Processus aléatoires Thomas Budzinski
ENS Paris, 2017-2018 Bureau V2
thomas.budzinski@ens.fr

TD 9 : Encore des martingales ! Corrigé

Mercredi 15 Novembre

Exercice 1 (Modèle de Wright-Fisher)
Un gène a deux allèles α et β. On considère une population qui se renouvelle entièrement à chaque génération, le nombre d'individus restant fixe et égal à $n \geq 2$. On suppose que pour tout $k \geq 0$, à la génération $k+1$, chacun des n individus choisit son parent uniformément parmi les n individus de la génération k, indépendamment les uns des autres. On note X_k le nombre d'individus portant l'allèle α à la génération k, et $\mathscr{F}_k = \sigma(X_0, \dots, X_k)$. On suppose $X_0 = a \in [\![1, n-1]\!]$.

1. Quelle est la loi de X_{k+1} conditionnellement à $\mathscr{F}_k$? En déduire que X est une martingale pour $(\mathscr{F}_k)_{k \geq 0}$.
2. Montrer que X converge p.s. vers une variable X_∞, et donner sa loi.
   On pose $\tau = \inf\{k | X_k = X_\infty\}$.
3. Calculer $\mathbb{E}\left[X_{k+1}(n - X_{k+1}) | \mathscr{F}_k\right]$ et trouver une martingale.
4. En déduire un encadrement de $\mathbb{E}[\tau]$. On pourra par exemple montrer

$$\mathbb{E}[\tau] = O(n \ln n).$$

<u>*Solution de l'exercice 1*</u>

1. Conditionnellement à $\mathscr{F}_k$, chaque individu de la génération porte l'allèle α avec probabilité $\frac{X_k}{n}$, indépendamment les uns des autres, donc X_{k+1} suit une loi binomiale de paramètres n et $\frac{X_k}{n}$. On en déduit $\mathbb{E}[X_{k+1} | \mathscr{F}_k] = n \times \frac{X_k}{n} = X_k$. De plus, le processus X est bien intégrable et adapté, donc X est bien une martingale.
2. La martingale X est bornée par n, donc elle converge p.s. et dans L^1 vers une variable X_∞. On va maintenant montrer que $X_\infty \in \{0, n\}$ p.s. Soit $j \in [\![1, n-1]\!]$. Pour tout k, on pose

$$q = \mathbb{P}\left(X_{k+1} = j | X_k = j\right) = \binom{n}{j} \left(\frac{j}{n}\right)^j \left(1 - \frac{j}{n}\right)^{n-j} < 1.$$

   On a pour tout $\ell \geq 0$:

$$\mathbb{P}\left(X_k = X_{k+1} = \dots = X_{k+\ell} = j\right) = \mathbb{P}\left(X_k = j\right) \prod_{i=0}^{\ell-1} \mathbb{P}\left(X_{k+i+1=j} | X_{k+i=j}\right) \leq q^\ell \xrightarrow[\ell \to +\infty]{} 0.$$

   On en déduit $\mathbb{P}\left(\forall i \geq k, X_i = j\right) = 0$ pour tous $j \in [\![1, n-1]\!]$ et $k \geq 0$, donc $\mathbb{P}\left(X_\infty = j\right) = 0$ (comme X est à valeurs entières et converge, elle est constante à partir d'un certain rang). On a donc $X_\infty \in \{0, n\}$ p.s. De plus, par convergence L^1 on a $\mathbb{E}\left[X_\infty\right] = \mathbb{E}[X_0] = a$, donc $\mathbb{P}\left(X_\infty = 0\right) = 1 - \frac{a}{n}$ et $\mathbb{P}\left(X_\infty = n\right) = \frac{a}{n}$.

3. On sait que l'espérance et la variance d'une variable de loi binomiale de paramètres n et p valent respectivement pn et $p(1-p)n$. On obtient donc

$$\begin{aligned}
\mathbb{E}[X_{k+1}(n-X_{k+1})|\mathscr{F}_k] &= n\mathbb{E}[X_{k+1}|\mathscr{F}_k] - \mathbb{E}[X_{k+1}|\mathscr{F}_k]^2 - \mathrm{Var}(X_{k+1}|\mathscr{F}_k) \\
&= nX_k - X_k^2 - \frac{1}{n}X_k(n-X_k) = \frac{n-1}{n}X_k(n-X_k).
\end{aligned}$$

On en déduit que $\left(\left(\frac{n}{n-1}\right)^k X_k(n-X_k)\right)_{k\geq 0}$ est une martingale, qu'on note M.

