

Modélisation Charge Sinistre M2 Actuariat

Chapitre III: Modèle Collectif

Pierre-Olivier Goffard

Université de Lyon 1
ISFA

`pierre-olivier.goffard@univ-lyon1.fr`

ISFA

October 18, 2021

I. Définition du modèle collectif

Le modèle collectif donne la somme agrégé des montants des sinistres sur une période d'exercice donnée à la maille portefeuille plutôt qu'à la maille contrat ou groupe homogène de contrat. On suppose que

- le nombre de sinistre est donné par une variable de comptage $N \in \mathbb{N}$, de loi notée

$$p_N(n) = \mathbb{P}(N = n), \quad n = 0, 1, \dots$$

- Les montants de sinistres U_1, U_2, \dots, U_N forme une suite de variables aléatoire positives **i.i.d.** de densité de probabilité f_U et indépendantes de N .

La charge totale est définie par

$$X = \sum_{i=1}^N U_i. \quad (1)$$

Ce modèle est moins précis que le modèle individuel mais bien plus commode à calibrer. Son problème principal est l'absence de formule fermée (à quelques exceptions près) pour sa fonction de répartition.

1. Distribution de X

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

$$d\mathbb{P}_X(x) = p_N(0)\delta_0(x) + f_X^+(x)d\lambda(x),$$

où $f_X^+(x)$ correspond à la densité de la partie absolument continue de X définie par

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

avec f_U^{*k} qui désigne le $k^{\text{ème}}$ produit de convolution de f_U avec elle-même (densité de la somme $U_1 + \dots + U_k$). Il s'agit d'une densité de probabilité défailante au sens où

$$\int f_X^+(x)d\lambda(x) = 1 - p_N(0).$$

L'expression de f_X^+ est problématique car elle implique une série infinie avec des produit de convolutions imbriqués (intégrale successive).

Exemple 1 (Modèle géométrique-Exponentiel)

Supposons que les sinistres soient de loi exponentielle $U \sim \text{Exp}(\beta)$ de densité

$$f_U(x) = \frac{e^{-x/\beta}}{\beta}, \quad x > 0.$$

et que la fréquence des sinistres soit de loi géométrique $N \sim \text{Geom}(p)$ de loi

$$p_N(k) = p^k(1-p), \quad k = 1, 2, \dots$$

On a

$$\begin{aligned} f_X^+(x) &= \sum_{k=1}^{+\infty} (1-p)p^k \frac{e^{-x/\beta} x^{k-1}}{\beta^k (k-1)!} \\ &= (1-p)e^{-x/\beta} \sum_{k=1}^{+\infty} p^k \frac{x^{k-1}}{\beta^k (k-1)!} \\ &= \frac{(1-p)p}{\beta} e^{-x/\beta} \sum_{k=0}^{+\infty} \frac{1}{k!} \left(\frac{px}{\beta}\right)^k \\ &= \frac{(1-p)p}{\beta} e^{-x(1-p)/\beta}. \end{aligned}$$

On doit très souvent avoir recours aux méthodes numériques pour calculer les quantités intéressantes pour la gestion des risques comme par exemple la *Value-at-Risk*

$$\text{VaR}_X(\alpha) = \inf\{x \geq 0 : F_X(x) \geq \alpha\}$$

ou la *Tail-Value-at-Risk*

$$\text{TVaR}_X(\alpha) = \mathbb{E}[X | X > \text{VaR}_X(\alpha)].$$

En réassurance, la prime d'un traité non-proportionnelle est tarifé par $\mathbb{E}[(X - c)_+]$ où $c > 0$ désigne le seuil de rétention, on voit parfois $\mathbb{E}[\min((X - c)_+, d)]$ où $d > 0$ est la limite.

2. Moments de X

Proposition 1 ($\mathbb{E}(X)$, $\mathbb{V}(X)$, et M_X)

- ① L'espérance de X est donnée par

$$\mathbb{E}(X) = \mathbb{E}(N) \times \mathbb{E}(U)$$

- ② La variance de X est donnée par

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

- ③ La fonction génératrice des moments $M_X(s) := \mathbb{E}(e^{sX})$ de X est donnée par

$$M_X(s) = G_N[M_U(s)],$$

où $G_N(s) := \mathbb{E}(s^N)$ est la fonction génératrice des probabilités de N

preuve:

①

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}[\mathbb{E}(X|N)] = \mathbb{E}\left[\mathbb{E}\left(\sum_{i=1}^N U_i | N\right)\right] = \mathbb{E}\left[\sum_{i=1}^N \mathbb{E}(U_i | N)\right] \\ &= \mathbb{E}[N\mathbb{E}(U)] = \mathbb{E}(N)\mathbb{E}(U). \end{aligned}$$

