Actuariat de l'Assurance Non-Vie # 1

A. Charpentier (Université de Rennes 1)

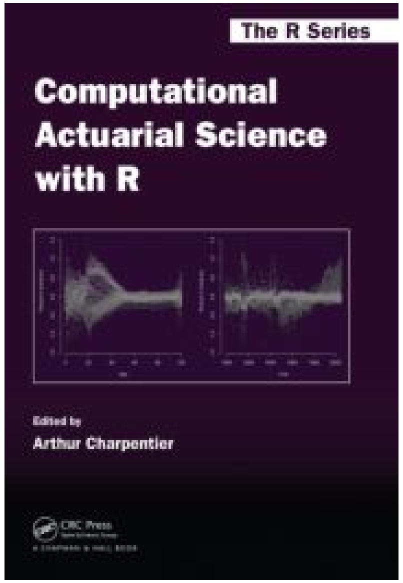
ENSAE ParisTech, 2016/2017.

http://freakonometrics.hypotheses.org

Rapide Introduction

A. Charpentier (Université de Rennes 1)

Professor of Actuarial Sciences, Mathematics Department, UQàM
(previously Economics Department, Univ. Rennes 1 & ENSAE Paristech
actuary in Hong Kong, IT & Stats FFSA)
PhD in Statistics (KU Leuven), Fellow Institute of Actuaries
MSc in Financial Mathematics (Paris Dauphine) & ENSAE
Editor of the freakonometrics.hypotheses.org's blog
Editor of Computational Actuarial Science, CRC



Quelques références

Denuit & Charpentier (2005). Mathématiques de l'Assurance Non-Vie. Economica.

Dhaene *et al.* (2005). Modern Actuarial Risk Theory. Springer Verlag.

Ohlsson & Johansson (2010) Non-life Insurance Pricing with Generalized Linear Models. Springer

de Jong & Heller (2008). Generalized Linear Models for Insurance Data. Cambridge University Press.

Denuit *et al.* (2007) Actuarial Modelling of Claims Counts. Wiley.

Cameron & Trividi (2013) Regression Analysis of Count Data Book. Cambridge University Press

Hilbe (2011) Negative Binomial Regression Cambridge University Press

McCullagh & Nelder (1989) Generalized Linear Models. CRC.

Wood (2006) Generalized Additive Models. CRC

Hastie *et al.* (1001) The Elements of Statistical Learning. Springer.

Charpentier. (2015). Computational Actuarial Pricing, with R. CRC.

Wüthrich & Merz. (2008). Stochastic Claims Reserving Methods in Insurance. Wiley.

Plan du Cours

- ■ Introduction Générale à la tarification en assurance non-vie
- ■■ Classification, régression logistique et arbres de classification
- ■■ Régression de Poisson et surdispersion (Binomiale Négative, Zero-Inflated)
- ■■ Tarification a posteriori, modèles de crédibilité
- ■■ Modélisation des coûts individuels, grands risques
- ■ Modèle collectif vs. modèle individuel, régression Tweedie
- ■■ Provisions pour Sinistres à Payer

Cf FSA: Applications of Statistical Techniques Module

La notion de 'prime pure'

Pour une variable de comptage

$$\mathbb{E}(N)=\sum_{n\in\mathbb{N}} n\mathbb{P}(N=n)=\sum_{n\in\mathbb{N}}\mathbb{P}(N>n)$$

Pour une variable continue positive

$$\mathbb{E}(X)=\int_{x\in\mathbb{R}_+} xf(x)dx=\int_{x\in\mathbb{R}_+}\mathbb{P}(X>x)dx$$

où $\mathbb{P}(X>x)=\overline{F}(x)=\int_x^\infty f(t)dt.$

La notion de 'prime pure'

Plus généralement,

$$\mathbb{E}(X) = \int_{x\in\mathbb{R}} x dF(x)$$

où F admet un nombre fini (ou dénombrable) de discontinuité $\{d_1 \leq d_2 \leq \cdots\}$.

$$dF(x) = \begin{cases} F(d_n) - F(d_n^-) \text{ si } x = d_n \\ \tilde{f}(x) \text{ sinon} \end{cases}$$

où

$$F(x) = \sum_{d_n \leq x} \mathbb{P}(X = d_n) + \int_{t\leq x} \tilde{f}(t)dt$$

Etant donné un risque X, la prime pure est $\pi_X = \mathbb{E}[X]$.

Calcul pour des lois classiques

Pour les lois discrètes classiques

- si $N \sim \mathcal{B}(n,p)$, $\mathbb{P}(N=k) = \binom{n}{k} p^k [1-p]^{n-k}$, alors $\mathbb{E}(X) = np$, et $\text{Var}(X) = np(1-p) < \mathbb{E}(X)$
- si $N \sim \mathcal{P}(\lambda)$, $\mathbb{P}(N=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$, alors $\mathbb{E}(X) = \lambda$, et $\text{Var}(X) = \lambda = \mathbb{E}(X)$
- si $N \sim NB(r,p)$, $\mathbb{P}(N=k) = \dfrac{\Gamma(r+k)}{k!\,\Gamma(r)} p^r [1-p]^k$, alors $\mathbb{E}(X) = \dfrac{r[1-p]}{p}$, et $\text{Var}(X) > \mathbb{E}(X)$

Calcul pour des lois classiques

Pour les lois continues classiques

- si $N \sim \mathcal{N}(\mu, \sigma^2)$, $\varphi(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right)$, alors $\mathbb{E}(X) = \mu$,
- si $N \sim LN(\mu, \sigma^2)$, $f(x) = \dfrac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\dfrac{[\ln(x)-\mu]^2}{2\sigma^2}\right)$, alors $\mathbb{E}(X) = \exp\left(\mu + \dfrac{\sigma^2}{2}\right)$,
- si $N \sim \mathcal{G}(\alpha, \beta)$, $f(x) = x^{\alpha-1}\dfrac{\beta^\alpha\, e^{-\beta\, x}}{\Gamma(\alpha)}$, alors $\mathbb{E}(X) = \dfrac{\alpha}{\beta}$,

Espérance mathématique d'une loi composée

Dans un modèle collectif, on s'intéresse à $S = \sum_{n=1}^{N} X_i$ si $N \geq 1$, 0 sinon.

Si les X_i sont i.i.d., indépendants de N, alors

$$\begin{aligned} \mathbb{E}[S] &= \mathbb{E}[\mathbb{E}(S|N)] = \sum_{k=1}^{+\infty} \mathbb{P}[N = k] \cdot \mathbb{E}\left[\sum_{i=1}^{k} X_i\right] \\ &= \left(\sum_{k=1}^{+\infty} \mathbb{P}[N = k] \cdot k\right) \mathbb{E}[X_1] = \mathbb{E}[N] \cdot \mathbb{E}[X_1]. \end{aligned}$$

où

- $\mathbb{E}[N]$ est la fréquence
- $\mathbb{E}[X_1]$ est le coût moyen

Espérance, Fair Price et Prime Pure

Pascal, Fermat ou Condorcet (XVIIIème siècle) proposaient d'évaluer le "*produit scalaire des probabilités et des gains*",

$$< \boldsymbol{p}, \boldsymbol{x} >= \sum_{i=1}^{n} p_i x_i = \mathbb{E}_{\mathbb{P}}(X),$$

selon la "*règle des parties*" $\Longrightarrow$ garantie un équilibre du système, en moyenne.