4. Soit $k \geq 0$. Si $\tau > k$, alors $1 \leq X_k \leq n-1$ donc $X_k(n-X_k) \geq n-1$ donc

$$\begin{aligned}
\mathbb{P}(\tau \geq k+1) &= \mathbb{P}(X_k(n-X_k) \geq n-1) \\
&= \mathbb{P}\left(M_k \geq (n-1)\left(\frac{n}{n-1}\right)^k\right) \\
&\leq \frac{1}{n-1}\left(\frac{n-1}{n}\right)^k \mathbb{E}[M_k] \\
&= \frac{1}{n-1}\left(\frac{n-1}{n}\right)^k \mathbb{E}[M_0] \\
&\leq \frac{n^2}{4(n-1)}\left(1-\frac{1}{n}\right)^k.
\end{aligned}$$

On a donc $\mathbb{P}(\tau \geq k+1) \leq 1 \wedge \frac{n^2}{4(n-1)}\left(1-\frac{1}{n}\right)^k$. On a $\frac{n^2}{4(n-1)}\left(1-\frac{1}{n}\right)^k = 1$ pour $k \approx n\ln n$. On écrit donc

$$\begin{aligned}
\mathbb{E}[\tau] &= \sum_{k\geq 0}\mathbb{P}(\tau \geq k+1) \\
&\leq \sum_{k=0}^{\lfloor n\ln n\rfloor} 1 + \sum_{k=1+\lfloor n\ln n\rfloor}^{+\infty} \frac{n^2}{4(n-1)}\left(1-\frac{1}{n}\right)^k \\
&\leq n\ln n + 1 + \frac{n^2}{4(n-1)}n\left(1-\frac{1}{n}\right)^{n\ln n} \\
&\leq n\ln n + 1 + \frac{n^3}{4(n-1)}\exp\left(-\frac{1}{n}n\ln n\right) \\
&= n\ln n + O(n).
\end{aligned}$$

Par ailleurs, pour tout $k \geq 0$ on a $X_k(n-X_k) \leq \mathbb{1}_{\tau>k}\frac{n^2}{4}$ donc

$$a(n-a) = \mathbb{E}[M_k] \leq \frac{n^2}{4}\left(\frac{n}{n-1}\right)^k \mathbb{P}(\tau > k).$$

On en déduit

$$\mathbb{E}[\tau] = \sum_{k\geq 0}\mathbb{P}(\tau \geq k+1) \geq \frac{4}{n^2}a(n-a)\sum_{k\geq 0}\left(1-\frac{1}{n}\right)^k = \frac{4a(n-a)}{n}.$$

Par exemple, on prenant $a = \frac{n}{2}$, on obtient $\mathbb{E}[\tau] \geq (1+o(1))n$.

Remarque On peut montrer qu'en faisant tendre a et n vers $+\infty$ avec $\frac{a}{n} \to x$, on a

$$\mathbb{E}[\tau] \sim -2(x\ln x + (1-x)\ln(1-x))n.$$

Voir par exemple ce lien (pages 4 et 5) pour une explication heuristique.

Exercice 2 (Une preuve de la loi forte des grands nombres)
Soit $(Z_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes. On suppose que $\mathbb{E}[Z_n] = 0$ pour tout $n \geq 1$ et $\sum_{n\geq 1} \frac{\mathrm{Var}(Z_n)}{n^2} < +\infty$. On pose, pour $n \geq 1$,

$$M_n = \sum_{j=1}^{n} \frac{Z_j}{j} \qquad \text{et} \qquad S_n = \sum_{j=1}^{n} Z_j\,.$$

