Chapitre 3 : Le modèle collectif

I - Définition

Modèle collectif : Montant agrégé des indemnisations à la maille portefeuille ou groupe homogène des contrats...

- Le nombre de sinistres sur la période d'exercice est une v.a de comptage N de loi

$$\mathbb{P}_N(n) = \mathbb{P}(N = n) \qquad , n = 0, 1, ..$$

- Les montants de sinistres $U_1, \dots, U_N$ forment une suite iid de v.a positives de densité f_U et indépendantes de N

La charge totale est donnée par :

$$X = \sum_{i=1}^{N} U_i$$

↳ La fonction de répartition de X n'est pas connue explicitement

1/ Distribution de X

La mesure de probabilité de X est mixte au sens où elle admet un atome de proba en 0 correspondant à l'événement $\{N=0\}$

On a

$$dP_X(x) = p_N(0)\,\delta_0(x) + f_X^+(x)\,d\lambda(x)$$

où f_X^+ est la densité de la partie absolument continue de la distribution de X donnée par

$$f_X^+(x) = \sum_{n=1}^{+\infty} p_N(n)\, f_U^{*k}(x) \quad , \quad x > 0$$

avec f_U^{*k} est le $k^{\text{ième}}$ produit de convolution de f_U avec elle même. Il s'agit de la densité de $U_1 + \dots + U_k$. f_X^+ est défaillante au sens où

$$\int_0^{\infty} f_X^+(x)\,dx = 1 - p_N(0)$$

Exemple 1 : (Modèle géométrique - exponentiel)

Supposons $U \sim \exp(\beta)$ de densité $f_U(x) = \frac{e^{-x/\beta}}{\beta}$, $x \geq 0$

et $N \sim Geom(p)$ de loi $\mathbb{P}_N(k) = p^k(1-p)$, $k = 0, 1, \dots$

$$S_X^+(x) = \sum_{k=1}^{\infty} p^k(1-p) \frac{e^{-x/\beta} x^{k-1}}{\beta^k (k-1)!}$$

$$= (1-p) e^{-x/\beta} \frac{p}{\beta} \sum_{k=1}^{\infty} \left(\frac{px}{\beta}\right)^{k-1} \frac{1}{(k-1)!}$$

$$= (1-p) e^{-x/\beta} \frac{p}{\beta} e^{\frac{px}{\beta}} = \frac{(1-p)p}{\beta} e^{-x\left(\frac{1-p}{\beta}\right)}$$

Pour tous les autres modèles $\Rightarrow$ Méthodes numériques pour calculer la Value at Risk

$$VaR_X(\alpha) = \inf\{x \geq 0;\ F_X \geq \alpha\}$$

ou la tail value at Risk

$$TVaR_X(\alpha) = \mathbb{E}[X \mid X > VaR_X(\alpha)]$$

ou la prime stop loss en réassurance

$$\Pi_c = \mathbb{E}[(X-c)_+] \quad \text{ou} \quad \Pi_{c,d} = \mathbb{E}[\min((X-c)_+, d)]$$

Proposition 1 :

1. L'espérance de X est donnée par
$$\mathbb{E}[X] = \mathbb{E}[N]\,\mathbb{E}[U]$$

2. La variance est donnée par
$$\mathbb{V}[X] = \mathbb{E}[N]\,\mathbb{V}[U] + \mathbb{E}[U]^2\,\mathbb{V}[N]$$

3. La fonction génératrice des moments est donnée par
$$M_X(s) = \mathbb{E}[e^{sX}] = G_N[M_U(s)] \quad , \quad s \in \mathbb{R}$$

où $G_N(s) = \mathbb{E}[s^N]$, $s \in \mathbb{R}$

Preuve :

(1)
$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid N]] = \mathbb{E}\Big[\mathbb{E}\Big[\sum_{i=1}^{N} U_i \mid N\Big]\Big]$$
$$= \mathbb{E}\Big[\sum_{i=1}^{N} \mathbb{E}[U_i \mid N]\Big]$$
$$= \mathbb{E}[N\,\mathbb{E}[U]] = \mathbb{E}[N]\,\mathbb{E}[U]$$

