# ASSURANCE NON-VIE

**Année universitaire 2004-2005**

# Théorie de la crédibilité

**Pierre-E. THEROND**

ptherond@jwa.fr

# 1. Introduction

Développée par les écoles suisse et scandinave, la théorie de la crédibilité repose sur les principes de l'inférence bayesienne et a notamment pour objectif la mise à jour des primes pures à partir de l'historique des sinistres des contrats.

Cette théorie s'est développée en parallèle de la statistique bayesienne dont elle n'a pas intégré tous les résultats. Nous présentons ici une introduction à cette théorie et insistons tout particulièrement sur le modèle de Bühlmann. Si le modèle de Bühlmann est rarement utilisable tel quel en pratique, il contient néanmoins tous les ingrédients théoriques permettant de se tourner vers le modèle de Bühlmann-Straub ou encore vers le modèle de crédibilité hiérarchique de Jewell. Ce dernier modèle est détaillé et appliqué dans le cas de la tarification du risque incendie pour les entreprises dans PARTRAT et BESSON [2004].

## 1.1. Prime de crédibilité

Intéressons-nous à une société d'assurance qui dispose d'un portefeuille de $n$ contrats homogènes auxquels sont associées les variables aléatoires $Y_{jt}$ pour $t = 1, \ldots, T$ et $j = 1, \ldots, n$ où $Y_{jt}$ représente le montant annuel de sinistres du contrat *j* pour la *t*-ème année. Faisons l'hypothèse que pour chaque contrat *j*, les montants de sinistres annuels $Y_{j1}, \ldots, Y_{jT}$ sont indépendants et identiquement distribués.

Supposons en outre que ce portefeuille n'ait pas fait l'objet d'une segmentation *a priori* et qu'il soit demandé en $t = 0$ le montant $p_{coll}$ à tous les assurés. L'assureur peut être tenté de personnaliser les primes *a posteriori* en intégrant l'information qui lui est fournie par l'historique des sinistres des contrats. Cette mise à jour des tarifs s'effectue donc « au mérite » en récompensant les contrats les moins sinistrés, ce qui a pour conséquence de limiter l'aléa moral.

Considérons le contrat *j* et plaçons-nous en $T$. L'assureur dispose du passé $\left(y_{j1}, \ldots, y_{jT}\right)$ où $y_{jt}$ représente la charge annuelle des sinistres engendrés par la police *j* lors de la *t*-ème année. En moyenne, le contrat *j* a engendré une charge annuelle moyenne de sinistres de

$$\overline{p}_j = \frac{y_{j1} + \ldots + y_{jT}}{T}.$$

Réclamer $\overline{p}_j$ à l'assuré du contrat *j* reviendrait à nier le principe de mutualisation qui est le fondement de l'activité d'assurance, aussi l'assureur peut avoir recours à un compromis qui le conduit à réclamer pour la période $[T, T+1]$ la prime de crédibilité

$$p_j = (1 - \alpha) p_{coll} + \alpha \overline{p}_j ,$$

où $\alpha$ est le facteur de crédibilité. Cette appellation provient du fait que $\alpha$ sera d'autant plus grand que l'on accorde de crédit à $\overline{p}_j$, *i. e.* au passé sinistre de l'assuré *j*.

## 1.2. Exemple : bons risques / mauvais risques[1]

Considérons un portefeuille regroupant exclusivement des bons risques et des mauvais risques. Notons $B$ l'événement « être un bon risque » et $B^c$ son complémentaire « être un mauvais risque ». Notons par ailleurs $N_k$ le nombre de sinistres déclarés au cours de l'année $k$ et supposons que :

$$\begin{cases} \mathbf{Pr}\left[N_k = 1 \mid B\right] = 0{,}2 = 1 - \mathbf{Pr}\left[N_k = 0 \mid B\right] \\ \mathbf{Pr}\left[N_k = 1 \mid B^c\right] = 0{,}8 = 1 - \mathbf{Pr}\left[N_k = 0 \mid B^c\right]. \end{cases}$$

Les assurés ont donc au plus un sinistre par année et les bons risques en ont en moyenne quatre fois moins que les mauvais risques.