1. Montrer que M_n converge p.s. quand n tend vers $+\infty$.
2. En exprimant S en fonction de M, en déduire $\frac{S_n}{n} \xrightarrow[n\to+\infty]{p.s.} 0$.
3. Soit X une variable aléatoire telle que $\mathbb{E}[X] = 0$. On considère une suite $(X_n)_{n\geq 1}$ de variables aléatoires i.i.d. de même loi que X. Pour tout $n \geq 1$, on pose $Y_n = X_n \mathbb{1}_{|X_n|\leq n}$. Montrer que :
   (i) $\mathbb{E}[Y_n] \xrightarrow[n\to\infty]{} \mathbb{E}[X]$,
   (ii) $\mathbb{P}\left(\exists n \geq 1, \forall j \geq n, X_j = Y_j\right) = 1$,
   (iii) $\sum_{n\geq 1} \frac{\mathrm{Var}(Y_n)}{n^2} < \infty$.
4. En déduire la loi forte des grands nombres.

Solution de l'exercice 2

1. Comme les X_i sont d'espérance nulle, on vérifie facilement que M est une martingale. De plus, par indépendance des X_i, pour tout n on a
$$\mathbb{E}[M_n^2] = \mathrm{Var}(M_n) = \sum_{i=j}^{n} \mathrm{Var}\left(\frac{Z_j}{j}\right) = \sum_{j=1}^{n} \frac{\mathrm{Var}(Z_j)}{j^2},$$
qui est bornée par l'hypothèse sur les variances. La martingale M est donc bornée dans L^2, donc elle converge p.s. vers une variable M_∞.
2. Pour tout $n \geq 1$, on a
$$\frac{S_n}{n} = \frac{1}{n}\sum_{j=1}^{n} j(M_j - M_{j-1}) = \frac{1}{n}\left(\sum_{j=1}^{n} jM_j - \sum_{j=1}^{n-1}(j+1)M_j\right) = M_n - \frac{1}{n}\sum_{j=1}^{n-1} M_j.$$
D'après le lemme de Cesaro, le membre de droite tend vers 0, d'où $\frac{S_n}{n} \to 0$ p.s.
3. (i) On a $\mathbb{E}[Y_n] = \mathbb{E}[X\mathbb{1}_{|X|\leq n}]$, avec $X\mathbb{1}_{|X|\leq n} \to X$ p.s. et $|X\mathbb{1}_{|X|\leq n}| \leq |X|$, donc $\mathbb{E}[Y_n] \to \mathbb{E}[X] = 0$ par convergence dominée.
   (ii) Pour tout $j \geq 1$, on a $\mathbb{P}(X_j \neq Y_j) = \mathbb{P}(|X_j| > j) = \mathbb{P}(|X| > j)$ donc
$$\sum_{j=1}^{\infty} \mathbb{P}(X_j \neq Y_j) = \sum_{j=1}^{\infty} \mathbb{P}(|X| > j) \leq \int_0^\infty \mathbb{P}(|X| > t)\,\mathrm{d}t = \mathbb{E}[|X|] < +\infty,$$
donc d'après le lemme de Borel-Cantelli, p.s., $X_j = Y_j$ pour j assez grand.
   (iii) Enfin, on a
$$\begin{aligned}\sum_{n=1}^{\infty} \frac{\mathrm{Var}(Y_n)}{n^2} \leq \sum_{n=1}^{\infty} \frac{\mathbb{E}[Y_n^2]}{n^2} &= \sum_{n=1}^{\infty} \mathbb{E}\left[\frac{|X|^2}{n^2}\mathbb{1}_{|X|\leq n}\right] = \mathbb{E}\left[|X|^2 \sum_{n\geq 1\vee |X|} \frac{1}{n^2}\right] \\ &\leq \mathbb{E}\left[|X|^2 \frac{c}{|X|}\right] = c\mathbb{E}[|X|] < +\infty,\end{aligned}$$
en utilisant à la fin $\sum_{n\geq 1\vee a} = O\left(\frac{1}{a}\right)$.

