

Feuille d'exercices n°2 : Chaînes de Markov : exemples et propriétés.

Exercice 1. [Propriété de Markov] On considère une chaîne de Markov de matrice de transition P , c'est-à-dire qui vérifie :

$$\mathbb{P}(X_0 = x_0, \dots, X_{t+1} = x_{t+1}) = \mathbb{P}(X_0 = x_0)P(x_0, x_1) \dots P(x_t, x_{t+1}), \quad t \in \mathbb{N}, x_0, \dots, x_{t+1} \in \Omega^{t+2}$$

1. Soit $A \subset \Omega^{t+1}$, $x, y \in \Omega$. Montrer que :

$$\mathbb{P}((X_s)_{0 \leq s \leq t} \in A, X_t = x, X_{t+1} = y) = \mathbb{P}((X_s)_{0 \leq s \leq t} \in A, X_t = x)P(x, y)$$

2. Soit $f : \Omega^{t+1} \rightarrow \mathbb{R}, g : \Omega^{r+1} \rightarrow \mathbb{R}, x \in \Omega$. Montrer que :

$$\mathbb{E}[f(X_0, \dots, X_t)\mathbf{1}_{\{X_t=x\}}g(X_t, \dots, X_{t+r})] = \mathbb{E}[f(X_0, \dots, X_t)\mathbf{1}_{\{X_t=x\}}]\mathbb{E}_x[g(X_0, \dots, X_r)]$$

3. On pose $\varphi : \Omega \rightarrow \mathbb{R}, x \mapsto \mathbb{E}_x[g(X_0, \dots, X_r)]$. Montrer enfin :

$$\mathbb{E}[f(X_0, \dots, X_t)g(X_t, \dots, X_{t+r})] = \mathbb{E}[f(X_0, \dots, X_t)\varphi(X_t)].$$

Correction. L'objectif de cet exercice est de montrer que l'on peut dériver des versions de plus en plus générales de la propriété de Markov en se basant sur la seule factorisation de la probabilité d'emprunter un chemin donné (donnée au début de l'énoncé), et qui constitue la définition d'une chaîne de Markov : la méthode est donc de développer les différentes probabilités sur les chemins.

$$\begin{aligned} & \mathbb{P}((X_s)_{0 \leq s \leq t} \in A, X_t = x, X_{t+1} = y) \\ &= \sum_{x_0, \dots, x_{t-1} : (x_0, \dots, x_{t-1}, x) \in A} \mathbb{P}(X_0 = x_0, \dots, X_{t-1} = x_{t-1}, X_t = x, X_{t+1} = y) \\ &= \sum_{x_0, \dots, x_{t-1} : (x_0, \dots, x_{t-1}, x) \in A} \mathbb{P}(X_0 = x_0, \dots, X_{t-1} = x_{t-1}, X_t = x)P(x, y) \\ &= \mathbb{P}((X_s)_{0 \leq s \leq t} \in A, X_t = x)P(x, y) \end{aligned}$$

$$\begin{aligned} & \mathbb{E}[f(X_0, \dots, X_t)\mathbf{1}_{\{X_t=x\}}g(X_t, \dots, X_{t+r})] \\ &= \sum_{x_0, \dots, x_{t+r}} f(x_0, \dots, x_{t-1}, x_t)g(x_t, x_{t+1}, \dots, x_{t+r})\mathbf{1}_{\{x_t=x\}}\mathbb{P}(X_0 = x_0, \dots, X_{t+r} = x_{t+r}) \\ &= \left(\sum_{x_0, \dots, x_{t-1}} f(x_0, \dots, x_{t-1}, x)\mathbb{P}(X_0 = x_0, \dots, X_{t-1} = x_{t-1}, X_t = x) \right) \cdot \\ & \quad \left(\sum_{x_{t+1}, \dots, x_{t+r}} g(x, x_{t+1}, \dots, x_{t+r})\mathbb{P}_x(X_1 = x_{t+1}, \dots, X_r = x_{t+r}) \right) \\ &= \mathbb{E}[f(X_0, \dots, X_t)\mathbf{1}_{\{X_t=x\}}]\mathbb{E}_x[g(X_0, \dots, X_r)] \end{aligned}$$

$$\begin{aligned}
& \mathbb{E}[f(X_0, \dots, X_t)g(X_t, \dots, X_{t+r})] \\
&= \mathbb{E}[f(X_0, \dots, X_t) \cdot \left(\sum_x \mathbf{1}_{\{X_t=x\}}\right) \cdot g(X_t, \dots, X_{t+r})] \\
&= \sum_x \mathbb{E}[f(X_0, \dots, X_t) \mathbf{1}_{\{X_t=x\}} g(X_t, \dots, X_{t+r})] \\
&= \sum_x \mathbb{E}[f(X_0, \dots, X_t) \mathbf{1}_{\{X_t=x\}}] \mathbb{E}_x[g(X_0, \dots, X_r)] \\
&= \sum_x \mathbb{E}[f(X_0, \dots, X_t) \mathbf{1}_{\{X_t=x\}}] \phi(x) \\
&= \mathbb{E}[f(X_0, \dots, X_t) \phi(X_t)]
\end{aligned}$$

Exercice 2. [Chaîne renversée] Soit P matrice de transition sur Ω fini, irréductible. On note π l'unique mesure de probabilité stationnaire de P . On pose, pour $x, y \in \Omega$,

$$Q(x, y) := \frac{\pi(y)}{\pi(x)} P(y, x).$$

1. Que vaut Q lorsque P est réversible par rapport à π ?
2. Vérifier que Q définit une matrice stochastique irréductible. Quelle est son unique mesure de probabilité stationnaire ? Calculer les puissances de la matrice Q .

Soit $n \geq 1$ un entier, et $(X_t)_{0 \leq t \leq n}$ une chaîne de Markov irréductible à valeurs dans Ω et de matrice de transition P et de mesure initiale π .

3. On pose $Y_t = X_{n-t}$, $t = 0 \dots n$. Montrer que $(Y_t)_{0 \leq t \leq n}$ est encore une chaîne de Markov de matrice de transition Q et de mesure initiale à préciser.