Supposons par ailleurs que les montants de sinistres sont des variables aléatoires indépendantes et identiquement distribuées de moyenne 1.

Quand un nouvel assuré arrive dans le portefeuille, comme l'assureur ne segmente pas son portefeuille *a priori* et qu'il ne dispose pas d'historique sinistre de cet individu, il va lui demander

$$p_{coll} = \mathbf{Pr}\left[N_k = 1 \mid B\right]\mathbf{Pr}\left[B\right] + \mathbf{Pr}\left[N_k = 1 \mid B^c\right]\mathbf{Pr}\left[B^c\right].$$

En supposant que les bons risques représentent 50 % des assurés, il vient

$$p_{coll} = 0{,}2 * 0{,}5 + 0{,}8 * 0{,}5 = 0{,}5.$$

Ce calcul revient à faire l'hypothèse que la probabilité que le nouvel assuré soit un bon risque est identique à celle observée dans le portefeuille. Il nous a permis de déterminer la prime pure *a priori*.

Intéressons-nous à présent à la mise à jour du tarif d'un assuré dont on dispose de trois années d'historiques $\left(N_1 = 1, N_2 = 1, N_3 = 0\right)$. Comme, conditionnellement à la qualité d'un risque, les variables aléatoires nombres annuels de sinistres $N_k$ sont indépendantes, on peut écrire

$$\begin{aligned} & \mathbf{Pr}\left[N_1 = 1, N_2 = 1, N_3 = 0\right] \\ = \ & \mathbf{Pr}\left[N_1 = 1, N_2 = 1, N_3 = 0 \mid B\right]\mathbf{Pr}\left[B\right] + \mathbf{Pr}\left[N_1 = 1, N_2 = 1, N_3 = 0 \mid B^c\right]\mathbf{Pr}\left[B^c\right] \\ = \ & \mathbf{Pr}\left[N_1 = 1 \mid B\right]\mathbf{Pr}\left[N_2 = 1 \mid B\right]\mathbf{Pr}\left[N_3 = 0 \mid B\right]\mathbf{Pr}\left[B\right] \\ & + \mathbf{Pr}\left[N_1 = 1 \mid B^c\right]\mathbf{Pr}\left[N_2 = 1 \mid B^c\right]\mathbf{Pr}\left[N_3 = 0 \mid B^c\right]\mathbf{Pr}\left[B^c\right] \\ = \ & 0{,}2 * 0{,}2 * 0{,}8 * 0{,}5 + 0{,}8 * 0{,}8 * 0{,}2 * 0{,}5 = 0{,}08. \end{aligned}$$

Ainsi en moyenne 8 % des assurés ont eu ce parcours sur les trois années.

[1] Cet exemple est repris de DENUIT [2003].

La tarification bayesienne va se reposer sur l'historique des sinistres pour réviser le montant de la prime pure. En effet, si l'assureur ne peut observer directement si tel assuré est un bon risque ou un mauvais risque, le passé sinistre va néanmoins lui fournir une information sur la qualité du risque. En effet,

$$\begin{aligned}\mathbf{Pr}[B \mid N_1 = 1, N_2 = 1, N_3 = 0] &= \frac{\mathbf{Pr}[B, N_1 = 1, N_2 = 1, N_3 = 0]}{\mathbf{Pr}[N_1 = 1, N_2 = 1, N_3 = 0]} \\ &= \frac{\mathbf{Pr}[N_1 = 1, N_2 = 1, N_3 = 0 \mid B]\mathbf{Pr}[B]}{\mathbf{Pr}[N_1 = 1, N_2 = 1, N_3 = 0]} \\ &= 0{,}2.\end{aligned}$$

Il y a donc 20 % de chances qu'un assuré ayant ce passé sinistre soit un bon risque. Rappelons que cette probabilité est de 50 % *a priori*.