4. On se place dans le cadre de la question 3. Pour tout $n \geq 1$, posons $Z_n = Y_n - \mathbb{E}[Y_n]$. On a $\mathrm{Var}(Z_n) = \mathrm{Var}(Y_n)$ pour tout $n \geq 0$ donc $(Z_n)_{n\geq 1}$ vérifie les hypothèses des questions 1 et 2, donc

$$\frac{1}{n}\sum_{j=1}^{n}(Y_j - \mathbb{E}[Y_j]) \xrightarrow[n\to\infty]{\text{p.s.}} 0\,.$$

Or $\mathbb{E}[Y_n] \xrightarrow[n\to\infty]{} \mathbb{E}[X] = 0$, donc d'après le lemme de Cesaro on a

$$\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}[Y_j] \xrightarrow[n\to\infty]{} 0, \qquad \text{d'où} \qquad \frac{1}{n}\sum_{j=1}^{n} Y_j \xrightarrow[n\to\infty]{\text{p.s.}} 0.$$

Enfin, p.s., il existe $J \geq 1$ tel que $Y_j = X_j$ pour tout $j \geq J$. On a alors

$$\frac{1}{n}\sum_{j=1}^{n} X_j = \frac{1}{n}\sum_{j=1}^{J-1} X_j + \frac{1}{n}\sum_{j\geq J} Y_j \xrightarrow[n\to\infty]{} 0,$$

ce qui achève la démonstration.

Exercice 3 (Vaisseau spatial perdu)
Le *Millenium Falcon* se trouve à une distance D_0 du Soleil mais ses commandes ne répondent plus : toutes les heures, Han Solo ne peut qu'entrer une distance r_n inférieure à la distance au Soleil dans l'ordinateur de bord, qui effectue alors un saut dans l'hyperespace de longueur r_n et de direction choisie uniformément dans la sphère S^2. On note D_n la distance du vaisseau au Soleil après n sauts et $\mathscr{F}_n$ la tribu engendrée par les n premiers sauts. Han Solo veut revenir dans le système solaire, c'est-à-dire à distance au plus d du soleil.

1. En utilisant des souvenirs de physique de prépa (théorème de Gauss), montrer que $\left(\frac{1}{D_n}\right)$ est une martingale.
2. En déduire que la probabilité que Han Solo revienne un jour dans le système solaire est inférieure ou égale à $\frac{d}{D_0}$.
3. A la place du pilote, feriez-vous plutôt de grands ou de petits sauts ?

<u>*Solution de l'exercice 3*</u> Toutes les justifications des interversions seront laissées en exercice.

1. Soit X_n la variable aléatoire à valeurs dans S^2 qui indique la direction du n-ième saut. Alors on veut montrer que $\mathbb{E}\left[\|S_n + r_{n+1}X_{n+1}\|^{-1}\,|\mathscr{F}_n\right] = \|S_n\|^{-1}$. Pour tout x dans $\mathbb{R}^3\backslash\{0\}$, on pose $f(x) = \|x\|^{-1}$. On veut donc montrer que pour tous $x \in \mathbb{R}^3$ et $r < \|x\|$:

$$\frac{1}{4\pi}\int_{S^2} f(x+ru)\mathrm{d}u = f(x).$$

Il suffit pour cela de vérifier que la dérivée par rapport à r du membre de gauche est nulle (par convergence dominée, il tend bien vers 0 quand $r \to 0$). Notons que le membre de gauche a une discontinuité en $r = \|x\|$, c'est pourquoi on impose $r < \|x\|$. En intervertissant dérivée et intégrale puis en appliquant le théorème de Gauss, on obtient

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}r}\int_{S^2} f(x+ru)\mathrm{d}u &= \int_{S^2}\nabla f(x+ru)\cdot u\,\mathrm{d}u\\ &= r\int_{B_1}\mathrm{div}(\nabla f(x+y))\,\mathrm{d}y\\ &= r\int_{B_1}\Delta f,\end{aligned}$$

où B_1 est la boule de rayon 1 autour de l'origine dans $\mathbb{R}^3$. Un simple calcul (ou, à nouveau, des souvenirs de physique de prépa) montre que le Laplacien Δf est nul sur $\mathbb{R}^3\backslash\{0\}$, donc $\frac{1}{D_n}$ est bien une martingale.

