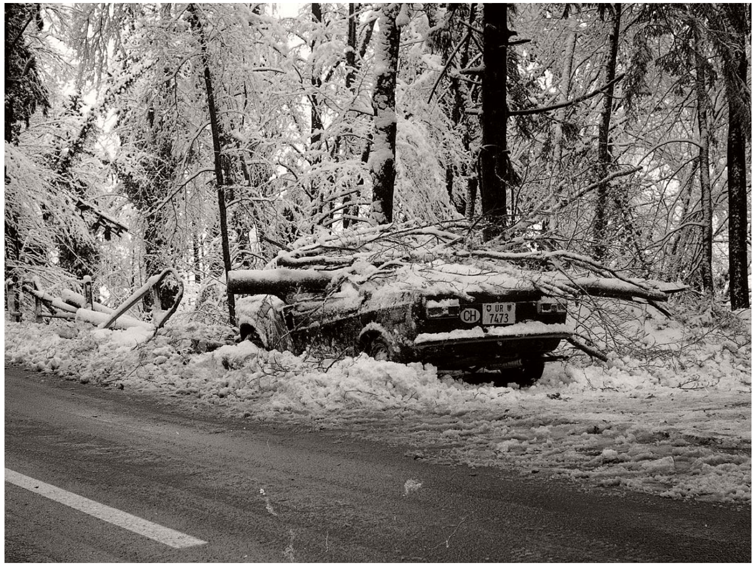
Actuariat de l'Assurance Non-Vie # 6

A. Charpentier (Université de Rennes 1)



ENSAE 2017/2018

credit: Arnold Odermatt

Tarification *a posteriori* ? (*experience rating*)

On distinguera

- modèles de crédibilité
- approches bonus-malus ou no-claim discount
- modèles de panels et approches longitudinales

Idée: introduire un aspect temporel, $N_t|\underline{\boldsymbol{N}}_{t-1}$, où $\underline{\boldsymbol{N}}_{t-1} = \{N_{t-1}, N_{t-2}, \cdots\}$

Préambule

Petit(s) rappel(s) de calculs de probabilité: pour tout vecteur (X, Y)

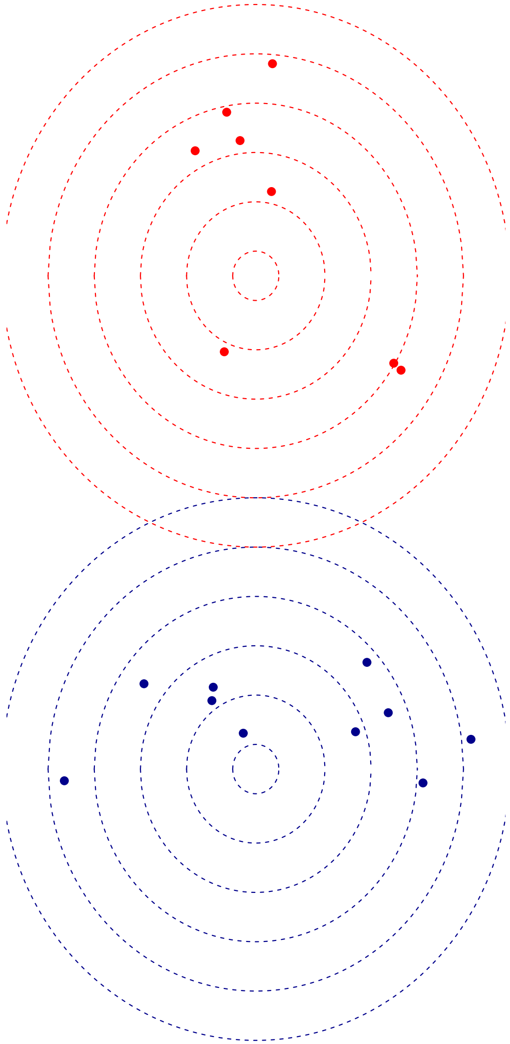
$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}(X|Y)]$$

$$\text{Var}[X] = \text{Var}[\mathbb{E}(X|Y)] + \mathbb{E}[\text{Var}(X|Y)]$$

e.g. pour une loi Poisson mélange, $N \sim \mathcal{P}(\Lambda)$,

$$\mathbb{E}[N] = \mathbb{E}[\mathbb{E}(N|\Lambda)] = \mathbb{E}[\Lambda]$$

$$\text{Var}[N] = \text{Var}[\mathbb{E}(N|\Lambda)] + \mathbb{E}[\text{Var}(N|\Lambda)] = \mathbb{E}[\Lambda] + \text{Var}[\Lambda]$$



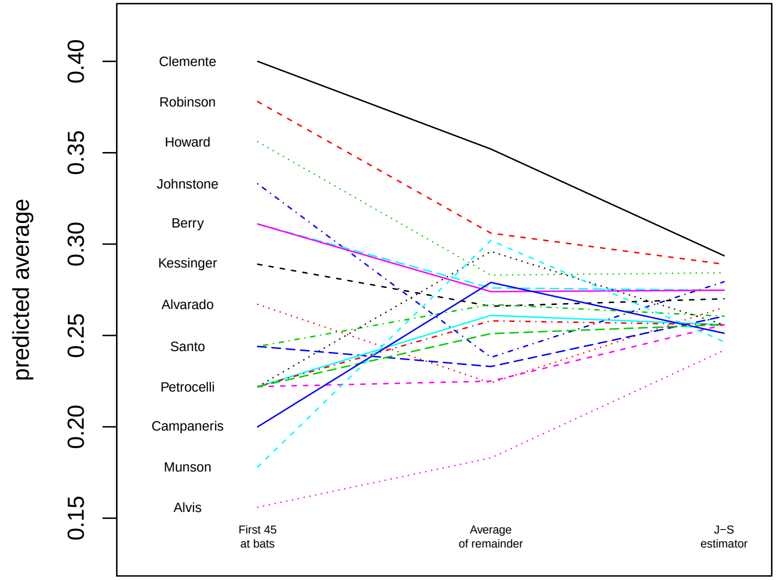
Préambule: Estimateur de Stein

Charles Stein essayait de prédire la moyenne à la batte au cours d'une saison, à partir d'observations passées (45 essais). L'estimateur proposé par Stein est

$$m + c(y_i - m)$$

où m est la moyenne de toutes les moyennes observées, c la constante de Stein, et y_i la valeur observée. La constante de Stein est donnée par

$$c = 1 - \frac{(k-3)s^2}{\sum_j (\overline{y}_j - m)^2}$$



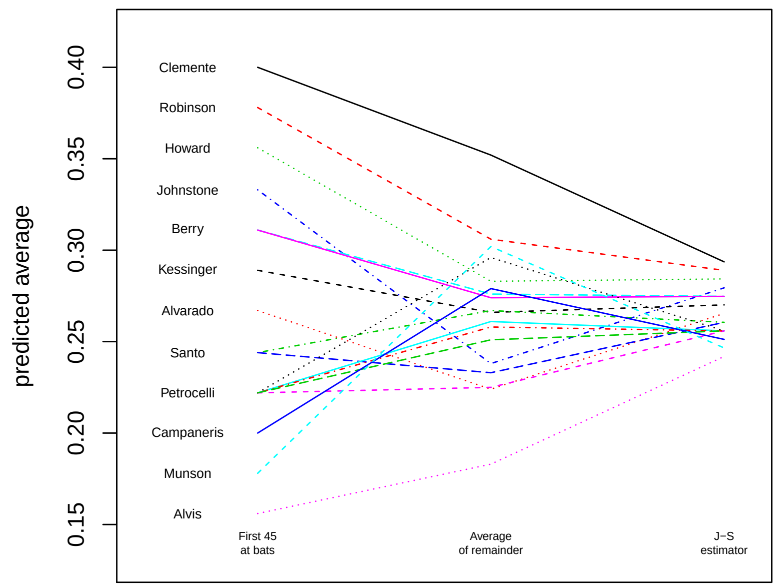
Préambule: Estimateur de Stein

$m + c(y_i - m)$ avec

$$c = 1 - \frac{(k-3)s^2}{\sum_j (\overline{y}_j - m)^2}$$

où k est le nombre de moyennes inconnues, s^2 le carré de l'écart-type pour la distribution considérée. On a une fonction correspondant qui correspond à un ratio de la variance, et de la variance interclasse.

Pour les données du baseball, $c = 0.212$ et $m = 0.265$.



Préambule: Statistique Bayésienne

Petit(s) rappel(s) de statistique bayésienne: la formule de Bayes nous dit que

$$\mathbb{P}[A|B]=\frac{\mathbb{P}[A\cap B]}{\mathbb{P}[B]}=\frac{\mathbb{P}[A]}{\mathbb{P}[B]}\cdot\mathbb{P}[B|A]\propto\mathbb{P}[A]\cdot\mathbb{P}[B|A]$$

On va utiliser la formule de Bayes sur $A=\Theta$ et $B=\{X_1,\cdots,X_n\}$

$$\underbrace{\pi(\theta|x_1,\cdots,x_n)}_{\text{loi a posteriori}}\propto\underbrace{\pi(\theta)}_{\text{loi a priori}}\cdot\underbrace{f(x_1,\cdots,x_n|\theta)}_{\text{vraisemblance}}$$

"it's time to adopt modern Bayesian data analysis as standard procedure in our scientific practice and in our educational curriculum. Three reasons:

1. Scientific disciplines from astronomy to zoology are moving to Bayesian analysis. We should be leaders of the move, not followers.

2. Modern Bayesian methods provide richer information, with greater flexibility and broader applicability than 20th century methods. Bayesian methods are intellectually coherent and intuitive.
   Bayesian analyses are readily computed with modern software and hardware.

3. Null-hypothesis significance testing (NHST), with its reliance on p values, has many problems.
   There is little reason to persist with NHST now that Bayesian methods are accessible to everyone.

My conclusion from those points is that we should do whatever we can to encourage the move to Bayesian data analysis." John Kruschke,

(quoted in Meyers & Guszcza (2013))

Inférence Bayésienne

Consider some Bernoulli sample $\boldsymbol{x} = \{x_1, x_2, \cdots, x_n\}$, where $x_i \in \{0,1\}$.

X_i's are i.i.d. $\mathcal{B}(p)$ variables, $f_X(x) = p^x[1-p]^{1-x}$, $x \in \{0,1\}$.

Standard frequentist approach

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^n x_i = \text{argmax}\Big\{\underbrace{\prod_{i=1}^n f_X(x_i)}_{\mathcal{L}(p;\boldsymbol{x})}\Big\}$$