2

$$\begin{aligned}
 \mathbb{V}(X) &= \mathbb{E}[\mathbb{V}(X|N)] + \mathbb{V}[\mathbb{E}(X|N)] = \mathbb{E}\left[\mathbb{V}\left(\sum_{i=1}^N U_i|N\right)\right] + \mathbb{V}\left[\mathbb{E}\left(\sum_{i=1}^N U_i|N\right)\right] \\
 &= \mathbb{E}\left[\sum_{i=1}^N \mathbb{V}(U_i|N)\right] + \mathbb{V}\left[\sum_{i=1}^N \mathbb{E}(U_i|N)\right] \\
 &= \mathbb{E}[N\mathbb{V}(U)] + \mathbb{V}[N\mathbb{E}(U)] = \mathbb{E}(N)\mathbb{V}(U) + \mathbb{V}(N)\mathbb{E}(U)^2.
 \end{aligned}$$

3

$$\begin{aligned}
 M_X(s) &= \mathbb{E}(e^{sX}) = \mathbb{E}\left(e^{s\sum_{i=1}^N U_i}\right) = \mathbb{E}\left[\mathbb{E}\left(\prod_{i=1}^N e^{sU_i}|N\right)\right] = \mathbb{E}\left[\prod_{i=1}^N \mathbb{E}(e^{sU})\right] \\
 &= \mathbb{E}\left[\mathbb{E}(e^{sU})^N\right] = \mathbb{E}\left[\mathbb{E}(M_U(s))^N\right] = G_N[M_U(s)].
 \end{aligned}$$

Exemple 2 (Modèle Poisson-Gamma)

Supposons que $N \sim \text{Pois}(\lambda)$ de loi

$$\mathbb{P}(N = n) = \frac{e^{-\lambda} \lambda^n}{n!}, \quad n = 0, 1, 2, \dots$$

et que $U \sim \text{Gamma}(\alpha, \beta)$ de densité

$$f_U(x) = \frac{e^{-x/\beta} x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)}$$

On a

- ① $\mathbb{E}(X) = \alpha\beta\lambda$
- ② $\mathbb{V}(X) = \alpha\beta^2\lambda + \alpha^2\beta^2\lambda$
- ③ $M_X(s) = \exp\left\{\lambda\left[\left(\frac{1}{1-\beta s}\right)^\alpha - 1\right]\right\}$

II. Modèle pour la fréquence des sinistres

Definition 1 (Les classiques)

- ① Loi de Poisson $N \sim \text{Pois}(\lambda)$, $\lambda > 0$, avec

$$p_N(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

- ② Loi binomial $N \sim \text{Bin}(n, p)$, $p \in [0, 1]$, $n \in \mathbb{N}^*$, avec

$$p_N(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n.$$

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

$$p_N(k) = \frac{\Gamma(\alpha + k)}{\Gamma(k+1)\Gamma(\alpha)} p^k (1-p)^\alpha, \quad k = 0, 1, 2, \dots$$

Remark 1

La loi de Poisson manque de flexibilité du fait de son seul paramètre λ . La loi binomial a le défaut d'admettre un support fini. La loi binomial négative est bien souvent le meilleur choix. On distingue ces distributions sur la base de leur dispersion

- *La loi de Poisson est équi-dispersée au sens où $\mathbb{E}(N) = \mathbb{V}(N)$*
- *La loi binomial est sous-dispersée avec $\mathbb{E}(N) \geq \mathbb{V}(N)$*
- *La loi binomiale négative est sur-dispersée avec $\mathbb{E}(N) \leq \mathbb{V}(N)$*

Les données de fréquence de sinistres sur des périodes d'exercice exhibent bien souvent un trop grand nombre de 0 qui rendent les distributions inopérantes. On peut néanmoins y remédier facilement en modifiant en 0 les lois de fréquence.

Definition 2 (Zero-inflated distribution)

Soit N une variable aléatoire de comptage de loi p_N , la variable N_0 de loi

$$p_{N_0}(k) = \begin{cases} p_0, & \text{pour } k = 0, \\ \frac{1-p_0}{1-p_N(0)} p_N(k), & \text{pour } k \geq 1. \end{cases}$$

où $p_0 \in [0, 1]$ est la version modifiée en 0 de la variable aléatoire N

Remark 2

Il s'agit d'une simple modification qui force la probabilité de ne pas observer de sinistres à p_0 . On estime ce paramètre sur la base de la proportion de 0 dans nos observations, puis on utilise les observations non nulle pour calibrer $p_{N_0|N_0>0}$.

Malgré ses défauts, le modèle de Poisson est souvent le point de départ de la modélisation de la fréquence des sinistres. Sur une période donnée, l'ensemble du portefeuille ou un assuré, déclare un nombre moyen de sinistre λ . Le modèle peut-être raffiné en supposant λ aléatoire

- Compenser l'erreur d'estimation associé à un manque de données
- Incorporer de l'information a priori
- Prendre en compte l'hétérogénéité du portefeuille d'assuré
- Accomoder une situation $\mathbb{V}(N) > \mathbb{E}(N)$

Definition 3 (Poisson Mélange)

Soit Λ une variable aléatoire positive, la variable aléatoire N suit une loi de Poisson mélange M-Pois(Λ), avec

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

autrement dit $N|\Lambda = \lambda \sim \text{Pois}(\lambda)$.

