Actuariat de l'assurance non-vie

Nicolas Baradel

9 janvier 2024

Table des matières

Introduction

Les assureurs prennent des risques pour les assurés, mais comment proposer une tarification individuelle ? Certains sinistres ne sont pas connus ou le montant n'est pas déterminé à la fin du contrat, comment évaluer la provision ? La solvabilité ?

Nous commencerons par introduire les Modèles Linéaires Généralisés afin de proposer une méthode de tarification individuelle. Il s'agit d'une tarification a priori, c'est à dire sans observer la sinistralité propre à l'assuré au court de sa souscription. Nous enchaînerons sur la tarification a posteriori, qui prend en compte cette observation, lorsqu'elle existe.

Nous poursuivrons sur la vie des sinistres après la couverture du contrat, certains n'étant pas encore déclarés ou leur coût demeure indéterminé. Nous commencerons par définir les triangles de liquidation, qui seront les données disponibles afin d'évaluer les provisions associées. Nous proposerons dans un premier temps la méthode de Chain Ladder - Mack, la plus utilisée. Nous proposerons une extension vers la méthode de Schnieper qui propose de séparer les sinistres non déclarés de ceux déclarés mais dont le coût n'est pas encore déterminé. Enfin, nous étudierons comment la méthode du bootstrap permet de tenir compte de l'incertitude et d'évaluer la solvabilité, puis nous concluerons sur la prise en compte de l'inflation.

Première partie

Tarification

On utilise souvent le modèle collectif en assurance, même pour une tarification individuelle (car dans ce cas, si le nombre de sinistres suit une loi de Bernoulli, on retrouve le modèle individuel). On rappelle le modèle collectif (pour plus de détails, voir [1]), on commence par définir ce qu'est une « somme aléatoire ».

Définition 0.1. *Soient N une variable aléatoire à valeurs dans $\mathbb{N}$ et $(Y_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires réelles. La variable aléatoire S définie par :*

$$S = \sum_{k=1}^{N} Y_k$$

est appelée « somme aléatoire ». C'est une somme finie (puisque N est finie), $\forall \omega \in \Omega$,

$$S(\omega) = \sum_{k=1}^{N(\omega)} Y_k(\omega)$$

et $S(\omega) = 0$ si $N(\omega) = 0$. Le terme de « somme aléatoire » vient du fait que, le nombre de Y_k (eux mêmes aléatoires) dans la somme définissant S est aléatoire et dépend de ω.

La variable aléatoire S représente la sinistralité totale, et son écriture en « somme aléatoire » est le modèle collectif. La variable aléatoire N représente le nombre total de sinistres et la suite $(Y_k)_{k\geq 1}$ le montant des sinistres. Si N et Y_1 sont des variables aléatoires de L^1, alors (voir [1, Première formule de Wald - Théorème 5.5]) $S \in L^1$ et :

$$\mathbb{E}(S) = \mathbb{E}(N)\mathbb{E}(Y_1).$$

Toutefois, l'assureur va tarifer individuellement les assurés : ils n'ont pas tous le même profil et donc les mêmes lois de sinistres.

Soit $n \geq 1$ le nombre de contrats d'assurance, et pour tout $i \in \{1, \ldots, n\}$, N^i le nombre de sinistres de l'assuré i et $(Y_k^i)_{k\geq 1}$ leurs coûts associés. On pose

$$S^i = \sum_{k=1}^{N^i} Y_k^i, \quad 1 \leq i \leq n,$$

le coût total de la sinistralité de l'individu i. Si, par exemple, $N^i \sim \mathcal{P}(\lambda_i)$ et que tous les assurés sont indépendants, alors le coût total des assurés :

$$S = \sum_{i=1}^{n} S^i$$

est encore un modèle collectif (voir [1, Théorème 5.20]). Dans la suite, nous allons nous intéresser à la sinistralité individuelle S_i de chaque assuré $i \in \{1, \ldots, n\}$.

L'espérance du coût des sinistres est ce qu'on appelle la *prime pure* qui est ici : $\mathbb{E}(S^i) = \mathbb{E}(N^i)\mathbb{E}(Y_1^i)$. Afin d'estimer $\mathbb{E}(S^i)$, nous allons estimer $\mathbb{E}(N^i)$ et $\mathbb{E}(Y_1^i)$. Etant donné que chaque individu aura ses caractéristiques propres, il convient d'écrire un modèle permettant d'avoir une prime individuelle. Si les caractéristiques observables sont représentées par une variable aléatoire X^i à valeurs dans $\mathbb{R}^d$, nous allons nous intéresser à $\mathbb{E}(N^i \mid X^i = x^i)$ et à $\mathbb{E}(Y_1^i \mid X^i = x^i)$ avec $x^i \in \mathbb{R}^d$. En conséquence, la prime pure de l'assuré i sera

$$\mathbb{E}(S^i \mid X^i = x^i) = \mathbb{E}(N^i \mid X^i = x^i)\mathbb{E}(Y_1^i \mid X^i = x^i).$$

Afin d'estimer pour chaque assuré son nombre moyen de sinistres $\mathbb{E}(N^i \mid X^i = x^i)$ et son coût moyen par sinistre $\mathbb{E}(Y_1^i \mid X^i = x^i)$, nous allons utiliser des modèles linéaires généralisés.

1 Modèles Linéaires Généralisés

Définition 1.1 (Famille exponentielle)**.** *Soit $(\mathbb{P}_\theta)_{\theta\in\mathbb{R}}$ une famille de mesures de probabilité. Le modèle fait parti de la famille exponentielle si cette famille est dominée et que sa densité peut s'écrire sous la forme :*

$$f(y; \theta, \phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right),$$

où a, b et c sont des fonctions mesurables, a est strictement positive et ϕ est un paramètre de dispersion.

Un premier résultat important est la caractérisation des deux premiers moments.

Lemme 1.2. *Si $Y \sim \mathbb{P}_\theta$ et si b est une fonction C^2, alors*

$$\mathbb{E}[Y] = b'(\theta),$$
$$Var(Y) = b''(\theta)a(\phi).$$

Démonstration. Si $Y \sim \mathbb{P}_\theta$, on a :

$$\mathbb{E}[\partial_\theta \log(f(Y; \theta, \phi))] = 0.$$

En effet, en notant ν la mesure dominante de la famille :

$$\begin{aligned}\mathbb{E}\left[\partial_\theta \log(f(Y;\theta,\phi))\right] = \mathbb{E}\left[\frac{\partial_\theta f(Y;\theta,\phi)}{f(Y;\theta,\phi)}\right] &= \int_{\mathbb{R}} \partial_\theta f(y;\theta,\phi) d\nu(y) \\ &= \partial_\theta \int_{\mathbb{R}} f(y;\theta,\phi) d\nu(y) = \partial_\theta(1) = 0.\end{aligned}$$

On en déduit :

$$\mathbb{E}\left[\frac{Y - b'(\theta)}{a(\phi)}\right] = 0,$$

c'est-à-dire $\mathbb{E}[Y] = b'(\theta)$. De plus, on a :

$$\mathbb{E}\left[(\partial_\theta \log(f(Y;\theta,\phi)))^2\right] = -\mathbb{E}\left[\partial_\theta^2 \log(f(Y;\theta,\phi))\right].$$

En effet,

$$\begin{aligned}\mathbb{E}\left[\partial_\theta^2 \log(f(Y;\theta,\phi))\right] &= \mathbb{E}\left[\partial_\theta\left(\frac{\partial_\theta f(Y;\theta,\phi)}{f(Y;\theta,\phi)}\right)\right] \\ &= \mathbb{E}\left[\frac{\partial_\theta^2 f(Y;\theta,\phi) f(Y;\theta,\phi) - (\partial_\theta f(Y;\theta,\phi))^2}{f(Y;\theta,\phi)^2}\right] \\ &= \int_{\mathbb{R}} \partial_\theta^2 f(y;\theta,\phi) d\nu(y) - \mathbb{E}\left[\left(\frac{\partial_\theta f(Y;\theta,\phi)}{f(Y;\theta,\phi)}\right)^2\right] \\ &= \partial_\theta^2 \int_{\mathbb{R}} f(y;\theta,\phi) d\nu(y) - \mathbb{E}\left[(\partial_\theta \log(f(Y;\theta,\phi)))^2\right] \\ &= -\mathbb{E}\left[(\partial_\theta \log(f(Y;\theta,\phi)))^2\right].\end{aligned}$$

□

Définition 1.3 (Fonction de lien canonique)**.** *On suppose que $b' : \mathbb{R} \to A$ où $A \subset \mathbb{R}$ est bijective. On a $\theta = b'^{-1}(\mu)$. La fonction b'^{-1} est la fonction de lien canonique. Comme la relation est bijective, si on a $\theta = g(\mu)$ avec une fonction g, alors $g = b'^{-1}$ est la fonction de lien canonique.*

La plupart des lois habituelles appartiennent à la famille exponentielle.

Exemple 1.4 (Loi de Bernoulli)**.** *La loi de Bernoulli de paramètre $p \in]0,1[$ fait parti de la famille exponentielle et sa densité (par rapport à la mesure $\delta_0 + \delta_1$) s'écrit :*

$$f(y;\theta,\phi) = (1-p)^{1-y} p^y = \exp\left(\frac{y \log\left(\frac{p}{1-p}\right) + \log(1-p)}{1} + 0\right).$$

On pose $\theta := \log\left(\frac{p}{1-p}\right)$ qui détermine la fonction de lien canonique. La densité se réécrit :

$$f(y;\theta,\phi) = \exp\left(\frac{y\theta - \log(1+e^\theta)}{1} + 0\right).$$

On identifie également $b(\theta) = \log(1+e^\theta)$, $a(\phi) = 1$ et $c(y,\phi) = 0$.

Exemple 1.5 (Loi de Poisson)**.** *La loi de Poisson de paramètre $\lambda > 0$ fait parti de la famille exponentielle et sa densité (par rapport à la mesure $\sum_{n\in\mathbb{N}} \delta_n$) s'écrit :*

$$f(y;\theta,\phi) = e^{-\lambda}\frac{\lambda^y}{y!} = \exp\left(y\log(\lambda) - \lambda - \log(y!)\right).$$

On pose $\theta := \log(\lambda)$ qui détermine la fonction de lien canonique. La densité se réécrit :

$$f(y;\theta,\phi) = \exp\left(y\theta - \exp(\theta) - \log(y!)\right).$$

On identifie également $b(\theta) = \exp(\theta)$, $a(\phi) = 1$ et $c(y,\phi) = -\log(y!)$. En pratique on utilise les fonctions de lien $g(\lambda) = \log(\lambda)$ et $g(\lambda) = \sqrt{\lambda}$.

Exemple 1.6 (Loi Binomiale Négative)**.** *La loi Binomiale Négative de paramètres $r > 0$ et $p \in]0,1[$ fait parti de la famille exponentielle et sa densité (par rapport à la mesure $\sum_{n\in\mathbb{N}} \delta_n$) s'écrit :*

$$f(y;\theta,\phi) = \frac{\Gamma(r+y)}{y!\Gamma(r)}p^r(1-p)^y.$$

Son espérance est $\mu := \frac{r(1-p)}{p}$. On paramétrise par (r,μ), la densité se réécrit :

$$\begin{aligned} f(y;\theta,\phi) &= \frac{\Gamma(r+y)}{y!\Gamma(r)}\left(\frac{r}{r+\mu}\right)^r\left(\frac{\mu}{r+\mu}\right)^y \\ &= \exp\left(y\log\left(\frac{\mu}{r+\mu}\right) + r\log\left(\frac{r}{r+\mu}\right) + \log\left(\frac{\Gamma(r+y)}{y!\Gamma(r)}\right)\right). \end{aligned}$$

On pose $\theta := \log\left(\frac{\mu}{r+\mu}\right)$ qui détermine la fonction de lien canonique. La densité se réécrit :

$$f(y;\theta,\phi) = \exp\left(y\theta + r\log\left(1-e^\theta\right) + \log\left(\frac{\Gamma(r+y)}{y!\Gamma(r)}\right)\right).$$

On identifie également $b(\theta) = -r\log\left(1-e^\theta\right)$, $a(\phi) = 1$ et $c(y,\phi) = \log\left(\frac{\Gamma(r+y)}{y!\Gamma(r)}\right)$. En pratique on utilise les fonctions de lien $g(\mu) = \log(\mu)$ et $g(\mu) = \sqrt{\mu}$.

Exemple 1.7 (Loi normale)**.** *La loi normale de paramètre $\mu \in \mathbb{R}$ et $\sigma^2 > 0$ fait parti de la famille exponentielle et sa densité (par rapport à la mesure de Lebesgue) s'écrit :*

$$f(y;\theta,\phi) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mu)^2}{2\sigma^2}} = \exp\left(\frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{y^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)\right).$$

On pose $\theta := \mu$ qui détermine la fonction de lien canonique. La densité se réécrit :

$$f(y;\theta,\phi) = \exp\left(\frac{y\theta - \frac{\theta^2}{2}}{\sigma^2} - \frac{y^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)\right).$$

On identifie également $b(\theta) = \frac{\theta^2}{2}$, $a(\phi) = \sigma^2$ et $c(y,\phi) = -\frac{y^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)$. On peut poser $\phi := \sigma^2$.

On suppose observer une variable aléatoire X à valeurs dans $\mathbb{R}^d$. On suppose de plus qu'il existe une fonction g bijective et $\beta_0 \in \mathbb{R}, \beta \in \mathbb{R}^d$ telle que :

$$g\left(\mathbb{E}(Y \mid X)\right) = \beta_0 + X'\beta.$$

La fonction g est appelée *fonction de lien*. Il est possible d'utiliser *la fonction de lien canonique* du modèle.

La quantité $\theta(X) := \beta_0 + X'\beta$ est appelée *le prédicteur linéaire*. Il vient :

$$\mathbb{E}(Y \mid X) = g^{-1}\left(\beta_0 + X'\beta\right).$$

La difficulté est d'estimer $(\beta_0, \beta) \in \mathbb{R}^{d+1}$.

1.1 Estimation par maximum de vraisemblance

On suppose que $Y \mid X$ suit la loi $\mathbb{P}_{\theta(X)}$. En particulier, si on observe $(y_i, x_i)_{1\leq i\leq n}$ variables aléatoires i.i.d., la vraisemblance, dans le cas de la fonction de lien canonique, s'écrit :

$$\begin{aligned}\ell\left(\beta_0, \beta; (y_i \mid x_i)_{1\leq i\leq n}\right) &= \sum_{i=1}^{n} \log\left(f(y_i; \beta_0 + x_i'\beta, \phi)\right) \\ &= \sum_{i=1}^{n} \frac{y_i(\beta_0 + x_i'\beta) - b(\beta_0 + x_i'\beta)}{a(\phi)} + c(y_i, \phi).\end{aligned}$$

On remarque que l'estimation par maximum de vraisemblance de (β_0, β) est indépendante de ϕ, qui pourra être estimé dans un second temps (c'est aussi vrai avec une autre fonction de lien).

Exemple 1.8 (Loi de Bernoulli)**.** *On obtient la log-vraisemblance :*

$$\ell\left(\beta_0, \beta; (y_i \mid x_i)_{1\leq i\leq n}\right) = \sum_{i=1}^{n} \log\left[\left(\frac{e^{\beta_0 + x_i'\beta}}{1 + e^{\beta_0 + x_i'\beta}}\right)^{y_i} \left(\frac{1}{1 + e^{\beta_0 + x_i'\beta}}\right)^{1-y_i}\right].$$

Exemple 1.9 (Loi de Poisson)**.** *On obtient la log-vraisemblance :*

$$\ell\left(\beta_0, \beta; (y_i \mid x_i)_{1\leq i\leq n}\right) \propto \sum_{i=1}^{n} \left(-e^{\beta_0 + x_i'\beta} + y_i\left(\beta_0 + x_i'\beta\right)\right).$$

Exemple 1.10 (Loi normale)**.** *On obtient la log-vraisemblance :*

$$\ell\left(\beta_0, \beta; (y_i \mid x_i)_{1\leq i\leq n}\right) \propto -\frac{n}{2} \log\left(\sigma^2\right) - \sum_{i=1}^{n} \frac{\left(y_i - (\beta_0 + x_i'\beta)\right)^2}{2\sigma^2}.$$

Les équations du premier ordre du problème :

$$(\widehat{\beta}_0, \widehat{\beta}) \in \underset{(\beta_0,\beta)\in\mathbb{R}^{d+1}}{\arg\max}\ \ell\left(\beta_0, \beta; (y_i \mid x_i)_{1\leq i\leq n}\right)$$

ne permettent pas d'en déduire, en général, une solution explicite (en dehors de cas particuliers, comme le modèle gaussien avec fonction de lien canonique qui permet de retrouver le modèle linéaire avec bruit gaussien). Nous passons par des méthodes numériques. Dans R, la fonction `glm` permet d'estimer des modèles linéaires généralisés. Il convient ensuite d'éliminer les facteurs non déterminants via des tests d'hypothèse.

1.2 Estimation de la dispersion

Dans certains modèles, la dispersion est fixée. Par exemple, dans le modèle avec une loi de Poisson, on a $a(\phi) = 1$. Dans ce cas, il est également intéressant de disposer d'un estimateur de la dispersion et s'assurer que cette propriété est bien vérifiée.

Dans un premier temps, nous définissons la *fonction de variance* du modèle $V : A \mapsto \mathbb{R}_+$. Il s'agit d'associer à la moyenne $\mu \in A$ la variance associée :

$$V(\mu) := \mathbb{E}\left[(Y-\mu)^2\right].$$

Par le Lemme 1.2, on a

$$V(\mu) = a(\phi)b''(b'^{-1}(\mu)).$$

La fonction de variance dépend de la famille de loi, mais pas de la fonction de lien choisie.

