

Théorie du risque

Nicolas Baradel

15 mai 2018

Table des matières

Introduction	3
I Préliminaires	3
1 Rappels sur les mesures de probabilité	3
2 Rappels sur la théorie de l'intégration	5
3 Les fonctions génératrices	5
II Modèles en assurance	7
4 Le modèle individuel	7
5 Le modèle collectif	8
5.1 Somme aléatoire et première formule de Wald	9
5.2 Principaux résultats	12
5.3 Discussions sur les résultats en Assurance	19
5.3.1 Formules de Wald	19
5.3.2 La loi de S et ses fonctionnelles	19
5.3.3 Le modèle individuel écrit par le modèle collectif	19
5.3.4 Incidence de la réassurance	21
III Comparaison abstraite des risques	24
6 Dominance stochastique d'ordre 1	24
7 Dominance stochastique d'ordre 2	26
IV Principe de prime et mesure de risque	26
8 Propriétés d'un principe de prime	27
9 Quelques principes de prime	27

10 Mesures de risque	28
V Théorie de la ruine	32
11 Processus de Poisson	32
12 Probabilité de ruine	35
VI Annexes	40
13 Preuve de l'exemple 5	40

Introduction

Tous les agents subissent des risques, qu'on assimile à des variables aléatoires. Qu'est ce qu'un risque ? Suis-je prêt à le conserver ou dois-je m'assurer ? De l'autre côté : suis-je prêt à l'acquérir ? À quel prix ?

Nous commencerons par rappeler quelques éléments et concepts de la théorie de la mesure et de l'intégration que nous serons amenés à utiliser, et introduirons les fonctions génératrices et caractéristiques. Nous enchaînerons ensuite sur les modèles d'aggrégation des risques en assurance. Le premier, individuel, aura pour but de regarder un à un les contrats où chaque risque a une probabilité de se réaliser avec un montant aléatoire. Le second, collectif, supposera un ensemble de risques homogènes où il y aura un nombre aléatoire de risques qui vont se produire. Nous continuerons par une étude abstraite des risques : existe-il des risques préférables à d'autres de manière générale ? Nous parlerons ensuite de principe de prime : quelles propriétés une application qui associe à un risque une prime devrait vérifier ? Enfin, nous terminerons sur la théorie de la ruine où il s'agit d'étudier la survie dans la temps d'un assureur qui subit continûment des sinistres par le modèle collectif et qui encaisse des primes continûment.

Première partie

Préliminaires

1 Rappels sur les mesures de probabilité

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace mesurable. Ω est l'espace sur lequel l'aléa se produit. \mathcal{A} est un ensemble de sous-ensembles de Ω sur lesquels est défini l'application \mathbb{P} . Cet ensemble \mathcal{A} est appelé *tribu*, contient Ω (on lui donnera une probabilité 1), contient les complémentaires (si $\mathbb{P}(A)$ existe alors $\mathbb{P}(A^c)$ existe et vaut $1 - \mathbb{P}(A)$), et les unions dénombrables (pour la définition de \mathbb{P}). L'application \mathbb{P} est appelée *mesure de probabilité*, qu'on appellera parfois plus simplement *probabilité* ou *loi*.

L'application

$$\begin{aligned} \mathbb{P} : \mathcal{A} &\rightarrow [0, 1] \\ A &\mapsto \mathbb{P}(A) \end{aligned}$$

vérifie par définition les deux propriétés

$$\begin{aligned} \mathbb{P}(\Omega) &= 1, \\ (A_n)_{n \in \mathbb{N}} \text{ disjoints} : \mathbb{P} \left(\bigcup_{n \in \mathbb{N}} A_n \right) &= \sum_{n \in \mathbb{N}} \mathbb{P}(A_n). \end{aligned}$$

Remarque : une *mesure* μ a la même définition qu'une *mesure de probabilité* sauf qu'on n'exige pas que sa masse totale soit 1, i.e. qu'on n'exige pas $\mu(\Omega) = 1$.

Lemme 1. Soit f une fonction mesurable positive et μ une mesure, alors

$$\begin{aligned} \mathbb{P} : \mathcal{A} &\rightarrow \mathbb{R}_+ \\ A &\mapsto \mathbb{P}(A) := \int_A f d\mu, \end{aligned}$$

est une mesure. C'est une mesure de probabilité si :

$$\mathbb{P}(\Omega) = \int_{\Omega} f d\mu = 1.$$

On appelle f la (une) **densité** par rapport à la mesure μ de la loi \mathbb{P} .

Lemme 2 (loi image). Soit X une variable aléatoire définie sur Ω et à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. La loi de X , notée \mathbb{P}_X et définie sur $\mathcal{B}(\mathbb{R})$, est :

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(\omega \in \Omega : X(\omega) \in B) = \mathbb{P}(X^{-1}(B)).$$

Définition 1 (mesure de Dirac). On appelle mesure de Dirac en $a \in \Omega$ la mesure

$$\delta_a : A \mapsto \delta_a(A) = \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{sinon.} \end{cases}$$

Définition 2 (loi discrète). Soit N une variable aléatoire. On dit que N est discrète si il existe $(k_n)_{n \in \mathbb{N}}$ tel que

$$\mathbb{P}_N \left(\bigcup_{n \in \mathbb{N}} \{k_n\} \right) = \sum_{n \in \mathbb{N}} \mathbb{P}_N(\{k_n\}) = 1.$$

Dans ce cas, la mesure est :

$$\mathbb{P}_N = \sum_{n \in \mathbb{N}} p_n \delta_{k_n}$$

avec $p_n := \mathbb{P}(N = k_n)$.

Exemple 1. La variable aléatoire constante $Y = c \in \mathbb{R}$ a pour loi

$$\mathbb{P}_Y = \delta_c.$$

La variable aléatoire Z de Bernoulli de paramètre $p \in [0, 1]$ a pour loi

$$\mathbb{P}_Z = p\delta_1 + (1 - p)\delta_0.$$

La variable aléatoire N de Poisson de paramètre $\lambda > 0$ a pour loi

$$\mathbb{P}_N = \sum_{n \in \mathbb{N}} \frac{e^{-\lambda} \lambda^n}{n!} \delta_n.$$

En assurance, ces lois représentent souvent le **nombre** de sinistres que l'assureur va subir pour un risque et sur une durée spécifique.

Définition 3 (loi absolument continue). Soit X une variable aléatoire. On dit que X est absolument continue (par rapport à la mesure de Lebesgue notée λ) si il existe une fonction f mesurable et positive telle que

$$\mathbb{P}_X(B) = \int_B f d\lambda.$$

Exemple 2. La variable aléatoire X de loi Gamma de paramètres $\alpha, \beta > 0$ a pour loi

$$\mathbb{P}_X(B) = \int_B \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbf{1}_{\mathbb{R}_+}(x) dx.$$

La variable aléatoire Y de loi log-normale de paramètre $\mu \in \mathbb{R}$ et $\sigma^2 > 0$ a pour loi

$$\mathbb{P}_Y(B) = \int_B \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(x) - \mu)^2}{2\sigma^2}\right) \mathbf{1}_{\mathbb{R}_+}(x) dx.$$

En assurance, ces lois représentent souvent le **coût** d'un sinistre que l'assureur va subir pour un risque spécifique.

Définition 4 (loi semi-discrète et semi-continue). *Soit X une variable aléatoire. On dit que X est semi-discrète et semi-continue si il existe $A \in \mathcal{B}(\mathbb{R})$ **discret** tel que*

$$\begin{aligned} \mathbb{P}_X(A) &\in]0, 1[, \\ X \mid \{X \in A^c\} &\text{ est } \mathbf{absolument\ continu}. \end{aligned}$$

Exemple 3. *Soit Z une variable aléatoire de Bernoulli de paramètre $p \in]0, 1[$ et X une variable aléatoire de loi exponentielle de paramètre $\lambda > 0$ indépendante de Z . Alors la variable aléatoire*

$$Y = ZX$$

est semi-discrète et semi-continue car pour $A = \{0\}$, on a $\mathbb{P}_Y(A) = 1 - p \in]0, 1[$ et $Y \mid \{Y > 0\}$ suit une loi exponentielle qui est absolument continue. Sa loi globale est

$$\mathbb{P}_Y(B) = (1 - p)\delta_0(B) + p \int_B \lambda e^{-\lambda x} \mathbb{1}_{[0, +\infty[}(x) dx.$$

La loi du coût total d'un assuré est souvent semi-discrète semi-continue. S'il n'a pas de sinistre, le coût total est 0, s'il a un sinistre, celui-ci est distribué continument.

2 Rappels sur la théorie de l'intégration

Ces deux théorèmes sont les plus importants pour les passages à la limite et nous les utiliserons souvent par la suite.

Théorème 1 (Convergence monotone). *Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables positives et croissantes ($\forall n \in \mathbb{N}, f_{n+1} \geq f_n$). On peut alors passer à la limite dans l'intégrale, i.e. :*

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n d\mu = \int_{\Omega} \lim_{n \rightarrow +\infty} f_n d\mu.$$

Théorème 2 (Convergence dominée). *Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables. On suppose que la suite converge vers une fonction f et qu'il existe $g \in L^1$ telle que, $\forall n \in \mathbb{N}, |f_n| \leq g$. Alors $f \in L^1$ et :*

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n d\mu = \int_{\Omega} \lim_{n \rightarrow +\infty} f_n d\mu = \int_{\Omega} f d\mu.$$

3 Les fonctions génératrices

Définition 5 (fonctions génératrices). *Soit $t \in \mathbb{R}$, la fonction génératrice des moments d'une variable aléatoire réelle X , notée g_X , est définie par*

$$\begin{aligned} g_X : \mathbb{R} &\rightarrow \overline{\mathbb{R}}_+ \\ t &\mapsto \mathbb{E}(e^{tX}). \end{aligned}$$

La fonction caractéristique de X , notée ϕ_X , est quant à elle définie par

$$\begin{aligned} \phi_X : \mathbb{R} &\rightarrow \mathbb{C} \\ t &\mapsto \mathbb{E}(e^{itX}). \end{aligned}$$

Soit $z \in \mathbb{C}$ tel que $|z| \leq 1$, la fonction génératrice des probabilités d'une variable aléatoire N à valeurs dans \mathbb{N} , notée P_N , est définie par

$$P_N : \mathbb{C} \rightarrow \mathbb{C} \\ z \mapsto \mathbb{E}(z^N).$$

Remarque 1. La fonction P_N peut-être défini sur un ensemble plus grand, cela dépend du rayon de convergence de la série, qui est au moins égal à 1.

Lemme 3. Si la fonction génératrice des moments d'une variable aléatoire réelle X est définie sur un intervalle ouvert contenant 0, tous ses moments existent.

Démonstration. Soit $I =]a, b[$ un intervalle contenant 0 sur lequel g_X est définie. Il contient un intervalle symétrique $[-t, t]$ pour $t > 0$ dans lequel g_X est finie, on a alors

$$g_{|X|}(t) = \mathbb{E}(e^{t|X|}) = \mathbb{E}(e^{tX} \mathbf{1}_{\{X \geq 0\}}) + \mathbb{E}(e^{-tX} \mathbf{1}_{\{X < 0\}}) \leq g_X(t) + g_X(-t) < +\infty.$$

On remarque que

$$e^{t|X|} \geq \sum_{k=0}^n \frac{t^k |X|^k}{k!}$$

puis, en appliquant l'opérateur espérance, on en déduit que les n premiers moments sont finis, et comme n est arbitraire, que tous les moments sont finis. \square

Théorème 3. Si la fonction génératrice des moments d'une variable aléatoire réelle X est définie sur un intervalle ouvert contenant 0, elle est dérivable dans cet intervalle, développable en série entière et sur cet intervalle,

$$g_X(t) = \sum_{n \in \mathbb{N}} \mathbb{E}(X^n) \frac{t^n}{n!},$$

en particulier, $g_X^{(n)}(0) = \mathbb{E}(X^n)$ et la suite des moments entiers caractérise la loi.

Démonstration. Admis. \square

Exemple 4. Soit $t \in \mathbb{R}$. Si $X \sim \mathcal{N}(0, 1)$,

$$g_X(t) = e^{\frac{t^2}{2}}, \\ \phi_X(t) = e^{-\frac{t^2}{2}}.$$

Puis si $X \sim \mathcal{N}(\mu, \sigma^2)$, avec $\mu \in \mathbb{R}$ et $\sigma > 0$, alors

$$g_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}, \\ \phi_X(t) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}.$$

Si $N \sim \mathcal{P}(\lambda)$,

$$g_N(t) = e^{\lambda(e^t - 1)}, \\ \phi_N(t) = e^{\lambda(e^{it} - 1)}, \\ P_N(t) = e^{\lambda(t - 1)}.$$

Si $X \sim G(\alpha, \beta)$ avec $\alpha, \beta > 0$,

$$g_X(t) = \left(\frac{\beta}{\beta - t} \right)^\alpha \mathbf{1}_{]-\infty, \beta[}(t) + \infty \mathbf{1}_{[\beta, +\infty[}(t), \\ \phi_X(t) = \left(\frac{\beta}{\beta - it} \right)^\alpha.$$

Remarque : la suite des moments seule caractérise la loi si la fonction génératrice des moments est définie au voisinage de 0, par le théorème précédent. Par contre, cette suite ne caractérise pas la loi si la fonction génératrice des moments n'est pas définie au voisinage de 0. Par exemple, c'est le cas de la loi log-normale.

Exemple 5. Soit X_0 une variable aléatoire de loi log-normale, de paramètres $\mu = 0, \sigma^2 = 1$, i.e. que X admet une densité par rapport à la mesure de Lebesgue, notée f_0 , définie par

$$f_0(x) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{\log(x)^2}{2}} \mathbf{1}_{]0,+\infty[}.$$

Soient $a \in [-1, 1]$ et X_a une variable aléatoire réelle dont la densité par rapport à la mesure de Lebesgue est définie par

$$f_a(x) = f_0(x) (1 + a \sin(2\pi \log(x))).$$

Alors,

1. La fonction f_a est bien une densité pour $a \in [-1, 1]$ et cette définition a un sens,
2. $\forall a \in [-1, 1], n \in \mathbb{N}, \mathbb{E}(X_a^n) = \mathbb{E}(X_0^n) = e^{\frac{n^2}{2}}$,
3. $\forall t > 0, g_{X_a}(t) = +\infty$.

L'exemple précédent montre que la suite des moments entiers n'est pas toujours suffisante pour caractériser la loi, mais dans ce cas, la fonction génératrice des moments n'est pas définie au voisinage de 0.

La preuve est reportée en Annexe.

Nous terminons cette section par un théorème très important que nous utiliserons souvent par la suite

Théorème 4. Soit X et Y deux variables aléatoires réelles. On a

$$\phi_X = \phi_Y \iff X \stackrel{loi}{=} Y.$$

Si g_X est finie sur un intervalle ouvert I contenant 0,

$$\forall t \in I, g_X(t) = g_Y(t) \iff X \stackrel{loi}{=} Y.$$

Si N et M sont deux variables aléatoires à valeurs dans \mathbb{N} ,

$$\forall z \in \mathbb{C}, |z| \leq 1, P_N(z) = P_M(z) \iff N \stackrel{loi}{=} M.$$

Deuxième partie

Modèles en assurance

4 Le modèle individuel

À chaque police d'assurance i correspond un risque individuel. L'assuré a une probabilité p_i de subir un sinistre. Si ce dernier a lieu, son coût est aléatoire et est noté X_i . Le modèle est dit « individuel » car le risque global se décompose comme la somme des risques individuels.

