# Chapitre I

# Introduction aux chaines de Markov

## I.1 Chaînes de Markov

Une chaîne de Markov est une suite de variables aléatoires $(X_n, n \in \mathbb{N})$ qui permet de modéliser l'évolution dynamique d'un système aléatoire : $X_n$ représente l'état du système à l'instant $n$. La propriété fondamentale des chaînes de Markov, dite propriété de Markov, est que son évolution future ne dépend du passé qu'au travers de sa valeur actuelle. Autrement dit, conditionnellement à $X_n$, $(X_0, \ldots, X_n)$ et $(X_{n+k}, k \in \mathbb{N})$ sont indépendants. Les applications des chaînes de Markov sont très nombreuses (réseaux, génétique des populations, mathématiques financières, gestion de stock, algorithmes stochastiques d'optimisation, simulation, ...).

nous donnons au paragraphe I.1.1 la définition et des propriétés élementaires des chaînes de Markov. Nous considérons les régimes stationnaires ou probabilité invariante des chaînes de Markov au paragraphe I.1.2. Nous caractérisons les propriétés des états d'une chaîne de Markov, et introduisons la notion de chaîne irréductible au paragraphe I.1.3. Intuitivement une chaîne de Markov irréductible partant d'un état donné peut visiter un autre état au bout d'un certain temps avec probabilité strictement positive. Nous présentons les théorèmes asymptotiques fondamentaux pour les chaînes de Markov irréductible dans le paragraphe I.1.4. Leur démonstration est reportée au paragraphe I.1.5. Le résultat principal est que pour les chaînes de Markov, par exemple à valeurs dans un espace fini, irréductible, la moyenne en temps $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$ converge p.s. vers la moyenne de $f$ par rapport à l'unique probabilité invariante $\pi$ : $(\pi, f)$. Ce résultat est l'analogue de la loi forte des grands nombres. Nous donnons des exemples importants d'utilisation des chaînes de Markov au paragraphe I.1.6.

### I.1.1 Définition et propriétés

Soit $E$ un espace discret, i.e. $E$ est un espace au plus dénombrable muni de la topologie discrète, où tous les points de $E$ sont isolés, et donc de la tribu $\mathcal{E} = \mathcal{P}(E)$.

**Définition I.1.1.** *Une matrice $P = (P(x, y), x, y \in E)$ est dite matrice stochastique si ses coefficients sont positifs et la somme sur une ligne des coefficients est égale à 1 :*

$$P(x, y) \geq 0 \quad et \quad \sum_{z \in E} P(x, z) = 1, \quad pour\ tous\ x, y \in E. \tag{I.1}$$

On rappelle la notation (**??**). On donne une définition des chaînes de Markov apparement plus faible que celle donnée en introduction, voir le théorème I.1.9 pour la propriété de Markov.

**Définition I.1.2.** *Soit $P$ une matrice stochastique sur $E$. Une suite de variables aléatoires $(X_n, n \in \mathbb{N})$ à valeurs dans $E$ est appelée chaîne de Markov de matrice de transition $P$ si pour tous $n \in \mathbb{N}$, $x \in E$, on a*

$$\mathbb{P}(X_{n+1} = x \mid X_n, \ldots, X_0) = \mathbb{P}(X_{n+1} = x \mid X_n) = P(X_n, x). \tag{I.2}$$

*On dit que la chaîne de Markov est issue de $\mu_0$ si la loi de $X_0$ est $\mu_0$.*

Comme l'espace d'état est discret l'équation (I.2) est équivalente à : pour tous $x_0, \ldots, x_n \in E$, tels que $\mathbb{P}(X_n = x_n, \ldots, X_0 = x_0) > 0$,

$$\mathbb{P}(X_{n+1} = x \mid X_n = x_n, \ldots, X_0 = x_0) = \mathbb{P}(X_{n+1} = x \mid X_n = x_n) = P(x_n, x).$$

Si $\mathbb{P}(X_0 = x) = 1$, autrement dit $\mu_0$ est la masse de Dirac en $x$, on dira plus simplement que la chaîne de Markov est issue de $x$.

La proposition suivante assure que la matrice de transition et la loi initiale caractérisent la loi de la chaîne de Markov.

**Proposition I.1.3.** *La loi d'une chaîne de Markov $(X_n, n \in \mathbb{N})$ est entièrement caractérisée par sa matrice de transition, $P$, et la loi de $X_0$, $\mu_0$. De plus, on a, pour tous $n \in \mathbb{N}^*$, $x_0, \ldots, x_n \in E$,*

$$\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) = \mu_0(x_0) \prod_{k=1}^{n} P(x_{k-1}, x_k). \tag{I.3}$$

*Démonstration.* Soit $n \in \mathbb{N}^*$, $x_0, \ldots, x_n \in E$. Si $\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) > 0$, une utilisation successive de la formule des probabilités conditionnelles donne

$$\begin{aligned}
&\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) \\
&\qquad = \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 | X_0 = x_0) \cdots \mathbb{P}(X_n = x_n | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) \\
&\qquad = \mu_0(x_0) \prod_{k=1}^{n} P(x_{k-1}, x_k).
\end{aligned}$$

Si $\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) = 0$, soit $\mathbb{P}(X_0 = x_0) = 0$ i.e. $\mu_0(x_0) = 0$ et donc (I.3) reste vrai ; soit il existe $m \in \{1, \ldots, n-1\}$ tel que $\mathbb{P}(X_0 = x_0, \ldots, X_{m-1} = x_{m-1}) > 0$ et $\mathbb{P}(X_0 = x_0, \ldots, X_m = x_m) = 0$. Dans ce dernier cas, on peut utiliser (I.3) avec $n = m$ et obtenir que

$$0 = \mathbb{P}(X_0 = x_0, \ldots, X_m = x_m) = \mu_0(x_0) \prod_{k=1}^{m} P(x_{k-1}, x_k).$$

On en déduit que (I.3) reste vrai avec les deux membres nuls. En conclusion (I.3) est toujours vérifié.

L'espace produit $E^{\mathbb{N}}$ est l'espace d'état de la chaîne de Markov. Il est muni de la tribu produit. En particulier comme la tribu sur $E$ est engendrée par les singletons, la tribu produit sur $\mathbb{E}^{\mathbb{N}}$ est, grâce à la définition **??**, engendrée par la collection $\mathcal{C} = \{(x_0, \ldots, x_n); n \in \mathbb{N}, x_0, \ldots, x_n \in E\}$. Le théorème de classe monotone et plus particulièrement le corollaire **??** assure que deux probabilités qui coïncident sur $\mathcal{C}$ sont égales. On déduit donc de (I.3) que la loi de la chaîne de Markov $(X_n, n \in \mathbb{N})$ est caractérisée par $P$ et $\mu_0$. □

Donnons quelques exemples de chaînes de Markov.

*Exemple* I.1.4. La marche aléatoire symétrique simple sur $\mathbb{Z}$, $S = (S_n, n \geq 0)$, est définie par $S_n = S_0 + \sum_{k=1}^{n} Z_k$, où $Z = (Z_n, n \geq 1)$ est une suite de variables aléatoires indépendantes et de même loi, $\mathbb{P}(Z_n = 1) = \mathbb{P}(Z_n = -1) = 1/2$, et $S_0$ est une variable aléatoire à valeurs dans $\mathbb{Z}$ indépendante de $Z$. On vérifie facilement que la marche aléatoire simple est une chaîne de Markov à valeurs dans $\mathbb{Z}$ de matrice de transition : $P(x, y) = 0$ si $|x - y| \neq 1$ et $P(x, y) = 1/2$ si $|x - y| = 1$ pour $x, y \in \mathbb{Z}$. $\diamondsuit$

*Exemple* I.1.5. On note $X_n$ l'état d'un stock de pièces détachées à l'instant $n$, $D_{n+1}$ la demande (aléatoire) formulée par des clients, et $q \in \mathbb{N}^*$ la quantité (déterministe) de pièces détachées fabriquées entre les instants $n$ et $n+1$. Alors à l'instant $n+1$, l'état du stock est $X_{n+1} = (X_n + q - D_{n+1})_+$, où $x_+$ désigne la partie positive de $x \in \mathbb{R}$. Si la demande $D = (D_n, n \in \mathbb{N}^*)$ est constituée de variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes et de même loi, et si $X_0$ est une variable aléatoire indépendante de $D$ à valeurs dans $\mathbb{N}$, alors $X = (X_n, n \in \mathbb{N})$ est une chaîne de Markov à valeurs dans $\mathbb{N}$ de matrice de transition : $P(x, y) = \mathbb{P}(D = k)$ si $y = x + q - k > 0$, et $P(x, 0) = \mathbb{P}(D \geq x + q)$. On donne quelques simulations de la chaîne de Markov $X$ pour différentes loi de $D_1$ dans le graphique I.1. $\diamondsuit$

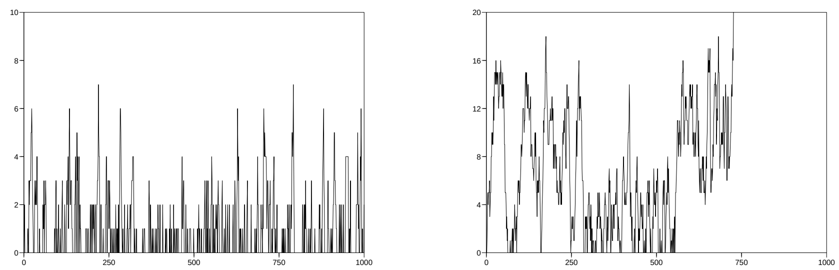


FIG. I.1 – Plusieurs réalisations de l'évolutaion aléatoire d'un stock de dynamique $X_{n+1} = (X_n + q - D_{n+1})_+$, avec $X_0 = 0$, $q = 3$ et où les variables aléatoires $(D_n, n \in \mathbb{N}^*)$ sont indépendantes de loi de Poisson de paramètre $\theta$ ($\theta = 4$ à gauche et $\theta = 3$ à droite).

Les deux exemples précédents sont des cas particuliers du lemme suivant. Sa démonstration est immédiate.

**Lemme I.1.6.** *Soit $U = (U_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires indépendantes et de même loi à valeurs dans un espace mesurable $F$. Soit $X_0$ une variable aléatoire à valeurs dans $E$ indépendante de $U$ et de loi $\mu_0$. Soit $f$ une fonction mesurable définie sur $E \times F$ à valeurs dans $E$. La suite $(X_n, n \in \mathbb{N})$ définie par $X_{n+1} = f(X_n, , U_{n+1})$ pour $n \in \mathbb{N}$ est une chaîne de Markov issue de $\mu_0$ et de matrice de transition $P$ définie par $P(x, y) = \mathbb{P}\Big(f(x, U_1) = y\Big)$ pour tous $x, y \in E$.*

Pour un choix judicieux de suite $U$ et de fonction $f$ dans le lemme précédent, il est aisé de vérifier que pour toute matrice stochastique $P$ sur $E$ et toute probabilité $\mu_0$ sur $E$, on peut construire une chaîne de Markov issue de $\mu_0$ et de matrice de transition $P$.

Il est facile de calculer des espérances ou des lois conditionnelles pour une chaîne de Markov à l'aide des puissances de sa matrice de transition. Nous introduisons d'abord quelques notations.

Soit $P$ et $Q$ deux matrices définies sur $E$. On note $PQ$ la matrice définie sur $E$ par $PQ(x, y) = \sum_{z \in E} P(x, z)Q(z, y)$. On pose $P^0$ la matrice identité et pour $k \geq 1$, $P^k = P^{k-1}P$ (on a aussi $P = PP^{k-1}$). Il est immédiat de vérifier que si $P$ et $Q$ sont des matrices stochastiques alors $PQ$ est une matrice stochastique.

On identifie une probabilité $\mu$ sur $E$ au vecteur $(\mu(x) = \mu(\{x\}), x \in \mathbb{E})$ de $\mathbb{R}^E$, et une fonction $f$ définie sur $E$ à valeurs dans $\mathbb{R}$ au vecteur $(f(x), x \in E)$. Pour une probabilité $\mu$ et une matrice stochastique $P$, on définit le vecteur $\mu P$ par $\mu P(y) = \sum_{x\in E} \mu(x)P(x, y)$ pour tout $y \in E$. Il est évident de vérifier, en utilisant (I.1), que $\mu P$ est également une probabilité. Pour une fonction $f$ positive où bornée et une matrice stochastique $P$, on définit la fonction $Pf$ par $Pf(x) = \sum_{y\in E} P(x, y)f(y)$.

La proposition suivante permet d'exprimer des espérances et des lois conditionnelles pour une chaîne de Markov en fonction de sa matrice de transition.

**Proposition I.1.7.** *Soit $(X_n, n \in \mathbb{N})$ une chaîne de Markov de matrice de transition $P$. On note $\mu_n$ la loi de $X_n$. Soit $f$ une fonction bornée. On a pour $n \in \mathbb{N}^*$*

1. $\mu_n = \mu_0 P^n$,
2. $\mathbb{E}[f(X_n)|X_0] = P^n f(X_0)$,
3. $\mathbb{E}[f(X_n)|X_0, \dots X_{n-1}] = Pf(X_{n-1})$,
4. $\mathbb{E}[f(X_n)] = (\mu_n, f) = (\mu_0 P^n, f) = (\mu_0, P^n f)$.