Exemple 1.11 (Loi de Bernoulli)**.** *On a* $b(\theta) = \log(1+e^\theta)$ *et* $a(\phi) = 1$. *On retrouve*

$$V(\mu) = \mu(1-\mu).$$

Exemple 1.12 (Loi de Poisson)**.** *On a* $b(\theta) = \exp(\theta)$ *et* $a(\phi) = 1$. *On retrouve*

$$V(\mu) = \mu.$$

Exemple 1.13 (Loi Binomiale Négative)**.** *On a* $b(\theta) = -r\log\left(1-e^\theta\right)$ *et* $a(\phi) = 1$. *On retrouve*

$$V(\mu) = \mu + \frac{\mu^2}{r}.$$

Exemple 1.14 (Loi normale)**.** *On a* $b(\theta) = \frac{\theta^2}{2}$ *et* $a(\phi) = \sigma^2$. *On retrouve*

$$V(\mu) = a(\phi) = \sigma^2.$$

On définit $V_\circ$ la fonction de variance normalisée par

$$V_\circ(\mu) := \frac{V(\mu)}{a(\phi)} = b''(b'^{-1}(\mu)).$$

Celle-ci ne dépend plus de la dispersion et est fonction uniquement de μ. En conséquence, on obtient la relation pour la dispersion :

$$a(\phi) = \mathbb{E}\left[\frac{(Y-\mu)^2}{V_\circ(\mu)}\right].$$

si on observe $(y_i, x_i)_{1\leq i\leq n}$ variables aléatoires i.i.d., un estimateur de la dispersion est

$$\widehat{a(\phi)} := \frac{1}{n-d-1}\sum_{i=1}^{n}\frac{(y_i-\widehat{\mu}_i)^2}{V_\circ(\widehat{\mu}_i)}.$$

Il est possible d'obtenir un intervalle de confiance asymptotique autour de la dispersion, ou d'effectuer un test asymptotique pour tester une valeur. La loi asymptotique de l'estimateur est :

$$\widehat{a(\phi)} \overset{loi}{\approx} a(\phi)\frac{\chi^2_{n-d-1}}{n-d-1}.$$

Exemple 1.15 (Loi de Poisson)**.** *Comme* $V_\circ(\mu) = V(\mu) = \mu$*,*

$$\widehat{a(\phi)} := \frac{1}{n-d-1}\sum_{i=1}^{n}\frac{(y_i - \widehat{\mu}_i)^2}{\widehat{\mu}_i}.$$

De plus, comme $a(\phi) = 1$*,*

$$\widehat{a(\phi)} \overset{loi}{\approx} \frac{\chi^2_{n-d-1}}{n-d-1}.$$

Nous pouvons tester si l'hypothèse H_0 *la dispersion est bien égale à 1 en vérifiant si* $\widehat{a(\phi)} \in \left[\frac{q^{\chi^2_{n-d-1}}_{\frac{\alpha}{2}}}{n-d-1}, \frac{q^{\chi^2_{n-d-1}}_{1-\frac{\alpha}{2}}}{n-d-1}\right]$ *pour un risque de première espèce* $\alpha \in]0,1[$*. Ou plus simplement, pour tester* H_0 *contre* $H_1 : a(\phi) > 1$*, la p-valeur du test est*

$$p := 1 - F_{\chi^2_{n-d-1}}\left((n-d-1)\widehat{a(\phi)}\right).$$

Si on rejette H_0 *car* $\widehat{a(\phi)}$ *est trop grand, on dit que le modèle est sur-dispersé, et s'il est trop petit, qu'il est sous-dispersé.*

Comme introduit dans l'exemple, la dispersion peut être contrainte par le modèle et, malgré cela, observer que l'estimation de $a(\phi)$ est significativement différente. Dans ce cas, on peut introduire un modèle surdispersé en mettant $a(\phi) = \phi$ libre. L'exemple fondamental est celui de la loi de Poisson. On fixe la fonction de variance :

$$V(\mu) = \phi\mu,$$

où ϕ est un paramètre à estimer. On introduit la quasi-densité :

$$f(y;\theta,\phi) = \exp\left(\frac{y\log(\lambda) - \lambda}{\phi} - \log(y!)\right).$$

Les estimations de (β_0, β) sont inchangées, mais les intervalles de confiance de ces paramètres sont élargis d'un facteur $\sqrt{\phi}$ ce qui change les seuils de significativité.

1.3 Estimation par maximum de vraisemblance pénalisé

Il est possible d'intégrer l'élimination des facteurs non déterminants au modèle en utilisant des méthodes de pénalisation. Soit $\lambda > 0$, on introduit :

$$\ell(\beta_0, \beta; (y_i \mid x_i)_{1\leq i\leq n}; \lambda) := \ell(\beta_0, \beta; (y_i \mid x_i)_{1\leq i\leq n}) - \lambda\|\beta\|_p^p,$$

où $\|\cdot\|_p$ est la norme $p \in \{1,2\}$ et λ est le facteur de pénalisation. Afin que le modèle soit invariant par facteur d'échelle dans l'estimation des paramètres, il est fondamental de centrer et réduire les variables explicatives non catégorielles x au préalable. Dans la maximisation, il y a un arbitrage entre l'apport dans la log-vraisemblance et la pénalisation.

Lorsque $p = 1$, on parle de pénalisation **Lasso** et lorsque $p = 2$ on parle de pénalisation **Ridge**. À $\lambda > 0$ fixé, on résout le problème :

$$(\widehat{\beta}_0, \widehat{\beta}) \in \underset{(\beta_0,\beta)\in\mathbb{R}^{d+1}}{\arg\max}\ \ell(\beta_0, \beta; (y_i \mid x_i)_{1\leq i\leq n}; \lambda).$$

La difficulté est alors de choisir un bon $\lambda > 0$. On procède par validation croisée. On découpe l'échantillon de manière aléatoire en $\kappa \geq 2$ parties de taille égale ou semblable. On peut choisir $\kappa = n$ le nombre d'observations, ce qui revient à découper individuellement. On note $\mathcal{P}_k$ la partie $k \in \{1, \ldots, \kappa\}$ et l'ensemble forme une partition de $\{1, \ldots, n\}$, c'est à dire, on a la réunion disjointe :

$$\bigcup_{k=1}^{\kappa} \mathcal{P}_k = \{1, \ldots, n\}.$$

Définition 1.16. *(Validation croisée) Soit $k \in \{1, \ldots, \kappa\}$ et $\lambda > 0$. L'estimation par validation croisée consiste à estimer le modèle sur l'échantillon observé $(y_i, x_i)_{i \notin \mathcal{P}_k}$. On en déduit $(\widehat{\beta}_0^{\lambda,k}, \widehat{\beta}^{\lambda,k})$. On calcule ensuite la log-vraisemblance, avec les paramètres estimés, sur l'échantillon de test :*

$$\mathfrak{l}_k(\lambda) := \sum_{i \in \mathcal{P}_k} \log\left(f(y_i; \widehat{\beta}_0^{\lambda,k} + x_i' \widehat{\beta}^{\lambda,k}, \phi) \right).$$

On introduit :

$$\mathfrak{l}(\lambda) := \sum_{k=1}^{\kappa} \mathfrak{l}_k(\lambda).$$

Le modèle sélectionné est celui qui résout le problème :

$$\widehat{\lambda} \in \arg\max_{\lambda > 0} \mathfrak{l}(\lambda).$$

Une fois $\widehat{\lambda}$ déterminé, on estime les paramètres sur l'ensemble des données :

$$(\widehat{\beta}_0, \widehat{\beta}) \in \arg\max_{(\beta_0, \beta) \in \mathbb{R}^{d+1}} \ell(\beta_0, \beta; (y_i \mid x_i)_{1 \leq i \leq n}; \widehat{\lambda}).$$

L'idéal est d'avoir une partition très fine : on estime sur un grand sous échantillon pour tester sur un petit échantillon. Mais plus la partition est fine, plus le temps de calcul est élevé, il y a souvent un arbitrage à faire. En pratique, le minimum étant d'avoir au moins les deux tiers de l'échantillon afin d'estimer les paramètres, et l'idéal étant d'avoir tout l'échantillon sauf un point qui servira de test.

La pénalisation permet d'avoir un modèle bien posé même si le nombre de variables explicatives dépasse le nombre d'observations. Dans ce cas, beaucoup de paramètres seront proches de 0 (pénalisation **Ridge**) ou égaux à 0 (pénalisation **Lasso**).

1.4 La prime pure

On note $(\widehat{\beta}_0^N, \widehat{\beta}^N)$ les paramètres estimés du modèle pour le nombre de sinistres et g_N sa fonction de lien. On note $(\widehat{\beta}_0^Y, \widehat{\beta}^Y)$ les paramètres estimés du modèle pour le coût des sinistres et g_Y sa fonction de lien. La prime pure estimée pour l'assuré i de variables explicatives x_i, notée $\widehat{P}_i$, est :

$$\widehat{P}_i := g_N^{-1}(\widehat{\beta}_0^N + x_i' \widehat{\beta}_0^N) g_Y^{-1}(\widehat{\beta}_0^Y + x_i' \widehat{\beta}_0^Y).$$

1.5 Exemple

Nous prenons un exemple avec des données simulées.

```
set.seed(0)
n <- 2*10^4
X <- data.frame(age = runif(n, 18, 100), rouge = rbinom(n, 1, .2))
X$jeune <- as.integer((X$age - runif(n, 18, X$age)) <= 3)

theta_N <- -2 - 0.03*X$age + 0.8*X$jeune + 0.6*rnorm(n)
mu_N <- exp(theta_N)

X$N <- rpois(n, mu_N)
```

La fonction qui permet d'estimer un modèle linéaire généralisé est `glm`.

```
(reg <- glm(N ~ . , data = X, family = poisson(link="log")))
summary(reg)
```

La famille de loi ainsi que sa fonction de lien sont renseignés avec l'argument `family`. Nous pouvons ensuite récupérer les coefficients (β_0, β) avec `reg$coefficients`, les valeurs de $(\widehat{\mu}_i)_{1\leq i\leq n}$ avec `reg$fitted.value`. Nous pouvons en déduire la dispersion et calculer la p-valeur du test.

```
mu_hat <- reg$fitted.values
dispersion <- sum((X$N - mu_hat)^2 / mu_hat) / reg$df.residual

pvaleur <- 1-pchisq(reg$df.residual*dispersion, reg$df.residual)
```

On observe une surdispersion, provoquée par le bruit `0.6*rnorm(n)`, on ne l'observe plus en le supprimant.

```
(reg_surdispersion <- glm(N ~ . , data = X, family = quasipoisson(
   link="log")))
summary(reg_surdispersion)
```

Pour utiliser la famille de Poisson en tenant compte de la surdispersion, on met en famille de loi `quasipoisson` au lieu de `poisson`. Pour utiliser une méthode de pénalisation et utiliser la validation croisée, on utilise la bibliothèque `glmnet`. On modifie légèrement le format des données pour pouvoir utiliser la fonction `glmnet`.

```
#Le package doit être installé au préalable
library(glmnet)
Xnet <- X; Xnet$N <- NULL
Xnet <- as.matrix(Xnet)

#Régressions pénalisées pour plusieurs lambda
reg_lasso <- glmnet(Xnet, X$N, family = poisson(link="log"))

#Affichage de l'évolution des coefficients en fonction de la pé
   nalisation
plot(reg_lasso)
```

Cette fonction ne réalise pas la validation croisée : elle estime les paramètres pour différents λ. Et nous pouvons afficher la valeur des paramètres en fonction de la pénalisation. Pour réaliser la validation croisée, nous passons par `cv.glmnet`.

```
#Validation croisée et sélection du lambda
reg_cv <- cv.glmnet(Xnet, X$N, family = poisson(link="log"), nfolds
   = 20)
```

Enfin, nous pouvons récupérer les coefficients estimés.

```
coef(reg_cv)
```

2 Tarification a posteriori

Malgré la différentiation des individus à travers leurs variables explicatives $x_i \in \mathbb{R}^d$, ceux-ci ont des caractéristiques propres inobservées. Ainsi, même si l'hypothèse que le nombre de sinistres de chaque individu survient à travers une loi de Poisson d'un paramètre qui lui est propre semble raisonnable, on observe souvent une variance supérieure à la moyenne. Cela se traduit par de la surdispertion dans un modèle linéaire généralisé avec la loi de Poisson ($a(\phi) > 1$). Une telle propriété apparaît naturellement en conséquence de caractéristiques individuelles inobservables. En effet, supposons que :

$$N_i \mid \{X_i = x_i, \Theta_i = \theta_i\} \overset{i.i.d.}{\sim} \mathcal{P}\left(\lambda(x_i, \theta_i)\right),$$

où seul x_i est observé, la caractérisque θ_i est inconnue et également propre à l'assuré. λ est ici une fonction à valeurs dans $\mathbb{R}_+$ Dans ce cas, l'espérance de N_i sachant x_i est :

$$\mathbb{E}\left[N_i \mid X_i = x_i\right] = \mathbb{E}\left[\mathbb{E}(N_i \mid X_i = x_i, \Theta_i)\right] = \mathbb{E}\left[\lambda(x_i, \Theta_i)\right].$$

Tandis que, par la formule de la variance totale, on a :

$$\begin{aligned} Var(N_i \mid X_i = x_i) &= \mathbb{E}\left[Var(N_i \mid X_i = x_i, \Theta_i)\right] + Var\left[\mathbb{E}(N_i \mid X_i = x_i, \Theta_i)\right] \\ &= \mathbb{E}\left[\lambda(x_i, \Theta_i)\right] + Var\left[\lambda(x_i, \Theta_i)\right] \end{aligned}$$

Si $\lambda(x_i, \Theta_i)$ n'est pas une variable aléatoire p.s. constante, on a $Var(N_i \mid X_i = x_i) > \mathbb{E}\left[N_i \mid X_i = x_i\right]$. Ce phénomène créer de la surdispersion. Un cas intéressant est lorsque $\lambda(x_i, \Theta_i)$ suit une loi Gamma. Dans ce cas la loi du nombre de sinistres de l'assuré pris individuellement suit une loi binomiale négative dont la variance est supérieure à la moyenne.

Lemme 2.1. *Soit $\Theta \sim \mathcal{G}(\alpha, \beta)$ et $N \mid \{\Theta = \theta\} \sim \mathcal{P}(\theta)$ avec $\alpha > 0, \beta > 0, \theta > 0$. Alors N suit une loi Binomiale Négative de paramètres*

$$\begin{aligned} r &= \alpha, \\ p &= \frac{\beta}{1+\beta}. \end{aligned}$$

Démonstration. Soit $n \in \mathbb{N}$. On a :

$$\begin{aligned}
\mathbb{P}(N = n) &= \mathbb{E}\left[\mathbb{E}\left(\mathbf{1}_{\{n\}}(N) \mid \Theta\right)\right] \\
&= \mathbb{E}\left[\exp(-\Theta)\frac{\Theta^n}{n!}\right] \\
&= \int_{\mathbb{R}_+} \exp(-\theta)\frac{\theta^n}{n!}\frac{\beta^\alpha}{\Gamma(\alpha)}\exp(-\beta\theta)\theta^{\alpha-1}d\theta \\
&= \frac{\beta^\alpha}{n!\Gamma(\alpha)}\int_{\mathbb{R}_+} \exp(-(1+\beta)\theta)\theta^{n+\alpha-1}d\theta \\
&\underset{[t=(1+\beta)\theta]}{=} \frac{\beta^\alpha}{n!\Gamma(\alpha)(1+\beta)^{n+\alpha}}\int_{\mathbb{R}_+} \exp(-t)t^{n+\alpha-1}dt \\
&= \frac{\Gamma(\alpha+n)}{n!\Gamma(\alpha)}\left(\frac{\beta}{1+\beta}\right)^\alpha\left(\frac{1}{1+\beta}\right)^n.
\end{aligned}$$

□

Toutefois, l'assureur va pouvoir observer au fil de l'eau le nombre de sinistres de l'assuré et leurs coûts. Cela va lui permettre d'obtenir de l'information sur la variable aléatoire inconnue Θ_i et sa sinistralité réelle.

2.1 Prime Baysienne

On suppose avoir à notre disposition la prime pure annuelle a priori de chaque assuré i :

$$\widehat{P}_i := g_N^{-1}(\widehat{\beta}_0^N + x_i'\widehat{\beta}_0^N)g_Y^{-1}(\widehat{\beta}_0^Y + x_i'\widehat{\beta}_0^Y) := \widehat{P}_i^N\widehat{P}_i^Y.$$

Toutefois, chaque assuré aura des caractéristiques individuelles inconnues. On suppose que son nombre de sinistres N_t^i sur l'intervalle de temps $[0,t]$ avec $t > 0$ suit une loi de Poisson de paramètre $\Theta_i t$ avec $\Theta_i \sim \mathcal{G}(\alpha_i, \beta_i)$. Plus précisement, on suppose que $(N_t^i)_{t\geq 0} \mid \{\Theta_i = \theta_i\}$ est un processus de Poisson de paramètre θ_i inconnu. Les paramètres (α_i, β_i) doivent être pris tels que

$$\mathbb{E}(N_1^i) = \mathbb{E}(\Theta_i) = \frac{\alpha_i}{\beta_i} = \widehat{P}_i^N.$$

En effet, sur l'intervalle [0, 1], $\mathbb{E}(N_1^i) = \mathbb{E}\left[\mathbb{E}(N_1^i \mid \Theta)\right] = \mathbb{E}\left[\Theta_i\right] = \frac{\alpha_i}{\beta_i}$ ce qui permet d'être consistant avec la définition de la prime pure. Pour fixer le second paramètre, dans le modèle linéaire généralisé, on a à notre disposition la fonction de variance $V : \mu \mapsto V(\mu)$, ce qui donne la relation :

$$Var(N_1^i) = \mathbb{E}(\Theta_i) + Var(\Theta_i) = \frac{\alpha_i}{\beta_i} + \frac{\alpha_i}{\beta_i^2} = V\left(\widehat{P}_i^N\right).$$

Nous avons la caractérisation :

$$\begin{cases} \frac{\alpha_i}{\beta_i} &= \widehat{P}_i^N, \\ \widehat{P}_i^N + \frac{\widehat{P}_i^N}{\beta_i} &= V\left(\widehat{P}_i^N\right). \end{cases}$$

qui se réécrit

$$\begin{cases} \alpha_i &= \frac{\left(\widehat{P}_i^N\right)^2}{\widehat{P}_i^N - V\left(\widehat{P}_i^N\right)}, \\ \beta_i &= \frac{\widehat{P}_i^N}{\widehat{P}_i^N - V\left(\widehat{P}_i^N\right)}. \end{cases}$$

Exemple 2.2 (Loi de Poisson)**.** *Dans le modèle surdispersé, on a* $V(\mu) = \phi\mu$ *avec* $\phi > 1$. *Ainsi,*

$$\widehat{P}_i^N + \frac{\widehat{P}_i^N}{\beta_i} = \widehat{\phi}\widehat{P}_i^N \iff \beta_i = \frac{1}{\widehat{\phi} - 1}.$$

On peut donc poser $\beta_i = \beta_0$ *commun pour tout* i.

Exemple 2.3 (Loi Binomiale Négative)**.** *Dans ce modèle, consitant avec cette approche (car la loi de* N_1^i *est une loi binomiale négative), on a* $V(\mu) = \mu + \frac{\mu^2}{r}$ *et on déduit*

$$\widehat{P}_i^N + \frac{\widehat{P}_i^N}{\beta_i} = \widehat{P}_i^N + \frac{(\widehat{P}_i^N)^2}{r} \iff \beta_i = \frac{r}{\widehat{P}_i^N}.$$

Cette fois-ci, on peut poser $\alpha_i = \alpha_0$ *qui est commun pour tout* i *avec* $\alpha_0 = r$.

Lemme 2.4. *On a, pour tout* $n \geq 0$ *et* $t \geq 0$ *:*

$$\Theta_i \mid \{N_t^i = n\} \sim \mathcal{G}\left(\alpha_i + n, \beta_i + t\right).$$

Démonstration. Il s'agit d'un calcul de loi a posteriori. La loi $\Theta \mid \{N_t^i = n\}$ est par construction absolument continue par rapport à la mesure de Lebesgue, pour $\theta > 0$, on a presque partout :

$$\begin{aligned} f(\theta \mid N_t^i = n) &\propto f(n \mid \Theta_i = \theta) f(\theta) \\ &\propto e^{-\theta t}\frac{(\theta t)^n}{n!}\frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)}e^{-\beta_i\theta}\theta^{\alpha_i - 1} \\ &\propto e^{-(\beta_i + t)\theta}\theta^{\alpha_i + n - 1}. \end{aligned}$$

On reconnait la loi $\mathcal{G}\left(\alpha_i + n, \beta_i + t\right)$.