From the central limit theorem

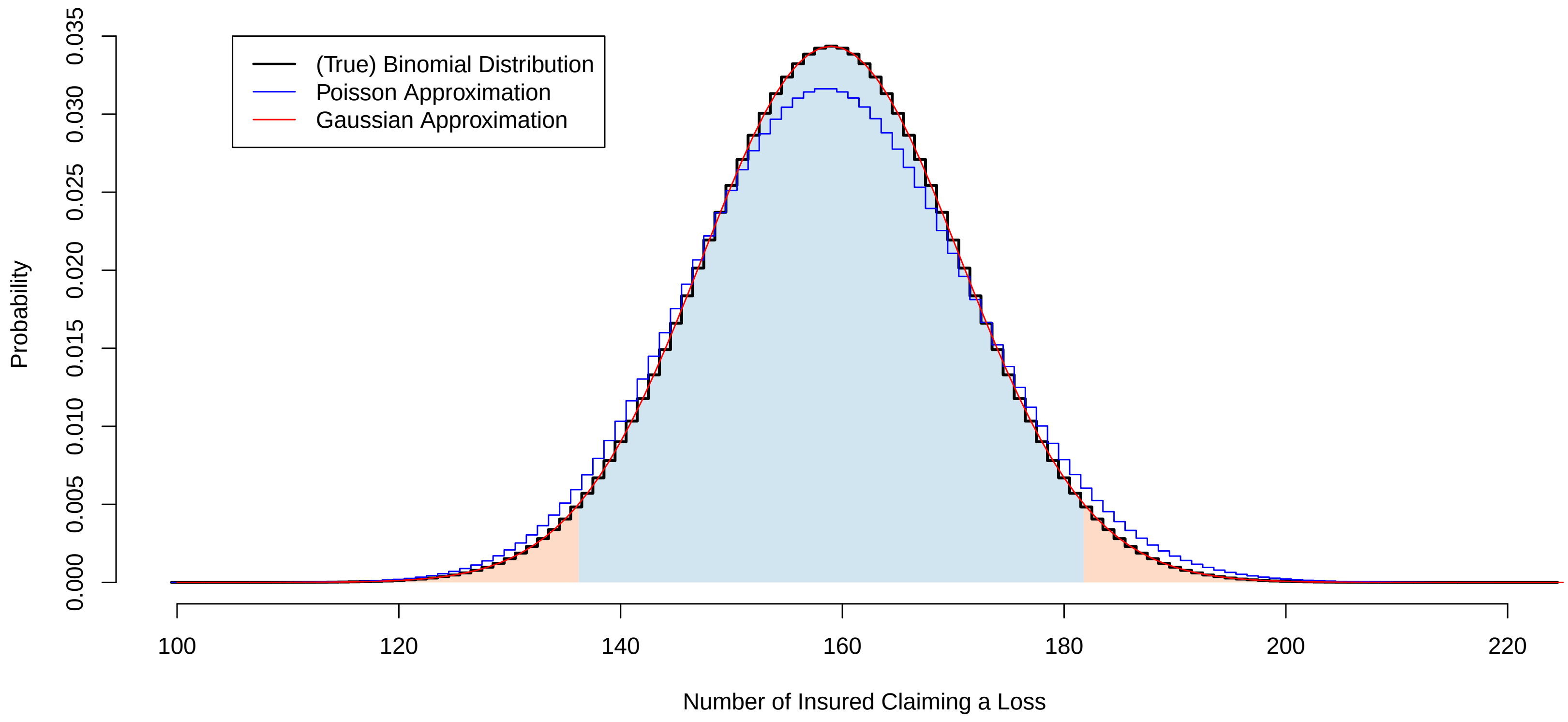
$$\sqrt{n}\frac{\widehat{p}-p}{\sqrt{p(1-p)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \text{ as } n \to \infty$$

we can derive an approximated 95% confidence interval

$$\left[\widehat{p} \pm \frac{1.96}{\sqrt{n}}\sqrt{\widehat{p}(1-\widehat{p})}\right]$$

Inférence Bayésienne

Example out of 1,047 contracts, 159 claimed a loss



Inférence Bayésienne

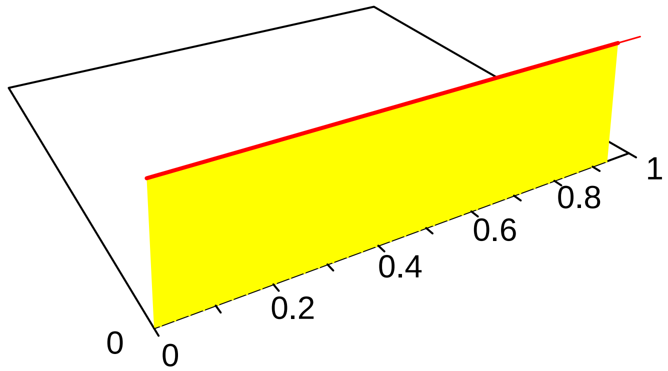
Consider sample $\boldsymbol{x} = \{0, 0, 0, 0, 0\}$.
Here the likelihood is

$$\begin{cases} (x_i|\theta) = \theta^{x_i}[1-\theta]^{1-x_i} \\ f(\boldsymbol{x}|\theta) = \theta^{\boldsymbol{x}^\mathsf{T}\mathbf{1}}[1-\theta]^{n-\boldsymbol{x}^\mathsf{T}\mathbf{1}} \end{cases}$$

and we need a priori distribution $\pi(\cdot)$ e.g. a beta distribution

$$\pi(\theta) = \frac{\theta^\alpha[1-\theta]^\beta}{B(\alpha,\beta)}$$

$$\pi(\theta|\boldsymbol{x}) = \frac{\theta^{\alpha+\boldsymbol{x}^\mathsf{T}\mathbf{1}}[1-\theta]^{\beta+n-\boldsymbol{x}^\mathsf{T}\mathbf{1}}}{B(\alpha+\boldsymbol{x}^\mathsf{T}\mathbf{1},\beta+n-\boldsymbol{x}^\mathsf{T}\mathbf{1})}$$



Lien entre Stein et l'approche Bayésienne

Considérons un échantillon $\{x_1, \cdots, x_p\}$ de loi $\mathcal{N}(\theta_i, \sigma^2)$ où la loi a priori est $\Theta_i \sim \mathcal{N}(\mu, \tau^2)$. La loi a posteriori est

$$\Theta_i | x_i \sim \mathcal{N}\left(\frac{\sigma^2}{\sigma^2+\tau^2} \cdot \mu + \frac{\tau^2}{\sigma^2+\tau^2} \cdot x_i, \frac{\sigma^2\tau^2}{\sigma^2+\tau^2}\right)$$

Notons que

$$\mathbb{E}[\overline{X}] = \mu \text{ et } \mathbb{E}\left[\frac{(p-3)\sigma^2}{\sum_i (X_i - \overline{X})^2}\right] = \frac{\sigma^2}{\sigma^2+\tau^2}$$

de telle sorte que l'espérance a posteriori peut s'approcher par

$$\left[\frac{(p-3)\sigma^2}{\sum_i (x_i - \overline{x})^2}\right] \cdot \overline{x} + \left[1 - \frac{(p-3)\sigma^2}{\sum_i (x_i - \overline{x})^2}\right] \cdot x_i.$$

Approche Bayésienne

The number of claims filed by a policyholder during one given period of time has a Poisson distribution, $\mathcal{P}(\theta)$. Assume that, in the population of potential policyholders, the parameter values are distributed as

$$\mathbb{P}(\theta = 1) = 70\%, \mathbb{P}(\theta = 2) = 20\%, \mathbb{P}(\theta = 3) = 10\%.$$

Assume moreover that the cost of the claims is $1,000$ euros (arbitrary). What is the experience rated premium for an insured in the second year if two losses occured in the first year ?

$$\mathbb{E}(N_2|N_1 = 2) = \mathbb{E}(\mathbb{E}(N_2|\Theta, N_1 = 2)) = \mathbb{E}(\Theta|N_1 = 2),$$

since the expectation of a Poisson variable is the parameter.

Note that the pure premium here is

$$\mathbb{E}(N_1) = \mathbb{E}(\mathbb{E}(N_1|\Theta)) = \mathbb{E}(\Theta) = \sum \theta \cdot \mathbb{P}(\Theta = \theta) = 1.4.$$

Approche Bayésienne

One gets

$$\begin{aligned}\mathbb{E}\left(N_2|N_1=2\right) &= \mathbb{E}\left(\Theta|N_1=2\right)=\sum\theta\mathbb{P}\left(\Theta=\theta|N_1=2\right)\\ &= \sum\theta\frac{\mathbb{P}\left(\Theta=\theta\right)}{\mathbb{P}\left(N_1=2\right)}\cdot\mathbb{P}\left(N_1=2|\Theta=\theta\right).\end{aligned}$$

Since

$$\begin{aligned}\mathbb{P}\left(N_1=2\right) &= \sum\mathbb{P}\left(N_1=2|\Theta=\theta\right)\mathbb{P}\left(\Theta=\theta\right)=\sum\frac{\theta^2}{2}\exp\left(-\theta\right)\mathbb{P}\left(\Theta=\theta\right)\\ &= \frac{1}{2}\exp\left(-1\right)\times 0.7+\frac{2^2}{2}\exp\left(-2\right)\times 0.2+\frac{3^2}{2}\exp\left(-3\right)\times 0.1=0.205,\end{aligned}$$

one gets

$$\mathbb{E}\left(N_2|N_1=2\right)=\sum\theta\frac{\mathbb{P}\left(\Theta=\theta\right)}{0.205}\frac{\theta^2}{2}\exp\left(-\theta\right)=1,482.$$

Approche Bayésienne

Assume now that the second year, 1 loss occured

$$\begin{aligned}\mathbb{E}\left(N_3|N_1=2,N_2=1\right) &= \mathbb{E}\left(\mathbb{E}\left(N_3|\Theta,N_1=2,N_2=1\right)\right)\\ &= \mathbb{E}\left(\Theta|N_1=2,N_2=1\right).\end{aligned}$$

The posterior distribution of Θ, given $\{N_1=2,N_2=1\}$ is then

$$\begin{aligned}\mathbb{P}\left(\Theta=\theta|N_1=2,N_2=1\right) &= \frac{\mathbb{P}\left(\Theta=\theta\right)}{\mathbb{P}\left(N_1=2,N_2=1\right)}\mathbb{P}\left(N_1=2,N_2=1|\Theta=\theta\right)\\ &\propto \mathbb{P}\left(\Theta=\theta\right)\mathbb{P}\left(N_1=2,N_2=1|\Theta=\theta\right)\end{aligned}$$

since the losses are assumed to be conditionnally independent, given $\Theta=\theta$, i.e.

$$\begin{aligned}\mathbb{P}\left(\Theta=1|N_1=2,N_2=1\right) &\propto \mathbb{P}\left(\Theta=1\right)\mathbb{P}\left(N_1=2,N_2=1|\Theta=1\right)\\ &= 0.7\times\exp\left(-1\right)\frac{1}{2}\exp\left(-1\right)\frac{1}{1}=0.0474\end{aligned}$$

Approche Bayésienne

and similarly

$$\mathbb{P}\left(\Theta = 2|N_1 = 2, N_2 = 1\right) \propto 0.0147$$

$$\mathbb{P}\left(\Theta = 3|N_1 = 2, N_2 = 1\right) \propto 0.0033$$

so that, the posterior distribution of Θ is proportitional to these numbers

$$\mathbb{P}\left(\Theta = 1|N_1 = 2, N_2 = 1\right) = \frac{0.0474}{0.0654} = 72.48\%$$

$$\mathbb{P}\left(\Theta = 1|N_1 = 2, N_2 = 1\right) = \frac{0.0147}{0.0654} = 22.48\%$$

$$\mathbb{P}\left(\Theta = 1|N_1 = 2, N_2 = 1\right) = \frac{0.0033}{0.0654} = 0.504\%$$

The expected value is then

$$\mathbb{E}\left(N_3|N_1 = 2, N_2 = 1\right) = \sum \theta \mathbb{P}\left(\Theta = \theta|N_1 = 2, N_1 = 1\right) = 1.3256$$

(which should be compared with the pure premium of 1.4).

Approche Bayésienne, cas Poisson-Gamma

Consider one car driver, and assume that his Poisson parameter, θ, can be considered to have been drawn from a Gamma distribution, with some parameters α and β, i.e.

$$f\pi(\theta) = \exp(-\theta\alpha)\,\theta^{\beta-1}\frac{\alpha^{\beta}}{\Gamma(\beta)},\ \theta > 0.$$

Recall that $\mathbb{E}(\Theta) = \beta/\alpha$.

Assume that the polcyholder had k claims in 4 years. What should be his experience rated premium for the fifth year ?

From the conditional independence, notice, first of all, that the number of claims in 4 years has Poisson distribution with parameter 4θ, i.e.

$$\mathbb{P}(N_1 + ... + N_4 = k|\Theta = \theta) = \exp(-4\theta)\frac{(4\theta)^k}{k!},\ k \in \mathbb{N},$$

(a sum of independent Poisson variates is Poisson).