On utilise les notations $\mathbb{P}_x$ et $\mathbb{E}_x$ quand $X_0$ est p.s. égal à $x$ (i.e. la loi de $X_0$ est la masse de Dirac en $x$). Ainsi on a $\mathbb{P}_x(A) = \mathbb{P}(A|X_0 = x)$ et $\mathbb{E}_x[Z] = \mathbb{E}[Z|X_0 = x]$. Avec ces notations, on peut réécrire la propriété 2 de la proposition : $\mathbb{E}_x[f(X_n)] = P^n f(x)$ pour tout $x \in E$.

*Démonstration.* Le point 1 découle de (I.3) en sommant sur $x_0, \dots, x_{n-1} \in E$. En sommant (I.3) sur $x_1, \dots, x_{n-1} \in E$, on obtient $\mathbb{P}(X_0 = x_0, X_n = x_n) = \mu_0(x_0)P^n(x_0, x_n)$. En mutlipliant par $f(x_n)$ et en sommant sur $x_n \in E$, on obtient $\mathbb{E}[f(X_n)\mathbf{1}_{\{X_0=x_0\}}] = \mu_0(x_0)P^n f(x_0)$. Le point 2 découle alors de la définition de l'espérance conditionnelle, voir le corollaire **??**. En multipliant (I.3) par $f(x_n)$ et en sommant sur $x_n \in E$, on obtient

$$\mathbb{E}[f(X_n)\mathbf{1}_{\{X_0=x_0,\dots,X_{n-1}=x_{n-1}\}}] = Pf(x_{n-1})\mathbb{P}(X_0 = x_0, \dots, X_{n-1} = x_{n-1}).$$

Le point 3 découle alors de la définition de l'espérance conditionnelle, voir le corollaire **??**. Le point 4 découle du point 1 pour la deuxième égalité et du point 2 pour la dernière. □

*Exemple* I.1.8. Soit $(U_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre $p \in ]0, 1[$. On désire calculer la probabilité $p_{\ell,n}$ d'observer une séquence de 1 consécutifs de longueur au moins $\ell$ dans la suite $(U_1, \dots, U_n)$. Pour cela, on indroduit la chaîne de Markov définie par $X_0 = 0$ et $X_{n+1} = (X_n + 1)\mathbf{1}_{\{U_{n+1}=1,X_n<\ell\}} + \ell\mathbf{1}_{\{X_n=\ell\}}$ pour $n \in \mathbb{N}$. Remarquons que dès que l'on observe une séquence de 1 consécutif de longueur $\ell$ la chaîne $X$ devient constante égale à $\ell$. En particulier, on a $p_{\ell,n} = \mathbb{P}(X_n = \ell) = P^n(0, \ell)$, où $P$ est la matrice de transition de la chaîne de Markov $(X_n, n \in \mathbb{N})$. On a $P(x, 0) = 1 - p$ et $P(x, x + 1) = p$ pour $x \in \{0, \dots, \ell - 1\}$, $P(\ell, \ell) = 1$ et tous les autres termes de la matrice sont nuls. On représente les valeurs de $p_{\ell,n}$ pour $n = 100$ et $p = 1/2$ dans le graphique I.2. On observe qu'avec une probabilité supérieure à $1/2$ on a une séquence de 1 consécutifs de longueur au moins 6 ($p_{6,100} \simeq 0.55$). ♢

Le théorème suivante assure que l'évolution future d'une chaîne de Markov ne dépend du passé qu'au travers de sa valeur présente. Cette propriété est appelée propriété de Markov.

**Théorème I.1.9** (Propriété de Markov des chaînes de Markov)**.** *Soit $(X_n, n \in \mathbb{N})$ une chaîne de Markov de matrice de transition $P$ issue de $\mu_0$. Soit $n_0 \in \mathbb{N}$. La loi de $(X_{n+n_0}, n \in \mathbb{N})$ sachant $(X_0, \dots, X_{n_0})$ est une chaîne de Markov de matrice de transition $P$ issue de $X_0$.*

*Démonstration.* On déduit de (I.3) que pour tous $n \geq 1$, $x_0, \dots, x_{n_0+n} \in E$

$$\mathbb{P}(X_{n_0} = x_{n_0}, \dots, X_{n_0+n} = x_{n_0+n}|X_0 = x_0, \dots, X_{n_0} = x_{n_0}) = \prod_{k=n_0+1}^{n_0+n} P(x_{k-1}, x_k).$$

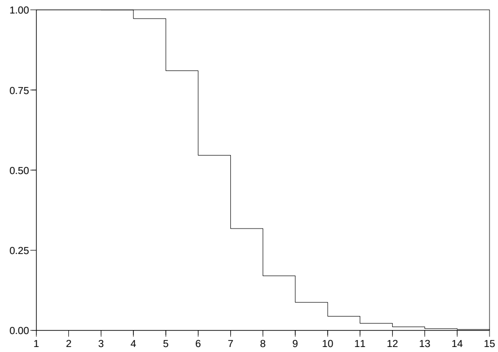


FIG. I.2 – Graphe de la fonction $x \mapsto \mathbb{P}(L_n \geq \lfloor x \rfloor)$, où $L_n$ est la longueur maximale des séquences de 1 consécutifs dans une suite de $n = 100$ variables aléatoires de Bernoulli indépendantes de paramètre $p = 1/2$.

Autrement dit, on a $\mathbb{P}(X_{n_0} = x_{n_0}, \ldots, X_{n_0+n} = x_{n_0+n} | X_0, \ldots, X_{n_0}) = F(X_{n_0})$ où $F(x) = \mathbb{P}(X_0 = x_{n_0}, \ldots, X_n = x_{n_0+n} | X_0 = x)$. On déduit donc de la proposition I.1.3 que conditionnellement à $(X_0, \ldots, X_{n_0})$, $(X_{n_0}, \ldots, X_{n_0+n})$ a même loi que $(X_0, \ldots, X_n)$ issu de $X_{n_0}$. Ceci étant vrai pour tout $n \geq 1$, on en déduit le résultat en utilisant la caractéristation de la loi d'une chaîne de Markov par sa matrice de transition et sa loi initiale. □

## I.1.2 Probabilités invariantes, réversibilité

Les probabilités invariantes jouent un rôle important dans l'étude des comportements asymptotiques des chaînes de Markov.

**Définition I.1.10.** *Une probabilité $\pi$ sur $E$ est appelée probabilité invariante, ou probabilité stationnaire, d'une chaîne de Markov de matrice de transition $P$ si $\pi = \pi P$.*

*Exercice* I.1.
Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov de matrice de transition $P$.

1. Montrer que $Z = (Z_n = X_{2n}, n \in \mathbb{N})$ est une chaîne de Markov. Préciser sa matrice de transition.
2. Vérifier que toute probabilité invariante pour $X$ est une probabilité invariante pour $Z$. En considérant un espace d'état à deux éléments, donner un contre-exemple pour la réciproque.

△

On considère $(X_n, n \in \mathbb{N})$ une chaîne de Markov de matrice de transtion $P$ issue d'une probabilité invariante $\pi$. On note $\mu_n$ la loi de $X_n$. On a $\mu_1 = \pi P = \pi$ et, en itérant, on obtient que $\mu_n = \pi$. Pour tout $n \in \mathbb{N}^*$, $X_n$ a même loi que $X_0$ : la loi de $X_n$ est donc constante ou stationnaire au cours du temps.

Supposons de plus que $\pi(x) > 0$ pour tout $x \in E$. Pour $x, y \in E$, on pose

$$Q(x, y) = \frac{\pi(y) P(y, x)}{\pi(x)} \cdot \tag{I.4}$$

Comme $\pi$ est une probabilité invariante, on a $\sum_{y \in E} \pi(y)P(y,x) = \pi(x)$ pour tout $x \in E$. En particulier la matrice $Q$ est une matrice stochastique. Pour tous $x, y \in E$, $n \geq 0$, on a

$$\mathbb{P}(X_n = y | X_{n+1} = x) = \frac{\mathbb{P}(X_n = y, X_{n+1} = x)}{\mathbb{P}(X_{n+1} = x)} = Q(x,y).$$

Autrement dit $\mathbb{P}(X_n = y | X_{n+1}) = Q(X_{n+1}, y)$ pour tout $y \in E$. Plus généralement, il est facile de vérifier que pour tous $k \in \mathbb{N}^*$, $y \in E$, on a $\mathbb{P}(X_n = y | X_{n+1}, \ldots, X_{n+k}) = Q(X_{n+1}, y)$. La matrice $Q$ s'interprète comme la matrice de transition de la chaîne de Markov $X$ après retournement du temps.

**Définition I.1.11.** *On dit qu'une chaîne de Markov de matrice de transition $P$, ou plus simplement la matrice $P$, est réversible par rapport à la probabilité $\pi$ si on a pour tous $x, y \in E$,*

$$\pi(x)P(x,y) = \pi(y)P(y,x). \tag{I.5}$$

En sommant (I.5) sur $x$, on en déduit le lemme suivant.

**Lemme I.1.12.** *Si une chaîne de Markov est réversible par rapport à la probabilité $\pi$, alors $\pi$ est une probabilité invariante.*

Les exemples I.1.38 et I.1.34 permettent d'illustrer le concept de chaînes de Markov réversibles.

Si $P$ est réversible par rapport à $\pi$, alors $Q$ définie par (I.4) est égal à $P$. Donc si une chaîne de Markov est réversible par rapport à une probabilité invariante, alors si elle est issue de cette probabilité invariante la chaîne et la chaîne après retournement du temps ont même loi. Plus précisément, si $(X_n, n \in \mathbb{N})$ est une chaîne de Markov réversible par rapport à $\pi$ et si la loi de $X_0$ est $\pi$, alors les vecteurs $(X_0, \ldots, X_{n-1}, X_n)$ et $(X_n, X_{n-1}, \ldots, X_0)$ ont même loi pour tout $n \in \mathbb{N}^*$.

### I.1.3 Irreductibilité, récurrence, transience, périodicité

**Définition I.1.13.** *Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov de matrice de transition $P$. On dit que $x$ est un état absorbant de la chaîne $X$ si $P(x,x) = 1$.*

En particulier si la chaîne de Markov atteint un de ses points absorbants, elle ne peut plus s'en échapper. Dans L'exemple I.1.35 sur le modèle de Wright-Fisher, on s'intéresse au temps d'atteinte des points absorbants.

*Exercice* I.2.
Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov de matrice de transition $P$. On utilise la convention $\inf \emptyset = +\infty$. On pose

$$\tau_1 = \inf\{k \geq 1; X_k \neq X_0\}.$$

1. Soit $x \in E$. Calculer la loi de $\tau_1$ conditionnellement à $X_0 = x$. Vérifier que, conditionnellement à $X_0 = x$, $\tau_1 = +\infty$ p.s. si $x$ est un point absorbant et sinon p.s. $\tau_1$ est fini.
2. Conditionnellement à $X_0 = x$, si $x$ n'est pas un point absorbant, calculer la loi de $X_{\tau_1}$.

On pose $S_0 = 0$, $Y_0 = X_0$ et on définit par récurrence sur $n \geq 1$ : $S_n = S_{n-1} + \tau_n$, et si $S_n < \infty$ : $Y_n = X_{S_n}$ et

$$\tau_{n+1} = \inf\{k \geq 1; X_{k+S_n} \neq X_{S_n}\}.$$

Soit $R = \inf\{n; \tau_n = +\infty\} = \inf\{n; S_n = +\infty\}$.

3. Montrer que si $X$ ne possède pas de points absorbants, alors p.s. $R = +\infty$.

On suppose que $X$ ne possède pas de points absorbants.

4. Montrer que $Y = (Y_n, n \in \mathbb{N})$ est une chaîne de Markov. Elle est appelée la chaîne trace associée à $X$. Montrer que sa matrice de transition $Q$ est définie par :

$$Q(x,y) = \frac{P(x,y)}{1-P(x,x)} \mathbf{1}_{\{x \neq y\}} \quad \text{pour } x, y \in E.$$

5. Soit $\pi$ une probabilité invariante pour $X$. On définit une mesure $\nu$ sur $E$ par

$$\nu(x) = \frac{\pi(x)(1-P(x,x))}{\sum_{y\in E} \pi(y)(1-P(y,y))}, \quad x \in E.$$

Vérifier que $\nu$ est une probabilité invariante pour $Y$.

△

**Définition I.1.14.** *On dit qu'une chaîne de Markov, ou sa matrice de transition, est irréductible si pour tous $x, y \in E$, la probabilité partant de $x$ d'atteindre $y$ est strictement positive, autrement dit : si pour tous $x, y \in E$, il existe $n = n_{x,y} \geq 1$ (dépendant a priori de $x$ et $y$) tel que $P^n(x,y) > 0$.*

La condition $P^n(x,y) > 0$ est équivalente à l'existence de $n \geq 1$, $x_0 = x$, $x_1, \ldots, x_n = y$ tels que $\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n | X_0 = x_0) = \prod_{k=1}^n P(x_{k-1}, x_k) > 0$.

Par exemple, la marche aléatoire symétrique simple sur $\mathbb{Z}$ de l'exemple I.1.4 et la chaîne de Markov de l'urne d'Ehrenfest de l'exemple I.1.38 sont irréductibles. Une chaîne possédant des états absorbants n'est pas irréductible (sauf si l'espace d'état est réduit à un point). Ainsi la chaîne de Markov du modèle de Wright-Fisher, voir l'exemple I.1.35 n'est pas irréductible.

*Exercice* I.3.
Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov de matrice de transition $P$ à valeurs dans $E$. On pose $Y_n = (X_{n-1}, X_n)$ pour $n \in \mathbb{N}^*$.