□

Corollaire 2.5. *La prime pure conditionnellement à l'observation* $\{N_t = n\}$ *est :*

$$\widetilde{P}_i := \mathbb{E}\left(N_{t+1}^i - N_t^i \mid N_t^i = n\right)\widehat{P}_i^Y = \left[(1 - z_t^i)\widehat{P}_i^N + z_t^i\frac{n}{t}\right]\widehat{P}_i^Y.$$

avec

$$z_t^i := \frac{t}{\beta_i + t}, \quad t \geq 0.$$

Démonstration. On a :

$$\begin{aligned} \mathbb{E}\left(N_{t+1}^i - N_t^i \mid N_t^i = n\right) &= \mathbb{E}\left[\mathbb{E}\left(N_{t+1}^i - N_t^i \mid N_t^i = n, \Theta_i\right) \mid N_t^i = n\right] \\ &= \mathbb{E}\left[\mathbb{E}\left(N_{t+1}^i - N_t^i \mid \Theta_i\right) \mid N_t^i = n\right] \\ &= \mathbb{E}\left(\Theta_i \mid N_t^i = n\right) = \frac{\alpha_i + n}{\beta_i + t} = \frac{\beta_i}{\beta_i + t}\frac{\alpha_i}{\beta_i} + \frac{t}{\beta_i + t}\frac{n}{t}, \end{aligned}$$

et on en déduit le résultat. □

```
# Cas Poisson sur-dispersé, suite du MLG simulé
# Calcul des paramètres initiaux de la loi Gamma
beta ← 1/(dispersion-1)
alpha ← beta * mu_hat

# Si tous les assurés sont observés 1 an avec la sinistralité X$N
beta_1 ← beta + 1
alpha_1 ← alpha + X$N
mu_tilde ← alpha_1/beta_1
```

2.2 Crédibilité non paramétrique

Le cadre bayésien exige une hypothèse de loi sur le paramètre inconnu. La prime qui était obtenue précédemment correspondait à l'espérance conditionnelle. Par exemple, pour le nombre de sinistres, si ce dernier suit un processus $(N_t^i)_{t\geq 0}$, et si en $t \geq 0$, on observe $\{N_t^i = n\}$, la prime pure unitaire par sinistre était par construction :

$$\mathbb{E}(N_{t+1}^i - N_t^i \mid N_t^i = n).$$

Ici, nous allons prendre une forme linéaire en l'observation N_t, ce qui permettra d'aboutir à une prime unitaire de la forme :

$$\left[(1 - z_t^i)\widehat{P}_i^N + z_t^i \frac{n}{t}\right].$$

Nous travaillerons directement sur le coût total des sinistres, mais ce qui suit pourra s'appliquer sur le nombre de sinistres.

Hypothèse 2.1. *Soit $(S_t^i)_{t\geq 0}$ le processus du coût total des sinistres de l'individu $i \geq 1$ et $x_i \in \mathbb{R}^d$ ses caractéristiques. On suppose que, à $i \geq 1$ fixé, la loi de $(S_{t+1}^i - S_t^i) \mid \{X_i = x_i, \Theta_i = \theta_i\}$ est i.i.d. pour tout $t \in \mathbb{N}$ et ont la même loi que $S_1^i \mid \{X_i = x_i, \Theta_i = \theta_i\}$.*

On peut remplacer le processus du coût total des sinistres par celui du nombre des sinistres si on s'intéresse au nombre moyen de sinistres, ou par la suite du coût des sinistres si on s'intéresse au coût unitaire moyen des sinistres.

Pour une catégorie $x_i \in \mathbb{R}^d$ fixée, afin d'alléger les notations, on introduit l'opérateur $\mathbb{E}_{x_i} := \mathbb{E}\left[\cdot \mid \{X_i = x_i\}\right]$, c'est à dire l'espérance conditionnellement à $\{X_i = x_i\}$. On pose :

$$\begin{aligned} P\left(\theta_i\right) &:= \mathbb{E}_{x_i}\left[S_1^i \mid \Theta_i = \theta_i\right], \\ \overline{S}_t^i &:= \frac{S_t^i}{t}. \end{aligned}$$

Par construction, $\widehat{P}_i = \mathbb{E}_{x_i}\left[P\left(\Theta_i\right)\right] = \mathbb{E}_{x_i}\left[S_1^i\right]$ est la tarification a priori. La crédibilité non paramétrique consiste à résoudre le problème :

$$(a^\star, b^\star) \in \underset{(a,b)\in\mathbb{R}^2}{\arg\min}\ \mathbb{E}_{x_i}\left[\left(P\left(\Theta_i\right) - a - b\overline{S}_t^i\right)^2\right],$$

puis on en déduira la prime pure a posteriori :

$$P_i^\star = a^\star + b^\star \overline{S}_t^i.$$

Les conditions du premier ordre donnent :

$$\left\{\begin{aligned} -2\mathbb{E}_{x_i}\left[P\left(\Theta_i\right) - a^\star - b^\star\overline{S}_t^i\right] &= 0 \\ -2\mathbb{E}_{x_i}\left[\overline{S}_t^i\left(P\left(\Theta_i\right) - a^\star - b^\star\overline{S}_t^i\right)\right] &= 0 \end{aligned}\right. \iff \begin{aligned} \mathbb{E}_{x_i}\left[P\left(\Theta_i\right) - a^\star - b^\star\overline{S}_t^i\right] &= 0 \\ \mathbb{E}_{x_i}\left[\overline{S}_t^i\left(P\left(\Theta_i\right) - a^\star - b^\star\overline{S}_t^i\right)\right] &= 0 \end{aligned}$$

La première équation donne : $a^\star = \mathbb{E}_{x_i}\left[P\left(\Theta_i\right) - b^\star\overline{S}_t^i\right]$. En substituant dans $a^\star$ la seconde, on obtient :

$$\begin{aligned} &\mathbb{E}_{x_i}\left[\overline{S}_t^i P\left(\Theta_i\right)\right] - \mathbb{E}_{x_i}\left[P\left(\Theta_i\right) - b^\star\overline{S}_t^i\right]\mathbb{E}_{x_i}\left[\overline{S}_t^i\right] - b^\star\mathbb{E}_{x_i}\left[\left(\overline{S}_t^i\right)^2\right] = 0 \\ \iff\ &Cov_{x_i}\left[\overline{S}_t^i, P\left(\Theta_i\right)\right] - b^\star Var_{x_i}(\overline{S}_t^i) = 0. \end{aligned}$$

On retrouve les résultats de la régression linéaire simple, c'est-à-dire :

$$a^\star = \mathbb{E}_{x_i}\left[P\left(\Theta_i\right)\right] - b^\star \mathbb{E}_{x_i}\left[\overline{S}_t^i\right],$$
$$b^\star = \frac{Cov_{x_i}\left[\overline{S}_t^i, P\left(\Theta_i\right)\right]}{Var_{x_i}(\overline{S}_t^i)}.$$

Comme

$$\mathbb{E}_{x_i}\left[\overline{S}_t^i\right] = \mathbb{E}_{x_i}\left[S_1^i\right] = \mathbb{E}_{x_i}\left[P(\Theta_i)\right],$$

on a

$$a^\star = (1-b^\star)\mathbb{E}_{x_i}\left[P\left(\Theta_i\right)\right].$$

puis

$$P_i^\star = a^\star + b^\star \overline{S}_t^i = (1-b^\star)\mathbb{E}_{x_i}\left[P\left(\Theta_i\right)\right] + b^\star \overline{S}_t^i.$$

Nous voyons apparaître à nouveau une moyenne entre la prime pure a priori et la moyenne empirique. Le numérateur de $b^\star$ se réécrit :

$$\begin{aligned}
Cov_{x_i}\left[\overline{S}_t^i, P\left(\Theta_i\right)\right] &= \frac{1}{t}\sum_{j=1}^{t} Cov_{x_i}\left[S_j^i - S_{j-1}^i, P(\Theta_i)\right] = Cov_{x_i}\left[S_1^i, P(\Theta_i)\right] \\
&= \mathbb{E}_{x_i}\left[S_1^i P(\Theta_i)\right] - \mathbb{E}_{x_i}\left[S_1^i\right]\mathbb{E}_{x_i}\left[P(\Theta_i)\right] \\
&= \mathbb{E}_{x_i}\left[\mathbb{E}_{x_i}\left[S_1^i P(\Theta_i) \mid \Theta_i\right]\right] - \mathbb{E}_{x_i}\left[\mathbb{E}_{x_i}\left[S_1^i \mid \Theta_i\right]\right]\mathbb{E}_{x_i}\left[P(\Theta_i)\right] \\
&= \mathbb{E}_{x_i}\left[P(\Theta_i)\mathbb{E}_{x_i}\left[S_1^i \mid \Theta_i\right]\right] - \mathbb{E}_{x_i}\left[P(\Theta_i)\right]^2 \\
&= Var_{x_i}\left[P(\Theta_i)\right].
\end{aligned}$$

Le dénominateur de $b^\star$ se réécrit :

$$\begin{aligned}
Var_{x_i}\left(\overline{S}_t^i\right) &= \mathbb{E}_{x_i}\left[Var_{x_i}\left(\overline{S}_t^i \mid \Theta_i\right)\right] + Var_{x_i}\left[\mathbb{E}_{x_i}\left(\overline{S}_t^i \mid \Theta_i\right)\right] \\
&= \frac{\mathbb{E}_{x_i}\left[Var_{x_i}\left(S_1^i \mid \Theta_i\right)\right]}{t} + Var_{x_i}\left[P\left(\Theta_i\right)\right].
\end{aligned}$$

Puis $b^\star$ se réécrit finalement

$$b^\star = \frac{Var_{x_i}\left[P(\Theta_i)\right]}{\frac{\mathbb{E}_{x_i}\left[Var_{x_i}\left(S_1^i|\Theta_i\right)\right]}{t} + Var_{x_i}\left[P\left(\Theta_i\right)\right]} = \frac{t}{t + \frac{\mathbb{E}_{x_i}\left[Var_{x_i}\left(S_1^i|\Theta_i\right)\right]}{Var_{x_i}\left[P(\Theta_i)\right]}} = \frac{t}{t + \frac{\mathbb{E}_{x_i}\left[Var_{x_i}\left(S_1^i|\Theta_i\right)\right]}{Var_{x_i}\left[\mathbb{E}_{x_i}\left(S_1^i|\Theta_i\right)\right]}}.$$

En posant $z_t^i := b^\star \in [0,1]$, on a :

$$P_i^\star = \left(1 - z_t^i\right)\mathbb{E}_{x_i}\left[P\left(\Theta_i\right)\right] + z_t^i \overline{S}_t^i.$$

Ci-dessus, par définition, $\widehat{P}_i$ est un estimateur de $\mathbb{E}_{x_i}\left[P\left(\Theta_i\right)\right]$. Dans le cas particulier du nombre de sinistres, cela s'écrit :

$$P_i^{\star N} = \left(1 - z_t^i\right)\widehat{P}_i^N + z_t^i \overline{N}_t^i.$$

Ceci est consistant avec le cas où $N_1 \mid \{\Theta_i = \theta\} \sim \mathcal{P}(\theta)$ et où $\Theta_i \sim \mathcal{G}(\alpha_i, \beta_i)$, on retrouve les résultats précédents :

$$z_t^i = \frac{t}{t + \frac{\mathbb{E}_{x_i}\left[Var_{x_i}\left(N_1^i|\Theta_i\right)\right]}{Var_{x_i}\left[\mathbb{E}_{x_i}\left(N_1^i|\Theta_i\right)\right]}} = \frac{t}{t + \frac{\mathbb{E}_{x_i}[\Theta_i]}{Var_{x_i}[\Theta_i]}} = \frac{t}{t + \beta_i}.$$

Deuxième partie

Sinitres tardifs

En assurance, les contrats couvrent les sinitres sur une année. L'assureur encaisse la prime immédiatement et est amené à payer par après les éventuels sinistres. Toutefois, à la fin des contrats, l'assureur ne sait pas exactement ce qu'il devra payer en conséquence de deux cas :

- Certains sinistres n'ont pas un coût final complètement déterminé,
- Certains sinistres ont eu lieu durant la période de couverture mais ne sont pas encore connus.

Ces deux cas amènent l'assureur à devoir estimer le coût réel de ses garanties passées échues car il ne peut pas connaître avec exactitude leur valeur. Ce sont les sinistres tardifs qu'on appelle en anglais les IBNR (Incurred But Not Reported), sachant que dans le premier cas le sinistre est connu, seul le montant est incertain.

Pour les contrats signés il y a $n \in \mathbb{N}^*$ années, l'assureur a pu observer le flux d'information pendant n années, ce qui lui permet :

- D'avoir une bonne connaissance des sinistres tardifs,
- De connaître plus précisément le coût des années précédentes.

3 Les triangles de liquidation

On se place sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On définit les variables aléatoire réelles $X_{i,j}$ pour $1 \leq i, j \leq n$ où $n \in \mathbb{N}^*$ est un nombre d'années.

- i correspond à l'année de survenance du sinistre, il s'agit de l'année de couverture du contrat
- j correspond à l'année de développement du sinistre, il s'agit de l'année à laquelle l'information arrive à l'assureur, relativement à i (par exemple, $j = 1$ correspond à l'année de couverture du contrat, $j = 2$ à un an après la couverture du contrat)
- $X_{i,j}$ est la variable aléatoire qui correspond au coût des sinistres survenus l'année i pour l'année de développement j.

La v.a. $X_{i,j}$ représente :

- Si $j = 1$, le coût estimé des sinistres survenus en l'année i et connus à la fin de l'année i,
- Si $j > 1$, la variation du coût estimé des sinistres survenus en l'année i et appris à la j-ème année après la survenance, c'est à dire en fin d'année $i + j - 1$,

Par contruction, en une année $k \in \mathbb{N}^*$, on observe les variables aléatoires $X_{i,j}$ telles que $i + j \leq k + 1$. On introduit la filtration :

$$\mathcal{F}_k := \sigma(X_{i,j} \mid i + j \leq k + 1).$$

Aujourd'hui, en année n, les variables observées sont celles qui sont $\mathcal{F}_n$-mesurables, c'est à dire les $X_{i,j}$ telles que $i + j \leq n + 1$ qui sont représentées par le triangle :

i,j	1	$\cdots$	n
1	$X_{1,1}$	$\cdots$	$X_{1,n}$
$\vdots$	$\vdots$	$\iddots$	
n	$X_{n,1}$		

On définit, pour $1 \leq i,j \leq n$,

$$C_{i,j} = \sum_{k=1}^{j} X_{i,k}.$$

Il s'agit de la somme cumulée et par construction, $C_{i,j}$ est $\mathcal{F}_k$-mesurable si $i+j \leq k+1$, comme l'est $X_{i,j}$. Il s'agit du coût total des sinistres survenus l'année i et vus en l'année de développement j. Si on suppose que les sinistres se développent en maximum n années, alors $C_{i,n}$ est le coût total des sinistres survenus pendant l'année i.

Nous avons le triangle cumulé :

i,j	1	$\cdots$	n
1	$C_{1,1}$	$\cdots$	$C_{1,n}$
$\vdots$	$\vdots$	$\iddots$	
n	$C_{n,1}$		

L'objectif principal est alors d'estimer la partie inconnue du triangle et d'en déduire une approximation du montant restant à payer pour les contrats passés. C'est-à-dire estimer $(C_{i,n})_{2\leq i \leq n}$.

3.1 L'exposition

L'exposition correspond à une mesure du risque du portefeuille.

On note $E_i \in \mathbb{R}_+$ l'exposition de l'année i. Plus l'exposition sera élevée, plus l'assureur aura souscrit de contrats et sera donc exposé à des sinistres tardifs.

Si tous les contrats sont identiques, l'exposition est le nombre de contrats en cours l'année i. Le montant total des primes est une mesure de l'exposition qui tient compte de la différence des risques des différents contrats.

Certains contrats sont signés en cours d'année, d'autres s'arrêtent en cours d'année. Il est alors possible de définir une exposition individuelle $(e_{k,i})_{k\geq 1}$ d'un contrat pour une année i et un individu k. Si le contrat a couvert une proportion $p_{k,i} \in [0,1]$ de l'année i, l'exposition associée au contrat (dans le cas où on compte le nombre de contrats) est :

$$e_{i,k} := p_{i,k}.$$

Si m est le nombre total de contrats, l'exposition totale de l'année i est :

$$E_i := \sum_{k=1}^{m} e_{i,k}.$$

L'exposition est une notion importante. En effet, si l'assureur a un nombre de contrats supérieur l'année $i+1$ par rapport à l'année i, on s'attend observer en moyenne des valeurs plus élevées pour $(X_{i,j})_{j\geq 1}$.

3.2 Risque de réserve

Le premier objectif est d'être capable d'estimer la partie inconnue du triangle en l'année n, c'est à dire les $X_{i,j}$ pour $i+j>n$. Cela permet d'estimer le coût total $(C_{i,n})_{1\leq i\leq n}$ de chaque année. Le montant à prévoir pour les évolution futur est $C_{i,n}-C_{i,n+1-i}$. La réserve est alors :

$$R := \sum_{i=1}^{n}(C_{i,n}-C_{i,n+1-i})$$

C'est une variable aléatoire pour laquelle nous mettrons en place des techniques afin d'estimer la moyenne compte tenu de l'information disponible, ainsi que le risque associé (car même si nous étions capable d'en déduire la moyenne exacte vu de n, il ne faut pas oublier qu'il s'agit d'une variable aléatoire).

3.3 Risque de souscription

Le risque de souscription pour la nouvelle année $n+1$ est la variable aléatoire $C_{n+1,n}$. Sachant l'exposition E_{n+1}, on peut estimer estimer le risque associé à la souscription pour l'année $n+1$ en s'intéressant à la distribution de $C_{n+1,n}$.

4 Modèles de Chain Ladder - Mack

Cette section s'appuie principalement sur [2].

4.1 Estimation des provisions

Il s'agit du modèle le plus utilisé en assurance. Il repose sur les hypothèses suivantes :

Hypothèse 4.1.

H1 Les v.a. $(C_{i_1,j})_{1\leq j\leq n}$ *et* $(C_{i_2,j})_{1\leq j\leq n}$ *sont indépendantes pour* $i_1\neq i_2$.

H2 Pour $1\leq j\leq n-1$, *il existe* $f_j\geq 0$ *tel que*

$$\mathbb{E}(C_{i,j+1}\mid \mathcal{F}_{i+j-1}) = f_j C_{i,j}, \quad 1\leq i\leq n.$$

H3 Pour $1\leq j\leq n-1$, *il existe* $\sigma_j\geq 0$ *tel que*

$$Var(C_{i,j+1}\mid \mathcal{F}_{i+j-1}) = \sigma_j^2 C_{i,j}, \quad 1\leq i\leq n.$$

Nous faisons donc trois hypothèses. La première, c'est que les années de survenance des sinistres sont indépendantes. La deuxième est que la variation du montant à payer, en moyenne pour l'année i lors du développement $j+1$, est proportionnelle à l'année précédente j. Enfin, la variance est proportionnelle à $C_{i,j}$ ce qui implicitement revient à avoir un grand portefeuille diversifié.

Pour $1\leq k\leq n$, on introduit la filtration $\mathcal{B}_k := \sigma(C_{i,j}\mid i+j\leq n+1, j\leq k)\subset \mathcal{F}_n$.

Remarque 4.1. *Sous H1, H2 et H3 vérifient également :*

$$\mathbb{E}(C_{i,j+1} \mid \mathcal{B}_j) = f_j C_{i,j},$$
$$Var(C_{i,j+1} \mid \mathcal{B}_j) = \sigma_j^2 C_{i,j}.$$

Démonstration. Par *H1*, à i fixé, les seules v.a. qui interviennent dans le conditionnement par $\mathcal{F}_{i+j-1}$ sont $(C_{i,k})_{1\leq k\leq j}$. Par cette même hypothèse, ce sont les seules variables qui interviennent également dans le conditionnement par $\mathcal{B}_j$ ce qui permet d'en déduire le résultat. □

On peut poser $\widehat{f}_{i,j} := \frac{C_{i,j+1}}{C_{i,j}}$ pour $i + j + 1 \leq n + 1$, i.e. pour les cas observables en année n.
Un estimateur de f_j est :

$$\widehat{f}_j := \frac{1}{n-j}\sum_{i=1}^{n-j}\widehat{f}_{i,j}.$$

Toutefois, il ne s'agit pas du meilleur estimateur linéaire de f_j avec les observations $\widehat{f}_{i,j}$ au sens de l'erreur quadratique moyenne.

