Université des Sciences et Technologies de Lille

U. F. R. de Mathématiques Pures et Appliquées

Chaînes de Markov

Daniel Flipo

Agrégation de Mathématiques

Sommaire

1 Définitions et exemples

Définition 1.1 *Soit* $(X_n)_{n\in\mathbf{N}}$ *une suite de variables aléatoires de* $(\Omega, \mathscr{A}, \mathbf{P})$ *dans un espace* E *fini ou dénombrable appelé espace des états.*

1) On dit que $(X_n)_{n\in\mathbf{N}}$ *est une chaîne de Markov si et seulement si*

$$\mathbf{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0) = \mathbf{P}(X_{n+1} = j \mid X_n = i)$$

pour tout $n \in \mathbf{N}$*, pour tout état* j *et pour toute suite d'états* $i_0, i_1, \ldots i_{n-1}, i$*, pour lesquels la probabilité conditionnelle a un sens, c.-à-d. tels que*

$$\mathbf{P}(X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0) > 0.$$

2) Si en plus la probabilité $\mathbf{P}(X_{n+1} = j \mid X_n = i)$ *ne dépend pas de* n*, c.-à-d. si*

$$\forall n \in \mathbf{N}, \quad \mathbf{P}(X_{n+1} = j \mid X_n = i) = \mathbf{P}(X_1 = j \mid X_0 = i)$$

on dit que la chaîne de Markov est homogène.

La propriété de Markov exprime que, si la valeur de X_n est connue à l'instant n, la loi des variables futures (X_{n+1}, X_{n+2} etc.) ne dépend pas du passé (les valeurs de X_{n-1}, X_{n-2} etc.). Vérifier à titre d'exercice que, si (X_n) est une chaîne de Markov homogène,

$$\begin{aligned} \mathbf{P}(X_{n+2} = k, X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0) \\ = \mathbf{P}(X_{n+2} = k \mid X_{n+1} = j)\ \mathbf{P}(X_{n+1} = j \mid X_n = i) \\ = \mathbf{P}(X_2 = k, X_1 = j \mid X_0 = i). \end{aligned} \tag{1}$$

Exemples : vérifier dans chacun des exemples suivants que X_n est une chaîne de Markov homogène et préciser sa matrice de transition.

1) *Promenades aléatoires* : soit $(Y_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes et de même loi à valeurs dans $\mathbf{Z}$ (ou $\mathbf{Z}^d$), soit X_0 une variable aléatoire à valeurs dans $\mathbf{Z}$ (ou $\mathbf{Z}^d$), indépendante des (Y_n), on pose $X_n = X_0 + \sum_{i=1}^n Y_i$ pour tout entier $n \geq 1$.

2) *Ruine du joueur* : deux joueurs A et B disposant respectivement de fortunes initiales a et b (entiers positifs) jouent à un jeu de hasard. La mise est de 1 € par partie ; les résultats des parties sont indépendants, à chaque partie A a la probabilité p de gagner, q de perdre et $r \geq 0$ de faire match nul ($0 < p < 1$, $0 < p < 1$ et $p + q + r = 1$). Le jeu se poursuit indéfiniment ou jusqu'à la ruine d'un des deux joueurs. On note X_n la fortune du joueur A après n parties.

3) *Séries de succès* : des candidats doivent répondre à une suite de questions de difficulté variable, les performances des différents candidats sont indépendantes. La probabilité pour chaque candidat de bien répondre à une question de niveau k est p_k, celle de donner une réponse fausse est $q_k = 1 - p_k$. Lorsqu'un candidat donne une réponse fausse, il est remplacé par le candidat suivant qui démarre au niveau 0. X_n représente le niveau atteint par le candidat en lice à l'instant n :

$$\forall n, k \in \mathbf{N}\ (n \geq k), \quad \begin{cases} \mathbf{P}(X_{n+1} = k+1 \mid X_n = k, X_{n-1} = k-1, \ldots X_{n-k} = 0) = p_k \\ \mathbf{P}(X_{n+1} = 0 \quad\ \mid X_n = k, X_{n-1} = k-1, \ldots X_{n-k} = 0) = q_k \end{cases}$$

4) *Modèle de diffusion gazeuse* : On considère une enceinte faite de deux compartiments séparés par une cloison poreuse. Au départ le compartiment de gauche contient a molécules de gaz de type A, celui de droite b molécules de gaz de type B. On modélise la diffusion au travers de la paroi en supposant qu'à chaque instant, il y a tirage au hasard d'une molécule dans chaque compartiment et échange des deux molécules tirées. La composition des deux urnes après le n-ième échange est complètement déterminée par la donnée de la variable X_n nombre de molécules de gaz A dans l'urne de gauche.

5) *File d'attente* : Soit $(T_n)_{n\in\mathbf{N}}$ la suite des instants (aléatoires) d'arrivée des clients à un guichet. Un seul client est servi à la fois. On note X_n le nombre de clients en attente ou en cours de service juste avant l'instant T_n et D_n le nombre de clients dont le service se termine dans l'intervalle $[T_n, T_{n+1}[$. On suppose les variables D_n indépendantes et de même loi donnée : pour tout $k \in \mathbf{N}$ $p_k = P(D_0 = k)$. On a, si $a^+ = \max(a, 0)$ désigne la partie positive du réel a,

$$\forall n \in \mathbf{N}, \quad X_{n+1} = (X_n + 1 - D_n)^+$$

6) *Gestion de stock* : on s'intéresse au nombre de pièces d'un même type en stock dans un entrepôt, à différents instants $(t_n)_{n\in\mathbf{N}}$, par exemple à chaque fin de journée ou de semaine. La demande pour ce type de pièces dans l'intervalle $[t_n, t_{n+1}[$ est une variable aléatoire entière D_n. La suite des v.a. $(D_n)_{n\in\mathbf{N}}$ est supposée indépendante et de même loi connue. La politique de gestion est la suivante : lorsque le niveau du stock à un des instants (t_n) descend en dessous d'un seuil s fixé on se réapprovisionne de façon à ramener le stock à son niveau maximal S déterminé par exemple par la taille de l'entrepôt ou les moyens financiers de l'entreprise. On admet que la livraison intervient sans délai, c'est-à-dire avant le début de la période suivante. La taille X_n du stock à l'instant t_n vérifie

$$\forall n \in \mathbf{N}, \quad \begin{cases} X_{n+1} = (X_n - D_n)^+ & \text{si} \quad s \leq X_n \leq S \\ X_{n+1} = (S - D_n)^+ & \text{si} \quad X_n < s \end{cases}$$

7) *Processus de branchement* : de nombreux exemples de chaînes de Markov interviennent en génétique (modèles de reproduction) et en physique (désintégrations atomiques). On suppose qu'à la fin de son existence chaque organisme i de la n-ième génération donne naissance à un nombre aléatoire $Y_{i,n}$ de descendants. Les variables aléatoires $(Y_{i,n})_{(i,n)\in\mathbf{N}^2}$ sont supposées indépendantes et de même loi. Le nombre X_n d'organismes de la n-ième génération vérifie

$$\forall n \in \mathbf{N}, \quad X_{n+1} = \sum_{i=1}^{X_n} Y_{i,n}$$

Remarque : si $(Y_n)_{n\in\mathbf{N}}$ est une suite de variables aléatoires indépendantes et de même loi à valeurs dans E et si $f : E \times E \mapsto E$ est une fonction quelconque, alors la suite $(X_n)_{n\in\mathbf{N}}$ définie par

$$\forall n \in \mathbf{N}, \quad X_{n+1} = f(X_n, Y_n) \quad \text{et} \quad X_0 \text{ donnée, indépendante des } (Y_n)_{n\in\mathbf{N}},$$

est une chaîne de Markov homogène. Ce résultat (à démontrer en exercice) fournit un moyen assez général pour établir qu'une suite de variables aléatoires est une chaîne de Markov homogène.

Définition 1.2 *On appelle matrice de transition de la chaîne de Markov homogène* (X_n) *la matrice* P *définie par*[1] *:* $\forall (i,j) \in E \times E \quad P(i,j) = \mathbf{P}(X_1 = j \mid X_0 = i)$.

Proposition 1.3 *La loi d'une chaîne de Markov homogène est complètement déterminée par la donnée de sa matrice de transition et de la loi de* X_0 *(appelée loi initiale)*[2] *:* $\forall i \in E,\ \mu(i) = \mathbf{P}(X_0 = i)$. *Pour tout entier* n *et tous* $i_0, i_1, \ldots, i_n$ *états de* E *:*

$$\mathbf{P}(X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0) = \mu(i_0)P(i_0,i_1)P(i_1,i_2)\ldots P(i_{n-1},i_n).$$

DÉMONSTRATION. Immédiate à partir de la formule

$$\mathbf{P}(\bigcap_{i=0}^{n} A_i) = \mathbf{P}(A_0)\,\mathbf{P}(A_1 \mid A_0)\,\mathbf{P}(A_2 \mid A_1 \cap A_0)\ldots\mathbf{P}(A_n \mid A_{n-1} \cap \ldots \cap A_0). \qquad (2)$$

et de la propriété de Markov homogène. □

Proposition 1.4 (Chapman-Kolmogorov) *Pour tout couple* (i,j) *d'états de* E *et pour tout couple* (n,m) *d'entiers naturels*

$$\mathbf{P}(X_{n+m} = j \mid X_0 = i) = \sum_{k\in E} \mathbf{P}(X_n = k \mid X_0 = i)\,\mathbf{P}(X_m = j \mid X_0 = k) \qquad (3)$$

1. Si E n'est pas fini, la matrice a un nombre infini de lignes et de colonnes...
2. Le plus souvent X_0 est connue de manière déterministe, dans ce cas la loi initiale est une mesure de Dirac.

