répartit comme suit :

- 10h de cours magistral ;
- 10h de travaux dirigés en salle machine ;
- Une évaluation (examen écrit et/ou projet).

L'objectif est de connaître les **techniques statistiques** utiles au provisionnement et **leurs limites**.

Le livre français de référence est [Partrat et al., 2007].

Objectif du cours

Aujourd'hui, Solvabilité 2 redéfinit le cadre prudentiel et promeut

- la mise en place d'un bilan économique,
- le calcul de mesures de risque sophistiquées.

Les provisions sont le socle même d'une évaluation prudentielle.

Nous désirons dans ce cours :

- ① **présenter les outils statistiques fondateurs** des derniers développements sur les méthodes de provisionnement,
- ② **poser les hypothèses** requises par chaque technique et en acquérir un regard critique.
- ③ **évaluer leurs principales \neq .**

1 Projection de triangles de liquidation : introduction

- Généralités
- Données de travail
- Résultats d'intérêt pour le praticien
- Traitement de phénomène parasite
- Prise en compte de réassurance

“RESERVING IS AN ART”

Utilité et visibilité du perfectionnement des méthodes

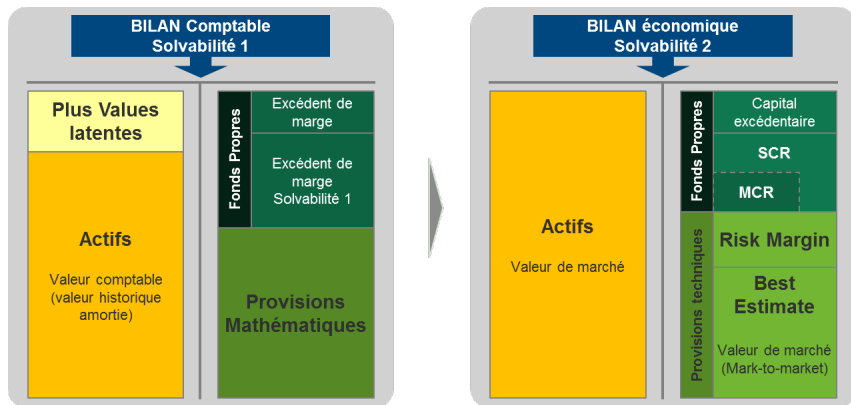
Suite la survenance d'un sinistre, l'assureur a une dette envers son assuré. Pour couvrir cet éventuel événement, cette dette figure au passif par la constitution d'une provision.

Toutes les techniques récentes visent à améliorer la fiabilité des calculs de provisionnement, utiles pour

- la certification du **best estimate**,
- la justification de la **suffisance du provisionnement**,

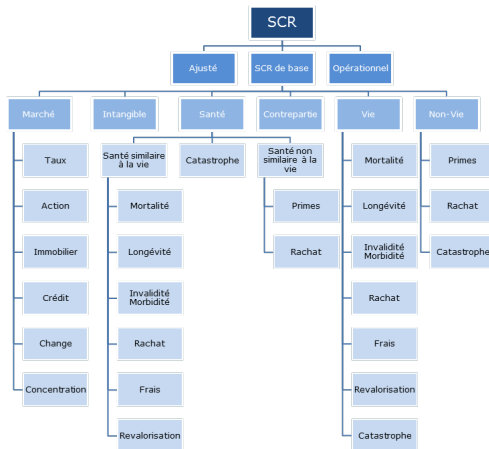
auprès des commissaires contrôleurs, auditeurs externes...

Contexte : bilan assureur ds Solva 2 - risque de réserve



Interactions actif-passif (PB, comportements, ...) \Rightarrow BEL varie !
Le SCR provision s'estime à la fin, sur la base de ce bilan.

SCR/BEL et structure modulaire - Solvency II



Formule stand. : matrice de corrél. agrège $SCR_k \Rightarrow$ SCR global.
 SCR prov. s'estime sur BEL \Rightarrow après décompos. des BEL/risque.

Problématique

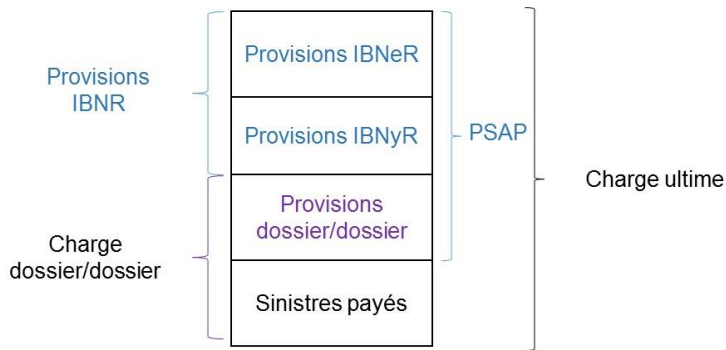
La problématique du reserving est très ancienne (+sieurs siècles).
Il existe maintenant

- une grande variété de méthodes (micro-level VS aggregated),
- des outils de gestion toujours plus réactifs et perfectionnés,
- beaucoup d'approximations pour le calcul de \neq marges :
 - marge de précaution, marge pour risque croissant,
 - marge pour dépréciation (des actifs)...

→ La prise en compte du caractère aléatoire de ces montants devient plus que nécessaire. Notamment sous la pression...

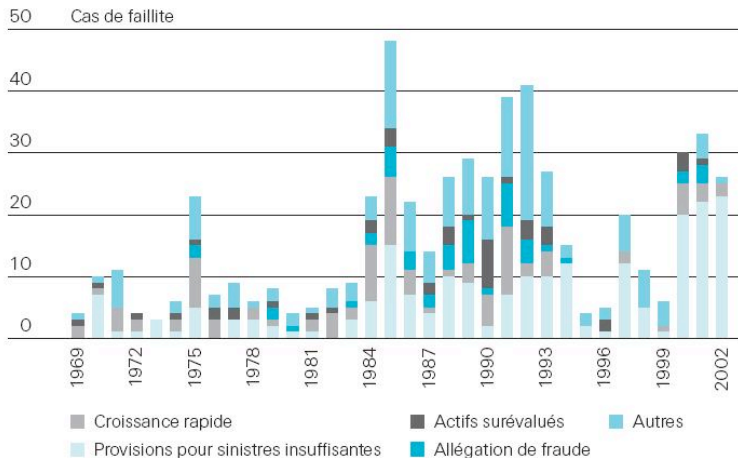
- ...des forces de marché,
- ...à cause de l'aversion au risque des administrations.

Décomposition de la charge d'un sinistre



Méthodes de provisionnement les + utilisées → rapport ASTIN.

Importance du calcul des PSAP - Faillites des assureurs dommages US (RC et accidents)



Source : A.M. Best: Best's Insolvency Study, Property/Casualty U.S. Insurers 1969-2002, mai 2004, p.34

Données type : triangle de liquidation

Au 31/12/l, les données sont répertoriés comme suit :

Année de survenance	Années de développement								
	0	1	...	j	...	$J - i$...	$J - 1$	J
0	$x_{0,0}$	$x_{0,1}$...	$x_{0,j}$	$x_{0,J-1}$	$x_{0,J}$
1	$x_{1,0}$	$x_{1,1}$...	$x_{1,j}$				$x_{1,J-1}$	
\vdots	\vdots	\vdots	...	\vdots	...		\ddots		
i	$x_{i,j}$...	$x_{i,J-i}$			
\vdots	\vdots	\vdots		\vdots					
$l - j$	$x_{l-j,j}$					
\vdots									
$l - 1$	$x_{l-1,0}$	$x_{l-1,1}$							
l	$x_{l,0}$								

Contexte général des études

L'hypothèse fondamentale de tous ces dvp est que les triangles de liquidation sont **fiables** : données déjà traitées en amont !

Considérations fondamentales à ne pas oublier :

- traitement de l'inflation (vision "as if"),
- traitement de données extrêmes (dans un souci de stabilité, séparation sinistres attritionnels/graves/CAT),
- traitement de données incomplètes : doit-on s'en servir ?

Méthodologies classiques pour évaluer les PSAP :

- 1 dossier-dossier (*case estimate*) par le gestionnaire sinistre.
- 2 et IBNR ou "tardifs".

Comparaison et performance des méthodes

Depuis 1991 en assurance, on peut utiliser des méthodes stats pour évaluer les sinistres survenus au cours 2 derniers exercices.

- La **convergence** de ces méthodes avec l'estimation dossier-dossier permet de vérifier leur **fiabilité**.
- Une **divergence** doit prévenir l'**instabilité** des données.

Différence : évaluation "historique" dossier par dossier = prospectif \neq évaluation par méthodes stats = données historiques.

Ces méthodes sont d'autant plus performantes que

- passé **régulier**, présent/futur **structurellement peu \neq du passé**,
- branche **peu volatile**, données **nombreuses** et fiables.

Nécessité d'une réflexion préalable

La qualité de l'estimation de la provision repose sur la qualité et la stabilité des données : comment ont-elles été générées ? quels sont les facteurs de risque d'un éventuel changement ?

① Les facteurs internes :

- évolution du portefeuille,
- politique de souscription, tarification et réassurance,
- politique de gestion des sinistres (cadence de règlement).

② Les facteurs externes :

- pratiques de marché,
- cycles économiques, inflation des montants de sinistres,
- évolution de la sinistralité (fréquence, sévérité),
- modifications réglementaires et comptables.

1 Projection de triangles de liquidation : introduction

- Généralités
- Données de travail
- Résultats d'intérêt pour le praticien
- Traitement de phénomène parasite
- Prise en compte de réassurance

Flexibilité des méthodes sur triangles

Outre la sinistralité historique, on peut intégrer des **informations exogènes** à l'estimation de la provision : indicateurs d'exposition (volume de prime, nb de contrats), plan de réassurance...

On peut utiliser les triangles de liquidation sur des quantités de **nature très** \neq :

- montants, montants moyens,
- primes, nombres de sinistres
- loss-ratios (sinistres / primes)...

→ **Quantités ramenées à des périodes** (année, semestre, ...).

Origine et développement

Les données sont rattachées à des **périodes d'origine**, soit

- une période de *survenance* (critère classique),
- une période de *souscription*,
- une période de *déclaration*.

Remarque : souvent maille annuelle pour éviter saisonnalité.

Notations : dans toute la suite,

- i : **année d'origine** (survenance)
- j : **délai règlement** (année de développement/déroulement)
- $x_{i,j}$: mesure de **sinistralité** à l'année d'origine i et au délai j .

Exemple de mesure de sinistralité : montants de sinistres.

Données type : triangle des montants non cumulés

Au 31/12/ l , les paiements sont répertoriés comme suit :

Année de survenance	Années de développement								
	0	1	...	j	...	$J-i$...	$J-1$	J
0	$x_{0,0}$	$x_{0,1}$...	$x_{0,j}$	$x_{0,J-1}$	$x_{0,J}$
1	$x_{1,0}$	$x_{1,1}$...	$x_{1,j}$				$x_{1,J-1}$	
\vdots	\vdots	\vdots	...	\vdots	...		\ddots		
i	$x_{i,j}$...	$x_{i,J-i}$			
\vdots	\vdots	\vdots		\vdots					
$l-j$	$x_{l-j,j}$					
\vdots									
$l-1$	$x_{l-1,0}$	$x_{l-1,1}$							
l	$x_{l,0}$								

Pour simplifier, on prendra ici $l = J = n$ ($l > J$ possible) : on observe donc les données dans $T = \{x_{ij} : i + j \leq n\}$.

La somme des indices des termes diagonaux vaut n (ou l, J).

Règlements calendaires

Les diagonales correspondent à des **règlements calendaires**.
L'année calendaire n a donné lieu au paiement total (ttes années d'origine confondues)

$$\sum_{i=0}^n x_{i, l-i}.$$

Remarque sur le nb de données : suite arithmétique de raison 1,

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}$$

Ex : $n = 6 \rightarrow 21$ obs. (short tail), $n = 16 \rightarrow 136$ obs (long tail)...

Construction du triangle des montants cumulés

Année de survenance	Années de déroulement							
	0	...	j	...	$J-i$...	$J-1$	J
0	$C_{0,0}$...	$C_{0,j}$	$C_{0,J-1}$	$C_{0,J}$
1	$C_{1,0}$...	$C_{1,j}$				$C_{1,J-1}$	
\vdots	\vdots	...	\vdots	...		\ddots		
i	$C_{i,j}$...	$C_{i,J-i}$			
\vdots	\vdots		\vdots					
$I-j$	$C_{I-j,j}$					
\vdots								
$I-1$	$C_{I-1,0}$	$C_{I-1,1}$						
I	$C_{I,0}$							

C_{ij} : *montant cumulé* pour l'année d'origine i jusqu'au délai j :

$$C_{i,j} = \sum_{h=0}^j x_{i,h} \quad \text{ou} \quad x_{i,j} = C_{i,j} - C_{i,j-1} \quad (j \geq 1).$$

Trois différentes approches

But : compléter la partie inférieure du triangle de liquidation.

- ① Techniques de **coût moyen** : souvent utilisées en réassurance (gros risques)

Charges ultimes = nb ultime sinistres \times coût moyen sinistres.

- ② Techniques basées sur les **loss-ratio** (LR) :

Charges ultimes = Primes \times Loss ratio ultime.

- ③ **Méthodes liquidatives** :

Raisonnement à partir des facteurs de développement.

Données pour la suite (stables)

Données d'agrégation de montants de dommages matériels en auto pour +sieurs compagnies au 31/12/1993 (source : FFSA).

Année de survenance	Années de déroulement					
	0	1	2	3	4	5
1988	3209	1163	39	17	7	21
1989	3367	1292	37	24	10	
1990	3871	1474	53	22		
1991	4239	1678	103			
1992	4929	1865				
1993	5217					

Hyp. : développement futur complet des sinistres sur 6 ans.

1 Projection de triangles de liquidation : introduction

- Généralités
- Données de travail
- Résultats d'intérêt pour le praticien
- Traitement de phénomène parasite
- Prise en compte de réassurance

Quantités cibles

A partir de ces données, on cherche à obtenir au **31/12/n** :

- une estimation de la **charge sinistres** pr année d'origine :

$$S_i = C_{i,n} \quad (i = 0, \dots, l).$$

- la **provision** à constituer pour chaque année d'origine :

$$R_i = C_{i,J} - C_{i,J-i} \quad (i = 0, \dots, l)$$

- la **provision totale** : $R = \sum_{i=1}^n R_i.$

⇒ **Détermination du résultat technique de la branche.**

Evaluation comptable réglementaire

Selon les normes IFRS, on évalue un portefeuille par somme des **cash-flows futurs actualisés**. Pour simplifier,

- ① on néglige l'actualisation ici,
- ② on obtient les cash-flows de l'année $(n + k)$ au titre des années d'origine 0 à n avec l'expression

$$CF_{n+k} = \sum_{i+j=n+k} x_{ij}.$$

⇒ On s'intéresse à la diagonale des triangles...

⇒ Cela correspond aux diagonales du **triangle inférieur** à estimer.

Cadence de règlement

Notion intimement liée aux **facteurs de développement** :

- on conduit les calculs avec les facteurs de développement,
- on interprète les résultats en cadences cumulées (*lag factor*).

Pr l'année d'origine i et son dvp en paiements cumulés au **31/12/n**,

	0	...	$J-i$...	j	$j+1$...	J
i	$C_{i,0}$...	$C_{i,J-i}$...	$C_{i,j}$	$C_{i,j+1}$...	$C_{i,J}$

- j^{e} **facteur de développement** : $f_{ij} = \frac{C_{i,j+1}}{C_{i,j}} \quad (j = 0, \dots, J-1)$;
- j^{e} **cadence cumulée** de règlement : $pc_{ij} = \frac{C_{i,j}}{C_{i,J}} \quad (j = 0, \dots, J)$.

Application à l'exemple slide 20

Ici, seule la 1^e année est complètement déroulée.

Pr l'année 1988, il ressort du triangle cumuls :

- les facteurs de développement :

j	0	1	2	3	4
	(0-1)	(1-2)	(2-3)	(3-4)	(4-5)
f_j	1.3624	1.0089	1.0039	1.0016	1.0047

- les cadences de règlement cumulés (en %) :

j	0	1	2	3	4	5
pc_j	72.0	98.1	99.0	99.4	99.5	100

Généralisation

On obtient, en généralisant, les formules suivantes :

pour les facteurs de développement : $f_j = pc_{j+1}/pc_j$,

et pour les cadences cumulées : $pc_j = (f_j f_{j+1} \dots f_{J-1})^{-1}$.

Logiquement,

$$R_i = C_{i,J} - C_{i,J-i} = \frac{1 - pc_{J-i}}{pc_{J-i}} C_{i,J-i} = (f_{J-i} \dots f_{J-1} - 1) C_{i,J-i}.$$

Impact de f_j : visualisation aisée de l'impact sur la provision.

Ex : augmentation des paiements la 2^{ème} année : soit $\beta > 1$,

$$f'_0 = \beta f_0 \quad \Rightarrow \quad pc'_0 = \frac{1}{f'_0 f_2 \dots f_{J-1}} = \frac{pc_0}{\beta}.$$

1 Projection de triangles de liquidation : introduction

- Généralités
- Données de travail
- Résultats d'intérêt pour le praticien
- **Traitement de phénomène parasite**
- Prise en compte de réassurance

Prise en compte de l'inflation

Idée : les règlements de sinistres sont “pollués” par l'inflation.

Cela peut considérablement

- ❶ biaiser l'analyse d'un triangle de paiements,
- ❷ biaiser la méthode de provisionnement (propagation erreur).

⇒ Il faudrait normalement une **étape d'actualisation des paiements qui tienne compte de cette inflation !**

On aimerait produire le calcul de la provision exacte, en considérant une inflation constante dans un premier temps...

⇒ Cette propriété est requise à tte bonne méthode de provision.,
ce qui caractérise un **triangle “régulier”**, défini par

- ① un vecteur $(S_0^{(0)}, S_1^{(0)}, \dots, S_I^{(0)})$ de **charges sinistres ultimes**,
en euros constants de l'année d'origine 0 ;
- ② un vecteur (p_0, \dots, p_{J-1}) de **cadences de règlement**
(**non-cumulées**) tel que $\sum_{j=0}^{J-1} p_j = 1$, avec $p_j \geq 0$;
- ③ un **taux d'inflation** τ annuel (ou autre unité de temps) constant.

On peut tjs exprimer les paiements non cumulés, $\forall i, j = 0, \dots, n$:

- en euros constants (de l'année 0) : $x_{ij}^{(0)} = p_j S_i^{(0)}$;
- ou en euros courants : $x_{ij} = p_j S_i^{(0)} (1 + \tau)^{i+j}$.

On en déduit les paiements cumulés en euros courants...

$$C_{ij} = \sum_{k=0}^j x_{ik} = S_i^{(0)} (1 + \tau)^i \underbrace{\sum_{k=0}^j p_k (1 + \tau)^k}_{B_j}.$$

...et les charges sinistres : $S_i = C_{i,J} = C_{i,n} = S_i^{(0)} (1 + \tau)^i B_J$,
conduisant aux provisions par année d'origine : $\forall i = 1, \dots, I$

$$\begin{aligned} R_i &= C_{i,J} - C_{i,J-i} \\ &= S_i^{(0)} (1 + \tau)^i (B_J - B_{J-i}) \\ &= S_i^{(0)} (1 + \tau)^i \sum_{k=J-i}^J p_k (1 + \tau)^k \end{aligned}$$

Et donc à la provision globale $R = \sum_{i=1}^I R_i$.