Correction. Cet exercice montre que la chaîne de Markov renversée en temps (à horizon fini donc) est encore une chaîne de Markov si on la considère sous sa mesure stationnaire, un résultat non intuitif dans la cas non réversible ; il précise aussi la matrice de transition de la chaîne renversée en temps.

Notons d'abord que l'irréductibilité de P garantit que tous les termes en π sont strictement positifs. Si P réversible par rapport à π , $\pi(y)P(y, x) = \pi(x)P(x, y)$ donc $Q(x, y) = P(x, y)$: en particulier, dans ce cas, Q reste stochastique et irréductible. Maintenant dans le cas général,

$$\sum_y Q(x, y) = \sum_y \frac{\pi(y)}{\pi(x)} P(y, x) = \frac{(\pi P)(x)}{\pi(x)} = 1.$$

Vérifions que π est stationnaire (ce sera nécessairement l'unique mesure de probabilité stationnaire) :

$$\sum_x \pi(x) Q(x, y) = \sum_x \pi(x) \frac{\pi(y)}{\pi(x)} P(y, x) = \pi(y) \sum_x P(y, x) = \pi(y).$$

Calculons maintenant les puissances de la matrice Q . On peut commencer par Q^2 pour se forger

l'intuition :

$$\begin{aligned}
 Q^2(x, z) &= \sum_y Q(x, y)Q(y, z) \\
 &= \sum_y \frac{\pi(y)}{\pi(x)} P(y, x) \frac{\pi(z)}{\pi(y)} P(z, y) \\
 &= \frac{\pi(z)}{\pi(x)} \sum_y P(z, y)P(y, x) \\
 &= \frac{\pi(z)}{\pi(x)} P^2(z, x)
 \end{aligned}$$

On postule alors que $Q^t(x, y) = \frac{\pi(y)}{\pi(x)} P(y, x)$, ce qui se vérifie aisément par récurrence sur x en faisant le même raisonnement.

Considérons maintenant la chaîne renversée $(Y_t)_{0 \leq t \leq n} = (X_{n-t})_{0 \leq t \leq n}$ maintenant :

$$\begin{aligned}
 \mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n) &= \mathbb{P}(X_0 = y_n, \dots, X_n = y_0) \\
 &= \pi(y_n) P(y_n, y_{n-1}) \dots P(y_1, y_0) \\
 &= \pi(y_n) Q(y_{n-1}, y_n) \frac{\pi(y_{n-1})}{y_n} \dots Q(y_0, y_1) \frac{\pi(y_0)}{\pi(y_1)} \quad \text{en remplaçant } P \text{ par } Q \\
 &= \pi(y_0) Q(y_0, y_1) \dots Q(y_{n-1}, y_n).
 \end{aligned}$$

ce qui montre que Y est une chaîne de Markov de mesure initiale π et de matrice de transition Q

Exercice 3. [Théorème de représentation] Soit Ω un ensemble fini, et \mathcal{E} un espace mesurable.

1. On suppose donnée une fonction mesurable $f : \Omega \times \mathcal{E} \rightarrow \Omega$ et une suite de variables aléatoires i.i.d. $(\xi_t)_{t \geq 0}$ à valeurs dans \mathcal{E} . Soit X_0 une variable aléatoire à valeurs dans Ω et indépendante des ξ_t . On définit par récurrence la suite $(X_t)_{t \geq 0}$ par

$$X_{t+1} = f(X_t, \xi_t).$$

Montrer que pour toute suite $(x_0, \dots, x_t) \in \Omega^{t+1}$,

$$\mathbb{P}\left(\bigcap_{0 \leq s \leq t} \{X_s = x_s\}\right) = \mathbb{P}(X_0 = x_0) P(x_0, x_1) \dots P(x_{t-1}, x_t),$$

pour une certaine matrice stochastique P que l'on précisera. En déduire que X est une chaîne de Markov de matrice de transition P .

2. On se donne réciproquement une matrice stochastique P sur $\Omega = \{1, 2, \dots, n\}$, et une suite de variables aléatoires $(\xi_t)_{t \in \mathbb{R}}$ i.i.d. uniformes dans $[0, 1]$. On définit une fonction

$$\begin{aligned}
 f : \Omega \times [0, 1] &\rightarrow \Omega \\
 (k, x) &\mapsto \sum_{i=1}^n \mathbf{1}_{x \geq \sum_{j=0}^{i-1} P(k, j)}.
 \end{aligned}$$

Montrer que $P(k, l) = \mathbb{P}(f(k, \xi_0) = l)$. En déduire que toute chaîne de Markov sur Ω peut être représentée en loi par une chaîne $(X_t)_{t \geq 0}$ définie par la relation de récurrence $X_{t+1} = f(X_t, \xi_t)$, avec une suite de variables i.i.d. $(\xi_t)_{t \geq 0}$.

Correction. Pour la première question, on calcule par récurrence la probabilité d'une trajectoire :

$$\begin{aligned} \mathbb{P}\left(\bigcap_{0 \leq s \leq t} \{X_s = x_s\}\right) &= \mathbb{P}\left(\{f(x_{t-1}, \xi_{t-1}) = x_t\} \cap \bigcap_{0 \leq s \leq t-1} \{X_s = x_s\}\right) \\ &= \mathbb{P}(f(x_{t-1}, \xi_{t-1}) = x_t) \mathbb{P}\left(\bigcap_{0 \leq s \leq t-1} \{X_s = x_s\}\right) \\ &= P(x_{t-1}, x_t) \mathbb{P}\left(\bigcap_{0 \leq s \leq t-1} \{X_s = x_s\}\right), \end{aligned}$$

puisque l'événement $\bigcap_{0 \leq s \leq t-1} \{X_s = x_s\}$ est entièrement déterminé par $X_0, \xi_0, \dots, \xi_{t-2}$, et est donc indépendant de l'événement $\{f(x_{t-1}, \xi_{t-1}) = x_t\}$ qui est déterminé par la valeur de ξ_{t-1} . Par récurrence sur t , on obtient bien la formule demandée, qui est équivalente à la définition d'une chaîne de Markov.