(2)
$$\mathbb{V}[X] = \mathbb{E}[\mathbb{V}[X \mid N]] + \mathbb{V}[\mathbb{E}[X \mid N]]$$
$$= \mathbb{E}\Big[\mathbb{V}\Big[\sum_{i=1}^{N} U_i \mid N\Big]\Big] + \mathbb{V}\Big[\mathbb{E}\Big[\sum_{i=1}^{N} U_i \mid N\Big]\Big]$$
$$= \mathbb{E}\Big[\sum_{i=1}^{N} \mathbb{V}(U_i \mid N)\Big] + \mathbb{V}\Big[\sum_{i=1}^{N} \mathbb{E}[U_i \mid N]\Big]$$

$$= \mathbb{E}\left[V[U]N\right] + V\left[N\mathbb{E}[U]\right]$$

$$= V(U)\mathbb{E}[N] + \mathbb{E}[U]^2 V[N]$$

③ $$\Lambda_X(s) = \mathbb{E}\left[e^{sX}\right] = \mathbb{E}\left[e^{s\sum_{i=1}^{N} U_i}\right]$$

$$= \mathbb{E}\left[\prod_{i=1}^{N} e^{sU_i}\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[\prod_{i=1}^{N} e^{sU_i} \,|\, N\right]\right]$$

$$= \mathbb{E}\left[\prod_{i=1}^{N} \mathbb{E}\left[e^{sU_i} \,|\, N\right]\right]$$

$$= \mathbb{E}\left[\prod_{i=1}^{N} \Lambda_U(s)\right] = \mathbb{E}\left[\Lambda_U(s)^N\right] = G_N(\Lambda_U(s))$$

<u>Exemple 2:</u> (Modèle poisson gamma)

Supposons $N \sim \mathcal{P}ois(\lambda)$ de loi $P_N(k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$, $k \geq 0$

et $U \sim \text{Gamma}(\alpha, \beta)$ de densité $f_U(x) = \dfrac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}$

On a $\mathbb{E}[X] = \lambda\alpha\beta$

$\mathbb{V}[X] = \lambda\alpha^2\beta^2 + \alpha\beta^2\lambda$

$\Lambda_X(s) = \exp\left\{\lambda\left(\left(\dfrac{1}{1-\beta}\right)^\alpha - 1\right)\right\}$

II - Modèle pour la fréquence des sinistres

Définition

1. $N \sim \mathcal{P}oiss(\lambda)$, $\lambda > 0$, avec $p_N(k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$, $k \geq 0$

2. $N \sim \mathcal{B}in(n,p)$, $n \in \mathbb{N}^*$, $p \in [0,1]$, avec $p_N(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, \dots, n$

3. $N \sim Neg\text{-}Bin(\alpha, p)$, avec $\alpha > 0$, $p \in [0,1]$

avec $$p_N(k) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(k+1)} (1-p)^\alpha p^k, \quad k \geq 0$$

NB : Loi de poisson $\Rightarrow$ 1 seul paramètre = peu de flexibilité

Loi binomial $\Rightarrow$ Support fini (pas pratique)

- Poisson = equi dispersé $\mathbb{E}[N] = V[N] = \lambda$
- Binomial = sous-dispersé, $\mathbb{E}[N] \geq V[N]$
- Binomial-negative = sur-dispersée $\mathbb{E}[N] \leq \mathbb{V}[N]$

Definition (Distribution zéro-Inflatée)

Soit N une va de comptage de loi p_N, la va. N_0 de loi

$$\mathbb{P}_{N_0}(k) = \begin{cases} p_0 \\ \dfrac{1-p_0}{1-p_N(0)} \, p_N(k), \text{ pour } k \geq 1 \end{cases}$$

Le modèle de poisson est souvent utilisé comme point de départ avec un nombre moyen de sinistre égal à λ.

On peut raffiner le modèle en supposant que $\Lambda = \lambda$ est aléatoire.