1 Projection de triangles de liquidation : introduction

- Généralités
- Données de travail
- Résultats d'intérêt pour le praticien
- Traitement de phénomène parasite
- Prise en compte de réassurance

Intégrer un programme de réassurance

- Réassurance proportionnelle, principe :
 - le **quote part** est facilement traitable et n'améliore pas la qualité de l'estimation du montant de réserves.
 - l'**excédent de plein** permet d'améliorer la base statistique nette de réassurance car données + homogènes.
- Réassurance non proportionnelle :
 - la réassurance non proportionnelle dont la priorité P est basse permet d'écrêter plusieurs sinistres et ainsi de rendre la triangulation sinistres + homogènes et l'estimation + précise.
 - avec P haute, permet par l'étude particulière des sinistres importants d'homogénéiser le triangle sinistre (+ de données). Approche fréquence - coût moyen très utile dans ce contexte.

Globalement, les données doivent être brutes d'autres effets, ces effets étant susceptibles de varier dans le temps...

2 Approches déterministes du provisionnement

- Méthode de Chain Ladder
- Autres approches autorégressives
- Méthodes basées sur des ratios
- Modèles factoriels
- Ajustement de courbes et lissages

Préambule

Les méthodes déterministes doivent impérativement être appliquées avant de recourir à des méthodes stochastiques.

- Elles ont l'avantage d'être simples et robustes.
- En fonction de la méthode choisie, elles s'appliquent à des triangles de toute nature :
 - paiements cumulés,
 - charges, ...
- Elles nécessitent moins d'hypothèses que les méthodes stochastiques.

2 Approches déterministes du provisionnement

- Méthode de Chain Ladder
- Autres approches autorégressives
- Méthodes basées sur des ratios
- Modèles factoriels
- Ajustement de courbes et lissages

Méthode Chain Ladder (CL) standard

Soient les **délais** j et $(j + 1)$ d'un triangle de paiements cumulés :

Année d' origine	Années		de		développement		
	0	...	j	$j + 1$	J
0	$C_{0,j}$	$C_{0,j+1}$	
\vdots			\vdots	\vdots			
i	$C_{i,j}$	$C_{i,j+1}$			
\vdots			\vdots	\vdots			
$I - j - 1$	$C_{I-j-1,j}$	$C_{I-j-1,j+1}$			
$I - j$	$C_{I-j,j}$				
\vdots							
I							

Hypothèse fondamentale : $\forall j = 0, \dots, J - 1$, les facteurs de développement individuels $C_{i,j+1}/C_{i,j}$ sont \perp de l'année d'origine i .

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

et nous pouvons donc déduire une **valeur commune** M de ces rapports en moyennant les facteurs de développement :

$$M = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}.$$

Dans la pratique, l'hypothèse n'est qu'approximativement vérifiée. On choisit pourtant ce facteur commun comme **facteur de développement** ($0 \leq j \leq J - 1$)

$$f_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}$$

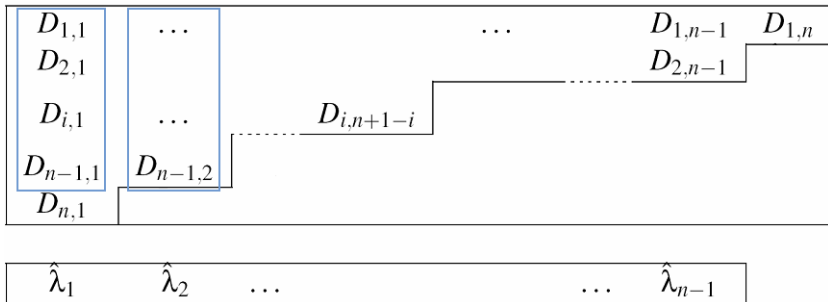


Fig. Estimation des facteurs de développement

$$\begin{aligned}
 \forall k = 1, \dots, n-1, \hat{\lambda}_k &= \frac{\sum_{i=1}^{n-k} D_{i,k+1}}{\sum_{i=1}^{n-k} D_{i,k}} \\
 &= \frac{1}{\sum_{i=1}^{n-k} w_{i,k}} \sum_{i=1}^{n-k} w_{i,k} \lambda_{i,k}
 \end{aligned}$$

Puis on en déduit les charges ultimes (rappel : $C_{i,n-i}$ est le terme diagonal)

$$S_i = C_{i,J} = C_{i,n} = C_{i,n-i} f_{n-i} \dots f_{n-1} = C_{i,n-i} \prod_{k=n-i}^{n-1} f_k,$$

qui conduisent aux **provisions par exercice** ($i = 1, \dots, I$) :

$$R_i = S_i - C_{i,n-i} \quad \Rightarrow \quad R = \sum_{i=1}^I R_i.$$

Si nécessaire, on peut retrouver le rectangle complet de liquidation avec les facteurs de développement

$$C_{i,j} = C_{i,n-i} f_{n-i} \dots f_{j-1} = C_{i,n-i} \prod_{k=n-i}^{j-1} f_k.$$

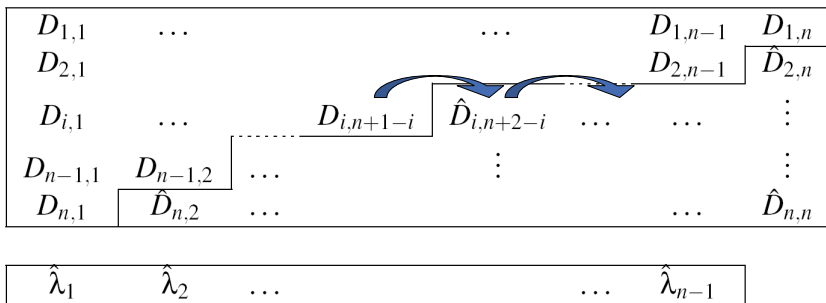


Fig. Estimation des règlements cumulés futurs

$$\hat{D}_{i,j} = (\hat{\lambda}_{n+1-i} \dots \hat{\lambda}_{j-1}) D_{i,n+1-i}$$

$$\hat{R}_i = \hat{D}_{i,n} - D_{i,n+1-i}$$

Exemple : données pédagogiques cumulées

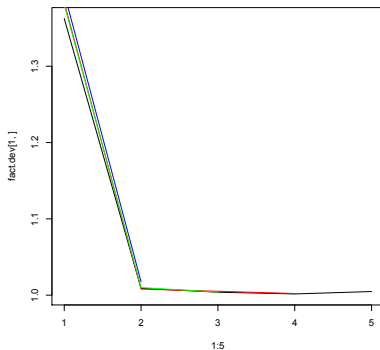
Année de survenance	Années de développement					
	0	1	2	3	4	5
1988	3209	4372	4411	4428	4435	4456
1989	3367	4659	4696	4720	4730	
1990	3871	5345	5398	5420		
1991	4239	5917	6020			
1992	4929	6794				
1993	5217					

Ce qui donne les **facteurs de développement**

j	0	1	2	3	4
	(0-1)	(1-2)	(2-3)	(3-4)	(4-5)
f_j	1.38	1.01	1.0043	1.0018	1.0047

Evolution des facteurs de développement

On trace l'allure des facteurs de dév. pr chaque année d'origine :



Et les cadences cumulées de règlement :

j	0	1	2	3	4	5
pc_j	70.8	97.8	98.9	99.3	99.5	100

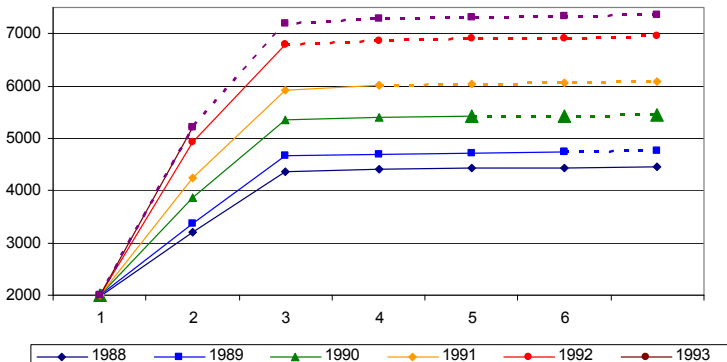
On en déduit le **triangle inférieur de liquidation** et les provisions

Exercice	i	0	1	2	3	4	5	Provisions
1988	0						4456	0
1989	1					4730	4752	22
1990	2				5420	5430	5456	36
1991	3			6020	6046	6057	6086	66
1992	4		6794	6872	6902	6914	6947	153
1993	5	5217	7204	7287	7318	7332	7367	2150
							Total	2427

La provision du dernier exercice représente 89% de la provision globale (caractéristique des branches short-tail).

Représentation des règlements cumulés estimés

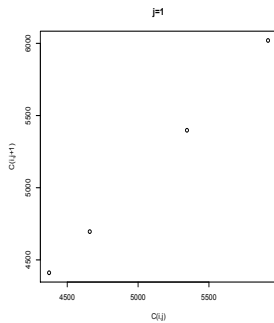
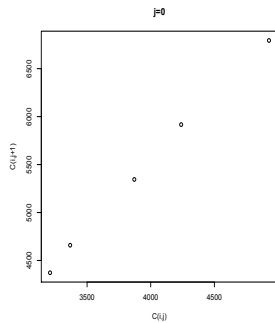
On voit une tendance en fonction de l'année d'origine...inflation ?

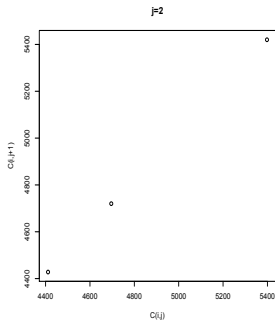


Validation de l'hypothèse : les C-C plot

Validation **empirique** car le modèle n'est pas stochastique.

Si pour j fixé, il existe f_j tel que $C_{i,j+1} = f_j C_{i,j}$,
 \Rightarrow les couples $(C_{i,j}, C_{i,j+1})_{i=0,\dots,l-j-1}$ alignés sur 1 droite à l'origine.





La validation essentielle repose sur le **triangle de dév.**, appelé **d-triangle**. Il permet d'effectuer des choix alternatifs car

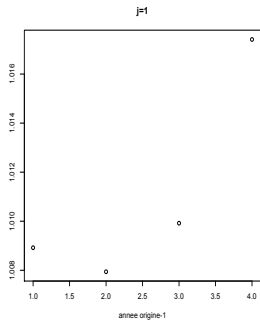
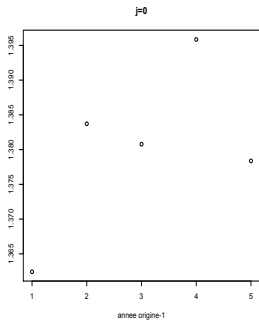
- on y détecte les **tendances** (trend) et les **saisonnalités**,
- il ressort les **valeurs extrêmes**...

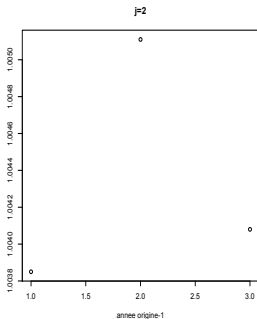
Exercice	i	0	1	2	3	4
1988	0	1.36242	1.00892	1.00385	1.00158	1.00474
1989	1	1.38372	1.00794	1.00511	1.00212	
1990	2	1.38078	1.00992	1.00408		
1991	3	1.39585	1.01741			
1992	4	1.37837				
Moyenne		1.38023	1.01105	1.00435	1.00185	
Ecart-type		0.01074	0.00374	0.00055	0.00027	
Coeff. variation		0.008	0.004	0.001	0.000	

⇒ Hyp. CL acceptable si f_{ij} “**constants**” à j fixé.

On y constate pour nos données de l'exemple “pédagogique” :

- très **faible volatilité**, décroissante avec le délai j .
- visualisation de la variation de ces facteurs :





- aucune structure non-aléatoire évidente pour $j = 1, 2, \dots$;
- mais peut-être un trend pour $j = 0$.

Chain Ladder pondéré : variante + flexible

CL standard \Rightarrow choix d'un même facteur de développement $\forall i$.

On pourrait choisir ce facteur comme une fonction des facteurs **individuels** $f_{0,j}, f_{1,j}, \dots, f_{l-j-1,j}$ de la colonne j .

La fonction la plus simple est celle de la **moyenne pondérée**,

$$f_j = \frac{\sum_{i=0}^{l-j-1} w_{i,j} f_{i,j}}{\sum_{i=0}^{l-j-1} w_{i,j}},$$

avec des pondérations $(w_{i,j})_{i=0,\dots,l-j-1}$ choisies en fonction de

- l'allure des facteurs individuels, leur fiabilité, leur volatilité,
- considérations externes : prudentielles, ...

Exemples de pondérations

- La **moyenne** : $f_j = \frac{1}{l-j} \sum_{i=0}^{l-j-1} f_{i,j}$.
- La **moyenne des k derniers** : par exemple pour $k = 3$,

$$f_j = \frac{1}{3} \sum_{i=l-j-3}^{l-j-1} f_{i,j} \quad \text{si } j = 1, \dots, l-3;$$

$$f_j = \frac{1}{2}(f_{0,j} + f_{1,j}) \quad \text{si } j = 1, \dots, l-2;$$

$$f_j = f_{0,j} \quad \text{si } j = l-1.$$

Rq : si l'on prend $\mathbf{w}_{i,j} = \mathbf{C}_{i,j}$, on retombe sur le CL standard.

Pondération par année calendaire

Le but est d'accorder **plus d'importance aux données récentes** qu'aux anciennes. Avec la pondération $w_{i,j} = i + j + 1$, on obtient

$$\sum_{i=0}^{l-j-1} w_{i,j} = \sum_{i=0}^{n-j-1} i + (n-j)(j+1) = \frac{(n-j)(n+j+3)}{2}.$$

Accentuer cet effet en prenant par ex. $w_{i,j} = (i + j + 1)^2$.

Rq : si l'on a détecté un trend dans le **d-triangle**, il peut être utile de le modéliser par un modèle de régression.

Bilan

La méthode de Chain Ladder pondérée est une méthode qui satisfait de bonnes propriétés.

Prop. : CL donne la provision exacte avec un triangle régulier.

- Elle est simple à mettre en oeuvre, universelle et robuste ; mais n'est pas très flexible ;
- Elle peut être utilisée avec des incréments négatifs, donc aussi adaptée aux recours.
- Les versions stochastiques des modèles pondérés nécessitent de développer de nouvelles formules pour la volatilité.

Elle est donc particulièrement adaptée à des triangles de charges (intégrant des recours) ou des triangles de paiement...

2 Approches déterministes du provisionnement

- Méthode de Chain Ladder
- **Autres approches autorégressives**
- Méthodes basées sur des ratios
- Modèles factoriels
- Ajustement de courbes et lissages

Alternatives

Les méthodes Chain Ladder supposent l'existence pour j fixé ($j = 0, \dots, J - 1$) d'un facteur f_j tel que

$$C_{i,j+1} = f_j C_{i,j}, \quad \text{pour } i = 1, \dots, I - j.$$

→ $(C_{i,j}, C_{i,j+1})$ alignés sur une droite passant par l'origine : parfois, cette **hypothèse est remise en cause**.

→ Une alternative serait alors de supposer que $C_{i,j+1}$ est une fonction affine de $C_{i,j}$.

⇒ On ajoute un effet constant (ordonnée à l'origine)...

Modélisation par London Chain

[Benjamin and Eagles, 1986]

- Méthode peu utilisée car risque de surparamétrisation.

Suppose l'existence, pour j fixé, de paramètres (f_j, a_j) tels que

$$C_{i,j+1} = f_j C_{i,j} + a_j \quad \forall i = 0, \dots, l-j-1$$

Encore une fois, les points devraient être sensiblement alignés.

Par analogie avec la régression linéaire, on a

$$(\hat{f}_j, \hat{a}_j) = \arg \min_{(f_j, a_j)} \sum_{i=0}^{l-j-1} (C_{i,j+1} - a_j - f_j C_{i,j})^2.$$

On annule les dérivées partielles, et en notant \bar{C}_j la moyenne de la colonne j donnée par

$$\bar{C}_j = \frac{1}{I-j} \sum_{i=0}^{I-j-1} C_{i,j}$$

on obtient les **estimateurs classiques** :

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

et

$$a_j^{LC} = \bar{C}_{j+1} - f_j^{LC} \bar{C}_j.$$

On remarque qu'il s'agit presque du **coefficient de corrélation linéaire de Pearson** entre les colonnes j et $(j + 1)$:

$$f_j^{LC} = \frac{\text{Cov}(\text{col } j, \text{col } j+1)}{\text{Var}(\text{col } j)}.$$

On obtient finalement le **montant cumulé** de la colonne $(j + 1)$ par

$$C_{i,j+1} = f_j^{LC} C_{i,j} + a_j^{LC} \quad \text{pour } i \geq I - j.$$

Rq : si l'on applique les MCO à CL classique $C_{i,j+1} = f_j C_{i,j}$, cela

conduit à un facteur $\frac{\sum_{i=0}^{I-j-1} C_{i,j} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}^2}$, \neq de f_j de CL.

Reprise de l'exemple pédagogique

En appliquant le modèle affine aux mêmes données, on obtient

j	0	1	2	3	4
f_j^{LC}	1.404	1.0405	1.0036	1.0103	1.0047
a_j^{LC}	-90.311	-147.27	3.742	-38.493	0

Exercice	i	0	1	2	3	4	5	Provisions
1988	0						4456	0
1989	1					4730	4752	22
1990	2				5420	5437	5463	43
1991	3			6020	6045	6069	6098	78
1992	4		6794	6922	6950	6983	7016	222
1993	5	5217	7234	7380	7410	7447	7483	2266
							Total	2631

Prop. le London Chain fournit la provision exacte d'un triangle régulier : les charges S_i^{LC} coïncident avec les charges exactes S_i .

Cette méthode n'implique aucune difficulté de mise en oeuvre puisqu'il s'agit d'une simple régression.

Problème : on estime $2(n - 1)$ paramètres à partir de $\frac{n(n+1)}{2}$ données, d'où un risque de **surparamétrisation**...

→ Un indicateur permettant de quantifier cet effet est donné par le ratio du nb de paramètres sur le nombre de données

$$\frac{4(n - 1)}{n(n + 1)} \quad (\text{environ } 1/2 \text{ pour } n = 5)$$

Méthode du London pivot

[Straub, 1988]

Objectif : **réduire le paramétrage** de la modélisation London Chain.

- Modèle intermédiaire entre Chain Ladder et London Chain.
- Suppose l'existence d'un param. **a** tel que ($j = 0, \dots, J - 1$)

$$C_{i,j+1} + a = f_j (C_{i,j} + a) \quad \forall i = 1, \dots, I - j$$

Ainsi, a est indépendant de j... \Rightarrow si l'hyp. est validée, les points $(C_{i,j+1}, C_{i,j})$ devraient être **alignés sur des droites concourantes**.

Leur point d'intersection se trouve en $(-a, -a)$, le fameux **point pivot** (dans la méthode CL, $(-a, -a) = (0, 0)$).

On estime les $(J + 1)$ paramètres par **MCO**, en minimisant

$$(\hat{f}_j, \hat{a}) = \arg \min_{(f_j, a)} \sum_{j=0}^{J-1} \sum_{i=0}^{I-j} [(C_{i,j+1} + a) - f_j(C_{i,j} + a)]^2.$$

→ *Mais ce problème n'a pas de solution analytique*, d'où un algorithme itératif qui provoque une utilisation peu courante...

Prop. Dans le cas régulier, cette méthode fournit la provision exacte.