En particulier la matrice des transitions en n étapes est la puissance n-ième de la matrice P des transitions en une étape[3] *:*

$$\forall n \in \mathbf{N},\ \forall (i,j) \in \mathrm{E}\times\mathrm{E}, \quad \mathbf{P}(\mathrm{X}_n = j \mid \mathrm{X}_0 = i) = \mathrm{P}^n(i,j) \tag{4}$$

Démonstration. Calculons $\mathbf{P}(\mathrm{X}_{n+m} = j \mid \mathrm{X}_0 = i)$ en décomposant l'événement $\{\mathrm{X}_{n+m} = j\}$ sur les événements $(\{\mathrm{X}_n = k\})_{k\in\mathrm{E}}$ qui forment une partition de Ω :

$$\begin{aligned}
\mathbf{P}(\mathrm{X}_{n+m} = j \mid \mathrm{X}_0 = i) &= \mathbf{P}\Big(\mathrm{X}_{n+m} = j \cap \Big(\bigcup_{k\in\mathrm{E}} \mathrm{X}_n = k\Big) \mid \mathrm{X}_0 = i\Big)\\
&= \sum_{k\in\mathrm{E}} \mathbf{P}(\mathrm{X}_{n+m} = j \cap \mathrm{X}_n = k \mid \mathrm{X}_0 = i)\\
&= \sum_{k\in\mathrm{E}} \frac{\mathbf{P}(\mathrm{X}_{n+m} = j, \mathrm{X}_n = k, \mathrm{X}_0 = i)}{\mathbf{P}(\mathrm{X}_0 = i)}\\
&= \sum_{k\in\mathrm{E}} \frac{\mathbf{P}(\mathrm{X}_{n+m} = j \mid \mathrm{X}_n = k, \mathrm{X}_0 = i)\, \mathbf{P}(\mathrm{X}_n = k, \mathrm{X}_0 = i)}{\mathbf{P}(\mathrm{X}_0 = i)}\\
&= \sum_{k\in\mathrm{E}} \mathbf{P}(\mathrm{X}_m = j \mid \mathrm{X}_0 = k)\, \mathbf{P}(\mathrm{X}_n = k \mid \mathrm{X}_0 = i)
\end{aligned}$$

car X_n est une chaîne de Markov homogène.
En particulier, pour $n = m = 1$ et pour tous $(i,j) \in \mathrm{E}\times\mathrm{E}$,

$$\mathbf{P}(\mathrm{X}_2 = j \mid \mathrm{X}_0 = i) = \sum_{k\in\mathrm{E}} \mathbf{P}(\mathrm{X}_1 = j \mid \mathrm{X}_0 = k)\, \mathbf{P}(\mathrm{X}_1 = k \mid \mathrm{X}_0 = i) = \sum_{k\in\mathrm{E}} \mathrm{P}(i,k)\,\mathrm{P}(k,j)$$

ce qui établit (4) dans le cas $n = 2$. Pour $m = 1$ et $n > 1$, l'égalité (3) s'écrit

$$\begin{aligned}
\forall (i,j) \in \mathrm{E}\times\mathrm{E} \quad \mathbf{P}(\mathrm{X}_{n+1} = j \mid \mathrm{X}_0 = i) &= \sum_{k\in\mathrm{E}} \mathbf{P}(\mathrm{X}_n = k \mid \mathrm{X}_0 = i)\, \mathbf{P}(\mathrm{X}_1 = j \mid \mathrm{X}_0 = k)\\
&= \sum_{k\in\mathrm{E}} \mathbf{P}(\mathrm{X}_n = k \mid \mathrm{X}_0 = i)\, \mathrm{P}(k,j)
\end{aligned}$$

donc si $\mathbf{P}(\mathrm{X}_n = k \mid \mathrm{X}_0 = i) = \mathrm{P}^n(i,k)$ alors $\mathbf{P}(\mathrm{X}_{n+1} = j \mid \mathrm{X}_0 = i) = \mathrm{P}^{n+1}(i,j)$ et la récurrence est établie. □

2 Généralisations de la propriété de Markov

Dans la section précédente nous avons travaillé sur les suites finies de variables aléatoires $(\mathrm{X}_k)_{k\leq m}$, il est intéressant de considérer la chaîne « globalement » en tant que variable aléatoire de $(\Omega, \mathscr{A})$ dans $\mathrm{E}^{\mathbf{N}}$ muni d'une tribu $\mathscr{F}$ adéquate. Si

3. Dans toute la suite la notation $\mathrm{P}^n(i,j)$ désigne le terme de la ligne i, colonne j de la matrice P^n, à ne pas confondre avec $\big(\mathrm{P}(i,j)\big)^n$!

on veut que $X = (X_n)_{n\in\mathbf{N}}$ soit mesurable de $(\Omega, \mathscr{A})$ dans $(E^{\mathbf{N}}, \mathscr{F})$ dès que toutes ses composantes X_n le sont de $(\Omega, \mathscr{A})$ dans $(E, \mathscr{P}(E))$, il ne faut pas choisir $\mathscr{F}$ trop grande (en particulier le choix $\mathscr{F} = \mathscr{P}(E^{\mathbf{N}})$ ne convient pas). On construit $\mathscr{F}$ à partir des « cylindres » de $E^{\mathbf{N}}$ définis ci-dessous.

2.1 Cylindres sur $E^{\mathbf{N}}$

Définition 2.1 *On appelle cylindre de* $E^{\mathbf{N}}$ *toute partie* C *de* $E^{\mathbf{N}}$ *de la forme* $C = B_0 \times B_1 \times \cdots \times B_n \times E \times E \times E \ldots$, *où* n *est un entier quelconque et où les* B_i *sont des parties quelconques*[4] *de* E.
On appelle tribu cylindrique sur $E^{\mathbf{N}}$ *la plus petite tribu* $\mathscr{F}$ *contenant les cylindres de* $E^{\mathbf{N}}$.

Proposition 2.2 *Les cylindres de* $E^{\mathbf{N}}$ *ont les propriétés suivantes :*

a) l'intersection de deux cylindres est un cylindre ;

b) la réunion de deux cylindres n'est pas toujours *un cylindre ;*

c) le complémentaire d'un cylindre est une réunion finie de cylindres ;

d) l'algèbre de Boole engendrée sur $E^{\mathbf{N}}$ *par les cylindres est la famille* $\mathscr{B}$ *dont les éléments sont les réunions finies de cylindres.*

Démonstration. a) Si $n \geq m$, on a

$$(B_0 \times B_1 \times \cdots \times B_n \times E \times \ldots) \cap (B'_0 \times B'_1 \times \cdots \times B_m \times E \times \ldots)$$
$$= B_0 \cap B'_0 \times \cdots \times B_m \cap B'_m \times B_{m+1} \times \cdots \times B_n \times E \times \ldots$$

b) Considérer la réunion des cylindres $B_0 \times E \times E \times \ldots$ et $E \times B'_1 \times E \times \ldots$ (faire un dessin dans $\mathbf{R}^3$ pour B_0 et B'_1 singletons de $\mathbf{N}$).

c) Soit $C = B_0 \times B_1 \times \cdots \times B_n \times E \times E \ldots$,

$$x \in C \iff \forall i \leq n, \quad x_i \in B_i \quad \text{donc}$$
$$x \in \overline{C} \iff \exists i \leq n, \quad x_i \in \overline{B_i} \quad \text{d'où} \quad \overline{C} = \bigcup_{i=0}^{n} \{E \times \cdots \times E \times \overline{B_i} \times E \times \ldots\}.$$

d) Montrons que la famille $\mathscr{B}$ est une algèbre de Boole : il est immédiat de vérifier que E (cylindre) appartient à $\mathscr{B}$ et que $\mathscr{B}$ est stable par réunion finie. Vérifions que $\mathscr{B}$ est stable par complémentation : soit $B = \cup_{j=1}^{k} C_j$ une réunion finie de cylindres, son complémentaire $\overline{B} = \cap_{j=1}^{k} \overline{C_j}$ et d'après le point c), chaque

4. E dénombrable, est muni de la tribu de toutes les parties de E. Si E était muni d'une tribu $\mathscr{E}$ plus petite, il faudrait bien sûr prendre les B_i dans $\mathscr{E}$.

$\overline{C_j}$ est une réunion de n_j cylindres : $\overline{C_j} = \cup_{m=1}^{n_j} C'_{j,m}$. Finalement, en posant $n = \max_{j\leq k}(n_j)$ et $C'_{j,m} = \emptyset$ pour $n_j < m \leq n$, on a

$$\overline{B} = \bigcap_{j=1}^{k} \overline{C_j} = \bigcap_{j=1}^{k} \left(\bigcup_{m=1}^{n_j} C'_{j,m} \right) = \bigcap_{j=1}^{k} \left(\bigcup_{m=1}^{n} C'_{j,m} \right) = \bigcup_{m=1}^{n} \left(\bigcap_{j=1}^{k} C'_{j,m} \right) = \bigcup_{m=1}^{n} C''_m$$

où les C''_m sont des intersections finies de cylindres donc des cylindres (point a) ; le complémentaire de B est bien une réunion finie de cylindres.

La famille $\mathscr{B}$ est donc une algèbre de Boole contenant les cylindres ; réciproquement toute algèbre de Boole contenant les cylindres contient par définition les réunions finies de cylindres, la proposition est démontrée. □

Proposition 2.3 *Si pour tout* $n \in \mathbf{N}$, X_n *est mesurable* $(\Omega, \mathscr{A}) \mapsto (E, \mathscr{P}(E))$ *et si* $\mathscr{F}$ *est la tribu cylindrique sur* $E^{\mathbf{N}}$*, alors* $X = (X_n)_{n\in\mathbf{N}}$ *est mesurable* $(\Omega, \mathscr{A}) \mapsto (E^{N}, \mathscr{F})$.

Démonstration. Comme $\mathscr{F}$ est la tribu engendrée par les cylindres, il suffit de vérifier que pour tout cylindre C, $X^{-1}(C)$ appartient à $\mathscr{A}$. Mais

$$X^{-1}(B_0 \times B_1 \times \cdots \times B_n \times E \times E \times E \ldots) = \bigcap_{i=0}^{n} X_i^{-1}(B_i) \in \mathscr{A}$$

car les X_n sont mesurables $(\Omega, \mathscr{A}) \mapsto (E, \mathscr{P}(E))$. □

Il reste à munir $(E^N, \mathscr{F})$ d'une probabilité, on pourrait s'en tirer en disant qu'il « suffit » de prendre la probabilité image de $\mathbf{P}$ (probabilité sur Ω) par X, mais ce serait bien hypocrite car $\mathbf{P}$ n'a jamais été définie précisément ! En fait comment construit-on $(\Omega, \mathscr{A}, \mathbf{P})$?

Qu'il s'agisse de construire un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ pour une suite dénombrable [5] de variables aléatoires indépendantes de lois données à valeurs dans $(E, \mathscr{E})$ ou d'une chaîne de Markov, on procède de la même manière :

- on se place sur l'espace produit E^N que l'on munit de sa tribu cylindrique $\mathscr{F}$;
- on essaie de prolonger les probabilités $\mathbf{P}_n$ définies « de façon naturelle » sur les produits finis E^n en une probabilité $\mathbf{P}$ sur E^N, l'existence et l'unicité de $\mathbf{P}$ sont assurées par le théorème de Kolmogorov (voir [4]) moyennant des conditions raisonnables de compatibilité des $\mathbf{P}_n$ (la trace sur E^n de la probabilité $\mathbf{P}_{n+1}$ définie sur E^{n+1} doit être identique à $\mathbf{P}_n$).

Finalement l'espace $(\Omega, \mathscr{A}, \mathbf{P})$ ainsi construit pour porter la suite $X = (X_n)_{n\in\mathbf{N}}$ est *identique* à l'espace d'arrivée $(E^N, \mathscr{F}, \mathbf{P})$ et X est l'application identité.

5. Le cas d'une suite finie de variables aléatoires indépendantes de lois données est trivial : on se place sur l'espace produit muni de la tribu produit et de la probabilité produit (indépendance oblige).

Dans le cas des chaînes de Markov homogènes les $\mathbf{P}_n$ sont définies à partir de la loi initiale λ et de la matrice de transition P (cf. prop. 1.3) et la condition de compatibilité est facile à vérifier.