- Compenser une erreur d'estimation
- Incorporer de l'information a priori
- Prendre en compte l'hétérogénéité dans le portefeuille
- Accomoder la situation $V(N) > \mathbb{E}[N]$

Définition (Poisson-mélange)

Soit une va positive, la va N est une loi Poisson-mélange Λ-Poisson(Λ) avec

$$\mathbb{P}(N=k) = \int_{\mathbb{R}^+} \frac{e^{-\lambda}\lambda^k}{k!} \, d\mathbb{P}_\Lambda(\lambda)$$

$$N \mid \Lambda = \lambda \sim \mathcal{P}ois(\lambda)$$

Proposition :

1. $\mathbb{E}[N] = \mathbb{E}[\Lambda]$
2. $\mathbb{V}[N] = \mathbb{E}[\Lambda] + \mathbb{V}[\Lambda]$
3. $G_N(s) = \mathbb{E}[s^N] = \Pi_\Lambda(s-1)$

Preuve

① $\mathbb{E}[N] = \mathbb{E}[\mathbb{E}[N \mid \Lambda]] = \mathbb{E}[\Lambda]$

② $\mathbb{V}[N] = \mathbb{V}[\mathbb{E}[N \mid \Lambda]] + \mathbb{E}[\mathbb{V}[N \mid \Lambda]] = \mathbb{V}(\Lambda) + \mathbb{E}[\Lambda]$

③ $G_N(s) = \mathbb{E}[s^N] = \mathbb{E}[\mathbb{E}[s^N \mid \Lambda]]$

$= \mathbb{E}[\exp(\Lambda(s-1))] = \Pi_\Lambda(s-1)$

Exemple 3 (Hétérogénéité dans le ptf)

Deux types de conducteurs :

- Les conducteurs prudent génèrent λ_1 sinistres en moyenne par période d'exercice
- — moins prudent — $\lambda_2 > \lambda_1$ —

Pour un conducteur $i \in \{1, \dots, n\}$, on ne connait pas sa classe d'appartenance le nombre de sinistre par ce conducteur

$$N_i = N_1 \mathbb{1}_{A_i} + N_2 \mathbb{1}_{A_i^c}$$

où $N_1 \sim \mathcal{P}(\lambda_1)$, $N_2 \sim \mathcal{P}ois(\lambda_2)$ et $A_i = \{\text{le conducteur } i \text{ est prudent}\}$ soit des va /évn idpt

On a $\mathbb{P}(N_i = m) = \mathbb{P}(N_i = m)\mathbb{P}(A_i) + \mathbb{P}(N_2 = m)(1 - \mathbb{P}(A_i))$

En supposant que $\mathbb{P}(A_i) = \mathbb{P}(A_j)$ $\forall i,j \in \{1, \dots, n\}$ alors le nombre totale de sinistre.

$$N = \sum_{i=1}^{n} N_i \sim N\text{-}\mathcal{P}ois(\lambda_1 \times k + \lambda_2(n-k)) \quad \text{où } k \sim \mathcal{B}in(n,p)$$

On part du principe qu'il y a une proportion p de bon conducteur dans la population. On peut mettre à jour la probabilité $\mathbb{P}(A_i)$ que i soit un bon conducteur via la règle de Bayes avec $\mathbb{P}(A_i | N_i = \ell) = \dfrac{\mathbb{P}(N_i = \ell | A_i)\mathbb{P}(A_i)}{\mathbb{P}(N_i = \ell | A_i)\mathbb{P}(A_i) + \mathbb{P}(N_i = \ell | A_i^c)\mathbb{P}(A_i^c)}$

↳ Ajustement de la prime via un système de Bonus-Malus

$$N \sim M\text{-}\mathcal{P}ois\left(\lambda_1 \sum_{i=1}^{n} \mathbb{1}_{A_i} + \lambda_2 \sum_{i=1}^{n} \mathbb{1}_{A_i^c}\right)$$

<u>Exemple</u>: (Incertitude autour du paramètre)

Si le nombre moyen de sinistre est donné par $\hat{\lambda} = \frac{1}{t} \sum_{i=1}^{t} N_i$ pour t période d'exercice avec $N_1, \dots, N_t$ les observations.

L'incertitude d'échantillonage peut être pris en compte en supposant que

$$\Lambda \sim \mathcal{U}nif\left(\left[\hat{\lambda} \pm q_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{t}}\right]\right)$$ où $q_{1-\alpha/2}$ est un quantile de la loi de Student à $t-1$ degré de liberté.