2 Approches déterministes du provisionnement

- Méthode de Chain Ladder
- Autres approches autorégressives
- Méthodes basées sur des ratios
- Modèles factoriels
- Ajustement de courbes et lissages

Méthodes de ratios : contexte d'utilisation

- Intègre 1 info. suppl. par rapport au triangle des paiements.
- Sous la forme d'un **indicateur d'exposition au risque E_i** par rapport à la sinistralité pour tte année d'origine i ($i = 0, \dots, l$) :
 - un montant de primes, un nombre de contrats,...
 - un nombre de sinistres déclarés ou réglés (si la définition d'un sinistre est constante au cours des années),

On utilisera ici les notations suivantes :

$L_{i,j} = C_{i,j}/E_i$ ($i, j = 0, \dots, n$) pour les ratio **individuels**,

$L_i = C_{i,n}/E_i = S_i/E_i$ ($i, j = 0, \dots, n$) pour le ratio **ultime**.

Exemple illustratif

Si l'indicateur d'exposition est une **prime** P_i , alors $L_{i,j-i}$ est la **part de prime de l'année d'origine i consommée** au **31/12/l**.

→ On parle de **loss ratio** (LR) pour L_i , ratio technique très utilisé.

Dans notre ex., les données d'expo. sont les primes acquises (CL appliqué ici donne pour le triangle inf. les primes manquantes) :

Exercice	i	0	1	2	3	4	5
1988	0	4563	4589	4590	4591	4591	4591
1989	1	4718	4674	4671	4672	4672	4672
1990	2	4836	4861	4861	4863	4863	4863
1991	3	5140	5168	5173	5175	5175	5175
1992	4	5633	5668	5671	5673	5673	5673
1993	5	6389	6425	6428	6431	6431	6431

Remarquez que les primes sont + vite collectées que le paiement des sinistres, les facteurs de dev. sont vite égaux à 1...

Exercice	i	S_i	P_i	Loss ratio ultime (%)
1988	0	4456	4591	97.1
1989	1	4752	4672	101.7
1990	2	5456	4863	112.2
1991	3	6086	5175	117.6
1992	4	6947	5673	122.5
1993	5	7367	6431	114.6

Naturellement, le *d-triangle* des facteurs individuels $(f_{i,j})_{i+j \leq n}$ associé à celui des paiements cumulés $(C_{i,j})_{i+j \leq n}$ est le même que celui du triangle des ratios $(L_{i,j})_{i+j \leq n}$, car

$$f_{i,j}^L = \frac{L_{i,j+1}}{L_{i,j}} = \frac{C_{i,j+1}/E_i}{C_{i,j}/E_i} = f_{i,j}.$$

⇒ Ainsi la provision déduite du triangle $(L_{i,j})_{i+j \leq n}$ par Chain Ladder **coincide** avec celle provenant du triangle $(C_{i,j})_{i+j \leq n}$.

Méthodes basées sur des ratios

- Le **Loss Ratio simple** : supposer constant les LR de toutes les années d'origine, et poser $L_i = L_0$ avec L_0 connu.
 - on cristallise le dernier LR connu : on déduit $S_i = E_i \times L_0$.
 - $R_i = S_i - C_{i,n-i}$: très simple mais résultats aberrants si la branche étudiée n'est pas hyper stable.
- Le **Loss Ratio complémentaire** : supposer les LR constants = à 1 param. L à estimer pr ttes années d'orig. ($L \sim C_{i,n}/E_i, \forall i$).
 - Par ex., $\sum_{i=0}^I C_{i,J} / \sum_{i=0}^I E_i \sim L$.
 - Ou bien pour estimer L , on peut tenter de détecter une tendance en déterminant $L^{(j)} = \sum_{i=0}^{I-j} C_{i,j} / \sum_{i=0}^{I-j} E_i$.
Puis l'analyse de la suite des $L^{(j)}$ (CV) conduit par extrap. au LR commun final (par rég. lin. par ex.).

Méthode de Bornhuetter-Ferguson

[Bornhuetter and Ferguson, 1972]

Principe : méthode toujours basée sur des ratios, ici les LR.

- Très utilisée en entreprise en complément de la méthode CL.
- Particulièrement adaptée au provisionnement des années récentes dans les triangles instables.

$$S_i = C_{i,n-i} + (1 - pc_{n-i})L_iP_i.$$

On substitue au LR inconnu L_i un LR “attendu” $\Phi_i \rightarrow \Phi_i$ déterminé par des considérations **exogènes**, par ex. un **benchmark marché**.

Rq : existe aussi en version stochastique.

⇒ Approche crédibilisée : a priori (via L_i) et expérience (via pc_{n-i}).

Reprise de l'exemple fil rouge

La réserve s'exprime donc comme suit :

$$R_i = S_i - C_{i,n-i} = (1 - pc_{n-i})P_i\Phi_i$$

Mise oeuvre toujours sur le même exemple :

Exercice	i	$(1 - pc_{n-i})$ (%)	P_i	Φ_i	R_i
1988	0	0	4591	97.1	0
1989	1	0.47	4672	100	22
1990	2	0.66	4863	105	34
1991	3	1.09	5175	110	62
1992	4	2.20	5673	120	150
1993	5	29.18	6431	125	2346
				Total	2614

Méthode Cape Cod (Bühlmann-Stanard)

- Basée sur une **segmentation a priori de l'ensemble des années d'origine** en groupes d'années “semblables”.
- On attribue à chacun de ces groupes un **Loss Ratio** et un vecteur de **cadences de règlement**.

En reprenant la notion de cadence cumulée :

$$pc_{n-i} = \frac{C_{i,n-i}}{C_{i,n}} \Rightarrow S_i = C_{i,J} = C_{i,n} = C_{i,n-i} + (1 - pc_{n-i})C_{i,n}.$$

D'où en utilisant le LR de la i^{eme} année d'origine $L_i = C_{i,n}/P_i$:

$$S_i = C_{i,n-i} + (1 - pc_{n-i})L_iP_i.$$

⇒ Mélange entre a priori (via L_i) et historique (via pc_{n-i}).

Notons $pc = (pc_0, \dots, pc_{n-1})$ le vecteur des **cadences cumulées** provenant par ex. d'une méthode CL avec pour l'exercice k ,

$$pc_{n-k} = \frac{C_{k,n-k}}{C_{k,n}}, \text{ et le loss ratio } L_k = \frac{C_{k,n}}{P_k} = \frac{C_{k,n-k}}{pc_{n-k} P_k}.$$

On obtient ensuite naturellement le loss-ratio L_A , **associé à un groupe $A \subset \{0, 2, \dots, n\}$** d'années d'origine semblables,

$$L_A = \sum_{k \in A} C_{k,n-k} / \sum_{k \in A} pc_{n-k} P_k.$$

La provision d'une année d'origine i de groupe A_i vaut

$$R_i = S_i - C_{i,n-i} = (1 - pc_{n-i}) P_i L_{A_i}.$$

Cas particuliers :

- si $A_i = \{i\}$ alors $L_{A_i} = \frac{C_{i,n-i}}{pc_{n-i}P_i}$.
- si à l'opposé $A_i = \{0, 1, 2, \dots, n\}$ alors $L_{A_i} = \frac{\sum_{k=0}^n C_{k,n-k}}{\sum_{k=0}^n pc_{n-k}P_k}$.

Reprise de l'exemple fil rouge : en utilisant les cadences du CL et avec $A_i = \{0, 2, \dots, I\}$, on obtient $L_{A_i} = 1.1141$ (111,41%) et :

Exercice	i	$(1 - pc_{n-i})$ (%)	P_i	R_i
1988	0	0	4591	0
1989	1	0.47	4672	24
1990	2	0.66	4863	36
1991	3	1.09	5175	63
1992	4	2.20	5673	139
1993	5	29.18	6431	2091
			Total	2353

Cette méthode, extension des méthodes Chain Ladder, est évidemment **sensible au choix des années d'origine semblables...**

Ce choix peut résulter :

- d'informations exogènes au triangle,
- de l'expérience du praticien,
- s'appuyer sur des graphiques de développement des LR.

Cette méthode fournit des **provisions toujours positives**, comme la méthode de Bornhuetter-Ferguson que nous avons présenté.

2 Approches déterministes du provisionnement

- Méthode de Chain Ladder
- Autres approches autorégressives
- Méthodes basées sur des ratios
- **Modèles factoriels**
- Ajustement de courbes et lissages

Les modèles factoriels : principales caractéristiques

Ces techniques s'appliquent

- aux **montants non-cumulés** x_{ij} (ou aux LR correspondants) ;
- sous la forme d'un produit de param. correspondant chacun à une **direction** du triangle de liquidation ($0 \leq i, j \leq n$) :

$$x_{i,j} = x_i y_j \lambda_{i+j}$$

où

- x_i est le paramètre de l'année d'origine,
- y_j est lié au délai de règlement,
- λ_{i+j} correspond à **l'année calendaire** (pour l'inflation par ex.).

- Approche privilégiée par la plupart des modèles stochastiques.
- 3n paramètres : inférieur au nb d'obs. à partir de $n > 4$!

Problème : sous cette forme, le modèle n'est pas identifiable :

$$[(x_i), (y_j), (\lambda_k)] \text{ et } [(a x_i), (\frac{y_j}{a}), (\lambda_k)] \text{ donnent les mêmes } x_{ij}...$$

⇒ Il faut donc imposer des contraintes qui

- rendent l'effet des facteurs identifiables,
- conservent l'interprétabilité des paramètres.

C'est l'objet du modèle de **De Vylder** que nous allons étudier maintenant...

Les moindres carrés de Vylder

[De Vylder, 1982]

On se place dans un cadre où l'on suppose que **l'inflation annuelle est constante** dans le triangle, car

- le triangle a été préalablement déflaté, ou
- la branche considérée a connu une inflation spécifique cste.

Alors on peut retenir la modélisation qui intègre directement cette inflation (avec $0 \leq i, j \leq n$)

$$x_{i,j} = x_i y_j$$

Contrainte supp. : $\sum_{j=0}^n y_j = 1$ pour rendre le modèle identifiable.

Les paramètres (x_i) et (y_j) sont toujours interprétables :

- $\sum_{j=0}^n x_{i,j} = x_i \sum_{j=0}^n y_j = x_i$: charge sinistres de l'année i ;
- $y_j = \frac{x_{ij}}{x_i}$ est la cadence (non-cumulée) de règlement en j^e année de développement.

Ainsi,

$$x_{ij} = x_i y_j \Rightarrow C_{i,j} = \sum_{k=0}^j x_{i,k} = x_i \sum_{k=0}^j y_k \Rightarrow \frac{C_{i,j+1}}{C_{i,j}} = \frac{\cancel{x_i} \sum_{k=0}^{j+1} y_k}{\cancel{x_i} \sum_{k=0}^j y_k}.$$

Les **facteurs de développement sont donc indépendants de i** ,
comme dans le modèle Chain Ladder !

Par contre, ces facteurs ne sont plus des facteurs individuels pondérés... On estime les $2n$ param. (x_i) et (y_j) par MCO minimisant

$$\Delta = \sum_{i=0}^n \sum_{j=0}^{n-i} w_{i,j} (x_{i,j} - x_i y_j)^2.$$

Cela revient à résoudre le système (en dérivant)

$$\begin{cases} x_i = \frac{\sum_{j=0}^{n-i} w_{i,j} x_{i,j} y_j}{\sum_{j=0}^{n-i} w_{i,j} y_j^2} & (i = 0, \dots, n) \\ y_j = \frac{\sum_{i=0}^{n-j} w_{i,j} x_{i,j} x_i}{\sum_{i=0}^{n-j} w_{i,j} x_i^2} & (j = 0, \dots, n) \end{cases}$$

Qui peut se réécrire $\begin{cases} x_i = f_i(y_0, \dots, y_n), \\ y_j = g_j(x_0, \dots, x_n). \end{cases}$

On peut déduire les réserves estimées R_i des valeurs $\hat{x}_{i,j}$ prévues par le modèle :

$$\hat{R}_i = \sum_{k=n-i+1}^n \hat{x}_{i,k} \Rightarrow \hat{R} = \sum_{i=1}^n \hat{R}_i.$$

Une méthode de validation intuitive consiste à comparer les valeurs **prévues** $\hat{x}_{i,j}$ avec les **valeurs observées** $x_{i,j}$ du triangle supérieur.

Prop. Dans le cas régulier, la méthode des moindres carrés de De Vylder produit la provision exacte.

Retour à notre exemple fil-rouge

Estimations par moindres carrés de De Vylder :

- Estimation des **facteurs** :

k	0	1	2	3	4	5
\hat{x}_k	4505.5	4761.0	5468.5	6019.3	6957.8	7372.1
\hat{y}_k	0.70767	0.27	0.01564	0.00427	0.00184	0.00466

- Estimation des **provisions** :

k	0	1	2	3	4	5	Total
\hat{R}_i	0	22	36	45	156	2155	2434

- Comparaisons avec valeurs observées :

Exercice	i	0	1	2	3	4	5
1988	0	3209 3188	1163 1217	39 52	17 19	7 8	21 21
1989	1	3367 3369	1292 1286	37 55	24 20	10 9	22 22
1990	2	3871 3870	1474 1477	53 63	22 23	10 10	26 26
1991	3	4239 4260	1678 1625	103 70	26 26	11 11	28 28
1992	4	4929 4924	1865 1879	81 81	30 30	13 13	32 32
1993	5	5217 5217	1990 1990	85 85	32 32	14 14	34 34

2 Approches déterministes du provisionnement

- Méthode de Chain Ladder
- Autres approches autorégressives
- Méthodes basées sur des ratios
- Modèles factoriels
- Ajustement de courbes et lissages

Les méthodes d'ajustement de courbe

L'intérêt de ces méthodes est qu'elles permettent

- de prendre en compte paramétriquement l'évolution des cadences de règlement au cours des années d'origine,
- d'effectuer des projections de celles-ci (problématique du facteur de queue, "tail factor"),
- inconvénient : parfois la mise en oeuvre (minimisation en moindres carrés non-linéaires).

Champ d'application : le plus souvent à des loss ratios cumulés (ou non-cumulés), mais aussi à des règlements (cumulés ou non).

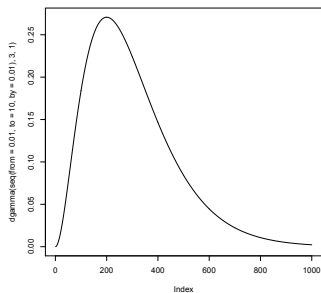
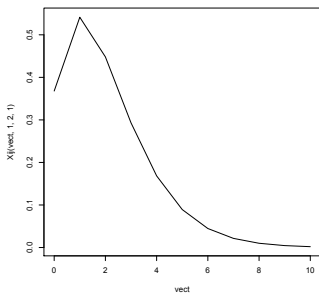
Exemple 1 : les courbes de Hoerl

Objectif : ajuster une fonction paramétrée par année d'origine sur les développements des paiements **non-cumulés** $x_{i,j}$.

- Méthode proposée en 1989 par Y. Kahane.
- On calibre une densité Gamma car allure proche des développements observés :

$$x_{i,j} = \alpha_i (1 + j)^{\beta_i} e^{-\gamma_i(1+j)}.$$

- α_i est le paramètre d'échelle,
- β_i et γ_i influent sur la forme et la longueur du déroulement de l'année d'origine i .



→ En pratique, β_i , γ_i peuvent ne pas dépendre de i .

→ **Minimisation des moindres carrés pondérés non-linéaires**
sur les données du triangle supérieur :

$$\Delta = \sum_{i+j \leq n} w_{i,j} \left[x_{i,j} - \alpha_i (1+j)^{\beta_i} e^{-\gamma_i (1+j)} \right]^2.$$

Estimation des paramètres

Pour simplifier, on se **ramène à une minimisation linéaire** par une transformation :

$$\log x_{i,j} = \log \alpha_i + \beta_i \log(1 + j) - \gamma_i(1 + j).$$

On cherche donc à minimiser

$$\Delta' = \sum_{i+j \leq n} w_{i,j} [\log x_{i,j} - \log \alpha_i + \beta_i \log(1 + j) - \gamma_i(1 + j)]^2.$$

Pb : 2 minimisations ne donnent pas exact. les mêmes solutions...

Astuce : on peut initialiser les paramètres pour la première minimisation en se servant de l'estimation de ceux de la seconde.

Exemple 2 : la méthode de Craighead

- S'applique sur les paiements ou loss ratios (LR) **cumulés** de chaque année d'origine.
- Ajuste une exponentielle décroissante semblable à la fonction de répartition d'une loi de Weibull à 2 paramètres :

$$L_{i,j} = L_i \left(1 - \exp \left[- \left(\frac{j+1}{\alpha_i} \right)^{\tau_i} \right] \right) \quad j \geq 0.$$

- L_i le ratio ultime de l'année i , inconnu ;
- $\alpha_i \sim$ **longueur du déroulement**, $\tau_i \sim$ **forme** de la courbe.
- peuvent ne pas dépendre de i (cf graph des dév. de $L_{i,j}$).

Estimation des paramètres

On détermine donc les triplets $(L_i, \alpha_i, \tau_i)_{i \geq 0}$ par minimisation de

$$\Delta = \sum_{i=0}^n \sum_{j=0}^{n-i} w_{i,j} \left[L_{i,j} - L_i \left(1 - \exp \left[- \left(\frac{j+1}{\alpha_i} \right)^{\tau_i} \right] \right) \right]^2.$$

Les **pondérations** permettent de moduler

- l'importance des années d'origine,
- la prépondérance des délais de règlement.

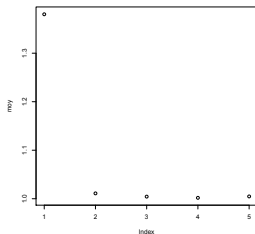
On déduit la charge ultime par $S_i = E_i L_i$.

Rq : méthode connue pour estimer des provisions avec une **forte volatilité**, en particulier pour les origines encore peu développées...

Lissage, interpolation, extrapolation

Recours aux méthodes de lissage dans les méthodes CL, pour **atténuer les irrégularités** des facteurs de dév. bruts $(f_j)_{j=1,\dots,J-1}$.

Très souvent, et surtout dans les branches short-tail, la suite $(f_j)_{j=1,\dots,J-1}$ est **décroissante dès l'origine** :



Le lissage de ces facteurs empiriques consiste à ajuster sur ces points une fonction $y = f(t)$, régulière et vérifiant $f(t) \geq 1$.

Les plus utilisées sont :

- la **puissance inverse** :
 - à 2 paramètres : $f(t) = 1 + a(1 + t)^{-b}$, $a, b \geq 0$,
 - à 3 paramètres : $f(t) = 1 + a(c + t)^{-b}$, $a, b \geq 0$,
- l'**exponentielle** négative : $f(t) = 1 + ae^{-bt}$, $b \geq 0$,
- la fonction de type **Weibull** (cf courbes de Hoerl) :

$$f(t) = 1 + a(1 + t)^b e^{-ct}, \quad a, b, c \geq 0.$$

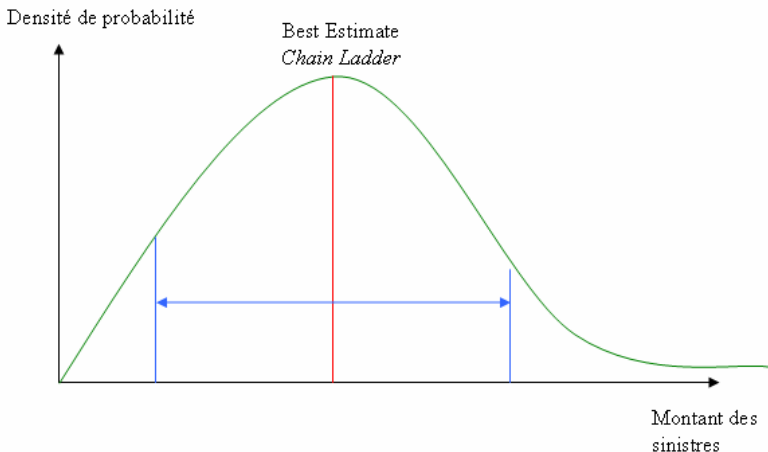
Intégration d'un tail factor : parfois, il manque des données dans la chronique des règlements pour certaines années d'origine (dissymétrie : délai règlement > nb d'années d'origine).

