Processus de branchement et descendance d'un individu

Les processus de branchement sont des modèles introduits pour étudier le développement d'une population, dans laquelle les individus se reproduisent *indépendamment* les uns des autres.

Ces modèles sont particulièrement utilisés en biologie (étude de la croissance d'une colonie de bactéries...) et en physique nucléaire, mais trouvent leur origine dans l'étude, au $19^{\text{ème}}$ siècle, des probabilités d'extinction des noms de familles illustres en Grande Bretagne (Francis Galton et Henry Watson, 1874). Leur problème était le suivant :

Si un homme a une probabilité p_0 de n'avoir aucun fils, p_1 d'avoir un fils, p_2 d'en avoir deux, etc ; si chacun de ses fils éventuels est dans le même cas, et ainsi de suite ; quelle est la probabilité pour qu'à terme cette branche de la famille s'éteigne ? Plus généralement, comment connaître la probabilité pour qu'il y ait exactement k individus à la génération n ?

1 Processus de Galton-Watson

Le modèle classique consiste à introduire une suite $(\mathrm{Z}_n)_{n\in\mathbb{N}}$ de variables aléatoires, Z_n modélisant le nombre d'individus de la population à la génération n. On fait ensuite les hypothèses suivantes sur le comportement des individus :

- Tous les individus, quel que soit la génération à laquelle ils appartiennent, se comportent *a priori* de la même façon (*i.e.* leur nombre de fils a la même loi de probabilité.
- Le comportement d'un individu donné n'est en rien influencé par celui d'autres individus, que ce soit de sa propre génération ou des générations antérieures.
- Ce comportement n'est pas non plus affecté par le nombre d'individus.

En particulier, les descendances de deux individus distincts sont des phénomènes indépendants.

C'est donc un modèle assez simplificateur : par exemple, on ne considère pas des critères du type « quelqu'un ayant peu de frères a tendance à avoir lui-même moins d'enfants que la moyenne ». On néglige également les interactions entre individus de la même génération, par exemple la limitation du nombre d'individus par les ressources naturelles disponibles, critère important en écologie par exemple lorsqu'on étudie la croissance d'une population animale.

On se donne donc une suite $(p_k)_{k\in\mathbb{N}}$ de nombres positifs ou nuls, telle que $\sum\limits_{k=0}^{+\infty} p_k = 1$. Chaque individu, au cours du processus, aura une probabilité p_k

d'avoir k enfants.

En particulier, si l'on suppose de plus que le processus démarre à partir d'un seul individu à la génération 0 ($\mathrm{Z}_0 = 1$), la loi de probabilité du nombre d'individus à la génération suivante sera décrite par les nombres p_k :

$$\mathrm{P}(\mathrm{Z}_1 = k) = p_k \quad \forall k \in \mathbb{N}$$

Les individus de la génération 1 deviennent à leur tour parents, toujours selon la même loi de probabilités, et ainsi de suite. Si l'on a k individus à la génération 1, la loi de probabilité de Z_2 sera alors la même que celle d'une somme de k copies indépendantes de la variable aléatoire Z_1.

À la génération n, si l'on sait que $\mathrm{Z}_1 = k$, la loi de Z_n sera celle de la somme de k copies indépendantes de Z_{n-1}, puisque les descendances des individus de la première génération sont indépendantes, et de même loi, qui n'est autre que la loi de probabilité de la descendance du premier ancêtre (celui de la génération 0).

N.B. De la même façon, sachant que $\mathrm{Z}_{n-1} = k$, la loi de probabilité de Z_n est celle de la somme de k copies indépendantes de la variable aléatoire Z_1.

C'est la propriété de Markov : le comportement de la variable Z_{n+1}, si l'on connaît Z_n, est absolument indépendant de $\mathrm{Z}_0, \mathrm{Z}_1, \ldots, \mathrm{Z}_{n-1}$, phénomène que l'on résume en disant que, connaissant le présent, le futur est indépendant du passé. La suite de variables aléatoires $(\mathrm{Z}_n)_{n\in\mathbb{N}}$ constitue donc ce que l'on appelle une *chaîne de Markov*.

Pour une introduction aux chaînes de Markov, on pourra consulter [2], qui contient par ailleurs une étude plus approfondie des processus des branchements, tout comme [3].

2 Généralités sur les fonctions génératrices

Une bonne façon de décrire une variable aléatoire X à valeurs entières est de donner sa fonction génératrice $\mathrm{F_X}$, définie par :

$$\mathrm{F_X}(s) = \sum_{k=0}^{+\infty} \mathrm{P}(\mathrm{X} = k) s^k$$

Une telle fonction contient toutes les informations disponibles sur la loi de la variable aléatoire X, puisque l'on peut retrouver tous les nombres $\mathrm{P}(\mathrm{X} = k)$ à partir de $\mathrm{F_X}$. En outre, la fonction génératrice $\mathrm{F_X}$ de la variable aléatoire X nous donne aisément, si elles existent, l'espérance et la variance de X.

Lemme 2.1 *La fonction* $\mathrm{F_X}$ *est convexe et strictement croissante sur* $[\,0\,;1\,]$*, elle vérifie* $\mathrm{F_X}(0) = \mathrm{P}(\mathrm{X} = 0) \geq 0$ *et* $\mathrm{F}(1) = 1$.

Démonstration : Rappelons que la fonction $\mathrm{F_X}$ est définie par

$$\mathrm{F_X} : s \mapsto \sum_{k=0}^{+\infty} \mathrm{P}(\mathrm{X} = k) s^k$$

où les nombre $\mathrm{P}(\mathrm{X} = k)$ sont tous positifs ou nuls, de somme 1.