Exemple 5 : (Influence Bayesienne de la loi de Poisson)

La stat Bayesienne permet de pallier à un manque d'observation via l'incorporant de l'information a priori sur le paramètre.

Pour un paramètre positif comme λ, un choix classique est la loi gamma $\Lambda \sim Gamma(\alpha,\beta)$. Notons $N_1,\dots,N_t$ les observations du nombre de sinistres en cours des périodes d'exercice.

La loi a posteriori de Λ avec

$$\mathbb{P}_{\Lambda|N}(\lambda|n) = \frac{\mathbb{P}_{N|\Lambda}(n|\lambda)\,\mathbb{P}_\Lambda(\lambda)}{\int \mathbb{P}_{N|\Lambda}(n|\lambda)\,d\mathbb{P}_\Lambda(\lambda)}$$

où $\mathbb{P}_{N|\Lambda}(n|\lambda)$ est la vraisemblance des données sachant que $\Lambda=\lambda$.

L'évolution du dénominateur pose problème (on évite son évaluation en utilisant un algo MCMC.)

Le choix de $N \sim Pois(\Lambda)$ et $\Lambda \sim Gamma(\alpha,\beta)$ donne

$$\int_0^\infty \mathbb{P}_{N|\Lambda}(n|\lambda)\,\mathbb{P}_\Lambda(\lambda)\,d\lambda = \int_0^\infty \prod_{i=1}^{t} \frac{e^{-\lambda}\lambda^{n_i}}{n_i!}\,\frac{e^{-\lambda/\beta}\lambda^{\alpha-1}}{\beta^\alpha\Gamma(\alpha)}\,d\lambda$$

$$= \int \frac{e^{-\lambda t}\lambda^{\sum n_i}}{\prod_{i=1}^{t} n_i!}\,\frac{e^{-\lambda/\beta}\lambda^{\alpha-1}}{\beta^\alpha\Gamma(\alpha)}\,d\lambda$$

$$= \frac{1}{\prod_{i=1}^{r} n_i!\,\beta^{\alpha}\,\Gamma(\alpha)} \int_0^{\infty} e^{-\lambda(t+1/\beta)}\,\lambda^{\sum n_i+\alpha-1}\,d\lambda$$

$$= \frac{1}{\prod_{i=1}^{r} n_i!\,\beta^{\alpha}\,\Gamma(\alpha)} \times \frac{\Gamma\left(\sum_{i=1}^{r} n_i+\alpha\right)}{(t+1/\beta)^{\sum n_i+\alpha}} \underbrace{\int \frac{(t+1/\beta)^{\sum n_i+\alpha}\, e^{-\lambda(\frac{1}{\beta}+t)}\,\lambda^{\sum n_i+\alpha-1}}{\Gamma\left(\sum_{i=1}^{r} n_i+\alpha\right)}\,d\lambda}_{1}$$

Exercice :

$N \sim \mathrm{Pois}(\Lambda)$, $\Lambda \sim \mathrm{Gamma}(\alpha,\beta)$, montrer que $N \sim \mathrm{Neg\text{-}Bin}(\cdot,\cdot)$

III - Modèles pour les montants de sinistres

1/ Les lois classiques

Les lois de type gamma ont le problème suivant

$$\mathbb{P}(U > x) \sim e^{-x/\beta} \quad , x \to +\infty$$

qui entraine une sous-estimation de la probabilité d'occurence de montants de sinistres élevés.

On préfèrera utiliser des distribution de probabilité à queue dite "lourde". En actuariat, la distribution de X est à queue lourde si $\underbrace{M_X(s)}_{=\mathbb{E}[e^{sX}]} = \infty \quad \forall s > 0$

Definition (loi lognormale)

La va $U \sim LN(\mu, \sigma^2)$ a pour densité

$$f_U(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right] , x > 0$$

et sa fonction de survie

$$\bar{F}_U(x) = 1 - \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)$$

On note que $U = \exp(\mu + \sigma Z)$, où $Z \sim \mathcal{N}(0,1)$

Proposition

Les moments de $U \sim LN(\mu, \sigma^2)$ sont donnés par

$$\mathbb{E}[U^k] = \exp\left(k\mu + k^2 \frac{\sigma^2}{2}\right), \quad k = 0, 1, \ldots$$

NB:

① La loi lognormale admet des moments de tout ordre, mais pas de fonction génératrice des moments.