En notant $p(B) = \mathbf{E}[N_k = 1 \mid B]$ la prime pure pour un bon risque et $p(B^c) = \mathbf{E}[N_k = 1 \mid B^c]$ la prime pure pour un mauvais risque, la tarification bayesienne consiste à exiger de notre assuré, pour la période $[3;4]$, le montant

$$\begin{aligned}p(1;1;0) &= \mathbf{Pr}[B \mid N_1 = 1, N_2 = 1, N_3 = 0]\,p(B) + \mathbf{Pr}[B^c \mid N_1 = 1, N_2 = 1, N_3 = 0]\,p(B^c) \\ &= 0{,}2 * 0{,}2 + 0{,}8 * 0{,}8 = 0{,}68.\end{aligned}$$

Cet assuré a donc vu sa prime évoluer de 0,5 pour la première année à 0,68 pour la quatrième.

## 2. Approche bayesienne

Considérons le contrat *j* auquel correspondent les montants annuels de sinistre $Y_{j1}, \ldots, Y_{jT}$ et le paramètre de risque $\Theta_j$. Ce paramètre de risque n'est pas observable mais pourra être partiellement recomposé à partir des sinistres observés.

Donnons-nous des lois entièrement spécifiées pour le paramètre de risque et pour les variables de sinistres conditionnellement au paramètre de risque :

- la loi marginale de $\Theta_j$ admet une densité $\mu(\theta_j)$,
- les variables $Y_{j1}, \ldots, Y_{jT}$ sont, conditionnellement à $\Theta_j$, indépendantes entre elles et distribuées identiquement avec pour densité $f(y_{jt} | \theta_j)$.

En remarquant que pour tous événements *A* et *B*, on a $\mathbf{Pr}[A \cap B] = \mathbf{Pr}[A \mid B]\mathbf{Pr}[B]$, la densité des variables $Y_{j1}, \ldots, Y_{jT}, \Theta_j$ est donnée par :

$$\mu(\theta_j)\prod_{t=1}^{T} f(y_{jt} \mid \theta_j).$$

## 2.1. Lois *a posteriori*

L'utilisation de la règle de Bayes va nous permettre, à partir de l'information disponible en date *T*, d'enrichir notre connaissance de $\Theta_j$ et de $Y_{j,T+1}$.

**Rappel :** *(Règle de Bayes). Quels que soient les événements A et B, on a :*

$$\mathbf{Pr}[B \mid A] = \frac{\mathbf{Pr}[A \mid B]\mathbf{Pr}[B]}{\mathbf{Pr}[A]}.$$

**Proposition 1 :** *Si le couple $(X, Y)$ à valeurs dans $\mathbf{R}^2$ admet une densité $h$, les densités conditionnelles existent et sont données par :*

$$f(x \mid y) = \frac{h(x,y)}{g(y)} \text{ et } g(y \mid x) = \frac{h(x,y)}{f(x)}.$$

*D'après la règle de Bayes, elles peuvent se réécrire :*

$$f(x \mid y) = \frac{g(y \mid x)f(x)}{\int g(y \mid x)f(x)dx} \text{ et } g(y \mid x) = \frac{f(x \mid y)g(y)}{\int f(x \mid y)g(y)dy}.$$

**Proposition 2 :** *La densité de la loi* a posteriori *du paramètre de risque $\Theta_j$ est donnée par :*

$$\mu(\theta_j \mid \sigma\{Y_{j1}, \ldots, Y_{jT}\}) = \frac{\mu(\theta_j)\prod_{t=1}^{T} f(y_{jt} \mid \theta_j)}{\int \mu(\theta)\prod_{t=1}^{T} f(y_{jt} \mid \theta)d\theta}.$$

On parle de densité du paramètre de risque *a posteriori* puisqu'il s'agit de la densité conditionnée par l'information fournie par les sinistres passés $\sigma\{Y_{j1}, \ldots, Y_{jT}\}$. Pareillement, cet historique de sinistres va nous permettre de préciser la distribution du risque futur $Y_{j,T+1}$.