Proposition 2

- 1 $\mathbb{E}(N) = \mathbb{E}(\Lambda)$
- 2 $\mathbb{V}(N) = \mathbb{E}(\Lambda) + \mathbb{V}(\Lambda)$
- 3 $G_N(s) = M_\Lambda(s-1)$

preuve:

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

Exemple 3 (Hétérogénéité du portefeuille)

On suppose l'existence de deux types de conducteurs,

- un conducteur prudent générant λ_1 accidents en moyenne par période
- un conducteur à risque générant $\lambda_2 > \lambda_1$ accidents en moyenne par période

Pour un conducteur donné $i \in \{1, \dots, n\}$, on ne connaît pas sa classe d'appartenance, le nombre de sinistres pour ce conducteur est donnée par

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

où $N_1 \sim \text{Pois}(\lambda_1)$, $N_2 \sim \text{Pois}(\lambda_2)$ et $A = \{\text{Conducteur prudent}\}$ sont des variables aléatoires/ événements indépendants. On a

$$\mathbb{P}(M_i = m) = \mathbb{P}(N_1 = m)\mathbb{P}(A_i) + \mathbb{P}(N_2 = m)(1 - \mathbb{P}(A_i))$$

Si on note $p_i = \mathbb{P}(A_i)$ et $\Lambda_i = \lambda_1 \mathbb{1}_{A_i} + \lambda_2 \mathbb{1}_{A_i^c}$ alors $M_i \sim \text{M-Pois}(\Lambda_i)$.

En supposant que les nombres de sinistres M_1, \dots, M_n pour chacun des assurés forme une suite de variables aléatoires iid M-Pois(Λ) ($p_i = p_j = p$ pour tout i, j) alors le nombre de sinistres pour l'ensemble du portefeuille vérifie

$$M = \sum_{i=1}^n M_i \sim \text{M-Pois}[\lambda_1 \times K + \lambda_2 \times (n - K)],$$

où $K \sim \text{Bin}(n, p)$. On part du postulat qu'on a en portefeuille une proportion p de bon risque. Au fur et à mesure, on peut mettre à jour la probabilité d'appartenir à telle ou telle classe de conducteur en utilisant le théorème de Bayes. Si le conducteur i subit $l \in \mathbb{N}$ sinistres, on remplacera $\mathbb{P}(A_i)$ par

$$\mathbb{P}(A_i | M_i = l) = \frac{\mathbb{P}(M_i = l | A_i) \mathbb{P}(A_i)}{\mathbb{P}(M_i = l | A_i) \mathbb{P}(A_i) + \mathbb{P}(M_i = l | A_i^c) \mathbb{P}(A_i^c)}.$$

Cela va permettre d'ajuster la prime via un mécanisme de Bonus-Malus par exemple. La distribution de M comme

$$M \sim \text{M-Pois} \left[\lambda_1 \sum_{i=1}^n \mathbb{1}_{A_i} + \lambda_2 \left(n - \sum_{i=1}^n \mathbb{1}_{A_i} \right) \right]$$

peut poser problème à l'évaluation. On constate ici comment un modèle de mélange permet de faire le pont entre modélisation individuelle et collective.

Exemple 4 (Incertitude sur le paramètre)

Si le nombre moyen de sinistres observés est $\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n N_i$, l'incertitude d'échantillonnage peut être pris en compte en définissant

$$\Lambda \sim \text{Unif} \left[\hat{\lambda} - t_{1-\alpha/2}(n-1) \frac{\hat{\sigma}}{\sqrt{n}}, \hat{\lambda} + t_{1-\alpha/2}(n-1) \frac{\hat{\sigma}}{\sqrt{n}} \right],$$

où $\hat{\sigma}$ représente l'écart type de la série d'observation de la fréquence des sinistres et $t_{1-\alpha/2}(n-1)$ désigne le quantile d'ordre $1-\alpha/2$ d'une loi de Student à $n-1$ degré de liberté.

Exemple 5 (Inférence Bayésienne de la loi de Poisson)

L'inférence Bayésienne permet souvent de pallier à un manque d'observation en incorporant de l'information a priori sous la forme d'une distribution a priori sur le paramètre. Pour un paramètre positif (comme λ) un choix classique est la loi gamma $\Lambda \sim \Gamma(\alpha, \beta)$. Notons $N = (N_1, \dots, N_n)$ la série d'observation de la fréquence des sinistres. Le paramètre λ est estimé par la moyenne ou le mode de la loi a posteriori

$$p_{\Lambda|N}(\Lambda|N) = \frac{p_{N|\Lambda}(k|\lambda)f_{\Lambda}(\lambda)}{\int p_{N|\Lambda}(k|\lambda)f_{\Lambda}(\lambda)d\lambda},$$

