Processus de branchement

Caroline Robet

5 septembre 2014

Table des matières

1 Survie ou Extinction ? 5

1.1 Définition d'un processus de branchement 5
1.2 Applications 8
1.3 Progéniture 10
1.4 Moments de l'arbre 11

2 Processus de branchement vu comme une marche aléatoire 14

3 Processus de branchement sur-critique 18

3.1 Loi limite 18
3.2 Population à lignée infinie 23

4 Théorème de temps de retour et de la progéniture totale 26

5 Processus de Poisson 28

6 Modèles plus avancés 35

6.1 Processus de branchement en temps continu 35
6.2 Distribution exponentielle de l'âge 37
6.3 Modèle binaire 41
6.4 Modèle de reproduction sexuée 44

Introduction

L'évolution d'une population et la reproduction sont des phénomènes difficiles à décrire précisément. Cependant, il existe de nombreuses modélisations mathématiques, plus ou moins complexes, qui permettent d'en appréhender, chacune à leur manière, les différents aspects.

Le modèle le plus simple est le modèle de *Maltus* (1798). C'est un modèle discret. Si on note N_i la population à l'instant t_i, alors on a la relation $N_{i+1} = N_i + rN_i$ où r s'interprète comme le taux d'accroissement de la population (différence entre le taux de fécondité et le taux de mortalité). Ce modèle prédit une évolution exponentielle de la population $N_i = N_0(1+r)^i$.

Un des principaux problème de ce modèle, c'est de ne pas prendre en compte les différents individus composant la population étudiée. Pour cela il faut faire appel aux probabilités, comme à travers les processus de branchement.

1 Survie ou Extinction ?

1.1 Définition d'un processus de branchement

Un processus de branchement est un modèle de population probabiliste discret. On considère une population qui évolue de génération en génération. On part de $Z_0 = 1$ et on note Z_n le nombre d'individus à la nième génération en supposant que, après avoir donné naissance, les individus de la (n-1)ième génération meurent.

A chaque génération $n \in \mathbb{N}$, chaque individu i engendre une portée d'individu de la génération suivante, de taille $X_{n+1,i}$ de loi $(p_i)_{i\geq 0}$ où $p_i = \mathbb{P}(avoir\ i\ enfants)$. On a la récurrence $Z_{n+1} = \sum_{i=1}^{Z_n} X_{n+1,i}$. Alors l'évolution de la population suit un processus de branchement si les tailles de chaque famille forment une famille de variables aléatoires mutuellement indépendantes identiquement distribuées. Ce modèle n'est pas spécifique à la reproduction sexuée et se rapproche plus de la division cellulaire.

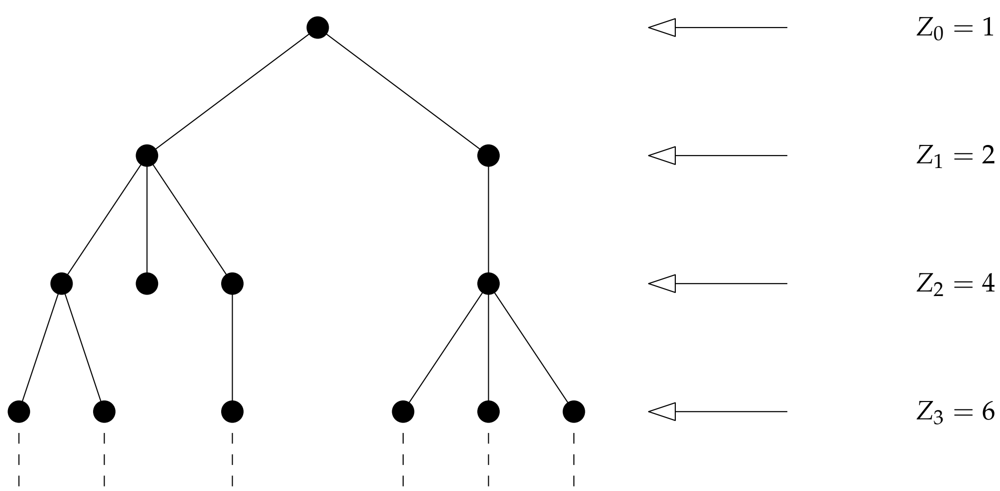

FIGURE 1 – processus de branchement

Théorème 1. *Soit un processus de branchement i.i.d de loi de reproduction X, et $\eta = \mathbb{P}(\exists n : \ Z_n = 0)$ la probabilité d'extinction de la lignée, alors on a :*

- *Lorsque $E(X) < 1$ alors $\eta = 1$ c'est le cas sous-critique.*
- *Lorsque $E(X) > 1$ alors $\eta < 1$ c'est le cas sur-critique*
- *Lorsque $E(X) = 1$ et $\mathbb{P}(X = 1) < 1$ alors $\eta = 1$ c'est la cas critique*

De plus, si G est la fonction génératrice des probabilités de X ($G(s) = E(s^X)$) alors la probabilité d'extinction η est la plus petite solution dans $[0, 1]$ de $\eta = G(\eta)$.

Preuve

- Posons $\eta_n = \mathbb{P}(Z_n = 0)$. Comme $\{Z_n = 0\} \subseteq \{Z_{n+1} = 0\}$, on peut déduire que (η_n) est croissante et comme elle est majorée par 1 (c'est une probabilité) elle est donc convergente vers $\eta \in [0, 1]$. On pose $G_n(s) = E(s^{Z_n})$ qui est la fonction génératrice des probabilités de la nième génération.

  $$G_n(s) = E(s^{Z_n}) = \sum_{i=0}^{+\infty} E(s^{Z_n}|Z_1 = i)\mathbb{P}(Z_1 = i) = \sum_{i=0}^{+\infty} p_i E(s^{Z_n}|Z_1 = i)$$

  On note alors $Z'_{n-1,j}$ le processus de branchement issu du j^{eme} individu de la première génération. pour obtenir

  $$\begin{aligned} G_n(s) &= \sum_{i=0}^{+\infty} p_i E(s^{Z'_{n-1,1}+Z'_{n-1,2}+\ldots+Z'_{n-1,i}}|Z_1 = i) = \sum_{i=0}^{+\infty} p_i \prod_{j=1}^{i} E(s^{Z'_{n-1,j}}) \\ &= \sum_{i=0}^{+\infty} p_i E(s^{Z'_{n-1,1}})^i = \sum_{i=0}^{+\infty} p_i G_{n-1}(s)^i \end{aligned}$$

  donc $G_n(s) = G(G_{n-1}(s))$ car $G(s) = E(s^{Z_1}) = \sum_{i=0}^{+\infty} p_i s^i$ d'où en évaluant pour s=0, on obtient $\eta_n = G(\eta_{n-1})$. Par continuité de G sur $[0, 1]$, on obtient par passage à la limite que $\eta = G(\eta)$

- *cas 1* :
  Lorsque $p_0 = 0$, on obtient $\eta = 0$ car $\eta = \mathbb{P}(\exists n | Z_n = 0)$. Si $p_0 = 0$ alors chaque individu a au moins un enfant et donc $Z_{n+1} \geq Z_n \geq ... \geq Z_0 = 1$. D'où $\forall n$, $Z_n \geq 1$ et donc $\eta = 0$.

- *cas 2*
  Lorsque $\mathbb{P}(X = 1) = 1$ on a alors $E(X) = 1$ et $\forall n$, $Z_n = 1$ p.s. Il n'y a donc rien à prouver.

- *cas 3*
  Puis, si $\mathbb{P}(X \leq 1) = 1$ mais $p = \mathbb{P}(X = 0) > 0$ alors

  $$\mathbb{P}(Z_n = 0) = 1 - \mathbb{P}(Z_n = 1) = 1 - \mathbb{P}(Z_1 = 1)^n = 1 - (1 - p)^n \underset{n \to +\infty}{\longrightarrow} 1$$

- *cas général*
  Enfin pour le reste de la preuve, on prend $\mathbb{P}(X \leq 1) < 1$. On suppose que $\psi \in [0, 1]$ vérifie $\psi = G(\psi)$. On va prouver que $\eta \leq \psi$ en montrant par récurrence que $\forall n$, $\eta_n \leq \psi$: Pour n=0, $\eta_0 = \mathbb{P}(Z_0 = 0) = 0 \leq \psi \in [0, 1]$ On suppose la propriété vraie au rang n, $\eta_n \leq \psi$ donc $G(\eta_n) \leq G(\psi)$ par croissance de G d'où $\eta_{n+1} \leq \psi$. Par limite, $\eta \leq \psi$ donc η est la plus petite solution dans [0,1] de $G(x) = x$.
  On remarque que la fonction $s \to G(s)$ est croissante et convexe avec $G''(s) = E(X(X-1)s^{X-2}) \geq 0$. Lorsque $\mathbb{P}(X \leq 1) < 1$, alors $G''(s) > 0$ donc la fonction est strictement convexe pour $s > 0$. Il existe donc au plus deux solutions dans [0,1] de $x = G(x)$.

  Si $G(0) > 0$, il y a une unique solution quand $G'(1) = E[X] < 1$, qui est $\eta = 1$. Il y a 2 solutions distinctes quand $G'(1) = E[X] > 1$, qui sont η et 1 donc $\eta < 1$. Il y a une unique solution quand $G'(1) = E[X] = 1$, qui est $\eta = 1$ sauf lorsque G(s)=s qui est le cas $\mathbb{P}(X = 1) = 1$ où alors on a vu que $\forall n$, $Z_n = 1$ et donc $\eta = 0$
  $\square$

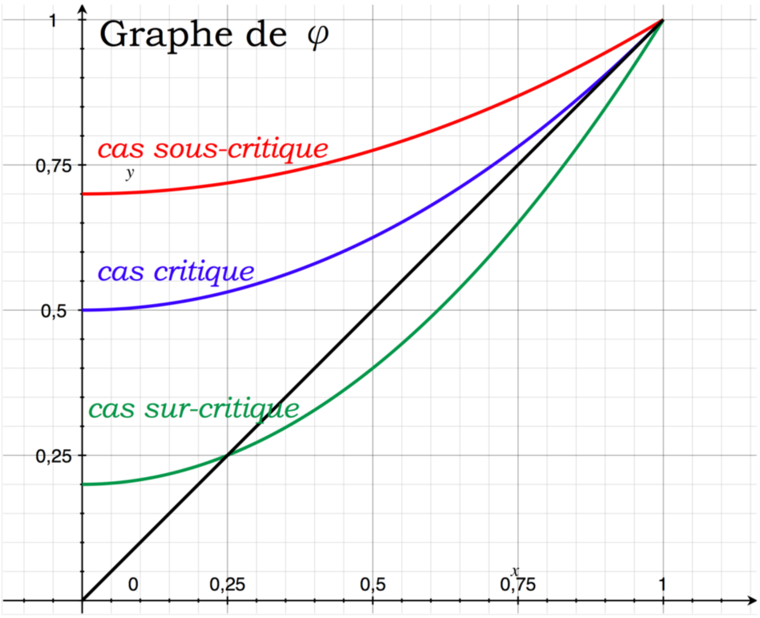

FIGURE 2 – Fonction génératrice de probabilité et probabilité d'extinction suivant l'espérance de la loi de reproduction

1.2 Applications

1) Lorsque la loi est donnée par $p_x = (1-p)\delta_0 + p\delta_2$, on parle de processus de branchement binaire. Trouver η en fonction de p

Si $p \leq \frac{1}{2}$, on a $E(X) = 2p \leq 1$ et comme $\mathbb{P}(X = 1) = 0$ alors $\eta = 1$.
Si $p > \frac{1}{2}$, on a $E(X) > 1$. On cherche η vérifiant $\eta = G(\eta)$ ici $G(x) = (1-p) + px^2$. Donc $\eta = 1 - p + p\eta^2$ on trouve alors $\eta = \frac{1-p}{p}$

2) Trouver la probabilité d'extinction pour une loi géométrique décalée, i.e $p_k = (1-p)^k p$

On a $E(X) = \sum\limits_{k=1}^{\infty} k(1-p)^k p = \frac{1-p}{p}$. Donc lorsque $E(X) \leq 1$ c'est à dire $p \geq \frac{1}{2}$ alors $\eta = 1$. Sinon on a

$$G(X) = E(s^X) = \sum_{k=0}^{\infty} (1-p)^k p s^k = p \sum_{k=0}^{\infty} (s(1-p))^k = \frac{p}{1 - s(1-p)}$$

. On calcule alors $\eta = \frac{p}{1-p}$.

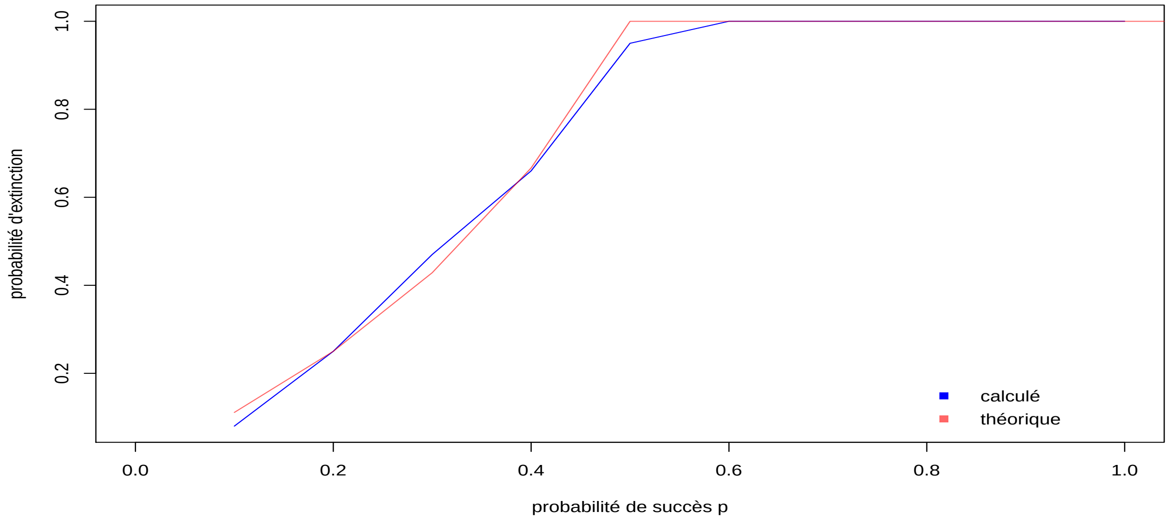

FIGURE 3 – Probabilité d'extinction d'une loi géométrique calculée à partir de la simulation d'une population en regardant la fréquence d'extinction de la population et probabilité théorique donnée par l'expression de η

3) Soit la loi donnée par $p_0 = 1 - b/p$ et $p_k = b(1-p)^{k-1}$. Trouver la probabilité d'extinction η

On a

$$E(x) = b\sum_{k=0}^{\infty} k(1-p)^{k-1} = -b(\sum_{k=0}^{\infty}(1-p)^k)' = -b(\frac{1}{1-(1-p)})' = b/p^2$$

.