On notera $(\Omega, \mathscr{A}, \mathbf{P}_\lambda)$ l'espace probabilisé adapté à la chaîne de Markov homogène de matrice de transition P donnée et de loi initiale λ (λ probabilité donnée sur E). Dans le cas où λ est une mesure de Dirac (cas où X_0 est connue de manière déterministe), on note $\mathbf{P}_x$ au lieu de $\mathbf{P}_{\delta_x}$ la probabilité construite sur$(\Omega, \mathscr{A})$.

2.2 Opérateurs de décalage sur $E^{\mathbf{N}}$

Définition 2.4 *On appelle opérateur de décalage (ou translation) sur $E^{\mathbf{N}}$ l'application θ de $E^{\mathbf{N}}$ dans lui-même définie par*

$$\forall (x_n) \in E^{\mathbf{N}} \quad \theta\big((x_0, x_1, \dots, x_n, \dots)\big) = (x_1, x_2, \dots, x_{n+1}, \dots)$$

On considère également ses itérés, définis par $\theta^{k+1} = \theta \circ \theta^k : \theta^k\big((x_n)\big) = \big((x_{n+k})\big)$.

La proposition suivante énonce la forme globale de la propriété de Markov homogène.

Proposition 2.5 *Soit $X = (X_n)_{n\in\mathbf{N}}$ une chaîne de Markov homogène définie sur $(\Omega, \mathscr{A}, \mathbf{P}_\lambda)$ à valeurs dans $(E^{\mathbf{N}}, \mathscr{F})$, et $\mathscr{A}_k = \sigma(X_0, \dots, X_k)$ la sous-tribu de $\mathscr{A}$ engendrée par les variables aléatoires $X_0, \dots, X_k$. Alors, la loi conditionnelle de $\theta^k(X)$ sachant $\mathscr{A}_k$ est la loi conditionnelle de X sachant X_k :*

$$\forall f \text{ mesurable} : (E^{\mathbf{N}}, \mathscr{F}) \mapsto (\mathbf{R}_+, \mathscr{B}(\mathbf{R}_+)), \quad \mathbf{E}_\lambda\big(f(\theta^k(X)) \mid \mathscr{A}_k\big) = \mathbf{E}_{X_k}\big(f(X)\big).$$

Démonstration. Il suffit de démontrer l'égalité ci-dessus pour les fonctions indicatrices (linéarité et passage à la limite croissante) et même pour les fonctions indicatrices de cylindres : ceci résulte par exemple du théorème d'unicité des mesures (la classe des cylindres est stable par intersection finie et contient $E^{\mathbf{N}}$). On peut aussi utiliser la proposition 2.2 et dire que si deux mesures coïncident sur les cylindres, elles coïncident sur les réunions finies de cylindres (formule de Poincaré et stabilité de la famille des cylindres par intersection finie) donc sur l'algèbre de Boole $\mathscr{B}$ engendrée et finalement aussi sur la tribu engendrée (cf. [1] th. I-4-7).

Calculons donc pour le cylindre $C = B_0 \times B_1 \times \cdots \times B_n \times E \times E \times E \dots$ l'espérance conditionnelle $\mathbf{E}_\lambda(\mathbf{1}_C(\theta^k(X)) \mid \mathscr{A}_k)$ sur l'événement $X_k = i_k, \dots, X_0 = i_0$:

$$\text{Sur } \{X_k = i_k, \dots, X_0 = i_0\}, \quad \mathbf{E}_\lambda(\mathbf{1}_C(\theta^k(X)) \mid \mathscr{A}_k) = \mathbf{P}_\lambda(X_{n+k} \in B_n, \dots, X_k \in B_0 \mid X_k = i_k, \dots, X_0 = i_0).$$

On décompose les B_j pour pouvoir appliquer la propriété de Markov :

$$\begin{aligned}
&\mathbf{P}_\lambda(X_{n+k}\in B_n,\dots,X_k\in B_0 \mid X_k=i_k,\dots,X_0=i_0)\\
&\qquad=\sum_{(j_0,\dots,j_n)\in B_0\times\cdots\times B_n}\mathbf{P}_\lambda(X_{n+k}=j_n,\dots,X_k=j_0\mid X_k=i_k,\dots,X_0=i_0)\\
&\qquad=\sum_{(j_0,\dots,j_n)\in B_0\times\cdots\times B_n}\mathbf{1}_{j_0=i_k}\,\mathbf{P}_\lambda(X_{n+k}=j_n,\dots,X_{k+1}=j_1\mid X_k=i_k,\dots,X_0=i_0)
\end{aligned}$$

soit en sommant en j_0 et en appliquant la propriété de Markov homogène [6] :

$$\begin{aligned}
&=\sum_{(j_1,\dots,j_n)\in B_1\times\cdots\times B_n}\mathbf{1}_{B_0}(i_k)\,\mathbf{P}_{i_k}(X_n=j_n,\dots,X_1=j_1)\\
&=\mathbf{1}_{B_0}(i_k)\,\mathbf{P}_{i_k}(X_n\in B_n,\dots,X_1\in B_1)=\mathbf{E}_{i_k}(\mathbf{1}_C(X))
\end{aligned}$$

Finalement, pour tous $i_0,\dots,i_k$ de E, on a sur $\{X_k=i_k,\dots,X_0=i_0\}$

$$\mathbf{E}_\lambda(\mathbf{1}_C(\theta^k(X))\mid\mathscr{A}_k)=\mathbf{E}_{X_k}(\mathbf{1}_C(X)).\qquad\square$$

2.3 Propriété de Markov forte

Rappelons quelques définitions classiques :

Définition 2.6 *Soit* $(\Omega,\mathscr{A},\mathbf{P})$ *un espace de probabilité.*
On appelle filtration adaptée à la suite $X=(X_n)_{n\in\mathbf{N}}$ *la suite croissante de sous-tribus* $(\mathscr{A}_n)$ *de* $\mathscr{A}$*, engendrées par les variables* $(X_k)_{k\le n}$ *:* $\mathscr{A}_n=\sigma(X_0,\dots,X_n)$.

On dit que la variable aléatoire T *à valeurs dans* $\overline{\mathbf{N}}=\mathbf{N}\cup\{+\infty\}$ *est un temps d'arrêt pour la filtration* $(\mathscr{A}_n)$ *si et seulement si pour tout* n *de* $\mathbf{N}$ *l'événement* $\{T=n\}$ *appartient à* $\mathscr{A}_n$.

On définit ensuite la tribu $\mathscr{A}_T$ *des événements antérieurs à* T *(temps d'arrêt) par*

$$\forall A\in\mathscr{A},\quad A\in\mathscr{A}_T\iff\forall n\in\mathbf{N},\ A\cap\{T=n\}\in\mathscr{A}_n.$$

On étend enfin la définition des opérateurs de décalage aux temps d'arrêt en posant $\theta^T=\theta^n$ *sur* $\{T=n\}$*, ce qui définit* θ^T *sur l'événement* $\{T<+\infty\}$.

Proposition 2.7 (propriété de Markov forte) *Soit* $X=(X_n)_{n\in\mathbf{N}}$ *une chaîne de Markov homogène définie sur* $(\Omega,\mathscr{A},\mathbf{P}_\lambda)$ *à valeurs dans* $(E,\mathscr{F})$ *et* T *un temps d'arrêt pour la filtration* $(\mathscr{A}_n=\sigma(X_0,\dots,X_n))$*, alors sur l'événement* $\{T<+\infty\}$

$$\forall f \text{ mesurable} : (E^{\mathbf{N}},\mathscr{F})\mapsto(\mathbf{R}_+,\mathscr{B}(\mathbf{R}_+)),\quad \mathbf{E}_\lambda\big(f(\theta^T(X))\mid\mathscr{A}_T\big)=\mathbf{E}_{X_T}\big(f(X)\big).$$

6. À faire en exercice en généralisant le calcul (1) p. 1 : utiliser la formule (2) p. 3.

DÉMONSTRATION. On calcule l'espérance conditionnelle $\mathbf{E}_\lambda(f(\theta^T(X)) \mid X_T, \ldots, X_0)$ sur l'événement $\{T < +\infty\}$ en la décomposant selon la valeur de T :

$$\mathbf{E}_\lambda\big(f(\theta^T(X)) \mid \mathscr{A}_T\big)\ \mathbf{1}_{T<+\infty} = \sum_{k\in\mathbf{N}} \mathbf{E}_\lambda\big(f(\theta^k(X)) \mid \mathscr{A}_k\big)\ \mathbf{1}_{T=k}$$

soit en utilisant la proposition 2.5 :

$$\begin{aligned} &= \sum_{k\in\mathbf{N}} \mathbf{E}_{X_k}\big(f(X)\big)\ \mathbf{1}_{T=k} \\ &= \mathbf{E}_{X_T}\big(f(X)\big)\ \mathbf{1}_{T<+\infty}. \end{aligned}$$

□

3 Classification des états

3.1 Classes d'états communicants

Définition 3.1 *Soient i et j deux états de* E*. On dit que l'état j est accessible à partir de i si et seulement si*

$$\exists n \geq 0, \quad P^n(i,j) = \mathbf{P}(X_n = j \mid X_0 = i) > 0.$$

On dit que les états i et j communiquent et on note $i \rightleftarrows j$ si et seulement si j est accessible à partir de i et i est accessible à partir de j.

Proposition 3.2 *La relation $i \rightleftarrows j$ est une relation d'équivalence sur* E*. L'espace* E *peut donc être partitionné en classes d'équivalence pour la relation $i \rightleftarrows j$, appelées classes d'états communicants.*

DÉMONSTRATION. La réflexivité est évidente (pour $n = 0$ $\mathbf{P}(X_0 = i \mid X_0 = i) = 1$), tout comme la symétrie. La transitivité résulte de la propriété de Chapman-Kolmogorov : si $P^n(i,k) > 0$ et si $P^m(k,j) > 0$,

$$P^{n+m}(i,j) = \mathbf{P}(X_{n+m} = j \mid X_0 = i) = \sum_{l\in E} P^n(i,l)\, P^m(l,j) \geq P^n(i,k)\, P^m(k,j) > 0.$$

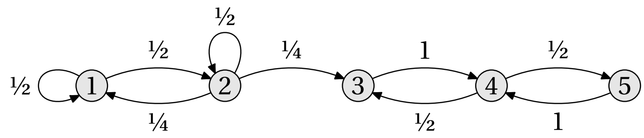
□

Définition 3.3 *Lorsque l'état* E *est réduit à une seule classe (cas où tous les états communiquent) on dit que la chaîne est irréductible.*

Pour rechercher les classes d'états communicants, il est commode de travailler sur le graphe de la chaîne plutôt que sur la matrice de transition : le graphe est obtenu en traçant pour tout couple d'états (i,j) un arc allant de i à j si et seulement si $P(i,j) > 0$. On peut ajouter une valeur à l'arc (la probabilité $P(i,j)$) dans ce cas la donnée du graphe valué est équivalente à la donnée de la matrice de transition.

Exemples :

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$



La chaîne comporte deux classes : $\{1,2\}$ et $\{3,4,5\}$.