où $p_{N|\Lambda}(k|\lambda)$ désigne la vraisemblance des données sachant que $\Lambda = \lambda$. L'évaluation du dénominateur de la vraisemblance a posteriori pose souvent des problèmes. Le choix gamma donne

$$\begin{aligned} \int p_{N|\Lambda}(k|\lambda)f_{\Lambda}(\lambda)d\lambda &= \int \frac{e^{-\lambda}\lambda^k}{k!} \frac{e^{-\lambda/\beta}\lambda^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} d\lambda \\ &= \frac{1}{k!\Gamma(\alpha)\beta^{\alpha}} \int e^{-\lambda\left(\frac{\beta+1}{\beta}\right)} \lambda^{k+\alpha-1} d\lambda \\ &= \frac{\Gamma(\alpha+k)}{k!\Gamma(\alpha)} \left(\frac{1}{1+\beta}\right)^{\alpha} \left(\frac{\beta}{1+\beta}\right)^k. \end{aligned}$$

II. Modèles pour les montants de sinistres

1. Les lois classiques

Les lois de type gamma ont le problème suivant

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

qui entraîne une sous estimation de la probabilité d'avoir des sinistres de forte intensité. On parle de distribution à queue légère. En actuariat, on considère fréquemment que les distributions sont à queues lourdes lorsqu'elle n'admettent pas de fonction génératrice des moments, c'est à dire

$$M_U(s) = \mathbb{E}(e^{sU}) = \infty \text{ pour tout } s > 0$$

Plusieurs lois sur \mathbb{R}_+ admettent une queue plus lourde permettant une modélisation plus appropriée des montants de sinistre.

Definition 4 (Loi lognormale)

La variable aléatoire U suit une loi lognormale $U \sim \text{LN}(\mu, \sigma^2)$ de densité

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

et de 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 est de loi normale centrée réduite.

Proposition 3

Les moments de $U \sim \text{LN}(\mu, \sigma^2)$ sont données par

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

preuve: Examen?

Remark 3

- 1 La loi lognormale admet des moments à tout ordre mais pas de fonction génératrice des moments.
- 2 Plus le paramètre σ est élevé plus la loi lognormale admet une queue lourde
- 3 La loi de $U_1 + U_2$ est inconnue, l'approximation de Fenton-Wilkinson [4] consiste à approcher la somme de variables lognormale par une variable de loi lognormale dont les moments d'ordre 1 et 2 coïncident avec ceux de la somme. Pour des méthodes plus sophistiquées voir par exemple Asmussen et al. [1]

Definition 5 (Loi de Weibull)

La variable aléatoire U suit une loi de Weibull $U \sim \text{Weib}(\alpha, \beta)$ de densité

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

et de fonction de survie

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

Remark 4

La valeur du paramètre de forme α régit le comportement dans la queue de la distribution. En effet,

- $\alpha = 1$ exponentielle
- $\alpha < 1$ sous-exponentielle
- $\alpha > 1$ sur-exponentielle

Proposition 4

Les moments de $U \sim \text{Weib}(\alpha, \beta)$ sont donnés par

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

preuve: Examen.

Definition 6 (Loi de Pareto)

La variable aléatoire U suit une loi de Pareto $U \sim \text{Par}(\alpha, \gamma)$ de densité

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

et

$$\bar{F}_U(x) = \begin{cases} \left(\frac{\gamma}{x}\right)^\alpha, & x \geq \gamma, \\ 1, & x < \gamma. \end{cases}$$

Remark 5

- 1 Si $\alpha \leq 1$ alors la loi de Pareto n'admet pas d'espérance, ni aucun autre moment!
- 2 $X = \gamma e^{Y/\alpha}$, où $Y \sim \text{Exp}(1)$

Proposition 5

Supposons que nous disposions d'un échantillon de montants de sinistres x_1, \dots, x_n , l'estimateur du maximum de vraisemblance de la loi de Pareto $\text{Par}(\alpha, \gamma)$ est donné par

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

preuve:

Soit $X = (x_1, \dots, x_n)$ les données disponibles, la vraisemblance du modèle s'écrit

$$p(X; \alpha, \gamma) = \prod_{i=1}^n \frac{\alpha \gamma^\alpha}{x_i^{\alpha+1}} \mathbb{1}_{x_i \geq \gamma} = \alpha^n e^{n\alpha \ln(\gamma)} \prod_{i=1}^n \frac{\mathbb{1}_{x_i \geq \gamma}}{x_i^{\alpha+1}}.$$

On remarque que nécessairement $\gamma \leq \min x_i$ sinon la vraisemblance s'annule. On remarque aussi que $p(X; \alpha, \gamma)$ est strictement croissante en $\gamma \in]0, \min x_i]$, on en déduit l'estimateur ML de γ avec $\hat{\gamma} = \min x_i$. Par suite, la log-vraisemblance est donnée par

$$\ln(p(X; \alpha, \hat{\gamma})) = n \ln(\alpha) + \alpha n \ln(\hat{\gamma}) - (\alpha + 1) \sum_{i=1}^n \ln(x_i).$$

puis

$$\frac{\partial}{\partial \alpha} \ln(p(X; \alpha, \hat{\gamma})) = 0 \Leftrightarrow \hat{\alpha} = \frac{n}{\sum_{i=1}^n \ln(x_i / \hat{\gamma})}.$$