Approche Bayésienne, cas Poisson-Gamma

Therefore,

$$\begin{aligned}\pi(\theta|N_1+...+N_4=k) &\propto \mathbb{P}(N_1+...+N_4=k|\Theta=\theta)\,f(\theta)\\ &\propto \exp(-\theta(4+\alpha))\cdot\theta^{k+\beta-1}\end{aligned}$$

which is a Gamma distribution with parameters $\alpha^* = (\alpha+4)$ and $\beta^* = k+\beta = \beta+4\overline{N}$ where $\overline{N}$ denotes the average number of claims per year (here $k/4$). The expected number of claims for the fifth year is then

$$\mathbb{E}(N_5|N_1+...+N_4=k) = \frac{\beta^*}{\alpha^*} = \frac{\beta+4\overline{N}}{\alpha+4},$$

where $t=4$ is the number of years. More generally, one gets

$$\begin{aligned}\mathbb{E}(N_{t+1}|N_1=n_1,...,N_t=n_t) &= \frac{\beta^*}{\alpha^*} = \frac{\beta+n\overline{n}}{\alpha+t} \text{ where } \overline{n} = \frac{1}{t}\sum_{i=1}^t n_i\\ &= \left[1-\frac{t}{\alpha+t}\right]\frac{\beta}{\alpha}+\frac{t}{\alpha+t}\overline{n}.\end{aligned}$$

De l'Approche Bayésienne ła Crédibilité

$$\mathbb{E}\left(N_{t+1}|\underline{\boldsymbol{N}}_t\right) = \underbrace{\left[1-\frac{t}{\alpha+t}\right]}_{1-z}\underbrace{\frac{\beta}{\alpha}}_{\mathbb{E}[N]}+\underbrace{\frac{t}{\alpha+t}}_{z}\underbrace{\overline{n}}_{\overline{n}}.$$

The first terme $\mathbb{E}[N]$ is the overall (unconditional) expected value, and the second term $\overline{n}$ is related to individual past experience.

The predicted number of claims is a weighted sum of those two terms, weight z being the credibility factor, increasing with t, and function of α that will be related to $\mathbb{E}[\text{Var}(N|\Theta)]$ and $\text{Var}[\mathbb{E}(N|\Theta)]$.

"*the problem of experience rating arises out of the necessity from the standpoint of equity to the individual risk, of striking a balance between class-experience on the one hand and risk experience on the other*". (Albert Whitney)

Formalisme des Méthodes de Crédibilité

Le but en crédibilité de précision consiste à calculer la "meilleure" prévision du nombre de sinistres de la prochaine année, $S_{i,t+1}$ pour chaque assuré. Si le niveau de risque du contrat i est connu, alors la meilleure prévision (au sens des moindres carrés) est

$$\mu(\theta_i) = \mathbb{E}[N_{i,t+1}|\Theta = \theta_i] = \int_0^{\infty} x \cdot dF(x|\theta_i) = \sum_{x=0}^{\infty} x \cdot \mathbb{P}[N_{i,t+1} = x|\theta_i]$$

Comme première approximation de la prime de risque, on peut utiliser la moyenne pondérée de toutes les primes de risque possibles:

$$m = \mathbb{E}[\mu(\Theta)] = \int_{-\infty}^{\infty} \mu(\theta)\, d\Pi(\theta).$$

Cette approximation sera la même pour tous les contrats; c'est la prime collective.

Formalisme des Méthodes de Crédibilité

Celle-ci est globalement adéquate, mais elle n'est pas nécessairement équitable. En termes statistiques, cela signifie qu'il existe une meilleure approximation des primes de risque lorsque des données sont disponibles.

La meilleure approximation de la prime de risque $\mu(\theta_i)$ est la fonction des observations $g^*(N_{i,1},\ldots,N_{i,t})$ minimisant l'erreur quadratique moyenne

$$g^* = \text{argmin}\left\{\mathbb{E}[\mu(\Theta) - g(N_{i,1},\ldots,N_{i,t})]^2\right\},$$

La fonction $g^*(S_{i1},\ldots,S_{in})$ est la prime bayésienne

$$B_{i,t+1} = \mathbb{E}[\mu(\Theta)|N_{i,1},\ldots,N_{i,t}] = \int_{-\infty}^{\infty}\mu(\theta)\,d\Pi(\theta|N_{i,1},\ldots,N_{i,t}),$$

où $\Pi(\theta|N_1,\ldots,N_t)$ est la distribution a posteriori des niveaux de risque.

Formalisme des Méthodes de Crédibilité

Pour certaines combinaisons de distributions $F(x|\theta)$ et $\Pi(\theta)$, la prime bayésienne est une fonction linéaire de la forme

$$B_{i,t+1} = z\overline{N}_i + (1-z)m, \text{ avec } \overline{S}_i = \frac{1}{t}\sum_{\tau=1}^{t} N_{i,\tau}$$

Une prime de cette forme est appelée prime de crédibilité et z est le facteur de crédibilité. Whitney (1918) et Bailey (1950) ont été les premiers à démontrer que certaines primes bayésiennes sont des primes de crédibilité avec un facteur de crédibilité toujours de la forme $z = n/(n+K)$, où K est une constante. Les résultats partiels de Bailey (1950) ont plus tard été complétés par Mayerson (1964) et unifiés par Jewell (1974).

Le modèle de Bühlman

Dans Bühlmann (1967), il est suggéré de restreindre l'approximation de la prime de risque aux seules fonctions linéaires des observations. La meilleure prévision s'avère alors être une prime de crédibilité

$$\pi_{i,t+1}^{B} = z\overline{S}_i + (1-z)m$$

$$\begin{aligned} \text{avec } z &= \frac{t}{t+s^2/a}, \\ m &= \mathbb{E}[\mu(\Theta)], \\ \nu = s^2 &= \mathbb{E}[\sigma^2(\Theta)] = \text{Var}[N_{i,t}|\Theta], \\ a &= \text{Var}[\mu(\Theta)]. \end{aligned}$$

Les quantités m, $\nu = s^2$ et a sont appelées les paramètres de structure puisqu'ils définissent la structure interne du portefeuille.

Le modèle de Bühlman

En analyse de variance, s^2 représente la dispersion intra (*within*) et a, la dispersion inter (*between*).

$$\begin{aligned}\text{Var}[N_{i,t}] &= \mathbb{E}[\text{Var}(N_{i,t}|\Theta_i)] + \text{Var}[\mathbb{E}(N_{i,t}|\Theta_i)] \\ &= \mathbb{E}[\sigma^2(\Theta_i)] + \text{Var}[\mu(\Theta_i)] = s^2 + a.\end{aligned}$$

En pratique, nous adoptons l'approche bayésienne empirique et nous estimons les paramètres de structure à partir des données du portefeuille.

$$\begin{aligned}\widehat{m} &= \overline{N} = \frac{1}{n}\sum_{i=1}^{n}\overline{N}_i \\ \widehat{s}^2 &= \frac{1}{n(t-1)}\sum_{i=1}^{n}\sum_{\tau=1}^{t}(N_{i,t}-\overline{N}_i)^2 \\ \widehat{a} &= \frac{1}{n-1}\sum_{i=1}^{n}(\overline{N}_i-\overline{N})^2 - \frac{1}{t}\widehat{s}^2.\end{aligned}$$

Calculs de Crédibilité

Si on dispose de données complètes

```
> install.packages("CASdatasets", repos = "http://cas.uqam.ca/pub/R/"
    , type="source")
> library(CASdatasets)
> data(Norberg)
> t(Norberg)
         year0 year1 year2 year3 year4 year5 year6 year7 year8 year9
risk1        0     0     0     0     0     0     0     0     0     0
risk2        0     0     0     0     0     0     0     0     0     0
risk3        1     0     1     0     0     0     0     0     0     0
risk4        0     0     0     0     0     0     0     0     0     0
risk5        0     0     0     0     0     1     0     0     1     0
risk6        0     0     0     0     0     0     0     0     0     0
```

...

```
risk19       0     0     0     0     0     0     0     0     0     1
risk20       0     0     0     0     0     0     0     0     0     0
```

Calculs de Crédibilité

```
> norberg=t(Norberg)
> T<- ncol(norberg)
> (m <- mean(norberg))
[1] 0.145
> (s2 <- mean(apply(norberg,1,var)))
[1] 0.1038889
> (a <- var(apply(norberg,1,mean))-s2/T)
[1] 0.02169006
> (Z <- T/(T+s2/a))
[1] 0.6761462
> Z*apply(norberg,1,mean)+(1-Z)*m
    risk1     risk2     risk3     risk4     risk5     risk6     risk7
0.0469588 0.0469588 0.1821880 0.0469588 0.1821880 0.0469588 0.1821880
    risk8     risk9    risk10    risk11    risk12    risk13    risk14
0.0469588 0.4526465 0.1145734 0.3174173 0.2498027 0.1145734 0.1145734
   risk15    risk16    risk17    risk18    risk19    risk20
0.0469588 0.0469588 0.3850319 0.1145734 0.1145734 0.0469588
```

Le modèle de Bühlmann

Ce modèle peut aussi s'écrire sous forme de mise à jour de prime,

$$\pi_{i,t+1}^{\text{B}} = \zeta_t \pi_{i,t}^{\text{B}} + [1 - \zeta_t]\overline{N}_i$$

Le modèle de Buhlmann-Straub

Le modèle de Bühlmann-Straub Bühlmann-Straub (1970) est une généralisation du modèle de Bühlmann permettant de tenir compte de l'exposition au risque des assurés. Ceci est particulièrement important dans les branches d'affaires où le volume des assurés varie beaucoup. Par exemple, pensons à l'assurance contre les accidents du travail, où les assurés sont des employeurs: une entreprise de 1000 employés a une exposition au risque beaucoup plus grande qu'une entreprise de 10 employés. Cette capacité du modèle de Bühlmann–Straub de départager les "gros" assurés des "petits" en fait le modèle de crédibilité le plus utilisé en pratique.

Le modèle de Buhlman-Straub

Ainsi, on attribue maintenant un poids (ou volume) w_{it} à chaque observation. L'hypothèse selon laquelle les obervations sont (conditionnellement) indépendantes et identiquement distribuées est remplacée par l'hypothèse suivante: pour tout $s, \tau = 1, \ldots, t$,

$$\begin{aligned}\mathbb{E}[N_{it}|\Theta_i] &= \mu(\Theta_i)\\ \text{Cov}(N_{i,s}, N_{i,t}|\Theta_i) &= \begin{cases}\dfrac{\sigma^2(\Theta_i)}{w_{it}}, & s = t\\ 0, & s \neq t.\end{cases}\end{aligned}$$

Or, pour que la relation soit vérifiée, les observations N_{it} doivent être homogènes, et correspondent souvent à des ratios. En général, les observations dans le modèle de Bühlmann–Straub sont des montants de sinistres divisés par le volume, ou le nombre de sinistres dividés par l'exposition totale, c'est-à-dire

$$N_{i,t} \leftarrow \frac{N_{i,t}}{w_{i,t}}.$$

Le modèle de Buhlman-Straub

Posons

$$
\begin{aligned}
N_{iw} &= \sum_{\tau=1}^{t} \frac{w_{i\tau}}{w_{i\cdot}} N_{i\tau}, \text{ où } w_{i\cdot} = \sum_{\tau=1}^{t} w_{i\tau}, \\
N_{ww} &= \sum_{i=1}^{n} \frac{w_{i\cdot}}{w_{\cdot\cdot}} N_{iw}, \text{ où } w_{\cdot\cdot} = \sum_{i=1}^{n} w_{i\cdot}, \\
N_{zw} &= \sum_{i=1}^{n} \frac{z_i}{z_{\cdot}} N_{iw}, \text{ où } z_{\cdot} = \sum_{i=1}^{n} z_i.
\end{aligned}
$$

Le modèle de Buhlman-Straub

La meilleure approximation linéaire de la prime de risque d'un assuré est la prime de crédibilité

$$\pi^{\text{BS}}_{i,t+1} = z_i N_{iw} + (1 - z_i)m, \text{ où } z_i = \frac{w_{i\cdot}}{w_{i\cdot} + s^2/a}.$$

Les estimateurs des paramètres de structure sont les suivants:

$$\begin{aligned}
\widehat{m} &= N_{zw} = \sum_{i=1}^{I} \frac{z_i}{z_\cdot} N_{iw}, \\
\widehat{s}^2 &= \frac{1}{I(t-1)} \sum_{i=1}^{I} \sum_{\tau=1}^{t} w_{i\tau}(N_{i\tau} - N_{iw})^2 \\
\widehat{a} &= \frac{w_{\cdot\cdot}}{w_{\cdot\cdot}^2 - \sum_{i=1}^{I} w_{i\cdot}^2} \left(\sum_{i=1}^{I} w_{i\cdot}(N_{iw} - N_{ww})^2 - (I-1)\widehat{s}^2 \right).
\end{aligned}$$

Le modèle de Buhlman-Straub

L'estimateur $\widehat{a}$ peut être négatif. En pratique, on choisit $\max(\widehat{a},0)$, qui est alors biaisé.

