Les théorèmes limites des processus de Galton-Watson

Adrien Joseph et Mickaël Launay

Juin 2006

Nous tenons à remercier Jean-François Le Gall qui a proposé ce sujet et encadré notre travail pour ses conseils et corrections.

Table des matières

1 Processus de Galton-Watson

On se donne une suite de réels positifs ou nuls $\langle p_k \rangle$ telle que $\sum_{k=0}^{\infty} p_k = 1$. Le processus de Galton-Watson se déroule de la manière suivante : on commence avec une particule. Elle a k enfants avec probabilité p_k. Chacun d'eux a de même k enfants avec probabilité p_k, indépendamment les uns des autres. Cela continue indéfiniment ou s'arrête lorsqu'il n'y a plus d'enfant (on parle alors d'extinction). Plus formellement, soit L une variable aléatoire telle que $\mathbf{P}(L = k) = p_k$ (L est appelée la loi de natalité du processus de Galton-Watson), et soit $\left(L_i^{(n)}\right)_{n,i\geq 1}$ des copies indépendantes de L. Le nombre de particules à l'instant n, noté Z_n, est défini récursivement de la manière suivante :

$$Z_0 = 1, \qquad Z_{n+1} = \sum_{i=1}^{Z_n} L_i^{(n+1)}.$$

On montre que $\langle Z_n \rangle_{n\in\mathbf{N}}$ est une chaîne de Markov de matrice de transition Q définie par : $Q(x,y) = \mu^{*x}(y)$, où μ désigne la loi de L et

$$\mu^{*0} = \delta_0, \qquad \mu^{*x} = \underbrace{\mu * \cdots * \mu}_{x \text{ fois}} \quad \text{si } x \geq 1.$$

L'espérance de L, notée m et appelée moyenne du processus de branchement, est :

$$m = \sum_{k=0}^{\infty} kp_k.$$

On appelle extinction l'événement $\{\exists n; Z_n = 0\}$. On note $\mathcal{F}_n$ la tribu engendrée par Z_i, $i \leq n$ et $\mathcal{L}_n$ la tribu engendrée par $L_i^{(k)}$, $i \geq 1$, $k \leq n$. On a alors $\mathcal{F}_n \subset \mathcal{L}_n$. L'arbre associé à un processus de branchement est obtenu en représentant les particules par des sommets, chacune étant liée par une arête à ses enfants.

Proposition 1.1 *Si $p_1 \neq 1$, alors, sur l'événement de non extinction, $Z_n \to \infty$ p.s..*

Preuve. Montrons que 0 est le seul état récurrent, ce qui entraîne le résultat de la proposition. Soit $x > 0$. Si $p_0 = 0$, comme $p_1 \neq 1$, il existe $k \geq 2$ tel que $p_k > 0$. Ainsi, $\mathbf{P}_x(Z_1 > x) > 0$, donc il existe $y > x$ tel que $Q(x,y) > 0$, et, en notant U le noyau potentiel, $U(x,y) > 0$. Or, comme $p_0 = 0$, $U(y,x) = 0$: x n'est pas récurrent. Supposons $p_0 > 0$. Revenir à une génération constituée de x particules exige que, dès la première étape, il n'y ait pas eu extinction, cet événement ayant une probabilité $1 - p_0^x < 1$ de se produire. En notant $H_x = \inf\{n \geq 1;\ X_n = x\}$, il vient $\mathbf{P}_x(H_x < \infty) < 1$: x est transitoire.

Proposition 1.2 *Lorsque $0 < m < \infty$, la suite $\langle \frac{Z_n}{m^n} \rangle_{n\in\mathbf{N}}$ est une martingale positive, qui converge p.s. vers une variable aléatoire W intégrable.*

Preuve. Par définition de Z_{n+1},

$$\mathbf{E}[Z_{n+1}|\mathcal{L}_n] = \mathbf{E}\Big[\sum_{i=1}^{\infty} \mathbf{1}_{\{i\leq Z_n\}} L_i^{(n+1)} |\mathcal{L}_n\Big] = \sum_{i=1}^{\infty} \mathbf{E}[\mathbf{1}_{\{i\leq Z_n\}} L_i^{(n+1)} |\mathcal{L}_n].$$

Or, $\mathbf{1}_{\{i\leq Z_n\}}$ est $\mathcal{L}_n$-mesurable, et $L_i^{(n+1)}$ est indépendante de $\mathcal{L}_n$, et son espérance vaut m, donc $\mathbf{E}[Z_{n+1}|\mathcal{L}_n] = mZ_n$. D'où :

$$\mathbf{E}[Z_{n+1}] = m\mathbf{E}[Z_n] = \cdots = m^{n+1}\mathbf{E}[Z_0] = m^{n+1} < \infty,$$

et

$$\mathbf{E}\Big[\frac{Z_{n+1}}{m^{n+1}}|\mathcal{L}_n\Big] = \frac{Z_n}{m^n}.$$

1.1 Fonction génératrice

La fonction génératrice de L est la fonction f définie par :

$$f(s) = \mathbf{E}\big[s^L\big] = \sum_{k=0}^{\infty} p_k s^k.$$

Remarquons que si P est une propriété telle que l'ensemble des arbres vérifiant P ait une probabilité s, alors $f(s)$ est la probabilité de l'ensemble des arbres dont tous les enfants vérifient P. En effet, un arbre a k enfants avec probabilité p_k, et ils vérifient tous P avec probabilité s^k. On définit de même la fonction conservatrice g telle que $g(s)$ soit la probabilité de l'ensemble des arbres qui ont au moins un fils satisfaisant la propriété P : $g(s) = \sum_{k=0}^{\infty} p_k(1-(1-s)^k)$. Il vient alors : $f(s) + g(1-s) = 1$.

Proposition 1.3 $\mathbf{E}\big[s^{Z_n}\big] = \underbrace{f \circ \cdots \circ f}_{n \text{ fois}}(s) = f^{(n)}(s).$

Preuve. On montre tout d'abord le résultat suivant :

$$\mathbf{E}\big[s^{Z_n} | \mathcal{F}_{n-1}\big] = f(s)^{Z_{n-1}}. \tag{1}$$

Les variables aléatoires $L_i^{(n)}$ étant indépendantes de la tribu $\mathcal{F}_{n-1}$, on a :

$$\mathbf{E}\big[s^{Z_n} | \mathcal{F}_{n-1}\big] = \mathbf{E}\Big[s^{\sum_{i=1}^{Z_{n-1}} L_i^{(n)}} | \mathcal{F}_{n-1}\Big] = \mathbf{E}\Big[\prod_{i=1}^{Z_{n-1}} s^{L_i^{(n)}} | \mathcal{F}_{n-1}\Big].$$

Or, Z_{n-1} est $\mathcal{F}_{n-1}$-mesurable et les $L_i^{(n)}$ sont des variables aléatoires indépendantes qui suivent la loi de L et sont indépendantes de la tribu $\mathcal{F}_{n-1}$. Par conséquent :

$$\mathbf{E}\big[s^{Z_n} | \mathcal{F}_{n-1}\big] = \prod_{i=1}^{Z_{n-1}} \mathbf{E}\big[s^{L_i^{(n)}}\big] = \mathbf{E}\big[s^L\big]^{Z_{n-1}} = f(s)^{Z_{n-1}}.$$

Ainsi, $\mathbf{E}\big[s^{Z_n}\big] = \mathbf{E}\big[f(s)^{Z_{n-1}}\big]$. Il suffit pour conclure de faire une récurrence sur n : c'est clair pour $n = 0$, et si c'est vrai pour n, alors :

$$\mathbf{E}\big[s^{Z_{n+1}}\big] = \mathbf{E}\big[f(s)^{Z_n}\big] = f^{(n)}(f(s)) = f^{(n+1)}(s).$$

Corollaire 1.4 *En notant $q = \mathbf{P}(\exists n;\ Z_n = 0)$, on a : $q = \lim_{n\to\infty} f^{(n)}(0)$. Si $m \leq 1$ et $p_1 \neq 1$, q vaut 1, tandis que si $m > 1$, q est l'unique solution de $f(t) = t$ dans $[0, 1[$.*

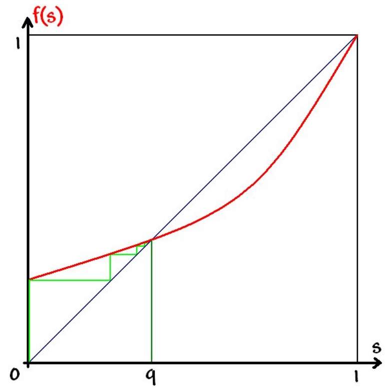
Preuve. Comme l'extinction est la réunion croissante des événements $\{Z_n = 0\}$, on a : $q = \lim_{n\to\infty} \mathbf{P}(Z_n = 0) = \lim_{n\to\infty} f^{(n)}(0)$.

Supposons $m \leq 1$ et $p_0 + p_1 \neq 1$ (le cas $p_0 + p_1 = 1$ se traite aisément). Soit h la fonction définie sur $[0, 1]$ par $h(t) = f(t) - t$. On a $h' = f' - 1$, et comme pour tout $t < 1$, $f'(t) < f'(1) = m \leq 1$, h est strictement décroissante sur $[0, 1]$. Or, $h(1) = 0$, donc

$h(t) > 0$ sur $[0,1[$, i.e. $f(t) > t$ pour tout $t \in [0,1[$. Notons alors $r = 1$ l'unique point fixe de f.

Prouvons que si $m > 1$, l'équation $f(t) = t$ admet une unique solution dans $[0,1]$. La fonction h' est strictement croissante, $h'(0) = p_0 - 1 < 0$ et $h'(1) = m - 1 > 0$, donc h' s'annule en un unique t_0. Ainsi, h décroît de 0 à t_0 et croît de t_0 à 1. Or, $h(0) = p_0 \geq 0$ et $h(1) = 0$, donc il existe un unique $r \in [0,1[$ tel $h(t) = 0$, i.e. tel que $f(t) = t$.

Puisque f est croissante et que $0 \leq f(0)$, la suite $\langle f^{(n)}(0)\rangle_{n\in\mathbf{N}}$ est croissante. On montre par récurrence qu'elle est de plus majorée par r. Par suite, elle converge vers un point fixe de f inférieur à r, i.e. vers r. Donc $q = r$.



Définition 1 *Un processus de branchement est qualifié de sous-critique quand $m < 1$, critique quand $m = 1$ et surcritique quand $m > 1$.*

Nous allons maintenant nous intéresser à la relation entre $\langle Z_n\rangle$ et $\langle Z_n^*\rangle$, où Z_n^* désigne le nombre de particules de la génération n qui engendrent une lignée infinie. Nous allons montrer que $\langle Z_n^*\rangle$ suit un processus de Galton-Watson, dont nous expliciterons la fonction génératrice. Nous étudierons également $\langle Z_n\rangle$ sous l'événement d'extinction.