D'où $\qquad \mathrm{F_X}(1) = \sum_{k=0}^{+\infty} \mathrm{P}(\mathrm{X} = k) = 1$, et $\mathrm{F_X}(0) = \mathrm{P}(\mathrm{X} = 0) \geq 0$

Chaque terme de la série définissant une fonction strictement croissante, la fonction F_X est donc elle-même strictement croissante. De plus, F_X étant une série entière de rayon de convergence au moins 1, on peut dériver terme à terme pour obtenir :

$$\forall s \in [\,0\,;1\,[\,, \quad F'_X(s) = \sum_{k=1}^{+\infty} k p_k s^{k-1}$$

Ceci montre que la fonction F'_X est elle-même croissante sur $[\,0\,;1\,[$, et donc la fonction F_X est convexe sur $[\,0\,;1\,]$.

Reste uniquement à démontrer que F_X est continue en 1. Il suffit pour cela de voir que, pour tout entier N et tout réel $s \in [\,0\,;1\,[$, on a :

$$\sum_{k=0}^{N} P(X = k)s^k \leq F_X(s) \leq 1$$

F_X étant croissante et majorée, elle admet une limite en 1. Passant à la limite en s, il vient

$$\sum_{k=0}^{N} P(X = k) \leq \lim_{s\to 1} F_X(s) \leq 1$$

Passant enfin à la limite en N, on conclut par le théorème des gendarmes

$$\lim_{s\to 1} F_X(s) = 1 = F_X(1) \qquad \square$$

N.B. Si $p_0 + p_1 < 1$, c'est-à-dire si l'un des nombres p_i est non nul pour un entier $i \geq 2$, alors F_X est strictement convexe. En effet, la série définissant F'_X coporte alors (au moins) un terme non constant, donc F'_X est alors strictement croissante, et F_X strictement convexe.

Le cas $p_0 + p_1 = 1$ correspond à un processus dans lequel un individu a au plus un enfant. La fonction F_X est alors une fonction affine (linéaire si $p_1 = 1$, cas peu intéressant...)

Théorème 2.2 *Si la variable aléatoire* X *possède une espérance (resp. une variance) finie, alors on obtient cette espérance (puis cette variance) à partir de la fonction génératrice de* X *par la relation :*

$$\mathbb{E}(X) = F'_X(1) \qquad \textit{(puis } Var(X) = F''_X(1) - \mathbb{E}(X)^2 + \mathbb{E}(X)\textit{)}$$

Démonstration : En effet, l'espérance $\mathbb{E}(X)$ de la variable aléatoire entière X, lorsqu'elle existe, est définie par :

$$\mathbb{E}(X) = \sum_{k=0}^{+\infty} P(X = k)k$$

Or, la fonction F_X est définie comme une série entière, à coefficient positifs et majorés (par 1 puisqu'ils sont de somme 1). Le rayon de convergence de la série est donc au moins 1, et sur l'intervalle $[\,0\,;1\,[$, on peut dériver terme à terme la série pour obtenir :

$$F'_X(s) = \sum_{k=1}^{+\infty} P(X = k)ks^{k-1}$$

(pour la valeur $s = 1$, la série peut ne pas converger, par exemple pour la suite $(p_k)_{k\in\mathbb{N}}$ définie par $p_{2^n} = \dfrac{1}{2^n}$.)

Lorsque l'espérance $\mathbb{E}(\mathrm{X})$ existe, on a la relation $\lim\limits_{s\to 1} \mathrm{F}'_{\mathrm{X}}(s) = \mathbb{E}(\mathrm{X})$. En effet, la fonction F'_{X} est alors croissante sur $[\,0\,;1\,[$, majorée (par la valeur en 1 de la série entière, c'est-à-dire $\mathbb{E}(\mathrm{X}) = \sum\limits_{k=1}^{+\infty} \mathrm{P}(\mathrm{X}=k)k$), donc admet une limite en 1.

Or, pour tout entier N et pour tout réel $s \in [\,0\,;1\,[$, on a l'encadrement :

$$\sum_{k=0}^{\mathrm{N}} \mathrm{P}(\mathrm{X}=k)ks^{k-1} \leq \mathrm{F}'_{\mathrm{X}}(s) \leq \sum_{k=1}^{+\infty} \mathrm{P}(\mathrm{X}=k)k$$

soit, en passant à la limite en s,

$$\sum_{k=0}^{\mathrm{N}} \mathrm{P}(\mathrm{X}=k)k \leq \lim_{s\to 1} \mathrm{F}'_{\mathrm{X}}(s) \leq \sum_{k=1}^{+\infty} \mathrm{P}(\mathrm{X}=k)k$$

Ceci étant vrai pour tout N, on a donc bien $\lim\limits_{s\to 1} \mathrm{F}'_{\mathrm{X}}(s) = \mathbb{E}(\mathrm{X})$. La fonction F_{X} étant continue sur l'intervalle $[\,0\,;1\,]$ et de classe $\mathcal{C}^1$ sur $[\,0\,;1\,[$, le théorème de limite des dérivées nous permet alors d'affirmer que F_{X} est dérivable en 1, de dérivée $\mathrm{F}'_{\mathrm{X}}(1) = \lim\limits_{s\to 1} \mathrm{F}'_{\mathrm{X}}(s)$.

On a bien $$\mathrm{F}'_{\mathrm{X}}(1) = \mathbb{E}(\mathrm{X})$$

Passons à la variance. Si elle existe, l'espérance est finie, et elle est égale à :

$$\mathrm{Var}(\mathrm{X}) = \sum_{k=0}^{+\infty} \mathrm{P}(\mathrm{X}=k)k^2 - \left(\sum_{k=0}^{+\infty} \mathrm{P}(\mathrm{X}=k)k\right)^2$$