1. Montrer que $Y = (Y_n, n \in \mathbb{N}^*)$ est une chaîne de Markov à valeurs dans $E^2$. Donner sa matrice de transition, $Q$.
2. Donner un exemple où $Q$ n'est pas irréductible alors que $P$ est irréductible. Changer l'espace d'état de $Y$ pour que $Y$ soit irréductible si $X$ est irréductible.
3. On suppose que $X$ possède une probabilité invariante, $\pi$. En déduire une probabilité invariante pour $Y$.

△

On utilise la convention $\inf \emptyset = +\infty$. On définit le temps d'atteinte de $x$ par

$$T^x = \inf\{n \geq 1; X_n = x\}.$$

**Définition I.1.15.** *Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov. On dit qu'un état $x \in E$ est :*

– *transient si $\mathbb{P}_x(T_x = +\infty) > 0$,*
– *récurrent si $\mathbb{P}_x(T_x = +\infty) = 0$.*

*Une chaîne de Markov est dite transiente (resp. récurrente) si tous les états sont transients (resp. récurrents).*

On pose $N_x = \sum_{n\in\mathbb{N}} \mathbf{1}_{\{X_n = x\}}$ le nombre de visites du point $x$. La proposition suivante permet de caractériser les points récurrents et transients.

**Proposition I.1.16.** *Soit $X$ une chaîne de Markov de matrice de transition $P$. On a les propriétés suivantes.*

1. *Si $x$ est transient alors on a $\mathbb{P}_x(N^x < +\infty) = 1$, $\sum_{n\in\mathbb{N}} P^n(x,x) < \infty$ et conditionnellement à $\{X_0 = x\}$, $N^x$ est de loi géométrique de paramètre $\mathbb{P}_x(T^x = \infty)$.*
2. *Si $x$ est récurrent alors on a $\mathbb{P}_x(N^x = +\infty) = 1$ et $\sum_{n\in\mathbb{N}} P^n(x,x) = \infty$.*
3. *Si $X$ est irréductible alors $X$ est soit transiente soit récurrente. Dans le premier cas on a $\mathbb{P}(N^x < +\infty) = 1$ pour tout $x \in E$ et dans le second $\mathbb{P}(N^x = +\infty) = 1$ pour tout $x \in E$.*

*Démonstration.* On pose $p = \mathbb{P}_x(T^x = +\infty) = \mathbb{P}_x(N^x = 1)$. On note $T_n^x = \inf\{k \geq 0; \sum_{r=0}^k \mathbf{1}_{\{X_r = x\}} = n\}$ le $n$-ème temps de passage en $x$. (On a $T_2^x = T^x$ quand $X_0 = x$). Remarquons que $\{T_n^x < +\infty\} = \{N^x \geq n\}$. En décomposant suivant les valeurs de $T_n^x$, il vient pour $n \in \mathbb{N}^*$ :

$$\begin{aligned}
\mathbb{P}_x(N^x > n) &= \sum_{r\geq n} \mathbb{P}(N^x > n, T_n^x = r) \\
&= \sum_{r\geq n} \mathbb{P}\left(\sum_{i=1}^r \mathbf{1}_{\{X_i = x\}} = n, X_r = x, \sum_{\ell\in\mathbb{N}} \mathbf{1}_{\{X_{r+\ell} = x\}} > 1\right) \\
&= \sum_{r\geq n} \mathbb{P}\left(\sum_{i=1}^r \mathbf{1}_{\{X_i = x\}} = n, X_r = x\right) \mathbb{P}_x\left(\sum_{\ell\in\mathbb{N}} \mathbf{1}_{\{X_\ell = x\}} > 1\right) \\
&= \mathbb{P}_x(N^x \geq n)\mathbb{P}_x(N^x > 1) \qquad \text{(I.6)} \\
&= \mathbb{P}_x(N^x > n-1)(1-p).
\end{aligned}$$

où l'on a utilisé la propriété de Markov à l'instant $r$ pour la troisième égalité. En itérant et en utilisant que $\mathbb{P}_x(N^x > 0) = 1$, on obtient $\mathbb{P}_x(N^x > n) = (1-p)^n$ pour $n \in \mathbb{N}$, puis $\mathbb{P}_x(N^x = n) = \mathbb{P}_x(N^x > n-1) - \mathbb{P}_x(N^x > n) = p(1-p)^{n-1}$ pour $n \in \mathbb{N}^*$. Remarquons enfin que $\mathbb{E}_x[N^x] = \sum_{n\in\mathbb{N}} \mathbb{P}_x(X_n = x) = \sum_{n\in\mathbb{N}} P^n(x,x)$.

Si $x$ est transient, on a $p > 0$ et, conditionnellement à $\{X_0 = x\}$, $N^x$ est de loi géométrique de paramètre $p > 0$. En particulier son espérance est finie.

Si $x$ est récurrent, on a $p = 0$ et donc $\mathbb{P}(N^x \in \mathbb{N}) = 0$. Ceci implique que $\mathbb{P}(N^x = +\infty) = 1$ et son espérance est infinie.

Il reste le point 3 à démontrer. Soit $x, y \in E$. Comme la chaîne est irréductible, il existe $n_1$ et $n_2$ tels que $P^{n_1}(y,x) > 0$ et $P^{n_2}(x,y) > 0$. En particulier on a pour tout $n \in \mathbb{N}$

$$P^{n+n_1+n_2}(y,y) \geq P^{n_1}(y,x)P^n(x,x)P^{n_2}(x,y) \text{ et } P^{n+n_1+n_2}(x,x) \geq P^{n_2}(x,y)P^n(y,y)P^{n_1}(y,x). \qquad \text{(I.7)}$$

Ceci implique que les deux séries $\sum_{n\in\mathbb{N}} P^n(x,x)$ et $\sum_{n\in\mathbb{N}} P^n(y,y)$ sont de même nature. Ainsi les deux états $x$ et $y$ sont soit tous les deux transients soit tous les deux récurrents. On en déduit qu'une chaîne irréductible est soit récurrente soit transiente.

Si la chaîne est transiente, on a en décomposant suivant la valeur de $T^x$ et en utilisant la propriété de Markov

$$\mathbb{P}(N^x = +\infty) = \sum_{n\in\mathbb{N}} \mathbb{P}(T^x = n)\mathbb{P}_x(N^x = +\infty) = 0.$$

Si la chaîne est récurrente, on a en décomposant suivant la valeur de $T^x$ et en utilisant la propriété de Markov

$$\mathbb{P}(N^x < +\infty) = \mathbb{P}(T^x = +\infty) + \sum_{n\in\mathbb{N}} \mathbb{P}(T^x = n)\mathbb{P}_x(N^x < +\infty) = \mathbb{P}(T^x = +\infty). \qquad \text{(I.8)}$$

Soit $y \in E$ distinct de $x$ et $N_x^y = \text{Card}\,(\{n < T^x; X_n = y\})$ le nombre de visites de $y$ avant de visiter $x$. On note $p = \mathbb{P}_y(T_x < T_y)$ la probabilité de visiter $x$ avant de retourner en $y$. Un calcul similaire à (I.6) assure que

$$\mathbb{P}_y(N_x^y > n) = \mathbb{P}_y(N_x^y \geq n)\mathbb{P}_y(N_x^y > 1) = \mathbb{P}_y(N_x^y > n-1)(1-p).$$

En itérant et en utilisant que $\mathbb{P}_y(N_x^y > 0) = 1$, on obtient $\mathbb{P}_y(N_x^y > n) = (1-p)^n$ pour $n \in \mathbb{N}$ et donc $\mathbb{P}_y(N_x^y = +\infty) = \mathbf{1}_{\{p=0\}}$. La chaîne étant irréductible, il existe $n > 0$ tel que $\mathbb{P}_y(X_n = x) > 0$. Comme $\{X_n = x\} \subset \{T^x < +\infty\}$, on en déduit que $\mathbb{P}_y(T^x = \infty) < 1$. Comme $\{N_x^y = +\infty\} = \{T^x = +\infty\}$, il vient $\mathbb{P}_y(N_x^y = +\infty) < 1$. Ceci implique $p > 0$ puis $\mathbb{P}_y(N_x^y = +\infty) = 0$ et $\mathbb{P}_y(T^x = +\infty) = 0$. Comme $\mathbb{P}(T^x = +\infty) = \sum_{y \in E} \mathbb{P}(X_0 = y)\mathbb{P}_y(T^x = +\infty) = 0$, on déduit de (I.8) que $\mathbb{P}(N^x < +\infty) = 0$. Ceci termine la démonstration du point 3. □

*Exercice* I.4.
On considère la marche aléatoire symétrique simple sur $\mathbb{Z}^d$, $d \geq 1$, définie par $S_0 = 0$, $S_n = \sum_{k=1}^n Z_k$ pour $n \in \mathbb{N}^*$, où les variables aléatoires $(Z_n, n \in \mathbb{N}^*)$ sont indépendantes de loi uniforme sur $\{z \in \mathbb{Z}^d; |z| = 1\}$, et $|\cdot|$ désigne la norme euclidienne sur $\mathbb{R}^d$. (Le cas $d = 1$ correspond à l'exemple I.1.4.)

1. Vérifier que la chaîne $(S_n, n \in \mathbb{N})$ est irréductible.
2. On suppose $d = 1$. Calculer $\mathbb{P}_0(S_{2n} = 0)$. En utilisant la formule de Stirling ($n! \sim \sqrt{2\pi} n^{n+1/2}\,\mathrm{e}^{-n}$ pour $n$ grand), montrer que la marche aléatoire est récurrente.
3. On suppose $d = 2$. Calculer $\mathbb{P}_0(S_{2n} = 0)$ et montrer que la marche aléatoire est récurrente. On pourra utiliser que $(1+x)^n(1+x)^n = (1+x)^{2n}$ pour calculer $\sum_{k=0}^n \frac{(n!)^2}{\left(k!(n-k)!\right)^2}$.
4. On suppose $d = 3$. Montrer que la marche aléatoire est transiente. On pourra utiliser, après l'avoir vérifié, que pour $n \geq 3$ :
$$\sum_{i+j+k=n} \left(\frac{n!}{3^n i! j! k!}\right)^2 \leq \left(\sum_{i+j+k=n} \frac{n!}{3^n i! j! k!}\right) \max_{i+j+k=n} \left(\frac{n!}{3^n i! j! k!}\right) \leq \frac{n!}{3^n (\lfloor n/3 \rfloor !)^3}.$$
5. Déduire de la question précédente que la marche aléatoire symétrique en dimension $d \geq 3$ est transiente.

△

Dans l'exemple I.1.4 de la marche aléatoire symétrique, remarquons que, si $S_0$ est pair, alors $S_n$ est pair si et seulement si $n$ est pair. On assiste à un phénomène périodique. En particulier $P^n(0,0)$ est nul si $n$ est impair. Ceci motive la définition suivante.

**Définition I.1.17.** *Soit $X$ une chaîne de Markov de matrice de transition $P$. La période d'un état $x \in E$ est le PGCD de $\{n \in \mathbb{N}^*; P^n(x,x) > 0\}$. Un état est dit apériodique si sa période est 1, sinon il est dit périodique. Une chaîne de Markov est dite apériodique si tous ses états sont apériodiques.*

*Exercice* I.5.
Montrer que pour la chaîne de Markov de l'urne d'Ehrenfest, voir l'exemple I.1.38, les états sont de période 2. △

*Exercice* I.6.
On note $J = \{n \in \mathbb{N}^*; \mathbb{P}_x(T^x = n) > 0\}$. Montrer que $I = \{n \in \mathbb{N}^*; P^n(x,x) > 0\}$ est stable par addition, et que c'est le plus petit sous-ensemble de $\mathbb{N}^*$ stable par addition contenant $J$. △

**Proposition I.1.18.** *Soit $X$ une chaîne de Markov de matrice de transition $P$. On a les propriétés suivantes.*

1. *Si $x \in E$ est apériodique, alors il existe $n_0 \in \mathbb{N}$ tel que $P^n(x,x) > 0$ pour tout $n \geq n_0$.*

*2. Si $X$ est irréductible, alors elle est apériodique si au moins un état est apériodique.*

*Démonstration.* Soit $x \in E$ apériodique. On note $I = \{n \in \mathbb{N}^*; P^n(x,x) > 0\}$. Comme $P^{n+m}(x,x) \geq P^n(x,x)P^m(x,x)$, on en déduit que $I$ est stable par addition. Par hypothèse, il existe $n_1, \ldots, n_K \in I$ premiers entre eux. Le théorème de Bézout assure qu'il existe $a_1, \ldots, a_K \in \mathbb{Z}$ tel que $\sum_{k=1}^K a_k n_k = 1$. On pose $n_+ = \sum_{k=1; a_k>0}^K a_k n_k$ et $n_- = \sum_{k=1; a_k<0}^K |a_k| \, n_k$, de sorte que $n_+, n_- \in I$ et $n_+ - n_- = 1$. Soit $n \geq n_-^2$. On considère la division euclidienne de $n$ par $n_-$ : il existe $r \in \{0, \ldots, n_- - 1\}$ et $q \geq n_-$ entier tels que

$$n = qn_- + r = qn_- + r(n_+ - n_-) = (q - r)n_- + rn_+.$$

Comme $q - r \geq 0$ et que $I$ est stable par addition, on a $n \in I$. Ceci démontre le point 1.

Si $P^n(x,x) > 0$ pour tout $n \geq n_0$, alors par définition $x$ est apériodique. Le point 2 découle alors du point 1 et de (I.7). □

*Exercice* I.7.
Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov irréductible. Montrer que s'il existe $x$ de période $d > 1$, alors on peut décomposer l'espace d'état $E$ en une partition à $d$ sous ensembles $C_1, \ldots, C_d = C_0$, tels que pour tout $k \in \{1, \ldots, d\}$,

$$\mathbb{P}(X_1 \in C_k \mid X_0 \in C_{k-1}) = 1.$$

△

Le lemme suivant sera utilisé au paragraphe I.1.5.

