UNIVERSITÉ LAVAL
École d'actuariat

ACT-2008
Mathématiques actuarielles IARD II
Notes de cours

David Beauchemin

5 mai 2017

Table des matières

Préface

Ce document contient l'ensemble des notes de cours ACT-2008 *mathématiques actuarielles IARD II*. Le document rassemble la matière vue en classe durant la session Hiver 2017 par David Beauchemin et provenant des notes manuscrites de Frédérick Guillot. De plus, on peut retrouver en annexe des résumés de la matière synthétisée par Samuel Lévesque.
Courriel de l'auteur : david.beauchemin.5@ulaval.ca

Chapitre 1

Introduction à la théorie de la crédibilité

On considère une cellule de tarification (un sous-portefeuille du portefeuille complet, mais qui a les mêmes caractéristiques *au sens des questions que l'assureur pose à l'émission du contrat*).
Afin d'avoir une tarification équitable, l'assureur veut :

1. Charger assez de primes pour payer les sinistres, les frais et dégager un profit. Ça revient à fixer le bon niveau GLOBAL des taux ;
2. Distribuer les primes collectées équitablement entre les assurés en fonction du risque (GLM + jugement) :
   - Structure de tarification
   - Tarification basée sur l'expérience ⇒ Théorie de la crédibilité

Remarque

En pratique, toutes les lignes d'affaires(auto, habitation, etc.) basent leur tarification sur une indication globale et sur une structure de tarification. Comme l'assureur ne peut généralement pas poser toutes les questions requises pour bien quantifier le risque, la plupart des cellules de tarification demeurent hétérogènes [1].
Dans certains autres cas, par exemple la CNESST (anciennement appeler la CSST), l'assureur base sa tarification sur la prime globale (ou bien sous une segmentation très restreinte/minimale). Dans ce cas-ci, il y aura aussi présence d'hétérogénéité. Comme un fort volume de sinistre est généré en

1. Distribution répartie de façon inégale.

assurance contre les accidents de travail, la CNESST peut se permettre d'incorporer une composante *basée sur l'expérience*.
La théorie de la crédibilité ajustera la prime en fonction de l'expérience des assurés de ces cellules hétérogènes. (*On définit une prime de globale selon des hypothèses et on ajuste avec l'observation du risque dans le temps*)

1.1 Exemple 1

La CNESST est l'organisme étatique qui assure dans le domaine des accidents en milieu de travail. Pour cet exemple, on s'intéresse au sinistre (accidents de travail) dans les usines.

1.1.1 Mise en situation

- Elle assure 10 assurés, $\Longrightarrow$ 10 usines qui emploient des travailleurs (usines possiblement différentes).
- Les actuaires procèdent alors à une analyse globale du niveau de taux avec une fréquence estimée $\lambda = 25\%$;
- Ils estiment également que la sévérité estimée $\mu = 100\ 000$ \$/réclamation ;
- La prime unique uniforme $\lambda = 25\ 000$ \$/employeur.
- On remarque que dans ce cas-ci, on suppose que la CNESST a décidé de ne pas faire d'analyse de structure de tarification. La prime ne varie pas selon les caractéristiques de l'employeur.

Observation des sinistres sur 10 ans :

Usine	N_1	S_1	N_2	S_2	...	N_{total}	S_{total}
1	0	0	1	125 000	...	6	125 000
2	0	0	1	184 000	...	3	600 000
3	0	0	0	0	...	2	180 000
4	0	0	0	0	...	2	190 750
5	0	0	0	0	...	1	3 000
6	0	0	0	0	...	2	120 000
7	0	0	0	0	...	0	0
8	0	0	0	0	...	0	0
9	1	85 000	1	82 500	...	10	850 000
10	0	0	0	0	...	0	0

Nous avons donc après 1 an selon nos observations que :

$$\widehat{\lambda}(1) = \frac{1}{10} = 10\%$$
$$\widehat{\mu}(1) = \frac{85\ 000\$}{1}$$

Est-ce que la prime est juste ? Est-elle trop élevée ? On ne peut pas conclure que la prime n'est pas adéquate, il manque de données pour assurer une crédibilité à notre expérience.

Après 2 ans, nous avons que :

$$\widehat{\lambda}(2) = \frac{4 \text{ sinistres}}{20 \text{ unités}} = 20\%$$
$$\widehat{\mu}(2) = \frac{476\ 500\$}{4 \text{ sinistres}} > 100\ 000\$$$

Il y a encore trop peu de données pour affirmer que la prime est trop élevée. Après 10 ans, nous avons que :

$$\widehat{\lambda}(10) = \frac{26 \text{ sinistres}}{100 \text{ unités}} = 26\%$$
$$\widehat{\mu}(10) = \frac{2\ 600\ 000\$}{26 \text{ sinistres}} = 100\ 000\$$$

Pour conclure, on note que la fréquence, $\widehat{\lambda} = 26\%$ converge vers la valeur estimée $\lambda = 25\%$ et que la gravité converge aussi vers la valeur théorique. Par contre, on peut remarquer que la fréquence et la gravité ne sont pas identiques pour chacune des entreprises. Un ajustement de prime devra être effectué selon le risque de l'entreprise.

Motivation pour l'utilisation de la crédibilité chez la CNESST

- La prime collective est adéquate globalement (l'analyse globale des taux est correcte).
- En revanche, la prime n'est pas équitable d'un assureur à l'autre.
- Besoin d'une technique de tarification basée sur l'expérience pour allouer adéquatement la prime entre les différents assurés.

Deux catégories de théorie de crédibilité :

1) Crédibilité de stabilité (*Limited fluctuation*) : On ne considère l'expérience de sinistre des assurés qu'à partir d'une *taille donnée*. Autrement dit, on peut prendre des conclusions seulement lorsque l'expérience devient stable dans le temps.

2) Crédibilité de précision(*Greatest acccuracy*) : On considère partiellement *l'expérience de sinistre* de chaque assuré pour ajuster sa prime. Autrement dit, on considère l'expérience *partiellement* pour prendre des conclusions. En donnant plus de poids à l'expérience à mesure que celle-ci devient stable.

Chapitre 2

Crédibilité de stabilité

2.1 Introduction par un exemple

On utilise l'exemple suivant pour introduire la théorie de la crédibilité de stabilité. Vous travaillez dans une compagnie d'assurance ayant 2 agents :

— Agent 1 : Ouvert depuis 10 ans ;

— Agent 2 : Nouvellement ouvert cette année.

On veut comparer les *taux de succès* $\left(S = \frac{\text{Nombre de ventes}}{\text{Nombre de soumissions}}\right)$:

Agent	Nombre de soumissions	Nombre de ventes	Taux de succès
1	5000	2 500	50%
2	50	25	50%

Question principale en crédibilité de stabilité :

À partir de quelle taille (n) est-ce que S est statistiquement stable ($\Rightarrow$ crédible) ?

Traitement mathématique :

On dit que la statistique (dans notre cas présent S) est *crédible* (ou *statistiquement stable*) à l'ordre (k, P) lorsque :

$$P\Big\{(1-k) \times E[s] \leq S \leq (1+k) \times E[s]\Big\} \geq P \tag{2.1}$$

Où k est petit (ex. : 5%) et P est près de 100 % (ex. : 95%).

On peut interpréter l'équation 2.1 comme étant la probabilité p% d'être à l'extérieur de notre intervalle autour de la statistique S.

Selon notre exemple initial :

$$\begin{aligned} S &= \frac{\text{Nombre de ventes}}{\text{Nombre de soumissions}} \\ &= \frac{\sum_{i=1}^{n} I_i}{n} \\ &\Rightarrow \left(\frac{\sum_{i=1}^{n} I_i}{n}\right) \sim \text{Bin(n,p)} \end{aligned}$$

À partir de quelle valeur de n (taille) est-ce que :

$$P\Big\{(1-k) \times E[s] \leq S \leq (1+k) \times E[s]\Big\} \geq P$$

Puisqu'on s'attend à un n grand, on utilise l'approximation normale. En assumant que le *vrai* taux de succès est de 50 % et que (k = 5%, p = 90 %), l'équation devient :

$$\begin{gathered} P\left(0.95 \times 0.5 \times n \leq \frac{\sum_{i=1}^{n} I_i}{n} \leq 1.05 \times 0.5 \times n\right) \geq P \\ P\left(0.475 \times n \leq \frac{\sum_{i=1}^{n} I_i}{n} \leq 0.525 \times n\right) \geq 0.9 \\ P\left(\frac{0.475n - 0.5n}{\sqrt{n \times 0.5 \times 0.5}} \leq \frac{\sum_{i=1}^{n} I_i}{n} \leq \frac{0.525n - 0.5n}{\sqrt{n \times 0.5 \times 0.5}}\right) \geq 0.9 \\ P\Big\{-0.05\sqrt{n} \leq Z \leq 0.05\sqrt{n}\Big\} \geq 0.9 \\ \Phi\big(0.05\sqrt{n}\big) - \Phi\big(-0.05\sqrt{n}\big) \geq 0.9 \\ \Phi\big(0.05\sqrt{n}\big) - \Big(1 - \Phi(-0.05\sqrt{n})\Big) \geq 0.9 \\ 2\Phi\big(0.05\sqrt{n}\big) - 1 \geq 0.9 \\ \Phi\big(0.05\sqrt{n}\big) \geq 0.95 \\ 0.05\sqrt{n} \geq \Phi^{-1}\big(0.95\big) \\ n \geq 1082.41 \end{gathered}$$

Pour conclure, à partir d'environ $\approx$ 1083 soumissions, on peut *croire* au taux de succès des agents de 50%.

2.2 Crédibilité complète d'ordre (k, p)

En crédibilité complète, un contrat (ou un portefeuille de contrat) est *crédible* si son expérience est stable. Intuitivement, cette stabilité va de pair avec la *taille* du contrat ou du portefeuille.
Ainsi, une crédibilité complète d'ordre (k, p) est attribuée à l'expérience S d'un contrat ou portefeuille si les paramètres de la distribution sont tels que

$$P\left(\frac{(1-k)E[S]-E[S]}{\sqrt{Var(S)}} \leq \frac{S-E[S]}{\sqrt{Var(S)}} \leq \frac{(1+k)E[S]-E[S]}{\sqrt{Var(S)}}\right) \geq P$$

est vérifié.

2.2.1 Exemple classique :

On considère le modèle classique du risque IARD :

$$S = \begin{cases} \sum_{i=1}^{n} X_i & \text{, si } N > 0 \\ 0 & \text{, si } N = 0 \end{cases} \tag{2.2}$$

S : Coûts totaux d'un portefeuille d'assurance IARD pour 1 an
N : Nombre de sinistres en 1 an $\sim$Poisson(λ)
X_i : Coût du sinistre *iid* avec $F_X(x), f_X(x), E[X], Var(X)$...
On sait que :

$$\begin{aligned} E[S] &= E[N]E[X] \\ &= \lambda E[X] \end{aligned}$$

ainsi que,

$$\begin{aligned} Var(S) &= E[N]Var(X) + Var(N)E^2[X] \\ &= \lambda Var(X) + \lambda E^2[X] \\ &= \lambda(Var(X) + E^2[X]) \\ &= \lambda E[X^2] \end{aligned}$$

On reprend l'équation *générale* 2.1 :

$$P\Big\{(1-k) \times E[s] \leq S \leq (1+k) \times E[s]\Big\} \geq P$$

Par conséquent, on cherche la taille n (λ) à partir de laquelle les *résultats* (S) sont statistiquement stables (*Also Know As* crédibles) d'année en année. À partir de l'équation 2.1, on effectue une approximation par la loi Normale

$$P\left(\frac{(1-k)E[S]-E[S]}{\sqrt{Var(S)}} \leq \frac{S-E[S]}{\sqrt{Var(S)}} \leq \frac{(1+k)E[S]-E[S]}{\sqrt{Var(S)}}\right) \geq P \tag{2.3}$$

À partir de la forme générale 2.1, on obtient :

$$P\left(\frac{k\lambda E[S]}{\sqrt{\lambda E[X^2]}} \leq Z \leq \frac{k\lambda E[X]}{\sqrt{\lambda E[X^2]}}\right) \geq P$$
$$2\Phi\left(k\sqrt{\lambda}\frac{E[X]}{\sqrt{E[X^2]}}\right) - 1 \geq P$$
$$\Phi\left(k\sqrt{\lambda}\frac{E[X]}{\sqrt{E[X^2]}}\right) \geq \frac{P+1}{2}$$
$$\left(\frac{k\sqrt{\lambda}E[X]}{\sqrt{E[X^2]}}\right) \geq \Phi^{-1}\left(\frac{P+1}{2}\right)$$
$$\lambda \geq \left(\frac{\sqrt{E[X^2]}}{E[X]k}\right)\Phi^{-1}\left(\frac{P+1}{2}\right)^2$$
$$\lambda \geq \left(\frac{\Phi^{-1}(\frac{P+1}{2})^2}{k}\right)\left(\frac{Var(X)+E^2[X]}{E^2[X]}\right)$$

On obtient l'équation finale suivante

$$\lambda \geq \left(\frac{\Phi^{-1}(\frac{P+1}{2})^2}{k^2}\right)(1+CV^2(X)) \tag{2.4}$$

Remarques :

1) Dans ce cas-ci, la *taille* du portefeuille est exprimée en *nombre de sinistres espérés annuellement (λ)* et non en nombre d'assurés, afin de rendre la théorie portable dans plusieurs lignes d'affaires (auto, habitations, ...).
2) Plus la distribution de X (sévérité) est volatile (plus CV(X) $\nearrow$) plus la *taille* doit être grande pour atteindre la crédibilité complète.
3) En assumant :

— k = 5%
— P = 90%
— La sévérité est non aléatoire (densité dégénérée) $\Rightarrow$ CV(X) = 0
On obtient,

$$\begin{aligned}\lambda &\geq \left(\frac{\Phi^{-1}(0.95)}{0.05}\right)^2 (1+0^2)\\ &\geq 1082.41\end{aligned}$$

2.3 Crédibilité partielle

Que fait-on lorsque la *taille* est inférieure au seuil de crédibilité complète ?

- Option 1 : Ne pas considérer les chiffres
- Option 2 : Whitney (1918) proposa de pondérer l'expérience individuelle avec un complément crédible (ex. prime collective).
  Soit :

$$\pi = Z \times \overline{S} + (1 - Z) \times m \tag{2.5}$$

— π est la statistique S crédibiliser
— Z est le facteur de pourcentage de crédibilité
— $\overline{S}$ est l'expérience observée par la statistique S
— m est le complément crédible

Sans trop développer de théorie, Whitney proposa diverses formes pour Z.

Z peut correspond à :

i)

$$Z = \min\left(\sqrt{\frac{\text{taille}(n)}{\text{taille de crédibilité complète}(n_0)}}; 1\right)$$

ii)

$$Z = \frac{n}{n+k}, \text{ où n = taille, k = constante arbitraite}$$

Retour sur l'exemple 2.1 :

On choisit la formule (i) pour Z, sachant que $n_0 = 1083$, alors

$$\begin{aligned} Z &= \min\left(\sqrt{\frac{n}{n_0}}; 1\right) \\ &= \min\left(\sqrt{\frac{50}{1083}}; 1\right) \end{aligned}$$

On suppose que le taux de succès de toutes les agences au Canada est de 45 %, assez crédible pour être choisi comme *complément*, car l'ensemble du Canada correspond à un échantillon de grande *taille*.

Agent	n	$\sum_{i=1}^{n} X_i$	S	Z	π
1	5 000	2 500	50%	100%	50%
2	50	25	50%	21.5%	46.1%

Remarques :

— Si la moyenne nationale avait été plus élevée, disons 55 %, alors $\pi_2 =$ 54%, donc on donne le *bénéfice du doute* à l'agent parce qu'il est trop peu crédible.

— Le but de la crédibilité partielle est d'incorporer autant d'expérience individuelle dans la prime, pour bien le faire,il faut déterminer les paramètres adéquatement.

Chapitre 3

Crédibilité Bayesienne

*Rappel : En ACT - 2000, la méthode de Bayes est vue comme une façon (1) un avis d'expert en plus (2) des données observées dans l'estimation statistique. Dans ce cas-ci, en crédibilité :

1) Avis d'expert : Prime collective
2) Données : expérience individuelle des assurés
3) Statistique : Souvent la prime.

3.1 Introduction par un exemple

Ex :

$$X|\Theta \sim \text{Bin}(1, \theta)$$

Où

- X est l'indicateur que la maison est inondée et
- θ est la probabilité d'inondation.

Avec le maximum de vraisemblance, on obtient :

$$\widehat{\Theta} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{3.1}$$

Que faire si on n'a pas de données ? On utilise l'avis d'experts.

$$X_i|\Theta_i \sim \text{Bern}(\theta_i) \tag{3.2}$$

$$\theta \sim \text{Beta}(\alpha, \beta) \tag{3.3}$$

La prime collective correspond à

$$\begin{aligned}
m &= E[\mu(\Theta)] \\
&= E[\theta] \\
&= \frac{\alpha}{\alpha+\beta}
\end{aligned}$$

On obtient la distribution a postériori suivante :

$$(\Theta|X_1, ..., X_n) \sim \text{Beta}\left(\alpha + \sum_{i=1}^{n} X_i, \beta - \sum_{i=1}^{n} X_i + n\right)$$

À partir de notre expression 3.3 et selon les différentes situations, on obtient :

— Aucune donnée : $E[x] = E[E[X|\Theta]] = E[\theta] = \frac{\alpha}{\alpha+\beta}$

— Entre-deux (Bayes) : $E[X|X_1, ..., X_n] = E[\theta|X_1, ..., X_n] = \frac{\alpha+\sum_{i=1}^{n} X_i}{\alpha+\beta+n}$

— Beaucoup de données (maximum de vraisemblance) : voir 3.1

Or,

$$\begin{aligned}
\frac{\alpha + \sum_{i=1}^{n} X_i}{\alpha+\beta+n} &= \frac{\alpha}{\alpha+\beta+n} + \frac{\sum_{i=1}^{n} X_i}{\alpha+\beta+n} \\
&= \frac{\alpha}{\alpha+\beta}\frac{\alpha+\beta}{\alpha+\beta+n} + \frac{n}{\alpha+\beta+n}\frac{\sum_{i=1}^{n} X_i}{n} \\
&= \frac{n}{\alpha+\beta+n}\overline{X} + \left(1 - \frac{n}{n+(\alpha+\beta)}\right)\left(\frac{\alpha}{(\alpha+\beta)}\right)
\end{aligned}$$

On obtient le résultat suivant :

$$\frac{\alpha + \sum_{i=1}^{n} X_i}{\alpha+\beta+n} = Z\overline{X} + (1-Z)m \tag{3.4}$$

où $Z = \frac{n}{n+k}$, $k = \alpha + \beta$ et $m = \frac{\alpha}{\alpha+\beta}$

3.2 Estimation Bayesienne(Rappel/Background)

— Observations $= \{x_1, x_2, ..., x_n\}$;

— où $X_i|\Theta \sim f_{X|\Theta}(x|\theta)$;

— et $\Theta \sim f_\Theta(\theta)$, la distribution a priori de Θ.

$$f_{\Theta|X_1,X_2,...,X_n}(\theta|x_1,x_2,...,x_n) = \frac{f_{\Theta,X_1,X_2,...,X_n}(\theta,x_1,x_2,...,x_n)}{f_{X_1,...,X_n}(x_1,...,x_n)}$$

Qui n'est pas une fonction de Θ. Par la suite, on crée une proportion afin que la somme des probabilités soit égale à 1. Soit :

$$\begin{aligned}&\propto f_{\Theta,X_1,X_2,...,X_n}(\theta,x_1,x_2,...,x_n)\\&= f_{X_1,X_2,...,X_n|\Theta}(x_1,x_2,...,x_n|\theta)f_\Theta(\theta)\end{aligned}$$

Donc la loi a posteriori de θ est

$$f_{\Theta|X_1,X_2,...,X_n}(\theta|x_1,x_2,...,x_n) \propto f_{X_1,X_2,...,X_n|\Theta}(x_1,x_2,...,x_n|\theta)f_\Theta(\theta) \tag{3.5}$$

et si $X_1,...,\perp,...,X_n$,

$$\begin{aligned}f_{\Theta|X_1,X_2,...,X_n}(\theta|x_1,x_2,...,x_n) &\propto f_{X_1,|\Theta}(x_1|\theta)\times...\times f_{X_n|\Theta}(x_n|\theta)f_\Theta(\theta)\\&= L(\Theta)f_\Theta(\theta)\end{aligned}$$

Ainsi,

$$\widehat{\Theta} = E[\Theta|X_1,...,X_n] \tag{3.6}$$

correspondant à l'espérance a posteriori de Θ.

Exemple

Un portefeuille d'assurance auto composé de 75 % d'assurés qui ne textent pas au volant et de 25 % qui textent au volant. Or, on *sait* que la probabilité d'accident est de 5 % pour ceux qui ne textent pas en conduisant, alors qu'elle est de 25 % pour ceux qui texte. La sévérité d'un accident est de 10 000 \$. On ne demande pas à l'assuré s'il texte ou non au volant.

— $S = \sum_{i=1}^{N} X_i$
— $(N_i|\Theta) \sim \text{Bern}(\Theta)$
— $P(X = 10\ 000) = 1$ [1]

$$P(\Theta=\theta) = \begin{cases} 75\% & ,\Theta = 5\% \\ 25\% & ,\Theta = 25\% \end{cases}$$

1. Il s'agit d'une fonction *dégénéré* *(degenate density function)*.

a)Prime collective :

La prime uniforme chargée initialement aux gens d'une même cellule de tarification. Soit,

$$\begin{aligned} m =& E[S] \\ =& E[N]E[X] \\ =& E[E[N|\Theta]] \times 10\ 000 \\ =& E[\Theta] \times 10\ 000 \\ =& (0.05 \times 0.75 + 0.25 \times 0.25) \times 10\ 000 \\ =& 1\ 000\$ \end{aligned}$$

Qui correspond à l'espérance a priori de S.

b) La loi a posteriori de Θ après 1 an d'expérience :

$$f_{\Theta|N_1}(\theta|n_1) \propto f_{N_1|\Theta}(n_1|\theta) f_\Theta(\theta)$$

Cas #1 ($n_1 = 0$) :

$$\begin{aligned} f_{\Theta|N_1}(\theta|n_1=0) \propto& f_{N_1|\Theta}(0|\theta) f_\Theta(\theta) \\ =& (1-\theta) \times f_\Theta(\theta) \\ =& \begin{cases} (1-0.05)75\% & ,\Theta = 5\% \\ (1-0.25)25\% & ,\Theta = 25\% \end{cases} \\ f_{\Theta|N_1}(\theta|0) =& \begin{cases} \frac{0.7125}{0.9} & ,\Theta = 5\% \\ \frac{0.1875}{0.9} & ,\Theta = 25\% \end{cases} \end{aligned}$$

Notes : On divise par 0.9 [2] pour *proportionaliser ($\propto$)* la densité et ainsi avoir une fonction de densité = 1. On utilise une proportion pour avoir une somme des probabilités = 1,

$$f_{\Theta|N_1}(\theta|0) = \begin{cases} 0.791666 & ,\Theta = 5\% \\ 0.2083333 & ,\Theta = 25\% \end{cases}$$

2. Qui correspond à l'addition des probabilités de la densité (0.7125+0.1875).

Cas #2 ($n_1 = 1$) :

$$\begin{aligned}
f_{\Theta|N_1}(\theta|1) \propto & f_{N_1|\Theta}(1|\theta) f_\Theta(\theta) \\
= & \theta \times f_\Theta(\theta) \\
= & \begin{cases} 0.05 \times 75\% & , \Theta = 5\% \\ 0.25 \times 25\% & , \Theta = 25\% \end{cases} \\
= & \begin{cases} \frac{0.0375}{0.1} & , \Theta = 5\% \\ \frac{0.0625}{0.1} & , \Theta = 25\% \end{cases} \\
f_{\Theta|N_1}(\theta|1) = & \begin{cases} 0.375 & , \Theta = 5\% \\ 0.625 & , \Theta = 25\% \end{cases}
\end{aligned}$$

c) Prime Bayesienne a posteriori après 1 an d'expérience :

On obtient,

$$\begin{aligned}
B_2 = E[S_2|S_1] &= E[S_2|N_1] \\
&= E[N_2|N_1]E[X] \\
&= E[\Theta|N_1] \times 10\ 000
\end{aligned}$$

Où $E[\Theta|N_1]$ correspond à l'espérance a posteriori de Θ. On reprendre les 2 cas énoncés plus haut.