Alors la somme $\sum\limits_{k\geq 0} \mathrm{P}(\mathrm{X}=k)k(k-1)$ converge elle aussi (elle est majorée par $\sum\limits_{k\geq 0} \mathrm{P}(\mathrm{X}=k)k^2$, qui converge), et tout comme pour la dérivée de F_{X}, on montre que la dérivée seconde de F_{X} tend vers $\sum\limits_{k=0}^{+\infty} \mathrm{P}(\mathrm{X}=k)k(k-1)$ lorsque s tend vers 1. Donc F_{X} est deux fois dérivable en 1 et l'on a :

$$\begin{aligned}
\mathrm{F}''_{\mathrm{X}}(1) &= \sum_{k=0}^{+\infty} \mathrm{P}(\mathrm{X}=k)k(k-1) \\
&= \sum_{k=0}^{+\infty} \mathrm{P}(\mathrm{X}=k)k^2 - \sum_{k=0}^{+\infty} \mathrm{P}(\mathrm{X}=k)k \\
&= \mathbb{E}\left(\mathrm{X}^2\right) - \mathbb{E}(\mathrm{X}) \\
\mathrm{F}''_{\mathrm{X}}(1) &= \mathrm{Var}(\mathrm{X}) + \mathbb{E}(\mathrm{X})^2 - \mathbb{E}(\mathrm{X})
\end{aligned}$$

c'est-à-dire $$\mathrm{Var}(\mathrm{X}) = \mathrm{F}''_{\mathrm{X}}(1) - \mathbb{E}(\mathrm{X})^2 + \mathbb{E}(\mathrm{X}) \qquad \square$$

Enfin, les fonctions génératrice possèdent la propriété intéressante suivante :

Propriété 2.3 *Soient* X *et* Y *deux variables aléatoires indépendantes, à valeurs dans* $\mathbb{N}$. *Leur somme* $\mathrm{X}+\mathrm{Y}$ *est encore une variable aléatoire à valeurs entières, et sa fonction génératrice est :*

$$\mathrm{F}_{\mathrm{X}+\mathrm{Y}} = \mathrm{F}_{\mathrm{X}} \times \mathrm{F}_{\mathrm{Y}}$$

Démonstration : Soit k un entier naturel. La somme $\mathrm{X}+\mathrm{Y}$ prend la valeur k si et seulement si les variables aléatoires X et Y prennent des valeurs i et j vérifiant $i+j=k$. C'est-à-dire que

$$\mathrm{P}(\mathrm{X}+\mathrm{Y}=k)=\sum_{i=0}^{k}\mathrm{P}(\mathrm{X}=i,\mathrm{Y}=k-i)$$

Les deux variables X et Y étant indépendantes, on a pour tout i la propriété $\mathrm{P}(\mathrm{X}=i,\mathrm{Y}=k-i)=\mathrm{P}(\mathrm{X}=i)\times\mathrm{P}(\mathrm{Y}=k-i)$, et donc

$$\mathrm{P}(\mathrm{X}+\mathrm{Y}=k)=\sum_{i=0}^{k}\mathrm{P}(\mathrm{X}=i)\times\mathrm{P}(\mathrm{Y}=k-i)$$

Utilisant cette égalité dans la formule définissant la fonction génératrice de $\mathrm{X}+\mathrm{Y}$, on obtient :

$$\begin{aligned}\mathrm{F}_{\mathrm{X}+\mathrm{Y}}(s) &= \sum_{k=0}^{+\infty}\mathrm{P}(\mathrm{X}+\mathrm{Y}=k)\,s^k\\ &= \sum_{k=0}^{+\infty}\sum_{i=0}^{k}\mathrm{P}(\mathrm{X}=i)\,s^i\times\mathrm{P}(\mathrm{Y}=k-i)\,s^{k-i}\\ &= \left(\sum_{i=0}^{+\infty}\mathrm{P}(\mathrm{X}=i)\,s^i\right)\left(\sum_{j=0}^{+\infty}\mathrm{P}(\mathrm{Y}=j)\,s^j\right)\end{aligned}$$

soit $$\mathrm{F}_{\mathrm{X}+\mathrm{Y}}(s)=\mathrm{F}_{\mathrm{X}}(s)\mathrm{F}_{\mathrm{Y}}(s)$$ □

Bien sûr, ce résultat se généralise par une récurrence immédiate à la somme de n variables aléatoires indépendantes.

N.B. On peut voir dans la fonction génératrice F_{X} l'expression de l'espérance de la variable aléatoire s^{X}, qui vaut s^k lorsque $\mathrm{X}=k$, c'est-à-dire avec la probabilité $\mathrm{P}(\mathrm{X}=k)$. Le résultat précédent est alors on ne peut plus naturel : si les deux variables aléatoires X et Y sont indépendantes, il en est de même pour s^{X} et s^{Y}. Et alors, l'espérance du produit est le produit des espérances :

$$\mathbb{E}\left(s^{\mathrm{X}+\mathrm{Y}}\right)=\mathbb{E}\left(s^{\mathrm{X}}\times s^{\mathrm{Y}}\right)=\mathbb{E}\left(s^{\mathrm{X}}\right)\times\mathbb{E}\left(s^{\mathrm{Y}}\right)$$

c'est-à-dire $$\mathrm{F}_{\mathrm{X}+\mathrm{Y}}(s)=\mathrm{F}_{\mathrm{X}}(s)\mathrm{F}_{\mathrm{Y}}(s)$$

3 Cas d'un processus de branchement

3.1 Fonctions génératrices des variables Z_n

Définition Soit $\mathrm{F}: s\mapsto\sum_{k=0}^{+\infty}p_k s^k$ la fonction génératrice de la variable aléatoire Z_1. On définit les itérées de la fonction F en posant pour tout s :

$$\begin{aligned}\mathrm{F}_0(s) &= s\\ \mathrm{F}_1(s) &= \mathrm{F}(s)\\ \mathrm{F}_n(s) &= \mathrm{F}\left(\mathrm{F}_{n-1}(s)\right)\quad\text{pour } n\geq 2\end{aligned}$$

Théorème 3.1 *Pour tout entier n, la variable aléatoire Z_n, qui représente le nombre d'individus à la génération n, a pour fonction génératrice F_n.*

Démonstration : Nous allons bien sur procéder par récurrence.

