Université Nice - Agrégation de Mathématiques

TRAVAUX PRATIQUES DE PROBABILITÉS NUMÉRO 7

PROCESSUS DE GALTON-WATSON

Références : Les références suivantes peuvent être consultées (les mots-clés à chercher dans les index étant **Galton-Watson** et **processus de branchements**). :

(1) Nicolas Bouleau, Probabilités de l'Ingénieur.
(2) Nicoals Bouleau, Processus Stochastiques, pages 54-59, pages 79-80.
(3) Didier Dacunha-Castelle et Marie Duflo, Tomes 1 et 2.
(4) Theodor Harris, Theory of Brancing processes, traduction française *a priori* disponible.
(5) Daniel Revuz, Probabilités, Lecture du Chapitre II
(6) Sheldon Ross, Introduction to Probability models.

1. Position du problème

Cette note vise à présenter quelques illustrations informatiques des processus de Galton-Watson. Rappelons brièvement que de tels processus permettent de d'écrire l'évolution des générations mâles descendant d'un même ancêtre, les individus mâles se reproduisant selon la même loi au cours du temps et de façon indépendante les uns des autres (l'hypothèse de reproduction porte sur la descendance mâle et non femelle pour des raisons historiques, le modèle ayant été introduit au XIXe siècle par deux Anglais, Galton et Watson, afin d'étudier l'extinction des noms de familles).

Néanmoins, d'autres applications sont envisagées : en physique nucléaire (multiplicateurs d'électrons) ou en génétique (développement d'un gène mutant). Dans cette perspective, l'ouvrage de T.E. Harris est riche de nombreux exemples, et dans une mesure plus modeste, la lecture II du Chapitre II du livre de D. Revuz est également très instructrice.

Rappelons brièvement que pour chaque génération n, le nombre d'individus mâles de la génération est modélisée par une v.a. notée X_n. La suite $(X_n)_{n\geq 0}$ se construit sur le modèle itératif suivant (typique des modèles de Markov, et par ailleurs régi par les hypothèses de reproduction faites précédemment) :

(1) On désigne par $(\xi_{k,n})_{k\geq 1,n\geq 0}$ une famille de v.a. indépendantes à valeurs dans $\mathbb{N}$ de même loi.
(2) Pour $n = 0$, $X_0 = 1$.
(3) Pour tout $n \geq 0$,

$$X_{n+1} = \sum_{k=1}^{X_n} \xi_{k,n} \text{ si } X_n \geq 1,$$
$$X_{n+1} = 0 \text{ si } X_n = 0.$$

La première égalité traduit le fait que les individus se reproduisent selon le même hasard et de façon indépendante. La seconde égalité signifie que l'état 0 est absorbant : si une génération est vide, elle a peu de chance d'avoir une descendance !

Remarquons que la loi commune des $(\xi_{k,n})_{k\geq 1,n\geq 0}$ est également la loi de X_1.

(4) Il semble assez pertinent de supposer $p_0 = \mathbb{P}\big\{X_1 = 0\big\} < 1$. Dans le cas contraire, l'ancêtre n'a aucune chance d'avoir une descendance mâle ! Ceci est alors assez peu intéressant.

(5) Si le nombre d'individus engendrés par un même individu mâle n'est pas trop élevé, il est raisonnable de supposer que la loi de X_1 admet une moyenne et une variance finies :
$$m = \mathbb{E}(X_1) = \sum_{n \geq 0} n p_n, \ \sigma^2 = \mathbb{V}(X_1) = \sum_{n\geq 0} (n-m)^2 p_n.$$

Objectif : Notre étude vise à étudier le comportement asymptotique de la descendance, i.e. le comportement en temps long de la suite $(X_n)_{n\geq 0}$. En particulier, est-il possible que la descendance s'éteigne en temps fini, et si oui avec quelle probabilité ? Inversement, que dire si la descendance perdure au cours du temps avec une probabilité non nulle ?

N.B. : L'hypothèse d'intégrabilité d'ordre 2 ne sera en réalité utile que pour la description fine du comportement asymptotique de la suite $(X_n)_{n\geq 0}$.