Cas 1

$$\begin{aligned}
E[S_2|S_1 = 0] = & E[\Theta|N_1 = 0] \times 10\ 000 \\
= & (0.791666 \times 0.05 + 0.2083 \times 0.25) \times 10\ 000 \\
= & 916.666\$ = \pi_2
\end{aligned}$$

Cas 2

$$\begin{aligned}
E[S_2|S_1 = 10\ 000] = & E[\Theta|N_1 = 1] \times 10\ 000 \\
= & (0.375 \times 0.05 + 0.625 \times 0.25) \times 10\ 000 \\
= & 1750\$ = \pi_2
\end{aligned}$$

Remarque : On peut écrire B_2 dans ce cas-ci comme :

$$B_2 = \pi_2 = 0.0833 \times \overline{X} + (1 - 0.0833) \times m \tag{3.7}$$

Où,

$\quad$ Z $= 0.08333$ (arbitraire) ;
$\quad \overline{X}$ $= 0$ \$ ou 100 000\$;
$\quad$ m $= 1\ 000$ \$.

On obtient une prime de crédibilité de combinaison linéaire entre l'expérience $\overline{X_1}$ et la prime collective m.

3.3 L'hétérogénéité résiduelle :

Dans le contexte de l'assurance IARD, on dira que le fait de ne pas poser toutes les questions à l'assuré concernant son risque introduit de l'hétérogénéité résiduelle dans la/les cellules de tarification. (Ex. Le fait de ne pas demander si l'assuré texte ou non au volant). En pratique, pour un portefeuille ou un sous-portefeuille (= cellule de tarification) composé des différents contrats d'assurés(soit I pour *Insured*), on *quantifie* le *niveau* de risque associé à *l'hétérogénéité résiduelle* par la *v.a.*Θ, pour i = 1,...,I.

Remarques :

1) Les Θ_i sont propres à chaque assuré i et ne changent pas avec le temps t ;
2) Les Θ_i sont des *v.a.* non observables ;
3) Par hypothèse, les Θ_i sont des *v.a.* *i.i.d* (dans la cellule de tarification) ;
4) Pour chaque assuré i, on observe les *v.a.* $S_{i,1}, ..., S_{i,n}$;
5) Chaque contrat est indépendant l'un de l'autre ;
6) Les observations $S_{i,t}$ d'un même contrat i sont indépendantes dans le temps t *conditionnellement à* Θ (contagion apparente).

Notes sur la contagion apparente :

$$S_{1,1} \not\perp ... \not\perp S_{1,5}$$
$$S_{2,1} \not\perp ... \not\perp S_{2,5}$$

Par contre :

$$(S_{1,1}|\Theta_1) \perp ... \perp (S_{1,5}|\Theta_1)$$
$$(S_{2,1}|\Theta_2) \perp ... \perp (S_{2,5}|\Theta_2)$$

Elles sont conditionnellement indépendants en sachant Θ.

3.4 Prévision Bayesienne (ou prime de crédibilité) :

Alors que la théorie de la crédibilité complète (limited fluctuations) visait à trouver la *taille* requise pour que les résultats d'un portefeuille (ou un sous-portefeuille) soient statistiquement stables, la crédibilité Bayesienne vise à calculer la *meilleure prévision* d'une quantité (sinistres totaux, nombre de sinistres, sévérités, loss ratio(LR),...) lors de la prochaine période.

Exemple : Sachant l'expérience passée d'un assuré, on pourrait vouloir calculer :
- Le nombre de sinistres prédit pour l'an prochain,
- sévérité prédite pour l'an prochain,
- le montant total des réclamations pour l'an prochain.

3.4.1 Prime de risque :

- Correspond à l'opposé de la prime collective (= m = E[S]),
- la prime que chaque contrat de la cellule de tarification devrait payer.

Autrement dit, si le niveau de risque Θ_i du contrat i est connu, alors la meilleure prévision est l'espérance :

$$\mu(\theta) = E[S_{i,t}|\Theta_i = \theta] = \int_0^\infty x f_{X|\Theta}(x|\theta)dx$$

Remarque : Puisque Θ n'est jamais observé, $\mu(\theta)$ est aussi inconnue et donc prévoir $S_{i,t+1}$ revient à prévoir $\mu(\theta_i)$.

3.4.2 Prime collective :

Comme *première approximation* (à t = 0) de la *vraie* prime de risque, on utilise souvent la moyenne pondérée des primes de risques possibles.

$$m = E[\mu(\theta)] = \int_{-\infty}^\infty \mu(\theta) f_\Theta(\theta)d\theta$$

Remarque :
Cette expression de l'approximation de la prime de risque sera la même pour tous les contrats.
On a aussi que :

$$\begin{aligned} m = E[\mu(\theta)] &= \int_{-\infty}^{\infty} \mu(\theta) f_{\Theta}(\theta) d\theta \\ &= \int_{-\infty}^{\infty} E[S_{i,t}|\theta] f_{\Theta}(\theta) d\theta \\ &= E[E[S_{i,t}|\Theta]] \\ &= E[S_{i,t}] \end{aligned}$$

Soit le montant moyen des sinistres dans le portefeuille.

3.4.3 Prime Bayesienne

D'après la section 3.4.2, la prime collective m est *globalement* adéquate, mais non équitable envers chaque assuré en raison de l'hétérogénéité résiduelle. Ainsi, la meilleure approximation de la prime de risque ($\mu(\theta_i)$) sachant l'expérience d'un contrat pendant n *périodes* est la prime Bayesienne, soit :

$$B_{i,n+1} = E[\mu(\Theta_i)|S_{i,1}, ..., S_{i,n}] = \int_{-\infty}^{\infty} \mu(\theta) f_{\Theta|S_{i,1},...,S_{i,n}}(\theta|s_{i,1}, ..., s_{i,n}) d\theta \tag{3.8}$$

Or, par le théorème de Bayes (3.2) :

$$\begin{aligned} f_{\Theta|S_{i,1},...,S_{i,n}}(\theta|S_{i,1}, ..., S_{i,n}) &\propto f_{S_{i,1},...,S_{i,n}|\Theta_i}(S_{i,1}, ..., S_{i,n}|\theta) f_{\Theta}(\theta) \\ &\overset{\perp}{=} f_{S_{i,1}|\Theta_i}(S_{i,1}|\theta) \times ... \times f_{S_{i,n}|\Theta_i}(S_{i,n}|\theta) \times f_{\Theta}(\theta) \\ &= \left[\prod_{t=1}^{n} f_{S_{i,t}|\Theta_i}(S_{i,t}|\theta) \times f_{\Theta}(\theta) \right] \end{aligned}$$

Remarques :
- La prime Bayesienne est la meilleure approximation de $S_{i,t+1}$ que l'on puisse calculer.
- L'ordre de survenance des sinistres n'est pas pris en compte dans la formule $B_{i,n+1}$.
- La prime collective peut s'interpréter comme la prime Bayesienne de

la première année :

$$\begin{aligned}B_{i,0+1} &= \int_{-\infty}^{\infty} \mu(\theta) f_{\Theta|\phi}(\theta|\phi) d\theta \\ &= \int_{-\infty}^{\infty} \mu(\theta) f_{\Theta}(\theta) d\theta \\ &= m\end{aligned}$$

Où ϕ est un ensemble vide.

Exemple

On ne s'intéresse qu'au nombre des sinistres du i^e contrat au cours de l'an t ($S_{i,t}$). On obtient :

$$(S_{i,t}|\Theta_i) \sim \text{Poisson}(\theta_i),\ \text{Où i} = 1, ..., I \text{ et t} = 1, ..., n$$

$$f_{\Theta|N_1}(\theta|0) = \begin{cases} 0.3 & , \Theta = 0.5 \\ 0.5 & , \Theta = 1 \\ 0.2 & , \Theta = 2 \end{cases}$$

a) Primes de risques possibles :

$$\mu(\theta) = \begin{cases} 0.5 & , \Theta = 0.5 \\ 1 & , \Theta = 1 \\ 2 & , \Theta = 2 \end{cases}$$

$$\begin{aligned}\mu(\theta) &= E[S_{i,t}|\Theta_i = \theta] \\ &= E[\Theta|\Theta_i = \theta] \\ &= \theta\end{aligned}$$

b) Prime collective :

$$\begin{aligned}\text{Équation}(\Delta) \Rightarrow m &= E\Big[E[S_{i,t}|\Theta_i]\Big] \\ &= E[\mu(\Theta_i)] \\ &= E[\Theta_i] \\ &= 0.3 \times 0.5 + 0.5 \times 1 + 0.2 \times 2 = 1.05\end{aligned}$$

c) Calculer la prime Bayesienne pour l'année 3 si $S_{i,1} = 2$ et $S_{i,2} = 1$:

1) Distribution a posteriori de Θ :

$$\begin{aligned} f_{\Theta|S_{i,1},S_{i,2}}(\theta|2,1) &\propto f_{S_{i,1},S_{i,2}|\Theta_i}(2,1|\theta) \times f_{\Theta}(\theta) \\ &= f_{S_{i,1}|\Theta}(2|\theta) \times f_{S_{i,2}|\Theta}(1|\theta) f_{\Theta}(\theta) \\ &= \frac{e^{-\theta}\theta^2}{2\,!}\frac{e^{-\theta}\theta^1}{1\,!} \times P(\Theta_i = \theta) \end{aligned}$$

On obtient

$$f_{\Theta|S_{i,1},S_{i,2}}(\theta|2,1) \propto \begin{cases} 0.00689774 & ,\Theta = 0.5 \\ 0.03383382 & ,\Theta = 1 \\ 0.01465251 & ,\Theta = 2 \end{cases}$$

On observe que la somme des différentes probabilités correspond à :

$$\sum_{\theta=i} = 0.0358407$$

On doit donc, normaliser la densité à l'aide de la constante de normalisation (0.0358407).

$$f_{\Theta|S_{i,1},S_{i,2}}(\theta|2,1) \propto \begin{cases} \frac{0.00689774}{0.0358407} & ,\Theta = 0.5 \\ \\ \frac{0.03383382}{0.0358407} & ,\Theta = 1 \\ \\ \frac{0.01465251}{0.0358407} & ,\Theta = 2 \\ \end{cases}$$

Ainsi,

$$f_{\Theta|S_{i,1},S_{i,2}}(\theta|2,1) \propto \begin{cases} 12.45\% & ,\Theta = 0.5 \\ 61.07\% & ,\Theta = 1 \\ 26.46\% & ,\Theta = 2 \end{cases}$$

2) Prime Bayesienne :

$$B_{i,3} = E\Big[\mu(\Theta_i)|S_{i,1} = 2, S_{i,2} = 1\Big] \tag{3.9}$$

À comparer avec l'équation (Δ) mentionnée plus haut.

Conclusion : Si n=0 (aucune observation) sur l'expérience du contrat i alors la prime Bayesienne $B_{i,0+1}$ coïncide avec la prime collective. Ainsi parce que $(S_{i,t}|\Theta_i) \sim \text{Poisson}(\Theta_i)$ on se rappelle que

$$\mu(\Theta_i) = E[S_{i,t}|\Theta_i = \theta] = \theta$$

Donc,

$$\begin{aligned}
B_{i,2+1} =& E\Big[\mu(\Theta_i)|S_{i,1} = 2, S_{i,2} = 1\Big] \\
=& 12.45\% \times 0.5 + 61.07\% \times 1 + 26.46\% \times 2 \\
=& 1.20
\end{aligned}$$

Remarque : Dans ce cas-ci, on a que

$$\begin{aligned}
1.20 =& Z \times \overline{S} + (1 - Z) \times m \\
1.20 =& Z\Big(\frac{2+1}{2}\Big) + (1 - Z) \times 1.05 \\
1.20 =& (33.333\%)\Big(\frac{2+1}{2}\Big) + (1 - 33.333\%) \times 1.05
\end{aligned}$$

La méthode Bayesienne a donc donné implicitement 33.333% de crédibilité à l'expérience de l'assuré i pour établir la prime crédibilisée.

3.5 Approche par la "distribution prédictive"

...est une approche alternative pour calculer $B_{i,n+1}$. Jusqu'à maintenant, on a vu que :

i)

$$\begin{aligned}
m &= E[\mu(\Theta_i)] \\
&= E\Big[E[S_{i,t}|\Theta_i]\Big] \\
&= E[S_{i,t}] \\
&= \int_0^\infty S \times f_{S_{i,t}}(s) ds \\
&= \int_0^\infty \int_{-\infty}^\infty S \times f_{S_{i,t}|\Theta}(s|\theta) f_\Theta(\theta) d\theta ds
\end{aligned}$$

ii)

$$\begin{aligned}
B_{i,n+1} &= E\Big[\mu(\Theta_i)|S_{i,1},...,S_{i,n}\Big] \\
&= E\Big[S_{i,t}|S_{i,1},...,S_{i,n}\Big] \\
&= \int_0^{\infty} S \times f_{S_{i,t}|S_{i,1},...,S_{i,n}}(s|S_{i,1},...,S_{i,n})ds \\
&= \int_0^{\infty}\int_{-\infty}^{\infty} S \times f_{S_{i,t}|\Theta}(s|\theta) f_{\Theta|S_{i,1},...,S_{i,n}}(\theta|S_{i,1},...,S_{i,n})d\theta ds
\end{aligned}$$

Preuve :

$$\begin{aligned}
f_{S_{i,t}|S_{i,1},...,S_{i,n}}(s|S_{i,1},...,S_{i,n}) &= \frac{f_{S_{i,t},S_{i,1},...,S_{i,n}}(s,S_{i,1},...,S_{i,n})}{f_{S_{i,1},...,S_{i,n}}(S_{i,1},...,S_{i,n})} \\
&= \frac{E_{\Theta_i}\Big[f_{S_{i,t},S_{i,1},...,S_{i,n}|\Theta_i}(s,S_{i,1},...,S_{i,n}|\theta_i)\Big]}{E_{\Theta_i}\Big[f_{S_{i,1},...,S_{i,n}|\Theta_i}(S_{i,1},...,S_{i,n}|\theta_i)\Big]} \\
&= \frac{\int_{-\infty}^{\infty} f_{S_{i,t},S_{i,1},...,S_{i,n}|\Theta_i}(s,S_{i,1},...,S_{i,n}|\theta_i) f_{\Theta_i}(\theta)d\theta}{\int_{-\infty}^{\infty} f_{S_{i,1},...,S_{i,n}|\Theta_i}(S_{i,1},...,S_{i,n}|\theta_i) f_{\Theta_i}(\theta)d\theta}
\end{aligned}$$

Puisque par hypothèse, $S_{i,1},...,S_{i,n}$ sont conditionnellement indépendant sachant Θ_i. Soit $\Big(S_{i,1}|\Theta_i\Big) \perp \Big(S_{i,2}|\Theta_i\Big) \perp ... \perp \Big(S_{i,n}|\Theta_i\Big)$. Donc

$$\begin{aligned}
&= \frac{\int_{-\infty}^{\infty} f_{S_{i,t}|\Theta_i}(s|\theta_i) f_{S_{i,1}|\Theta_i}(S_{i,1}|\theta_i)...f_{S_{i,n}|\Theta_i}(S_{i,n}|\theta_i) f_{\Theta_i}(\theta)d\theta}{\int_{-\infty}^{\infty} f_{S_{i,1}|\Theta_i}(S_{i,1}|\theta_i)...f_{S_{i,n}|\Theta_i}(S_{i,n}|\theta_i) f_{\Theta_i}(\theta)d\theta} \\
&= \int_{-\infty}^{\infty} f_{S_{i,t}|\Theta_i}(s|\theta_i)\left(\frac{f_{S_{i,1}|\Theta_i}(S_{i,1}|\theta_i)...f_{S_{i,n}|\Theta_i}(S_{i,n}|\theta_i) f_{\Theta_i}(\theta)}{\int_{-\infty}^{\infty} f_{S_{i,1}|\Theta_i}(S_{i,1}|\theta_i)...f_{S_{i,n}|\Theta_i}(S_{i,n}|\theta_i) f_{\Theta_i}(\theta)d\theta}\right)d\theta
\end{aligned}$$

Où $\int_{-\infty}^{\infty} f_{S_{i,1}|\Theta_i}(S_{i,1}|\theta_i)...f_{S_{i,n}|\Theta_i}(S_{i,n}|\theta_i) f_{\Theta_i}(\theta)d\theta$ est la constante de normalisation.

$$= \int_{-\infty}^{\infty} f_{S_{i,t}|\Theta_i}(s|\theta_i)\left(\frac{f_{S_{i,1}...S_{i,n}|\Theta_i}(S_{i,1}...S_{i,n}|\theta_i) f_{\Theta_i}(\theta)}{\int_{-\infty}^{\infty} f_{S_{i,1}...S_{i,n}|\Theta_i}(S_{i,1}...S_{i,n}|\theta_i) f_{\Theta_i}(\theta)d\theta}\right)d\theta$$

$$\text{Bayes} = \int_{-\infty}^{\infty} f_{S_{i,t}|\Theta_i}(s|\theta_i) f_{\Theta_i|S_{i,1}...S_{i,n}}(\theta_i|S_{i,1}...S_{i,n})d\theta$$

Où $f_{\Theta_i|S_{i,1}...S_{i,n}|\Theta_i}(\theta_i|S_{i,1}...S_{i,n})$ est la densité a posteriori de Θ.
On obtient le résultat final représentant la densité prédictive

$$f_{S_{i,t}|S_{i,1},...,S_{i,n}}(s|S_{i,1},...,S_{i,n}) = \int_{-\infty}^{\infty} f_{S_{i,t}|\Theta_i}(s|\theta_i) f_{\Theta_i|S_{i,1}...S_{i,n}|\Theta_i}(\theta_i|S_{i,1}...S_{i,n})d\theta \tag{3.10}$$

Analogie :

$$f_{S_{i,t}}(s) = \int_{-\infty}^{\infty} f_{S_{i,t}|\Theta_i}(s|\theta_i) f_{\Theta_i}(\theta_i) d\theta$$

Où $f_{\Theta_i}(\theta_i)$ est la densité a priori et $f_{S_{i,t}}(s)$ est la densité marginale.

$$\begin{aligned} B_{i,n+1} &= E[S_{i,t}|S_{i,1}...S_{i,n}] \\ &= \int_0^{\infty} s f_{S|S_{i,1}...S_{i,n}}(s|S_{i,1}...S_{i,n}) ds \\ m = E[S_{i,t}] &= \int_0^{\infty} s f_{S_{i,t}}(s) ds \end{aligned}$$

3.6 Crédibilité Bayesienne "linéaire" :

Dans plusieurs cas, en combinant *certaines* distributions usuelles, la prime Bayesienne peut se réécrire sous la forme :

$$B_{i,n+1} = Z \times \overline{S} + (1 - Z) \times m \tag{3.11}$$

Cette forme est appelée *prime de crédibilité* et Z $\varepsilon[0,1]$ est appelé *facteur de crédibilité*. Seulement quelques combinaisons de distributions résultent en une prime de crédibilité :

- $S|\Theta \sim$ Poisson, $\Theta \sim$ Gamma ;
- $S|\Theta \sim$ exp, $\Theta \sim$ Gamma ;
- $S|\Theta \sim$ Normale, $\Theta \sim$ Normale ;
- $S|\Theta \sim$ Bern, $\Theta \sim$ Beta ;
- $S|\Theta \sim$ Géom, $\Theta \sim$ Beta ;

Exemple pratique

Contexte :

$$S_{i,t} = \sum_{j=1}^{N_{i,t}} X_{i,j}$$

Soit le montant total payé à l'assuré i durant l'année t. De plus,

$$\left(N_{i,t}|\Theta_i\right) \sim \text{Poisson}\left(\lambda_{i,t} \times \theta_i\right)$$

Soit le nombre de sinistre de l'assurée de i durant l'année t. Où $\lambda_{i,t}$ correspond à la prévision de E[N] basé sur les réponses aux questions posées à l'assuré et θ_i correspond au facteur aléatoire qui tient compte de l'effet de toutes les questions non posées.

GLM Poisson :

- $\lambda_{i,t} = 0.05 g(\widehat{a}ge_{i,t}) \times h(sexe_{i,t})$
- $$g(\widehat{a}ge_{i,t}) \begin{cases} 1\% & , 16 \leq \text{âge} \leq 24 \\ 0.7\% & , \text{âge} \geq 24 \end{cases}$$
- $$h(sexe_{i,t}) \begin{cases} 1\% & , \text{sexe} = \text{f} \\ 1.4\% & , \text{sexe} = \text{h} \end{cases}$$

On regarde maintenant pour la sévérité.

$$\Theta_i \sim \Gamma(\alpha, \alpha)$$

Soit le paramètre de la fréquence $\Big(N_{i,t}|\Theta_i\Big)$

$$\Bigg(X_{i,t}|\psi_i\Bigg) \sim \exp\Big(\beta_{i,t} \times \psi_i\Big) \tag{3.12}$$

Où $\beta_{i,t}$ corresponds à $\frac{1}{\text{sévérité esperée selon les questions posées à l'assurée}}$ et ψ_i correspond à l'hétérogénéité résiduelle de sévérité.

GLM exponentielle :

- $\beta_{i,t} = \frac{1}{5\ 000 \times \varphi(valeur_{i,t})}$
- $$\varphi(valeur_{i,t}) \begin{cases} 0.7\% & , \text{ valeur} \leq 10\ 000 \\ 1\% & , 10\ 000 < \text{ valeur} < 30\ 000 \\ 2\% & , \text{ valeur} \geq 30\ 000 \end{cases}$$
- On obtient donc, $\psi_i \sim \Gamma(b+1, b)$. L'idée est que $E\Bigg[\frac{1}{\psi}\Bigg] = 1$.

a)

$$\begin{aligned}
E[N_{i,t}] &= E\Big[E[N_{i,t}|\Theta_i\Big] \\
&= E[\lambda_{i,t} \times \Theta_i] \\
&= \lambda_{i,t} E[\Theta_i] \\
&= \lambda_{i,t} \frac{\alpha}{\alpha} \\
&= \lambda_{i,t} \\
Var(N_{i,t}) &= E\Big[Var(N_{i,t}|\Theta_i)\Big] + \mathrm{Var}(E[N_{i,t}|\Theta_i]) \\
&= E[\lambda_{i,t} \times \Theta_i] + \mathrm{Var}(\lambda_{i,t} \times \Theta_i) \\
&= \lambda_{i,t} E[\Theta_i] + \lambda_{i,t}^2 \mathrm{Var}(\Theta_i) \\
&= \lambda_{i,t} \frac{\alpha}{\alpha} + \lambda_{i,t}^2 \frac{\alpha}{\alpha^2} \\
&= \lambda_{i,t} + \frac{\lambda_{i,t}^2}{\alpha}
\end{aligned}$$

Où α est le paramètre de *surdispersion* qui accroit la variance par rapport à celle de la poisson à cause de l'hétérogénéité résiduelle.
Remarques :

- $$\alpha \to \infty \Rightarrow \mathrm{Var}(\Theta) \to 0$$

  Il n'y a pas d'hétérogénéité résiduelle. On a posé toutes les questions qui affectent N.