Proposition 1. Soient Z_i des variables aléatoires de loi de Bernoulli de paramètre p_i indépendantes et X_i des variables aléatoires indépendantes entre elles et des Z_i . On pose

$$S = \sum_{i=1}^n Z_i X_i.$$

- Si les X_i sont positives ou dans L^1 , on a

$$\mathbb{E}(S) = \sum_{i=1}^n p_i \mathbb{E}(X_i),$$

et $S \in L^1$ si $\forall 1 \leq i \leq n, X_i \in L^1$.

- De plus, si tous les $X_i \in L^2$,

$$\text{Var}(S) = \sum_{i=1}^n p_i \text{Var}(X_i) + \sum_{i=1}^n p_i(1 - p_i) \mathbb{E}(X_i)^2.$$

Démonstration. Pour l'espérance, par indépendance,

$$\mathbb{E}(S) = \mathbb{E} \left(\sum_{i=1}^n Z_i X_i \right) = \sum_{i=1}^n \mathbb{E}(Z_i X_i) = \sum_{i=1}^n p_i \mathbb{E}(X_i),$$

Pour la variance, par indépendance,

$$\text{Var}(S) = \text{Var} \left(\sum_{i=1}^n Z_i X_i \right) = \sum_{i=1}^n \text{Var}(Z_i X_i),$$

et nous avons

$$\begin{aligned} \text{Var}(Z_i X_i) &= \mathbb{E}(Z_i^2 X_i^2) - \mathbb{E}(Z_i X_i)^2 = \mathbb{E}(Z_i X_i^2) - p_i^2 \mathbb{E}(X_i)^2 = p_i \mathbb{E}(X_i^2) - p_i^2 \mathbb{E}(X_i)^2 \\ &= p_i \text{Var}(X_i) + p_i \mathbb{E}(X_i)^2 - p_i^2 \mathbb{E}(X_i)^2 \\ &= p_i \text{Var}(X_i) + p_i(1 - p_i) \mathbb{E}(X_i)^2. \end{aligned}$$

□

5 Le modèle collectif

Sur une branche d'assurance, dans une classe d'individus homogènes, on peut supposer que ces derniers auront des sinistres dont la loi est proche, idéalement ils seront indépendants et de même loi. Le défaut du précédent modèle est qu'il ne tient pas compte de la possibilité qu'un individu ait plusieurs accidents, tout du moins ce n'est pas naturel. Nous allons nous placer dans un modèle dit « collectif » où l'assureur aura un nombre aléatoire de sinistres, généralement noté N , et une séquence de sinistres aléatoires X_1, \dots, X_N . Les théorèmes de passage à la limite de la théorie de l'intégration (convergence monotone et convergence dominée) seront fréquemment utilisés.

5.1 Somme aléatoire et première formule de Wald

Nous commençons par un petit résultat qui nous sera utile et dont plusieurs démonstrations auront le même modèle.

Lemme 4. *Soit N une variable aléatoire à valeurs dans \mathbb{N} ou \mathbb{N}^* . On a alors l'égalité suivante :*

$$\mathbb{E}(N) = \sum_{n \in \mathbb{N}} \mathbb{P}(N > n) = \sum_{n \in \mathbb{N}^*} \mathbb{P}(N \geq n).$$

Démonstration. Remarquons dans un premier temps que :

$$N = \sum_{n \in \mathbb{N}} \mathbb{1}_{\{N > n\}} = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \mathbb{1}_{\{N > k\}}.$$

En effet, $\forall \omega \in \Omega$;

$$\sum_{n \in \mathbb{N}} \mathbb{1}_{\{N(\omega) > n\}} = \sum_{n=0}^{N(\omega)-1} \mathbb{1}_{\{N(\omega) > n\}} = N(\omega).$$

Ensuite, si on pose

$$Y_n = \sum_{k=0}^n \mathbb{1}_{\{N > k\}},$$

la suite (Y_n) est alors positive et croissante de limite N . Par le théorème de convergence monotone :

$$\begin{aligned} \mathbb{E}(N) &= \mathbb{E} \left(\lim_{n \rightarrow +\infty} \sum_{k=0}^n \mathbb{1}_{\{N > k\}} \right) = \lim_{n \rightarrow +\infty} \mathbb{E} \left(\sum_{k=0}^n \mathbb{1}_{\{N > k\}} \right) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \mathbb{P}(N > k) \\ &= \sum_{n=0}^{+\infty} \mathbb{P}(N > n) = \sum_{n=1}^{+\infty} \mathbb{P}(N \geq n). \end{aligned}$$

□

Définition 6. *Soient N une variable aléatoire à valeurs dans \mathbb{N} et $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles. La variable aléatoire S définie par :*

$$S = \sum_{n=1}^N X_n$$

est appelée « somme aléatoire ». C'est une somme finie (puisque N est finie), $\forall \omega \in \Omega$,

$$S(\omega) = \sum_{n=1}^{N(\omega)} X_n(\omega)$$

et $S(\omega) = 0$ si $N(\omega) = 0$. Le terme de « somme aléatoire » vient du fait que, le nombre de X_n (eux mêmes aléatoires) dans la somme définissant S est aléatoire et dépend de ω .

Dans la définition précédente, S est annoncée comme une variable aléatoire. Cette affirmation est la conséquence du lemme suivant.

Lemme 5. Soient N une variable aléatoire à valeurs dans \mathbb{N} et $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles. Nous avons l'égalité suivante :

$$S = \sum_{n=1}^N X_n = \sum_{n=1}^{+\infty} X_n \mathbf{1}_{\{N \geq n\}}.$$

Démonstration. Soit $\omega \in \Omega$, alors

$$\sum_{n=1}^{+\infty} X_n(\omega) \mathbf{1}_{\{N(\omega) \geq n\}} = \sum_{n=1}^{N(\omega)} X_n(\omega) \mathbf{1}_{\{N(\omega) \geq n\}} = \sum_{n=1}^{N(\omega)} X_n(\omega) = S(\omega)$$

S est la limite d'une suite de variables aléatoires, c'est donc une variable aléatoire. \square

Proposition 2 (Première formule de Wald). Soient N une variable aléatoire à valeurs dans \mathbb{N} et $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires **réelles positives de même espérance**. On suppose de plus que $N \perp\!\!\!\perp (X_n)_{n \in \mathbb{N}^*}$. On a alors

$$\mathbb{E}(S) = \mathbb{E}(N)\mathbb{E}(X)$$

Démonstration. $\forall n \in \mathbb{N}^*$, soit :

$$Y_n = \sum_{k=1}^n X_k \mathbf{1}_{\{N \geq k\}}$$

Par le lemme 5, $Y_n \rightarrow S$. Par le théorème de convergence monotone et le lemme 4 :

$$\begin{aligned} \mathbb{E}(S) &= \mathbb{E} \left(\lim_{n \rightarrow +\infty} \sum_{k=1}^n X_k \mathbf{1}_{\{N \geq k\}} \right) = \lim_{n \rightarrow +\infty} \mathbb{E} \left(\sum_{k=1}^n X_k \mathbf{1}_{\{N \geq k\}} \right) \\ &= \lim_{n \rightarrow +\infty} \sum_{k=1}^n \mathbb{E}(X_k) \mathbb{E}(\mathbf{1}_{\{N \geq k\}}) = \mathbb{E}(X) \lim_{n \rightarrow +\infty} \sum_{k=1}^n \mathbb{P}(N \geq k) \\ &= \mathbb{E}(X)\mathbb{E}(N). \end{aligned}$$

\square

Nous pouvons maintenant passer au cas réel L^1 .

Théorème 5 (Première formule de Wald). Soient N une variable aléatoire L^1 à valeurs dans \mathbb{N} et $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires **réelles L^1 de même espérance**. On suppose de plus que $N \perp\!\!\!\perp (X_n)_{n \in \mathbb{N}^*}$. Alors $S \in L^1$ et :

$$\mathbb{E}(S) = \mathbb{E}(N)\mathbb{E}(X)$$

Démonstration. Soit :

$$S^* = \sum_{n=1}^N |X_n|.$$

Alors $\mathbb{E}(S^*) = \mathbb{E}(N)\mathbb{E}(|X|)$ par la proposition 2 et $S^* \in L^1$.
 $\forall n \in \mathbb{N}^*$, on définit :

$$Y_n = \sum_{k=1}^n X_k \mathbb{1}_{\{N \geq k\}}$$

On a alors $Y_n \rightarrow S$ et $\forall n \in \mathbb{N}^*$, $|Y_n| \leq S^*$. La suite Y_n converge vers S et est dominée par $S^* \in L^1$. Par le théorème de convergence dominée, $S \in L^1$ et :

$$\begin{aligned} \mathbb{E}(S) &= \mathbb{E} \left(\lim_{n \rightarrow +\infty} \sum_{k=1}^n X_k \mathbb{1}_{\{N \geq k\}} \right) = \lim_{n \rightarrow +\infty} \mathbb{E} \left(\sum_{k=1}^n X_k \mathbb{1}_{\{N \geq k\}} \right) \\ &= \lim_{n \rightarrow +\infty} \sum_{k=1}^n \mathbb{E}(X_k) \mathbb{E}(\mathbb{1}_{\{N \geq k\}}) = \mathbb{E}(X) \lim_{n \rightarrow +\infty} \sum_{k=1}^n \mathbb{P}(N \geq k) \\ &= \mathbb{E}(X) \mathbb{E}(N). \end{aligned}$$

□

Nous allons donner une autre manière de démontrer la formule qui nous sera également utile par la suite. Celle-ci, plutôt que de passer à la limite dans une somme « tronquée » passe par un système complet d'évènements profitant du caractère dénombrable des valeurs de N .

Commençons par redonner la définition de l'espérance conditionnellement à un évènement, même si cela n'est pas nécessaire pour la suite, cela offre une représentation intuitive du résultat suivant.

Définition 7. Soit A un élément de la tribu de probabilité non nulle et X une variable aléatoire **positive ou réelle et intégrable**. On définit l'espérance de X sachant l'évènement A par

$$\mathbb{E}(X|A) = \frac{1}{\mathbb{P}(A)} \int \mathbb{1}_A X d\mathbb{P} = \frac{1}{\mathbb{P}(A)} \mathbb{E}(\mathbb{1}_A X).$$

Cette quantité est un réel, on prendra garde à ne pas confondre avec l'espérance conditionnelle (où A serait une tribu).

Lemme 6. Soit (A_n) une suite d'éléments disjoints de la tribu de réunion Ω dont aucun A_n n'est négligeable (on appelle cela un système complet d'évènements). Soit X une variable aléatoire positive ou réelle et intégrable. Alors

$$\mathbb{E}(X) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n) \mathbb{E}(X|A_n) = \sum_{n \in \mathbb{N}} \int \mathbb{1}_{A_n} X d\mathbb{P} = \sum_{n \in \mathbb{N}} \mathbb{E}(\mathbb{1}_{A_n} X)$$

Démonstration. Pour le cas X positive, on pose $B_m = \cup_{n=0}^m A_n$ puis $Y_m = \mathbb{1}_{B_m} X$. On remarque que les Y_m sont positives, croissantes et de limite X . On a, par convergence monotone,

$$\mathbb{E}(X) = \mathbb{E} \left(\lim_{m \rightarrow +\infty} \mathbb{1}_{B_m} X \right) = \lim_{m \rightarrow +\infty} \mathbb{E}(\mathbb{1}_{B_m} X) = \lim_{m \rightarrow +\infty} \mathbb{E} \left(\sum_{n=0}^m \mathbb{1}_{A_n} X \right) = \sum_{n \in \mathbb{N}} \mathbb{E}(\mathbb{1}_{A_n} X).$$

L'autre partie des égalités se déduit par la définition de $\mathbb{E}(X|A_n)$. Pour le cas réel et $X \in L^1$, on remarque que $|Y_m| \leq |X|$ puis on obtient le résultat par convergence dominée de la même manière. □

Le résultat suivant permet d'offrir un autre schéma de preuve pour les calculs des moments de S .

Proposition 3. Soient N une variable aléatoire à valeurs dans \mathbb{N} et $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires **réelles**. On suppose de plus que $N \perp\!\!\!\perp (X_n)_{n \in \mathbb{N}^*}$. Soit g une fonction mesurable telle que $g(S) \geq 0$ p.s. ou telle que $g(S) \in L^1$. Alors

$$\mathbb{E}(g(S)) = \sum_{n \in \mathbb{N}} \mathbb{P}(N = n) \mathbb{E} \left(g \left(\sum_{k=1}^n X_k \right) \right).$$

Démonstration. On utilise le lemme 6 avec $A_n = \{N = n\}$,

$$\begin{aligned} \mathbb{E}(g(S)) &= \sum_{n \in \mathbb{N}} \mathbb{E}(\mathbf{1}_{\{N=n\}} g(S)) \\ &= \sum_{n \in \mathbb{N}} \mathbb{E} \left(\mathbf{1}_{\{N=n\}} g \left(\sum_{k=1}^n X_k \right) \right) \\ &= \sum_{n \in \mathbb{N}} \mathbb{E}(\mathbf{1}_{\{N=n\}}) \mathbb{E} \left(g \left(\sum_{k=1}^n X_k \right) \right) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(N = n) \mathbb{E} \left(g \left(\sum_{k=1}^n X_k \right) \right). \end{aligned}$$

□

Corollaire 1. La proposition 2 et le théorème 5 peuvent se démontrer avec la proposition 3.

Démonstration. On prend g l'identité, i.e. $g : x \mapsto x$, et on a :

$$\begin{aligned} \mathbb{E}(S) &= \sum_{n \in \mathbb{N}} \mathbb{P}(N = n) \mathbb{E} \left(\sum_{k=1}^n X_k \right) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(N = n) n \mathbb{E}(X) \\ &= \mathbb{E}(N) \mathbb{E}(X). \end{aligned}$$

□

5.2 Principaux résultats

Théorème 6 (Formules de Wald). Soient N une variable aléatoire \mathbf{L}^2 à valeurs dans \mathbb{N} et $(X_n)_{n \in \mathbb{N}^*}$ une suite i.i.d. de variables aléatoires réelles \mathbf{L}^2 . On suppose de plus que $N \perp\!\!\!\perp (X_n)_{n \in \mathbb{N}^*}$. Alors $S \in L^2$ et :

- $\mathbb{E}(S) = \mathbb{E}(N) \mathbb{E}(X)$,
- $\text{Var}(S) = \mathbb{E}(N) \text{Var}(X) + \text{Var}(N) \mathbb{E}(X)^2$.