Exemple 6

Si on suppose être en présence de valeurs extrêmes, une procédure de détection du seuil de valeurs extrêmes consiste à grapher la séquence

$$\hat{\alpha} = \frac{k}{\sum_{i=1}^k \ln(x_{n-i+1:n}/x_{n-k:n})}, \quad k = 1, \dots, n-1.$$

où $x_{1:n}, \dots, x_{n:n}$ représente les statistiques d'ordres de l'échantillon initial (soit l'échantillon initial trier par ordre croissant). Le seuil correspond à $x_{n-k:n}$ pour k l'indice pour lequel l'estimateur précédent se stabilise.

2. Les modèles composites

Les modèles composites ont été introduit en modélisation des montants de sinistres pour calibrer le 'ventre' et la queue de la distribution séparément. Les données montrent souvent une occurrence forte de sinistres de faible intensité et une occurrence non négligeable de sinistres de forte intensité.

Definition 7 (Modèles composites)

Soient f_1, f_2 deux densités de probabilité sur \mathbb{R}_+ . Un modèle composite est défini par une densité de probabilité

$$f(x) = \begin{cases} r \frac{f_1(x)}{F_1(\theta)} & \text{pour } 0 \leq x \leq \theta, \\ (1-r) \frac{f_2(x)}{1-F_2(\theta)} & \text{pour } x > \theta. \end{cases}$$

Le paramètre θ est souvent appelé *attachement point* ou *splicing point* auquel les conditions de régularités suivante sont imposées

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

Cela facilite notamment l'utilisation d'algorithmes d'optimisation pour inférer les paramètres via la méthode du maximum de vraisemblance.

Exemple 7 (Modèle Weibull-Pareto)

Le modèle composite Weib-Par($r, k, \beta, \alpha, \theta$) mélange une 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 une loi de Pareto

$$f_2(x) = \begin{cases} \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, & x \geq \theta, \\ 0, & x < \theta, \end{cases}$$

voir Scollnick and Sun [10]. Les conditions de régularité $f(\theta^-) = f(\theta^+)$ et $f'(\theta^-) = f'(\theta^+)$ impose de fixer la valeur de deux paramètres (par rapport aux autres). Par exemple, on peut fixer r et β avec

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

En effet, la continuité de f en θ est équivalente à

$$\frac{r}{1-r} = \frac{f_2(\theta)F_1(\theta)}{f_1(\theta)}$$

puis on note que

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

et

$$f_2'(x) = -f_2(x) \frac{\alpha+1}{x}.$$

La dérivabilité en θ entraîne que

$$\frac{r}{1-r} \frac{f_1'(\theta)}{F_1(\theta)} = f_2'(\theta) \Leftrightarrow \beta = \left(\frac{k}{k+\alpha} \right)^{1/k} \theta$$

puis

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

Technique de résolution qui pourrait bien s'avérer utile le jour de l'examen ...

Remark 6

Les hypothèses de régularité ne sont pas requise pour l'inférence. Deux solutions

- 1 *Fixer le seuil θ via le Hill estimator, et estimer séparément f_1 et f_2 à l'aide des données inférieures et supérieures à θ respectivement*
 - ↪ *Le choix du seuil ne prend pas en compte la forme de la distribution pour le ventre de la distribution*
- 2 *Estimation simultanée par le maximum de vraisemblance. En pratique, le seuil θ peut être choisi parmi les observations. On considère tour à tour chaque observation comme candidat pour le seuil, inférer les paramètres de f_1 et f_2 de la même manière que dans 1 puis comparer l'adéquation de chaque modèle via des critères d'information comme le BIC.*

3. Les modèles flexibles

Definition 8 (Mélange de Erlang)

La variable aléatoire X admet une loi Erlang mélange $M\text{-Erl}(\boldsymbol{\pi}, \beta)$ si sa densité est donnée par

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

Il s'agit d'un modèle de mélange avec un variable aléatoire latente $Z \sim \pi$ de comptage de loi

$$\mathbb{P}(Z = z_i), \quad z_i \in \mathbb{N}^* \quad i = 1, \dots, m.$$

On a en fait $X \sim \text{Gamma}(Z, \beta)$.

Theorem 1

Les lois mélange de Erlang sont denses dans l'ensemble des distributions de \mathbb{R}_+ .

On peut théoriquement approcher n'importe quel variable aléatoire à support sur \mathbb{R}_+ par la limite d'une suite de variables aléatoires de loi Erlang mélange, voir Tijms [12, p. 163-164] ou plus accessible et axé sur la modélisation des sinistres via la loi Erlang mélange Lee et Lin [6].