Un estimateur alternatif du paramètre a est, quant à lui, toujours positif:

$$\tilde{a} = \frac{1}{I-1}\sum_{i=1}^{I} z_i(N_{iw}-N_{zw})^2.$$

Cet estimateur, dit de Bichsel–Straub, est en fait un pseudo-estimateur de Vylder (1981) dans la mesure où il dépend de paramètres inconnus. Il est évalué itérativement par la méthode du point fixe. On peut démontrer que si $\widehat{a}<0$, alors $\tilde{a}$ converge vers 0.

Enfin, il n'est pas difficile de vérifier que lorsque tous les poids sont égaux et que le nombre d'années d'expérience est le même pour tous les assurés, alors le modèle de Bühlmann–Straub est en tous points équivalent à celui de Bühlmann.

Modèle de Bühlman et Approche Bayésienne

Let us get back on the very first example, where we've been computing $\mathbb{E}\left(N_{t+1}|N_1,...,N_t\right)$, while Bühlmann suggested to use

$$z\overline{n}+(1-z)\,m$$

where $m=\mathbb{E}\left(X\right)=\mathbb{E}\left(\mathbb{E}\left(X|\Theta\right)\right)$, $z=\dfrac{n}{n+k}$, with $k=\nu/a$, where $\nu=\mathbb{E}\left(\text{Var}\left(X|\Theta\right)\right)$ and $a=\text{Var}\left(\mathbb{E}\left(X|\Theta\right)\right)$.

Let us compute those quantities, and compare them with our initial results.

Modèle de Bühlman et Approche Bayésienne

The 3 values for $\mathbb{E}(X|\Theta=\theta)$ are 1, 2 and 3, so that the variance is

$$a = \text{Var}(\mathbb{E}(X|\Theta)) = 0.7\times 1^2 + 0.2\times 2^2 + 0.1\times 3^2 - 1.4^2 = 0.44,$$

(the *between* variance) and because the 3 values for $Var(X|\Theta)$ are 1 2 and 3, the mean is

$$\nu = \mathbb{E}(\text{Var}(X|\Theta)) = 0.7\times 1 + 0.2\times 2 + 0.1\times 3 = 1.4,$$

(the *within* variance). One gets, based on two observations (which yield to 1 and 2)

$$z = \frac{2}{2+1.4/0.44} = 0.386$$

so that the estimated premium is

$$0.386\frac{3}{2} + (1-0.386)\times 1.4 = 1.439.$$

(to be compared with 1.3256 obtained with the Bayesian approach).

Modèle de Bühlman et Approche Bayésienne

In the Poisson-Gamma model,

$$m = \mathbb{E}\left(\mathbb{E}\left(X|\Theta\right)\right) = \mathbb{E}\left(\Theta\right) = \beta/\alpha$$

$$a = \text{Var}\left(\mathbb{E}\left(X|\Theta\right)\right) = \text{Var}\left(\Theta\right) = \frac{\beta}{\alpha^2}$$

while

$$\nu = \mathbb{E}\left(\text{Var}\left(X|\Theta\right)\right) = \mathbb{E}\left(\Theta\right) = \frac{\beta}{\alpha},$$

so that Bühlmann's premium is

$$z\overline{n} + (1-z)\frac{\beta}{\alpha} \text{ where } z = \frac{t}{t+k},\ k = \frac{\nu}{a} = \alpha$$

ie.

$$\left[1 - \frac{t}{\alpha+t}\right]\frac{\beta}{\alpha} + \frac{t}{\alpha+t}\overline{n}.$$

Note that here, Bühlmann's premium is equal to Bayes's pure premium.

Loi de la Famille Exponentielle et Loi Conjuguée

En fait, ce résultat est beaucoup plus général. Si la loi de N est une loi de la famille exponentielle, et que la loi a priori de Θ est sa loi conjuguée, alors la loi a posteriori est dans la même famille (à un changement de paramètres près),

- si $N \sim \mathcal{P}(\Theta)$ avec $\Theta \sim \mathcal{G}(\alpha, \beta)$, alors

$$\Theta|\underline{\boldsymbol{N}_t} \sim \mathcal{G}\left(\alpha + \sum_{i=1}^{t} n_i, \beta + t\right)$$

- si $N \sim \mathcal{B}(n, \Theta)$ avec $\Theta \sim \mathcal{B}(\alpha, \beta)$, alors

$$\Theta|\underline{\boldsymbol{N}_t} \sim \mathcal{B}\left(\alpha + \sum_{i=1}^{t} n_i,\ \beta + \sum_{i=1}^{t}(1 - n_i)\right)$$

Loi de la Famille Exponentielle et Loi Conjuguée

- si $N \sim \mathcal{G}(\Theta)$ avec $\Theta \sim \mathcal{B}(\alpha, \beta)$, alors

$$\Theta|\underline{\boldsymbol{N}_t} \sim \mathcal{B}\left(\alpha + t, \beta + \sum_{i=1}^{t} n_i\right)$$

- si $N \sim BN(r, \Theta)$ avec $\Theta \sim \mathcal{B}(\alpha, \beta)$, alors

$$\Theta|\underline{\boldsymbol{N}_t} \sim \mathcal{B}\left(\alpha + \sum_{i=1}^{t} n_i, \beta + +rt\right)$$

- si $\boldsymbol{N} \sim \mathcal{M}(\boldsymbol{\Theta})$ avec $\boldsymbol{\Theta} \sim \mathcal{D}(\boldsymbol{\alpha})$, alors

$$\Theta|\underline{\boldsymbol{N}_t} \sim \mathcal{D}\left(\boldsymbol{\alpha} + \sum_{i=1}^{t} \boldsymbol{n}_i\right)$$

Calculs de Crédibilité

Pour l'implémentation, considérons

```
> library(actuar)
> data(hachemeister)
> hachemeister
      state ratio.1 ratio.2 ratio.3 ratio.4 ratio.5 ratio.6
[1,]      1    1738    1642    1794    2051    2079    2234
[2,]      2    1364    1408    1597    1444    1342    1675
[3,]      3    1759    1685    1479    1763    1674    2103
[4,]      4    1223    1146    1010    1257    1426    1532
[5,]      5    1456    1499    1609    1741    1482    1572
      ratio.7 ratio.8 ratio.9 ratio.10 ratio.11 ratio.12
[1,]     2032    2035    2115     2262     2267     2517
[2,]     1470    1448    1464     1831     1612     1471
[3,]     1502    1622    1828     2155     2233     2059
[4,]     1953    1123    1343     1243     1762     1306
[5,]     1606    1735    1607     1573     1613     1690
```

Calculs de Crédibilité

```
      weight.1 weight.2 weight.3 weight.4 weight.5 weight.6
[1,]      7861     9251     8706     8575     7917     8263
[2,]      1622     1742     1523     1515     1622     1602
[3,]      1147     1357     1329     1204      998     1077
[4,]       407      396      348      341      315      328
[5,]      2902     3172     3046     3068     2693     2910
      weight.7 weight.8 weight.9 weight.10 weight.11 weight.12
[1,]      9456     8003     7365      7832      7849      9077
[2,]      1964     1515     1527      1748      1654      1861
[3,]      1277     1218      896      1003      1108      1121
[4,]       352      331      287       384       321       342
[5,]      3275     2697     2663      3017      3242      3425
```

Calculs de Crédibilité

Le calcul des coefficients donne, dans le modèle de Bühlmann

```
> cm(~state, hachemeister, ratios = ratio.1:ratio.12)

Structure Parameters Estimators

  Collective premium: 1671

  Between state variance: 72310
  Within state variance: 46040
```

Calculs de Crédibilité

On peut retrouver ces niveaux à la main

```
> H <-  hachemeister[,2:13]
> n <- ncol(H)
> (m <- mean(H))
[1] 1671.017
> (s2 <- mean(apply(H,1,var)))
[1] 46040.47
> (a <- var(apply(H,1,mean))-s2/n)
[1] 72310.02
```

Calculs de Crédibilité

et dans le modèle de Bühlmann–Straub

```
> cm(~state, hachemeister, ratios = ratio.1:ratio.12,
      weights = weight.1:weight.12)

Structure Parameters Estimators

  Collective premium: 1684

  Between state variance: 89639
  Within state variance: 139120026
```

Calculs de Crédibilité

Pour calculer les primes de crédibilité

```
> fit <- cm(~state, hachemeister, ratios = ratio.1:ratio.12,
      weights = weight.1:weight.12)
> predict(fit)
[1] 2055 1524 1793 1443 1603
```

Le détail des quantités estimées est donné dans la sortie

```
> summary(fit)

Structure Parameters Estimators

  Collective premium: 1684

  Between state variance: 89639
  Within state variance: 139120026
```

Calculs de Crédibilité

```
Detailed premiums

  Level: state
    state Indiv. mean Weight Cred. factor Cred. premium
    1     2061        100155 0.9847       2055
    2     1511         19895 0.9276       1524
    3     1806         13735 0.8985       1793
    4     1353          4152 0.7279       1443
    5     1600         36110 0.9588       1603
```

Modèle(s) Hiérarchique(s)

On note les données N_{ijt}, où l'indice $i = 1, \ldots, I$ identifie la cohorte, l'indice $j = 1, \ldots, J_i$ identifie l'assuré et l'indice $t = 1, \ldots, n_{ij}$ identifie l'observation dans une période. À chacune de ces observations correspond un poids w_{ijt}, connu.

On modélise l'hétérogénéité à l'aide des variables aléatoires Φ_i et Θ_{ij} représentant les niveaux des risque des cohortes et des assurés, dans l'ordre. Les hypothèses du modèle sont les suivantes:

1. Les variables aléatoires $\Phi_1, \ldots, \Phi_I$ sont i.i.d.
2. Les variables aléatoires $\Theta_{i1}, \ldots, \Theta_{i,J_i}$ sont conditionnellement i.i.d. sachant Φ_i, $i = 1, \ldots, I$.
3. Les variables aléatoires $N_{ij1}, \ldots, N_{ijn_{ij}}$ sont conditionnellement i.i.d. sachant Θ_{ij} (et Φ_i), $j = 1, \ldots, I_j$. Leur variance est finie.