– Le résultat est évident pour $n = 0$ ou 1 : la variable Z_0 est égale à 1, avec une probabilité 1, donc a pour fonction génératrice $F_{Z_0}(s) = s$.
Quant à la variable Z_1, elle a par définition pour fonction génératrice $F = F_1$.
– Soit donc $n \geq 2$, et supposons le résultat établi pour tout entier $m < n$.
Sous la condition $Z_1 = i$, nous avons vu que la variable aléatoire Z_n est équivalente à la somme de i copies indépendantes de Z_{n-1}. Par hypothèse de récurrence, Z_{n-1} a pour fonction génératrice F_{n-1}, et la propriété 2.3 nous permet d'affirmer que la somme de i copies indépendantes de Z_{n-1} a pour fonction génératrice $(F_{n-1})^i$.
Soit i un entier tel que $P(Z_1 = i) > 0$. Sous la condition $Z_1 = i$, la fonction génératrice de Z_n est donc $(F_{n-1})^i$. C'est-à-dire que l'on peut dons écrire pour tout entier i

$$\sum_{k=0}^{+\infty} P\left(Z_n = k | Z_1 = i\right) s^k = \left(F_{n-1}(s)\right)^i$$

(Si $P(Z_1 = i) = 0$, on ne peut pas définir la probabilité conditionnelle $P\left(Z_n = k | Z_1 = i\right)$). La formule des probabilités conditionnelles $P(A, B) = P(A|B) \times P(B)$ nous permet alors de décomposer la probabilité $P\left(Z_n = k\right)$ comme suit :

$$\begin{aligned} P\left(Z_n = k\right) &= \sum_{\substack{i=0 \\ P(Z_1=i)\neq 0}}^{+\infty} P\left(Z_n = k, Z_1 = i\right) \\ &= \sum_{\substack{i=0 \\ P(Z_1=i)\neq 0}}^{+\infty} P\left(Z_n = k | Z_1 = i\right) \times P\left(Z_1 = i\right) \end{aligned}$$

La fonction génératrice de Z_n s'écrit donc :

$$\begin{aligned} \sum_{k=0}^{+\infty} P\left(Z_n = k\right) s^k &= \sum_{k=0}^{+\infty} \sum_{\substack{i=0 \\ P(Z_1=i)\neq 0}}^{+\infty} P\left(Z_n = k | Z_1 = i\right) \times P\left(Z_1 = i\right) s^k \\ &= \sum_{\substack{i=0 \\ P(Z_1=i)\neq 0}}^{+\infty} P\left(Z_1 = i\right) \left(\sum_{k=0}^{+\infty} P\left(Z_n = k | Z_1 = i\right) s^k\right) \\ &= \sum_{\substack{i=0 \\ P(Z_1=i)\neq 0}}^{+\infty} P\left(Z_1 = i\right) \left(F_{n-1}(s)\right)^i \end{aligned}$$

Lorsque $P\left(Z_1 = i\right) = 0$, on a aussi $P\left(Z_1 = i\right) \left(F_{n-1}(s)\right)^i = 0$, et donc

$$\sum_{k=0}^{+\infty} P\left(Z_n = k\right) s^k = \sum_{i=0}^{+\infty} P\left(Z_1 = i\right) \left(F_{n-1}(s)\right)^i$$

On reconnaît dans cette dernière formule l'expression de $F\left(F_{n-1}(s)\right)$, c'est-à-dire de $F_n(s)$. Et la variable aléatoire Z_n a donc bien pour fonction génératrice F_n.
– On conclut par récurrence : pour tout n, Z_n a pour fonction génératrice F_n.

$\square$

En pratique, ceci nous permet, dans des cas simples, de calculer explicitement la probabilité pour qu'il y ait k individus à la génération n.

Exemple : Supposons que la loi de la variable aléatoire Z_1 soit de la forme :

$$P(Z_1 = i) = \frac{1}{2^{i+1}}$$

(La somme des probabilités fait bien $\sum_{k=0}^{+\infty} 1/2^{k+1} = 1$.) Alors la variable Z_1 a pour fonction génératrice :

$$F : s \mapsto \sum_{k=0}^{+\infty} \frac{s^k}{2^{k+1}} = \frac{1}{2}\frac{1}{1-s/2} = \frac{1}{2-s}$$

Pour calculer les itérées de F, on procède comme pour les suites récurrentes définies par une homographie ; on cherche d'abord les points fixes de F :

$$\begin{aligned} F(s) = s &\iff s(2-s) = 1 \\ &\iff s^2 - 2s + 1 = 0 \\ &\iff s = 1 \end{aligned}$$

F possède une racine double, on va donc chercher à exprimer la quantité $\dfrac{1}{F(s)-1}$ en fonction de $\dfrac{1}{s-1}$.

Or
$$\frac{1}{F(s)-1} = \frac{1}{\dfrac{1}{2-s} - 1} = \frac{2-s}{s-1} = \frac{1}{s-1} - 1$$

Par récurrence il vient, pour tout n, $\dfrac{1}{F_n(s)-1} = \dfrac{1}{s-1} - n$.

et donc
$$F_n(s) = \frac{n-(n-1)s}{(n+1)-ns}$$

Il suffit alors de développer cette expression en série entière pour trouver, quels que soient k et n, la probabilité pour qu'il y ait k individus à la génération n.

On trouve
$$\begin{aligned} F_n(s) &= \frac{n-1}{n} + \frac{1}{n(n+1)} \cdot \frac{1}{1 - \dfrac{n}{n+1}s} \\ &= \frac{n-1}{n} + \frac{1}{n(n+1)} \sum_{k=0}^{+\infty} \left(\frac{ns}{n+1}\right)^k \end{aligned}$$

Pour $k > 0$, il vient
$$P(Z_n = k) = \frac{n^{k-1}}{(n+1)^{k+1}}$$

3.2 Probabilité d'extinction

Une des première questions que l'on peut se poser à propos d'un tel processus est de savoir si oui ou non il va continuer indéfiniment. Par exemple, si le premier individu n'a pas d'enfants, ce processus s'arrête tout de suite. Pour la même raison, le processus peut s'arrêter à chaque génération.

La question est alors de savoir selon quelle probabilité la descendance va s'éteindre. Dans toute la suite de ce texte, on supposera donc $p_0 > 0$ (sinon, tout individu a un nombre strictement positif d'enfants, et le processus ne peut pas s'arrêter).