L'espérance mathématique est un prix "*juste*" (Feller (1943)),

- moindres carrés, $\mathbb{E}(X) = \text{argmin}\{\|X - c\|_{\ell^2}, c \in \mathbb{R}\}$
- loi des grands nombres, $\dfrac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(X)$,
- théorème central limite, $\dfrac{X_1 + \ldots + X_n}{n} \overset{\mathcal{L}}{\sim} \mathcal{N}\left(\mathbb{E}(X), \dfrac{Var(X)}{\sqrt{n}}\right)$,
- probabilité de ruine, $\lim\limits_{n\to\infty} \mathbb{P}\left(\dfrac{X_1 + \ldots + X_n}{n} > \pi\right) = 1$ pour $\pi < \mathbb{E}(X)$.

Prime Pure Sans Segmentation

$$\begin{cases} \mathbb{E}(Y) = \underset{m\in\mathbb{R}}{\text{argmin}}\{\|Y-m\|_{\ell_2}\} = \underset{m\in\mathbb{R}}{\text{argmin}}\{\mathbb{E}\left([Y-m]^2\right)\} \\ \text{Var}(Y) = \underset{m\in\mathbb{R}}{\min}\{\mathbb{E}\left([Y-m]^2\right)\} = \mathbb{E}\left([Y-\mathbb{E}(Y)]^2\right) \end{cases}$$

$$\begin{cases} \overline{y} = \underset{m\in\mathbb{R}}{\text{argmin}}\{\sum_{i=1}^n \frac{1}{n}[y_i-m]^2\} \\ s^2 = \underset{m\in\mathbb{R}}{\min}\{\sum_{i=1}^n \frac{1}{n}[y_i-m]^2\} = \sum_{i=1}^n \frac{1}{n}[y_i-\overline{y}]^2 \end{cases}$$

Éxcédant Moyen de Sinistre

L'excédant moyen de sinistre (encore appelé durée de vie moyenne restante en assurance sur la vie) est défini par

$$\begin{aligned} e_X(x) &= \mathbb{E}[X - x | X > x] \\ &= \frac{1}{\overline{F}_X(x)} \int_x^{+\infty} (s - x) dF_X(s), \quad x \geq 0. \end{aligned}$$

Éxcédant Moyen de Sinistre

Loi de probabilité	$e_X(x)$
$\mathcal{E}(\lambda)$	$\dfrac{1}{\theta}$
$\mathcal{G}(\alpha,\beta)$	$\dfrac{\alpha}{\beta}\dfrac{1-\Gamma(\alpha+1,\beta x)}{1-\Gamma(\alpha,\beta x)}-x$
$LN(\mu,\sigma^2)$	$\exp(\mu+\frac{\sigma^2}{2})\dfrac{\overline{\Phi}\left(\dfrac{\ln(x)-\mu-\sigma^2}{\sigma}\right)}{\overline{\Phi}\left(\dfrac{\ln(x)-\mu}{\sigma}\right)}-x$
$\mathcal{P}areto(\alpha,x_0)$	$\dfrac{x_0+x}{\alpha-1}$

Primes Stop-Loss ou Contrat avec Franchise

Etant donné un risque X, la prime stop-loss pour une franchise $d \geq 0$ est définie par

$$\pi_X(d) = \mathbb{E}[(X-d)_+] = e_X(d)\overline{F}(d).$$

Un traité de réassurance stop-loss (ou excédent de perte) consiste à faire prendre en charge par le réassureur la partie de la charge totale S des sinistres qui dépasse une certaine somme d. La portion réassurée, notée S^R, est donc définie par

$$S^R = (S-d)_+ = \begin{cases} 0, \text{ si } S \leq d, \\ S-d, \text{ si } S > d. \end{cases}$$

La prime pure que la cédante devra verser au réassureur pour un tel contrat, appelée prime stop-loss, est donnée par

$$\mathbb{E}[S^R] = \mathbb{E}[(S-d)_+].$$

Etant donné un risque X, la prime stop-loss pour une rétention $t \geq 0$ est définie par

$$\pi_X(t) = \mathbb{E}[(X - t)_+].$$

La fonction π_X est encore appelée la transformée stop-loss de la variable aléatoire X.

La transformée stop-loss peut s'exprimer

$$\pi_X(t) = \int_{x=t}^{+\infty} \overline{F}_X(x)dx.$$

Supposons $\mathbb{E}[X] < +\infty$. La transformée stop-loss π_X possède les propriétés suivantes:

(i) elle est décroissante et convexe.

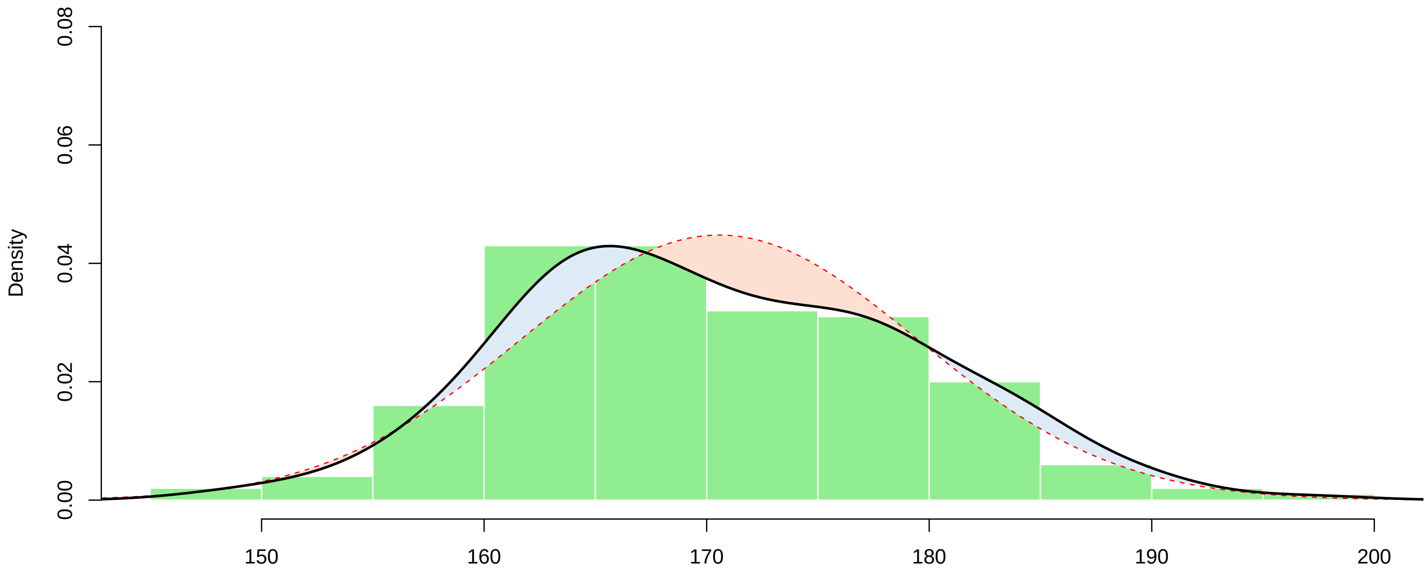
(ii) $\lim\limits_{t \to +\infty} \pi_X(t) = 0$ et $\lim\limits_{t \to -\infty} \{\pi_X(t) + t\} = \mathbb{E}[X]$.