2. Soit $T = \inf\{n | D_n \leq d\}$. Le temps T est un temps d'arrêt (éventuellement infini), donc on peut appliquer le théorème d'arrêt à $T \wedge t$ et à la martingale $\frac{1}{D_n}$:

$$\mathbb{E}\left[\frac{1}{D_{T\wedge t}}\right] = \frac{1}{D_0}.$$

On en déduit

$$\begin{aligned}
\mathbb{P}(T \leq t) &= \mathbb{P}(D_{T\wedge t} \leq d) \\
&= \mathbb{P}\left(\frac{1}{D_{T\wedge t}} \geq \frac{1}{d}\right) \\
&\leq \left(\frac{1}{d}\right)^{-1} \mathbb{E}\left[\frac{1}{D_{T\wedge t}}\right] \\
&= \frac{d}{D_0}.
\end{aligned}$$

Ceci est valable pour tout $t > 0$ donc $\mathbb{P}(T < +\infty) \leq \frac{d}{D_0}$.

3. On veut que l'inégalité de la question précédente soit la plus serrée possible. Le seul endroit où on n'a pas égalité ci-dessus est dans l'inégalité de Markov (avant-dernière ligne du dernier calcul). Pour que l'inégalité de Markov soit serrée, il faut que $\frac{1}{D_{T\wedge t}}$ ne puisse pas être "beaucoup" plus grande que $\frac{1}{d}$. Il faut donc faire de petits sauts à l'approche du système solaire. On peut vérifier (exercice !) que pour tout $\varepsilon > 0$, si le saut à chaque étape n est inférieur ou égal à $D_n - d + \varepsilon$, alors on a $\mathbb{P}(T < +\infty) \geq \frac{d-\varepsilon}{D_0}$.

Remarque L'hypothèse "les sauts sont plus petits que la distance au Soleil" peut paraître arbitraire. En supprimant cette hypothèse, le processus $\left(\frac{1}{D_n}\right)_{n\geq 0}$ n'est plus forcément une martingale mais une *surmartingale*, c'est-à-dire que $E\left[\frac{1}{D_{n+1}} | \mathscr{F}_n\right] \leq \frac{1}{D_n}$. La raison est que, au sens des distributions, le Laplacien de $x \to \|x\|^{-1}$ sur $\mathbb{R}^3$ est (à une constante) multiplicative près) $-\delta_0$, donc est négatif. Le théorème d'arrêt pour les surmartingales donne

$$\mathbb{E}\left[\frac{1}{D_{T\wedge t}}\right] \leq \frac{1}{D_0}.$$

L'inégalité étant dans le bon sens, le résultat de la question 2 reste vrai. Cela montre que faire des sauts trop grands ne peut qu'aggraver la situation de notre vaisseau.

Exercice 4 (Processus de Galton–Watson surcritique)
Soit μ une loi sur $\mathbb{N}$ telle que $\sum_i i\mu(i) = m > 1$ et $\sum_i i^2\mu(i) < +\infty$. Soient $(Z_{n,i})_{n,i\in\mathbb{N}}$ des variables i.i.d. de loi μ. On définit le processus X par $X_0 = 1$ et, pour tout $n \geq 0$,

$$X_{n+1} = \sum_{i=1}^{X_n} Z_{n,i}.$$

1. Que peut décrire le processus X ?
2. On pose $p_n = \mathbb{P}(X_n = 0)$ et $p = \mathbb{P}(\exists n, X_n = 0)$. Montrer une formule de récurrence de la forme $p_{n+1} = f(p_n)$, et en déduire que $p < 1$.
3. On pose $M_n = m^{-n} X_n$. Montrer que M est une martingale. En déduire que M_n converge p.s. vers une variable M_∞.
4. Trouver une relation de récurrence sur $\mathbb{E}[M_n^2]$, et en déduire que $M_n \to M_\infty$ dans L^2.
5. On note $q = \mathbb{P}(M_\infty = 0)$. Donner une équation sur q. En déduire que $q = p$. Qu'est-ce-que cela signifie sur la croissance de X_n ?