Démonstration. Le premier point est un cas particulier du théorème 5. On a

$$S^2 = \sum_{i=1}^N \sum_{j=1}^N X_i X_j$$

On pose $Z_k := X_k X_k$ et $Z'_\ell = X_i X_j$ pour $i \neq j$. S^2 peut se réécrire :

$$S^2 = \sum_{k=1}^N Z_k + \sum_{\ell=1}^{N^2-N} Z'_\ell.$$

Nous avons pour tout k, ℓ , $\mathbb{E}(Z_k) = \mathbb{E}(X^2)$ et $\mathbb{E}(Z'_\ell) = \mathbb{E}(X)^2$. Par application du théorème 5, $S^2 \in L^1$, i.e. $S \in L^2$ et

$$\mathbb{E}(S^2) = \mathbb{E}(N)\mathbb{E}(X^2) + (\mathbb{E}(N^2) - \mathbb{E}(N)) \mathbb{E}(X)^2.$$

Enfin,

$$\begin{aligned} \text{Var}(S) &= \mathbb{E}(S^2) - \mathbb{E}(S)^2 = \mathbb{E}(N)\mathbb{E}(X^2) + (\mathbb{E}(N^2) - \mathbb{E}(N)) \mathbb{E}(X)^2 - \mathbb{E}(N)^2 \mathbb{E}(X)^2 \\ &= \mathbb{E}(N) (\mathbb{E}(X^2) - \mathbb{E}(X)^2) + (\mathbb{E}(N^2) - \mathbb{E}(N)^2) \mathbb{E}(X)^2 \\ &= \mathbb{E}(N) \text{Var}(X) + \text{Var}(N) \mathbb{E}(X)^2. \end{aligned}$$

□

Exemple 6. Soient N une variable aléatoire de loi $\mathcal{P}(\lambda)$ et $(X_n)_{n \in \mathbb{N}^*}$ une suite i.i.d. de variables aléatoires réelles de L^2 . Les formules de Wald s'écrivent simplement

- $\mathbb{E}(S) = \lambda \mathbb{E}(X)$,
- $\text{Var}(S) = \lambda \text{Var}(X) + \lambda \mathbb{E}(X)^2 = \lambda (\text{Var}(X) + \mathbb{E}(X)^2) = \lambda \mathbb{E}(X^2)$.

Théorème 7. Soient N une variable aléatoire à valeurs dans \mathbb{N} et $(X_n)_{n \in \mathbb{N}^*}$ une suite i.i.d. de variables aléatoires **réelles**. Soit $t \in \mathbb{R}$, on note g la fonction génératrice des moments, i.e. $g_S(t) := \mathbb{E}(e^{tS})$. Si $g_X(t) < +\infty$ (sauf si $N = 0$) et si $g_N(\log[g_X(t)]) < +\infty$ on a

$$g_S(t) = g_N(\log[g_X(t)])$$

sinon $g_S(t) = +\infty$.

Démonstration. On applique la proposition 3 avec, $g : x \mapsto e^{tx}$ où $t \in \mathbb{R}$ est fixé. On a

$$\begin{aligned} \mathbb{E}(e^{tS}) &= \sum_{n \in \mathbb{N}} \mathbb{P}(N = n) \mathbb{E} \left(e^{t(\sum_{k=1}^n X_k)} \right) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(N = n) \prod_{k=1}^n \mathbb{E}(e^{tX_k}) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(N = n) g_X(t)^n \\ &= \mathbb{E}(g_X(t)^N) = \mathbb{E}(e^{N \log(g_X(t))}) \\ &= g_N(\log[g_X(t)]). \end{aligned}$$

□

Théorème 8 (fonction caractéristique). Soient N une variable aléatoire à valeurs dans \mathbb{N} et $(X_n)_{n \in \mathbb{N}^*}$ une suite i.i.d. de variables aléatoires **réelles**. Soit $t \in \mathbb{R}$, on note ϕ la fonction caractéristique, i.e. $\phi_S(t) := \mathbb{E}(e^{itS})$. On a

$$\phi_S(t) = P_N(\phi_X(t)).$$

Démonstration. Ce résultat se démontre de la même manière que pour la fonction génératrice des moments en remarquant que le module de toute fonction caractéristique est majoré par 1, ce qui assure que $g(S) \in L^1$. \square

Proposition 4. Soient S_1, \dots, S_n des sommes aléatoires définies par

$$S_i = \sum_{k=1}^{N_i} X_k^i$$

où toutes les variables aléatoires impliquées sont indépendantes entre elles. Pour tout i, k , on note P_{N_i} la fonction génératrice des probabilités de N_i et \mathbb{P}_X la loi **identique** de tous les X_k^i . Alors la variable aléatoire

$$S := \sum_{i=1}^n S_i$$

est une somme aléatoire et a la représentation en loi

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

où les X_k ont la loi \mathbb{P}_X et nous avons l'égalité en loi $N = N_1 + \dots + N_n$.

Démonstration. Comme les N_i sont indépendants, nous avons $P_{\sum_{i=1}^n N_i} = \prod_{i=1}^n P_{N_i}$. Soit $t \in \mathbb{R}$, il vient

$$\phi_S(t) = \phi_{\sum_{i=1}^n S_i}(t) = \prod_{i=1}^n \phi_{S_i}(t) = \prod_{i=1}^n P_{N_i}(\phi_X(t)) = P_{\sum_{i=1}^n N_i}(\phi_X(t)) = P_N(\phi_X(t))$$

\square

Proposition 5 (loi mélange). Soient Y_1, \dots, Y_n des variables aléatoires de lois respectives $\mathbb{P}_{Y_1}, \dots, \mathbb{P}_{Y_n}$. On appelle « loi mélange » la variable aléatoire de la forme

$$X := \sum_{i=1}^n Z_i Y_i$$

où $Z = (Z_1, \dots, Z_n)$ est un vecteur indépendant des Y_i dont toutes les composantes sont nulles sauf une qui vaut 1 avec $p_i := \mathbb{P}(Z_i = 1)$ et $\sum_{i=1}^n p_i = 1$. Alors

$$\mathbb{P}_X = \sum_{i=1}^n p_i \mathbb{P}_{Y_i}.$$

Démonstration. Soit A un ensemble de la tribu. Alors

$$\begin{aligned}
\mathbb{P}(X \in A) &= \mathbb{P}\left(\sum_{i=1}^n Z_i Y_i \in A\right) \\
&= \mathbb{P}\left(\left\{\sum_{i=1}^n Z_i Y_i \in A\right\} \cap \left\{\bigcup_{j=1}^n \{Z_j = 1\}\right\}\right) \\
&= \mathbb{P}\left(\bigcup_{j=1}^n \left\{\left\{\sum_{i=1}^n Z_i Y_i \in A\right\} \cap \{Z_j = 1\}\right\}\right) \\
&= \sum_{j=1}^n \mathbb{P}(\{Y_j \in A\} \cap \{Z_j = 1\}) \\
&= \sum_{j=1}^n p_j \mathbb{P}_{Y_j}(A).
\end{aligned}$$

□

Corollaire 2. On a

$$F_X = \sum_{i=1}^n p_i F_{Y_i}.$$

Démonstration. Il suffit d'appliquer la proposition précédente avec $A =]-\infty, t]$. □

Lemme 7 (premiers moments de la loi mélange). *Soit X une loi mélange définie comme dans la proposition 5 où toutes les variables sont dans L^2 . Nous avons alors*

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^n p_i \mathbb{E}(Y_i), \\
\mathbb{E}(X^2) &= \sum_{i=1}^n p_i \mathbb{E}(Y_i^2), \\
\text{Var}(X) &= \sum_{i=1}^n p_i \text{Var}(Y_i) + \sum_{i=1}^n p_i (\mathbb{E}(Y_i) - \mathbb{E}(X))^2.
\end{aligned}$$

La dernière égalité peut s'interpréter comme la somme d'une variance intra-classe et d'une variance inter-classe.

Démonstration. Pour l'espérance,

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^n Z_i Y_i\right) = \sum_{i=1}^n \mathbb{E}(Z_i Y_i) = \sum_{i=1}^n \mathbb{E}(Z_i) \mathbb{E}(Y_i) = \sum_{i=1}^n p_i \mathbb{E}(Y_i).$$

Pour les moments d'ordre 2,

$$\mathbb{E}(X^2) = \mathbb{E}\left(\sum_{i=1}^n Z_i^2 Y_i^2\right) = \mathbb{E}\left(\sum_{i=1}^n Z_i Y_i^2\right) = \sum_{i=1}^n p_i \mathbb{E}(Y_i^2).$$

Puis,

$$\begin{aligned}
\text{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\
&= \sum_{i=1}^n p_i \mathbb{E}(Y_i^2) - \mathbb{E}(X)^2 \\
&= \sum_{i=1}^n p_i \text{Var}(Y_i) + \sum_{i=1}^n p_i \mathbb{E}(Y_i)^2 - \mathbb{E}(X)^2 \\
&= \sum_{i=1}^n p_i \text{Var}(Y_i) + \sum_{i=1}^n p_i (\mathbb{E}(Y_i) - \mathbb{E}(X))^2.
\end{aligned}$$

Pour obtenir la dernière égalité, il suffit de remarquer que

$$\sum_{i=1}^n p_i (\mathbb{E}(Y_i) - \mathbb{E}(X))^2 = \sum_{i=1}^n p_i \mathbb{E}(Y_i)^2 - 2 \sum_{i=1}^n p_i \mathbb{E}(Y_i) \mathbb{E}(X) + \mathbb{E}(X)^2 = \sum_{i=1}^n p_i \mathbb{E}(Y_i)^2 - \mathbb{E}(X)^2.$$

□

Lemme 8 (fonction caractéristique de la loi mélange). *On reprend les variables aléatoires définies dans la proposition 5. On a*

$$\phi_X = \sum_{i=1}^n p_i \phi_{Y_i}.$$

Démonstration. Soit $t \in \mathbb{R}$,

$$\begin{aligned}
\phi_X(t) &= \mathbb{E} \left(e^{it \sum_{j=1}^n Z_j Y_j} \right) \\
&= \int \int e^{it \sum_{j=1}^n z_j y_j} d\mathbb{P}_Z(z) d\mathbb{P}_Y(y) \\
&= \int \sum_{j=1}^n p_j e^{it y_j} d\mathbb{P}_Y(y) \\
&= \sum_{j=1}^n \int p_j e^{it y_j} d\mathbb{P}_Y(y) \\
&= \sum_{j=1}^n \int p_j e^{it y_j} d\mathbb{P}_{Y_j}(y) \\
&= \sum_{j=1}^n p_j \phi_{Y_j}(t).
\end{aligned}$$

□

Théorème 9. *Soient S_1, \dots, S_n des sommes aléatoires définies par*

$$S_i = \sum_{k=1}^{N_i} X_k^i$$

où toutes les variables aléatoires impliquées sont indépendantes entre elles et, pour tout i, k , $N_i \sim \mathcal{P}(\lambda_i)$ et. On note \mathbb{P}_{X^i} la loi identique des $X_k^i, k \geq 1$. On pose $\lambda := \sum_{i=1}^n \lambda_i$. Alors la variable aléatoire

$$S := \sum_{i=1}^n S_i$$

est une somme aléatoire dont la loi est la même que

$$S' \stackrel{\text{loi}}{=} \sum_{k=1}^M Y_k,$$

où $M \sim \mathcal{P}(\lambda)$ et les Y_k suivent la loi mélange $\sum_{i=1}^n \frac{\lambda_i}{\lambda} \mathbb{P}_{X^i}$.

Démonstration.

$$\begin{aligned} \phi_S(t) &= \phi_{\sum_{i=1}^n S_i}(t) = \prod_{i=1}^n \phi_{S_i}(t) \\ &= e^{\sum_{i=1}^n \lambda_i (\phi_{X^i}(t) - 1)} \\ &= e^{\lambda [\sum_{i=1}^n \frac{\lambda_i}{\lambda} \phi_{X^i}(t) - 1]} \\ &= e^{\lambda (\phi_Y(t) - 1)} \\ &= P_M(\phi_Y(t)) \\ &= \phi_{S'}(t). \end{aligned}$$

□

Définition 8 (produit de convolution). Soient \mathbb{P}_X et \mathbb{P}_Y la loi respective de deux variables aléatoires réelles X et Y indépendantes. La loi de la somme est appelée produit de convolution et est notée $\mathbb{P}_X * \mathbb{P}_Y$ et on a par définition, pour tout A de la tribu,

$$\mathbb{P}_X * \mathbb{P}_Y(A) := \mathbb{P}_{X+Y}(A) = \int \int \mathbf{1}_A(x+y) d\mathbb{P}_X(x) d\mathbb{P}_Y(y)$$

Proposition 6. Soient \mathbb{P}_X et \mathbb{P}_Y la loi respective de deux variables aléatoires réelles X et Y indépendantes. On suppose qu'elles sont toutes les deux absolument continues par rapport à la mesure de Lebesgue, notée λ , de densités respectives f_X et f_Y . Alors $X + Y$ est également absolument continue par rapport à λ et de densité h définie pour tout z par

$$h(z) = \int_{\mathbb{R}} f_X(t) f_Y(z-t) d\lambda(t) = \int_{\mathbb{R}} f_X(z-t) f_Y(t) d\lambda(t).$$

Démonstration. Soit A un élément de la tribu. On a

$$\begin{aligned} \mathbb{P}_{X+Y}(A) &= \int \int \mathbf{1}_A(x+y) d\mathbb{P}_X(x) d\mathbb{P}_Y(y) \\ &= \int \int \mathbf{1}_A(x+y) f_X(x) f_Y(y) d\lambda(x) d\lambda(y). \end{aligned}$$

On fait le changement de variable $(t, z) = (x, x+y)$ dont le jacobien est 1, nous avons

$$\begin{aligned} \mathbb{P}_{X+Y}(A) &= \int \int \mathbf{1}_A(z) f_X(t) f_Y(z-t) d\lambda(z) d\lambda(t) \\ &= \int_A \left(\int f_X(t) f_Y(z-t) d\lambda(t) \right) d\lambda(z). \end{aligned}$$

L'autre égalité s'obtient en inversant les rôles de X et de Y .

□

Ces relations permettent de définir par récurrence la loi de la somme de n variables aléatoires X_1, \dots, X_n . Dans le cas particulier où $\mathbb{P}_{X_1} = \dots = \mathbb{P}_{X_n} = \mathbb{P}_X$, on notera la loi de $X_1 + \dots + X_n$ par \mathbb{P}_X^{*n} . On notera $\mathbb{P}_X^{*0} := \delta_0$ la loi qui représente la somme d'aucune variable aléatoire.

Proposition 7. En notant $p_n := \mathbb{P}(N = n)$, la loi de S est :

$$\mathbb{P}_S = \sum_{n \in \mathbb{N}} p_n \mathbb{P}_X^{*n}.$$

Si $p_0 > 0$, cette loi a une masse ponctuelle en 0, quelle que soit la loi de X . Si X est absolument continue, \mathbb{P}_X^{*n} l'est aussi pour $n > 0$ et la loi de $S|\{N > 0\}$ est continue.