⇒ Le lissage permet de projeter les facteurs de développement.

- 3 Approches stochastiques de provisionnement
 - Introduction
 - Le modèle de Mack
 - Munich Chain Ladder
 - Approches factorielles et extension par GLM
 - Incertitude d'estimation sur la provision

Objectif dans l'idéal : la loi de probabilité de la provision. Sinon

- 1 ...estime la volatilité des provisions (**fiabilité** de l'estim.).
- 2 ...détermine certaines mesures de risque.



Modélisation stochastique du rectangle de liquidation

Les paramètres sont *estimés à partir des données du triangle supérieur T* .

→ Données supposées être des **v.a.r.**, observées au **31/12/l**.

Les méthodes stochastiques s'appuient sur les déterministes, et

- permettent d'**explicitier/valider les hyp.** (analyse des résidus),
- de détecter des irrégularités potentielles dans le triangle,
- d'**évaluer la variabilité des estimations** de ceux-ci,
- donc d'en construire des intervalles de confiance,
- de **simuler** (Monte Carlo) la sinistralité d'exercices futurs.

Avantage : on peut, avec une hyp. de loi sur les données d'entrée, remonter à l'estimation de la loi de proba. de la provision totale R ...

Cela implique beaucoup de calculs utiles en gestion des risques :

- Value-at-Risk,
- probabilité d'insuffisance...

ATTENTION au problème de mauvaise spécification du modèle !

→ Etant donné l'importance pratique de la méthode CL, on a d'abord cherché une méthode stochastique visant à reproduire les évaluations de CL : notamment le **best estimate de la provision**.

Ainsi ont été développés :

- le modèle récursif de Mack,
- le modèle log-Poisson de Renshaw et Verrall.

Pour ces 2 modèles qui reproduisent exactement les provisions CL en espérance, voici quelques caractéristiques importantes :

- Le modèle de Mack :

- basé sur les **montants cumulés** ;
- Extension : le modèle “Munich chain ladder”.
→ Permet de faire CV les provisions chain ladder obtenues sur le triangle de charges et celui des paiements dans une même branche.

- Le modèle Log-Poisson :

- est basé sur les **incréments** (montants non cumulés) ;
- se trouve dans la lignée des modèles factoriels ;
- type de modèle très largement développé depuis dans le cadre de l'application du bootstrap aux modèles GLM.

Notations en environnement stochastique

- Les v.a.r. d'incrément X_{ij} , celles de montants cumulés C_{ij} ;
- Parmi les n^2 variables du rectangle de liquidation, celles du triangle supérieurs sont observées et notées $(X_{i,j})_{i+j \leq n}$;
- La provision pour la i^{eme} année d'origine ($i = 1, \dots, l$) s'écrit

$$R_i = C_{i,J} - C_{i,J-i}, \quad \text{ou alors} \quad \mathbf{R}_i = \sum_{j=J-i+1}^J \mathbf{X}_{i,j}.$$

- La provision globale : $\mathbf{R} = \sum_{i=1}^n \mathbf{R}_i$ et F_R sa FdR,
- $\mathbf{M}_R(\mathbf{s}) = \mathbb{E}[e^{s^R}]$: fonction génératrice des moments (FGM).

Quantités d'intérêt de la loi de R

- L'espérance notée $\mathbb{E}[R]$,
- La variance $\mathbb{V}[R]$ et l'écart-type $\sigma(R)$: $\sigma(R) = \sqrt{\mathbb{V}[R]}$,
- Le coefficient d'asymétrie ou skewness, noté $\gamma_1(R)$:

$$\gamma_1(R) = \frac{\mu_3(R)}{\sigma^3(R)} \quad \text{avec } \mu_3(R) = \mathbb{E}[(R - \mathbb{E}[R])^3].$$

- Pour les moments d'ordre supérieur, on utilisera la FGM ;
- Quantile d'ordre $(1 - \eta)$ de R (déterminer capital éco.) :

$$\text{VaR}_\eta(R) = q_{1-\eta}(R) = F_R^{-1}(1 - \eta).$$

C'est la provision suffisante dans $100(1 - \eta)\%$ des cas.

Incertitude d'un estimateur

Il y a plusieurs sources d'erreur dans l'estimation d'un phénomène...

On décompose l'erreur globale en 3 termes. Supposons que le vrai modèle sous-jacent aux observations est donné par

$$X_{ij} = u(i, i; \theta) + \epsilon_{ij}.$$

On estime de notre côté le modèle

$$X_{ij} = v(i, i; \phi) + \epsilon_{ij},$$

qui donne les prévisions $\hat{X}_{ij} = v(i, i; \hat{\phi})$.

L'erreur de prévision du modèle (résidu) vaut $X_{ij} - \hat{X}_{ij}$, qui se décompose en

$$X_{ij} - \hat{X}_{ij} = [v(i, i; \phi) - v(i, i; \hat{\phi})] + \epsilon_{ij} + [u(i, i; \theta) - v(i, i; \phi)].$$

On parle de :

- 1^{er} terme : parameter error → incertitude sur l'estimation du paramètre sachant le modèle bien spécifié ;
- 2^{ème} : process error → aléa des futures réalisations (variance), même si modèle bien spécifié et param. bien estimés ;
- 3^{ème} : erreur de modèle (mis-specification error) → difficilement quantifiable, mais à ne pas négliger !

Solvency 2 standards says : *Undertakings are not required to quantify the degree of model error in quantitative terms, or to re-calculate the value of the capital requirement using a more accurate method in order to demonstrate that the difference between the result of the chosen method and the result of a more accurate method is immaterial. Instead, it is sufficient if there is reasonable assurance that the model error included in simplification is immaterial.*

- 3 **Approches stochastiques de provisionnement**
 - Introduction
 - **Le modèle de Mack**
 - Munich Chain Ladder
 - Approches factorielles et extension par GLM
 - Incertitude d'estimation sur la provision

Le modèle de Mack

[Mack, 1993], [Mack, 1994b], [Mack, 1994a], [Mack, 2000]

→ Ce modèle “stochastise” le modèle Chain Ladder standard,
→ en y ajoutant des indicateurs de risque de prédiction déduits d'une **hypothèse de volatilité** au sein du triangle. Il est :

- non-paramétrique et conditionnel :
 - aucune hyp. de loi faite sur les composantes du triangle,
 - conditionnel car les espérances sont prises connaissant les réalisations fournies par le triangle supérieur T,
- s'applique à des quantités cumulés :
 - montants, nombres de sinistres, paiements, charges ;
 - cette caractéristique le distingue des méthodes factorielles.

Hypothèses fondamentales

Le modèle de Mack sous-entend les hypothèses suivantes :

- ① **H1** : indépendance des exercices d'origine :

$$(C_{i_1,j})_{j=1,\dots,J} \perp\!\!\!\perp (C_{i_2,j})_{j=1,\dots,J} \quad \text{pour } i_1 \neq i_2.$$

- ② **H2 (Markov ou AR(1))** : $\exists f_j$ t.q. pr $i = 0, \dots, I; j = 0, \dots, J - 1$:

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

→ H1 non réaliste si changements dans la **gestion des sinistres**, **inflation** (affecte par effet calendaire + sieurs exercices d'origine).

→ H2 implique l'**absence de corrélation** entre facteurs successifs.

Interprétations :

- pour chaque survenance i , le même facteur \hat{f}_j quantifie l'↗ du montant de l'année j à l'année $j + 1$ (comme CL) ;
- ce facteur est \neq des facteurs individuels $f_{i,j}$;
- la base du modèle de Mack est de supposer que les facteurs individuels de développement sont des réalisations d'une v.a. d'espérance inconnue f_j .

Lien direct avec la méthode Chain Ladder :

Cette espérance est estimée par l'estimateur \hat{f}_j de CL.

Une remarque sur H2

En réécrivant l'hypothèse H2 de Mack,

$$\mathbb{E} \left[\frac{C_{i,j+1}}{C_{i,j}} \mid C_{i,1}, \dots, C_{i,j} \right] = f_j,$$

l'espérance de $\frac{C_{i,j+1}}{C_{i,j}}$ ne dépend pas de l'information passée.

En particulier, elle ne dépend pas de $C_{i,j} / C_{i,j-1}$.

⇒ On ne peut donc pas utiliser Mack dans 1 branche où l'on sait que $\frac{C_{i,j+1}}{C_{i,j}}$ sera – élevé si $\frac{C_{i,j}}{C_{i,j-1}}$ est + élevé que pr les autres années de survénance (et inversement). Pas de compensation donc.

Propriétés importantes

Sous H1 et H2 et conditionnellement au triangle

$$T = \{C_{i,j} : i + j \leq l \text{ et } i \leq l\}$$

des v.a. observables, on a un certain nb de résultats intéressants.

Propriété 1 : pour $i \geq 1$, $\mathbb{E}[C_{i,j} \mid T] = (f_{j-1} \dots f_{l-i}) C_{i,l-i}$.

L'hypothèse H2 donne même plus généralement :

$$\mathbb{E}[C_{i,j+1} \mid C_{i,1}, \dots, C_{i,n-i}] = f_j C_{i,j}.$$

On montre que les facteurs CL $\hat{f}_j = \sum_{i=0}^{l-j-1} C_{i,j+1} / \sum_{i=0}^{l-j-1} C_{i,j}$ sont des estimateurs sans biais des param. f_j , non corrélés (cf articles).

L'absence de biais se transmet aux valeurs calculées de montants à l'aide des facteurs chain-ladder standards :

$$\hat{C}_{i,n} = \hat{C}_{i,J} = C_{i,l-i} \hat{f}_{l-i} \dots \hat{f}_{J-1}$$

Plus généralement : $\hat{C}_{i,j} = C_{i,l-i} \hat{f}_{l-i} \dots \hat{f}_{j-1}$.

Du coup on a

$$\hat{R}_i = \hat{C}_{i,J} - C_{i,l-i}$$

et

$$\hat{R} = \sum_{i=1}^n \hat{R}_i$$

Au final, on estime donc **sans biais** la provision globale par la méthode de Mack.

Incertitude de l'estimateur

→ Pour mesurer l'incertitude de l'estimateur $\hat{\theta}$ d'un **PARAMETRE** θ , on utilise l'**erreur quadratique moyenne** (Mean Squared Error) :

$$MSE(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2].$$

Ici, on s'intéresse à la **MSEP** (Mean Square Error of Prediction) de la provision :

$$MSEP(\hat{R}_i) = \mathbb{E}[(\hat{R}_i - R_i)^2],$$

Dans notre contexte (conditionnel à T), on considère

$$MSEP(\hat{C}_{i,J}) = \mathbb{E}[(\hat{C}_{i,J} - C_{i,J})^2 \mid T].$$

Seul $C_{i,J}$ est aléatoire **grâce au conditionnement** ($\hat{C}_{i,J}$ étant estimé à partir des \hat{f}_j estimés sans biais et des obs.), donc

$$MSEP(\hat{C}_{i,J}) = \mathbb{V}[C_{i,J} | T] + [\mathbb{E}[C_{i,J} | T] - \hat{C}_{i,J}]^2.$$

On appelle le 1er terme la **process variance** (process error), alors que le 2e est l'**erreur d'estimation** (parameter error).

⇒ Pour estimer la MSEP d'une provision individuelle, on a donc besoin de spécifier la forme de $\mathbb{V}[\mathbf{C}_{i,J} | \mathbf{T}]$.

Remarque : $R_i - \hat{R}_i = C_{i,J} - \hat{C}_{i,J} \Rightarrow MSEP(\hat{R}_i) = MSEP(\hat{C}_{i,J})$.

Estimateur de la volatilité

Pour estimer cette MSEP, Mack introduit donc une hypothèse de plus sur la variabilité au sein du triangle :

H3 : pour $j = 0, \dots, J - 1$; il existe un paramètre σ_j^2 tel que :

$$\mathbb{V}[C_{i,j+1} | C_{i,1}, \dots, C_{i,j}] = \sigma_j^2 C_{ij} \quad (i = 0, \dots, I)$$

Peut aussi être transcrit sur les facteurs individuels de dév. :

$$\mathbb{V}[f_{i,j} | C_{i,1}, \dots, C_{i,j}] = \frac{\sigma_j^2}{C_{ij}} \quad (i = 0, \dots, I).$$

Interprétation : + montants sont grands, + variance des f_{ij} diminue.

Sous H1, H2 et H3 ; et en suivant la même démarche que pour les résultats sur \hat{f}_j , l'estimateur $\hat{\sigma}_j^2$ est un estimateur sans-biais de σ_j^2 :

$$\hat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} c_{i,j} \left(\frac{c_{i,j+1}}{c_{i,j}} - \hat{f}_j \right)^2 \quad j < J-1$$

- Cet estimateur compare les facteurs $f_{i,j} = \frac{c_{i,j+1}}{c_{i,j}}$ du d-triangle aux facteurs communs \hat{f}_j donnés par la méthode CL.
- Pour $j = J-1$ (un seul terme de variance impossible), prendre

$$\hat{\sigma}_j^2 = \min \left\{ \frac{(\hat{\sigma}_{J-2}^2)^2}{\hat{\sigma}_{J-3}^2}, \min(\hat{\sigma}_{J-2}^2, \hat{\sigma}_{J-3}^2) \right\}.$$

- σ^2 peut être vu comme la somme normalisée de résidus de Pearson standardisés.

Pour ce qui concerne l'incertitude sur la provision estimée, on a :

Théorème. en posant $\hat{C}_{i,l-i} = C_{i,l-i}$ et sous les hypothèses H1 à H3, $MSEP(\hat{R}_i)$ peut être estimée pour $i = 1, \dots, l$, par

$$M\hat{S}EP(\hat{R}_i) = \hat{C}_{i,J}^2 \sum_{j=l-i}^{J-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2} \left[\frac{1}{\hat{C}_{i,j}} + \frac{1}{\sum_{k=0}^{l-j-1} C_{k,j}} \right].$$

Rq : 1er terme = Process Risk, 2eme terme = Parameter Risk.

Pour estimer $\mathbf{MSEP}(\hat{R})$, on utilise l'estimateur suivant (qui intègre de la corrélation entre les estimateurs précédents) :

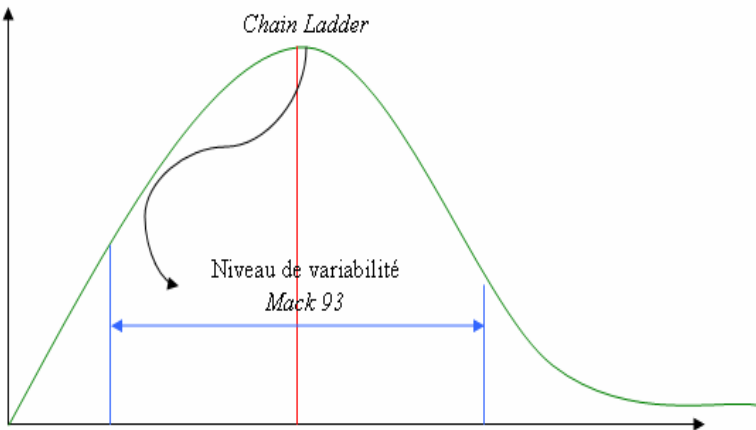
$$M\hat{S}EP(\hat{R}) = \sum_{i=1}^l \left\{ M\hat{S}EP(\hat{R}_i) + \hat{C}_{i,J} \left(\sum_{k=i+1, k < l}^l \hat{C}_{k,J} \right) \sum_{j=l-i}^{J-1} \frac{2\hat{\sigma}_j^2}{\hat{f}_j^2 \sum_{h=0}^{l-j-1} C_{h,j}} \right\}.$$

Densité de probabilité

Best Estimate
Chain Ladder

Niveau de variabilité
Mack 93

Montant des
sinistres



Comment vérifier les hypothèses de Mack ?

A t'on des moyens pratiques de vérifier la validité des hypothèses ?

- **H1** : le principal obstacle à l'⊥ entre années d'origine serait un effet diagonal. On doit donc tester l'existence d'un tel effet.
- **H2** : test de corrélation sur les rangs des obs. car H2 :

$$\text{Corr}(f_{i,j-1}, f_{i,j}) = 0 \Leftrightarrow \text{Corr}\left(\frac{C_{i,j}}{C_{i,j-1}}, \frac{C_{i,j+1}}{C_{i,j}}\right) = 0 \quad (i = 0, \dots, l).$$

- **H3** : forme de la variance et de l'espérance à vérifier :
 - pr l'espérance : graphe des $C_{i,j+1}$ en fonction des $C_{i,j}$.
 - pr la variance : caractère aléatoire des résidus.

H1 : effet diagonal et $\perp\!\!\!\perp$ d'années d'origine

L' $\perp\!\!\!\perp$ entre années d'origine peut être remise en cause par

- changement dans le traitement des données,
- évolution de l'inflation...

⇒ **affecte plusieurs années de la même façon.**

Remarque : souvent, cette hyp. est violée à cause d'une agrégation trop grande des données... cf ex. commentaire

Notons D_j les éléments de la diagonale j :

$$D_j = \{C_{j,0}, C_{j-1,1}, \dots, C_{0,j}\} \quad 0 \leq j \leq J$$

Introduisons les facteurs de développement (qui dépendent des éléments de D_j , ces derniers figurant au dénominateur...) :

$$A_j = \left\{ \frac{C_{j,1}}{C_{j,0}}, \dots, \frac{C_{0,j+1}}{C_{0,j}} \right\}.$$

→ Si les éléments de D_j sont + grands que d'habitude, alors les éléments de A_j seront + petits.

Pour chaque colonne k , on marque les facteurs de développement

- d'un **G** s'ils sont plus grands que la médiane de la colonne,
- d'un **P** s'ils sont plus petits que cette médiane,
- on notera une **étoile qui correspond à la médiane lorsque le nombre d'éléments de la colonne est impair.**

Pour chaque diagonale ($1 \leq j \leq J - 1$), on compte alors

- G_j le nombre d'éléments G de la diagonale considérée,
- P_j le nombre d'éléments P de cette même diagonale.

→ **S'il n'y a pas de changement entre année calendaire, le nombre de G_j et de P_j doit être similaire** puisque chaque facteur de développement a une chance de 50% d'excéder la médiane. Autrement dit, $Z_j = \min(P_j, G_j) \simeq (G_j + P_j)/2$.

Pour élaborer un test, on doit déterminer les 2 premiers moments de Z_j . Dans le cas de l'hyp. H_0 (non-effet d'une année calendrier),

$$P_j \sim \mathcal{B}(n_j = G_j + P_j, 0.5).$$

idem pour G_j .