2. Présentation des résultats à illustrer

Dans la suite de ces travaux pratiques, nous appellerons probabilité d'extinction la quantité donnée par $\mathbb{P}\{\lim_{n\to+\infty} X_n = 0\}$. Nous verrons qu'en réalité :
$$\mathbb{P}\Big\{\lim_{n\to+\infty} X_n = 0\Big\} = \mathbb{P}\big\{\exists\ n \geq 1,\ X_n = 0\big\} = \lim_{n\to+\infty} \mathbb{P}\big\{X_n = 0\big\}.$$

Dans ce contexte, la séance a pour objectif d'illustrer les théorèmes suivants, dont les preuves sont données en annexe :

Théorème 1 : Le comportement asymptotique de la descendance est régi par les cas suivants :

(1) S'il existe $c \in \mathbb{N}$ tel que $\mathbb{P}\big\{X_1 = c\big\} = 1$. Alors, deux cas sont possibles :
   (a) Si $c = 0$, alors la descendance s'éteint avec probabilité 1.
   (b) Si $c \geq 1$, alors la descendance ne s'éteint presque-sûrement jamais.

(2) Si X_1 n'est pas une v.a. dégénérée. Alors,
   (a) Si $m \leq 1$, alors la descendance s'éteint avec probabilité 1.
   (b) Si $m > 1$, alors la descendance s'éteint avec probabilité < 1.
   (c) Dans les deux cas, la probabilité d'extinction est le plus petit point fixe de la fonction génératrice de X_1, notée φ, et définie par :
$$\forall\ s \in [0,1],\ \varphi(s) = \mathbb{E}\big(s^{X_1}\big).$$
   Cette fonction admet en fait au moins un point fixe (égal à 1) et au plus deux.

La démonstration est proposée en Annexe C.

Au cours du Chapitre sur la Théorie des Martingales, nous montrerons :

Théorème 2 : Le comportement asymptotique de la descendance peut être précisé de la façon suivante :

(1) Si la loi de X_1 n'est pas dégénérée et $m \leq 1$, alors :
$$\mathbb{P}-\text{p.s.},\ X_n \to 0,$$
i.e., avec probabilité 1, la suite $(X_n)_{n\geq 1}$ ne prend qu'un nombre fini de valeurs non nulles.

(2) Si $m > 1$ et X_1 est de carré intégrable, alors il existe une v.a. W positive de carré intégrable et de moyenne 1 telle que
$$\mathbb{P}-\text{p.s.},\ m^{-n} X_n \to W.$$

Une démonstration de ce résultat n'utilisant pas la théorie des martingales est toutefois proposée à titre culturel en Annexe D.

Enfin, nous admettrons le résultat suivant (néanmoins accessible à notre niveau) :

Théorème : La probabilité d'extinction coïncide avec $\mathbb{P}\{W=0\}$. En particulier,

(1) Sur l'ensemble $\{W=0\}$, $X_n \to 0$,

(2) Sur l'ensemble $\{W>0\}$, $X_n \to +\infty$.

3. Données Numériques

L'illustration informatique se concentre sur les lois géométriques. De telles lois se prêtent bien aux calculs, et ont été utilisées (voir l'ouvrage de Harris, Chapitre 1, Paragraphe 7), sous une forme généralisée, dans des études démographiques.

3.1. **Choix de X_1.** On choisit X_1 de loi géométrique sur $\mathbb{N}$ de paramètre c :

$$\forall\ k \geq 0,\ \mathbb{P}\{X_1 = k\} = c(1-c)^k.$$

N.B. : Afin d'avoir un choix uniforme de c sur l'ensemble des machines, nous proposons de prendre, dans le cadre des applications informatiques, les valeurs $c = 0.6$ et $c = 0.4$.

3.2. **Caractéristiques du modèle.** En référence à l'étude théorique menée dans les paragraphes précédents, il s'agit de calculer dans un premier temps les quantités nécessaires à la description de l'évolution du modèle. Dans cette perspective, l'exercice suivant a pour but de déterminer la moyenne et la fonction génératrice de X_1, et d'exhiber un éventuel point fixe non trivial de cette fonction (dans le cadre de la séance de T.P., nous admettrons les résultats énoncés dans l'exercice).

Exercice 1 : Démontrer les résultats suivants :

(1) Démontrer que la moyenne de X_1 est donnée par la relation $\mathbb{E}(X_1) = \dfrac{1-c}{c}$.
En déduire les cas d'évolution suivants :
   (a) Si $c \geq 1/2$, alors, avec probabilité 1, toute descendance s'éteint en temps fini.
   (b) Si $c < 1/2$, alors, il existe un événement de mesure non nulle sur lequel les descendances ne s'éteignent jamais.