Démonstration. Soit A un élément de la tribu. On a alors

$$\begin{aligned} \mathbb{P}_S(A) &= \mathbb{P} \left(\sum_{k=1}^N X_k \in A \right) = \mathbb{P} \left(\left\{ \sum_{k=1}^N X_k \in A \right\} \cap \left\{ \bigcup_{n \in \mathbb{N}} \{N = n\} \right\} \right) \\ &= \mathbb{P} \left(\bigcup_{n \in \mathbb{N}} \left\{ \left\{ \sum_{k=1}^N X_k \in A \right\} \cap \{N = n\} \right\} \right) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P} \left(\left\{ \sum_{k=1}^n X_k \in A \right\} \cap \{N = n\} \right) \\ &= \sum_{n \in \mathbb{N}} p_n \mathbb{P}_X^{*n}(A). \end{aligned}$$

□

Exemple 7. Lorsque la loi de la somme des X_k reste dans la même famille, on peut connaître facilement \mathbb{P}^{*n} . C'est le cas si les $X_k \sim \mathcal{E}(\beta)$, alors $\mathbb{P}^{*n} = \text{Ga}(n, \beta)$. Si de plus $N \sim G_0(p)$, loi géométrique issue de 0, définie par $\mathbb{P}(N = n) = p(1-p)^n$, on a pour tout ensemble A de la tribu,

$$\mathbb{P}_S(A) = p\delta_0(A) + (1-p) \int_A \mathbf{1}_{\mathbb{R}^+}(x) p\beta e^{-p\beta x} d\lambda(x),$$

i.e. S a une probabilité p de valoir 0 et une probabilité $(1-p)$ de suivre une loi $\mathcal{E}(p\beta)$ (loi mélange).

Démonstration.

$$\begin{aligned} \mathbb{P}_S(A) &= \sum_{n \in \mathbb{N}} p_n \mathbb{P}_X^{*n}(A) \\ &= p\delta_0(A) + \sum_{n=1}^{+\infty} p(1-p)^n \int_A \mathbf{1}_{\mathbb{R}^+}(x) \frac{\beta^n}{(n-1)!} x^{n-1} e^{-\beta x} d\lambda(x) \\ &= p\delta_0(A) + p(1-p)\beta \int_A \sum_{n=1}^{+\infty} \mathbf{1}_{\mathbb{R}^+}(x) (1-p)^{(n-1)} \frac{\beta^{n-1}}{(n-1)!} x^{n-1} e^{-\beta x} d\lambda(x) \\ &= p\delta_0(A) + p(1-p)\beta \int_A \mathbf{1}_{\mathbb{R}^+}(x) \sum_{n=1}^{+\infty} \left[\frac{[(1-p)\beta x]^{n-1}}{(n-1)!} \right] e^{-\beta x} d\lambda(x) \\ &= p\delta_0(A) + p(1-p)\beta \int_A \mathbf{1}_{\mathbb{R}^+}(x) e^{(1-p)\beta x} e^{-\beta x} d\lambda(x) \\ &= p\delta_0(A) + (1-p) \int_A \mathbf{1}_{\mathbb{R}^+}(x) p\beta e^{-p\beta x} d\lambda(x). \end{aligned}$$

□

Comme nous arrivons à une loi finale que l'on peut reconnaître, et que la fonction génératrice des moments g_S existe dans un voisinage de 0, nous aurions pu la calculer et en déduire le résultat par ce moyen.

Il n'est souvent pas possible d'arriver à un résultat simple pour la loi de S .

5.3 Discussions sur les résultats en Assurance

5.3.1 Formules de Wald

Les formules de Wald énoncées permettent d'obtenir les deux premiers moments d'une somme aléatoire, éléments souvent important. Pour rappel, lorsque tout le monde est indépendant et L^2 ,

- $\mathbb{E}(S) = \mathbb{E}(N)\mathbb{E}(X)$,
- $\mathbb{E}(N)Var(X) + Var(N)\mathbb{E}(X)^2$.

La première formule est la même si $N = n$ constante, mais la seconde a un terme additionnel $Var(N)\mathbb{E}(X)^2$ qui va dépendance de la variance de N , c'est une valeur supplémentaire. Ces quantités n'auront jamais à être estimée par Monte Carlo lorsque les lois de N et des X_k seront connues. Dans ce cas particulier, elles permettent de vérifier que la simulation est à priori bien implémentée.

5.3.2 La loi de S et ses fonctionnelles

Nous avons obtenu l'expression de la loi de S en fonction des convolutions à tout ordre, qui s'écrit comme une loi mélange infinie. Cependant, cette formule est très lourde, en pratique il faut la tronquer, et s'il faut calculer numériquement toutes les convolutions (sans avoir une propriété où on connaît tout de suite la loi) et si λ n'est pas très faible, c'est inexploitable.

Il existe différentes méthodes pour l'approximer. L'une d'elles repose sur une discrétisation de la loi de S par l'algorithme de Panjer. Cet algorithme est obsolète aujourd'hui compte tenu de la puissance de calcul, mais peut avoir un intérêt pédagogique, il est traité en Annexe (*pas encore*). Une autre méthode consiste à approximer S par une loi normale en utilisant ses moments par ceux des formules de Wald. C'est une méthode très grossière et à proscrire. Avec la puissance de calcul d'aujourd'hui, il est parfois possible d'estimer précisément les quantités souhaitées, par exemple un quantile, en simulant S par Monte Carlo. L'algorithme est : Simuler (N_1, \dots, N_n) , puis simuler pour chacun d'entre eux (X_1, \dots, X_{N_i}) , puis calculer S_1, \dots, S_n et estimer la quantité souhaitée. Si la somme des X_i suit une loi connue, simuler immédiatement S_i après la simulation de N_i .

5.3.3 Le modèle individuel écrit par le modèle collectif

Nous nous plaçons dans L^2 . Dans le modèle individuel, chaque sinistre individuel était

$$S_i = Z_i X_i,$$

où $Z_i \sim \mathcal{B}(p_i)$ indépendante de X_i . La sinistralité totale est :

$$S = \sum_{i=1}^n S_i.$$

On peut écrire

$$S_i = \sum_{k=1}^{Z_i} X_k^i.$$

Nous avons exactement la même variable aléatoire avec la même loi. Si les X_i sont indépendants et de **même loi** et si tous les p_i sont **identiques** égaux à p , par la proposition 4, nous avons l'égalité en loi

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

où en loi $N = \sum_i Z_i$, i.e. $N \sim \mathcal{B}(n, p)$. Cela traduit des risques homogènes. Cependant, la richesse du modèle individuel était le caractère hétérogène des risques, que ce soit des p_i ou des X_k . Le théorème 9 donne un cadre permettant d'agréger des risques hétérogènes. On peut poser, pour chaque sinistre individuel

$$S_i^c = \sum_{k=1}^{N_i} X_k^i,$$

où $N_i \sim \mathcal{P}(p_i)$, alors en moyenne il y aura le même nombre de sinistres, mais chaque individu peut avoir plusieurs sinistres. Cela permet de rendre plus cohérent une sinistralité individuelle automobile par exemple, mais moins le cas d'une assurance en cas de décès, ce dernier ne peut être qu'unique. En définissant $S^c := \sum_{i=1}^n S_i^c$, par le théorème 9, nous avons l'égalité en loi

$$S^c = \sum_{k=1}^M Y_k,$$

où M est la loi de Poisson de paramètre $\lambda = \sum_{i=1}^n p_i$ et les Y_k sont les lois mélanges définies dans le théorème 9, i.e. en reprenant les notations du théorème,

$$\mathbb{P}_Y = \sum_{i=1}^n \frac{p_i}{\lambda} \mathbb{P}_{X_i}.$$

Nous pouvons calculer les deux premiers moments de S et de S^c . Pour S , le modèle individuel, nous avons par la proposition 1,

$$\begin{aligned} \mathbb{E}(S) &= \sum_{i=1}^n p_i \mathbb{E}(X_i), \\ \text{Var}(S) &= \sum_{i=1}^n p_i \text{Var}(X_i) + \sum_{i=1}^n p_i (1 - p_i) \mathbb{E}(X_i)^2. \end{aligned}$$

Pour le modèle collectif, par les formules de Wald,

$$\begin{aligned} \mathbb{E}(S^c) &= \lambda \mathbb{E}(Y), \\ \text{Var}(S^c) &= \lambda \text{Var}(Y) + \lambda \mathbb{E}(Y)^2 = \lambda \mathbb{E}(Y^2). \end{aligned}$$

Par le lemme 7, nous avons

$$\begin{aligned} \mathbb{E}(Y) &= \sum_{i=1}^n \frac{p_i}{\lambda} \mathbb{E}(X_i), \\ \mathbb{E}(Y^2) &= \sum_{i=1}^n \frac{p_i}{\lambda} \mathbb{E}(X_i^2). \end{aligned}$$

En remplaçant,

$$\begin{aligned}\mathbb{E}(S^c) &= \sum_{i=1}^n p_i \mathbb{E}(X_i), \\ \text{Var}(S^c) &= \sum_{i=1}^n p_i \mathbb{E}(X_i^2).\end{aligned}$$

On remarque que $\mathbb{E}(S) = \mathbb{E}(S^c)$. Pour la variance, on peut réécrire

$$\begin{aligned}\text{Var}(S) &= \sum_{i=1}^n p_i (\mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2) + \sum_{i=1}^n p_i (1 - p_i) \mathbb{E}(X_i)^2 \\ &= \sum_{i=1}^n p_i \mathbb{E}(X_i^2) - \sum_{i=1}^n p_i^2 \mathbb{E}(X_i)^2\end{aligned}$$

Ainsi $\text{Var}(S^c) > \text{Var}(S)$ si X_i n'est pas centré. La version collective est dite « plus prudente ».

5.3.4 Incidence de la réassurance

L'assureur subissant $S = \sum_{k=1}^N X_k$ peut choisir de se réassurer afin de transférer une partie de son risque.

Définition 9 (Traité quote-part). *Dans un traité quote-part, l'assureur conserve une part α des sinistres qu'il subit tandis que le réassureur paie quant à lui une part $1 - \alpha$. C'est à dire, l'assureur subit :*

$$S_\alpha := \alpha S = \sum_{k=1}^N \alpha X_k.$$

Le nombre de sinistres est identique, seule la loi du montant est modifiée par une homothétie. Nous allons maintenant voir un premier contrat de réassurance non proportionnelle.

Définition 10 (Traité « Excess-of-Loss »). *Dans un traité « Excess-of-Loss », l'assureur subit les sinistres nets $\min(X_k, P)$ tandis que le réassureur paie $\max(X_k - P, 0)$. On a bien $X_k = \min(X_k, P) + \max(X_k - P, 0)$.*

Remarque : la plupart du temps, le réassureur prend en charge un montant maximal L appelé « portée ». Ainsi, il ne prend à sa charge que $\min(\max(X_k - P, 0), L)$ et donc réassure la partie $[P, P + L]$ qui est une « tranche de réassurance ». P est appelé la *priorité* et L la *portée*.

Cela introduit deux nouveaux modèles collectifs,

$$S_a = \sum_{k=1}^N \min(X_k, P) \quad \text{et} \quad S_r = \sum_{k=1}^N \max(X_k - P, 0).$$

Dans le premier, certains sinistres valent exactement P (avec probabilité $1 - F_X(P)$), dans le second, certains valent exactement 0 (avec probabilité $F_X(P)$). Nous pouvons poser

$$\bar{N} := \sum_{k=1}^N \mathbb{1}_{\{X_k > P\}},$$

et **en loi**, $\bar{X}_k = (X_1 - P) | \{X_1 > P\}$.

Lemme 9. *Sous les notations précédentes, $\bar{N} \sim \mathcal{P}(\lambda(1 - F_X(P)))$ et, en posant*

$$\bar{S} = \sum_{k=1}^{\bar{N}} \bar{X}_k$$

on a, en loi, $\bar{S} = S_r$.

Démonstration. Remarquons que $\bar{N} = \sum_{k=1}^{\bar{N}} \mathbb{1}_{\{X_k > P\}}$ s'écrit comme un modèle collectif, nous allons utiliser la caractérisation par la fonction caractéristique. De plus,

$$\phi_{\mathbb{1}_{\{X_k > P\}}}(t) = F_X(P) + (1 - F_X(P))e^{it}.$$

Nous avons donc

$$\begin{aligned} \phi_{\bar{N}}(t) &= P_N \left(\phi_{\mathbb{1}_{\{X_k > P\}}}(t) \right) \\ &= e^{\lambda(F_X(P) + (1 - F_X(P))e^{it} - 1)} \\ &= e^{\lambda(1 - F_X(P))(e^{it} - 1)}, \end{aligned}$$

ainsi $\bar{N} \sim \mathcal{P}(\lambda(1 - F_X(P)))$. De plus, si $1 - F_X(P) > 0$ (sinon le résultat est trivial), en posant $(X - P)^+ := \max(X - P, 0)$,

$$\begin{aligned} \phi_{(X-P)^+}(t) &= \int e^{it \max(x-P, 0)} d\mathbb{P}_X(x) \\ \phi_{(X-P)^+}(t) &= \int \mathbb{1}_{\{x \leq P\}} d\mathbb{P}_X(x) + \int e^{it(x-P)} \mathbb{1}_{\{x > P\}} d\mathbb{P}_X(x) \\ \phi_{(X-P)^+}(t) &= F_X(P) + \int_P^{+\infty} e^{it(x-P)} d\mathbb{P}_X(x), \end{aligned}$$

et en utilisant les lois conditionnelles,

$$\begin{aligned} \phi_{\bar{X}}(t) &= \int \frac{e^{it \max(x-P, 0)} \mathbb{1}_{\{x > P\}}}{1 - F_X(P)} d\mathbb{P}_X(x) \\ \phi_{\bar{X}}(t) &= \frac{1}{1 - F_X(P)} \int_P^{+\infty} e^{it(x-P)} d\mathbb{P}_X(x) \\ \phi_{\bar{X}}(t) &= \frac{\phi_{(X-P)^+}(t) - F_X(P)}{1 - F_X(P)}. \end{aligned}$$

Enfin,

$$\begin{aligned} \phi_{\bar{S}}(t) &= P_{\bar{N}}(\phi_{\bar{X}}(t)) \\ &= e^{\lambda(1 - F_X(P))(\phi_{\bar{X}}(t) - 1)} \\ &= e^{\lambda(\phi_{(X-P)^+}(t) - F_X(P) - (1 - F_X(P)))} \\ &= e^{\lambda(\phi_{(X-P)^+}(t) - 1)} \\ &= P_N(\phi_{(X-P)^+}(t)) = \phi_{S_r}(t). \end{aligned}$$

□

On peut démontrer un résultat analogue avec N qui suit une loi binomiale négative de paramètre (r, p) . Dans ce cas, \bar{N} suit une loi binomiale négative de paramètre $\left(r, \frac{p}{p + (1-p)(1 - F_X(P))}\right)$. Nous terminons par une représentation de l'espérance du sinistre restant à payer pour l'assureur après réassurance.