Solution de l'exercice 4

1. Supposons qu'une population évolue de la manière suivante : à chaque génération n, les individus se reproduisent indépendamment des générations précédentes et les uns des autres, de telle manière que le nombre d'enfants d'un individu a pour loi μ. Alors le processus X décrit le nombre d'individus à la génération n.
2. Dire que $X_{n+1} = 0$ revient à dire qu'il existe i tel que le premier individu a eu i enfants (ce qui arrive avec proba $\mu(i)$), et chacun de ces i enfants n'a pas de descendant à la génération n (ce qui arrive avec proba p_n pour chaque enfant). Par conséquent, on a

$$p_{n+1} = \sum_i \mu(i) p_n^i = f(p_n),$$

avec $f(x) = \sum_i \mu(i)x^i$. On sait de plus que $p = \lim_{n\to+\infty} p_n$, donc p est un point fixe de f. De plus, f est croissante (les $\mu(i)$ sont positifs), donc si p' est un point fixe de f, on a par récurrence $p_n \le p'$ pour tout n, donc $p \le p'$. On en déduit que p est le plus petit point fixe de f, donc montrer que $p < 1$ revient à montrer que f admet un point fixe strictement inférieur à 1. Or, on a $f(1) = 1$ et $f'(1) = m > 1$, donc $f(x) < x$ pour x assez proche de 1. Mais on a aussi $f(0) \ge 0$, donc par le théorème des valeurs intermédiaires f admet un point fixe dans $[0, 1[$, donc $p < 1$.
3. Soit $\mathscr{F}_n$ la tribu engendrée par les $Z_{k,i}$ pour $k \le n-1$. Alors X_n ne dépend que des $Z_{k,i}$ avec $k \le n-1$ et $i \in \mathbb{N}$, donc X est $(\mathscr{F}_n)$-adapté, donc M aussi. De plus, comme M est positif, on peut faire le calcul suivant sans savoir M_n et M_{n+1} sont intégrables :

$$\begin{aligned}
\mathbb{E}\left[M_{n+1}|\mathscr{F}_n\right] &= m^{-(n+1)}\mathbb{E}\left[X_{n+1}|\mathscr{F}_n\right] \\
&= m^{-(n+1)}\mathbb{E}\left[\sum_{i=1}^{X_n} Z_{n,i}\middle|\mathscr{F}_n\right] \\
&= m^{-(n+1)}\sum_{i=1}^{X_n}\mathbb{E}\left[Z_{n,i}|\mathscr{F}_n\right] \\
&= m^{-(n+1)}\sum_{i=1}^{X_n} m \\
&= m^{-(n+1)} m X_n \\
&= M_n.
\end{aligned}$$

En particulier, on en déduit que $\mathbb{E}[M_n] = \mathbb{E}[M_0] < +\infty$ pour tout n, et M est une martingale positive, donc elle converge p.s..
4. Notons σ^2 la variance de la loi μ. Pour tout n, on a

$$\begin{aligned}
\mathbb{E}\left[M_{n+1}^2|\mathscr{F}_n\right] &= m^{-2(n+1)}\mathbb{E}\left[\left(\sum_{i=1}^{X_n} Z_{n,i}\right)^2\middle|\mathscr{F}_n\right] \\
&= m^{-2(n+1)}\sum_{i,j=1}^{X_n}\mathbb{E}\left[Z_{n,i}Z_{n,j}\right] \\
&= m^{-2(n+1)}\left(\sum_{i=1}^{X_n}\mathbb{E}\left[Z_{n,i}^2\right] + \sum_{i\ne j}\mathbb{E}\left[Z_{n,i}\right]\mathbb{E}\left[Z_{n,j}\right]\right) \\
&= m^{-2(n+1)}\left((m^2+\sigma^2)X_n + m^2 X_n(X_n-1)\right) \\
&= M_n^2 + \frac{\sigma^2}{m^{2(n+1)}}M_n.
\end{aligned}$$

En prenant l'espérance des deux côtés, on obtient

$$\mathbb{E}\left[M_{n+1}^2\right] = \mathbb{E}\left[M_n^2\right] + \frac{\sigma^2}{m^{2(n+1)}}\mathbb{E}\left[M_n\right] = \mathbb{E}\left[M_n^2\right] + \frac{\sigma^2}{m^{n+2}}.$$

Comme $\sum_n \frac{\sigma^2}{m^{n+2}} < +\infty$, on en déduit que $\mathbb{E}\left[M_n^2\right]$ est borné, donc M est bornée dans L^2, donc elle converge dans L^2.

