Modélisation Charge Sinistre M2 Actuariat

Chapitre II: Modèle individuel

Pierre-Olivier Goffard

Université de Lyon 1
ISFA
pierre-olivier.goffard@univ-lyon1.fr

ISFA
September 16, 2021

I. Définition de X_{ind}

Le modèle individuel consiste à modéliser la charge totale de sinsitre sur une période en sommant les montants de sinistres pour chaque contrat. On définit ainsi la variable aléatoire

$$X_{\text{ind}} = \sum_{i=1}^{n} I_i W_i, \tag{1}$$

où

- n correspond au nombre de contrats en portefeuille
- $I_i \sim \text{Ber}(p_i)$, de loi
  $$\mathbb{P}(I_i = 1) = p_i \text{ et } \mathbb{P}(I_i = 1) = 1 - p_i = q_i$$
  indique si un sinistre a été reporté pour le contrat i ou non.
- W_i est une variable aléatoire positive égale au cumul des indemnisations générées par le contrat i.

Ce modèle permet de tenir compte de l'hétérogénéité du portefeuille et garantit une mesure très précise du risque. En effet, la probabilité d'observer un sinistre diffère pour chaque contrat, de même que la loi du montants des indemnisation. Deux problèmes se pose néanmoins

1. La loi de X_{ind} est inaccessible sauf par simulation
2. La calibration de modèle paramétriques pour les différentes variables aléatoires nécessite un historique de données important (voir prohibitif) et ce pour chaque contrat.

II. Moments de X_{ind}

Les moments et la fonction génératrices des moments de X_{ind} peuvent s'exprimer en fonction de $p_1, \dots, p_n$ et de l'espérance de $W_1, \dots, W_n$ à condition d'imposer une condition d'indépendance entre les contrats.

Proposition 1 (Variance et fonction génératrice des moments)

Posons $q_i = 1 - p_i$ et supposons que $I_1, \dots, I_n, W_1, \dots, W_n$ soient des ***v.a.****indépendantes, alors*

1. *L'espérance de X_{ind} est donnée par*

$$\mathbb{E}(X_{ind}) = \sum_{i=1}^{n} p_i \mathbb{E}(W_i).$$

2. *La variance de X_{ind} est donnée par*

$$\mathbb{V}(X_{ind}) = \sum_{i=1}^{n} \left[p_i q_i \mathbb{E}(W_i)^2 + p_i \mathbb{V}(W_i) \right]. \tag{2}$$

3. *La fonction génératrice des moments de X_{ind} est donnée par*

$$m_{X_{ind}}(s) = \mathbb{E}\left(e^{sX_{ind}}\right) = \prod_{i=1}^{n} \left[p_i m_{W_i}(s) + q_i \right], \text{ pour } s \in \mathbb{R}.$$

preuve:

1. $$\mathbb{E}(X_{\text{Ind}}) = \sum_{i=1}^{n} \mathbb{E}(I_i W_i) = \sum_{i=1}^{n} \mathbb{E}(I_i)\mathbb{E}(W_i) = \sum_{i=1}^{n} p_i \mathbb{E}(W_i)$$

2. $$\mathbb{V}(X_{\text{Ind}}) = \sum_{i=1}^{n} \mathbb{V}(I_i W_i) = \sum_{i=1}^{n} \{\mathbb{V}[\mathbb{E}(I_i W_i|I_i)] + \mathbb{E}[\mathbb{V}(I_i W_i|I_i)]\}. \tag{3}$$

   on a

   $$\mathbb{V}[\mathbb{E}(I_i W_i|I_i)] = \mathbb{V}[I_i \mathbb{E}(W_i|I_i)] = \mathbb{V}(I_i)\mathbb{E}(W_i)^2 = p_i q_i \mathbb{E}(W_i)^2 \tag{4}$$

   d'une part, et

   $$\mathbb{E}[\mathbb{V}(I_i W_i|I_i)] = \mathbb{V}(W_i I_i|I_i = 1)\mathbb{P}(I_i = 1) + \mathbb{V}(W_i I_i|I_i = 0)\mathbb{P}(I_i = 0) = p_i \mathbb{V}(W_i|I_i) = p_i \mathbb{V}(W_i) \tag{5}$$

   d'autre part. On obtient (2) en remplaçant (4) et (5) dans (3).