Lemme 10. Soit X une variable aléatoire réelle **positive** et $P \in \mathbb{R}_+ \cup \{+\infty\}$. Alors,

$$\mathbb{E}(\min(X, P)) = \int_0^P (1 - F_X(x)) dx.$$

On remarque que dans le cas particulier $P = +\infty$ on a :

$$\mathbb{E}(X) = \int_0^{+\infty} (1 - F_X(x)) dx.$$

Démonstration.

$$\mathbb{E}(\min(X, P)) = \int_0^{+\infty} (x \mathbb{1}_{\{x \leq P\}} + P \mathbb{1}_{\{x > P\}}) d\mathbb{P}_X(x)$$

On remarque que, pour $x \geq 0$,

$$x \mathbb{1}_{\{x \leq P\}} + P \mathbb{1}_{\{x > P\}} = \int_0^x \mathbb{1}_{\{t \leq P\}} dt.$$

Comme tout est positif, nous allons pouvoir utiliser Fubini, on a :

$$\begin{aligned} \mathbb{E}(\min(X, P)) &= \int_0^{+\infty} \left(\int_0^x \mathbb{1}_{\{t \leq P\}} dt \right) d\mathbb{P}_X(x) = \int_0^{+\infty} \left(\int_0^{+\infty} \mathbb{1}_{\{t \leq x\}} \mathbb{1}_{\{t \leq P\}} dt \right) d\mathbb{P}_X(x) \\ &= \int_0^{+\infty} \mathbb{1}_{\{t \leq P\}} \left(\int_0^{+\infty} \mathbb{1}_{\{t \leq x\}} d\mathbb{P}_X(x) \right) dt = \int_0^{+\infty} \mathbb{1}_{\{t \leq P\}} (1 - F_X(t)) dt \\ &= \int_0^P (1 - F_X(t)) dt. \end{aligned}$$

□

Avant de terminer cette section, donnons une idée des propriétés que doit respecter un traité de réassurance.

Définition 11. Si S est la sinistralité, le réassureur remboursera $I(S)$ où I est définie sur \mathbb{R}_+ . On s'attend à ce qu'une telle fonction soit issue de 0, i.e. $I(0) = 0$, croissante, i.e. pour tout $s, h \geq 0$,

$$I(s + h) - I(s) \geq 0.$$

D'un autre côté, lorsque la charge augmente, le réassureur ne remboursera pas plus que la nouvelle charge, ceci se traduit par

$$I(s + h) - I(s) \leq h.$$

On remarque que si I admet des dérivées à gauche ou à droite notées $I'_g(s)$ et $I'_d(s)$, cela se traduit par $0 \leq I'_g(s) \leq 1$ et $0 \leq I'_d(s) \leq 1$.

On remarque que $0 \leq I(s + h) - I(s) \leq h$ et qu'on a aussi $0 \leq I(s) - I(s - h) \leq h$ dès que $s - h \geq 0$. En faisant tendre h vers 0, on observe que la fonction I est automatiquement continue.

Exemple 8. Le traité Quote-Part est représenté par la fonction $I : s \mapsto (1 - \alpha)s$ avec $\alpha \in [0, 1]$. Le traité « Stop-Loss » est représenté par $I : s \mapsto (s - P)^+$.

Troisième partie

Comparaison abstraite des risques

Si on dispose de deux variables aléatoires, X et Y , qui représentent chacune un risque monétaire, le choix de préférer supporter X ou Y dépendra des préférences de l'individu, par exemple, un assureur. Si l'on suit l'axiomatique de Von Neumann-Morgenstern, l'individu comparera son espérance d'utilité en utilisant sa fonction d'utilité personnelle. Toutefois si, avec probabilité un, $X = 1$ et $Y = 2$, et que l'individu préfère plus que moins, qui se traduit par une fonction d'utilité croissante, alors l'individu préférera subir X plutôt que Y . Et cela, sans même se soucier de l'aversion au risque. Si par contre, $X = 1/2$ et $Y \sim \mathcal{B}(1/2)$, on ne peut plus parler aussi aisément de plus ou de moins. On voit qu'on peut écrire $Y = X + \epsilon$ où ϵ est centrée et indépendante de X . Ainsi, Y est X mais « plus aléatoire » et pour un assureur qui a des préférences croissantes et averses au risque, se traduisant par une fonction d'utilité concave, on s'attend à ce que l'assureur préfère subir X , peu importe la spécification de sa fonction d'utilité.

La dominance stochastique d'ordre 1 répondra au premier exemple tandis que la dominance stochastique d'ordre 2 répondra au second exemple.

6 Dominance stochastique d'ordre 1

Définition 12. Soient X et Y deux variables aléatoires réelles. X domine stochastiquement à l'ordre 1 Y , qu'on écrit $X \geq_1 Y$, si, pour toute fonction **croissante**¹ v ,

$$\mathbb{E}(v(X)) \geq \mathbb{E}(v(Y))$$

pour lesquels les deux espérances ont du sens (mais peuvent éventuellement valoir $+\infty$ ou $-\infty$).

Remarque 2. Tout individu admettant une fonction d'utilité croissante préférera subir Y plutôt que X pour le principe de l'espérance d'utilité.

Démonstration. Soit u la fonction d'utilité croissante de l'individu. La fonction $z \mapsto v(z) = -u(-z)$ est croissante et

$$\mathbb{E}(v(X)) \geq \mathbb{E}(v(Y)) \Leftrightarrow \mathbb{E}(u(-X)) \leq \mathbb{E}(u(-Y)).$$

□

Proposition 8. La relation \geq_1 est un pré-ordre, i.e. elle est réflexive et transitive.

Démonstration. La réflexivité, i.e. que pour toute variable aléatoire, $X \geq_1 X$ est une conséquence immédiate de la définition. Si $X \geq_1 Y$ et $Y \geq_1 Z$ alors $X \geq_1 Z$ en appliquant la transitivité de \geq . □

Proposition 9. Si $X \geq_1 Y$ et $Y \geq_1 X$ alors X et Y ont la même loi.

Démonstration. Par hypothèse, en appliquant la définition, on a pour tout fonction mesurable croissante positive v ,

$$\mathbb{E}(v(X)) = \mathbb{E}(v(Y)).$$

En particulier, pour $t \in \mathbb{R}$, la fonction $v : z \mapsto \mathbb{1}_{]t, +\infty[}(z)$ est croissante et

$$\mathbb{P}(X > t) = \mathbb{P}(Y > t)$$

□

1. Une fonction croissante est mesurable

Compte tenu de la définition, on remarque que seule la loi de X et de Y importe, indépendamment de leurs dépendances. Le résultat suivant, associant les lois directement, aurait pu servir de définition.

Proposition 10. *Soient X et Y deux variables aléatoires réelles. Alors*

$$X \geq_1 Y \Leftrightarrow F_X \leq F_Y.$$

Démonstration. Si $X \geq_1 Y$, pour $t \in \mathbb{R}$, la fonction $v : z \mapsto \mathbb{1}_{]t, +\infty[}(z)$ est croissante et

$$\mathbb{P}(X > t) \geq \mathbb{P}(Y > t).$$

Réciproquement, supposons $v \geq 0$. Si $F_X \leq F_Y$, comme v est mesurable et positive, la suite étagée de la forme

$$v_n = \sum_{k=1}^{2^{2n+1}} \frac{k-1}{2^n} \mathbb{1}_{A_k^n}$$

où $A_k^n = v^{-1}(\left] \frac{k-1}{2^n}, \frac{k}{2^n} \right])$ pour $k \leq 2^{2n}$ et $A_{2^{2n+1}}^n = v^{-1}(]2^n, +\infty[)$ est positive et converge simplement vers v . Nous pouvons réécrire

$$v_n = \sum_{k=1}^{2^{2n+1}} \frac{1}{2^n} \sum_{j=k+1}^{2^{2n+1}} \mathbb{1}_{B_j^n}$$

avec

$$B_j := \bigcup_{k=j}^{2^{2n+1}} A_k^n = f^{-1}\left(\left] \frac{k-1}{2^n}, +\infty \right[\right).$$

Comme f est croissante, les B_j sont de la forme $]t_j^n, +\infty[$. Par hypothèse, $\mathbb{P}(X > t) \geq \mathbb{P}(Y > t)$ pour tout t , ainsi,

$$\mathbb{E}(v_n(X)) \geq \mathbb{E}(v_n(Y)).$$

Par le théorème de convergence monotone, nous avons le résultat. Pour le cas où v n'est pas considérée positive, on écrit $v = v^+ - v^-$. On a donc $\mathbb{E}(v^+(X)) \geq \mathbb{E}(v^+(Y))$. Pour v^- , on construit de la même manière, mais en remarquant que la fonction est décroissante,

$$v_n^- = \sum_{k=1}^{2^{2n+1}} \frac{1}{2^n} \sum_{j=k+1}^{2^{2n+1}} \mathbb{1}_{B_j^n},$$

ainsi cette fois les B_j sont de la forme $] - \infty, t_j^n]$. Puisque $\mathbb{P}(X \leq t) \leq \mathbb{P}(Y \leq t)$, on obtient

$$\mathbb{E}(v_n^-(X)) \leq \mathbb{E}(v_n^-(Y)).$$

De même, on passe à la limite par convergence monotone puis le résultat final se déduit en utilisant $v = v^+ - v^-$. \square

Proposition 11. *Si $X \geq Y$ p.s. alors $X \geq_1 Y$. Réciproquement, si $X \geq_1 Y$ alors on peut construire deux variables aléatoires X' et Y' dont les lois sont respectivement les mêmes que celles de X et de Y telles que $X' \geq Y'$ p.s.*

Démonstration. Pour la première partie, puisque $X \geq Y$, en appliquant une fonction croissante v et en utilisant la croissance de l'espérance, nous avons $\mathbb{E}(v(X)) \geq \mathbb{E}(v(Y))$. Réciproquement, on définit

$$F_X^{\leftarrow}(\alpha) := \inf \{x \in \mathbb{R} : F_X(x) \geq \alpha\}.$$

Puisque $X \geq_1 Y$, on a $F_X \leq F_Y \Rightarrow F_X^{\leftarrow} \geq F_Y^{\leftarrow}$. Soit $U \sim \mathcal{U}([0, 1])$, on pose $X' = F_X^{\leftarrow}(U)$ et $Y' = F_Y^{\leftarrow}(U)$. Alors X' et Y' ont respectivement la même loi que X et Y et on a $X' \geq Y'$. \square

7 Dominance stochastique d'ordre 2

Définition 13. Soient X et Y deux variables aléatoires réelles. X domine stochastiquement à l'ordre 2 Y , qu'on écrit $X \succeq_2 Y$, si, pour toute fonction **croissante** et **convexe** v ,

$$\mathbb{E}(v(X)) \geq \mathbb{E}(v(Y))$$

pour lesquels les deux espérances ont du sens (mais peuvent éventuellement valoir $+\infty$ ou $-\infty$).

Remarque 3. Soient X et Y deux variables aléatoires réelles. On a l'implication :

$$X \succeq_1 Y \implies X \succeq_2 Y.$$

Remarque 4. Tout individu admettant une fonction d'utilité croissante et concave préférera subir Y plutôt que X pour le principe de l'espérance d'utilité.

Démonstration. Soit u la fonction d'utilité croissante et concave de l'individu. La fonction $z \mapsto v(z) = -u(-z)$ est croissante et convexe et

$$\mathbb{E}(v(X)) \geq \mathbb{E}(v(Y)) \Leftrightarrow \mathbb{E}(u(-X)) \leq \mathbb{E}(u(-Y)).$$

□

Proposition 12. Soient X et Y deux variables aléatoires réelles. Alors

$$X \succeq_2 Y \Leftrightarrow \forall t \in \mathbb{R}, \mathbb{E}((X - t)^+) \geq \mathbb{E}((Y - t)^+).$$

Démonstration. La fonction $x \mapsto (x - t)^+$ est croissante et convexe ce qui donne un sens de l'équivalence. L'autre sens est admis. □

Proposition 13. Soient X et Y deux variables aléatoires réelles. Alors $X \succeq_2 Y$ équivaut à l'existence d'une variable aléatoire Z telle que $Y + Z = X$ en loi et $\mathbb{E}(Z|Y) \geq 0$ p.s.

Démonstration. Si il existe Z telle que définie dans l'énoncé, alors, pour tout fonction v croissante et convexe, en utilisant l'inégalité de Jensen pour les espérances conditionnelles,

$$\mathbb{E}(v(X)) = \mathbb{E}(v(Y + Z)) = \mathbb{E}[\mathbb{E}(v(Y + Z)) | Y] \geq \mathbb{E}(v(Y + \mathbb{E}(Z|Y))) \geq \mathbb{E}(v(Y)).$$

La réciproque est admise.

□

Proposition 14. Soient X et Y deux variables aléatoires réelles. Si $\mathbb{E}(X) \geq \mathbb{E}(Y)$ et s'il existe une constante $c \in \mathbb{R}$ telle que $F_Y \leq F_X$ sur $] -\infty, c[$ puis $F_Y \geq F_X$ sur $]c, +\infty[$, alors $X \succeq_2 Y$.

Démonstration. Admis. □

Quatrième partie

Principe de prime et mesure de risque

Lors que l'assurance est face à un risque X positif de mesure de probabilité \mathbb{P}_X , il propose à l'assuré de le prendre en charge contre une prime H . Intuitivement, $H \geq \mathbb{E}(X)$ et plus le risque associé sera élevé, plus H sera élevée. La variance est une mesure du risque, mais elle ne le caractérise pas, c'est pourquoi nous allons proposer différents « principes de prime ».

Définition 14 (Principe de prime). *Un principe de prime H est une fonctionnelle qui, à une variable aléatoire réelle associe un montant de prime non aléatoire, i.e.*

$$H : L^0 \rightarrow \bar{\mathbb{R}}_+ \\ X \mapsto H(X).$$

Si la prime est infinie on dit que le risque est inassurable.

8 Propriétés d'un principe de prime

Ci-dessous figurent quelques propriétés souhaitables pour un « bon » principe de prime.

- Profitabilité : $H(X) \geq \mathbb{E}(X)$,
- Inférieur à la perte maximale : $H(X) \leq \text{ess sup } X$,
- Invariant par translation : $H(X + a) = H(X) + a$,
- Homogène : $H(aX) = aH(X)$ ($a \geq 0$),
- Sous-additive : $H(X + Y) \leq H(X) + H(Y)$,
- Indépendance : Le risque ne dépend que de la loi de X ,
- Monotonie : $X \leq Y$ p.s. $\Rightarrow H(X) \leq H(Y)$.

Nous allons voir différents principes de prime que nous allons discuter.

9 Quelques principes de prime

Un petit lemme préliminaire nous servira à vérifier si certaines propriétés sont vérifiées ou non.

Lemme 11. *Soit X une variable aléatoire telle que $0 \leq X \leq \theta$ p.s. pour $\theta > 0$. Alors, en posant $\mu := \mathbb{E}(X)$,*

$$\text{Var}(X) \leq \theta\mu - \mu^2 \leq \frac{\theta^2}{4}.$$

La première borne supérieure est atteinte pour X vérifiant $\mathbb{P}(X = \theta) = \frac{\mu}{\theta} = 1 - \mathbb{P}(X = 0)$ et la seconde dans le cas particulier $\mu = \frac{\theta}{2}$, i.e. $\mathbb{P}(X = \theta) = \mathbb{P}(X = 0) = \frac{1}{2}$.