Hétérogénéité

Hétérogénéité

Considérons la taille d'un groupe d'assurés,

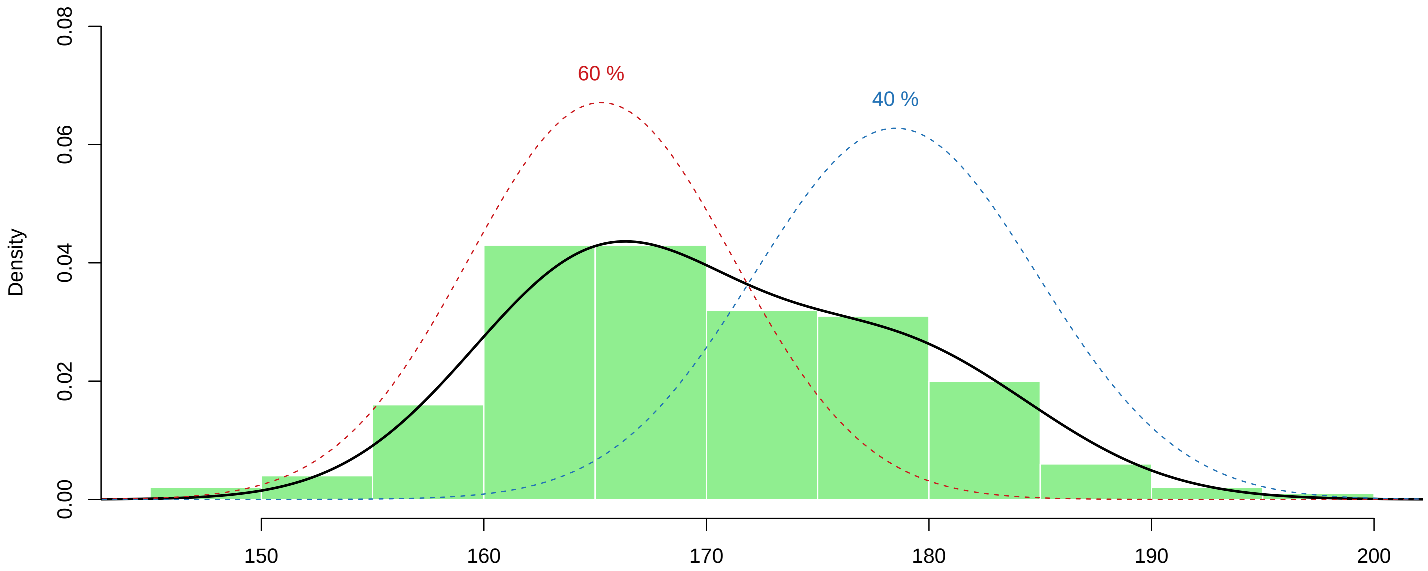
```
> Davis=read.table("http://socserv.socsci.mcmaster.ca/jfox/Books/
    Applied-Regression-2E/datasets/Davis.txt")
> Davis[12,c(2,3)] <- Davis[12,c(3,2)]
> Y <- Davis$height
```



Hétérogénéité, effet aléatoire

Comme $Y \not\sim \mathcal{N}(\mu, \sigma^2)$, on peut considérer un mélange pour Y

$$Y \sim \begin{cases} \mathcal{N}(\mu_1, \sigma_1^2) \text{ avec probabilié } p_1, \text{ i.e. variable latente } \Theta = 1 \\ \mathcal{N}(\mu_2, \sigma_2^2) \text{ avec probabilité } p_2, \text{ i.e. variable latente } \Theta = 2 \end{cases}$$



Hétérogénéité, effet aléatoire

On fait alors du maximum de vraisemblance 'à la main',

$$f(y) = pf_{\theta_1}(x) + [1-p]f_{\theta_2}(x)$$

```
> logdf <- function(x,parameter){
+ p  <- parameter[1]
+ m1 <- parameter[2]
+ s1 <- parameter[4]
+ m2 <- parameter[3]
+ s2 <- parameter[5]
+ return(log(p*dnorm(x,m1,s1)+(1-p)*dnorm(x,m2,s2))) }
> logL <- function(parameter) -sum(logdf(X,parameter))
> Amat <- matrix(c(1,-1,0,0,0,0,
+ 0,0,0,0,1,0,0,0,0,0,0,0,0,1), 4, 5)
> bvec <- c(0,-1,0,0)
> constrOptim(c(.5,160,180,10,10), logL, NULL, ui = Amat, ci = bvec)$
   par
[1]   0.5996263 165.2690084 178.4991624   5.9447675   6.3564746
```

Hétérogénéité, effet aléatoire

ou, plus efficace, on utilise un algorithme EM,

```
> library(mixtools)
> mix <- normalmixEM(Y)
 number of iterations= 335
> (param12 <- c(mix$lambda[1],mix$mu,mix$sigma))
[1] 0.4002202 178.4997298 165.2703616 6.3561363 5.9460023
```

(sans spécifier le nombre de classes... ni contraintes, e.g. $\sigma_1^2 = \sigma_2^2$)