3. Examen :)

III. Distribution de X_{ind}

L'étude de la distribution de X_{ind} nécessite de faire la même hypothèse d'indépendance que précédemment. Il faut ensuite conditionné par rapport à toute les situations possibles par exemple que seul le premier contrat subisse un sinistre, ou que seuls le premier et le deuxième subissent une sinistre. Pour chaque cas il faut alors calculer des produits de convolution.
Par exemple si les trois premiers contrats sont sinistrés alors

$$X_{\text{ind}}\Big|\bigcap_{k=1}^{3} I_k = 1 \cap \bigcap_{l=4}^{n} I_l = 0 = W_1 + W_2 + W_3$$

La densité de la somme de deux variables aléatoires indépendante et positives $W_1 + W_2$ est donnée par

$$f_{W_1+W_2}(x) = (f_{W_1} * f_{W_2})(x) = \int_0^x f_{W_1}(y) f_{W_2}(x-y)\mathrm{d}y$$

puis

$$f_{W_1+W_2+W_3}(x) = ((f_{W_1} * f_{W_2}) * f_{W_3})(x) = \int_0^x \int_0^z f_{W_1}(y) f_{W_2}(z-y) f_{W_3}(x-z)\mathrm{d}y\mathrm{d}z$$

Les calculs deviennent rapidement prohibitifs... De Pril [1] proposa en 1989 une méthode de calcul récursive en faisant les deux approximation suivantes

- Regroupement des contrats en classe homogène $\{(i,j)\ ,\ i=1,\ldots,a \text{ et } j=1,\ldots,b\}$ contenant n_{ij}. Un contrat qui appartient à la classe (i,j)
  - a une probabilité d'être sinistré p_i
  - a un cumul d'indemnisation distribué comme W_j
- Les **v.a.** W_j, $j=1,\ldots,b$ sont discrètes de loi de probabilité

$$p_{W_j}(w)=\mathbb{P}(W_j=w),\ w=1,\ldots,m_j,\ \text{et } j=1,\ldots b,$$

où m_j correspond au montant maximal pour un contrat de la classe j.

On note que $n=\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}$.

	Classe montant des sinistres				
Classe occurence sinistre	**1**	…	**j**	…	b
1	n_{11}		n_{1j}		n_{1b}
2	n_{21}		n_{2j}		n_{2b}
⋮					
i	n_{i1}		n_{ij}		n_{ib}
⋮					
a	n_{a1}		n_{aj}		n_{ab}

La charge totale de sinistre est une variable aléatoire discrète dont La loi de probabilité

$$p_{X_{\text{ind}}}(x) = \mathbb{P}(X_{\text{ind}} = x), x = 0, 1, \ldots, m, \text{ où } m = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} m_i,$$

peut-être évaluée via la formule récursive suivante

Theorem 1 (Recurrence de De Pril)

La loi de probabilité de X_{Ind} est donnée par

$$p_{X_{ind}}(0) = \prod_{i=1}^{a} (q_i)^{n_i}$$

où $n_i = \sum_{j=1}^{b} n_{ij}$ et

$$x.p_{X_{ind}}(x) = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} v_{ij}(x),$$

avec

$$v_{ij}(x) = \frac{p_i}{q_i} \sum_{l=1}^{x \wedge m_j} p_{W_j}(l) \left[l p_{X_{Ind}}(x-l) - v_{ij}(x-l) \right], \ x = 1, \ldots m, \ et \ v_{ij}(0) = 0.$$

preuve: En TD!

La fonction génératrice des probabilité de X_{Ind} est donnée par

$$G_{X_{\text{Ind}}}(s) = \prod_{i=1}^{a} \prod_{i=1}^{b} \left[q_i + p_i G_{W_j}(s) \right]^{n_{ij}}. \tag{6}$$

L'évaluation en 0 donne directement

$$p_{X_{\text{ind}}}(0) = \prod_{i=1}^{a} (q_i)^{n_i}.$$

En prenant le log et en dérivant dans l'équation (6), il vient

$$G^{(1)}_{X_{\text{Ind}}}(s) = \sum_{i=1}^{a} \sum_{i=1}^{b} n_{ij} \frac{p_i G^{(1)}_{W_j}(s)}{q_i + p_i G_{W_j}(s)} G_{X_{\text{Ind}}}(s) := \sum_{i=1}^{a} \sum_{i=1}^{b} n_{ij} V_{ij}(s)$$

or on sait que par définition de la fonction génératrice des probabilités,

$$G_{X_{\text{Ind}}}(s) = \sum_{x=0}^{m} p_{X_{\text{Ind}}}(x) s^x, \text{ et } G^{(1)}_{X_{\text{Ind}}}(s) = \sum_{x=1}^{m} x p_{X_{\text{Ind}}}(x) s^{x-1}.$$

Cette remarque permet d'identifier les termes dans chacune des sommes définissant $G^{(1)}_{X_{\text{Ind}}}(s)$ avec

$$xp_{X_{\text{Ind}}}(x) = \frac{G^{(x)}_{X_{\text{Ind}}}(0)}{(x-1)!} = \sum_{i=1}^{a}\sum_{i=1}^{b} n_{ij}\frac{V^{(x-1)}_{ij}(0)}{(x-1)!},\ \text{avec } x = 1,\ldots,m.$$