Vérifier que l'espace d'états associé à la ruine du joueur comporte trois classes (à préciser).

Vérifier que la promenade aléatoire sur $\mathbf{Z}^2$ telle qu'elle est définie à l'exercice 5 comporte deux classes à préciser.

En revanche la promenade aléatoire sur $\mathbf{Z}^2$ définie par

$$(X_{n+1}, Y_{n+1}) = \begin{cases} (X_n + 1, Y_n) & \text{avec probabilité } \frac{1}{4} \\ (X_n - 1, Y_n) & \text{avec probabilité } \frac{1}{4} \\ (X_n, Y_n + 1) & \text{avec probabilité } \frac{1}{4} \\ (X_n, Y_n - 1) & \text{avec probabilité } \frac{1}{4} \end{cases}$$

est irréductible.

3.2 Récurrence et transience

Définition 3.4 *Un état i de E est dit récurrent si et seulement si, partant de i, la chaîne X revient P_i-presque sûrement à l'état i. Un état non récurrent est dit transient. On pose*

$$\tau_i(\omega) = \begin{cases} \inf\{n \geq 1 \mid X_n(\omega) = i\} & \textit{ou} \\ +\infty & \textit{sur} \quad \{\forall n \geq 1,\ X_n(\omega) \neq i\} \end{cases}$$

τ_i est un temps d'arrêt pour la filtration ($\mathscr{A}_n = \sigma(X_0, \dots, X_n)$), il est appelé temps de retour à i lorsque la chaîne part de i et temps d'atteinte de i sinon. On a l'équivalence suivante :

$$\forall i \in E, \quad i \textit{ récurrent} \iff \mathbf{P}_i(\tau_i < +\infty) = 1$$
$$\forall i \in E, \quad i \textit{ transient} \iff \mathbf{P}_i(\tau_i < +\infty) < 1$$

Exemple : on montre (cf. exercice 3) que pour la promenade aléatoire sur $\mathbf{Z}$ définie comme somme de variables aléatoires de Bernoulli $\mathbf{P}_0(\tau_0 < +\infty) = 1 - |p - q|$.

Ainsi, pour $p = q$ tous les états sont récurrents, tandis que pour $p \neq q$ tous les états sont transients (le calcul fait pour l'état 0 vaut bien sûr pour tout autre état).

On s'intéresse également à la variable aléatoire N_i nombre de passages de la chaîne par l'état i après l'instant 0 :

$$\mathrm{N}_i(\omega) = \sum_{n\in\mathbf{N}^*} \mathbf{1}_{\mathrm{X}_n(\omega)=i}$$

Nous aurons besoin du résultat intuitif suivant (expliquer sa signification) :

Proposition 3.5 *Soient i et j deux états quelconques de* E*, on a*

$$\forall n \in \mathbf{N}, \quad \mathbf{P}_j(\mathrm{N}_i \geq n+1) = \mathbf{P}_j(\tau_i < +\infty)\, \mathbf{P}_i(\mathrm{N}_i \geq n). \tag{5}$$

Démonstration. On considère l'instant τ_i de premier passage de la chaîne par l'état i. L'événement $\{\mathrm{N}_i \geq n+1\}$ peut s'écrire[7] :

$$\{\mathrm{N}_i \geq n+1\} = \{\tau_i < +\infty \cap \mathrm{N}_i \circ \theta^{\tau_i} \geq n\}$$

où $\mathrm{N}_i \circ \theta^{\tau_i}$ désigne le nombre de visites à l'état i après l'instant τ_i. La propriété de Markov forte appliquée à l'instant τ_i donne :

$$\begin{aligned}\mathbf{P}_j(\mathrm{N}_i \geq n+1) &= \mathbf{P}_j(\mathrm{N}_i \circ \theta^{\tau_i} \geq n \mid \tau_i < +\infty)\, \mathbf{P}_j(\tau_i < +\infty)\\ &= \mathbf{P}_i(\mathrm{N}_i \geq n)\, \mathbf{P}_j(\tau_i < +\infty) \quad \text{car} \quad \mathrm{X}_{\tau_i} = i \text{ sur } \{\tau_i < +\infty\}. \qquad \square\end{aligned}$$

Remarque : il est possible de contourner le recours à la propriété de Markov forte dans la démonstration de (5). On remarque que $\{\mathrm{N}_i \geq 1\} = \{\tau_i < +\infty\}$ et on décompose l'événement $\{\mathrm{N}_i \geq n+1\}$ en fonction des valeurs prises par τ_i, instant de premier passage par i :

$$\forall n \geq 0,\ \{\mathrm{N}_i \geq n+1\} = \Big\{\mathrm{N}_i \geq n+1 \cap \big(\bigcup_{k\in\mathbf{N}^*} \tau_i = k\big)\Big\} = \bigcup_{k\in\mathbf{N}^*} \{\mathrm{N}_i \geq n+1 \cap \tau_i = k\}$$

les événements $\{\tau_i = k\}_{k\in\mathbf{N}^*}$ étant deux à deux disjoints,

$$\begin{aligned}\forall n \geq 0,\ \mathbf{P}_j(\mathrm{N}_i \geq n+1) &= \sum_{k\in\mathbf{N}^*} \mathbf{P}_j(\mathrm{N}_i \geq n+1 \cap \tau_i = k)\\ &= \sum_{k\in\mathbf{N}^*} \mathbf{P}_j(\mathrm{N}_i \geq n+1 \cap (\mathrm{X}_k = i, \mathrm{X}_{k-1} \neq i, \ldots, \mathrm{X}_1 \neq i))\\ &= \sum_{k\in\mathbf{N}^*} \mathbf{P}_j(\mathrm{N}_i \geq n+1 \mid \mathrm{X}_k = i, \mathrm{X}_{k-1} \neq i, \ldots, \mathrm{X}_1 \neq i)\\ &\qquad\qquad \cdot \mathbf{P}_j(\mathrm{X}_k = i, \mathrm{X}_{k-1} \neq i, \ldots, \mathrm{X}_1 \neq i)\\ &= \sum_{k\in\mathbf{N}^*} \mathbf{P}_i(\mathrm{N}_i \geq n)\, \mathbf{P}_j(\tau_i = k) \text{ (prop. de Markov homogène)}\\ &= \mathbf{P}_j(\tau_i < +\infty)\, \mathbf{P}_i(\mathrm{N}_i \geq n). \quad \square\end{aligned}$$

7. Dans cette formule, θ opère sur Ω, ce qui suppose $\Omega = \mathrm{E}^{\mathbf{N}}$, voir la construction de $(\Omega, \mathscr{A}, \mathbf{P}_\lambda)$ page 6.

Proposition 3.6 *Les conditions suivantes sont équivalentes :*

a) l'état i est récurrent ($\mathbf{P}_i(\tau_i < +\infty) = 1$) ;

b) la chaîne X revient $\mathbf{P}_i$-p.s. une infinité de fois à l'état i : $\mathbf{P}_i(\mathrm{N}_i = +\infty) = 1$;

c) la série $\sum_{n\in\mathbf{N}} \mathrm{P}^n(i,i)$ diverge.

Les conditions suivantes sont équivalentes :

a') l'état i est transient ($\mathbf{P}_i(\tau_i < +\infty) < 1$) ;

b') la variable aléatoire N_i est $\mathbf{P}_i$-p.s. finie ($\mathbf{P}_i(\mathrm{N}_i = +\infty) = 0$) et elle suit une loi géométrique[8] sur $\mathbf{N}$: $\forall n \in \mathbf{N}, \quad \mathbf{P}_i(\mathrm{N}_i \geq n) = \big(\mathbf{P}_i(\tau_i < +\infty)\big)^n$;

c') la variable aléatoire N_i est $\mathbf{P}_i$-intégrable : $\mathbf{E}_i(\mathrm{N}_i) = \sum_{n\geq 1} \mathrm{P}^n(i,i) < +\infty$.

Démonstration. Appliquons l'égalité (5) dans le cas $j = i$.

Cas i récurrent : $\mathbf{P}_i(\tau_i < +\infty) = 1$, l'égalité (5) s'écrit pour $j = i$:

$$\forall n \in \mathbf{N}, \quad \mathbf{P}_i(\mathrm{N}_i \geq n+1) = \mathbf{P}_i(\mathrm{N}_i \geq n) = \mathbf{P}_i(\mathrm{N}_i \geq 1) = \mathbf{P}_i(\tau_i < +\infty) = 1$$

donc $\mathbf{P}_i(\mathrm{N}_i = +\infty) = 1$ et l'implication (a $\Rightarrow$ b) est établie.

Cas i transient : $\alpha_i = \mathbf{P}_i(\tau_i < +\infty) < 1$, l'égalité (5) donne ($\{\mathrm{N}_i \geq 1\} = \{\tau_i < +\infty\}$) :

$$\forall n \in \mathbf{N}, \quad \mathbf{P}_i(\mathrm{N}_i \geq n+1) = \alpha_i\ \mathbf{P}_i(\mathrm{N}_i \geq n) = (\alpha_i)^n\ \mathbf{P}_i(\mathrm{N}_i \geq 1) = (\alpha_i)^{n+1}$$

donc $\mathbf{P}_i(\mathrm{N}_i = +\infty) = \lim \downarrow (\alpha_i)^n = 0$, N_i suit une loi géométrique sur $\mathbf{N}$; rappelons que, pour toute variable Z à valeurs dans $\mathbf{N}$, on a

$$\mathbf{E}(\mathrm{Z}) = \sum_{n\in\mathbf{N}} n\,\mathbf{P}(\mathrm{Z} = n) = \sum_{n\geq 1} \mathbf{P}(\mathrm{Z} \geq n), \quad \text{d'où}$$
$$\mathbf{E}_i(\mathrm{N}_i) = \sum_{n\geq 1} \mathbf{P}_i(\mathrm{N}_i \geq n) = \frac{\alpha_i}{1-\alpha_i} < +\infty.$$

L'égalité $\mathbf{E}_i(\mathrm{N}_i) = \sum_{n\in\mathbf{N}} \mathrm{P}^n(i,i)$ découle immédiatement du théorème de Tonelli :

$$\mathbf{E}_i(\mathrm{N}_i) = \mathbf{E}_i\Big(\sum_{n\in\mathbf{N}^*} \mathbf{1}_{\mathrm{X}_n = i}\Big) = \sum_{n\in\mathbf{N}^*} \mathbf{E}_i(\mathbf{1}_{\mathrm{X}_n=i}) = \sum_{n\in\mathbf{N}^*} \mathbf{P}_i(\mathrm{X}_n = i) = \sum_{n\in\mathbf{N}^*} \mathrm{P}^n(i,i).$$

Ainsi les implications (a' $\Rightarrow$ b' $\Rightarrow$ c') sont établies, (b $\Rightarrow$ c) également (si la loi de N_i charge $+\infty$, N_i ne peut être intégrable et $\sum_{n\in\mathbf{N}} \mathrm{P}^n(i,i)$ diverge). Les implications (c $\Rightarrow$ a) (contraposée de (a' $\Rightarrow$ c')) et (c' $\Rightarrow$ a') (contraposée de (a $\Rightarrow$ c)) sont établies aussi. □

8. Elle charge 0 contrairement aux lois géométriques usuelles : pour tout n positif ou nul, $\mathbf{P}_i(\mathrm{N}_i = n) = (1-\alpha_i)(\alpha_i)^n$ avec $\alpha_i = \mathbf{P}_i(\tau_i < +\infty)$.