Soit A un arbre suivant un processus de Galton-Watson de fonction génératrice f, et A' le sous-arbre constitué des particules qui génèrent une infinité de descendants. Notons $Y_{i,j}^{(n)}$ l'indicatrice valant 1 lorsque le j-ème enfant de la i-ème particule de la génération n engendre une lignée infinie et

$$\overline{q} = 1 - q, \qquad L_i^{*(n+1)} = \sum_{j=1}^{L_i^{(n+1)}} Y_{i,j}^{(n)}, \qquad Z_{n+1}^* = \sum_{i=1}^{Z_n} L_i^{*(n+1)}.$$

Proposition 1.5 *Supposons que $m > 1$.*

La loi de A' sous l'événement non extinction est la même que celle de l'arbre de Galton-Watson avec la fonction génératrice suivante :

$$f^*(s) = \frac{f(q + \overline{q}s) - q}{\overline{q}}. \tag{2}$$

La loi de A sous l'événement extinction est la même que celle de l'arbre de Galton-Watson avec la fonction génératrice suivante :

$$\widetilde{f}(s) = \frac{f(qs)}{q}. \tag{3}$$

La fonction génératrice conjointe de $L - L^*$ et de L^* est :

$$\mathbf{E}\big[s^{L-L^*}t^{L^*}\big] = f(qs + \overline{q}t). \tag{4}$$

La fonction génératrice conjointe de $Z_n - Z_n^*$ et de Z_n^* est :

$$\mathbf{E}\big[s^{Z_n-Z_n^*}t^{Z_n^*}\big] = f^{(n)}(qs + \overline{q}t). \tag{5}$$

Preuve. Commençons par démontrer (4). On écrit

$$\mathbf{E}\big[s^{L-L^*}t^{L^*}\big] = \mathbf{E}\big[\mathbf{E}\big[s^{L-L^*}t^{L^*}|L\big]\big] = \mathbf{E}\Big[\mathbf{E}\Big[s^{\sum_{j=1}^{L}(1-Y_j)}t^{\sum_{j=1}^{L}Y_j}|L\Big]\Big].$$

Comme les Y_i sont indépendantes et identiquement distribuées, on a :

$$\mathbf{E}\big[s^{L-L^*}t^{L^*}\big] = \mathbf{E}\big[\big(\mathbf{E}\big[s^{1-Y_1}t^{Y_1}\big]\big)^L\big].$$

Or, $\mathbf{P}(Y_1 = 0) = q$ et $\mathbf{P}(Y_1 = 1) = \overline{q}$, donc

$$\mathbf{E}\big[s^{L-L^*}t^{L^*}\big] = \mathbf{E}\big[(qs + \overline{q}t)^L\big] = f(qs + \overline{q}t).$$

On montre de la même manière (5) en remarquant que $\mathbf{E}\big[s^{Z_n}\big] = f^{(n)}(s)$.

Notons $\mathcal{F}_{n,i}$ la tribu engendrée par $L_k^{*(n)}(k \neq i)$ et $L_k^{*(m)}(m < n, k \geq 1)$. Afin de montrer (2), nous devons prouver que la fonction génératrice de $L_i^{*(n)}$ sachant $L_i^{*(n)} \neq 0$ et sachant $\mathcal{F}_{n,i}$ est f^*. Or, sur l'événement $L_i^{*(n)} \neq 0$,

$$\mathbf{E}\big[s^{L_i^{*(n)}}|\mathcal{F}_{n,i}\big] = \mathbf{E}\big[s^{L_i^{*(n)}}|L_i^{*(n)} \neq 0\big] = \frac{\mathbf{E}\big[s^{L^*}\mathbf{1}_{\{L^*\neq 0\}}\big]}{\overline{q}},$$

et comme

$$\frac{\mathbf{E}\big[s^{L^*}(\mathbf{1} - \mathbf{1}_{\{L^*=0\}})\big]}{\overline{q}} = \frac{\mathbf{E}\big[1^{L-L^*}s^{L^*}\big] - \mathbf{P}(L^* = 0)}{\overline{q}},$$

on a d'après (4) :

$$\mathbf{E}\big[s^{L_i^{*(n)}}|\mathcal{F}_{n,i}\big] = \mathbf{E}\big[s^{L_i^{*(n)}}|L_i^{*(n)} \neq 0\big] = \frac{f(q + \overline{q}s) - q}{\overline{q}} = f^*(s).$$

De même, (3) vient du fait que sur l'événement extinction :

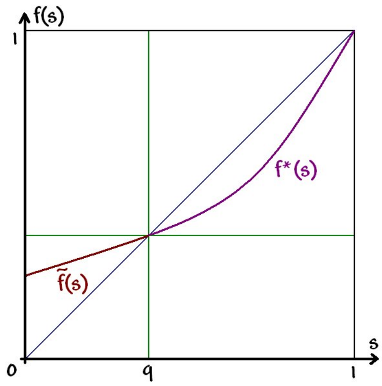
$$\mathbf{E}\big[s^{L_i^{(n)}}|\mathcal{F}_{n,i}\big] = \mathbf{E}\big[s^{L_i^{(n)}}|L_i^{*(n)} = 0\big] = \frac{\mathbf{E}\big[s^{L}\mathbf{1}_{\{L^*=0\}}\big]}{q},$$

et puisque

$$\frac{\mathbf{E}\big[s^{L}\mathbf{1}_{\{L^*=0\}}\big]}{q} = \frac{\mathbf{E}\big[s^{L-L^*}0^{L^*}\big]}{q},$$

on obtient par (4) :

$$\mathbf{E}\big[s^{L_i^{(n)}}|\mathcal{F}_{n,i}\big] = \mathbf{E}\big[s^{L_i^{(n)}}|L_i^{*(n)} = 0\big] = \frac{f(qs + \overline{q}0)}{q} = \widetilde{f}(s).$$



1.2 Hérédité

Définition 2 *On dit qu'une propriété est héréditaire si tous les arbres finis ont cette propriété et si, lorsqu'un arbre possède cette propriété, il en est de même pour ses enfants.*

Proposition 1.6 *Soit P une propriété héréditaire et F l'ensemble des arbres vérifiant P. Alors $\mathbf{P}(F) \in \{q, 1\}$. Si $m > 1$, cette propriété a une probabilité 0 ou 1 conditionnellement à l'événement non extinction.*

Preuve. Pour tout arbre A dont la racine possède k enfants, notons $A^{(1)}, \ldots, A^{(k)}$ les arbres issus de ses enfants. Alors, par définition d'une propriété héréditaire,

$$\mathbf{P}(F) = \mathbf{E}\big[\mathbf{P}(A \in F | Z_1)\big] \leq \mathbf{E}\big[\mathbf{P}(A^{(1)} \in F, \ldots, A^{(Z_1)} \in F | Z_1)\big].$$

Comme $A^{(1)}, \ldots, A^{(Z_1)}$ sont indépendantes et identiquement distribuées sachant Z_1, on a : $\mathbf{P}(F) \leq \mathbf{E}\big[\mathbf{P}(F)^{Z_1}\big] = f(\mathbf{P}(F))$. De plus, comme tous les arbres éteints (qui sont finis) vérifient P, $\mathbf{P}(F) \geq q$. D'où $\mathbf{P}(F) \in \{q, 1\}$. Si $m > 1$ et $\mathbf{P}(F) = q$, alors l'ensemble des arbres éteints, qui est inclus dans F, a la même probabilité que F, donc les arbres vérifiant P sont éteints p.s. : P a une probabilité nulle sur l'événement non extinction.

Corollaire 1.7 *Si $0 < m \leq 1$ et $p_1 \neq 1$, alors $W = 0$ p.s., tandis que si $1 < m < \infty$, alors $W = 0$ p.s. ou $W > 0$ p.s. sur l'ensemble de non extinction, i.e. $\mathbf{P}(W = 0) \in \{q, 1\}$.*

Preuve La propriété $W = 0$ étant héréditaire, la proposition 1.6 s'applique.

Théorème 1.8 *Si $1 < m < \infty$, alors il existe une suite $\langle c_n \rangle$ de réels positifs telle que*

$$\lim_{n\to\infty} \frac{Z_n}{c_n} \textit{ existe p.s. dans } [0, \infty[, \tag{6}$$

$$\mathbf{P}\Big(\lim_{n\to\infty} \frac{Z_n}{c_n} = 0\Big) = q, \tag{7}$$

$$\lim_{n\to\infty} \frac{c_{n+1}}{c_n} = m. \tag{8}$$

Preuve. Soit $s_0 \in]q,1[$. Posons $s_{n+1} = f^{-1}(s_n)$ pour tout $n \geq 0$. Alors, puisque, sur $]q,1[$, f est croissante et $f(t) < t$, la suite $\langle s_n \rangle$ est croissante et converge vers l'unique point fixe de f sur $]q,1]$, i.e. 1. D'après (1), $\langle s_n^{Z_n} \rangle_{n\in\mathbf{N}}$ est une martingale et pour tout $n \in \mathbf{N}$, $0 \leq s_n^{Z_n} \leq 1$, donc $\langle s_n^{Z_n} \rangle_{n\in\mathbf{N}}$ converge p.s. et dans L^1 vers une variable aléatoire notée $Y \in [0,1]$. En particulier, on a : $\mathbf{E}[Y] = \mathbf{E}\big[s_0^{Z_0}\big] = s_0$. En posant, $c_n = -\frac{1}{\log s_n}$, on obtient ainsi : $s_n^{Z_n} = e^{-Z_n/c_n}$, donc $\frac{Z_n}{c_n}$ admet une limite dans $[0,\infty]$ presque sûrement. Comme $\lim_{x\to 1} f(x) = 1$, $\log f(s) \sim_{x\to 1} f(s) - 1$. Puisque $f'(1) = m$, on en déduit que :

$$\lim_{s\to 1} \frac{\log f(s)}{\log s} = m, \text{ puis } \lim_{n\to\infty} -\frac{1}{\log s_n} \frac{1}{-\frac{1}{\log f(s_n)}} = m,$$

ce qui prouve (8), et donc que la propriété $\lim \frac{Z_n}{c_n} = 0$ est héréditaire. Par la proposition 1.6, $\mathbf{P}(\lim \frac{Z_n}{c_n} = 0) \in \{q, 1\}$. Or, si $\lim \frac{Z_n}{c_n} = 0$ p.s., alors $\lim s_n^{Z_n} = 1$ p.s., donc $\mathbf{E}[Y] = 1$, ce qui contredit $s_0 < 1$. D'où (7). De même, la propriété $\limsup \frac{Z_n}{c_n} < \infty$ est héréditaire, donc a pour probabilité q ou 1. Or, si elle a pour probabilité q, alors, puisque $\lim \frac{Z_n}{c_n}$ existe p.s., $\mathbf{P}(\lim \frac{Z_n}{c_n} < \infty) = q$, donc $\mathbf{P}(Y > 0) = q$. Par suite, $\mathbf{E}[Y] = \mathbf{E}[Y \mathbf{1}_{\{0<Y\leq 1\}}] \leq \mathbf{P}(Y > 0)$, donc $\mathbf{E}[Y] \leq q$, ce qui est faux. Donc $\mathbf{P}(\limsup \frac{Z_n}{c_n} < \infty) = 1$, et comme $\lim \frac{Z_n}{c_n}$ existe p.s., il vient : $\lim_{n\to\infty} \frac{Z_n}{c_n}$ existe p.s. dans $[0,\infty[$.

Nous verrons dans la partie consacrée au cas surcritique que l'on peut prendre $c_n = m^n$ si et seulement si $\mathbf{E}[L \log^+ L] < \infty$.

1.3 Arbres biaisés

Les arbres sont des objets qui vont nous intéresser tout au long de l'exposé ; explicitons ce que l'on entend par arbres :

Définition 3 *Un ensemble A de suites finies d'entiers naturels est un arbre lorsque la suite vide appartient à A (appelée racine de A) et lorsque $\langle i_1, i_2, \ldots, i_n \rangle \in A$ implique que pour tout $k \in [1,n]$, $\langle i_1, i_2, \ldots, i_k \rangle \in A$, et que pour tout $j \in [1, i_n]$, $\langle i_1, i_2, \ldots, i_{n-1}, j \rangle \in A$. Dans ce cas, $\langle i_1, i_2, \ldots, i_n \rangle$ est le i_n-ème enfant du i_{n-1}-ème enfant du ... du i_1-ème enfant de la racine. Si $\langle i \rangle \in A$, on défini l'arbre issu du sommet i par $A^{(i)} = \{\langle i_1, i_2, \ldots i_n \rangle ; \langle i, i_1, i_2, \ldots i_n \rangle \in A\}$. La hauteur de A est le supremum des longueurs des suites appartenant à A. Pour tout n, on appelle troncature de A à ses n premiers niveaux l'ensemble $A \upharpoonright n = \{\langle i_1, i_2, \ldots, i_k \rangle \in A; k \leq n\}$. C'est un arbre de hauteur au plus n. A est dit localement fini lorsque toutes ses troncatures sont des ensembles finis. On note $\mathcal{A}$ l'ensemble des arbres localement finis.*

On souhaite construire la mesure de probabilité $\mathbf{GW}$ sur l'ensemble $\mathcal{A}$ permettant de traduire sous forme d'arbre un processus de Galton-Watson. On rappelle que la racine de l'arbre représentant un processus de Galton-Watson correspond à la particule initiale, et que chaque particule est liée par une arête à ses enfants.