La seconde question demande de construire pour toute matrice stochastique P une fonction $f : \Omega \times [0, 1] \rightarrow \Omega$ qui lui correspond; vérifions que la formule proposée convient. Si $U = \xi_0$ est uniforme sur $[0, 1]$, alors la probabilité pour que U tombe dans $[\sum_{j=0}^{l-1} P(k, j), \sum_{j=0}^l P(k, j))$ est la taille de cet intervalle, c'est-à-dire $P(k, l)$. Or,

$$\begin{aligned} \left\{U \in \left[\sum_{j=0}^{l-1} P(k, j), \sum_{j=0}^l P(k, j)\right)\right\} &= \left\{U \geq \sum_{j=0}^{l-1} P(k, j) \text{ si et seulement si } i \in \{1, \dots, l\}\right\} \\ &= \left\{\sum_{i=1}^n \mathbf{1}_{U \geq \sum_{j=0}^{i-1} P(k, j)} = l\right\} \\ &= \{f(k, U) = l\}. \end{aligned}$$

On a donc bien $P(k, l) = \mathbb{P}[f(k, \xi_0) = l]$.

Exercice 4. [Chaîne image] Soit $(X_t)_{t \geq 0}$ une chaîne de Markov à valeurs dans Ω et de matrice P , et $f : \Omega \rightarrow \Omega'$ une fonction surjective. On pose $Y_t = f(X_t)$, et on suppose que, pour tout $y_1, y_2 \in \Omega'$, la probabilité de transition $P(x, f^{-1}\{y_2\})$ est la même pour tout $x \in f^{-1}\{y_1\}$. On note alors cette quantité $Q(y_1, y_2)$.

1. Vérifier que Q est une matrice stochastique sur Ω' .
2. Montrer que, pour tout $t \in \mathbb{N}$ et $(y_s)_{0 \leq s \leq t} \in (\Omega')^{t+1}$,

$$\mathbb{P}\left(\bigcap_{0 \leq s \leq t} \{Y_s = y_s\}\right) = \mathbb{P}\left(\bigcap_{0 \leq s \leq t-1} \{Y_s = y_s\}\right) Q(y_{t-1}, y_t)$$

et en déduire que $(Y_t)_{t \geq 0}$ est une chaîne de Markov sur Ω' de matrice de transition Q .

On appelle mesure image de π par f la mesure de probabilité $f(\pi)$ sur Ω' définie par $f(\pi)(\{y\}) = \pi(f^{-1}(\{y\}))$, $y \in \Omega'$.

3. Si π est une mesure de probabilité stationnaire pour P , montrer que $f(\pi)$ est encore stationnaire pour Q .

4. Si P est réversible par rapport à π , montrer que Q est réversible par rapport à $f(\pi)$.

Correction. Soit $y_1 \in \Omega'$, et x_1 n'importe quel élément de $f^{-1}(\{y_1\})$ (il y en a au moins un puisque f est surjective). On a

$$\begin{aligned}
\sum_{y_2 \in \Omega'} Q(y_1, y_2) &= \sum_{y_2 \in \Omega'} P(x_1, f^{-1}(\{y_2\})) \\
&= \sum_{y_2 \in \Omega'} \sum_{x_2 \in \Omega} P(x_1, x_2) \mathbf{1}_{f(x_2)=y_2} \\
&= \sum_{x_2 \in \Omega} P(x_1, x_2) \left(\sum_{y_2 \in \Omega'} \mathbf{1}_{f(x_2)=y_2} \right) \\
&= \sum_{x_2 \in \Omega} P(x_1, x_2) = 1
\end{aligned}$$

Maintenant, puisque l'événement $\bigcap_{0 \leq s \leq t-2} \{Y_s = y_s\}$ ne dépend que des variables X_0, \dots, X_{t-2} , on obtient en utilisant la propriété de Markov pour la chaîne $(X_t)_{t \geq 0}$:

$$\begin{aligned}
\mathbb{P} \left(\bigcap_{0 \leq s \leq t} \{Y_s = y_s\} \right) &= \sum_{\substack{x_{t-1} \in f^{-1}\{y_{t-1}\} \\ x_t \in f^{-1}\{y_t\}}} \mathbb{P} \left(\{X_t = x_t\} \cap \{X_{t-1} = x_{t-1}\} \cap \bigcap_{0 \leq s \leq t-2} \{Y_s = y_s\} \right) \\
&= \sum_{\substack{x_{t-1} \in f^{-1}\{y_{t-1}\} \\ x_t \in f^{-1}\{y_t\}}} P(x_{t-1}, x_t) \mathbb{P} \left(\{X_{t-1} = x_{t-1}\} \cap \bigcap_{0 \leq s \leq t-2} \{Y_s = y_s\} \right) \\
&= \sum_{x_{t-1} \in f^{-1}\{y_{t-1}\}} Q(y_{t-1}, y_t) \mathbb{P} \left(\{X_{t-1} = x_{t-1}\} \cap \bigcap_{0 \leq s \leq t-2} \{Y_s = y_s\} \right) \\
&= Q(y_{t-1}, y_t) \mathbb{P} \left(\bigcap_{0 \leq s \leq t-1} \{Y_s = y_s\} \right).
\end{aligned}$$

Autrement dit,

$$\mathbb{P} \left(Y_t = y_t \mid \bigcap_{0 \leq s \leq t-1} \{Y_s = y_s\} \right) = Q(y_{t-1}, y_t),$$

soit la définition d'une chaîne de Markov de matrice de transition Q .