- Si $E(X) = b/p^2 < 1$ alors $\eta = 1$.
- Si $E(x) = b/p^2 = 1$ alors comme $\mathbb{P}(X = 1) = b = p^2 < 1$ donc $\eta = 1$.
- Si $E(X) = b/p^2 > 1$ alors on cherche $G(\eta) = \eta$ d'où

$$(1-b/p) + \sum_{k=1}^{\infty} b(1-p)^{k-1}\eta^k = \eta$$

,

$$(1-b/p) + \eta b\frac{1}{1-(1-p)\eta} = \eta$$

et $\eta = \frac{1-b/p}{q}$

- *Nous montrons par récurrence que la fonction génératrice de probabilité* G_n *de* Z_n *est* $G_n(s) = \begin{cases} 1 - \mu^n \frac{1-\eta}{\mu^n-\eta} + \frac{\mu^n(\frac{1-\eta}{\mu^n-\eta})^2 s}{1-\frac{\mu^n-1}{\mu^n-\eta}} & lorsque\ b \neq p^2 \\ \frac{nq-(nq-p)s}{p+nq-nps} & lorsque\ b = p^2 \end{cases}$

En déduire que $\mathbb{P}(Z_n > 0, \exists m > n\ tel\ que\ Z_m = 0) = \begin{cases} \mu^n \frac{1-\eta}{\mu^n-\eta} & lorsque\ b < p^2 \\ \frac{p}{p+nq} & lorsque\ b = p^2 \\ \frac{(1-\eta)\eta}{\mu^n-\eta} & lorsque\ b > p^2 \end{cases}$

On a

$$\begin{aligned} \mathbb{P}(Z_n > 0, \exists m > n\ tel\ que\ Z_m = 0) &= \sum_{j=n+1}^{\infty} \mathbb{P}(Z_{j-1} > 0\ et\ Z_j = 0) \\ &= \sum_{j=n+1}^{\infty} (\mathbb{P}(Z_j = 0) - \mathbb{P}(Z_{j-1} = 0)) \\ &= \mathbb{P}(Z_\infty = 0) - \mathbb{P}(Z_n = 0) = \mathbb{P}(Z_\infty = 0) - G_n(0) \end{aligned}$$

- Si $b < p^2$ alors $\mathbb{P}(Z_\infty = 0) = 1$ et $G_n(0) = 1 - \mu^n \frac{1-\eta}{\mu^n-\eta}$ donc $\mathbb{P}(Z_n > 0, \exists m > n\ tel\ que\ Z_m = 0) = \mu^n \frac{1-\eta}{\mu^n-\eta}$
- Si $b = p^2$ alors $\mathbb{P}(Z_\infty = 0) = 1$ et $G_n(0) = \frac{nq}{p+nq}$ donc $\mathbb{P}(Z_n > 0, \exists m > n\ tel\ que\ Z_m = 0) = \frac{p}{p+nq}$
- Si $b > p^2$ alors $\mathbb{P}(Z_\infty = 0) = \eta$ et $G_n(0) = 1 - \mu^n \frac{1-\eta}{\mu^n-\eta}$ donc $\mathbb{P}(Z_n > 0, \exists m > n\ tel\ que\ Z_m = 0) = \frac{(1-\eta)\eta}{\mu^n-\eta}$

1.3 Progéniture

Par la suite, on va souvent s'intéresser à la probabilité de survie qui est noté $\zeta = 1-\eta$ qui correspond à la probabilité que le processus de branchement survive pour toujours : $\zeta = \mathbb{P}(Z_n > 0\ \forall n \geq 0)$.

Définition 2. *La progéniture totale du processus de branchement est notée* T *et est définie par* $T = \sum_{n=0}^{+\infty} Z_n$. *On pose* $G_T(s)$ *la fonction génératrice de probabilité de* T : $G_T(s) = E(s^T)$

Théorème 3. *Pour un processus de branchement i.i.d de loi de reproduction X et de fonction génératrice de probabilité G, la fonction génératrice de probabilité de la progéniture totale vérifie $G_T(s) = sG(G_T(s))$*

Preuve
On a $Z_0 = 1$. Si $Z_1 = i$ alors on appelle T_j, $j \in \{1 \dots i\}$ la progéniture totale du j^{ieme} enfant de la première génération. On obtient $T = 1 + \sum_{j=1}^{i} T_j$ où 1 correspond à l'individu initial. Donc

$$\begin{aligned} G_T(s) = E(s^T) &= \sum_{i=0}^{+\infty} E(s^T|Z_1 = i)\mathbb{P}(Z_1 = i) = \sum_{i=0}^{+\infty} p_i E(s^T|Z_1 = i) \\ &= \sum_{i=0}^{+\infty} p_i E(s^{1+T_1+..+T_i}|Z_1 = i) = \sum_{i=0}^{+\infty} s\ p_i E(s^{T_1+..+T_i}|Z_1 = i) \end{aligned}$$

Or on a $(T_j)_{j\in\{1..i\}}$ est une suite i.i.d de variables de loi T donc

$$G_T(s) = \sum_{i=0}^{+\infty} s\ p_i E(s^{T_1})^i = \sum_{i=0}^{+\infty} s\ p_i G_T(s)^i = sG(G_T(s))$$

□

Application
4) Dans le cas d'un branchement binaire donné par l'application 1, trouver G_T
On veut

$$G_T(s) = sG(G_T(s)) = s(\ (1-p) + p[G_T(s)]^2\ ).$$

Donc $G_T(s)_1 = \frac{1-\sqrt{1-4s^2pq}}{2sp}$ ou $G_T(s)_2 = \frac{1+\sqrt{1-4s^2pq}}{2sp} > 1$ pour s petit donc impossible car on doit avoir $G_T(s) \in [0,1]$ pour tout $s \in [0,1]$ donc on garde $G_T(s)_1$.

1.4 Moments de l'arbre

Nous calculons la taille moyenne d'une génération dans un processus de branchement pour en déduire la progéniture totale moyenne.

Théorème 4. *Pour tout* $n \geq 0$ *et si* $\mu = E(Z_1)$ *est l'espérance du nombre d'enfants alors on a :*

$$E(Z_n) = \mu^n$$

Preuve

On procède par récurrence. On a la propriété pour n=0 car $E(Z_0) = 1 = \mu^0$. Puis, $Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i}$ avec $(X_{n,i})_i$ i.i.d de moyenne μ et indépendant de Z_{n-1} donc

$$\begin{aligned}
E(Z_n) = E(E(Z_n|Z_{n-1})) &= E\Big(E\Big(\sum_{i=1}^{Z_{n-1}} X_{n,i}|Z_{n-1}\Big)\Big) \\
&= E\Big(\sum_{j\geq 0} \mathbb{P}(Z_{n-1} = j) \sum_{i=1}^{j} E(X_{n,i})\Big) = \sum_{j\geq 0} \mathbb{P}(Z_{n-1} = j) j\mu \\
&= \mu E(Z_{n-1}) = \mu^n
\end{aligned}$$

Théorème 5. *Soit* $n \geq 0$. *On pose* $\mu = E(Z_1)$ *la loi de reproduction d'un individu et on suppose* $\mu < 1$ *alors*

$$\mathbb{P}(Z_n > 0) \leq \mu^n$$

Preuve

D'après l'inégalité de Markov, on a $\mathbb{P}(Z_n \geq a) \leq \frac{E(Z_n)}{a}$. On a de plus Z_n qui est entier donc $\mathbb{P}(Z_n > 0) = \mathbb{P}(Z_n \geq 1) \leq E(Z_n) = \mu^n$.

Remarque : En régime sous-critique ($\mu < 1$), la probabilité de survie décroit exponentiellement avec n.

Théorème 6. *Pour un processus de branchement de loi de reproduction i.i.d* X*, on a pour progéniture moyenne lorsque* $\mu < 1$, $E(T) = \frac{1}{1-\mu}$

Preuve

Comme $T = \sum_{n=0}^{+\infty} Z_n$, on a $E(T) = \sum_{n=0}^{+\infty} E(Z_n) = \sum_{n=0}^{+\infty} \mu^n = \frac{1}{1-\mu}$ $\square$

2 Processus de branchement vu comme une marche aléatoire

Dans les processus de branchement, on étudie souvent le nombre de descendants de chaque génération. Pour les besoins des graphes aléatoires, il est souvent commode d'utiliser un mode de construction différent d'un processus de branchement par recherche séquentielle du nombre d'enfants pour chaque individu de la population.

Nous donnons maintenant la représentation d'une marche aléatoire d'un processus de branchement. Soient X_1, X_2,... des variables aléatoires de loi X. On définit

$$\begin{cases} S_0 = 1 \\ S_i = S_{i-1} + X_i - 1 = X_1 + ... + X_i - (i-1) \end{cases}$$

On note $T = \min\{t \colon S_t = 0\} = \min\{t \colon X_1 + ... + X_t = t - 1\}$. Si un tel t n'existe pas alors on définit $T = +\infty$.

On va montrer que T correspond bien à la progéniture totale d'un processus de branchement. Pour cela, on note X_1 le nombre d'enfants de l'individu initial et on définit $S_1 = S_0 + X_1 - 1 = X_1$. On a alors S_1 individus non-explorés, c'est-à-dire les individus dont on n'a pas encore exploré combien d'enfants ils ont.

Lemme 7. *Le processus aléatoire $(S_i)_{i=0}^{+\infty}$ défini ci-dessus a la même distribution que le processus aléatoire $(S'_i)_{i=0}^{+\infty}$ où S'_i désigne le nombre d'individus non explorés dans l'exploration d'un processus de branchement après l'exploration successive de i individus.*

Preuve
Par récurrence sur i :

Pour $i = 0$, ok car on part d'un individu.

On suppose la propriété vraie pour S_{i-1}.
- Si $S_{i-1} = 0$, nous avons fini car alors tous les individus ont été exploré une fois et le nombre total d'individus explorés est égale à la taille de l'arbre qui est T par définition.

- Si $S_{i-1} > 0$. Alors on choisit arbitrairement un individu non exploré et on appelle X_i le nombre de ses enfants donc le nouveau nombre d'individus non exploré est égal à $S_i = S_{i-1} + X_i - 1$ car on a rajouté les X_i enfants et on a enlevé l'individu étudié. $\square$

On note H=$(X_1, ..., X_T)$ l'histoire du processus jusqu'au temps T. La suite $(x_1, .., x_t)$ est une histoire possible si et seulement si la suite vérifie

$$\begin{cases} s_0 = 1 \\ s_i = x_1 + ... + x_i - (i-1) \quad avec\ s_i > 0\ \forall i < t \\ s_t = 0 \end{cases}$$

alors lorsque $t < \infty$, $\mathbb{P}(H = (x_1, ..., x_t)) = \prod_{i=1}^{t} p_{x_i}$. Il faut noter que cette formule détermine la loi d'un processus de branchement conditionné à l'extinction. Nous allons utiliser la perspective de la marche aléatoire pour décrire la loi d'un processus de branchement conditionné à l'extinction.

Définition 8. *On dit que p et p' est une paire conjuguée si* $p'_x = \eta^{x-1} p_x$ *où* η *désigne la probabilité d'extinction de la loi* $(p_x)_{x=0}^{\infty}$.

La relation entre un processus de branchement sur-critique conditionné à l'extinction et son processus de branchement conjugué est :

Théorème 9. *Soient p et p' des distributions conjuguées. Le processus de branchement de distribution p conditionné à l'extinction a la même distribution que le processus de branchement de distribution p'.*

Preuve
Il suffit de montrer que pour toute histoire finie H=$(x_1, ..., x_t)$ la probabilité est la même pour un processus de branchement conditionné à l'extinction et un processus de branchement de loi p'.

$$\begin{aligned}\mathbb{P}(H=(x_1,...,x_t)|extinction) &= \frac{\mathbb{P}(H=(x_1,...,x_t)\cap extinction)}{\mathbb{P}(extinction)}\\ &= \frac{\mathbb{P}(H=(x_1,...,x_t)}{\eta}\end{aligned}$$

Puis on a montré :

$$\mathbb{P}(H=(x_1,...,x_t)) = \prod_{i=1}^{t} p_{x_i} = \prod_{i=1}^{t} p'_{x_i}\eta^{-(x_i-1)} = \eta^{t-\sum\limits_{i=1}^{t} x_i}\prod_{i=1}^{t} p'_{x_i}$$

Par définition de t on a $x_1+..+x_t = t-1$, donc $\mathbb{P}(H=(x_1,...,x_t)) = \eta\prod\limits_{i=1}^{t} p'_{x_i}$ d'où $\mathbb{P}(H=(x_1,...,x_t)|extinction) = \mathbb{P}'(H=(x_1,...,x_t))$ où $\mathbb{P}'$ est la distribution d'un processus de branchement de probabilité p'. $\square$

Application *Soit $G_d(s)=E'(s^{X_1})$ la fonction génératrice de probabilité du processus de branchement dual. On peut prouver que $G_d(s)=\frac{1}{\eta}G(\eta s)$*
$G_d(s)=E'(s^{X_1})=\sum\limits_{i=0}^{\infty}p'_i s^i=\sum\limits_{i=0}^{\infty}\eta^{i-1}p_i s^i=\frac{1}{\eta}\sum\limits_{i=0}^{\infty}p_i(s\eta)^i=\frac{1}{\eta}G(\eta s)$

Théorème 10. *Pour un processus de branchement de loi de reproduction i.i.d X de moyenne $\mu=E(X)>1$, on a*

$$\mathbb{P}(k\leq T<+\infty)\leq \tfrac{e^{-Ik}}{1-e^{-I}} \text{ avec } I=\sup_{t\leq 0}(t-ln(E(e^{tX})))$$

Ce théorème peut être reformulé en disant que lorsque la progéniture totale est grande alors le branchement va survivre avec une grande probabilité.