(2) A propos de la fonction génératrice,
   (a) Démontrer que la fonction génératrice de X_1, notée φ, est donnée par la relation :
$$\forall\ s \in [0,1],\ \varphi(s) = \frac{c}{1-(1-c)s}.$$
   (b) En déduire, dans le cas $c < 1/2$, que la probabilité d'extinction de la descendance est donnée par la relation :
$$a = \frac{c}{1-c}.$$
   On rappelle, vus les résultats théoriques exhibés dans les paragraphes précédents, que cette probabilité est l'unique point fixe de φ appartenant à l'intervalle $[0,1[$.
   (c) Vérifier, dans le cas $c \geq 1/2$, que la fonction φ n'a pas de point fixe dans $[0,1[$. Noter que cela est cohérent avec les résultats annoncés par le théorème 1.

(3) *Application Numérique.* Dans le cas $c = 0.4$, vérifier que la probabilité d'extinction est donnée par $a = 2/3$.

4. Simulation de la Chaîne

On cherche désormais à simuler une réalisation de la chaîne de Galton Watson, pour une valeur c dans $]0,1[$ quelconque. Le principe, assez simple, est régi par l'écriture de la chaîne $(X_n)_{n\geq 1}$ exhibée dans les paragraphes précédents. Dans ce contexte, nous proposons la démarche suivante :

(1) La génération 0 ne compte qu'un individu (l'ancêtre).

(2) La génération 1 compte x_1 individus, x_1 désignant une réalisation d'une loi géométrique de paramètre c.

(3) Plus généralement, si la génération n compte x_n individus, alors, la génération $n+1$ compte x_{n+1} indvidus, x_{n+1} désignant la somme des termes des réalisations d'un x_n-échantillon de loi géométrique de paramètre c. Bien entendu, si x_n est nul, alors x_{n+1} est également nul.

Exercice 2 : A partir du principe précédent,

(1) Donner une méthode pour simuler une v.a. de loi géométrique de paramètre c sur $\mathbb{N}$. (On pourra remarquer que la partie entière supérieure d'une v.a. de loi exponentielle de paramtre $-\ln(1-c)$ suit une loi géométrique de paramètre c sur $\mathbb{N}^*$.)

(2) Proposer un algorithme pour simuler les effectifs des 20 premières générations d'une chaîne de Galton Watson régie par une loi géométrique de paramètre c.

(3) Programmer cet algorithme en Scilab.

(4) Proposer, à partir du programme précédent, une représentation graphique de l'évolution de la descendance d'un ancêtre au cours des 20 premières générations. On pourra choisir les valeurs $c = 0.6$ et $c = 0.4$.

(5) Tracer sur un même dessin plusieurs réalisations de la chaîne de Galton-Watson, et commenter les différents types de comportement asymptotique obtenus. Traiter le cas $c = 0.6$ puis le cas $c = 0.4$

Les illustrations graphiques sont proposées en Annexe A. La figure 1 représente, dans le cas $c = 0.6$, et sur une durée de 20 générations, 100 réalisations de la chaîne de Galton-Watson. Cette figure illustre la prévision théorique de la chaîne de Galton-Watson : avec probabilité 1, les descendances s'éteignent en temps fini.

La figure 2 s'attache au cas $c = 0.4$. Nous avons, là encore, représenté, sur une durée de 20 générations, 100 réalisations du processus. Ce graphique confirme l'évolution théorique de la chaîne de Galton-Watson : la probabilité d'extinction est manifestement strictement plus petite que 1. Il est néanmoins impossible de déterminer graphiquement la valeur de la probabilité d'extinction, ou du moins une estimation de cette valeur.

5. Renormalisation de la Chaîne

La figure 2 suggère que les descendances ne s'éteignant pas ont un comportement asymptotique de type exponentiel. Rappelons à ce sujet que les résultats théoriques établis dans les sections précédentes montrent que la suite $\left(m^{-n}X_n\right)_{n\geq 1}$ converge p.s. vers une v.a. W, nulle sur l'ensemble d'extinction de la chaîne, et strictement positive sur son complémentaire.

Afin de vérifier graphiquement cette évolution, nous vous proposons l'exercice suivant :

Exercice 3 : Représenter dans le cas $c = 0.4$ une centaine de réalisations du processus $\left(m^{-n}X_n\right)_{n\geq 1}$ sur une durée de 20 générations. On rappelle que dans le cas $c = 0.4$, la valeur de m est $3/2$.

La représentation graphique est disponible en Annexe A.