4. Pour tout $s, t = 1, \ldots, n_{ij}$,

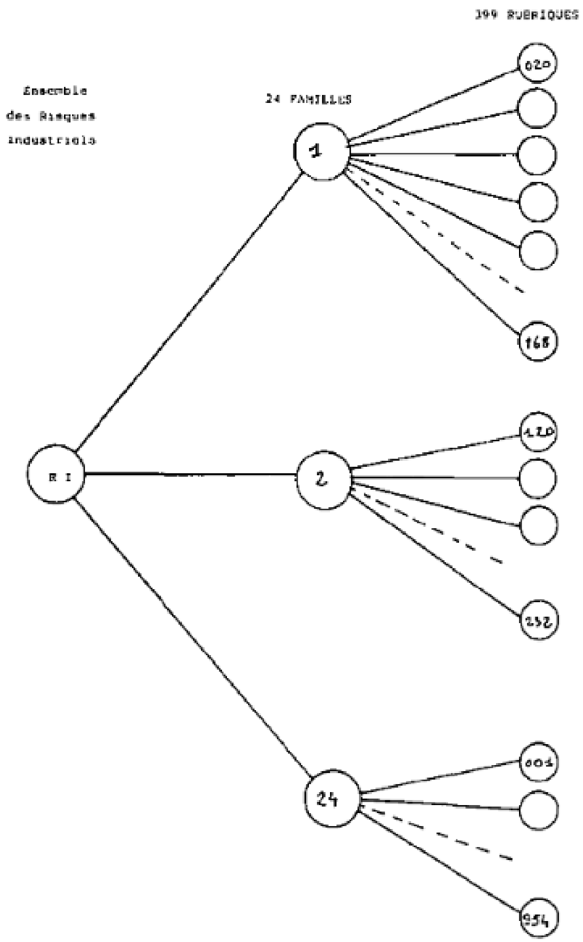
$$\begin{aligned}\mathbb{E}[N_{ijt}|\Theta_{ij},\Phi_i] &= \mu(\Theta_{ij},\Phi_i)\\ \text{Cov}(N_{ijs},N_{ijt}|\Theta_{ij},\Phi_i) &= \begin{cases}\frac{\sigma^2(\Theta_{ij},\Phi_i)}{w_{ijt}}, & s=t\\ 0, & s\neq t.\end{cases}\end{aligned}$$

On définit également les paramètres de structure suivants: la moyenne collective

$$\begin{aligned} m &= \mathbb{E}[\mathbb{E}[\mu(\Theta_{ij},\Phi_i)]|\Phi_i], \text{la variance intra-assuré moyenne}\\ s^2 &= \mathbb{E}[\mathbb{E}[\sigma^2(\Theta_{ij},\Phi_i)]|\Phi_i], \text{la variance inter-assuré (ou intra-cohorte) moyenne}\\ a &= \mathbb{E}[\text{Var}\mu(\Theta_{ij},\Phi_i)]|\Phi_i]\text{et, enfin, la variance inter-cohortes}\\ b &= \text{Var}\mathbb{E}[\mu(\Theta_{ij},\Phi_i)]|\Phi_i].\end{aligned}$$

Modèle(s) Hiérarchique(s)

Utilisé par Cohen, Dupin & Lévi (1986) sur le risque incendie entreprise,



Typologie des Rubriques Industrielles en 24 familles

Famille Industrielle	Facteur de Credibilite (‰)	Taux de Credibilite (‰)	Complement de Fluctuation a 75% (‰)	Nombre de Rubriques	Signe de $\hat{W}$
Alimentation	96,6	1,140	0,21	43	positif
Textiles	93,4	1,426	0,46	60	positif
Vêtements	89,2	1,836	0,51	29	positif
Bois	91,4	1,966	0,30	19	négatif
Métaux 1	98,1	0,599	0,17	1	sans objet
Métaux 2	74,9	0,985	0,44	14	positif
Métaux 3	93,5	0,451	0,18	6	positif
Electricité	94,6	1,104	0,45	11	positif
Ciment, céramique	89,3	0,719	0,31	17	positif
Chimie minérale	76,3	0,873	0,31	18	négatif
Corps gras	38,2	1,027	1,01	12	positif
Chimie	95,9	0,846	0,17	30	positif
Résines	88,9	2,285	0,47	6	positif
Combustibles	87,1	0,403	0,20	22	positif
Fabrication papier	80,4	1,263	0,75	5	négatif
Façonnage papier	90,3	1,255	0,39	12	positif
Spectacle	83,2	0,850	0,33	10	positif
Garages	95,2	0,489	0,12	5	positif
Magasins divers	93,6	1,075	0,50	30	positif
Magasins de vente	95,1	1,459	0,29	10	positif
Magasins	93,3	1,169	0,48	1	sans objet
Risques agricoles	25,4	0,993	0,94	9	négatif
Ensembles immobiliers	92,3	0,255	0,06	2	positif
Divers	95,1	0,668	0,22	27	positif
Portefeuille		1,047	0,10	399	

Prévision dans les Modèle(s) Hiérarchique(s)

On cherche à estimer les primes de risque des contrats $\mu(\Theta_{ij}, \Phi_i)$. Or, la prime de crédibilité d'un assuré dans le modèle hiérarchique est de la même forme qu'auparavant, sauf que le complément de crédibilité $(1 - z_i)$ est maintenant attribué à la prime de crédibilité de la cohorte. Cette dernière est une moyenne pondérée de l'expérience de la cohorte et de celle du portefeuille. Ainsi, la meilleure prévision linéaire de $\mu(\Theta_{ij}, \Phi_i)$ est donnée par les équations récursives suivantes:

$$\begin{aligned} \widehat{\pi}_{ijk}^{\text{H}} &= z_{ij} N_{ijw} + (1 - z_{ij}) \widehat{\pi}_i^{\text{H}}, \\ \widehat{\pi}_i^{\text{H}} &= z_i N_{izw} + (1 - z_i) m, \end{aligned}$$

Prévision dans les Modèle(s) Hiérarchique(s)

avec les facteurs de crédibilité

$$\begin{aligned} z_{ij} &= \frac{w_{ij\cdot}}{w_{ij\cdot}+s^2/a}, \\ w_{ij\cdot} &= \sum_{t=1}^{n_{ij}} w_{ijt}, \\ z_i &= \frac{z_{i\cdot}}{z_{i\cdot}+a/b}, \\ z_{i\cdot} &= \sum_{j=1}^{J_i} z_{ij} \end{aligned}$$

et les moyennes pondérées

$$\begin{aligned} N_{ijw} &= \sum_{t=1}^{n_{ij}} \frac{w_{ijt}}{w_{ij\cdot}} N_{ijt}, \\ N_{izw} &= \sum_{j=1}^{J_i} \frac{z_{ij}}{z_{i\cdot}} N_{ijw}. \end{aligned}$$

Estimation des paramètres de structure

Le modèle hiérarchique à deux niveaux compte quatre paramètres de structure à estimer. En premier lieu, l'estimateur de la moyenne collective m est

$$\widehat{m} = N_{zzw} = \sum_{i=1}^{I} \frac{z_i}{z_\cdot} N_{izw}.$$

L'estimateur de la dispersion intra-assuré moyenne s^2 est une généralisation de l'équation précédante:

$$\widehat{s}^2 = \frac{1}{\sum_{i=1}^{I} \sum_{j=1}^{J_i} (n_{ij} - 1)} \sum_{i=1}^{I} \sum_{j=1}^{J_i} \sum_{t=1}^{n_{ij}} w_{ijt} (N_{ijt} - N_{ijw})^2.$$

Il existe trois grands types d'estimateurs des composants de variance a et b, tous supportés par la fonction `cm`. Tout d'abord, soit

$$\begin{aligned} A_i &= \sum_{j=1}^{J_i} w_{ij\cdot}(N_{ijw}-N_{iww})^2-(J_i-1)s^2 \\ c_i &= w_{i\cdot\cdot}-\sum_{j=1}^{J_i}\frac{w_{ij\cdot}^2}{w_{i\cdot\cdot}} \\ B &= \sum_{i=1}^{I} z_{i\cdot}(N_{izw}-\overline{X}_{zzw})^2-(I-1)a \\ d &= z_{\cdot\cdot}-\sum_{i=1}^{I}\frac{z_{i\cdot}^2}{z_{\cdot\cdot}}, \end{aligned}$$

avec

$$\overline{X}_{zzw}=\sum_{i=1}^{I}\frac{z_{i\cdot}}{z_{\cdot\cdot}}N_{izw}.$$

On remarquera au passage la différence entre les moyennes pondérées.

On peut démontrer que $\mathbb{E}[A_i] = c_i a$ et $\mathbb{E}[B] = db$. Alors, les estimateurs de Bühlmann–Gisler sont

$$\begin{aligned}\widehat{a} &= \frac{1}{I}\sum_{i=1}^{I}\max\left(\frac{A_i}{c_i},0\right)\\ \widehat{b} &= \max\left\{\tfrac{B}{d},0\right\},\end{aligned}$$

les estimateurs de Ohlsson sont

$$\begin{aligned}\widehat{a}^{\mathsf{T}} &= \frac{\sum_{i=1}^{I} A_i}{\sum_{i=1}^{I} c_i}\\ \widehat{b}^{\mathsf{T}} &= \frac{B}{d}\end{aligned}$$

et les (pseudo-)estimateurs itératifs sont

$$\begin{aligned}\tilde{a} &= \frac{1}{\sum_{i=1}^{I}(J_i-1)}\sum_{i=1}^{I}\sum_{j=1}^{J_i} z_{ij}(N_{ijw}-N_{izw})^2\\ \tilde{b} &= \frac{1}{I-1}\sum_{i=1}^{I} z_i(N_{izw}-N_{zzw})^2,\end{aligned}$$

Le modèle de Bühlmann–Straub est un cas spécial du modèle hiérarchique.

Évaluation Numérique

La fonction `cm` présente une interface unifiée pour les divers modèles de crédibilité présentés jusqu'à présent reconnaissant le lien de filiation mentionné ci-dessus. Pour traiter un modèle hiérarchique, il suffit d'en décrire la structure de classification dans le premier argument de `cm()`.

Afin de donner un exemple facile à reproduire, nous regroupons les états 1 et 3 du jeu de données de Hachemeister dans une cohorte, et les états 2, 4 et 5 dans une seconde cohorte:

```
> X <- cbind(cohort = c(1, 2, 1, 2, 2), hachemeister)
> X[, 1:7]
     cohort state ratio.1 ratio.2 ratio.3 ratio.4 ratio.5
[1,]      1     1    1738    1642    1794    2051    2079
[2,]      2     2    1364    1408    1597    1444    1342
[3,]      1     3    1759    1685    1479    1763    1674
[4,]      2     4    1223    1146    1010    1257    1426
[5,]      2     5    1456    1499    1609    1741    1482
```

Ceci démontre également que les données n'ont pas à être triées par niveau. Le modèle ajusté avec les estimateurs itératifs (??) et (??) est:

```
> fit <- cm(~cohort + cohort:state, data = X, ratios = ratio.1:ratio
    .12,
     weights = weight.1:weight.12, method = "iterative")
> fit
Call:
cm(formula = ~cohort + cohort:state, data = X, ratios = ratio.1:ratio
    .12,
    weights = weight.1:weight.12, method = "iterative")

Structure Parameters Estimators

  Collective premium: 1746

  Between cohort variance: 88981
  Within cohort/Between state variance: 10952
  Within state variance: 139120026
```

Les estimateurs de Bühlmann–Gisler et de Ohlsson sont obtenus

```
> predict(fit)
$cohort
[1] 1949 1543

$state
[1] 2048 1524 1875 1497 1585
```

La fonction `summary` affiche toujours l'ensemble des principaux résultats:

```
> summary(fit)

Structure Parameters Estimators

  Collective premium: 1746

  Between cohort variance: 88981
  Within cohort/Between state variance: 10952
  Within state variance: 139120026
```

```

Detailed premiums

  Level: cohort
    cohort Indiv. mean Weight Cred. factor Cred. premium
    1      1967        1.407  0.9196       1949
    2      1528        1.596  0.9284       1543

  Level: state
    cohort state Indiv. mean Weight Cred. factor Cred. premium
    1      1     2061        100155 0.8874       2048
    2      2     1511         19895 0.6103       1524
    1      3     1806         13735 0.5195       1875
    2      4     1353          4152 0.2463       1497
    2      5     1600         36110 0.7398       1585
```

Crédibilité et Régression

En passant du modèle de crédibilité de base de Bühlmann au modèle de Bühlmann–Straub, c'est l'hypothèse de variance égale qui a été remplacée par celle de variance variable. Toutefois, on a conservé l'hypothèse de prime de risque constante dans le temps. Or, celle-ci peut ne pas être satisfaite dans un portefeuille soumis, par exemple, à une hausse structurelle des coûts dans le temps. L'utilisation du modèle de Bühlmann–Straub dans un tel contexte résulterait à coup sûr en une sous-estimation des coûts. C'est pour remédier à cette situation que Hachemeister (1975) a proposé le modèle de crédibilité de régression avec coefficients aléatoires.