Preuve
Si T=s alors $S_s=0$ et $X_1+...+X_s=s-1\leq s$
donc $\mathbb{P}(k\leq T<+\infty)\leq\sum\limits_{s=k}^{\infty}\mathbb{P}(S_s=0)\leq\sum\limits_{s=k}^{\infty}\mathbb{P}(X_1+...+X_s\leq s)$

On utilise ensuite le théorème de Cramer :
$\forall a \leq E(X_1)$, on a $\mathbb{P}(\sum\limits_{i=1}^{n} X_i \leq na) \leq e^{-nI(a)}$ avec I(a)=$\sup\limits_{t\leq 0}(ta - ln(E(e^{tX})))$
Il suffit de prendre a=1 pour obtenir $\mathbb{P}(\sum\limits_{i=1}^{s} X_i \leq s) \leq e^{-sI}$.
Donc $\mathbb{P}(k \leq T < +\infty) \leq \sum\limits_{s=k}^{\infty} e^{-sI} = \frac{e^{-Ik}}{1-e^{-I}}$ $\square$

3 Processus de branchement sur-critique

Dans cette partie, nous allons prouver un résultat de convergence pour la nième génération. Dans le cas sous-critique, on a $Z_n \longrightarrow 0$ car $\eta = 1$

Lemme 11. *Le processus ne peut pas se stabiliser dans le cas sur-critique. On a alors* $\lim_{n\to\infty} Z_n = 0$ *ou* $\lim_{n\to\infty} Z_n = \infty$ *et*

$$\mathbb{P}(\lim_{n\to\infty} Z_n = 0) = \eta = 1 - \mathbb{P}(\lim_{n\to\infty} Z_n = \infty)$$

Preuve
Montrons que tous les états $k \geq 1$ sont transients. C'est le cas si en posant $R_k = \mathbb{P}(\exists j \geq 1, Z_{n+j} = k | Z_n = k)$ on a $R_k < 1$. Si $p_0 = 0$ alors $R_k = p_1 < 1$. Si $p_0 > 0$ alors $R_k \leq 1 - \mathbb{P}(\text{passer de l'état k à l'état 0}) \leq 1 - p_0^k < 1$. D'où $\mathbb{P}(Z_n = k \text{ pour une infinité de n}) = 0$ d'où $Z_n \longrightarrow 0$ ou $Z_n \longrightarrow \infty$

3.1 Loi limite

Théorème 12 (Convergence d'un processus sur-critique)**.** *Pour un processus de branchement sur-critique de loi de reproduction X , on a :*

$$\tfrac{Z_n}{\mu^n} \xrightarrow{p.s} W_\infty \textit{ variable aléatoire qui est finie p.s .}$$

Preuve
On utilise le théorème de la convergence des martingales : si (M_n) est une martingale bornée dans L^1 et positive alors $M_n \xrightarrow{p.s} M_\infty$ qui est finie p.s.
Si on prend $M_n = \frac{Z_n}{\mu^n}$, on montre que M_n est une martingale. En effet, $(\frac{Z_n}{\mu^n})$ est adaptée à $\sigma(Z_0, .., Z_n)$, $E(\frac{Z_n}{\mu^n}) = \frac{E(Z_n)}{\mu^n} = 1$ et $E(\frac{Z_{n+1}}{\mu^{n+1}} | F_n) = \frac{1}{\mu^{n+1}} E(\sum_{i=1}^{Z_n} X_{n,i} | F_n) = \frac{1}{\mu^{n+1}} Z_n \mu = \frac{Z_n}{\mu^n}$. $\square$

Malheureusement, nous ne connaissons pas beaucoup de chose à propos de la loi limite W_∞.

Théorème 13. *La fonction génératrice de probabilité de W_∞, $G_W(s) = E(s^{W_\infty})$ vérifie la relation pour $s \in [0,1]$ $G_W(s) = G(G_W(s^{\frac{1}{\mu}}))$.*

Preuve
En effet $E(s^{\frac{Z_n}{\mu^n}}) \longrightarrow E(s^{W_\infty}) = G_W(s)$ car $\frac{Z_n}{\mu^n}$ converge p.s vers W_∞ et on applique la convergence dominée pour une martingale positive pour obtenir la convergence des espérances. Or

$$E(s^{\frac{Z_n}{\mu^n}}) = E(E(s^{\frac{Z_n}{\mu^n}}|Z_1)) = E(E(s^{Z_{n-1}/\mu^n})^{Z_1}) = G(E(s^{\frac{Z_{n-1}}{\mu^n}}))$$

$$E(s^{\frac{Z_n}{\mu^n}}) = G\Big[E\Big((s^{\frac{1}{\mu}})^{\frac{Z_{n-1}}{\mu^{n-1}}}\Big)\Big] \longrightarrow G\Big(G_W(s^{\frac{1}{\mu}})\Big)$$

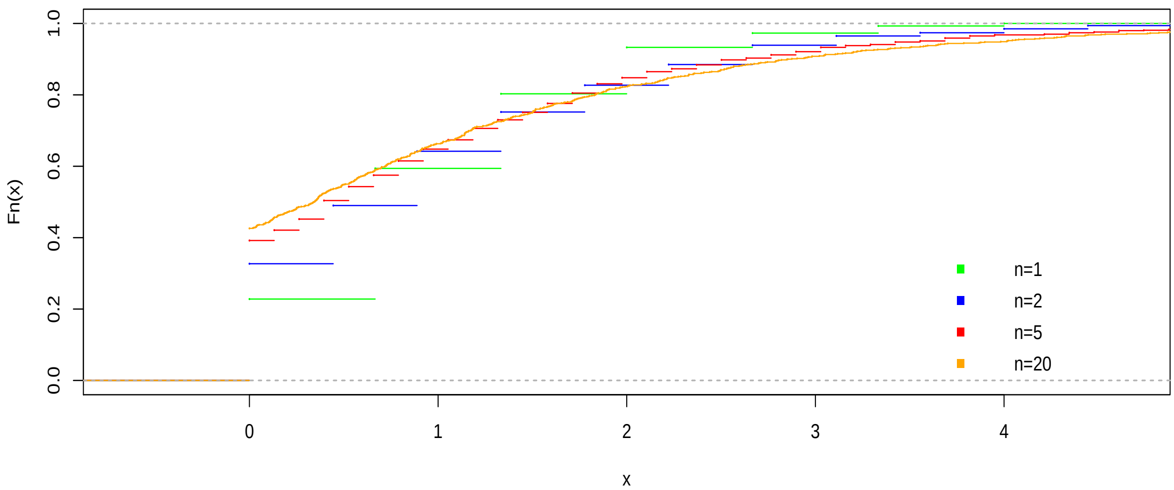

FIGURE 4 – fonction de répartition de la loi limite W_∞ calculée grâce à la fonction de répartition empirique de $\frac{Z_n}{\mu^n}$ pour différents n

Théorème 14. *Pour un processus de branchement iid de loi de reproduction X de moyenne $\mu > 1$, on a :*

1. *Si $E(X\log(X)) < \infty$, alors $\mathbb{P}(W_\infty = 0) = \eta$ et $E(W_\infty) = 1$.*
2. *Si $E(X\log(X)) = \infty$, alors $\mathbb{P}(W_\infty = 0) = 1$.*

Remarque : Dans le cas 1, on a précisément $\{W_\infty = 0\} = \{Z_n \underset{n\to\infty}{\longrightarrow} 0\}$.
Pour prouver ce théorème, nous allons devoir montrer les lemmes et corollaires suivants :

Lemme 15. *Soit X une variable aléatoire positive avec $0 < E(X) = m < \infty$ alors $\forall a > 0$ $\int_0^a \frac{1}{u^2}(E(e^{-uX/m}) - e^{-u})du < \infty$ si et seulement si $E(X|$ $log(X)|) < \infty$.*

Preuve
On commence par écrire

$$E(e^{-uX/m}) - e^{-u} = E(e^{-uX/m} - 1 + \frac{uX}{m}) + 1 - u - e^{-u}$$

Pour $u \geq 0$, on a $0 \leq e^{-u} - 1 + u \leq \frac{u^2}{2}$ d'où $0 \leq \int_0^a \frac{e^{-u}-1+u}{u^2}du \leq \int_0^a \frac{1}{2} = \frac{a}{2}$ donc la deuxième partie de l'égalité ne pose pas problème. Étudions désormais $E(e^{-uX/m} - 1 + \frac{uX}{m})$, pour montrer que $\int_0^a \frac{E(e^{-uX/m}-1+\frac{uX}{m})}{u^2}du < \infty$ si et seulement si $E(X|\log(X)|) < \infty$.
Soit $F(x) = \mathbb{P}(X \leq mx)$, on a :

$$\begin{aligned}
\int_0^a \frac{E(e^{-uX/m} - 1 + \frac{uX}{m})}{u^2}du &= \int_0^a u^{-2}(\ \int_0^\infty (e^{-ux} - 1 + ux)dF(x)\ )du \\
&= \int_0^\infty (\ \int_0^a u^{-2}(e^{-ux} - 1 + ux)du\ )dF(x) \\
&= \int_0^\infty (\ x \int_0^{ax} v^{-2}(e^{-v} - 1 + v)dv\ )dF(x)
\end{aligned}$$

puis

$$\begin{aligned}
\int_0^a u^{-2}(e^{-ux} - 1 + ux)du &= x\int_0^{ax} v^{-2}(e^{-v} - 1 + v)dv \\
&= x(\int_0^1 v^{-2}(e^{-v} - 1 + v)dv\ +\ \int_1^{ax} v^{-2}(e^{-v} - 1 + v)dv)
\end{aligned}$$

On a $\int_0^1 v^{-2}(e^{-v} - 1 + v) \leq \frac{1}{2}$ et $\int_1^{ax} v^{-2}(e^{-v} - 1) < \infty$ il nous reste donc $\int_1^{ax} \frac{1}{v} = \log(ax) = \log(a) + \log(x)$ d'où $\int_1^{ax} v^{-2}(e^{-v} - 1 + v)dv \underset{x\to\infty}{\sim} \log(x)$ $\int_0^a u^{-2}(e^{-ux} - 1 + ux)du \underset{x\to\infty}{\sim} x\log(x)$ d'où le résultat. $\square$

Corollaire 16. *Soit* $G(s) = \sum\limits_{j=0}^{\infty} p_j s^j$ *la fonction génératrice de probabilité.*
Si $0 < m = G'(1) < \infty$ *alors* $\int_0^1 u^{-2}(G(e^{-u/m}) - e^{-u}du < \infty$ *ssi* $\sum\limits_{j=0}^{\infty} p_j\, j\, \log(j) < \infty$

Preuve
On prend X une variable aléatoire positive avec G sa fonction génératrice.
On applique le lemme précédent car $E(e^{-uX/m}) = G(e^{-u/m})$ et
$E(X|\log(X)|) = \sum\limits_{j=0}^{\infty} p_j\, j\, \log(j)$. On prend a=1 et on a l'égalité. □

Corollaire 17. *Soit G la fonction génératrice de probabilité et*
$0 < m = G'(1) < \infty$*. Soit la fonction A définie par :*

$$A(u) = \begin{cases} m - \frac{1-G(1-u)}{u} \ \textit{pour}\ 0 < u \leq 1 \\ 0 \ \textit{pour}\ u = 0 \end{cases}$$

alors A(u) est positive et croissante et de plus
$\forall r, c \in]0,1[,\ \sum\limits_{n=0}^{\infty} A(cr^n) < \infty$ *ssi* $\sum\limits_{j=0}^{\infty} j\, \log(j)\, p_j < \infty$

Preuve
On a $\lim\limits_{u \to 0} A(u) = 0$ car $\lim\limits_{u \to 0} \frac{1-G(1-u)}{u} = \lim\limits_{u \to 0} \frac{G(1)-G(1-u)}{u} = G'(1) = m$.
G est convexe donc $A(u) \geq 0$ et $A'(u) \geq 0$ ou alors on utilise le fait que A vérifie $A(u) = \sum\limits_{j=1}^{\infty} p_j (\sum\limits_{r=0}^{j-1} (1-(1-u)^r)\ ) \geq 0$ nous donne A est positive et croissante.

Ensuite, par monotonie de A (comparaison série-$\int$), on a $\sum\limits_{n=0}^{\infty} A(cr^n) < \infty$ ssi $\int_0^{\infty} A(cr^t)dt < \infty$. On pose $v = cr^t$ d'où $\frac{dv}{v} = ln(r)dt$ pour obtenir $\int_0^{\infty} A(cr^t)dt = \frac{1}{ln(r)} \int_c^0 \frac{A(v)}{v} dv$ donc on doit montrer que $\int_0^1 \frac{A(v)}{v} dv < \infty$.
Comme $\frac{A(v)}{v} = v^{-2}(G(1-v) - 1 + mv)$ on prend ensuite $1 - v = e^{-u/m}$ pour obtenir $\frac{A(v)}{v} = v^{-2}(\ G(e^{-u/m}) - e^{-u} + m(1 - e^{-u/m} - \frac{u}{m}) + (e^{-u} - 1 + u)\ )$

Maintenant $\frac{x}{1-e^{-x}}$ est borné et supérieur à 1 sur tout intervalle fini et $|1-x-e^{-x}|x^{-2} \leq \frac{1}{2}$ pour $x \geq 0$. Donc $\int_0^1 \frac{A(v)}{v}dv < \infty$ ssi $\int_0^c (G(e^{-u/m}) - e^{-u})u^{-2}du < \infty$. Donc ok par le corollaire 16 $\square$

Preuve du théorème 14

Étape 1 : On veut montrer que $E(X\log(X)) < \infty$ entraîne $E(W_\infty) = 1$.