6. Estimation de la probabilité d'extinction

Nous savons que la probabilité d'extinction est donnée par :

$$a = \lim_{n \uparrow +\infty} \mathbb{P}\big\{X_n = 0\big\}.$$

Dans ce contexte, nous vous proposons l'exercice suivant :

Exercice 4 : Dans le cas $c = 0.4$,

(1) Supposons que la suite $\big(X_1^{(k)}, ..., X_{20}^{(k)}\big)_{1 \leq k \leq 500}$ désigne 500 vecteurs aléatoires indépendants, chacun suivant la loi du vecteur $(X_1, ..., X_{20})$ désignant les 20 premières générations de la chaîne de Galton-Watson. Proposer à partir de ces 500-échantillons un estimateur empirique du vecteur $\big(\mathbb{P}\{X_1 = 0\}, ..., \mathbb{P}\{X_{20} = 0\}\big)$.

(2) Simuler avec Scilab une réalisation du 500-échantillon précédent. En déduire une réalisation de l'estimateur empirique du vecteur $\big(\mathbb{P}\{X_1 = 0\}, ..., \mathbb{P}\{X_{20} = 0\}\big)$.

La convergence attendue est-elle visible ?

Annexe A - Figures

Figure 1. Cas $c = 0.6$

Figure 2. Cas $c = 0.4$

Figure 3. Cas $c = 0.4$. Renormalisation

Annexe B - Programmes Scilab

```
///////////////////////////////////////////////////////////
// Simulation d'une realisation
// de la chaine de Galton Watson (1ere methode)
///////////////////////////////////////////////////////////

function galton=galton(c,n)
// c parametre de la loi geometrique, n nombre de generations,
x(1)=grand(1,1,"nbn",1,c);
// simulation de la premiere generation
for k=2:n,
x(k)=sum(grand(1,x(k-1),"nbn",1,c));
// simulation de la k ieme generation
end;
galton=x;
// renvoyer en fin de boucle le vecteur des generations
endfunction

///////////////////////////////////////////////////////////
// Simulation d'une realisation
// de la chaine de Galton Watson (2eme methode)
///////////////////////////////////////////////////////////

function galton=galton(c,n)
// c parametre de la loi geometrique, n nombre de generations,
if n==1 then galton=grand(1,1,"nbn",1,c),
// simulation de la premiere generation
else,
galton=sum(grand(1,galton(c,n-1),"nbn",1,c));
end;
endfunction

///////////////////////////////////////////////////////////
// Simulation de m realisations
// de la chaine de Galton Watson (3eme methode)
///////////////////////////////////////////////////////////

function galton=galton(c,n,m);
// c parametre de la loi geometrique, n nombre de generations,
// m nombre de realisations
x=grand(1,m,"nbn",1,c);
// simulation de m realisations de la premiere generation
for k=2:n,
loc2=[0,cumsum(grand(1,sum(x(k-1,:)),"nbn",1,c))];
// simulation du nombre total de descendants
loc3=1+cumsum(x(k-1,:));
```

```
x(k,:)=loc2(loc3)-[0,loc2(loc3([1:m-1]))];
// repartition de la simulation entre les m realisations
end;
galton=x;
endfunction
```

Annexe C - Preuve du Théorème 1

Probabilité d'extinction - définition. Désignons pour tout $n \geq 1$ la quantité $\mathbb{P}\{X_n = 0\}$ par :

$$a_n = \mathbb{P}\big\{X_n = 0\big\}.$$

Dans ce cadre, le lemme suivant nous indique que la suite $(a_n)_{n\geq 1}$ joue un rôle crucial pour étudier l'éventuelle extinction de la descendance :

Lemme C.1 : La suite $(a_n)_{n\geq 1}$ est une suite croissante majorée par 1, et la probabilité que la descendance s'éteigne en temps fini est donnée par :

$$\mathbb{P}\big\{\lim_{n\to+\infty} X_n = 0\big\} = \mathbb{P}\bigg(\bigcup_{n\geq 1}\big\{X_n = 0\big\}\bigg) = \lim_{n\to+\infty} a_n.$$

Preuve : La suite $(X_n)_{n\geq 1}$ est à valeurs entières. En particulier,

$$\big\{\lim_{n\to+\infty} X_n = 0\big\} \subset \bigcup_{n\geq 1}\{X_n = 0\}.$$

Par ailleurs, on sait par hypothèses que $\forall\ n \geq 0, X_n = 0 \Rightarrow X_{n+1} = 0$. On en déduit que :

$$\forall\ n \geq 1,\ \big\{X_n = 0\big\} \subset \big\{\sup_{k\geq n} X_k = 0\big\}.$$