On peut montrer pour la loi du min que $\mathbb{E}[Z_j] = \frac{n_j}{2} C_{m_j}^{n_j-1} \frac{n_j}{2^{n_j}}$, et

$$\mathbb{Var}[Z_j] = \frac{n_j(n_j - 1)}{4} - C_{m_j}^{n_j-1} \frac{n_j(n_j - 1)}{2^{n_j}} + \mathbb{E}[Z_j] - (\mathbb{E}[Z_j])^2$$

où m_j est la partie entière de $(n_j - 1)/2$.

Pour considérer ttes les diag., on introduit $Z = Z_2 + \dots + Z_{n-1}$, avec $\mathbb{E}[Z] = \sum \mathbb{E}[Z_j]$ et $\mathbb{Var}[Z] = \sum \mathbb{Var}[Z_j]$.

Comme Z est **somme de v.a. i.i.d.**, on applique le T.C.L. et **asymptotiquement Z suit une loi normale.**

⇒ On peut construire un IC : nous ne rejetons pas l'hyp. d'effet non significatif (au seuil 5%) d'une année calendaire dès lors que

$$\mathbb{E}[Z] - 2 \sqrt{\mathbb{Var}[Z]} \leq Z \leq \mathbb{E}[Z] + 2 \sqrt{\mathbb{Var}[Z]}.$$

H2 : non-corrélation des coef. de dév. successifs

Un test de corrélation est nécessaire pour valider cette hypothèse : en l'occurrence ce sera le **test de Spearman**.

En effet, on veut valider le fait que le ratio $\frac{C_{i,k+1}}{C_{i,k}}$ soit le même quelque soit la valeur précédente $\frac{C_{i,k}}{C_{i,k-1}}$, qui peut être plus grande ou plus petite que “d’habitude”.

Pour cela, pour chaque colonne k ,

- 1 on calcule le rang r_{ik} de chaque facteur individuel de développement,
- 2 on retire ensuite le dernier coefficient de la colonne (en bas), et on retrié en notant s_{ik} le nouveau rang des facteurs de développement restants ;

③ on définit alors la **statistique** (rho de Spearman) :

$$T_k = 1 - 6 \sum_{i=1}^{n-k} \frac{(r_{ik} - s_{ik})^2}{(n-k)^3 - n + k} \quad (2 \leq k \leq n-2)$$

On peut alors montrer que

- $-1 \leq T_k \leq 1$,
- sous l'hypothèse de non-corrélation :

$$\mathbb{E}[T_k] = 0 \quad \text{Var}[T_k] = \frac{1}{n-k-1}.$$

On veut tester l'hypothèse sur l'ensemble du triangle et non pas par paires de colonnes d'où l'utilisation de la statistique globale :

$$T = \sum_{k=2}^{n-2} \frac{n-k-1}{\sum_{k=2}^{n-2} n-k-1} T_k = \sum_{k=2}^{n-2} \frac{n-k-1}{(n-2)(n-3)/2} T_k$$

C'est une moyenne pondérée des T_k , par des poids inversement proportionnels à la variance !

Ainsi $\mathbb{E}[T] = 0$ et $\mathbb{V}ar[T] = \frac{2}{(n-2)(n-3)}$.

Si $n - k \geq 10$, la distribution de T_k est **approx. gaussienne**.
Or, T est la somme de variables **i.i.d.** T_k .

Par le T.C.L., T est donc aussi asymptotiquement normale !

Etant donné que le test n'est qu'approximatif et qu'on veut détecter des corrélations dans le triangle entier, on ne considère pas un IC à 95% \Rightarrow celui-ci n'aurait que peu de chance de rejeter l'hyp. de non-corrélation.

On choisit donc un intervalle à 50%, ce qui signifie qu'on ne rejettera pas l'hypothèse de non-corrélation dès lors que

$$\frac{-0.67}{\sqrt{(n-2)(n-3)/2}} \leq T \leq \frac{0.67}{\sqrt{(n-2)(n-3)/2}}$$

Construction d'intervalles de confiance

H1-H3 ne permettent pas de déterminer la description complète de la distribution R des provisions.

Mais on a des **IC en supposant une loi pour la provision** R , dont espérance/variance seront **fixés = aux estimations des moments** :

- si $R \sim \mathcal{N}(\mu, \sigma)$ (IC au seuil 5%) :

$$IC_{95\%}(R) = [\hat{R} - 1.96 \text{ SEP}(\hat{R}); \hat{R}_i + 1.96 \text{ SEP}(\hat{R})].$$

- si $R \sim \mathcal{LN}(\mu, \sigma)$ ($\mu = \ln \hat{R} - \frac{\sigma^2}{2}$, $\sigma^2 = \ln(1 + M\hat{SEP}(\hat{R})/\hat{R}^2)$) :

$$IC_{95\%}(R) = [\exp(\mu - 1.96\sigma); \exp(\mu + 1.96\sigma)],$$

- 3 **Approches stochastiques de provisionnement**
 - Introduction
 - Le modèle de Mack
 - **Munich Chain Ladder**
 - Approches factorielles et extension par GLM
 - Incertitude d'estimation sur la provision

Amélioration : le Munich Chain Ladder

[Quarg and Mack, 2008]

La méthode de Munich Chain Ladder est relativement récente, puisqu'elle a vu le jour en 2003 dans un colloque ASTIN.

Particularité : permet le traitement du problème crucial de non-convergence

- des provisions CL issues du triangle des paiements d'un côté,
- et de celui des charges d'autre part.

On ne la présente pas ici, mais il faut savoir que cela \exists ...

- 3 **Approches stochastiques de provisionnement**
 - Introduction
 - Le modèle de Mack
 - Munich Chain Ladder
 - **Approches factorielles et extension par GLM**
 - Incertitude d'estimation sur la provision

Modèles factoriels stochastiques

Ces modèles sont basés sur les **paiements non-cumulés**.

Données stockées ds le vecteur X des incréments $(X_{ij})_{i,j=0,\dots,n}$, qui satisfont une hypothèse fondamentale d'indépendance :

$$H : \forall i, j = 0, \dots, n, \text{ les } X_{ij} \text{ sont } \perp\!\!\!\perp .$$

Ceci implique que

- la compensation (recours) de règlements de sinistres d'un délai au suivant pour une même année d'origine est exclu ;
- cette $\perp\!\!\!\perp$ est invalidée par des variations annuelles d'inflation des montants, introduisant des corrélations entre X_{ij} .

En contrepartie,

- on a des tests statistiques qui jugent du **caractère réaliste de l'hypothèse d' \perp** sachant les réalisations du triangle sup. T
- on peut **parfois estimer toutes les caractéristiques de la loi de R** , y compris sa distribution, avec ce type de modèle.

→ Ces modèles sont majoritairement **paramétriques** : on définit une loi de proba. pour les v.a.r. X_{ij} .

→ Tous les X_{ij} sont munis du **même type de loi** (mais auront un paramètre \neq évidemment, noté θ : paramètre de tendance)..

→ Certains modèles font parfois intervenir un paramètre additionnel (dit de dispersion), ϕ .

Avantage : on peut intégrer une variable exogène adaptée à la problématique en pratique (car modèle de régression).

→ Modèles explicatifs très proches de modèles de tarif. IARD.

Notations :

- $\mu_{ij} = \mathbb{E}[X_{ij}] = \mathbb{E}[X_{ij} | i, j]$: espérance de X_{ij} . Des facteurs de risque (covariables) influent cette moyenne (cf plus loin).
- $f(\mu_{ij}, \phi)$ est la densité de la loi des X_{ij} : $X_{ij} \sim f(\mu_{ij}, \phi)$.

On rappelle le calcul des provisions (la provision du premier exercice ($i=0$) est nulle puisqu'on a tout observé !) :

$$R_i = \sum_{j=J-i+1}^J X_{ij} \qquad R = \sum_{i=1}^I R_i$$

Rq : théoriquement, on pourrait trouver la loi de R par convolution des lois des X_{ij} , car sous (H) c'est une somme de v.a.r. $\perp\!\!\!\perp$.

En pratique, la convolution de $\frac{n(n+1)}{2}$ variables est complexe. On peut **passer par la FGM** de R car

$$M_R(s) = \mathbb{E}[e^{s(\sum_{i=1}^n R_i)}] = \prod_{i=1}^n M_{R_i}(s) = \prod_{i=1}^n \prod_{i+j>n} M_{X_{i,j}}(s).$$

La FGM nous permet de générer les moments, avec

$$\mathbb{E}[R] = \sum_{i=1}^n \sum_{i+j>n} \mathbb{E}[X_{i,j}] \quad \mathbb{V}ar[R] = \sum_{i=1}^n \sum_{i+j>n} \mathbb{V}ar[X_{i,j}]$$

et ainsi de suite pour les moments d'ordre supérieur...

Modélisation des X_{ij} : l'usage des GLM

[McCullagh and Nelder, 1989]

On a recours aux modèles GLM, **très populaire en actuariat...**

Ces modèles économétriques permettent de

- modéliser des réponses de \neq **domaines de déf.** : $\mathbb{R}, \mathbb{R}^+, \mathbb{N}$;
- intégrer toute type d'**information exogène** influant sur réponse,
- **quantifier l'impact des facteurs de risque.**

Nécessite d'introduire 2 hypothèses fondamentales :

- individus (ici les incréments $X_{i,j}$) \perp entre eux ;
- **variables explicatives** (survenances et délais ici) \perp 2 à 2.

Composants d'un GLM

- 1 La **loi de la réponse aléatoire** X_{ij} : par hyp. elle \in à une *distribution de la famille exponentielle*.

- 2 Le **prédicteur** θ_{ij} est linéaire et déterministe : par ex.

$$\theta_{ij} = \mu + \alpha_2 \mathbb{1}_{i=2} + \dots + \alpha_n \mathbb{1}_{i=n} + \beta_2 \mathbb{1}_{j=2} + \dots + \beta_n \mathbb{1}_{j=n}$$

- 3 La **fonction de lien** g : monotone, dérivable, inversible t.q.

$$g(\mathbb{E}[X_{ij} | i, j]) = g(\mu_{ij}) = \theta_{ij}.$$

Modèle linéaire : $g = \text{Id}$ $\theta_{ij} = \mu + \alpha_i + \beta_j$ $X_{ij} \sim \mathcal{N}(\theta_{ij}, \phi = \sigma^2)$.

Les variables explicatives

- **L'effet période** (année d'origine) : considéré comme catégoriel (ordinaire ou non), à n modalités $(1, \dots, n)$.
 - Une des modalités sert de référence (ex. 1),
 - les autres sont représentées par les indicatrices d'appartenance à une modalité, d'où $(n - 1)$ param. $\alpha_2, \dots, \alpha_n$.
- **L'effet ancienneté** (ou délai) : naturellement quantitatif, mais considéré ici comme qualitatif/catégoriel. On a le choix entre
 - un facteur à n modalités ($\beta_1 = 0$, modalité de réf. par ex.) ;
 - une variable quantitative à valeurs entières : effet sur les $X_{i,j}$ modélisé par fonctions paramétrées de j (βj , $\beta \ln(j + 1)$, ...)
- **L'effet année calendaire** (diagonal) : mobiliserait $2(n - 1)$ paramètres μ_{i+j} .
 - Pr un triangle à inflation cste, paramétrage se limite : $\mu_{i+j} = \mu$.
 - Inexistant en pratique dans les modélisations.

Caractérisation d'un GLM par sa fonction variance

Parfois, on caractérise un GLM par sa fonction variance plutôt que par sa distribution. La fonction variance $V(\cdot)$ est définie par

$$\text{Var}(X_{ij}) = \phi V(\mu_{ij}),$$

où $\mu_{ij} = \mathbb{E}[X_{ij}]$.

Voici quelques fonctions variance pour des modèles classiques :

- $V(\mu_{ij}) = 1$ pour le modèle gaussien,
- $V(\mu_{ij}) = \mu_{ij}$ pour le modèle log-Poisson,
- $V(\mu_{ij}) = \mu_{ij}^2$ pour le modèle Gamma.

Exemples de GLM utilisés en provisionnement

- **Gaussienne** : $X_{ij} \sim \mathcal{N}(\theta_{ij}, \phi)$, $\mu_{ij} = \theta_{ij}$, $\phi = \sigma^2$, $V(\mu_{ij}) = 1$
 - si lien multiplicatif : $\theta_{ij} = x_i y_j \Rightarrow$ moindres carrés de De Vylder
 - si lien additif et application aux log-incréments \Rightarrow extension vers le modèle lognormal.
- **Loi Gamma** pour X_{ij} : $\mathbb{E}[X_{i,j}] = \mu_{ij}$, $\phi = 1$, $V(\mu_{ij}) = \mu_{ij}^2$;
 - on retombe sur le premier modèle de Mack (1991) ;
 - donne des résultats similaires à CL.
- **Loi de Poisson** pour X_{ij} : $\mathbb{E}[X_{i,j}] = \mu_{ij}$, $\phi = 1$, $V(\mu_{ij}) = \mu_{ij}$;
 - avec un lien *log* \rightarrow modèle log-Poisson (cf [Verrall, 1991]),
 - les provisions estimées coïncident avec celles de CL.
 - résultat similaire avec le log-Poisson surdispersé ($\phi > 1$), sauf que $V(\mu_{ij}) = \mu_{ij}(1 + \phi\mu_{ij})$.

Exemple de la régression Lognormale

Dans cet esprit des modèles factoriels, la régression lognormale est l'un des modèles pionniers.

- Tout premier modèle stochastique, né de Kremer en 1982.
- Introduit pr l'analyse de provisions de sinistres (X_{ij} **positif**).
- On y utilise le **modèle linéaire gaussien** suite à une transfo.

On suppose que

$$X_{ij} \sim \mathcal{LN}(\theta_{ij}, \sigma^2),$$

où σ^2 est un paramètre de dispersion.

→ Par définition, on sait que $Y_{i,j} = \ln(X_{i,j}) \sim \mathcal{N}(m_{ij}, \sigma^2)$ avec

$$\mathbb{E}[Y_{i,j}] = m_{ij} = \mu + \alpha_i + \beta_j.$$

(De manière équivalente, $Y_{i,j} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$, $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$)

En utilisant la transformée de Laplace, on obtient

$$\begin{cases} \theta_{ij} = \mathbb{E}[X_{i,j}] = e^{m_{ij} + \sigma^2/2} = e^{\mu + \alpha_i + \beta_j + \sigma^2/2}, \\ \text{Var}[X_{i,j}] = \mathbb{E}[X_{i,j}]^2 (e^{\sigma^2} - 1) = e^{2m_{ij} + \sigma^2} (e^{\sigma^2} - 1). \end{cases}$$

Le modèle de Mack vu comme un GLM

Il s'agit en fait de supposer une distribution gaussienne pour les montants cumulés...

$$D_{i,j+1} | D_{i,j} \sim N(\lambda_j D_{i,j}, \sigma_j^2 D_{i,j})$$

$$\lambda_{i,j} | D_{i,j} \sim N(\lambda_j, \frac{\sigma_j^2}{D_{i,j}})$$

$$\ln(\lambda_j) = \gamma_j$$

- Hypothèse sur la distribution des montants cumulés,
- Modèle récursif,
- γ_j estimés par maximum de vraisemblance,
- Estimation séparée de σ_j^2 .

Le modèle de Poisson surdispersé

Il s'agit de considérer une variable aléatoire surdispersée.
Pour un modèle de Poisson classique,

$$X \sim \mathcal{P}(\lambda) \quad \Rightarrow \quad \mathbb{E}[X] = \mathbb{V}(ar(X) = \lambda.$$

Ici, il s'agit de considérer le modèle

$$X \sim \mathcal{P}_{surd}(\lambda, \phi) \quad \Leftrightarrow \quad \frac{X}{\phi} \sim \mathcal{P}\left(\frac{\lambda}{\phi}\right)$$

et donc

$$\mathbb{E}[X] = \lambda \quad \text{et} \quad \mathbb{V}ar(X) = \phi\lambda = \phi\mathbb{E}[X].$$

- 3 **Approches stochastiques de provisionnement**
 - Introduction
 - Le modèle de Mack
 - Munich Chain Ladder
 - Approches factorielles et extension par GLM
 - Incertitude d'estimation sur la provision

Estimation de la réserve globale

Dans ce type de modèle, on a donc

$$\widehat{\mathbb{E}[R]} = \sum_{i=1}^I \sum_{j=J-i+1}^J \widehat{\mathbb{E}[X_{i,j}]} = \sum_{i=1}^n \sum_{j=n-i+1}^n \hat{\mu}_{ij}$$

C'est le **best estimate** de la provision.

On estime par **max. de vraisemblance** sur les observations du triangle supérieur, en notant dans la suite

- $L[\theta; (x_{ij})]$ la vraisemblance de l'échantillon ;
- $\theta = (\delta, (\alpha_i)_{i=1,\dots,n}, (\beta_j)_{j=1,\dots,n})$ le vecteur des paramètres à estimer (qui intègre éventuellement ϕ).

On note $\hat{\theta} = (\hat{\delta}, (\hat{\alpha}_i)_{i=1,\dots,n}, (\hat{\beta}_j)_{j=1,\dots,n})$ l'EMV (avec toutes ses propriétés asymptotiques classiques).

Par **invariance fonctionnelle** de la technique du max. de vraisemblance, on en déduit les EMV de $\mu_{ij} = \mathbb{E}[X_{ij}]$:

$$\hat{\mu}_{ij} = g^{-1}(\hat{\delta}, \hat{\alpha}_i, \hat{\beta}_j).$$

Ainsi, l'EMV de $\mathbb{E}[R_i]$ est donné par $\hat{\mathbb{E}}[R_i] = \sum_{j=J-i+1}^J \hat{\mu}_{ij}$;

et celui de $\mathbb{E}[R]$ vaut $\hat{\mathbb{E}}[R] = \sum_{i=1}^I \sum_{j=J-i+1}^J \hat{\mu}_{ij}$.

Risque d'estimation - évaluation

Voir l'article de [?].

Lorsqu'on veut estimer l'incertitude sur l'estimation de R , on a besoin de **connaître les covariances** entre les différents μ_{ij} car

$$\begin{aligned}\mathbb{V}ar\left[\hat{\mathbb{E}}[R_i]\right] &= \mathbb{V}ar\left[\sum_{j=n-i+1}^n \hat{\mu}_{ij}\right] \\ &= \sum_j \mathbb{V}ar[\hat{\mu}_{ij}] + \sum_{j_1} \sum_{j_2} \text{Cov}[\hat{\mu}_{ij_1}, \hat{\mu}_{ij_2}]. \\ \mathbb{V}ar\left[\hat{\mathbb{E}}[R]\right] &= \mathbb{V}ar\left[\sum_{i=1}^n \hat{\mathbb{E}}[R_i]\right] \\ &= \sum_{i=1}^n \sum_{j=n-i+1}^n \mathbb{V}ar[\hat{\mu}_{ij}] \\ &\quad + \sum_{i_1} \sum_{i_2} \sum_{j_1=n-i_1+1}^n \sum_{j_2=n-i_2+1}^n \text{Cov}[\hat{\mu}_{i_1 j_1}, \hat{\mu}_{i_2 j_2}].\end{aligned}$$

Pb : le calcul exact de ces valeurs est svt problématique...

→ On pourra alors approcher ces quantités par \neq méthodes.

① Techniques de **reéchantillonnage** :

- bootstrap,
- ...

② **Expressions analytiques asymptotiques** :

- basées sur la méthode Delta,
- nécessite suffisamment de données.

→ La méthode Delta ne fait intervenir que des produits matriciels, d'où sa facilité d'utilisation...