Proposition 3.7 *La récurrence et la transience sont des propriétés de classe : si les états i et j communiquent, alors i et j sont tous deux récurrents ou tous deux transients.*

DÉMONSTRATION. Si les états i et j communiquent, il existe des entiers $n \geq 1$ et $m \geq 1$ tels que , $\mathrm{P}^n(i,j) > 0$ et $\mathrm{P}^m(j,i) > 0$. P^{n+k+m} est le produit des trois matrices positives P^n, P^k, P^m, d'où :

$$\forall k \in \mathbf{N}, \quad \mathrm{P}^{n+k+m}(i,i) \geq \mathrm{P}^n(i,j)\,\mathrm{P}^k(j,j)\,\mathrm{P}^m(j,i)$$

soit en sommant sur k :

$$\sum_{k\in\mathbf{N}^*} \mathrm{P}^k(i,i) \geq \sum_{k\in\mathbf{N}^*} \mathrm{P}^{n+k+m}(i,i) \geq \mathrm{P}^n(i,j)\,\mathrm{P}^m(j,i) \sum_{k\in\mathbf{N}^*} \mathrm{P}^k(j,j).$$

D'après la proposition 3.6 un état i est transient ou récurrent selon que la série $\sum_{k\in\mathbf{N}^*} \mathrm{P}^k(i,i)$ converge ou diverge ; l'inégalité ci-dessus prouve que la convergence de la série $\sum_{k\in\mathbf{N}^*} \mathrm{P}^k(i,i)$ implique celle de $\sum_{k\in\mathbf{N}^*} \mathrm{P}^k(j,j)$ et que la divergence de $\sum_{k\in\mathbf{N}^*} \mathrm{P}^k(j,j)$ implique celle de $\sum_{k\in\mathbf{N}^*} \mathrm{P}^k(i,i)$, donc les deux séries sont toujours de même nature et les états i et j aussi. □

Proposition 3.8 *Tous les états d'une même classe récurrente sont visités $\mathbf{P}_j$-p.s. une infinité de fois à partir de n'importe quel état j de la classe : soient i et j deux états appartenant à la même classe récurrente, alors*

$$\mathbf{P}_j(\tau_i < +\infty) = \mathbf{P}_j(\mathrm{N}_i = +\infty) = 1.$$

DÉMONSTRATION. Comme les états i et j communiquent, il existe un entier n tel que $\mathrm{P}^n(i,j) > 0$. Si $\mathbf{P}_j(\tau_i = +\infty) > 0$, la probabilité de ne pas repasser une infinité de fois par i en partant de i est minorée par le produit $\mathrm{P}^n(i,j)\,\mathbf{P}_j(\tau_i = +\infty) > 0$, ce qui contredit le fait que i est récurrent. On a donc $\mathbf{P}_j(\tau_i = +\infty) = 0$ ou encore $\mathbf{P}_j(\tau_i < +\infty) = 1$ et en reportant ceci dans l'égalité (5) on obtient pour tout n, $\mathbf{P}_j(\mathrm{N}_i \geq n+1) = \mathbf{P}_i(\mathrm{N}_i \geq n) \geq \mathbf{P}_i(\mathrm{N}_i = +\infty) = 1$ (i est récurrent) puis par passage à la limite décroissante $\mathbf{P}_j(\mathrm{N}_i = +\infty) = 1$. □

Proposition 3.9 *La probabilité de sortir d'une classe récurrente est nulle ; plus précisément si i est un état récurrent et $\mathrm{C}(i)$ sa classe*

$$\forall j \notin \mathrm{C}(i), \quad \forall n \in \mathbf{N}, \quad \mathbf{P}_i(\mathrm{X}_n = j) = \mathrm{P}^n(i,j) = 0.$$

DÉMONSTRATION. Soit $j \notin \mathrm{C}(i)$, supposons qu'il existe un n tel que $\mathrm{P}^n(i,j) > 0$; dans ce cas, pour tout m, $\mathrm{P}^m(j,i) = 0$ sinon les états i et j communiqueraient. Mais alors la probabilité de non retour à i partant de i est non nulle car minorée par $\mathrm{P}^n(i,j) > 0$, ce qui contredit le fait que i est récurrent. □

Proposition 3.10 *Toute chaîne de Markov homogène sur un espace d'états fini a au moins un état récurrent. En particulier, toute chaîne irréductible sur un espace d'états fini est récurrente.*

DÉMONSTRATION. Montrons que pour tout état i transient et pour tout j, l'espérance du nombre de passages par l'état i, $\mathbf{E}_j(\mathrm{N}_i)$ est finie : rappelons que pour toute variable Z entière positive $\mathbf{E}(\mathrm{Z}) = \sum_{n\geq 1} \mathbf{P}(\mathrm{Z} \geq n)$, en utilisant l'égalité (5) on obtient

$$\begin{aligned}
\mathbf{E}_j(\mathrm{N}_i) = \sum_{n\in\mathbf{N}} \mathbf{P}_j(\mathrm{N}_i \geq n+1) &= \sum_{n\in\mathbf{N}} \mathbf{P}_j(\tau_i < +\infty)\, \mathbf{P}_i(\mathrm{N}_i \geq n) \\
&= \mathbf{P}_j(\tau_i < +\infty) \sum_{n\in\mathbf{N}} \mathbf{P}_i(\mathrm{N}_i \geq n) \\
&= \mathbf{P}_j(\tau_i < +\infty)\, \big(1 + \mathbf{E}_i(\mathrm{N}_i)\big).
\end{aligned}$$

donc $\mathbf{E}_j(\mathrm{N}_i) < +\infty$ pour tout état i transient.

Si tous les états de E (fini) étaient transients on aurait aussi

$$\mathbf{E}_j\Big(\sum_{i\in\mathrm{E}} \mathrm{N}_i\Big) = \sum_{i\in\mathrm{E}} \mathbf{E}_j(\mathrm{N}_i) < +\infty \quad \text{(Tonelli).}$$

ce qui est absurde puisque $\sum_{i\in\mathrm{E}} \mathrm{N}_i$ est le nombre total de visites aux états de E c'est-à-dire card($\mathbf{N}$) = $+\infty$. □

4 Théorèmes limites

Les deux questions que l'on est amené à se poser sont les suivantes :

1. Y-a-t-il convergence en loi des X_n lorsque $n \to +\infty$?

$$\mathbf{P}_i(\mathrm{X}_n = j) = \mathrm{P}^n(i,j) \xrightarrow[n\to+\infty]{} \mu(i,j)?$$

2. Y-a-t-il convergence des fréquences de passage par chaque état j ?

$$f_n^j(\omega) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\mathrm{X}_k(\omega)=j} \xrightarrow[n\to+\infty]{\mathbf{P}_i-\text{p.s.}} \lambda(i,j)?$$

Remarquons tout de suite que si les deux convergences ont lieu, les limites $\mu(i,j)$ et $\lambda(i,j)$ sont les mêmes : la fréquence $f_n^j(\omega)$ est majorée par 1, fonction $\mathbf{P}_i$-intégrable, donc par convergence dominée $\mathbf{E}_i(f_n^j) \to \mathbf{E}_i(\lambda(i,j)) = \lambda(i,j)$ et par linéarité $\mathbf{E}_i(f_n^j) = \frac{1}{n}\sum_{k=1}^{n} \mathrm{P}^k(i,j) \to \mu(i,j)$ (si une suite u_n converge, la suite de ses moyennes de Césaro converge vers la même limite).

4.1 Cas j transient

Lorsque j est un état transient, on a vu dans la démonstration de la proposition 3.10 que pour tout i, $\mathbf{E}_i(\mathrm{N}_j)$ est finie, mais

$$\mathbf{E}_i(\mathrm{N}_j) = \mathbf{E}_i\Big(\sum_{n\in\mathbf{N}^*} \mathbf{1}_{\mathrm{X}_n=j}\Big) = \sum_{n\in\mathbf{N}^*} \mathbf{E}_i(\mathbf{1}_{\mathrm{X}_n=j}) = \sum_{n\in\mathbf{N}^*} \mathrm{P}_i(\mathrm{X}_n=j) = \sum_{n\in\mathbf{N}^*} \mathrm{P}^n(i,j).$$

Cette série converge, son terme général tend donc vers 0.

Le nombre de visites à l'état transient j étant fini $\mathbf{P}_i$-p.s. pour tout i (même $\mathbf{P}_i$-intégrable, cf. ci-dessus), la suite $(\mathbf{1}_{\mathrm{X}_k(\omega)=j})_{k\in\mathbf{N}}$ est donc nulle pour tout k assez grand (dépendant de ω) et donc

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\mathrm{X}_k(\omega)=j} \xrightarrow[n\to+\infty]{\mathbf{P}_i-\text{p.s.}} 0.$$

Proposition 4.1 *Pour tout état i de E et pour tout état j transient on a*

$$\mathbf{P}_i(\mathrm{X}_n=j) = \mathrm{P}^n(i,j) \xrightarrow[n\to+\infty]{} 0 \quad \textit{et} \quad \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\mathrm{X}_k(\omega)=j} \xrightarrow[n\to+\infty]{\mathbf{P}_i-\text{p.s.}} 0.$$

4.2 Cas des chaînes irréductibles récurrentes

La probabilité de sortir d'une classe récurrente est nulle (cf. proposition 3.9). Si les états i et j sont récurrents, le seul cas non trivial est celui où ils sont dans la même classe, ce cas se ramène à l'étude d'une chaîne irréductible récurrente.

Le dernier cas, i transient et j récurrent, est laissé en exercice (voir exercice 7).

4.2.1 Mesures invariantes

Une façon commode de définir une mesure λ sur un espace d'états E dénombrable est de se donner un *vecteur ligne* $\lambda = (\lambda_i)_{i\in\mathrm{E}}$ où $\lambda_i = \lambda(\{i\})$.

Définition 4.2 *Soit P la matrice de transition d'une chaîne de Markov homogène ; on dit qu'une mesure π sur E est invariante par P si et seulement si elle est* positive *et π vérifie l'équation matricielle $\pi\mathrm{P} = \pi$.*

La proposition suivante justifie l'intérêt porté aux mesures invariantes.