En particulier,

$$\bigcup_{n\geq 1}\{X_n = 0\} \subset \big\{\lim_{n\to+\infty} X_n = 0\big\}.$$

Il en résulte que :

$$\mathbb{P}\big\{\lim_{n\to+\infty} X_n = 0\big\} = \mathbb{P}\bigg(\bigcup_{n\geq 1}\big\{X_n = 0\big\}\bigg).$$

Finalement, de l'inclusion :

$$\forall\ n \geq 1,\ \big\{X_n = 0\big\} \subset \big\{X_{n+1} = 0\big\},$$

on déduit que la suite $(a_n)_{n\geq 0}$ est croissante (et évidemment majorée par 1). De plus, de la théorie de la mesure, il vient :

$$\mathbb{P}\bigg(\bigcup_{n\geq 1}\big\{X_n = 0\big\}\bigg) = \lim_{n\to+\infty} a_n.$$

Fonction génératrice. Notre développement s'appuie sur l'introduction de la fonction génératrice de la v.a. X_1 :

$$\forall\ s \in [0,1],\ \varphi(s) = \mathbb{E}\big(s^{X_1}\big) = \sum_{n\geq 0} p_n s^n.$$

De façon plus générale, pour $n \geq 1$, nous introduisons la fonction génératrice de X_n :

$$\forall\ s \in [0,1],\ \varphi_n(s) = \mathbb{E}\big(s^{X_n}\big).$$

Dans ces conditions, il est immédiat que $\varphi = \varphi_1$. Rappelons que sous ces notations :

$$\forall\ n \geq 1,\ \mathbb{P}\big\{X_n = 0\big\} = \varphi_n(0).$$

Le lemme suivant donne une relation entre φ et φ_n :

Lemme C.2 : Pour tout $n \geq 1$, pour tout $s \in [0,1]$, $\varphi_n(s) = \varphi^{\circ n}(s)$.

Preuve : Fixons $n \geq 1$. On sait que $\forall\ s \in [0,1],\ \varphi_{n+1}(s) = \mathbb{E}\big(s^{X_{n+1}}\big)$.

En utilisant la relation entre X_{n+1} et X_n, il vient :

$$\forall\ s \in [0,1],\ \varphi_{n+1}(s) = \mathbb{E}\Big(\mathbf{1}_{\{X_n>0\}} \prod_{k=1}^{X_n} s^{\xi_{k,n+1}} + \mathbf{1}_{\{X_n=0\}}\Big).$$

Il est clair que la v.a. X_n est indépendante du vecteur $\big(\xi_{k,n+1}\big)_{1\le k\le n}$. Il vient alors :

$$\begin{aligned}
\forall\ s \in [0,1],\ \varphi_{n+1}(s) &= \mathbb{E}\Big(\mathbf{1}_{\{X_n>0\}} \prod_{k=1}^{X_n} s^{\xi_{k,n+1}} + \mathbf{1}_{\{X_n=0\}}\Big) \\
&= \sum_{j\ge 1} \mathbb{E}\Big(\mathbf{1}_{\{X_n=j\}} \prod_{k=1}^{j} s^{\xi_{k,n+1}}\Big) + \mathbb{P}\big\{X_n = 0\big\} \\
&= \sum_{j\ge 1} \mathbb{P}\big\{X_n = j\big\}\mathbb{E}\Big(\prod_{k=1}^{j} s^{\xi_{k,n+1}}\Big) + \mathbb{P}\big\{X_n = 0\big\} \\
&= \sum_{j\ge 1} \mathbb{P}\big\{X_n = j\big\} \prod_{k=1}^{j} \mathbb{E}\big(s^{X_1}\big) + \mathbb{P}\big\{X_n = 0\big\} \\
&= \sum_{j\ge 0} \mathbb{P}\big\{X_n = j\big\}\big(\varphi(s)\big)^j = \varphi_n\big(\varphi(s)\big).
\end{aligned}$$

Une récurrence montre sans difficulté que $\forall\ n \ge 1,\ \varphi_n = \varphi^{\circ n}$.