On remarque que $\phi_n(u) = E(e^{-uW_n})$ avec $W_n = \frac{Z_n}{\mu^n}$ est croissante en n pour u fixé : cela provient de l'inégalité de Jensen et du fait que W_n est une martingale : $E(e^{-uW_{n+1}}|Z_n) \geq e^{-uE(W_{n+1}|Z_n)} = e^{-uW_n}$ d'où en prenant l'espérance des 2 côtés, on obtient : $E(e^{-uW_{n+1}}) \geq E(e^{-uW_n})$
On pose $\psi_{n+1}(u) = \frac{\phi_{n+1}(u)-\phi_n(u)}{u}$ pour $u > 0$. Or $\phi_{n+1}(u) = G(\phi_n(\frac{u}{m}))$ donc

$$\psi_{n+1}(u) = \frac{G(\phi_n(u/m)) - G(\phi_{n-1}(u/m))}{u} \leq \sup_{[0,1]}||G'||\frac{|\phi_n(u/m) - \phi_{n-1}(u/m)|}{u}$$

or G est convexe donc $\sup_{[0,1]}||G'|| = G'(1) = m$ et ϕ est croissante en n donc on peut enlever les valeurs absolues.

$$\psi_{n+1}(u) \leq m\frac{\phi_n(u/m) - \phi_{n-1}(u/m)}{u} = \psi_n(u/m)$$

Par récurrence, $\psi_{n+1}(u) \leq \psi_1(u/m^n)$ or $\phi_0(u) = E(e^{-u}) = e^{-u}$ donc

$$0 \leq \frac{\phi(u) - \phi_0(u)}{u} = \frac{\phi(u) - e^{-u}}{u} = \lim_{n\to\infty} \frac{\phi_{n+1}(u) - \phi_0(u)}{u} = \sum_{n=1}^{\infty} \psi_n(u)$$

On doit montrer que si $E(X\log(X)) < \infty$ alors $g(u) = \sum_{n=1}^{\infty} \psi_n(u) < \infty$ pour tout $u > 0$ et $\lim_{u\to 0} g(u) = 0$. Cela montrera que

$$\lim_{u\to 0}\frac{1-\phi(u)}{u} = \lim_{u\to 0}(\frac{1-e^{-u}}{u} + \frac{e^{-u}-\phi(u)}{u}) = (e^x)'|_0 + \lim_{u\to 0} g(u) = 1 = \phi'(0) = E(W_\infty)$$

On utilise la monotonie de ψ_1.

$$0 \leq g(u) = \sum_{n=1}^{\infty} \psi_n(u) \leq \sum_{n=1}^{\infty} \psi_1(u/m^{n-1}) \leq \int_0^\infty \psi_1(u/m^t)dt$$

On pose $v = \frac{u}{m^t}$, on a pour u petit, $0 \leq g(u) \leq \int_0^u \frac{\psi_1(v)}{v} dv$. Par le corollaire 16, on a $0 \leq g(u) \leq \int_0^u \frac{\psi_1(v)}{v} dv < \infty$ donc $g(u) \longrightarrow 0$

ÉTAPE 2 : On veut montrer que $E(X\log(X)) = \infty$ entraîne $E(W_\infty) = 0$ ($\Leftrightarrow \mathbb{P}(W_\infty = 0) = 1$) avec $E(W_\infty) = \lim\limits_{u\to 0} \frac{1-\phi(u)}{u}$.

On suppose que $E(W_\infty) > 0$ alors il existe α et $\beta \in]0,1[$ tel que $\frac{1-\phi(u)}{u} \geq 2\beta$ pour $u \in [0,\alpha]$. Soit $\lambda_n(u) = \frac{1-\phi_n(u)}{u}$ comme $\phi_n(u) \to \phi(u)$, il existe n_0 tel que pour $n \geq n_0$ on ait $\lambda_n(\alpha) > \beta$. Or $\forall n$, $\lambda_n(u)$ est une fonction décroissante en $u > 0$. Cela nous donne $\lambda_n(u) > \beta$ pour $0 \leq u \leq \alpha$ et $n \geq n_0$. Comme $\phi_{n+1}(u) = G(\phi_n(u/m))$, on obtient

$$\lambda_{n+1}(u) = \frac{1 - G(\phi_n(u/m))}{u} = \lambda_n(u/m)\ [\ 1 - \frac{1}{m} A(\frac{u}{m}\lambda_n(u/m))\ ]$$

avec A défini dans le corollaire 17. Comme A est une fonction croissante positive de u dans [0,1], on obtient

$$\lambda_{n+1}(u) \leq \lambda_n(u/m)\ [\ 1 - \frac{1}{m} A(\frac{\beta u}{m})\ ] \leq \lambda_n(u/m)\ e^{-\frac{1}{m}A(\frac{\beta u}{m})}$$

Par itération,

$$\lambda_{n_0+k} \leq e^{\frac{-1}{m}\sum\limits_{r=1}^{k} A(\frac{\beta u}{m^r})} \lambda_{n_0}(u/m^k) \leq e^{\frac{-1}{m}\sum\limits_{r=1}^{k} A(\frac{\beta u}{m^r})}$$

car $\lambda_{n_0}(u/m^k) \leq \lim\limits_{x\to 0} \lambda_{n_0}(x) = E(W_{n_0}) = 1$.

Par le corollaire 17, si $E(X\log(X)) = \infty$ alors $\sum_{r=1}^{\infty} A(\frac{\beta u}{m^r}) = \infty$ d'où en faisant tendre k vers l'infini dans l'itération , on obtient $\lim\limits_{n\to\infty} \lambda_n(u) = 0$ pour $0 \leq u \leq \infty$ cela contredit $\lambda_n(u) > \beta > 0$ donc E(W)=0.

3.2 Population à lignée infinie

Nous continuons par étudier le nombre de personnes avec une lignée infinie de descendants. On note $Z_n^{(\infty)}$ les personnes de la nième génération dont la lignée survit indéfiniment.

Théorème 18. *Conditionné à la survie, le processus* $(Z_n^{(\infty)})_n$ *est encore un processus de branchement de distribution* $p^{(\infty)}$ *donnée par*

$$\begin{cases} p_0^{(\infty)} = 0 \\ p_k^{(\infty)} = \frac{1}{\zeta} \sum\limits_{j=k}^{\infty} \binom{k}{j} \eta^{j-k}(1-\eta)^k p_j \end{cases}$$

De plus, comme on a $\mu^{(\infty)} = E(Z_1^{(\infty)}) = \mu = E(Z_1)$*, le processus de branchement est sur-critique avec la même loi que* $(Z_n)_n$ *lui-même.*

Preuve

- On note A_∞ l'événement $Z_n \to +\infty$. On prouve par récurrence sur $n \geq 0$ que la distribution de $(Z_k^{(\infty)})_{k\in(0..n)}$ conditionnellement à A_∞ est égale à celle du processus de branchement $(\widehat{Z_k})_{k\in(0..n)}$ de loi de reproduction donnée par $p^{(\infty)}$.

Pour n=0, $Z_0^{(\infty)} = 1$ car un seul individu existe initialement et sa lignée ne meurt pas car on est sur A_∞. Or par convention $\widehat{Z_0} = 1$. L'initialisation est donc faite. On suppose désormais que la distribution $(Z_k^{(\infty)})_{k\in(0..n)}$ conditionné à A_∞ est celle de $(\widehat{Z_k})_{k\in(0..n)}$. Nous devons maintenant montrer que alors la distribution $(Z_k^{(\infty)})_{k\in(0..n+1)}$ conditionné à A_∞ est celle de $(\widehat{Z_k})_{k\in(0..n+1)}$. Par l'hypothèse de récurrence, on doit montrer que la loi de $Z_{n+1}^{(\infty)}$ sachant $(Z_k^{(\infty)})_{k\in(0..n)}$ est la même que $\widehat{Z_{n+1}}$ sachant $(\widehat{Z_k})_{k\in(0..n)}$.
La loi de $\widehat{Z_{n+1}}$ sachant $(\widehat{Z_k})_{k\in(0..n)}$ est une somme indépendante de $\widehat{Z_n}$ variables aléatoires de loi $p^{(\infty)}$. Maintenant la loi de $Z_{n+1}^{(\infty)}$ sachant $(Z_k^{(\infty)})_{k\in(0..n+1)}$ est égale à la loi de $Z_{n+1}^{(\infty)}$ sachant $Z_n^{(\infty)}$. Chaque individu de la génération n et de descendance infinie donne naissance à un nombre aléatoire i.i.d d'enfants de descendance infinie avec la même loi que $Z_1^{(\infty)}$ conditionné par A_∞.

- On doit maintenant montrer que $\mathbb{P}(Z_1^{(\infty)} = k|A_\infty) = p_k^\infty$
Pour k=0, conditionnellement à A_∞ on a $Z_1^\infty \geq 1$ donc on a bien l'égalité.
Pour $k \geq 1$, on va conditionner par Z_1 qui est $\geq Z_1^{(\infty)}$:
$\mathbb{P}(Z_1^{(\infty)} = k|A_\infty) = \zeta^{-1}\mathbb{P}(Z_1^{(\infty)} = k) = \zeta^{-1} \sum\limits_{j\geq k} \mathbb{P}(Z_1^{(\infty)} = k|Z_1 = j)\mathbb{P}(Z_1 = j)$
or pour déterminer $\mathbb{P}(Z_1^{(\infty)} = k|Z_1 = j)$ il suffit de choisir k enfants parmi les j qui auront une lignée infinie de probabilité $1-\eta$ et les (j-k) autres enfants

auront une descendance qui sera vouée à l'extinction de probabilité η donc

$$\mathbb{P}(Z_1^{(\infty)} = k|Z_1 = j) = \binom{k}{j}\eta^{j-k}(1-\eta)^k$$

d'où

$$\mathbb{P}(Z_1^{(\infty)} = k|A_\infty) = \zeta^{-1}\sum_{j\geq k}\binom{k}{j}\eta^{j-k}(1-\eta)^k p_j = p_k^\infty$$

- Enfin, on veut prouver l'égalité des moyennes.

$$\begin{aligned}
\mu^\infty &= \sum_{k=0}^{\infty} k p_k^\infty = \sum_{k=0}^{\infty} k\zeta^{-1}\sum_{j\geq k}\binom{k}{j}\eta^{j-k}(1-\eta)^k p_j \\
&= \zeta^{-1}\sum_{j=0}^{\infty} p_j \sum_{k=0}^{j} k\binom{k}{j}\eta^{j-k}(1-\eta)^k = \zeta^{-1}\sum_{j=0}^{\infty} p_j E(Bin(j,\zeta)) \\
&= \zeta^{-1}\sum_{j=0}^{\infty} p_j j\zeta = \sum_{j=0}^{\infty} p_j j = \mu
\end{aligned}$$

Théorème 19 (Proportion de particules avec une descendance infinie). *Conditionnellement à la survie,* $\frac{Z_n^{(\infty)}}{Z_n} \xrightarrow{p.s} \zeta$

Preuve

On va prouver le théorème avec $\mu = E(Z_1) < \infty$. On se place dans le cas où $E(Z_1\log(Z_1)) < \infty$. On a $Z_1 \geq Z_1^{(\infty)} \geq 1$ car on est conditionné à la survie d'où $Z_1\log(Z_1) \geq Z_1^{(\infty)}\log(Z_1^{(\infty)}) \geq 0$. On peut donc passer à l'espérance pour avoir $E(Z_1^{(\infty)}\log(Z_1^{(\infty)})) < \infty$. On a montré que $E(Z_1^{(\infty)}) = \mu$ qui nous permet d'obtenir l'existence de $W^{(\infty)}$ tel que $Z_n^{(\infty)}\mu^{-n} \longrightarrow W^{(\infty)}$. Comme $\mathbb{P}(W^{(\infty)} > 0) = 1$ car la probabilité d'extinction $\eta^{(\infty)}$ du processus de branchement $(Z_n^{(\infty)})_n$ est nulle et que $E(Z_1^{(\infty)}\log(Z_1^{(\infty)})) < \infty$ entraine $\mathbb{P}(W^{(\infty)} = 0) = \eta^{(\infty)}$. Donc $\frac{Z_n^{(\infty)}}{Z_n} = \frac{Z_n^{(\infty)}}{\mu^n}\frac{\mu^n}{Z_n}$ converge p.s vers une limite finie et positive R.

De plus, grâce à $(p_k^{(\infty)})_k$, on a la loi de $Z_n^{(\infty)}$ sachant Z_n qui est $Bin(Z_n,\zeta)$. Or si X$\sim$ $Bin(k,\zeta)$ alors par la loi des grands nombres $\frac{X}{k} \xrightarrow{proba} \zeta$. Par unicité de la limite $R = \zeta$

4 Théorème de temps de retour et de la progéniture totale

Théorème 20 (progéniture totale)**.** *Pour un processus de branchement avec une loi de reproduction i.i.d X, on a*
$\mathbb{P}(T = n) = \frac{1}{n}\mathbb{P}(X_1 + ... + X_n = n - 1)$ *où* $(X_i)_i \sim X$

On va prouver une version plus générale, qui dit que :

$$\mathbb{P}(T_1 + .. + T_k = n) = \frac{k}{n}\mathbb{P}(X_1 + ... + X_n = n - k)$$

où T_1,...,T_k sont k variables i.i.d de loi T. (On peut penser $T_1 + .. + T_k$ comme la progéniture totale d'un processus de branchement commençant avec k individus.) On note $\mathbb{P}_k$ la loi de la marche aléatoire commençant en k, $(Y_i)_i$ les pas i.i.d de la marche aléatoire et $S_n = k + Y_1 + ... + Y_n$ la position de la marche démarrant en k après n pas. On pose $T_0 = \inf\{n\colon S_n = 0\}$ le premier temps de retour à l'origine de la marche.

Théorème 21. *Pour une marche aléatoire de pas* $(Y_i)_{i\geq 1}$ *i.i.d satisfaisant* $\mathbb{P}(Y_i \geq -1) = 1$*, la distribution de* T_0 *est donnée par*

$$\mathbb{P}(T_0 = n) = \frac{k}{n}\mathbb{P}_k(S_n = 0)$$

.