On s'attend bien sûr à ce que la mesure image π' de π par f soit une mesure Q -invariante si

π est P -invariante. On le vérifie :

$$\begin{aligned}
(\pi'Q)(y_2) &= \sum_{y_1 \in \Omega'} \pi'(y_1) Q(y_1, y_2) \\
&= \sum_{y_1 \in \Omega'} \sum_{x_1 \in \Omega} \mathbf{1}_{\{f(x_1)=y_1\}} \pi(x_1) Q(y_1, y_2) \\
&= \sum_{y_1 \in \Omega'} \sum_{x_1 \in \Omega} \mathbf{1}_{\{f(x_1)=y_1\}} \pi(x_1) P(x_1, f^{-1}\{y_2\}) \\
&= \sum_{x_1 \in \Omega} \pi(x_1) P(x_1, f^{-1}\{y_2\}) \sum_{y_1 \in \Omega'} \mathbf{1}_{\{f(x_1)=y_1\}} \\
&= \sum_{x_1 \in \Omega} \pi(x_1) P(x_1, f^{-1}\{y_2\}) \\
&= (\pi P)(f^{-1}\{y_2\}) = \pi(f^{-1}\{y_2\}) = \pi'(y_2).
\end{aligned}$$

Pour la réversibilité maintenant, supposant P réversible on montre que Q l'est de même; il s'agit d'écrire la chaîne d'égalités :

$$\begin{aligned}
\pi'(y_1)Q(y_1, y_2) &= \pi(f^{-1}\{y_1\})Q(y_1, y_2) \\
&= \sum_{x_1 \in \Omega} \mathbf{1}_{\{f(x_1)=y_1\}} \pi(x_1) P(x_1, f^{-1}\{y_2\}) \\
&= \sum_{x_1, x_2 \in \Omega} \mathbf{1}_{\{f(x_1)=y_1\}} \mathbf{1}_{\{f(x_2)=y_2\}} \pi(x_1) P(x_1, x_2) \\
&= \sum_{x_1, x_2 \in \Omega} \mathbf{1}_{\{f(x_1)=y_1\}} \mathbf{1}_{\{f(x_2)=y_2\}} \pi(x_2) P(x_2, x_1) \text{ par réversibilité}
\end{aligned}$$

Puis on remonte dans l'autre sens...

Exercice 5. [Urne d'Erhenfest¹] Soit $n \geq 1$. Le graphe $G = (V, E)$ défini par

$$V = \{0, 1\}^n \text{ et } E = \{\{x, y\} \in V^2 : \sum_{1 \leq i \leq n} |x(i) - y(i)| = 1\}$$

s'appelle l'hypercube de dimension n . On considère la marche aléatoire simple $(X_t)_{t \in \mathbb{N}}$ sur ce graphe. L'application somme des coordonnées $f : x \in V \mapsto \sum_{1 \leq i \leq n} x(i) \in \{0, \dots, n\}$ appliquée à $(X_t)_{t \in \mathbb{N}}$ donne $(Y_t)_{t \in \mathbb{N}} := (f(X_t))_{t \in \mathbb{N}}$.

1. Représenter le graphe obtenu pour $n = 2$ et $n = 3$, et justifier ainsi le nom d'hypercube pour le graphe $G = (V, E)$.
2. Quel est le degré des sommets de G ?
3. Donner la matrice de transition P de la chaîne $(X_t)_{t \geq 0}$. Est-elle irréductible ? Montrer que P admet une unique mesure de probabilité stationnaire π , et la préciser.

1. Interprétation de ce modèle, introduit en 1907 par les époux Ehrenfest pour illustrer certains des paradoxes apparus dans les fondements de la mécanique statistique : on dispose de 2 urnes qui comprennent au total n boules, et, à chaque instant, une boule choisie au hasard parmi les n boules est changée d'urne; on veut alors comprendre la répartition des boules dans les 2 urnes.

4. Vérifier que $(Y_t)_{t \geq 0}$ est encore une chaîne de Markov. Donner sa matrice de transition Q . Est-elle irréductible ?
5. Déterminer la mesure image $f(\pi)$ et vérifier que Q est réversible par rapport à $f(\pi)$.
6. En déduire à l'aide du cours la valeur de $g(k) = \mathbb{E}_k[\tau_k^+(Y)]$. Calculer la limite de

$$\frac{1}{n} \log(g(k))$$

lorsque $n \rightarrow \infty$ et $k = k(n)$ est une suite telle que $\frac{k(n)}{n} \rightarrow \alpha \in (0, 1)$. Donner enfin un équivalent de $g(\frac{n}{2})$ (on pourra s'aider de l'équivalent de Stirling $n! \sim (\frac{n}{e})^n \sqrt{2\pi n}$).

Correction. On obtient pour $n = 2$ et 3 respectivement le carré avec 4 sommets et 4 arêtes, et le cube avec 8 sommets et 12 arêtes. Les degrés des sommets valent tous n : puisque le graphe est régulier, la matrice de transition $P(x, y) = \frac{\mathbf{1}_{x \sim y}}{n}$ est réversible par rapport à la mesure uniforme donnée par $\pi(x) = \frac{1}{|V|} = \frac{1}{2^n}$. Par ailleurs, la chaîne est irréductible : étant données deux suites $x = (x(1), \dots, x(n))$ et $y = (y(1), \dots, y(1))$, on peut passer de l'une à l'autre par des transitions de la chaîne en modifiant $x(1)$, puis $x(2)$, etc. Ceci implique en particulier l'unicité de la mesure invariante.

Vérifions que la fonction f satisfait aux hypothèses de l'exercice précédent. La fonction est bien surjective vers $\{0, \dots, n\}$, et si $f(x_1) = k_1$, alors

$$P(x_1, f^{-1}(k_2)) = \begin{cases} 0 & \text{si } |k_1 - k_2| \neq 1, \\ \frac{k_1}{n} & \text{si } k_1 - k_2 = 1, \\ \frac{n-k_1}{n} & \text{si } k_1 - k_2 = -1. \end{cases}$$

En effet, pour passer de x_1 à une autre suite x_2 avec une valeur de moins égale à 1, il faut choisir l'une des $k_1 = f(x_1)$ coordonnées de x_1 égale à 1 et la changer en 0, donc $P(x_1, f^{-1}(k_1 - 1)) = \frac{k_1}{n}$; l'autre calcul est similaire. On en déduit que $P(x_1, f^{-1}(k_2))$ ne dépend que de $k_1 = f(x_1)$, et c'est ce que l'on voulait montrer. Le processus $(Y_t = f(X_t))_{t \geq 0}$ est donc une chaîne de Markov de matrice de transition

$$Q(k, k-1) = \frac{k}{n} \quad ; \quad Q(k, k+1) = \frac{n-k}{n}.$$

Cette chaîne est clairement irréductible sur $\{0, \dots, n\}$.