Proposition 4.3 *a) Soit $(\mathrm{X}_n)_{n\in\mathbf{N}}$ une chaîne de Markov homogène. Si π est une* probabilité *invariante et si à un instant k, la loi de X_k est π, alors à tout instant ultérieur $m \geq n$, X_m est aussi de loi π.*

b) *Si* E *est un espace d'états fini et si pour tout couple* (i,j) *de* E^2*,* $\mathrm{P}^n(i,j) \to \mathrm{L}_i(j)$*, alors* L_i *est une probabilité invariante par* P*.*

DÉMONSTRATION. a) Si π est la loi de X_k, il est immédiat de vérifier que $\pi\mathrm{P}$ est la loi de X_{k+1}.

b) Si la suite de matrices P^n converge vers la matrice L, alors P^{n+1} converge aussi vers L et comme $\mathrm{P}^{n+1} = \mathrm{P}^n\mathrm{P}$, la matrice limite L vérifie $\mathrm{L} = \mathrm{LP}$ (linéarité du passage à la limite, E étant fini), donc chacune de ses lignes L_i vérifie $\mathrm{L}_i = \mathrm{L}_i\mathrm{P}$. □

Les mesures limites possibles sont donc à rechercher parmi les mesures invariantes.

Théorème 4.4 *Toute chaîne de Markov homogène récurrente irréductible admet une mesure invariante strictement positive sur* E *et toutes les mesures invariantes sont proportionnelles.*

Ce résultat découle des deux lemmes suivants, le premier établit l'existence d'une mesure invariante, le second son unicité à un facteur multiplicatif près.

Lemme 4.5 (existence) *Soit* X *une chaîne de Markov homogène de matrice* P*, irréductible récurrente. On fixe un état* k *et on pose*

$$\forall j \in \mathrm{E}, \quad \nu_j^k = \mathbf{E}_k\Big(\sum_{0 \le n < \tau_k} \mathbf{1}_{\mathrm{X}_n = j}\Big)$$

ν_j^k *représente le nombre moyen de passages par l'état* j *entre deux passages par l'état* k*. Le vecteur ligne* ν^k *a les propriétés suivantes :*

a) $\nu_k^k = 1$ *;*

b) ν^k *est invariant par* P *:* $\nu^k\mathrm{P} = \nu^k$ *;*

c) $\forall j \in \mathrm{E}, \quad 0 < \nu_j^k < +\infty$.

DÉMONSTRATION. a) est évident : pour $n = 0$, $\mathbf{1}_{\mathrm{X}_n = k} = 1$ $\mathbf{P}_k$-p.s. et les autres termes de la somme sont tous nuls ($n < \tau_k$).

b) Calculons le j-ième terme $(\nu^k\mathrm{P})_j$ du vecteur ligne $\nu^k\mathrm{P}$:

$$\begin{aligned}
\sum_{i \in \mathrm{E}} \nu_i^k \mathrm{P}(i,j) &= \sum_{i \in \mathrm{E}} \mathbf{E}_k\Big(\sum_{0 \le n < \tau_k} \mathbf{1}_{\mathrm{X}_n = i}\Big) \mathrm{P}(i,j) \\
&= \sum_{i \in \mathrm{E}} \mathbf{E}_k\Big(\sum_{n \in \mathbf{N}} \mathbf{1}_{n < \tau_k} \mathbf{1}_{\mathrm{X}_n = i}\Big) \mathrm{P}(i,j) \\
&= \sum_{i \in \mathrm{E}} \sum_{n \in \mathbf{N}} \mathbf{E}_k\Big(\mathbf{1}_{n < \tau_k} \mathbf{1}_{\mathrm{X}_n = i}\Big) \mathrm{P}(i,j) \quad \text{(Tonelli)}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i\in E}\sum_{n\in\mathbf{N}} \mathbf{P}_k(n<\tau_k, X_n=i)\,P(i,j)\\
&= \sum_{i\in E}\sum_{n\in\mathbf{N}} \mathbf{P}_k(X_n=i, X_n\neq k, X_{n-1}\neq k,\dots,X_1\neq k)\,P(i,j)\\
&= \sum_{\substack{i\in E\\ i\neq k}}\sum_{n\in\mathbf{N}} \mathbf{P}_k(X_n=i, X_{n-1}\neq k,\dots,X_1\neq k)\,P(i,j)
\end{aligned}
$$

or $P(i,j) = \mathbf{P}(X_{n+1}=j \mid X_n=i, X_{n-1}\neq k,\dots,X_1\neq k)$, d'où

$$
= \sum_{\substack{i\in E\\ i\neq k}}\sum_{n\in\mathbf{N}} \mathbf{P}_k(X_{n+1}=j, X_n=i, X_{n-1}\neq k,\dots,X_1\neq k)
$$

Pour $i\neq k$, l'événement $\{X_n=i, X_{n-1}\neq k,\dots,X_1\neq k\}$ s'écrit aussi $\{\tau_k>n, X_n=i\}$ donc $\bigcup_{i\neq k}\{X_n=i, X_{n-1}\neq k,\dots,X_1\neq k\} = \{\tau_k>n\}$, soit en permutant les sommes (Tonelli) :

$$
\begin{aligned}
&= \sum_{n\in\mathbf{N}} \mathbf{P}_k(X_{n+1}=j, \tau_k>n)\\
&= \sum_{n\in\mathbf{N}} \mathbf{E}_k(\mathbf{1}_{X_{n+1}=j}\,\mathbf{1}_{\tau_k>n}) = \mathbf{E}_k\Big(\sum_{n\in\mathbf{N}} \mathbf{1}_{X_{n+1}=j}\,\mathbf{1}_{\tau_k>n}\Big) \quad \text{(Tonelli)}\\
&= \mathbf{E}_k\Big(\sum_{0\le n<\tau_k} \mathbf{1}_{X_{n+1}=j}\Big) = \mathbf{E}_k\Big(\sum_{1\le m\le\tau_k} \mathbf{1}_{X_m=j}\Big) = \nu^k_j.
\end{aligned}
$$

La dernière égalité résulte du fait que, la chaîne étant récurrente irréductible, τ_k est $\mathbf{P}_k$-presque sûrement fini, on peut donc remplacer dans la somme, le terme d'indice $m=0$ par le terme d'indice $m=\tau_k$ puisque les deux indicatrices valent $\mathbf{1}_{j=k}$ $\mathbf{P}_k$-presque sûrement.

c) La chaîne (X_n) étant irréductible, tous les états communiquent et en particulier il existe deux entiers n et m tels que $P^n(i,k)>0$ et $P^m(k,i)>0$. D'après b), ν^k est invariant par P, donc aussi par toutes les puissances de P et comme d'après a), $\nu^k_k=1$:

$$
1=\nu^k_k = \sum_{j\in E}\nu^k_j P^n(j,k) \ge \nu^k_i P^n(i,k)
$$

donc, tout i, ν^k_i est fini et aussi strictement positif car

$$
\nu^k_i = \sum_{j\in E}\nu^k_j P^m(j,i) \ge \nu^k_k P^m(k,i) \ge P^m(k,i) > 0. \qquad \square
$$

Lemme 4.6 (unicité) *Soit X une chaîne de Markov homogène irréductible, de matrice de transition P et λ une mesure invariante par P telle que $\lambda_k=1$. Alors $\lambda\ge\nu^k$ où ν^k est la mesure définie dans le lemme 4.5. Si en plus X est récurrente, alors $\lambda=\nu^k$.*

DÉMONSTRATION. Écrivons l'équation d'invariance pour λ en isolant le terme d'indice k pour lequel $\lambda_k = 1$:

$$\forall j \in \mathrm{E}, \quad \lambda_j = \sum_{i\in \mathrm{E}} \lambda_i \mathrm{P}(i,j) = \sum_{i\neq k} \lambda_i \mathrm{P}(i,j) + \mathrm{P}(k,j).$$

On itère le procédé en isolant à chaque fois le terme en λ_k :

$$\begin{aligned}
\lambda_j &= \sum_{i\neq k}\Big(\sum_{i_1\neq k} \lambda_{i_1} \mathrm{P}(i_1,i) + \mathrm{P}(k,i)\Big)\mathrm{P}(i,j) + \mathrm{P}(k,j)\\
&= \sum_{i\neq k}\sum_{i_1\neq k} \lambda_{i_1} \mathrm{P}(i_1,i)\mathrm{P}(i,j) + \sum_{i\neq k} \mathrm{P}(k,i)\mathrm{P}(i,j) + \mathrm{P}(k,j)\\
&= \dots\\
&= \sum_{i\neq k}\sum_{i_1\neq k}\dots\sum_{i_n\neq k} \lambda_{i_n} \mathrm{P}(i_n,i_{n-1})\dots\mathrm{P}(i_1,i)\mathrm{P}(i,j)\\
&\quad + \sum_{i\neq k}\sum_{i_1\neq k}\dots\sum_{i_{n-1}\neq k} \mathrm{P}(k,i_{n-1})\mathrm{P}(i_{n-1},i_{n-2})\dots\mathrm{P}(i_1,i)\mathrm{P}(i,j) + \cdots\\
&\quad + \sum_{i\neq k}\sum_{i_1\neq k} \mathrm{P}(k,i_1)\mathrm{P}(i_1,i)\mathrm{P}(i,j) + \sum_{i\neq k} \mathrm{P}(k,i)\mathrm{P}(i,j) + \mathrm{P}(k,j)
\end{aligned}$$

on minore λ_j en négligeant la première somme (la seule à contenir des termes en λ) :

$$\begin{aligned}
\lambda_j &\geq \sum_{i\neq k}\sum_{i_1\neq k}\dots\sum_{i_{n-1}\neq k} \mathrm{P}(k,i_{n-1})\mathrm{P}(i_{n-1},i_{n-2})\dots\mathrm{P}(i_1,i)\mathrm{P}(i,j) + \cdots\\
&\quad + \sum_{i\neq k}\sum_{i_1\neq k} \mathrm{P}(k,i_1)\mathrm{P}(i_1,i)\mathrm{P}(i,j) + \sum_{i\neq k} \mathrm{P}(k,i)\mathrm{P}(i,j) + \mathrm{P}(k,j)
\end{aligned}$$

on interprète les produits grâce à la prop. 1.3, dans chaque somme on part de k pour arriver à j en évitant l'état k entre-temps, d'où :

$$\begin{aligned}
\lambda_j &\geq \mathbf{P}_k(\mathrm{X}_{n+1} = j, \tau_k \geq n+1) + \cdots + \mathbf{P}_k(\mathrm{X}_2 = j, \tau_k \geq 2) + \mathbf{P}_k(\mathrm{X}_1 = j, \tau_k \geq 1)\\
&\geq \sum_{m=1}^{n+1} \mathbf{P}_k(\mathrm{X}_m = j, \tau_k \geq m) = \mathbf{E}_k\Big(\sum_{m=1}^{n+1} \mathbf{1}_{\mathrm{X}_m=j}\, \mathbf{1}_{\tau_k\geq m}\Big) = \mathbf{E}_k\Big(\sum_{m=1}^{\min(n+1,\tau_k)} \mathbf{1}_{\mathrm{X}_m=j}\Big).
\end{aligned}$$

On fait tendre n vers $+\infty$: pour tout $j \neq k$, les termes d'indices $m = 0$ et $m = \tau_k$ (sur $\{\tau_k < +\infty\}$) de la dernière somme sont nuls, l'espérance croît donc vers ν^k_j (Beppo-Levi) et comme $\lambda_k = \nu^k_k = 1$, pour tout j, $\lambda_j \geq \nu^k_j$ et l'inégalité $\lambda \geq \nu^k$ est établie.