**Lemme I.1.19.** *Soit $X = (X_n, n \in \mathbb{N})$ et $Y = (Y_n, n \in \mathbb{N})$ deux chaînes de Markov indépendantes irréductibles à valeurs dans $E$. Alors $Z = ((X_n, Y_n), n \in \mathbb{N})$ est une chaîne de Markov à valeurs dans $E^2$. Si $\pi$ (resp. $\nu$) est une probabilité invariante pour $X$ (resp. $Y$), alors $\pi \otimes \nu$ est une probabilité invariante pour $Z$. Si $X$ ou $Y$ est apériodique, alors $Z$ est irréductible.*

*Démonstration.* Il est élémentaire de vérifier que $Z$ est une chaîne de Markov de matrice de transition $P(z, z') = P_1(z_1, z_1')P_2(z_2, z_2')$ où $P_1$ (resp. $P_2$) est la matrice de transition de $X$ (resp. $Y$) et $z = (z_1, z_2), z' = (z_1', z_2') \in E^2$.

Si $\pi$ (resp. $\nu$) est une probabilité invariante pour $X$ (resp. $Y$), alors on a, pour $z = (z_1, z_2) \in E^2$.

$$\pi \otimes \nu P(z') = \sum_{z_1, z_2 \in E} \pi(z_1)\nu(z_2)P((z_1, z_2), (z_1', z_2')) = \sum_{z_1, z_2 \in E} \pi(z_1)P_1(z_1, Z_1')\nu(z_2)P_2(z_2, z_2') = \pi \otimes \nu(z).$$

Et donc $\pi \otimes \nu$ est une probabilité invariante pour $Z$.

Supposons que $X$ est apériodique et montrons que $Z$ est irréductible. Soit $z = (z_1, z_2), z' = (z_1', z_2') \in E^2$. Les chaînes $X$ et $Y$ étant irréductibles, il existe $n_1, n_2, n_3 \in \mathbb{N}^*$ tels que $P_1^{n_1}(z_1, z_1') > 0$, $P_2^{n_2}(z_2, z_2') > 0$ et $P_2^{n_3}(z_2', z_2') > 0$. La proposition I.1.18 assure que $P^{kn_3+n_2-n_1}(z_1', z_1') > 0$ pour $k \in \mathbb{N}^*$ suffisament grand. Ainsi, on obtient que

$$\begin{aligned} P^{kn_3+n_2}(z, z') &= P_1^{kn_3+n_2}(z_1, z_1')P_2^{kn_3+n_2}(z_2, z_2') \\ &\geq P_1^{n_1}(z_1, z_1')P_1^{kn_3+n_2-n_1}(z_1', z_1')P_2^{n_2}(z_2, z_2')P_2^{n_3}(z_2', z_2')^k > 0. \end{aligned}$$

La chaîne $Z$ est donc irréductible. □

### I.1.4 Théorèmes asymptotiques

Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov à valeurs dans $E$. On rappelle la définition du temps d'atteinte de $x$ : $T^x = \inf\{n \geq 1; X_n = x\}$. Notons que, comme $T^x \geq 1$, on a $\mathbb{E}_x[T^x] \geq 1$. On pose

$$\pi(x) = \frac{1}{\mathbb{E}_x[T^x]} \in [0, 1]. \tag{I.9}$$

Pour une chaîne de Markov irréductible transiente on a, pour tout $x \in E$, $\mathbb{P}_x(T^x = +\infty) > 0$ et donc $\pi = 0$.

**Définition I.1.20.** *Un état $x$ récurrent est dit récurrent positif si $\pi(x) > 0$ et récurrent nul si $\pi(x) = 0$. Une chaîne est dite récurrente positive (resp. nulle) si tous ses états sont récurrents positifs (resp. nuls).*

On dit qu'un évènement $A$ est presque sûr (p.s.) pour une chaîne de Markov $X$ de matrice de transition $P$ si $\mathbb{P}_x(A) = 1$ pour tout $x \in E$ et donc si $\mathbb{P}(A) = 1$ quelque soit la loi initiale de la chaîne de Markov.

Les deux théorèmes fondamentaux suivants sur le comportement asymptotique d'une chaîne de Markov irréductible seront démontrés au paragraphe I.1.5.

**Théorème I.1.21.** *Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov irréductible à valeurs dans $E$.*

1. *La chaîne de Markov $X$ est soit transiente soit récurrente nulle soit récurrente positive.*
2. *Si la chaîne de Markov est transiente ou récurrente nulle, elle ne possède pas de probabilité invariante. De plus on a $\pi = 0$.*
3. *Pour tout $x \in E$, on a*

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{X_k = x\}} \xrightarrow[n\to+\infty]{p.s.} \pi(x). \tag{I.10}$$

**Théorème I.1.22** (Théorème ergodique)**.** *Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov irréductible à valeurs dans $E$ récurrente positive.*

1. *Le vecteur $\pi$ défini par (I.9) est une probabilité et c'est l'unique probabilité invariante de $X$. De plus on a $\pi(x) > 0$ pour tout $x \in E$.*
2. *Pour toute fonction $f$ définie sur $E$, telle que $f \geq 0$ ou $(\pi, |f|) < \infty$, on a*

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \xrightarrow[n\to+\infty]{p.s.} (\pi, f). \tag{I.11}$$

3. *Si de plus $X$ est apériodique alors on a $X_n \xrightarrow[n\to+\infty]{loi} \pi$ et donc $\lim_{n\to+\infty} \mathbb{P}(X_n = x) = \pi(x)$ pour tout $x \in E$.*

En particulier pour une chaîne de Markov récurrente positive la moyenne temporelle converge p.s. vers la moyenne spatialle par rapport à la probabilité invariante. Le théorème ergodique indique que le comportement de la chaîne de Markov en régime stationnaire correspond au comportement asymptotique.

Si $X$ est transiente, alors on a $\lim_{n\to+\infty} \mathbb{P}(X_n = x) = 0$ pour tout $x \in E$. On peut aussi montrer que si $X$ est récurrente nulle on a $\lim_{n\to+\infty} \mathbb{P}(X_n = x) = 0$ pour tout $x \in E$.

*Exercice* I.8.
Suite de l'exemple I.1.5. On suppose que $q = 1$ et que la demande peut prendre seulement les trois valeurs suivantes 0,1 et 2 et avec des probabilités strictement positives.

1. Vérifier que la chaîne de Markov $X$ est irréductible.
2. Calculer la probabilité invariante si $\mathbb{E}[D_1] > 1$ (i.e. si $p_2 > p_0$).
3. En déduire que si $\mathbb{E}[D_1] > 1$, alors $\frac{1}{n}\sum_{k=1}^n X_k$ possède une limite p.s. et la calculer.

△

Nous donnons le corollaire suivant pour le cas $E$ fini.

**Corollaire I.1.23.** *Une chaîne de Markov irréductible à valeurs dans un espace $E$* ***fini*** *est positive récurrente : $\pi$ défini par (I.9) est son unique probabilité invariante, $\pi(x) > 0$ pour tout $x \in E$ et on a (I.11) pour toute fonction $f$.*

*Démonstration.* En sommant (I.10) sur $x$, on obtient que $\sum_{x\in E}\pi(x) = 1$. Ainsi la chaîne est récurrente positive d'après le théorème I.1.21, point 1 et 2. La suite découle du théorème I.1.22. □

La convergence des moyennes empiriques, voir (I.11), pour les chaînes de Markov irréductible récurrente positive est à comparer avec la loi forte des grands nombres ; mais ici les variables aléatoires $(X_n, n \in \mathbb{N})$ sont dépendantes. Remarquons également que la probabilité qui apparaît à la limite, dans le membre de droite de (I.11) n'est pas liée à la loi initiale de la chaîne de Markov. Une propriété importante des chaînes de Markov irréductible est cet oubli de la condition initiale. Tout comme le théorème central limite (TCL) pour la loi forte des grands nombres, on peut estimer la vitesse de convergence dans (I.11) en étudiant la convergence en loi de $\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^n f(X_k) - (\pi, f)\right)$. Sous des hypothèses raisonnables, on peut montrer que la limite existe et est gaussienne. Mais la variance de la loi gaussienne limite est en général difficile à calculer, voir la remarque qui suit. Contrairement au TCL, on ne peut pas l'estimer à partir des observations, ni donc construire un intervalle de confiance comme dans la méthode de Monte-Carlo.

*Remarque* I.1.24. Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov irréductible, de matrice de transtion $P$, à valeurs dans $E$ fini. Le corollaire I.1.23 assure qu'il existe une unique probabilité invariante $\pi$. Soit $f$ une fonction définie sur $E$ à valeurs dans $\mathbb{R}$. On pose $I_n(f) = \frac{1}{n}\sum_{k=1}^n f(X_k)$. On a donc p.s. $\lim_{n\to+\infty} I_n(f) = (\pi, f)$. Calculons la variance de $\sqrt{n}I_n(f)$. On a

$$\mathrm{Var}(\sqrt{n}I_n(f)) = \frac{1}{n}\sum_{k,\ell=1}^n \mathrm{Cov}(f(X_k), f(X_\ell)) = \frac{1}{n}\sum_{k=1}^n \left(\mathrm{Var}(f(X_k)) + 2\sum_{j=1}^{n-k}\mathrm{Cov}(f(X_k), f(X_{k+j}))\right).$$

Comme la limite d'une moyenne temporelle revient à intégrer par rapport à la probabilité invariante, intuitivement $\mathrm{Var}(\sqrt{n}I_n(f))$ converge vers $\sigma(f)^2 = \mathrm{Var}_\pi(f(X_0)) + 2\sum_{j\in\mathbb{N}^*}\mathrm{Cov}_\pi(f(X_0), f(X_j))$ où l'indice $\pi$ signifie que la loi de $X_0$ est $\pi$. Il vient, en utilisant le point 3 de la proposition I.1.7,

$$\sigma(f)^2 = \Big(\pi, \big(f - (\pi, f)\big)^2\Big) + \mathbb{E}_\pi\left[\big(f(X_0) - (\pi, f)\big)P\sum_{j\in\mathbb{N}} P^j\big(f - (\pi, f)\big)(X_0)\right].$$

La fonction $F = \sum_{j\in\mathbb{N}} P^j\big(f - (\pi, f)\big)$ est une solution formelle de l'équation suivante, dite équation de Poisson :

$$F(x) - PF(x) = f(x) - (\pi, f), \quad \text{pour tout } x \in E. \tag{I.12}$$

On obtient alors

$$\sigma(f)^2 = \big(\pi, (F - PF)^2\big) + 2\big(\pi, (F - PF)PF\big) = (\pi, F^2) - (\pi, (PF)^2). \tag{I.13}$$

En fait, on peut montrer le résultat suivant[1]. Pour toute fonction $f$ il existe, à une constante additive près, une unique solution $F$ de l'équation de Poisson (I.12) et on a la convergence en loi suivante :

$$\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} f(X_k) - (\pi, f)\right) \xrightarrow[n\to+\infty]{\textbf{loi}} \mathcal{N}(0, \sigma(f)^2),$$

où $\sigma(f)^2$ est donné par (I.13). En général calculer la solution d'une équation de Poisson est aussi difficile que de calculer la probabilité invariante. Des résultats similaires existent pour les chaînes irréductibles récurrentes positives à valeurs dans $E$ discret dénombrable. $\Diamond$

### I.1.5 Démonstration des théorèmes asymptotiques

Soit $X = (X_n, n \geq 0)$ une chaîne de Markov à valeurs dans un espace discret $E$. On suppose que $X$ est **irréductible**.

On reprend les notations du paragraphe précédent. En particulier $\pi$ est défini par (I.9). Le lemme suivant assure que s'il existe une probabilité invariante alors elle est égale à $\pi$.

**Lemme I.1.25.** *Si (I.10) est vérifié et s'il existe une probabilité invariante $\nu$, alors $\nu = \pi$.*

*Démonstration.* Supposons qu'il existe une probabilité invariant $\nu$. Après avoir remarqué que le membre de gauche de (I.10) est borné par 1, par convergence dominée, on obtient, en prenant l'espérance dans (I.10) avec la probabilité $\nu$ comme loi initiale que

$$\frac{1}{n}\sum_{k=1}^{n} \nu P^k(x) \xrightarrow[n\to+\infty]{} \pi(x).$$

Comme $\nu$ est une probabilité invariante, on a $\nu P^k(x) = \nu(x)$ pour tout $x \in E$. Donc, il vient $\nu = \pi$. $\square$

Le lemme suivant concerne les chaînes transientes.

**Lemme I.1.26.** *Soit $X = (X_n, n \geq 0)$ une chaîne de Markov irréductible transiente. Alors on a $\pi = 0$, (I.10) est vérifiée, et $X$ ne possède pas de probabilité invariante.*

*Démonstration.* La proposition I.1.16 assure que, pour tout $x \in E$, $\mathbb{P}_x(T^x = +\infty) > 0$ et donc $\mathbb{E}_x[T^x] = +\infty$. Ceci implique par définition que $\pi = 0$. Le point 3 de la proposition I.1.16 assure que la somme $\sum_{k\in\mathbb{N}} \mathbf{1}_{\{X_k=x\}}$ est p.s. finie. On en déduit donc que (I.10) est vérifiée. Comme $\pi = 0$ et que (I.10) est vérifiée, on déduit du lemme I.1.25 qu'il n'existe pas de probabilité invariante. $\square$

On suppose dorénavant que $X$ est **récurrente**.