- $$\alpha < \infty \Rightarrow \mathrm{Var}(\Theta) > 0$$

  Il y a présence d'hétérogénéité résiduelle. Il reste des questions à poser pour prévoir N.

b)

$$
\begin{aligned}
E[X_{i,t}] &= E\Big[E[X_{i,t}|\psi_i]\Big] \\
&= E\Big[\frac{1}{\beta_{i,t} \times \psi_i}\Big] \\
&= \frac{1}{\beta_{i,t}} E\Big[\frac{1}{\psi_i}\Big] \\
&= \frac{1}{\beta_{i,t}} \\
&= 5\ 000 \varphi(valeur_{i,t}) \\
Var(X_{i,t}) &= E\Big[\mathrm{Var}(X_{i,t}|\psi_i)\Big] + \mathrm{Var}\Big(E[X_{i,t}|\psi_i]\Big) \\
&= E\Big[\frac{1}{\beta_{i,t}^2 \psi_i^2}\Big] + \mathrm{Var}\Big(\frac{1}{\beta_{i,t}\psi_i}\Big) \\
&= \frac{1}{\beta_{i,t}^2} E\Big[\frac{1}{\psi_i^2}\Big] + \frac{1}{\beta_{i,t}^2}\mathrm{Var}\Big(\frac{1}{\psi_i}\Big) \\
&= \frac{1}{\beta_{i,t}^2} E\Big[\frac{1}{\psi_i^2}\Big]^2 \\
&= \frac{1}{\beta_{i,t}^2} \times 1
\end{aligned}
$$

c)Loi postériori de Θ_i

$$\begin{aligned}
f_{\Theta|N_{i,1}...N_{i,n}}(\theta|n_{i,1}...n_{i,n}) &\propto \left[\prod_{t=1}^{n} f_{N_{i,t}|\Theta}(n_{i,t}|\theta)\right] f_{\Theta}(\theta) \\
&= \left(\prod_{t=1}^{n} \frac{e^{-\lambda_{i,t}\theta_i}(\lambda_{i,t}\theta)^{n_t}}{n_t}\right) \frac{\alpha^{\alpha}\theta^{\alpha-1}e^{-\alpha\theta}}{\Gamma(\alpha)} \\
&= e^{-\theta\sum_{t=1}^{n}\lambda_{i,t}}\theta^{\sum_{t=1}^{n}n_t} \frac{\left(\prod_{t=1}^{n}\lambda_{i,t}^{n_t}\right)}{\left(\prod_{t=1}^{n}n_t\,!\right)} \frac{\alpha^{\alpha}\theta^{\alpha-1}e^{-\alpha\theta}}{\Gamma(\alpha)} \\
&\propto e^{-\theta\sum_{t=1}^{n}\lambda_{i,t}}\theta^{\sum_{t=1}^{n}n_t}\theta^{\alpha-1}e^{-\alpha\theta} \\
&= \theta^{(\alpha+\sum_{t=1}^{n}n_t)-1}e^{-\theta(\alpha+\sum_{t=1}^{n}\lambda_{i,t})} \\
&\propto \frac{\left(\alpha+\sum_{t=1}^{n}\lambda_{i,t})^{\alpha+\sum_{t=1}^{n}n_t}\right)}{\Gamma\left(\alpha+\sum_{t=1}^{n}n_t\right)}\theta^{(\alpha+\sum_{t=1}^{n}n_t)-1}e^{-\theta(\alpha+\sum_{t=1}^{n}\lambda_{i,t}}
\end{aligned}$$

Conclusion, la loi a posteriori de Θ_i

$$\left(\Theta_i|N_{i,1}...N_{i,n}\right) \sim \Gamma\left(\alpha+\sum_{t=1}^{n}N_{i,t}, \alpha+\sum_{t=1}^{n}\lambda_{i,t}\right)$$

Ainsi,

$$\begin{aligned}
B_{i,n+1}^{(n)} &= E[\mu_n(\theta_i)|N_{i,1}...N_{i,n}] \\
&= E[E[N_{i,t}|\Theta_i]|N_{i,1}...N_{i,n}] \\
&= E[\lambda_{i,t}\Theta_i|N_{i,1}...N_{i,n}] \\
&= \lambda_{i,t}E[\Theta_i|N_{i,1}...N_{i,n}] \\
&= \lambda_{i,n+1}\frac{\alpha+\sum_{t=1}^{n}N_{i,t}}{\alpha+\sum_{t=1}^{n}\lambda_{i,t}}
\end{aligned}$$

or,

$$\begin{aligned}
B_{i,n+1}^{(n)} &= \lambda_{i,n+1}\frac{\sum_{t=1}^{n}N_{i,t}}{\sum_{t=1}^{n}\lambda_{i,t}}\left(\frac{\sum_{t=1}^{n}\lambda_{i,t}}{\alpha+\sum_{t=1}^{n}\lambda_{i,t}}\right)+\left(\frac{\alpha}{\alpha+\sum_{t=1}^{n}\lambda_{i,t}}\right) \\
&= \lambda_{i,n+1}\frac{\sum_{t=1}^{n}N_{i,t}}{\sum_{t=1}^{n}\lambda_{i,t}}Z+(1-Z)\lambda_{i,t}
\end{aligned}$$

On a une prime de crédibilité.

d) Loi a posteriori de Ψ_i

Idée de la preuve : (On utilise la densité conditionnelle 3.12)

$$
\begin{aligned}
f_{\Psi_i|X_{i,1}...X_{i,n}}(\psi|X_{i,1}...X_{i,n}) \propto & \left[\prod_{i=1}^{n} f_{X_{i,t}|\psi_i}\psi_i\right] f_{\psi_i}(\psi) \\
= & \left[\prod_{i=1}^{k} \beta_{i,t} \times \psi_i e^{\beta_{i,t}\times\psi_i x_t}\right] \frac{\beta^{\beta+1}\psi^{\beta+1-1}e^{-\beta\psi}}{\Gamma(\beta+1)}
\end{aligned}
$$

Où $k = N_{i,1}...N_{i,n}$

$$
\begin{aligned}
& \propto \psi^{(k+\beta+1)-1} e^{-\psi(\beta+\sum_{t=1}^{k}\beta_{i,t}x_t)} \\
\Rightarrow & \Big(\psi|X_{i,1}...X_{i,n}\Big) \sim \Gamma\Big(\beta+k+1, \beta+\sum_{t=1}^{k}\beta_{i,t}x_t\Big)
\end{aligned}
$$

La prime Bayesienne :

$$
\begin{aligned}
E\Big[\mu_{X|X_i,...,X_n}(\psi_i)\Big] &= E\Big[E\big[X_{i,t+1}|\psi_i\big]|X_{i,1},...,X_{i,n}\Big] \\
&= E\left[\frac{1}{\beta_{i,n+1}\psi_i}|X_{i,1},...,X_{i,n}\right] \\
&= \frac{1}{\beta_{i,n+1}} E\left[\frac{1}{\psi_i}|X_{i,1},...,X_{i,n}\right] \\
B_{i,k+1}^{(x)} &= \frac{1}{\beta_{i,n+1}} \frac{\beta+\sum_{t=1}^{k}\beta_{i,t}x_t}{\beta+k} \\
&= \frac{1}{\beta_{i,n+1}} \frac{\sum_{t=1}^{k}\beta_{i,t}x_t}{k} \times \frac{k}{\beta+k} + \frac{\beta}{\beta+k} \times \frac{1}{\beta_{i,n+1}}
\end{aligned}
$$

Où

- $\frac{1}{\beta_{i,n+1}} \frac{\sum_{t=1}^{k}\beta_{i,t}x_t}{k}$ corresponds à la sévérité ajustée selon la sinistralité réelle ;
- $\frac{k}{\beta+k} =$ Z ;
- $\frac{\beta}{\beta+k} =$ 1 - Z ;
- $\frac{1}{\beta_{i,n+1}}$ corresponds à la sévérité *collective* de la classe.

$$
\begin{aligned}
\Rightarrow B_{i,n+1}^{(s)} &= B_{i,n+1}^{(n)} \times B_{i,k+1}^{(x)} \\
&= \lambda_{i,n+1} \times \frac{\alpha + \sum_{t=1}^{n} N_{i,t}}{\alpha + \sum_{t=1}^{n} \lambda_{i,t}} \times \mu_{i,n+1} \times \frac{\beta + \sum_{t=1}^{k} \frac{x_{i,t}}{\mu_{i,t}}}{\beta + k} \\
&= \lambda_{i,n+1} \times \frac{\alpha + \sum_{t=1}^{k} N_{i,t}}{\alpha + \sum_{t=1}^{k} \lambda_{i,t}} \times \mu_{i,n+1} \times \frac{\beta + \sum_{t=1}^{N_{i,1}} \frac{x_{i,t,1}}{\mu_{i,1}} + ... + \sum_{t=N_{i,n+1}}^{N_{i,n}} \frac{x_{i,t,n}}{\mu_{i,n}}}{\beta + \sum_{t=1}^{n} N_{i,t}}
\end{aligned}
$$

Chapitre 4

Modèle de Bühlmann

Le modèle de crédibilité Bayesienne du chapitre 3 est le plus précis des modèles possibles au niveau de la prévision, mais il comportent les deux limitations suivantes :

1) La prime Bayesienne est linéaire (équation 3.11) uniquement dans quelques cas.
2) *Pire* : La prime Bayesienne repose sur des hypothèses subjectives pour la distribution de $S_{i,t}|\theta_i$ et de θ_i.

Idée du modèle de Bühlmann

Approximer la prime Bayesienne par la meilleure droite (équation 2.5) possible, tout en faisant tomber les hypothèses sur les distributions de $S_{i,t}|\theta_i$ et de θ_i.

4.1 Notations & backgroud :

a) $S^2 = E[\text{var}(S_{i,t}|\theta_i)] = E[\sigma^2(\Theta_i)]$
b) $a = \text{var}(E[S_{i,t}|\theta_i]) = \text{var}(\mu(\theta_i)$
c) Delta de Kronecker :

$$\delta_{i,j} \begin{cases} 1 & , \text{ i = j} \\ 0 & , \text{ i} \neq \text{j} \end{cases}$$

Autrement dit, $\delta_{i,j}$ est une indicatrice.

d)

$$\begin{aligned}
\text{cov}(S_{i,t}, S_{i,u}) &= E[\text{cov}(S_{i,t}, S_{i,u})|\Theta_i] + \text{cov}(E[S_{i,t}], E[S_{i,u}]) \\
&= E_\theta[\delta_{t,u} \times var(S_{i,t}|\Theta_i)] + \text{cov}_{\theta_i}(\mu(\theta_i), \mu(\theta_i)) \\
&= \delta_{t,u} E_\theta[\sigma^2(\theta_i)] + \text{var}_{\theta_i}(\mu(\theta_i)) \\
&= \delta_{t,u} \times S^2 + a
\end{aligned}$$

$$\text{cov}(S_{i,t}, S_{i,u}) = \delta_{t,u} \times S^2 + a \tag{4.1}$$

e)

$$\begin{aligned}
\text{cov}(\mu(\theta_i), S_{i,t}) &= E[\text{cov}(\mu(\theta_i), S_{i,t})] + \text{cov}(E[\mu(\theta_i)|\theta_i], E[S_{i,t}|\theta_i]) \\
&= E[0] + \text{cov}(\mu(\theta_i), \mu(\theta_i)) \\
&= \text{var}(\mu(\theta_i)) \\
&= a
\end{aligned}$$

$$\text{cov}(\mu(\theta_i), S_{i,t}) = a \tag{4.2}$$

4.2 Hypothèses du modèle Bühlmann

On considère un portefeuille (ou un sous-portefeuille, exemple : une cellule de tarification) composé de I contrats et observé pendant n années.
Le i-iéme contrat possède un niveau de risque associé à l'hétérogénéité résiduelle quantifiée par la v.a. non observée θ_i et qui génère les observations $S_{i,1}...S_{i,n}$, i = 1...I.
Chaque contrat est indépendant quant à son niveau d'hétérogénéité résiduelle ($\theta_1 \perp \theta_2 \perp ... \perp \theta_I$).
Les v.a. $S_{i,t}$ représentant les sinistres totaux payés pour l'assuré i durant l'année t, et sont tel que :

$$\begin{aligned}
E[\text{var}(S_{i,t}|\theta_i)] &= \mu(\Theta_i) \\
\text{cov}(S_{i,t}, S_{i,u}|\theta_i) &= \delta_{t,u} \times \sigma^2(\theta_i)
\end{aligned}$$

Ainsi, l'idée du modèle de Bühlmann est de trouver la meilleure approximation de $\mu(\theta_i)$ de l'assuré i qui est linéaire selon les observations passées $S_{i,1}...S_{i,n}$ de cet assuré. En mathématique, ceci veut dire de trouver β_0 et β_1 qui minimisent les moindres carrées espérés, c'est-à-dire minimiser

$$\phi = E[\big(\mu(\theta_i) - \{\beta_0 + \beta_1 \overline{S}\}\big)^2]$$

Où $\beta_0 = (1-Z)m$ et $\beta_1 = Z$

Problème : Trouver β_0 et β_1 qui minimisent

$$\phi = E[\big(\mu(\theta_i) - \{\beta_0 + \beta_1 \overline{S}\}\big)^2]$$

Où $\overline{S} = \frac{\sum_{t=1}^{n} S_{i,t}}{n}$

$$\begin{aligned}
\frac{\partial \phi}{\partial \beta_0} = 0 \Rightarrow & E[\big(\mu(\theta_i) - \{\beta_0 + \beta_1 \overline{S}\}\big)^2] = 0 \\
0 = & E[\mu(\theta_i)] - \beta_0 - \beta_1 E[\overline{S}_i]
\end{aligned}$$

$$\beta_0 = E[\mu(\theta_i)] - \beta_1 E[\overline{S}_i] \tag{4.3}$$

$$\begin{aligned}
\frac{\partial \phi}{\partial \beta_1} = 0 \Rightarrow & E[\big(\mu(\theta_i) - \{\beta_0 + \beta_1 \overline{S}\}\big)^2] = 0 \\
0 = & E[\mu(\theta_i)\overline{S}_i] - \beta_0 E[\overline{S}_i] - \beta_1 E[\overline{S}_i \overline{S}_i] \\
0 = & E[\mu(\theta_i)\overline{S}_i] - \Big[E[\mu(\theta_i)] - \beta_1 E[\overline{S}_i]\Big] E[\overline{S}_i] - \beta_1 E[\overline{S}_i \overline{S}_i]
\end{aligned}$$

$$\begin{aligned}
E[\mu(\theta_i)\overline{S}_i] - E[\mu(\theta_i)]E[\overline{S}_i] &= \beta_1 \Big[E[\overline{S}_i \overline{S}_i] - E[\overline{S}_i]E[\overline{S}_i]\Big] \\
\mathrm{cov}(\mu(\theta_i), \overline{S}_i) &= \beta_1 \mathrm{cov}(\overline{S}_i, \overline{S}_i) \\
\mathrm{cov}(\mu(\theta_i), \frac{\sum_{t=1}^{n} S_{i,t}}{n}) &= \beta_1 \mathrm{cov}\left(\frac{\sum_{t=1}^{n} S_{i,t}}{n}, \frac{\sum_{u=1}^{n} S_{i,u}}{n}\right) \\
\frac{1}{n}\sum_{t=1}^{n} \mathrm{cov}(\mu(\theta_i), S_{i,t}) &= \frac{\beta_1}{n^2}\sum_{t=1}^{n}\sum_{u=1}^{n} \mathrm{cov}\Big(S_{i,t}, S_{i,u}\Big)
\end{aligned}$$

Selon l'équation 4.1 et l'équation 4.2 de la section 4.1

$$
\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} a &= \frac{\beta_1}{n^2}\sum_{t=1}^{n}\sum_{u=1}^{n}(a+\delta_{t,u}-S^2) \\
\frac{n}{n}a &= \frac{\beta_1}{n^2}(n^2a+nS^2) \\
&= \beta_1 a + \frac{\beta_1 S^2}{n} \\
&= \beta_1(a+\frac{S^2}{n}) \\
&= \beta_1 \times \frac{an+S^2}{n} \\
\beta_1 &= \frac{a\times n}{an+S^2} \\
\beta_1 &= \frac{n}{n+\frac{S^2}{a}} \\
\beta_1 &= \frac{n}{n+k}
\end{aligned}
$$

Où $k = \frac{S^2}{a} = \frac{E[\text{var}(S_{i,t}|\theta_i)]}{\text{var}(E[S_{i,t}|\theta_i)]}$

$$
\begin{aligned}
\beta_0 &= E[\mu(\theta_i)] - \beta_1 E[\overline{S}_i] \\
&= E[E[S_{i,t}|\theta_i)]] - \beta_1 E[\frac{\sum_{t=1}^{n} S_{i,t}}{n}] \\
&= E[S_{i,t}] - \beta_1 \frac{1}{n} \times nE[S_{i,t}] \\
&= m - \frac{n}{n+k} \times m \\
&= m(1-\frac{n}{n+k})
\end{aligned}
$$

Conclusion : la meilleure approximation linéaire de $\mu(\theta_1)$ qui minimalise ϕ est :

$$
\begin{aligned}
\prod_{i,n+1}^{B} &= \beta_0 + \beta_1\overline{S}_i \\
&= m(1-\frac{n}{n+k}) + \frac{n}{n+k}\overline{S}_i \\
&= Z\overline{S}_i + (1-Z)m,
\end{aligned}
$$

Où

$$\begin{aligned}
Z &= \frac{n}{n+k} \\
k &= \frac{S^2}{a} \\
\overline{S}_i &= \frac{\sum_{t=1}^{n} S_{i,t}}{n} \\
m &= E[S_{i,t}] \text{ (=Prime collective)}
\end{aligned}$$

4.3 Approche 1 : Paramétrique

Dans un premier temps, on peut considérer que les distributions de $S_{i,t}|\theta_i$ et de θ_i sont connues (comme en crédibilité Bayesienne) mais on veut le modèle de Bühlmann pour trouver la prime de crédibilité approximative linéairement.
Exemple :

$$\begin{aligned}
(S_{i,t}|\theta_i) &\sim \text{Bern}(\theta_i) \\
\theta_i &\sim \text{U}]a, b[, 0 < a < b < 1
\end{aligned}$$

Dans ce cas, la prime Bayesienne est très complexe à calculer. Par contre, la prime de Bühlmann est très simple et donne une excellente approximation de l'équation 3.8.

i) $\prod_{i,n+1}^{B} = Z\overline{S}_i + (1-Z)m$ où $Z = \frac{n}{n+\frac{S^2}{a}}$

ii) $m = E[\mu(\theta_i)] = E[E[S_{i,t}|\theta_i]] = E[\theta_i] = \frac{a+b}{2}$

iii)

$$\begin{aligned}
S^2 &= E[\sigma^2(\theta_i)] \\
&= E[\text{var}(S_{i,t}|\theta_i)] \\
&= E[\theta_i(1-\theta_i)] \\
&= E[\theta_i]E[\theta_i^2] \\
&= \frac{a+b}{2} - \frac{a^2+ab+b^2}{3}
\end{aligned}$$

iv) $a = \text{var}(\mu(\theta_i)) = \text{var}(E[S_{i,t}|\theta_i]) = \text{var}(\theta_i) = \frac{(b-a)^2}{12}$

v) $\prod_{i,n+1}^{B} = Z\overline{S}_i + (1-Z)m$

Où

$$
\begin{aligned}
Z &= \frac{n}{n+k} \\
k &= \frac{S^2}{a} = \frac{\frac{a+b}{2} - \frac{a^2+ab+b^2}{3}}{\frac{(b-a)^2}{12}} \\
m &= \frac{a+b}{2}
\end{aligned}
$$

4.4 Approche non paramétrique

En pratique, les lois de $S_{i,t}|\Theta_i$ et de θ_i sont rarement connues. Dans ce cas, pourvu que l'on soit en mesure d'estimer les paramètres de structure (m, S^2 et a), on pourra obtenir une prime de crédibilité linéaire (2.5) et très précise avec le modèle de Bühlmann.

$$
E^{cred}[X|\Theta] = \widehat{Z}(\overline{X|\theta}) + (1 - \widehat{Z})\widehat{m} \tag{4.4}
$$

4.4.1 Estimateur de m :

$$
\begin{aligned}
m &\stackrel{\text{def}}{=} E[\mu(\Theta_i)] = \text{moyenne du portefeuille de la cellule} \\
&= \frac{1}{I} \times \frac{1}{n} \sum_{i=1}^{I} \sum_{t=1}^{n} S_{i,t}
\end{aligned}
$$

Remarque :

$$
\begin{aligned}
E[\widehat{m}] &= \frac{1}{I} \times \frac{1}{n} \sum_{i=1}^{I} \sum_{t=1}^{n} E[S_{i,t}] \\
&= \frac{1}{I} \times \frac{1}{n} \times I \times n \times m \\
&= m \Rightarrow \text{Sans biais}
\end{aligned}
$$

4.4.2 Estimateur de S^2 :

$$\begin{aligned}S^2 &\stackrel{\text{def}}{=} E[\sigma^2(\Theta_i)] = \text{moyenne des variances de chaque contrat}\\ &= \frac{1}{I}\sum_{i=1}^{I}\widehat{\sigma^2}(\Theta_i)\\ &= \frac{1}{n-1}\times\frac{1}{I}\sum_{i=1}^{I}\sum_{t=1}^{n}\left(S_{i,t}-\overline{S}_i\right)^2\end{aligned}$$

Remarque :

$$\begin{aligned}E[\widehat{S}^2] &= \frac{1}{n-1}\times\frac{1}{I}\sum_{i=1}^{I}\sum_{t=1}^{n}E\Big[\left(S_{i,t}-\overline{S}_i\right)^2\Big]\\ &= \frac{1}{I}\times\frac{1}{n}\sum_{i=1}^{I}\sum_{t=1}^{n}\frac{n-1}{n}S^2\\ &= S^2 \Rightarrow \text{Sans biais}\end{aligned}$$

4.4.3 Estimateur de a :

$$a \stackrel{\text{def}}{=} \text{Var}(\mu(\Theta_i)) = \text{variance de } \mu(\Theta_i) \text{ d'un assuré à l'autre (between)}$$

Intuitivement, on pourrait penser à

$$\widehat{a} = \frac{1}{I-1}\sum_{i=1}^{I}\left(S_{i,t}-\overline{S}_i\right)^2$$

mais il ne s'agit pas d'un bon estimateur, car il est biaisé.

Par contre, on remarque intuitivement que

$$\begin{aligned}
E[\widehat{a}] &= \frac{1}{I-1}\sum_{i=1}^{I}\left(E\left[(\overline{S}_i - \overline{S})^2\right]\right) \\
&= \frac{1}{I-1}\sum_{i=1}^{I}\left(\mathrm{Var}(\overline{S}_i - \overline{S}) + E[\overline{S}_i - \overline{S}]^2\right) \\
&= \frac{1}{I-1}\sum_{i=1}^{I}\left(\mathrm{Var}(\overline{S}_i) + \mathrm{Var}(\overline{S}) - 2\mathrm{Cov}(\overline{S}_i, \overline{S})\right)
\end{aligned}$$

On développe la Covariance et la variance,

$$\begin{aligned}
\mathrm{Cov}(\overline{S}_i, \overline{S}) &= \mathrm{Cov}\Big(\overline{S}_i, \frac{1}{I}\sum_{j=1}^{I}\overline{S}_j\Big) \\
&= \frac{1}{I}\sum_{j=1}^{I}\mathrm{Cov}\Big(\overline{S}_i, \overline{S}_j\Big) \\
&= \frac{1}{I}\sum_{j=1}^{I}\delta_{i,j}\mathrm{Var}(\overline{S}_i)
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(\overline{S}) &= \mathrm{Var}\Big(\frac{1}{I}\sum_{j=1}^{I}\overline{S}_j\Big) \\
&= \frac{I}{I^2}\mathrm{Var}\Big(\overline{S}_i\Big) \\
&= \frac{\mathrm{Var}\Big(\overline{S}_i\Big)}{I}
\end{aligned}$$

On obtient ainsi,

$$\begin{aligned}
E[\widehat{a}] &= \frac{1}{I-1}\sum_{i=1}^{I}\left(\mathrm{Var}(\overline{S}_i) + \frac{\mathrm{Var}\left(\overline{S}_i\right)}{I} - \frac{2}{I}\mathrm{Var}(\overline{S}_i)\right)\\
&= \frac{1}{I-1}\sum_{i=1}^{I}\left(\frac{I-1}{I}\mathrm{Var}(\overline{S}_i)\right)\\
&= \mathrm{Var}(\overline{S}_i)\\
&= E[\mathrm{var}(S_{i,t}|\theta_i)] + \mathrm{Var}(E[(S_{i,t}|\theta_i)])\\
&= E\left[\frac{\sigma^2(\theta_i)}{n}\right] + \mathrm{Var}(\mu(\theta_i))\\
&= \frac{S^2}{n} + a
\end{aligned}$$

On obtient ainsi un résultat asymptotiquement sans biais.

$$E[\widehat{a}] = a + \frac{S^2}{n}$$

Solution alternative

Des alternatives ont été proposées,

- $\widehat{a}' = \widehat{a} - \frac{S^2}{n}$ Par contre, cet estimateur peut prendre des valeurs négatives (impossible pour une variance). On obtient l'estimateur suivant

$$\begin{aligned}
\widehat{a}' &= \widehat{a} - \frac{\widehat{S}^2}{n}\\
&= \frac{1}{I-1}\sum_{i=1}^{I}\left(S_{i,t} - \overline{S}_i\right)^2 - \frac{\widehat{S}^2}{n}
\end{aligned}$$

- $\widehat{a}'' = \max\{\widehat{a}', 0\}$ Cet estimateur est biaisé et illogique, car $Z = \frac{n}{n+\frac{S^2}{a}} \not\forall$,si $\widehat{a}'' = 0$.