On suppose maintenant que, conditionnellement au niveau de risque Θ_i, les observations $\boldsymbol{X}_i = (N_{i1}, \ldots, N_{in})^\mathsf{T}$ d'un assuré obéissent au modèle de régression linéaire classique. On pose donc que

$$\begin{aligned}\mathbb{E}[\boldsymbol{X}_i|\Theta_i=\theta_i] &= \boldsymbol{Y}\boldsymbol{\beta}(\theta_i)\\ \text{Cov}(\boldsymbol{X}_i,\boldsymbol{X}_i^\mathsf{T}|\Theta_i=\theta_i) &= \sigma^2(\theta_i)\boldsymbol{W}_i^{-1},\end{aligned}$$

où $\boldsymbol{Y}$ est une matrice de schéma $n \times (p+1)$, $\boldsymbol{\beta}(\theta_i) = (\beta_0(\theta_i), \cdots, \beta_p(\theta_i))^\mathsf{T}$ est un vecteur de coefficients aléatoires et $\boldsymbol{W}$ est une matrice de poids connus. Nous ne considérons, ici, que des matrices de poids diagonales de la forme

$$\boldsymbol{W}_i = \begin{bmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_{in} \end{bmatrix}.$$

Individuellement, nous devons estimer les coefficients d'un modèle de régression pondérée. On sait que, dans ce cas, la meilleure estimation linéaire sans biais de $\boldsymbol{\beta}(\theta_i)$ est

$$\boldsymbol{b}_i = (\boldsymbol{Y}^\mathsf{T}\boldsymbol{W}_i\boldsymbol{Y})^{-1}\boldsymbol{Y}^\mathsf{T}\boldsymbol{W}_i\boldsymbol{X}_i.$$

La grande contribution de Hachemeister (1975) aura été de démontrer que l'estimation ci-dessus peut être améliorée en tirant profit de l'information présente dans l'expérience des autres assurés du portefeuille. Ainsi, l'estimateur de crédibilité du vecteur de coefficients $\boldsymbol{\beta}(\theta_i)$ est

$$\begin{aligned} \boldsymbol{b}_i^a &= \boldsymbol{Z}_i \boldsymbol{b}_i + (\boldsymbol{I} - \boldsymbol{Z}_i)\boldsymbol{m} \\ \text{où } \boldsymbol{Z}_i &= \boldsymbol{A}^{-1}(\boldsymbol{A} + s^2(\boldsymbol{Y}^\top \boldsymbol{W}_i \boldsymbol{Y})^{-1}). \end{aligned}$$

Les quantités $\boldsymbol{m} = \mathbb{E}[\boldsymbol{\beta}(\theta)]$, $s^2 = \mathbb{E}[\sigma^2(\Theta)]$ et

$$\boldsymbol{A} = \text{Cov}(\boldsymbol{\beta}(\theta), \boldsymbol{\beta}(\theta)^\top)$$

forment l'ensemble des paramètres de structure (scalaires et matriciels) du modèle.

En définitive, la meilleure estimation des cofficients de régression d'un assuré est une combinaison des coefficients estimés à partir de l'expérience individuelle, $\boldsymbol{b}_i$, et des coefficients moyens dans le portefeuille, $\boldsymbol{m}$.

Prévision et Estimation des paramètres de structure

Soit $\boldsymbol{y}_{n+1}$ un vecteur des valeurs “futures” de la matrice de schéma $\boldsymbol{Y}$. Alors, la meilleure prévision linéaire de la prime de risque correspondant à ce vecteur $\boldsymbol{y}_{n+1}$ est

$$\pi^{\mathrm{R}}_{i,n+1} = \boldsymbol{y}_{n+1}^{\mathsf{T}} \boldsymbol{b}_i^a.$$

Nous utilisons les estimateurs des paramètres de structure proposés par Goovaerts *et al.* (1987). Ils sont directement inspirés des estimateurs du modèle de Bühlmann–Straub. On a

$$\begin{aligned}
\widehat{\boldsymbol{m}} &= (\boldsymbol{Z}_i)^{-1} \textstyle\sum_{i=1}^{I} \boldsymbol{Z}_i \boldsymbol{b}_i \\
\widehat{s}^2 &= \tfrac{1}{I(n-p)} \textstyle\sum_{i=1}^{I} (\boldsymbol{X}_i - \boldsymbol{Y}\boldsymbol{b}_i)^{\mathsf{T}} \boldsymbol{W} (\boldsymbol{X}_i - \boldsymbol{Y}\boldsymbol{b}_i) \\
\widehat{\boldsymbol{A}} &= \tfrac{1}{I-1} \textstyle\sum_{i=1}^{I} \boldsymbol{Z}_i (\boldsymbol{b}_i - \widehat{\boldsymbol{m}})(\boldsymbol{b}_i - \widehat{\boldsymbol{m}})^{\mathsf{T}}.
\end{aligned}$$

L'estimateur $\widehat{\boldsymbol{A}}$ est évalué itérativement. Comme la matrice doit être symétrique, $\widehat{\boldsymbol{A}}$ est remplacée à chaque itération par $(\widehat{\boldsymbol{A}} + \widehat{\boldsymbol{A}}^{\mathsf{T}})/2$.

Évaluation Numérique

On peut ajuster un modèle de crédibilité de régression avec la fonction `cm` en ajoutant les arguments `regformula` et `regdata`. Le premier est une formule de la forme ˜ décrivant le modèle de régression et le second est un *data frame* contenant la matrice de schéma $\boldsymbol{Y}$.

Par exemple, pour ajuster le simple modèle de régression linéaire

$$N_{it} = \beta_0 + \beta_1 t + \varepsilon_t, \quad t = 1, \ldots, 12$$

aux données de Hachemeister (1975) il suffit de faire

```
> fit <- cm(~state, hachemeister, regformula = ~time,
     regdata = data.frame(time = 1:12), ratios = ratio.1:ratio.12,
     weights = weight.1:weight.12)
> fit
Call:
cm(formula = ~state, data = hachemeister, ratios = ratio.1:ratio.12,
    weights = weight.1:weight.12, regformula = ~time, regdata = data.
    frame(time = 1:12))
```

```

Structure Parameters Estimators

  Collective premium: 1469 32.05

  Between state variance:  24154 2700.0
                            2700  301.8
  Within state variance: 49870187
```

Comme précédemment, les primes de crédibilité sont évaluées par un appel à `predict()`. Dans le présent contexte, il faut également fournir le vecteur $\boldsymbol{y}$ des valeurs futures de la matrice de schéma. Ici, cela se réduit à:

```
> predict(fit, newdata = data.frame(time = 13))
[1] 2437 1651 2073 1507 1759
```

Systèmes Bonus-Malus

“L’expression bonus malus ou coefficient de réduction-majoration désigne une méthode de pondération de l’appréciation du risque par la sinistralité surtout utilisée pour les assurances auto”, wikipedia

Systèmes Bonus-Malus par classes

Consider the following bonus-malus scheme,

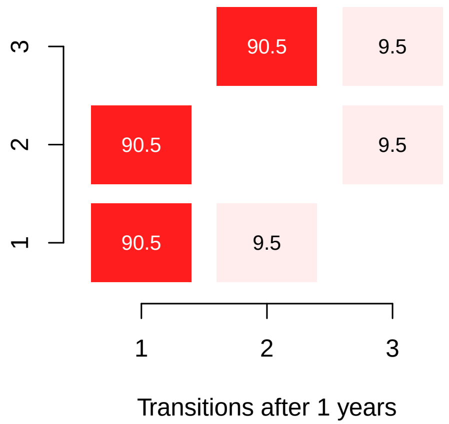
class	premium	claim	no claim
3	P_3	3	2
2	P_2	3	1
1	P_1	2	1

If claims occurrence is driven by an homogeneous Poisson process (with intensity λ), the class at time t is a Markov process. If p denote the probability to have no claim over a year, $p = e^{-\lambda}$, the transition probability matrix is here

$$M = \begin{pmatrix} p & 1-p & 0 \\ p & 0 & 1-p \\ 0 & p & 1-p \end{pmatrix}$$

Systèmes Bonus-Malus par classes

```
> load(
"http://freakonometrics.free.fr/bonusmalus.R")
> M=Mat_Fictif(.1)
> M
          [,1]       [,2]       [,3]
[1,] 0.9048374 0.09516258 0.00000000
[2,] 0.9048374 0.00000000 0.09516258
[3,] 0.0000000 0.90483742 0.09516258
> pwr=function(M,k){
+   N=M
+   if(k==0) N=diag(1,ncol(M))
+   if(k>1) for(i in 2:k) N=N%*%M
+   return(N)
+ }
```



Systèmes Bonus-Malus par classes

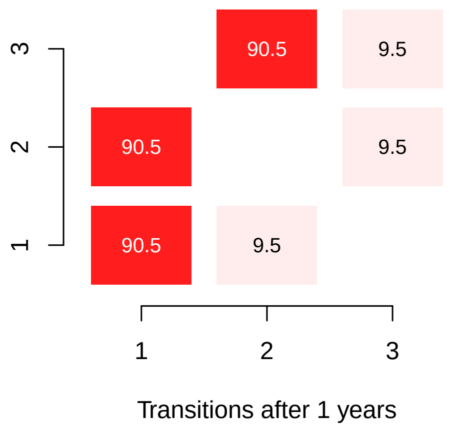
Here, the invariant measure, solution of

$$\mu = \mu M$$

is

$$\mu \propto (\kappa^2, \kappa, 1)^\mathsf{T}, \text{ where } \kappa = \frac{p}{1-p},$$

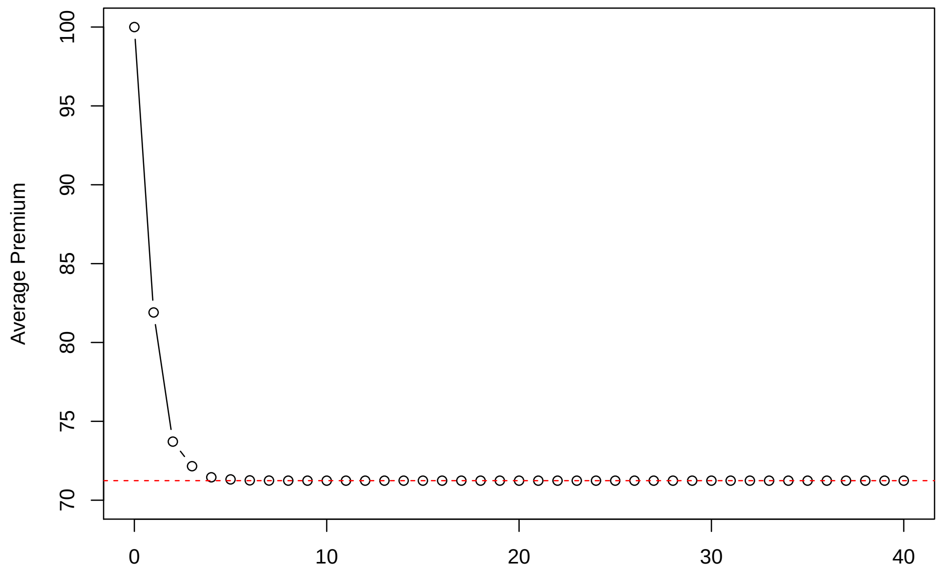
```
> pwr(Mat_Fictif(.1),50)[3,]
[1] 0.895871238 0.094219601 0.009909162
> mu=eigen(t(Mat_Fictif(.1)))$vectors[,1]
> mu/sum(mu)
[1] 0.895871238 0.094219601 0.009909162
```



Systèmes Bonus-Malus par classes

The stationnary average premium is

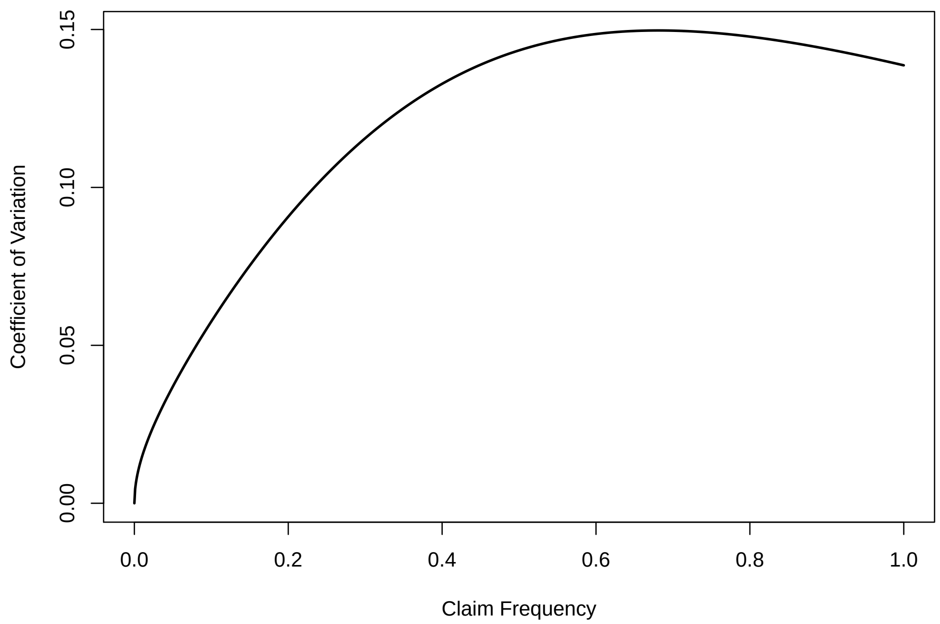
$$P_\infty = \frac{\kappa^2 P_1 + \kappa P_2 + P_3}{\kappa^2 + \kappa + 1}.$$