Avec les poids positifs $p := (p_{i,j})$ (et non tous nuls, à déterminer ultérieurement), on redéfinit :

$$\widehat{f}_j := \frac{\sum_{i=1}^{n-j} p_{i,j}\widehat{f}_{i,j}}{\sum_{i=1}^{n-j} p_{i,j}}.$$

Lemme 4.2. *Pour tout* $1 \leq j \leq n-1$, *on a :*

$$\mathbb{E}\left(\widehat{f}_{i,j} \mid \mathcal{B}_j\right) = f_j.$$

Démonstration. Par *H2* :

$$\mathbb{E}\left(\widehat{f}_{i,j} \mid \mathcal{B}_j\right) = \frac{1}{C_{i,j}}\mathbb{E}\left(C_{i,j+1} \mid \mathcal{B}_j\right) \overset{H2}{=} f_j.$$

□

Proposition 4.3. *Sous* $H2$, *pour tout* $1 \leq j \leq n-1$, *pour tous les poids* p *qui sont* $\mathcal{B}_j$*-mesurable, l'estimateur* $\widehat{f}_j$ *est un estimateur sans biais de* f_j *conditionnellement à l'information disponible* $\mathcal{B}_j$, *c'est-à-dire :*

$$\mathbb{E}\left(\widehat{f}_j \mid \mathcal{B}_j\right) = f_j.$$

En particulier, il est sans biais :

$$\mathbb{E}\left(\widehat{f}_j\right) = f_j.$$

Démonstration.

$$\mathbb{E}(\widehat{f}_j \mid \mathcal{B}_j) = \mathbb{E}\left(\frac{\sum_{i=1}^{n-j} p_{i,j}\widehat{f}_{i,j}}{\sum_{i=1}^{n-j} p_{i,j}} \mid \mathcal{B}_j\right) = \frac{\sum_{i=1}^{n-j} p_{i,j}\mathbb{E}\left(\widehat{f}_{i,j} \mid \mathcal{B}_j\right)}{\sum_{i=1}^{n-j} p_{i,j}} = f_j.$$

□

Il reste à choisir les poids optimaux. On va s'appuyer sur le résultat suivant.

Lemme 4.4. *Soit $\mathcal{B}$ une tribu. Soient $(X_1, \ldots, X_n)$ n variables aléatoires indépendantes de L^2 de variance conditionnelle $Var(X_i \mid \mathcal{B}) = \sigma_i^2$. Soit :*

$$\widehat{\theta}_n := \sum_{i=1}^{n} p_i X_i,$$

avec $\sum_{i=1}^{n} p_i = 1$ qui sont $\mathcal{B}$-mesurables. Le choix des $(p_i)_{1 \leq i \leq n}$ qui minimisent $Var(\widehat{\theta}_n \mid \mathcal{B})$ est :

$$p_i \propto \frac{1}{\sigma_i^2}.$$

Démonstration. On souhaite minimiser sur p, on introduit le langrangien :

$$Var(\widehat{\theta}_n \mid \mathcal{B}) + \lambda \left(\sum_{i=1}^{n} p_i - 1 \right) = \sum_{i=1}^{n} p_i^2 \sigma_i^2 + \lambda \left(\sum_{i=1}^{n} p_i - 1 \right).$$

La condition du premier ordre donne :

$$2 p_i \sigma_i^2 + \lambda = 0 \Longleftrightarrow p_i = -\frac{\lambda}{2\sigma_i^2}, \quad 1 \leq i \leq n,$$

c'est-à-dire que :

$$p_i \propto \frac{1}{\sigma_i^2}, \quad 1 \leq i \leq n.$$

On vérifie aisément qu'il s'agit bien d'un minimum. □

Proposition 4.5. *Sous H1 - H2 - H3, le meilleur estimateur linéaire de f_j en $(\widehat{f}_{i,j})_{1 \leq i \leq n-j}$ pour l'erreur quadratique moyenne conditionellement à $\mathcal{B}_j$ est*

$$\widehat{f}_j := \frac{\sum_{i=1}^{n-j} C_{i,j} \widehat{f}_{i,j}}{\sum_{i=1}^{n-j} C_{i,j}},$$

c'est-à-dire :

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

Démonstration. Pour $1 \leq j \leq n-1$ fixé, l'erreur quadratique moyenne est :

$$\mathbb{E}\left[(\widehat{f}_j - f_j)^2 \mid \mathcal{B}_j \right] = \left[(\mathbb{E}(\widehat{f}_j \mid \mathcal{B}_j) - f_j) \right]^2 + Var\left[\widehat{f}_j \mid \mathcal{B}_j \right].$$

Comme l'estimateur est sans biais, cela revient à minimiser la variance. Par *H3* :

$$Var(\widehat{f}_{i,j} \mid \mathcal{B}_j) = \frac{\sigma_j^2}{C_{i,j}}, \quad 1 \leq i \leq n-j.$$

Par le Lemme 4.4, on en déduit que :

$$p_{i,j} \propto C_{i,j}, \quad 1 \leq i \leq n-j.$$

□

On peut en déduire la variance des $\widehat{f}_j$.

Lemme 4.6. *Pour tout $1 \leq j \leq n-1$, on a*

$$Var(\widehat{f}_j \mid \mathcal{B}_j) = \frac{\sigma_j^2}{\sum_{i=1}^{n-j} C_{i,j}}.$$

Démonstration. Par *H1* et *H3*, pour tout $1 \leq j \leq n-1$,

$$Var(\widehat{f}_j \mid \mathcal{B}_j) \overset{H1}{=} \frac{\sum_{i=1}^{n-j} Var(C_{i,j+1} \mid \mathcal{B}_j)}{\left(\sum_{i=1}^{n-j} C_{i,j}\right)^2} \overset{H3}{=} \frac{\sum_{i=1}^{n-j} \sigma_j^2 C_{i,j}}{\left(\sum_{i=1}^{n-j} C_{i,j}\right)^2} = \frac{\sigma_j^2}{\sum_{i=1}^{n-j} C_{i,j}}.$$

□

Le lemme suivant implique en particulier que les estimateurs $\widehat{f}_j$ ne sont pas corrélés entre eux.

Lemme 4.7. *Soient $1 \leq k \leq n-1$ et $j_1 < j_2 < \ldots < j_k$, on a*

$$\mathbb{E}\left(\prod_{i=1}^{k} \widehat{f}_{j_i} \;\middle|\; \mathcal{B}_{j_1}\right) = \prod_{i=1}^{k} \mathbb{E}(\widehat{f}_{j_i}) = \prod_{i=1}^{k} f_{j_i}.$$

Démonstration. Pour $k=1$ la propriété est vraie puisque $\mathbb{E}(\widehat{f}_{j_1} \mid \mathcal{B}_{j_1}) = f_{j_1}$ par le lemme 4.3. Puis, pour $k \geq 2$,

$$\mathbb{E}\left[\prod_{i=1}^{k} \widehat{f}_{j_i} \;\middle|\; \mathcal{B}_{j_1}\right] = \mathbb{E}\left[\mathbb{E}\left(\prod_{i=1}^{k} \widehat{f}_{j_i} \;\middle|\; \mathcal{B}_{j_k}\right) \;\middle|\; \mathcal{B}_{j_1}\right] = \mathbb{E}\left[\left(\prod_{i=1}^{k-1} \widehat{f}_{j_i}\right) \mathbb{E}(\widehat{f}_{j_k} \mid \mathcal{B}_{j_k}) \;\middle|\; \mathcal{B}_{j_1}\right] = \mathbb{E}\left[\prod_{i=1}^{k-1} \widehat{f}_{j_i} \;\middle|\; \mathcal{B}_{j_1}\right] f_{j_k},$$

ce qui permet de conclure par récurrence. □

On a le lemme suivant :

Lemme 4.8. *Soient $1 \leq i \leq n$ et $n+1-i \leq j \leq n$.*

$$\mathbb{E}(C_{i,j} \mid \mathcal{F}_n) = \left(\prod_{k=n+1-i}^{j-1} f_k\right) C_{i,n+1-i}.$$

Démonstration. Pour $j = n+1-i$, on est sur la diagonale et $\mathbb{E}(C_{i,n+1-i} \mid \mathcal{F}_n) = C_{i,n+1-i}$. Soit $n+2-i \leq j \leq n$. On a $i+j-2 \geq n$. En utilisant la formule des espérances conditionnelles emboitées, par *H2*, on a

$$\mathbb{E}(C_{i,j} \mid \mathcal{F}_n) = \mathbb{E}(\mathbb{E}(C_{i,j} \mid \mathcal{F}_{i+j-2}) \mid \mathcal{F}_n) = f_{j-1}\mathbb{E}(C_{i,j-1} \mid \mathcal{F}_n).$$

Puis, par récurrence, on en déduit le résultat. □

Proposition 4.9. *Soit $1 \leq i \leq n$ et $n-i+1 \leq j \leq n$. L'estimateur*

$$\widehat{C}_{i,j} := \left(\prod_{k=n+1-i}^{j-1} \widehat{f}_k\right) C_{i,n+1-i}$$

est un estimateur sans biais de $C_{i,j}$ conditionnellement à $\mathcal{B}_{n+1-i}$.

Démonstration. En appliquant le Lemme 4.8,

$$\mathbb{E}\left(\widehat{C}_{i,j} \;\middle|\; \mathcal{B}_{n+1-i}\right) = \mathbb{E}\left(\prod_{k=n+1-i}^{j-1} \widehat{f}_k \;\middle|\; \mathcal{B}_{n+1-i}\right) C_{i,n+1-i} = \left(\prod_{k=n+1-i}^{j-1} f_k\right) C_{i,n+1-i}.$$

□

Le résultat précédent permet d'en déduire un estimateur sans biais des provisions associées à chaque année de sinistralité qui est :

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

Et enfin, un estimateur du besoin en provision totale :

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

Remarque 4.10. *Le modèle est homogène par changement d'unité ou de monnaie. En effet, si on multiplie tous les triangles par $\alpha > 0$, on remarque qu'on a les mêmes estimateurs et la même estimation des provisions. En effet, si, pour tout i, j, on pose $C^{\alpha}_{i,j} = \alpha C_{i,j}$, on a $\widehat{f}^{\alpha}_j = \widehat{f}_j$ et donc $\widehat{R}^{\alpha} = \alpha \widehat{R}$.*

4.2 Mesure de l'erreur sur l'estimation des provisions

Bien que les $\widehat{f}_j$, les $\widehat{C}_{i,n}$ et les $\widehat{R}_i$ soient des estimateurs centrés, ils demeurent des variables aléatoires. Nous allons donc calculer une mesure de l'erreur : l'erreur quadratique moyenne entre $\widehat{R}_i$ et R_i dans un premier temps puis entre $\widehat{R}$ et R dans un second temps.

Nous aurons besoin d'un estimateur des σ_j^2.

Lemme 4.11. *Supposons que nous connaissons les $(f_j)_{1 \leq j \leq n-1}$. Pour $1 \leq j \leq n-1$, l'estimateur :*

$$\widehat{\sigma}_j^2 := \frac{1}{n-j} \sum_{i=1}^{n-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right)^2 = \frac{1}{n-j} \sum_{i=1}^{n-j} C_{i,j} \left(\widehat{f}_{i,j} - f_j\right)^2$$

est un estimateur sans biais de σ_j^2 conditionnellement à $\mathcal{B}_j$.

Démonstration. On rappelle *H3*, pour $1 \leq j \leq n-1$:

$$Var(C_{i,j+1} \mid \mathcal{B}_j) = \sigma_j^2 C_{i,j}.$$

C'est à dire que :

$$\begin{aligned}
& \mathbb{E}\left[(C_{i,j+1} - \mathbb{E}(C_{i,j+1} \mid \mathcal{B}_j))^2 \mid \mathcal{B}_j\right] = \sigma_j^2 C_{i,j} \\
\iff & \mathbb{E}\left[(C_{i,j+1} - f_j C_{i,j})^2 \mid \mathcal{B}_j\right] = \sigma_j^2 C_{i,j} \\
\iff & \mathbb{E}\left[C_{i,j}\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right)^2 \;\middle|\; \mathcal{B}_j\right] = \sigma_j^2.
\end{aligned}$$

□

Comme nous n'avons aucune hypothèse supplémentaire, l'estimateur est la moyenne pondérée associée. Mais nous ne connaissons pas les f_j, nous devons les remplacer par leurs estimateurs respectifs.

Lemme 4.12. *Pour $1 \leq j \leq n-2$, l'estimateur :*

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

est un estimateur sans biais de σ_j^2 conditionnellement à $\mathcal{B}_j$.

Démonstration. On décompose la somme en trois éléments.

$$\sum_{i=1}^{n-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j \right)^2 = \sum_{i=1}^{n-j} C_{i,j} \left[\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j \right)^2 + \left(\widehat{f}_j - f_j \right)^2 - 2 \left(\frac{C_{i,j+1}}{C_{i,j}} - f_j \right) \left(\widehat{f}_j - f_j \right) \right].$$

Pour le premier, par le Lemme 4.11,

$$\mathbb{E}\left[\sum_{i=1}^{n-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - f_j \right)^2 \,\middle|\, \mathcal{B}_j \right] = (n-j)\sigma_j^2.$$

Pour le deuxième, par le Lemme 4.6,

$$\mathbb{E}\left[\sum_{i=1}^{n-j} C_{i,j} \left(\widehat{f}_j - f_j \right)^2 \,\middle|\, \mathcal{B}_j \right] = \sum_{i=1}^{n-j} C_{i,j} \mathbb{E}\left[\left(\widehat{f}_j - f_j \right)^2 \,\middle|\, \mathcal{B}_j \right] = Var\left(\widehat{f}_j \,\middle|\, \mathcal{B}_j \right) \sum_{i=1}^{n-j} C_{i,j} = \sigma_j^2.$$

Pour le troisième, on effectue le développement.

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{n-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - f_j \right) \left(\widehat{f}_j - f_j \right) \mid \mathcal{B}_j \right] \\
&= \sum_{i=1}^{n-j} C_{i,j} \mathbb{E}\left[\frac{C_{i,j+1}}{C_{i,j}} \widehat{f}_j - \widehat{f}_j f_j - \left(\frac{C_{i,j+1}}{C_{i,j}} - f_j \right) f_j \mid \mathcal{B}_j \right] \\
&\overset{H2}{=} \sum_{i=1}^{n-j} \mathbb{E}\left[C_{i,j+1} \widehat{f}_j \mid \mathcal{B}_j \right] - \sum_{i=1}^{n-j} C_{i,j} \mathbb{E}\left[\widehat{f}_j f_j \mid \mathcal{B}_j \right] \\
&= \sum_{i=1}^{n-j} \mathbb{E}\left[\frac{C_{i,j+1}^2 + C_{i,j+1} \sum_{k \neq i}^{n-j} C_{k,j+1}}{\sum_{k=1}^{n-j} C_{k,j}} \mid \mathcal{B}_j \right] - f_j^2 \sum_{i=1}^{n-j} C_{i,j} \\
&\overset{H1}{=} \sum_{i=1}^{n-j} \frac{\mathbb{E}\left[C_{i,j+1}^2 \mid \mathcal{B}_j \right] + \mathbb{E}\left[C_{i,j+1} \mid \mathcal{B}_j \right] \mathbb{E}\left[\sum_{k \neq i}^{n-j} C_{k,j+1} \mid \mathcal{B}_j \right]}{\sum_{k=1}^{n-j} C_{k,j}} - f_j^2 \sum_{i=1}^{n-j} C_{i,j} \\
&\overset{H2}{=} \sum_{i=1}^{n-j} \frac{\mathbb{E}\left[C_{i,j+1}^2 \mid \mathcal{B}_j \right] + f_j^2 C_{i,j} \sum_{k \neq i}^{n-j} C_{k,j}}{\sum_{k=1}^{n-j} C_{k,j}} - f_j^2 \sum_{i=1}^{n-j} C_{i,j} \\
&= \sum_{i=1}^{n-j} \frac{\mathbb{E}\left[C_{i,j+1}^2 \mid \mathcal{B}_j \right] - f_j^2 C_{i,j}^2 + f_j^2 C_{i,j} \sum_{k=1}^{n-j} C_{k,j}}{\sum_{k=1}^{n-j} C_{k,j}} - f_j^2 \sum_{i=1}^{n-j} C_{i,j} \\
&= \sum_{i=1}^{n-j} \frac{Var\left[C_{i,j+1} \mid \mathcal{B}_j \right]}{\sum_{k=1}^{n-j} C_{k,j}} \overset{H3}{=} \sum_{i=1}^{n-j} \frac{\sigma_j^2 C_{i,j}}{\sum_{k=1}^{n-j} C_{k,j}} = \sigma_j^2.
\end{aligned}$$

Enfin, en combinant, on obtient le résultat.

$$\mathbb{E}\left(\widehat{\sigma}_j^2 \mid \mathcal{B}_j \right) = \frac{(n-j)\sigma_j^2 + \sigma_j^2 - 2\sigma_j^2}{n-j-1} = \sigma_j^2.$$

□

On remarque qu'on ne peut pas avoir d'estimateur de σ_{n-1}^2. On pourra utiliser une extrapolation, on suppose pour la suite avoir à disposition un tel estimateur. Par exemple, [2] propose :

$$\widehat{\sigma}_{n-1}^2 := \min(\widehat{\sigma}_{n-2}^4/\widehat{\sigma}_{n-3}^2, \widehat{\sigma}_{n-3}^2, \widehat{\sigma}_{n-2}^2).$$

Lemme 4.13. *Soient* $1 \leq i \leq n$ *et* $n+1-i \leq j \leq n$.

$$Var(C_{i,j} \mid \mathcal{F}_n) = \left(\sum_{k=n+1-i}^{j-1} \left[\left(\prod_{\ell=k+1}^{j-1} f_\ell^2 \right) \sigma_k^2 \left(\prod_{\ell=n+1-i}^{k-1} f_\ell \right) \right] \right) C_{i,n+1-i}$$

Démonstration. Pour $j = n+1-i$, on est sur la diagonale et $Var(C_{i,n+1-i} \mid \mathcal{F}_n) = 0$. Soit $n+2-i \leq j \leq n$. On a $i+j-2 \geq n$. En utilisant la formule de la variance totale, par *H2*, *H3* et le lemme 4.8, on a

$$\begin{aligned} Var(C_{i,j} \mid \mathcal{F}_n) &= \mathbb{E}(Var(C_{i,j} \mid \mathcal{F}_{i+j-2}) \mid \mathcal{F}_n) + Var(\mathbb{E}(C_{i,j} \mid \mathcal{F}_{i+j-2}) \mid \mathcal{F}_n) \\ &= \sigma_{j-1}^2 \mathbb{E}(C_{i,j-1} \mid \mathcal{F}_n) + f_{j-1}^2 Var(C_{i,j-1} \mid \mathcal{F}_n) \\ &= \sigma_{j-1}^2 \left(\prod_{\ell=n+1-i}^{j-2} f_\ell \right) C_{i,n+1-i} + f_{j-1}^2 Var(C_{i,j-1} \mid \mathcal{F}_n) \end{aligned}$$

Puis, par récurrence, on en déduit le résultat. □

Lemme 4.14. *L'erreur quadratique moyenne des provisions pour une année i est*

$$eqm(\widehat{R}_i \mid \mathcal{F}_n) := \mathbb{E}\left[(\widehat{R}_i - R_i)^2 \mid \mathcal{F}_n \right].$$