② Plus $\sigma \nearrow$ plus la loi lognormale est à queue lourde.

③ La loi de $U_1 + U_2$, pour $U_1, U_2 \overset{iid}{\sim} LN(\mu, \sigma^2)$, est inconnue.

Définition (loi de Weibull)

$U \sim Weib(\alpha, \beta)$ de densité $f_U(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-(x/\beta)^\alpha}$

et de fonction de survie $\bar{F}_U(x) = e^{-(x/\beta)^\alpha}$

NB: α régit la queue de la distribution

- $\alpha = 1 \Rightarrow Exp$
- $\alpha < 1 \Rightarrow$ Sous exponentiel
- $\alpha > 1 \Rightarrow$ Sur exponentiel

Proposition

Les moments de $U \sim Weil(\alpha, \beta)$ sont donnés par

$$\mathbb{E}[U^k] = \beta^k \Gamma(1 + k/\alpha), \quad k \geq 0$$

Définition

$U \sim Par(\alpha, \gamma)$ de densité

$$f_U(x) = \begin{cases} \frac{\alpha \gamma^\alpha}{x^{\alpha+1}}, & si \ x \geq \gamma \\ 0, & sinon \end{cases}$$

et $\mathbb{P}(U > x) = \bar{F}_U(x) = \begin{cases} \left(\frac{\gamma}{x}\right)^\alpha, & si \ x > \gamma \\ 1, & si \ x \leq \gamma \end{cases}$

NB: 1. Si $\alpha < 1$, $\mathbb{E}[U] = \infty$

2. $X = \gamma e^{Y/\alpha}$, $Y \sim \mathcal{E}(1)$

Proposition

Soit $x_1, \dots, x_n$ un échantillon iid de loi de Pareto, l'estimateur du maximum de vraisemblance est donné par

$$\hat{\gamma} = \min_{i=1,\dots,n} x_i \qquad et \qquad \hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(x_i/\hat{\gamma})}$$

Exemple

En présence de valeurs extrêmes, une procédure de choix pour le seuil de valeurs extrême consiste à grapher la suite

$$\hat{\alpha}_k = \frac{k}{\sum_{i=j+1}^{n} \ln\left(\frac{x_{n-i+1}}{x_{n-k,n}}\right)} \quad , k = 1, \dots, n-1$$

où $x_{1,n}, \dots, x_{n,n}$ désigne les statistiques d'ordre de l'échantillon $x_1, \dots, x_n$.

2/ Les modèles composites

Définition

Soient f_1 et f_2 deux densités de probabilité sur $\mathbb{R}^+$, le modèle composite est donnée par

$$f(x) = \begin{cases} r \dfrac{f_1(x)}{F_1(\theta)} & , x \leq \theta \\ (1-r) \dfrac{f_2(x)}{F_2(\theta)} & , x > \theta \end{cases}$$

θ est le paramètre de seuil. On impose les conditions de régularité suivantes.

$$f(\theta^-) = f(\theta^+) \quad \text{et} \quad f'(\theta^-) = f'(\theta^+)$$

Exemple (Important)

Modèle composite Weib-Par $(r, k, \beta, \alpha, \theta)$ comme mélange de la loi de Weibull de densité

$$f_1(x) = \frac{k}{\beta}\left(\frac{x}{\beta}\right)^{k-1} e^{-(x/\beta)^k} \quad , x > 0$$

et de la loi de Pareto

$$f_2(x) = \begin{cases} \dfrac{\alpha\theta^\alpha}{x^{\alpha+1}} , & x > \theta \\ 0 , & \text{si } x \leq \theta \end{cases}$$