On pose $N_n^x = \sum_{k=0}^{n} \mathbf{1}_{\{X_k=x\}}$ le nombre de visites du point $x$ jusqu'à l'instant $n$. Le point 3 de la proposition I.1.16 assure que p.s. $\lim_{n\to+\infty} N_n^x = +\infty$. On peut donc définir p.s. les temps de passage en $x$ et les inter-temps de passage par : $S_0^x = 0$ et, pour $n \in \mathbb{N}$,

$$S_{n+1}^x = \inf\{k \geq 0; N_k^x = n+1\} \quad \text{et} \quad T_n^x = S_{n+1}^x - S_n^x.$$

[1] S. P. Meyn et R. L. Tweedie. *Markov chains and stochastic stability.* Springer-Verlag, 1993.

Remarquons que $\{X_0 = x\} = \{S_1^x = 0\} = \{T_0^x = 0\}$. On introduit les excursions hors de l'état $x$, c'est-à-dire les variables aléatoires $(Y_n, n \in \mathbb{N})$ à valeurs dans l'espace des trajectoires de longueur finie $E^{\text{traj}} = \cup_{k\geq 1}\{k\} \times E^{k+1}$ définies par

$$Y_n = (T_n^x, X_{S_n^x}, X_{S_n^x+1}, \dots, X_{S_{n+1}^x}). \tag{I.14}$$

Le lemme suivant est la clef de la démonstration que nous avons choisi de présenter des théorèmes asymptotiques pour les chaînes de Markov.

**Lemme I.1.27.** *Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov irréductible récurrente. Les variables aléatoires $(Y_n, n \in \mathbb{N})$, définies par (I.14) sont indépendantes. Les variables $(Y_n, n \in \mathbb{N}^*)$ ont même loi que $Y_1$ conditionnellement à $\{X_0 = x\}$.*

*Démonstration.* Pour $y = (r, x_0, \dots, x_r) \in E^{\text{traj}}$, on note $t_y = r$ et $y(k) = x_k$ pour $0 \leq k \leq r$. Soit $n \geq 1$ et $y_0, \dots, y_n \in E^{\text{traj}}$. On dit que la suite d'excursions $(y_0, \dots, y_n)$ est compatible si $y_0(t_{y_0}) = x$ et $y_k(t_{y_k}) = y_k(0) = x$ pour tout $1 \leq k \leq n$ (c'est-à-dire toutes les excursions se terminent en $x$, et toutes, sauf éventuellement la première, débutent en $x$). Si la condition de compatibilité n'est pas vérifiée, alors on a $\mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n) = 0$ ainsi que $\mathbb{P}(Y_0 = y_0)\prod_{k=1}^n \mathbb{P}_x(Y_1 = y_k) = 0$.

On suppose que la condition de compatibilité est vérifiée. En utilisant la propriété de Markov à l'instant $s = \sum_{k=0}^{n-1} t_{y_k}$ et en remarquant que sur l'événement $\{Y_0 = y_0, \dots, Y_n = y_n\}$ on a $S_n^x = s$ et donc $X_s = x$, on obtient

$$\mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n) = \mathbb{P}(Y_0 = y_0, \dots, Y_{n-1} = y_{n-1})\mathbb{P}_x(Y_1 = y_n).$$

Si la suite $(y_0, \dots, y_n)$ est compatible, alors il en est de même des suite $(y_0, \dots, y_k)$ pour $1 \leq k \leq n$. En itérant le calcul précédent, on obtient que

$$\mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n) = \mathbb{P}(Y_0 = y_0)\prod_{k=1}^n \mathbb{P}_x(Y_1 = y_k). \tag{I.15}$$

On a vu que cette relation est également vraie si la suite $(y_0, \dots, y_n)$ n'est pas compatible. On a donc obtenu l'égalité (I.15) pour tout $n \geq 1$ et $y_0, \dots, y_n \in E^{\text{traj}}$. On conclut alors à l'aide des définitions **??** et **??**. $\square$

Nous pouvons maintenant démontrer (I.10) pour les chaînes irréductibles récurrentes. Ce résultat et le lemme I.1.26 impliquent le point 3 du théorème I.1.21.

**Proposition I.1.28.** *Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov irréductible récurrente. Alors (I.10) est vérifiée.*

*Démonstration.* Comme $T_0^x$ est fini et que $S_n^x = \sum_{k=0}^{n-1} T_k^x$, où les variables aléatoires $(T_n^x, n \in \mathbb{N}^*)$ sont positives indépendantes de même loi que $T^x$ conditionnellement à $\{X_0 = x\}$, on déduit de la loi forte des grands nombres, et plus précisément de la proposition **??** que

$$\frac{S_n^x}{n} \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}_x[T^x]. \tag{I.16}$$

Par construction, on a $S_{N_n^x}^x \leq n < S_{N_n^x+1}^x$. Ainsi on a $\dfrac{N_n^x}{N_n^x+1}\dfrac{N_n^x+1}{S_{N_n^x+1}^x} \leq \dfrac{N_n^x}{n} \leq \dfrac{N_n^x}{S_{N_n^x}^x}$. Comme p.s. $\lim_{n\to+\infty} N_n^x = +\infty$, on déduit de (I.16) que p.s.

$$\lim_{n\to+\infty} \frac{N_n^x}{n} = \frac{1}{\mathbb{E}_x[T^x]} = \pi(x).$$

Ceci termine la démonstration de (I.10). $\square$

Le lemme suivant et le point 3 de la proposition I.1.16 impliquent le point 1 du théorème I.1.21.

**Lemme I.1.29.** *Soit* $X = (X_n, n \in \mathbb{N})$ *une chaîne de Markov irréductible récurrente. Alors elle est soit récurrente positive soit récurrente nulle.*

*Démonstration.* Après avoir remarqué que le membre de gauche de (I.10) est borné par 1, par convergence dominée, on obtient, en prenant l'espérance dans (I.10) avec la masse de Dirac en $x$ comme loi initiale que pour tout $x \in E$,

$$\frac{1}{n}\sum_{k=1}^{n} P^k(x,x) \xrightarrow[n\to+\infty]{} \pi(x).$$

Par ailleurs, comme la chaîne de Markov est irréductible, on déduit de (I.7) que si la limite ci-dessus est nulle pour un $x$ donné, alors elle est nulle pour tout $x \in E$. On en déduit donc que soit $\pi = 0$ soit $\pi(x) > 0$ pour tout $x \in E$. □

Le lemme suivant et le lemme I.1.26 impliquent le point 2 du théorème I.1.21. Sa démonstration découle du lemme I.1.25 et du fait que $\pi = 0$ pour les chaînes récurrentes nulles.

**Lemme I.1.30.** *Soit* $X = (X_n, n \in \mathbb{N})$ *une chaîne de Markov irréductible récurrente nulle. Alors elle ne possède pas de probabilité invariante.*

Le théorème I.1.21 est donc démontré.

On suppose dorénavant que $X$ est **récurrente positive**.

**Proposition I.1.31.** *Soit* $X = (X_n, n \in \mathbb{N})$ *une chaîne de Markov irréductible récurrente positive. Le vecteur* $\pi$ *défini par (I.9) est une probabilité et pour toute fonction* $f$ *définie sur* $E$*, telle que* $f \geq 0$ *ou* $(\pi, |f|) < \infty$*, on a (I.11).*

*Démonstration.* On reprend les notations du lemme I.1.27. Soit $f$ une fonction réelle positive finie définie sur $E$. On pose, pour $y = (r, x_0, \dots, x_r) \in E^{\text{traj}}$,

$$F(y) = \sum_{k=1}^{r} f(x_k).$$

Le lemme I.1.27 assure que les variables aléatoires $(F(Y_n), n \in \mathbb{N}^*)$ sont positives, indépendantes et de même loi que $F(Y_1)$ conditionnellement à $\{X_0 = x\}$. Comme $F(Y_0)$ est positif et fini, on déduit de la loi forte des grands nombres, voir la proposition **??**, que p.s. $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n} F(Y_k) = \mathbb{E}_x[F(Y_1)]$.

Par construction, on a $\displaystyle\frac{1}{S_n^x}\sum_{i=1}^{S_n^x} f(X_i) = \frac{n}{S_n^x}\frac{1}{n}\sum_{k=1}^{n} F(Y_k)$ (au moins pour $n \geq 2$). On déduit de (I.16) que p.s.

$$\lim_{n\to\infty} \frac{1}{S_n^x}\sum_{i=1}^{S_n^x} f(X_i) = \pi(x)\mathbb{E}_x[F(Y_1)].$$

Par construction, on a $S^x_{N_n^x} \leq n < S^x_{N_n^x+1}$. Pour $n \geq 2$, on a les inégalités

$$\frac{S^x_{N_n^x}}{S^x_{N_n^x+1}}\frac{1}{S^x_{N_n^x}}\sum_{i=1}^{S^x_{N_n^x}} f(X_i) \leq \frac{1}{n}\sum_{i=1}^{n} f(X_i) \leq \frac{S^x_{N_n^x+1}}{S^x_{N_n^x}}\frac{1}{S^x_{N_n^x+1}}\sum_{i=1}^{S^x_{N_n^x+1}} f(X_i).$$

Comme p.s. $\lim_{n\to\infty} N_n^x = +\infty$ et d'après (I.16) $\lim_{n\to\infty} S_n^x = +\infty$ p.s. ainsi que $\lim_{n\to\infty} S_n^x/S_{n+1}^x = 1$ p.s., on en déduit que

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \xrightarrow[n\to+\infty]{\textbf{p.s.}} \pi(x)\mathbb{E}_x[F(Y_1)]. \tag{I.17}$$

En choisissant $f = \mathbf{1}_{\{y\}}$, on déduit de (I.10) que

$$\pi(y) = \pi(x)\mathbb{E}_x\Big[\sum_{k=1}^{T^x} \mathbf{1}_{\{X_k=y\}}\Big]. \tag{I.18}$$

En sommant sur $y \in E$, il vient par convergence monotone, $\sum_{y\in E}\pi(y) = \pi(x)\mathbb{E}_x[T^x] = 1$. On en déduit que $\pi = (\pi(x), x \in E)$ est une probabilité. Enfin remarquons que grâce à (I.18), et par convergence monotone,

$$\pi(x)\mathbb{E}_x[F(Y_1)] = \sum_{y\in E} f(y)\pi(x)\mathbb{E}_x\Big[\sum_{k=1}^{T^x}\mathbf{1}_{\{X_k=y\}}\Big] = \sum_{y\in E} f(y)\pi(y) = (\pi, f).$$

On déduit alors (I.11) de (I.17).

Enfin si $f$ est de signe quelconque et $(\pi, |f|)$ est fini, on utilise (I.11) avec $f_+(x) = \max(f(x), 0)$ et $f_-(x) = \max(-f(x), 0)$ et on en fait la différence pour obtenir (I.11). La condition $(\pi, |f|) < +\infty$ assure que $(\pi, f_+) - (\pi, f_-)$ a un sens et est égal à $(\pi, f)$. □

La proposition suivante et la proposition I.1.31 terminent la démonstration des points 1 et 2 du théorème ergodique I.1.22.

**Proposition I.1.32.** *Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov irréductible récurrente positive. Le vecteur $\pi$ défini par (I.9) est l'unique probabilité invariante de $X$.*

*Démonstration.* Vérifions que $\pi$ est une probabilité invariante. Soit $\nu$ la loi de $X_0$. On pose

$$\bar{\nu}_n(x) = \frac{1}{n}\sum_{k=1}^{n} \nu P^k(x).$$

Par convergence dominée, on déduit de (I.10), en prenant l'espérance, que $\lim_{n\to\infty}\bar{\nu}_n(x) = \pi(x)$ pour tout $x \in E$. Soit $f$ bornée. Par convergence dominée, en prenant l'espérance dans (I.11), il vient

$$(\bar{\nu}_n, f) \xrightarrow[n\to\infty]{} (\pi, f).$$

En choisissant $f(\cdot) = P(\cdot, y)$, on a $(\bar{\nu}_n, f) = \bar{\nu}_n P(y) = \frac{n+1}{n}\bar{\nu}_{n+1}(y) - \frac{1}{n}\,\nu P(y)$. Par passage à la limite, il vient

$$\pi P(y) = \pi(y).$$

On en déduit donc que $\pi$ est une probabilité invariante. Le lemme I.1.25 assure que c'est la seule. □

La proposition suivante termine la démonstration du théorème ergodique I.1.22.

