Processus de Galton-Watson [1]

Leçons : 223, 226 [2], 229, 243, 260, 264, *206, 241, 244, 253*

Merci Laura ! [3]

Soit X une variable aléatoire intégrable à valeurs dans $\mathbb{N}$.
On note, pour $n \in \mathbb{N}$, $p_n = \mathbb{P}(X = n)$ et $m = \mathbb{E}[X] = \sum_{n=0}^{\infty} n p_n < \infty$.
Soit $\left(X_{i,j}\right)_{i,j \in \mathbb{N}}$ une famille de variables aléatoires indépendantes et identiquement distribuées, suivant la loi $\mathbb{P}_X$.
On définit la suite $(Z_n)_{n \in \mathbb{N}}$ de la façon suivante :

$$\begin{cases} Z_0 = 1 \\ \forall n \in \mathbb{N}, Z_{n+1} = \sum_{i=1}^{Z_n} X_{i,n} \end{cases}$$

L'idée est alors de modéliser avec (Z_n) la taille d'une population ; plus précisément, Z_n symbolisera le nombre d'individus à la $n^{\text{ème}}$ génération, et pour $i \in [\![1, Z_n]\!]$, $X_{i,n}$ représentera le nombre de descendants que l'individu de la $n^{\text{ème}}$ génération portant le numéro i aura engendré (les individus de la population qu'on considère génèrent des enfants tous seuls).
On va étudier la suite (Z_n), et particulièrement, on va répondre à la question “Que vaut $\mathbb{P}\left(\exists n \in \mathbb{N}, Z_n = 0\right)$? (Quelle est la probabilité que la population considérée s'éteigne ?)” [4]

Lemme

On a : $\forall n \in \mathbb{N}^*, \forall i \in \mathbb{N}, Z_n \perp\!\!\!\perp X_{i,n}$.

Démonstration :

Soit $n \in \mathbb{N}^*$; Z_n ne dépend que de Z_{n-1} et de la famille $(X_{i,n-1})_{i \in \mathbb{N}}$.
Ainsi, par une récurrence immédiate, il vient : Z_n ne dépend que de la famille $\left(X_{i,j}\right)_{i \geqslant 0, j < n}$.
Et, par indépendance des variables $X_{i,j}$, on obtient que $\forall i \in \mathbb{N}, Z_n \perp\!\!\!\perp X_{i,n}$. ■

1. On sautera l'aspect modélisation et le premier lemme. Dans la 229, on ne montre pas que π_∞ est le plus petit point fixe de G, mais on détaille la stricte convexité par la suite. Dans le dernier théorème, on ne construit que les tableaux de variations, et oralement.
2. Attention à bien justifier ce développement dans cette leçon. C'est la suite $(\pi_n)_{n \in \mathbb{N}}$ qui permet de ne pas être de mauvaise foi, évitez l'argument “On regarde la suite $(Z_n, n)_{n \in \mathbb{N}}$” qui est quand même moins élégant ; ou encore pire : “$Z_{n+1} = f_n(Z_n)$” qui est carrément hors-sujet.
3. Un lien vers la page personnelle de Laura GAY.
4. On aurait pu se poser la question “Que vaut $\mathbb{E}[Z_n]$? (Quel est le nombre moyen d'individus à la $n^{\text{ème}}$ génération ?)”

Théorème

On a : $\forall n \in \mathbb{N}, \mathbb{E}[Z_n] = m^n$.

En effet, soit $n \in \mathbb{N}$;

$$\begin{aligned} \mathbb{E}[Z_{n+1}] &= \mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n} \middle| Z_n\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{\infty} \mathbb{1}_{i \leqslant Z_n} X_{i,n} \middle| Z_n\right]\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{\infty} \mathbb{1}_{i \leqslant Z_n} \mathbb{E}\left[X_{i,n} | Z_n\right]\right] \text{ (par Fubini-Tonelli, car } \mathbb{1}_{i \leqslant Z_n} X_{i,n} \geqslant 0 \text{ ps et par } Z_n\text{-mesurabilité de } \mathbb{1}_{i \leqslant Z_n}) \\ &= \mathbb{E}\left[\sum_{i=1}^{Z_n} \mathbb{E}\left[X_{i,n}\right]\right] \text{ (d'après le lemme, on a l'indépendance entre } X_{i,n} \text{ et } Z_n) \\ &= \mathbb{E}\left[\sum_{i=1}^{Z_n} m\right] = m\mathbb{E}[Z_n] \end{aligned}$$

On conclut par récurrence, en utilisant le fait que $Z_0 = 1$.