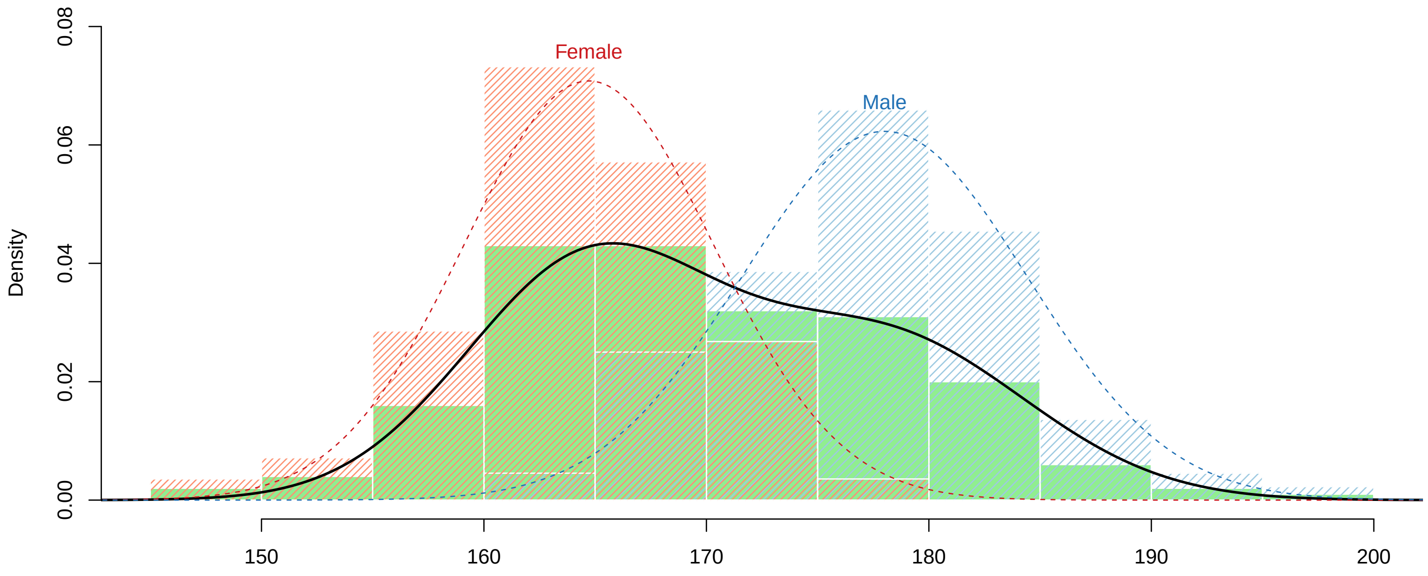
Remarque

```
> sd(Y)
[1] 8.932228
```

Hétérogénéité, effet fixe

On peut utiliser une covariable pour modéliser $Y|X = x$

$$Y \sim \begin{cases} \mathcal{N}(\mu_M, \sigma_M^2) \text{ si } X = \text{`male'}, \text{ avec } p_M \\ \mathcal{N}(\mu_F, \sigma_F^2) \text{ si } X = \text{`female'}, \text{ avec } p_F \end{cases}$$



Hétérogénéité, effet fixe

```
> (pM <- mean(sex=="M"))
[1] 0.44
> (paramF <- fitdistr(X[sex=="F"],"normal")$estimate)
      mean          sd
164.714286   5.633808
> (paramM <- fitdistr(X[sex=="M"],"normal")$estimate)
      mean          sd
178.011364   6.404001
```

ou si on veut une variance identique

```
> s=sqrt((sum((height[sex=="M"]-paramM[1])^2)+sum((height[sex=="F"]-
    paramF[1])^2))/(nrow(Davis)-2))
> s
[1] 6.015068
```

Hétérogénéité d'un portefeuille en information parfaite

Considérons une assurance voyage, avec les indemnitées

- pays A, 250 € avec probabilité 10%
- pays B, 250 € avec probabilité 20%

La dépense sera, conditionnelle à la destination

$$S_A = \begin{cases} 0, & \text{avec la probabilité } 0.9, \\ 250€, & \text{avec la probabilité } 0.1 \end{cases}$$

et

$$S_B = \begin{cases} 0, & \text{avec la probabilité } 0.8, \\ 250€, & \text{avec la probabilité } 0.2. \end{cases}$$

Hétérogénéité d'un portefeuille en information parfaite

Si la destination n'est pas observée

$$S_{AB} = \begin{cases} 0, & \text{avec la probabilité } 0.85, \\ 250€, & \text{avec la probabilité } 0.15. \end{cases}$$

La prime pure

- sans segmentation sera $\mathbb{E}(S_{AB}) = 37.5€$
- avec segmentation sera $\mathbb{E}(S_A) = 25€$ à destination de A et $\mathbb{E}(S_B) = 50€$ à destination de B

S'il y a deux compagnies qui se font concurence, une qui segmente, l'autre pas...

Hétérogénéité et mélange de lois

Supposons qu'il existe un effet aléatoire Θ représentant le niveau de risque (inconnu) d'un assuré pris au hasard dans le portefeuille. On suppose que X sachant $\Theta = \theta$ admet pour loi F_θ. Alors

$$\mathbb{P}(X \leq x) = \mathbb{E}[\mathbb{P}(X \leq x|\Theta)] = \int_\Omega F_\theta(x) d\Pi(\theta)$$

(moyenne des F_θ pondérée par $d\Pi$).

Mélange de lois de Poisson et sur-dispersion

La variable aléatoire de comptage N est de loi de Poisson mélange de moyenne λ et de niveau de risque relatif Θ lorsque

$$\begin{aligned}\mathbb{P}[N=k] &= \mathbb{E}\left[\exp(-\lambda\Theta)\frac{(\lambda\Theta)^k}{k!}\right]\\ &= \int_0^{+\infty}\exp(-\lambda\theta)\frac{(\lambda\theta)^k}{k!}dF_\Theta(\theta),\quad k\in\mathbb{N},\end{aligned}\tag{1}$$

où F_Θ est la fonction de répartition de Θ, telle que $\mathbb{E}[\Theta]=1$. On notera $\mathcal{MP}oi(\lambda,F_\Theta)$ ou $\mathcal{MP}oi(\lambda,\Theta)$.

Mélange de lois de Poisson et sur-dispersion

Considérons l'exemple classique des bons et mauvais conducteur,

- pour les bons, $N \sim \mathcal{P}oi(\lambda\theta_1)$
- pour les mauvais, $N \sim \mathcal{P}oi(\lambda\theta_2)$

avec $\theta_2 > 1 > \theta_1$.

Si la proportion de bons risques est ϱ,

$$\Theta = \begin{cases} \theta_1, & \text{avec une probabilité } \varrho, \\ \theta_2, & \text{avec une probabilité } 1-\varrho, \end{cases}$$

où θ_1, θ_2 et ϱ sont contraints par

$$\mathbb{E}[\Theta] = \varrho\theta_1 + (1-\varrho)\theta_2 = 1.$$

Mélange de lois de Poisson et sur-dispersion

La probabilité qu'une police (dont on ne sait pas s'il s'agit d'un bon ou d'un mauvais risque) donne lieu à k sinistres durant la période de référence est alors de

$$\mathbb{P}[N=k]=\varrho\exp(-\lambda\theta_1)\frac{(\lambda\theta_1)^k}{k!}+(1-\varrho)\exp(-\lambda\theta_2)\frac{(\lambda\theta_2)^k}{k!},$$

Si Θ devient une variable aléatoire continue, de densité de probabilité f_Θ alors

$$\mathbb{P}[N=k]=\int_0^{+\infty}\exp(-\lambda\theta)\frac{(\lambda\theta)^k}{k!}f_\Theta(\theta)d\theta,\quad k\in\mathbb{N}. \tag{2}$$

Mélange de lois de Poisson et sur-dispersion

Soit $N \sim \mathcal{MP}oi(\lambda, \Theta)$, alors

$$\begin{aligned}\mathbb{E}[N] &= \int_0^{+\infty} \left(\sum_{k=0}^{+\infty} k \exp(-\lambda\theta) \frac{(\lambda\theta)^k}{k!} \right) dF_\Theta(\theta) \\ &= \lambda \int_0^{+\infty} \theta dF_\Theta(\theta) = \lambda.\end{aligned}$$

et

$$\begin{aligned}\text{Var}[N] &= \int_0^{+\infty} \left(\sum_{k=0}^{+\infty} k^2 \exp(-\lambda\theta) \frac{(\lambda\theta)^k}{k!} \right) dF_\Theta(\theta) - \lambda^2 \\ &= \int_0^{+\infty} (\lambda\theta + \lambda^2\theta^2) dF_\Theta(\theta) - \lambda^2 = \lambda + \lambda^2 \text{Var}[\Theta].\end{aligned}$$

Comme

$$\text{Var}[N] = \mathbb{E}[N] + \lambda^2 \text{Var}[\Theta] > \mathbb{E}[N]$$

pour autant que Θ ne soit pas constant.