Systèmes Bonus-Malus par classes

It is also possible to define the coefficient of variation of the premium as the ratio

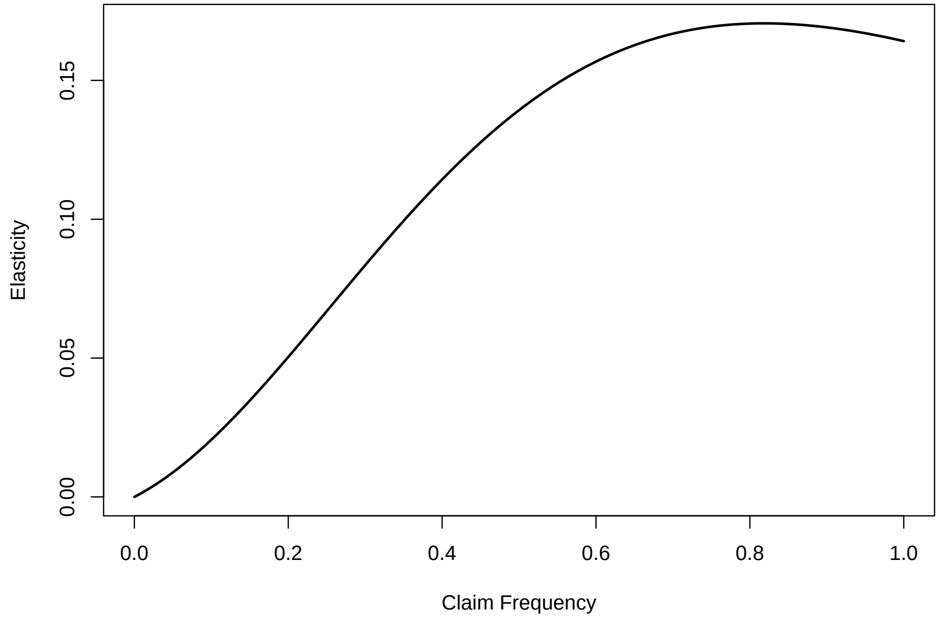
$$\frac{\sqrt{\text{Var}(\text{premium})}}{\mathbb{E}[\text{premium}]}, \text{ e.g. when premium} = P_\infty$$



Systèmes Bonus-Malus par classes

or the price elasticity of the premium as the ratio

$$\lambda \mapsto \frac{\partial \log P(\lambda)}{\partial \log \lambda}, \text{ e.g. when } P = P_\infty$$



Systèmes Bonus-Malus par classes: Hong Kong

HONG KONG
Table B-9. Hong Kong System

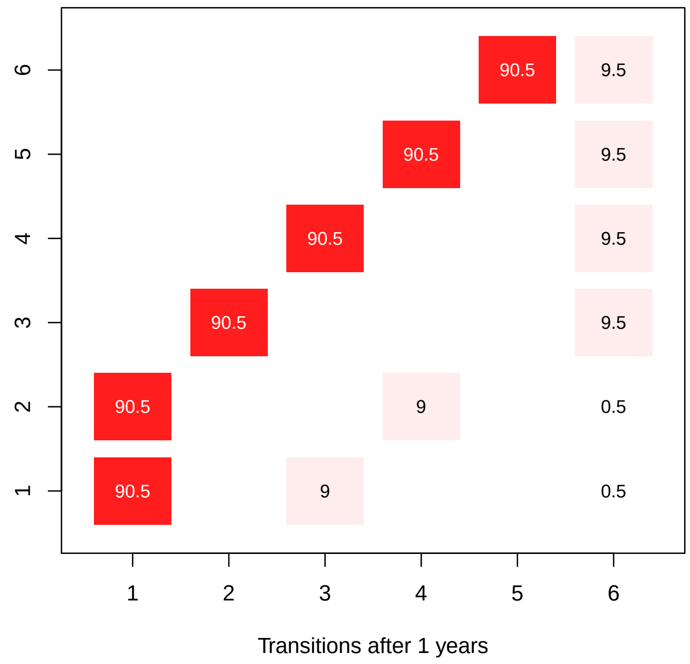
Class	Premium	Class After 0 Claims	Class After 1 Claims	Class After ≥2 Claims
6	100	5	6	6
5	80	4	6	6
4	70	3	6	6
3	60	2	6	6
2	50	1	4	6
1	40	1	3	6

Starting class: 6.

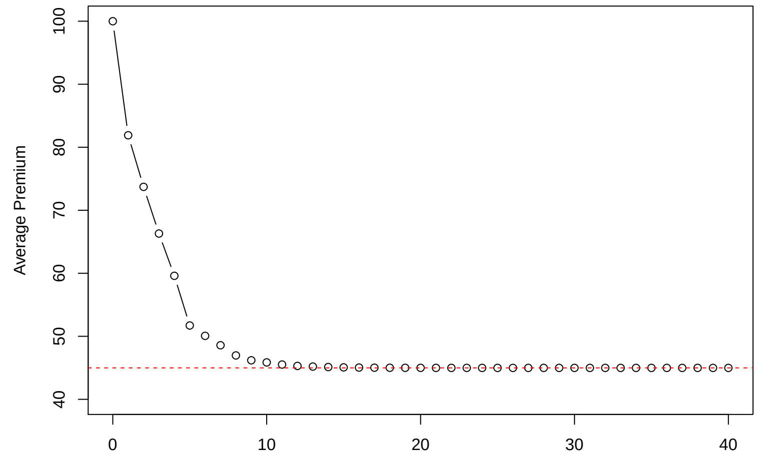
via Lemaire (1995)

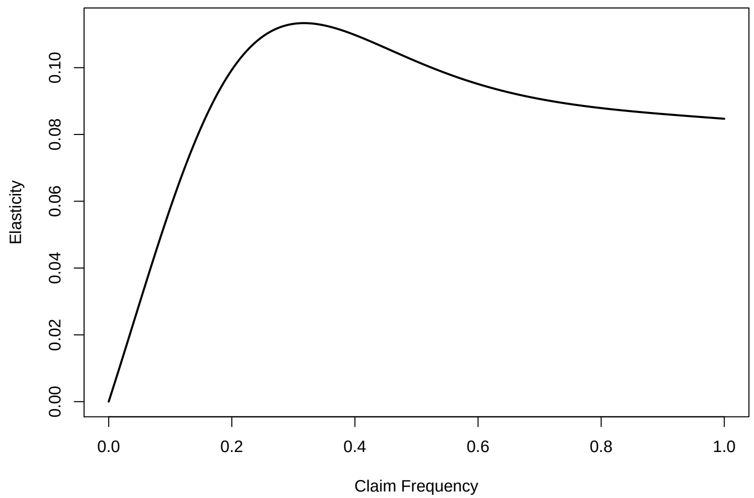
Systèmes Bonus-Malus par classes: Hong Kong

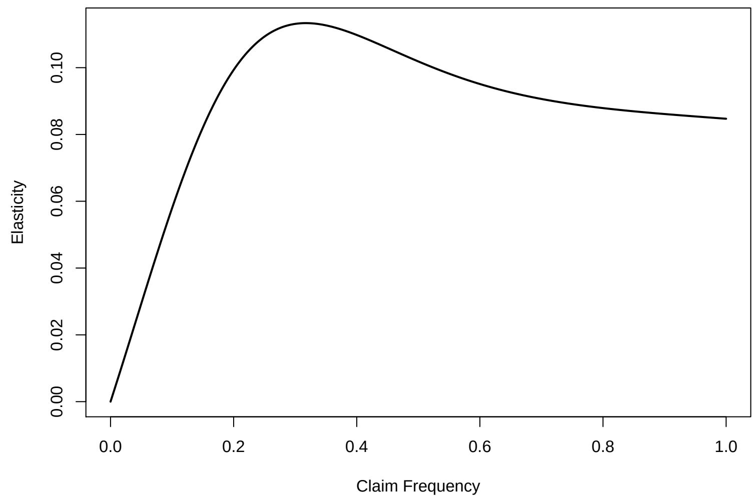
```
> M=Mat_Hong_Kong(.1)
> M
      [,1]  [,2]  [,3]  [,4]  [,5]  [,6]
[1,] 0.905 0.000 0.090 0.000 0.000 0.005
[2,] 0.905 0.000 0.000 0.090 0.000 0.005
[3,] 0.000 0.905 0.000 0.000 0.000 0.095
[4,] 0.000 0.000 0.905 0.000 0.000 0.095
[5,] 0.000 0.000 0.000 0.905 0.000 0.095
[6,] 0.000 0.000 0.000 0.000 0.905 0.095
```



Systèmes Bonus-Malus par classes: Hong Kong







Systèmes Bonus-Malus par classes: Allemagne

GERMANY

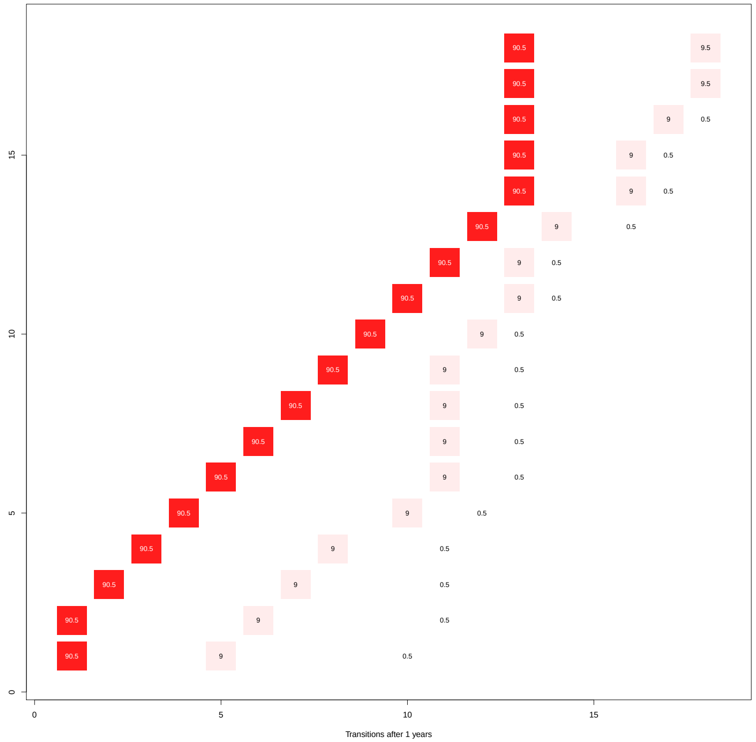
Table B-7. Old German System (Early 1980s)

Class	Premium	Class After 0 Claims	1	2	3	≥4
18	200	13	18	18	18	18
17	200	13	18	18	18	18
16	175	13	17	18	18	18
15	175	13	16	17	18	18
14	125	13	16	17	18	18
13	100	12	14	16	17	18
12	85	11	13	14	16	18
11	70	10	13	14	16	18
10	65	9	12	13	14	18
9	60	8	11	13	14	18
8	55	7	11	13	14	18
7	50	6	11	13	14	18
6	45	5	11	13	14	18
5	40	4	10	12	13	18
4	40	3	8	11	13	18
3	40	2	7	11	13	18
2	40	1	6	11	13	18
1	40	1	5	10	12	18

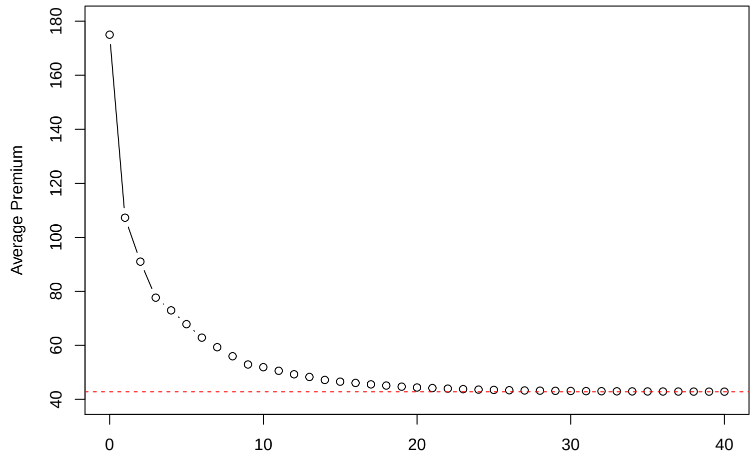
Starting class: 15, or 14 if the new entrant has held a valid driver's license for at least three years.