**Proposition I.1.33.** *Soit $X = (X_n, n \in \mathbb{N})$ une chaîne de Markov irréductible apériodique récurrente positive. On note $\pi$ son unique probabilité invariante. Alors on a $X_n \xrightarrow[n\to+\infty]{\textbf{loi}} \pi$.*

*Démonstration.* Soit $Y = (Y_n, n \in \mathbb{N})$ une chaîne de Markov indépendante de $X$, de même matrice de transition et de loi initiale $\pi$. Le lemme I.1.19 assure que la chaîne $Z = ((X_n, Y_n), n \in \mathbb{N})$ est irréductible et que $\pi \otimes \pi$ est une probabilité invariante. En particulier $Z$ est récurrente positive. Soit $x \in E$. On pose $T = \inf\{n \in \mathbb{N}^*; X_n = Y_n = x\}$. On a alors pour $y \in E$,

$$\mathbb{P}(X_n = y) = \mathbb{P}(X_n = y, T \leq n) + \mathbb{P}(X_n = y, T > n) \leq \mathbb{P}(X_n = y, T \leq n) + \mathbb{P}(T > n).$$

En décomposant suivant les événements $\{T = k\}$, en utilisant que sur $\{T = k\}$ on a $X_k = x = Y_k$, la propriété de Markov à l'instant $k$ et le fait que $X$ et $Y$ ont même matrice de transition, il vient $\mathbb{P}(X_n = y, T \leq n) = \mathbb{P}(Y_n = y, T \leq n)$. On a donc obtenu

$$\mathbb{P}(X_n = y) \leq \mathbb{P}(Y_n = y, T \leq n) + \mathbb{P}(T > n) \leq \mathbb{P}(Y_n = y) + \mathbb{P}(T > n).$$

En intervertissant le rôle de $X_n$ et $Y_n$ dans les inégalités précédentes, on en déduit que

$$|\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)| \leq \mathbb{P}(T > n).$$

Comme $Z$ est récurrente, p.s. $T$ est fini et donc on a $\lim_{n\to+\infty} |\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)| = 0$. Il suffit alors de chosisir $Y$ de loi intiale $\pi$ et de constater que $\mathbb{P}(Y_n = y) = \pi(y)$ pour obtenir $\lim_{n\to+\infty} |\mathbb{P}(X_n = y) - \pi(y)| = 0$. On conclut à l'aide de la proposition **??**. □

### I.1.6 Applications

*Exemple* I.1.34. **Algorithme de Metropolis-Hastings**

Soit $E$ un espace discret et $\pi$ une probabilité donnée sur $E$. Le but de l'algorithme de Metropolis-Hastings[2] est de simuler une variable aléatoire de loi $\pi$, ou tout du moins ayant une loi proche de $\pi$. Quitte à réduire l'espace d'états, on peut supposer que $\pi(x) > 0$ pour tout $x \in E$.

Pour cela on considère une matrice de transition irréductible $Q$ sur $E$ telle que pour tous $x, y \in E$, si $Q(x, y) = 0$ alors $Q(y, x) = 0$. Cette matrice est appelée matrice de sélection.

Pour $x, y \in E$ tels que $Q(x, y) > 0$, soit $(\rho(x, y), \rho(y, x)) \in ]0, 1]^2$ tels que

$$\rho(x, y)\pi(x)Q(x, y) = \rho(y, x)\pi(y)Q(y, x). \tag{I.19}$$

La fonction $\rho$ est considérée comme étant une probabilité d'acceptation. On peut construire une telle fonction $\rho$ en posant

$$\rho(x, y) = \gamma\left(\frac{\pi(y)Q(y, x)}{\pi(x)Q(x, y)}\right), \quad \text{pour tous} \quad x, y \in E \quad \text{tels que} \quad Q(x, y) > 0, \tag{I.20}$$

où $\gamma$ est une fonction à valeurs dans $]0, 1]$ telle que $\gamma(u) = u\gamma(1/u)$. En général, pour des raisons de vitesse de convergence, on choisit $\gamma(u) = \min(1, u)$ pour l'algorithme de Metropolis. Le cas $\gamma(u) = u/(1 + u)$ correspond à l'algorithme de Boltzmann ou de Barker.

Soit $X_0$ une variable aléatoire à valeurs dans $E$ de loi $\mu_0$. À l'étape $n$, les variables aléatoires $X_0, \ldots, X_n$ sont construires. On génére alors une variable aléatoire $Y_{n+1}$ de loi $Q(X_n, \cdot)$. Avec probabilité $\rho(X_n, Y_{n+1})$, on accepte cette transition et on pose $X_{n+1} = Y_{n+1}$. Si on rejette cette transition, on pose $X_{n+1} = X_n$. Il est immédiat de vérifier que $X = (X_n, n \in \mathbb{N})$ est une chaîne de Markov de matrice de transition $P$ définie par : pour $x, y \in E$,

$$P(x, y) = \begin{cases} Q(x, y)\rho(x, y) & \text{si } x \neq y, \\ 1 - \sum_{z\neq x} P(x, z) & \text{si } x = y. \end{cases}$$

[2] W. Hastings : Monte carlo sampling methods using Markov chains and their applications. *Biometrika*, 57 :97–109, 1970.

La condition (I.19) implique que la chaîne de Markov $X$ est réversible par rapport à la probabilité $\pi$. De plus elle est irréductible car $Q$ est irréductible. En particulier $X$ est récurrente positive de probabilité invariante $\pi$.

S'il existe $x, y \in E$ tels que $Q(x, y) > 0$ et $\rho(x, y) < 1$, alors on a $P(x, x) > 0$ et la chaîne $X$ est apériodique. La propriété 3 du théorème I.1.22 assure alors que $X$ converge en loi vers $\pi$.

Supposons que $\pi$ soit connue à une constante près. (C'est souvent le cas en physique statistique où $E$ désigne l'ensemble des états d'un système et la probabilité que le système soit dans l'état $x$ est proportionnelle à $\exp(-H(x)/T)$, où $H(x)$ est l'énergie de l'état $x$ et $T$ la température. Dans ce cadre, les probabilités sont appelées loi de Boltzmann, et souvent la constante de normalisation n'est pas calculable.) Si on utilise une probabilité d'acceptation donnée par (I.20), c'est le cas des algorithmes de Metropolis et de Boltzmann, alors seul le rapport $\pi(x)/\pi(y)$ intervient. En particulier, pour la simulation de $X$, il n'est pas nécessaire de calculer la constante de normalisation de $\pi$. Si les probabilités données par $Q(x, \cdot)$ sont facilement simulables, cet algorithme permet de simuler la chaîne de Markov $X$. Pour estimer $(\pi, f)$, on peut utiliser l'approximation $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$ qui d'après la propriété 2 du théorème I.1.22 converge p.s. vers $(\pi, f)$. L'inconvénient de cette méthode, au demeurant très utilisée, est que l'on ne peut pas en général donner d'intervalle de confiance pour $(\pi, f)$. ♢

*Exemple* I.1.35. **Le modèle de Wright Fisher.**

Le modèle d'évolution de population de Wright-Fisher a été introduit par Fisher en 1922[3] et par Wright[4] en 1931. On considère une population de taille $N$ constante au cours du temps. On suppose qu'à l'instant $n$ les $N$ individus meurent et donnent naissance à $N$ enfants. On suppose que la reproduction est asexuée et aléatoire : si on note $a_i^{n+1} \in \{1, \ldots, N\}$ le parent de l'individu $i$ de la génération $n+1$, vivant à la génération $n$, alors les variables aléatoires $(a_i^{n+1}, i \in \{1, \ldots, N\}, n \in \mathbb{N})$ sont indépendantes et de même loi uniforme sur $\{1, \ldots, N\}$. Tout se passe comme si chaque individu choisissait de manière indépendante son parent dans la génération précédente.

On suppose que la population est divisée deux, une partie possédant le gène $A$ et l'autre le gène $a$. On note $X_n$ le nombre d'individus portant le gène $A$ dans la population à l'instant $N$. Comme l'évolution de la population à l'instant $n+1$ ne dépend des générations passées qu'au travers de la génération $n$, il est clair que $(X_n, n \in \mathbb{N})$ est une chaîne de Markov. Si $X_n = i$, chaque enfant de la génération $n+1$ a une probabilité $i/N$ d'avoir un parent possédant le gène $A$ et donc d'avoir le gène $A$. Chaque enfant choisissant son parent de manière indépendante, on en déduit que conditionnellement à $X_n = i$, la loi de $X_{n+1}$ est une loi binomiale de paramètre $(N, i/N)$. La matrice de transition, $P$, est donc :

$$P(i,j) = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}, \quad i, j \in \{0, \ldots, N\}.$$

Les états 0 et $N$ sont des états absorbants. On note $\tau$ le premier instant de disparition de la diversité

$$\tau = \inf\{n; X_n \in \{0, N\}\},$$

avec la convention $\inf \emptyset = +\infty$. On a le résultat suivant concernant ce modèle qui assure que l'un des deux gènes disparaît p.s. en temps fini et que la probabilité que toute la population finisse par posséder le gène $A$ est égale à la proportion initiale du gène $A$.

*Lemme* I.1.36. *La variable aléatoire $\tau$ est p.s. finie. De plus, on a $\mathbb{P}(X_\tau = N | X_0) = X_0/N$.*

*Démonstration.* Si on pose $p = \min_{x \in [0,1]} x^N + (1-x)^N = 2^{-N+1}$, il vient pour tout $i \in \{1, \ldots, N-1\}$,

$$\mathbb{P}_i(\tau = 1) = \mathbb{P}_i(X_1 \in \{0, N\}) = \left(\frac{i}{N}\right)^N + \left(1 - \frac{i}{N}\right)^N \geq p.$$

[3]R. A. Fisher : On the dominance ratio. *Proc. Roy. Soc. Edinburgh*, 42 :321–341, 1922.
[4]S. Wright : Evolution in Mendelian populations. *Genetics*, 16 :97–159, 1931.

On pose $q = 1 - p$. Pour tout $i \in \{1, \dots, N-1\}$, on a $\mathbb{P}_i(\tau > 1) \leq q$. En utilisant la propriété de Markov pour $X$, il vient pour $k \geq 2$,

$$\begin{aligned}
\mathbb{P}_i(\tau > k) &= \sum_{j=1}^{N-1} \mathbb{P}_i(X_k \notin \{0, N\}, X_{k-1} = j) \\
&= \sum_{j=1}^{N-1} \mathbb{P}(X_k \notin \{0, N\} | X_{k-1} = j, X_0 = i) \mathbb{P}_i(X_{k-1} = j) \\
&= \sum_{j=1}^{N-1} \mathbb{P}_j(X_1 \notin \{0, N\}) \mathbb{P}_i(X_{k-1} = j) \\
&\leq q \sum_{j=1}^{N-1} \mathbb{P}_i(X_{k-1} = j) = q\mathbb{P}_i(\tau > k-1).
\end{aligned}$$

En itérant, on en déduit que $\mathbb{P}_i(\tau > k) \leq q^{k-1}\mathbb{P}_i(\tau > 1) \leq q^k$ et donc $\mathbb{P}_i(\tau = +\infty) = 0$ car $q < 1$. Ceci implique que p.s. $\tau$ est fini.

On remarque que

$$X_n = X_\tau \mathbf{1}_{\{\tau \leq n\}} + X_n \mathbf{1}_{\{\tau > n\}}.$$

Comme $\tau$ est fini p.s., on a donc p.s. $\lim_{n\to\infty} X_n = X_\tau$. Comme $0 \leq X_n \leq N$ pour tout $n \in \mathbb{N}$, il vient par convergence dominée

$$\lim_{n\to\infty} \mathbb{E}[X_n | X_0] = \mathbb{E}[X_\tau | X_0].$$

Comme, conditionnellement à $X_{n-1}$, $X_n$ suit la loi de Bernoulli de paramètre $(N, X_{n-1}/N)$, on a

$$\mathbb{E}[X_n | X_{n-1}, X_0] = \mathbb{E}[X_n | X_{n-1}] = N \frac{X_{n-1}}{N} = X_{n-1}.$$

On en déduit que $\mathbb{E}[X_n|X_0] = \mathbb{E}[\mathbb{E}[X_n|X_{n-1}, X_0]|X_0] = \mathbb{E}[X_{n-1}|X_0]$. En itérant, il vient $\mathbb{E}[X_n|X_0] = \mathbb{E}[X_0|X_0] = X_0$. On obtient donc $\mathbb{E}[X_\tau|X_0] = X_0$. De plus comme $X_\tau \in \{0, N\}$, on a $\mathbb{E}[X_\tau|X_0] = N\mathbb{P}(X_\tau = N|X_0)$. On en déduit donc que $\mathbb{P}(X_\tau = N|X_0) = X_0/N$. □

*Remarque* I.1.37. Il est intéressant d'étudier le temps moyen où disparaît la diversité : $t_i = \mathbb{E}_i[\tau]$, pour $i \in \{0, \dots, N\}$. Bien sûr, on a $t_0 = t_N = 0$. Pour $1 \leq i \leq N-1$, on a

$$\begin{aligned}
t_i &= \sum_{j \in \{0,\dots,N\}} \mathbb{E}_i[\tau \mathbf{1}_{\{X_1 = j\}}] \\
&= 1 + \sum_{j \in \{0,\dots,N\}} \mathbb{E}[\inf\{k \geq 0; X_{k+1} \in \{0, N\}\} \mathbf{1}_{\{X_1 = j\}}] \\
&= 1 + \sum_{j \in \{0,\dots,N\}} \mathbb{E}_j[\tau] \mathbb{P}_i(X_1 = j) \\
&= 1 + PT(i),
\end{aligned}$$

où l'on a utilisé la propriété de Markov à l'instant 1 pour la troisième égalité et $T$ est la fonction définie par $T(i) = t_i$. Comme 0 et $N$ sont des états absorbants, on a $t_i = \sum_{j\in\{0,\dots,N\}} P(i,j)t_j = 0$ pour $i \in \{0, N\}$. Si on note $e_0$ (resp. $e_N$) le vecteur de $\mathbb{R}^{N+1}$ ne comportant que des 0 sauf un 1 en première (resp. dernière) position, et $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^{N+1}$, on a

$$T = PT + \mathbf{1} - e_0 - e_N.$$

Le calcul des temps moyens où disparaît la diversité se fait donc en résolvant un système linéaire. Pour $N$ grand, on a l'approximation suivante[5] pour $x \in [0,1]$,

$$\mathbb{E}_{\lfloor Nx \rfloor}[\tau] \sim -2N\left(x\log(x) + (1-x)\log(1-x)\right).$$

où $\lfloor z \rfloor$ désigne la partie entière de $z$. Cette approximation est illustrée par le graphique I.3. ♢

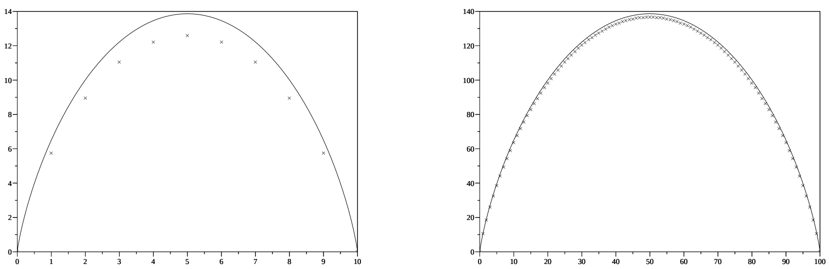

FIG. I.3 – Temps moyen de disparition de la diversité ($k \mapsto \mathbb{E}_k[\tau]$) et son approximation continue, $Nx \mapsto 2N\left(x\log(x) + (1-x)\log(1-x)\right)$, pour $N = 10$ (à gauche) et $N = 100$ (à droite).