La mesure image ν de π par f est donnée pour tout $k \in \{0, \dots, n\}$ par

$$\begin{aligned} \nu(k) &= \pi(\{x \in \{0, 1\}^n \mid f(x) = k\}) = \pi(\{x \in \{0, 1\}^n \mid x \text{ contient } k \text{ valeurs } 1\}) \\ &= \frac{1}{2^n} |\{x \in \{0, 1\}^n \mid x \text{ contient } k \text{ valeurs } 1\}| = \frac{1}{2^n} \binom{n}{k}; \end{aligned}$$

c'est la mesure d'une loi binomiale $\mathcal{B}(n, \frac{1}{2})$. On vérifie alors que ν est réversible par rapport à Q (on peut aussi appliquer les questions du dernier exercice). Comme $Q(k, l) = 0$ si $|k - l| \neq 1$, il suffit de vérifier que $\nu(k) Q(k, k+1) = \nu(k+1) Q(k+1, k)$, et c'est un calcul simple avec les coefficients binomiaux :

$$\begin{aligned} \nu(k) Q(k, k+1) &= \frac{1}{2^n} \frac{n-k}{n} \binom{n}{k} = \frac{1}{2^n} \binom{n-1}{k}; \\ \nu(k+1) Q(k+1, k) &= \frac{1}{2^n} \frac{k+1}{n} \binom{n}{k+1} = \frac{1}{2^n} \binom{n-1}{k}. \end{aligned}$$

On sait du cours que $g(k)\nu(k) = 1$, donc $g(k) = \frac{1}{\nu(k)} = \frac{2^n}{\binom{n}{k}}$. On utilise ensuite la formule de Stirling

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{k!(n-k)!} = \frac{\sqrt{2\pi n}}{\sqrt{2\pi k}\sqrt{2\pi(n-k)}} \frac{\left(\frac{n}{e}\right)^n}{\left(\frac{k}{e}\right)^k \left(\frac{n-k}{e}\right)^{n-k}} (1 + o(1)) \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \frac{n^n}{k^k (n-k)^{n-k}} (1 + o(1)) \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \exp(n \log n - k \log k - (n-k) \log(n-k)) (1 + o(1)). \end{aligned}$$

On en déduit l'estimée suivante lorsque $n \rightarrow \infty$ et $\frac{k}{n} \rightarrow \alpha$:

$$\begin{aligned} \frac{1}{n} \log \binom{n}{k} &= -\frac{\log(2\pi)}{2n} + \frac{1}{2n} \log \left(\frac{n}{k(n-k)} \right) + \log n - \frac{k}{n} \log k - \frac{n-k}{n} \log(n-k) + o\left(\frac{1}{n}\right) \\ &= -\left(\frac{k}{n} \log \frac{k}{n}\right) - \left(\frac{n-k}{n} \log \frac{n-k}{n}\right) + O\left(\frac{\log n}{n}\right) \\ &= -\alpha \log \alpha - (1-\alpha) \log(1-\alpha) + o(1). \end{aligned}$$

Ainsi, posant $I(\alpha) = -\alpha \log \alpha - (1-\alpha) \log(1-\alpha)$, qui est une fonction positive sur $(0, 1)$, on obtient

$$\lim_{n \rightarrow \infty, \frac{k}{n} \rightarrow \alpha} \frac{\log g(k)}{n} = \log 2 - I(\alpha).$$

Ce calcul signifie que $g(k)$ se comporte comme $e^{n(\log 2 - I(\alpha))}$ dans l'échelle logarithmique. Notons que $\log(2) - I(\alpha) > 0$ pour tout $\alpha \neq \frac{1}{2}$.

Pour $k = \frac{n}{2}$, on a $\alpha = \frac{1}{2}$ et $I(\alpha) = \log 2$, donc l'équivalent logarithmique est nul. En reprenant le calcul découlant de la formule de Stirling, on obtient

$$\begin{aligned} \binom{n}{\frac{n}{2}} &= 2^n \sqrt{\frac{2}{\pi n}} (1 + o(1)) \\ g\left(\frac{n}{2}\right) &= \sqrt{\frac{\pi n}{2}} (1 + o(1)). \end{aligned}$$

On constate la différence entre des temps de retour moyens polynomiaux en n pour $k = \frac{n}{2}$, et exponentiels en n pour $k = \alpha n$ avec $\alpha \neq 1/2$. Notons que dans le cas de la ruine du joueur (avec $P(0, 1) = P(n, n-1) = 1$), qui est une autre marche aléatoire sur $\{0, \dots, n\}$ à laquelle il est légitime de comparer ce modèle, on obtenait des temps de retour quadratiques en n .

Exercice 6. [Ruine du joueur] On considère $(X_t)_{t \in \mathbb{N}}$ une chaîne de Markov sur $\Omega = \{0, \dots, n\}$ de matrice de transition

$$P(i, j) = \begin{cases} \frac{1}{2} \mathbf{1}_{\{|j-i|=1\}} & \text{si } i \in \{1, \dots, n-1\} \\ \mathbf{1}_{i=j} & \text{si } i \in \{0, n\}. \end{cases}$$

Il s'agit de la marche aléatoire formée par les gains d'un joueur qui joue à un jeu équilibré : le joueur gagne ou perd 1 à chaque tour de jeu et s'arrête lorsqu'il son gain atteint 0 ou n . On s'intéresse au temps aléatoire $\tau = \min\{t \geq 0, X_t \in \{0, n\}\} \in \mathbb{N} \cup \{\infty\}$ qui modélise la durée du jeu, et aussi à la variable aléatoire X_τ , définie sur l'événement $\{\tau < \infty\}$.

1. Soit $k \in \Omega \setminus \{0, n\}$. On note $\theta = \theta_1$ l'opérateur de shift, et $\tau \circ \theta$ la variable aléatoire $\tau(\theta(X))$. Observer que pour tout entier $t \in \mathbb{N}$,

$$\{\tau \circ \theta \geq t\} \cap \{\tau \neq 0\} = \{\tau \geq t + 1\}$$

et en déduire l'égalité suivante à l'aide de la propriété de Markov :

$$\mathbb{P}_k(\tau \geq t + 1 \mid X_1 = k + 1) = \mathbb{P}_{k+1}(\tau \geq t).$$

2. Noter que pour tout $n \in \mathbb{N} \cup \{\infty\}$, $n = \sum_{t \geq 1} \mathbf{1}_{t \leq n}$, et en déduire que

$$\mathbb{E}_k[\tau \mid X_1 = k + 1] = 1 + \mathbb{E}_{k+1}[\tau].$$

Pour la seconde identité, noter qu'on ne sait pas si la variable aléatoire τ est p.s. finie, mais comme elle est à valeurs dans $\mathbb{N} \cup \{\infty\}$, son espérance est bien définie, possiblement égale à $+\infty$.