Mélange de lois de Poisson et sur-dispersion

Tout mélange de Poisson implique donc une surdispersion des données. La fonction génératrice des probabilités de $N \sim \mathcal{MP}oi(\lambda, \Theta)$ et la tranformée de Laplace de Θ sont liées par la formule

$$M_N(z) = \int_0^{+\infty} \exp(\lambda\theta(z-1))f_\Theta(\theta)d\theta = L_\Theta(\lambda(1-z)). \quad (3)$$

Si nous considérons que $\Theta \sim \mathcal{G}(\alpha, \alpha)$,

$$M_N(z) = \left(1 + \frac{\lambda(1-z)}{\alpha}\right)^{-\alpha},$$

qui est la fonction génératrice des probabilités associée à la loi $BN(\alpha, \alpha/(\alpha+\lambda))$.

Les mélanges de Poisson sont identifiables, i.e. si $N_1 \sim \mathcal{MP}oi(\lambda, \Theta_1)$ et $N_2 \sim \mathcal{MP}oi(\lambda, \Theta_2)$ alors

$$N_1 \stackrel{\mathcal{L}}{=} N_2 \Rightarrow \Theta_1 \stackrel{\mathcal{L}}{=} \Theta_2.$$

Mélange de lois de Poisson et sur-dispersion

Les mélanges de Poisson jouissent d'une propriété trs importante établie par Shaked (1980) et connue comme le “Shaked's Two Crossings Theorem".

Si $N \sim \mathcal{MP}oi(\lambda, \Theta)$ alors il existe deux valeurs entières $0 \leq k_0 < k_1$ telle que

$$\begin{aligned}
\mathbb{P}[N=k] &\geq \exp(-\lambda)\frac{\lambda^k}{k!} \text{ pour } k=0,1,\ldots,k_0,\\
\mathbb{P}[N=k] &\leq \exp(-\lambda)\frac{\lambda^k}{k!} \text{ pour } k=k_0+1,\ldots,k_1,\\
\mathbb{P}[N=k] &\geq \exp(-\lambda)\frac{\lambda^k}{k!} \text{ pour } k\geq k_1+1.
\end{aligned}$$

Mélange de lois de Poisson et sur-dispersion

```
> library(stats4); library(MASS)
> FREQUENCE=c(7840,1317,239,42,14,4,4,1)
> x=rep(0:7,FREQUENCE)
> table(x)
x
   0    1    2    3    4    5    6    7
7840 1317  239   42   14    4    4    1
> mean(x)
[1] 0.2143537
> var(x)
[1] 0.2889314
```

Mélange de lois de Poisson et sur-dispersion

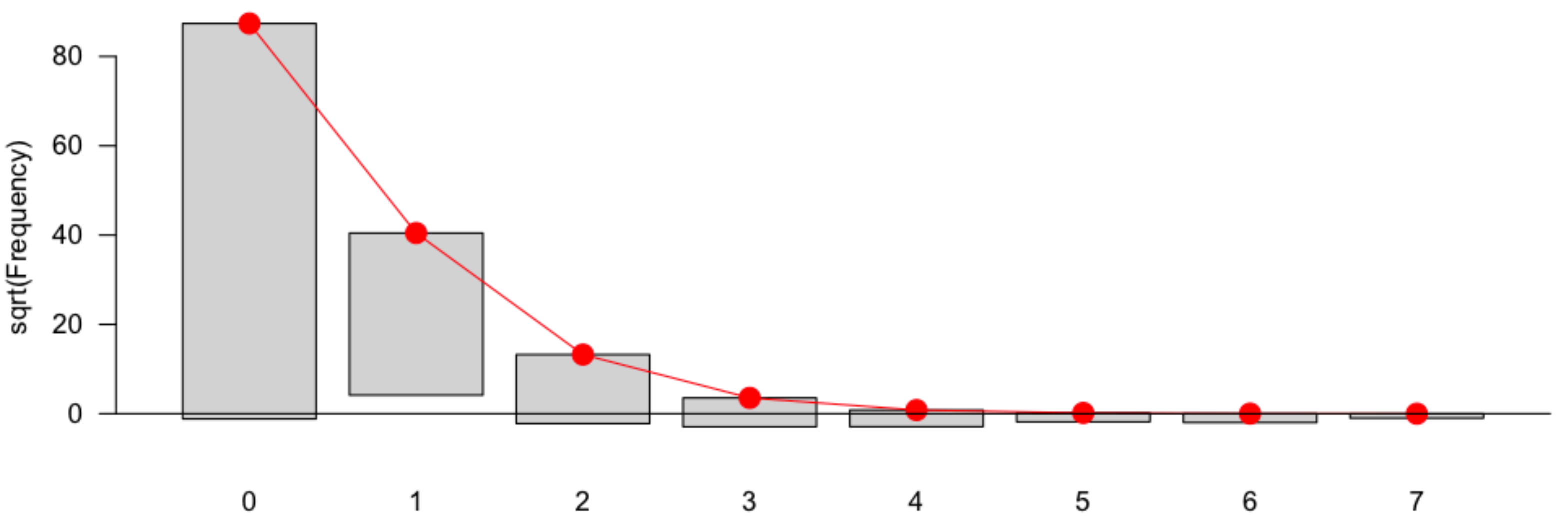
```
> fitdistr(x,"poisson")
     lambda
  0.21435366
 (0.00475989)
> fitdistr(x,"negative binomial")
      size            mu
  0.701486138   0.214355746
 (0.062786579) (0.005438638)

> library(vcd)
> model.poisson=goodfit(x,type="poisson",method="ML")
```

Mélange de lois de Poisson et sur-dispersion

```
> model.poisson

Observed and fitted values for poisson distribution
with parameters estimated by `ML'

 count observed      fitted
     0     7840    7635.622
     1     1317    1636.724
     2      239    175.4188
     3       42    12.53389
     4       14    0.671671
     5        4    0.028795
     6        4    0.001028
     7        1    0.000031
```

Mélange de lois de Poisson et sur-dispersion

```
> summary(model.poisson)

   Goodness-of-fit test for poisson distribution

                     X^2 df     P(> X^2)
Likelihood Ratio 302.484  6 2.401523e-62
> plot(model.poisson)
```

