Processeur de Galton * – Watson †

Lecons 223, 226, 229, 260, 261, 264

Salim Rostam

23 décembre 2014

Soit X une variable aléatoire intégrable à valeurs dans $\mathbb{N}$; on note $p_n := \mathbb{P}(X = n)$ et $m := \mathbb{E}[X] = \sum_{n=1}^{+\infty} np_n$. Soit $(X_{i,j})$ une famille de variables aléatoire i.i.d de même loi que X. Soit (Z_n) la suite définie par :

$$Z_0 := 1$$

$$\forall n \in \mathbb{N}, Z_{n+1} := \sum_{i=1}^{Z_n} X_{i,n}$$

L'idée est de modéliser avec (Z_n) la taille d'une population : à l'instant n, il y a Z_n individus et chaque individu i a un nombre $X_{i,n}$ de descendants.

Ce développement étudie la suite (Z_n), en particulier s'il existe un n tel que $Z_n = 0$, *i.e.* la population s'éteint. Remarquons une première chose qui va nous servir dans la suite.

Propriété. *Pour $n \in \mathbb{N}^*$, la variable Z_n est indépendante de $X_{i,j}$ pour tout i et pour tout $j < n$.*

Démonstration. En effet, on peut remarquer que Z_n ne dépend que de Z_{n-1} et des $(X_{i,n-1})_i$ donc par récurrence, Z_n ne dépend que des $(X_{i,j})_{i\geq 0, j<n}$ et l'on conclut par le caractère i.i.d. de la suite des $X_{i,j}$. □

1 Espérance de Z_n

On va montrer la proposition suivante, qui va nous donner une première idée de l'évolution de la taille de la population, *i.e.* de la suite Z_n.

Proposition. *On a $\mathbb{E}[Z_n] = m^n$.*

On va raisonner par récurrence ; on remarque tout d'abord que $Z_0 = 1$ donc $\mathbb{E}[Z_0] = 1 = m^0$. Pour $n \in \mathbb{N}$, supposons que $\mathbb{E}[Z_n] = m^n$. On a :

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\right]$$

*Sir Francis Galton (1822–1911) ; a notamment étudié la statistique des patronymes de l'Angleterre victorienne.

†Henry William Watson (1827–1903).

Pour permuter espérance et somme, on va conditionner par Z_n, ce qui aura pour effet de « fixer » la variable Z_n :

$$\begin{aligned}\mathbb{E}[Z_{n+1}] &= \mathbb{E}[\mathbb{E}[Z_{n+1}|Z_n]] \\ &= \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\middle| Z_n\right]\right]\end{aligned}$$

donc comme Z_n est fixée (voir annexe A.1 pour plus de détails) on peut permuter, pour obtenir :

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}\left[\sum_{i=1}^{Z_n} \mathbb{E}[X_{i,n}|Z_n]\right]$$

Or, par indépendance (cf. propriété énoncée au début) on a $\mathbb{E}[X_{i,n}|Z_n] = \mathbb{E}[X_{i,n}]$ et comme cette dernière espérance vaut m on obtient finalement :

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}[mZ_n] = m\mathbb{E}[Z_n]$$

d'où le résultat par récurrence.

2 Étude de la probabilité d'extinction

2.1 Réécriture

Notons P_{ext} la probabilité d'extinction de la population, *i.e.* $P_{\text{ext}} := \mathbb{P}(\exists n \in \mathbb{N}, Z_n = 0)$. On remarque que si $Z_n = 0$ alors $Z_{n+1} = 0$, autrement dit la suite d'évènements $(\{Z_n = 0\})_{n\in\mathbb{N}}$ est croissante ; ainsi :

$$P_{\text{ext}} = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{Z_n = 0\}\right) = \lim_{n\to+\infty}\uparrow \mathbb{P}(Z_n = 0)$$

On a donc montré que la suite $(\mathbb{P}(Z_n = 0))_n$ converge vers P_{ext} ; en obtenant des renseignements sur cette suite, on obtiendra donc des renseignements sur la probabilité d'extinction P_{ext}, le but étant de savoir si elle est égale à 1 ou non.

2.2 Relation de récurrence

Notons G la fonction génératrice de la variable X : pour $s \in [0, 1]$, on a $G(s) = \sum_{n=0}^{+\infty} p_n s^n$; G est dérivable sur $[0, 1[$ (c'est une série entière) et continue sur $[0, 1]$ (par exemple par convergence uniforme). On va montrer le fait suivant, où G_n désigne la fonction génératrice de Z_n.