5. On dit qu'un individu x est à *descendance lente* si le nombre de descendants de x après n générations est $o(m^n)$. En particulier, un individu dont la descendance s'éteint est à descendance lente, et q est la probabilité que l'individu de départ soit à descendance lente.

   Dire que l'individu de départ a une descendance lente revient à dire qu'il existe i tel qu'il a i enfants, et chacun d'eux a une descendance lente. De même que dans la question 2, la probabilité que cela arrive vaut $\sum_i \mu(i)q^i = f(q)$, donc $q = f(q)$. Or, μ est une série entière à coefficients positifs, et il y a au moins un $i \geq 2$ tel que $\mu(i) > 0$, donc f est strictement convexe, donc elle a au plus deux points fixes. Comme p et 1 sont deux points fixes de f, on a donc soit $q = p$, soit $q = 1$. Mais dans le second cas, on a $M_\infty = 0$ p.s.. C'est absurde car M_n converge vers M_∞ dans L^2, donc aussi dans L^1, et $\mathbb{E}[M_n] = \mathbb{E}[M_0] = 1$ pour tout n. On a donc $q = p$. Cela signifie que presque sûrement, soit le processus X s'éteint, soit X_n est asymptotiquement équivalent à m^n fois une variable aléatoire strictement positive.

Remarque On a utilisé la convergence L^2 pour montrer une convergence L^1. Il est naturel de se demander si la convergence L^1 de M reste vraie si μ n'est plus de carré intégrable. Le théorème de Kesten–Stigum affirme que M converge dans L^1 vers M_∞ si et seulement si

$$\sum_i i \log i \, \mu(i) < +\infty.$$

Pour une preuve du théorème de Kesten–Stigum, voir par exemple le chapitre 12.2 de Probability on trees and networks, de Lyons et Peres.

Exercice 5 (Propriété de Liouville)
Soit $G = (V, E)$ un graphe infini, connexe et localement fini (i.e. où chaque sommet n'a qu'un nombre fini de voisins) et $h : V \to \mathbb{R}$. On dit que h est *harmonique sur* G si pour tout $x \in V$, on a

$$h(x) = \frac{1}{\deg(x)} \sum_{y \sim x} h(y),$$

où la somme est effectuée sur les voisins y de x et où $\deg(x)$ est le nombre de ces voisins. On dit que G vérifie la *propriété de Liouville* si toute fonction bornée harmonique sur G est constante.

1. Montrer que si h est harmonique et (X_n) est une marche aléatoire simple sur G, alors $(h(X_n))_{n\geq 0}$ est une martingale.
2. Montrer que si la marche aléatoire simple sur G est récurrente (i.e. si elle visite presque sûrement tous les points une infinité de fois), alors G vérifie la propriété de Liouville.
3. Soient $x, y \in \mathbb{Z}^d$ tels que $\sum_{i=1}^d x_i \equiv \sum_{i=1}^d y_i \pmod 2$. Montrer qu'il existe (X_n) et (Y_n) deux marches aléatoires simples (non indépendantes !) issues respectivement de x et y telles que p.s., pour n assez grand, $X_n = Y_n$.
4. En déduire que $\mathbb{Z}^d$ vérifie la propriété de Liouville.
5. Donner un exemple de graphe (connexe, localement fini) ne vérifiant pas la propriété de Liouville.

<u>*Indication :*</u> Pour la question 3, commencer par le cas $d = 1$ puis essayer d'adapter à d quelconque.

<u>*Solution de l'exercice 5*</u>

1. Soit $\mathscr{F}_n = \sigma(X_0, \ldots, X_n)$. Conditionnellement à $\mathscr{F}_n$, le sommet X_{n+1} est uniforme parmi les voisins de X_n, donc $\mathbb{E}\left[h(X_{n+1})|\mathscr{F}_n\right]$ est la moyenne de h sur les voisins de X_n, c'est-à-dire X_n car h est harmonique.
2. Soit h harmonique bornée sur G. Soient $x, y \in V$ et (X_n) une marche aléatoire simple issue de x. Alors $(h(X_n))_{n\geq 0}$ est une martingale, donc elle converge p.s. Or, elle visite une infinité de fois x et y, donc elle prend une infinité de fois les valeurs $h(x)$ et $h(y)$, donc $h(x) = h(y)$, et ce pour tous x et y. La fonction h est donc constante, donc G est Liouville.