Si on suppose la chaîne X récurrente, on sait d'après le lemme 4.5 que ν_k est une mesure invariante. On pose $\mu = \lambda - \nu_k$, c'est également une mesure positive (cf. ci-dessus), invariante comme différence de mesures invariantes. Mais comme $\lambda_k =$

$\nu_k^k = 1$, $\mu_k = 0$ et comme la chaîne est irréductible tous les états communiquent : pour tout état j, il existe un entier n tel que $P^n(j,k) > 0$, μ est invariante par P^n, d'où :

$$\forall j \in E, \quad 0 = \mu_k = \sum_{i\in E} \mu_i P^n(i,k) \geq \mu_j P^n(j,k) \quad \text{donc} \quad \mu_j = 0 \quad \text{soit} \quad \lambda = \nu_k. \quad \square$$

Remarque : l'hypothèse λ mesure *positive* est essentielle pour l'unicité : la promenade aléatoire symétrique sur $\mathbf{Z}$ est récurrente irréductible mais l'équation $\lambda P = \lambda$ admet pour solutions les vecteurs ligne de la forme $(\lambda_i = A + Bi)_{i\in\mathbf{Z}}$ (voir exercice 3), c'est la condition $\lambda \geq 0$ sur $\mathbf{Z}$ qui impose $B = 0$ et fait que cette chaîne n'a que les mesures constantes comme mesures (positives) invariantes.

Il reste à voir si la masse totale de ces mesures invariantes (toutes proportionnelles) est finie ou non. Si elle l'est, il existera une probabilité invariante pour la chaîne. La masse totale de ν^k est :

$$\begin{aligned}
\nu^k(E) = \sum_{j\in E} \nu_j^k &= \sum_{j\in E} \mathbf{E}_k\Big(\sum_{0\leq n<\tau_k} \mathbf{1}_{X_n=j}\Big) \\
&= \mathbf{E}_k\Big(\sum_{0\leq n<\tau_k} \sum_{j\in E} \mathbf{1}_{X_n=j}\Big) \qquad \text{(Tonelli)} \\
&= \mathbf{E}_k\Big(\sum_{0\leq n<\tau_k} 1\Big) = \mathbf{E}_k(\tau_k).
\end{aligned}$$

Définition 4.7 *Pour tout état i récurrent, le temps τ_i de retour à i est fini $\mathbf{P}_i$-p.s. et deux cas se présentent :*
– *soit τ_i est aussi $\mathbf{P}_i$-intégrable, on dit alors que i est récurrent positif,*
– *soit τ_i est non intégrable ($\mathbf{E}_i(\tau_i) = +\infty$), on dit alors que i est récurrent nul.*

Le théorème suivant établit que la récurrence positive (resp. nulle) est une propriété de classe et donne une condition nécessaire et suffisante pour qu'une chaîne de Markov homogène irréductible soit récurrente positive.

Théorème 4.8 *Soit X une chaîne de Markov homogène irréductible, les trois propositions suivantes sont équivalentes :*

a) tous les états sont récurrents positifs,

b) il existe au moins un état récurrent positif,

c) X admet une probabilité invariante π.

Si l'une de ces conditions est réalisée, π est unique : $\quad \forall i \in E, \quad \pi_i = \dfrac{1}{\mathbf{E}_i(\tau_i)}$.

Démonstration. L'implication $a \Rightarrow b$ est triviale.
Montrons $b \Rightarrow c$: Si k est un état récurrent positif, $\mathbf{E}_k(\tau_k) = \nu^k(E)$ est fini, il suffit

de poser pour tout état i, $\pi_i = \nu_i^k / \nu^k(E)$ pour obtenir une probabilité π invariante. On a pour tout état k récurrent positif, $\pi_k = 1/\mathbf{E}_k(\tau_k)$ ce qui établit la formule finale.

Montrons $c \Rightarrow a$: soit π une probabilité invariante et k un état quelconque. Il existe au moins un état j tel que $\pi_j > 0$ et π, invariante par P, l'est aussi par toutes ses puissances P^n ; comme j et k communiquent il existe un n tel que $P^n(j,k) > 0$ et $\pi_k = \sum_{i\in E} \pi_i P^n(i,k) \geq \pi_j P^n(j,k) > 0$. En divisant π par la constante $\pi_k > 0$, on obtient une nouvelle mesure λ invariante par P et telle que $\lambda_k = 1$. On peut alors appliquer le lemme d'unicité (4.6) à λ, d'où $\lambda \geq \nu^k$ et on en déduit que k est récurrent positif :

$$\mathbf{E}_k(\tau_k) = \sum_{i\in E} \nu_i^k \leq \sum_{i\in E} \lambda_i = \sum_{i\in E} \frac{\pi_i}{\pi_k} = \frac{1}{\pi_k} < +\infty \quad (\pi \text{ probabilité}). \qquad \square$$

Corollaire 4.9 *Toute chaîne de Markov homogène irréductible à espace d'états fini est récurrente positive : elle admet une unique probabilité invariante définie par*

$$\forall i \in E, \quad \pi_i = \frac{1}{\mathbf{E}_i(\tau_i)}.$$

Démonstration. Une chaîne de Markov homogène irréductible sur un espace d'états fini est toujours récurrente (cf. prop. 3.10), elle admet donc des mesures invariantes (théorème 4.4) qui sont nécessairement de masse totale finie, d'où l'existence d'une probabilité invariante. Son expression est donnée par le théorème précédent. $\square$

L'exemple classique de chaîne récurrente nulle est donné par la promenade aléatoire symétrique ($p = q$) sur $\mathbf{Z}$.

Comme exemple de chaîne récurrente positive sur un espace dénombrable on peut citer la série de succès sur $\mathbf{N}$ avec pour tout i, $p_i = p$, $q_i = 1 - p$ (constants) ou les promenades aléatoires sur $\mathbf{N}$ avec barrière réfléchissante en 0 telles que $p < q$.

Remarque : l'existence de mesures invariantes pour une chaîne irréductible n'implique nullement que la chaîne soit récurrente : seule l'existence d'une *probabilité* invariante permet de conclure à la récurrence (positive). Il est facile de vérifier que les promenades aléatoires sur $\mathbf{Z}$ telles que $p > q$ sont irréductibles, transientes et qu'elles admettent pour mesures invariantes les mesures de la forme :

$$\forall i \in \mathbf{Z}, \quad \lambda_i = A + B\left(\frac{q}{p}\right)^i \quad A \geq 0, B \geq 0.$$

On remarquera aussi que dans cet exemple les mesures invariantes ne sont pas toutes proportionnelles...

Définition 4.10 *Soit* X *une chaîne de Markov homogène de matrice de transition* P*. On appelle mesure réversible pour* P *toute mesure* λ *vérifiant*

$$\forall (i,j) \in \mathrm{E}^2, \quad \lambda_i \, \mathrm{P}(i,j) = \lambda_j \, \mathrm{P}(j,i)$$

La notion de mesure réversible, liée au retournement du temps (non traité ici, voir [5]), peut faciliter la recherche d'une probabilité invariante : les équations ci-dessus sont beaucoup plus faciles à résoudre que les équations d'invariance et fournissent des mesures invariantes, comme le montre la proposition suivante.

Proposition 4.11 *Toute mesure* P*-réversible est* P*-invariante.*

DÉMONSTRATION. Soit λ P-réversible, calculons $(\lambda\mathrm{P})_j$:

$$(\lambda\mathrm{P})_j = \sum_{i\in\mathrm{E}} \lambda_i \, \mathrm{P}(i,j) = \sum_{i\in\mathrm{E}} \lambda_j \, \mathrm{P}(j,i) = \lambda_j \sum_{i\in\mathrm{E}} \mathrm{P}(j,i) = \lambda_j. \qquad \square$$

Exemple d'application : considérons la promenade aléatoire sur **N** avec barrière en 0, sa matrice de transition est donnée par

$$\mathrm{P}(0,0) = 1-r, \mathrm{P}(0,1) = r, \quad \forall i \geq 1, \mathrm{P}(i,i+1) = p > 0, \mathrm{P}(i,i-1) = q = 1-p > 0.$$

Les équations de réversibilité s'écrivent

$$\lambda_0 r = \lambda_1 q \quad \forall i \geq 1, \quad \lambda_i p = \lambda_{i+1} q.$$

La résolution est immédiate, les mesures réversibles sont de la forme

$$\forall i \geq 1, \quad \lambda_i = \frac{r}{q}\left(\frac{p}{q}\right)^{i-1} \lambda_0.$$

La chaîne est irréductible si et seulement si $r > 0$; dans ce cas, si en plus $p < q$, on a une probabilité invariante et la chaîne est récurrente positive.

4.2.2 Période d'un état

Le fait qu'une chaîne de Markov homogène soit irréductible récurrente positive, assure l'existence d'une unique probabilité invariante, mais pas la convergence des matrices P^n comme le montre l'exemple trivial suivant :

$$\text{Si } \mathrm{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \forall n \in \mathbf{N}, \quad \mathrm{P}^{2n} = \mathrm{I} \quad \text{et} \quad \mathrm{P}^{2n+1} = \mathrm{P}.$$

D'où la définition suivante :

Définition 4.12 *On appelle période d'un état i l'entier*

$$d(i) = PGCD\{n \geq 1 \mid \mathrm{P}^n(i,i) > 0\}$$

Dans l'exemple ci-dessus les deux états sont de période 2. La période est également une propriété de classe :

Proposition 4.13 *Tous les états d'une même classe ont même période.*

DÉMONSTRATION. Soient i et j deux états communicants, montrons que $d(j)$ divise $d(i)$, ce qui suffit par symétrie pour établir que $d(i) = d(j)$. Comme i et j communiquent, il existe deux entiers l et m tels que $\mathrm{P}^l(i,j) > 0$ et $\mathrm{P}^m(j,i) > 0$. Si n est tel que $\mathrm{P}^n(i,i) > 0$, alors $\mathrm{P}^{m+n+l}(j,j) \geq \mathrm{P}^m(j,i)\mathrm{P}^n(i,i)\mathrm{P}^l(i,j) > 0$, donc $d(j)$ divise $m+n+l$. Mais comme $\mathrm{P}^{m+l}(j,j) \geq \mathrm{P}^m(j,i)\mathrm{P}^l(i,j) > 0$, $d(j)$ divise aussi $m+l$ et par différence $d(j)$ divise n. $d(j)$ divise donc tous les entiers tels que $\mathrm{P}^n(i,i) > 0$, donc aussi leur PGCD $d(i)$. □

Définition 4.14 *On dit qu'une classe est apériodique si et seulement si tous ses états sont de période* 1*.*

Remarque : Si pour un état i, $\mathrm{P}(i,i) > 0$, alors la classe de i est apériodique.

Exemple de classe périodique : outre l'exemple trivial donné ci-dessus, les promenades aléatoires sur $\mathbf{Z}$, $\mathrm{S}_n = \mathrm{S}_0 + \sum_{i=1}^n \mathrm{Y}_i$ où les Y_i valent ± 1 sont des exemples typiques de chaînes irréductibles de période 2.