3. Soit $k \in \Omega$. Posons $h(k) = \mathbb{E}_k[\tau]$. Expliciter $h(0)$ et $h(n)$, et donner l'équation de récurrence satisfaite par h . Résoudre ce système (on pourra poser $\ell(k) = h(k+1) - h(k)$).
4. Soit $k \in \Omega \setminus \{0, n\}$. Justifier l'identité :

$$\{(\theta \circ X)_{\tau \circ \theta} = n\} \cap \{\tau \neq 0\} \cap \{\tau < \infty\} = \{X_\tau = n\} \cap \{\tau \neq 0\} \cap \{\tau < \infty\}$$

puis en déduire à l'aide de la propriété de Markov que

$$\mathbb{P}_k(\tau < \infty \text{ et } X_\tau = n \mid X_1 = k + 1) = \mathbb{P}_{k+1}(\tau < \infty \text{ et } X_\tau = n).$$

Soit $k \in \Omega$. Posons $g(k) = \mathbb{P}_k(\tau < \infty \text{ et } X_\tau = n)$. Expliciter $g(0)$ et $g(n)$, et donner l'équation de récurrence satisfaite par g . Résoudre ce système.

5. Pour cette dernière question, on modifie les probabilités de transition depuis l'état 0 en supposant que $P(0, 1) = 1$ (autrement dit, le joueur est un addict et il se remet à jouer immédiatement après avoir perdu). On pose $\tau^0 = \min\{t \geq 0, X_t = n\} \in \mathbb{N} \cup \{\infty\}$, et $h^0(k) = \mathbb{E}_k(\tau^0)$. Donner $h^0(0) - h^0(1)$ et $h^0(n)$. Montrer que h^0 vérifie la même équation de récurrence que h , et la résoudre.

Correction. Répéter aux élèves combien cet exercice est important. À l'aide de l'opérateur de shift $\theta = \theta_1$ défini par

$$(\theta \circ X)_t = X_{t+1}.$$

on peut exprimer $\tau = \tau(X)$ peut être exprimé en fonction de $\tau \circ \theta = \tau \circ \theta(X)$ comme suit

$$\{\tau \geq t + 1\} = \{\tau \circ \theta \geq t\} \cap \{\tau \neq 0\}$$

puis on écrit que

$$\begin{aligned} \mathbb{P}_k(\tau \geq t + 1 \mid X_1 = k + 1) &= \mathbb{P}_k(\tau \circ \theta \geq t, \tau \neq 0 \mid X_1 = k + 1) \\ &= \mathbb{P}_k(\tau \circ \theta \geq t \mid X_1 = k + 1) \quad \text{car } \mathbb{P}_k(\tau \neq 0 \mid X_1 = k + 1) = 1 \\ &= \mathbb{P}_{k+1}(\tau \geq t) \quad \text{par Markov} \end{aligned}$$

Maintenant, pour tout $n \in \mathbb{N} \cup \{\infty\}$, $n = \sum_{t \geq 1} \mathbf{1}_{n \geq t}$ et donc, pour toute variable aléatoire $N \in \mathbb{N} \cup \{+\infty\}$,

$$\mathbb{E}[N] = \sum_{t \geq 1} \mathbb{P}(N \geq t),$$

et en particulier pour τ sous la mesure $\mathbb{P}_k[\cdot | X_1 = k + 1]$:

$$\begin{aligned}
\mathbb{E}_k[\tau | X_1 = k + 1] &= \sum_{t=1}^{\infty} \mathbb{P}_k(\tau \geq t | X_1 = k + 1) \\
&= 1 + \sum_{t=1}^{\infty} \mathbb{P}_k(\tau \geq t + 1 | X_1 = k + 1) \\
&= 1 + \sum_{t=1}^{\infty} \mathbb{P}_{k+1}(\tau \geq t) \\
&= 1 + \mathbb{E}_k[\tau],
\end{aligned}$$

ces égalités ayant lieu dans $\mathbb{R}^+ \cup \{\infty\}$. On a de même $\mathbb{P}_k(\tau < +\infty | X_1 = k - 1) = \mathbb{P}_{k-1}(\tau < \infty)$ et $\mathbb{E}_k[\tau | X_1 = k - 1] = \mathbb{E}_{k-1}[\tau] + 1$ pour $k \notin \{0, n\}$.

Pour $k \notin \{0, n\}$ on a

$$\begin{aligned}
h(k) &= \mathbb{E}_k[\tau] = \mathbb{E}_k[\tau | X_1 = k + 1] \mathbb{P}_k[X_1 = k + 1] + \mathbb{E}_k[\tau | X_1 = k - 1] \mathbb{P}_k[X_1 = k - 1] \\
&= \frac{1}{2} (1 + \mathbb{E}_{k+1}[\tau] + 1 + \mathbb{E}_{k-1}[\tau]) \\
&= 1 + \frac{h(k+1) + h(k-1)}{2}
\end{aligned}$$

et les conditions au bord sont $h(0) = h(n) = 0$. Si $\ell(k) = h(k+1) - h(k)$, alors pour $k \in \{1, \dots, n-1\}$,

$$\ell(k) = h(k+1) - h(k) = 2h(k) - 2 - h(k-1) - h(k) = \ell(k-1) - 2.$$

On a donc $\ell(k) = \ell(0) - 2k$, et par ailleurs la somme des $\ell(k)$ est nulle, donc

$$0 = \sum_{k=0}^{n-1} \ell(k) = n\ell(0) - n(n-1),$$

soit $\ell(0) = n - 1$ et $\ell(k) = n - 1 - 2k$. On en déduit que

$$h(k) = h(k) - h(0) = \sum_{j=0}^{k-1} \ell(j) = (n-1)k - (k-1)k = (n-k)k.$$

Les quantités k et $n - k$ jouent bien un rôle symétrique dans cette expression comme attendu.