On a la décomposotition :

$$eqm(\widehat{R}_i \mid \mathcal{F}_n) = \left(\widehat{R}_i - \mathbb{E}(R_i \mid \mathcal{F}_n) \right)^2 + Var(R_i \mid \mathcal{F}_n).$$

avec

$$\begin{aligned} \left(\widehat{R}_i - \mathbb{E}(R_i \mid \mathcal{F}_n) \right)^2 &= \left(\prod_{k=n+1-i}^{n-1} \widehat{f}_k - \prod_{k=n+1-i}^{n-1} f_k \right)^2 C_{i,n+1-i}^2, \\ Var(R_i \mid \mathcal{F}_n) &= \left(\sum_{k=n+1-i}^{n-1} \left[\left(\prod_{\ell=k+1}^{n-1} f_\ell^2 \right) \sigma_k^2 \left(\prod_{\ell=n+1-i}^{k-1} f_\ell \right) \right] \right) C_{i,n+1-i}. \end{aligned}$$

Démonstration.

$$\begin{aligned} \mathbb{E}\left[(\widehat{R}_i - R_i)^2 \mid \mathcal{F}_n \right] &= \mathbb{E}\left[(\widehat{R}_i - \mathbb{E}(R_i \mid \mathcal{F}_n))^2 \mid \mathcal{F}_n \right] + \mathbb{E}\left[(R_i - \mathbb{E}(R_i \mid \mathcal{F}_n))^2 \mid \mathcal{F}_n \right] \\ &\quad - 2\mathbb{E}\left[(\widehat{R}_i - \mathbb{E}(R_i \mid \mathcal{F}_n))(R_i - \mathbb{E}(R_i \mid \mathcal{F}_n)) \mid \mathcal{F}_n \right] \\ &= \left(\widehat{R}_i - \mathbb{E}(R_i \mid \mathcal{F}_n) \right)^2 + Var(R_i \mid \mathcal{F}_n). \end{aligned}$$

En remarquant que $\left(\widehat{R}_i - \mathbb{E}(R_i \mid \mathcal{F}_n) \right)^2 = \left(\widehat{C}_{i,n} - \mathbb{E}(C_{i,n} \mid \mathcal{F}_n) \right)^2$ et que $Var(R_i \mid \mathcal{F}_n) = Var(C_{i,n} \mid \mathcal{F}_n)$, par la définition de $\widehat{C}_{i,n}$ et par le lemme 4.13, on obtient le résultat. □

Proposition 4.15. *L'erreur quadratique moyenne peut être estimée par :*

$$\widehat{eqm}(\widehat{R}_i \mid \mathcal{F}_n) = \widehat{C}_{i,n}^2 \sum_{j=n+1-i}^{n-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \left(\frac{1}{\widehat{C}_{i,j}} + \frac{1}{\sum_{k=1}^{n-j} C_{k,j}} \right).$$

Démonstration. On reprend l'expression de $eqm(\widehat{R}_i \mid \mathcal{F}_n)$ du lemme 4.14. Le membre de droite s'estime en remplaçant les (f_k) et les (σ_k^2) par leurs estimateurs respectifs, et on déduit

$$\begin{aligned}
\widehat{Var}(R_i \mid \mathcal{F}_n) &= \left(\sum_{k=n+1-i}^{n-1} \left[\left(\prod_{\ell=k+1}^{n-1} \widehat{f}_\ell^2 \right) \widehat{\sigma}_k^2 \left(\prod_{\ell=n+1-i}^{k-1} \widehat{f}_\ell \right) \right] \right) C_{i,n+1-i} \\
&= \left(\sum_{k=n+1-i}^{n-1} \left[\left(\prod_{\ell=k+1}^{n-1} \widehat{f}_\ell \right) \frac{\widehat{\sigma}_k^2}{\widehat{f}_k} \right] \right) \widehat{C}_{i,n} \\
&= \left(\sum_{k=n+1-i}^{n-1} \frac{\widehat{C}_{i,n}}{\widehat{C}_{i,k}} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \right) \widehat{C}_{i,n} \\
&= \widehat{C}_{i,n}^2 \sum_{k=n+1-i}^{n-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \frac{1}{\widehat{C}_{i,k}}
\end{aligned}$$

Pour la seconde partie, qui est l'erreur d'estimation des paramètres, on ne peut pas simplement remplacer les (f_k) par les $(\widehat{f}_k)$ car cela conduirait à une erreur nulle. On admet qu'on peut approximer cette quantité par (voir [2, Théorème 3])

$$\widehat{C}_{i,n}^2 \sum_{j=n+1-i}^{n-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \frac{1}{\sum_{k=1}^{n-j} C_{k,j}},$$

ce qui permet de conclure. □

La première partie de l'erreur quadratique moyenne correspond au fait que, vu d'aujourd'hui, $C_{i,n}$ est une variable aléatoire. Si nous connaissions les vrais paramètres, elle serait inconnue d'aujourd'hui et cela correspond à sa variance. La seconde partie correspond à une erreur d'estimation des paramètres.

Proposition 4.16. *L'erreur quadratique moyenne des provisions totales peut être estimée par :*

$$\widehat{eqm}(\widehat{R}) = \sum_{i=2}^{n} \left(\widehat{eqm}(\widehat{R}_i) + \widehat{C}_{i,n} \left(\sum_{k=i+1}^{n} \widehat{C}_{k,n} \right) \sum_{j=n+1-i}^{n-1} \frac{2\widehat{\sigma}_j^2 / \widehat{f}_j^2}{\sum_{\ell=1}^{n-j} C_{\ell,j}} \right).$$

Démonstration. Admis (voir [2, Corollaire après Théorème 3]) □

La première partie (somme des erreurs quadratiques moyennes) correspond à la somme des différentes années, indépendantes. La seconde est associée à l'erreur des paramètres : ce sont les mêmes paramètres qui interviennent, la diversification est amoindrie.

i	$C_{i,1}$	$C_{i,2}$	$C_{i,3}$	$C_{i,4}$	$C_{i,5}$	$C_{i,6}$	$C_{i,7}$	$C_{i,8}$	$C_{i,9}$	$C_{i,10}$
1	357848	1124788	1735330	2218270	2745596	3319994	3466336	3606286	3833515	3901463
2	352118	1236139	2170033	3353322	3799067	4120063	4647867	4914039	5339085	
3	290507	1292306	2218525	3235179	3985995	4132918	4628910	4909315		
4	310608	1418858	2195047	3757447	4029929	4381982	4588268			
5	443160	1136350	2128333	2897821	3402672	3873311				
6	396132	1333217	2180715	2985752	3691712					
7	440832	1288463	2419861	3483130						
8	359480	1421128	2864494							
9	376686	1363294								
10	344014									

4.3 Exemple

Nous avons le triangle cumulé suivant (tiré de [2]).
Le code suivant permet de l'importer directement.

```
C ← read.table("https://nicolasbaradel.fr/enseignement/ressources/R
   /donneesMACK.txt", sep = "&")
n ← nrow(C)
```

Si le triangle ci-dessus est importé dans une matrice ou data.frame qui s'appelle C dans le langage R, alors le code suivant permet de calculer $(\widehat{f}_j)_{1\leq j\leq n-1}$.

```
f_ ← function(C)
{
    f ← numeric(n−1)
    for(j in 1:(n−1))
        f[j] ← sum(C[1:(n−j), j + 1])/sum(C[1:(n−j), j])
    return(f)
}
f ← f_(C)
```

Le code suivant permet de calculer $(\widehat{\sigma}_j^2)_{1\leq j\leq n-2}$. On y utlise l'extrapolation de [2] pour $\widehat{\sigma}_{n-1}^2$.

```
s2_ ← function(C, f)
{
    s2 ← numeric(n−1)
    for(j in 1:(n−2))
        s2[j] ← sum( C[1:(n−j), j] * (C[1:(n−j), j + 1]/C[1:(n−j),
           j] − f[j])^2 )/(n − j − 1)
    s2[n−1] ← min(s2[n−2]^2/s2[n−3], s2[n−3], s2[n−2])
    return(s2)
}
s2 ← s2_(C, f)
```

On obtient le Tableau 1.
On peut ensuite calculer les $\widehat{C}_{i,j}$ pour $i+j>n+1$ avec le code R suivant.

```
hatC_ ← function(C, f)
{
    for(j in 1:(n−1))
        C[(n−j+1):n, j+1] ← f[j]*C[(n−j+1):n, j]
```

j	1	2	3	4	5	6	7	8	9
$\widehat{f}_j$	3.491	1.747	1.457	1.174	1.104	1.086	1.054	1.077	1.018
$\widehat{\sigma}_j^2$	160280	37737	41965	15183	13731	8186	447	1147	447

Table 1 – Estimateurs des paramètres de Mack - Chain Ladder

```
    return(C)
}
C ← hat_C(C)
```

Et en déduire les provisions.

```
R_ ← function(C)
        return(C[, n] − rev(C[row(C) + col(C) == n + 1]))
R ← R_(C)
```

On obtient le tableau en Figure 1.

i	$\widehat{C}_{i,n}$	$\widehat{R}_i$
1	3901463	0
2	5433719	94634
3	5378826	469511
4	5297906	709638
5	4858200	984889
6	5111171	1419459
7	5660771	2177641
8	6784790	3920296
9	5642265	4278971
10	4969824	4625810

Figure 1 – Ultimes et provisions par année.

La provision totale est $\widehat{R} = 18680848$. Nous pouvons maintenant calculer l'erreur quadratique moyenne de chaque année avec le code suivant. La matrice `C0` ci-dessous, qui ne contient pas la diagonale, est construite pour obtenir efficacement les sommes de $C_{i,j}$ qui interviennent dans les formules des erreurs quadratiques moyennes.

```
C0 ← C; C0[row(C) + col(C) >= n+1] ← 0
eqm_R ← numeric(n)
for(i in 2:n)
{
    J ← (n+1−i):(n−1)
    eqm_R[i] ← C[i, n]^2*sum( (s2[J]/f[J]^2) *
                  (1/C[i,J] + 1/colSums(C0[, J, drop = FALSE])) )
}
```

En exprimant sa racine carrée, pour chaque année, en pourcentage de l'estimateur des provisions, on obtient :

i	$\frac{\sqrt{\widehat{EQM}(\widehat{R}_i)}}{\widehat{R}_i}(\%)$
1	0.0
2	79.8
3	25.9
4	18.8
5	26.5
6	29.0
7	25.6
8	22.3
9	22.7
10	29.5

Enfin, le code R suivant permet de calculer l'erreur quadratique moyenne de l'estimation des provisions :

```
eqm <- eqm_R[n]
for(i in 2:(n-1))
{
    J <- (n+1-i):(n-1)
    eqm <- eqm + eqm_R[i] + C[i, n]*sum(C[(i+1):n, n]) *
            2*sum((s2[J]/f[J]^2)/colSums(C0[, J, drop = FALSE]))
}
```

En prenant la racine carrée de l'erreur quadratique moyenne et en divisant par $\widehat{R}$, nous obtenons 13,1%.

5 Modèle de Schnieper

Cette section s'appuie principalement sur [3].

5.1 Estimation des provisions

Cette fois-ci nous allons utiliser l'exposition. Le défaut de Chain Ladder - Mack est le caractère multiplicatif pur. Les sinistres latents sont proportionnels à ce qui est connu. L'idée du modèle de Schnieper est la suivante :

— Il y a les sinistres qui ont déjà eu lieu et qui sont inconnus de l'assureur, ils sont proportionnels à l'exposition, il s'agit des *vrais IBNR* ;

— Il y a les sinistres connus mais dont le montant estimé peut être amené à évoluer (coût final inconnu), cette évolution est proportionnelle à la charge actuelle, il s'agit de révisions de charges de sinitre : les *IBNER (Incurred But Not Enough Reported).*

On reprend les variables aléatoires $(C_{i,j})_{1\leq i,j\leq n}$ qui représentent le coût total cumulé. Nous définissons de plus :

— Les variables aléatoires $(N_{i,j})_{1\leq i,j\leq n}$ qui représentent le montant des nouveaux sinistres survenus en année i et déclarés en année de développement j ;

— Les variables aléatoires $(D_{i,j})_{1\leq i,j\leq n}$ qui représentent la baisse du coût estimé des sinistres déclarés les années précédentes.

Par construction $D_{i,1} = 0$ pour tout $1 \leq i \leq n$ et $D_{i,j} \leq C_{i,j-1}$ pour $1 \leq i \leq n$ et $2 \leq j \leq n$.

Le coût incrémental des sinistres s'écrit :

$$\begin{array}{ll} C_{i,1} = N_{i,1}, & 1 \leq i \leq n \\ C_{i,j+1} = C_{i,j} + N_{i,j+1} - D_{i,j+1}, & 1 \leq i \leq n,\ 1 \leq j \leq n-1 \end{array} \tag{1}$$

Parmi C, N, D, il suffit d'en connaître deux pour en déduire le troisième, l'un d'entre-eux est redondant linéairement. On introduit la filtration :

$$\mathcal{F}_k := \sigma(N_{i,j}, D_{i,j} \mid i + j \leq k + 1).$$

Nous avons donc deux triangles de données, dont l'agrégation en suivant (1) donne C.

Le modèle repose sur les hypothèses suivantes :

Hypothèse 5.1.

H1 Les v.a. $(N_{i_1,j}, D_{i_1,j})_{1\leq j\leq n}$ *et* $(N_{i_2,j}, D_{i_2,j})_{1\leq j\leq n}$ *sont indépendantes pour* $i_1 \neq i_2$.

H2 Pour $1 \leq j \leq n$, *il existe* $\lambda_j \geq 0$ *et pour* $1 \leq j \leq n-1$, *il existe* $\delta_j \leq 1$ *tels que*

$$\begin{aligned} \mathbb{E}(N_{i,j} \mid \mathcal{F}_{i+j-2}) &= \lambda_j E_i, \\ \mathbb{E}(D_{i,j+1} \mid \mathcal{F}_{i+j-1}) &= \delta_j C_{i,j}. \end{aligned}$$

H3 Pour $1 \leq j \leq n-1$, *il existe* $\sigma_j^2 \geq 0$ *et* $\tau_j^2 \geq 0$ *tels que*

$$\begin{aligned} Var(N_{i,j} \mid \mathcal{F}_{i+j-2}) &= \sigma_j^2 E_i, \\ Var(D_{i,j+1} \mid \mathcal{F}_{i+j-1}) &= \tau_j^2 C_{i,j}. \end{aligned}$$

Remarque 5.1. *Par la formule des espérances conditionnelles emboîtées et de la variance totale, en combinant* $H2$ *et* $H3$*, on a également :*

$$\begin{aligned} \mathbb{E}(N_{i,j}) &= \lambda_j E_i, \\ Var(N_{i,j}) &= \sigma_j^2 E_i. \end{aligned}$$

Pour $1 \leq k \leq n$, on introduit la filtration $\mathcal{B}_k = \sigma(N_{i,j}, D_{i,j} \mid i + j \leq n + 1, j \leq k) \subset \mathcal{F}_n$.

Remarque 5.2. *Sous H1, H2 et H3 vérifient également :*

$$\begin{aligned} \mathbb{E}(N_{i,j} \mid \mathcal{B}_{j-1}) &= \lambda_j E_i, \\ \mathbb{E}(D_{i,j+1} \mid \mathcal{B}_j) &= \delta_j C_{i,j}, \\ Var(N_{i,j} \mid \mathcal{B}_{j-1}) &= \sigma_j^2 E_i, \\ Var(D_{i,j+1} \mid \mathcal{B}_j) &= \tau_j^2 C_{i,j}. \end{aligned}$$

Démonstration. Voir la démonstration de la Remarque 4.1. □

On peut poser $\widehat{\lambda}_{i,j} := \frac{N_{i,j}}{E_i}$ pour $i + j \leq n + 1$, i.e. pour les cas observables en année n ; et $\widehat{\delta}_{i,j} := \frac{D_{i,j+1}}{C_{i,j}}$ pour $i + j + 1 \leq n + 1$.

En introduisant des poids comme dans le modèle Chain Ladder, un estimateur de λ_j et un estimateur de δ_j sont :

$$\begin{aligned} \widehat{\lambda}_j &:= \frac{\sum_{i=1}^{n-j+1} p_{i,j} \widehat{\lambda}_{i,j}}{\sum_{i=1}^{n-j+1} p_{i,j}}, & 1 \leq j \leq n, \\ \widehat{\delta}_j &:= \frac{\sum_{i=1}^{n-j} p_{i,j} \widehat{\delta}_{i,j}}{\sum_{i=1}^{n-j} p_{i,j}}, & 1 \leq j \leq n-1. \end{aligned}$$

Lemme 5.3. *On a :*

$$\mathbb{E}\left(\widehat{\lambda}_{i,j} \mid \mathcal{B}_{j-1}\right) = \lambda_j, \qquad 1 \leq j \leq n$$
$$\mathbb{E}\left(\widehat{\delta}_{i,j} \mid \mathcal{B}_{j}\right) = \delta_j, \qquad 1 \leq j \leq n-1.$$

Démonstration. Par *H2* :

$$\mathbb{E}\left(\widehat{\lambda}_{i,j} \mid \mathcal{B}_{j-1}\right) = \frac{1}{E_i}\mathbb{E}\left(N_{i,j} \mid \mathcal{B}_{j-1}\right) = \lambda_j,$$

et

$$\mathbb{E}\left(\widehat{\delta}_{i,j} \mid \mathcal{B}_{j}\right) = \frac{1}{C_{i,j}}\mathbb{E}\left(D_{i,j+1} \mid \mathcal{B}_{j}\right) = \delta_j.$$

□

Proposition 5.4. *Sous H2, pour tout $1 \leq j \leq n$, pour tous les poids qui sont $\mathcal{B}_{j-1}$-mesurable pour l'estimateur $\widehat{\lambda}_j$, et pour tout $1 \leq j \leq n-1$ pour tous les poids qui sont $\mathcal{B}_j$-mesurable pour l'estimateur $\widehat{\delta}_j$, on a :*

$$\mathbb{E}\left(\widehat{\lambda}_{j} \mid \mathcal{B}_{j-1}\right) = \lambda_j,$$
$$\mathbb{E}\left(\widehat{\delta}_{j} \mid \mathcal{B}_{j}\right) = \delta_j.$$

En particulier, il est sans biais :

$$\mathbb{E}\left(\widehat{\lambda}_{j}\right) = \lambda_j,$$
$$\mathbb{E}\left(\widehat{\delta}_{j}\right) = \delta_j.$$

Démonstration. Analogue à la proposition 4.3. □

Proposition 5.5. *Sous $H1-H2-H3$, le meilleur estimateur linéaire de λ_j en $(\widehat{\lambda}_{i,j})_{1\leq i\leq n-j+1}$ conditionnellement à $\mathcal{B}_{j-1}$ et le meilleur estimateur linéaire de δ_j en $(\widehat{\delta}_{i,j})_{1\leq i\leq n-j}$ conditionnellement à $\mathcal{B}_j$ pour l'erreur quadratique moyenne sont*

$$\widehat{\lambda}_j := \frac{\sum_{i=1}^{n-j+1} E_i\widehat{\lambda}_{i,j}}{\sum_{i=1}^{n-j+1} E_i},$$
$$\widehat{\delta}_j := \frac{\sum_{i=1}^{n-j} C_{i,j}\widehat{\delta}_{i,j}}{\sum_{i=1}^{n-j} C_{i,j}},$$

c'est-à-dire :

$$\widehat{\lambda}_j := \frac{\sum_{i=1}^{n-j+1} N_{i,j}}{\sum_{i=1}^{n-j+1} E_i},$$
$$\widehat{\delta}_j := \frac{\sum_{i=1}^{n-j} D_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}.$$

Démonstration. Comme dans la démonstration de la Proposition 4.5, étant donné que les estimateurs sont sans biais, il suffit de minimiser la variance. Les poids optimaux sont donnés par le Lemme 4.4.