Exemple

Soit le portefeuille suivant où I = 3 et n = 6.

	Années d'expériences						
Contrat	1	2	3	4	5	6	$\overline{S}_i$
1	0	1	2	1	2	0	1
2	3	4	2	1	4	4	3
3	3	3	2	1	2	1	2

On obtient les calculs suivants :

$$\begin{aligned}
\overline{S} &= \widehat{m} = 2 \\
\widehat{S}^2 &= \frac{1}{3}\frac{1}{6-1}\sum_{i=1}^{3}\sum_{t=1}^{6}\left(S_{i,t} - \overline{S}_i\right)^2 = \frac{16}{15} \\
\widehat{a} &= \frac{1}{3-1}\sum_{i=1}^{3}\left(S_{i,t} - \overline{S}_i\right)^2 - \frac{\frac{16}{15}}{6} = \frac{37}{45} \\
\widehat{K} &= \frac{\widehat{S}^2}{\widehat{a}} \approx 1.30 \\
\Pi_{i,7}^{B} &= Z\overline{S}_i + (1-Z)m \\
&= Z\overline{S}_i + (1-Z) \times 2 \\
Z &= \frac{n}{n+k} \\
&= \frac{6}{6+1.30} \\
&= 0.82
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\Pi_{1,6+1}^{B} &= 0.82 \times 1 + (1-0.82) \times 2 = 1.18 \\
\Pi_{2,6+1}^{B} &= 0.82 \times 3 + (1-0.82) \times 2 = 2.82 \\
\Pi_{3,6+1}^{B} &= 0.82 \times 2 + (1-0.82) \times 2 = 2
\end{aligned}$$

Chapitre 5

Bühlmann-Straub

En pratique, le modèle de Bùhlmann est impossible à appliquer, particulièrement en raison de la notion d'unité d'exposition au risque, laquelle découle du fait que :

— Les contrats ont des dates effectives différentes.
— Certains assurés peuvent annuler leur police en cours de contrats.

Exemple 1 :

L'assuré A achète une police le 01/04/2014 et la conserve pendant 1 an. On s'intére à l'exposition du risque en 2015

Assuré i	2015 UA	1990 $S_{i,1}$	...	2014 $S_{i,n-1}$	2015 $S_{i,n}$
$\vdots$	$\frac{9}{12}$	...	...	...	...
A	$\frac{3}{12}$	0	...	2 000\$	1 500\$

On remarque que pour 2015, l'exposition était de $\frac{3}{12}$ pour un sinistre de 1 500\$, ce qui représente un risque plus grand que la situation en 2014 pour ce même assuré avec une exposition de 1 pour un sinistre de 2 000\$.
Le modèle de Bühlmann ne permet que de considérer des cas parfaits où l'exposition de chaque assuré dans chaque année est de 1. On utilise alors un modèle plus général, le modèle Bühlmann-Straub.

Exemple 2

L'assuré B achète une police 1 an le 1er juillet 2015 et annule son contrat le 1er février 2017. On assume que l'assuré a renouvelé son contrat entre 2016 et 2017.

Année	UA
2000	0
2001	0
$\vdots$	$\vdots$
2013	0
2014	0
2015	$\frac{1}{2}$
2016	$\frac{1}{2}+\frac{1}{2}=1$
2017	$\frac{1}{2}$
2018	0

assuré i	2000 $S_{i,1}$	...	2016 $S_{i,n}$	2017 $S_{i,n}$
1	0	...	...	...
2	0	...	...	...
$\vdots$	0	...	...	...
B	0	...	1 000$	2 000$
$\vdots$	0	...	...	...
I	0	...	...	...

On remarque la même situation d'exposition au risque. Le modèle de Bûlhmann ne permet pas d'introduire cette *situation*.

Contexte plus général en pratique

Dans un contexte plus général, *l'historique* prend la forme suivante

		Sinistres totaux	Exposition	Prime pure
Contrat	Hétérogénéité	t = 1 ... t =n	t = 1 ... t =n	t = 1 ... t =n
1	θ_1	$S_{1,1} \dots S_{1,n}$	$w_{1,1} \dots w_{1,n}$	$X_{1,1} \dots X_{1,n}$
$\vdots$	$\vdots \dots \vdots$	$\vdots \dots \vdots$	$\vdots \dots \vdots$	$\vdots \dots \vdots$
I	θ_I	$S_{I,1} \dots S_{I,n}$	$w_{I,1} \dots w_{I,n}$	$X_{I,1} \dots X_{I,n}$

Remarques :

1) Avec Bûhlman-Straub, on modélisera $X_{i,t}$ et non $S_{i,t}$ (on s'intéresse toujours au ratio)
2) Lorsque $w_{i,t} = 1; \forall i = 1, ...I$ et $t = 1, ..., n$, alors tous les résultats du modèle de Bühlmann-Straub seront égaux à ceux du modèle de Bühlmann.
3) Formellement, $w_{i,t}$ = pourcentage de l'année t où l'assuré i a été exposé au risque.
4) Formellement, $X_{i,t} = \frac{S_{i,t}}{w_{i,t}}$ = Prime pure de l'assuré i durant l'année t, soit le coût moyen *annualisé* de l'assuré i durant l'année t.
5) Occassionnellement, l'actuaire peut définir $W_{i,t}$ comme le total des primes acquises pour le contrat i durant l'année t, dans ce cas $X_{i,t} = \frac{S_{i,t}}{w_{i,t}}$ devient le loss ratio (au lieu de la prime pure)

Hypothèse de Bühlmann-Straub

1) Les contrats sont indépendants $\Leftrightarrow \Theta_i \perp \Theta_2 \perp ... \perp \Theta_I$
2)

$$\begin{aligned}
E[X_{i,t}|\Theta_i] &= E\Big[\frac{S_{i,t}}{w_{i,t}} \mid \theta_i\Big] \\
&= E\Big[\frac{\sum_{k=1}^{w_{i,t}} S_{i,t}^{(k)}}{w_{i,t}} \mid \theta_i\Big] \\
&= \frac{\sum_{k=1}^{w_{i,t}} E[S_{i,t}^{(k)} \mid \theta_i]}{w_{i,t}} \\
&= \frac{w_{i,t}\mu(\theta_i)}{w_{i,t}} \\
&= \mu(\theta_i)
\end{aligned}$$

3)

$$\begin{aligned}
\mathrm{Var}(X_{i,t}|\Theta_i) &= \mathrm{Var}\Big(\frac{S_{i,t}}{w_{i,t}} \mid \theta_i\Big) \\
&= \frac{1}{w_{i,t}^2}\mathrm{Var}\Big(\sum_{k=1}^{w_{i,t}} S_{i,t}^{(k)} \mid \theta_i\Big) \\
&= \frac{w_{i,t}}{w_{i,t}^2}\sigma^2(\theta_i) \\
&= \frac{\sigma^2(\theta_i)}{w_{i,t}}
\end{aligned}$$

Analogie avec ACT-2000

$$\begin{aligned}
E[\overline{X}] &= n \times E[X] = E[X] \\
\mathrm{Var}(\overline{X}) &= \frac{\mathrm{Var}(X)}{n}
\end{aligned}$$

5.1 Notations supplémentaires :

- $w_{i\bullet} = \sum_{t=1}^{n} w_{i,t}$
- $w_{\bullet\bullet} = \sum_{i=1}^{I} \sum_{t=1}^{n} w_{i,t}$
- $Z_{\bullet} = \sum_{i=1}^{I} Z_i$
- $X_{iW} = \sum_{i=1}^{n} \left(\frac{w_{i,t}}{w_{i\bullet}}\right) \times X_{i,t}$

  Ce qui correspond à

$$\begin{aligned}
X_{iW} &= \frac{\text{Pourcentage de temps où l'assuré a été exposé au risque dans l'année } t}{\text{Pourcentage de temps où l'assuré a été exposé au risque entre l'année } \mathit{1} \text{ et l'année } n} \\
&= \text{Pourcentage de l'exposition au risque attribuable à l'an } t \text{ pour l'assuré i}
\end{aligned}$$

- $$\begin{aligned} X_{WW} &= \sum_{i=1}^{I}\sum_{t=1}^{n}\left(\frac{w_{i,t}}{w_{\bullet\bullet}}\right) \times X_{i,t} \\ &= \sum_{i=1}^{t}\frac{W_{i\bullet}}{W_{\bullet\bullet}}\sum_{t=1}^{n}\frac{W_{i,t}}{W_{i\bullet}} \times X_{i,t} \\ &= \sum_{i=1}^{I}\frac{W_{i\bullet}}{W_{\bullet\bullet}} \times X_{iw} \end{aligned}$$

- $X_{ZW} = \sum_{i=1}^{I}\left(\frac{Z_i}{Z_\bullet}\right) \times X_{iW}$

Covariances

- $\mathrm{Cov}(X_{i,t}, X_{k,u}) = \delta_{i,k}\left(a + \frac{\delta_{t,u} \times S^2}{w_{i,t}}\right)$
- $\mathrm{Cov}(\mu(\Theta_i), X_{k,u}) = \delta_{i,k} \times a$
- $\mathrm{Cov}(X_{i,t}, X_{kw}) = \delta_{i,k} \times \left(a + \frac{S^2}{w_{i\bullet}}\right)$

5.2 Prime de crédibilité linéaire

Comme précédemment (section 4.2), on cherche β_0 et β_1 qui minimisent

$$\phi = E\left[\left(\mu(\theta_i) - \{\beta_0 - \beta_1 X_{iW}\}\right)^2\right]$$

Solution :

$$\Pi_{i,n+1}^{B-S} = Z_i \times X_{iW} + (1 - Z_i) \times m$$

avec $Z_i = \frac{w_{i\bullet}}{w_{i\bullet} + \frac{S^2}{a}}$

5.3 Estimation des paramètres de structure :

5.3.1 Estimation de m

$$
\begin{aligned}
\widehat{m} &= X_{ww} \\
&= \sum_{i=1}^{I} \left(\frac{w_{i\bullet}}{w_{\bullet\bullet}}\right) \times X_{iw} \\
&= \frac{\sum_{i=1}^{I} w_{i\bullet} \sum_{t=1}^{n} \left(\frac{w_{i,t}}{w_{i\bullet}}\right) \times X_{i,t}}{w_{\bullet\bullet}} \\
&= \frac{\sum_{i=1}^{I} \sum_{t=1}^{n} w_{i,t} \left(\frac{S_{i,t}}{w_{i,t}}\right)}{\sum_{i=1}^{I} \sum_{t=1}^{n} w_{i,t}} \\
&= \frac{\sum_{i=1}^{I} \sum_{t=1}^{n} S_{i,t}}{\sum_{i=1}^{I} \sum_{t=1}^{n} w_{i,t}} \\
&= \frac{\sum \text{sinistre}}{\sum \text{UA}} \\
&= \overline{PP}
\end{aligned}
$$

Remarque : On peut démontrer que l'estimateur linéaire de m a variance minimale n'est pas X_{ww} mais bien ;

$$
\widehat{m}' = X_{Zw} = \sum_{i=1}^{I} \left(\frac{Z_i}{Z_\bullet}\right) \times X_{iw}
$$

5.3.2 Estimation de S^2

En généralisant l'estimateur du modèle de Bühlmann, on a :

$$
\widehat{S}^2 = \frac{1}{I} \frac{1}{n-1} \sum_{i=1}^{I} \sum_{t=1}^{n} w_{i,t} \Big(X_{i,t} - X_{iw}\Big)^2
$$

5.3.3 Estimation de a

— L'équivalent biaisé en Bühlmann-Straub : $\widehat{a} = \sum_{i=1}^{I} \Big(X_{iw} - X_{ww}\Big)^2 w_{i\bullet}$

— Estimateur $\widehat{a}'$ avec correction par biais :

$$\widehat{a}' = \Big(\frac{w_{\bullet\bullet}}{w_{\bullet\bullet}^2 - \sum_{i=1}^{x} w_{i\bullet}^2}\Big)(\widehat{a} - (I-1)S^2)$$

Cet estimateur peut-être négatif.

5.3.4 Autre estimateur de a

$$\widetilde{a} = \frac{1}{I-1}\sum_{i=1}^{I} Z_i (X_{iw} - X_{Zw})^2 \tag{5.1}$$

Remarques : $\widetilde{a} = \phi(Z_i)$, or $Z_i = \frac{w_{i\bullet}}{w_{i\bullet} + \frac{S^2}{\widetilde{a}}}$ Donc, $Z_i = \phi(\widetilde{a})$ Alors, $\widetilde{a}$ est un pseudoestimateur qui doit être estimé numériquement.

Chapitre 6

Réserves en IARD

6.1 But et motivation

Un des rôles de l'assurance IARD corporatifs est l'évaluation des provisions actuarielles (réserve).
Contrairement à l'assurance vie, les réserves IARD sont définies par l'Autorité des Marchés Financiers comme :

> Le processus d'évaluation du montant total nécessaire pour acquitter tous les paiements futurs associés aux sinistres déjà survenus en date d'évaluation (ex. au 31 décembre).

Cette évaluation est très sérieuse, car les réserves IARD représentent ~ 75 % du passif total de ces assureurs, leurs valeurs affectent directement le profit de la compagnie.
Pour cette raison l'*AMF* requiert notamment que les réserves soient signées par l'actuaire désigné, qui doit obligatoirement être *fellow* de la *CAS* et de l'*ICA*.
Bien que les réserves doivent être déposées à l'AMF au 31 décembre de chaque année, plusieurs assureurs les réévaluent au trimestre ou même au mois.

Remarques

1. Habituellement, les réserves IARD sont évaluées par ligne d'affaires. (Assurance auto, habitation, bien commercial, RC entreprise, ferme, aviation, ...)
   Ligne d'affaires = produits que l'assureur vend
2. Il est très important de bien évaluer le niveau des réserves IARD

a) Si les réserves sont sous-évaluées :
   — On surévalue la santé financière de la compagnie
   — On expose la compagnie au risque de défaut sur ses paiements futurs
   — On expose l'assureur à la ruine
b) Si les réserves sont surévaluées
   — Les dépenses sont plus élevées
   — Le profit est diminué
   — Le montant d'impôt à payer est diminué
   — Le surplus est diminué
   — La valeur de la compagnie est moindre, donc la valeur de l'action est sous-évaluée.

6.2 Définitions

Différence entre les réserves vie et IARD

Réserves vie

$$VP_{@u}\Big(\text{Prestations à payer pour les sinistres futurs}\Big) - VP_{@u}\Big(\text{Primes futures à recevoir}\Big)$$

Réserves IARD

$$\sum\left(\text{Prestations à payer pour le développement sur les dossiers ouverts en u}\right) - 0$$

Conséquence

En IARD, **LA** problématique derrière l'évaluation des réserves est surtout la modélisation mathématique du développement.

a) Dossier de sinistre (Claim file)

Dès qu'un sinistre est déclaré à l'assureur, celui-ci ouvre un *dossier* dans son système informatique. Une réclamation correspond à une ouverture d'un dossier. Ce dossier contient plusieurs informations sur la réclamation, tel que

— La date de perte

— La date de réclamation
— Le montant et date de chaque paiement
— Des informations qualitatives

b) Case reserves (réserves automatiques)

Une estimation des coûts totaux associés à un dossier de sinistres. Cette estimation est faite dès l'ouverture du dossier.

c) Gross IBNR (Incurred but not reported)

Comporte quatre composantes :

96 % Total des provisions pour le développement des sinistres déjà ouverts
1 % Les provisions pour dossiers fermés pouvant rouvrir
1 % Les provisions pour sinistres encourus, mais non rapportés
1 % Les provisions pour les sinistres rapportés, mais non codifiés dans le système informatique.

d) Réserve totale

Réserve totale = Case reserve (10 %) + Gross IBNR (90 %)

e) Développement

Changement temporel de la somme des paiements faits par l'assureur pour tous ses assurés.

f) Paid age-to-age

Le développement entre deux dates données.

g) Age-to-ultimate development

Développement des sinistres qui ont atteint un âge donné jusqu'à l'ultime moment où ces dossiers sont fermés.

h) Loss development factor ($LDF_{j-1,j}$)

Aussi appeler le *Paid loss development factor*.

$$LDF_{j-1,j} = \frac{\text{Somme des paiements faits par l'assureur sur tous les dossiers de sinistres à t = j}}{\text{Somme des paiements fait par l'assureur sur tous les dossiers des sinistres à t = j - 1}}$$

i) Sinistres en suspend (SS)

Somme des *cases reserves* révisées pour tous les sinistres qui ne sont pas encore fermés à une date donnée.

j) Sinistres payés (SP)

Somme des montants d'indemnités payés aux réclamants à une date donnée. Inclus les frais de règlement aussi.

k) Sinistres encourus (SE)

Somme des montants engendrés par des sinistres qui ont donné (SP), ou qui donneront (SS) lieu à des paiements d'indemnités à une date donnée.

$$SE = SP + SS$$

l) Incurred loss development factor ($ILDF_{j-1,j}$)

$$\begin{aligned} ILDF_{j-1,j} &= \frac{\sum SE_{@j}}{\sum SE_{@j-1}} \\ &= \frac{C_j}{C_{j-1}} \\ &= \frac{\text{Encouru cumulatif à } t = j}{\text{Encouru cumulatif à } t = j-1} \end{aligned}$$

m) Notations en triangles cumulatifs ($\triangleleft$)

On débute par un exemple pour illustrer la présentation des données.
Ex. :

AA/âge	12 mois	24 mois	36 mois	48 mois	60 mois
2012	$C_{2012,12}$	$C_{2012,24}$	$C_{2012,36}$	$C_{2012,48}$	$C_{2012,60}$
2013	$C_{2013,12}$	$C_{2013,24}$	$C_{2013,36}$	$C_{2013,48}$	?
2014	$C_{2014,12}$	$C_{2014,24}$	$C_{2014,36}$	?	?
2015	$C_{2015,12}$	$C_{2012,24}$	?	?	?
2016	$C_{2016,12}$	?	?	?	?

Où AA est l'année d'accident qui correspond à l'année de la date de survenance du sinistre.

Notation académique

Afin d'alléger la notation, on utilise la notation simplifiée suivante :

$AA_i/\widehat{a}ge_j$	1	2	3	4	5
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$	?
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	?	?
4	$C_{4,1}$	$C_{4,2}$	?	?	?
5	$C_{5,1}$	?	?	?	?

Ainsi $C_{i,j}$ est le montant total payé par l'assureur pour tous les sinistres survenus lors de l'année i (AA_i) et développés pendant j années ($\widehat{a}ge_j$).

6.3 Méthode du Chain-Ladder

Cette méthode est la plus employée en pratique. Elle suppose

$$C_{i,j+1} = C_{i,j} \times LDF_j$$

Si on dispose d'un triangle de n années de données cumulatives, alors la méthode de C-L revient à effectuer n-1 régressions linéaires simples sans ordonnées à l'origine.
Ainsi, par la théorie de la régression linéaire simple, on peut montrer que

$$\widehat{LDF}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}; \forall j = 1, 2, ..., n-1$$

Exemple :

$AA_i/\widehat{a}ge_j$	1	2	3	4	5
1	100	150	175	180	200
2	110	168	192	205	?
3	115	169	202	?	?
4	125	185	?	?	?
5	150	?	?	?	?

On obtient les LDF suivants :

$$\begin{aligned}
\widehat{LDF}_4 &= \frac{200}{180} = 1.11 \\
\widehat{LDF}_3 &= \frac{180+205}{175+192} = 1.049046 \\
\widehat{LDF}_2 &= \frac{175+192+202}{150+168+169} = 1.168738 \\
\widehat{LDF}_1 &= \frac{150+168+169+185}{100+110+115+125} = 1.493
\end{aligned}$$

Selon C-L

$$\begin{aligned}
\widehat{C}_{2,5} &= C_{2,4} \times \widehat{LDF}_4 = 205 \times 1.111 = 227.77 \\
\widehat{C}_{3,4} &= C_{3,3} \times \widehat{LDF}_3 = 202 \times 1.049046 = 211.9074 \\
\widehat{C}_{3,5} &= \widehat{C}_{3,4} \times \widehat{LDF}_4 \\
&= \left(C_{3,3} \times \widehat{LDF}_3 \right) \times \widehat{LDF}_4 \\
&= 202 \times 1.04906 \times 1.111 \\
&= 235.4526 \\
\widehat{C}_{4,5} &= \widehat{C}_{4,2} \times \widehat{LDF}_2 \times \widehat{LDF}_3 \times \widehat{LDF}_4
\end{aligned}$$

On note une forme de récurrence dans les développements.
Cette forme de récurrence nous permet d'estimer les informations dans le triangle de donnée précédent,

$AA_i/\widehat{a}ge_j$	1	2	3	4	5
1	100	150	175	180	200
2	110	168	192	205	227.77
3	115	169	202	211.9074	235.4526
4	125	185	216.21653	226.8211	251.7714
5	150	223.95	261.7389	274.57612	304.7794

En résumé, selon C-L

$$\widehat{C}_{i,n} = C_{i,n-i+1} \times LDF_{n-i+1} \times ... \times LDF_{n-1}; i = 1, 2, ...n \tag{6.1}$$

Réserve selon C-L

La réserve C-L pour l'AA i est notée $\widehat{R_i}$ et correspond à la différence entre les sinistres *ultimes* ($\widehat{C}_{i,n}$) et les sinistres payés en date d'évaluation ($\widehat{C}_{i,n-i+1}$).

$$\widehat{R_i} = \widehat{C}_{i,n} - \widehat{C}_{i,n-i+1} \tag{6.2}$$

De notre exemple précédant, on obtient

i	$\widehat{C}_{i,n}$	$\widehat{C}_{i,n-i+1}$	$\widehat{R_i}$
1	200	200	0
2	227.77	205	22.78
3	235.45	202	33.45
4	251.7714	185	66.7714
5	304.7794	150	154.7795

Finalement, la réserve totale inscrite au passif des états financiers est notée $\widehat{R}$ et est telle que

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

Dans notre exemple, elle correspond à

$$\begin{aligned} R &= \sum_{i=1}^{n} \widehat{R}_i \\ &= 0 + 22.78 + 33.45 + 66.7714 + 154.7795 \\ &= 277.7809 \end{aligned}$$

Remarques sur C-L

1) C-L suppose que les AA (=lignes du triangle) sont indépendantes.
2) C-L assume que l'âge des sinistres est la seule variable explicative du développement.
3) C-L suppose aussi $LDF_{i,j} = LDF_j$, n'est pas fonction de l'AA i.
4) Avec Chain-Ladder, la réserve $\widehat{R}_n$ de l'AA la plus récente est sujette à une forte incertitude.

6.4 Méthode de Bornhuetter-Ferguson

L'objectif de cette méthode est de permettre une meilleure *stabilité* dans l'estimation des réserves $\widehat{R}_i$ pour les AA i les plus récentes. En d'autres mots, on veut corriger le désavantage principal de Chain-Ladder avec Bornhuetter-Ferguson .