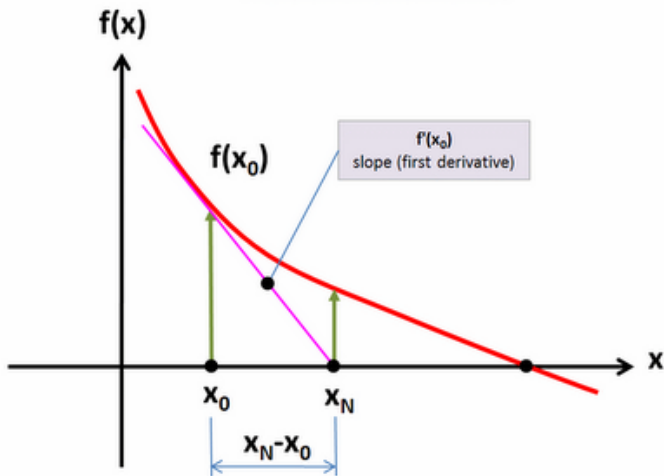
→ Ces deux méthodes donnent de plus des intervalles de confiance pour la provision globale $\mathbb{E}[\mathbf{R}]$!

1ère approche : la méthode Delta

Par définition, $\hat{\theta} = (\hat{\mu}, (\hat{\alpha}_i), (\hat{\beta}_j))$ est solution du système qui annule les dérivées partielles d'ordre 1 de la log-vraisemblance :

$$\begin{cases} \frac{\partial \ln L}{\partial \mu} = 0 \\ \frac{\partial \ln L}{\partial \alpha_i} = 0 & (i = 1, \dots, n) \\ \frac{\partial \ln L}{\partial \beta_j} = 0 & (j = 1, \dots, n). \end{cases}$$

Ce système se résout par l'utilisation d'algo. t.q. **Newton-Raphson** (inspirée de la formule de Taylor-Lagrange) : $\hat{\theta}_{k+1} = \hat{\theta}_k - \frac{L'(\hat{\theta}_k)}{L''(\hat{\theta}_k)}$.



Objectif : trouver le 0 d'une fonction...Déplacement dans le sens opposé du gradient.

Du coup (cf formule Newton-Raphson), on se retrouve à inverser la matrice d'information de Fisher (Hessienne en la densité des obs.) :

$$I(\theta) = \left(\frac{\partial^2 \ln L}{\partial \theta^2} \right).$$

L'inversion de cette dernière matrice n'est rien d'autre que la matrice de covariance asymptotique de l'estimateur $\hat{\theta}$!

Conséquences :

- sur la diag. la variance des estimateurs : $\mathbb{V}(\hat{\mu})$, $\mathbb{V}(\hat{\alpha}_i)$, $\mathbb{V}(\hat{\beta}_j)$;
- les autres termes sont les covariances entre ces estimateurs :

$$\text{Cov}[\hat{\mu}\hat{\alpha}_i]$$

$$\text{Cov}[\hat{\mu}, \hat{\beta}_j]$$

$$\text{Cov}[\hat{\beta}_j, \hat{\alpha}_i]$$

Sous certaines conditions de régularité de la vraisemblance, on sait que $\hat{\theta}$ possède de bonnes propriétés :

- il est **asymptotiquement efficace** (sans biais et de variance minimale),
- il est **asymptotiquement gaussien**.

On déduit de tous ces résultats les variances asymptotiques des best estimates des provisions...

Pour cela on utilise la méthode Delta, utile lors de l'introduction d'une fonction appliquée à l'estimateur dont nous disposons.

Application de la méthode Delta

Exemple du modèle additif

① Par invariance fonctionnelle de l'EMV : $\hat{\eta}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$.

② On définit l'application $\eta : \mathbb{R}^{2n-1} \rightarrow \mathbb{R}^{n^2}$ telle que

$$\theta \rightarrow \eta(\theta) = (\eta_{ij}).$$

③ La **jacobienne** J_η de cette application vaut pour l'indice (i,j)

$$\frac{\partial \eta_{ij}}{\partial \mu} = 1, \quad \frac{\partial \eta_{ij}}{\partial \alpha_k} = \begin{cases} 1 & \text{si } k = i \\ 0 & \text{sinon} \end{cases}, \quad \frac{\partial \eta_{ij}}{\partial \beta_l} = \begin{cases} 1 & \text{si } l = j \\ 0 & \text{sinon} \end{cases}.$$

- ④ Par la méthode **Delta**, on sait que $\hat{\eta} = (\hat{\eta}_{ij})$ est asymptotiquement normal, centré en η et de matrice de cov.

$$\Sigma(\hat{\eta}) = J_{\eta} \Gamma^{-1}(\theta) J_{\eta}^T$$

où J_{η}^T est la transposée de J_{η} .

On suppose la fonction de lien **g** telle que $\mu_{ij} = g(\eta_{ij})$ avec :

- g dérivable ;
- g strictement monotone.

Notons **D la jacobienne de la transformation** de η_{ij} en μ_{ij} par **g**. Cette matrice est diagonale d'éléments $g'(\eta_{ij})$.

Par suite, on obtient l'estimateur $\hat{\mu} = (\hat{\mu}_{ij})$, asymptotiquement gaussien noté $\hat{\mu} \sim \mathcal{N}(\mu, \Sigma(\hat{\mu}))$ avec $\Sigma(\hat{\mu}) = D \Sigma(\hat{\eta}) D^T$.

Puisque

$$\mathbb{E}[R_i] = \sum_{j=n-i+2}^n \mu_{ij},$$

on déduit que la **jacobienne** J_μ de cette transformation a en ligne i

$$\frac{\partial \mathbb{E}[R_i]}{\partial \mu_{kl}} = \begin{cases} 0 & \text{si } k \neq i \\ 0 & \text{si } k = i, l \leq n - i \\ 1 & \text{si } k = i, l > n - i \end{cases}.$$

On en déduit que pour tout $i = 1, \dots, n$, les variables aléatoires

$$\hat{\mathbb{E}}[R_i] \sim \mathcal{N}(\mathbb{E}[R_i], \Sigma(\hat{\mathbb{E}}[R_i])),$$

avec $\Sigma(\hat{\mathbb{E}}[R_i]) = J_\mu \Sigma(\hat{\mu}) J_\mu^T$.

On procède identiquement pour avoir l'estimateur $\hat{\mathbb{E}}[R]$ de $\mathbb{E}[R]$ car

$$\mathbb{E}[R] = \sum_{i=1}^n \mathbb{E}[R_i].$$

De la normalité asymptotique de $\hat{\mathbb{E}}[R]$, on déduit un **intervalle de confiance** pour $\mathbb{E}[R]$ à un certain niveau de confiance $1 - \alpha$:

$$IC_{1-\alpha} = \left[\hat{\mathbb{E}}[R] - q_{1-\alpha/2} \Sigma(\hat{\mathbb{E}}[R]); \hat{\mathbb{E}}[R] + q_{1-\alpha/2} \Sigma(\hat{\mathbb{E}}[R]) \right].$$

Distributions prédictives et mesures de risque

- Introduction
- Distribution prédictive par convolution
- Inversion de la fonction génératrice des moments
- Utilisation des moments
- Techniques de rééchantillonnage

Généralités

Comme nous l'avons vu, il ne suffit pas de connaître les moments de R . On aimerait connaître des quantiles de la loi de R .

Il faut donc estimer la loi de R : c'est la distribution prédictive.

On dispose globalement de trois moyens de l'obtenir :

- 1 la **convolution** des incréments X_{ij} , supposés \perp ;
- 2 l'approximation par spécification paramétrique d'une loi :
estimer les moments de R pour en déduire une **distribution prédictive approchée basée sur ces moments** ;
- 3 les simulations **Monte-Carlo** après rééchantillonnage (bootstrap) des résidus d'une modélisation.

La technique de convolution : additivité

Il est rare de pouvoir obtenir la loi prédictive de R par convolution directe si la loi générique des $X_{i,j}$ est additive.
C'est pourtant le cas avec les modèles suivants : Poisson, Poisson surdispersée, gaussien.

→ En effet, si les $X_{i,j} \sim \mathcal{P}(\mu_{ij})$ alors idem pour $R = \sum_i \sum_j X_{ij}$:

$$R \sim \mathcal{P}(\mu_R) \quad \text{avec} \quad \mu_R = \sum_{i=1}^n \sum_{j=n-i+1}^n \mu_{ij}.$$

→ On peut aussi utiliser l'**approximation normale** $\mathcal{N}(\mu_R, \mu_R)$ dans le cas où $\mu_R \geq 50$:

$$\mathbb{P}(R \leq r) = \Phi\left(\frac{r - \mu_R}{\sqrt{\mu_R}}\right).$$

Ainsi, on peut obtenir le quantile d'ordre $(1 - \eta)$ de R par

$$q_{1-\eta}(R) \simeq q_{1-\eta}^{(P)}(R) = \mu_R + \sqrt{\mu_R} q_{1-\eta},$$

où $q_{1-\eta}$ est la quantile de la loi normale centrée réduite.

On déduit des propriétés d'invariance des EMV que

$$\hat{q}_{1-\eta}^{(P)}(R) = \hat{\mu}_R + \sqrt{\hat{\mu}_R} q_{1-\eta}.$$

→ On peut encore appliquer ce raisonnement à la distribution de **Poisson surdispersée**. Ainsi,

$$X_{ij} \sim \mathcal{P}_{\text{surd}}(\mu_{ij}, \phi) \quad \Rightarrow \quad R \sim \mathcal{P}_{\text{surd}}(\mu_R, \phi),$$

avec $\mu_R = \sum_{i=1}^n \sum_{i+j > n} \mu_{ij}$.

Dans ce cas, $\frac{R}{\phi} \sim \mathcal{P}\left(\frac{\mu_R}{\phi}\right)$ et donc

$$\mathbb{P}(R \leq r) = \Phi\left(\frac{r - \mu_R}{\sqrt{\phi \mu_R}}\right),$$

puis

$$\hat{q}_{1-\eta}^{\text{surd}}(R) = \hat{\mu}_R + \sqrt{\hat{\phi} \hat{\mu}_R} q_{1-\eta}.$$

4

Distributions prédictives et mesures de risque

- Introduction
- Distribution prédictive par convolution
- Inversion de la fonction génératrice des moments
- Utilisation des moments
- Techniques de rééchantillonnage

Autre méthode : inversion de la FGM

Rappelons la densité d'une loi \in famille exponentielle :

$$f_X(x_{ij}; \theta, \phi) = \exp \left\{ \frac{x_{ij}\theta - b(\theta)}{a(\phi)} + c(x_{ij}, \phi) \right\}.$$

Et la FGM correspondante : $M_X(s) = \exp \{[b(\theta + s\phi) - b(\theta)]/\phi\}$.
La fonction génératrice des cumulants (FGC) est définie par

$$C_X(s) = \ln M_X(s) = \frac{1}{\phi} [b(\theta + s\phi) - b(\theta)].$$

→ **Idée** : appliquer ces résultats au triangle, sous réserve que la distribution des X_{ij} **admette une FGM** et qu'on ait **l'hyp. d'⊥**.

En effet, on peut déduire la FGM de R grâce à l'⊥ :

$$M_R(s) = \prod_{i=1}^n M_{R_i}(s) = \prod_{i=1}^n \prod_{i+j>n} M_{X_{ij}}(s).$$

En se remémorant que $\theta_{ij} = \theta(\mu_{ij}) = b'^{-1}(\mu_{ij})$,

$$M_R(s) = \exp \left\{ \frac{1}{\phi} \sum_{i=1}^n \sum_{i+j>n} (b[\theta(\mu_{ij}) + s\phi] - b(\theta(\mu_{ij}))) \right\}.$$

D'où FGC : $C_R(s) = \frac{1}{\phi} \sum_{i=1}^n \sum_{i+j>n} (b[\theta(\mu_{ij}) + s\phi] - b(\theta(\mu_{ij})))$.

On a un système à 2 inconnues et 2 équations : FGC et FGM servent à estimer $\hat{\mu}_{ij}$ et $\hat{\phi}$ par **maximum de vraisemblance**.

4

Distributions prédictives et mesures de risque

- Introduction
- Distribution prédictive par convolution
- Inversion de la fonction génératrice des moments
- **Utilisation des moments**
- Techniques de rééchantillonnage

Distributions approchées

Cette technique nécessite le **calcul des premiers moments de la charge sinistres** pour approcher la distribution de cette charge.

En actuariat, on utilise souvent une des 4 méthodes suivantes :

- approximation Normal-Power,
- transformation d'Esscher,
- loi Gamma traduite,
- loi Gamma Bowers.

Ces techniques sont utilisées pour obtenir des approximations fermées de la FdR de R . **Nous ne les développons pas ici.**

4

Distributions prédictives et mesures de risque

- Introduction
- Distribution prédictive par convolution
- Inversion de la fonction génératrice des moments
- Utilisation des moments
- Techniques de rééchantillonnage

Rappel sur le bootstrap

[Efron and Tibshirani, 1993]

Méthode de rééchantillonnage qui substitue à des calculs stats complexes l'usage des **simu Monte Carlo** sur l'échantillon originel

- Permet de réduire le biais d'un estimateur,
- Et permet d'en estimer sa variance.

Introduit par Efron en 1979, largement utilisé aujourd'hui.

→ Justification théorique plutôt complexe.

→ Mise en oeuvre simple car requiert uniquement des simu.

Ici on appliquera le bootstrap au cas de v.a. i.i.d.

Concept du bootstrap

On considère une variable aléatoire X dont la fonction de répartition, F_X , est inconnue.

On observe un échantillon i.i.d. (X_1, \dots, X_n) . Puisqu'on estime à partir d'un échantillon aléatoire, il y a une incertitude : l'estimateur

$$\theta(X_1, \dots, X_n)$$

est donc lui-même aléatoire (ex pour $\theta(X_1, \dots, X_n)$: moyenne).

Q : de combien varie $\theta(X_1, \dots, X_n)$ lorsque (X_1, \dots, X_n) varie ?

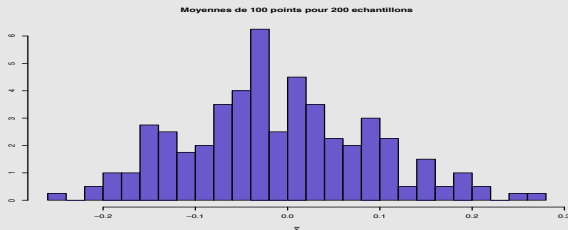
- quelle est la variance de $\theta(X_1, \dots, X_n)$?
- quelle est la distribution de $\theta(X_1, \dots, X_n)$?

Ex. : comment approcher la distribution de $\theta(\hat{F}_n)$?

Exemple (Échantillon normal)

Soit X_1, \dots, X_{100} un échantillon normal $\mathcal{N}(\theta, 1)$. Sa moyenne θ est estimée par

$$\hat{\theta} = \frac{1}{100} \sum_{i=1}^{100} X_i$$



Variation compatible avec la loi (connue) $\hat{\theta} \sim \mathcal{N}(\theta, 1/100)$

Ici $\theta(\hat{F}_n) = \hat{\theta}$ est l'EMV de θ , donc on sait que $\hat{\theta} \sim \mathcal{N}(\theta, 1/100)$.

Pourquoi le bootstrap ?

On peut estimer ce param. (ici moyenne), mais on ne connaît pas **l'incertitude de cette estimation** (sauf cas particulier comme ici car on pris une gaussienne et l'EMV).

Ceci est également lié aux difficultés suivantes :

- on observe **un seul** échantillon en général,
- la loi de l'échantillon est souvent **inconnue**,
- l'évaluation de la variation moyenne de $\theta(X_1, \dots, X_n)$ est essentielle pour la construction d'IC et de tests comme

$$H_0 : \theta \leq 0$$

⇒ Le bootstrap est souvent utilisé dans le cas de petits échantillons pour évaluer l'incertitude des estimateurs.

Notion fondamentale : CV de la FdR empirique

Par le théorème de **Glivenko-Cantelli**, un estimateur sans biais et convergent de F est la fonction de répartition empirique :

$$F_X^n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{X_i \leq x} \quad \longrightarrow \quad F(x) = \mathbb{P}(X \leq x).$$

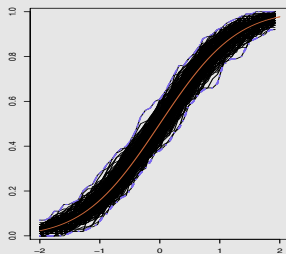
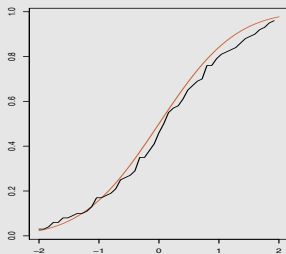
On distingue en général

- le bootstrap **non-paramétrique** : loi de F inconnue.
- le bootstrap **paramétrique** : loi connue, paramètre inconnu.

Rq : bootstrap + efficace si on connaît F et on cherche son param.,
mais on peut perdre en robustesse si mauvaise spécification !

CV de la FdR empirique

Exemple (Échantillon normal)



Estimation de la fonction de répartition F à partir d'un échantillon normal de 100 points et variation de cette estimation sur 200 échantillons normaux

Principe

Pour **toute fonctionnelle de la FdR**, de la forme

$$\theta(F) = \int h(x) dF(x),$$

on utilise l'approximation

$$\begin{aligned}\theta(\hat{F}) &= \theta(\hat{F}_n) \\ &= \int h(x) d\hat{F}_n(x) \\ &= \frac{1}{n} \sum_{i=1}^n h(X_i)\end{aligned}$$

qui correspond donc à un **estimateur des moments**.

Comment approcher la distribution de $\theta(\hat{F}_n)$?

Comme $\theta(\hat{F}_n) = \theta(X_1, \dots, X_n)$ où $X_1, \dots, X_n \stackrel{i.i.d.}{\sim} F$, remplace F par \hat{F}_n :

$$\theta(\hat{F}_n) \simeq \theta(X_1^*, \dots, X_n^*) \quad \text{avec } X_1^*, \dots, X_n^* \stackrel{i.i.d.}{\sim} \hat{F}_n.$$

→ \hat{F}_n étant connue, on peut simuler suivant \hat{F}_n , donc approcher la loi de $\theta(X_1^*, \dots, X_n^*)$ au lieu de celle de $\theta(\hat{F}_n) = \theta(X_1, \dots, X_n)$.

→ La loi de \hat{F}_n donne une probabilité de $1/n$ à chaque point de l'échantillon $\{x_1, \dots, x_n\}$:

$$\mathbb{P}^{\hat{F}_n}(X^* = x_i) = 1/n,$$

donc il suffit d'opérer des tirages avec remise dans (X_1, \dots, X_n) .

Distribution bootstrap d'un estimateur

Ici, B désigne le nombre d'échantillons bootstrap.

① Pour $b = 1, \dots, B$,

- générer un échantillon X_1^b, \dots, X_n^b suivant \hat{F}_n
- construire l'image correspondante :

$$\hat{\theta}^b = \theta(X_1^b, \dots, X_n^b)$$

② Utiliser l'échantillon

$$\hat{\theta}^1, \dots, \hat{\theta}^B$$

pour approcher la distribution de $\theta(X_1, \dots, X_n)$.

Mise en place pratique - Distribution de la médiane

Ici, B désigne toujours le nombre d'échantillons bootstrap. Je veux estimer la distrib. de l'estimateur de la médiane de mon échant..

① Pour $b = 1, \dots, B$,

- je génère un échantillon X_1^b, \dots, X_n^b suivant \hat{F}_n
- je calcule la médiane empirique sur mon échantillon bootstrap :

$$\hat{\theta}^b = \theta(X_1^b, \dots, X_n^b)$$

② J'ai un échantillon de médianes $\hat{\theta}^1, \dots, \hat{\theta}^B$ pour approcher la distrib. de la médiane théo. inconnue $\theta(X_1, \dots, X_n)$.