On note, pour $n \in \mathbb{N}$, $\pi_n = \mathbb{P}(Z_n = 0)$; et $\pi_\infty = \mathbb{P}(\exists n \in \mathbb{N}, Z_n = 0)$ la probabilité d'extinction.
Comme $Z_n = 0 \Rightarrow Z_{n+1} = 0$, la suite d'événements $(\{Z_n = 0\})_{n\in\mathbb{N}}$ est croissante et on a bien :

$$\pi_\infty = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{Z_n = 0\}\right) = \lim_{n\to\infty} \pi_n$$

Si $p_0 = 0$, alors on a $\forall n \in \mathbb{N}^*, Z_n \geqslant 1$ ps et $\pi_\infty = 0$.
Si $p_0 = 1$, alors on a $\forall n \in \mathbb{N}^*, Z_n = 0$ ps et $\pi_\infty = 1$.
On suppose donc désormais $p_0 \in]0,1[$.

Proposition

On définit la série génératrice de X par $G : s \mapsto \mathbb{E}\left[s^X\right] = \sum_{k=0}^{\infty} p_k s^k$.
On a les résultats suivants :

1. G est bien définie sur $[0,1]$ et y est de classe $\mathcal{C}^1$.
2. (a) G est strictement croissante sur $]0,1[$.
   (b) G est convexe sur $]0,1[$.
   (c) G est strictement convexe sur $]0,1[\Leftrightarrow p_0 + p_1 < 1$.

Démonstration :

1. $\forall k \in \mathbb{N}, s \mapsto p_k s^k$ est de classe $\mathcal{C}^1$ sur $[0,1]$, la série $\sum_{k\geqslant 0} p_k 1^k$ converge (vers 1), et la série de fonctions $\sum_{k\geqslant 1} k p_k s^{k-1}$ converge normalement (car X est intégrable) donc uniformément sur $[0,1]$.
   Par conséquent, la série $\sum_{k\geqslant 0} p_k s^k$ converge uniformément vers G, de classe $\mathcal{C}^1$ sur $[0,1]$.
2. La série entière $\sum_{k\geqslant 0} p_k s^k$ ayant un rayon de convergence $\geqslant 1$, on a :

$$\forall s \in [0,1[, G'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1} \text{ et } G''(s) = \sum_{k=2}^{\infty} k(k-1) p_k s^{k-2}$$

   Comme $p_0 < 1$, on a : $\exists k_0 > 0, p_{k_0} > 0$.
   (a) Ainsi : $\forall s \in]0,1[, G'(s) \geqslant k_0 p_{k_0} s^{k_0 - 1} > 0$ et G est strictement croissante sur $]0,1[$.
   (b) Aussi : $\forall s \in]0,1[, G''(s) \geqslant k_0 (k_0 - 1) p_{k_0} s^{k_0 - 2} \geqslant 0$ et G est convexe sur $]0,1[$.
   (c) Si $p_0 + p_1 = 1$, alors on a $k_0 = 1$ et G est affine donc n'est pas strictement convexe sur $]0,1[$.
   Si $p_0 + p_1 < 1$, alors on peut avoir $k_0 > 1$ et $G'' > 0$ sur $]0,1[$ d'où la stricte convexité. ■

Proposition

Pour $n \in \mathbb{N}$, on définit la série génératrice de Z_n par $G_n : s \mapsto \mathbb{E}\left[s^{Z_n}\right] = \sum_{k=0}^{\infty} \mathbb{P}(Z_n = k) s^k$.
Comme précédemment, on peut montrer que G_n est bien définie sur $[0,1]$.
On a : $\forall n \in \mathbb{N}^*, G_n = \underbrace{G \circ \ldots \circ G}_{n \text{ fois}}$ sur $[0,1]$.

Démonstration :
Soit $n \in \mathbb{N}$, $s \in [0,1]$,

$$\begin{aligned}
G_{n+1}(s) = \mathbb{E}\left[s^{Z_{n+1}}\right] = \mathbb{E}\left[s^{\sum_{i=1}^{Z_n} X_{i,n}}\right] = \mathbb{E}\left[\prod_{i=1}^{Z_n} s^{X_{i,n}}\right] &= \mathbb{E}\left[\sum_{j=0}^{\infty} \mathbb{1}_{Z_n=j} \prod_{i=1}^{j} s^{X_{i,n}}\right] \\
&= \sum_{j=0}^{\infty} \mathbb{E}\left[\mathbb{1}_{Z_n=j} \prod_{i=1}^{j} s^{X_{i,n}}\right] \text{(par Fubini-Tonelli, car les termes sont tous positifs)} \\
&= \sum_{j=0}^{\infty} \mathbb{E}\left[\mathbb{1}_{Z_n=j}\right] \prod_{i=1}^{j} \mathbb{E}\left[s^{X_{i,n}}\right] \text{(car les variables en jeu ici sont toutes indépendantes)} \\
&= \sum_{j=0}^{\infty} \mathbb{P}\left(Z_n=j\right) \mathbb{E}\left[s^{X}\right]^j = \sum_{j=0}^{\infty} \mathbb{P}\left(Z_n=j\right) G(s)^j = G_n(G(s))
\end{aligned}$$

On conclut par récurrence, en utilisant le fait que $Z_1 = X_{0,0}$ suit la loi de X. ■

Proposition

La probabilité d'extinction π_∞ est le plus petit point fixe de G sur l'intervalle $[0,1]$.