**Proposition 3 :** *La densité prédictive sur le risque futur $Y_{j,T+1}$ est donnée par :*

$$f(y_{j,T+1} \mid \sigma\{Y_{j1}, \ldots, Y_{jT}\}) = \frac{\int \mu(\theta)\prod_{t=1}^{T+1} f(y_{jt} \mid \theta)d\theta}{\int \mu(\theta)\prod_{t=1}^{T} f(y_{jt} \mid \theta)d\theta}.$$

La prime de crédibilité (ou de Bayes) pour la période $[T, T+1]$ est donc donnée par

$$p(y_{j1}, \ldots, y_{jT}) = \mathbf{E}[Y_{j,T+1} \mid Y_{j1} = y_{j1}, \ldots, Y_{jT} = y_{jT}] = \int y_{j,T+1} f(y_{j,T+1} \mid \sigma\{Y_{j1}, \ldots, Y_{jT}\})dy_{j,T+1}.$$

Notons que la prime de Bayes est solution de

$$\min_p \mathbf{E}\left[(Y_{j,T+1} - p)^2 \mid Y_{j1} = y_{j1}, \ldots, Y_{jT} = y_{jT}\right].$$

Cette prime est donc optimale au sens des moindres carrés.

Le principal inconvénient de ces formules réside dans la grande dépendance de ces lois au choix de la loi *a priori* $\mu(\theta_j)$. Toutefois, cette forte dépendance aux lois a priori peut présenter un avantage lorsque l'on travaille avec peu de données et donc lorsque l'approche fondée sur l'historique des sinistres n'est pas stable et que l'on souhaite donc donner un poids plus important au tarif *a priori*.

## 2.2. Mises à jour des primes

L'approche bayesienne devient très intéressante lorsqu'il s'agit de mettre à jour les primes puisqu'elle nous fournit des formules de mise à jour simples.

**Proposition 4 :** *La loi jointe de* $\Theta_j, Y_{j,T+1}$ *sachant* $\sigma\{Y_{j1},\ldots,Y_{jT}\}$ *est donnée par :*

$$f\left(y_{j,T+1} \mid \theta_j\right)\mu\left(\theta_j \mid \sigma\{Y_{j1},\ldots,Y_{jT}\}\right).$$

La proposition suivante exprime la densité de la loi du paramètre de risque conditionnellement à la connaissance d'une charge annuelle de sinistres supplémentaire.

**Proposition 5 :** *La nouvelle densité de la loi* a posteriori *du paramètre de risque* $\Theta_j$ *est donnée par :*

$$\mu\left(\theta_j \mid \sigma\{Y_{j1},\ldots,Y_{j,T+1}\}\right) = \frac{f\left(y_{j,T+1} \mid \theta_j\right)\mu\left(\theta_j \mid \sigma\{Y_{j1},\ldots,Y_{jT}\}\right)}{\int f\left(y_{j,T+1} \mid \theta\right)\mu\left(\theta \mid \sigma\{Y_{j1},\ldots,Y_{jT}\}\right)d\theta}.$$

Cette mise à jour permet d'exprimer simplement la mise à jour de la densité prédictive du risque futur.

**Proposition 6 :** *La nouvelle densité prédictive sur le risque futur* $Y_{j,T+2}$ *est donnée par :*

$$f\left(y_{j,T+2} \mid \sigma\{Y_{j1},\ldots,Y_{j,T+1}\}\right) = \int f\left(y_{j,T+2} \mid \theta\right)\mu\left(\theta \mid \sigma\{Y_{j1},\ldots,Y_{j,T+1}\}\right)d\theta .$$

Remarquons qu'il existe des choix coordonnés de la loi *f* et de la loi a priori $\mu$ qui permettent des mises à jour explicites des lois *a posteriori* successives (*cf.* DROESBEKE et al. [2002] et GOURIEROUX [1999]).

# 3. Modèle de Bühlmann

Comme les variables $Y_{j1},\dots,Y_{jT}$ sont, conditionnellement à $\Theta_j$, indépendantes entre elles et distribuées identiquement, la proposition suivante nous donne la meilleure approximation de $\mathbf{E}\left[Y_{j,T+1} \mid \theta_j\right]$ par une fonction affine.