Trois grandes étapes :

1. Estimation des sinistres ultimes $\Rightarrow \widehat{C}_{i,n}$
   En pratique, on utilise un taux de sinistre *(Loss ratio)* qui est *assumé* comme *connu*, et on multiplie ce *loss ratio* par les primes acquises (PA_i). Soit,
   $$\widehat{C}_{i,n}^{B-F} = PA_i \times LR_i \,;\, i = 1, 2, ..., n$$
   Où LR_i est un avis d'expert.
2. Calcul des facteurs de développements $(\widehat{LDF}_i)$
   Avec la même méthode/formule qu'avec la Chain-ladder
   $$\widehat{LDF}_j = \frac{\sum_{i=1}^{n-1} C_{i,j+1}}{\sum_{i=1}^{n-1} C_{i,j}}$$
3. Calcul de la réserve
   On débute avec la définition de $\widehat{R}_i^{C-L}$,
   $$\begin{aligned}\widehat{R}_i &= \widehat{C}_{i,n} - C_{i,n-i+1} \\ &= \widehat{C}_{i,n} \times \left(1 - \frac{C_{i,n-i+1}}{\widehat{C}_{i,n}}\right) \\ &\overset{(6.1)}{=} \widehat{C}_{i,n} \times \left(1 - \frac{C_{i,n-i+1}}{\widehat{C}_{i,n-i+1} \times \prod_{j=n-i+1}^{n-1} \widehat{LDF}_j}\right)\end{aligned}$$

$$\widehat{R}_i^{B-F} = \widehat{C}_{i,n}^{B-F} \times \left(1 - \frac{1}{\prod_{j=n-i+1}^{n-1} \widehat{LDF}_j}\right) \tag{6.3}$$

Exemple

AA	PA	LR	$\widehat{C}_{i,n}^{BF} = U$	$1 - \frac{1}{\prod LDF_i}$
2010	330	0.6	198	0
2011	350	0.65	228	0.1
2012	365	0.7	256	0.14
2013	385	0.75	289	0.26
2014	400	0.8	320	0.51

À l'aide de l'équation 6.3

$$\begin{aligned}
R_1^{B-F} &= \widehat{C}_{1,5}^{BF} \times \left(1 - \frac{1}{\prod_{j=5-1+1}^{5-1} \widehat{LDF}_i}\right) \\
&= \widehat{C}_{1,5}^{BF} \times \left(1 - \frac{1}{\prod_{j=5}^{4} \widehat{LDF}_i}\right) \\
&= 198 \times 0 \\
&= 0 \\
R_2^{B-F} &= \widehat{C}_{2,5}^{BF} \times \left(1 - \frac{1}{\prod_{j=5-2+1}^{5-1} \widehat{LDF}_i}\right) \\
&= \widehat{C}_{2,5}^{BF} \times \left(1 - \frac{1}{\widehat{LDF}_4}\right) \\
&= 228 \times 0.1 \\
&= 22.8 \\
R_3^{B-F} &= \widehat{C}_{3,5}^{BF} \times \left(1 - \frac{1}{\prod_{j=5-3+1}^{5-1} \widehat{LDF}_i}\right) \\
&= \widehat{C}_{3,5}^{BF} \times \left(1 - \frac{1}{\widehat{LDF}_3 \widehat{LDF}_4}\right) \\
&= 256 \times 0.14 \\
&= 35.84
\end{aligned}$$

$$
\begin{aligned}
R_4^{B-F} &= \widehat{C}_{4,5}^{BF} \times \left(1 - \frac{1}{\prod_{j=5-4+1}^{5-1} \widehat{LDF}_i}\right) \\
&= \widehat{C}_{4,5}^{BF} \times \left(1 - \frac{1}{\widehat{LDF}_2 \widehat{LDF}_3 \widehat{LDF}_4}\right) \\
&= 289 \times 0.26 \\
&= 75.14 \\
R_5^{B-F} &= \widehat{C}_{5,5}^{BF} \times \left(1 - \frac{1}{\prod_{j=5-5+1}^{5-1} \widehat{LDF}_i}\right) \\
&= \widehat{C}_{5,5}^{BF} \times \left(1 - \frac{1}{\widehat{LDF}_1 \widehat{LDF}_2 \widehat{LDF}_3 \widehat{LDF}_4}\right) \\
&= 320 \times 0.51 \\
&= 163.2
\end{aligned}
$$

AA	R_i
2010	0
2011	22.8
2012	35.84
2013	75.14
2014	163.2
$\sum_{i=2010}^{2014} R_i =$	296.98

Terminologie

- $\frac{1}{\prod_{j=n-i+1}^{n-1} \widehat{LDF}_i}$ est appeler le pourcentage des sinistres développés. À partir de l'exemple précédant, le pourcentage des sinistres développés de l'année 2011 est de 90 %.
- $1 - \frac{1}{\prod_{j=n-i+1}^{n-1} \widehat{LDF}_i}$ est appeler le pourcentage des sinistres non développés. À partir de l'exemple précédant, le pourcentage des sinistres non développés de l'année 2011 est de 10 %.

Remarques

1. L'avantage de Bornhuetter-Ferguson est qu'elle procure une meilleure stabilité dans les réserves $\widehat{R}_i$ des AA i plus récentes.

2. Le désavantage souvent cité de Bornhuetter-Ferguson est l'utilisation de données *externes* (PA_i et LR_i) qui peuvent introduire de la *subjectivité*.
3. Dans la méthode de Bornhuetter-Ferguson, c'est un peu comme si on faisait l'hypothèse que l'âge des sinistres n'affecte pas le développement.

$$\bullet B - F \Rightarrow \widehat{C}_{i,n}^{B-F} = PA_i \times LR_i \equiv a_i + C_{i,n-i+1} \times 0$$

$$\bullet C - L \Rightarrow \widehat{C}_{i,n}^{C-L} = C_{i,n-i+1} \times \Big(\prod_{j=n-i+1}^{n-1} \widehat{LDF}_j \Big) = 0 + C_{i,n-i+1} \times \beta_i$$

6.5 Méthode de Mack

Bien que la méthode Chain-ladder et la méthode Bornhuetter-Ferguson soient très efficaces pour estimer les réserves espérées, elles ont la limitation de ne pas pouvoir quantifier la variance de ces réserves. La méthode de Mack permet ceci en ajustant un cadre stochastique à la méthode Chain-Ladder.

Trois hypothèses de Mack

1. Les AA sont indépendantes, c'est à dire $\{C_{1,1}, ..., C_{1,n}\} \perp \{C_{2,1}, ..., C_{2,n}\} \perp ...$
2. E $\Big[C_{i,k+1} | C_{i,1}, ..., C_{i,k}\Big] = C_{i,k} \times LDF_k \, ; \, i = 1, ..., n$ et $k = 1, ..., n-1$
3. Var $\Big(C_{i,k+1} | C_{i,1}, ..., C_{i,k}\Big) = C_{i,k} \times \sigma_k^2 \, ; \, i = 1, ..., n$ et $k = 1, ..., n-1$

Définition

— $H_i = \{C_{i,1}, ..., C_{i,n-i+1}\}$: Les données $C_{i,j}$ disponible dans le triangle pour l'AA i.

— $D = \{C_{i,j} : i + j \leq n + 1\}$: Le triangle de données complet.

Sous cette spécification, on peut montrer que le meilleur estimateur de LDF_j est les mêmes qu'avec Chain-Ladder.

$$\widehat{LDF}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}$$

Idée de la preuve précédante

$$\begin{aligned}
E[C_{i,j}|D] &\overset{H1}{=} E[C_{i,j}|H_i] \\
&= E\Big[E[C_{i,j}|C_{i,j},...,C_{i,j-1}]\Big|H_i\Big] \\
&\overset{H2}{=} E[C_{i,j-1}|H_i] \times LDF_{j-1} \\
&= \text{à appliquer plusieurs fois...} \\
&= C_{i,n-i+1} \times LDF_{n-i+1} \times ... \times LDF_{j-1}
\end{aligned}$$

Ce qui coïncide avec Chain-Ladder.

Conséquences des hypothèses

Une première conséquence des hypothèses de Mack est

$$\begin{aligned}
E[\widehat{LDF}_k|D_k] &\overset{\text{déf}}{=} E\left[\frac{\sum_{i=1}^{n-k} C_{i,k+1}}{\sum_{i=1}^{n-k} C_{i,k}}\middle|D_k\right] \\
&= \frac{E[\sum_{i=1}^{n-k} C_{i,k+1}|D_k]}{\sum_{i=1}^{n-k} C_{i,k}} \\
&= \frac{\sum_{i=1}^{n-k} E[C_{i,k+1}|D_k]}{\sum_{i=1}^{n-k} C_{i,k}} \\
&= \frac{\sum_{i=1}^{n-k} C_{i,k} \times LDF_k}{\sum_{i=1}^{n-k} C_{i,k}} \\
&= \left(\frac{\sum_{i=1}^{n-k} C_{i,k}}{\sum_{i=1}^{n-k} C_{i,k}}\right) \times LDF_k \\
E[\widehat{LDF}_k|D_k] &= LDF_k \text{ (sans biais)}
\end{aligned}$$

De plus, pour $k < j$, on a que

$$\begin{aligned}
E[\widehat{LDF}_k \times \widehat{LDF}_j] &= E\Big[E[\widehat{LDF}_k \times \widehat{LDF}_j|D_j]\Big] \\
&= E\Big[LDF_k \times E[\widehat{LDF}_j|D_j\Big] \\
&= E[\widehat{LDF}_k] \times E[\widehat{LDF}_j] \\
\Rightarrow \text{cov}(\widehat{LDF}_k, \widehat{LDF}_j) &= E[\widehat{LDF}_k \times \widehat{LDF}_j] - E[\widehat{LDF}_k] \times E[\widehat{LDF}_j] \\
&= 0 \rightarrow \text{(non corrélés)}
\end{aligned}$$

Cette non-corrélation est essentielle, elle permet d'écrire le développement suivant

$$
\begin{aligned}
E[\widehat{R}_i] &= E[\widehat{C}_{i,n} - \widehat{C}_{i,n-i+1}] \\
&= E[C_{i,n-i+1} \times \widehat{LDF}_{n-i+1} \times ... \times \widehat{LDF}_{n-1} - \widehat{C}_{i,n-i+1}] \\
&= C_{i,n-i+1} E[\widehat{LDF}_{n-i+1} \times ... \times \widehat{LDF}_{n-1}] - \widehat{C}_{i,n-i+1} \\
&= C_{i,n-i+1} E[\widehat{LDF}_{n-i+1}] \times ... \times E[\widehat{LDF}_{n-1}] - \widehat{C}_{i,n-i+1} \\
&= C_{i,n-i+1} LDF_{n-i+1} \times ... \times LDF_{n-1} - \widehat{C}_{i,n-i+1} \\
&= R_i
\end{aligned}
$$

Sous Mack, $\widehat{R}_i$ est un estimateur sans biais de la réserve de l' AA i.

Remarque

$$
\begin{aligned}
E[\widehat{R}] &= E\Big[\sum_{i=1}^{n} \widehat{R}_i\Big] \\
&= \sum_{i=1}^{n} E\Big[\widehat{R}_i\Big] \\
&= \sum_{i=1}^{n} R_i \\
&= R \text{ (sans biais)}
\end{aligned}
$$

Estimation de la volatilité (la variance) des réserves

Rappel

L'hypothèse numéro 3 : Var $\Big(C_{i,j+1}|C_{i,1},...,C_{i,j}\Big) = C_{i,j} \times \sigma_j^2$

Proposition de Mack pour estimer $\widehat{\sigma}_j$

À partir du triangle de données, on peut trouver les relations suivantes :

$$\begin{aligned}\widehat{LDF}_2 &= \frac{\sum_{i=1}^{3} C_{i,3}}{\sum_{i=1}^{3} C_{i,2}} \\ &= \sum_{i=1}^{3}\left(\frac{C_{i,2}}{\sum_{k=1}^{3} C_{k,2}}\right)\frac{C_{i,3}}{C_{i,2}}\end{aligned}$$

On observe que les *age to age factor* correspondent à

$$\begin{aligned}A_{to}A_{1,2} &= \frac{C_{1,3}}{C_{1,2}} \\ A_{to}A_{2,2} &= \frac{C_{2,3}}{C_{2,2}} \\ A_{to}A_{3,2} &= \frac{C_{3,3}}{C_{3,2}}\end{aligned}$$

On peut donc faire la relation entre les *LDF* et les *age to age factor*,

$$\widehat{LDF}_2 = \sum_{i=1}^{3} W_i A_{to}A_{i,2}$$

où $W_i = \frac{C_{i,2}}{\sum_{k=1}^{3} C_{k,2}}$.
Remarque : $\sum_{i=1}^{3} W_i = 1$

L'analogue pour $\widehat{\sigma}_2^2$

$$\widehat{\sigma}_2^2 = \frac{1}{n-j-1}\sum_{i=1}^{n-j} C_{i,2}\left(\frac{C_{i,3}}{C_{i,2}} - \widehat{LDF}_2\right)^2$$

Où $\frac{1}{n-j-1}$ est la correction pour être sans biais et $\frac{C_{i,3}}{C_{i,2}}$ est le *age to age ratio* ($A_{to}A_{i,3}$).

Cas général

$$\widehat{\sigma}_j^2 = \frac{1}{n-j-1}\sum_{i=1}^{n-j} C_{i,j}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{LDF}_j\right)^2 \tag{6.4}$$

Preuve

I)

$$
\begin{aligned}
E\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{LDF}_j\right)^2 \middle| D_j\right] = {} & E\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - LDF_j\right)^2 \middle| D_j\right] \\
& - 2E\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - LDF_j\right) \times (\widehat{LDF}_j - LDF_j) \middle| D_j\right] \\
& + E\left[\left(\widehat{LDF}_j - LDF_j\right)^2 \middle| D_j\right]
\end{aligned}
$$

II)

$$
\begin{aligned}
E\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - LDF_j\right)^2 \middle| D_j\right] &= \operatorname{Var}\left(\frac{C_{i,j+1}}{C_{i,j}} \middle| D_j\right) = \frac{\sigma_j^2}{C_{i,j}} \\
E\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - LDF_j\right) \times (\widehat{LDF}_j - LDF_j) \middle| D_j\right] &= \operatorname{Cov}\left(\frac{C_{i,j_1}}{C_{i,j}}, \widehat{LDF}_j \middle| D_j\right) \\
&= \frac{\sigma_j^2}{\sum_{i=1}^{n-j} C_{i,j}}
\end{aligned}
$$

III)

$$E[\sigma_j^2] = \sigma_j^2 \text{ en substituant II et I}$$

Par conséquent, l'erreur quadratique moyenne de la réserve pour l'AA i est

$$MSE(\widehat{R}_i) = E\big[(\widehat{R}_i - R_i)^2 | D\big]$$

et est estimée par

$$MSE(\widehat{R}_i) = \widehat{C}_{i,n}^2 \sum_{k=n-i+1}^{n-1} \frac{\widehat{\sigma}_k^2}{(LDF_k)^2} \left(\frac{1}{\widehat{C}_{i,k}} + \frac{1}{\sum_{j=1}^{n-k} C_{j,k}}\right) \qquad (6.5)$$

Remarque

Lorsque k = n - 1, soit le dernier terme de la somme, on a que

$$\widehat{\sigma}_{n-1}^2 = \frac{1}{n-(n-1)-1}\left(\frac{C_{i,n}}{C_{i,n-1}} - \widehat{LDF}_{n-1}\right)^2$$

On note que $\frac{1}{n-(n-1)-1} = \frac{1}{0}$. Cette situation étant problématique, deux solutions ont été proposées :

1) On pose $\widehat{\sigma}_{n-1}^2 = \widehat{\sigma}_{n-2}^2$
2) Il est d'usage, en pratique, d'utiliser la forme suivante si $\widehat{\sigma}_j^2$ n'est pas complètement plat pour j grand,

$$\widehat{\sigma}_{n-1}^2 = \min\left(\frac{\widehat{\sigma}_{n-2}^4}{\widehat{\sigma}_{n-3}^2}, \min\left(\widehat{\sigma}_{n-3}^2, \widehat{\sigma}_{n-2}^2\right)\right)$$

I.C. pour $\widehat{R}_i$

Deux options fournies par Mack,

1) On suppose que

$$R_i \sim N\left\{MU_i = E[\widehat{R}_i]; SIGMA_i^2 = MSE(\widehat{R}_i)\right\}$$
$$\Rightarrow IC^N(\widehat{R}_i) = \left[\widehat{R}_i - Z_{\frac{\alpha}{2}} \times \sqrt{MSE(\widehat{R}_i)}; \widehat{R}_i + Z_{\frac{\alpha}{2}} \times \sqrt{MSE(\widehat{R}_i)}\right]$$

2) On suppose que

$$R_i \sim LN\left\{MU_i = \ln(\widehat{R}_i) - \frac{SIGMA_i^2}{2}; SIGMA_i^2 = \ln\left(1 + \left(\frac{\sqrt{MSE(\widehat{R}_i)}}{\widehat{R_i}}\right)^2\right)\right\}$$
$$\Rightarrow IC^{LN}(\widehat{R}_i) = \left[e^{MU_i - Z_{\frac{\alpha}{2}} \times \sqrt{SIGMA_i^2}}; e^{MU_i + Z_{\frac{\alpha}{2}} \times \sqrt{SIGMA_i^2}}\right]$$

MSE et I.C. pour $\widehat{R} = \sum_{i=1}^n \widehat{R}_i$

La volatilité de la réserve totale est donnée par

$$\begin{aligned} MSE(\widehat{R}) &= E\left[(\widehat{R} - R)^2 | D\right] \\ &= E\left[\left(\sum_{i=2}^n \widehat{R}_i - \sum_{i=2}^n R_i)^2\right) \Big| D\right] \end{aligned}$$

et est estimée par

$$MSE(\widehat{R}) = \sum_{i=2}^{n} \left(\widehat{MSE}(\widehat{R}_i) + \widehat{C}_{i,n} \times \Big(\sum_{j=i+1}^{n} \widehat{C}_{j,n} \Big) \times \sum_{k=n-i+1}^{n-1} \frac{\frac{2\widehat{\sigma}_k^2}{\widehat{LDF}_k^2}}{\sum_{j=1}^{n-k} C_{j,k}} \right)$$

Deux options fournies pour I.C.

1)

$$R \sim N\Big\{ MU = \widehat{R}; SIGMA^2 = MSE(\widehat{R}) \Big\}$$

$$\Rightarrow IC^N(\widehat{R}) = \left[\widehat{R} - Z_{\frac{\alpha}{2}} \times \sqrt{MSE(\widehat{R})}; \widehat{R} + Z_{\frac{\alpha}{2}} \times \sqrt{MSE(\widehat{R})} \right]$$

2)

$$R \sim LN\left\{ MU = \ln(\widehat{R}) + \frac{SIGMA^2}{2}; SIGMA^2 = \ln\left(1 + \left(\frac{\sqrt{MSE(\widehat{R}_i)}}{\widehat{R_i}} \right)^2 \right) \right\}$$

$$\Rightarrow IC^{LN}(\widehat{R}) = \left[e^{MU - Z_{\frac{\alpha}{2}} \times \sqrt{SIGMA^2}}; e^{MU + Z_{\frac{\alpha}{2}} \times \sqrt{SIGMA^2}} \right]$$

6.6 Méthode basée sur les GLM

Bien que la méthode de Mack introduise la notation de variabilité dans les réserves IARD. Les GLM permettent de faire plus :

1) Ils ne se limitent pas à la distribution Normale ou Log-Normale, mais admettent toutes les distributions de la famille exponentielle (Normale, Log-Normale, Binomiale, Poisson, exponentielle, ...)

$$f_Y(y) = \exp\left\{ \frac{y \times \theta - b(\theta)}{\phi \times w_i} + c(y, \phi, w_i) \right\}$$

2) Ils permettent d'incorporer des variables explicatives plus élaborées que l'âge j dans le modèle pour prédire le développement. Par exemple :
   - Changement de la vice-présidence aux réclamations
   - Changement de loi ou de gouvernement
   - Événement catastrophique
3) Ils permettent de faire des réserves granulaires au lieu d'agréger les données par AA et âge dans un triangle.

Remarques avec les GLM

1) Avec les GLM, on travaille avec le triangle de sinistre incrémental $y_{i,j}$ au lieu de cumulatif ($C_{i,j}$).

$$Y_{i,j} = C_{i,j} - C_{i,j-1}$$

   Qui correspond au montant de variation des sinistres de l'AA i entre les âges (j-1) et j. On calcule donc un triangle des accroissements/incréments où Y_i sont indépendants.
2) En pratique, on choisit une loi dans la famille exponentielle selon la variance observée dans les $Y_{i,j}$.

Loi	Fonction de variance	Utilisation en réserves
Normale	$\sigma^2 \neq f(u)$	Très rare, car la variance diminue avec l'âge j
Poisson	μ	Souvent
Gamma	var $= f(\mu^2)$	Moins souvent
Inverse Gaussienne	var $= f(\mu^3)$	Très rare

   Remarque

   Pour la Poisson, si la paramétrisation est adéquate, le GLM Poisson coïncide avec C-L.
3) Le modèle classique
   - $E[Y_{i,j}] = \mu_{i,j} = e^{\mu+\alpha_i+\beta_j}$
   - Contrainte d'identification, soit poser $\alpha_1 = \beta_1 = 0$. La *base* est pour l'AA = 1 et l'âge = 1.

$$E[Y_{i,j}] = e^{\mu\alpha_i} + e^{\beta_j}$$

   Où le premier terme est effet de l'AA i et le deuxième terme est l'effet de l'âge j.
4) Le modèle plus élaboré

$$X_{i,j} = \begin{cases} 1 & , \text{ si condition particulière} \\ 0 & , \text{ sinon} \end{cases}$$

Exemple de condition particulière : Si l'ancien vice-président aux sinistres était en fonction dans l'année de calendrier (i + j - 1). On obtient ainsi un modèle plus élaboré pour la condition particulière :

$$\begin{aligned} E[Y_{i,j}] &= \mu_{i,j} \\ &= e^{\mu+\alpha_i+\beta_j+\gamma\times X_{i,j}} \end{aligned}$$

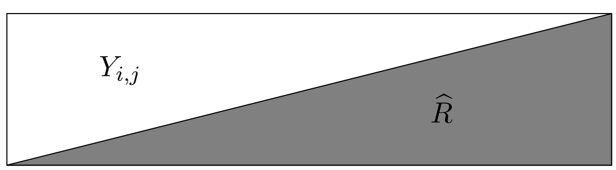
Où $\gamma \times X_{i,j}$ est l'effet de la situation particulière. Ainsi dans le futur la valeur de cette expression est de zéro.

5) Les paramètres tels que $\Theta = (\mu, \alpha_2, ..., \alpha_n, \beta_1, ..., \beta_n)$ sont estimés via la méthode du maximum de vraisemblance, soit en maximisant

$$\phi(\Theta) = \sum_{i,j \in \triangleleft} \ln\big(f_{Y_{i,j}}(y_{i,j})\big)$$

6) Une fois les paramètres estimés, on calcule les réserves $\widehat{R}_1, ..., \widehat{R}_n$, ou encore $\widehat{R}$ en sommant les $\widehat{Y}_{i,j}$ du *triangle inférieur.*

Exemple



$\leftarrow \widehat{R}_1 = 0$

$\leftarrow \widehat{R}_2 = \widehat{Y}_{2,n}$

$\leftarrow \widehat{R}_3 = \widehat{Y}_{3,n-1} + \widehat{Y}_{3,n}$

$\leftarrow \widehat{R}_n = \widehat{Y}_{n,2} + ... + \widehat{Y}_{n,n}$

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

$$\begin{aligned} \widehat{Y}_{2,n} &= e^{\widehat{\mu}+\widehat{\alpha}_2+\widehat{\beta}_n} \\ \widehat{Y}_{3,n-1} &= e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_{n-1}} \\ \widehat{Y}_{3,n} &= e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_n} \end{aligned}$$

$$\begin{aligned} \widehat{Y}_{2,n} &= E\left[e^{\widehat{\mu}+\widehat{\alpha}_2+\widehat{\beta}_n}\right] \\ &\neq e^{E[\widehat{\mu}]+E[\widehat{\alpha}_2]+E[\widehat{\beta}_n]} \end{aligned}$$

Car $E[g(X_1, ..., X_n)]$ n'égale pas toujours égale à $g(E[X_1], ..., E[X_n])$.