Soit $(\Omega, \mathbf{P})$ l'espace mesuré sur lequel les variables aléatoires $L_i^{(n)}$ sont définies. Notons tout d'abord $\langle p_n \rangle_{n\in\mathbf{N}}$ la suite croissante des nombres premiers, et posons

$$I(\emptyset) = 1, \qquad I(\langle i_1, \dots, i_j \rangle) = \prod_{k=1}^{j} p_{i_k}^{i_k} \quad \text{si } j \geq 1.$$

On définit la fonction mesurable $\phi : \Omega \to \mathcal{A}$ par :

$$\phi\big((L_i^{(n)})_{i,n\geq 1}\big) = \{\langle i_1, \dots, i_n \rangle ; n \in \mathbf{N}, 1 \leq i_{j+1} \leq L^{(j+1)}_{I(\langle i_1,\dots,i_j\rangle)} (0 \leq j \leq n-1)\}.$$

La mesure de probabilité $\mathbf{GW}$ est alors la mesure image de $\mathbf{P}$ par ϕ.

Introduisons la notion de loi biaisée. Supposons que, pour un arrêt de bus donné, l'écart entre l'arrivée de deux bus successifs suit une loi sur les entiers notée μ, de moyenne $0 < m = \sum_{k=0}^{\infty} k p_k < \infty$, avec $p_k = \mu(k)$. L'attente (plus précisément le double de l'attente) d'une personne arrivant à l'arrêt de bus en milieu de journée et souhaitant prendre le bus suit alors la loi biaisée : la probabilité d'attendre $\frac{k}{2}$ minutes est $\frac{kp_k}{m}$. On constate que l'attente moyenne vaut $\frac{1}{2}(m + \frac{\sigma^2}{m}) \geq \frac{m}{2}$, où $\sigma^2 = \sum_{k=0}^{\infty} k^2 p_k - m^2$.

Donnons un second exemple : c'est la fête de l'école et, comme chaque année, une tombola est organisée. Chaque famille est invitée à mettre un billet dans l'urne. Si la proportion des familles ayant k enfants est p_k, la probabilité pour qu'une famille de k enfants gagne est p_k. Cette méthode n'est pas très juste parce qu'elle offre autant de chance de gagner à une famille nombreuse qu'à une famille n'ayant qu'un enfant dans l'école. Une autre façon de procéder est de mettre un billet dans l'urne pour chaque enfant. La probabilité pour que la famille gagnante ait k enfants est alors la probabilité biaisée $\frac{kp_k}{m}$.

Proposition 1.9 *Soit $\widehat{L}$ une variable suivant la loi biaisée de L, i.e.*

$$\mathbf{P}(\widehat{L} = k) = \frac{kp_k}{m}.$$

Plus généralement, si X désigne une variable aléatoire positive intégrable non nulle p.s., on dit que $\widehat{X}$ suit la distribution biaisée de X si, pour tout intervalle $A \subset [0, \infty[$, on a :

$$\mathbf{P}(\widehat{X} \in A) = \frac{\mathbf{E}[X \mathbf{1}_A(X)]}{\mathbf{E}[X]}.$$

Ceci est équivalent à : pour tout fonction mesurable $f : [0, \infty[\to [0, \infty[$,

$$\mathbf{E}[f(\widehat{X})] = \frac{\mathbf{E}[X f(X)]}{\mathbf{E}[X]}.$$

Preuve. Montrons cette équivalence. Posons

$$\mathcal{M} = \Big\{ B \in \mathcal{B}([0, \infty[) ; \mathbf{P}(\widehat{X} \in B) = \frac{\mathbf{E}[X \mathbf{1}_B(X)]}{\mathbf{E}[X]} \Big\}.$$

$\mathcal{M}$ est une classe monotone qui contient l'ensemble $\mathcal{C}$ des intervalles de $[0, \infty[$. Comme $\mathcal{C}$ est stable par intersections finies et engendre $\mathcal{B}([0, \infty[)$, $\mathcal{M}$ est $\mathcal{B}([0, \infty[)$ tout entier. Ensuite, si f est une fonction mesurable positive, alors elle est limite croissante de fonctions étagées positives, et en appliquant le théorème de convergence monotone, on a le résultat.

Pour la réciproque, prendre $f = 1_A$.

Une manière de biaiser **GW** relativement à Z_n pour tout n est de choisir uniformément un sommet de la génération n. La probabilité limite obtenue est appelée $\widehat{\mathbf{GW}}$. Aussi $\widehat{\mathbf{GW}}$ correspond-elle à l'idée intuitive de choisir un arbre par les feuilles : les arbres ayant un plus grand W ont plus de chance d'être choisis. Afin de pouvoir définir rigoureusement cette mesure de probabilité, on va en construire une autre, notée $\widehat{\mathbf{GW}}_*$, définie sur l'ensemble $\mathcal{A}_*$ des arbres infinis munis d'un chemin distingué infini :

$$\mathcal{A}_* = \{(a, \langle c_n \rangle_{n\in\mathbf{N}^*}); a \in \mathcal{A}, \forall k \in \mathbf{N}^* \langle c_1, c_2, \ldots, c_k \rangle \in a\}.$$

Puisque $\mathbf{E}[Z_n] = m^n$, on souhaite que, pour tout $a \in \mathcal{A}$, cette dernière mesure satisfasse la relation suivante :

$$\widehat{\mathbf{GW}}_*[a; s]_n = \frac{1}{m^n}\mathbf{GW}[a]_n, \tag{9}$$

où $[a]_n$ désigne la partie de $\mathcal{A}$ constituée des arbres dont les n premiers niveaux coïncident avec ceux de a et où $[a; s]_n$ désigne la partie de $\mathcal{A}_*$ formée des arbres dont les n premiers niveaux coïncident avec ceux de a et munis d'un chemin distingué infini passant par s, s étant un sommet appartenant au n-ième niveau de a. On remarque que si $\widehat{\mathbf{GW}}_*$ satisfait la relation (9), alors, en notant $W_n(a) = \frac{Z_n}{m^n}$, sa projection sur l'ensemble $\mathcal{A}$, que l'on notera $\widehat{\mathbf{GW}}$, vérifie la relation fondamentale suivante :

$$\widehat{\mathbf{GW}}[a]_n = W_n(a)\mathbf{GW}[a]_n. \tag{10}$$

On constate que, une fois cette relation établie, on a, pour tout k :

$$\widehat{\mathbf{GW}}(Z_n = k) = \frac{k}{m^n}\mathbf{GW}(Z_n = k) = k\frac{\mathbf{P}_{GW}(Z_n = k)}{\mathbf{E}_{GW}[Z_n]},$$

i.e.

$$\widehat{\mathbf{GW}}(Z_n = k) = \mathbf{GW}(\widehat{Z}_n = k). \tag{11}$$

Avant de définir proprement $\widehat{\mathbf{GW}}_*$, voyons ce que (9) implique. Soit a un arbre de hauteur supérieure ou égale à $n+1$ et dont la racine a k enfants. Notons $a^{(1)}, a^{(2)}, \ldots, a^{(k)}$ les arbres issus de ces enfants. Soit $s = \langle i_1, i_2, \ldots i_{n+1} \rangle$ un sommet de a appartenant au $(n+1)$-ème niveau. Il appartient aussi à un unique $a^{(i)}$ (en fait $i = i_1$). Par définition de la mesure de probabilité **GW**, on a :

$$\mathbf{GW}[a]_{n+1} = p_k \prod_{j=1}^{k} \mathbf{GW}[a^{(j)}]_n = kp_k \frac{1}{k}\Big(\prod_{j\neq i} \mathbf{GW}[a^{(j)}]_n\Big)\mathbf{GW}[a^{(i)}]_n.$$

Par (9), on en déduit que la mesure de probabilité $\widehat{\mathbf{GW}}_*$ doit vérifier la relation de récurrence suivante :

$$\widehat{\mathbf{GW}}_*[a; s]_{n+1} = \frac{kp_k}{m}\frac{1}{k}\Big(\prod_{j\neq i} \mathbf{GW}[a^{(j)}]_n\Big)\widehat{\mathbf{GW}}_*[a^{(i)}; \widetilde{s}]_n, \tag{12}$$

où $\widetilde{s}$ désigne la suite $\langle i_2, \ldots i_{n+1} \rangle$. Réciproquement, si une mesure de probabilité $\widehat{\mathbf{GW}}_*$ vérifie (12), alors, par récurrence, elle vérifie aussi (9). Ces remarques vont nous permettre de construire $\widehat{\mathbf{GW}}_*$.

Par définition d'une loi biaisée, $\mathbf{P}(\widehat{L} = k) = \frac{kp_k}{m}$. Soit $\widehat{L}_n$ des copies indépendantes de $\widehat{L}$. La construction d'un arbre $\widehat{A}$ de Galton-Watson biaisé commence avec la donnée d'une particule initiale s_0. Elle donne naissance à $\widehat{L}_1$ particules (on remarque que, presque sûrement, s_0 a au moins un enfant). Prenons un de ses enfants au hasard, que l'on note s_1. Les autres engendrent indépendamment les uns des autres des arbres de Galton-Watson classiques. La particule s_1, quant à elle, donne naissance à $\widehat{L}_2$ particules. De même, on prend une de ces nouvelles particules au hasard, que l'on note s_2. Les autres particules sont les racines d'arbres de Galton-Watson ordinaires. On recommence ce processus indéfiniment. Notons que, comme $\widehat{L} \geq 1$, les arbres biaisés sont toujours infinis.

Définissons la mesure de probabilité $\widehat{\mathbf{GW}}_*$ sur l'ensemble $\mathcal{A}_*$. Notons $(\Omega, \mathbf{P})$ l'espace mesuré sur lequel les variables aléatoires $\big(\widehat{L}_n\big)_{n\geq 1}$, $(U_n)_{n\geq 1}$ et $\big(L_i^{(n)}\big)_{n\geq 2, i\geq 1}$ sont définies. On suppose que les variables aléatoires $\widehat{L}_n$ sont des copies indépendantes de $\widehat{L}$, les couples $(\widehat{L}_n, U_n)$ sont indépendants et vérifient :

$$\mathbf{P}(U_n = k | \widehat{L}_n) = \frac{1}{\widehat{L}_n} \mathbf{1}_{\{1 \leq k \leq \widehat{L}_n\}}.$$

Les variables aléatoires $(L_i^{(n)})$ sont enfin supposées indépendantes des couples $(\widehat{L}_n, U_n)$. On définit la fonction mesurable $\psi : \Omega \to \mathcal{A}_*$ par :

$$\begin{aligned} & \psi\big((L_i^{(n)})_{n\geq 2, i\geq 1}, (U_n)_{n\geq 1}, (\widehat{L}_n)_{n\geq 1}\big) \\ = \ & \Big(\bigcup_{k\geq 1} \Big\{ \langle U_1, \ldots, U_{k-1}, i_k, \ldots, i_n \rangle ; n \geq k-1, i_k \in \big\{1, \ldots, U_k - 1, U_k + 1, \ldots, \widehat{L}_k\big\}, \\ & 1 \leq i_{j+1} \leq L^{(j+1)}_{I(\langle U_1, \ldots, U_{k-1}, i_k, \ldots, i_j \rangle)} (k \leq j \leq n-1) \Big\}, \langle U_n \rangle_{n \in \mathbf{N}^*} \Big). \end{aligned}$$

La mesure de probabilité $\widehat{\mathbf{GW}}_*$ est alors la mesure image de $\mathbf{P}$ par ψ. On déduit de cette construction et de $\mathbf{E}[Z_n] = m^n$ la relation (9).