On sait maintenant que $\tau < \infty$ \mathbb{P}_k p.s. quelque soit $k \in \Omega$. Pour calculer $\mathbb{P}_k(X_\tau = n | X_1 = k + 1)$, Avec cette notation, on observe alors que

$$\{X_\tau = n\} \cap \{\tau \neq 0\} \cap \{\tau < \infty\} = \{(\theta \circ X)_{\tau \circ \theta} = n\} \cap \{\tau \neq 0\} \cap \{\tau < \infty\}$$

ce qui entraîne (on note que $\{\tau < \infty\}$ est un événement presque sûr sous \mathbb{P}_k quelque soit k) :

$$\begin{aligned}
\mathbb{P}_k(X_\tau = n | X_1 = k + 1) &= \mathbb{P}_k(X_\tau = n, \tau \neq 0 | X_1 = k + 1) \\
&= \mathbb{P}_k((\theta \circ X)_{\tau \circ \theta} = n, \tau \neq 0 | X_1 = k + 1) \\
&= \mathbb{P}_k((\theta \circ X)_{\tau \circ \theta} = n | X_1 = k + 1) \\
&= \mathbb{P}_{k+1}((\theta \circ X)_{\tau \circ \theta} = n | X_1 = k + 1) \\
&= \mathbb{P}_{k+1}(X_\tau = n)
\end{aligned}$$

On en tire l'équation de récurrence

$$\begin{aligned}
 g(k) &= \mathbb{P}_k(X_\tau = n) \\
 &= \mathbb{P}_k(X_\tau = n \mid X_1 = k+1) \mathbb{P}_k(X_1 = k+1) + \mathbb{P}_k(X_\tau = n \mid X_1 = k-1) \mathbb{P}_k(X_1 = k-1) \\
 &= \frac{1}{2} (\mathbb{P}_{k+1}(X_\tau = n) + \mathbb{P}_{k-1}(X_\tau = n)) \\
 &= \frac{g(k+1) + g(k-1)}{2}
 \end{aligned}$$

pour $k \notin \{0, n\}$; c'est la même équation que précédemment, mais on a maintenant $g(0) = 0$ et $g(n) = 1$. Pour résoudre cette équation, notons qu'elle peut se réécrire $g(k+1) - g(k) = g(k) - g(k-1)$; la fonction g a donc des accroissements constants et c'est la fonction affine $g(k) = \frac{k}{n}$.

Pour la fonction h^0 , la même équation de récurrence que pour h est satisfaite : les preuves sont identiques à précédemment. Par ailleurs, les valeurs au bord sont :

$$h^0(n) = 0 \quad ; \quad h^0(0) = 1 + h^0(1).$$

Si l'on pose $\tilde{\ell}(k+1) = h^0(k+1) - h^0(k)$, alors on a comme précédemment $\tilde{\ell}(k) = \tilde{\ell}(k-1) - 2$, avec cette fois-ci la condition initiale $\tilde{\ell}(0) = -1$. Il suit $\tilde{\ell}(k) = -1 - 2k$, puis

$$-h^0(0) = h^0(n) - h^0(0) = \sum_{k=0}^{n-1} \tilde{\ell}(k) = -n^2,$$

donc $h^0(0) = n^2$, et finalement

$$h^0(k) = n^2 + \sum_{j=0}^{k-1} \tilde{\ell}(j) = n^2 - k^2.$$

Exercice 7. [Temps d'atteinte et de retour d'espérance finie] Soit $(X_t)_{t \in \mathbb{N}}$ une chaîne de Markov irréductible sur Ω fini. On fixe $y \in \Omega$ dans ce qui suit.

1. Montrer qu'il existe $t_0 \in \mathbb{N}$, et $\delta \in]0, 1]$ tel que pour tout $x \in \Omega$, $\mathbb{P}_x(\tau_y \leq t_0) \geq \delta$.
2. Soit $x \in \Omega$. En déduire que pour $k \in \mathbb{N}$, $\mathbb{P}_x(\tau_y > kt_0) \leq (1 - \delta)^k$.
3. Conclure que $\mathbb{E}_x[\tau_y] \leq t_0/\delta$.
4. En déduire que $\mathbb{E}_x[\tau_y^+] \leq \mathbf{1}_{\{x=y\}} + t_0/\delta$.

Correction. Voir le poly pour la correction. On notera que les estimées obtenues sont uniformes en x par construction (mais le t_0 dépend de y).

Exercice 8. [Chaîne induite] Soit $(X_t)_{t \geq 0}$ une chaîne de Markov irréductible à valeurs dans Ω et de matrice de transition P . On pose

$$\begin{cases} \theta_0 = 0 \\ \theta_{t+1} = \min\{s \in \mathbb{N} : s > \theta_t, X_s \neq X_{\theta_t}\} \end{cases} \quad t \geq 0$$

puis $Y_t = X_{\theta_t}$ pour tout entier $t \in \mathbb{N}$.

1. Justifier que $\mathbb{P}(\theta_t < \infty) = 1$ pour tout entier $t \in \mathbb{N}$.
2. Montrer que $(Y_t)_{t \geq 0}$ est encore une chaîne de Markov, et exprimer sa matrice de transition à l'aide de P .

Exercice 9. [Dernier site occupé] Soit $n \geq 3$. On considère la marche aléatoire $(X_t)_{t \geq 0}$ sur le n -cycle :

$$G = (V, E) \quad ; \quad V = \{1, 2, \dots, n\} \quad ; \quad E = \{\{1, 2\}, \{2, 3\}, \dots, \{n-1, n\}, \{n, 1\}\}.$$

On convient dans ce qui suit que les additions sont modulo n . On note Y le dernier site visité par la marche aléatoire, formellement défini par

$$\{Y = y\} = \{\tau_y = \max_{1 \leq x \leq n} \tau_x\}.$$

1. Expliquer pourquoi

$$\{Y = y\} = \{\tau_{y-1} < \tau_{y+1} < \tau_y\} \cup \{\tau_{y+1} < \tau_{y-1} < \tau_y\}.$$

2. Montrer que, pour tout $y \in \{1, \dots, n\}$,

$$\mathbb{P}_{y-1}(\tau_{y+1} < \tau_y) = \frac{1}{n-1}.$$

On pourra chercher un système d'équations pour la fonction $f(k) = \mathbb{P}_k(\tau_{y+1} < \tau_y)$ puis résoudre ce système.