Remarquons tout d'abord que la probabilité pour qu'une variable aléatiore entière X soit nulle se calcule à partir de sa fonction génératrice F_X par :

$$P(X = 0) = F_X(0)$$

En effet $$F_X(0) = \sum_{k=0}^{+\infty} P(X = k)\, 0^k = P(X = 0)$$

Revenant à notre processus, ceci signifie que la probabilité pour que la descendance ait disparu à la génération n est $F_n(0)$.

La suite $(F_n(0))_{n\in\mathbb{N}}$ est donc croissante (si la génération n n'existe pas, il en va de même pour la génération $n+1$), et bien sûr majorée par 1. Elle tend donc vers une limite $\pi \leq 1$, qui représente la *probabilité d'extinction* du processus :

$$\begin{aligned}\pi &= \lim_{n\to\infty} F_n(0)\\ &= \lim_{n\to\infty} P(Z_i = 0 \text{ pour un } 1 \leq i \leq n)\\ &= P(Z_i = 0 \text{ pour un entier } i)\end{aligned}$$

Exemple : Dans l'exemple précédemment décrit ($P(Z_1 = k) = 1/2^{k+1}$ pour tout entier k), on obtient une formule explicite pour $F_n(0)$, qui n'est autre que le coefficient constant de la série entière définissant F_n :

$$F_n(0) = \frac{n-1}{n} + \frac{1}{n(n+1)} = \frac{n}{n+1}$$

En particulier, on a ici $$\pi = \lim_{n\to\infty} F_n(0) = 1$$

Théorème 3.2 *π est le plus petit point fixe de la fonction* F.

Démonstration : C'est un résultat classique sur les suites récurrentes. π est la limite de la suite $(F_n(0))_{n\in\mathbb{N}}$, suite qui peut être définie par récurrence par les relations :

$$\begin{cases} F_0(0) = 0 \\ F_{n+1}(0) = F(F_n(0)) \quad \forall n \in \mathbb{N}\end{cases}$$

Nous avons déjà vu que la suite $(F_n(0))_{n\in\mathbb{N}}$ est croissante (ce qui se retrouve aisément en disant que la fonction F est elle-même croissante, et comme on sait que $F_1(0) = p_0 > 0 = F_0(0)$, on obtient la croissance de la suite par récurrence).

Comme l'intervalle $[0;1]$ est stable par la fonction F (celle-ci est croissante, donc à valeurs dans $[F(0);F(1)] \subset [0;1]$), la suite $(F_n(0))_{n\in\mathbb{N}}$ est majorée, et donc elle converge vers une limite π.

Passant à la limite en n dans l'équation $F_{n+1}(0) = F(F_n(0))$, on obtient alors $F(\pi) = \pi$, c'est-à-dire que π est comme annoncé un point fixe de F.

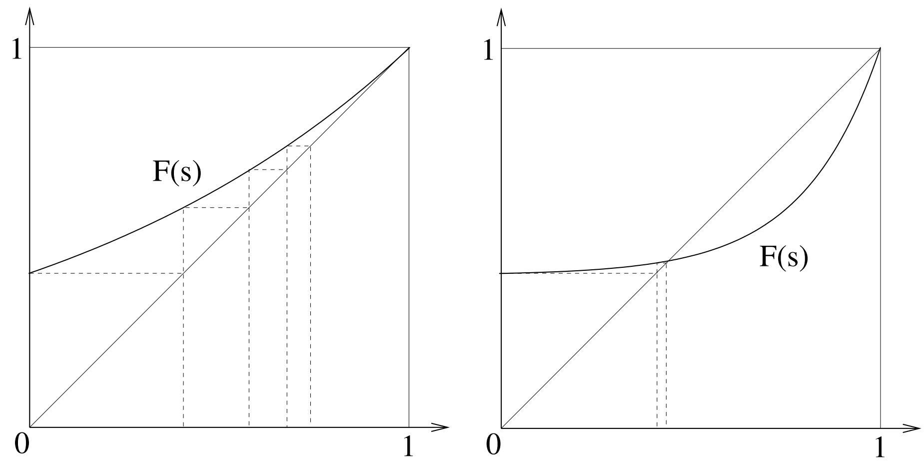
Soit x_0 le plus petit point fixe de F, dans l'intervalle $[0;1]$. Comme la fonction F est croissante, on a par récurrence :

- $F_0(0) = 0 \leq x_0$
- Si $F_n(0) \leq x_0$ alors $F(F_n(0)) \leq F(x_0)$, c'est-à-dire que $F_{n+1}(0) \leq x_0$.
- Par le principe de récurrence, on en déduit que pour tout n, $F_n(0) \leq x_0$.
- Par passage à la limite, on en déduit que $\pi \leq x_0$.

π est donc le plus petit point fixe de F dans l'intervalle $[0;1]$.

□

La probabilité d'extinction du processus est donc intimement liée aux propriétés de la fonction F. On sait que celle-ci est strictement convexe sur $[\,0\,;1\,]$, et vérifie $F(1) = 1$. Deux cas se présentent alors : soit la courbe de F est au dessus (strictement) de la droite d'équation $y = x$ sur l'intervalle $[\,0\,;1\,[$, soit elle la traverse en un autre point. Ces deux situations, ainsi que les premiers termes de la suite $(F_n(0))_{n\in\mathbb{N}}$, sont représentées par les deux graphes suivant.



Graphiquement, on s'aperçoit que, dans le premier cas, la suite $(F_n(0))_{n\in\mathbb{N}}$ tend vers 1, tandis qu'elle tend vers la première intersection de la courbe avec la droite d'équation $y = x$ dans le second cas. Dans les deux cas, la limite π est bien le plus petit point fixe de F.

Nous allons maintenant voir comment distinguer ces deux cas à partir de la fonction F, c'est-à-dire connaissant la loi de reproduction des individus.