Démonstration. On a

$$\text{Var}(X) = \int_0^\theta x^2 d\mathbb{P}_X(x) - \mu^2.$$

Comme $0 \leq x \leq \theta$, on a $x^2 \leq x\theta$ d'où

$$\text{Var}(X) \leq \int_0^\theta x\theta d\mathbb{P}_X(x) - \mu^2 = \theta\mu - \mu^2.$$

Puis on vérifie qu'il s'agit bien de la borne supérieure en remarquant que si $\mathbb{P}(X = \theta) = \frac{\mu}{\theta} = 1 - \mathbb{P}(X = 0)$, on a $\mathbb{E}(X) = \mu$ et $\text{Var}(X) = \theta\mu - \mu^2$. Enfin, pour la seconde inégalité, on remarque que le maximum pour $\mu \in [0, \theta]$ de $\text{Var}(X)$ est atteint dans le cas particulier $\mu = \frac{\theta}{2}$. \square

Nous donnons ci-après quelques principes de prime et leurs propriétés.

- Principe de la prime pure : $H(X) = \mathbb{E}(X)$. Ce principe vérifie toutes les propriétés, cependant, le niveau de la prime apparaît comme insuffisant puisque nous ne tenons pas compte du risque.
- Principe de l'espérance : $H(X) = (1 + \theta)\mathbb{E}(X)$ avec $\theta > 0$. Ce principe ne vérifie pas l'invariance par translation ni d'être toujours inférieur à la perte maximale. Il y a une prise en compte du risque mais celle-ci est directement liée à l'espérance.
- Principe de la variance : $H(X) = \mathbb{E}(X) + \alpha Var(X)$ avec $\alpha > 0$. Ce principe ne vérifie pas l'homogénéité. Il ne vérifie pas non plus d'être inférieur à la perte maximale, pour cela il suffit de considérer X telle que $\mathbb{P}(X = 0) = \mathbb{P}(X = \theta) = 1/2$ et de résoudre $H(X) > \text{ess sup } X = \theta$ qui donne $\theta > \frac{2}{\alpha}$ convient. La monotonie non plus avec le même exemple en considérant $Y = \theta$. La sous-additivité est vérifiée dans le cas d'indépendance (c'est même une additivité).
- Principe de l'écart-type : $H(X) = \mathbb{E}(X) + \alpha\sigma(X)$ avec $\alpha > 0$. Cette fois-ci l'homogénéité est satisfaite mais la sous-additivité ne l'est plus en général même pour des variables aléatoires indépendantes. On remarque que la prime n'est pas toujours inférieure à la perte maximale on faisant la même construction que pour le principe de la variance (pour certains α).

Une manière économique d'approcher une prime est d'utiliser l'invariance d'utilité. Si u est la fonction d'utilité de l'assureur, alors on cherche la quantité $H(X)$ telle que

$$\mathbb{E}[u(R + H(X) - X)] = u(R),$$

où R sont les réserves. À noter que le montant de $H(X)$ dépend aussi de R et ainsi deux sinistres peuvent avoir des primes différentes en fonction de la valeur de R .

Ces approches basiques montrent qu'il est difficile de construire un principe de prime simple et cohérent. L'approche moderne tend plus à utiliser la notion de coût du capital : chaque nouveau risque assuré X va apporter un supplément en capital qu'on peut noter $\Delta SCR(X)$ et celui-ci a un coût économique. La prime sera généralement de la forme

$$H(X) = \mathbb{E}(X) + \text{Prime de risque}(X) + \text{Coût du capital} \times \Delta SCR(X)$$

où X est considérée comme le sinistre avec les frais de l'assureur associé. Cela se rapproche des différents principes, en considérant X comme son coût total : le sinistre, les frais et le coût de l'immobilisation du capital, et une prime de risque associée au risque pris.

10 Mesures de risque

Afin d'être capable de rembourser le risque assuré, un assureur doit immobiliser du capital. Il utilise la diversification afin de diminuer le risque sans toutefois le supprimer. Le capital immobilisé est à la fois une garantie de bonne capacité à rembourser les assurés sinistrés et de ne pas être un risque financier trop important pour la collectivité. Toutefois, associer une quantité monétaire à un risque représenté par une variable aléatoire n'est pas facile. On introduit la notion de mesure de risque qui a pour but d'associer à une variable aléatoire X une quantité monétaire à immobiliser $\rho(X)$.

Définition 15 (Mesure de risque). *Une mesure de risque ρ est une fonctionnelle qui, à une variable aléatoire réelle, associe un montant de prime non aléatoire, i.e.*

$$\begin{aligned} \rho : L^0 &\rightarrow \mathbb{R} \cup \{+\infty\} \\ X &\mapsto \rho(X), \end{aligned}$$

qui vérifie les trois propriétés :

- Monotonie : si $X \leq Y$ p.s., $\rho(X) \leq \rho(Y)$,
- Invariance par translation : si $m \in \mathbb{R}$, $\rho(X + m) = \rho(X) + m$,
- Normalisée : $\rho(0) = 0$.

On peut remarquer que $\rho(X - \rho(X)) = 0$ et que $\rho(m) = m$.

Exemple 9. L'application $\rho : X \mapsto \mathbb{E}(X)$ est une mesure de risque. L'application $\rho : X \mapsto \text{ess sup}_{\omega \in \Omega} X(\omega)$ est également une mesure de risque.

Lemme 12. On a

$$|\rho(X) - \rho(Y)| \leq \text{ess sup } |X - Y|.$$

Démonstration. On a $X \leq Y + \text{ess sup } |X - Y|$ p.s., d'où $\rho(X) \leq \rho(Y) + \text{ess sup } |X - Y|$. En inversant X et Y on obtient l'autre inégalité. \square

Définition 16. Une mesure de risque est cohérente si elle satisfait :

- Homogénéité positive : pour $\beta \geq 0$, $\rho(\beta X) = \beta\rho(X)$,
- Sous-additivité : $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

Lemme 13. Si une mesure de risque vérifie deux propriétés parmi la convexité, l'homogénéité positive et la sous-additivité, alors elle vérifie la troisième.

Démonstration. 1. Si $\rho(X + Y) \leq \rho(X) + \rho(Y)$ et si $\rho(\beta X) = \beta\rho(X)$ alors

$$\rho(\beta X + (1 - \beta)Y) \leq \rho(\beta X) + \rho((1 - \beta)Y) = \beta\rho(X) + (1 - \beta)\rho(Y).$$

2. Si $\rho(X + Y) \leq \rho(X) + \rho(Y)$ et $\rho(\beta X + (1 - \beta)Y) \leq \beta\rho(X) + (1 - \beta)\rho(Y)$ alors, on considérant $Y = 0$, $\rho(\beta X) \leq \beta\rho(X)$. De plus

$$0 = \rho(\beta X - \beta\rho(X)) \leq \rho(\beta X) + \rho(-\beta\rho(X)) = \rho(\beta X) - \beta\rho(X),$$

c'est-à-dire $\beta\rho(X) \leq \rho(\beta X)$.

3. Enfin, si $\rho(\beta X) = \beta\rho(X)$ et si $\rho(\beta X + (1 - \beta)Y) \leq \beta\rho(X) + (1 - \beta)\rho(Y)$, alors

$$\rho(X + Y) = \rho\left(\beta\frac{X}{\beta} + (1 - \beta)\frac{Y}{1 - \beta}\right) \leq \beta\rho\left(\frac{X}{\beta}\right) + (1 - \beta)\rho\left(\frac{Y}{1 - \beta}\right) = \rho(X) + \rho(Y).$$

\square

Définition 17 (Value-at-Risk). La Value-at-Risk d'ordre α d'une variable aléatoire X est le quantile d'ordre α correspondant, i.e.

$$\text{VaR}(X, \alpha) = \inf \{x \in \mathbb{R} : F_X(x) \geq \alpha\}.$$

On peut également définir la Value-at-Risk comme le quantile non pas à droite sur les pertes, mais à gauche sur les gains. X , dans la définition ci-dessus, correspond aux pertes. Attention, si X représente les gains, il faut adapter la définition de la VaR et de la mesure de risque. Dans Solvabilité II, le niveau α est fixé à 99.5%.

Lemme 14. La Value-at-Risk est une mesure de risque homogène.

Démonstration. On remarque que $VaR(0, \alpha) = 0$. Soit $X \leq Y$ p.s., alors $F_Y \leq F_X$. Ainsi :

$$\{x \in \mathbb{R} : F_Y(x) \geq \alpha\} \subset \{x \in \mathbb{R} : F_X(x) \geq \alpha\},$$

d'où $VaR(X, \alpha) \leq VaR(Y, \alpha)$. De plus,

$$\begin{aligned} VaR(X + m, \alpha) &= \inf \{x \in \mathbb{R} : F_{X+m}(x) \geq \alpha\} = \inf \{x \in \mathbb{R} : F_X(x - m) \geq \alpha\} \\ &= m + \inf \{x \in \mathbb{R} : F_X(x) \geq \alpha\} = m + VaR(X, \alpha). \end{aligned}$$

De même pour $\beta > 0$:

$$\begin{aligned} VaR(\beta X, \alpha) &= \inf \{x \in \mathbb{R} : F_{\beta X}(x) \geq \alpha\} = \inf \{x \in \mathbb{R} : F_X(x/\beta) \geq \alpha\} \\ &= \beta \inf \{x \in \mathbb{R} : F_X(x) \geq \alpha\} = \beta VaR(X, \alpha). \end{aligned}$$

□

En fait, on a le résultat plus général suivant :

Lemme 15. *Soit g une fonction injective de \mathbb{R} dans E . Si g est **croissante**, on a*

$$VaR(g(X), \alpha) = g(VaR(X, \alpha)).$$

Démonstration. C'est le même principe que dans la précédente preuve : on note g^{-1} l'inverse de g . Si g est croissante :

$$\begin{aligned} VaR(g(X), \alpha) &= \inf \{x \in \mathbb{R} : F_{g(X)}(x) \geq \alpha\} = \inf \{x \in E : F_X(g^{-1}(x)) \geq \alpha\} \\ &= g(\inf \{x \in \mathbb{R} : F_X(x) \geq \alpha\}) = g(VaR(X, \alpha)). \end{aligned}$$

□

Toutefois, la VaR n'est pas sous additive, ce qui implique qu'elle n'est également pas convexe, puisqu'elle est positivement homogène.

Exemple 10. *Soit X une variable aléatoire telle que $\mathbb{P}(X > x) = \frac{1}{1+x}$ pour $x \geq 0$. Soit Y une autre variable aléatoire de même loi et indépendante de X . Alors, $\forall \alpha \in]0, 1[$:*

$$VaR(X + Y, \alpha) > VaR(X, \alpha) + VaR(Y, \alpha),$$

c'est un exemple dans lequel la VaR n'est pas sous-additive.

Démonstration. On remarque que $VaR(X, \alpha) + VaR(Y, \alpha) = VaR(X, \alpha) + VaR(X, \alpha) = VaR(2X, \alpha)$. On a $F_{2X}(x) = 1 - \frac{2}{2+x}$ pour $x \geq 0$. On peut montrer que :

$$\mathbb{P}(X + Y \leq x) = 1 - \frac{2}{2+x} - \frac{2 \log(1+x)}{(2+x)^2}.$$

Ainsi, $F_{X+Y} < F_{2X}$. Comme F_{X+Y} et F_{2X} sont continues et inversibles, cela implique que $F_{X+Y}^{-1} > F_{2X}^{-1}$. On en déduit que

$$VaR(X + Y, \alpha) > VaR(2X, \alpha),$$

c'est-à-dire que

$$VaR(X + Y, \alpha) > VaR(X, \alpha) + VaR(Y, \alpha),$$

□

Exemple 11. Soit X et Y deux variables aléatoires indépendantes et de même loi telles que : $\mathbb{P}(X = \gamma) = p$ et $\mathbb{P}(X = 0) = 1 - p$ avec $p \in]0, 1[$ et $\gamma > 0$.

Démonstration. On remarque que, par indépendance, $\mathbb{P}(X + Y = 0) = (1 - p)^2$, $\mathbb{P}(X + Y = \gamma) = 2p(1 - p)$, $\mathbb{P}(X + Y = 2\gamma) = p^2$. Ainsi, on a, pour $\alpha \in]0, 1[$,

$$VaR(X + Y, \alpha) = 0 \times \mathbf{1}_{\{0 < \alpha \leq (1-p)^2\}} + \gamma \mathbf{1}_{\{(1-p)^2 < \alpha \leq 1-p^2\}} + 2\gamma \mathbf{1}_{\{1-p^2 < \alpha < 1\}}.$$

De plus,

$$VaR(2X, \alpha) = 0 \times \mathbf{1}_{\{0 \leq \alpha \leq 1-p\}} + 2\gamma \mathbf{1}_{\{1-p < \alpha < 1\}}.$$

On remarque que pour $\alpha \in](1 - p)^2, 1 - p]$, on a $VaR(X + Y, \alpha) > VaR(2X, \alpha) = VaR(X, \alpha) + VaR(Y, \alpha)$. \square

On introduit une autre mesure de risque, l'Expected Shortfall :

Définition 18 (Expected Shortfall). L'Expected Shortfall d'ordre α d'une variable aléatoire X est définie par

$$ES(X, \alpha) = \frac{1}{1 - \alpha} \int_{\alpha}^1 VaR(X, p) dp.$$

Lorsque X est une variable aléatoire continue, on a :

$$ES(X, \alpha) = \mathbb{E}(X \mid X \geq VaR(X, \alpha)).$$

On remarque que l'Expected Shortfall est une mesure de risque homogène : elle hérite ces propriétés de la VaR. On ne s'intéresse maintenant qu'aux variables aléatoires continues.

Lemme 16. On a la représentation, en notant \mathcal{F} la tribu :

$$ES(X, \alpha) = \sup_{A \in \mathcal{F}} \{\mathbb{E}(X \mid A) : \mathbb{P}(A) \geq 1 - \alpha\}$$

Démonstration. Admis. \square

Proposition 15. Soit \mathbb{Q} une mesure de probabilité sur $\mathcal{B}([0, 1])$ et soit l'application :

$$\rho(X) = \mathbb{E}^{\mathbb{Q}}(VaR(X, u)),$$

alors ρ est une mesure de risque homogène.

Démonstration. Cela découle du fait que $VaR(X, \alpha)$ est une mesure de risque homogène et de la linéarité de l'espérance. \square

On remarque que si on prend $\mathbb{Q} := \delta_{\alpha}$ on obtient la $VaR(X, \alpha)$ et que si on prend $d\mathbb{Q} := \frac{1}{1-\alpha} \mathbf{1}_{[\alpha, 1]} d\lambda$, on obtient l' $ES(X, \alpha)$.

Corollaire 3. L'Expected Shortfall est sous-additive.

Démonstration. On a

$$\begin{aligned} \sup_{A \in \mathcal{F}} \{\mathbb{E}(X + Y \mid A) : \mathbb{P}(A) \geq 1 - \alpha\} &= \sup_{A \in \mathcal{F}} \{\mathbb{E}(X \mid A) + \mathbb{E}(Y \mid A) : \mathbb{P}(A) \geq 1 - \alpha\} \\ &\leq \sup_{A \in \mathcal{F}} \{\mathbb{E}(X \mid A) : \mathbb{P}(A) \geq 1 - \alpha\} + \sup_{A \in \mathcal{F}} \{\mathbb{E}(Y \mid A) : \mathbb{P}(A) \geq 1 - \alpha\}, \end{aligned}$$

d'où le résultat par application du Lemme 16. \square

Cinquième partie

Théorie de la ruine

Le modèle de Lundberg, présenté en 1903, donne le montant résiduel des capitaux de l'assureur en comparant les recettes et les coûts dans un cadre spécifique. Partant d'un capital initial, l'assurance sera déclarée en ruine à la première date où ses capitaux deviennent négatifs.