♢

*Exemple* I.1.38. **L'urne d'Ehrenfest**

Le modèle d'urne d'Ehrenfest[6] a été introduit en 1907 pour décrire la diffusion d'un gaz placé dans deux récipients mis en communication.

On considère un système de $N$ molécules de gaz qui sont réparties dans deux recipients $A$ et $B$ identiques. À chaque instant on choisit au hasard une particule et on la change de recipient. On note $X_n$ le nombre de particules dans le recipient $A$ à l'instant $n$ et $X_0$ est la configuration initiale. La suite $X = (X_n, n \in \mathbb{N})$ représente l'évolution du nombre de molécules de gaz dans le récipient $A$. L'état d'équilibre correspond à $N_0$ molécules dans $A$, avec $N_0 = N/2$ si $N$ est pair et, par convention, $N_0 = (N-1)/2$ sinon. Les états extrêmes, où presque toutes les particules sont soit dans le récipient $A$ soit dans le récipient $B$ sont possibles dans ce modèle. Pourtant, ils sont improbables dans la réalité. Pour expliquer ce paradoxe, nous étudions, à l'aide des résultats sur les chaînes de Markov, les quantités suivantes : le temps moyen passé au voisinage de l'état d'équiblibre, la moyenne du temps de retour à l'équilibre et du temps de retour à un état extrême, ainsi les temps de transition d'un état extrême à l'état d'équilibre et vice-versa.

Remarquons que la suite $X$ est une chaîne de Markov irréductible. Sa matrice de transition $P$ est définie par $P(x,y) = 0$ si $|x-y| \neq 1$ et $P(x,x+1) = (N-x)/N$, $P(x,x-1) = x/N$ pour $x \in \{0,\dots,N\}$. Vérifions que $P$ est réversible par rapport à la loi binomiale de paramètre $(N,1/2)$, que l'on note $\pi_N$. Comme $P(x,y) = 0$ si $|x-y| \neq 1$, il suffit de vérifier que $\pi(x)P(x,x+1) = \pi(x+1)P(x+1,x)$ pour $x \in \{0,\dots,N-1\}$. Pour $x \in \{0,\dots,N-1\}$, on a

$$\pi(x+1)P(x+1,x) = \binom{N}{x+1}2^{-N}\frac{x+1}{N} = \binom{N}{x}2^{-N}\frac{N-x}{N} = \pi(x)P(x,x+1).$$

En particulier, d'après lemme I.1.12 et le corollaire I.1.23, on en déduit que $\pi_N$ est l'unique probabilité invariante de $X$.

[5]voir par exemple le livre de W. J. Ewens. *Mathematical population genetics.*, Springer-Verlag, second edition, 2004.

[6]T. Ehrenfest et P. Ehrenfest : The conceptual foundations of the statistical approach in mechanics, *Ithaca, NY, Cornell Univ. Press* (1959)

Le temps moyen passé par le système dans $I \subset \{0,\ldots,N\}$, $\displaystyle\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{\{X_k\in I\}}$ converge vers $\pi_N(I)$ quand $n$ tend vers l'infini, d'après le corollaire I.1.23. Après avoir remarqué que $\pi_N$ est la loi de la somme de $N$ variables aléatoires indépendantes de même loi de Bernoulli de paramètre $1/2$, on déduit du théorème central limite que pour $N$ grand, $\pi_N(I_{a,N})$, avec $I_{a,N} = [N_0 \pm \frac{1}{2}a\sqrt{N}]$ et $a > 0$, converge, quand $N$ tend vers l'infini, vers $\mathbb{P}(G \in [-a,a])$ où $G$ est une gaussienne de loi $\mathcal{N}(0,1)$. Pour $a$ de l'ordre de quelques unités, cette probabilité est très proche de 1, voir les tableaux du paragraphe **??**. En particulier, le temps moyen passé par le système dans le voisinage de l'état d'équilibre $I_{a,N}$ est très proche de 1, pour $a$ de l'ordre de quelques unités. Ainsi, on observe très rarement des états loin de l'équilibre $N_0$ à plus de quelques unités fois $\sqrt{N}$. Toujours après avoir remarqué que $\pi_N$ est la loi de la somme de $N$ variables aléatoires indépendantes de même loi de Bernoulli de paramètre $1/2$, on déduit du théorème de grande déviation **??** et de (**??**) avec $p = 1/2$ que pour $\varepsilon \in [0,1]$, quand $N$ tend vers l'infini, on a $\frac{2}{N}\log(\pi_N([0, N(1-\varepsilon)/2])) \sim -(1+\varepsilon)\log(1+\varepsilon) - (1-\varepsilon)\log(1-\varepsilon)$. Ainsi la probabilité d'observer des états loin de l'équilibre $N_0$ à plus de quelques unités fois $N$ décroit exponentiellement vite vers 0 avec $N$.

Le temps moyen de retour à l'état d'équilibre $N_0$ est donné, d'après (I.9), par $1/\pi_N(N_0)$. Un calcul élémentaire, utilisant la formule de Stirling, voir le lemme **??**, donne l'équivalent suivant du temps moyen de retour à l'état d'équilibre $1/\pi_N(N_0) \sim \sqrt{\pi N/2}$. En revanche le temps moyen de retour dans l'état extrême 0 est $1/\pi_N(0) = 2^N$. En particulier, le temps moyen de retour dans l'état extrême 0 n'est pas du tout du même ordre de grandeur que le temps moyen de retour dans l'état d'équilibre.

Dans le même état d'esprit, le lemme qui suit montre que le temps moyen de transition d'un état extrême vers l'état d'équilibre est bien plus faible que le temps moyen de transition de l'état d'équilibre vers un état extrême. On rappelle la définition du temps d'atteinte de l'état $m$ : $T^m = \inf\{n \geq 1; X_n = m\}$. On note $t_{k,m} = \mathbb{E}_k[T^m]$ le temps moyen d'atteinte de l'état $m$ partant de l'état $k$. Remarquons que $t_{m,m} = 1/\pi_N(m)$. Par symmétrie, on a $t_{N-k,N-m} = t_{k,m}$. Il suffit donc de calculer $t_{k,m}$ pour $k < m$.

*Lemme* I.1.39. *Pour $k,m \in \{0,\ldots,N\}$ et $k < m$, on a*

$$t_{k,m} = \frac{N}{2}\int_0^1 \frac{du}{u}\,(1-u)^{N-m}(1+u)^k\left[(1+u)^{m-k} - (1-u)^{m-k}\right], \tag{I.21}$$

*ainsi que les équivalents suivants quand $N$ tend vers l'infini :*

$$t_{0,N_0} \sim \frac{N}{4}\log(N) \quad \textit{et} \quad t_{N_0,0} \sim 2^N.$$

*Démonstration.* En décomposant suivant la valeur de $X_1$ (voir aussi les calculs de la remarque I.1.37) il est immédiat d'obtenir $t_{0,m} = 1 + t_{1,m}$, $t_{m-1,m} = 1 + \frac{m-1}{N}t_{m-2,m}$ et pour $1 \leq k < m-1$, on a

$$t_{k,m} = 1 + \frac{k}{N}t_{k-1,m} + \frac{N-k}{N}t_{k+1,m}.$$

On pose $b_k = N\int_0^1 (1+u)^k(1-u)^{N-k-1}\,du$. Il est élémentaire de vérifier que $t_{k,m} = \sum_{i=k}^{m-1} b_i$ et d'obtenir ainsi (I.21).

On suppose $N$ pair, de sorte que $N = 2N_0$. Pour $r \in \mathbb{N}^*$, on pose

$$I_r = \int_0^1 \frac{du}{u}\left(1 - (1-u)^r\right).$$

Un changement de variable élémenataire permet d'obtenir

$$t_{0,N_0} = N_0 I_{2N_0} - N_0\int_0^1 \frac{du}{u}\left(1-(1-u^2)^{N_0}\right) = N_0 I_{2N_0} - \frac{N_0}{2}I_{N_0}.$$

Pour étudier $I_r$, on écrit

$$I_r = \int_0^1 \frac{du}{u} \left(1 - \mathrm{e}^{-ru}\right) + \int_0^1 \frac{du}{u} \left(\mathrm{e}^{-ru} - (1-u)^r\right).$$

Il est immédiat de vérifier que, pour $u \in [0,1]$, $1 - u \leq \mathrm{e}^{-u} \leq 1 - u + u^2/2$. En particulier on a pour $u \in [0,1]$ :

$$0 \leq \mathrm{e}^u(1-u) \leq 1 \quad \text{et} \quad 0 \leq 1 - \mathrm{e}^u(1-u) \leq \mathrm{e}^u \frac{u^2}{2}.$$

On déduit du lemme **??** que pour $r \geq 2$

$$\begin{aligned}
\int_0^1 \frac{du}{u} \left|\mathrm{e}^{-ru} - (1-u)^r\right| = \int_0^1 \frac{du}{u} \mathrm{e}^{-ru} \left|1 - (\mathrm{e}^u(1-u))^r\right| &\leq r \int_0^1 \frac{du}{u} \mathrm{e}^{-ru} \left|1 - \mathrm{e}^u(1-u)\right| \\
&\leq \frac{r}{2} \int_0^1 du\, u\, \mathrm{e}^{-(r-1)u} \\
&\leq \frac{r}{2(r-1)^2} \int_0^{r-1} u\, \mathrm{e}^{-u}\, du \leq 1.
\end{aligned}$$

Par ailleurs, on a

$$\int_0^1 \frac{du}{u} \left(1 - \mathrm{e}^{-ru}\right) = \int_0^r \frac{du}{u} \left(1 - \mathrm{e}^{-u}\right) = \int_0^1 \frac{du}{u} \left(1 - \mathrm{e}^{-u}\right) + \log(r) - \int_1^r \frac{du}{u} \mathrm{e}^{-u} = \log(r) + O(1).$$

On en déduit que $I_r = \log(r) + O(1)$ et donc

$$t_{0,N_0} = N_0 I_{2N_0} - \frac{N_0}{2} I_{N_0} = \frac{N}{4} \log(N) + O(N).$$

Pour $N$ impair, le calcul est similaire.

On étudie maintenant $t_{N_0,0}$. On suppose $N$ pair. Il vient

$$\begin{aligned}
t_{N_0,0} &= N_0 \int_0^1 \frac{du}{u} \; \left((1+u)^N - 1\right) + N_0 \int_0^1 \frac{du}{u} \; \left(1 - (1-u)^N\right) \\
&= N_0 \int_0^1 \frac{du}{u} \; \left((1 + \frac{u}{N})^N - 1\right) + N_0 \int_1^N \frac{du}{u} \; (1 + \frac{u}{N})^N - N_0 \log(N) + N_0 I_N.
\end{aligned}$$

Comme $(1+\frac{u}{N})^N \leq \mathrm{e}^u$, le théorème de convergence dominée assure que $\int_0^1 \frac{du}{u} \; \left((1+\frac{u}{N})^N - 1\right)$ converge quand $N$ tend vers l'infini vers $\int_0^1 \frac{du}{u} \; (\mathrm{e}^u - 1)$ qui est fini. À l'aide d'une intégration par partie, on obtient que quand $N$ tend vers l'infini,

$$\int_1^N \frac{du}{u} \; (1 + \frac{u}{N})^N \sim 2^{N+1}/N.$$

On a vu que $I_r = \log(r) + O(1)$ et donc, il vient

$$t_{N_0,0} \sim N_0 2^{N+1}/N = 2^N.$$

Pour $N$ impair, le calcul est similaire. □

◊

*Exemple* I.1.40. **Modèle de file d'attente, modèle d'évolution d'un stock.** On considère une file d'attente $Y = (Y_n, n \in \mathbb{N})$, où $Y_0$ est le nombre intial de personnes dans la file d'attente et $Y_n$ est le nombre de personnes dans la file d'attente quand le service du $n$-ème client se termine. On a $Y_{n+1} = (Y_n - 1 + E_{n+1})_+$, où $E_{n+1}$ est le nombre de personnes arrivant durant le service du $n$-ème client.

Si l'on considère la chaîne $(X_n, n \in, \mathbb{N})$ qui décrit le nombre de personnes dans la file d'attente aux instants où arrive un nouveau client, on obtient l'équation $X_{n+1} = (X_n + 1 - D_{n+1})_+$, avec $D_{n+1}$ le nombre de personnes servies entre l'arrivée du $n$-ème et $n+1$-ème client. On retrouve alors la même équation d'évolution que dans le modèle de l'exemple I.1.5 d'évolution d'un stock.