**Proposition 7 :** *La meilleure approximation, au sens des moindres carrés, de* $\mathbf{E}\left[Y_{j,T+1} \mid \theta_j\right]$ *par une fonction affine des observations* $Y_{j1},\dots,Y_{jT}$ *est donnée par :*

$$p\left(Y_{j1},\dots,Y_{jT}\right)=\left(1-\frac{\mathbf{Cov}\left(Y_{j1},Y_{j2}\right)}{\mathbf{Var}\left(\overline{Y}_{jT}\right)}\right)\mathbf{E}\left[Y_{j1}\right]+\frac{\mathbf{Cov}\left(Y_{j1},Y_{j2}\right)}{\mathbf{Var}\left(\overline{Y}_{jT}\right)}\overline{Y}_j,$$

*où* $\overline{Y}_j=\frac{1}{T}\sum_{t=1}^{T}Y_{jt}$. *De plus, l'erreur d'approximation commise est donnée par :*

$$e_T^2=\mathbf{Var}\left[\mathbf{E}\left[Y_{j,T+1}\mid\theta_j\right]-p\left(Y_{j1},\dots,Y_{jT}\right)\right]=\left(1-\frac{\mathbf{Cov}\left(Y_{j1},Y_{j2}\right)}{\mathbf{Var}\left(\overline{Y}_{jT}\right)}\right)\mathbf{Cov}\left(\mathbf{E}\left[Y_{j,T+1}\mid\theta_j\right],Y_{j,T+1}\right).$$

*Démonstration : Cf.* PARTRAT et al. |2004]. ❑

## 3.1. Prime de crédibilité

Le modèle de crédibilité de Bühlmann (*cf.* BÜHLMANN [1967] et [1969]) utilise cette approximation linéaire puisqu'il propose de retenir pour la période $[T,T+1]$ une prime de la forme

$$p\left(Y_{j1},\dots,Y_{jT}\right)=c_0+c_1Y_{j1}+\dots+c_TY_{jT},$$

où les $c_t$ sont choisis de manière à minimiser l'écart quadratique moyen

$$\mathbf{E}\left[\left\{\mathbf{E}\left[Y_{j,T+1}\mid\Theta_j\right]-c_0-c_1Y_{j1}-\dots-c_TY_{jT}\right\}^2\right].$$

Ce modèle repose sur les hypothèses suivantes :

(H1) Pour $j=1,\dots,n$, les vecteurs $\left(\Theta_j,Y_{j1},\dots,Y_{jT}\right)$ sont indépendants.

(H2) Les variables $\Theta_1,\dots,\Theta_n$ sont indépendantes et identiquement distribuées.

(H3) Conditionnellement à $\Theta_j=\theta$, les variables aléatoires $Y_{j1},\dots,Y_{jT}$ sont indépendantes et identiquement distribuées de fonction de répartition $G_\theta$.

(H4) Les deux premiers moments $m(\theta)=\mathbf{E}\left[Y_{jt}\middle|\Theta_j=\theta\right]=\int_0^\infty x\,dG_\theta(x)$ et $\sigma^2(\theta)=\mathbf{Var}\left[Y_{jt}\middle|\Theta_j=\theta\right]=\int_0^\infty\{x-m(\theta)\}^2\,dG_\theta(x)$ existent et sont finis.

Dans la suite, on posera $m=\mathbf{E}\left[Y_{jt}\right]$ le montant de la prime *a priori*.

Sous ces hypothèses, la covariance entre les montants annuels de sinistres vaut

$$\begin{aligned}\mathbf{Cov}\left[Y_{jt}, Y_{js}\right] &= \mathbf{E}\left[\mathbf{Cov}\left[Y_{jt}, Y_{js} \mid \Theta_j\right]\right] + \mathbf{Cov}\left[\mathbf{E}\left[Y_{jt} \mid \Theta_j\right], \mathbf{E}\left[Y_{js} \mid \Theta_j\right]\right] \\ &= \delta_{s,t}\mathbf{E}\left[\sigma^2(\Theta_j)\right] + \mathbf{Cov}\left[m(\Theta_j), m(\Theta_j)\right] \\ &= \delta_{s,t}\Sigma^2 + M^2.\end{aligned}$$

Par ailleurs,

$$\begin{aligned}\mathbf{Cov}\left[Y_{jt}, m(\Theta_j)\right] &= \mathbf{E}\left[\mathbf{Cov}\left[Y_{jt}, m(\Theta_j) \mid \Theta_j\right]\right] + \mathbf{Cov}\left[\mathbf{E}\left[Y_{jt} \mid \Theta_j\right], \mathbf{E}\left[m(\Theta_j) \mid \Theta_j\right]\right] \\ &= Var\left[m(\Theta_j)\right] = M^2.\end{aligned}$$