Je peux par ex. prendre la moyenne (**principe de Monte Carlo**) :

$$\theta_{bootstrap}(X) = \frac{1}{B} \sum_{b=1}^B \hat{\theta}^b.$$

Précision de l'estimateur Bootstrap

L'estimateur Bootstrap converge à la vitesse \sqrt{B} .

Cela veut dire que sa variance est donnée par

$$\text{Var}(\theta) = \frac{1}{B} \sum_{b=1}^B (\hat{\theta}^b - \bar{\theta})^2$$

Variation moyenne par procédure bootstrap

La variance d'un estimateur $\theta(X_1, \dots, X_n)$ est donnée par

$$\mathbb{V}ar(F) = \mathbb{E}_F \left[\left(\theta(X_1, \dots, X_n) - \mathbb{E}_F[\theta(X_1, \dots, X_n)] \right)^2 \right].$$

Cette variance est approchée dans le cas du bootstrap par

$$\mathbb{V}ar(\hat{F}_n) = \mathbb{E}_{\hat{F}_n} \left[\left(\theta(X_1, \dots, X_n) - \mathbb{E}_{\hat{F}_n}[\theta(X_1, \dots, X_n)] \right)^2 \right],$$

elle-même approchée par l'estimateur des moments

$$\hat{\mathbb{V}ar}(\hat{F}_n) = \frac{1}{B} \sum_{b=1}^B \left(\theta(X_1^b, \dots, X_n^b) - \bar{\theta} \right)^2 \quad \text{avec } \bar{\theta} = \frac{1}{B} \sum_{b=1}^B \theta(X_1^b, \dots, X_n^b)$$

Bootstrap et provisionnement par GLM

[England and Verrall, 2001b], [England and Verrall, 2001a]

On rappelle certaines notations :

$$X_{ij} \quad f(x_{i,j}) \quad \mu_{ij} = \mathbb{E}[X_{ij}] \quad \text{Var}[X_{ij}] = \phi V(\mu_{ij})$$

Covariables GLM : **années d'origine** et **délais de règlement**.

La technique consiste à appliquer le bootstrap aux résidus d'un modèle pertinent déjà calibré.

On applique le bootstrap **aux résidus** car les données X_{ij} sont \perp mais pas i.i.d. ! \rightarrow rééchantillonnage des résidus i.i.d. $(r_{ij})_{i+j \leq n}$.

Ici, résidus de **Pearson** :

$$r_{ij}^{(P)} = \frac{x_{i,j} - \widehat{\mathbb{E}[X_{ij}]}}{\sqrt{\widehat{\text{Var}(X_{ij})}}} = \frac{x_{i,j} - \hat{\mu}_{ij}}{\sqrt{\widehat{\text{Var}(X_{ij})}}}.$$

On peut considérer plus simplement

$$r_{ij}^{(P^*)} = \frac{x_{i,j} - \hat{\mu}_{ij}}{\sqrt{V(\hat{\mu}_{ij})}}$$

Ensuite, on construit l'**observation bootstrappée** x_{ij}^* , donnée par

$$x_{i,j}^* = \hat{\mu}_{ij} + r_{ij}^* \sqrt{V(\hat{\mu}_{ij})},$$

où r_{ij}^* est le résidu de Pearson bootstrappé.

Remarque : pas obligé d'estimer le param. ϕ (s'agit d'une constante de normalisation que l'on ne réinjecte pas ensuite...)

Quelques remarques importantes

→ La loi des résidus ne doit PAS dépendre des param. du modèle (on les appelle “pivots”) ! OK pr lognormale, Poisson et Gamma.

→ La dispersion ϕ , si constante pour le jeu de données, peut être omise ds le calcul des résidus et de l’observation bootstrappée.

→ En calibrant un GLM, 2 observ. seront = aux prévisions (car modèle construit sur une seule observ. !) : X_{0J} et $\hat{\mu}_{0J}$, ainsi que X_{J0} et $\hat{\mu}_{J0}$ \Rightarrow ces résidus n’étant pas réalistes, il ne faudrait pas les considérer dans le tirage des résidus bootstrappés...

→ Rigoureusement, les résidus de Pearson ne sont pas i.i.d. \Rightarrow on utilise donc parfois les résidus de Pearson **standardisés** (pas tjs !).

Application : distribution de la RESERVE MOYENNE

Exemple d'application au param. $\mathbb{E}[R]$ d'un triangle d'incrémentes $(x_{ij})_{i+j \leq n}$, une fois choisi le modèle GLM $\{\mu, (\alpha_i), (\beta_j)\}, \mu_{ij}, V, \phi$.

- 1 Estimation des coef. de régression, puis des prévisions $\hat{\mu}_{ij}$:

$$\hat{\mathbb{E}}[R] = \hat{R} = \sum_{i+j > n} \hat{\mu}_{ij}.$$

- 2 Calcul des résidus de Pearson sur le triangle supérieur (on peut aussi calculer param. dispersion $\phi \Rightarrow$ permet rescaler résidus en tenant compte du nb d'obs. et nb de param.) :

$$\hat{\phi} = \frac{1}{M - p} \sum_{i+j \leq n} \left(r_i^{(P)} \right)^2 \quad (M : \text{nb obs.}, p : \text{nb param.})$$

③ Procédure bootstrap : pour $b = 1, \dots, B$,

- obtention d'un échantillon bootstrap de résidus $(r_{ij}^{*b})_{i+j \leq n}$ par rééchantillonnage de l'ensemble des résidus initiaux ;
- détermination du triangle d'incrémentes bootstrappés $(x_{ij}^{*b})_{i+j \leq n}$

$$x_{ij}^{*b} = \hat{\mu}_{ij} + r_{ij}^{*b} \sqrt{\hat{\phi}V(\hat{\mu}_{ij})}.$$

- **pour le modèle retenu** et ces nouvelles données du triangle T, re-estimation du GLM et prévisions des $(\hat{\mu}_{ij}^{*b})_{i+j > n}$ pour

$$\hat{E}[R]^{*b} = \hat{R}^{*b} = \sum_{i+j > n} \hat{\mu}_{ij}^{*b};$$

- stockage de \hat{R}^{*b} ; puis $b \leftarrow b + 1$ et recommencer !

④ Utilisation du B-échantillon bootstrap $(\hat{R}^{*b})_{b=1, \dots, B}$ pour

- estimer le biais de $\hat{E}[R]$, IC pour $\hat{E}[R]$...

Exemple : bootstrap sur modèle de Mack (hyp. : réserve gaussienne)

En formulant l'algo. sur le modèle de Mack, on fera les étapes :

- ➊ Transition triangle d'incrémentes vers triangle cumulé.
- ➋ Application de CL standard : détermination des facteurs de développement et de la provision correspondante \hat{R} .
- ➌ Recalcul des cumulés initialement observés ($i + j \leq n$) par application rétrospective des facteurs de dév. à la diagonale calendaire ultime du triangle supérieur.
- ➍ Calcul des résidus de Pearson de ce modèle par

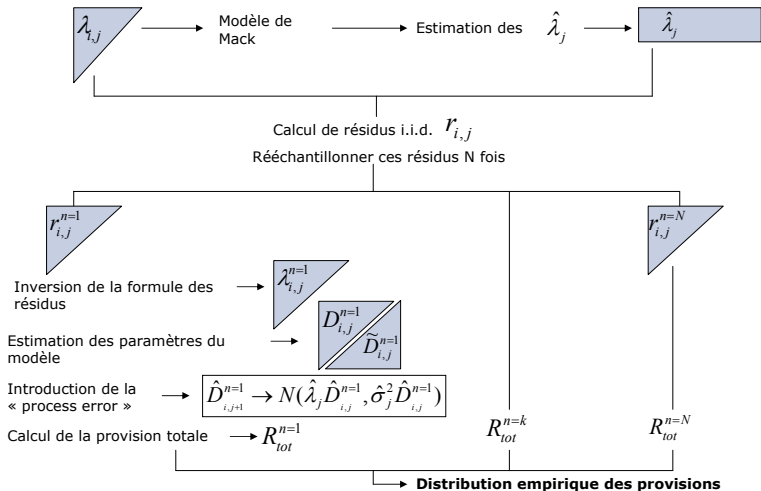
$$r_{ij}^{(P)} = \frac{X_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}}}$$

- ⑤ (Ajustement possible des résidus pr correction d'1 biais de prédiction, pour avoir une variance unitaire)

$$\sqrt{\frac{M}{M-p}} r_{ij}^{(P)}, \quad i+j \leq n, Mnbobs., knbparamGLM$$

- ⑥ Construction d'échantillons bootstrap des résidus ajustés de Pearson.
- ⑦ Reconstruction des triangles supérieurs d'incréments bootstrappés.
- ⑧ Calcul des nouveaux facteurs de développement.
- ⑨ Construction du triangle inférieur de montants cumulés bootstrappés.
- ⑩ Déduction de la **provision pr chaque échantillon bootstrap**.

Schématisation de la procédure (Mack)



Bootstrap sur modèle log-Poisson surdispersé

Prop. : les prévisions par GLM Poisson surdispersé coïncide exactement avec celles du CL \Rightarrow on peut appliquer le même algo !
Mack/ODP donnent la même estimation de la provision moyenne, mais les hypothèses de volatilité sont \neq et donc la SEP est \neq !

Algorithme simple avec des allers-retours cumulés-incréments :

- 1 Transition triangle d'incrémentes vers triangle cumulé, et application de CL standard : détermination des facteurs de développement et de la provision correspondante \hat{R} .
- 2 Calcul par application rétrospective des facteurs de dév. à la diagonale calendaire ultime des valeurs cumulées prévues en triangle supérieur. Déduction des incréments prévus.

- ③ Calcul des résidus de Pearson et de $\hat{\phi}$:

$$r_{ij}^{(P)} = \frac{X_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}}} \quad \text{et} \quad \hat{\phi} = \frac{1}{M - p} \sum_{i,j} (r_{ij}^{(P)})^2$$

- ④ Construction d'échantillons bootstrap des résidus ajustés, puis calcul des triangles sup. d'incréments bootstrappés.
- ⑤ Calcul des nouveaux facteurs de développement (en effet, on sait que le modèle ODP donne la même provision en moyenne que le CL, donc c'est plus rapide de faire comme pour Mack) : construction triangle inf. de montants cumulés bootstrappés.

⇒ Dédution **provision pour chaque échantillon bootstrap.**

Bootstrap - MSEP de la provision totale

Définissons l'erreur standard associée au bootstrap par

$$SE^b(\hat{\mu}) = \sqrt{\frac{1}{B} \sum_{b=1}^B (\hat{\mu}^b - \hat{\mu})^2},$$

avec $\hat{\mu}$ le MLE, et $\hat{\mu}^b$ l'estimateur d'un échantillon bootstrap.

Puis on en déduit l'erreur de prévision globale par la formule :

$$SEP^b(\hat{\mu}) = \sqrt{\hat{\phi}\hat{\mu} + \frac{M}{M-p} SE^b(\hat{\mu})^2}.$$

En général, $M = n(n-1)$ et $p = 2n-1$ (GLM).

5 Cas particuliers

Provisionnement multivarié, multi-branches

Prise en compte de la corrélation entre les branches : cf travaux de C. Genest.

Prise en compte des recours

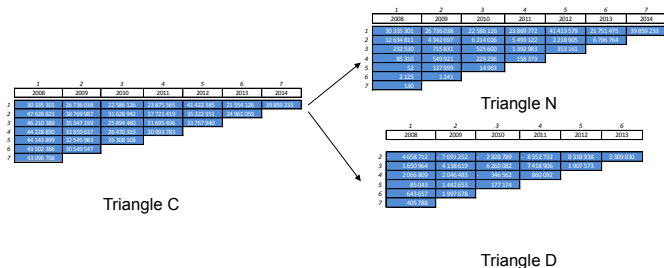
Facteurs de queue pour branche a developpement long

Quantification de l'incertitude

Cas de reouverture de sinistre

Méthode de Schnieper (1991)

Principe : décomposition des IBNeR (triangle D) et IBNyR (triangle N)...



$$C_{i,j} = C_{i,j-1} - D_{i,j} + N_{i,j}$$

Pour obtenir les triangles d'IBNeR et d'IBNyR, une base de données détaillée pour tous les sinistres et leurs évolutions est nécessaire : il faut pouvoir disposer de l'évolution de chaque sinistre de manière individuelle, contrairement à la méthode de

Sinistres graves : méthode de DHV

Même principe : décomposition des IBNR en IBNyR et IBNeR.

Projection des IBNeR : variante de Chain Ladder sinistre par sinistre.

Projection des IBNyR : méthode fréquence-coût, coût à partir des IBNeR et fréquence par Bornhuetter-Ferguson.

Rq : voir le mémoire d'actuariat d'Eric Gettler.

- 6 Calcul de SCR provisionnement à un an
 - Contexte et notions fondamentales
 - Le modèle de Merz & Wüthrich

SCR et ruine économique

Le SCR (**Solvency Capital Requirement**) est un montant à provisionner permettant de se prémunir contre une ruine économique pour l'année à venir dans 99,5% des cas.

Il est défini à partir des fonds propres (FP). + précisément, il faut que

$$FP - SCR > 0$$

dans un an dans 99,5% des cas.

Les provisions évaluées en Best Estimate évoluent dans le temps : cette variation vient modifier les FP dispo, donc impacte le SCR.

Le SCR est donc indirectement lié à un quantile de la distribution des pertes, et **directement lié au quantile de la distribution des FP** (qu'il faut donc pouvoir estimer...).

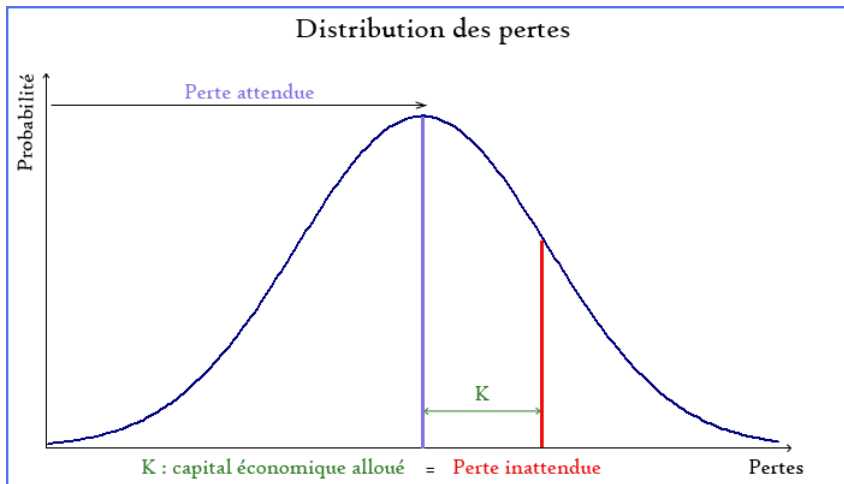
Notons L la variable aléatoire des pertes.

Pour évaluer le $SCR(L)$, on cherche ainsi

$$q_{0.995}(L) = \inf\{I : \mathbb{P}(L \leq I) \geq 99.5\% \}.$$

La loi demande de provisionner $BE + SCR$, donc en tout $q_{0.995}(L)$!
Dans ce cadre, le SCR correspond au capital économique.

SCR ~ capital économique



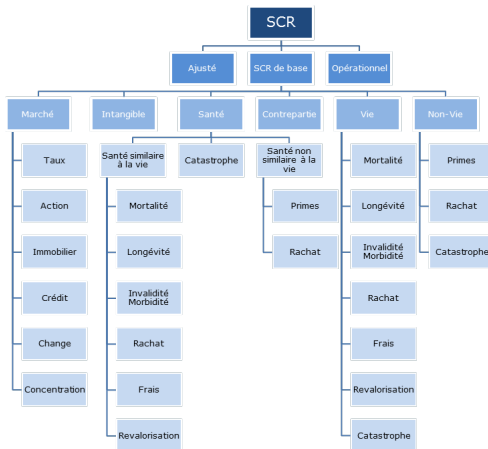
Formule standard ou modèle interne

Pour la mise en place de la réforme européenne Solvabilité 2 début 2016, on a recours à plusieurs méthodes pour évaluer le SCR :

- utiliser la **formule standard** fournie par les autorités ;
- développer un **modèle interne partiel** : i.e. utiliser un modèle interne pour certains modules de la matrice des risques, voire utiliser formule standard en modifiant des paramètres (USP)
- développer un **modèle interne** “total”.

L'Autorité de Contrôle Prudentiel et de Résolution (ACPR) est chargée de valider ces \neq approches au sein des entreprises.

Modules de risque sous Solvency II



Une matrice de corrélation (linéaire) permet d'agréger les SCR de chaque sous-module pour obtenir le SCR global.

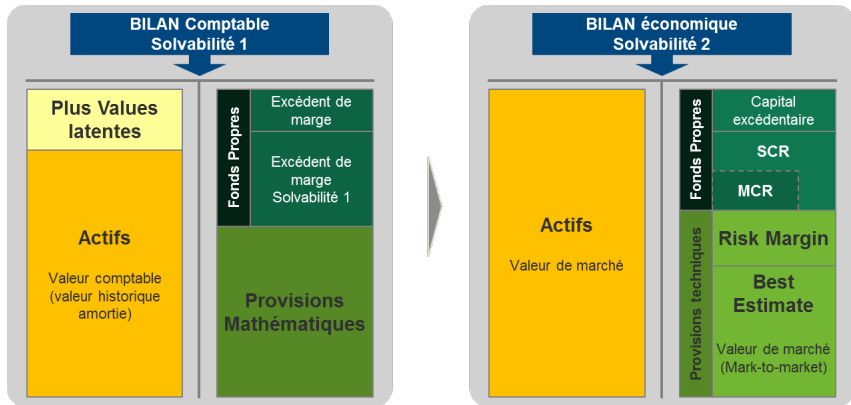
Bilan SII d'une compagnie

Dans le cadre de Solva 2, les passifs d'assurance non-vie doivent être évalués à leur juste valeur, définie comme la somme de

- la meilleure estimation (**Best Estimate**) : correspond à la valeur actuelle probable des flux de trésorerie futurs ;
- la marge pour risque (**Risk Margin**) : garantit que le montant global des provisions techniques soit équivalent à la somme que les assureurs devraient payer s'ils transféraient leurs droits/obligations contractuels à une autre entreprise.

Risque de provisionnement : incertitude sur l'estimation du Best Estimate (BE). Le BE varie de façon + ou – importante ds le temps, dû aux interactions actif-passif.

Schéma bilantiel



Le SCR : une notion prospective

Notion complexe car ce n'est pas le SCR aujourd'hui. Il faut donc étudier la déformation du bilan dans un an.

En effet, le SCR provisionnement est lié à la déformation du Best Estimate dans l'année à venir.

Il faut donc être capable de quantifier l'impact de cette déformation sur les FP dans un an : pour cela, on introduit le **Claims Development Result (CDR)** (cf ci-après).

- 6 Calcul de SCR provisionnement à un an
 - Contexte et notions fondamentales
 - Le modèle de Merz & Wüthrich

Le modèle de Merz & Wüthrich

[Wüthrich and Merz, 2008], [Merz and Wüthrich, 2008],
[Wüthrich et al., 2009], [Verrall and Wüthrich, 2012]

Ce modèle s'est imposé comme une **référence de marché**.