Théorème

Soit $Y = g(X_1, ..., X_n)$; une statistique fonction de la *v.a* $X_1, ..., X_n$. En développant la fonction $g(\bullet)$ en série de Taylor multidimensionnelle, et en tronquant après l'ordre 2, on trouve que

$$E[y] \approx g(E[X_1], ..., E[X_n]) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 g(E[X_1], ..., E[X_n])}{\partial X_i \partial X_j} \times \mathrm{Cov}(X_i, X_j) \tag{6.6}$$

$$\mathrm{Var}(Y) \approx \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial}{\partial x_i} g(E[X_1], ..., E[X_n])\right)\left(\frac{\partial}{\partial x_j} g(E[X_1], ..., E[X_n])\right) \times \mathrm{Cov}(X_i, X_j) \tag{6.7}$$

Exemple

À partir des équations 6.6 et 6.7 et en posant n = 5, on développe les équations suivantes :

$$\begin{aligned}\widehat{R}_3 &= \widehat{Y}_{3,4} + \widehat{Y}_{3,5} \\ &= e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5} \\ &= g(\widehat{\mu}, \widehat{\alpha}_3, \widehat{\beta}_4, \widehat{\beta}_5)\end{aligned}$$

On développe l'expression de l'espérance de $\widehat{R}_3$

$$
\begin{aligned}
E[\widehat{R}_3] \cong\ & e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5} \\
& + \frac{1}{2}\Bigg(\Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\Big]\mathrm{Var}(\widehat{\mu}) \\
& + 2\Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\Big] \times \mathrm{Cov}(\widehat{\mu}, \widehat{\alpha}_3) \\
& + 2\Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4}\Big] \times \mathrm{Cov}(\widehat{\mu}, \widehat{\beta}_4) \\
& + 2\Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\Big] \times \mathrm{Cov}(\widehat{\mu}, \widehat{\beta}_5) \\
& + \Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\Big] \times \mathrm{Cov}(\widehat{\alpha}_3, \widehat{\alpha}_3) \\
& + 2\Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4}\Big] \times \mathrm{Cov}(\widehat{\alpha}_3, \widehat{\beta}_4) \\
& + 2\Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\Big] \times \mathrm{Cov}(\widehat{\alpha}_3, \widehat{\beta}_5) \\
& + \Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4}\Big] \times \mathrm{Var}(\widehat{\beta}_4) \\
& + 2\Big[0\Big] \times \mathrm{Cov}(\widehat{\beta}_4, \widehat{\beta}_5) \\
& + 2\Big[e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\Big] \times \mathrm{Var}(\widehat{\beta}_5)\Bigg)
\end{aligned}
$$

On développe l'expression de la variance de $\widehat{R}_3$

$$
\begin{aligned}
\operatorname{Var}(\widehat{R}_3) = & \left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right)^2 \times \operatorname{Var}(\widehat{\mu}) \\
& + 2\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right)^2 \operatorname{Cov}(\widehat{\mu}, \widehat{\alpha}_3) \\
& + 2\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right)\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4}\right) \times \operatorname{Cov}(\widehat{\mu}, \widehat{\beta}_4) \\
& + 2\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right)\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right) \times \operatorname{Cov}(\widehat{\mu}, \widehat{\beta}_5) \\
& + \left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right)^2 \times \operatorname{Var}(\widehat{\alpha}_3) \\
& + 2\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right)\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4}\right) \times \operatorname{Cov}(\widehat{\alpha}_3, \widehat{\beta}_4) \\
& + 2\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right)\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right) \times \operatorname{Cov}(\widehat{\alpha}_3, \widehat{\beta}_5) \\
& + \left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4}\right)^2 \times \operatorname{Var}(\widehat{\beta}_4) \\
& + 2\left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_4} + e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right) \times \operatorname{Cov}(\widehat{\beta}_4, \widehat{\beta}_5) \\
& + \left(e^{\widehat{\mu}+\widehat{\alpha}_3+\widehat{\beta}_5}\right)^2 \times \operatorname{Var}(\widehat{\beta}_5)
\end{aligned}
$$

Cas de la loi Poisson

Dans ce cas, on suppose que $Y_{i,j} \sim Poisson(\mu_{i,j})$ où $\mu_{i,j} = E[Y_{i,j}] = \text{Var}(Y_{i,j})$ et $\mu_{i,j} = e^{\mu+\alpha_i+\beta_j}$ avec $\alpha_1 = \beta_1 = 0$. Alors,

$$\begin{aligned}
\phi(\Theta) &= \sum_{i,j\in\triangleleft} \ln\big(f_{Y_{i,j}}(y_{i,j})\big) \\
&= \sum_{i,j\in\triangleleft} \ln\left(\frac{e^{-\mu_{i,j}}\mu_{i,j}^{y_{i,j}}}{y_{i,j}\,!}\right) \\
&= \sum_{i,j\in\triangleleft} \left(y_{i,j}\ln(\mu_{i,j}) - \mu_{i,j} - \ln(y_{i,j}\,!)\right)
\end{aligned}$$

Log-Vraisemblance proportionnelle

$$\begin{aligned}
\phi(\Theta) &\simeq \sum_{i,j\in\triangleleft} \left(y_{i,j}\ln(\mu_{i,j}) - \mu_{i,j}\right) \\
&= \sum_{i,j\in\triangleleft} \left(y_{i,j}(\mu+\alpha_i+\beta_j) - e^{\mu+\alpha_i+\beta_j}\right)
\end{aligned}$$

En pratique, il est impossible de trouver les paramètres tels que $(\widehat{\mu}, \widehat{\alpha_2}, ..., \widehat{\alpha_n}, \widehat{\beta}_1, ..., \widehat{\beta_n})$ qui maximisent $\phi(\Theta)$ de manière exacte. On utilise donc l'algorithme de Newton-Raphson.

$$\Theta = \begin{bmatrix} \mu \\ \alpha_2 \\ \vdots \\ \alpha_n \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

Avec la méthode de Newton-Raphson : $\Theta^{(i+1)} = \Theta^{(i)} + I(\Theta^{(i)})^{-1}S(\Theta^{(i)})$
Rappel du cours de modèle linéaire : Dans le cas Poisson, on a trouvé que :

$$S(\Theta) = \mathbb{X}^T\mathbb{W}\,;\ \text{Où } \mathbb{W} = \mathbb{Y} - \widehat{\mathbb{Y}}$$

De plus, on a

$$I(\Theta) = \mathbb{X}^{\intercal}\mathbb{H}\mathbb{X}$$

$$\mathbb{H} = \begin{bmatrix} e^{X_1\beta} & \cdots & 0 & \cdots & 0 \\ 0 & e^{X_2\beta} & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & e^{X_{n-1}\beta} & 0 \\ 0 & \cdots & 0 & \cdots & e^{X_n\beta} \end{bmatrix}$$

Remarque

Dans les formules précédentes de $S(\Theta)$ et $I(\Theta)$; on utilise la formulation sur les GLM acquise dans le cours de modèle linéaire (ACT-2003), soit

$$Y_{i,j} \sim \text{Poisson}(e^{\mu+\alpha_i+\beta_i}) \sim \text{Poisson}(e^{\mathbb{X}\underset{\sim}{\beta}})$$

Où,

$$\underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

Exemple

Si n = 5, on obtient

$$\begin{aligned} & X_{i,0} = 1 & & \\ & X_{i,1} = 1\{i = 2\} & & X_{i,5} = 1\{j = 2\} \\ & X_{i,2} = 1\{i = 3\} & & X_{i,6} = 1\{j = 3\} \\ & X_{i,3} = 1\{i = 4\} & & X_{i,7} = 1\{j = 4\} \\ & X_{i,4} = 1\{i = 5\} & & X_{i,8} = 1\{j = 5\} \end{aligned}$$

On exprime les données dans la table suivante :

i	j	$Y_{i,j}$	$X_{i,0}$	$X_{i,1}$	$X_{i,2}$	$\cdots$	$X_{i,8}$
1	1	$\vdots$	1	1	0	$\cdots$	0
1	2	$\vdots$	1	1	0	$\cdots$	0
1	3	$\vdots$	1	1	0	$\cdots$	0
1	4	$\vdots$	1	1	0	$\cdots$	0
1	5	$\vdots$	1	1	0	$\cdots$	1
2	1	$\vdots$	1	0	1	$\cdots$	0
2	2	$\vdots$	1	0	1	$\cdots$	0
2	3	$\vdots$	1	0	1	$\cdots$	0
2	4	$\vdots$	1	0	1	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	1	0	0	$\vdots$	$\vdots$
4	1	$\vdots$	1	0	0	$\cdots$	$\vdots$
4	2	$\vdots$	1	0	0	$\cdots$	$\vdots$
5	1	$\vdots$	1	0	0	$\cdots$	$\vdots$

Suite de l'exemple Excel avec n = 5

$$\begin{aligned}\widehat{R} = 0 & \\ & + \widehat{Y}_{2,5} \\ & + \widehat{Y}_{3,4} + \widehat{Y}_{3,5} \\ & + \widehat{Y}_{4,3} + \widehat{Y}_{4,4} + \widehat{Y}_{4,5} \\ & + \widehat{Y}_{5,2} + \widehat{Y}_{5,3} + \widehat{Y}_{5,4} + \widehat{Y}_{5,5}\end{aligned}$$

On définit $\mathbb{X}^*$ comme suit

$$\mathbb{X}^* = \begin{bmatrix} X_0^* \\ X_1^* \\ X_2^* \\ X_3^* \\ X_4^* \\ X_5^* \\ X_6^* \\ X_7^* \\ X_8^* \\ X_9^* \\ X_{10}^* \end{bmatrix} = \begin{array}{c} \begin{array}{ccccccccc} \mu & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \beta_2 & \beta_3 & \beta_4 & \beta_5 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$\begin{aligned} \widehat{R} &= e^{X_1^* \widehat{\beta}} + e^{X_2^* \widehat{\beta}} + ... + e^{X_{10}^* \widehat{\beta}} \\ &= g(\beta_0, \beta_1, ..., \beta_{10}) \end{aligned}$$

On obtient ainsi

$$E[\widehat{R}] = g(\widehat{\underset{\sim}{\beta}}) + \frac{1}{2} \sum_{i=0}^{8} \sum_{j=0}^{8} \left(\frac{\partial^2 \widehat{R}}{\partial \underset{\sim}{\beta}^2} \right)_{i+1,j+1} \times I^{-1}(\widehat{\underset{\sim}{\beta}})_{i+1,j+1} \tag{6.8}$$

Où

$$\begin{aligned} \frac{\partial^2 \widehat{R}}{\partial \widehat{\underset{\sim}{\beta}}^2} &= [X_1^*]_{i+1} [X_1^*]_{j+1} \times e^{X_1 \widehat{\underset{\sim}{\beta}}} + \cdots + [X_{10}^*]_{i+1} [X_{10}^*]_{j+1} \times e^{X_{10} \widehat{\underset{\sim}{\beta}}} \\ &= (\mathbb{X}^*)^\intercal \mathbb{W} \mathbb{X}^* \end{aligned}$$

Où

$$\mathbb{W} = \begin{bmatrix} e^{X_1 \widehat{\underset{\sim}{\beta}}} & 0 & \cdots & 0 \\ 0 & e^{X_2 \widehat{\underset{\sim}{\beta}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & e^{X_9 \widehat{\underset{\sim}{\beta}}} & 0 \\ 0 & \cdots & 0 & e^{X_{10} \widehat{\underset{\sim}{\beta}}} \end{bmatrix}$$

Pour la variance de $\widehat{R}$, on obtient

$$\operatorname{Var}(\widehat{R}) = \sum_{i=0}^{8}\sum_{j=0}^{8}\left(\frac{\partial \widehat{R}}{\partial \underset{\sim}{\beta}}\right)_{i+1} \times I^{-1}(\widehat{\underset{\sim}{\beta}})_{i+1,j+1}\left(\frac{\partial \widehat{R}}{\partial \underset{\sim}{\beta}}\right)_{j+1} \tag{6.9}$$

Qu'on peut réécrire en vecteur de la façon suivante,

$$\operatorname{Var}(\widehat{R}) = \left(\frac{\partial \widehat{\mathbb{R}}}{\partial \underset{\sim}{\beta}}\right)^{\intercal} I^{-1}(\widehat{\underset{\sim}{\beta}})\left(\frac{\partial \widehat{\mathbb{R}}}{\partial \underset{\sim}{\beta}}\right) \tag{6.10}$$

Note :

$$\begin{aligned}\left(\frac{\partial \widehat{\mathbb{R}}}{\partial \underset{\sim}{\beta}}\right) &= [X_1^*] \times e^{X_1\widehat{\underset{\sim}{\beta}}} + \cdots + [X_{10}^*] \times e^{X_{10}\widehat{\underset{\sim}{\beta}}} \\ &= (\mathbb{X}^*)^{\intercal}\mathbb{M}\end{aligned}$$

Où

$$\mathbb{M} = \begin{bmatrix} e^{X_1^*\underset{\sim}{\beta}} \\ e^{X_2^*\underset{\sim}{\beta}} \\ e^{X_3^*\underset{\sim}{\beta}} \\ e^{X_4^*\underset{\sim}{\beta}} \\ \vdots \\ e^{X_{10}^*\underset{\sim}{\beta}} \end{bmatrix}$$

Ce qui permet de redéfinir l'équation 6.10 de la façon suivante

$$\begin{aligned}\operatorname{Var}(\widehat{R}) &= \big((\mathbb{X}^*)^{\intercal}\mathbb{M}\big)^{\intercal} \times I^{-1}(\widehat{\underset{\sim}{\beta}}) \times \big((\mathbb{X}^*)^{\intercal}\mathbb{M}\big) \\ &= \big((\mathbb{X}^*)^{\intercal}\widehat{\mathbb{Y}}\big)^{\intercal} \times I^{-1}(\widehat{\underset{\sim}{\beta}}) \times \big((\mathbb{X}^*)^{\intercal}\widehat{\mathbb{Y}}\big)\end{aligned}$$

Où $\widehat{\mathbb{Y}}$ correspond à la matrice des valeurs des incrémentations estimées.

Tests d'hypothèses avec les GLM

H_0 : Un *sous-modèle* de M_1 (noté M_0) est acceptable.
H_1 : Il est nécessaire d'utiliser le modèle *complet* (noté M_1)
Pour tester H_0 contre H_1, on obtient la statistique

$$\chi^2_{obs} = 2 \times \Big(\ell(H_1) - \ell(H_0)\Big) \tag{6.11}$$

Et on rejette H_0 au niveau de confiance $100 \times (1 - \alpha)\%$ si [1]

$$\chi^2_{obs} \geq \chi^2_{\alpha}\Big(\text{Nombre de paramètres}(\Xi) \text{ de } (M_1) \text{ - Nombre de paramètres}(\Xi) \text{ de } (M_0)\Big) \tag{6.12}$$

Exemple

Tableau des données :

i / j	1	2	3	4
1	100	48	23	10
2	115	55	27	
3	115	56		
4	115			

A)

$$\begin{aligned} E[Y_{i,j}] &= e^{\mu+\alpha_i+\beta_j} \text{ ; } \alpha_1 = \beta_1 = 0 \\ \widehat{\mu} &= 4.60 \\ \Xi &= 7 \\ \ell(\widehat{\beta}) &= 2249.283 \end{aligned}$$

i / j	1	2
1	ϕ	ϕ
2	0.14	-0.73
3	0.15	-1.46
4	0.14	-2.30

1. Où Ξ corresponds au nombre de paramètres.

B)

$$
\begin{aligned}
E[Y_{i,j}] &= e^{\mu+\alpha_i+\beta\times(j-1)}\,;\ \alpha_1 = 0\\
\widehat{\mu} &= 4.60\\
\widehat{\beta} &= -0.74\\
\Xi &= 5\\
\ell(\widehat{\beta}) &= 2249.283\\
\widehat{\alpha}_2 &= 0.15\\
\widehat{\alpha}_3 &= 0.15\\
\widehat{\alpha}_4 &= 0.14
\end{aligned}
$$

Test

À l'aide de l'équation 6.11 on effectue le test d'hypothèse suivant
H_0 : Linear developpment $\beta_j = \beta \times (j-1)$
H_1 : No linear development

$$
\begin{aligned}
\chi^2_{obs} &= 2 \times \Big(2249.283 - 2248.234\Big)\\
&= 0.098087\\
\chi^2_{0.05}(7-5) &= 5.99
\end{aligned}
$$

Pour rejeter H_0, on doit satisfaire l'équation 6.12,

$$
\begin{aligned}
\chi^2_{obs} &\geq \chi^2_{\alpha}\Big(\Xi(M_1) - \Xi(M_0)\Big)\\
0.098087 &\overset{?}{\geq} 5.99
\end{aligned}
$$

On ne peut pas rejeter H_0.

C)

$$
\begin{aligned}
E[Y_{i,j}] &= e^{\mu+\alpha\times(i-1)+\beta\times(j-1)} \\
\widehat{\mu} &= 4.64 \\
\widehat{\alpha} &= 0.05 \\
\widehat{\beta} &= -0.73 \\
\Xi &= 3 \\
\ell(\widehat{\beta},\widehat{\alpha},\widehat{\mu}) &= 2248.644
\end{aligned}
$$

Test

H_0 : Linear trend and developpment ($e^{\mu+\alpha(i-1)+\beta\times(j-1)}$)
H_1 : H_0 is false ($e^{\mu+\alpha_i+\beta_j}$)

$$
\begin{aligned}
\chi^2_{obs} &= 2\times\Big(2249.283 - 2248.664\Big) \\
&= 1.024 \\
\chi^2_{0.05}(7-3) &= 9.488
\end{aligned}
$$

Pour rejeter H_0, on doit satisfaire l'équation 6.12,

$$
\begin{aligned}
\chi^2_{obs} &\geq \chi^2_{\alpha}\Big(\Xi(M_1)-\Xi(M_0)\Big) \\
1.024 &\overset{?}{\geq} 9.488
\end{aligned}
$$

On ne peut pas rejeter H_0.

Alternative au test d'hypothèse

H_0 : Linear trend and developpment ($e^{\mu+\alpha(i-1)+\beta\times(j-1)}$)
H_1 : H_0 is false ($e^{\mu+\alpha_i+\beta\times(j-i)}$)

D) On sait qu'il y a eu une nouvelle loi sur le règlement des sinistres à partir de l'année de calendrier 3. $I_{i,j}$ = indicateur de nouvelle loi à i, j.

i / j	1	2	3	4
1	0	0	1	1
2	0	1	1	
3	1	1		
4	1			

$$E[Y_{i,j}] = e^{\mu+\alpha\times(i-1)+\beta\times(j-1)} + \alpha I_{i,j}$$

Pour tester si la nouvelle loi a un impact sur le développement, on utilise le H_0 suivant

$$\begin{aligned}
H_0 &= e^{\mu+\alpha\times(i-1)+\beta\times(j-1)} \\
\widehat{\mu} &= 4.64 \\
\widehat{\alpha} &= 0.05 \\
\widehat{\beta} &= -0.73 \\
\Xi &= 3 \\
\ell(\widehat{\beta}, \widehat{\alpha}, \widehat{\mu}, \widehat{\alpha}) &= 2248.644
\end{aligned}$$

et le H_1 suivant

$$\begin{aligned}
H_1 &= e^{\mu+\alpha\times(i-1)+\beta\times(j-1)+\alpha I_{i,j}} \\
\widehat{\mu} &= 4.64 \\
\widehat{\alpha} &= 0.05 \\
\widehat{\beta} &= -0.73 \\
\widehat{\alpha} &= -001644 \\
\Xi &= 4 \\
\ell(\widehat{\beta}, \widehat{\alpha}, \widehat{\mu}, \widehat{\alpha}) &= 2248.649
\end{aligned}$$

Test

En utilisant les hypothèses H_0 et H_1 précédentes,

$$\begin{aligned}\chi^2_{obs} &= 2 \times \Big(2248.649 - 2248.644\Big)\\ &= 0.009305\\ \chi^2_{0.05}(4-3) &= 3.84\end{aligned}$$

Pour rejeter H_0, on doit satisfaire l'équation 6.12,

$$\begin{aligned}\chi^2_{obs} &\geq \chi^2_{\alpha}\Big(\Xi(M_1) - \Xi(M_0)\Big)\\ 0.009305 &\overset{?}{\geq} 3.84\end{aligned}$$

On ne peut pas rejeter H_0.

6.7 Méthode de réserve IARD basée sur la théorie de la crédibilité

Rappel sur C-L

1. Le modèle est bon pour les vieilles AA. Il y a beaucoup de données disponibles, autrement dit, d'expérience individuelle est *crédible*.
2. Par contre, il est moins bon pour les AA récentes. Il y a moins de données, autrement dit, l'expérience individuelle est moins volumineuse et donc moins crédible

L'idée : Accorder plus de crédibilité à C-L si AA vieille, donc on possède beaucoup de données, et de donner plus de crédibilité à B-F si AA récente.

Rappel

$$\begin{aligned}
\widehat{R}_i^{C-L} &= \widehat{C}_{i,n}^{C-L} - C_{i,n-i+1} \\
&= C_{i,n-i+1} \times \Big(\prod_{j=n-i-1}^{n-1} LDF_j \Big) - C_{i,n-i+1} \\
&= C_{i,n-i+1} \times \Big(\prod_{j=n-i-1}^{n-1} LDF_j - 1 \Big)
\end{aligned}$$

$$\begin{aligned}
\widehat{R}_i^{B-F} &= \widehat{C}_{i,n}^{B-F} - C_{i,n-i+1} \\
&= PA_i \times LR_i - C_{i,n-i+1} \\
&= U_i - \frac{U_i}{\prod_{j=n-i-1}^{n-1} LDF_j} \\
&= U_i \left(1 - \frac{1}{\prod_{j=n-i-1}^{n-1} LDF_j} \right) ; \text{ posons } \beta_{n-i+1} = \frac{1}{\prod_{j=n-i-1}^{n-1} LDF_j} \\
&= U_i \left(1 - \beta_{n-i+1} \right)
\end{aligned}$$

L'idée est de combiner $\widehat{R}_i^{C-L}$ et $\widehat{R}_i^{B-F}$.

$$\begin{aligned}
\widehat{R}_i^{Cred} &= C_i \widehat{R}_i^{C-L} + (1 - C_i) \widehat{R}_i^{B-F} \\
&= C_i \times C_{i,n-i+1} \left(\prod_{j=n-i-1}^{n-1} LDF_j - 1 \right) + (1 - C_i) U_i (1 - \beta_{n-i+1}) \\
&= C_i \times C_{i,n-i+1} \left(\prod_{j=n-i-1}^{n-1} LDF_j - 1 \right) + (1 - C_i) U_i (1 - \beta_{n-i+1}) \\
&= C_i \times \widehat{C}_{i,n}^{C-L} - C_i C_{i,n-i+1} + (1 - C_i) U_i - (1 - C_i) U_i (1 - \beta_{n-i+1}) \\
&= [C_i \times \widehat{C}_{i,n}^{C-L} + (1 - C_i) U_i] - \beta_{n-i+1} \left[\frac{C_i C_{i,n-i+1}}{\beta_{n-i+1}} + (1 - C_i) U_i \right]
\end{aligned}$$

On développe β_{n-i+1} afin de réduire l'expression précédante,

$$\begin{aligned}
\beta_{n-i+1} &= \frac{1}{\prod_{j=n-i-1}^{n-1} LDF_j} \\
\frac{1}{\beta_{n-i+1}} &= \prod_{j=n-i-1}^{n-1} LDF_j \\
\frac{C_{i,n-i+1}}{\beta_{n-i+1}} &= C_{i,n-i+1} \prod_{j=n-i-1}^{n-1} LDF_j \\
\frac{C_i C_{i,n-i+1}}{\beta_{n-i+1}} &= C_i C_{i,n-i+1} \prod_{j=n-i-1}^{n-1} LDF_j \\
\frac{C_i C_{i,n-i+1}}{\beta_{n-i+1}} &= C_i \widehat{C}_{i,n}^{C-L}
\end{aligned}$$

À partir de cette *identité* et de l'expression suivante, on obtient le résultat suivant

$$\begin{aligned}
\widehat{R}_i^{Cred} &= [C_i \times \widehat{C}_{i,n}^{C-L} + (1-C_i)U_i] - \beta_{n-i+1}\left[C_i C_{i,n}^{C-L} + (1-C_i)U_i\right] \\
&= [C_i \times \widehat{C}_{i,n}^{C-L} + (1-C_i)U_i](1-\beta_{n-i+1})
\end{aligned}$$

Que l'on peut réécrire de la façon suivante

$$\widehat{R}_i^{Cred} = \left[C_{i,n-i+1} + \left(1 - \frac{1}{\prod_{j=n-i+1}^{n-1} \widehat{LDF}_j}\right) PA_i \times LR_i\right]\left(1 - \frac{1}{\prod_{j=n-i+1}^{n-1} \widehat{LDF}_j}\right) \tag{6.13}$$

Au début des années 1980, Benktander et Hovinen ont trouvé le raisonnement suivant :

i) $U_i(1-\beta_{n-i+1}) = R_i^{B-F}$

ii) $\widehat{R}_i^{B-F} = U_i - C_{i,n-i+1} \Rightarrow \widehat{R}_i^{B-F} = U_i^{B-H} - \widehat{C}_{i,n}\beta_{n-i+1}$

À partir de i) $\rightarrow$ ii), on obtient le développement suivant

$$\begin{aligned}
\widehat{C}_{i,n}^{B-F}(1-\beta_{n-i+1}) &= U_i^{B-H} - \beta_{n-i+1}\widehat{C}_{i,n}^{C-L} \\
U_i^{B-H} &= \beta_{n-i+1}\widehat{C}_{i,n}^{C-L} + (1-\beta_{n-i+1})\widehat{C}_{i,n}^{B-F} \\
\widehat{C}_{i,n}^{B-H} &= \beta_{n-i+1}\widehat{C}_{i,n}^{C-L} + (1-\beta_{n-i+1})\widehat{C}_{i,n}^{B-F}
\end{aligned}$$

Où β_{n-i+1} correspond à C_i et $(1-\beta_{n-i+1})$ correspond à $1-C_i$.