Soit $\Psi : \mathbf{R}^{\mathbf{T}} \to \mathbf{R}_+$, la fonction qui à $c$ associe $\Psi(c) = \mathbf{E}\left[\left\{m(\Theta_j) - c_0 - c_1 Y_{j1} - \ldots - c_T Y_{jT}\right\}^2\right]$. Déterminons les $c_t$ qui minimisent $\Psi$. On a :

$$\frac{\partial}{\partial c_0}\Psi(c) = 0 \Leftrightarrow -2\,\mathbf{E}\left[m(\Theta_j) - c_0 - \sum_{t=1}^{T} c_t Y_{jt}\right] = 0$$

et

$$\frac{\partial}{\partial c_s}\Psi(c) = 0 \Leftrightarrow -2\,\mathbf{E}\left[Y_{js}\left(m(\Theta_j) - c_0 - \sum_{t=1}^{T} c_t Y_{jt}\right)\right] = 0 \text{ pour } s = 1, \ldots, T.$$

On dispose ainsi d'un système de $T+1$ équations à $T+1$ inconnues

$$\begin{cases} c_0 = m\left(1 - \sum_{t=1}^{T} c_t\right) \\ \mathbf{E}\left[Y_{js}\, m(\Theta_j)\right] - c_0\mathbf{E}\left[Y_{js}\right] - \sum_{t=1}^{T} c_t \mathbf{E}\left[Y_{js} Y_{jt}\right] = 0 \text{ pour } s = 1, \ldots, T, \end{cases}$$

$$\Leftrightarrow \begin{cases} c_0 = m\left(1 - \sum_{t=1}^{T} c_t\right) \\ \mathbf{Cov}\left[Y_{js}, m(\Theta_j)\right] + \mathbf{E}\left[Y_{js}\right]\mathbf{E}\left[m(\Theta_j)\right] - c_0\mathbf{E}\left[Y_{js}\right] - \sum_{t=1}^{T} c_t\left(\mathbf{Cov}\left[Y_{js}, Y_{jt}\right] + \mathbf{E}\left[Y_{js}\right]\mathbf{E}\left[Y_{jt}\right]\right) = 0, \end{cases}$$

$$\Leftrightarrow \begin{cases} c_0 = m\left(1 - \sum_{t=1}^{T} c_t\right) \\ \mathbf{Cov}\left[Y_{js}, m(\Theta_j)\right] + m^2 - c_0 m - \sum_{t=1}^{T} c_t\left(\mathbf{Cov}\left[Y_{js}, Y_{jt}\right] + m^2\right) = 0, \end{cases}$$

$$\Leftrightarrow \begin{cases} c_0 = m\left(1 - \sum_{t=1}^{T} c_t\right) \\ \mathbf{Cov}\left[Y_{js}, m(\Theta_j)\right] - \sum_{t=1}^{T} c_t \mathbf{Cov}\left[Y_{js}, Y_{jt}\right] = 0, \end{cases}$$

$$\Leftrightarrow \begin{cases} c_0 = \dfrac{\Sigma^2}{\Sigma^2 + TM^2} m \\ c_t = \dfrac{M^2}{\Sigma^2 + TM^2} \text{ pour } t = 1, \ldots, T. \end{cases}$$

On remarque que le terme $c_t$ ne dépend pas de $t$, ce qui signifie que l'ancienneté du sinistre n'est pas prise en compte : dans le modèle de Bühlmann, un sinistre récent aura le même poids qu'un sinistre plus ancien pour la détermination de la prime de crédibilité. Pour corriger cette hypothèse peu réaliste, une extension du modèle de Bühlmann, le modèle de Bühlmann-Straub permet d'intégrer des pondérations aux contrats. Celle-ci prend en compte le fait que les classes de contrats sont souvent d'importance inégales, notamment par le nombre de contrats qu'elles regroupent.