Théorème. *Pour $n \in \mathbb{N}$ on a $G_{n+1} = G_n \circ G$ (sur $[0, 1]$).*

Démonstration. On fait comme pour le calcul de l'espérance ; soit $s \in [0, 1]$.

$$\begin{aligned}
G_{n+1}(s) &= \mathbb{E}\left[s^{Z_{n+1}}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[s^{Z_{n+1}}\middle| Z_n\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\prod_{i=1}^{Z_n} s^{X_{i,n}}\middle| Z_n\right]\right] \\
&= \mathbb{E}\left[\prod_{i=1}^{Z_n} \mathbb{E}\left[s^{X_{i,n}}\middle| Z_n\right]\right] \text{ (voir annexe A.2)} \\
&\overset{\perp\!\!\!\perp}{=} \mathbb{E}\left[\prod_{i=1}^{Z_n} \mathbb{E}\left[s^{X_{i,n}}\right]\right] \\
&= \mathbb{E}\left[G(s)^{Z_n}\right] \\
G_{n+1}(s) &= G_n \circ G
\end{aligned}$$

□

Ainsi, comme $G_0 = 1$ on en déduit que $G_n = \underbrace{G \circ \cdots \circ G}_{n \text{ fois}}$ et donc que :

$$G_{n+1} = G \circ G_n$$

En évaluant en 0 on obtient, comme $G_n(0) = \mathbb{P}(Z_n = 0)$:

$$\mathbb{P}(Z_{n+1} = 0) = G(\mathbb{P}(Z_n = 0))$$

et on a notre relation de récurrence pour la suite $(\mathbb{P}(Z_n = 0))_n$!

2.3 Convergence

Comme la limite de la suite $(\mathbb{P}(Z_n = 0))$ est dans $[0, 1]$ et comme G est continue sur $[0, 1]$ on en déduit le théorème suivant.

Théorème. *La probabilité d'extinction P_{ext} est un point fixe de G sur* $[0, 1]$.

Pour pouvoir déterminer ce point fixe, on va utiliser la précision suivante.

Corollaire. *La probabilité d'extinction P_{ext} est le plus point fixe de G sur l'intervalle* $[0, 1]$.

Démonstration. Soit ℓ un point fixe de G sur $[0, 1]$. Comme $Z_0 = 1$, on a $\mathbb{P}(Z_0 = 1) = 0$ et donc :

$$\mathbb{P}(Z_0 = 0) \leq \ell \tag{1}$$

On a remarqué que G est dérivable sur $[0, 1[$; comme les coefficients de la série entière G sont positifs, on en déduit que $G' \geq 0$ sur $[0, 1[$ donc que G est *croissante* sur $[0, 1]$. Ainsi, on peut composer par G dans l'inégalité (1) pour trouver $G(\mathbb{P}(Z_0 = 0)) \leq G(l)$, et donc d'après ce qui a été fait précédemment et comme l est un point fixe on obtient :

$$\mathbb{P}(Z_1 = 0) \leq l$$

En réitérant, on obtient par récurrence :

$$\forall n \in \mathbb{N}, \mathbb{P}(Z_n = 0) \leq l$$

et donc en passant à la limite :

$$P_{\text{ext}} \leq l$$

Ceci pour tout point fixe de G sur $[0, 1]$ donc comme P_{ext} est un point fixe de G, c'est bien le plus petit. □

2.4 La population va-t-elle presque sûrement s'éteindre ?

D'après la section précédente, il suffit donc de savoir si G possède un point fixe dans $[0, 1[$. Pour cela, on remarque que $G'' \geq 0$ sur $[0, 1[$ (toujours car les p_n sont positifs) donc G est *convexe* sur $[0, 1]$.

2.4.1 Cas $m > 1$

Rappelons que $G(1) = 1$; de plus, $G'(s) = \sum_{n=1}^{+\infty} np_n s^{n-1}$ donc $G'(1) = \sum_{n=1}^{+\infty} np_n = m$. Ainsi, par convexité la courbe représentative de G est au-dessous de la droite passant par $(1, G(1) = 1)$ et de coefficient directeur $G'(1) = m > 1$. Autrement dit, on est dans la situation de la figure 1.