Par *H3* :

$$Var(\widehat{\lambda}_{i,j} \mid \mathcal{B}_{j-1}) = \frac{\sigma_j^2}{E_i}, \quad Var(\widehat{\delta}_{i,j} \mid \mathcal{B}_j) = \frac{\tau_j^2}{C_{i,j}}.$$

Pour $\widehat{\lambda}_j$,

$$p_{i,j} \propto E_i,$$

tandis que, pour $\widehat{\delta}_j$,

$$p_{i,j} \propto C_{i,j}.$$

□

On peut en déduire la variance des $\widehat{\lambda}_j$ et des $\widehat{\delta}_j$.

Lemme 5.6. *On a*

$$Var(\widehat{\lambda}_j \mid \mathcal{B}_{j-1}) = \frac{\sigma_j^2}{\sum_{i=1}^{n-j+1} E_i},$$
$$Var(\widehat{\delta}_j \mid \mathcal{B}_j) = \frac{\tau_j^2}{\sum_{i=1}^{n-j} C_{i,j}}.$$

Démonstration. On a par *H1* et *H3*,

$$Var(\widehat{\lambda}_j \mid \mathcal{B}_{j-1}) \overset{H1}{=} \frac{\sum_{i=1}^{n-j+1} Var(N_{i,j} \mid \mathcal{B}_{j-1})}{\left(\sum_{i=1}^{n-j+1} E_i\right)^2} \overset{H3}{=} \frac{\sum_{i=1}^{n-j+1} \sigma_j^2 E_i}{\left(\sum_{i=1}^{n-j+1} E_i\right)^2} = \frac{\sigma_j^2}{\sum_{i=1}^{n-j+1} E_i},$$

$$Var(\widehat{\delta}_j \mid \mathcal{B}_j) \overset{H1}{=} \frac{\sum_{i=1}^{n-j} Var(D_{i,j+1} \mid \mathcal{B}_j)}{\left(\sum_{i=1}^{n-j} C_{i,j}\right)^2} \overset{H3}{=} \frac{\sum_{i=1}^{n-j} \tau_j^2 C_{i,j}}{\left(\sum_{i=1}^{n-j} C_{i,j}\right)^2} = \frac{\tau_j^2}{\sum_{i=1}^{n-j} C_{i,j}}.$$

□

Le résultat suivant, qui implique une absence de corrélation entre les estimateurs impliqués, nous permettra d'établir un estimateur sans biais de l'estimation des provisions.

Lemme 5.7. *Soient $1 \leq k \leq n-1$ et $1 \leq j_1 < j_2 < \ldots < j_k$, on a*

$$\mathbb{E}\left(\prod_{i=1}^{k}(1-\widehat{\delta}_{j_i}) \;\middle|\; \mathcal{B}_{j_1}\right) = \prod_{i=1}^{k}\mathbb{E}(1-\widehat{\delta}_{j_i}) = \prod_{i=1}^{k}(1-\delta_{j_i}).$$

De plus, on a

$$\mathbb{E}\left(\widehat{\lambda}_{j_1}\prod_{i=1}^{k}(1-\widehat{\delta}_{j_i}) \;\middle|\; \mathcal{B}_{j_1-1}\right) = \mathbb{E}(\widehat{\lambda}_{j_1})\prod_{i=1}^{k}\mathbb{E}(1-\widehat{\delta}_{j_i}) = \lambda_{j_1}\prod_{i=1}^{k}(1-\delta_{j_i}).$$

Démonstration. La première égalité se démontre de la même manière que la Proposition 4.7. La seconde :

$$\begin{aligned}
\mathbb{E}\left[\widehat{\lambda}_{j_1}\prod_{i=1}^{k}(1-\widehat{\delta}_{j_i}) \;\middle|\; \mathcal{B}_{j_1-1}\right] &= \mathbb{E}\left[\mathbb{E}\left(\widehat{\lambda}_{j_1}\prod_{i=1}^{k}(1-\widehat{\delta}_{j_i}) \;\middle|\; \mathcal{B}_{j_k}\right) \;\middle|\; \mathcal{B}_{j_1-1}\right] \\
&= \mathbb{E}\left[\widehat{\lambda}_{j_1}\prod_{i=1}^{k-1}(1-\widehat{\delta}_{j_i})\mathbb{E}(1-\widehat{\delta}_{j_k} \mid \mathcal{B}_{j_k}) \;\middle|\; \mathcal{B}_{j_1-1}\right] \\
&= \mathbb{E}\left[\widehat{\lambda}_{j_1}\prod_{i=1}^{k-1}(1-\widehat{\delta}_{j_i}) \;\middle|\; \mathcal{B}_{j_1-1}\right](1-\delta_{j_k}),
\end{aligned}$$

puis on conclut par récurrence. □

Lemme 5.8. *Soient* $1 \le i \le n$ *et* $n-i+1 \le j \le n$.

$$\mathbb{E}(C_{i,j} \mid \mathcal{F}_n) = \left(\prod_{k=n+1-i}^{j-1} (1-\delta_k) \right) C_{i,n+1-i} + E_i \sum_{k=n+2-i}^{j} \lambda_k \left(\prod_{\ell=k}^{j-1} (1-\delta_\ell) \right).$$

Démonstration. Pour $j = n+1-i$, on est sur la diagonale et $\mathbb{E}(C_{i,n+1-i} \mid \mathcal{F}_n) = C_{i,n+1-i}$. Soit $n+2-i < j \le n$. On a $i+j-2 \ge n$. En utilisant la formule des espérances conditionnelles emboitées, par *H2*, on a

$$\mathbb{E}(C_{i,j} \mid \mathcal{F}_n) = \mathbb{E}(\mathbb{E}(C_{i,j-1} + N_{i,j} - D_{i,j} \mid \mathcal{F}_{i+j-2}) \mid \mathcal{F}_n) = (1-\delta_{j-1})\mathbb{E}(C_{i,j-1} \mid \mathcal{F}_n) + E_i \lambda_j.$$

Puis, par récurrence, on en déduit le résultat.

□

Proposition 5.9. *Soient* $1 \le i \le n$ *et* $n-i+1 \le j \le n$. *L'estimateur*

$$\widehat{C}_{i,j} := \left(\prod_{k=n+1-i}^{j-1} (1-\widehat{\delta}_k) \right) C_{i,n+1-i} + E_i \sum_{k=n+2-i}^{j} \widehat{\lambda}_k \left(\prod_{\ell=k}^{j-1} (1-\widehat{\delta}_\ell) \right)$$

est un estimateur sans biais de $C_{i,j}$ *conditionnellement à* $\mathcal{B}_{n+1-i}$.

Démonstration. En appliquant le Lemme 5.7,

$$\begin{aligned}
\mathbb{E}\left(\widehat{C}_{i,j} \mid \mathcal{B}_{n+1-i}\right) &= \mathbb{E}\left(\prod_{k=n+1-i}^{j-1} (1-\widehat{\delta}_k) \,\middle|\, \mathcal{B}_{n+1-i} \right) C_{i,n+1-i} \\
&\quad + E_i \sum_{k=n+2-i}^{j} \mathbb{E}\left[\mathbb{E}\left(\widehat{\lambda}_k \left(\prod_{\ell=k}^{j-1} (1-\widehat{\delta}_\ell) \right) \,\middle|\, \mathcal{B}_{k-1} \right) \,\middle|\, \mathcal{B}_{n+1-i} \right] \\
&= \left(\prod_{k=n+1-i}^{j-1} (1-\delta_k) \right) C_{i,n+1-i} + E_i \sum_{k=n+2-i}^{j} \lambda_k \left(\prod_{\ell=k}^{j-1} (1-\delta_\ell) \right).
\end{aligned}$$

□

Remarque 5.10. *Le modèle est invariant par changement d'unité ou de monnaie. En effet, si on multiplie tous les triangles ainsi que les expositions* $(E_i)_{1 \le i \le n}$ *par* α *>0, on remarque qu'on a les mêmes estimateurs et la même estimation des provisions. Si on ne multiplie pas* E_i *par* α *(c'est, par exemple, le nombre de contrats), nous n'avons plus les mêmes estimateurs de* $\widehat{\lambda}_j$ *mais l'estimation des provisions est inchangée.*

5.2 Mesure de l'erreur

Lemme 5.11. *Supposons que nous connaissons les* $(\lambda_j)_{1 \le j \le n}$ *et les* $(\delta_j)_{1 \le j \le n-1}$. *Pour* $1 \le j \le n-1$, *les estimateurs :*

$$\begin{aligned}
\widehat{\sigma}_j^2 &:= \frac{1}{n-j+1} \sum_{i=1}^{n-j+1} E_i \left(\frac{N_{i,j}}{E_i} - \lambda_j \right)^2 &&= \frac{1}{n-j+1} \sum_{i=1}^{n-j+1} E_i \left(\widehat{\lambda}_{i,j} - \lambda_j \right)^2, \\
\widehat{\tau}_j^2 &:= \frac{1}{n-j} \sum_{i=1}^{n-j} C_{i,j} \left(\frac{D_{i,j+1}}{C_{i,j}} - \delta_j \right)^2 &&= \frac{1}{n-j} \sum_{i=1}^{n-j} C_{i,j} \left(\widehat{\delta}_{i,j} - \delta_j \right)^2.
\end{aligned}$$

sont des estimateurs sans biais de σ_j^2 *conditionnellement à* $\mathcal{B}_{j-1}$ *et de* τ_j^2 *conditionnellement à* $\mathcal{B}_j$.

Démonstration. On rappelle *H3*, pour $1 \leq j \leq n-1$:

$$Var(N_{i,j} \mid \mathcal{B}_{j-1}) = \sigma_j^2 E_i,$$
$$Var(D_{i,j+1} \mid \mathcal{B}_j) = \tau_j^2 C_{i,j}.$$

C'est à dire que :

$$\begin{aligned} & \mathbb{E}\left[(N_{i,j} - \mathbb{E}(C_{i,j} \mid \mathcal{B}_{j-1}))^2 \mid \mathcal{B}_{j-1}\right] = \sigma_j^2 E_i \\ \iff & \mathbb{E}\left[(N_{i,j} - \lambda_j E_i)^2 \mid \mathcal{B}_{j-1}\right] = \sigma_j^2 C_{i,j} \\ \iff & \mathbb{E}\left[E_i \left(\frac{N_{i,j}}{E_i} - \lambda_j\right)^2 \mid \mathcal{B}_{j-1}\right] = \sigma_j^2. \end{aligned}$$

De même on montre que

$$\mathbb{E}\left[C_{i,j} \left(\frac{D_{i,j+1}}{C_{i,j}} - \delta_j\right)^2 \mid \mathcal{B}_j\right] = \tau_j^2.$$

□

Lemme 5.12. *Pour* $1 \leq j \leq n-2$, *les estimateurs :*

$$\widehat{\sigma}_j^2 := \frac{1}{n-j} \sum_{i=1}^{n-j+1} E_i \left(\frac{N_{i,j}}{E_i} - \widehat{\lambda}_j\right)^2,$$
$$\widehat{\tau}_j^2 := \frac{1}{n-j-1} \sum_{i=1}^{n-j} C_{i,j} \left(\frac{D_{i,j+1}}{C_{i,j}} - \widehat{\delta}_j\right)^2.$$

sont des estimateurs sans biais de σ_j^2 *conditionnellement à* $\mathcal{B}_{j-1}$ *et de* τ_j^2 *conditionnellement à* $\mathcal{B}_j$.

Démonstration. Il suffit d'adapter la démonstration du Lemme 4.12. □

Lemme 5.13. *Soient* $1 \leq i \leq n$ *et* $n+1-i \leq j \leq n$.

$$Var(C_{i,j} \mid \mathcal{F}_n) = \sum_{k=n+1-i}^{j-1} \left[\left(\prod_{\ell=k+1}^{j-1} (1-\delta_\ell)^2\right) \left[\sigma_{k+1}^2 E_i + \tau_k^2 \left(\left[\prod_{\ell=n+1-i}^{k-1} (1-\delta_\ell)\right] C_{i,n+1-i} + E_i \sum_{\ell=n+2-i}^{k} \lambda_\ell \prod_{m=\ell}^{k-1} (1-\delta_m)\right)\right]\right]$$

Démonstration. Pour $j = n+1-i$, on est sur la diagonale et $Var(C_{i,n+1-i} \mid \mathcal{F}_n) = 0$. Soit $n+2-i \leq j \leq n$. On a $i+j-2 \geq n$. En utilisant la formule de la variance totale, par *H2*, *H3*, on a

$$\begin{aligned} Var(C_{i,j} \mid \mathcal{F}_n) &= \mathbb{E}(Var(C_{i,j} \mid \mathcal{F}_{i+j-2}) \mid \mathcal{F}_n) + Var(\mathbb{E}(C_{i,j} \mid \mathcal{F}_{i+j-2}) \mid \mathcal{F}_n) \\ &= \mathbb{E}(Var(C_{i,j-1} + N_{i,j} - D_{i,j} \mid \mathcal{F}_{i+j-2}) \mid \mathcal{F}_n) + Var(\mathbb{E}(C_{i,j-1} + N_{i,j} - D_{i,j} \mid \mathcal{F}_{i+j-2}) \mid \mathcal{F}_n) \\ &= \sigma_j^2 E_i + \tau_{j-1}^2 \mathbb{E}(C_{i,j-1} \mid \mathcal{F}_n) + (1-\delta_{j-1})^2 Var(C_{i,j-1} \mid \mathcal{F}_n) \end{aligned}$$

Enfin, par le lemme 5.8,

$$Var(C_{i,j} \mid \mathcal{F}_n) = \sigma_j^2 E_i + \tau_{j-1}^2 \left(\left[\prod_{\ell=n+1-i}^{j-2} (1-\delta_\ell)\right] C_{i,n+1-i} + E_i \sum_{\ell=n+2-i}^{j-1} \lambda_\ell \prod_{m=\ell}^{j-2} (1-\delta_m)\right) + (1-\delta_{j-1})^2 Var(C_{i,j-1} \mid \mathcal{F}_n)$$

Puis, par récurrence, on en déduit le résultat. □

5.3 Exemple

Nous avons le triangle cumulé suivant (tiré de [3]) auquel on associe l'exposition.

i	E_i	$C_{i,1}$	$C_{i,2}$	$C_{i,3}$	$C_{i,4}$	$C_{i,5}$	$C_{i,6}$	$C_{i,7}$
1	10224	7.5	28.9	52.6	84.5	80.1	76.9	79.5
2	12752	1.6	14.8	32.1	39.6	55	60	
3	14875	13.8	42.4	36.3	53.3	96.5		
4	17365	2.9	14	32.5	46.9			
5	19410	2.9	9.8	52.7				
6	17617	1.9	29.4					
7	18129	19.1						

Celui-ci se décompose en la matrice N :

i	$N_{i,1}$	$N_{i,2}$	$N_{i,3}$	$N_{i,4}$	$N_{i,5}$	$N_{i,6}$	$N_{i,7}$
1	7.5	18.3	28.5	23.4	18.6	0.7	5.1
2	1.6	12.6	18.2	16.1	14	10.6	
3	13.8	22.7	4	12.4	12.1		
4	2.9	9.7	16.4	11.6			
5	2.9	6.9	37.1				
6	1.9	27.5					
7	19.1						

et D :

i	$D_{i,1}$	$D_{i,2}$	$D_{i,3}$	$D_{i,4}$	$D_{i,5}$	$D_{i,6}$	$D_{i,7}$
1	0	-3.1	4.8	-8.5	23	3.9	2.5
2	0	-0.6	0.9	8.6	-1.4	5.6	
3	0	-5.9	10.1	-4.6	-31.1		
4	0	-1.4	-2.1	-2.8			
5	0	0	-5.8				
6	0	0					
7	0						

Nous importons les triangles ci-dessus dans des matrices ou des data.frame qui s'appellent respectivement `C`, `N` et `D` dans le langage R, et l'exposition dans un vecteur `E`.

```r
E <- read.table("https://nicolasbaradel.fr/enseignement/ressources/R
   /donneesSCHNIEPER_E.txt")[[1]]
D <- read.table("https://nicolasbaradel.fr/enseignement/ressources/R
   /donneesSCHNIEPER_D.txt", sep = "&")
N <- read.table("https://nicolasbaradel.fr/enseignement/ressources/R
   /donneesSCHNIEPER_N.txt", sep = "&")
n <- nrow(N)
C <- N
for(j in 1:(n-1))
    C[, j+1] <- C[, j] + N[, j + 1] - D[, j+1]
```

Le code suivant permet de calculer $(\widehat{\lambda}_j)_{1\leq j\leq n}$ et $(\widehat{\delta}_j)_{1\leq j\leq n-1}$.

```r
lambda_ ← function(N, E)
{
    lambda ← numeric(n)
    for(j in 1:n)
        lambda[j] ← sum(N[1:(n-j+1), j])/sum(E[1:(n-j+1)])
    return(lambda)
}
lambda ← lambda_(N, E)

delta_ ← function(C, D)
{
    delta ← numeric(n-1)
    for(j in 1:(n-1))
        delta[j] ← sum(D[1:(n-j), j + 1])/sum(C[1:(n-j), j])
        return(delta)
}
delta ← delta_(C, D)
```

Le code suivant permet de calculer $(\widehat{\sigma}_j^2)_{1\leq j\leq n}$ et $(\widehat{\tau}_j^2)_{1\leq j\leq n-1}$.

```r
s2 ← numeric(n)
for(j in 1:(n-1))
    s2[j] ← sum( E[1:(n-j+1)] * (N[1:(n-j+1), j]/E[1:(n-j+1)]
                - lambda[j])^2 )/(n - j)
s2[n] ← 0

tau2 ← numeric(n-1)
for(j in 1:(n-2))
    tau2[j] ← sum( C[1:(n-j), j] * (D[1:(n-j), j + 1]/C[1:(n-j), j]
                - delta[j])^2 )/(n - j - 1)
tau2[n-1] ← 0
```

j	1	2	3	4	5	6	7
$\widehat{\lambda}_j \times 10^3$	0.450	1.059	1.396	1.150	1.181	0.492	0.499
$\widehat{\delta}_j$	-0.359	0.072	-0.048	-0.054	0.070	0.033	
$\widehat{\sigma}_j$	0.0538	0.0737	0.1089	0.0795	0.0560	0.0575	0.0000
$\widehat{\tau}_j$	0.3874	1.2686	1.1768	3.4603	0.3034	0.0000	

On peut ensuite estimer les $C_{i,j}$ pour $i + j > n + 1$ avec le code R suivant.

```r
for(j in 2:n)
    C[(n-j+2):n, j] ← lambda[j]*E[(n-j+2):n] +
                          (1-delta[j-1])*C[(n-j+2):n, j-1]
```

Et en déduire les provisions.

```r
R_ ← function(C)
{
    return(C[, n] - rev(C[row(C) + col(C) == n + 1]))
```

```
}
R ← R_(C)
```

i	$\widehat{C}_{i,n}$	$\widehat{R}_i$
1	79.5	0
2	64.4	4.4
3	101.3	4.8
4	79.8	32.9
5	113.0	60.3
6	106.6	77.2
7	123.4	104.3

La provision totale est $\widehat{R} = 283,9$. En comparaison, avec la méthode de Chain Ladder Mack on obtient $\widehat{R} = 464,2$.

5.4 Le cas particulier du nombre de sinistres au-dessus d'un seuil

En réassurance, on s'intéresse parfois uniquement aux sinistres au-dessus d'un seuil qui joue le rôle de franchise. En assurance, on est parfois amené à séparer les petits sinistres des grands sinistres, également via un seuil, et d'estimer le nombre de sinistres au-dessus de ce seuil.