Preuve
Nous allons prouver $\mathbb{P}(T_0 = n) = \frac{k}{n}\mathbb{P}_k(S_n = 0)$ $\forall k \geq 0$ par récurrence sur $n \geq 1$.
-Lorsque n=1, les deux cotés de l'égalité sont nuls lorsque $k \geq 2$ et $k = 0$ et sont égales à $\mathbb{P}(Y_1 = -1)$ lorsque k=1.
-Lorsque $n \geq 2$, on remarque que les 2 cotés sont nuls pour k=0. On prend maintenant $k \geq 1$.
On a $\mathbb{P}_k(T_0 = n) = \sum\limits_{s=-1}^{\infty} \mathbb{P}_k(T_0 = n|Y_1 = s)\mathbb{P}(Y_1 = s)$.
Par la propriété de Markov,

$$\mathbb{P}_k(T_0 = n|Y_1 = s) = \mathbb{P}_{k+s}(T_0 = n - 1) = \frac{k+s}{n-1}\mathbb{P}_{k+s}(S_{n-1} = 0)$$

par hypothèse de récurrence avec $k \geq 1$ et $s \geq -1$ donc

$$\mathbb{P}_k(T_0 = n) = \sum_{s=-1}^{\infty} \frac{k+s}{n-1} \mathbb{P}_{k+s}(S_{n-1} = 0)\mathbb{P}(Y_1 = s)$$

or $\mathbb{P}_{k+s}(S_{n-1} = 0) = \mathbb{P}_k(S_n = 0|Y_1 = s)$

$$\begin{aligned}
\mathbb{P}_k(T_0 = n) &= \sum_{s=-1}^{\infty} \frac{k+s}{n-1} \mathbb{P}_k(S_n = 0|Y_1 = s)\mathbb{P}(Y_1 = s) \\
&= \frac{1}{n-1} \sum_{s=-1}^{\infty} (k+s)\mathbb{P}_k(Y_1 = s|S_n = 0)\mathbb{P}_k(S_n = 0) \\
&= \frac{\mathbb{P}_k(S_n = 0)}{n-1} \left(\sum_{s=-1}^{\infty} k\mathbb{P}_k(Y_1 = s|S_n = 0) + \sum_{s=-1}^{\infty} s\mathbb{P}_k(Y_1 = s|S_n = 0) \right) \\
&= \frac{\mathbb{P}_k(S_n = 0)}{n-1} \big(k + \mathbb{E}_k(Y_1|S_n = 0)\big)
\end{aligned}$$

On remarque que $\mathbb{E}_k(Y_i|S_n = 0)$ est indépendant de i donc

$$\mathbb{E}_k(Y_1|S_n = 0) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_k(Y_i|S_n = 0) = \frac{1}{n} \mathbb{E}_k(\sum_{i=1}^{n} Y_i|S_n = 0)$$

De plus, $S_n = k + \sum Y_i$ donc $\mathbb{E}_k(Y_1|S_n = 0) = \frac{-k}{n}$ donc

$$\mathbb{P}_k(T_0 = n) = \frac{1}{n-1}\mathbb{P}_k(S_n = 0)(k - \frac{k}{n}) = \frac{k}{n}\mathbb{P}_k(S_n = 0)$$

5 Processus de Poisson

Dans cette partie, les processus de branchement de Poisson sont étudiés que l'on notera par $\mathbb{P}^*_\lambda$. X^* est une variable de Poisson et T^* est la progéniture totale. Pour une variable de Poisson X^* de moyenne λ, la fonction génératrice de probabilité est égale à : $G^*_\lambda(s) = E^*_\lambda(s^{X*}) = \sum_{i=0}^{\infty} s^i e^{-\lambda}\frac{\lambda^i}{i!} = e^{\lambda(s-1)}$ donc la probabilité d'extinction vérifie : $\eta_\lambda = e^{\lambda(\eta_\lambda - 1)}$.

Par des simulations avec le logiciel R, le graphique ci-dessous a été construit à partir de simulations de générations de population suivant une loi de reproduction $P(\lambda)$. Comme l'arbre s'éteint au bout de peu de générations lorsqu'il y a extinction, la probabilité d'extinction a été calculée en regardant si Z_{15} était nulle pour 100 processus de Poisson de même loi.

Puis en faisant varier le paramètre λ, il est possible de visualiser la probabilité d'extinction en fonction du paramètre et de mettre en avant le changement de comportement par rapport à la valeur $\lambda = 1$. La courbe théorique où η_λ est solution de l'équation précédente est également représentée.

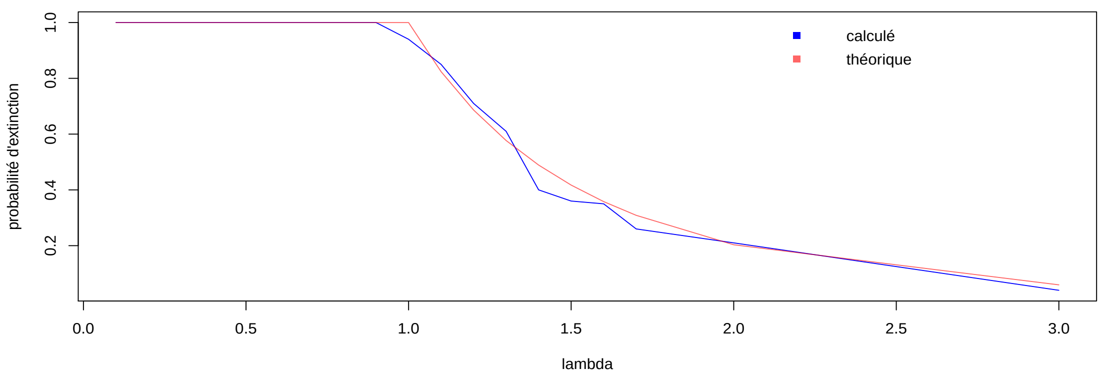

FIGURE 5 – probabilité d'extinction du processus de Poisson

On rappelle que H=$(X^*_1, ..., X^*_T)$ est l'histoire du processus de branchement. Alors conditionnellement à l'extinction, un processus de branchement a pour loi p' donné par $p'_i = \eta_\lambda^{i-1} p_i = \eta_\lambda^{i-1} e^{-\lambda}\frac{\lambda^i}{i!} = \frac{1}{\eta_\lambda} e^{-\lambda}\frac{(\lambda\eta_\lambda)^i}{i!} = e^{\lambda(1-\eta_\lambda)} e^{-\lambda}\frac{(\lambda\eta_\lambda)^i}{i!}$
D'où $p'_i = e^{-\lambda\eta_\lambda}\frac{(\lambda\eta_\lambda)^i}{i!}$.

Cette distribution est encore une loi de Poisson de moyenne $\mu_\lambda = \lambda\eta_\lambda$ et on a $\mu_\lambda e^{-\mu_\lambda} = \lambda\eta_\lambda e^{-\lambda\eta_\lambda} = \lambda e^{\lambda(\eta_\lambda - 1)} e^{-\lambda\eta_\lambda} = \lambda e^{-\lambda}$

Définition 22. *On dit que $\mu < 1 < \lambda$ est une paire conjuguée si $\mu e^{-\mu} = \lambda e^{-\lambda}$*

Théorème 23 (Principe de dualité de Poisson)**.** *Soit $\mu < 1 < \lambda$ conjugués. Le processus de branchement de Poisson de moyenne λ conditionné à l'extinction a la même loi qu'un processus de Poisson de moyenne μ.*

Théorème 24. *Pour un processus de branchement de loi de reproduction i.i.d X avec X une loi de Poisson de moyenne λ, on a :*

$$\mathbb{P}^*_\lambda(T^* = n) = \tfrac{(\lambda n)^{n-1}}{n!} e^{-\lambda n}$$

Preuve

On utilise le théorème de la progéniture qui nous donne pour $X_i^* \sim P(\lambda)$

$$\mathbb{P}^*_\lambda(T^* = n) = \frac{1}{n}\mathbb{P}^*_\lambda(X_1^* + ... + X_n^* = n - 1)$$

On a $\sum X_i^* \sim P(n\lambda) \sim Y$ avec $\mathbb{P}(Y = n-1) = e^{-n\lambda}\frac{(\lambda n)^{n-1}}{(n-1)!}$. On en déduit $\mathbb{P}^*_\lambda(T^* = n) = \frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n}$. $\square$

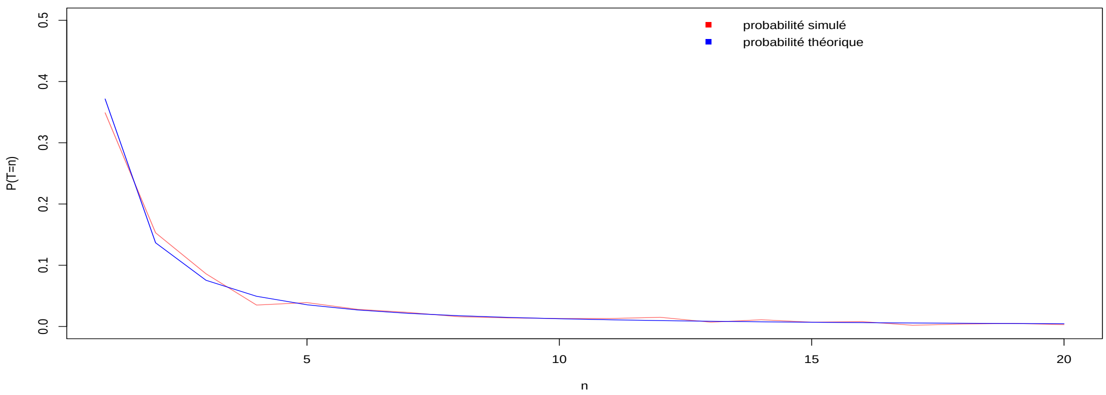

FIGURE 6 – Probabilité que la progéniture totale vaille n

Sur le graphique, deux courbes sont tracées. Elles correspondent à $\mathbb{P}(T^* = n)$ en fonction de n. Dans le cas de la courbe calculée, de nombreuses populations finies p.s. ont été générées par un cas sous-critique. Il a fallu ensuite calculé la probabilité que la progéniture vaille une valeur n à partir d'un échantillon de taille 1000. Alors que la courbe théorique représente $\frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n}$ en fonction de n. On constate que la courbe calculée est très proche de la théorique.

Définition 25.
- Un arbre étiqueté sur n est un arbre de taille n où chacun des sommets a un nom dans 1,..,n et chaque nom apparait une seule fois.
- Une arrête d'un arbre étiqueté est une paire v1,v2 où v1 et v2 sont les noms de 2 sommets reliés dans l'arbre.
- L'ensemble d' arrêtes d'un arbre de taille n est la collection de ses (n-1) arrêtes.
- Deux arbres étiquetés sont égaux si et seulement si ils ont le même ensemble d' arrêtes.

Théorème 26 (Théorème de Cayley)**.** *Le nombre d'arbre étiqueté de taille n est égale à* n^{n-2}

Preuve

Les sommets vont être appelés en terme de mots. La racine est le mot $\emptyset$. Les enfants de la racine sont les mots 1, 2, ..., $d_\emptyset$ où pour un mot w, on note d_w le nombre de ses enfants. Les enfants de 1 sont 11,12,... $1d_1$. Un arbre est ainsi représenté de manière unique par l'ensemble de ses mots. Par exemple, le mot 123 représente le 3ème enfant du second enfant du premier enfant de la racine.
Deux arbres sont les mêmes si et seulement si ils sont représentés par le même ensemble de mots. Nous obtenons un processus de branchement lorsque les variables $(d_w)_w$ sont égales à une collection i.i.d de variables aléatoires. Pour un mot w, nous notons $|w|$ sa longueur. La longueur d'un mot w est le nombre de pas auquel le mot est relié à la racine et est égale à sa génération.
On note $\mathcal{T}$ l'arbre d'un processus de branchement de Poisson de moyenne 1 et on calcule alors la probabilité d'avoir un arbre donné t :

$$\mathbb{P}(\mathcal{T} = t) = \prod_{w \in t} \mathbb{P}(\xi = d_w) \quad \text{avec} \quad \xi \sim P(1)$$

et d_w le nombre d'enfants du mot w dans l'arbre t. Comme $\mathbb{P}(\xi = d_w) = \frac{e^{-1}}{d_w!}$, on obtient $\mathbb{P}(T = t) = \prod\limits_{w \in t} \frac{e^{-1}}{d_w!} = \frac{e^{-n}}{\prod\limits_{w \in t} d_w!}$ car n est le nombre de sommets de t.

Conditionnellement à avoir une progéniture totale $T^* = n$, nous introduisons l'étiquetage suivant : La racine porte l'étiquette 1 et les autres sommets ont une étiquette entre 2 et n. Si $\mathcal{T}$ est donné alors, il y a précisément $\prod\limits_{w \in t} d_w!$ manières possibles de donner des noms à l'arbre qui donne naissance au même arbre labellisé (on permute les noms des frères et sœurs pour les points extrémaux sinon on permute les branches des points de même génération). De plus la probabilité pour que l'arbre soit entièrement étiqueté selon un de ces étiquetages décrits est $\frac{1}{(n-1)!}$.

Pour un arbre étiqueté ℓ, on choisit t_ℓ un arbre quelconque par lequel on peut obtenir l'arbre labellisé ℓ en par étiquetage des sommets. Alors la probabilité d'obtenir un arbre étiqueté donné ℓ est :

$$\mathbb{P}(\mathcal{L} = \ell) = \mathbb{P}(\mathcal{T} = t_\ell) \frac{\prod\limits_{w \in t} d_w!}{(|l| - 1)!}$$

En effet, on choisit un arbre quelconque de taille $|l|$, puis on regarde la probabilité d'avoir un arbre labellisé ℓ précisément. On divise donc par $(|l| - 1)!$ qui est l'ensemble des arbres labellisés possibles et on multiplie par

$\prod_{w \in t_\ell} d_w!$ qui est l'ensemble des arbres équivalents à ℓ. D'où

$$\mathbb{P}(\mathcal{L} = \ell) = \frac{e^{-|l|}}{\prod_{w \in t_\ell} d_w!} \frac{\prod_{w \in t_\ell} d_w!}{(|l|-1)!} = \frac{e^{-|l|}}{(|l|-1)!}$$

De plus, conditionnellement à $T^* = n$, la probabilité d'un arbre labellisé de taille n donné est :

$$\mathbb{P}(\mathcal{L} = \ell |\ |\mathcal{L}| = n) = \frac{\mathbb{P}(\mathcal{L} = \ell)}{\mathbb{P}(|\mathcal{L}| = n)} = \frac{\mathbb{P}(\mathcal{L} = \ell)}{\mathbb{P}(T^* = n)}$$

or $\mathbb{P}(T^* = n) = \frac{e^{-n} n^{n-2}}{(n-1)!}$ d'après le théorème précédent. On en déduit donc

$$\mathbb{P}(\mathcal{L} = \ell |\ |\mathcal{L}| = n) = \frac{e^{-n}}{(n-1)!} \frac{(n-1)!}{e^{-n} n^{n-2}} = \frac{1}{n^{n-2}}.$$

On obtient une probabilité uniforme pour l'ensemble des arbres labellisés. Donc le nombre d'arbres labellisés est égale à n^{n-2}.