Théorème 3.3 *Si la variable aléatoire* Z_1 *possède une espérance* m*, et que* $m \leq 1$*, alors* 1 *est le seul point fixe de* F *sur l'intervalle* $[\,0\,;1\,]$*, et la probabilité d'extinction est* 1.

Sinon, F *admet exactement un point fixe sur* $[\,0\,;1\,[$*, qui est la probabilité d'extinction du processus.*

Autrement dit, si chaque individu a en moyenne un enfant ou moins, la descendance s'arrêtera un jour avec une probabilité 1. Si au contraire chaque individu a en moyenne strictement plus d'un enfant, le processus peut bien sûr s'éteindre, mais il y a aussi des chances pour que la population survive indéfiniment.

N.B. Le premier cas est appelé *sous-critique* (*critique* lorsque $m = 1$), le second est appelé *sur-critique.*

Démonstration : Remarquons tout d'abord que, lorsque l'espérance m de Z_1 est bien définie, sa fonction génératrice F est dérivable en 1, et l'on a $F'(1) = m$.

- Dans le cas sous-critique ($m < 1$), comme F est convexe, la courbe de F est au-dessus de sa tangente au point de coordonnées $(1, 1)$, qui est elle-même strictement au-dessus de la droite d'équation $y = x$ sur l'intervalle $[\,0\,;1\,[$.
  Donc F n'admet pas de point fixe sur cet intervalle, et 1 est le seul point fixe de F. La probabilité d'extinction du processus est donc forcément 1.

– Dans le cas critique ($m = 1$), comme on a supposé $p_0 > 0$, on a aussi forcément un entier $n \geq 2$ tel que $p_n > 0$ (sinon $p_0 + p_1 = 1$, et l'espérance de Z_1 est $p_1 < 1$). En particulier F est alors strictement convexe, donc strictement au-dessus de sa tangente au point de coordonnées $(1, 1)$, qui est alors la droite d'équation $y = x$.

– Enfin, dans le cas sur-critique, l'espérance m est strictement supérieure à 1, voire non définie (quand la somme $\sum_{k \geq 0} \mathrm{P}(\mathrm{Z}_1 = k)k$ est infinie). La courbe représentative de F admet alors au point $(1, 1)$ une tangente de pente $m > 1$ (resp. verticale).
  Comme F est alors strictement convexe, la fonction $x \mapsto \dfrac{1 - \mathrm{F}(x)}{1 - x}$ est strictement croissante sur $[\,0\,;1\,[$, et continue car F l'est. Elle prend en 0 la valeur $1 - p_0 \leq 1$, et en 1 a pour limite $m > 1$ (resp. $+\infty$). Par le théorème des valeurs intermédiaires, elle prend donc la valeur 1 en un point de l'intervalle $[\,0\,;1\,[$, ce point étant unique comme la fonction est strictement croissante. Ce point est alors l'unique point fixe de F sur l'intervalle $[\,0\,;1\,[$.
  Comme la probabilité d'extinction du processus est le plus petit point fixe de F sur $[\,0\,;1\,]$, c'est nécessairement celui-ci, puisqu'il est inférieur à 1.

$\square$

Exemple : Reprenant toujours le même exemple ($\mathrm{P}(\mathrm{Z}_1 = k) = 1/2^{k+1}$ pour tout entier k), on a $\mathrm{F}(s) = \dfrac{1}{2 - s}$, et donc $m = \mathrm{F}'(1) = 1$. Nous sommes donc dans le cas critique, et la probabilité d'extinction est bien égale à 1.

4 Croissance dans le cas sur-critique

4.1 Généralités

Nous avons vu que, dans les cas sous-critique et critique, le processus s'éteint avec une probabilité 1, alors que cette probabilité d'extinction est strictement inférieure à 1 dans le cas sur-critique.

La question qui se pose alors est de savoir ce qui se passe, dans le cas sur-critique, lorsque le processus ne s'éteint pas. Nous nous placerons désormais dans le cas où la variable Z_1 a une espérance $m > 1$ (espérance finie).

Tout d'abord, on a vu que la fonction génératrice d'une variable aléatoire permet de calculer son espérance, grâce à la relation $\mathbb{E}(\mathrm{X}) = \mathrm{F}'_{\mathrm{X}}(1)$.

Propriété 4.1 *Z_n a pour espérance m^n, où m est l'espérance de Z_1.*

Démonstration : Comme Z_n a pour fonction génératrice la n-ième itérée de F, on peut calculer par récurrence

$$\mathrm{F}'_n(s) = \mathrm{F}' \circ \mathrm{F}_{n-1}(s) \times \mathrm{F}'_{n-1}(s)$$

Et donc $$\mathrm{F}'_n(1) = \mathrm{F}'\left(\mathrm{F}_{n-1}(1)\right) \times \mathrm{F}'_{n-1}(1) = \mathrm{F}'(1) \times \mathrm{F}'_{n-1}(1)$$

Par récurrence, on montre alors aisément que

$$\mathbb{E}(\mathrm{Z}_n) = \mathrm{F}'_n(1) = \mathrm{F}'(1)^n = m^n$$ $\square$

Ceci nous renseigne déjà de façon assez précise sur le comportement de la suite Z_n : celle-ci a « en moyenne » une croissance exponentielle. Par un effet de vases communiquant, on peut donc s'attendre à ce que, lorsque le processus ne s'éteint pas, la population grandisse indéfiniment. C'est ce que nous allons montrer maintenant.

4.2 États transitoires

Dans toute cette partie, k est un entier non nul désormais fixé. Nous allons montrer qu'avec une probabilité égale à 1, la suite $(Z_n)_{n\in\mathbb{N}^*}$ ne prend la valeur k qu'un nombre fini de fois.