Le modèle précédemment développé s'adapte parfaitement à ce cas. En particulier :

- Les variables aléatoires $(N_{i,j})_{1\leq i,j\leq n}$ représentent le nombre de nouveaux sinistres qui dépassent un seuil fixé. Il peut s'agir de sinitres complètement nouveaux dont le coût estimé est supérieur au seuil, comme de sinistres déclarés antérieurement dont la valeur l'année précédente était sous le seuil.
- Les variables aléatoires $(D_{i,j})_{1\leq i,j\leq n}$ représentent le nombre de sinistres au-dessus du seuil en année de développement $j-1$ dont le coût estimé chute sous le seuil.

Nous avons à nouveau la matrice $(C_{i,j})_{1\leq i,j\leq n}$ qui représente cette fois-ci le nombre de sinistres survenus en année i et au-dessus du seuil en année de développement j. Par construction nous avons cette fois-ci la contrainte supplémentaire $D_{i,j} \geq 0$.

Nous avons l'hypothèse suivante :

Hypothèse 5.2.

H2' $N_{i,j} \mid \mathcal{F}_{i+j-2} \sim \mathcal{P}(\lambda_j E_i)$ *et* $D_{i,j+1} \mid \mathcal{F}_{i+j-1} \sim \mathcal{B}(C_{i,j}, \delta_j)$

Remarque 5.14. *L'hypothèse H2' implique les hypothèses H2 et H3.*

Démonstration. Pour *H2* il suffit d'appliquer l'espérance. Pour *H3*, il suffit d'appliquer la variance et de poser $\sigma_j^2 := \lambda_j$ et $\tau_j^2 := \delta_j(1-\delta_j)$. □

Proposition 5.15. *Sous H2', les estimateurs précédemment définis :*

$$\widehat{\lambda}_j := \frac{\sum_{i=1}^{n-j+1} N_{i,j}}{\sum_{i=1}^{n-j+1} E_i},$$

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

sont également les estimateurs du maximum de vraisemblance.

Démonstration. **Étape 1**. Pour $1 \leq j \leq n$, on observe $(N_{i,j} \mid \mathcal{F}_{i+j-2})_{1 \leq i \leq n-j+1}$ qui suivent indépendamment la loi $\mathcal{P}(\lambda_j E_i)$. La log-vraisemblance, notée $\ell(\lambda_j)$, s'écrit :

$$\ell(\lambda_j) = -\sum_{i=1}^{n-j+1} \lambda_j E_i + \sum_{i=1}^{n-j+1} N_{i,j} \log(\lambda_j E_i) - \sum_{i=1}^{n-j+1} \log(N_{i,j}!)$$
$$\ell(\lambda_j) \propto -\sum_{i=1}^{n-j+1} \lambda_j E_i + \sum_{i=1}^{n-j+1} N_{i,j} \log(\lambda_j)$$

On a :

$$\ell'(\lambda_j) = -\sum_{i=1}^{n-j+1} E_i + \frac{1}{\lambda_j} \sum_{i=1}^{n-j+1} N_{i,j}.$$

On en déduit la condition du premier ordre :

$$\ell'(\widehat{\lambda}_j) = 0 \Longleftrightarrow \widehat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} N_{i,j}}{\sum_{i=1}^{n-j+1} E_i}.$$

De plus,

$$\ell''(\widehat{\lambda}_j) = -\frac{1}{\widehat{\lambda}_j^2} \sum_{i=1}^{n-j+1} N_{i,j} < 0.$$

Étape 2. Pour $1 \leq j \leq n-1$, on observe $(D_{i,j+1} \mid \mathcal{F}_{i+j-1})_{1 \leq i \leq n-j}$ qui suivent indépendamment la loi $\mathcal{B}(C_{i,j}, \delta_j)$. La log-vraisemblance, notée $\ell(\delta_j \mid \mathcal{B}_j)$, s'écrit :

$$\ell(\delta_j \mid \mathcal{B}_j) = \sum_{i=1}^{n-j} \log \binom{C_{i,j}}{D_{i,j+1}} + \sum_{i=1}^{n-j} D_{i,j+1} \log(\delta_j) + \sum_{i=1}^{n-j} (C_{i,j} - D_{i,j+1}) \log(1 - \delta_j)$$
$$\ell(\delta_j \mid \mathcal{B}_j) \propto \sum_{i=1}^{n-j} D_{i,j+1} \log(\delta_j) + \sum_{i=1}^{n-j} (C_{i,j} - D_{i,j+1}) \log(1 - \delta_j)$$

On a :

$$\ell'(\delta_j \mid \mathcal{B}_j) = \frac{1}{\delta_j} \sum_{i=1}^{n-j} D_{i,j+1} - \frac{1}{1-\delta_j} \sum_{i=1}^{n-j} (C_{i,j} - D_{i,j+1}).$$

On en déduit la condition du premier ordre :

$$\begin{aligned}
\ell'(\widehat{\delta}_j \mid \mathcal{B}_j) = 0 &\Longleftrightarrow (1 - \widehat{\delta}_j) \sum_{i=1}^{n-j} D_{i,j+1} - \widehat{\delta}_j \sum_{i=1}^{n-j} (C_{i,j} - D_{i,j+1}) = 0 \\
&\Longleftrightarrow \sum_{i=1}^{n-j} D_{i,j+1} - \widehat{\delta}_j \sum_{i=1}^{n-j} C_{i,j} = 0 \\
&\Longleftrightarrow \widehat{\delta}_j := \frac{\sum_{i=1}^{n-j} D_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}
\end{aligned}$$

De plus,

$$\ell''(\widehat{\delta}_j \mid \mathcal{B}_j) = -\frac{1}{\widehat{\delta}_j^2}\sum_{i=1}^{n-j} D_{i,j+1} - \frac{1}{(1-\widehat{\delta}_j)^2}\sum_{i=1}^{n-j}(C_{i,j} - D_{i,j+1}) < 0$$

□

Proposition 5.16. *Les estimateurs* $\widehat{\lambda}_j$ *et* $\widehat{\delta}_j$ *sont efficaces.*

Démonstration. On calcule l'information de Fischer, notée $I(\lambda_j)$ associée à la log-vraisemblance $\ell(\lambda_j)$:

$$I(\lambda_j) = -\mathbb{E}\left[\ell''(\lambda_j) \mid \mathcal{B}_j\right] = \frac{1}{\lambda_j^2}\sum_{i=1}^{n-j+1} \lambda_j E_i = \frac{\sum_{i=1}^{n-j+1} E_i}{\lambda_j}.$$

Or, comme $\sigma_j^2 = \lambda_j$ par *H2'*, par le Lemme 5.6, on sait que :

$$Var(\widehat{\lambda}_j \mid \mathcal{B}_j) = \frac{\lambda_j}{\sum_{i=1}^{n-j+1} E_i} = I^{-1}(\lambda_j).$$

L'estimateur $\widehat{\lambda}_j$ est efficace.

On calcule l'information de Fischer, notée $I(\delta_j)$ associée à la log-vraisemblance $\ell(\delta_j)$:

$$I(\delta_j \mid \mathcal{B}_j) = -\mathbb{E}\left[\ell''(\delta_j \mid \mathcal{B}_j) \mid \mathcal{B}_j\right] = \frac{1}{\delta_j^2}\sum_{i=1}^{n-j} \delta_j C_{i,j} + \frac{1}{(1-\delta_j)^2}\sum_{i=1}^{n-j}(C_{i,j} - \delta_j C_{i,j}) = \frac{\sum_{i=1}^{n-j} C_{i,j}}{\delta_j(1-\delta_j)}.$$

Or, comme $\tau_j^2 = \delta_j(1-\delta_j)$ par *H2'*, par le Lemme 5.6, on sait que :

$$Var(\widehat{\delta}_j \mid \mathcal{B}_j) = \frac{\delta_j(1-\delta_j)}{\sum_{i=1}^{n-j} C_{i,j}} = I^{-1}(\delta j \mid \mathcal{B}_j).$$

L'estimateur $\widehat{\delta}_j$ est efficace.

□

On peut obtenir la loi des $C_{i,j}$. Nous allons utiliser le lemme suivant.

Lemme 5.17. *Soit N une variable aléatoire de loi $\mathcal{P}(\lambda)$ avec $\lambda > 0$ et D une variable aléatoire telle que $D \mid N \sim \mathcal{B}(N, p)$ avec $p > 0$. Alors*

$$N - D \sim \mathcal{P}(\lambda(1-p)).$$

Démonstration. Soit $k \in \mathbb{N}$.

$$
\begin{aligned}
\mathbb{P}(N-D=k) &= \mathbb{P}\left(\{N-D=k\}\cap\left\{\bigcup_{n\in\mathbb{N}}\{N=n\}\right\}\right)\\
&= \sum_{n\in\mathbb{N}}\mathbb{P}\left(\{N-D=k\}\cap\{N=n\}\right)\\
&= \sum_{n\in\mathbb{N}}\mathbb{P}\left(\{D=n-k\}\mid\{N=n\}\right)\mathbb{P}\left(N=n\right)\\
&= \sum_{n\geq k}\frac{n!}{k!(n-k)!}p^{n-k}(1-p)^k\frac{e^{-\lambda}\lambda^n}{n!}\\
&= \frac{e^{-\lambda}(1-p)^k}{k!}\sum_{n\geq 0}\frac{p^n\lambda^{n+k}}{n!}\\
&= \frac{e^{-\lambda}\left[\lambda(1-p)\right]^k}{k!}\sum_{n\geq 0}\frac{(p\lambda)^n}{n!}\\
&= \frac{e^{-\lambda(1-p)}\left[\lambda(1-p)\right]^k}{k!}.
\end{aligned}
$$

□

Lemme 5.18. *Pour tout* $1 \leq i, j \leq n$*, on a*

$$C_{i,j} \sim \mathcal{P}\left(\lambda_{i,j}\right),$$

avec

$$\lambda_{i,j} := E_i \sum_{k=1}^{j}\lambda_k\left(\prod_{\ell=k}^{j-1}(1-\delta_\ell)\right).$$

Démonstration. On raisonne par récurrence. On fixe $1 \leq i \leq n$. On a $C_{i,1} = N_{i,1} \sim \mathcal{P}(\lambda_1 E_i)$. On suppose par hypothèse de récurrence que $C_{i,j}$ suit une loi de Poisson de paramètre $\lambda_{i,j} := \sum_{k=1}^{j}\lambda_k E_i\left(\prod_{\ell=k}^{j-1}(1-\delta_\ell)\right)$

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

Sous *H2'*, par le lemme précédent,

$$C_{i,j} - D_{i,j+1} \sim \mathcal{P}\left(\lambda_{i,j}(1-\delta_j)\right),$$

puis

$$C_{i,j+1} \sim \mathcal{P}\left(\lambda_{j+1}E_i + \lambda_{i,j}(1-\delta_j)\right).$$

□

On peut également obtenir la loi des $C_{i,j}$ conditionnellement à $\mathcal{F}_n$, pour $i+j > n+1$, et en particulier, de $C_{i,n}$. Nous allons avoir besoin du lemme suivant.

Lemme 5.19. *Soit* $N \in \mathbb{N}^*$*. Soit* $D_1 \sim \mathcal{B}(N, p_1)$*, puis, par réccurence,* $D_{n+1} \mid \{\sum_{i=1}^{n} D_i = d\} \sim \mathcal{B}(N-d, p_{n+1})$ *pour* $n \geq 1$*. Alors*

$$N - \sum_{i=1}^{n} D_i \sim \mathcal{B}\left(N, \prod_{i=1}^{n}(1-p_i)\right).$$

Démonstration. On procède par récurrence. Il est clair que $N - D_1 \sim \mathcal{B}(N, 1-p_1)$. Soit $n \geq 1$, on suppose que

$$N - \sum_{i=1}^{n} D_i \sim \mathcal{B}\left(N, \prod_{i=1}^{n}(1-p_i)\right).$$

Soit $k \in \{0, \dots, N\}$. On note $1 - \pi_n := \prod_{i=1}^{n}(1-p_i)$. Alors

$$\begin{aligned}
\mathbb{P}\left[N - \sum_{i=1}^{n+1} D_i = k\right] &= \sum_{d=0}^{N} \mathbb{P}\left[N - \sum_{i=1}^{n+1} D_i = k \;\middle|\; \sum_{i=1}^{n} D_i = d\right] \mathbb{P}\left(\sum_{i=1}^{n} D_i = d\right) \\
&= \sum_{d=0}^{N} \mathbb{P}\left[D_{n+1} = N - d - k \;\middle|\; \sum_{i=1}^{n} D_i = d\right] \mathbb{P}\left(N - \sum_{i=1}^{n} D_i = N - d\right) \\
&= \sum_{d=0}^{N-k} \frac{(N-d)!}{k!(N-d-k)!} p_{n+1}^{N-d-k}(1-p_{n+1})^k \frac{N!}{d!(N-d)!} \pi_n^d (1-\pi_n)^{N-d} \\
&= \frac{N!}{k!(N-k)!}[(1-p_{n+1})(1-\pi_n)]^k \sum_{d=0}^{N-k} \frac{(N-k)!}{(N-d-k)!d!} p_{n+1}^{N-d-k} \pi_n^d (1-\pi_n)^{N-d-k} \\
&= \frac{N!}{k!(N-k)!}(1-\pi_{n+1})^k \left(\pi_n + p_{n+1}(1-\pi_n)\right)^{N-k} \\
&= \frac{N!}{k!(N-k)!}(1-\pi_{n+1})^k \pi_n^{N-k}.
\end{aligned}$$

□

Lemme 5.20. *Pour tout* $1 \leq i, j \leq n$ *tel que* $i + j > n + 1$*, on a*

$$C_{i,j} \mid \mathcal{F}_n \sim \mathcal{B}(C_{i,n-i+1}, \delta_{i,j}) + \mathcal{P}\left(\lambda_{i,j}\right),$$

où la somme des lois est indépendante avec

$$\begin{aligned}
\delta_{i,j} &:= \prod_{j=n-i+1}^{j-1} (1-\delta_k), \\
\lambda_{i,j} &:= \sum_{k=n-i+2}^{j} \lambda_k E_i \left(\prod_{\ell=k}^{j-1}(1-\delta_\ell)\right).
\end{aligned}$$

Démonstration. On démontre par récurrence. On a

$$C_{i,n-i+2} = C_{i,n-i+1} - D_{i,n-i+2} + N_{i,n-i+2}.$$

On remarque que $C_{i,n-i+1} - D_{i,n-i+2} \mid \mathcal{F}_n \sim \mathcal{B}(C_{i,n-i+1}, 1 - \delta_{i,n-i+1})$ et que $N_{i,n-i+2} \mid \mathcal{F}_n \sim \mathcal{P}(\lambda_{n-i+2} E_i)$. □

Lemme 5.21. *Sous H1 et H2', nous avons :*

$$\widehat{\lambda}_j \mid \mathcal{B}_{j-1} \sim \frac{\mathcal{P}\left(\lambda_j \sum_{i=1}^{n-j+1} E_i\right)}{\sum_{i=1}^{n-j+1} E_i}$$

$$\widehat{\delta}_j \mid \mathcal{B}_j \sim \frac{\mathcal{B}\left(\sum_{i=1}^{n-j} C_{i,j}, \delta_j\right)}{\sum_{i=1}^{n-j} C_{i,j}}$$

Démonstration. Par définition, on a :

$$\widehat{\lambda}_j := \frac{\sum_{i=1}^{n-j+1} N_{i,j}}{\sum_{i=1}^{n-j+1} E_i},$$
$$\widehat{\delta}_j := \frac{\sum_{i=1}^{n-j} D_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}.$$

Par *H1*, les $N_{i,j}$ sont indépendants entre eux, tout comme les $D_{i,j+1}$. Par *H2'*, comme $N_{i,j} \mid \mathcal{B}_{j-1} \sim \mathcal{P}(\lambda_j E_i)$ et $D_{i,j+1} \mid \mathcal{B}_j \sim \mathcal{B}(C_{i,j}, \delta_j)$, on obtient le résultat. □

Cette fois-ci, on ne définit plus des provisions, mais le nombre estimé de nouveaux sinistres qui seront au-dessus du seuil : $\widehat{M}_i := \widehat{C}_{i,n} - C_{i,n-i+1}$.

5.5 Exemple

Nous avons le triangle cumulé suivant auquel on associe l'exposition.

i	E_i	$C_{i,1}$	$C_{i,2}$	$C_{i,3}$	$C_{i,4}$	$C_{i,5}$
1	20	10	13	11	12	13
2	25	7	8	9	9	
3	32	10	8	9		
4	39	18	15			
5	42	19				

Celui-ci se décompose en la matrice N :

i	$N_{i,1}$	$N_{i,2}$	$N_{i,3}$	$N_{i,4}$	$N_{i,5}$
1	10	5	0	1	1
2	7	6	3	1	
3	10	3	2		
4	18	5			
5	19				

et D :

i	$D_{i,1}$	$D_{i,2}$	$D_{i,3}$	$D_{i,4}$	$D_{i,5}$
1	0	2	2	0	0
2	0	5	2	1	
3	0	5	1		
4	0	8			
5	0				

Nous importons les données puis nous calculons les $(\widehat{\lambda}_j)_{1\leq j\leq n}$ et $(\widehat{\delta}_j)_{1\leq j\leq n-1}$. Nous pouvons

j	1	2	3	4	5
$\widehat{\lambda}_j$	0.405	0.164	0.065	0.044	0.05
$\widehat{\delta}_j$	0.444	0.172	0.050	0.000	

en déduire le nombre moyen de sinistres au dessus du seuil à venir : Le nombre moyen de sinistres à venir est $\widehat{M} = 7,97$.

i	$\widehat{C}_{i,n}$	$\widehat{M}_i$
1	13	0
2	10.25	1.25
3	11.57	2.57
4	17.88	2.88
5	20.26	1.26

6 Méthode du boostrap

La méthode du boostrap s'appuie sur un rééchantillonnage de l'erreur. Il permet ensuite de calculer, par exemple, des intervalles de confiance sur différents estimateurs.

Le cas le plus simple est celui de l'estimation d'un intervalle de confiance pour la moyenne empirique. Si on observe : $(X_1, \ldots, X_n)$, un échantillon i.i.d. de L^2, en notant $\mu := \mathbb{E}(X_1)$ et $\sigma := \sqrt{Var(X_1)}$, on sait par le théorème central limite que

$$\sqrt{n}\frac{(\bar{X}_n - \mu)}{\sigma} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1),$$

ce qui permet d'en déduire un intervalle de confiance **asymptotique**. L'idée du bootstrap est d'utiliser la variabilité intrinsèque de l'échantillon afin d'en construire une distribution estimée de l'estimateur autour de μ.

Dans ce cas, on rééchantillonne $(X_1, \ldots, X_n)$, c'est à dire qu'on introduit :

$$(X_1^m, \ldots, X_n^m) \overset{i.i.d.}{\sim} \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i},\ 1 \leq m \leq N.$$

On construit ainsi N nouveaux échantillons de taille n où chaque X_i^m suit de manière indépendante la loi uniforme sur l'échantillon observé $(X_1, \ldots, X_n)$. On calcule ensuite les moyennes empiriques associées $(\widehat{\mu}^m)_{1\leq m\leq N}$ ce qui donne une distribution de l'estimateur.

Par exemple, si on observe $(5, 5, 7, 4, 11, 4, 10, 8, 6, 7, 12, 5)$, on a $\widehat{\mu} = 7$ et $\widehat{\sigma} \approx 2.73$. On en déduit sous le TCL l'intervalle au niveau 95% : [5.46, 8.54].

En revanche, si on applique la procédure du boostrap un grand nombre de fois, et qu'on prend l'intervalle de confiance sur la distribution de $(\widehat{\mu}^m)_{1\leq m\leq N}$, on obtient au niveau 95% : [5.58, 8.50].