Théorème 27.
*Pour un processus de branchement de loi de reproduction i.i.d $X \sim P(\lambda)$, on a lorsque $n \longrightarrow \infty$ $\mathbb{P}^*_\lambda(T^* = n) = \frac{1}{\lambda\sqrt{2\pi n^3}} e^{-I_\lambda n}(1 + o(1))$
où $I_\lambda = \lambda - 1 - ln(\lambda)$*

Corollaire 28 (Différentiabilité de la probabilité de survie)**.** *Soit η_λ la probabilité d'extinction d'un processus de branchement de Poisson de moyenne $\lambda > 1$ alors on a $\frac{d}{d\lambda}\zeta_\lambda = \frac{\eta_\lambda(\lambda-\mu_\lambda)}{\lambda(1-\mu_\lambda)}$ où μ_λ est le dual de λ ($\mu_\lambda = \eta_\lambda\lambda$)*

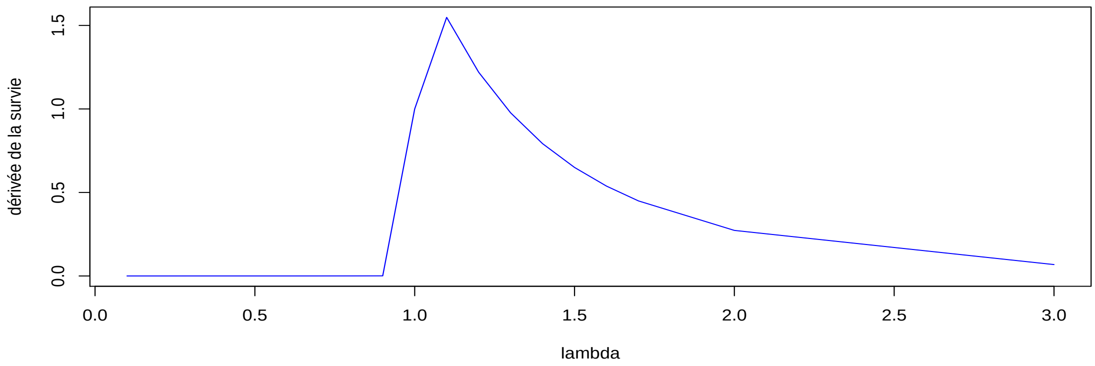

FIGURE 7 – Dérivée de la survie en fonction de λ

Remarque : La courbe a été lissée par le logiciel R. En réalité, la dérivée est nulle pour $\lambda < 1$ et elle admet une valeur infinie par limite à droite en 1.

Preuve

La fonction η_λ qui est notée $\eta(\lambda)$ est décroissante et satisfait

$$\eta(\lambda) = \mathbb{P}^*_\lambda(T^* < \infty) = \sum_{n=1}^{\infty} \mathbb{P}^*_\lambda(T^* = n) = \sum_{n=1}^{\infty} e^{-\lambda n} \frac{(\lambda n)^{n-1}}{n!}$$

Puis,

$$\begin{aligned} 0 \leq \frac{d}{d\lambda}\zeta(\lambda) = \frac{-d}{d\lambda}\eta(\lambda) &= -\sum_{n=1}^{\infty}(e^{-\lambda n})' \frac{(\lambda n)^{n-1}}{n!} - \sum_{n=1}^{\infty} e^{-\lambda n}(\frac{(\lambda n)^{n-1}}{n!})' \\ &= \sum_{n=1}^{\infty} n e^{-\lambda n} \frac{(\lambda n)^{n-1}}{n!} - \sum_{n=2}^{\infty} e^{-\lambda n} \frac{n^{n-1}(n-1)\lambda^{n-2}}{n!} \\ &= \sum_{n=1}^{\infty} e^{-\lambda n} \frac{(\lambda n)^{n-1}}{(n-1)!} - \sum_{n=2}^{\infty} e^{-\lambda n} \frac{n^{n-2}\lambda^{n-2}}{(n-2)!} \end{aligned}$$

De plus, $\mathbb{E}^*_\lambda(T^*|T^* < \infty) = \frac{1}{\mathbb{P}^*_\lambda(T^*<\infty)} \sum_{n=1}^{\infty} n e^{-\lambda n} \frac{(\lambda n)^{n-1}}{n!} = \frac{1}{\eta(\lambda)} \sum_{n=1}^{\infty} e^{-\lambda n} \frac{(\lambda n)^{n-1}}{(n-1)!}$

D'où, $\frac{d}{d\lambda}\zeta(\lambda) = \eta(\lambda)\mathbb{E}^*_\lambda(T^*|T^* < \infty) - \sum_{n=2}^{\infty} e^{-\lambda n} \frac{n^{n-2}\lambda^{n-2}}{(n-2)!}$

or

$$\begin{aligned}
\sum_{n=2}^{\infty} e^{-\lambda n} \frac{(\lambda n)^{n-2}}{(n-2)!} &= \sum_{n=1}^{\infty} (n-1) e^{-\lambda n} \frac{(\lambda n)^{n-2}}{(n-1)!} \\
&= \sum_{n=1}^{\infty} \frac{1}{\lambda} e^{-\lambda n} \frac{(\lambda n)^{n-1}}{(n-1)!} - \sum_{n=1}^{\infty} e^{-\lambda n} \frac{(\lambda n)^{n-2}}{(n-1)!} \\
&= \frac{\eta(\lambda)}{\lambda} \mathbb{E}^*_\lambda(T^* | T^* < \infty) - \frac{1}{\lambda} \sum_{n=1}^{\infty} e^{-\lambda n} \frac{(\lambda n)^{n-1}}{n!} \\
&= \frac{\eta(\lambda)}{\lambda} \mathbb{E}^*_\lambda(T^* | T^* < \infty) - \frac{1}{\lambda} \mathbb{P}^*_\lambda(T^* < \infty) \\
&= \frac{\eta(\lambda)}{\lambda} \mathbb{E}^*_\lambda(T^* | T^* < \infty) - \frac{\eta(\lambda)}{\lambda}
\end{aligned}$$

donc $\frac{d}{d\lambda}\zeta(\lambda) = \eta(\lambda)\mathbb{E}^*_\lambda(T^*|T^* < \infty) - \frac{\eta(\lambda)}{\lambda}\mathbb{E}^*_\lambda(T^*|T^* < \infty) + \frac{\eta(\lambda)}{\lambda}$. De plus par le principe de dualité, le processus de branchement de moyenne λ conditionné à l'extinction a la même distribution que le processus de branchement de moyenne μ_λ donc $\mathbb{E}^*_\lambda(T^*|T^* < \infty) = \frac{1}{1-\mu_\lambda}$ donc

$$\frac{d}{d\lambda}\zeta(\lambda) = \frac{\eta(\lambda)}{1-\mu_\lambda} + \frac{\eta(\lambda)}{\lambda} = \frac{\eta_\lambda(\lambda - \mu_\lambda}{\lambda(1-\mu_\lambda)}$$

6 Modèles plus avancés

6.1 Processus de branchement en temps continu

Un des défauts du processus de branchement tel qu'il est présenté jusqu'à présent est son caractère discret. Il existe un modèle plus général qui permet de prendre en compte le fait que deux individus d'une même génération n'engendrent pas leurs familles simultanément. Pour modéliser ceci, on introduit une nouvelle variable aléatoire, l'âge. On suppose que la famille des âges est mutuellement indépendante identiquement distribuée de densité f_A et indépendante des tailles de famille. De plus l'âge est une variable continue et positive.

On note $Z(t)$ le nombre d'individus à l'instant t, et $G_t(s) = \mathbb{E}(s^{Z(t)})$ la fonction génératrice de $Z(t)$. La fonction génératrice dépend du temps. On suppose de plus que $Z(0) = 1$: il y a un unique individu initial. Chaque individu vit pendant une période, égale à son âge, avant d'être remplacé par la famille qu'il engendre. On note A l'âge de l'individu initial. Tout comme dans le modèle discret, on cherche à obtenir une expression de G_t afin de déterminer $Z(t)$.

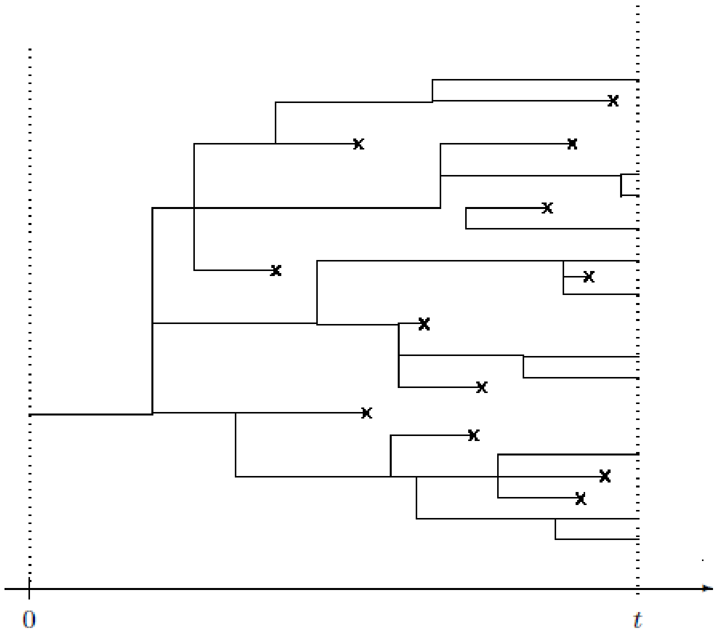

FIGURE 8 – Exemple d'arbre suivant un processus de branchement en temps continu

Théorème 29. *Pour tout t dans $\mathbb{R}^+$, pour tout s dans $[0,1]$, on a :*

$$G_t(s) = \int_0^t G(G_{t-u}(s)) f_A(u)\,du + \int_t^\infty s f_A(u)\,du$$

Preuve
On a l'égalité suivante :

$$G_t(s) = \mathbb{E}[s^{Z(t)}] = \mathbb{E}[\mathbb{E}[s^{Z(t)}|A]] = \int_{\mathbb{R}_+} \mathbb{E}[s^{Z(t)}|A = u] f_A(u)\mathrm{d}u.$$

Si $A = u$, à l'instant u l'individu initial meurt et est remplacé par N descendants. N est une variable aléatoire de fonction génératrice G. Chacun des N individus engendre une famille comme son ancêtre l'a fait. La mort de l'ancêtre a pour effet de remplacer le processus par une somme de N copies du processus décalées dans le temps de u.

- Si $u > t$, alors $Z(t) = 1$ et donc pour tout s dans $[0,1], \mathbb{E}[s^{Z(t)}|A = u] = s$.
- Si $u < t$, alors $Z(t) = Y_1 + Y_2 + ... + Y_N$ une somme de N copies indépendantes de la génération $Z(t-u)$. On obtient $\mathbb{E}[s^{Z(t)}|A = u] = G(G_{t-u}(s))$.

Donc en séparant l'intégrale au temps t, on obtient l'égalité du théorème.

Théorème 30. *Soit t dans $\mathbb{R}_+$, on note $m(t) = \mathbb{E}[Z_t]$. Sous réserve d'existence, m vérifie l'équation suivante :*

$$\forall t \in \mathbb{R}_+, \quad m(t) = \mu \int_0^t m(t-u) f_A(u)\, du + \int_0^\infty f_A(u)\, du$$

avec $\mu = G'(1) = E(Z_1)$.

Preuve

on a :

$$\forall t \in \mathbb{R}_+, \quad \mathbb{E}[Z_t] = \lim_{s \nearrow 1} \frac{\partial G_t}{\partial s}(s).$$

Par théorème de dérivation sous le signe somme,

$$\mathbb{E}[Z_t] = \lim_{s \nearrow 1} \int_0^t \frac{\partial G}{\partial s}(G_{t-u}(s)) \frac{\partial G_{t-u}}{\partial s}(s) f_A(u) \mathrm{d}u + \int_t^\infty f_A(u) \mathrm{d}u.$$

On conclut avec le théorème de convergence dominée car

$$\lim_{s \nearrow 1} \frac{\partial G}{\partial s}(G_{t-u}(s)) = \lim_{s \nearrow 1} \frac{\partial G}{\partial s}(s) = \mu \quad \text{et} \lim_{s \nearrow 1} \frac{\partial G_{t-u}}{\partial s}(s) = m(t-u).$$

□

6.2 Distribution exponentielle de l'âge

Dans le cas particulier où f_A est la densité d'une loi exponentielle, il est possible de déterminer une équation différentielle plus simple, vérifiée par G_t.