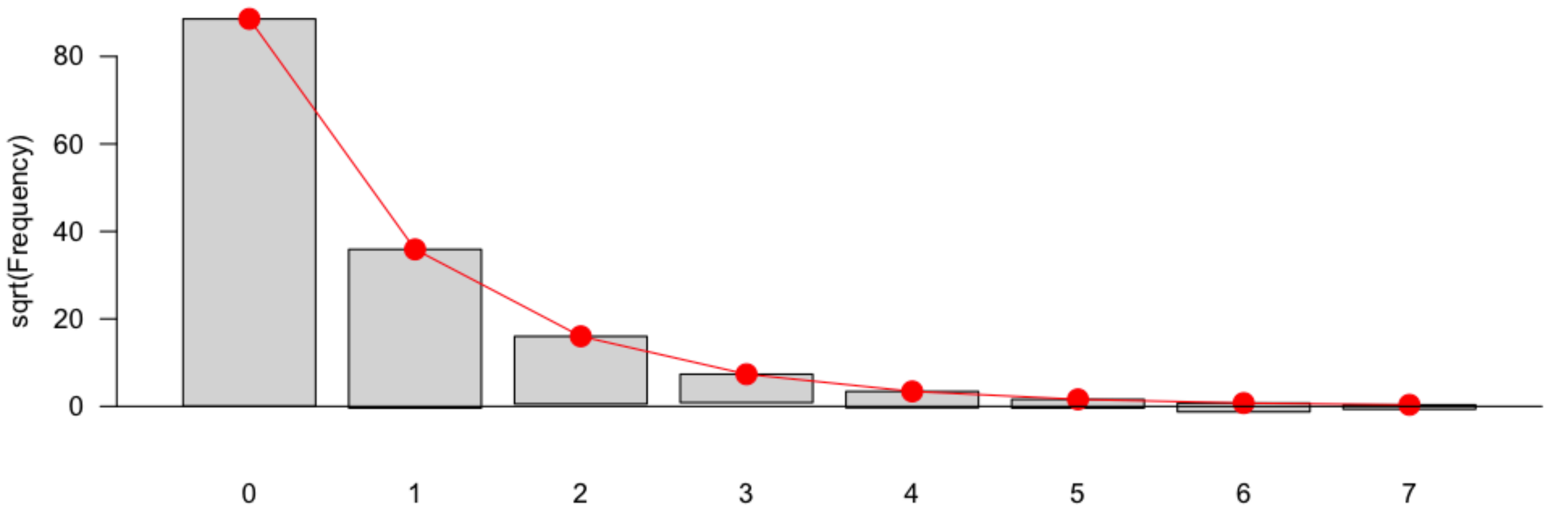


Mélange de lois de Poisson et sur-dispersion

```
> model.nb=goodfit(x,type="nbinomial",method="ML")
> model.nb

Observed and fitted values for nbinomial distribution
with parameters estimated by 'ML'

 count observed       fitted
     0     7840 7847.0055590
     1     1317 1288.3613321
     2      239  256.5374241
     3       42   54.0688144
     4       14   11.7105586
     5        4    2.5772550
     6        4    0.5732035
     7        1    0.1284389
```

Mélange de lois de Poisson et sur-dispersion

```
> summary(model.nb)

   Goodness-of-fit test for nbinomial distribution

                     X^2 df    P(> X^2)
Likelihood Ratio 17.00285  5 0.00449439
> plot(model.nb)
```



Segmentation en assurance

On qualifie de segmentation toute technique que l'assureur utilise pour différencier la prime, et éventuellement aussi la couverture, en fonction d'un certain nombre de caractéristiques spécifiques du risque à assurer, et ce afin de parvenir à une meilleure concordance entre les coûts qu'une personne déterminée met à charge de la collectivité des preneurs d'assurance et la prime que cette personne doit payer pour la couverture offerte. Dans certains cas, cela peut impliquer que l'assureur refuse le risque à assurer.

Segmentation et espérance conditionnelle

Si $N \sim \mathcal{MP}oi(\lambda, \Theta)$, la fonction de répartition de Θ sachant que $N = n$, notée $F_\Theta(\cdot|n)$,

$$
\begin{aligned}
F_\Theta(t|n) &= \frac{\mathbb{P}[\Theta \leq t, N = n]}{\mathbb{P}[N = n]} \\
&= \frac{\int_0^t \exp(-\lambda\theta)\frac{(\lambda\theta)^n}{n!} dF_\Theta(\theta)}{\int_{\mathbb{R}+} \exp(-\lambda\theta)\frac{(\lambda\theta)^n}{n!} dF_\Theta(\theta)} \\
&= \frac{\int_0^t \exp(-\lambda\theta)\theta^n dF_\Theta(\theta)}{\int_{\mathbb{R}+} \exp(-\lambda\theta)\theta^n dF_\Theta(\theta)}.
\end{aligned}
$$

On peut alors en calculer l'espérance, $\mathbb{E}[\Theta|N = n] = \displaystyle\int_{\mathbb{R}+} \theta dF_\Theta(\theta|n)$, qui donne

$$
\mathbb{E}[\Theta|N = n] = \frac{\int_0^t \exp(-\lambda\theta)\theta^{n+1} dF_\Theta(\theta)}{\int_{\theta\in\mathbb{R}+} \exp(-\lambda\theta)\theta^n dF_\Theta(\theta)}.
$$

Segmentation et espérance conditionnelle

On vérifie aisément que l'espérance conditionnelle possède les propriétés suivantes: quelles que soient les variables aléatoires X_1, X_2 et X_3, et la constante réelle c,

(i) $\mathbb{E}[c|X_1 = x_1] = c$ quel que soit $x_1 \in \mathbb{R}$.

(ii) $\mathbb{E}[X_1 + X_2|X_3 = x_3] = \mathbb{E}[X_1|X_3 = x_3] + \mathbb{E}[X_2|X_3 = x_3]$ quel que soit $x_3 \in \mathbb{R}$.

(iii) $\mathbb{E}[cX_1|X_2 = x_2] = c\mathbb{E}[X_1|X_2 = x_2]$, quel que soit $x_2 \in \mathbb{R}$.

(iv) quelle que soit la fonction g, $\mathbb{E}[g(X_1, X_2)|X_2 = x_2] = \mathbb{E}[g(X_1, x_2)|X_2 = x_2]$ quel que soit $x_2 \in \mathbb{R}$.

(v) si X_1 et X_2 sont indépendantes alors $\mathbb{E}[X_1|X_2 = x_2] = \mathbb{E}[X_1]$.

Segmentation et espérance conditionnelle

A moins que X_1 et X_2 ne soient indépendantes, la loi de X_1 sachant $X_2 = x_2$ dépend de x_2. En particulier, $\mathbb{E}[X_1|X_2 = x_2]$ est une fonction de x_2, i.e.

$$\mathbb{E}[X_1|X_2 = x_2] = h(x_2).$$

On pourrait ainsi s'intéresser à la variable aléatoire

$$h(X_2) = \mathbb{E}[X_1|X_2].$$

La variable aléatoire $\mathbb{E}[X_1|X_2]$ a même moyenne que X_1:

$$\mathbb{E}\Big[\mathbb{E}[X_1|X_2]\Big] = \mathbb{E}[X_1].$$

Segmentation et espérance conditionnelle

Lorsque X_1 et X_2 sont continues, cette égalité s'établit comme suit:

$$\begin{aligned}
\mathbb{E}\Big[\mathbb{E}[X_1|X_2]\Big] &= \int_{x_2\in\mathbb{R}} \mathbb{E}[X_1|X_2=x_2]f_2(x_2)dx_2 \\
&= \int_{x_2\in\mathbb{R}} \left\{\int_{x_1\in\mathbb{R}} x_1 f_{1|2}(x_1|x_2)dx_1\right\} f_2(x_2)dx_2 \\
&= \int_{x_2\in\mathbb{R}}\int_{x_1\in\mathbb{R}} x_1 f_{\boldsymbol{X}}(x_1,x_2)dx_1dx_2 \\
&= \int_{x_1\in\mathbb{R}} x_1 f_1(x_1)dx_1 = \mathbb{E}[X_1].
\end{aligned}$$

Le raisonnement est similaire dans les cas discret et mixte.