Dans notre cas, il s'agit de *repérer* ce qui est (ou sera supposé) i.i.d. Nous allons dans un premier temps regarder le cas du modèle de Mack puis dans un second temps celui de Schnieper.

6.1 Modèle de Mack

Dans le modèle de Mack, nous avions comme hypothèse : $\mathbb{E}(C_{i,j+1} \mid \mathcal{B}_j) = f_j C_{i,j}$ et $Var(C_{i,j+1} \mid \mathcal{B}_j) = \sigma_j^2 C_{i,j}$.

On définit le résidu de Pearson de la manière suivante :

$$r_{i,j} := \frac{\frac{C_{i,j+1}}{C_{i,j}} - f_j}{\frac{\sigma_j}{\sqrt{C_{i,j}}}} = \frac{C_{i,j+1} - f_j C_{i,j}}{\sigma_j\sqrt{C_{i,j}}}.$$

ce qui implique que

$$\mathbb{E}(r_{i,j} \mid \mathcal{B}_j) = 0, \quad Var(r_{i,j} \mid \mathcal{B}_j) = 1.$$

On va faire l'hypothèse supplémentaire que $r_{i,j}$ a la même loi pour $1 \leq i \leq n$ et $1 \leq j \leq n-1$. Comme les paramètres f_j et σ_j sont inconnus, on les remplace par leurs estimateurs, on définit pour $i + j + 1 \leq n + 1$:

$$\widehat{r}_{i,j} := \frac{C_{i,j+1} - \widehat{f}_j C_{i,j}}{\widehat{\sigma}_j \sqrt{C_{i,j}}}.$$

C'est l'échantillon sur lequel nous allons appliquer le bootstrap. On simule $(\widehat{r}^m_{i,j})_{1 \leq m \leq N}$ en tirant uniformément et indépendamment dans l'échantillon $(\widehat{r}_{i,j})_{i+j \leq n}$. On recalcule le triangle supérieur avec la relation :

$$\begin{aligned} C^m_{i,1} &:= C_{i,1}, \\ C^m_{i,j+1} &:= \widehat{f}_j C^m_{i,j} + \widehat{r}^m_{i,j} \widehat{\sigma}_j \sqrt{C^m_{i,j}}. \end{aligned}$$

Cela permet d'en déduire un nouveau triangle sur lequel on estime à nouveau les paramètres f_j et σ_j, on note $\widehat{f}^m_j$ et $\widehat{\sigma}^m_j$ les estimateurs correspondants. Ensuite, pour la partie inconnue du triangle, on la simule selon :

$$\begin{aligned} C^m_{i,n-i+1} &:= C_{i,n-i+1}, \\ C^m_{i,j+1} &:= \widehat{f}^m_j C^m_{i,j} + \widehat{r}^m_{i,j} \widehat{\sigma}^m_j \sqrt{C^m_{i,j}}. \end{aligned}$$

On peut ensuite calculer :

$$R^m_i = C^m_{i,n} - C_{i,n-i+1}, \ 1 \leq i \leq n, \ 1 \leq m \leq N.$$

Ces simulations tiennent à la fois compte de l'erreur d'estimation des paramètres (première étape de reconstruction du triangle où on estime à nouveau les paramètres) et du bruit (seconde étape où on simule la partie inconnue du triangle).

6.2 Modèle de Schnieper

Nous allons cette fois-ci définir des résidus pour chacun des triangles. Le triangle des nouveaux sinistres, nous avions comme hypothèse : $\mathbb{E}(N_{i,j} \mid \mathcal{B}_{j-1}) = \lambda_j E_i$ et $Var(N_{i,j} \mid \mathcal{B}_{j-1}) = \sigma_j^2 E_i$.

On définit le résidu de Pearson de la manière suivante :

$$n_{i,j} := \frac{\frac{N_{i,j}}{E_i} - \lambda_j}{\frac{\sigma_j}{\sqrt{E_i}}} = \frac{N_{i,j} - \lambda_j E_i}{\sigma_j \sqrt{E_i}}.$$

De manière analogue, compte tenu des hypothèse sur $D_{i,j+1}$, on définit :

$$d_{i,j} := \frac{\frac{D_{i,j+1}}{C_{i,j}} - \delta_j}{\frac{\tau_j}{\sqrt{C_{i,j}}}} = \frac{D_{i,j+1} - \delta_j C_{i,j}}{\tau_j \sqrt{C_{i,j}}}.$$

On va faire l'hypothèse supplémentaire que les $n_{i,j}$ ont la même loi pour $1 \leq i \leq n$ et $1 \leq j \leq n$ et de même pour les $d_{i,j}$. Comme les paramètres λ_j et σ_j sont inconnus, on les remplace par leurs estimateurs, on définit pour $i + j \leq n + 1$:

$$\widehat{n}_{i,j} := \frac{N_{i,j} - \widehat{\lambda}_j E_i}{\widehat{\sigma}_j \sqrt{E_i}}.$$

Et de même :

$$\widehat{d}_{i,j} := \frac{D_{i,j+1} - \widehat{\delta}_j C_{i,j}}{\widehat{\tau}_j \sqrt{C_{i,j}}}.$$

On simule $(\widehat{n}_{i,j}^m)_{1 \le m \le N}$ et $(\widehat{d}_{i,j}^m)_{1 \le m \le N}$ en tirant uniformément et indépendamment respectivement dans l'échantillon $(\widehat{n}_{i,j})_{i+j \le n+1}$ et $(\widehat{d}_{i,j})_{i+j \le n}$. On recalcule le triangle supérieur avec la relation :

$$\begin{aligned} C_{i,1}^m &:= \widehat{\lambda}_1 E_i + \widehat{n}_{i,j}^m \widehat{\sigma}_1 \sqrt{E_i}, \\ C_{i,j+1}^m &:= \widehat{\lambda}_{j+1} E_i + \widehat{n}_{i,j}^m \widehat{\sigma}_{j+1} \sqrt{E_i} + (1 - \widehat{\delta}_j) C_{i,j}^m + \widehat{d}_{i,j}^m \widehat{\tau}_j \sqrt{C_{i,j}^m}. \end{aligned}$$

Cela permet d'en déduire un nouveau triangle sur lequel on estime à nouveau les paramètres λ_j et δ_j. On note respectivement $\widehat{\lambda}_j^m$ et $\widehat{\delta}_j^m$ les estimateurs correspondants. Ensuite, pour la partie inconnue du triangle, on la simule selon :

$$\begin{aligned} C_{i,n-i+1}^m &:= C_{i,n-i+1}, \\ C_{i,j+1}^m &:= \widehat{\lambda}_{j+1}^m E_i + \widehat{n}_{i,j}^m \widehat{\sigma}_{j+1}^m \sqrt{E_i} + (1 - \widehat{\delta}_j^m) C_{i,j}^m + \widehat{d}_{i,j}^m \widehat{\tau}_j^m \sqrt{C_{i,j}^m}. \end{aligned}$$

On peut ensuite calculer :

$$R_i^m = C_{i,n}^m - C_{i,n-i+1}, \ 1 \le i \le n, \ 1 \le m \le N.$$

Ces simulations tiennent à la fois compte de l'erreur d'estimation des paramètres (première étape de reconstruction du triangle où on estime à nouveau les paramètres) et du bruit (seconde étape où on simule la partie inconnue du triangle).

Pour le cas des **nombres** au-dessus d'un seuil, sous les hypothèses plus fortes de distribution, nous ne passons plus par les rédidus. Nous simulons directement les valeurs selon les lois supposées. Plus précisément, pour le triangle supérieur, nous simulons :

$$\begin{aligned} C_{i,1}^m &\sim \mathcal{P}\left(\widehat{\lambda}_1 E_i\right), \\ C_{i,j+1}^m &\sim C_{i,j}^m + \mathcal{P}\left(\widehat{\lambda}_{j+1} E_i\right) - \mathcal{B}\left(C_{i,j}^m, \widehat{\delta}_j\right). \end{aligned}$$

Cela permet à nouveau d'en déduire un nouveau triangle sur lequel on estime une nouvelle fois les paramètres λ_j et δ_j. On note à nouveau respectivement $\widehat{\lambda}_j^m$ et $\widehat{\delta}_j^m$ les estimateurs correspondants. Ensuite, pour la partie inconnue du triangle, on la simule selon :

$$\begin{aligned} C_{i,n-i+1}^m &:= C_{i,n-i+1}, \\ C_{i,j+1}^m &:= C_{i,j}^m + \mathcal{P}\left(\widehat{\lambda}_{j+1}^m E_i\right) - \mathcal{B}\left(C_{i,j}^m, \widehat{\delta}_j^m\right). \end{aligned}$$

On peut ensuite calculer le nombre de sinistres manquant au-dessus du seuil :

$$M_i^m = C_{i,n}^m - C_{i,n-i+1}, \ 1 \le i \le n, \ 1 \le m \le N.$$

6.3 Exemples

Nous regardons un exemple pour le modèle de Chain Ladder Mack et un pour celui de Schnieper.

6.3.1 Modèle de Chain Ladder Mack

Nous reprenons le triangle de l'Exemple 4.3. Nous calculons les résidus.

```
r_ ← function(C, f, s)
{
    r ← numeric(0)
    for(j in 1:(n−1))
        r ← c(r, (C[1:(n−j), j+1] − f[j]*C[1:(n−j), j])/
             (s[j]*sqrt(C[1:(n−j), j])))
    return(r)
}
s ← sqrt(s2)
r ← r_(C, f, s)
```

Puis nous définissons deux fonctions. La première permet de simuler, par bootstrap, la partie supérieure du triangle sur laquelle nous pouvons réestimer les paramètres, par après. La seconde permet de simuler la partie inférieure et d'en déduire une simulation des réserves, par après.

```
Cbootstrap_ ← function(C, f, s, r)
{
    Cm ← C
    for(j in 1:(n−1))
    {
        rm ← sample(r, n−j, replace = TRUE)
        Cm_j ← Cm[1:(n−j), j]
        Cm[1:(n−j), j+1] ← Cm_j*f[j] + s[j]*sqrt(Cm_j)*rm
    }
    return(Cm)
}

hatCbootstrap_ ← function(C, fm, sm, r)
{
    Cm ← C
    for(j in 1:(n−1))
    {
        rm ← sample(r, j, replace = TRUE)
        Cm_j ← Cm[(n−j+1):n, j]
        Cm[(n−j+1):n, j+1] ← Cm_j*fm[j] + sm[j]*sqrt(Cm_j)*rm
    }
    return(Cm)
}
```

Nous appliquons maintenant l'algorithme du bootstrap pour estimer la distribution des réserves.

```
N ← 10000
Rm ← numeric(N)
for(m in 1:N)
{
    Cm ← Cbootstrap_(C, f, s, r)
    fm ← f_(Cm)
    sm ← sqrt(s2_(Cm, fm))

    Cm ← hatCbootstrap_(C, fm, sm, r)
    Rm[m] ← sum(R_(Cm))
}
```

On obtient la distribution des réserves $(R^m)_{1 \leq m \leq N}$. Sa moyenne (sur un jeu de simulation) est $\bar{R}^m = 18907182$ qui est supérieure de 1,0% à l'estimation directe par Chain Ladder Mack ($\widehat{R} = 18680848$). Son écart-type relatif est de 10,8%, inférieur à celui obtenu par Chain Ladder Mack (13,1%).

6.3.2 Modèle de Schnieper

Nous reprenons le triangle de l'Exemple 5.3. Nous calculons les résidus.

```
rn_ ← function(N, E, lambda, s)
{
    r ← numeric(0)
    for(j in 1:(n-1))
        r ← c(r, (N[1:(n-j+1), j] - lambda[j]*E[1:(n-j+1)])/(s[j]*
           sqrt(E[1:(n-j+1)])))
    return(r)
}
s ← sqrt(s2)
rn ← rn_(N, E, lambda, s)

rd_ ← function(D, C, delta, tau)
{
    r ← numeric(0)
    for(j in 1:(n-2))
        r ← c(r, (D[1:(n-j), j+1] - delta[j]*C[1:(n-j), j])/(tau[j
           ]*sqrt(C[1:(n-j), j])))
    return(r)
}
tau ← sqrt(tau2)
rd ← rd_(D, C, delta, tau)
```

Puis nous définissons deux fonctions. La première permet de simuler, par bootstrap, la partie supérieure du triangle N. La seconde fait de même pour D et calcul C. La troisième permet de compléter la partie inférieure et d'en déduire une simulation des réserves, par après.

Etant donné que le coût des nouveau sinistre peut être négatif tout comme la variation du montant des sinistres peut être supérieure au montant total cumulé, nous les bornons.

```
Nbootstrap_ ← function(E, lambda, s, rn)
```

```
{
    Nm <- matrix(0, n, n)
    Nm[, 1] <- lambda[1]*E + s[1]*sqrt(E)*sample(rn, n, replace =
       TRUE)
    for(j in 1:(n-1))
    {
        rnm <- sample(rn, n-j, replace = TRUE)
        Em_j <- E[1:(n-j)]
        Nm[1:(n-j), j+1] <- lambda[j+1]*Em_j + s[j]*sqrt(Em_j)*rnm
    }
    return(Nm*(Nm >= 0))
}
DCbootstrap_ <- function(Nm, delta, tau, rd)
{
    Cm <- Nm
    Dm <- matrix(0, n, n)
    for(j in 1:(n-1))
    {
        rdm <- sample(rd, n-j, replace = TRUE)
        Cm_j <- Cm[1:(n-j), j]
        Dm[1:(n-j), j+1] <- delta[j]*Cm_j + tau[j]*sqrt(Cm_j)*rdm
        Dm[1:(n-j), j+1][Dm[1:(n-j), j+1] > Cm_j] <- 0
        Cm[1:(n-j), j+1] <- Cm_j + Nm[1:(n-j), j+1] - Dm[1:(n-j), j
           +1]
     }
    return(list(Dm = Dm, Cm = Cm))
}
hatCbootstrap_ <- function(C, E, lambdam, deltam, sm, taum, rn, rd)
{
    Cm <- C
    for(j in 1:(n-1))
    {
        rnm <- sample(rn, j, replace = TRUE)
        rdm <- sample(rd, j, replace = TRUE)
        Cm_j <- Cm[(n-j+1):n, j]
        Cm[(n-j+1):n, j+1] = lambdam[j+1]*E[(n-j+1):n] + sm[j]*sqrt
           (E[(n-j+1):n])*rnm + Cm_j*(1-deltam[j]) + taum[j]*sqrt(
           Cm_j)*rdm
        Cm[Cm < 0] <- 0
    }
    return(Cm)
}
```

Nous appliquons maintenant l'algorithme du bootstrap pour estimer la distribution des réserves.

```
N <- 10**6
Rm <- numeric(N)
```

```
tau2_ ← function(D, C, delta)
{
    tau2 ← numeric(n−1)
    for(j in 1:(n−2))
        tau2[j] ← sum( C[1:(n−j), j] * (D[1:(n−j), j + 1]/C[1:(n−j)
           , j] − delta[j])^2 , na.rm = TRUE)/(sum(C[1:(n−j), j] !=
            0) − 1)
    return(tau2)
}

for(m in 1:N)
{
    Nm ← Nbootstrap_(E, lambda, s, rn)
    lambdam ← lambda_(Nm, E)
    sm ← sqrt(s2_(Nm, E, lambdam))

    L ← DCbootstrap_(Nm, delta, tau, rd)
    deltam ← delta_(L$Dm, L$Cm)
    taum ← sqrt(tau2_(L$Dm, L$Cm, deltam))

    Cm ← hatCbootstrap_(C, E, lambdam, deltam, sm, taum, rn, rd)
    Rm[m] ← sum(R_(Cm))
}
```

On obtient la distribution des réserves $(R^m)_{1\leq m\leq N}$. Sa moyenne (sur un jeu de simulation) est $\bar{R}^m = 302,1$ qui est supérieure de 6,4% à l'estimation directe par Schnieper ($\widehat{R} = 283,9$). À noter que dans certaines simulations, $N^m_{i,j}$ ou $C^m_{i,j}$ pouvaient être négatifs, et ont été mis à zéro ce qui peut augmenter légèrement l'estimation. Son écart-type relatif est de 35,0%.

Enfin, pour le cas des nombres :

```
Nbootstrap_ ← function(E, lambda)
{
    Nm ← matrix(0, n, n)
    Nm[, 1] ← rpois(n, lambda[1]*E)
    for(j in 1:(n−1))
        Nm[1:(n−j), j+1] ← rpois(n−j, lambda[j+1]*E[1:(n−j)])
    return(Nm)
}

DCbootstrap_ ← function(Nm, delta)
{
    Cm ← Nm
    Dm ← matrix(0, n, n)
    for(j in 1:(n−1))
    {
        Cm_j ← Cm[1:(n−j), j]
        Dm[1:(n−j), j+1] ← rbinom(n−j, Cm_j, delta[j])
        Cm[1:(n−j), j+1] ← Cm_j + Nm[1:(n−j), j+1] − Dm[1:(n−j), j
           +1]
```

```
    }
    return(list(Dm = Dm, Cm = Cm))
}

hatCbootstrap_ ← function(C, E, lambdam, deltam)
{
    Cm ← C
    for(j in 1:(n−1))
    {
        Cm_j ← Cm[(n−j+1):n, j]
        Cm[(n−j+1):n, j+1] = Cm_j + rpois(j, lambdam[j+1]*E[(n−j+1)
           :n]) − rbinom(j, Cm_j, deltam[j])
        Cm[Cm < 0] ← 0
    }
    return(Cm)
}

N ← 10**6
Rm ← numeric(N)

for(m in 1:N)
{
    Nm ← Nbootstrap_(E, lambda)
    lambdam ← lambda_(Nm, E)

    L ← DCbootstrap_(Nm, delta)
    deltam ← delta_(L$Cm, L$Dm)
    deltam[is.nan(deltam)] ← 0

    Cm ← hatCbootstrap_(C, E, lambdam, deltam)
    Mm[m] ← sum(R_(Cm))
}
```

On obtient la distribution du nombre de sinistres moyen à venir au dessus du seuil. $(M^m)_{1\leq m\leq N}$. Sa moyenne (sur un jeu de simulations) est $\widehat{M}^m = 7,94$. À noter que dans certaines simulations, $C_{i,j}^m$ peut être nuls et ont été et δ a été mis à zéro. Son écart-type relatif est de 130,2% (10,34 de manière absolue).

7 Prise en compte de l'inflation

L'inflation joue un rôle important. Les sinistres qui ont lieu dans une année i sont réglés à différents moments. De plus, le coût des sinistres sur chaque année est touché par l'inflation. Il convient donc de déflater le triangle, afin de travailler en **euros constants**.

Si on suppose que $X_{i,j}$ (le coût incrémental) correspond au nouveau coût réel ou la charge estimée à venir, évaluée en année $i+j$, et que $(I_m)_{m\geq 1}$ est l'indice des prix associé au secteur de notre triangle, alors on introduit :

$$X'_{i,j} := X_{i,j} \frac{I_{n+1}}{I_{i+j}}.$$

On en déduit le nouveau triangle cumulé dans le cas de Chain Ladder Mack (et de manière analogue dans le cas de Schnieper). On fait ensuite les estimations sur les triangles déflatés. Enfin, pour compléter le triangle, on applique l'inflation en multipliant par $\frac{I_{i+j}}{I_{n+1}}$.

Références

[1] Nicolas Baradel. Cours de théorie du risque. `https://nicolasbaradel.fr/enseignement/ressources/cours_theorie_du_risque.pdf`.

[2] Thomas Mack. Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin : The Journal of the IAA*, 23(2) :213–225, 1993.

[3] R Schnieper. Separating true ibnr and ibner claims 1. *ASTIN Bulletin : The Journal of the IAA*, 21(1) :111–127, 1991.