Probabilité d'extinction et fonction génératrice. Le corollaire suivant donne le lien entre la suite $(a_n)_{n\ge1}$ et la fonction génératrice :

Corollaire C.3 : La suite $(a_n)_{n\ge1}$ satisfait la relation de récurrence :

$$\forall\ n \ge 1,\ a_{n+1} = \varphi\big(a_n\big).$$

Preuve : On applique le lemme C.2 avec $s = 0$. Il vient $\forall\ n \ge 1,\ \varphi_{n+1}(0) = \varphi\big(\varphi_n(0)\big)$.

En utilisant les notations introduites avant le lemme C.1, on déduit $\forall\ n \ge 1,\ a_{n+1} = \varphi\big(a_n\big)$.

On déduit immédiatement le corollaire suivant, crucial dans notre développement :

Corollaire C.4 : En désignant par a la limite de la suite croissante $(a_n)_{n\ge1}$:

$$a = \lim_{n\to+\infty} a_n,$$

il vient par continuité de φ :

$$a = \varphi(a).$$

Etude de l'équation. Il convient donc d'étudier les solutions de l'équation $\varphi(x) = x$. Pour cela, on affirme que :

Proposition C.5 : Si la v.a. X_1 n'est pas réduite à une constante, alors l'équation :

$$\varphi(x) = x,$$

admet au moins 1 pour solution, mais admet en fait au plus deux solutions. Dans ce cadre, a désigne la plus petite de ces solutions (a est bien-sûr la solution si celle-ci est unique).

Preuve : Notons que dans tous les cas, 1 est effectivement solution de l'équation.

Traitons maintenant le cas : $p_0 + p_1 = 1$ (notons que $0 < p_0 < 1$, car X_1 est supposée non constante). Dans ce cas,

$$\forall\ s \in [0,1],\ \varphi(s) = p_0 + (1 - p_0)s.$$

Alors, l'équation $\varphi(x) = x$ admet une unique solution, qui est bien entendu 1. En particulier, $a = 1$.

Supposons maintenant : $p_0 + p_1 < 1$. On sait que :

$$\forall\ s \in [0,1],\ \varphi(s) = \sum_{n\geq 0} p_n s^n.$$

On sait que φ est indéfiniment dérivable sur $[0,1[$ (et même deux fois dérivable sur $[0,1]$ si l'on suppose X_1 de carré intégrable). En particulier, pour tout $s \in [0,1[$:

$$\varphi'(s) = \sum_{n\geq 1} n p_n s^{n-1},$$
$$\varphi''(s) = \sum_{n\geq 2} n(n-1) p_n s^{n-2}.$$

De l'hypothèse, $p_0 + p_1 < 1$, on en déduit que :

$$\forall\ s \in]0,1[,\ \varphi'(s) > 0,\ \varphi''(s) > 0.$$

On en déduit que la fonction φ est strictement croissante et strictement convexe sur $[0,1]$.

En particulier, la fonction $x \in [0,1[\mapsto \frac{\varphi(1)-\varphi(x)}{1-x} = \frac{1-\varphi(x)}{1-x}$ est strictement croissante, et donc injective. Il y a donc au plus un unique élément de $[0,1[$ solution de l'équation considérée.

Finalement, il y a au plus deux solutions :

(1) Si 1 est l'unique solution de l'équation, alors $a = 1$.

(2) Montrons que s'il existe $b < 1$, tel que $\varphi(b) = b$, alors $a = b$.
Preuve : On sait que $a_1 = \mathbb{P}\big\{X_1 = 0\big\} = \varphi(0)$. En particulier, en utilisant la relation $0 \leq b$, et la croissance de φ, on déduit,

$$a_1 = \varphi(0) \leq \varphi(b) = b.$$

En s'appuyant sur une récurrence, il vient $\forall\ n \geq 1,\ a_n \leq b$. Finalement, $a \leq b$. En utilisant la relation $\varphi(a) = a$, on déduit qu'effectivement $a = b$. a est la plus petite des deux solutions de l'équation.

Conclusion. Voici finalement le théorème caractérisant l'évolution asymptotique de la descendance :

Théorème : Le comportement asymptotique de la descendance est régi par les cas suivants :

(1) S'il existe $c \in \mathbb{N}$ tel que $\mathbb{P}\big\{X_1 = c\big\} = 1$. Alors, deux cas sont possibles :
  (a) Si $c = 0$, alors $a = 1$ (a *priori* exclu par l'hypothèse $p_0 < 1$).
  (b) Si $c \geq 1$, alors $a = 0$.

