TD2 : THÉORIE DE L'ALGORITHME DE PANJER

TDs Modélisation Charge Sinistre – 2019-2020
Romain Gauchon

Le but de ce TD est d'étudier la loi de la variable aléatoire $S = \sum_{k=1}^{N} X_k$.

Considérons la variable aléatoire,

$$S = \sum_{k=1}^{N} X_k \tag{1}$$

avec

- $X_1, X_2, X_3, \dots$ des variables aléatoires iid positives de mêmes lois que X.
- N est une variable aléatoire entière dont la loi est donnée par

  $$p_N(n) = \mathbb{P}(N = n),\ n = 0, 1, 2, \dots.$$

  N est supposée indépendante des X_k.

Une interprétation possible de la variable aléatoire S est la suivante:

Supposons qu'on considère un portefeuille non-vie (par exemple un portefeuille automobile), étudié sur un an. La variable N compte le nombre total de sinistres pendant l'année. Chaque sinistre entraine une perte X_k pour l'assureur. La variable aléatoire S correspond donc à la perte totale de l'assureur sur une année. Une société d'assurance va étudier la variable aléatoire S afin de pouvoir mieux connaitre et contrôler son risque, permettant le calcul des SCR par exemple.

En théorie des files d'attente, N peut modéliser le nombre de clients coincé dans la file d'attente, et X_k leurs temps d'attentes respectifs. La variable aléatoire S représente le temps d'attente total à la caisse.

Panjer a proposé un algorithme récursif permettant de calculer la loi de S [1]. Cet algorithme permet de calculer de manière exacte

$$p_S(k) = \mathbb{P}(S = k),\ k = 0, 1, 2, \dots$$

lorsque les variables aléatoires X_k sont discrètes. Pour la suite, nous supposerons donc que les variables aléatoires X_k ont la même loi que la variable aléatoire discrète X, dont la loi est donnée par

$$p_X(k) = \mathbb{P}(X = k),\ \ k = 0, 1, 2, \dots .$$

Pour fonctionner, l'algorithme nécessite une hypothèse supplémentaire sur la loi de N. On dit que la variable aléatoire N appartient à la famille de Panjer si il existe $a, b \in \mathbb{R}$ tel que

$$p_N(k) = \left(a + \frac{b}{k}\right) p_N(k-1), k = 1, 2, \dots . \tag{2}$$

Les lois binomiales, binomiales négatives et de Poisson sont les seules lois appartenant la famille de Panjer.

Ce TD consiste à redémontrer l'algorithme de Panjer.

1. Supposons que N suit une loi binomiale négative Neg-Bin(α, p), avec $\alpha > 1$ et $p \in [0, 1]$. La loi de N est donnée par

$$p_N(k) = \binom{\alpha + k - 1}{k} (1-p)^\alpha p^k,\ \ k = 0, 1, \dots .$$

   La loi binomiale négative est souvent utilisée car elle est surdispersée (la variance est plus grande que l'espérance). La construction des variables binomiales négatives repose sur une somme de variables aléatoires géométriques. En effet, si $M_1, \dots, M_\alpha$ sont des variables aléatoires iid suivant une loi géométrique Geom(p):

$$\mathbb{P}(M_1 = k) = (1-p)p^k,\ \ k \in \mathbb{N}.$$

   Alors, la variable aléatoire $N = \sum_{l=1}^{\alpha} M_l$ suit une loi binomiale négative Neg-Bin(α, p).

   (a) Montrer que la fonction génératrice des moments M_1 vérifie

$$\Phi_{M_1}(t) = \frac{1-p}{1-pe^t}. \tag{3}$$

   En déduire l'espérance et la variance de M_1.

   (b) Calculer l'espérance, la variance et la fonction génératrice des moments de N.

   (c) Montrer que la fonction génératrice des probabilités $\mathcal{G}_N(t) = \mathbb{E}(t^N)$ de N vérifie

$$\mathcal{G}_N(t) = \left(\frac{1-p}{1-pt}\right)^\alpha$$

   (d) Montrer que N appartient à la famille de Panjer, en montrant qu'il existe a et b tel que l'équation (2) soit vérifiée. Exprimer a et b en fonction de α et p.

2. Montrer que
$$\mathbb{E}\left(X_1\Big|\sum_{k=1}^{n} X_k = j\right) = \frac{j}{n},\ j,n \in \mathbb{N}.$$
3. Montrer que
$$\mathbb{P}\left(X_1 = i\Big|\sum_{k=1}^{n} X_k = j\right) = \frac{p_X(i)p_X^{*(n-1)}(j-i)}{p_X^{*(n)}(j)},\ j \leq k \text{ and } n \geq 1,$$
avec $p_X^{*(n)}(j) = \mathbb{P}\left(\sum_{k=1}^{n} X_k = j\right)$.
4. Montrer que
$$\mathbb{P}(S=0) = \mathcal{G}_N\left[p_X(0)\right],$$
où $\mathcal{G}_N(t) = \mathbb{E}\left(t^N\right)$ désigne la fonction génératrice des probabilités de N.
5. L'objectif de cette question est de montrer que
$$p_S(j) = \sum_{i=0}^{j}\left(a + b\frac{i}{j}\right)p_X(i)p_S(j-i),\ \text{for } j \geq 1. \tag{4}$$
On suppose que N vérifie (2) et appartient à la famille de Panjer.
   (a) Montrer que, pour $j \geq 1$,
$$p_S(j) = a\sum_{n=1}^{+\infty} p_X^{*(n)}(j)p_N(n-1) + b\sum_{n=1}^{+\infty}\frac{1}{n}p_X^{*(n)}(j)p_N(n-1). \tag{5}$$
   (b) Montrer que
$$a\sum_{n=1}^{+\infty} p_X^{*(n)}(j)p_N(n-1) = a\sum_{i=0}^{j} p_X(i)p_S(j-i). \tag{6}$$
   Indice: Remarquer que $p_X^{(n)}(j) = \sum_{i=0}^{j} p_X(i)p_X^{(n-1)}(j-i)$, et jouer sur les interversions de sommes.
   (c) Montrer que
$$b\sum_{n=1}^{+\infty}\frac{1}{n}p_X^{*(n)}(j)p_N(n-1) = b\sum_{i=0}^{j}\frac{i}{j}p_X(i)p_S(j-i). \tag{7}$$
   Indice: Utiliser la question 2 pour remplacer $\frac{1}{n}$, puis utiliser la question 3 et le fait que
$$\mathbb{E}\left(X_1\Big|\sum_{k=1}^{n} X_k = j\right) = \sum_{i=0}^{j} i\mathbb{P}\left(X_1 = i\Big|\sum_{k=1}^{n} X_k = j\right).$$
Additionner les équations (6) et (7) montre (5).

References

[1] Harry H. Panjer. Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, 12(1):22–26, 006 1981.