N.B. On résume cette propriété en disant que k est un état *transitoire* pour la chaine de Markov $(Z_n)_{n\in\mathbb{N}^*}$. A l'opposé, si la suite $(Z_n)_{n\in\mathbb{N}^*}$ prenait la valeur k une infinité de fois, et ce avec une probabilité 1, on parlerait d'état *récurrent.*

Pour ceci, nous allons étudier en réalité un deuxième processus de branchement, légèrement différent : on suppose qu'il y a k individus à la génération 0 (à part cette hypothèse, tout fonctionne comme précédemment). Pour distinguer ce processus du précédent, on notera Z'_n le nombre d'individus à la génération n.

Nous allons définir une suite de nombres positifs $(u_n)_{n\in\mathbb{N}^*}$ par :

$$u_n = P\left(Z'_n = k \text{ et } Z'_i \neq k \text{ pour tout } 0 < i < n\right)$$

C'est-à-dire que u_n est la probabilité pour que la suite $(Z'_n)_{n\in\mathbb{N}}$ prenne la valeur k pour la première fois au rang n (génération 0 exclue).

Toutes ces possibilités s'excluant, on a forcément $u = \sum\limits_{n=1}^{+\infty} u_n \leq 1$.

Lemme 4.2 *La probabilité pour que la suite $(Z'_n)_{n\in\mathbb{N}}$ ne prenne pas la valeur k est non nulle, donc $u < 1$.*

Démonstration : Rappelons que l'on a supposé $p_0 > 0$. La probabilité pour que le processus s'éteigne dès la première génération est donc non nulle, et dans ce cas la suite $(Z'_n)_{n\in\mathbb{N}}$ ne prendra jamais la valeur k.

□

N.B. Même sans l'hypothèse $p_0 > 0$, on a toujours ce résultat. Car, si $p_0 = 0$, la suite $(Z'_n)_{n\in\mathbb{N}}$ est croissante. Il suffit donc qu'il y ait strictement plus de k individus à la génération 1 pour que la suite ne prenne plus jamais la valeur k. Or l'hypothèse $m > 1$ implique que tout individu peut avoir strictement plus d'un enfant, donc l'événement $Z_1 > k$ arrive avec une probabilité non nulle.

Soit alors U la série génératrice de la suite $(u_n)_{n\in\mathbb{N}^*}$:

$$U(s) = \sum_{k=1}^{+\infty} u_n s^n$$

On définit de la même manière les nombres $u_n^{(r)}$ $(n, r \in \mathbb{N}^*)$ comme la probabilité pour que la suite $(Z'_n)_{n\in\mathbb{N}^*}$ prenne la valeur k pour la r-ième fois au rang n.

Théorème 4.3 *Pour tout entier r strictement positif, la série génératrice de la suite* $\left(u_n^{(r)}\right)_{n\in\mathbb{N}^*}$ *est la fonction* U^r.

Démonstration : Par récurrence sur r. Le résultat est évident pour $r = 1$ par définition de la fonction U.

Supposons le résultat établi jusqu'au rang $r-1$. Soit n un entier naturel non nul. Alors l'événement « la suite $(\mathrm{Z}'_n)_{n\in\mathbb{N}^*}$ prend la valeur k pour la r-ième fois au rang n » peut se décomposer de la façon suivante :

- La suite $(\mathrm{Z}'_n)_{n\in\mathbb{N}^*}$ prend la valeur k pour la première fois au rang 1, puis elle prend la valeur k exactement $r-1$ fois jusqu'au rang n.
  Ceci arrive avec une probabilité $u_1 \times u_{n-1}^{(r-1)}$. En effet, si la valeur k est atteinte au rang 1, la suite $(\mathrm{Z}'_n)_{n\geq 2}$ a après cela exactement la même loi que le processus de branchement partant de k individus. La probabilité pour qu'elle prenne la valeur k au rang n pour la $r-1$-ième fois est donc $u_{n-1}^{(r-1)}$.
- La suite $(\mathrm{Z}'_n)_{n\in\mathbb{N}^*}$ prend la valeur k pour la première fois au rang 2, puis elle prend la valeur k exactement $r-1$ fois jusqu'au rang n.
  Ceci arrive avec une probabilité $u_2 \times u_{n-2}^{(r-1)}$.
- Et ainsi de suite, jusqu'au cas où la suite prend la valeur k pour la première fois au rang $n-1$…

Toutes ces possibilités s'excluent, et donc notre événement a pour probabilité la somme de ces probabilités :

$$u_n^{(r)} = u_1 \times u_{n-1}^{(r-1)} + u_2 \times u_{n-2}^{(r-1)} + \cdots + u_{n-1} \times u_1^{(r-1)}$$

On reconnait dans l'expression précédente le produit de Cauchy des suites $(u_n)_{n\in\mathbb{N}}$ et $\left(u_n^{(r-1)}\right)_{n\in\mathbb{N}}$.

Par hypothèse de récurrence, la série génératrice de la suite $\left(u_n^{(r-1)}\right)_{n\in\mathbb{N}}$ est la fonction U^{r-1}. Et par définition, celle de la suite $(u_n)_{n\in\mathbb{N}}$ est la fonction U. La série génératrice de la suite $\left(u_n^{(r)}\right)_{n\in\mathbb{N}}$ est donc U^r.

On conclut par le principe de récurrence.

□

N.B. On vient de voir l'intérêt de considérer un processus de branchement qui commence avec exactement k individus : lorsque la suite $(\mathrm{Z}'_n)_{n\in\mathbb{N}^*}$ prend la valeur k, elle a par la suite la même loi de probabilité qu'à partir de la première génération. Par exemple, la génération qui vient juste après cette valeur k a la même loi que la génération 1, et ainsi de suite. Ce ne serait bien sûr pas le cas si l'on considérait un processus commençant par un seul individu.

Nous sommes désormais en mesure de montrer le

Théorème 4.4 *La probabilité pour que la suite* $(\mathrm{Z}'_n)_{n\in\mathbb{N}^*}$ *prenne la valeur* k *une infinité de fois est nulle.*

Nous repoussons la démonstration de ce théorème pour énoncer son corollaire, qui constitue le résultat principal de cette partie.

Corollaire 4.5 *Il en va de même pour la suite* $(\mathrm{Z}_n)_{n\in\mathbb{N}^*}$.