**Proposition 8 :** *Dans le modèle de Bühlmann, la prime pour l'année $T+1$ est donnée par*

$$p(Y_{j1}, \ldots, Y_{jT}) = \frac{\Sigma^2}{\Sigma^2 + TM^2} m + \frac{TM^2}{\Sigma^2 + TM^2} \overline{Y}_j .$$

Il s'agit bien d'une prime de crédibilité car cette prime *a posteriori* est une moyenne pondérée de la prime *a priori* $m$ et de la prime observée $\overline{Y}_{jT}$.

Le facteur de crédibilité $\alpha = \dfrac{TM^2}{\Sigma^2 + TM^2}$ croit avec l'hétérogénéité du portefeuille $M^2$ alors qu'il diminue avec $\Sigma^2$, *i. e.* avec les variations dues au hasard.

## 3.2. Estimation des paramètres

Le fait que l'assureur utilise une tarification *a posteriori* ne modifie pas la sinistralité mais seulement la répartition de la charge entre les différents assurés. Aussi dans le cas du modèle de Bühlmann, on a :

$$\mathbf{E}\left[p(Y_{j1}, \ldots, Y_{jT})\right] = \frac{\Sigma^2}{\Sigma^2 + M^2T} m + \frac{M^2T}{\Sigma^2 + M^2T} \mathbf{E}\left[\overline{Y}_j\right] = m .$$

Pour pourvoir utiliser la modèle, l'assureur va devoir estimer les paramètres $m, M^2$ et $\Sigma^2$ sur les sinistres observés. On dispose des estimateurs sans biais suivants :

$$\hat{m} = \frac{1}{n} \sum_{j=1}^{n} \frac{1}{T} \sum_{t=1}^{T} Y_{jt} = \overline{\overline{Y}} ,$$

$$\hat{\Sigma}^2 = \frac{1}{n(T-1)} \sum_{j=1}^{n} \sum_{t=1}^{T} \left(Y_{jt} - \overline{Y}_j\right)^2 ,$$

et

$$\hat{M}^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(\overline{Y}_j - \overline{\overline{Y}}\right)^2 - \frac{\hat{\Sigma}^2}{T} .$$

Remarquons que l'estimateur du facteur de crédibilité $\dfrac{\hat{M}^2 T}{\hat{\Sigma}^2 + \hat{M}^2 T}$ est, en général, biaisé car ce n'est pas une fonction linéaire des estimateurs sans biais $\hat{\Sigma}^2$ et de $\hat{M}^2$.

# 4. Modèle de Bühlamm-Straub

Le modèle de Bühlmann-Straub est une extension du modèle de Bühlmann qui intègre des pondérations sur les observations.

Ce modèle repose sur les hypothèses suivantes :

(H1) Pour $j = 1, \ldots, n$, les vecteurs $\left(\Theta_j, Y_{j1}, \ldots, Y_{jT}\right)$ sont indépendants.
(H2) Les variables $\Theta_1, \ldots, \Theta_n$ sont indépendantes et identiquement distribuées.
(H3) Conditionnellement à $\Theta_j = \theta$, les variables aléatoires $Y_{j1}, \ldots, Y_{jT}$ sont indépendantes et identiquement distribuées de fonction de répartition $G_\theta$.
(H4) Les variables aléatoires $Y_{j1}, \ldots, Y_{jT}$ sont de carré intégrable et $\mathbf{E}\left[Y_{jt} \middle| \Theta_j = \theta\right] = m(\theta)$ et $\mathbf{Var}\left[Y_{jt} \middle| \Theta_j = \theta\right] = \dfrac{1}{\omega_{jt}} \sigma^2(\theta)$ où $\omega_{jt}$ traduit l'importance de la police *j* dans le portefeuille, au cours de la *t*-ème année.

## 4.1. Prime de crédibilité

**Proposition 9 :** *Dans le modèle de Bühlmann-Straub, la prime pour l'année* $T+1$ *est donnée par*

$$p\left(Y_{j1}, \ldots, Y_{jT}\right) = \frac{\Sigma^2}{\Sigma^2 + M^2 \omega_{j\bullet}} m + \frac{M^2 \omega_{j\bullet}}{\Sigma^2 + M^2 \omega_{j\bullet}} \frac{1}{\omega_{j\bullet}} \sum_{t=1}^{T} \omega_{jt} Y_{jt},$$

*où* $\omega_{j\bullet} = \sum_{t=1}^{T} \omega_{jt}$.