Les sinistres S à l'instant $t \geq 0$ sont donnés par

$$S_t = \sum_{k=1}^{N_t} X_k$$

où S est un **processus de Poisson composé**, i.e. que N est un processus de Poisson et que les X_k sont i.i.d. indépendants de N , ils seront de plus supposés positifs. À t fixé, S_t est une somme aléatoire, et S croît à mesure que N croît.

Les recettes notées P , sont supposées arriver continument avec un facteur constant c , i.e.

$$P_t = ct.$$

L'assureur démarre avec un capital initial u à la date 0. Si on note U_t le capital de l'assureur en date t , alors

$$U_t = u + P_t - S_t = u + ct - S_t.$$

Avant toute chose, rappelons quelques définitions et résultats associés au processus $(S_t)_{t \geq 0}$.

11 Processus de Poisson

Définition 19. (*Processus de Poisson*) Un Processus $(N_t)_{t \geq 0}$ est un processus de Poisson de paramètre $\lambda > 0$ si, pour tout $0 \leq s \leq t$,

- $N_0 = 0$ p.s.
- $t \mapsto N_t$ est càdlàg p.s.
- $N_t - N_s$ est indépendant de \mathcal{F}_s ,
- $N_t - N_s \sim \mathcal{P}(\lambda(t - s))$,

où $\mathcal{F}_s := \sigma(N_u, u \leq s)$, i.e. la tribu engendrée par l'observation de $(N_u)_{0 \leq u \leq s}$, i.e « le passé ».

Le théorème suivant donne une autre caractérisation des processus de Poisson.

Théorème 10. Un processus $(N_t)_{t \geq 0}$ est un processus de Poisson de paramètre $\lambda > 0$ si, et seulement si pour tout $0 \leq s \leq t$,

- $N_0 = 0$ p.s.
- En définissant $\tau_0 = 0$, $\tau_1 := \inf \{s \geq 0, N_s > 0\}$, $\tau_{i+1} := \inf \{s \geq \tau_i, N_s - N_{\tau_i} > 0\}$ pour $i \geq 0$, alors $(\tau_{i+1} - \tau_i)_{i \geq 0}$ forment une suite i.i.d. de loi exponentielle de paramètre λ .
- Les sauts sont de taille 1 : $\forall i \geq 1, N_{\tau_{i+1}} = N_{\tau_i} + 1$.

Démonstration. Admis. □

Continuons avec quelques propriétés qui permettent de mieux appréhender le processus de Poisson. Nous avons d'abord le lemme suivant :

Lemme 17. *Le processus de Poisson compensé de paramètre $\lambda > 0$, noté \tilde{N} et défini pour tout $t \geq 0$ par*

$$\tilde{N}_t := N_t - \lambda t,$$

est une martingale.

Démonstration. Comme N_t est un processus de Poisson, l'intégrabilité de \tilde{N}_t est immédiate. On a, pour $s \leq t$

$$\mathbb{E}(\tilde{N}_t | \mathcal{F}_s) = \mathbb{E}(N_t | \mathcal{F}_s) - \lambda t = N_s - \lambda t + \mathbb{E}(N_t - N_s | \mathcal{F}_s) = N_s - \lambda t + \lambda(t - s) = \tilde{N}_s.$$

□

Lemme 18. *Soit $(X_t)_{t \geq 0}$ un processus p.s. croissant ou p.s. décroissant et une variable aléatoire Z telle que*

$$X_t \xrightarrow[t \rightarrow t_0]{\mathbb{P}} Z,$$

alors,

$$X_t \xrightarrow[t \rightarrow t_0]{p.s.} Z.$$

Démonstration. On suppose que $X_t \xrightarrow[t \rightarrow t_0]{\mathbb{P}} Z$. Par construction, si X est un processus décroissant sur Ω , en regardant la limite à gauche, $X_t \geq Z$ pour $t < t_0$ avec probabilité 1 sinon la convergence en probabilité ne peut pas avoir lieu. En effet, si pour un $t' < t_0$, $\mathbb{P}(X_{t'} - Z < 0) > 0$, par décroissance du processus X ,

$$\lim_{t \rightarrow t_0} \mathbb{P}(|X_t - Z| > 0) \geq \lim_{t \rightarrow t_0} \mathbb{P}(X_t - Z < 0) \geq \mathbb{P}(X_{t'} - Z < 0) > 0,$$

ce qui contredit la convergence en probabilité. Soit Ω' de probabilité 1 tel que X_t est décroissant et $X_t \geq Z$. Ainsi, pour $\omega \in \Omega'$, $X_t(\omega)$ est décroissante et minorée par $Z(\omega)$ lorsque $t \rightarrow t_0$, elle converge vers un certain $Y(\omega)$, et comme la convergence p.s. vers Y implique la convergence en probabilité vers Y , par unicité de la limite, $Y = Z$ p.s.

On montre de même pour la limite à droite et de même lorsque X est croissant. □

Proposition 16. *Soit $(N_t)_{t \geq 0}$ un processus de Poisson. Alors*

- $\forall t > 0, N_t - N_{t-} = 0$ p.s.
- $\mathbb{P}(\forall t > 0, N_t - N_{t-} \leq 1) = 1$,
- *Les temps de saut sont des temps d'arrêts totalement inaccessibles.*

Démonstration. Soit $t > 0$. On remarque que $\mathbb{P}(N_t - N_{t-h} > 0) = 1 - \exp(-\lambda h)$ qui tend vers 0 lorsque $h \rightarrow 0$. Ainsi,

$$N_t - N_{t-h} \xrightarrow[h \rightarrow 0]{\mathbb{P}} 0,$$

et comme $N_t - N_{t-h}$ est décroissant, par le lemme 18, on en déduit

$$N_t - N_{t-h} \xrightarrow[h \rightarrow 0]{p.s.} 0.$$

i.e. que $N_t - N_{t-} = 0$ p.s.

Pour le deuxième point, on fixe n et $m \in \mathbb{N}$

$$\begin{aligned}
\mathbb{P}(\exists t \in]m, m+1], N_t - N_{t-} > 1) &\leq \mathbb{P}\left(\bigcup_{k=0}^{n-1} \{N_{(k+1)/n} - N_{k/n} > 1\}\right) \\
&\leq \sum_{k=0}^{n-1} \mathbb{P}(N_{(k+1)/n} - N_{k/n} > 1) \\
&\leq n\mathbb{P}(N_{1/n} - N_0 > 1) \\
&\leq n(1 - e^{-\lambda/n} - e^{-\lambda/n}\lambda/n) \\
&\leq n(1 - e^{-\lambda/n}) - e^{-\lambda/n}\lambda \\
&\leq \lambda + o(1) - e^{-\lambda/n}\lambda
\end{aligned}$$

et nous observons lorsque $n \rightarrow +\infty$ que,

$$\mathbb{P}(\exists t \in]m, m+1], N_t - N_{t-} > 1) = 0.$$

De plus,

$$\begin{aligned}
\mathbb{P}(\exists t > 0, N_t - N_{t-} > 1) &= \mathbb{P}\left(\bigcup_{m \in \mathbb{N}} \{\exists t \in]m, m+1], N_t - N_{t-} > 1\}\right) \\
&\leq \sum_{m \in \mathbb{N}} \mathbb{P}(\exists t \in]m, m+1], N_t - N_{t-} > 1) = 0.
\end{aligned}$$

Puis, par complémentarité, on en déduit le résultat.

Enfin, pour le dernier point, soit τ un temps de saut du processus, si par l'absurde il existe une suite de temps d'arrêt $\tau_n < \tau$ qui converge vers (annonce) τ ,

$$\mathbb{E}(N_{t \wedge \tau} - N_{t \wedge \tau-}) = \mathbb{E}(\tilde{N}_{t \wedge \tau} - \tilde{N}_{t \wedge \tau-}) = \lim_{n \rightarrow +\infty} \mathbb{E}(\tilde{N}_{t \wedge \tau} - \tilde{N}_{t \wedge \tau_n}).$$

Or, par définition de τ et application du théorème de convergence monotone,

$$\lim_{t \rightarrow +\infty} \mathbb{E}(N_{t \wedge \tau} - N_{t \wedge \tau-}) = \mathbb{E}(N_\tau - N_{\tau-}) = 1.$$

tandis que, par le théorème d'arrêt,

$$\forall n \geq 1, \mathbb{E}(\tilde{N}_{t \wedge \tau} - \tilde{N}_{t \wedge \tau_n}) = 0,$$

ce qui mène à une contradiction. □

Enfin, donnons une construction d'un processus de Poisson qui découle du théorème 10.

Lemme 19. Soient $(T_i)_{i \geq 1}$ une suite i.i.d. de loi exponentielle de paramètre λ . On pose pour $j \geq 1$,

$$\tau_j := \sum_{i=1}^j T_i.$$

Alors le processus $(N_t)_{t \geq 0}$ défini pour tout $t \geq 0$ par

$$N_t := \sum_{j \geq 1} \mathbb{1}_{\{\tau_j \leq t\}}$$

est un processus de Poisson de paramètre λ .

Démonstration. On vérifie facilement les hypothèses du théorème 10. □

Des propriétés analogues s'obtiennent pour le processus de Poisson composé (S_t) par héritage des propriétés de (N_t) . On supposera que la suite (X_k) est dans L^1 , dans ce cas ce sera $\tilde{S}_t := S_t - \mathbb{E}(X_1)\lambda t$ qui sera une martingale.

12 Probabilité de ruine

On pose $\mathbb{E}(X_1) = \mu$.

Proposition 17 (Loi des grands nombres sur S). *Avec les notations précédentes, on a*

$$\frac{S_t}{t} \xrightarrow[t \rightarrow +\infty]{p.s., L^1} \lambda\mu.$$

Démonstration. On pose n l'opérateur « partie entière », i.e. pour tout $t \geq 0$, on a

$$n(t) \leq t < n(t) + 1$$

avec la fonction n à valeurs dans \mathbb{N} . Soit $n \in \mathbb{N}$, on a

$$\frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^n (S_i - S_{i-1})$$

où les $(S_i - S_{i-1})_{i \geq 1}$ sont i.i.d. de loi S_1 , par la loi des grands nombres,

$$\frac{S_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s., L^1} \lambda\mu. \tag{1}$$

De plus, S_t est croissant, on a,

$$\frac{S_{n(t)}}{n(t) + 1} \leq \frac{S_t}{t} \leq \frac{S_{n(t)+1}}{n(t)}.$$

En adaptant légèrement (1), on en déduit par encadrement,

$$\frac{S_t}{t} \xrightarrow[t \rightarrow +\infty]{p.s.} \lambda\mu.$$

Pour la convergence L^1 ,

$$\begin{aligned} \left| \frac{S_t}{t} - \lambda\mu \right| &\leq \left| \frac{S_{n(t)}}{t} - \frac{n(t)\lambda\mu}{t} \right| + \left| \frac{S_t - S_{n(t)}}{t} - \lambda\mu \left(1 - \frac{n(t)}{t} \right) \right| \\ &\leq \frac{n(t)}{t} \left| \frac{S_{n(t)}}{n(t)} - \lambda\mu \right| + \left| \frac{S_{n(t)+1} - S_{n(t)}}{t} \right| + \left| \lambda\mu \left(1 - \frac{n(t)}{t} \right) \right|. \end{aligned}$$

Par (1), le premier terme tend vers 0 dans L^1 . Le second terme a la même loi que $\frac{S_1}{t}$ qui tend également vers 0 dans L^1 . Enfin, le dernier terme est une suite numérique qui tend vers 0, c'est également le cas dans L^1 , et finalement,

$$\frac{S_t}{t} \xrightarrow[t \rightarrow +\infty]{L^1} \lambda\mu.$$

□

Corollaire 4. Avec les notations précédentes, on a

$$\frac{U_t}{t} \xrightarrow[p.s., L^1]{t \rightarrow +\infty} c - \lambda\mu.$$

Démonstration. On a

$$\left| \frac{U_t}{t} - (c - \lambda\mu) \right| \leq \left| \frac{u}{t} \right| + \left| \frac{S_t}{t} - \lambda\mu \right|.$$

Puis on en déduit la convergence p.s. et L^1 . \square

On remarque que pour être profitable, l'assureur doit avoir $c \geq \lambda\mu$, i.e. plus que sa sinistralité moyenne.

Nous avons également un théorème central limite pour S .

Proposition 18 (Théorème central limite sur S). Avec les notations précédentes, si $\mathbb{E}(X_1^2) < +\infty$,

$$\sqrt{t} \frac{\frac{1}{t}S_t - \lambda\mu}{\sqrt{\lambda\mathbb{E}(X_1^2)}} = \frac{S_t - \lambda t\mu}{\sqrt{\lambda t\mathbb{E}(X_1^2)}} \xrightarrow[t \rightarrow +\infty]{Loi} \mathcal{N}(0, 1).$$

Démonstration. La suite $(S_n - S_{n-1})_{n \geq 1}$ est i.i.d. et L^2 , par le théorème central limite,

$$\sqrt{n} \frac{\frac{1}{n}S_n - \lambda\mu}{\sqrt{\lambda\mathbb{E}(X_1^2)}} = \frac{S_n - \lambda n\mu}{\sqrt{\lambda n\mathbb{E}(X_1^2)}} \xrightarrow[n \rightarrow +\infty]{Loi} \mathcal{N}(0, 1).$$

On a la décomposition :

$$\frac{S_{n(t)} - \lambda n(t)\mu}{\sqrt{\lambda n(t)\mathbb{E}(X_1^2)}} \sqrt{\frac{n(t)}{t}} + \frac{S_t - S_{n(t)} - \lambda(t - n(t))\mu}{\sqrt{\lambda t\mathbb{E}(X_1^2)}}$$

On a montré que le premier terme converge vers une $\mathcal{N}(0, 1)$. On a $\frac{\lambda(t - n(t))\mu}{\sqrt{\lambda t\mathbb{E}(X_1^2)}} \leq \frac{\lambda\mu}{\sqrt{\lambda t\mathbb{E}(X_1^2)}}$ qui converge vers 0. On a $\frac{S_t - S_{n(t)}}{\sqrt{\lambda t\mathbb{E}(X_1^2)}} \leq \frac{S_{n(t)+1} - S_{n(t)}}{\sqrt{\lambda t\mathbb{E}(X_1^2)}}$ et :

$$\mathbb{E} \left(\frac{S_{n(t)+1} - S_{n(t)}}{\sqrt{\lambda t\mathbb{E}(X_1^2)}} \right) = \frac{\lambda\mu}{\sqrt{\lambda t\mathbb{E}(X_1^2)}} \xrightarrow[t \rightarrow +\infty]{} 0,$$

comme la convergence L^1 implique la convergence en probabilité, par le théorème de Slutsky, on en déduit le résultat. \square

Définition 20 (probabilité de ruine). L'assureur est dit ruiné si à un instant t , $U_t < 0$. La probabilité de ruine sur $[0, T]$ est définie par

$$\psi(u, T) = \mathbb{P}(\exists t \in [0, T], U_t < 0)$$

et la probabilité de ruine par

$$\psi(u) = \mathbb{P}(\exists t \geq 0, U_t < 0).$$

La date de la ruine est notée

$$\tau = \inf \{t \geq 0, U_t < 0\}.$$

On notera λ le paramètre du processus de Poisson et μ l'espérance des X_k supposés intégrables.