L'inférence des lois de mélange s'effectue souvent via un algorithme itératif dit d'Espérance-Maximisation (EM). L'idée est d'initialiser la valeur des paramètres $\theta^{(0)} = (\pi_1^{(0)}, \dots, \pi_m^{(0)}, \beta^{(0)})$, puis chercher à maximiser la vraisemblance des données $X = (x_1, \dots, x_n)$ définie par

$$p(X, Z; \theta) = \prod_{k=1}^n p(x_k, Z; \theta) = \frac{e^{-x_k/\beta} x_k^{Z-1}}{\beta^Z (Z-1)!},$$

approchée par son espérance sous la loi de conditionnelle de la variable latente $Z|X; \theta^{(0)}$ sachant les données et la valeur du paramètre.

① Etape Espérance (E): Calcul de

$$Q(\theta|\theta^{(0)}) = \mathbb{E}_{Z|X; \theta^{(0)}} \left\{ \ln [p(X, Z; \theta^{(0)})] \right\} = \sum_{i=1}^m \sum_{k=1}^n \ln [p(x_k, z_i; \theta^{(0)})] \mathbb{P}(Z = z_i | x_k; \theta^{(0)})$$

2 Etape Maximisation (M):

$$\theta^{(1)} = \begin{cases} \operatorname{argmax}_{\theta} Q(\theta|\theta^{(0)}) \\ \text{sous la contrainte} \end{cases} \quad \sum_{i=1}^m \pi_i = 1$$

On note que

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

Le Lagrangien associé à notre problème d'optimisation est donné par

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

puis $(\pi_1^{(1)}, \dots, \pi_1^{(n)}, \beta^{(1)}, \lambda)$ sont solutions du système d'équations suivant

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

Comme

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

alors on en déduit

$$\pi_i^{(1)} = \frac{\sum_{k=1}^n \mathbb{P}(Z_i = z_i | x_k; \theta^{(0)})}{n}, \quad i = 1, \dots, m, \quad \text{et} \quad \beta^{(1)} = \frac{\frac{1}{n} \sum_{k=1}^n x_k}{\sum_{i=1}^m \pi_i^{(1)} z_i}$$

On ré-itére les étapes E et M jusqu'à convergence de l'algorithme, on peut montrer que la procédure converge bien vers un maximum local.

Remark 7

En pratique, le calibrage consiste à considérer successivement $Z = 1$, $Z \in \{1, 2\}, \dots, Z \in \{1, 2, \dots, m\}$ puis à choisir le modèle qui minimise par exemple le BIC

$$BIC(i) = -2 \ln(p(X; \pi_1, \dots, \pi_i, \beta)) - (i + 1) \times \ln(n), \quad i = 1, \dots, m.$$

L'algorithme d'optimisation est lancé pour différentes valeurs initiales différentes de façon à rechercher un optimum global.

Les lois mélange de Erlang sont en fait un cas particuliers d'une famille de lois plus générales sur \mathbb{R}_+ appelées, lois phase type, voir Asmussen et al. [2].

Remark 8

Le principale défaut des lois mélange de Erlang est le suivant

$$\bar{F}_X(x) \sim e^{-x/\beta}, \text{ pour } x \rightarrow \infty$$

Une solution est de modéliser le ventre de la distribution par une Erlang mélange dans le cadre d'un modèle composite avec une queue de distributions modéliser par une Pareto, voir Reynkens et al. [9].

III. Méthodes d'approximation

1. Approximation normale et gamma

Definition 9 (Approximation normale)

Supposons que $N \sim \text{Pois}(\lambda)$, on note $\mu = \mathbb{E}(U)$ et $\sigma^2 = \mathbb{V}(U)$. On a donc

$$\mathbb{E}(X) = \lambda\mu \text{ et } \mathbb{V}(X) = \lambda(\mu^2 + \sigma^2).$$

L'approximation normale consiste à approcher la loi de

$$Z = \frac{X - \lambda\mu}{\sqrt{\lambda(\sigma^2 + \mu^2)}}$$

par une loi normale centrée réduite. La fonction de répartition de X est donnée par

$$\mathbb{P}(X \leq x) \approx \Phi\left(\frac{x - \lambda\mu}{\sqrt{\lambda(\sigma^2 + \mu^2)}}\right)$$

Cette approximation est d'autant meilleure que $\lambda \rightarrow \infty$.

Definition 10 (Approximation Gamma)

L'approximation gamma consiste à approcher X par une variable aléatoire de loi gamma translaturée $V \sim T - \Gamma(\alpha, \beta, w_0)$ définie par

$$V = w_0 + W$$

où $W \sim \Gamma(\alpha, \beta)$ en faisant correspondre les moments de X et V jusqu'à l'ordre 3. En notant $\mu_3(X) := \mathbb{E}\{[X - \mathbb{E}(X)]^3\}$, on a

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

Remark 9

Chacune des deux méthodes peut-être utilisées dans un contexte où des données sur la fréquence et les montants de sinistre sont disponibles. On remplace simplement $\mathbb{E}(X)$ et $\mathbb{V}(X)$ ou $(\lambda, \mu$ et $\sigma^2)$ par leur contrepartie empirique. A noter que le moment centré d'ordre 3 est une quantité très volatile et qu'un estimateur fiable nécessite une grande quantité de données.