Segmentation et espérance conditionnelle

Epinglons à présent une caractéristique remarquable de l'espérance conditionnelle, qui peut être prise comme définition générale de ce concept.

Quelle que soit la fonction $h : \mathbb{R} \rightarrow \mathbb{R}$, nous avons

$$\mathbb{E}\Big[h(X_2)\,\{X_1 - \mathbb{E}[X_1|X_2]\}\Big] = 0.$$

La Propriété précédante nous permet d'écrire

$$\begin{aligned}
&\mathbb{E}\Big[h(X_2)\,\{X_1 - \mathbb{E}[X_1|X_2]\}\Big] \\
= \;& \mathbb{E}\Big[\mathbb{E}\Big[h(X_2)\,\{X_1 - \mathbb{E}[X_1|X_2]\}\Big|X_2\Big]\Big] \\
= \;& \mathbb{E}\Big[h(X_2)\mathbb{E}\Big[\{X_1 - \mathbb{E}[X_1|X_2]\}\Big|X_2\Big]\Big] = 0,
\end{aligned}$$

ce qui achève la vérification du résultat annoncé.

On peut voir $X_1 - \mathbb{E}[X_1|X_2]$ comme un résidu (c'est-à-dire comme la partie de X_1 que ne parvient pas à expliquer X_2).

Ceci garantit que $\mathbb{E}[X_1|X_2]$ est le meilleur prédicteur de X_1 au sens des moindres carrés, comme l'indique le résultat suivant.

La variable aléatoire $h^*(X_2) = \mathbb{E}[X_1|X_2]$ est celle minimisant $\mathbb{E}[(X_1 - h(X_2))^2]$ sur toutes les fonctions h.

Ecrivons

$$
\begin{aligned}
\mathbb{E}[(X_1 - h(X_2))^2] &= \mathbb{E}[(X_1 - \mathbb{E}[X_1|X_2] + \mathbb{E}[X_1|X_2] - h(X_2))^2] \\
&= \underbrace{\mathbb{E}[(X_1 - \mathbb{E}[X_1|X_2])^2]}_{\text{indépendant de } h} + \mathbb{E}[(\mathbb{E}[X_1|X_2] - h(X_2)^2] \\
&\quad +2\underbrace{\mathbb{E}[(X_1 - \mathbb{E}[X_1|X_2])(\mathbb{E}[X_1|X_2] - h(X_2))]}_{=0 \text{ par définition de l'espérance conditionnelle}}
\end{aligned}
$$

sera minimum lorsque $h(X_2) = \mathbb{E}[X_1|X_2]$.

Segmentation et espétance conditionnelle

Il n'y a bien entendu aucune raison de se limiter à une seule variable conditionnante, et on peut considérer un vecteur $\boldsymbol{X}$ de dimension n et définir $\mathbb{E}[X_1|X_2,\ldots,X_n]$.

Considérons un vecteur aléatoire $\boldsymbol{X}$ de dimension n. L'espérance conditionelle $\mathbb{E}[X_1|X_2,\ldots,X_n]$ de X_1 sachant $X_2,\ldots,X_n$ est la variable aléatoire $h^*(X_2,\ldots,X_n)$ telle que l'égalité

$$\mathbb{E}\Big[h(X_2,\ldots,X_n)\left\{X_1-h^*(X_2,\ldots,X_n)\right\}\Big]=0, \tag{4}$$

est vérifiée pour toute fonction $h:\mathbb{R}^{n-1}\rightarrow\mathbb{R}$.

Transfert de risque sans segmentation

Tentons à présent de formaliser le problème de la segmentation. Considérons un assuré soumis à un risque S, prélevé au hasard au sein d'un portefeuille d'assurance. Supposons que toutes les caractéristiques de l'assuré influençant le risque soient reprises dans un vecteur aléatoire $\mathbf{\Omega} = (\Omega_1, \Omega_2, \ldots)$. La notation “oméga" rappelle que le vecteur du mÍme nom contient toute l'information à propos de l'assuré, qu'elle soit observable ou non par l'assureur.

On peut imaginer que l'assureur ne tienne compte en aucune manière des caractéristiques $\mathbf{\Omega}$ de l'assuré, et lui réclame donc une prime pure de montant $\mathbb{E}[S]$, la mÍme que celle qu'il réclame à tous les assurés du portefeuille. Dans ce cas, la situation est telle que présentée au Tableau suivant

Transfert de risque sans segmentation

	Assurés	Assureur
Dépense	$\mathbb{E}[S]$	$S - \mathbb{E}[S]$
Dépense moyenne	$\mathbb{E}[S]$	0
Variance	0	$\text{Var}[S]$

Table 1: Situation des assurés et de l'assureur en l'absence de segmentation.

L'assureur prend donc l'entièreté de la variance des sinistres $\text{Var}[S]$ à sa charge, que celle-ci soit due à l'hétérogénéité du portefeuille, ou à la variabilité intrinsèque des montants des sinistres.

Segmentation avec information parfaite

A l'autre extrême, supposons que l'assureur incorpore toute l'information $\boldsymbol{\Omega}$ dans la tarification. On serait alors dans la situation décrite au Tableau suivant

	Assurés	Assureur
Dépense	$\mathbb{E}[S\|\boldsymbol{\Omega}]$	$S-\mathbb{E}[S\|\boldsymbol{\Omega}]$
Dépense moyenne	$\mathbb{E}[S]$	0
Variance	$\text{Var}\Big[\mathbb{E}[S\|\boldsymbol{\Omega}]\Big]$	$\text{Var}\Big[S-\mathbb{E}[S\|\boldsymbol{\Omega}]\Big]$

Table 2: Situation des assurés et de l'assureur dans le cas où la segmentation est opérée sur base de $\boldsymbol{\Omega}$.

Contrairement au cas précédent, la prime payée par un assuré prélevé au hasard dans le portefeuille est à présent une variable aléatoire: $\mathbb{E}[S|\boldsymbol{\Omega}]$ dépend des caractéristiques $\boldsymbol{\Omega}$ de cet assuré.