Définissons maintenant la mesure de probabilité $\widehat{\mathbf{GW}}$ sur l'ensemble $\mathcal{A}$. Pour cela, on pose tout d'abord, pour tout arbre $a \in \mathcal{A}$,

$$ch(a) = \{\langle u_n \rangle_{n \in \mathbf{N}^*} ; \forall i \geq 1, \langle u_1, \ldots, u_i \rangle \in a\},$$

puis, pour toute partie F de $\mathcal{A}$,

$$F_* = \bigcup_{a \in F} \Big(\bigcup_{c \in ch(a)} \{(a, c)\} \Big).$$

Pour toute partie F de $\mathcal{A}$ telle que F_* soit mesurable, on peut alors définir

$$\widehat{\mathbf{GW}}(F) = \widehat{\mathbf{GW}}_*(F_*).$$

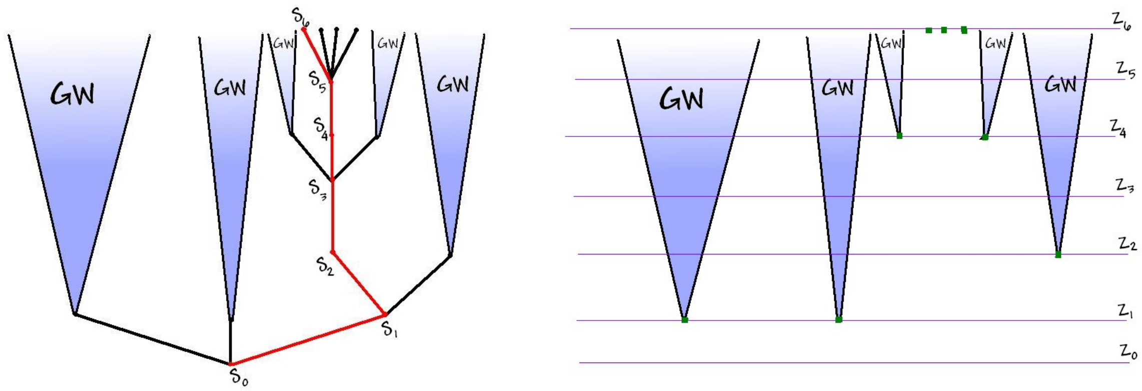
On obtient ainsi la relation fondamentale (10) entre les deux probabilités $\widehat{\mathbf{GW}}$ et $\mathbf{GW}$.

Les sommets hors du chemin distingué $\langle s_0, s_1, \ldots\rangle$ de l'arbre biaisé forment un processus de branchement avec immigration. En général, un tel processus est défini par deux lois : une loi de natalité et une loi d'immigration. Le processus démarre avec aucune particule, et à chaque génération $n \geq 1$, il y a une immigration de Y_n particules, les Y_n étant des variables aléatoires indépendantes suivant la loi d'immigration. De plus, chacune de ces particules est la particule initiale d'un processus de Galton-Watson suivant la loi de natalité, ces processus étant indépendants les uns des autres. Le nombre Z_n de particules à l'instant n est alors défini récursivement de la manière suivante :

$$Z_0 = 0, \qquad Z_{n+1} = \sum_{i=1}^{Z_n} L_i^{(n+1)} + Y_{n+1},$$

où les $L_i^{(n)}$ sont des variables aléatoires indépendantes suivant la loi de natalité. Dans notre cas, la $\widehat{\mathbf{GW}}$-loi de $Z_n - 1$ est la même que celle de la taille de la n-ème génération d'un processus de branchement avec immigration ayant pour loi de natalité celle de L et pour loi d'immigration celle de $\widehat{L} - 1$.

Pour chaque arbre biaisé $(a, s) \in \mathcal{A}_*$, on dispose d'un chemin $s = \langle s_0, s_1, \ldots\rangle$ partant de la racine dont les sommets ont un nombre d'enfants selon la loi de la variable aléatoire biaisée $\widehat{L}$. Si, dans un tel arbre, on efface les sommets constituant s, le graphe obtenu apparaît comme celui d'une population se développant selon une natalité L et une immigration $\widehat{L} - 1$:



2 Cas surcritique

2.1 Premiers résultats

Lemme 2.1 *Soit μ une mesure finie et ν une mesure de probabilité sur une tribu $\mathcal{F}$. Soit $\langle \mathcal{F}_n \rangle$ une filtration de $\mathcal{F}$ dont la réunion engendre $\mathcal{F}$. On suppose que $(\mu \upharpoonright \mathcal{F}_n)$ est absolument continue par rapport à $(\nu \upharpoonright \mathcal{F}_n)$, de dérivée de Radon-Nikodym X_n. Posons $X = \limsup_{n\to\infty} X_n$. Alors :*

$$\mu \ll \nu \quad \Longleftrightarrow \quad X < \infty \quad \mu\text{-p.s.} \quad \Longleftrightarrow \quad \int X d\nu = \int d\mu,$$

$$\mu \perp \nu \quad \Longleftrightarrow \quad X = \infty \quad \mu\text{-p.s.} \quad \Longleftrightarrow \quad \int X d\nu = 0.$$

Preuve. On montre tout d'abord que $\langle (X_n, \mathcal{F}_n) \rangle_{n\in\mathbf{N}}$ est une martingale positive relativement à la mesure ν. Soit $A \in \mathcal{F}_n \subset \mathcal{F}_{n+1}$.

$$\mathbf{E}[1_A X_{n+1}] = \int_A X_{n+1} d\nu = \int_A d\mu = \int_A X_n d\nu = \mathbf{E}[1_A X_n].$$

De plus, X_n est $\mathcal{F}_n$-mesurable. D'où le résultat par la propriété caractéristique. Par conséquent, $\langle X_n \rangle_{n\in\mathbf{N}}$ converge ν-p.s. vers $X \in L^1(d\nu)$, et en particulier $X < \infty$ ν-p.s..

Montrons que le lemme découle de la décomposition suivante de μ en une partie ν-absolument continue et une partie ν-singulière :

$$\forall A \in \mathcal{F} \quad \mu(A) = \int_A X d\nu + \mu(A \cap \{X = \infty\}). \tag{13}$$

En effet : si $\mu \ll \nu$, alors $X < \infty$ μ-p.s.; si $X < \infty$ μ-p.s., alors par (13) on a : $\int X d\nu = \int d\mu$; et si $\int X d\nu = \int d\mu$, alors, par (13), $X < \infty$ μ-p.s., et $\mu \ll \nu$. Si maintenant $\mu \perp \nu$, alors il existe un événement $A \in \mathcal{F}$ tel que $\mu(A^c) = \nu(A) = 0$, ce qui, d'après (13), implique que $\mu(A \cap \{X = \infty\}) = \mu(A)$, d'où $X = \infty$ μ-p.s.; si $X = \infty$ μ-p.s., alors par (13), $\int X d\nu = 0$; et si $\int X d\nu = 0$, alors par (13), $X = \infty$ μ-p.s., d'où $\mu \perp \nu$ en considérant $\{X = \infty\}$.

Afin de prouver (13), supposons tout d'abord que $\mu \ll \nu$ avec une dérivée de Radon-Nikodym $\widetilde{X}$. Alors $X_n = \mathbf{E}[\widetilde{X} | \mathcal{F}_n]$, et $\langle (X_n, \mathcal{F}_n) \rangle_{n\in\mathbf{N}}$ est une martingale fermée qui converge ν-p.s. vers $\widetilde{X}$. Donc $X = \widetilde{X}$ ν-p.s., et (13) est la définition de la dérivée de Radon-Nikodym.

Dans le cas général, notons ρ la mesure de probabilité définie par : $\rho = (\mu + \nu)/C$, où $C = \int d(\mu + \nu)$. Alors $\mu, \nu \ll \rho$, donc on peut appliquer ce que nous avons montré avec les variables $U_n = d(\mu \upharpoonright \mathcal{F}_n)/d(\rho \upharpoonright \mathcal{F}_n)$ et $V_n = d(\nu \upharpoonright \mathcal{F}_n)/d(\rho \upharpoonright \mathcal{F}_n)$. Posons $U = \limsup_{n\to\infty} U_n$ et $V = \limsup_{n\to\infty} V_n$. Alors $\langle U_n \rangle_{n\in\mathbf{N}}$ et $\langle V_n \rangle_{n\in\mathbf{N}}$ convergent respectivement vers U et V ρ-p.s.. De plus, $U_n + V_n = C$ ρ-p.s., donc $\rho(\{U = V = 0\}) = 0$. Par conséquent :

$$\frac{U}{V} = \frac{\lim U_n}{\lim V_n} = \lim_{n\to\infty} \frac{U_n}{V_n} = \lim_{n\to\infty} X_n = X \quad \rho\text{-p.s..}$$

On déduit alors (13) du résultat suivant : pour tout $A \in \mathcal{F}$,

$$\mu(A) = \int_A U d\rho = \int_A XV d\rho + \int_A 1_{\{V=0\}} U d\rho.$$

Proposition 2.2 *Soit $X, X_1, X_2, \ldots$ des variables aléatoires indépendantes et de même loi. Alors p.s. :*

$$\limsup_{n\to\infty} \frac{X_n}{n} = \begin{cases} 0 & \text{si } \mathbf{E}[X] < \infty, \\ \infty & \text{si } \mathbf{E}[X] = \infty. \end{cases}$$

Preuve.

1. Supposons tout d'abord que $\mathbf{E}[X] < \infty$. Soit $\epsilon > 0$. Posons $A_n = \{\frac{X_n}{n} \geq \epsilon\}$. Alors $\mathbf{P}(A_n) = \mathbf{P}(X \geq n\epsilon)$, donc

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \mathbf{E}\Big[\sum_{n=1}^{\infty} \mathbf{1}_{\{n \leq \frac{X}{\epsilon}\}}\Big] \leq \mathbf{E}\Big[\frac{X}{\epsilon}\Big] < \infty.$$

   Par le lemme de Borel-Cantelli, on a $\mathbf{P}\big(\limsup_{n\to\infty}(\{\frac{X_n}{n} \geq \epsilon\})\big) = 0$, donc : $\mathbf{P}(\limsup_{n\to\infty} \frac{X_n}{n} < \epsilon) = 1$. D'où :

$$\mathbf{P}\Big(\bigcap_{k\geq 0}\{\limsup_{n\to\infty} \frac{X_n}{n} < 2^{-k}\}\Big) = 1, \text{ i.e. p.s. } \limsup_{n\to\infty} \frac{X_n}{n} = 0.$$

2. Si $\mathbf{E}[X] = \infty$. Soit $A > 0$. En posant $B_n = \{\frac{X_n}{n} \geq A\}$, $\mathbf{P}(B_n) = \mathbf{P}(X \geq An)$, donc

$$\sum_{n=1}^{\infty} \mathbf{P}(B_n) = \mathbf{E}\Big[\sum_{n=1}^{\infty} \mathbf{1}_{\{n \leq \frac{X}{A}\}}\Big] \geq \mathbf{E}\Big[\frac{X}{A} - 1\Big] = \infty.$$

   Comme les événements B_n sont indépendants, le lemme de Borel-Cantelli dans le cas d'indépendance assure que $\mathbf{P}\big(\limsup_{n\to\infty}(\{\frac{X_n}{n} \geq A\})\big) = 1$. Par conséquent, $\mathbf{P}(\limsup_{n\to\infty} \frac{X_n}{n} \geq A) = 1$, et ce pour tout $A > 0$. Ainsi

$$\mathbf{P}\Big(\bigcap_{k>0}\big\{\limsup_{n\to\infty} \frac{X_n}{n} \geq k\big\}\Big) = 1, \text{ i.e. p.s. } \limsup_{n\to\infty} \frac{X_n}{n} = \infty.$$

Proposition 2.3 *Soit $X, X_1, X_2, \ldots$ des variables aléatoires indépendantes et de même loi. Alors, pour tout $c \in]0,1[$, p.s. :*

$$\sum_{n=1}^{\infty} e^{X_n} c^n < \infty \quad \text{si } \mathbf{E}[X] < \infty, \qquad \sum_{n=1}^{\infty} e^{X_n} c^n = \infty \quad \text{si } \mathbf{E}[X] = \infty.$$

Preuve.