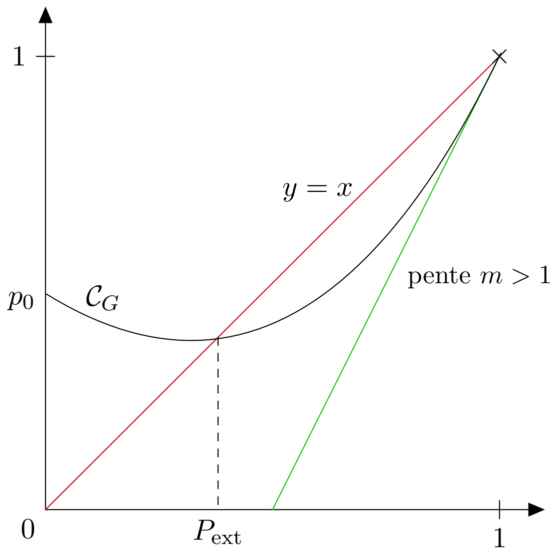

Figure 1 – Cas $m > 1$

La courbe se confondant localement avec la tangente, on en déduit que $G - x < 0$ au voisinage de 1^- donc comme $G - x$ vaut $p_0 \geq 0$ en 0, par le théorème des valeurs intermédiaires on en déduit que $G - x$ s'annule sur $[0, 1[$; ainsi, par ce qui précède on a $P_{\text{ext}} < 1$.

2.4.2 Cas $m < 1$

On est cette fois dans la situation de la figure 2 ; la courbe représentative de G étant au-dessus de la droite verte, on en déduit que le seul point fixe de G sur $[0, 1]$ est en $x = 1$ donc finalement $P_{\text{ext}} = 1$.

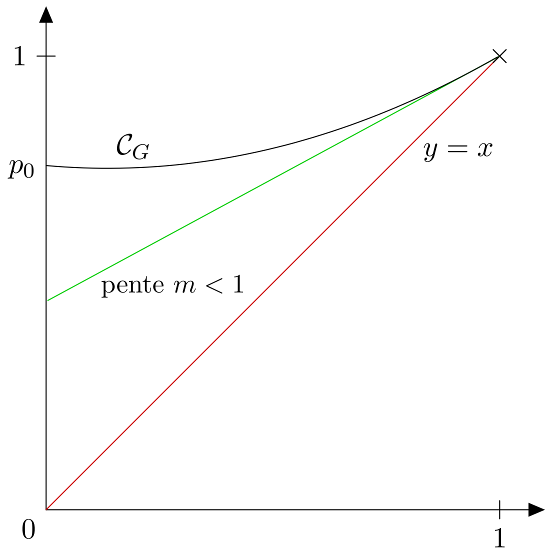

Figure 2 – Cas $m < 1$

2.4.3 Cas $m = 1$

Cette fois, les droites verte et rouge coïncident ; c'est la situation de la figure 3 (où l'on n'a par conséquent tracé que la droite rouge).

Pour pouvoir conclure, on va utiliser le fait que si $p_0 + p_1 < 1$, alors $G'' > 0$ sur $[0, 1[$ donc que G est strictement convexe.

Remarque. Si $p_0 + p_1 = 1$, alors comme $m = 1$ on a $p_1 = 1$ donc $G = \text{id}$ et p.s. $X = 1$. Autrement dit, $Z_n = 1$ p.s. donc on a bien $P_{\text{ext}} = 0$:-) On supposera donc dans la suite $p_0 + p_1 < 1$.

La fonction G étant strictement convexe, la courbe représentative de G est strictement au dessus de la droite représentative de $y = x$ privée du point $(1, 1)$. Ainsi, G possède un unique point fixe sur $[0, 1]$, atteint pour $x = 1$ et donc finalement $P_{\text{ext}} = 1$.

Références

[1] TOULOUSE Paul S., *Thèmes de probabilités et statistique.* Dunod, 1999.

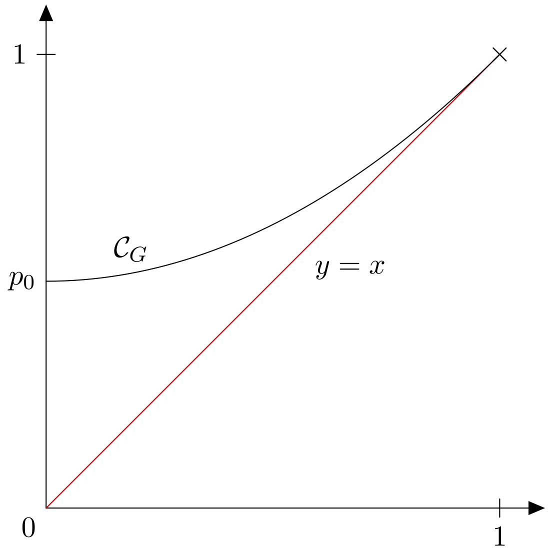

FIGURE 3 – Cas $m = 1$ avec $p_0 + p_1 < 1$

A Espérance conditionnelle et permutations

A.1 Somme

On veut justifier le fait suivant :

$$\mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\middle| Z_n\right] = \sum_{i=1}^{Z_n} \mathbb{E}\left[X_{i,n}\middle| Z_n\right]$$