Théorème 31 (Distribution exponentielle de l'âge)**.** *On suppose que la densité de l'âge est :* $f_A : t \mapsto \lambda e^{-\lambda t}\mathbb{1}_{\mathbb{R}_+}(t)$.
Alors la fonction G_t *vérifie l'équation différentielle suivante :*

$$\frac{\partial}{\partial t}G_t(s) = \lambda[G(G_t(s)) - G_t(s)].$$

Preuve
On part de l'équation du théorème 29 et on dérive par rapport à t. Pour tout $t \in \mathbb{R}_+$, pour tout $s \in [0,1]$, on a :

$$\frac{\partial}{\partial t}G_t(s) = G(G_0(s))f_A(t) + \int_0^t \frac{\partial}{\partial t}G(G_{t-u}(s))f_A(u)\mathrm{d}u - sf_A(t)$$

Or $G_0(s) = s$ et $\frac{\partial}{\partial t}G(G_{t-u}(s)) = -\frac{\partial}{\partial u}G(G_{t-u}(s))$. Donc on a :

$$\frac{\partial}{\partial t}G_t(s) = (G(s) - s)f_A(t) - \int_0^t \frac{\partial}{\partial u}G(G_{t-u}(s))f_A(u)\mathrm{d}u$$
$$\frac{\partial}{\partial t}G_t(s) = (G(s) - s)f_A(t) - [G(G_{t-u}(s))f_A(u)]_{u=0}^t + \int_0^t G(G_{t-u}(s))\frac{\partial}{\partial u}f_A(u)\mathrm{d}u$$

En utilisant l'expression de f_A, on peut simplifier :

$$\begin{aligned}\frac{\partial}{\partial t}G_t(s) &= (G(s) - s)f_A(t) - G(G_0(s))f_A(t) + G(G_t(s))f_A(0) - \lambda\int_0^t G(G_{t-u}(s))f_A(u)\mathrm{d}u \\ &= -sf_A(t) + \lambda G(G_t(s)) - \lambda\int_0^t G(G_{t-u}(s))f_A(u)\mathrm{d}u\end{aligned}$$

D'après le théorème 29, on a :

$$\int_0^t G(G_{t-u}(s))f_A(u)\mathrm{d}u = G_t(s) - s\int_t^\infty f_A(u)\mathrm{d}u = G_t(s) - s\frac{f_A(t)}{\lambda}$$

Pour finir, on obtient :

$$\frac{\partial}{\partial t}G_t(s) = \lambda[G(G_t(s)) - G_t(s)].$$

Plusieurs simulations de processus de branchement à temps continu ont été réalisées :
-Tout d'abord, le processus créé est celui où la mort du parent entraine la naissance de ces enfants. Une matrice permet de rassembler les dates de naissance et de mort des individus au cours du temps.
La date de naissance des individus correspond à la date de la mort de son parent et la date de décès correspond à la date de naissance à laquelle on a rajouté une variable aléatoire suivant une loi exponentielle de moyenne 70. On peut alors trouver la population en fonction du temps t en regardant quels individus ont leur date de naissance $< t <$ date de mort.

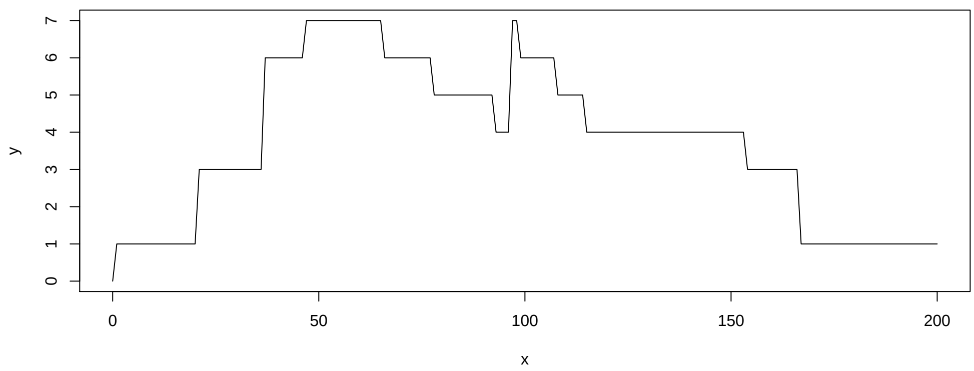

Figure 9 – Population avec modèle de non-coexistence des parents et enfants

Ce modèle n'est pas forcément très réaliste car la loi exponentielle est une loi sans vieillissement donc certains individus vivent 200 ans avec ce modèle. De plus, tous les enfants naissent en même temps et seulement à la mort de leur parent.

-Ensuite, j'ai voulu créer un modèle où les enfants naissent au fur et à mesure de la vie de leur parent et donc plusieurs générations peuvent coexister. Pour cela, j'ai changé l'expression de la date de naissance en prenant la date de naissance du parent auquel on ajoute une loi exponentielle de moyenne 30. Si le résultat est plus grand que la date de mort du parent alors on donne pour date de naissance à l'enfant la date de mort de son parent comme précédemment.

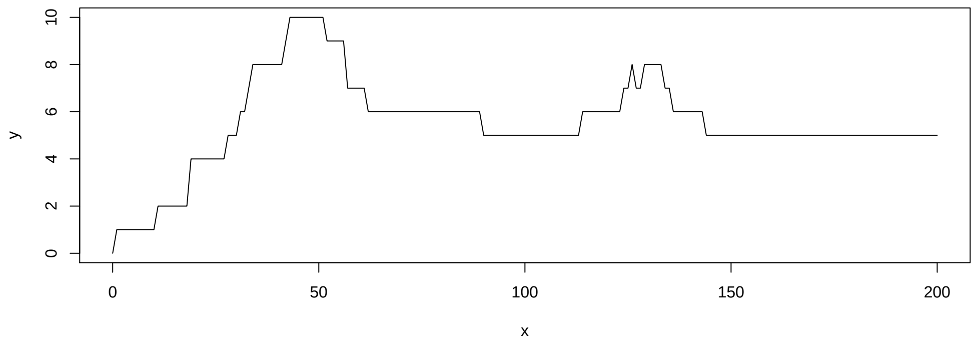

FIGURE 10 – Population avec modèle de coexistence des parents et enfants

Cependant ces deux modèles doivent être programmés au fur et à mesure des générations. Le processus n'a pas pu être généralisé.

-Enfin, la population est modélisée en utilisant les propriétés de loi exponentielle. En effet, on s'intéresse à chercher parmi les individus vivants lequel va mourir en premier pour donner vie à ses enfants. Si on a p individus à l'instant t alors on cherche $\min(X_1, ..., X_p)$ où les X_i sont des lois exponentielles de paramètre λ et correspondent à l'âge des individus. Or $\min(X_1, ..., X_p)$ suit une loi exponentielle de paramètre $p\lambda$.
On crée donc une matrice où on retient dans une ligne les temps où la population change. Dans l'autre ligne, on indique le nombre d'individus à partir de ce temps qui est le nombre d'individus à l'instant précédent auquel on enlève la personne décédée et on rajoute ses enfants.

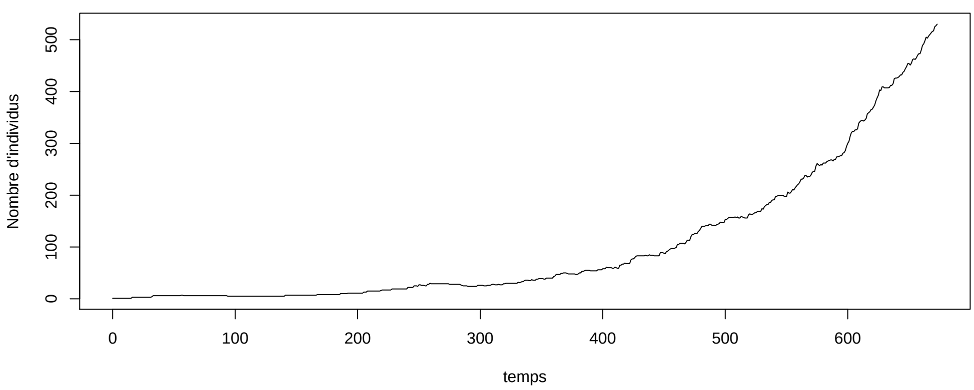

FIGURE 11 – Population avec modèle du minimum

6.3 Modèle binaire

On reprend le modèle binaire vu dans le cas discret. On est dans le cas de la division cellulaire. A l'instant t=0, il n'y a qu'une seule cellule ($Z(0) = 1$). On suppose que l'âge suit une loi $\mathcal{E}(1)$. Chaque cellule peut soit mourir, soit se dupliquer, de manière équiprobable. On a :

$$\forall s \in [0,1], G(s) = \frac{1}{2}(1+s^2).$$

L'équation différentielle définissant G_t s'écrit donc :

$$\frac{\partial}{\partial t}G_t(s) = \frac{1}{2}(G_t(s)-1)^2.$$

Théorème 32 (Expression de G_t)**.**

$$\forall t \in \mathbb{R}_+, \quad \forall s \in [0,1], G_t(s) = \frac{2s + t(1-s)}{2+t(1-s)}$$

Preuve
La fonction $t \mapsto \frac{1}{2}(t-1)^2$ est localement lipschitzienne sur $\mathbb{R}_+$, donc d'après le théorème de Cauchy-Lipschitz, il existe une unique solution maximale. La fonction constante égale à 1 est solution de l'équation, donc si une solution passe par la valeur 1, c'est la fonction constante. On considère donc φ une solution, et on suppose qu'elle ne prend pas la valeur 1. On peut alors écrire :

$$\frac{\varphi'}{(\varphi(t)-1)^2} = \frac{1}{2}$$

$$\frac{1}{\varphi(t)-1} - \frac{1}{\varphi(0)-1} = -\frac{t}{2}$$

Or $\forall s \in [0,1]$, $G_0(s) = \int_0^\infty s\lambda e^{-\lambda t} du = s$

$$\frac{1}{\varphi(t)-1} = \frac{1}{s-1} - \frac{t}{2}$$

$$\frac{1}{\varphi(t)-1} = \frac{2+t(1-s)}{2(s-1)}$$

$$\varphi(t) = \frac{2s+t(1-s)}{2+t(1-s)}$$

Par unicité, on a le résultat. $\square$

A partir de la fonction génératrice, on peut remonter aux lois des $Z(t)$ car $G_t(s) = E(s^{Z(t)}) = \sum_{n=0}^{\infty} \mathbb{P}(Z(t) = n)s^n$. Donc $G_t^{(n)}(0) = n!\ \mathbb{P}(Z(t) = n)$
On en déduit les résultats suivants sur le processus :

Théorème 33. $\forall t \in \mathbb{R}_+, \quad \forall n \in \mathbb{N}$ *:*

$$\mathbb{P}(Z(t) = n) = \begin{cases} \frac{t}{2+t} & si \quad n = 0 \\ \\ \frac{4t^{n-1}}{(2+t)^{n+1}} & si \quad n \geq 1 \end{cases}$$

Preuve
Le calcul des dérivées successives de G se fait assez facilement.
On montre que $G_t'(s) = \frac{4}{[2+t(1-s)]^2}$. Alors on a par récurrence,

$$G_t^{(n)}(s) = \frac{4t^{n-1}n!}{[2 + t(1 - s)]^{n+1}}$$

On évalue en 0.

Théorème 34 (Espérance et variance de Z(t))**.**

$$\forall t \in \mathbb{R}_+, \quad \mathbb{E}[Z(t)] = 1 \quad et \quad var(Z(t)) = t$$

Preuve
Pour tout t dans $\mathbb{R}_+, s \mapsto G_t(s)$ est de classe $\mathcal{C}^\infty$ sur $[0, 1]$ et on a :

$$\frac{\partial}{\partial s} G_t(s) = \frac{4}{[2 + t(1 - s)]^2}.$$

Elle admet donc en 1 une dérivée à gauche finie qui vaut $\mathbb{E}[Z(t)] = 1$. De même

$$\frac{\partial^2}{\partial s^2} G_t(s) = \frac{8t}{[2 + t(1 - s)]^3}.$$

Ainsi $\mathrm{var}(Z(t)) = t$.

Théorème 35 (Fonction de répartition de $Z(t)$)**.**

$$\forall t \in \mathbb{R}_+, \quad \forall k \in \mathbb{N}^* \quad \textit{on a :} \quad \mathbb{P}(Z(t) \geq k) = \frac{2t^{k-1}}{(2+t)^k}$$

Preuve
Soit k dans $\mathbb{N}^*$, on a :

$$\mathbb{P}(Z(t) \geq k) = \sum_{n=k}^{\infty} \frac{4t^{n-1}}{(2+t)^{n+1}} = \frac{4}{(2+t)^2} \sum_{n=k-1}^{\infty} \frac{t^n}{(2+t)^n}$$
$$\mathbb{P}(Z(t) \geq k) = \frac{4}{(2+t)^2} \frac{(\frac{t}{2+t})^{k-1}}{1-\frac{t}{2+t}} = \frac{2t^{k+1}}{(2+t)^k}$$

□

On n'a pas calculé la fonction de répartition F à proprement parler mais $1-F$. C'est parfois plus facile à manier.

On peut maintenant s'intéresser aux comportements asymptotiques de notre modèle.

Théorème 36 (Comportement asymptotique)**.**
La population s'éteint presque sûrement.
De plus, conditionnellement à l'événement "$\{Z(t) > 0\}$", $\frac{Z(t)}{t}$ converge en loi vers $\mathcal{E}(2)$.