On écrit ensuite

$$V_{ij}(s) = \frac{p_i G^{(1)}_{W_j}(s)}{q_i + p_i G_{W_j}(s)} G_{X_{\text{Ind}}}(s) \Leftrightarrow V_{ij}(s) = \frac{p_i}{q_i}\left(G^{(1)}_{W_j}(s) G_{X_{\text{Ind}}}(s) - G_{W_j}(s) V_{ij}(s)\right).$$

On applique ensuite la formule de Leibniz pour édriver $V_{ij}(s)$ à l'ordre $x-1$ et évaluer en 0 pour obtenir

$$\begin{aligned} v_{ij}(x) \quad &:= \quad \frac{V^{(x-1)}_{ij}(0)}{(x-1)!} = \frac{p_i}{q_i}\sum_{l=0}^{x-1}\Big[(l+1)p_{W_j}(l+1)p_{X_{\text{Ind}}}(x-1-l) - p_{W_j}(l)v_{ij}(x-l)\Big] \\ &= \quad \frac{p_i}{q_i}\sum_{l=1}^{x}\Big[lp_{W_j}(l)p_{X_{\text{Ind}}}(x-l) - p_{W_j}(l-1)v_{ij}(x-l+1)\Big] \qquad (7) \end{aligned}$$

On note que $p_{W_j}(0) = 0$ et on fixe $v_{ij}(0) = 0$ par convention ce qui permet d'écrire

$$v_{ij}(x) = \frac{p_i}{q_i} \sum_{l=1}^{x \wedge m_j} p_{W_j}(l) \left[l p_{X_{\mathbf{Ind}}}(x-l) - v_{ij}(x-l) \right].$$

On peut tout de même s'assurer de la cohérence de la formule sur un cas simple. Soit un portefeuille de contrats d'assurance décès avec versement d'un capital 1 en cas de décès. La probabilité p_i pour $i = 1, \ldots, a$ correspond à la probabilité de décès dans la classe i, qui peuvent s'interpréter comme des classes d'âge. De plus on a

$$\mathbb{P}(W_j = 1) = 1, \text{ pour tout } j = 1, \ldots b.$$

La probabilité $p_{X_{\text{ind}}}(0)$ correspond a aucun décès dans le portefeuille. La probabilité $p_{X_{\text{ind}}}(1)$ correspond à un unique décès. La probabilité d'un décès dans la classe i est donnée par

$$\binom{n_i}{1} q_i^{n_i - 1} p_i, \text{ pour } i = 1, \ldots, a.$$

soit la probabilité qu'une v.a. de loi binomial soit égale à 1. La probabilité qu'on ait un décès dans la classe i et pas de décès dans aucune autre classe (évènement noté A_i) est donnée par

$$n_i q_i^{n_i - 1} p_i \prod_{\substack{k \neq i \\ k = 1, \ldots, a}} q_k^{n_k}, \text{ pour } i = 1, \ldots, a.$$

On a enfin

$$p_{X_{\text{Ind}}}(1) = \mathbb{P}\left(\cup_{i=1}^{a} A_i\right) = \sum_{i=1}^{a} n_i q_i^{n_i - 1} p_i \prod_{\substack{k \neq i \\ k = 1, \ldots, a}} q_k^{n_k}.$$

En appliquant la formule, on retrouve aisément

$$v_{ij}(1) = \frac{p_i}{q_i} \sum_{l=1}^{1\wedge 1} (l p_{W_i}(l) p_{X_{\text{Ind}}}(1-l) - p_{W_i}(l-1) v_{ij}(1-l)) = \frac{p_i}{q_i} p_{X_{\text{Ind}}}(0)$$

puis

$$p_{X_{\text{Ind}}}(1) = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \frac{p_i}{q_i} p_{X_{\text{Ind}}}(0) = \sum_{i=1}^{a} n_i q_i^{n_i - 1} p_i \prod_{\substack{k \neq i \\ k=1,\ldots,a}} q_k^{n_k}.$$

Si le portefeuille de contrats est parfaitement homogène, c'est à dire que

- Les $I_1, \ldots, I_n$ sont **i.i.d.**$(p_1 = p_2 = \ldots p_n = p)$
- Les $W_1, \ldots, W_n$ sont **i.i.d.**et indépendants des $I_1, \ldots, I_n$

alors

$$X_{ind} = \sum_{i=1}^{N} W_i$$

où $N \sim \mathrm{Bin}(n, p)$.
Le modèle individuel devient alors un modèle collectif.

Références bibliographiques I

Mes notes s'inspirent des notes de Stéphane Loisel [2]

Nelson De Pril.
The aggregate claims distribution in the individual model with arbitrary positive claims.
ASTIN Bulletin, 19(1):9–24, 1989.

S. Loisel.
Notes de cours: Modélisation charge sinistre.