Pour cela, on écrit :

$$\mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\middle| Z_n\right] = \mathbb{E}\left[\sum_{i=1}^{+\infty} \mathbf{1}_{i\leq Z_n} X_{i,n}\middle| Z_n\right]$$

On peut permuter par le théorème de Fubini (l'intégrande est positive) puis utiliser le fait que $\mathbf{1}_{i\leq Z_n}$ est Z_n-mesurable pour le sortir de l'espérance conditionnelle :

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\middle| Z_n\right] &= \sum_{i=1}^{+\infty} \mathbb{E}\left[\mathbf{1}_{i\leq Z_n} X_{i,n}\middle| Z_n\right] \\
&= \sum_{i=1}^{+\infty} \mathbf{1}_{i\leq Z_n} \mathbb{E}[X_{i,n}|Z_n] \\
\mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\middle| Z_n\right] &= \sum_{i=1}^{Z_n} \mathbb{E}[X_{i,n}|Z_n]
\end{aligned}$$

et c'est ce que l'on voulait !

Remarque. Si l'on détaille cette manipulation, on n'a en fait pas besoin de l'espérance conditionnelle, comme le montre le calcul suivant.

$$\begin{aligned} \mathbb{E}[Z_{n+1}] = \mathbb{E}\left[\sum_{i=1}^{Z_n} X_{i,n}\right] = \mathbb{E}\left[\sum_{i=1}^{+\infty} \mathbf{1}_{i\leq Z_n} X_{i,n}\right] \\ = \sum_{i=1}^{+\infty} \mathbb{E}\left[\mathbf{1}_{i\leq Z_n} X_{i,n}\right] \overset{\perp\!\!\!\perp}{=} \sum_{i=1}^{+\infty} \mathbb{E}[\mathbf{1}_{i\leq Z_n}]\mathbb{E}[X_{i,n}] = m\sum_{i=1}^{+\infty} \mathbb{E}[\mathbf{1}_{i\leq Z_n}] \\ = m\mathbb{E}\left[\sum_{i=1}^{+\infty} \mathbf{1}_{i\leq Z_n}\right] = m\mathbb{E}\left[\sum_{i=1}^{Z_n} 1\right] = m\mathbb{E}[Z_n] \end{aligned}$$

A.2 Produit

La manipulation précédente avec l'indicatrice ne va plus fonctionner avec le produit [1]. Pour cela (et pour la justification d'avant également), on utilise le lemme suivant.

Lemme. *Soient (X, Y) une variable aléatoire à valeurs dans l'espace probabilisable $(E \times F, \mathcal{E} \otimes \mathcal{F})$ et $f \in \mathcal{L}^1(E \times F, \mathcal{E} \otimes \mathcal{F}, \mathbb{P}_{(X,Y)})$. On suppose que X est $\mathcal{B}$-mesurable et que Y et $\mathcal{B}$ sont* indépendantes. *La fonction $\hat{f}$ définie par*

$$\forall x \in E, \hat{f}(x) := \mathbb{E}[f(x, Y)]$$

est $\mathcal{E}$-mesurable et on a :

$$\mathbb{E}[f(X, Y)|\mathcal{B}] = \hat{f} \circ X \ \ \mathbb{P}-\text{p.s.}$$

Démonstration. OUVRARD Jean-Yves, *Probabilités 2* (troisième édition). Cassini, 2009. (Proposition 11.22, p. 156–157.) □

Remarque. Pour se souvenir des hypothèses, on peut se dire que $\hat{f}(X) := \hat{f} \circ X$ est X-mesurable, ce qui est bon signe car $\mathbb{E}[\dots|\mathcal{B}]$ est une variable $\mathcal{B}$-mesurable.

Dans notre cas, on fixe s et on utilise le lemme avec :
- $f(z, x) = \prod_{i=1}^{z} s^{x_i}$;
- $X = Z_n$;
- $Y = (X_{i,n})_i$;
- $\mathcal{B} = \sigma(X)$.

Ainsi, $\hat{f}(z) = \mathbb{E}\left[\prod_{i=1}^{z} s^{X_{i,n}}\right] \overset{\perp\!\!\!\perp}{=} \prod_{i=1}^{z} \mathbb{E}[s^{X_{i,n}}] = G(s)^z$. Par nos hypothèses d'indépendances, on peut donc appliquer le théorème pour en déduire que $\mathbb{E}[s^{Z_{n+1}}|Z_n] = G(s)^{Z_n}$ et donc $G_{n+1} = G_n \circ G$.

1. On ferait apparaître des $s^{\mathbf{1}_{i\leq Z_n}}$; je ne sais pas vous, mais moi je ne permute pas des intégrales et des produits infinis tous les jours !