Définition 21 (coefficient de chargement). On note ρ le coefficient de chargement, défini par

$$\rho = \frac{c}{\lambda\mu} - 1.$$

La condition de profit net s'écrit

$$c > \lambda\mu \Leftrightarrow \rho > 0.$$

Les primes perçues s'écrivent

$$P_t = ct = (1 + \rho)\lambda\mu t = (1 + \rho)\mathbb{E}(S_t).$$

Théorème 11. On a

$$\psi(u) = 1 \iff \rho \leq 0.$$

Démonstration. Admis. □

Lemme 20. En notant Z_i la durée entre deux sauts du processus N_t , nous avons

$$\{\exists t \geq 0, U_t < 0\} = \left\{ \exists n \geq 1 : u + \sum_{i=1}^n (cZ_i - X_i) < 0 \right\} \cup \{u < 0\}.$$

En particulier,

$$\psi(u) = \mathbb{P} \left(\left\{ \exists n \geq 1 : u + \sum_{i=1}^n (cZ_i - X_i) < 0 \right\} \cup \{u < 0\} \right).$$

Démonstration. On procède par double inclusion. Puisque le membre de droite correspond à celui de gauche évalué aux temps de saut seulement, il est inclus dans celui de gauche. Réciproquement, si pour tout $n, u + \sum_{i=1}^n (cZ_i - X_i) \geq 0$, alors pour $\sum_{i=1}^n Z_i \leq t < \sum_{i=1}^{n+1} Z_i$, $u + ct - S_t = u + ct - \sum_{i=1}^n X_i \geq u + c \sum_{i=1}^n Z_i - \sum_{i=1}^n X_i \geq 0$. □

Théorème 12. Soit $u \geq 0$. On suppose que pour un $t > 0, g_X(t) < +\infty$. Dans ce cas,

$$\psi(u) \leq e^{-Ru}$$

où R , appelé le coefficient d'ajustement, est l'unique solution strictement positive de

$$\mathbb{E} [e^{R(X_1 - cZ_1)}] = 1 \tag{2}$$

c'est-à-dire de

$$g_X(R) = 1 + \frac{cR}{\lambda} \iff g_X(R) - 1 - (1 + \rho)\mu R = 0$$

Démonstration. Étape 1 : on montre que l'équation (2) admet une unique solution strictement positive. Par le théorème 3, g_X est dérivable et en particulier $g'_X(0) = \mathbb{E}(X)$. On pose

$$q(r) := g_X(r) - 1 - (1 + \rho)\mu r.$$

On remarque que $q(0) = 0, q'(0) = -\rho\mu < 0$. De plus, par Jensen,

$$q(r) \geq e^{r\mu} - 1 - (1 + \rho)\mu r.$$

On a

$$q(r) \xrightarrow[r \rightarrow +\infty]{} +\infty$$

En combinant ces différents points, en remarquant que q est convexe et en construisant un tableau de variations, combiné à la continuité de q , par le théorème des valeurs intermédiaires, on remarque que $q(r) = 0$ admet une unique solution strictement positive.

Étape 2 : On pose

$$\psi_n(u) := \mathbb{P} \left(\left\{ \exists k \in \{1 \leq k \leq n\} : u + \sum_{i=1}^k (cZ_i - X_i) < 0 \right\} \cup \{u < 0\} \right).$$

Par construction, $(\psi_n(u))_{n \geq 1}$ est une suite croissante de limite $\psi(u)$ puisque $\psi_n(u)$ est la probabilité d'une union dont $\psi(u)$ est la limite. Il suffit de montrer en particulier que, pour tout $n \geq 1$,

$$\psi_n(u) \leq e^{-Ru}. \quad (3)$$

On procède par récurrence. Pour $n = 1$, en utilisant l'inégalité de Markov, avec $u \geq 0$,

$$\psi_1(u) = \mathbb{P}(X_1 - cZ_1 > u) = \mathbb{P}(e^{R(X_1 - cZ_1)} > e^{Ru}) \leq e^{-Ru} \mathbb{E}(e^{R(X_1 - cZ_1)}) = e^{-Ru},$$

car $\mathbb{E}[e^{R(X_1 - cZ_1)}] = 1$ par définition de R . On suppose (3) vraie au rang n . Nous affirmons que :

$$\psi_{n+1}(u) = \mathbb{E}(\psi_n(u + cZ_1 - X_1)) \quad (4)$$

que nous prouverons à l'étape suivante. On notera $\mathbb{P}_{(Z,X)}$ la mesure de probabilité commune des (Z_i, X_i) (qui sont indépendants et « $d\mathbb{P}_{(Z,X)} = d\mathbb{P}_Z d\mathbb{P}_X$ »). Nous avons,

$$\begin{aligned} \psi_{n+1}(u) &= \mathbb{E}(\psi_n(u + cZ_1 - X_1)) \\ &\leq \mathbb{E}(e^{-R(u + cZ_1 - X_1)}) \\ &\leq e^{-Ru} \mathbb{E}(e^{R(X_1 - cZ_1)}) \\ &\leq e^{-Ru} \end{aligned}$$

où la dernière égalité s'obtient en utilisant la définition de R .

Étape 3 : Nous prouvons maintenant (4). On pose $E = \mathbb{R}_+^2$ de sorte que $(Z_i, X_i) \in E$ et $\bar{\psi}_n(u) := 1 - \psi_n(u)$ qui représente la probabilité de survie (i.e. qu'il n'y ait pas de ruine). On note $\mathbb{P}_{(Z,X)}^{\otimes n}$ la mesure produit de n fois $\mathbb{P}_{(Z,X)}$, i.e. « $d\mathbb{P}_{(Z,X)}^{\otimes n}((z_1, x_1), \dots, (z_n, x_n)) = d\mathbb{P}_{(Z,X)}(z_1, x_1) \dots d\mathbb{P}_{(Z,X)}(z_n, x_n)$ ». Il vient

$$\begin{aligned} \bar{\psi}_{n+1}(u) &= \mathbb{P} \left(\bigcap_{j=1}^{n+1} \left\{ u + \sum_{i=1}^j cZ_i - X_i \geq 0 \right\} \right) \\ &= \int_{E^{n+1}} \mathbb{1}_{\{\bigcap_{j=1}^{n+1} \{u + \sum_{i=1}^j cZ_i - X_i \geq 0\}\}} d\mathbb{P}_{(Z,X)}^{\otimes n+1}((z_1, x_1), \dots, (z_n, x_n)) \\ &= \int_E \left[\int_{E^n} \mathbb{1}_{\{u + cz_1 - x_1 \geq 0\}} \mathbb{1}_{\{\bigcap_{j=2}^{n+1} \{u + cz_1 - x_1 + \sum_{i=2}^j cZ_i - X_i \geq 0\}\}} d\mathbb{P}_{(Z,X)}^{\otimes n}((z_2, x_2), \dots, (z_{n+1}, x_{n+1})) \right] d\mathbb{P}_{(Z,X)}(z_1, x_1) \\ &= \int_E \bar{\psi}_n(u + cz_1 - x_1) d\mathbb{P}_{(Z,X)}(z_1, x_1) \\ &= \mathbb{E}(\bar{\psi}_n(u + cZ_1 - X_1)) \end{aligned}$$

ce qui conclut la preuve on utilisant $\bar{\psi}_n = 1 - \psi_n$ et la linéarité de l'espérance. \square

Nous avons le petit résultat suivant :

Lemme 21. *Nous avons la relation*

$$\begin{aligned} \sum_{k=N_s+1}^{N_t} X_k &\perp\!\!\!\perp \mathcal{F}_s, \\ \sum_{k=N_s+1}^{N_t} X_k &\stackrel{\mathcal{L}}{=} \sum_{k=1}^{N_t-N_s} X_k, \end{aligned}$$

où $\mathcal{F}_s := \sigma(N_u, u \leq s, X_k, j \leq N_s)$.

Démonstration. Soit $t \in \mathbb{R}$, nous avons

$$\begin{aligned} \mathbb{E} \left(e^{it \sum_{k=N_s+1}^{N_t} X_k} \middle| \mathcal{F}_s \right) &= \mathbb{E} \left(e^{it \sum_{k=1}^{N_t-N_s} X_{N_s+k}} \middle| \mathcal{F}_s \right) \\ &= \mathbb{E} \left(e^{it \sum_{k=1}^{N_t} X_k} \right) \\ &= P_{N_t-N_s} \circ \phi_X(t) \\ &= P_{N_t-s} \circ \phi_X(t) \end{aligned}$$

□

Nous en déduisons le résultat suivant.

Corollaire 5. *Le processus $(e^{-RU_t})_{t \geq 0}$ est une \mathcal{F} -martingale.*

Démonstration. Nous avons que, pour tout t , e^{-RU_t} est intégrable et adapté. De plus,

$$\begin{aligned} \mathbb{E} \left(e^{-RU_t} \middle| \mathcal{F}_s \right) &= \mathbb{E} \left(e^{-RU_s - R(U_t - U_s)} \middle| \mathcal{F}_s \right) \\ &= e^{-RU_s - cR(t-s)} \mathbb{E} \left(e^{R \sum_{k=N_s+1}^{N_t} X_k} \middle| \mathcal{F}_s \right) \\ &= e^{-RU_s - cR(t-s)} \mathbb{E} \left(e^{R \sum_{k=1}^{N_t-N_s} X_k} \right) \\ &= e^{-RU_s - cR(t-s)} e^{\lambda(t-s)(g_X(R)-1)} \end{aligned}$$

On rappelle que par définition, $g_X(R) = 1 + \frac{cR}{\lambda}$. On en déduit

$$\mathbb{E} \left(e^{-RU_t} \middle| \mathcal{F}_s \right) = e^{-RU_s}.$$

□

Théorème 13. *Soit $u \geq 0$. On suppose que pour un $t > 0$, $g_X(t) < +\infty$. On note R le coefficient d'ajustement. Alors*

$$\psi(u) = \frac{e^{-Ru}}{\mathbb{E} \left(e^{-RU_\tau} \middle| \{\tau < +\infty\} \right)}.$$

En particulier, on retrouve

$$\psi(u) \leq e^{-Ru}.$$

Démonstration. Par le lemme 5, $(e^{-RU_t})_{t \geq 0}$ est une martingale. Par le théorème d'arrêt,

$$\mathbb{E}(e^{-RU_{\tau \wedge n}}) = e^{-Ru},$$

c'est à dire,

$$\mathbb{E}(e^{-RU_{\tau}} \mathbf{1}_{\{\tau \leq n\}}) + \mathbb{E}(e^{-RU_n} \mathbf{1}_{\{\tau > n\}}) = e^{-Ru}.$$

On remarque que $e^{-RU_n} \mathbf{1}_{\{\tau > n\}} \leq \mathbf{1}_{\{\tau > n\}} \leq 1$. De plus comme $\rho > 0$, on a $U_n \rightarrow +\infty$ p.s.; par convergence dominée,

$$\mathbb{E}(e^{-RU_n} \mathbf{1}_{\{\tau > n\}}) \rightarrow 0.$$

La suite $(e^{-RU_{\tau}} \mathbf{1}_{\{\tau \leq n\}})_{n \geq 0}$ est croissante, par convergence monotone,

$$\mathbb{E}(e^{-RU_{\tau}} \mathbf{1}_{\{\tau < +\infty\}}) = e^{-Ru}$$

Enfin, par définition, $\psi(u) = \mathbb{P}(\{\tau < +\infty\})$. Il vient,

$$\psi(u) \mathbb{E}(e^{-RU_{\tau}} \mid \{\tau < +\infty\}) = e^{-Ru}.$$

□

Sixième partie

Annexes

13 Preuve de l'exemple 5

Démonstration. 1. Soit $a \in [-1, 1]$, $f_a \geq 0$ et, avec le changement de variable $u(x) = \log(x)$, on a

$$\int_0^{+\infty} f_a(x) dx = \int_{\mathbb{R}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} (1 + a \sin(2\pi u)) du = \int_{\mathbb{R}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du + a \int_{\mathbb{R}} \sin(2\pi u) \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du.$$

Le premier membre s'intègre à 1, le second est bien intégrable et la fonction intégrée est symétrique (impaire), elle s'intègre donc à 0.

2. Comme la loi de X_0 est définie sur \mathbb{R}_+ , pour $t \leq 0$, $g_X(t)$ existe. Soit $t > 0$. Un calcul rapide, en effectuant le changement de variable $u(x) = \log(x)$, puis en reprenant les résultats de l'Exemple 4, donne

$$\mathbb{E}(X_0^n) = e^{\frac{n^2}{2}}.$$

Calculons maintenant $\mathbb{E}(X_a^n)$. Nous avons, en effectuant le changement de variable $u(x) = \log(x)$, et en remarquant que $\sin(2\pi(u+n)) = \sin(2\pi u)$,

$$\begin{aligned} \mathbb{E}(X_a^n) &= \int_{\mathbb{R}} \frac{e^{nu} e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} (1 + a \sin(2\pi u)) du = e^{\frac{n^2}{2}} + a \int_{\mathbb{R}} \sin(2\pi u) \frac{e^{nu} e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du \\ &= e^{\frac{n^2}{2}} + a \int_{\mathbb{R}} \sin(2\pi u) \frac{e^{-\frac{(u-n)^2}{2} + \frac{n^2}{2}}}{\sqrt{2\pi}} du = e^{\frac{n^2}{2}} + a e^{\frac{n^2}{2}} \int_{\mathbb{R}} \sin(2\pi(u+n)) \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du \\ &= e^{\frac{n^2}{2}} = \mathbb{E}(X_0^n) \end{aligned}$$

3. On pose

$$u_n(t) := \mathbb{E}(X_a^n) \frac{t^n}{n!} = \frac{e^{\frac{n^2}{2}} t^n}{n!},$$

on remarque que

$$\frac{u_{n+1}(t)}{u_n(t)} = \frac{te^{n+\frac{1}{2}}}{n+1} \xrightarrow{n \rightarrow +\infty} +\infty,$$

la série associée diverge très grossièrement ($u_n(t) \rightarrow +\infty$). Et comme la variable aléatoire est positive,

$$e^{tX_a} \geq \sum_{k=0}^n X_a^k \frac{t^k}{k!},$$

En appliquant l'opérateur espérance et en utilisant la définition de u_n ,

$$g_{X_a}(t) \geq \sum_{k=0}^n u_k(t).$$

Puis, en faisant tendre n vers l'infini, on obtient $g_{X_a}(t) = +\infty$ (pour $t > 0$). □