Les conditions de régularité $f(\theta^-) = f(\theta^+)$ et $f'(\theta^-) = f'(\theta^+)$ conduisent à fixer deux paramètres sur cinq. Par exemple, on peut fixer

$$r = \frac{\frac{\alpha}{\theta}\left[1 - \exp\left(-\frac{k+\alpha}{k}\right)\right]}{\frac{k}{\theta} + \frac{k}{\theta}\exp\left(-\frac{k+\alpha}{k}\right)} \quad \text{et} \quad \beta = \left(\frac{k}{k+\alpha}\right)^{1/k}\theta$$

$$f(\theta^-) = f(\theta^+) \iff \frac{r}{1-r} = \frac{f_2(\theta)\, F_1(\theta)}{f_1(\theta)}$$

On note

$$f_1'(x) = f_1(x)\left[\frac{k-1}{x} - k\frac{x^{k-1}}{\beta^k}\right] , \quad x > 0$$

et $f_2'(x) = -f_2(x)\dfrac{\alpha+1}{x}$

$$f'(\theta^-) = f'(\theta^+) \Rightarrow \frac{r}{1-r} \frac{f_1'(\theta)}{f_1(\theta)} = f_2'(\theta) \Leftrightarrow \beta = \left(\frac{k}{k+\alpha}\right)^{1/k} \theta$$

puis $$r = \frac{f_2(\theta) F_1(\theta)}{f_2(\theta) F_1(\theta) + f_1(\theta)} = \frac{\frac{\alpha}{\theta}\left[1 - \exp\left(-\frac{k+\alpha}{k}\right)\right]}{\frac{\alpha}{\theta} + \frac{k}{\theta}\exp\left(-\frac{k+\alpha}{k}\right)}$$

<u>NB</u> : Pour l'inférence

1. fixer θ, puis estimer séparément les paramètres de f_1 et f_2 avec les montants de sinistres $\leq \theta$ et $\geq \theta$ respectivement

   ↳ le choix du seuil ne prend pas en compte la forme de la distribution du ventre de la queue de la distribution.

2. Estimation simultanée de tous les paramètres du modèle composite via le maximum de vraisemblance

3. Les modèles flexibles

Définition (Mélange de Erlang)

$X \sim M\text{-}Erl(\Pi, \beta)$ a pour densité

$$f_X(x) = \sum_{i=1}^{m} \Pi_i \frac{e^{-x/\beta} x^{z_i - 1}}{\beta^{z_i} (z_i - 1)!}$$

Il s'agit d'un modèle de mélange avec une va Z à valeurs entières $z_1, \dots, z_n$ de loi Π avec $\mathbb{P}(Z = z_i) = \Pi_i$, $i = 1, \dots, n$

$X \sim Gamma(Z, \beta)$

Théorème

Les lois mélange de Erlang sont denses dans l'ensemble des distribution de $\mathbb{R}^+$.

Conséquence : On peut en théorie approcher n'importe quelle loi sur $\mathbb{R}^+$.

L'inférence des mélanges de Erlang s'effectue via un algo itératif de type Expectation - Maximisation (EM).
On initialise la valeur des paramètres avec $\Theta^{(0)} = (\Pi_1^{(0)}, \dots, \Pi_m^{(0)}, \beta^{(0)})$.
On souhaite maximiser la vraisemblance des données $x = (x_1, \dots, x_n)$

définie par $p(x,z,\theta) = \prod_{i=1}^{n} \frac{e^{-x_k/\beta} x_k^{z-1}}{\beta^z (z-1)!}$

approchée par l'espérance conditionnelle sachant z.

① Etape Espérance (E)

on calcul $Q(\theta|\theta^{(0)}) = \mathbb{E}_{z|x,\theta^{(0)}}\left[\ln(p(x,z,\theta^{(0)}))\right]$

$$= \sum_{i=1}^{n}\sum_{k=1}^{n} \ln(p(x_k, z_i, \theta^{(0)}))\mathbb{P}(Z_i = z_i | x_k, \theta^{(0)})$$

② Etape Maximisation (M)