3. Dans le cas $d = 1$, l'idée est de faire démarrer deux marches aléatoires indépendantes $\widetilde{X}$ et $\widetilde{Y}$ issues de x et y. Par récurrence de la marche aléatoire simple sur $\mathbb{Z}$ (plus la condition de parité), le temps $\tau_1 = \inf\{n | \widetilde{X}_n = \widetilde{Y}_n\}$ est fini p.s. On prend alors $X = \widetilde{X}$ ainsi que $Y_n = \widetilde{Y}_n$ pour $n \leq \tau_1$ et $Y_n = \widetilde{X}_n$ pour $n \geq \tau_1$.

   Dans le cas général, il faut "coupler les coordonnées une par une" : on démarre deux marches indépendantes de x et y, et on note τ_1 le premier temps où leurs coordonnées selon e_1 coïncident. À partir de τ_1, on applique la stratégie du cas $d = 1$ pour que les coordonnées selon e_1 restent les mêmes pour $n \geq \tau_1$. Puis on attend τ_2, le premier temps où les coordonnées selon e_2 coïncident, et ainsi de suite.

   Les détails sont laissés en exercice, vous pouvez aussi venir me voir au bureau V2.

4. Soit h harmonique bornée sur G et soient x, y, X et Y comme dans la question précédente. Alors $(h(X_n))_{n\geq 0}$ et $(h(Y_n))_{n\geq 0}$ sont deux martingales bornées, donc elles convergent p.s. et dans L^1 vers respectivement X_∞ et Y_∞. Comme $X_n = Y_n$ pour n assez grand, on a $X_\infty = Y_\infty$ p.s. Par convergence L^1, on peut donc écrire

$$h(x) = \mathbb{E}[X_\infty] = \mathbb{E}[Y_\infty] = h(y),$$

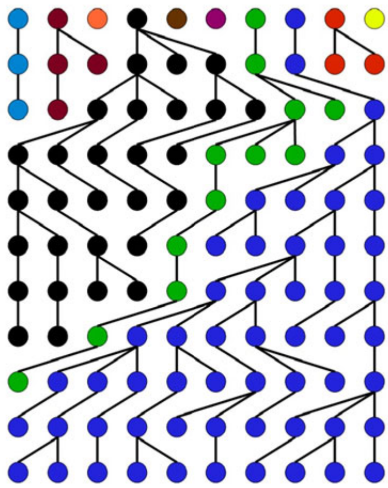
   donc h est constante sur $\{x \in \mathbb{Z}^d | \sum_{i=1}^d x_i \equiv 0 \pmod 2\}$, et il est facile d'en conclure que h est constante sur V.

5. Considérer l'arbre binaire infini. La marche aléatoire est transiente (car la "hauteur" augmente de 1 avec proba $\frac{2}{3}$ et diminue de 1 avec proba $\frac{1}{3}$), donc à partir d'un certain rang elle reste soit dans la moitié gauche de l'arbre, soit dans la moitié droite. On peut alors poser

$$h(x) = \mathbb{P}\left(\text{la marche aléatoire issue de } x \text{ finit dans la moitié gauche de l'arbre}\right).$$

   On vérifie facilement que h est harmonique et bornée (par 1). De plus, si x est "très haut" dans la moitié gauche de l'arbre, il a une très faible probabilité de redescendre jusqu'à la racine, donc $h(x)$ est proche de 1. Les détails sont laissés en exercice.

Exercice 6
Que représente la jolie image ci-dessous ?



Solution de l'exercice 6 Il s'agit d'une illustration du modèle de Wright–Fisher, mais où les individus de départ sont tous différents (et pas seulement de 2 types). Les générations les plus anciennes sont vers le haut, et chaque individu de la génération $k + 1$ est relié à son parent dans la génération k. Image issue d'un article de Bastien Mallein sur Culture Math.