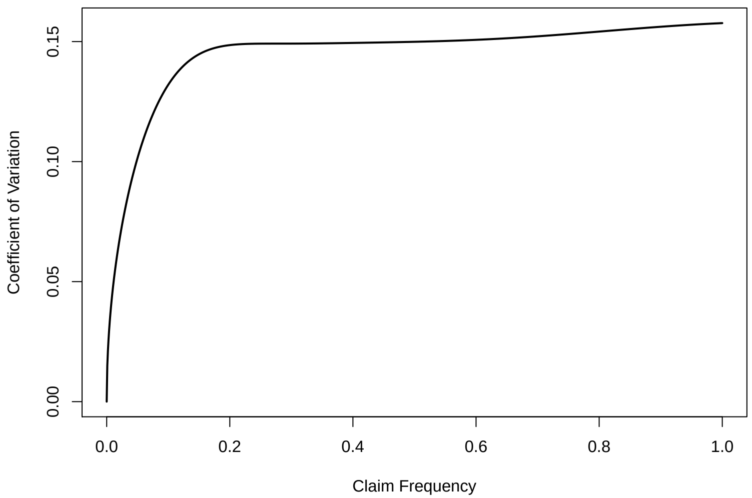
Systèmes Bonus-Malus par classes: Allemagne

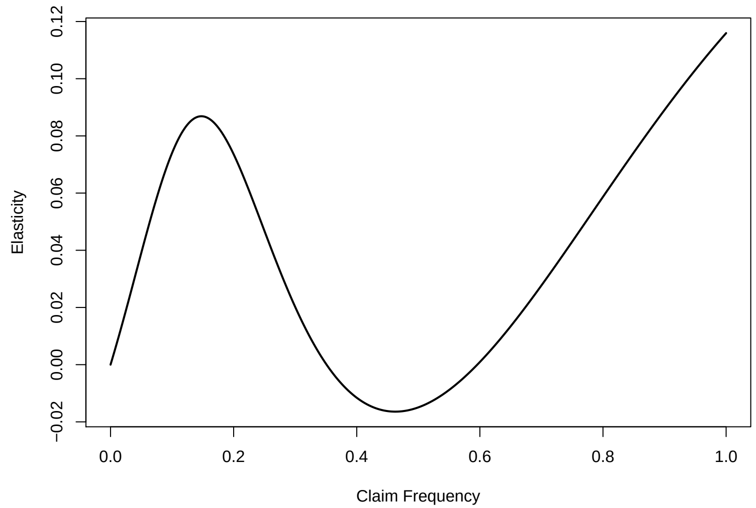
```
> M=Mat_Germany(.1)
      [,1]  [,2]  [,3]  [,4]  [,5]  [,6]
 [1,] 0.905 0.000 0.000 0.000 0.090 0.000
 [2,] 0.905 0.000 0.000 0.000 0.000 0.090
 [3,] 0.000 0.905 0.000 0.000 0.000 0.000
 [4,] 0.000 0.000 0.905 0.000 0.000 0.000
 [5,] 0.000 0.000 0.000 0.905 0.000 0.000
 [6,] 0.000 0.000 0.000 0.000 0.905 0.000
 [7,] 0.000 0.000 0.000 0.000 0.000 0.905
 [8,] 0.000 0.000 0.000 0.000 0.000 0.000
 [9,] 0.000 0.000 0.000 0.000 0.000 0.000
[10,] 0.000 0.000 0.000 0.000 0.000 0.000
[11,] 0.000 0.000 0.000 0.000 0.000 0.000
[12,] 0.000 0.000 0.000 0.000 0.000 0.000
[13,] 0.000 0.000 0.000 0.000 0.000 0.000
```



Systèmes Bonus-Malus par classes: Allemagne







Systèmes Bonus-Malus par classes

Table B-6. New Finnish System

Class	Premium	Class After 0 Claims	1	2	3	≥4
17	100	16	17	17	17	17
16	100	14	17	17	17	17
15	100	13	17	17	17	17
14	95	13	17	17	17	17
13	90	12	16	17	17	17
12	85	11	16	17	17	17
11	80	10	14	17	17	17
10	75	9	13	17	17	17
9	70	8	12	17	17	17
8	65	7	12	17	17	17
7	60	6	11	16	17	17
6	55	5	10	14	17	17
5	50	4	9	13	17	17
4	45	3	8	13	17	17
3	40	2	7	12	17	17
2	35	1	6	11	16	17
1	30	1	5	10	14	17

Starting class: 15.

Systèmes Bonus-Malus par classes

Table B-11. New Italian System (1991)

Class	Premium	Class After 0 Claims	1	2	3	4
18	200	17	18	18	18	18
17	175	16	18	18	18	18
16	150	15	18	18	18	18
15	130	14	17	18	18	18
14	115	13	16	18	18	18
13	100	12	15	18	18	18
12	94	11	14	17	18	18
11	88	10	13	16	18	18
10	82	9	12	15	18	18
9	78	8	11	14	17	18
8	74	7	10	13	16	18
7	70	6	9	12	15	18
6	66	5	8	11	14	17
5	62	4	7	10	13	16
4	59	3	6	9	12	15
3	56	2	5	8	11	14
2	53	1	4	7	10	13
1	50	1	3	6	9	12

Starting Class: 14. For more than four claims, the pattern continues (three-class penalty per claim).

Bonus Malus en France, par coefficient majorateur/minorateur

Comment fonctionne la clause de bonus-malus ?

La cotisation d'assurance évolue chaque année en fonction des accidents survenus. L'automobiliste qui ne cause pas d'accidents bénéficie d'un bonus, et celui qui est responsable d'un accident est pénalisé d'un malus. Bonus et malus sont exprimés par des coefficients de réduction ou de majoration, compris entre 0,50 et 3,50.

Bonus : chaque année sans sinistre engageant la responsabilité de l'assuré entraîne une réduction de 5 % de ce coefficient. Pour calculer le nouveau coefficient, il suffit de multiplier celui de l'année précédente par 0,95. Le maximum est fixé à 0,50, ce qui correspond à un bonus de 50 %.

Malus : tout accident dont l'assuré est totalement responsable entraîne une majoration de 25 % du coefficient précédent (12,5 % en cas de partage de responsabilité). On obtient le nouveau coefficient en multipliant le précédent par 1,25 (1,125 en cas de responsabilité partagée). Le malus disparaît après deux années d'assurance consécutives sans accident.

Cf. Kelle (2000) et Denuit *et al.* (2006).

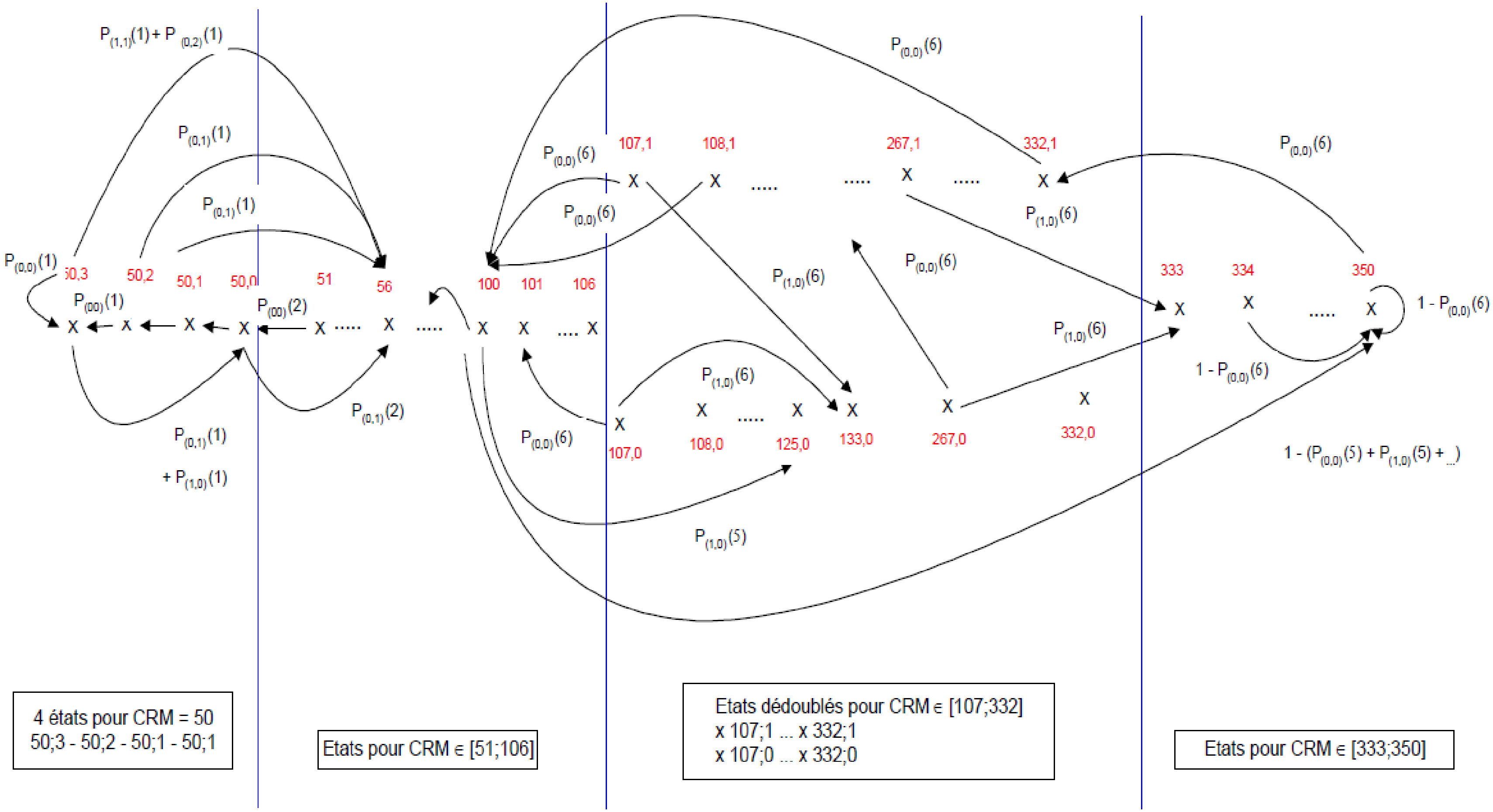
Bonus Malus en France, par coefficient majorateur/minorateur

Le système français n'est toutefois pas Markovien à l'ordre 1,

"Unconducteur ayant un coefficient 0,50 depuis au moins trois ans n'est pas pénalisé pour sonpremier sinistre (franchise de malus) [...] L'expression bonus malus ou coefficient de réduction-majoration désigne une méthode de pondération de l'appréciation du risque par la sinistralité surtout utilisée pour les assurances auto."

On peut toutefsoi bricoler en créant des niveaux 50.0, 50.1, etc.

Systèmes Bonus-Malus en France



via Kelle (2000)

Modèles Longitudinaux, Multiniveaux, etc

Dans les modèles longitudinaux, on cherche à modélier $y_{i,t}$, pour un individu i, à une certaine date t.

à poursuivre...