En conclusion, pour que l'hypothèse de Benktander - Hoviden soit respectée, autrement dit, à la coïncidence des diagonales du triangle, il faut que

$$C_i = \beta_{n-i+1} = \frac{1}{\prod_{j=n-i+1}^{n-1}} = \text{ La proportion de ce qui a été fait} \tag{6.14}$$

Remarque

En pratique, cette méthode est appelée Bornhuetter - Ferguson itérée.
En répétant m fois ce raisonnement d'itération, on trouve :

$$\widehat{C}_{i,n}^{B-H} = \begin{cases} C_{i,n-i+1} + (1-\beta_{n-i+1})\widehat{C}_{i,n}^{B-H(m-1)} & ,\text{si } m \geq 1 \\ C_{i,n-i+1} + (1-\beta_{n-i+1})U_i & ,\text{si } m = 0 \end{cases}$$

Où $U_i = PA_i \times LR_i$

$$\widehat{R}_i^{B-H(m)} = \widehat{C}_{i,n}^{B-H(m)} - C_{i,n-i+1}$$

6.8 Actualisation des réserves IARD

Il faut se rappeler qu'on actualise selon l'année de calendrier (Calendar Year (CY))

i	0	12	24	36	48	72	84	0
2005								
2006								
2007								
2008								2015
2009							2015	2016
2010						2015	2016	2017
2011					2015	2016	2017	2018
2012				2015	2016	2017	2018	2019
2013			2015	2016	2017	2018	2019	2020
2014		2015	2016	2017	2018	2019	2020	2021

Étapes

1) Développer le triangle avec l'une des méthodes vues en classe : C-L, B-F, Mack, GLM ou B-H.

2) Calculer les sinistres incrémentaux $Y_{i,j}$ dans la partie inférieure du triangle
3) Sommer les $Y_{i,j}$ de l'étape 2 par année de calendrier (CY)
4) Actualiser par CY

CY	$\sum Y_i$	Duration	Annual yield curve rate	Actualized figure
2015	CF_{15}	0.5	1 %	$CF_{15}(1+0.01)^{-0.5}$
2016	CF_{16}	1.5	2 %	$CF_{16}(1+0.02)^{-1.5}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2021	CF_{21}	6.5	5 %	$CF_{21}(1+0.05)^{-6.5}$

$\widehat{R} = \sum$ Actualized figure

Notes : Pour simplifier les calculs, on suppose une distribution uniforme [0, 1] des montants de la réserve et on utilise le point milieu pour l'actualisation.

Annexe A

Résumé examen intra

A.1 La théorie de la crédibilité

Il est possible d'utiliser des observations afin de prédire les réalisations futures d'une variable aléatoire. Ce concept est particulièrement utile afin de traiter de l'hétérogénéité résiduelle intrinsèque à une cellule de tarification. Cela est d'autant plus vrai si l'assureur utilise des cellules de tarification comportant très peu de critères, car l'hétérogénéité résiduelle sera nettement plus importante (CNESST par exemple)
Par contre, lorsque ces observations ne sont pas assez nombreuses, il ne serait pas juste de ne se fier qu'à elles, car elles ne représentent pas nécessairement bien la réalité. C'est ici qu'intervient la théorie de la crédibilité

A.1.1 La crédibilité de stabilité

La crédibilité de stabilité (ou américaine ou limited fluctuations) consiste à déterminer à partir de quel nombre d'observations il est raisonnable de juger l'estimation empirique crédible.
On dit alors que S est statistiquement stable ou crédible à l'ordre (k, p) si

$$\boxed{P((1-k)E[S] \leq S \leq (1+k)E[S]) \geq p}$$

Dans le cas où nous nous intéressions à une suite d'observations tirées d'une loi Bernouilly, il nous serait possible d'estimer la valeur du paramètre p en calculant l'espérance empirique de la probabilité de succès. Dans ce cas précis, on obtient que le nombre magique d'observations à partir duquel on peut juger crédible notre estimation de p est 1082,41, nombre qui revient dans plusieurs contextes en théorie de la crédibilité.

Voir exemple classique dans les notes (Cas d'assurance sous le modèle fréquence sévérité).
Un raccourci important qui revient assez souvent est le cas où $N \sim$ Poisson. Dans ce cas, on a l'identité suivante, où X représente la sévérité des réclamations individuelles :

$$\boxed{\begin{array}{c} \lambda \geq \left(\dfrac{\Phi^{-1}\left(\frac{p+1}{2}\right)}{k} \right)^2 \left(1 + CV(X)^2\right) \\ CV(X) = \dfrac{\sqrt{Var(X)}}{E(X)} \end{array}}$$

Dans le contexte où $N \sim$ Poisson, si le montant de la sévérité est non aléatoire ($CV(X) = 0$), on retrouve encore une fois

$$\lambda \geq 1082,41$$

A.1.2 La crédibilité partielle

Lorsque le nombre d'observations n'est pas suffisant pour juger l'expérience aléatoire crédible, on utilise la crédibilité partielle afin d'obtenir une bonne estimation de la prime crédibilisée en faisant une pondération de l'expérience aléatoire et de la valeur théorique de l'espérance. Whitney propose donc d'écrire la statistique S crédibiliser comme

$$\pi = Z\overline{S} + (1 - Z)M$$

Le modèle principal pour l'expression de Z

1. $$Z = \min\left(\sqrt{\frac{\text{taille}}{\text{taille de crédibilité complète}}}, 1\right)$$

   Aucune théorie mathématique ne supporte cette hypothèse.

   $$Z = \min\left(\sqrt{\frac{n}{1083}}, 1\right)$$

2. $$Z = \frac{n}{n + K}$$

   oùK est une constante arbitraire.

A.2 La crédibilité Bayesienne

A.2.1 Idée de base

L'idée de la crédibilité Bayesienne est que les expériences que nous avons pu observer de la variable aléatoire nous donnent une bonne approximation de l'espérance des coûts à venir, mais si et seulement si le nombre d'observations est suffisant pour que la statistique soit jugée crédible. On a donc que, dans plusieurs cas, selon le MLE,

$$\widehat{\Theta} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Il faut également prendre en compte "l'avis d'expert" dans notre modèle, soit les hypothèses initiales quant aux paramètres de la distribution observée. On note donc la "distribution a priori"

$$(X|\Theta) \sim Bern(\Theta)$$

$$\Theta \sim Beta(\alpha, \beta)$$

par exemple et la distribution a posteriori, soit la distribution ajustée aux données observées

$$(\Theta|X_1...X_n) \sim Beta\left(\alpha + \sum_{i=1}^{n} X_i,\ \beta - \sum_{i=1}^{n} X_i + n\right)$$

et

$$E[\Theta|X_1...X_n] = \frac{\alpha + \sum_{i=1}^{n} X_i}{\alpha + \beta + n}$$

Dans certains cas, il est possible de réécrire $E[X|X_1...X_n]$ sous la forme d'une prime de crédibilité. Dans de tels cas (assez rares), on a donc que :

$$E[X|X_1...X_n] = Z\overline{X} + (1 - Z)M$$

A.2.2 Estimation Bayesienne

Initialement, on a

$$(X_i|\Theta) \sim f_{X|\Theta}(x;\theta)$$

$$\Theta \sim f_{\Theta}(\theta)$$

qui génèrent une série d'observations $(x_1, ..., x_n)$.

On obtient donc, par une simple application du théorème de Bayes que

$$f_{\Theta|X_1,...,X_n}(\theta; x_1, ..., x_n) = \frac{f_{\Theta,X_1,...,X_n}(\theta, x_1, ..., x_n)}{f_{X_1,...,X_n}(x_1, ..., x_n)}$$

où

$$f_{X_1,...,X_n}(x_1, ..., x_n) = \int_{-\infty}^{\infty} f_{\Theta,X_1,...,X_n}(\theta, x_1, ..., x_n) d\theta$$

qui est une constante. On retrouve donc que

$$\begin{aligned} f_{\Theta|X_1,...,X_n}(\theta; x_1, ..., x_n) &\propto f_{\Theta,X_1,...,X_n}(\theta, x_1, ..., x_n) \\ &= f_{X_1,...,X_n|\Theta}(x_1, ..., x_n; \theta) f_\Theta(\theta) \end{aligned}$$

Et, par le lien de proportionnalité, on est souvent en mesure de retrouver la loi de la distribution a posteriori en faisant ressortir des constantes.

A.2.3 Hétérogénéité résiduelle

Dans un portefeuille d'assurance, ou une cellule de tarification, on retrouve toujours de l'hétérogénéité résiduelle due au fait que "toutes les questions n'ont pas été posées". On modélise cette différence entre les assurés par la v.a. θ, qui, par hypothèse, n'est pas une fonction du temps.
Sous cette hypothèse, on a donc que les observations $S_{i,t}$ sont indépendantes dans le temps t, *conditionnellement à* Θ_i, c'est à dire

$$\begin{aligned} (S_{1,1}|\Theta_1) \perp\!\!\!\perp ... \perp\!\!\!\perp (S_{1,5}|\Theta_1) \\ (S_{2,1}|\Theta_2) \perp\!\!\!\perp ... \perp\!\!\!\perp (S_{2,5}|\Theta_2) \end{aligned}$$

A.2.4 Prévision Bayesienne

A.2.4.1 Prime de risque

Tout d'abord, on note

$$\boxed{\mu(\theta) = E[S_{i,t}|\Theta_i = \theta] \stackrel{def}{=} \int_0^{\infty} x f_{S|\Theta}(x; \theta)}$$

L'espérance des coûts de chaque contrat sachant le degré de risque individuel (Θ_i).

A.2.4.2 Prime collective

Comme première estimation de la prime de risque, on utilise souvent la moyenne pondérée des risques individuels, soit le risque moyen pour la cellule de tarification. On note cette prime :

$$\boxed{\begin{aligned} m = E[\mu(\theta)] &\overset{def}{=} \int_{-\infty}^{\infty} \mu(\theta) f_{\Theta}(\theta) \\ &= E[E[S_{i,t}|\Theta]] \\ &= E[S_{i,t}] \end{aligned}}$$

A.2.4.3 Prime Bayésienne

Toutefois, bien que la prime collective soit globalement juste pour l'ensemble du portefeuille, elle ne l'est pas lorsqu'on observe chaque assuré indépendamment. C'est pourquoi on s'intéresse à la prime Bayésienne, qui tient compte de l'expérience passée de chaque assuré. On note cette prime :

$$\boxed{\begin{aligned} B_{i,n+1} &= E[\mu(\theta)|S_{i,1},...,S_{i,n}] \\ &= \int_{-\infty}^{\infty} \mu(\theta) f_{\Theta|S_{i,1},...,S_{i,n}}(\theta; s_{i,1},...,s_{i,n}) \end{aligned}}$$

où, par le théorème de Bayes

$$\begin{aligned} f_{\Theta|S_{i,1},...,S_{i,n}}(\theta; s_{i,1},...,s_{i,n}) &\propto f_{S_{i,1},...,S_{i,n}|\Theta_i}(s_{i,1},...,s_{i,n};\theta) f_{\Theta_i}(\theta) \\ &= \prod_{t=1}^{n} f_{S_{i,t}|\Theta_i}(s_{i,t};\theta) f_{\Theta}(\theta) \end{aligned}$$

A.2.5 Approche par la *distribution prédictive*

Lorqu'on se retrouve avec un cas qui est "moins doux" et qui ne donne pas de forme analytique facilement utilisable, on peut recourir à cette méthode numérique afin de calculer $B_{i,n+1}$
Tout d'abord, on a que

$$\begin{aligned} B_{i,n+1} &= E[\mu(\Theta_i)|S_{i,1},...,S_{i,n}] \\ &= \int_{0}^{\infty} s f_{S_{i,t}|S_{i,1},...,S_{i,n}}(s|s_1,...,s_n) ds \end{aligned}$$

où

$$f_{S_{i,t}|S_{i,1},...,S_{i,n}}(s|s_1,...,s_n) = \frac{f_{S_{i,t},S_{i,1},...,S_{i,n}}(s,s_1,...,s_n)}{f_{S_{i,1},...,S_{i,n}}(s_1,...,s_n)}$$

$$= \int_{-\infty}^{\infty} f_{S_{i,t}|\Theta_i}(s|\theta)\left(\frac{f_{S_{i,1}|\Theta}(s_1|\theta) * ... * f_{S_{i,n}|\Theta}(s_n|\theta)}{\int_{-\infty}^{\infty} f_{S_{i,1}|\Theta}(s_1|\phi) * ... * f_{S_{i,n}|\Theta}(s_n|\phi)d\phi}\right)d\theta$$

$$\overset{Bayes}{=} \int_{-\infty}^{\infty} f_{S_{i,t}|\Theta_i}(s|\theta) f_{\Theta|S_{i,1},...,S_{i,n}}(\theta|s_1,...,s_n)d\theta$$

Et on peut donc évaluer le tout numériquement, et ce, même pour des cas très complexes (Burr-Inverse Gaussienne par exemple).

A.2.6 Crédibilité Bayesienne linéaire

Dans certains cas, on peut réécrire la Bayesienne comme étant

$$B_{i,n} = X * \overline{S} + (1 - Z)M$$

Et on appelle cette prime la prime de crédibilité.
Voici quelques exemples où on peut obtenir une prime de crédibilité :

$$\begin{array}{lr}
(S|\Theta) \sim Poisson & \text{et } \Theta \sim Gamma \\
(S|\Theta) \sim Exponentielle & \text{et } \Theta \sim Gamma \\
(S|\Theta) \sim Normale & \text{et } \Theta \sim Normale \\
(S|\Theta) \sim Bernouilly & \text{et } \Theta \sim Beta \\
(S|\Theta) \sim Geometrique & \text{et } \Theta \sim Beta
\end{array}$$

Voir exemple pratique dans les notes pour avoir un exemple concret sur cette matière.

A.3 Modèle de Bühlmann

Le modèle de crédibilité Bayesien est le plus précis, mais il comporte deux limitations majeures :
1. La prime Bayesienne n'est linéaire que dans un nombre très limité de cas
2. la prime Bayesienne nous force à poser des hypothèses subjectives quant aux distributions de $S_{i,t}|\Theta_i$ et Θ_i.

A.3.1 Notation et identités à connaitre

Tout d'abord, on note les identités suivantes :

$$\boxed{\begin{aligned} S^2 &= E[Var(S_{i,t}|\Theta_i)] = E[\sigma^2(\Theta_i)] \\ a &= Var(E[S_{i,t}|\Theta_i]) = Var(\mu(\Theta_i)) \\ \delta_{i,j} &= \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \\ Cov(S_{i,t}, S_{i,u}) &= a + \delta_{t,u} S^2 \\ Cov(\mu(\Theta_i), S_{i,t}) &= a \end{aligned}}$$

A.3.2 Hypothèses du modèle de Bühlmann

1. Chaque contrat est indépendant quant à son niveau de risque individuel

$$\Theta_1 \perp\!\!\!\perp ... \perp\!\!\!\perp \Theta_I$$

2. Les variables aléatoires $S_{i,t}$ sont telles que

$$\begin{aligned} E[S_{i,t}|\Theta_i] &= \mu(\Theta_i) \\ Cov(S_{i,t}, S_{i,u}|\Theta_i) &= \delta_{t,u}\sigma^2(\Theta_i) \end{aligned}$$

L'idée du modèle de Bühlmann est de trouver β_0 et β_1 qui minimisent les "moindres carrés espérés", c'est à dire :

$$\phi = E\left[\left(\mu(\Theta_i) - (\beta_0 + \beta_1 \overline{S}_1)\right)^2 \right]$$

En dérivant partiellement ϕ par rapport à β_0 et à β_1 et en égalisant les deux équations à zéro, on obtient que

$$\boxed{\begin{aligned} &\beta_0 = E[\mu(\Theta_i)] - \beta_1 E[\overline{S}_i] = m\left(1 - \frac{n}{n+k}\right) \\ &\beta_1 = \frac{n}{n + \frac{S^2}{a}} \end{aligned}}$$

Et on peut alors conclure que la meilleure approximation linéaire de $\mu(\Theta_i)$

qui minimise ϕ est

$$\boxed{\begin{aligned} \Pi^B_{i,n+1} &= Z\overline{S}_i + (1-Z)M \\ Z &= \frac{n}{n+k} \\ \overline{S}_i &= \frac{\sum_{t=1}^n S_{i,t}}{n} \\ k &= \frac{S^2}{a} \\ m &= E[S_{i,t}] \end{aligned}}$$

A.3.3 Approche paramétrique

Lorsqu'on connaît les distributions de $(S_{i,t}|\Theta_i)$ et de Θ_i, il est facile d'obtenir directement des expressions pour S^2 et a. On peut donc directement utiliser ces quantités dans la formule de la prime de Bühlmann afin de calculer la prime crédibilisée.

A.3.4 Approche non paramétrique

En pratique, les lois de$(S_{i,t}|\Theta_i)$ et de Θ_i sont rarement connues, c'est pourquoi nous devons estimer les valeurs de S^2 et de a et se servant des observations que nous avons à notre disposition.

A.3.4.1 m

Comme nous l'aurions fait si nous y étions bêtement allé avec logique, nous avons que

$$\boxed{\begin{aligned} m &\overset{def}{=} E[\mu(\Theta_i)] \\ \widehat{m} &= \frac{\sum_{i=1}^I \sum_{t=1}^n S_{i,t}}{In} \end{aligned}}$$

est un estimateur non-biaisé pour m.

A.3.4.2 S^2

Par définition, on a que

$$\boxed{\begin{aligned} &S^2 \stackrel{def}{=} E[\sigma^2(\Theta_i)] \\ \\ &\overline{S}^2 = \frac{\sum_{i=1}^{I} \widehat{\sigma}^2(\Theta_i)}{I} = \frac{\sum_{i=1}^{I} \sum_{t=1}^{n} (S_{i,t} - \overline{S}_i)^2}{I(n-1)} \end{aligned}}$$

est un estimateur non-biaisé pour S^2
Attention : Puisque, dans le cas non paramétrique, la variance est estimée, il faut diviser par $n-1$ et non n !

A.3.4.3 a

Dans ce cas, il faut **faire attention**, car l'estimateur que nous pourions sembler logique :

$$\boxed{\widehat{a} = \frac{\sum_{i=1}^{I} (\overline{S}_i - \overline{S})^2}{I-1}}$$

est biaisé !
En effet, on omettant le développement, on a que

$$E[\widehat{a}] = a + \frac{S^2}{n}$$

$\widehat{a}$ est donc asymptotiquement sans biais, mais, à l'examen, afin d'éviter le problème du biais, on utilisera plutôt afin d'estimer la valeur de a :

$$\boxed{\begin{aligned} &\widehat{a}' = \widehat{a} - \frac{S^2}{n} \\ \\ &\widehat{a}' = \frac{\sum_{i=1}^{I} (\overline{S}_i - \overline{S})^2}{I-1} - \frac{S^2}{n} \end{aligned}}$$

Malgré le fait qu'elle pourrait nous retourner des valeurs négatives.

A.4 Modèle de Bühlmann-Straub

Un gros point négatif du modèle de Bühlmann est le fait qu'on assume que tous les contrats sont en vigueur pour toutes les années, et ce, durant l'année complète et que l'exposition au risque est donc la même pour tous les contrats, ce qui n'est évidemment pas le cas dans un contexte pratique.

Dans ce contexte, on note $W_{i,t}$ le degré d'exposition au risque de l'assuré i au temps t et $X_{i,t} = \frac{S_{i,t}}{W_{i,t}}$ le ratio sinistres/exposition pour l'assuré i au temps t.
L'actuaire pourrait aussi définir $W_{i,t}$ comme le montant de primes acquises par l'assuré i au temps t afin de modéliser le loss ratio au lieu de la prime pure.

A.4.1 Hypothèses du modèle de Bühlmann-Straub

1. Les contrats sont indépendants ($\Theta_1 \perp\!\!\!\perp ... \perp\!\!\!\perp \Theta_I$)
2. $E[W_{i,t}|\Theta_i] = \mu(\Theta_i)$
3. $Var(X_{i,t}|\Theta_i) = \frac{\sigma^2(\Theta_i)}{W_{i,t}}$

A.4.2 Notation supplémentaire

$$W_{i\bullet} = \sum_{t=1}^{n} W_{i,t}$$

$$W_{\bullet\bullet} = \sum_{i=1}^{I} \sum_{t=1}^{n} W_{i,t}$$

$$Z_{\bullet} = \sum_{i=1}^{I} Z_i$$

$$X_{i,W} = \sum_{t=1}^{n} \left(\frac{W_{i,t}}{W_{i\bullet}} \right) X_{i,t}$$

$$\begin{aligned} X_{WW} &= \sum_{i=1}^{I} \sum_{t=1}^{n} \left(\frac{W_{i,t}}{W_{\bullet\bullet}} \right) X_{i,t} \\ &= \sum_{i=1}^{I} \left(\frac{W_{i\bullet}}{W_{\bullet\bullet}} \right) \sum_{t=1}^{n} \left(\frac{W_{i,t}}{W_{i\bullet}} \right) X_{i,t} \\ &= \sum_{i=1}^{I} \left(\frac{W_{i\bullet}}{W_{\bullet\bullet}} \right) X_{i,W} \end{aligned}$$

$$X_{Z,W} = \sum_{i=1}^{I} \left(\frac{Z_i}{Z_{\bullet}} \right) X_{i,W}$$

A.4.3 Covariances

Identités à connaitre afin de faire les preuves, si besoin.

$$\boxed{\begin{aligned} Cov(X_{i,t}, X_{k,u}) &= \delta_{i,k}\left(a + \frac{\delta_{t,u}S^2}{W_{i,t}}\right) \\ Cov(\mu(\Theta_i), X_{k,u}) &= \delta_{i,k}a \\ Cov(X_{i,t}, X_{k,W}) &= \delta_{i,k}\left(a + \frac{S^2}{W_{i\bullet}}\right) \end{aligned}}$$

A.4.4 Prime de crédibilité linéaire

Comme pour le modèle de Bühlmann, on cherche β_0 et β_1 qui minimisent

$$\phi = E\left[(\mu(\Theta_i) - (\beta_0 + \beta_1 X_{i,W}))^2\right]$$

On trouve finalement que

$$\Pi^{BS}_{i,n+1} = Z_i X_{i,W} + (1 - Z_i)M$$

$$Z_i = \frac{W_{i\bullet}}{W_{i\bullet} + \frac{S^2}{a}}$$

A.4.4.1 $\widehat{m}$

Tout d'abord, on a

$$\boxed{\begin{aligned} \widehat{m} &= \sum_{i=1}^{I}\left(\frac{W_{i\bullet}}{W_{\bullet\bullet}}\right)X_{i,W} \\ &= X_{W,W} \\ &= \frac{\sum \text{Sinistres}}{\sum UA} \end{aligned}}$$

Toutefois, l'estimateur linéaire de m a variance minimale n'est pas $X_{W,W}$, mais bien

$$\boxed{\widehat{m}' = X_{Z,W} = \sum_{i=1}^{I} \left(\frac{Z_i}{Z_\bullet} \right) X_{i,W}}$$

A.4.4.2 $\widehat{S}^2$

Épargnons le développement, on a que

$$\boxed{\widehat{S}^2 = \frac{\sum_{i=1}^{I} \sum_{t=1}^{n} W_{i,t}(X_{i,t} - X_{i,W})^2}{I(n-1)}}$$

Notons encore une fois que, puisqu'on ne connait pas les ditrbutions, la variance est estimée et on divise par $(n-1)$ et non seulement par n.