Démonstration :
La proposition précédente donne : $\forall n \in \mathbb{N}, \forall s \in [0,1], G_{n+1}(s) = G\left(G_n(s)\right)$.
En évaluant en 0, on obtient la relation : $\pi_{n+1} = G\left(\pi_n\right)$, puis par continuité de G sur $[0,1]$, on obtient que π_∞ est un point fixe de G. Reste à montrer que c'est le plus petit.
Soit $u \in [0,1]$ un point fixe de G. On va montrer par récurrence que $\forall n \in \mathbb{N}^*, \pi_n \leqslant u$.

- On a : $\pi_1 = G\left(\pi_0\right) = G\left(\mathbb{P}\left(Z_0 = 0\right)\right) = G(0) \leqslant G(u) = u$, car G est croissante.
- Si $\pi_n \leqslant u$, alors $\pi_{n+1} = G(\pi_n) \leqslant G(u) = u$, toujours par croissance de G.

Par conséquent, $\forall n \in \mathbb{N}, \pi_n \leqslant u$, puis par passage à la limite : $\pi_\infty \leqslant u$. ■

Théorème

Si $m \leqslant 1$, alors $\pi_\infty = 1$.
Si $m > 1$, alors π_∞ est l'unique point fixe de G sur $]0,1[$.

Démonstration :
On rappelle qu'on a deux cas :

- Si $p_0 + p_1 = 1$, et alors G est une fonction affine ; comme $p_0 > 0$, la droite représentative de G coupe en un unique point la droite d'équation $y = x$. Nécessairement, ce point d'intersection a pour coordonnées $(1,1)$.
- Sinon, G est strictement convexe sur $]0,1[$; il en va alors de même pour $x \mapsto G(x) - x$ qui s'annule donc au plus deux fois [5].

Dans tous les cas, on a : $G(x) - x$ s'annule au plus 2 fois sur $[0,1]$; aussi $G'(0) = p_1$ et $G'(1) = \sum_{n=1}^{\infty} n p_n = m$.

Supposons $m > 1$.
Alors $G' - 1$ est une fonction croissante de $p_1 - 1 < 0$ (car $p_0 > 0$) à $m - 1 > 0$, donc elle s'annule en un point $\alpha \in]0,1[$.
La fonction $G-\mathrm{Id}$ est alors décroissante sur $[0,\alpha]$ puis croissante sur $[\alpha,1]$.
Comme $G(0) - 0 = p_0 > 0$ et $G(1) - 1 = 0$, il existe un point dans l'intervalle $]0,\alpha]$ où $G-\mathrm{Id}$ s'annule.
π_∞ est donc l'unique point fixe de G sur l'intervalle $]0,1[$ (car G en a au plus 2).

x	0		π_∞		α		1
$G'(x)-1$	p_1-1		$-$		0	$+$	$m-1$
$G(x)-x$	p_0	$\searrow$	0	$\searrow$		$\nearrow$	0

5. Effectivement, si cette fonction s'annule en trois points distincts, par le théorème de Rolle, sa dérivée s'annulera en deux points distincts, $a < b$. Mais $x \mapsto G(x) - x$ est convexe, donc sa dérivée est croissante, donc nulle sur $[a,b]$. Ceci implique que $G' - 1$ n'est pas strictement croissante, et donc que $x \mapsto G(x) - x$ n'est pas strictement convexe. Attention à l'idée reçue qui consiste à penser que "strictement convexe" équivaut à "dérivée seconde strictement positive" (pour les fonctions deux fois dérivables) – pour vous en convaincre, considérez $x \mapsto x^4$.

Supposons $m \leqslant 1$.

Alors $G' - 1$ est une fonction croissante sur $[0, 1]$, négative ou nulle en 1 ; donc négative sur $[0, 1]$.
Donc $G - \mathrm{Id}$ est décroissante sur $[0, 1]$, et s'annule en 1.
Comme cette fonction admet au plus 2 annulations, elle ne s'annule qu'en 1 (car sinon elle s'annulerait sur un intervalle non-réduit à un singleton).
Par conséquent, $\pi_\infty = 1$.

x	0		1
$G'(x) - 1$	$p_1 - 1$	$-$	$m - 1$
$G(x) - x$	p_0	$\searrow$	0

■