1. Si $\mathbf{E}[X] < \infty$, alors, p.s., $\limsup_{n\to\infty} \frac{X_n}{n} = 0$, donc, en posant $\alpha = -\frac{\ln c}{2}$, il existe $n_0(\omega)$ tel que pour tout $n \geq n_0$, $\frac{X_n}{n} \leq \alpha$, donc $X_n \leq \alpha n$, et $e^{X_n} c^n \leq (e^{\alpha + \ln c})^n$, avec $\alpha + \ln c < 0$. D'où $\sum_{n=1}^{\infty} e^{X_n} c^n < \infty$ p.s.
2. Supposons $\mathbf{E}[X] = \infty$. Alors, p.s., $\limsup_{n\to\infty} \frac{X_n}{n} = \infty$, donc il existe une sous-suite $\big(X_{n_k}(\omega)\big)_{k\in\mathbf{N}}$ telle que $\lim_{k\to\infty} \frac{X_{n_k}}{n_k} = \infty$, donc

$$\exists k_0(\omega), \forall k \geq k_0, \frac{X_{n_k}}{n_k} \geq -\ln c, \text{ et } e^{X_{n_k}} c^{n_k} \geq 1.$$

   D'où $\sum_{n=1}^{\infty} e^{X_n} c^n = \infty$ p.s.

2.2 Processus d'immigration

Théorème 2.4 *Soit Z_n la population d'un processus de Galton-Watson avec immigration Y_n. Soit Y une variable aléatoire de même loi que les Y_n. On suppose que $m > 1$. Alors si $\mathbf{E}[\log^+ Y] < \infty$, $\langle \frac{Z_n}{m^n} \rangle_{n \in \mathbf{N}^*}$ converge p.s., tandis que si $\mathbf{E}[\log^+ Y] = \infty$ alors, pour tout $c > 0$, $\limsup_{n \to \infty} \frac{Z_n}{c^n} = \infty$ p.s..*

Preuve. Supposons que $\mathbf{E}[\log^+ Y] = \infty$. Par la proposition 2.2, il existe p.s. une sous-suite $(n_k)_{k \in \mathbf{N}}$ telle que $\lim \log(Y_{n_k}^{1/n_k}) = \infty$, donc $\lim \frac{Y_{n_k}^{1/n_k}}{c} = \infty$, et donc $\lim \frac{Y_{n_k}}{c^{n_k}} = \infty$. Comme $Z_n \geq Y_n$, on a le résultat.

Supposons dorénavant que $\mathbf{E}[\log^+ Y] < \infty$. Soit $\mathcal{G}_n = \sigma(Z_1, \ldots, Z_n)$ et $Z_{n,k}$ le nombre de descendants au niveau n des particules immigrées à l'instant k. Le nombre total de sommets à l'instant n est $\sum_{k=1}^{n} Z_{n,k}$. Notons $Q_{(y_k)}$ la loi conditionnelle de la suite $\langle Z_n \rangle_{n \geq 1}$ sachant que $Y_k = y_k$ pour tout $k \geq 1$ (il est aisé de la construire directement puisqu'il s'agit de la probabilité correspondant à un processus de Galton-Watson avec immigration de y_k particules à la génération k pour tout $k \geq 1$). Supposons que $\sum_{k=1}^{\infty} \frac{y_k}{m^k} < \infty$. Alors, sous la probabilité $Q_{(y_k)}$, $\langle \frac{Z_n}{m^n} \rangle_{n \geq 1}$ est une sous-martingale relativement à la filtration $\mathcal{G}_n$. En effet, comme $Z_{n,k}$ représente la $(n-k)$-ème génération d'un processus de Galton-Watson classique débutant néanmoins avec y_k particules, il vient :

$$\begin{aligned}
\mathbf{E}\Big[\frac{Z_{n+1}}{m^{n+1}} | \mathcal{G}_n\Big] &= \sum_{k=1}^{n+1} \frac{1}{m^k} \mathbf{E}\Big[\frac{Z_{n+1,k}}{m^{n+1-k}} | \mathcal{G}_n\Big] \geq \sum_{k=1}^{n} \frac{1}{m^k} \mathbf{E}\Big[\frac{Z_{n+1,k}}{m^{n+1-k}} | \mathcal{G}_n\Big], \\
\mathbf{E}\Big[\frac{Z_{n+1}}{m^{n+1}} | \mathcal{G}_n\Big] &\geq \sum_{k=1}^{n} \frac{1}{m^k} \frac{Z_{n,k}}{m^{n-k}} = \frac{Z_n}{m^n}.
\end{aligned}$$

Sous $Q_{(y_k)}$, cette sous-martingale est bornée dans L^1 puisque

$$\mathbf{E}\Big[\frac{Z_n}{m^n}\Big] = \mathbf{E}\Big[\frac{1}{m^n} \sum_{k=1}^{n} Z_{n,k}\Big] = \sum_{k=1}^{n} \frac{1}{m^k} \mathbf{E}\Big[\frac{Z_{n,k}}{m^{n-k}}\Big] = \sum_{k=1}^{n} \frac{y_k}{m^k} \leq \sum_{k=1}^{\infty} \frac{y_k}{m^k} < \infty.$$

Finalement, lorsque $\sum_{k=1}^{\infty} \frac{y_k}{m^k} < \infty$, $\langle \frac{Z_n}{m^n} \rangle_{n \geq 1}$ converge $Q_{(y_k)}$-p.s. On sait de plus par la proposition 2.3 (par hypothèse $\mathbf{E}[\log^+ Y] < \infty$ et $m > 1$) que $\sum_{k=1}^{\infty} e^{\log^+ Y_k} \big(\frac{1}{m}\big)^k < \infty$ p.s., donc $\sum_{k=1}^{\infty} \frac{Y_k}{m^k} < \infty$ p.s. En notant $C = \{\langle \frac{Z_n}{m^n} \rangle_{n \geq 1} \text{ converge}\}$, on obtient ainsi :

$$\mathbf{P}(C) = \mathbf{E}[\mathbf{P}(C | \langle Y_k \rangle_{k \geq 1})] = \mathbf{E}[Q_{(Y_k)}(C)] = 1.$$

Interprétation. Le choix de la base du logarithme ne changeant $\mathbf{E}[\log^+ Y]$ qu'à une constante multiplicative près, il n'a aucune importance dans l'énoncé du théorème. Pour comprendre le sens de ce théorème, il peut alors être intéressant de prendre ces logarithmes en base m. En effet, si on considère un arbre de Galton-Watson classique, l'espérance de Z_n est m^n, et $\log_m(Z_n) \simeq n$: ce logarithme permet d'évaluer l'âge de la population. Aussi l'immigration Y peut-elle être considérée comme l'arrivée d'individus issus d'une population qui se développe depuis $\log_m(Y)$ générations. De ce point de vue, $\mathbf{E}[\log^+ Y] < \infty$ signifie que les populations immigrantes ont un âge moyen fini : lorsque n sera devenu grand devant $\mathbf{E}[\log^+ Y]$, le nombre de nouveaux arrivants sera négligeable devant la population déjà en place et l'immigration ne perturbera presque plus le développement de la population. En revanche, si $\mathbf{E}[\log^+ Y] = \infty$, l'âge moyen des populations immigrantes est infini. L'immigration est en permanence en mesure de déborder la population en place, et changer m en c, c'est-à-dire modifier la natalité de la population et subir une transition démographique, ne permet pas d'y échapper.

2.3 Théorème de Kesten-Stigum

Théorème 2.5 *Supposons que* $1 < m < \infty$. *Alors :*

$$\mathbf{P}(W = 0) = q \qquad \Longleftrightarrow \qquad \mathbf{E}[W] = 1 \qquad \Longleftrightarrow \qquad \mathbf{E}[L \log^+ L] < \infty.$$

Ces assertions sont aussi équivalentes à la convergence dans L^1 *de la suite* $\langle Z_n \rangle$ *vers* W.

Preuve. Soit $\mathcal{F}_n$ la tribu engendrée par les n premiers niveaux des arbres. Alors $\langle \mathcal{F}_n \rangle$ est une filtration, et (10) s'écrit aussi :

$$\frac{d(\widehat{\mathbf{GW}} \upharpoonright \mathcal{F}_n)}{d(\mathbf{GW} \upharpoonright \mathcal{F}_n)} = W_n. \tag{14}$$

Afin de définir W pour tout arbre $a \in \mathcal{A}$, on pose $W(a) = \limsup_{n \to \infty} W_n(a)$ si a est infini, $W(a) = 0$ sinon. De (14) et du lemme 2.1, on a le résultat fondamental suivant :

$$\int W d\mathbf{GW} = 1 \quad \Longleftrightarrow \quad \widehat{\mathbf{GW}} \ll \mathbf{GW} \quad \Longleftrightarrow \quad W < \infty \quad \widehat{\mathbf{GW}}\text{-p.s.}, \tag{15}$$

$$W = 0 \quad \mathbf{GW}\text{-p.s.} \quad \Longleftrightarrow \quad \widehat{\mathbf{GW}} \perp \mathbf{GW} \quad \Longleftrightarrow \quad W = \infty \quad \widehat{\mathbf{GW}}\text{-p.s.}. \tag{16}$$

Par conséquent, si $\mathbf{E}[L \log^+ L] < \infty$, i.e. si $\mathbf{E}[\log^+ \widehat{L}] < \infty$, alors $\mathbf{E}[\log^+(\widehat{L} - 1)] < \infty$, donc, par le théorème 2.4, $\frac{Z_n - 1}{m^n}$ converge $\widehat{\mathbf{GW}}$-p.s., i.e., comme $m > 1$, $\frac{Z_n}{m^n}$ converge $\widehat{\mathbf{GW}}$-p.s., donc $W < \infty$ $\widehat{\mathbf{GW}}$-p.s.. Ainsi, par (15), $\int W d\mathbf{GW} = 1$. En revanche, si $\mathbf{E}[L \log^+ L] = \infty$, alors $\mathbf{E}[\log^+(\widehat{L} - 1)] = \infty$, donc, par le théorème 2.4, $\limsup \frac{Z_n - 1}{m^n} = \infty$ $\widehat{\mathbf{GW}}$-p.s., i.e., comme $m > 1$, $W = \infty$ $\widehat{\mathbf{GW}}$-p.s., et $W = 0$ $\mathbf{GW}$-p.s. par (16).

Supposons maintenant que $\langle Z_n \rangle$ converge vers W dans L^1. On a alors convergence de la suite $\langle \mathbf{E}[Z_n] \rangle$ vers $\mathbf{E}[W]$, donc $\mathbf{E}[W] = 1$. Réciproquement, on suppose que $\mathbf{E}[W] = 1$. On écrit $|Z_n - W| = Z_n + W - Z_n \wedge W$. Comme $|Z_n \wedge W| \leq |W|$, le théorème de convergence dominée assure que $\langle \mathbf{E}[Z_n \wedge W] \rangle$ converge vers 1. Aussi la suite $\langle \mathbf{E}[|Z_n - W|] \rangle$ converge-t-elle vers 0, ce qui achève la preuve.

3 Autres cas

3.1 Cas sous-critique

Définition 4 *Soit $\langle \nu_n \rangle$ une suite de mesures de probabilité sur les entiers naturels. On dit que $\{\nu_n\}$ est tendue si pour tout $\epsilon > 0$, il existe $k \in \mathbf{N}$ tel que pour tout $n \in \mathbf{N}$, $\nu_n(\{0, \ldots, k\}) \geq 1 - \epsilon$.*

Définition 5 *Soit $\langle \nu_n \rangle_{n \in \mathbf{N}}$ une suite de mesures de probabilités sur les réels (ou sur les entiers naturels). On dit que $\langle \nu_n \rangle_{n \in \mathbf{N}}$ tend en probabilité vers l'infini si, pour tout $A > 0$, $\lim_{n \to \infty} \nu_n(]A, \infty]) = 1$.*

Lemme 3.1 *Soit $\langle \nu_n \rangle$ une suite de mesures de probabilité sur les entiers naturels de moyenne respective a_n. Soit $\widehat{\nu}_n$ la mesure biaisée de ν_n, i.e. $\widehat{\nu}_n(k) = k\nu_n(k)/a_n$. Si $\{\widehat{\nu}_n\}$ est tendue, alors $\sup a_n < \infty$, tandis que si $\langle \widehat{\nu}_n \rangle_{n \in \mathbf{N}}$ tend vers l'infini en probabilité, alors la suite $\langle a_n \rangle$ n'est pas bornée.*

Preuve.