A.4.4.3 $\widehat{a}$

Comme avec le modèle de Bühlmann, l'estimateur logique de a est biaisé ! Il faut donc faire un ajustement au biais afin d'avoir un estimateur non biaisé pour a.
On a

$$\boxed{\widehat{a} = \sum_{i=1}^{I} W_{i\bullet}(X_{i,W} - X_{W,W})^2}$$

qui est notre estimateur biaisé de a.
Après ajustement, on obtient plutôt

$$\boxed{\widehat{a}' = \left(\frac{W_{\bullet\bullet}}{W_{\bullet\bullet}^2 - \sum_{i=1}^{I} W_{i\bullet}^2} \right) \left(\widehat{a} - (I-1)\widehat{S}^2 \right)}$$

Lorsqu'on a accès à un ordinateur, il est également possible d'utiliser

$$\boxed{\tilde{a} = \frac{\sum_{i=1}^{I} Z_i(X_{i,W} - X_{Z,W})^2}{I-1}}$$

, mais cette méthode nous oblige à résoudre le tout par la méthode du point milieu, car $\tilde{a}$ est fonction de Z_i et vice versa.

Annexe B

Résumé examen final

But et notation

- Évaluation des réserves : Rôle de l'actuaire **corporatif**
- Réserve IARD : Montant nécessaire pour payer les **sinistres déjà survenus à la date d'évaluation** (Sinistres dont le développement n'est pas complet)
- Les réserves doivent être signées par un **actuaire désigné**
- Les réserves **doivent** être évaluées au 31 décembre, mais plusieurs compagnies les évaluent au **trimestre** ou au **mois**

Il est important de bien évaluer les réserves, car

Si les réserves sont surévaluées :

- Dépenses $\nearrow$
- Profits $\searrow$
- Impôts à payer $\searrow$
- Surplus de l'assureur $\searrow$
- Valeur de la compagnie (prix de l'action) $\searrow$

Si les réserves sont sous-évaluées

- Surévalue la santé financière de la compagnie
- Expose l'assuré à ne pas être payé en cas de réclamation
- Expose l'assureur à la ruine

Définitions

Dossier de sinistre :

- Un dossier est ouvert à chaque fois qu'un assuré fait une réclamation.
- Ce dossier contient toutes les informations relatives à la réclamation (date d'accident, date de réclamation, montants et moments de

chaque paiement, information qualitative).

Case Reserves :

- Une *Case Reserve* est la meilleure estimation d'un montant de sinistre avant même qu'un paiement soit fait (**expert en sinistre** ou **modèle prédictif**).
- Les *Case Reserves* sont la somme des Case Reserves individuelles.

Gross IBNR (ou Bulk Reserve)

- IBNR = Incurred but not Reported
- Contiens les provisions pour :
  - Développement futur des sinistres
  - Sinistres fermés pouvant rouvrir
  - Sinistres encourus, mais non rapportés (Pure IBNR)
  - Sinistres rapportés, mais non codés dans le système informatique

Réserve totale :

$$\text{Réserve totale} = \text{Case Reserves} + \text{Gross IBNR}$$

Développement :

- Changement temporel de la somme des paiements effectués sur tous les dossiers de sinistre (Prestations payées durant une période)

Paid Age-to-Age

- Développement entre deux dates données (on suit généralement après chaque année ou mois d'âge successifs)

Age-to-ultimate Development

- Développement des sinistres ayant atteint un certain âge jusqu'à l'ultime.

Paid Loss Development Factor (PLDF$_{j-1,j}$)

$$\text{PLDF}_{j-1,j} = \frac{\text{Somme des paiements faits par l'assureur sur tous les dossiers de sinistres à t=j}}{\text{Somme des paiements faits par l'assureur sur tous les dossiers de sinistres à t=j-1}}$$

Sinistres en suspend (SS)

- Somme des *Case Reserves* qui ne sont pas encore fermées à une date donnée.

Sinistres payés (SP)

- Somme des indemnités versées aux réclamants (**inclus les frais de règlement**)

Sinistres encourus (SE)

- Somme des montants engendrés par un sinistre (Passé + Futur)

$$SE = SP + SS$$

Incurred Loss Development Factor (ILDF$_{j-1,j}$)

$$\text{ILDF}_{j-1,j} = \frac{\sum SE@j}{\sum SE@j-1} = \frac{C_j}{C_{j-1}} = \frac{\text{Encouru cumulatif @ t=j}}{\text{Encouru cumulatif @ t=j-1}}$$

Notation en triangles cumulatifs

Il est commode de noter $C_{i,j}$, le total des sinistres survenus dans l'année i développés pendant j années de la façon suivante :

i/j	Age 1	Age 2	Age 3	Age 4	Age 5
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$	
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$		
4	$C_{4,1}$	$C_{4,2}$			
5	$C_{5,1}$				

$\Rightarrow$ Année de calendrier 3

Méthode de Chain-Ladder

Notions mathématiques

Pour la méthode de Chain-Ladder, on assume tout simplement que LDF_j, le taux de développement pour l'année j est le même pour toutes les années de sinistres. On a donc

$$C_{i,j+1} = C_{i,j} \times LDF_j$$

Avec

$$\widehat{LDF}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}, \forall j = 1, ..., n-1$$

Pour le calcul de la réserve IBNR, on s'intéresse à la différence entre les coûts à l'ultime, soit les coûts totaux lorsque les sinistres seront pleinement développés, et les prestations qui ont été payé jusqu'à la date d'évaluation. Il est donc crucial de bien calcul $C_{i,n}$. Pour Chain-Ladder, on a

$$\widehat{C}_{i,n} = (\widehat{LDF}_{n+1-i} \times ... \times \widehat{LDF}_{n-1}) \times C_{i,n+1-i}$$

Et la réserve pour l'année i est

$$\widehat{R}_i = \widehat{C}_{i,n} - C_{i,n+1-i}, \forall i = 2, ..., n$$

Et la réserve totale est

$$\widehat{R} = \sum_{i=2}^{n} \widehat{R}_i$$

Remarques

- Cette méthode suppose que les années d'accident sont indépendantes entre elles.
- On assume aussi implicitement que l'âge des sinistres est la seule variable explicative du développement des sinistres.
- On fait l'hypothèse simplificatrice que $LDF_{i,j} = LDF_j$. Ceci n'est pas nécessairement vrai, car plusieurs facteurs peuvent venir influencer la vitesse de développement de l'année i :
  - Changements internes dans les procédures de la compagnie
  - Changement de loi
  - ...
- Avec Chain-Ladder, la réserve de l'année d'accident la plus récente est sujette à une forte incertitude, car elle dépend seulement des paiements effectués dans l'année de calendrier la plus récente.

Méthode de Bornhuetter-Ferguson

Notions mathématiques

La méthode de B-F se fait en trois grandes étapes

Étape 1 : Estimation des sinistres ultimes

On assume que le **taux de sinistre** (Loss Ratio) est connu pour chaque année d'accident. Ainsi, soit LR_i le taux de sinistre de l'année i et PA_i les primes acquises pour l'année d'accident i, on a

$$\boxed{\widehat{C}_{i,n}^{(B-F)} = \widehat{LR_i} \times PA_i}$$

Étape 2 : Calcul des facteurs de développement

Le calcul des facteurs de développement est exactement le même que pour Chain-Ladder, c'est à dire

$$\widehat{LDF}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}$$

Étape 3 : Calcul de la réserve

Pour B-F, on a

$$\boxed{\widehat{R}_i^{(B-F)} = \widehat{C}_{i,n}^{(B-F)} \times \left(1 - \frac{1}{\prod_{j=n+1-i}^{n-1} \widehat{LDF}_j}\right)}$$

Remarques

1. L'avantage de cette méthode est qu'elle permet une meilleure stabilité des réserves des années de survenance récentes.
2. L'inconvénient majeur de cette méthode est qu'elle requiert l'utilisation de données externes (LR_i, PA_i) qui introduisent de la subjectivité.
3. Dans cette méthode, c'est comme si on assumait que l'âge n'affecte pas le développement du sinistre

$$\begin{aligned}
\widehat{C}_{i,n}^{(B-F)} &= PA_i \times LR_i = a_i \Rightarrow \text{ordonnée à l'origine} \\
\widehat{C}_{i,n}^{(C-L)} &= C_{i,n+1-i} \times \left(\prod_{j=n+1-i}^{n-1} LDF_j\right) \\
&= C_{i,n+1-i} \times B_i \Rightarrow \text{Pente}
\end{aligned}$$

Il serait donc intéressant de mélanger C-L et B-F afin d'obtenir une régression linéaire avec pente et ordonnée à l'origine.

Méthode de Mack

Essentiellement, la méthode de Mack est simplement la méthode de Chain-Ladder, mais avec un cadre statistique plus solide qui permet une estimation de la variance des réserves qui peut être très utile afin d'avoir une meilleure idée du risque auquel s'expose la compagnie.

Hypothèses du modèle de Mack

1. $\{C_{1,1}, ..., C_{1,n}\} \perp\!\!\!\perp \{C_{2,1}, ..., C_{2,n}\} \perp\!\!\!\perp ... \perp\!\!\!\perp \{C_{n,1}, ..., C_{n,n}\}$
2. $E(C_{i,k+1}|C_{i,1}, ..., C_{i,k}) = LDF_k$
3. $Var(C_{i,k+1}|C_{i,1}, ..., C_{i,k}) = \sigma_k^2 \times C_{i,k}$

Estimation des LDF

Notons tout d'abord H_i les données historiques pour l'année d'accident i et D le triangle de données complet. Comme pour le modèle Chain-Ladder, on a

$$\widehat{LDF}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}$$

De plus, le cadre théorique du modèle de Mack nous permet de prouver que :

$$E(C_{i,j}|D) = C_{i,n+1-i} \times LDF_{n+1-i} \times ... \times LDF_{j-1}$$

On a aussi, avec D_k, le triangle tronqué à l'âge k,

$$E(C_{i,k+1}|D_k) \stackrel{H2}{=} C_{i,k} \times LDF_k$$

et

$$E(\widehat{LDF}_k|D_k) = \frac{\sum_{i=1}^{n-k} C_{i,k} LDF_k}{\sum_{i=1}^{n-k} C_{i,k}} = LDF_k$$

Un autre résultat important que l'on peut prouver par la méthode de Mack est que

$$\mathrm{Cov}(\widehat{LDF}_k, \widehat{LDF}_j) = 0$$

et

$$E(\widehat{R}_i) = R_i$$

Estimation des σ_j^2

L'estimateur non-biaisé de σ^2 est

$$\boxed{\widehat{\sigma}_j^2 = \frac{1}{n-j-1} \sum_{i=1}^{n-j} C_{i,j} \times \left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{LDF}_j \right)^2}$$

Qui correspond à l'équation 6.4.
De plus, comme on peut le voir, le calcul de $\widehat{\sigma}^2$ est problématique pour $j = n-1$. On utilise donc l'approximation suivante :

$$\boxed{\widehat{\sigma}_{n-1}^2 = min\left(\frac{\widehat{\sigma}_{n-2}^4}{\widehat{\sigma}_{n-3}^2}, min\left(\widehat{\sigma}_{n-3}^2, \widehat{\sigma}_{n-2}^2\right)\right)}$$

Estimation de l'erreur quadratique moyenne

Tout d'abord, puisque $\widehat{R}_i$ est un estimateur non biaisé de R, on a $Var(\widehat{R}_i) = MSE(\widehat{R}_i)$

$$\boxed{\begin{array}{l} MSE(\widehat{R}_i) = E\left((\widehat{R}_i - R_i)^2|D\right) \\ MSE(\widehat{R}_i) = \widehat{C}_{i,n}^2 \times \displaystyle\sum_{k=n-i+1}^{n-1} \frac{\widehat{\sigma}_k^2}{\widehat{LDF}_k^2} \times \left(\frac{1}{\widehat{C}_{i,k}} + \frac{1}{\sum_{j=1}^{n-k} C_{j,k}}\right) \end{array}}$$

Intervalles de confiance

En ayant maintenant estimé l'erreur quadratique moyenne des réserves calculées, il est possible de poser des hypothèses quant à la loi suivie par la réserve afin de construire une intervalle de confiance.

Hypothèse de normalité

Puisqu'on assume ici

$$\widehat{R}_i \sim N(E[\widehat{R}_i], \widehat{MSE}(\widehat{R}_i))$$

Ce qui impose

$$\boxed{I.C.^N(\widehat{R}_i) = \left[\widehat{R}_i - Z_{\frac{\alpha}{2}} \times \sqrt{\widehat{MSE}(\widehat{R}_i)}, \widehat{R}_i + Z_{\frac{\alpha}{2}} \times \sqrt{\widehat{MSE}(\widehat{R}_i)}\right]}$$

Hypothèse de log-normalité

On suppose ici

$$R_i \sim LN(\mu_i, \sigma_i^2)$$

Avec

$$E[R_i] = \widehat{R}_i = e^{\mu_i + \frac{\sigma_i^2}{2}}$$
$$MSE(R_i) = E(R_i)^2 \times \left(e^{\sigma_i^2} - 1\right)$$

$$\Leftrightarrow \mu_i = ln(\widehat{R}_i) - \frac{\sigma_i^2}{2}$$
$$\Leftrightarrow \sigma_i^2 = ln\left(1 + \left(\frac{\sqrt{\widehat{MSE}(\widehat{R}_i)}}{\widehat{R}_i}\right)^2\right)$$

Ce qui impose

$$\boxed{I.C.^{LN}(\widehat{R}_i) = \left[e^{\mu_i - Z_{\frac{\alpha}{2}}\sigma_i}, e^{\mu_i + Z_{\frac{\alpha}{2}}\sigma_i}\right]}$$

Variabilité de la réserve totale

On a tout simplement

$$\begin{aligned} MSE(\widehat{R}) &= E\left[(\widehat{R} - R)^2 | D\right] \\ &= E\left(\left(\sum_{i=2}^{n} \widehat{R}_i - \sum_{i=2}^{n} R_i\right)^2 | D\right) \end{aligned}$$

$$\boxed{\widehat{MSE}(\widehat{R}) = \sum_{i=2}^{n}\left(\widehat{MSE}(\widehat{R}_i) + \widehat{C}_{i,n} \times \left(\sum_{j=i+1}^{n} \widehat{C}_{j,n}\right) \times \sum_{k=n-i+1}^{n-1} \frac{\frac{2\widehat{\sigma}_k^2}{\widehat{LDF}_k^2}}{\sum_{j=1}^{n-k} C_{j,k}}\right)}$$

Méthode GLM

- La méthode GLM ne se limite pas à la loi normale. En effet, on peut utiliser n'importe quelle loi de la famille exponentielle. (Normale, Log-Normale, Binomiale, Poisson, Gamma, Exponentielle, Inverse Gaussienne,...)
- Permets d'incorporer au modèle des variables explicatives autres que l'âge et l'année d'accident. (Changement de VP, Changement de loi, catastrophe naturelle)
- Permets de faire des réserves granulaires au lieu d'agréger les données par âge et année d'accident.

Remarques

1. Avec les GLM, on utilise le triangle des sinistres incrémentaux ($Y_{i,j}$) et non cumulatifs ($C_{i,j}$).
2. En pratique, on utilise une loi de la famille exponentielle selon la variance observée dans les $Y_{i,j}$.

Loi	Fonction de variance	Utilisation en réserves
Normale	$\sigma^2 \neq f(u)$	Très rare, car la variance diminue avec l'âge j
Poisson	μ	Souvent
Gamma	var $= f(\mu^2)$	Moins souvent
Inverse Gaussienne	var $= f(\mu^3)$	Très rare

Le modèle de base

Si on ne se sent pas exotique, on peut passer un modèle GLM sans variables explicatives autres que l'âge et l'année d'accident. Dans ce cas, on a donc

$$\alpha_1 = \beta_1 = 0$$
$$E[Y_{i,j}] = \mu_{i,j} = e^{\mu+\alpha_i+\beta_j}$$

Où on obtient ici $\tilde{\beta} = [\mu, \alpha_2, ..., \alpha_n, \beta_2, ..., \beta_n]$ par maximum de vraisemblance.

Un modèle plus élaboré

Si on se sent aventurier, on peut ajouter d'autres variables explicatives à notre modèle de calcul des réserves. On pourrait par exemple ajouter la variable indicatrice d'un changement de VP $X_{i,j}$ et on obtiendrait ainsi :

$$E[Y_{i,j}] = \mu_{i,j} = e^{\mu+\alpha_i+\beta_j+\gamma X_{i,j}}$$

Théorème de Taylor multidimensionnel (nice...)

Soit $Y = g(X_1, X_2, ..., X_n)$ une fonction statistique des variables aléatoires $(X_1, ..., X_n)$, on a

$$\boxed{E[Y] \approx g(E[X_1], ..., E[X_n]) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 g(E[X_1], .., E[X_n])}{\partial x_i \partial x_j}\text{Cov}(X_i, X_j)}$$

et

$$\boxed{Var(Y) \approx \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial}{\partial x_i}\, g(E[X_1], ..., E[X_n])\right)\left(\frac{\partial}{\partial x_j}\, g(E[X_1], ..., E[X_n])\right)\text{Cov}(X_i, X_j)}$$

Malheureusement, puisque Y peut facilement être une quantité souvent utilisée telle $\widehat{R}_i$, il faut le sentir.

Estimation des coefficients

Tel que vu en ACT-2003, la meilleure façon d'estimer les coefficients nécessaires au calcul des réserves IARD est le maximum de vraisemblance. Dans le cas Poisson, on aurait donc

$$\begin{aligned}
\mu_{i,j} &= e^{\mu+\alpha_i+\beta_j} \\
l(\Theta) &= \sum_{(i,j)\in\triangleright} ln\left(\frac{e^{-\mu_{i,j}}(\mu_{i,j})^{y_{i,j}}}{(y_{i,j})!}\right) \\
&= \sum_{(i,j)\in\triangleright} (y_{i,j}\ln(\mu_{i,j}) - \mu_{i,j} - \ln(y_{i,j}!)) \\
&\propto \sum_{(i,j)\in\triangleright} (y_{i,j}\ln(\mu_{i,j}) - \mu_{i,j})
\end{aligned}$$

On utilise le log-vraisemblance proportionnelle pour éviter les problèmes numériques que pourraient poser le calcul de 2000! par exemple. Maintenant que la fonction à optimiser est posée, il ne reste plus qu'à la maximiser par la méthode de Newton-Raphson. On a donc

$$\Theta = \begin{bmatrix} \mu \\ \alpha_2 \\ ... \\ \alpha_n \\ \beta_2 \\ ... \\ \beta_n \end{bmatrix}$$

$$\boxed{\Theta^{(i+1)} = \Theta^{(i)} + I(\Theta^{(i)})^{-1}S(\Theta^{(i)})}$$

avec

$$\boxed{\begin{aligned} &S(\Theta) = X^T W, W = Y - \widehat{Y} \\ &I(\Theta) = X^T H X \end{aligned}}$$

$$H = \begin{bmatrix} e^{X_1\beta} & 0 & ... & 0 \\ 0 & e^{X_2\beta} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & e^{X_n\beta} \end{bmatrix}$$

Quantités importantes dans le cas Poisson (Sûrement à l'examen)

Maintenant que nous avons estimé les coefficients, on peut se laisser aller. Ainsi, soit X^*, la matrice schéma servant à prédire la partie inférieure du triangle de données, on a :

$$\widehat{R} = e^{X_1^*\beta} + e^{X_2^*\beta} + ... + e^{X_n^*\beta}$$

$$\boxed{E[\widehat{R}] = g(\widehat{\beta}) + \frac{1}{2}\sum_{i=0}^{8}\sum_{j=0}^{8}\left(\frac{\partial^2 R}{\partial\widehat{\beta}^2}\right)_{i+1,j+1} I^{-1}(\widehat{\beta})_{i+1,j+1}}$$

Avec

$$\frac{\partial^2 R}{\partial\widehat{\beta}^2} = (X^*)^T W X^*$$

$$W = \begin{bmatrix} e^{X_1^*\widehat{\beta}} & 0 & ... & 0 \\ 0 & e^{X_2^*\widehat{\beta}} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & e^{X_n^*\widehat{\beta}} \end{bmatrix}$$

$$\boxed{\begin{aligned} Var(\widehat{R}) &= \sum_{i=0}^{8}\sum_{j=0}^{8}\left(\frac{\partial\widehat{R}}{\partial\widehat{\beta}}\right)_{i+1} I^{-1}(\widehat{\beta})_{i+1,j+1}\left(\frac{\partial\widehat{R}}{\partial\widehat{\beta}}\right)_{j+1} \\ &= \left(\frac{\partial\widehat{R}}{\partial\widehat{\beta}}\right)^T I^{-1}(\widehat{\beta})\left(\frac{\partial\widehat{R}}{\partial\widehat{\beta}}\right) \end{aligned}}$$

De plus, il est important de noter que

$$\left(\frac{\partial\widehat{R}}{\partial\widehat{\beta}}\right)_{i+1} = [X_1^*]_{i+1}e^{X_1^*\beta} + [X_2^*]_{i+1}e^{X_2^*\beta} + ... + [X_n^*]_{i+1}e^{X_n^*\beta}$$

$$\boxed{\left(\frac{\partial\widehat{R}}{\partial\widehat{\beta}}\right) = (X^*)^T M}$$

$$M = \begin{bmatrix} e^{X_1^*\beta} \\ e^{X_2^*\beta} \\ ... \\ e^{X_n^*\beta} \end{bmatrix}$$

Tests d'hypothèses avec les GLM (#TBT)

Cette section est exactement la même matière que vue en ACT-2003. On a donc les cas suivants :

1. H_0 : Un sous-modèle de M1 (noté M0) est acceptable
2. H_1 : Il est nécessaire d'utiliser le modèle complet (noté M1)

Pour tester H_0 contre H_1, on utilise la statistique :

$$\chi^2_{\text{obs}} = 2\left(l(H_1) - l(H_0)\right)$$

et on rejette H_0 au niveau $100(1-\alpha)\%$ si

$$\chi^2_{\text{obs}} \geq \chi^2_{\alpha}(\Delta\text{Nombre de paramètres})$$

Méthode de réserves IARD basée sur la théorie de la crédibilité

Comme on a vu précédemment, Chain-Ladder est inefficace pour estimer la réserve des années d'accident récentes alors que Bornhutter-Ferguson est plutôt inefficace pour estimer les vieilles années d'accident. La méthode basée sur la théorie de la crédibilité cherche donc à estimer la réserve pour l'année d'accident i en faisant une pondération entre les réserves obtenues par C-L et B-F de sorte à surpondérer la méthode la plus adéquate.
En évitant une majeure partie du développement mathématique, on retrouve donc

$$\boxed{\begin{aligned} \widehat{R}_i^{\text{cred}} &= c_i \widehat{R}_i^{\text{C-L}} + (1-c_i)\widehat{R}_i^{\text{B-F}} \\ &= ... \\ &= \left[c_i \widehat{C}_{i,n}^{C-L} + (1-c_i)U_i\right](1-\beta_{n-i+1}) \end{aligned}}$$

De plus, Benktander et Hovinen ont posé les hypothèses suivantes :

1. $U_i(1-\beta_{n-i+1}) = \widehat{R}_i^{B-F}$
2. $\widehat{R}_i^{B-F} = U_i - C_{i,n-i+1}$

Qui nous permettent de faire un lien entre les deux méthodes et donc d'obtenir

$$\boxed{\widehat{C}_{i,n}^{B-H} = \beta_{n-i+1}\widehat{C}_{i,n}^{C-L} + (1-\beta_{n-i+1})\widehat{C}_{i,n}^{B-F}}$$

Avec

$$\boxed{c_i = \beta_{n-i+1} = \frac{1}{\prod_{j=n-i+1}^{n-1} LDF_j}}$$

On appelle aussi cette méthode la méthode Bornhuetter-Ferguson itérée

Actualisation des réserves IARD

Cette section est très simple. L'idée est simplement que l'on doit actualiser les flux monétaires selon les années de calendrier et non par année d'accident. On suit donc les étapes suivantes :

1. Développer le triangle (partie inférieure) avec la méthode souhaitée (C-L, B-F, Mack, GLM, B-H).
2. Calculer les sinistres incrémentaux $Y_{i,j}$ dans la partie inférieure du triangle,
3. Sommer les $Y_{i,j}$ de l'étape 2 par CY.
4. Actualiser par CY avec la structure à terme du problème.