- L'un des 1^{er} modèles développés pour la mesure de ce risque.
- Très utilisé pr sa simplicité d'implémentation (formule fermée) ;
- Repris dans les textes réglementaires (cf QIS... etc).

→ **Extension de Mack, permettant le calcul du SCR provisions à un an.** Pour intégrer cette dimension de projection temporelle, il s'appuie sur un processus auto-régressif AR(1).

Le Claims Development Result (CDR)

Le CDR correspond à la \neq entre la prévision du montant final des paiements vus en date I et vus en date $I + 1$ (un an plus tard).

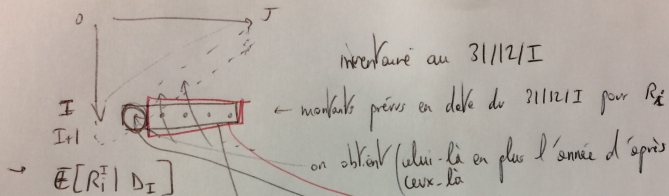
→ Permet de comparer la robustesse de la prévision avec ce qu'il s'est effectivement passé par la suite.

Ainsi, on introduit les notations suivantes : pour la provision de l'année de survenance i vue en $t = I$,

$$R_i^I = C_{i,J} - C_{i,I-i},$$

et vue en $t = I + 1$:

$$R_i^{I+1} = C_{i,J} - C_{i,I-i+1}.$$



$\boxed{i = I} :$

$$CDR_I(I+1) = \underbrace{E[R_I^I | D_I]}_{\text{}} - \left[\underbrace{\cancel{X_{I,I+1}}}_{X_{I,1}} + E[R_I^{I+1} | D_{I+1}] \right]$$

CDR - 1Y vs Ultimate

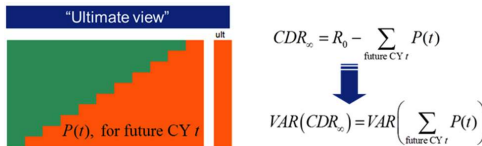


Figure 1: Ultimate View

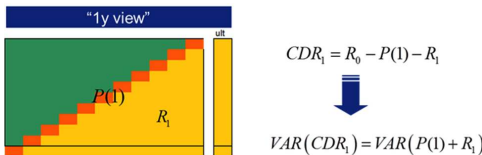


Figure 2: One Year View

- R_0 la provision de départ, connue,
- $P(t)$ les paiements de l'année calendaire t (mais run-off),
- R_1 la provision de cloture, après avoir observé $P(1)$.

Hypothèses et information en $t = I$ et $t = I + 1$

On définit les données observées en $t = I$ par le triangle supérieur T . Dans la suite, on note cette information D_I où

$$D_I = \{C_{i,j} : i + j \leq I \text{ et } i \leq I\}$$

En $t = I + 1$, cette information devient

$$D_{I+1} = \{C_{i,j} : i + j \leq I + 1 \text{ et } i \leq I\} = D_I \cup \{C_{i,I-I+1} : i \leq I\}$$

Les hypothèses de ce modèle sont les mêmes que celles de Mack.

Evaluation de provisions

Pour évaluer la provision, il faut estimer le montant moyen final des paiements :

$$\mathbb{E}[C_{i,J} | D_I] = C_{i,I-I} \prod_{j=I-I}^{J-1} f_j,$$

et

$$\mathbb{E}[C_{i,J} | D_{I+1}] = C_{i,I-I+1} \prod_{j=I-I+1}^{J-1} f_j$$

où les f_j sont les facteurs de développement.

Problème : il faut connaître les f_j . On les estimera par CL.

Estimation CL en $t = l$ et $t = l + 1$

En $t = l$ (donc sachant D_l), pour $j = 0, \dots, J - 1$ on a :

$$\hat{f}_j^l = \frac{\sum_{i=0}^{l-j-1} C_{i,j+1}}{\sum_{i=0}^{l-j-1} C_{i,j}} = \frac{\sum_{i=0}^{l-j-1} C_{i,j+1}}{S_j^l} \quad \text{avec } S_j^l = \sum_{i=0}^{l-j-1} C_{i,j}.$$

En $t = l + 1$, on obtient :
$$\hat{f}_j^{l+1} = \frac{\sum_{i=0}^{l-j} C_{i,j+1}}{\sum_{i=0}^{l-j} C_{i,j}} = \frac{\sum_{i=0}^{l-j} C_{i,j+1}}{S_j^{l+1}}.$$

→ Ces estimateurs sont sans biais, et pour tout $j \neq l$,

$$\text{Corr}(\hat{f}_j^l, \hat{f}_l^l) = \text{Corr}(\hat{f}_j^{l+1}, \hat{f}_l^{l+1}) = 0$$

Donc pour $j \geq I - i$,

$$\hat{C}_{i,j}^I = C_{i,I-i} \hat{f}_{I-i}^I \dots \hat{f}_{j-2}^I \hat{f}_{j-1}^I : \text{estimateur non-biaisé de } \mathbb{E}[C_{i,j} | D_I],$$

$$\hat{C}_{i,j}^{I+1} = C_{i,I-i+1} \hat{f}_{I-i+1}^{I+1} \dots \hat{f}_{j-1}^{I+1} : \text{estimateur non-biaisé de } \mathbb{E}[C_{i,j} | D_{I+1}].$$

Comme dans le modèle de Mack, on a

$$MSEP_{C_{i,j} | D_I}(\hat{C}_{i,j}^I) = \mathbb{E}[(C_{i,j} - \hat{C}_{i,j}^I)^2 | D_I].$$

Or $\hat{C}_{i,j}^I$ est D_I -mesurable, donc

$$MSEP_{C_{i,j} | D_I}(\hat{C}_{i,j}^I) = \mathbb{V}ar(C_{i,j} | D_I) + (\mathbb{E}[C_{i,j} | D_I] - \hat{C}_{i,j}^I)^2.$$

C'est la décomposition en **process variance** et **estimation error** (cette dernière serait nulle si on connaissait les f_j !).

Retour au CDR

Pour l'année comptable $(I, I + 1]$, le vrai CDR de l'année de survenance i (où $I \in [1, I]$) est donné par

$$\begin{aligned} CDR_i(I + 1) &= \mathbb{E}[R_i^I | D_I] - (X_{i,I-i+1} + \mathbb{E}[R_i^{I+1} | D_{I+1}]) \\ &= \mathbb{E}[C_{i,I} | D_I] - \mathbb{E}[C_{i,I} | D_{I+1}] \end{aligned}$$

avec $X_{i,I-i+1} = C_{i,I-i+1} - C_{i,I-i}$.

Pour obtenir le *CDR* global, on agrège sur les années de survenance :

$$CDR(I + 1) = \sum_{i=1}^I CDR_i(I + 1).$$

Par définition, on a un processus centré :

$$\mathbb{E}[CDR_i(I+1) | D_I] = 0.$$

En effet, si les facteurs CL f_j étaient connus, le vrai CDR attendu vu en date I vaudrait 0 ; et on aurait

$$MSEP_{CDR_i(I+1)|D_I}(0) = \mathbb{V}ar(CDR_i(I+1) | D_I) = \mathbb{E}[C_{i,J} | D_I]^2 \frac{1}{C_{i,I-I}} \frac{\sigma_{I-I}^2}{f_{I-I}^2}$$

Problème : les facteurs CL f_j ne sont pas connus ! \Rightarrow Le CDR n'est pas observable...

Estimation du CDR

On estime le CDR_i par

$$C\hat{D}R_i(I+1) = \hat{R}_i^I - (X_{i,I-I+1} + \hat{R}_i^{I+1}) = \hat{C}_{i,J}^I - \hat{C}_{i,J}^{I+1}$$

avec

$$\hat{R}_i^I = \hat{C}_{i,J}^I - C_{i,I-I} \quad : \text{estimateur non biaisé de } \mathbb{E}[R_i^I | D_I]$$

$$\hat{R}_i^{I+1} = \hat{C}_{i,J}^{I+1} - C_{i,I-I+1} \quad : \text{estimateur non biaisé de } \mathbb{E}[R_i^{I+1} | D_{I+1}]$$

Et le CDR global par $C\hat{D}R(I+1) = \sum_{i=1}^I C\hat{D}R_i(I+1)$.

Incertitude autour de cette estimation

Rappel :

→ Pour quantifier l'incertitude de l'estimateur \hat{R}_i d'une **v.a.** R_i , on utilise $MSE(\hat{R}_i) = \mathbb{E}[(\hat{R}_i - \mathbb{E}[R_i])^2]$.

→ La **MSEP** est donnée par $MSEP(\hat{R}_i) = \mathbb{E}[(\hat{R}_i - R_i)^2]$.

Adapté à notre contexte, on obtient

$$MSEP_{\hat{C}DR_i(I+1) | D_I}(0) = \mathbb{E}[(\hat{C}DR_i(I+1) - 0)^2 | D_I]$$

$$MSEP_{\hat{C}DR_i(I+1) | D_I}(\hat{C}DR_i(I+1)) = \mathbb{E}[(\hat{C}DR_i(I+1) - \hat{C}DR_i(I+1))^2 | D_I]$$

La première quantité correspond à une vision prospective : en moyenne le vrai CDR devrait valoir 0 dans un an.

La seconde est rétrospective : on regarde la distance entre vrai CDR et CDR observable.

Pour quantifier ces MSEP, on a recours à l'estimateur de la variance comme défini dans Mack :

$$\hat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2 \quad j \leq J-1.$$

Grâce à cet estimateur, on donne dans la suite l'expression des MSEP qui nous intéressent.

Notations et résultats

Soient

- $\hat{\Delta}_{i,J}^l = \frac{1}{S_{l-i}^l} \left(\frac{\hat{\sigma}_{l-i}}{\hat{f}_{l-i}^l} \right)^2 + \sum_{j=l-i+1}^{J-1} \left(\frac{C_{l-j,j}}{S_j^{l+1}} \right)^2 \frac{1}{S_j^l} \left(\frac{\hat{\sigma}_j}{\hat{f}_j^l} \right)^2$

- $\hat{\Phi}_{i,J}^l = \sum_{j=l-i+1}^{J-1} \left(\frac{C_{l-j,j}}{S_j^{l+1}} \right)^2 \frac{1}{C_{l-j,j}} \left(\frac{\hat{\sigma}_j}{\hat{f}_j^l} \right)^2$

- $\hat{\Psi}_i^l = \frac{1}{C_{l-i,i}} \left(\frac{\hat{\sigma}_{l-i}}{\hat{f}_{l-i}^l} \right)^2$

- $\hat{\Gamma}_{i,J}^l = \hat{\Phi}_{i,J}^l + \hat{\Psi}_i^l \geq \hat{\Phi}_{i,J}^l.$

On peut montrer que l'estimateur de la variance s'écrit

$$\hat{\text{Var}}(CDR_i(I+1) | D_I) = (\hat{C}_{i,J}^I)^2 \hat{\Psi}_i^I$$

et au final on obtient les estimateurs suivants :

$$M\hat{S}EP_{C\hat{D}R_i(I+1) | D_I}(0) = (\hat{C}_{i,J}^I)^2 [\hat{\Gamma}_{i,J}^I + \hat{\Delta}_{i,J}^I]$$

$$M\hat{S}EP_{C\hat{D}R_i(I+1) | D_I}(C\hat{D}R_i(I+1)) = (\hat{C}_{i,J}^I)^2 [\hat{\Phi}_{i,J}^I + \hat{\Delta}_{i,J}^I].$$

Puis il suffit d'agréger les années de survenance pour en déduire le CDR global. Le calcul fait intervenir des termes de covariance.

Agrégation pour obtenir le CDR total

On obtient :

$$\begin{aligned}
 M\hat{SEP}_{CDR(I+1)|D_I} &= M\hat{SEP}_{\sum_i CDR_i(I+1)|D_I} \left(\sum_{i=1}^I C\hat{D}R_i(I+1) \right) \\
 &= \sum_{i=1}^I M\hat{SEP}_{CDR_i(I+1)|D_I} (C\hat{D}R_i(I+1)) + 2 \sum_{k>i>0} \hat{C}_{i,J}^I \hat{C}_{k,J}^I [\hat{\Phi}_{i,J}^I + \hat{\Lambda}_{i,J}^I]
 \end{aligned}$$

où

$$\hat{\Lambda}_{i,J}^I = \frac{C_{i,I-i}}{S_{I-i}^{I+1}} \frac{1}{S_{I-i}^I} \left(\frac{\hat{\sigma}_{I-i}}{\hat{f}_{I-i}^I} \right)^2 + \sum_{j=I-i+1}^{J-1} \left(\frac{C_{I-j,j}}{S_j^{I+1}} \right)^2 \frac{1}{S_j^I} \left(\frac{\hat{\sigma}_j}{\hat{f}_j^I} \right)^2$$

Et en introduisant

$$\hat{E}_{i,J}^I = \hat{\Phi}_{i,J}^I + \frac{1}{S_{I-i}^{I+1}} \left(\frac{\hat{\sigma}_{I-i}}{\hat{f}_{I-i}^I} \right)^2 \geq \hat{\Phi}_{i,J}^I,$$

on a

$$\begin{aligned} M\hat{S}EP_{\hat{C}\hat{D}R(I+1)|D_I}(0) &= M\hat{S}EP_{\sum_i \hat{C}\hat{D}R_i(I+1)|D_I}(0) \\ &= \sum_{i=1}^I M\hat{S}EP_{\hat{C}\hat{D}R_i(I+1)|D_I}(0) + 2 \sum_{k>i>0} \hat{C}_{i,J}^I \hat{C}_{k,J}^I [\hat{E}_{i,J}^I + \hat{\Lambda}_{i,J}^I] \end{aligned}$$

Application

Voir le papier de Merz-Wuthrich, Modelling the Claims Development Result for Solvency purposes

Le calcul du SCR risque de provisionnement passe donc par le postulat d'une hypothèse sur la forme de la distribution du CDR : l'hypothèse la plus communément utilisée est l'hypothèse log-normale.

Preuves

Voir si on ajoute des preuves dans Mack, GLM, Merz...

Expliquer que la prediction variance est la somme de la process variance et de l'estimation variance...

7 Extensions de ces modèles

- Technique bayésienne et méthodes *MCMC*²
- Provisionnement ligne à ligne : censure et covariable
- Corrélation : agrégation de triangles de réserves par copule
- Tendence et effets calendaires - modele de Gluck and Venter

Valeurs négatives

Parfois le triangle contient des valeurs négatives...

[?]

Agrégation de triangles

Un groupe peut avoir plusieurs branches d'activité, chacune étant lié à une provision. Comment calculer une provision globale au niveau groupe, sachant que certaines business lines sont corrélées... ?

[?]

Bornhuetter-Ferguson stochastique

Blabla

7 Extensions de ces modèles

- Technique bayésienne et méthodes *MCMC*²
- Provisionnement ligne à ligne : censure et covariable
- Corrélation : agrégation de triangles de réserves par copule
- Tendence et effets calendaires - modele de Gluck and Venter

7 Extensions de ces modèles

- Technique bayésienne et méthodes *MCMC*²
- **Provisionnement ligne à ligne : censure et covariable**
- Corrélation : agrégation de triangles de réserves par copule
- Tendence et effets calendaires - modele de Gluck and Venter

7 Extensions de ces modèles

- Technique bayésienne et méthodes *MCMC*²
- Provisionnement ligne à ligne : censure et covariable
- **Corrélation : agrégation de triangles de réserves par copule**
- Tendence et effets calendaires - modele de Gluck and Venter

7 Extensions de ces modèles

- Technique bayésienne et méthodes *MCMC*²
- Provisionnement ligne à ligne : censure et covariable
- Corrélation : agrégation de triangles de réserves par copule
- Tendence et effets calendaires - modele de Gluck and Venter

CONCLUSION

Conclusion

Les méthodes actuelles permettent de quantifier l'incertitude :

- ces méthodes évoluent pour répondre à la non réalisation d'une ou plusieurs conditions de performance.
- les résultats entre ces différentes approches peuvent différer assez largement !
- Toutes les techniques présentées nécessitent
 - des données fiables et plutôt nombreuses,
 - un passé régulier, et un présent/futur structurellement proche.
- La branche doit être peu volatile.

On apporte des modifications pour concilier avec des observations qui sortent de l'ordinaire : lissage, pondération, coût moyen, dépendance des coefficients, méthode complémentaire...

Certaines questions opérationnelles ne sont pas traitées dans le cadre de ce cours : on pense notamment à

- les recours : comment les traiter ?
- les données absentes ou non triangulaires
- le changement du programme de réassurance, d'une politique de gestion des sinistres...

Certains mémoires abordent ces questions et y apportent des solutions pratiques.

Bibliographie



Benjamin, S. and Eagles, L. M. (1986).
Reserves in Lloyd's and the London Market.
J. Inst. Actuar., 113(2) :197–256.



Bornhuetter, R. and Ferguson, R. E. (1972).
The actuary and IBNR.
Casualty Actuar. Soc., 59 :181–195.



De Vylder, F. (1982).
Estimation of IBNR claims by credibility theory.
Insur. Math. Econ., 1(1) :35–40.



Efron, B. and Tibshirani, R. J. (1993).
An introduction to the bootstrap.
CRC Monographs on Statistics & Applied Probability. Chapman and Hall, London.



England, P. D. and Verrall, R. J. (2001a).
A flexible framework for stochastic claims reserving.
Casualty Actuar. Soc., 88 :1–38.



England, P. D. and Verrall, R. J. (2001b).

Analytic and bootstrap estimates of prediction errors in claims reserving.

IME, 15 :281–293.



Mack, T. (1993).

Distribution-free calculation of the standard error of chain-ladder reserve estimates.

ASTIN Bull., 23 :213–225.



Mack, T. (1994a).

Measuring the variability of chain-ladder reserve estimates.

Casualty Actuar. Soc., 23 :101–182.



Mack, T. (1994b).

Which stochastic model is underlying the chain-ladder method ?

IME, 15 :133–138.



Mack, T. (2000).

Credible claims reserves : The Benktander method.

ASTIN Bull., 20 :333–347.



McCullagh, P. and Nelder, J. A. (1989).

Generalized linear models, 2nd ed.

Monographs on Statistics and Applied Probability. Chapman and Hall, London.



Merz, M. and Wüthrich, M. V. (2008).

Modelling the claims development result for solvency purposes. CAS E-Forum (Fall.



Partrat, C., Lecoecur, E., Nessi, J. M., Nisipasu, E., and Reiz, O. (2007).
Provisionnement technique en assurance non-vie. Perspectives actuarielles modernes.

Assurance Audit Actuariat. Economica, Paris.



Quarg, G. and Mack, T. (2008).

Munich Chain Ladder : A Reserving Method that Reduces the Gap between IBNR Projections Based on Paid Losses and IBNR Projections Based on Incurred Losses.
Variance, 2(2) :266–299.



Straub, E. (1988).

Non-Life Insurance Mathematics.
Springer Verlag, Berlin.



Verrall, R. J. (1991).

On the estimation of reserves from loglinear models.
Insur. Math. Econ., 10(1) :75–80.



Verrall, R. J. and Wüthrich, M. V. (2012).

Reversible jump Markov chain Monte Carlo method for parameter reduction in claims reserving.

North Am. Actuar. J., 16(2) :240–259.



Wüthrich, M. and Merz, M. (2008).

Stochastic Claims Reserving Methods in Insurance.
Wiley Interscience. John Wiley & Sons, Hoboken, NJ, USA.



Wüthrich, M. V., Merz, M., and Lysenko, N. (2009).

Uncertainty of the claims development result in the chain ladder method.

Scand. Actuar. J., 2009(1) :63–84.