$\theta^{(1)} = \underset{\theta}{\operatorname{argmax}}\ Q(\theta|\theta^{(0)})$ sous la contrainte $\sum_{i=1}^{n} \pi_i = 1$

on note que

$$\mathbb{P}(Z = z_i | x; \theta^{(0)}) = \frac{p(x, Z = z_i, \theta)}{p(x,\theta)} = \frac{\pi_i^{(0)} \prod_{k=1}^{n} \frac{e^{-x_k/\beta^{(0)}} x_k^{z_i - 1}}{(z_i-1)!\, \beta^{(0)z_i}}}{\sum_{i=1}^{m} \pi_i \prod_{k=1}^{n} \frac{e^{-x_k/\beta^{(0)}} x_k^{\beta_i - 1}}{(z_i-1)!\, \beta^{(0)z_i}}}$$

$$L(\pi_1, \dots, \pi_m, \beta) = \sum_{i=1}^{m}\sum_{k=1}^{n} \ln\left[p(x_k, z_i, \theta)\mathbb{P}(Z = z_i | x_k, \theta^{(0)})\right] - \lambda\left(\sum_{i=1}^{m} \pi_i - 1\right)$$

$(\pi_1^{(1)}, \dots \pi_m^{(1)}, \beta^{(1)})$ sont solutions du système

$$\begin{cases} \dfrac{\partial L}{\partial \pi_i}(\pi_i, \dots, \pi_n, \beta) = 0 \\ \dfrac{\partial L}{\partial \beta}(\pi_1, \dots \pi_m, \beta) = 0 \\ \displaystyle\sum_{i=1}^{m} \pi_i = 1 \end{cases}$$

$X \sim M\text{-}Erl(2, \beta)$

$\mathbb{E}[X] = \mathbb{E}[Z]\beta$

On résout le système pour trouver

$$\pi_i^{(1)} = \frac{\sum_{k=1}^{n} \mathbb{P}(Z_i = z_i \mid x_k, \Theta^{(0)})}{n} \quad i = 1, n \qquad \text{et} \quad \beta^{(1)} = \frac{\frac{1}{n}\sum_{k=1}^{n} x_k}{\sum_{i=1}^{m} \pi_i^{(1)} z_i}$$

On ré-itère les étapes E et M jusqu'à convergence

__NB__ : En pratique, le calibrage de la M-Erl consiste à considérer successivement $z = 1$, $z \in \{1, 2\}, \dots z \in \{1, 2, \dots, n\}$ puis à choisir le modèle qui minimise le BIC, défini par $BIC(i) = -2\ln(p(x; \pi_1, \dots; \pi_i; \beta)) + (i+1)\ln(n)$, $i = 1, \dots, n$

__NB__ : Les lois mélanges de Erlang sont des distributions à queue légère. Une solution est d'utiliser la distribution M-Erl comme composante pour la partie attritionnelle dans un modèle composite.

Autre solution : phases types généralisées (cf. google scholar)

III - Méthode d'approximation

1/ Approximation normale et gamma

Définition (Approximation normale)

Supposons que $N \sim \mathcal{P}ois(\lambda)$, on note $\mu = \mathbb{E}[U]$ et $\sigma^2 = \mathbb{V}[U]$

On a donc

$$\mathbb{E}[X] = \lambda \, \mathbb{E}[U] \quad \text{et} \quad \mathbb{V}[X] = \lambda(\mu^2 + \sigma^2)$$

L'approximation normale consiste à approcher

$$Z = \frac{X - \lambda\mu}{\sqrt{\lambda(\mu^2+\sigma^2)}} \sim \mathcal{N}(0,1)$$

La fonction de répartition de X est approché par

$$\mathbb{P}(X \leqslant x) = \phi\left(\frac{x - \lambda\mu}{\sqrt{\lambda(\mu^2+\sigma^2)}} \right) \quad \text{où } \phi \text{ est la f.d.r } \mathcal{N}(0,1)$$

Cette approximation est d'autant meilleur que $\lambda \to \infty$

Définition (Approximation Gamma)