*Démonstration : Cf.* PARTRAT et al. [2004]. ❑

On remarque que dans le cas particulier où $\omega_{jt} = 1$ pour tout *j* et tout *t*, on retrouve la prime de crédibilité du modèle de Bühlmann.

Le facteur de crédibilité pour la police *j* est

$$\alpha_j = \frac{M^2 \omega_{j\bullet}}{\Sigma^2 + M^2 \omega_{j\bullet}}.$$

Ce facteur tend vers 1 si $\omega_{j\bullet} \to +\infty$. Par ailleurs, comme dans le modèle de Bühlmann, il croît avec l'hétérogénéité du portefeuille $M^2$ alors qu'il diminue avec $\Sigma^2$.

### 4.2. Estimation des paramètres

On dispose des estimateurs sans biais suivants :

$$\hat{m} = \frac{1}{\omega_{\bullet\bullet}} \sum_{j=1}^{n} \sum_{t=1}^{T} \omega_{jt} Y_{jt} ,$$

$$\hat{\Sigma}^2 = \frac{1}{n(T-1)} \sum_{j=1}^{n} \sum_{t=1}^{T} \omega_{jt} \left(Y_{jt} - \overline{Y}_j\right)^2 ,$$

et

$$\hat{M}^2 = \frac{\omega_{\bullet\bullet}}{\omega_{\bullet\bullet}^2 - \sum_{j=1}^{n} \omega_{j\bullet}^2} \left\{ \sum_{j=1}^{n} \omega_{j\bullet} \left(\overline{Y}_j - \overline{\overline{Y}}\right)^2 - (n-1)\hat{\Sigma}^2 \right\},$$

où $\omega_{\bullet\bullet} = \sum_{j=1}^{n} \omega_{j\bullet} = \sum_{j=1}^{n} \sum_{t=1}^{T} \omega_{jt}$ .

### 4.3. Choix des pondérations

La prime de crédibilité dépendra fortement de la pondération affectée aux polices. Il n'y a pas de règle universelle pour déterminer ces pondérations : elles dépendront essentiellement du contexte. Par exemple, les pondérations peuvent se fonder sur le nombre de contrats (lorsque le modèle de Bühlmann est appliqué à des classes de contrat), le nombre de sinistres ou encore la charge de sinistres.

## Bibliographie


BÜHLMANN H. [1967] « Experience rating and credibility, I ». *ASTIN Bulletin*, vol. 4.

BÜHLMANN H. [1969] « Experience rating and credibility, II ». *ASTIN Bulletin*, vol. 5.

DENUIT M. [2003] *Théorie de la crédibilité*. Notes de cours, ISFA.

DROESBEKE J.J., FINE J., SAPORTA G. [2002] *Méthodes bayesiennes en statistique*. Technip.

GOURIEROUX CH. [1999] *Statistique de l'assurance*. Economica.

PARTRAT CH., BESSON J.L. [2004] *Assurance non-vie. Modélisation, simulation*. Economica.

# Exercices

**Exercice 1 :** Une compagnie d'assurance couvre deux contrats depuis trois ans. Elle dispose des montants annuels de sinistres suivants :

| **Contrat** | **Année** | | |
|---|---|---|---|
| | **1** | **2** | **3** |
| **1** | 5 | 8 | 11 |
| **2** | 11 | 13 | 12 |

Déterminez, selon le modèle de Bühlmann, les primes *a posteriori* pour la quatrième année.

**Exercice 2 :** Un portefeuille de 340 assurés a produit 240 déclarations de vol au cours d'une année. La répartition de ces sinistres est la suivante :

| **Nombre de sinistres** | **Nombre d'assurés** |
|---|---|
| 0 | 200 |
| 1 | 80 |
| 2 | 50 |
| 3 | 10 |

En supposant que le nombre de sinistres déclarés par un assuré suit une loi de Poisson dont le paramètre peut varier d'un assuré à l'autre, déterminez le facteur de crédibilité (selon le modèle de Bühlmann) pour un assuré de ce portefeuille. Déduisez en l'augmentation de la prime d'un assuré qui aurait déclaré deux sinistres.