3. En déduire la valeur de $\mathbb{P}_x(Y = y)$, et reconnaître la distribution de la variable aléatoire Y . Commenter ce résultat.

Correction. On note que les quantités $\{\tau_z, z \in \{1, \dots, n\}\}$ sont deux à deux distinctes. On a :

$$\begin{aligned} \{Y = y\} &= \{\tau_y = \max_x \tau_x\} = \bigcap_{x \neq y} \{\tau_y > \tau_x\} \\ &= \{\tau_y > \tau_{y-1}\} \cap \{\tau_y > \tau_{y+1}\} \\ &= \{\tau_y > \tau_{y-1} > \tau_{y+1}\} \cup \{\tau_y > \tau_{y+1} > \tau_{y-1}\}. \end{aligned}$$

Le seul point non trivial dans cette suite d'égalités est la seconde ligne. Clairement, si $\tau_y > \tau_x$ pour tout $x \neq y$, alors $\tau_y > \tau_{y+1}$ et $\tau_y > \tau_{y-1}$. Réciproquement, si ces deux inégalités sont vérifiées, alors on ne peut pas avoir $\tau_x < \tau_y$ pour un $x \in V$, car dans ce cas on aurait soit

$$\tau_x < \tau_{x+1} < \dots < \tau_{y-1} < \tau_y,$$

soit

$$\tau_x < \tau_{x-1} < \dots < \tau_{y+1} < \tau_y,$$

puisque la marche aléatoire se déplace par sauts de taille 1.

Posons $f(k) = \mathbb{P}_k[\tau_{y+1} < \tau_y]$. On a bien sûr $f(y+1) = 1$ et $f(y) = 0$, et pour les autres valeurs de k ,

$$f(k) = \frac{f(k-1) + f(k+1)}{2}$$

en utilisant la propriété de Markov. La solution de ce systèmes d'équations est

$$f(y) = 0, f(y-1) = \frac{1}{n-1}, f(y-2) = \frac{2}{n-1}, \dots, f(y+2) = \frac{n-2}{n-1}, f(y+1) = 1$$

puisque l'on a $n-1$ pas entre y et $y+1$ si l'on se déplace cycliquement et par pas décroissant d'une unité, et puisque $f(k+1) - f(k) = f(k) - f(k-1)$, c'est-à-dire que les accroissements de f sont constants. En particulier, on obtient bien

$$\mathbb{P}_{y-1}[\tau_{y+1} < \tau_y] = \frac{1}{n-1}.$$

Pour calculer $\mathbb{P}_x[Y = y]$, on remarque que les deux événements $\{\tau_{y-1} < \tau_{y+1} < \tau_y\}$ et $\{\tau_{y+1} < \tau_{y-1} < \tau_y\}$ sont disjoints, donc

$$\mathbb{P}_x[Y = y] = \mathbb{P}_x[\tau_{y-1} < \tau_{y+1} < \tau_y] + \mathbb{P}_x[\tau_{y+1} < \tau_{y-1} < \tau_y]$$

Propriété qu'on peut aussi obtenir en décomposant comme suit selon les valeurs du temps d'arrêt τ_{y-1} , dont on rappelle qu'il est fini. Maintenant,

$$\begin{aligned} \mathbb{P}_x[\tau_{y-1} < \tau_{y+1} < \tau_y] &= \sum_{t \geq 0} \mathbb{P}_x[\tau_{y-1} = t, \tau_{y-1} < \tau_{y+1} < \tau_y] \\ &= \sum_{t \geq 0} \mathbb{P}_x[\tau_{y-1} = t, t < \tau_{y+1}, \tau_{y+1} < \tau_y] \\ &= \sum_{t \geq 0} \mathbb{P}_x[\tau_{y-1} = t, t < \tau_{y+1}, \tau_{y+1} \circ \theta_t < \tau_y \circ \theta_t] \\ &= \sum_{t \geq 0} \mathbb{P}_x[\tau_{y-1} = t, t < \tau_{y+1}] \mathbb{P}_{y-1}[\tau_{y+1} < \tau_y] \quad \text{par la propriété de Markov} \\ &= \sum_{t \geq 0} \mathbb{P}_x[\tau_{y-1} = t, t < \tau_{y+1}] \mathbb{P}_{y-1}[\tau_{y+1} < \tau_y] \\ &= \frac{1}{n-1} \sum_{t \geq 0} \mathbb{P}_x[\tau_{y-1} = t, t < \tau_{y+1}] \mathbb{P}_{y-1}[\tau_{y+1} < \tau_y] \\ &= \mathbb{P}_x[\tau_{y-1} < \tau_{y+1}] \mathbb{P}_{y-1}[\tau_{y+1} < \tau_y] \end{aligned}$$

(on utilise en fait ce qu'on appelle la propriété de Markov forte), mais on connaît la valeur du dernier facteur :

$$\mathbb{P}_x[\tau_{y-1} < \tau_{y+1} < \tau_y] = \frac{1}{n-1} \mathbb{P}_x[\tau_{y-1} < \tau_{y+1}].$$

On a alors symétriquement

$$\mathbb{P}_x[\tau_{y+1} < \tau_{y-1} < \tau_y] = \mathbb{P}_x[\tau_{y+1} < \tau_{y-1}] \mathbb{P}_{y+1}[\tau_{y-1} < \tau_y] = \frac{1}{n-1} \mathbb{P}_x[\tau_{y+1} < \tau_{y-1}].$$

En additionnant les deux, on conclut que pour $y \neq x$, $\mathbb{P}_x[Y = y] = \frac{1}{n-1}$. Ainsi, le dernier site visité par la marche aléatoire simple sur $\mathbb{Z}/n\mathbb{Z}$ est uniforme sur l'ensemble des sites distincts du site de départ, ce qui n'est pas spécialement intuitif. Noter aussi que la preuve ne nécessite pas de comprendre qui de $y-1$ ou de $y+1$ est atteint en premier.