Démonstration : Si n_1 désigne le rang de la première génération de taille exactement k, alors la loi de $(\mathrm{Z}_n)_{n>n_1}$ est celle d'un processus de branchement à partir de k individus, c'est-à-dire la loi de $(\mathrm{Z}'_n)_{n\in\mathbb{N}^*}$. La probabilité pour que la suite $(\mathrm{Z}_n)_{n>n_1}$ prenne ensuite une infinité de fois la valeur k est donc nulle.

Et donc la probabilité pour que la suite $(\mathrm{Z}_n)_{n\in\mathbb{N}^*}$ prenne la valeur k une infinité de fois est nulle.

$\square$

Démonstration du théorème : Soit r un entier naturel. Nous avons vu (théorème 4.3) que la série génératrice des $u_n^{(r)}$ est la fonction U^r, c'est-à-dire que

$$\forall s \in [\,0\,;1\,], \quad \sum_{n=1}^{+\infty} u_n^{(r)} s^n = \mathrm{U}(s)^r$$

En particulier, pour la valeur $s = 1$, on obtient la relation :

$$\sum_{n=1}^{+\infty} u_n^{(r)} = \mathrm{U}(1)^r = u^r$$

où u est la valeur de U en 1, c'est-à-dire $\sum\limits_{n=1}^{+\infty} u_n$.

Mais la somme $\sum\limits_{n=1}^{+\infty} u_n^{(r)}$ est la somme en n des probabilités pour que la suite $(\mathrm{Z}_n)_{n\in\mathbb{N}^*}$ prenne la valeur k pour la r-ième fois au rang n, c'est donc la probabilité pour que la suite $(\mathrm{Z}_n)_{n\in\mathbb{N}^*}$ prenne (au moins) r fois la valeur k. Et nous avons montré (lemme 4.2) que $u < 1$, et donc

$$\lim_{r\to+\infty} u^r = 0$$

La probabilité pour que la suite prenne la valeur k une infinité de fois étant inférieure à la probabilité pour qu'elle prenne la valeur k au moins r fois, et ce pour tout entier r, cette probabilité est donc nulle.

$\square$

N.B. De ce résultat, on peut déduire sans trop d'efforts la propriété suivante, valable pour tout k :

$$\lim_{n\to+\infty} \mathrm{P}\,(\mathrm{Z}_n = k) = 0$$

En effet, la négation de cette propriété revient à affirmer qu'il existe un réel $\varepsilon > 0$ tel qu'il existe une infinité d'entiers n vérifiant $\mathrm{P}\,(\mathrm{Z}_n = k) \geq \varepsilon$. Mais alors, on a pour tout entier m

$$\mathrm{P}\,(\exists n \geq m, \mathrm{Z}_n = k) \geq \varepsilon$$

(il suffit de considérer un entier $n \geq m$ tel que $\mathrm{P}\,(\mathrm{Z}_n = k) \geq \varepsilon\ldots$) Les événements « il existe $n \geq m$ tel que $\mathrm{Z}_n = k$ » sont des événements décroissants en m (s'il existe un tel n plus grand que $m+1$ alors cet entier n convient pour m également). La conjonction d'événement décroissants tous de probabilité supérieure à ε est encore de probabilité supérieure à ε (propriété générale d'une mesure de probabilité, cf [1] par exemple). C'est-à-dire que l'on a

$$\mathrm{P}\,(\forall m, \exists n \geq m, \mathrm{Z}_n = k) \geq \varepsilon$$

Mais la propriété « pour tout m, il existe $n \geq m$ tel que $\mathrm{Z}_n = k$ » signifie exactement qu'il y a une infinité d'entiers n tels que $\mathrm{Z}_n = k$, ce qui est de probabilité nulle d'après le corollaire 4.5 !

4.3 Croissance de la population

Nous venons de voir que la probabilité pour que la suite $(\mathrm{Z}_n)_{n\in\mathbb{N}^*}$ prenne la valeur k une infinité de fois est nulle, ce pour tout entier k. À l'inverse, cela signifie que la probabilité pour qu'elle prenne la valeur k un nombre fini de fois est 1.

Si N désigne un entier naturel, il y a donc une probabilité égale à 1 pour que la suite $(\mathrm{Z}_n)_{n\in\mathbb{N}^*}$ prenne une valeur comprise entre 1 et N un nombre fini de fois seulement. Ceci étant valable pour tout entier N, cela signifie qu'avec une probabilité 1 :

- Soit la suite prend la valeur 0, auquel cas le processus s'éteint.
- Soit la suite ne prend une valeur inférieure ou égale à N qu'un nombre fini de fois, et ce pour tout entier N, ce qui signifie que la suite $(\mathrm{Z}_n)_{n\in\mathbb{N}^*}$ tend vers $+\infty$.

N.B. On a utilisé au passage le fait que la conjonction d'un nombre dénombrable d'événements de probabilité 1 est encore de probabilité 1. C'est ce qui permet de passer de « pour tout N, la probabilité pour que la suite prenne une valeur comprise entre 1 et N un nombre fini de fois seulement est égale à 1 » à « la probabilité pour que pour tout N la suite prenne une valeur comprise entre 1 et N un nombre fini de fois seulement est égale à 1 ».

Comme l'on connaît déjà la probabilité d'extinction π, qui est le plus petit point fixe de la fonction génératrice de Z_1 ; la probabilité pour que la suite $(\mathrm{Z}_n)_{n\in\mathbb{N}^*}$ tende vers $+\infty$ est donc $1-\pi$. On a montré le

Théorème 4.6 $\mathrm{Z}_n \underset{n\to\infty}{\longrightarrow} +\infty$ *avec une probabilité* $1-\pi$*, où* π *est la probabilité d'extinction.*

Références

[1] P. Billingsley, *Probability and measure*, John Wiley & Sons, 1995.

[2] W. Feller, *An introduction to probability theory and its applications*, John Wiley & Sons, 1957.

[3] T. Harris, *The theory of branching processes*, Springer, 1963.