(2) Si X_1 n'est pas une v.a. dégénérée. Alors,
  (a) Si $m \leq 1$, alors $a = 1$.
  (b) Si $m > 1$, alors $a < 1$. a est l'unique solution dans $[0,1[$ de l'équation $\varphi(x) = x$. De plus, $a = 0 \Leftrightarrow p_0 = 0$.

Preuve : Le point 1 est assez clair. En revanche, nous allons montrer le point 2.

Supposons pour cela que X_1 n'est pas dégénérée.

Plaçons-nous dans le cas $m \leq 1$.

1er cas. Si $p_0 + p_1 = 1$, alors l'équation $\varphi(x) = x$ admet 1 pour unique solution. En particulier,

$a = 1$.

2ème cas. Si $p_0 + p_1 < 1$, alors la fonction φ est strictement croissante et strictement convexe. Rappelons que φ est dérivable en 1, de dérivée m :

$$\lim_{s \to 1,\ s<1} \frac{\varphi(1) - \varphi(s)}{1 - s} = m \leq 1.$$

Or, la fonction $s \in [0, 1[\mapsto \frac{\varphi(1)-\varphi(s)}{1-s} = \frac{1-\varphi(s)}{1-s}$ est strictement croissante.

Donc,

$$\forall\ s \in [0, 1[,\ \frac{1 - \varphi(s)}{1 - s} < 1.$$

En particulier, l'équation $\varphi(x) = x$ n'admet qu'une solution : $x = 1$. On en déduit $a = 1$.

Plaçons nous dans le cas $m > 1$.

Remarquons qu'une telle hypothèse implique $p_0 + p_1 < 1$. La fonction φ est donc strictement croissante et strictement convexe. φ est dérivable en 1, de dérivée m :

$$\lim_{s \to 1,\ s<1} \frac{\varphi(1) - \varphi(s)}{1 - s} = m > 1.$$

Or, la fonction $s \in [0, 1] \mapsto \frac{\varphi(1)-\varphi(s)}{1-s} = \frac{1-\varphi(s)}{1-s}$ si $s < 1$ et m si $s = 1$ est continue.

On sait que :

$$\frac{1 - \varphi(0)}{1 - 0} = 1 - p_0.$$

On en déduit :

(1) Si $p_0 > 0$, alors la fonction continue $s \in [0, 1] \mapsto \frac{1-\varphi(s)}{1-s}$ si $s < 1$ et m si $s = 1$, vaut $1 - p_0 < 1$ en 0 et $m > 1$ en 1. Par le théorème des valeurs intermédiaires, il existe au moins une solution dans $]0, 1[$ de l'équation :

$$\varphi(x) = x.$$

La proposition C.5 affirme qu'il y a en fait une unique dans $]0, 1[$, et que a est égal à cette solution.

(2) Si $p_0 = 0$, alors $\varphi(0) = 0$. Et donc, $a = 0$.

7. Annexe D - Preuve du Théorème 2

Preuve du point (1) : Sous les hypothèses du point (1), nous savons que $\lim_{n \to +\infty} \mathbb{P}\big\{X_n = 0\big\} = 1$. On en déduit que $\lim_{n \to +\infty} \mathbb{P}\big\{\sup_{k \geq n} X_k = 0\big\} = 1$. Ceci implique le résultat.

Preuve du point (2) : La preuve est décomposée en plusieurs étapes. On suppose bien-sûr que $m > 1$:

Étude des moments. Les lemmes D.1 et D.2 nous renseignent sur l'évolution des moments de $(X_n)_{n \geq 0}$:

Lemme D.1 : Pour tout $n \geq 1$, l'égalité suivante est vérifiée :

$$\mathbb{E}(X_{n+1}) = m\mathbb{E}(X_n).$$

En particulier, pour tout $n \geq 1$, $\mathbb{E}(X_n) = m^n$.

Preuve : De la définition de la suite $(X_n)_{n\geq 1}$, on sait que pour tout $n \geq 1$:

$$\begin{aligned}
\mathbb{E}(X_{n+1}) &= \mathbb{E}\Big[\mathbf{1}_{\{X_n\geq 1\}}\Big(\sum_{k=1}^{X_n}\xi_{k,n+1}\Big)\Big] \\
&= \sum_{j\geq 1}\mathbb{E}\Big[\mathbf{1}_{\{X_n=j\}}\Big(\sum_{k=1}^{j}\xi_{k,n+1}\Big)\Big] \\
&= \sum_{j\geq 1}\Big[\mathbb{P}\{X_n=j\}\mathbb{E}\Big(\sum_{k=1}^{j}\xi_{k,n+1}\Big)\Big] = m\sum_{j\geq 1}j\mathbb{P}\{X_n=j\} = m\mathbb{E}(X_n).
\end{aligned}$$

On en déduit bien évidemment que $\forall\ n\geq 1,\ \mathbb{E}(X_n)=m^n$.