L'approximation gamma de X consiste à approcher le modèle collectif par une loi gamma translatée. $V \sim T\text{-}Gamma(\alpha, \beta, w_0)$ avec

$$V = w_0 + W, \quad \text{où } W \sim Gamma(\alpha, \beta)$$

On fait correspondre les moments de X et V jusqu'à l'ordre 3. On a

$$\beta = \frac{\mu_3(X)}{2\,\mathbb{V}(X)}, \quad \alpha = \frac{4\,\mathbb{V}[X]^3}{\mu_3(X)^2} \quad \text{et} \quad w_0 = \mathbb{E}[X] - 2\frac{\mathbb{V}[X]}{\mu_3(X)}$$

où $\mu_3(X) = \mathbb{E}[(X - \mathbb{E}[X])^3]$

$$\begin{cases} \mathbb{E}[X] = \mathbb{E}[V] \to \text{exprimer en fc de } w_0, \alpha \text{ et } \beta \\ \mathbb{V}[X] = \mathbb{V}[V] \\ \mu_3(X) = \mu_3(V) \end{cases}$$

2/ Algorithme de Panjer

Définition

La distribution d'une v.a de comptage N appartient à la famille de Panjer si

$$p_N(n) = \left(a + \frac{b}{n}\right) p_N(n-1) \quad , \quad n \geq 1 \quad \text{où } a < 1 \text{ et } b \in \mathbb{R}$$

Proposition

Si N est dans la famille de Panjer alors

- $N \sim \mathcal{P}ois(\lambda)$ avec $a = 0$ et $\lambda = b$
- $N \sim Neg\text{-}Bin(\alpha, p)$ avec $p = a$ et $\alpha = 1 + bp^{-1}$
- $N \sim Bin(n, p)$ avec $p = a(a-1)^{-1}$ et $n = 1 - ba^{-1}$

Définition

Si U est une va de comptage alors X est aussi une va de comptage dont la loi vérifie

$$p_X(k) = \begin{cases} G_N(p_U(0)) & , \text{pour } k = 0 \\ \dfrac{1}{1 - a p_U(0)} \displaystyle\sum_{j=1}^{k} \left(a + b\frac{j}{k}\right) p_U(j) p_X(k-j) & \end{cases}$$

NB : (Discrétisation de la loi des montants de sinistres)

Si U est une va continue, il est possible d'utiliser l'algo de Panjer après avoir discrétisé la loi des montants de sinistres.

On schéma de discrétisation classique est le suivant

$$p_U(kh) = \begin{cases} F_U(h/2) & , \text{si } k=0 \\ F_U(kh+h/2) - F_U(kh-h/2), & k \geq 1 \end{cases}$$

3/ Fast Fourier Transform (FFT)

La méthode FFT est motivée par deux faits

1. Accessibilité de la transformée de Fourier $\mathbb{E}[e^{isX}]$
2. Temps de calcul trop long avec Panjer

La procédure FFT est exacte si les montants de sinistres sont des v.a discrètes à support fini. Supposons que $X \in \mathbb{N}$ avec $X \leq n$

Sa transformée de Fourier est donnée par

$$\Pi_X(is) = \mathbb{E}[e^{isX}] = \sum_{k=0}^{n} p_X(k)\, e^{isk} \quad , \; s \in \mathbb{R}$$

On évalue la transformée de Fourier de X sur une grille de point

$$s_j = \frac{2\pi j}{n+1} \quad , \; j = 0 \dots, n$$

Les $p_X(k)$ sont solutions de

$$\begin{cases} \Pi_X(s_0) = \sum_{k=0}^{n} p_X(k)\, e^{is_0 k} \\ \\ \Pi_X(s_n) = \sum_{k=0}^{n} p_X(k) e^{is_n k} \end{cases}$$

qui se ré-écrit matriciellement

$$\Pi_X^T = F\, p_X^T \qquad \text{où} \; \begin{cases} \Pi_X = (\Pi_X(s_0), \dots, \Pi_X(s_n)) \\ p_X = (p_X(0), \dots, p_X(n)) \end{cases}$$

et

$$F = \begin{bmatrix} e^{is_0 \times 0} & 1 \; \dots & e^{is_n \times 0} \\ e^{is_0} & & \vdots \\ \vdots & & \vdots \\ e^{ins_0} & \dots & e^{ins_n} \end{bmatrix} = \left(e^{iks_\ell}\right)_{k,\ell = 0,\dots,n}$$

On a $F^{-1} = \dfrac{1}{n+1} \left(e^{-i\ell s_k}\right)_{k,\ell=0,\dots,n}$