Preuve

On montre que $\lim\limits_{t\to\infty} G_t(0) = 1$, $\{\text{extinction}\} = \bigcup_{t\in\mathbb{R}_+}\{Z(t) = 0\}$.
Par inclusion, $\mathbb{P}(\{\text{extinction}\}) \geq \mathbb{P}(\bigcup_{n\in\mathbb{N}}\{Z(n) = 0\}$. On peut calculer cette dernière quantité qui vaut 1. Une probabilité étant toujours inférieure à 1,

on a : $\mathbb{P}(\{\text{extinction}\}) = 1$. Pour tout $x \in \mathbb{R}_+^*$, on a :

$$\mathbb{P}\left(\frac{Z(t)}{t} \geq x \,|\, Z(t) > 0\right) = \mathbb{P}(Z(t) \geq tx | Z(t) > 0) = \frac{\mathbb{P}(Z(t) \geq tx \quad \text{et} \quad Z(t) > 0)}{\mathbb{P}(Z(t) > 0)}$$

$$= \frac{\mathbb{P}(Z(t) \geq tx)}{\mathbb{P}(Z(t) > 0)} = \frac{2t^{\lfloor tx \rfloor}}{(2+t)^{\lfloor tx \rfloor + 1}} \frac{2+t}{2} = \left(\frac{t}{2(2+t)}\right)^{\lfloor tx \rfloor}$$

$$= e^{-\lfloor tx \rfloor \ln 2(1+\frac{2}{t})} = e^{-\lfloor tx \rfloor \ln 2} e^{-\lfloor tx \rfloor \ln (1+\frac{2}{t})}$$

Ainsi $\quad \mathbb{P}\left(\frac{Z(t)}{t} \geq x | Z(t) > 0\right) \underset{n \to +\infty}{\longrightarrow} = e^{-2x}$

En notant F la fonction de répartition de $\frac{Z(t)}{t}$ conditionnellement à $\{Z(t) > 0\}$, on a montré que $F(x) = 1 - e^{-2x}$ si $x > 0$. On montre facilement que F est nulle si $x \leq 0$. C'est donc la fonction de répartition d'une loi $\mathcal{E}(2)$. $\square$

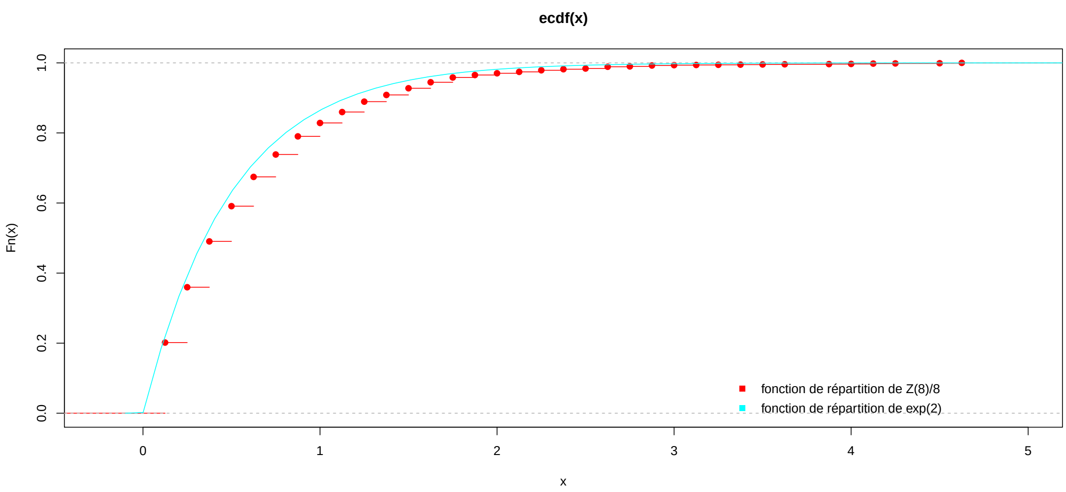

FIGURE 12 – Convergence vers la loi $\mathcal{E}(2)$

6.4 Modèle de reproduction sexuée

Le modèle de Daley sert à décrire le mécanisme de reproduction d'une population sexuée. On définit F_n le nombre de femelles et M_n le nombre de mâles à la génération n et Z_n le nombre de couples de cette même génération. On définit, ensuite, le nombre de naissances donné par un couple (tous les couples suivent la même loi de reproduction G) par : X le nombre de femelles et Y le nombre de mâles. Les variables X et Y sont indépendantes de l'indice de génération.

Définition 37. *Soit une suite* $(X_{n,i}, Y_{n,i})_{(n,i)\in\mathbb{N}\times\mathbb{N}^*}$ *de variables aléatoires indépendantes et identiquement distribuées selon le loi du couple* (X, Y)*, à valeurs entières.*
Soit $L : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ *croissante en* x *et en* y*, à valeurs entières, telle que :*

$$\forall (x, y) \in (\mathbb{R}_+)^2,\ L(x, y) \leq xy.$$

Soit $(Z_n)_{n\in\mathbb{N}}$ *une suite telle que :*

$$\begin{cases} Z_0 = N \quad (N \geq 1) & \\ (F_{n+1}, M_{n+1}) = (0, 0) & si \quad Z_n = 0, \\ (F_{n+1}, M_{n+1}) = \sum_{i=1}^{Z_n} (X_{n,i}, Y_{n,i}) & sinon \\ Z_n = L(F_n, M_n) & n \geq 1 \end{cases}$$

Alors $(Z_n)_{n\in\mathbb{N}}$ *suit un modèle de Daley.*

Théorème 38. *La suite* $(Z_n)_{n\in\mathbb{N}}$ *est une chaîne de Markov homogène à valeur dans* $\mathbb{N}$*. L'état 0 est absorbant.*

On sait que $G(z_1, z_2) = \mathbb{E}(z_1^X z_2^Y)$ avec G la fonction génératrice du couple (X, Y). On a donc $\mathbb{E}(z_1^{F_{n+1}} z_2^{M_{n+1}} | Z_n = j) = (G(z_1, z_2))^j$, avec $j \in \mathbb{N}$
Exemple de loi de reproduction :
Soient $T = X + Y$ le nombre d'individu dans la portée d'un couple donné, $H(z) = \sum_{i=0}^{\infty} p_i z^i$ sa fonction génératrice et $(p, q) \in (]0, 1[)^2$ où p est la probabilité que le nouveau-né soit une femelle. On obtient ainsi :$G(z_1, z_2) = H(pz_1 + qz_2)$. De plus, on peut supposer que le nombre de mâles et de femelles d'une portée sont indépendants donc : $G(z_1, z_2) = G_X(z_1)G_Y(z_2)$.
Exemples (cas particulier de la définition) :

- Soit $L(x, y) = x \min(1, y)$,on a donc :

$$Z_n = \begin{cases} F_n & si\, M_n \geq 1 \\ 0 & si\, M_n = 0 \end{cases}$$

  Ce modèle représente une situation de promiscuité totale des individus et un pouvoir reproductif illimité des mâles. On obtient alors ce genre de représentations au fur et à mesure des générations :

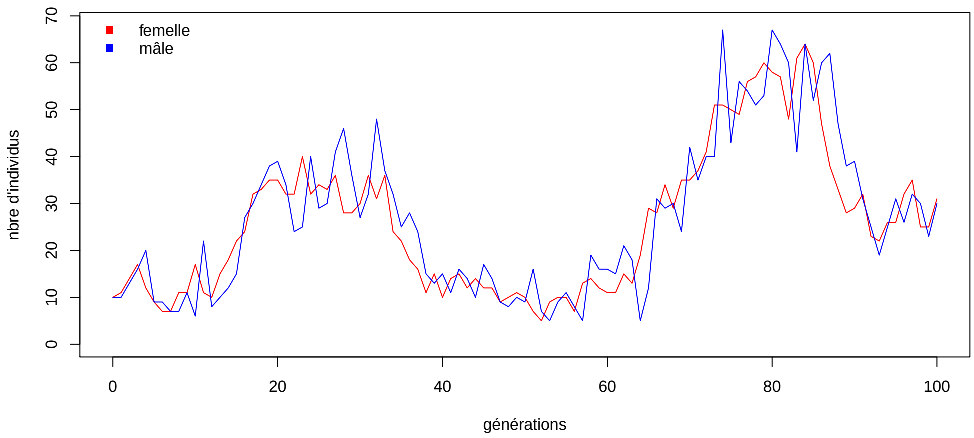

FIGURE 13 – Représentation de l'évolution de la population suivant leur sexe

– Soit $L(x, y) = \min(x, dy)$ avec $d \in \mathbb{N}^*$,on a donc :

$$Z_n = \begin{cases} F_n & \text{si } dM_n \geq F_n \\ dM_n & \text{sinon} \end{cases}$$

Ce modèle est appelé le modèle des couples fidèles.

Enfin, voici une simulation en temps continu d'une population sexuée. Pour cela, l'utilisation des propriétés de minimum et de non-vieillissement des lois exponentielles ont été utilisées. En effet, j'ai construit une matrice à 3 lignes. La première ligne indique les temps où la population change, la deuxième ligne indique le nombre de femelles présentes à partir de ce temps et jusqu'au temps suivant et la troisième ligne indique le nombre de mâles. Les temps sont simulés à partir du $\min(X_1, \ldots, X_f, Y_1, ..., Y_m)$ (qui sont les ages des f femelles et m mâles présents au temps précédent) qui suit une loi $\mathcal{E}(f\lambda_f + m\lambda_m)$ avec $1/\lambda_f$ et $1/\lambda_m$ les espérances de vie pour les femelles et les mâles. On ajoute alors au temps précédent une variable aléatoire de loi $\mathcal{E}(f\lambda_f + m\lambda_m)$ pour trouver le temps à partir duquel une personne meurt.

On tire ensuite une Bernoulli de paramètre de succès la proportion de femelles présentes pour déterminer si un homme ou une femme est mort. Si c'est une femelle, le nouveau nombre de femelle est alors : le nombre précédent où on enlève la personne décédée et auquel on rajoute une variable suivant une loi de poisson de paramètre le nombre moyen d'enfants femelles qu'à une femelle, et le nouveau nombre de mâle est le nombre précédent auquel on

rajoute les enfants mâles de la femelle. Si c'est un mâle qui meurt alors on enlève juste un mâle au nombre de mâle précédent.
Le graphique ci-dessous se place dans le cas de la population française sans immigration avec un taux de natalité de 2,1 enfants par femmes dont $48,8\%$ des bébés sont des filles et $51,2\%$ des garçons. Et l'espérance de vie est de 85 ans pour les femmes et 78 ans pour les hommes. On commence à l'an 0 avec 50 femmes et 50 hommes.

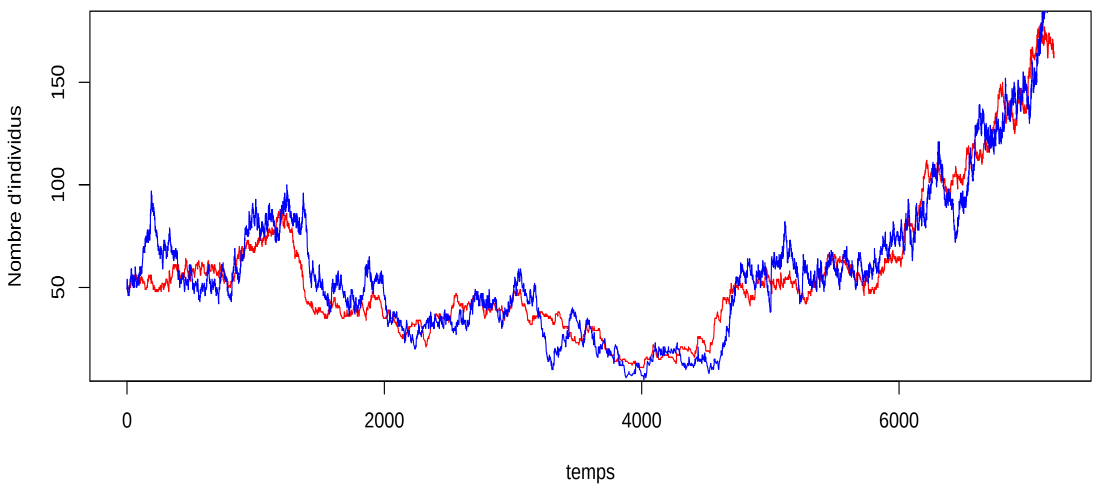

FIGURE 14 – Évolution de la population au cours du temps avec en rouge les femmes et en bleu les hommes

On peut ensuite rajouter le fait que les femelles donnent naissance au cours de leur vie et donc peuvent coexister avec leurs enfants. On va supposer que chaque femme donne naissance en même temps à tous ses enfants. Pour cela, la ligne des femelles a été séparé en deux lignes. L'une pour les femelles n'ayant pas encore donné naissance et l'autre pour le nombre de femelles déjà mère. On obtient ce genre de résultat où ici la population s'éteint.

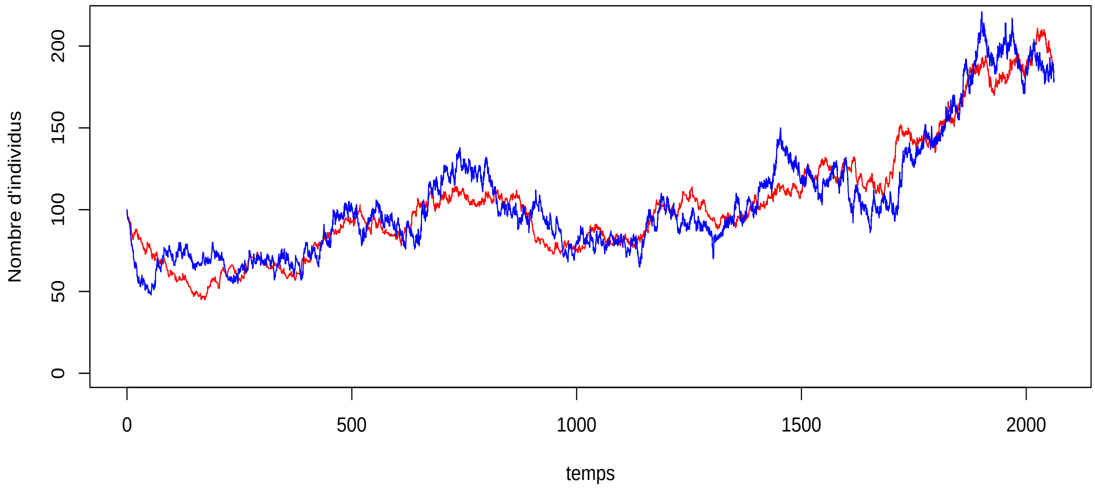

FIGURE 15 – Évolution de la population au cours du temps avec en rouge les femmes et en bleu les hommes

Conclusion

Pour conclure, on peut remarquer par ce travail que la modélisation d'une population peut très vite devenir complexe. Le modèle discret de Galton-Watson nous permet de décrire de manière assez simple l'évolution d'une population jusqu'à l'extinction ou alors jusqu'à un nombre infini d'individus. Ce modèle est utile pour, par exemple, regarder l'évolution d'un nom de famille en regardant le nombre de fils de chaque homme.

Ensuite, le processus de branchement en temps continu nous permet de prendre en compte les durées de vie différentes des individus. Il se rapproche davantage d'un modèle représentant une population. Cependant, on peut encore l'améliorer en ajoutant la notion de sexe et d'accouplement dans la population. On obtient ainsi une population avec de la reproduction sexuée.

J'aimerai enfin remercier mon maitre de stage Nicolas Pétrélis qui m'a encadrée pendant ces six semaines. Cela m'a confirmée dans ma préférence pour les probabilités et m'a permis de découvrir d'un peu plus près le quotidien d'un chercheur.

Références

[1] REMCO VAN DER HOFSTAD, *Random Graphs and Complex Networks*

[2] K. ATHREYA et P. NEY, *Branching Processes*

[3] Geoffrey GRIMMET et David STIRZAKER, *Probability and Random Processes*, (seconde édition), 1992

[4] Olivier GARET, *Chaînes de Galton-Watson*, septembre 2010

[5] Textes d'agrégation externe de mathématiques rendus public par le jury, *Épreuve de modélisation*, session 2008