Lemme D.2 : Pour tout $n \geq 1$, l'égalité suivante est vérifiée :

$$\mathbb{E}\Big(\frac{X_{n+1}}{m}-X_n\Big)^2 = \frac{\sigma^2}{m^2}m^n.$$

Preuve : On a effectivement :

$$\begin{aligned}
\mathbb{E}\Big(\frac{X_{n+1}}{m}-X_n\Big)^2 &= \mathbb{E}\Big[\mathbf{1}_{\{X_n\geq 1\}}\Big(\frac{1}{m}\sum_{k=1}^{X_n}\xi_{k,n+1}-X_n\Big)^2\Big] \\
&= \sum_{j\geq 1}\mathbb{E}\Big[\mathbf{1}_{\{X_n=j\}}\Big(\frac{1}{m}\sum_{k=1}^{j}\xi_{k,n+1}-j\Big)^2\Big] \\
&= \frac{1}{m^2}\sum_{j\geq 1}\Big[\mathbb{P}\{X_n=j\}\mathbb{E}\Big(\sum_{k=1}^{j}(\xi_{k,n+1}-m)\Big)^2\Big] \\
&= \frac{1}{m^2}\sum_{j\geq 1}\Big(\mathbb{P}\{X_n=j\}\sum_{k=1}^{j}\mathbb{E}(\xi_{k,n+1}-m)^2\Big) \\
&= \frac{1}{m^2}\sum_{j\geq 1}\mathbb{P}\{X_n=j\}j\sigma^2 = \frac{\sigma^2}{m^2}\mathbb{E}(X_n).
\end{aligned}$$

En utilisant le lemme D.1, il vient $\mathbb{E}\Big(\frac{X_{n+1}}{m}-X_n\Big)^2 = \frac{\sigma^2}{m^2}m^n$.

Le corollaire suivant est immédiat :

Corollaire D.3 : La suite $(X_n)_{n\geq 1}$ vérifie la propriété suivante :

$$\forall\ n\geq 1,\ \mathbb{E}\Big(\frac{X_{n+1}}{m^{n+1}}-\frac{X_n}{m^n}\Big)^2 = \frac{\sigma^2}{m^{n+2}}.$$

Convergence de la suite $(m^{-n}X_n)_{n\geq 1}$. La proposition suivante nous donne l'évolution de la suite $(m^{-n}X_n)_{n\geq 1}$:

Proposition D.4 : La suite $\big(m^{-n}X_n\big)_{n\geq 1}$ converge presque-sûrement et dans L^2 vers une v.a. W de carré intégrable et de moyenne 1.

Preuve : On sait que $\forall\ n\geq 1,\ \mathbf{E}\Big(\frac{X_{n+1}}{m^{n+1}}-\frac{X_n}{m^n}\Big)^2 = \frac{\sigma^2}{m^{n+2}}$.

Le critère de Cauchy dans L^2 assure alors la convergence de la suite $(m^{-n}X_n)_{n\geq 1}$ dans L^2 vers une v.a. W de carré intégrable. Le lemme D.1 montre que W est de moyenne 1. Par ailleurs, en utilisant l'égalité précédente et en appliquant l'inégalité de Markov, il vient pour $m^{-1} < \rho^2 < 1$ (remarquons qu'un tel ρ existe car $m > 1$) :

$$\sum_{n \geq 1} \mathbb{P}\left\{ \left| \frac{X_{n+1}}{m^{n+1}} - \frac{X_n}{m^n} \right| > \rho^n \right\} \leq \sum_{n \geq 1} \frac{\sigma^2}{(\rho^2 m)^n m^2} < \infty.$$

Le lemme de Borel-Cantelli (voir le cours) montre que la série $\sum_{n \geq 1} |m^{-(n+1)} X_{n+1} - m^{-n} X_n|$ converge p.s. Ceci implique que la suite $(m^{-n} X_n)_{n \geq 1}$ converge p.s. vers une v.a.r. notée W'. En remarquant que convergence p.s. et convergence L^2 impliquent la convergence en probabilité, on en déduit que $W = W'$ p.s.

Finalement, $\mathbb{P}-$p.s., $m^{-n} X_n \to W$.