Ces exemples sont des cas particuliers du modèle plus général suivant. On considère la chaîne de Markov $X = (X_n, n \in \mathbb{N})$ définie pour $n \in \mathbb{N}$ par

$$X_{n+1} = (X_n + Z_{n+1})_+, \tag{I.22}$$

où les variables aléatoires $(Z_n, n \in, \mathbb{N})$ sont indépendantes, de même loi, intégrables et indépendante de $X_0$.

Le lemme suivant assure, sous quelques hypothèses techniques supplémentaires, que si $\mathbb{E}[Z_1] > 0$ alors $X$ est transiente, si $\mathbb{E}[Z_1] = 0$ alors $X$ est récurrente nulle, si $\mathbb{E}[Z_1] < 0$ alors $X$ est récurrente positive. Le graphique de droite (resp. gauche) de la figure I.1 donne un exemple de simulation dans le cas $\mathbb{E}[Z_1] = 0$ (resp. $\mathbb{E}[Z_1] < 0$).

*Lemme* I.1.41. *On considère la chaîne de Markov $X = (X_n, n \in \mathbb{N})$ définie par (I.22). On suppose que $\mathbb{P}(Z_1 > 0) > 0$, $\mathbb{P}(Z_1 < 0) > 0$ et le PGCD de $\{|x|; \mathbb{P}(Z_1 = x) > 0\}$ est 1. Alors $X$ est irréductible.*

*On suppose de plus que $Z_1$ est intégrable. On a alors les propriétés suivantes.*

1. *Si $\mathbb{E}[Z_1] > 0$, alors $X$ est transiente.*
2. *Si $\mathbb{E}[Z_1] = 0$ et $Z_1$ de carré intégrable, alors $X$ est récurrente nulle.*
3. *Si $\mathbb{E}[Z_1] < 0$ et s'il existe $\lambda > 0$ tel que $\mathbb{E}[\mathrm{e}^{\lambda Z_1}] < +\infty$, alors $X$ est récurrente nulle.*

*Démonstration.* Comme $\mathbb{P}(Z_1 < 0) > 0$, on en déduit que pour tout $n \in \mathbb{N}^*$, la chaîne $X$ issue de $n$ peut atteindre 0 avec probabilité strictement positive. Pour montrer que $X$ est irréductible, il suffit donc montrer que la chaîne $X$ issue de 0 peut atteindre tout $n \in \mathbb{N}^*$ avec une probabilité strictement positive.

Soit $n_+$ le PGCD de $J_+ = \{x > 0; \mathbb{P}(Z_1 = x) > 0\}$ et $n_-$ le PGCD de $J_- = \{x > 0; \mathbb{P}(Z_1 = -x) > 0\}$. On note $I = \{n; \text{il existe } k \text{ tel que } P^k(0, n) > 0\}$. Il est immédiat de vérifier que $I$ est un sous-groupe de $\mathbb{N}^*$ stable par addition qui contient $J_+$ et tel que pout tout $n \in I$, $k \in J_-$ et $k < n$, on a $n - k \in I$.

On déduit de la démonstration du point 1 de la proposition I.1.18 que pour $n_0$ suffisament grand, pour tout $n \geq n_0$, on a $nn_+ \in J_+$ et $nn_- \in J_-$. Par hypothèse $n_+$ et $n_-$ sont premiers entre eux. Donc, en prenant des nombres premiers $m_+ \geq n_0$ et $m_- \geq n_0$ et premiers avec $n_+$ et $n_-$, on a $m_+n_+$ et $m_-n_-$ premier entre eux, et donc d'après le théorème de Bézout, il existe $a, b \in \mathbb{Z}$ tels que $am_+n_+ + bm_-n_- = -1$. Remarquons que $m_+n_+ \in J_+$ et $m_-n_- \in J_-$. Si $a < 0$ et $b > 0$, on a $|a|m_+n_+ - bm_-n_- = 1$ et donc $1 \in I$ puis $I = \mathbb{N}^*$. Si $a > 0$ et $b < 0$, pour tout $n$, on a pour $m \geq \max(n, n_0)$, $n = mn_+ + (an_+m_+ - |b|n_-m_-)(mn_+ - n)$ et donc $n \in I$. On a donc obtenu $I = \mathbb{N}^*$. Ceci assure que $X$ est irréductible.

On considère la marche aléatoire auxiliaire définie par $S_0 = X_0$ et $S_n = S_0 + \sum_{k=1}^n Z_k$. Il est immédiat de vérifier que $X_n \geq S_n$ et que sur $\{T^0 > n\}$ on a $X_n = S_n$.

On suppose $\mathbb{E}[Z_1] > 0$. La loi forte des grands nombres assure que p.s. $\lim_{n\to\infty} S_n/n = \mathbb{E}[X_1]$ et donc p.s. $\lim_{n\to\infty} S_n = +\infty$. On en déduit que p.s. $\lim_{n\to\infty} X_n = +\infty$ et donc $X$ est transiente.

On suppose $\mathbb{E}[Z_1] < 0$ et qu' il existe $\lambda_0 > 0$ tel que $\mathbb{E}[\mathrm{e}^{\lambda_0 Z_1}] < +\infty$. On considère $\Lambda$ la log-Laplace de $Z_1$ : $\Lambda(\lambda) = \log(\mathbb{E}[\mathrm{e}^{\lambda Z_1}])$. En particulier $\Lambda$ est bornée sur $[0, \lambda_0]$. Le point 4 du lemme **??** assure que la log-Laplace $\Lambda$ de $Z_1$ est dérivable sur $]0, \lambda[$. On déduit de (**??**) que $\Lambda'(0+) = \mathbb{E}[Z_1] < 0$. Comme $\Lambda(0) = 0$, il existe $\lambda_1 > 0$ tel que $\Lambda(\lambda_1) < 0$. On déduit de (**??**) avec $x = 0$ que

$$\mathbb{P}_0(T^0 > n) \leq \mathbb{P}_0(S_n > 0) \leq \mathrm{e}^{n\Lambda(\lambda_1)}.$$

Comme $\Lambda(\lambda_1) < 0$, on obtient que $\sum_{n\in\mathbb{N}} \mathbb{P}_0(T^0 > n) < +\infty$. On déduit du théorème de Fubini que $\sum_{n\in\mathbb{N}^*} \mathbb{P}_0(T^0 > n) = \mathbb{E}_0[T^0]$. Ainsi $\mathbb{E}_0[T^0]$ est fini, ce qui assure que la chaîne est récurrente positive.

Il reste le cas le plus délicat. On suppose $\mathbb{E}[Z_1] = 0$ et $Z_1$ de carré intégrable. Montrons par l'absurde que la chaîne est récurrente. On suppose la chaîne $X$ transiente. Comme sur $\{T^0 > n\}$ on a $X_n = S_n$, il vient

$$\{T^0 = +\infty\} \overset{\text{p.s.}}{=} \{T^0 = +\infty, \lim_{n\to\infty} X_n = +\infty\} \subset \{\lim_{n\to\infty} S_n = +\infty\}.$$

La loi du 0-1 de Kolmogorov assure que l'événement $\{\lim_{n\to\infty} S_n = +\infty\}$, qui est dans la tribu asymptotique engendrée par $(Z_n, n \in \mathbb{N}^*)$, est de probabilité égale à 0 ou 1. Remarquons que $\mathbb{P}(S_n \geq 0) = \mathbb{P}(S_n/\sqrt{n} \geq 0)$ et que le théorème de la limite centrale assure que cette quantité converge vers $\mathbb{P}(G \geq 0)$ où $G$ est de loi gaussienne $\mathcal{N}(0, \mathrm{Var}(Z_1))$. On en déduit que

$$\mathbb{P}(\lim_{n\to\infty} S_n = +\infty) \leq \lim_{n\to\infty} \mathbb{P}(S_n \geq 0) = 1/2,$$

puis que $\mathbb{P}(\lim_{n\to\infty} S_n = +\infty) = 0$ et donc $\mathbb{P}(T^0 = +\infty) = 0$. Ceci est absurde, donc $X$ est récurrente.

Montrons que $X$ est récurrente nulle. Soit $K > 0$ et $\varepsilon > 0$. On a pour $n \geq K^2/\varepsilon^2$

$$\mathbb{P}(X_n \geq K) \geq \mathbb{P}(\frac{S_n}{\sqrt{n}} \geq \frac{K}{\sqrt{n}}) \geq \mathbb{P}(\frac{S_n}{\sqrt{n}} \geq \varepsilon).$$

On déduit du théorème central limite que $\liminf_{n\to\infty} \mathbb{P}(X_n \geq K) \geq \mathbb{P}(G \geq \varepsilon)$. Comme $\varepsilon > 0$ est arbitraire, on obtient $\liminf_{n\to\infty} \mathbb{P}(X_n \geq K) \geq 1/2$ pour tout $K$. Supposons que la chaîne soit récurrente positive. On déduit de (I.11) que p.s. $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{X_k \geq x\}} = \sum_{x\geq K} \pi(x)$. Le théorème de convergence dominée assure alors que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \mathbb{P}(X_k \geq x) = \sum_{x\geq K} \pi(x).$$

On obtient donc que $\sum_{x\geq K} \pi(x) \geq 1/2$ pour tout $K > 0$. Ceci est absurde. On en déduit que $X$ est récurrente nulle. □

◊

# Bibliographie


[1] A. Barbour, L. Holst, and S. Janson. *Poisson approximation.* Oxford Studies in Probability. Clarendon Press, 1992.

[2] N. Bartoli and P. Del Moral. *Simulation et algorithmes stochastiques.* Cépaduès, 2001.

[3] P. Bickel and K. Doksum. *Mathematical statistics. Basic ideas and selected topics.* Holden-Day Series in Probability and Statistics. Holden-Day, San Francisco, 1977.

[4] P. Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, 1968.

[5] A. A. Borovkov. *Mathematical statistics.* Gordon and Breach Science Publishers, Amsterdam, 1998.

[6] N. Bouleau. *Probabilités de l'ingénieur*, volume 1418 of *Actualités Scientifiques et Industrielles.* Hermann, Paris, 1986. Variables aléatoires et simulation.

[7] L. Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. Corrected reprint of the 1968 original.

[8] P. Brémaud. *Introduction aux probabilités.* Springer Verlag, 1984.

[9] P. Brémaud. *Markov chains. Gibbs fields, Monte Carlo simulation, and queues.* Springer texts in applied mathematics. Springer, 1998.

[10] D. Dacunha-Castelle and M. Duflo. *Exercices de probabilités et statistiques. Tome 1.* Collection Mathématiques Appliquées pour la Maîtrise. Masson, Paris, 1982. Problèmes à temps fixe.

[11] D. Dacunha-Castelle and M. Duflo. *Probabilités et statistiques. Tome 1.* Collection Mathématiques Appliquées pour la Maîtrise. Masson, Paris, 1982. Problèmes à temps fixe.

[12] D. Dacunha-Castelle and M. Duflo. *Probabilités et statistiques. Tome 2.* Collection Mathématiques Appliquées pour la Maîtrise. Masson, Paris, 1983. Problèmes à temps mobile.

[13] J.-F. Delmas and B. Jourdain. *Modèles aléatoires*, volume 57 of *Mathématiques & Applications.* Springer, Berlin, 2006. Applications aux sciences de l'ingénieur et du vivant.

[14] M. Duflo. *Algorithmes stochastiques*, volume 23 of *Mathématiques & Applications.* Springer, Berlin, 1996.

[15] W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, New York, 1968.

[16] G. S. Fishman. *Monte Carlo.* Springer Series in Operations Research. Springer-Verlag, New York, 1996. Concepts, algorithms, and applications.

[17] C. Gouriéroux and A. Monfort. *Statistique des modèles économétriques.* Economica, Paris, 2ème edition, 1995.

[18] G. R. Grimmett and D. R. Stirzaker. *Probability and random processes.* Oxford University Press, New York, third edition, 2001.

[19] D. Lamberton and B. Lapeyre. *Introduction au calcul stochastique appliqué à la finance.* Ellipses Édition Marketing, Paris, second edition, 1997.

[20] L. Mazliak, P. Priouret, and P. Baldi. *Martingales et chaînes de Markov.* Hermann, 1998.

[21] A. Monfort. *Cours de Statistique Mathématique.* Economica, Paris, 1982.

[22] J. Neveu. *Bases mathématiques du calcul des probabilités.* Masson, 1964.

[23] J. Neveu. *Introduction aux probabilités.* École Polytechnique, 1988.

[24] G. Pagès and C. Bouzitat. *En passant par hasard : Probabilités de tous les jours.* Vuibert, 1999.

[25] V. V. Petrov. *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability.* The Clarendon Press Oxford University Press, New York, 1995. Sequences of independent random variables, Oxford Science Publications.

[26] S. Port. *Theoretical probability for applications.* Wiley & Sons, 1993.

[27] W. Rudin. *Real and complex analysis.* Mc Graw-Hill, 3rd edition, 1986.

[28] G. Saporta. *Probabilités, Statistique et Analyse des Données.* Technip, 1990.

[29] L. Schwartz. *Analyse III : calcul intégral.* Hermann, 1993.

[30] L. Schwartz. *Analyse IV : théorie de la mesure.* Hermann, 1993.

[31] Y. Sinai. *Probability theory : an introductory course.* Springer Verlag, 1992.

[32] D. Stirzaker. *Elementary probability.* Cambridge Univ. Press, 1994.

[33] B. Ycart. *Modèles et algorithmes markoviens*, volume 39 of *Mathématiques & Applications.* Springer, Berlin, 2002.