1. Si $\{\widehat{\nu}_n\}$ est tendue, il existe $k \in \mathbf{N}$ tel que pour tout $n \in \mathbf{N}$, $\widehat{\nu}_n(\{0, \ldots, k\}) \geq \frac{1}{2}$, i.e.

$$\frac{1}{a_n} \sum_{i=0}^{k} i\nu_n(i) \geq \frac{1}{2},$$

   donc $a_n \leq 2\sum_{i=0}^{k} i\nu_n(i)$, donc $a_n \leq k(k+1)$. Ainsi, $\sup a_n < \infty$.

2. Supposons maintenant que $\langle \widehat{\nu}_n \rangle_{n \in \mathbf{N}}$ tende vers l'infini en probabilité. Il existe une suite d'entiers naturels $\langle n_k \rangle_{k>1}$ strictement croissante telle que

$$\forall k > 1, \quad \widehat{\nu}_{n_k}(\{1, \ldots, k-1\}) \leq 2^{-k-1}, \text{ i.e. } \quad 2^{k+1} \sum_{i=1}^{k-1} i\nu_{n_k}(i) \leq a_{n_k}.$$

   S'il y a un nombre infini d'entiers $k' > 1$ tels qu'il existe $i \in \{1, \ldots, k'-1\}$ tel que $\nu_{n_{k'}}(i) \geq \frac{1}{2k'}$, alors $2^{k'}/k' \leq a_{n_{k'}}$, donc $\langle a_n \rangle$ n'est pas bornée. Sinon, il existe un nombre infini d'entiers $k' > 1$ tels que pour tout $i \in \{1, \ldots, k'-1\}$, $\nu_{n_{k'}}(i) < \frac{1}{2k'}$, donc $\nu_{n_{k'}}(\{1, \ldots, k'-1\}) < \frac{1}{2}$, puis $\nu_{n_{k'}}([k', \infty[) \geq \frac{1}{2}$. Par l'inégalité de Markov, on en déduit que $a_{n_{k'}} \geq \frac{k'}{2}$, donc $\langle a_n \rangle$ n'est pas bornée.

Lemme 3.2 *Soit $\langle \nu_n \rangle$ une suite de mesures de probabilité sur les entiers naturels. Si $\langle \nu_n \rangle$ converge étroitement vers une mesure de probabilité ν, alors $\{\nu_n\}$ est tendue.*

Preuve. Soit $\epsilon > 0$. Il existe $k_1 \in \mathbf{N}$ tel que $\nu(\{0, \ldots, k_1\}) \geq 1 - \frac{\epsilon}{2}$. Comme $\langle \nu_n \rangle$ converge étroitement vers ν, $\lim_{n \to \infty} \nu_n(\{0, \ldots, k_1\}) = \nu(\{0, \ldots, k_1\})$, donc il existe $n_0 \in \mathbf{N}$ tel que pour tout $n \geq n_0$, $\nu_n(\{0, \ldots, k_1\}) \geq 1 - \epsilon$. Il existe de plus $k_2 \geq k_1$ tel que pour tout $i \in \{0, \ldots, n_0 - 1\}$, $\nu_i(\{0, \ldots, k_2\}) \geq 1 - \epsilon$. Finalement, $\nu_n(\{0, \ldots, k_2\}) \geq 1 - \epsilon$ pour tout $n \in \mathbf{N}$: $\{\nu_n\}$ est tendue.

Théorème 3.3 *Soit Z_n la population d'un processus de Galton-Watson avec immigration Y_n. Soit Y une variable aléatoire de même loi que les Y_n. On suppose que $m < 1$. Si $\mathbf{E}[\log^+ Y] < \infty$, $(Z_n)_{n\in\mathbf{N}^*}$ converge en loi vers une certaine variable aléatoire finie p.s., tandis que si $\mathbf{E}[\log^+ Y] = \infty$, alors $(Z_n)_{n\in\mathbf{N}^*}$ converge en probabilité vers l'infini.*

Preuve. Soit $\mathcal{G}$ la tribu engendrée par les Y_n. Soit $Z_{n,k}$ le nombre de descendants au niveau n des particules immigrées à l'instant k. Ainsi, à l'instant n, le nombre total de sommets est $\sum_{k=1}^n Z_{n,k}$. Les v.a. $Z_{n,k}$ pour n fixé sont indépendantes et la loi de $Z_{n,k}$ ne dépend que de $n-k$. De plus, pour tous $i_1, \ldots, i_n$ tels que $i_k \geq k$, les variables aléatoires $Z_{i_1,1}, \ldots, Z_{i_n,n}$ sont indépendantes. Donc Z_n a même loi que $T_n = \sum_{k=1}^n Z_{2k-1,k}$. $(T_n)_{n\in\mathbf{N}^*}$ est une suite croissante. Notons Z_∞ sa limite. Prouvons que Z_∞ est finie p.s. ou infinie p.s.. Notons $\mathcal{K}_n = \sigma(Z_{2k-1,k};\ k \geq n)$. Comme pour tout k, $Z_{2k-1,k}$ est finie p.s., on a : $Z_\infty = \infty$ si et seulement si pour tout n, $\lim_{k\to\infty} \sum_{j=n}^k Z_{2j-1,j} = \infty$, donc l'événement $\{Z_\infty = \infty\}$ appartient à $\mathcal{K}_n$ pour tout n. Or, les variables $Z_{2k-1,k}$ sont indépendantes. On en déduit par la loi du tout ou rien que $\mathbf{P}(Z_\infty = \infty) \in \{0, 1\}$, i.e. Z_∞ est finie p.s. ou infinie p.s.. Nous allons montrer que $Z_\infty < \infty$ p.s. si et seulement si $\mathbf{E}[\log^+ Y] < \infty$.

Supposons que $\mathbf{E}[\log^+ Y] < \infty$. On a :

$$\mathbf{E}[Z_\infty|\mathcal{G}] = \sum_{k=1}^\infty m^{k-1}\mathbf{E}\Big[\frac{Z_{2k-1,k}}{m^{k-1}}|\mathcal{G}\Big] = \sum_{k=1}^\infty Y_k m^{k-1} \leq \sum_{k=1}^\infty e^{\log^+ Y_k} m^{k-1}.$$

Comme ($m < 1$), la proposition 2.3 assure que la suite $\langle \sum_{k=1}^n e^{\log^+ Y_k} m^{k-1}\rangle$ converge p.s. : Z_∞ est finie p.s..

Supposons $Z_\infty < \infty$ p.s.. Écrivons $Z_{2k-1,k} = \sum_{i=1}^{Y_k} \zeta_k(i)$, où $\zeta_k(i)$ désigne la population de la $(k-1)$-ème génération d'un processus de Galton-Watson ordinaire, les $\zeta_k(i)$ étant indépendantes et identiquement distribuées. Par conséquent, $Z_\infty = \sum_{k=1}^\infty \sum_{i=1}^{Y_k} \zeta_k(i) < \infty$ p.s., donc seul un nombre fini de $\zeta_k(i)$ sont non nulles. Quitte à conditionner par la suite $\langle Y_k\rangle_{k\geq 1}$, le lemme de Borel-Cantelli dans le cas d'indépendance assure alors que $\sum_k \sum_{i\leq Y_k} \mathbf{P}(\zeta_k(i) \geq 1) < \infty$, et comme $\zeta_k(i)$ a même loi que $\widetilde{Z}_{k-1}$ (où $\widetilde{Z}_i$ est la i-ème génération d'un processus de Galton-Watson classique), $\sum_{k=1}^\infty Y_k \mathbf{GW}(\widetilde{Z}_{k-1} \geq 1) < \infty$ p.s. Or $\mathbf{GW}(\widetilde{Z}_{k-1} \geq 1) \geq \mathbf{P}(L > 0)^{k-1}$, donc : $\sum_{k=1}^\infty e^{\log Y_k}\mathbf{P}(L > 0)^k < \infty$ p.s. Par la proposition 2.3, il vient : $\mathbf{E}[\log^+ Y] < \infty$.

Par conséquent, si $\mathbf{E}[\log^+ Y] < \infty$, i.e. si $Z_\infty < \infty$ p.s., alors, puisque $(T_n)_{n\in\mathbf{N}^*}$ converge presque sûrement (et donc en loi) vers Z_∞ et que Z_n et T_n suivent la même loi, $(Z_n)_{n\in\mathbf{N}^*}$ converge en loi vers Z_∞. Si $\mathbf{E}[\log^+ Y] < \infty$, i.e. si $Z_\infty = \infty$ p.s., alors $(Z_n)_{n\in\mathbf{N}^*}$ converge en probabilité vers l'infini. En effet, pour tout $A > 0$, $\mathbf{P}(Z_n > A) = \mathbf{E}[\mathbf{1}_{\{T_n > A\}}]$, donc par le théorème de convergence dominée, $\lim_{n\to\infty} \mathbf{P}(Z_n > A) = 1$.

Théorème 3.4 *Pour tout processus de Galton-Watson tel que $0 < m < \infty$, la suite $\langle \mathbf{P}(Z_n > 0)/m^n\rangle$ décroît et la suite $\langle \mathbf{E}[Z_n|Z_n > 0]\rangle$ croît. De plus :*

$$\lim_{n\to\infty} \frac{\mathbf{P}(Z_n > 0)}{m^n} > 0 \quad\Longleftrightarrow\quad \sup_{n\in\mathbf{N}} \mathbf{E}[Z_n|Z_n > 0] < \infty.$$

Si $m < 1$, ces deux affirmations sont aussi équivalentes à $\mathbf{E}[L\log^+ L] < \infty$.

Preuve. Pour tout arbre $a \in \mathcal{A}$ tel que $Z_n > 0$, notons $\xi_n(a)$ l'enfant le plus à gauche de la racine qui a au moins un descendant à la génération n. Si $Z_n > 0$, désignons par $H_n(a)$ le nombre de descendants de $\xi_n(a)$ à la génération n ; sinon, on pose $H_n(a) = 0$. On a alors pour tout enfant x de la racine :

$$\mathbf{P}(H_n = k|Z_n > 0) = \mathbf{P}(H_n = k|Z_n > 0, \xi_n = x) = \mathbf{P}(Z_{n-1} = k|Z_{n-1} > 0).$$

Comme $H_n \leq Z_n$, cela montre que $\langle \mathbf{E}[Z_n|Z_n > 0] \rangle$ croît. On a aussi :

$$\frac{\mathbf{P}(Z_n > 0)}{m^n} = \frac{\mathbf{P}(Z_n > 0)}{\mathbf{E}[Z_n \mathbf{1}_{\{Z_n>0\}}]} = \frac{1}{\mathbf{E}[Z_n|Z_n > 0]}.$$

Par conséquent, la suite $\langle \frac{\mathbf{P}(Z_n>0)}{m^n} \rangle$ décroît et sa limite est non nulle si et seulement si $\sup_{n\in\mathbf{N}} \mathbf{E}[Z_n|Z_n > 0] < \infty$.