2. Algorithme de Panjer

Definition 11 (Famille de Panjer)

La distribution d'une variable de comptage N appartient à la famille de Panjer si elle vérifie

$$p_N(n) = \left(a + \frac{b}{n}\right) p_N(n-1), \quad n = 1, 2, \dots$$

où $a < 1$ et $b \in \mathbb{R}$.

Proposition 6 (Sundt & Jewell [11])

Si N est dans la famille de Panjer alors

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

Definition 12 (Algorithme de Panjer [8])

Si U est une variable aléatoire de comptage alors X est aussi une variable aléatoire de comptage dont la loi de probabilité vérifie

$$p_X(k) = \begin{cases} G_N[p_U(0)] & \text{pour } k = 0 \\ \frac{1}{1 - ap_U(0)} \sum_{j=1}^k \left(a + b \frac{j}{k}\right) p_U(j) p_X(k-j) & \text{pour } k \geq 1. \end{cases}$$

La preuve sera faite en TD et l'application dans le devoir maison en python!

Remark 10 (Discrétisation de la loi des montants)

L'hypothèse des montants de sinistres à valeurs entières est un peu contraignante. L'algorithme de Panjer fournit une méthode d'approximation via une discrétisation de la loi des montants. Un schéma de discrétisation classique est le suivant

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

3. Fast Fourier Transform

La méthode *Fast Fourier Transform* est motivée par deux faits

- l'accessibilité des transformées (de Laplace ou de Fourier) de X
- le besoin d'optimiser les temps de calculs trop long dans le cadre de l'algorithme de Panjer

La procédure retrouve exactement la distribution de X si les montants de sinistres sont des variables aléatoires discrètes à support fini. X est alors discrètes à support sur \mathbb{N} que l'on va tronquer à l'ordre n (support sur $\{0, \dots, n\}$ sans perte de généralité) et sa transformée de Fourier est donnée par

$$M_X(is) = \sum_{k=0}^n p_X(k) e^{isk}, \quad s \in \mathbb{R}.$$

La méthode FFT permet d'évaluer les $p_X(k)$ sur la base de l'évaluation de la transformée de Fourier sur une grille de points définie par

$$s_k = \frac{2\pi k}{n+1}, \quad k = 0, \dots, n.$$

puis la résolution du système linéaire suivant

$$\begin{cases} M_X(is_0) = \sum_{i=1}^n e^{is_0 k} p_X(k) \\ \vdots \\ M_X(is_n) = \sum_{i=1}^n e^{is_n k} p_X(k) \end{cases}$$

ce qui se ré-écrit matriciellement

$$M_X^T = F p_X^T$$

où $M_X = (M_X(is_0), \dots, M_X(is_n))$, $p_X = (p_X(0), \dots, p_X(n))$ et

$$F = (e^{k is_l})_{k,l=0,\dots,n} = \begin{pmatrix} e^{0 \times is_0} & 1 & \dots & e^{0 \times is_n} \\ e^{is_0} & e^{is_1} & \dots & e^{is_n} \\ \vdots & \ddots & \ddots & \vdots \\ e^{nis_0} & e^{is_1} & \dots & e^{nis_n} \end{pmatrix}$$

La matrice F s'inverse aisément (matrice transconjugée ou adjointe) avec

$$F^{-1} = \frac{1}{n+1} (e^{-lis_k})_{k,l=0,\dots,n}.$$

Une comparaison entre l'algorithme de Panjer et la méthode FFT est effectuée dans le papier de Embrecht et Frei [3].

4. Développement polynomial

Des méthodes plus récentes s'affranchissent de la discrétisation des sinistres, on pourra consulter par exemple Goffard and Laub [5]. On note

$$\gamma_{(r,m)}(x) = \frac{e^{-x/m} x^{r-1}}{m^r \Gamma(r)}$$

la densité de la loi gamma et

$$Q_n(x) = (-1)^n \binom{n+r-1}{n-i}^{-1/2} \sum_{i=0}^n \binom{n+r-1}{n-i} \frac{(-x)^i}{i!} \quad n \geq 0,$$

Une suite de polynômes orthogonaux au sens où

$$\int Q_n(x) Q_m(x) \gamma_{(r,m)}(x) dx = \begin{cases} 1, & \text{si } n = m, \\ 0, & \text{sinon.} \end{cases}$$

$(Q_n)_{n \geq 0}$ forme un système orthonormal de fonctions pour l'espace des fonction de carré intégrable contre la mesure gamma, noté $\mathcal{L}_{\gamma_{(r,m)}}^2$.