Segmentation avec information parfaite

Comme la variable aléatoire $S - \mathbb{E}[S|\boldsymbol{\Omega}]$ est centrée, le risque assumé par l'assureur la variance du résultat financier de l'opération d'assurance, i.e.

$$\begin{aligned}
\text{Var}\Big[S - \mathbb{E}[S|\boldsymbol{\Omega}]\Big] &= \mathbb{E}\Big[\Big(S - \mathbb{E}[S|\boldsymbol{\Omega}]\Big)^2\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\Big[\Big(S - \mathbb{E}[S|\boldsymbol{\Omega}]\Big)^2\Big|\boldsymbol{\Omega}\Big]\Big] \\
&= \mathbb{E}\Big[\text{Var}[S|\boldsymbol{\Omega}]\Big].
\end{aligned}$$

On assiste dans ce cas à un partage de la variance totale de S (c'est-à-dire du risque) entre les assurés et l'assureur, matérialisé par la formule

$$\text{Var}[S] = \underbrace{\mathbb{E}\Big[\text{Var}[S|\boldsymbol{\Omega}]\Big]}_{\to\text{assureur}} + \underbrace{\text{Var}\Big[\mathbb{E}[S|\boldsymbol{\Omega}]\Big]}_{\to\text{assurés}}.$$

Segmentation avec information parfaite

Ainsi, lorsque toutes les variables pertinentes $\boldsymbol{\Omega}$ ont été prises en compte, l'intervention de l'assureur se limite à la part des sinistres due exclusivement au hasard; en effet, $\text{Var}[S|\boldsymbol{\Omega}]$ représente les fluctuations de S dues au seul hasard. Dans cette situation idéale, l'assureur mutualise le risque et il n'y a donc aucune solidarité induite entre les assurés du portefeuille: chacun paie en fonction de son propre risque.

Segmentation avec information imparfaite

Bien entendu, la situation décrite au paragraphe précédent est purement théorique puisque parmi les variables explicatives $\boldsymbol{\Omega}$ nombreuses sont celles qui ne peuvent pas être observées par l'assureur. En assurance automobile par exemple, l'assureur ne peut pas observer la vitesse à laquelle roule l'assuré, son agressivité au volant, ni le nombre de kilomètres qu'il parcourt chaque année.

Dès lors, l'assureur ne peut utiliser qu'un sous-ensemble $\boldsymbol{X}$ des variables explicatives contenues dans $\boldsymbol{\Omega}$, i.e. $\boldsymbol{X} \subset \boldsymbol{\Omega}$. La situation est alors semblable à celle décrite au Tableau suivant

Segmentation avec information imparfaite

	Assuré	Assureur
Dépense	$\mathbb{E}[S\vert\boldsymbol{X}]$	$S-\mathbb{E}[S\vert\boldsymbol{X}]$
Dépense moyenne	$\mathbb{E}[S]$	0
Variance	$\text{Var}\Big[\mathbb{E}[S\vert\boldsymbol{X}]\Big]$	$\mathbb{E}\Big[\text{Var}[S\vert\boldsymbol{X}]\Big]$

Table 3: Situation de l'assuré et de l'assureur dans le cas où la segmentation est opérée sur base de $\boldsymbol{X}\subset\boldsymbol{\Omega}$.

Il est intéressant de constater que

$$\begin{aligned}\mathbb{E}\Big[\text{Var}[S|\boldsymbol{X}]\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[\text{Var}[S|\boldsymbol{\Omega}]\Big|\boldsymbol{X}\Big]\Big]+\mathbb{E}\Big[\text{Var}\Big[\mathbb{E}[S|\boldsymbol{\Omega}]\Big|\boldsymbol{X}\Big]\Big] \\ &= \underbrace{\mathbb{E}\Big[\text{Var}[S|\boldsymbol{\Omega}]\Big]}_{\text{mutualisation}}+\underbrace{\mathbb{E}\Big\{\text{Var}\Big[\mathbb{E}[S|\boldsymbol{\Omega}]\Big|\boldsymbol{X}\Big]\Big\}}_{\text{solidarité}}. \qquad (5)\end{aligned}$$

Segmentation avec information imparfaite

On peut interpréter cette décomposition du risque pris en charge par la compagnie comme suit: l'assureur, lorsque tous les facteurs de risque ne sont pas pris en compte, intervient pour réparer les conséquences fâcheuses du hasard (premier terme traduisant la mutualisation du risque), mais doit aussi prendre en charge les variations de la prime pure exacte $\mathbb{E}[S|\boldsymbol{\Omega}]$ qui ne sont pas expliquées par les facteurs de risque $\boldsymbol{X}$ intégrés au tarif (second terme traduisant la solidarité induite par une personnalisation imparfaite du montant de la prime). En d'autres mots, en plus de contrer les mauvais coups du sort, l'assureur doit également supporter la variabilité des sinistres due aux caractéristiques des assurés, non prises en compte par le tarif.

Segmentation avec information imparfaite

Dans une tarification segmentée en fonction de $\boldsymbol{X} \subset \boldsymbol{\Omega}$, le partage de la variance de S s'effectue comme suit:

$$\begin{aligned}
\text{Var}[S] &= \mathbb{E}\Big[\text{Var}[S|\boldsymbol{X}]\Big] + \text{Var}\Big[\mathbb{E}[S|\boldsymbol{X}]\Big] \\
&= \underbrace{\underbrace{\mathbb{E}\Big[\text{Var}[S|\boldsymbol{\Omega}]\Big]}_{\text{mutualisation}} + \underbrace{\mathbb{E}\Big[\text{Var}\Big[\mathbb{E}[S|\boldsymbol{\Omega}]\Big|\boldsymbol{X}\Big]\Big]}_{\text{solidarité}}}_{\rightarrow\text{assureur}} \\
&\quad + \underbrace{\text{Var}\Big[\mathbb{E}[S|\boldsymbol{X}]\Big]}_{\rightarrow\text{assurés}}.
\end{aligned}$$

Tarification *a priori* et *a posteriori*

Toute l'idée qui sous-tend la tarification d'expérience est que l'historique des sinistres révèle les caractéristiques non observables des assurés. Plus précisément, si on note $\boldsymbol{S}$ les informations concernant la sinistralité passée des assurés disponibles pour l'assureur, l'information contenue dans $(\boldsymbol{X}, \boldsymbol{S})$ devient comparable à $\boldsymbol{\Omega}$, tant et si bien que $\mathbb{E}[S|\boldsymbol{X}, \boldsymbol{S}]$ devrait converger vers $\mathbb{E}[S|\boldsymbol{\Omega}]$ (dans un sens à préciser).

Lectures sur les modèles prédictifs en tarification

Wendler & Modlin Basic Ratemaking. pdf

Barnes Predictive Modeling—You Mean Actuarial Wizardry?. html

Guven Predictive Modeling. html

Myers Beyond GLMs. html

Holler, Sommers & Trahair Something Old, Something New in Classification Ratemaking With a Novel Use of GLMs for Credit Insurance. pdf

Newton Multi-Year Policy Pricing. pdf

Werner & Guven GLM Basic Modeling: Avoiding Common Pitfalls . pdf

Mildenhal A Systematic Relationship Between Minimum Bias and Generalized Linear Models. pdf

FFSA La segmentation : un fondement essentiel de l'assurance. html

Galloy Jusqu'où faut-il segmenter?. html

Charpentier, Denuit & Elie Segmentation et Mutualisation, les deux faces d'une même pièce. pdf

Charpentier La loi des petits nombres. pdf