Supposons que $m < 1$. Soit μ_n la loi de Z_n sous $\mathbf{P}(\ .\ |Z_n > 0)$. La variable aléatoire $\widehat{Z}_n$ a même loi que $\widehat{\mu}_n$ puisque :

$$\widehat{\mu}_n(k) = \frac{k\mu_n(k)}{\sum_{j=1}^{\infty} j\mu_n(j)} = \frac{k\frac{\mathbf{P}(Z_n=k)}{\mathbf{P}(Z_n>0)}}{\sum_{j=1}^{\infty} j\frac{\mathbf{P}(Z_n=j)}{\mathbf{P}(Z_n>0)}} = \frac{k\mathbf{P}(Z_n = k)}{\sum_{j=1}^{\infty} j\mathbf{P}(Z_n = j)}.$$

D'après (11), $\langle \widehat{\mu}_n \rangle$ décrit l'évolution d'un processus de Galton-Watson avec immigration $\widehat{L}-1$. Supposons que $\sup_{n\in\mathbf{N}} \mathbf{E}[Z_n|Z_n > 0] < \infty$, c'est-à-dire que les moyennes des μ_n sont bornées. Alors, par le lemme 3.1, les lois $\widehat{\mu}_n$ ne tendent pas en probabilité vers l'infini. Par le théorème 3.3, $\mathbf{E}[L \log^+ L] < \infty$. Réciproquement, si $\mathbf{E}[L \log^+ L] < \infty$, alors, d'après le théorème 3.3, $\langle \widehat{\mu}_n \rangle$ converge étroitement, donc est tendue par le lemme 3.2. Le lemme 3.1 assure alors que $\sup_{n\in\mathbf{N}} \mathbf{E}[Z_n|Z_n > 0] < \infty$.

3.2 Cas critique

Définition 6 *On dit qu'une variable aléatoire X suit une loi géométrique de paramètre $p \in [0, 1[$ lorsque pour tout $k \geq 1$, $\mathbf{P}(X = k) = (1 - p)p^{k-1}$.*

Lemme 3.5 *Soit A une variable aléatoire indépendante des variables aléatoires B et C définies sur $\mathbf{N}^*$. Supposons que la fonction $x \mapsto \mathbf{P}(C \leq x)/\mathbf{P}(B \leq x)$ est croissante, que $\mathbf{P}(A \geq B) > 0$ et que $\mathbf{P}(A \geq C) > 0$. Alors la loi de A sachant $A \geq B$ est stochastiquement dominée par la loi de A sachant $A \geq C$, et l'hypothèse sur B et C est satisfaite si ces variables suivent une loi géométrique de paramètres respectifs b et c avec $b \leq c$.*

Preuve. Nous devons montrer que $\mathbf{P}(A > x|A \geq B) \leq \mathbf{P}(A > x|A \geq C)$ pour tout x. Soit F la fonction de répartition de A : $F(b) = \mathbf{P}(A \leq b)$. Elle est croissante et continue à droite. Notons μ_F la mesure de Stiljes associée : $\mu_F([a, b]) = F(b)-F(a^-) = \mathbf{P}(a \leq A \leq b)$. On a : $\{A \geq B\} = \bigcup_{b=1}^{\infty} \{A \geq b\} \cap \{B = b\}$. Par indépendance de A et B, on en déduit que

$\mathbf{P}(A \geq B) = \sum_{b=1}^{\infty} \mathbf{P}(A \geq b)\mathbf{P}(B = b)$. Or, $\mathbf{P}(A \geq b) = \mu_F([b,\infty[) = \int_{y \in \mathbf{R}} \mathbf{1}_{\{b \leq y\}} \mu_F(dy)$. On en déduit par convergence monotone et par le théorème de Fubini-Tonelli que :

$$\begin{aligned}
\mathbf{P}(A \geq B) &= \sum_{b=1}^{\infty} \int_{y \in \mathbf{R}} \mathbf{1}_{\{b \leq y\}} \mathbf{P}(B = b) \mu_F(dy) \\
&= \int_{y \in \mathbf{R}} \sum_{b=1}^{\infty} \mathbf{1}_{\{b \leq y\}} \mathbf{E}[\mathbf{1}_{\{B=b\}}] \mu_F(dy) \\
&= \int_{y \in \mathbf{R}} \mathbf{E}\Big[\sum_{b=1}^{\infty} \mathbf{1}_{\{b \leq y\} \cap \{B=b\}}\Big] \mu_F(dy).
\end{aligned}$$

Or, $\{B \leq y\} = \bigcup_{b=1}^{\infty} \{b \leq y\} \cap \{B = b\}$. D'où $\sum_{b=1}^{\infty} \mathbf{1}_{\{b \leq y\} \cap \{B=b\}} = \mathbf{1}_{\{B \leq y\}}$. Finalement, $\mathbf{P}(A \geq B) = \int_{y \in \mathbf{R}} \mathbf{P}(B \leq y) \mu_F(dy)$. On montre de même que : $\mathbf{P}(A > x, A \geq B) = \int_{y > x} \mathbf{P}(B \leq y) \mu_F(dy)$. Par conséquent, pour tout $x \in \mathbf{R}$, on a :

$$\begin{aligned}
\mathbf{P}(A > x | A \geq B) &= \frac{\mathbf{P}(A > x, A \geq B)}{\mathbf{P}(A \geq B)} = \frac{\int_{y > x} \mathbf{P}(B \leq y) \mu_F(dy)}{\int_{y \in \mathbf{R}} \mathbf{P}(B \leq y) \mu_F(dy)} \\
&= \left(1 + \frac{\int_{y \leq x} \mathbf{P}(B \leq y) \mu_F(dy)}{\int_{y > x} \mathbf{P}(B \leq y) \mu_F(dy)}\right)^{-1} \\
&\leq \left(1 + \frac{\int_{y \leq x} \mathbf{P}(C \leq y) \frac{\mathbf{P}(B \leq x)}{\mathbf{P}(C \leq x)} \mu_F(dy)}{\int_{y > x} \mathbf{P}(C \leq y) \frac{\mathbf{P}(B \leq x)}{\mathbf{P}(C \leq x)} \mu_F(dy)}\right)^{-1} \\
&= \left(1 + \frac{\int_{y \leq x} \mathbf{P}(C \leq y) \mu_F(dy)}{\int_{y > x} \mathbf{P}(C \leq y) \mu_F(dy)}\right)^{-1} \\
&= \mathbf{P}(A > x | A \geq C).
\end{aligned}$$

Montrons que l'hypothèse est satisfaite dans le cas géométrique, i.e. que :

$$\forall k \in \mathbf{N}, \qquad \frac{1 - c^k}{1 - b^k} \leq \frac{1 - c^{k+1}}{1 - b^{k+1}}.$$

Il suffit d'établir la croissance sur l'intervalle [0,1[de la fonction $f : x \mapsto \frac{1-x^{k+1}}{1-x^k}$. On a :

$$\begin{aligned}
\frac{f'(x)}{f(x)} &= k\frac{x^{k-1}}{1 - x^k} - (k+1)\frac{x^k}{1 - x^{k+1}} \\
&= \frac{x^{k-1}}{(1 - x^k)(1 - x^{k+1})}\big(k(1 - x^{k+1}) - (k+1)x(1 - x^k)\big).
\end{aligned}$$

Il s'agit donc de prouver que la fonction g définie sur [0,1] par

$$g(x) = k + x^{k+1} - (k+1)x$$

est positive. Puisqu'elle décroît et que $g(1) = 0$, on a le résultat.

Théorème 3.6 *Si $m = 1$ et $p_1 < 1$, et en posant $\sigma^2 = \mathrm{Var}(L) = \mathbf{E}[L^2] - 1 \leq \infty$, alors :*

$$\lim_{n\to\infty} n\mathbf{P}(Z_n > 0) = \frac{2}{\sigma^2}. \tag{17}$$

Preuve. Soit D_n une variable aléatoire indépendante de Z_1 qui suit une loi géométrique de paramètre $1 - \mathbf{P}(Z_{n-1} > 0)$. Alors son espérance vaut $\frac{1}{\mathbf{P}(Z_{n-1}>0)}$ et pour tout $k \geq 1$, $\mathbf{P}(D_n \leq k) = 1 - (1 - \mathbf{P}(Z_{n-1} > 0))^k$. Ainsi :

$$\begin{aligned}
\mathbf{P}(Z_1 = k \text{ et } Z_n > 0) &= \mathbf{P}(Z_1 = k)\big(1 - (1 - \mathbf{P}(Z_{n-1} > 0))^k\big) \\
&= \mathbf{P}(Z_1 = k)\mathbf{P}(D_n \leq k) \\
&= \mathbf{P}(Z_1 = k \text{ et } D_n \leq Z_1),
\end{aligned}$$

puis $\mathbf{P}(Z_n > 0) = \mathbf{P}(D_n \leq Z_1)$. Par conséquent, Z_1 sachant que $Z_n > 0$ et Z_1 sachant que $Z_1 \geq D_n$ suivent la même loi. De plus, le paramètre de D_{n-1} est inférieur ou égal à celui de D_n. Par le lemme 3.5, on en déduit que la loi conditionnelle de Z_1 sachant $Z_n > 0$ croît stochastiquement avec n. Soit Y_n le nombre d'enfants de la racine qui ont un descendant à la génération n. Puisque $\lim_{n\to\infty} \mathbf{P}(Z_n > 0) = 0$, on a : $\lim_{n\to\infty} \mathbf{P}(Y_n = 1 | Z_n > 0) = 1$. De plus, pour tout $k \geq 1$,

$$\begin{aligned}
\mathbf{P}(Z_1 = k | Z_n > 0) &= \frac{\mathbf{P}(Z_1 = k \text{ et } Z_n > 0)}{\mathbf{P}(Z_n > 0)} \\
&= \frac{\mathbf{P}(Z_1 = k)\big(1 - (1 - \mathbf{P}(Z_{n-1} > 0))^k\big)}{\mathbf{P}(Z_n > 0)},
\end{aligned}$$

donc

$$\mathbf{P}(Z_1 = k | Z_n > 0) \sim_{n\to\infty} k\mathbf{P}(Z_1 = k)\frac{\mathbf{P}(Z_{n-1} > 0)}{\mathbf{P}(Z_n > 0)}.$$

Montrons maintenant que $\frac{\mathbf{P}(Z_{n-1}>0)}{\mathbf{P}(Z_n>0)} = 1$. Posons $v_n = \mathbf{P}(Z_n > 0)$. On a :

$$1 - v_n = \mathbf{E}[\mathbf{E}[\mathbf{1}_{\{Z_n=0\}} | Z_1]] = \mathbf{E}[\mathbf{P}(Z_{n-1} = 0)^{Z_1}] = f(\mathbf{P}(Z_{n-1} = 0)),$$

i.e. $v_n = 1 - f(1 - v_{n-1})$. Or, comme $m = 1$, $\lim_{n\to\infty} v_n = 0$ et $1 - f(1 - x) \sim_{x\to 0} x$. Par conséquent, $1 - f(1 - v_{n-1}) \sim_{n\to\infty} v_{n-1}$, puis $v_n \sim_{n\to\infty} v_{n-1}$. On obtient finalement que $\lim_{n\to\infty} \mathbf{P}(Z_1 = k | Z_n > 0) = k\mathbf{P}(Z_1 = k)$. De $\lim_{n\to\infty} \mathbf{P}(Z_n > 0) = 0$, on déduit aussi que la distribution de l'enfant de la racine le plus à gauche ayant un descendant à la génération n sachant que $Z_n > 0$ tend vers une loi uniforme parmi les particules de la première génération. Puisque, pour tout $k \in \mathbf{N}^*$, $\mathbf{P}(Z_1 \geq k | Z_n > 0)$ croît avec n, il vient aussi, par convergence monotone :

$$\lim_{n\to\infty} \uparrow \mathbf{E}[Z_1 | Z_n > 0] = \lim_{n\to\infty} \uparrow \sum_{k=1}^{\infty} \mathbf{P}(Z_1 \geq k | Z_n > 0) = \sum_{k=1}^{\infty} \lim_{n\to\infty} \uparrow \mathbf{P}(Z_1 \geq k | Z_n > 0).$$

Or, toujours par convergence monotone, on a, pour tout $k \geq 1$:

$$\begin{aligned}
\lim_{n\to\infty} \uparrow \mathbf{P}(Z_1 \geq k | Z_n > 0) &= \lim_{n\to\infty} \uparrow \sum_{j=k}^{\infty} \mathbf{P}(Z_1 = j | Z_n > 0) \\
&= \sum_{j=k}^{\infty} \lim_{n\to\infty} \uparrow \mathbf{P}(Z_1 = j | Z_n > 0) \\
&= \sum_{j=k}^{\infty} j \mathbf{P}(Z_1 = j) = \mathbf{P}(\widehat{L} \geq k).
\end{aligned}$$