Proposition 7

Si $f_X^+/\gamma_{(r,m)} \in \mathcal{L}_\gamma^2$ alors

$$f_X^+(x) \approx \sum_{i=0}^K p_i \gamma_{(r+i,m)}(x).$$

On a également

$$\bar{F}_X(x) \approx \sum_{i=0}^K p_i \bar{F}_{\gamma_{(r+i,m)}}(x),$$

et

$$\mathbb{E}[(X-a)_+] \approx \sum_{i=0}^K p_i [m(r+i) \bar{F}_{\gamma_{(r+i+1,m)}}(a) - a \bar{F}_{\gamma_{(r+i,m)}}(a)].$$

Les coefficients du développement sont données par

$$p_i = \sum_{k=1}^K q_k (-1)^{i+k} / [i!(k-i)!] \sqrt{k! \Gamma(k+r) / \Gamma(r)}, \quad i \leq K,$$

où $q_k = \mathbb{E}[Q_k(X)]$, $k \geq 0$.

preuve:

Comme $f_X^+/\gamma(r,m) \in \mathcal{L}_\gamma^2$ alors

$$\frac{f_X^+(x)}{\gamma(r,m)(x)} = \sum_{k=0}^{+\infty} q_k Q_k(x),$$

avec

$$q_k = \left\langle \frac{f_X^+(x)}{\gamma(r,m)}, Q_k \right\rangle = \int \frac{f_X^+(x)}{\gamma(r,m)(x)} Q_k(x) \gamma(r,m)(x) dx = \mathbb{E}[Q_k(X)]$$

La densité f_X^+ est alors approchée par simple troncature à l'ordre K

$$f_X^+(x) \approx \sum_{k=0}^K q_k Q_k(x) \gamma(r,m)(x),$$

En ré-injectant l'expression des polynômes Q_n dans l'équation et en ré-arrangeant les sommes, on trouve que

$$f_X^+(x) \approx \sum_{i=0}^K p_i \gamma(r+i,m)(x),$$

où

$$p_i = \sum_{k=1}^K q_k (-1)^{i+k} / [i!(k-i)!] \sqrt{k! \Gamma(k+r) / \Gamma(r)}, \quad i \leq K.$$

On intègre l'expression ci dessus pour obtenir les approximations de F_X et $\mathbb{E}(X-a)_+$

Remark 11

- 1 m est choisit pour s'assurer que L'hypothèse $\frac{f_X^+(x)}{\gamma_{(r,m)}(x)}$ est vérifié, r est ensuite obtenu par comparaison des moments d'ordres 1
- 2 L'hypothèse $\frac{f_X^+(x)}{\gamma_{(r,m)}(x)}$ restreint l'application de la méthode aux distributions ayant une fonction génératrice des moments bien définie. On peut s'en affranchir en approchant la version tiltée de f_X^+ donnée par

$$f_{\theta}(x) = \frac{e^{-\theta x} f_X^+(x)}{\int_0^{\infty} e^{-\theta x} f_X^+(x) dx}$$

Références bibliographiques I

Mes notes s'inspirent des notes de Stéphane Loisel [7]



Søren Asmussen, Pierre-Olivier Goffard, and Patrick J Laub.
Orthonormal polynomial expansions and lognormal sum densities.
Risk and Stochastics, 2019.



Søren Asmussen, Olle Nerman, and Marita Olsson.
Fitting phase-type distributions via the em algorithm.
Scandinavian Journal of Statistics, pages 419–441, 1996.



Paul Embrechts and Marco Frei.
Panjer recursion versus fft for compound distributions.
Mathematical Methods of Operations Research, 69(3):497–508, 2009.



Lawrence Fenton.
The sum of log-normal probability distributions in scatter transmission systems.
IRE Transactions on Communications Systems, 8(1):57–67, 1960.



Pierre-Olivier Goffard and Patrick J. Laub.
Orthogonal polynomial expansions to evaluate stop-loss premiums.
Journal of Computational and Applied Mathematics, 370:112648, may 2020.

Références bibliographiques II



Simon CK Lee and X Sheldon Lin.

Modeling and evaluating insurance losses via mixtures of erlang distributions.
North American Actuarial Journal, 14(1):107–130, 2010.



S. Loisel.

Notes de cours: Modélisation charge sinistre.



H. H. Panjer.

Recursive evaluation of a family of compound distributions.
ASTIN Bulletin, 12(1):22–26, 1981.



Tom Reynkens, Roel Verbelen, Jan Beirlant, and Katrien Antonio.

Modelling censored losses using splicing: A global fit strategy with mixed erlang and extreme value distributions.
Insurance: Mathematics and Economics, 77:65–77, 2017.



David PM Scollnik and Chenchen Sun.

Modeling with weibull-pareto models.
North American Actuarial Journal, 16(2):260–272, 2012.

Références bibliographiques III



Bjørn Sundt and William S Jewell.

Further results on recursive evaluation of compound distributions.

ASTIN Bulletin: The Journal of the IAA, 12(1):27–39, 1981.



Henk C Tijms and Henk C Tijms.

Stochastic models: an algorithmic approach, volume 994.

John Wiley & Sons Chichester, 1994.