D'où

$$\lim_{n\to\infty} \uparrow \mathbf{E}[Z_1 | Z_n > 0] = \sum_{k=1}^{\infty} \mathbf{P}(\widehat{L} \geq k) = \mathbf{E}\big[\widehat{L}\big] = \frac{\mathbf{E}[L^2]}{\mathbf{E}[L]} = \sigma^2 + 1.$$

Notons u_n^n la particule située le plus à gauche de la génération n lorsque $Z_n > 0$, $u_{n-1}^n, \ldots, u_0^n$ les ancêtres de u_n^n, u_i^n appartenant à la i-ème génération. Soit $\widetilde{X}_i$ le nombre de descendants de u_i^n à la génération n qui ne sont pas descendants de u_{i+1}^n. Soit X_i le nombre d'enfants de u_i^n situés à la droite de u_{i+1}^n. Alors $Z_n = 1 + \sum_{i=0}^{n-1} \widetilde{X}_i$, et comme chaque X_i est à l'origine d'un processus de Galton-Watson indépendant de même loi que le processus d'origine (donc de moyenne égale à 1), $\mathbf{E}[\widetilde{X}_i | Z_n > 0] = \mathbf{E}[X_i | Z_n > 0]$. D'où :

$$\frac{1}{n\mathbf{P}(Z_n > 0)} = \frac{1}{n}\mathbf{E}[Z_n | Z_n > 0] = \frac{1}{n} + \frac{1}{n}\sum_{i=0}^{n-1} \mathbf{E}[X_i | Z_n > 0].$$

Nous avons vu que la loi de X_i sous $\mathbf{P}(\ .\ | Z_n > 0)$ tend vers la loi d'une variable aléatoire de loi uniforme sur $\{0, 1, 2, \ldots, \widehat{L} - 1\}$, donc $\lim_{n\to\infty} \mathbf{E}[X_i | Z_n > 0] = \frac{1}{2}\mathbf{E}[\widehat{L} - 1] = \frac{\sigma^2}{2}$. On en déduit finalement (17) par Césaro.

Références

[1] R. Lyons, Y. Peres, *Probability on trees and networks* (chapitre 11), 1997-2005, `http ://mypage.iu.edu/~rdlyons/prbtree/book.pdf`.

Appendice

Construction et interprétation intuitives des mesures $\widehat{\mathbf{GW}}_n$ et $\widehat{\mathbf{GW}}$

Soit une suite d'entiers $p_0, p_1, \dots$ telle que $\sum_{k=0}^{\infty} p_k = 1$. On construit alors un arbre de la façon suivante : on attribue à la racine k enfants avec probabilité p_k, puis à chacun de ses enfants on attribue à nouveau k enfants avec probabilité p_k, etc. On construit ainsi un arbre fini ou infini. Si s est un sommet d'un arbre a, on note $h(s)$ sa hauteur, c'est-à-dire sa distance à la racine, et $e_a(s)$ son nombre d'enfants. On appelle Z_n le nombre de sommets s de l'arbre tels que $h(s) = n$. On note également $m = \sum_{k=0}^{\infty} kp_k$ le nombre d'enfants moyen d'un sommet. Posons $W_n = Z_n/m^n$ et $W = \limsup_{n\to\infty} W_n$. Dans la suite, L sera une variable aléatoire telle que $\mathbf{P}[L = k] = p_k$.

Soit a un arbre et $n \in \mathbf{N}$. Notons $[a]_n$ l'ensemble des arbres dont les n premiers étages coïncident avec a. Sur l'espace des arbres on considère la tribu $\mathcal{A}$ engendrée par les $[a]_n$ pour tout arbre a et tout entier naturel n. On peut définir sur cette tribu la probabilité $\mathbf{GW}$ par :

$$\mathbf{GW}([a]_n) = \prod_{\{s\in a,\ h(s)<n\}} p_{e_a(s)}.$$

Pour le théorème de Kesten-Stigum, nous nous plaçons dans le cas où les arbres deviennent infinis avec une probabilité non nulle. Pour cela, il faut et il suffit que $m > 1$. Choisir au hasard un arbre selon $\mathbf{GW}$, consiste à prendre une racine puis à voir comment se développe sa descendance étant donnée la suite $(p_k)_{k\in\mathbf{N}}$. Le but du théorème étant d'étudier le comportement de ces arbres à l'infini, il est naturel de vouloir introduire une nouvelle mesure de probabilité (que nous noterons plus tard $\widehat{\mathbf{GW}}$) dont le principe serait de choisir l'arbre par les feuilles (à l'infini) plutôt que par la racine. De manière informelle, si deux ensembles d'arbres A_1 et A_2 ont la même probabilité selon $\mathbf{GW}$ et si, pour les arbres de A_1, Z_n croît comme m^n, tandis que pour ceux de A_2, Z_n croît comme $2m^n$, alors on devra avoir $\widehat{\mathbf{GW}}(T_2) = 2\widehat{\mathbf{GW}}(T_1)$.

Afin de construire $\widehat{\mathbf{GW}}$, commençons par introduire une variable aléatoire, notée $\widehat{L}$, dite décalée, dont la loi est définie de manière intuitive par :

$$\mathbf{P}[\widehat{L} = k] = \text{ probabilité pour un sommet d'avoir } k \text{ frères.}$$

Remarquons que, ici, le mot frère est employé dans un sens général : mes frères sont les fils de mon père, donc je suis mon propre frère. De même, dans la suite, mon père fait parti de mes oncles, mes frères et moi même faisons partie de mes cousins, etc.

La probabilité d'avoir k frères est d'autant plus importante que la famille est nombreuse (i.e. que k est grand). Par exemple, s'il y a autant de familles à deux enfants que de familles à un enfant ($p_1 = p_2 = 1/2$), alors il y a deux fois plus de sommets qui ont deux frères que de sommets qui ont un frère ($\mathbf{P}[\widehat{L} = 1] = 1/3$ et $\mathbf{P}[\widehat{L} = 2] = 2/3$).

Plus généralement, la probabilité d'avoir k frères est proportionnelle à kp_k, soit après normalisation :

$$\mathbf{P}[\widehat{L} = k] = \frac{kp_k}{m}.$$

Probabilité décalée d'ordre 1. Dans un premier temps, cherchons à comprendre comment les probabilités sont modifiées si, au lieu de choisir l'arbre par la racine, nous le choisissons par un des sommets de la première génération. Si nous prenons au hasard un sommet parmi ceux de la première génération, il a k frères avec probabilité kp_k/m. Autrement dit, la racine a k enfants avec probabilité kp_k/m. Puis chacun de ses enfants a une descendance classique selon $\mathbf{GW}$. On en déduit une nouvelle façon de construire un arbre : on donne à la racine une descendance selon la loi de $\widehat{L}$ et à tous les autres sommets une descendance selon la loi de L. On peut alors définir une nouvelle probabilité par :

$$\widehat{\mathbf{GW}}_1([a]_n) = \frac{lp_l}{m} \prod_{\{s \in a,\ 0 < h(s) < n\}} p_{e_a(s)},$$

où l est le nombre d'enfants de la racine de a. Cette probabilité $\widehat{\mathbf{GW}}_1$ a décalé d'un étage la probabilité $\mathbf{GW}$, puisque tirer un arbre par la racine avec $\widehat{\mathbf{GW}}_1$ revient à tirer un arbre par la première génération avec $\mathbf{GW}$.

Probabilité décalée d'ordre n. Par le même type d'argument, nous allons maintenant définir une probabilité $\widehat{\mathbf{GW}}_n$ qui correspond au choix d'un sommet dans la n-ième génération. Soit s un sommet choisi au hasard parmi les sommets de la n-ième génération. Son père a k enfants avec probabilité kp_k/m. Ses frères (autres que lui-même) ont une descendance classique selon $\mathbf{GW}$. Son grand-père a k enfants avec probabilité kp_k/m. Ce fait n'est pas évident car s étant choisi parmi les individus de la n-ième génération avec une probabilité uniforme, le père de s n'est pas pris parmi les individus de la $(n-1)$-ième génération avec une probabilité uniforme. Le père de s est néanmoins choisi uniformément parmi les oncles de s qui ont le même nombre d'enfants que lui. Comme le nombre d'enfants du grand-père de s et le nombre d'enfants de son père sont indépendants, cela justifie le résultat. Ses oncles différents de son père ont une descendance classique selon $\mathbf{GW}$. Et ainsi de suite : tous les ancêtres directs de s ont k enfants avec probabilité kp_k/m, tandis que les autres sommets ont une descendance en $\mathbf{GW}$.

Cette constatation justifie une nouvelle construction d'arbres : on donne à la racine s_0 une descendance selon la loi de $\widehat{L}$, puis on choisit une racine s_1 uniformément parmi les enfants de la racine. On donne à s_1 une descendance selon la loi de $\widehat{L}$. Puis on choisit une racine s_2 uniformément parmi les enfants de s_1. Et ainsi de suite : on donne à s_{n-2} une descendance selon la loi de $\widehat{L}$, puis on choisit une racine s_{n-1} uniformément parmi les enfants de s_{n-2}. On donne à s_{n-1} une descendance selon la loi de $\widehat{L}$, puis on choisit une racine s_n uniformément parmi les enfants de s_{n-1}. Tous les autres sommets ont une descendance en $\mathbf{GW}$.

Pour définir $\widehat{\mathbf{GW}}_n$, remarquons qu'avec la méthode précédente, on peut tirer le même arbre en étant passé par un chemin $s_0, s_1, \ldots, s_n$ différent. En fait, cette méthode fournit la probabilité de tirer le sommet s_n parmi les Z_n sommets de n-ième génération que possède l'arbre. Comme ces Z_n sommets ont tous la même probabilité d'être choisis, il s'ensuit :

$\widehat{\mathbf{GW}}_n([a]_k) =$ (nombre de sommets distingués qui permettent la construction du même arbre) $\prod$(probabilités intervenant lors de la construction des k premiers niveaux de a).

Les probabilités intervenant lors de la construction des n premiers niveaux de a sont :

1. pour un s_i $(i < n)$ qui a k enfants : kp_k/m pour avoir eu k enfants et $1/k$ pour le choix de s_{i+1} parmi eux, soit un produit de p_k/m
2. pour un autre sommet qui a k enfants : p_k.

On trouve alors, en notant $n \wedge k = \min(n, k)$:

$$\widehat{\mathbf{GW}}_n([a]_k) = Z_{n\wedge k} \frac{1}{m^{n\wedge k}} \prod_{\{s \in a,\ h(s)<n\}} p_{e(s)},$$

$$\widehat{\mathbf{GW}}_n([a]_k) = W_{n\wedge k} \mathbf{GW}([a]_k).$$

Probabilité décalée d'ordre infini. Nous pouvons désormais passer à la construction de $\widehat{\mathbf{GW}}$. On construit un arbre selon $\widehat{\mathbf{GW}}$ comme on construit un arbre avec $\widehat{\mathbf{GW}}_n$, mais la suite des s_n est prolongée jusqu'à l'infini. Ceci donne donc un sens plus précis à ce que l'on entendait au début par choisir un arbre par les feuilles. Pour trouver l'expression de $\widehat{\mathbf{GW}}([a]_n)$, on peut effectuer les mêmes calculs que nous avions fait pour $\widehat{\mathbf{GW}}_n$, et on trouve l'expression souhaitée si on fait tendre n vers l'infini dans $\widehat{\mathbf{GW}}_n([a]_k) = W_{n\wedge k} \mathbf{GW}([a]_k)$. On peut également comprendre cela en se disant que si on évalue $[a]_n$, nous n'avons pas plus d'informations sur l'arbre en étant à l'infini que nous en avions en connaissant ses n premiers niveaux, soit $\widehat{\mathbf{GW}}([a]_n) = \widehat{\mathbf{GW}}_n([a]_n)$.

On obtient finalement la relation fondamentale suivante : $\widehat{\mathbf{GW}}([a]_n) = W_n \mathbf{GW}([a]_n)$.