Proposition 4.15 *Soit i un état de période* 1*, alors*

a) *il existe un entier $\mathrm{N}(i)$ tel que pour tout $n \geq \mathrm{N}(i)$, $\mathrm{P}^n(i,i) > 0$;*

b) *pour tout état j communiquant avec i, il existe un entier $\mathrm{N}(i,j)$ tel que pour tout $n \geq \mathrm{N}(i,j)$, $\mathrm{P}^n(i,j) > 0$.*

DÉMONSTRATION. a) Soit $\mathscr{D}(i) = \{n \geq 1 \mid \mathrm{P}^n(i,i) > 0\}$, cet ensemble est (presque) un idéal : si n et m appartiennent à $\mathscr{D}(i)$, pour tous p et q de $\mathbf{N}$, $pn + qm$ appartient aussi à $\mathscr{D}(i)$[9]. En effet, $\mathrm{P}^{pn+qm}(i,i) \geq (\mathrm{P}^n(i,i))^p\ (\mathrm{P}^m(i,i))^q$ (produit de matrices positives).

Montrons d'abord qu'il existe un entier m tel que m et $m+1$ appartiennent à $\mathscr{D}(i)$: le PGCD de $\mathscr{D}(i)$ valant 1, d'après l'égalité de Bezout il existe une suite finie[10] $(n_1,\ldots,n_k)$ d'entiers de $\mathscr{D}(i)$ et $(a_1,\ldots,a_k) \in \mathbf{Z}^k$ tels que $\sum_{j=1}^k a_j n_j = 1$. On sépare les a_j selon leur signe en posant $a_j^+ = \max(a_j,0)$ et $a_j^- = \max(-a_j,0)$, l'égalité de

9. Si la propriété était vraie pour p et q dans $\mathbf{Z}$, $\mathscr{D}(i)$ serait un « vrai » idéal de $\mathbf{Z}$.

10. Le PGCD $d_k(i)$ de $\{n \geq 1 \mid n \leq k,\ \mathrm{P}^n(i,i) > 0\}$ décroît vers $d(i)$ lorsque $k \to +\infty$, comme c'est une suite d'entiers $d_k(i) = d(i)$ pour k assez grand.

Bezout s'écrit alors $\sum_{j=1}^{k} a_j^+ n_j = 1 + \sum_{j=1}^{k} a_j^- n_j$. En posant $m = \sum_{j=1}^{k} a_j^- n_j$ on a $m+1 = \sum_{j=1}^{k} a_j^+ n_j$ et ces deux entiers sont dans $\mathscr{D}(i)$.

Il suffit maintenant de prendre $\mathrm{N}(i) = m^2$: pour tout $n \geq m^2$ la division de n par m s'écrit : $n = qm + r$ avec $0 \leq r < m$ et $q \geq m > r$, on peut donc écrire q sous la forme $q = r + p$ avec $p \in \mathbf{N}$ d'où $n = (r+p)m + r = r(m+1) + pm$ donc n appartient à $\mathscr{D}(i)$.

b) Si i communique avec j, j est également de période 1 et il existe un entier s tel que $\mathrm{P}^s(i,j) > 0$; en posant $\mathrm{N}(i,j) = \mathrm{N}(j) + s$ on a pour tout $l \geq \mathrm{N}(i,j)$, $l = n + s$ avec $n \geq \mathrm{N}(j)$, d'où $\mathrm{P}^l(i,j) \geq \mathrm{P}^s(i,j)\mathrm{P}^n(j,j) > 0$. □

Pour les chaînes à espace d'états fini, la proposition précédente admet une forme plus forte et d'énoncé plus simple :

Proposition 4.16 *Soit* X *une chaîne de Markov homogène irréductible sur un espace d'états fini et* P *sa matrice de transition. Si* X *est apériodique, les matrices* P^n *sont toutes strictement positives pour n assez grand. Réciproquement, si il existe un n pour lequel* P^n *est strictement positive, toutes les matrices* $(\mathrm{P}^m)_{m \geq n}$ *sont aussi strictement positives et* X *est apériodique.*

DÉMONSTRATION. Posons $\mathrm{N} = \max(\max_{i \in \mathrm{E}} \mathrm{N}(i), \max_{(i,j) \in \mathrm{E}^2} \mathrm{N}(i,j))$. Si X est apériodique, N est fini d'après la proposition précédente et tous les termes des matrices P^n, sont strictement positifs pour $n \geq \mathrm{N}$.

Réciproquement, si pour un entier n, P^n est strictement positive, il est clair que toutes les matrices $(\mathrm{P}^m)_{m \geq n}$ le sont aussi (écrire le produit $\mathrm{P}^{n+1} = \mathrm{P} \cdot \mathrm{P}^n$) et le fait que $\mathrm{P}^n(i,i)$ et $\mathrm{P}^{n+1}(i,i)$ soient strictement positifs implique que $\mathscr{D}(i)$ contient n et $n+1$, donc aussi leur différence, d'où $d(i) = 1$ et la chaîne X est apériodique. □

Ce résultat ne s'étend évidemment pas aux chaînes irréductibles sur un espace d'états dénombrable : penser à la chaîne des séries de succès, elle est apériodique irréductible mais pour tout état i, $\mathrm{P}^n(i,i) > 0 \Rightarrow n \geq i+1$ car pour revenir en i il est nécessaire de passer par l'état 0.

4.2.3 Convergence en loi

Théorème 4.17 (Convergence en loi) *Soit* X *une chaîne de Markov homogène irréductible apériodique pour laquelle il existe une probabilité invariante* π*. Alors, pour toute loi initiale* λ, $\mathbf{P}_\lambda(\mathrm{X}_n = j) \xrightarrow[n \to +\infty]{} \pi_j$*. En particulier, pour tout état i,* $\mathrm{P}^n(i,j) \xrightarrow[n \to +\infty]{} \pi_j$*.*

La démonstration proposée ici repose sur un argument de couplage emprunté à [5].

Démonstration. Soit X la chaîne de Markov homogène de matrice de transition P et de loi initiale λ. On considère une seconde chaîne de Markov homogène Y de même matrice de transition P mais de loi initiale π invariante par P, indépendante de X ; une telle chaîne Y existe : pour construire le couple (X, Y) il suffit de se placer sur l'espace produit $E^{\mathbf{N}} \times E^{\mathbf{N}}$ muni de la tribu produit des tribus cylindriques et de la probabilité produit $\mathbf{P}_\lambda \times \mathbf{P}_\pi$. Soit $(W_n) = (X_n, Y_n)$ la chaîne « couplée », c'est une chaîne de Markov homogène de loi initiale $\lambda \times \pi$ et de matrice de transition définie par $\tilde{P}((i,k)(j,l)) = P(i,j)P(k,l)$.

Cette chaîne est irréductible, en effet pour tout n, $\tilde{P}^n((i,k)(j,l)) = P^n(i,j)P^n(k,l)$ (indépendance) et d'après la proposition 4.15, ce produit est strictement positif pour n assez grand car X est irréductible apériodique (Y qui a même matrice de transition l'est donc aussi).

Il est immédiat de vérifier que la probabilité $\tilde{\pi}$ définie sur $E \times E$ par $\tilde{\pi}_{(i,k)} = \pi_i \pi_k$ est invariante par $\tilde{P}$. D'après le théorème 4.8, la chaîne couple W est donc récurrente positive.

Fixons un état a quelconque de E. Le temps d'atteinte T de l'état $\{a\} \times \{a\}$ par la chaîne récurrente W est donc fini $\tilde{\mathbf{P}}_{\lambda\times\pi}$-p.s. On définit maintenant une nouvelle suite (Z_n) en remplaçant X_n par Y_n à partir l'instant T :

$$Z_n(\omega) = \begin{cases} X_n(\omega) & \text{sur } \{T(\omega) > n\} \\ Y_n(\omega) & \text{sur } \{T(\omega) \leq n\} \end{cases}$$

$Z = (Z_n)$ est une chaîne de Markov homogène de loi initiale λ et de matrice de transition P (calculer $\mathbf{P}(Z_{n+1} = j \mid Z_n = i \cap \ldots \cap Z_0 = i_0)$ en distinguant trois cas selon que Z_n passe pour la première fois par $\{a\}$ après l'instant n, à l'instant n ou avant).

X et Z sont donc deux chaînes de Markov homogènes de même loi initiale λ et de même matrice de transition P, elles ont donc même loi (cf. prop. 1.3). Calculons la différence

$$\mathbf{P}_\lambda(X_n = j) - \pi_j = \tilde{\mathbf{P}}_{\lambda\times\pi}(X_n = j) - \tilde{\mathbf{P}}_{\lambda\times\pi}(Y_n = j) = \tilde{\mathbf{P}}_{\lambda\times\pi}(Z_n = j) - \tilde{\mathbf{P}}_{\lambda\times\pi}(Y_n = j)$$

en décomposant sur $\{T > n\}$ et $\{T \leq n\}$ pour utiliser le fait que $Z_n = Y_n$ sur $\{T \leq n\}$:

$$\begin{aligned} \mathbf{P}_\lambda(X_n = j) - \pi_j &= \tilde{\mathbf{P}}_{\lambda\times\pi}(Z_n = j, T > n) + \tilde{\mathbf{P}}_{\lambda\times\pi}(Z_n = j, T \leq n) \\ &\quad - \tilde{\mathbf{P}}_{\lambda\times\pi}(Y_n = j, T > n) - \tilde{\mathbf{P}}_{\lambda\times\pi}(Y_n = j, T \leq n) \\ &= \tilde{\mathbf{P}}_{\lambda\times\pi}(Z_n = j, T > n) - \tilde{\mathbf{P}}_{\lambda\times\pi}(Y_n = j, T > n) \end{aligned}$$

Finalement, $\left|\mathbf{P}_\lambda(X_n = j) - \pi_j\right| = \left|\tilde{\mathbf{P}}_{\lambda\times\pi}(Z_n = j, T > n) - \tilde{\mathbf{P}}_{\lambda\times\pi}(Y_n = j, T > n)\right|$ est majorée par $\tilde{\mathbf{P}}_{\lambda\times\pi}(T > n)$ qui décroît vers $\tilde{\mathbf{P}}_{\lambda\times\pi}(T = +\infty) = 0$ lorsque n croît vers $+\infty$.

$\square$

4.2.4 Théorème ergodique

Théorème 4.18 *Soit* X *une chaîne de Markov homogène irréductible de loi initiale* λ *quelconque. Alors*[11] *:*

$$\forall j \in E, \quad \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{X_k = j} \xrightarrow[n\to+\infty]{\mathbf{P}_\lambda-\text{p.s.}} \frac{1}{\mathbf{E}_j(\tau_j)}$$

De plus, si X *est récurrente positive de probabilité invariante* π*, pour toute fonction* f $E \to \mathbf{R}$ *bornée*[12] *:*

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \xrightarrow[n\to+\infty]{\mathbf{P}_\lambda-\text{p.s.}} \sum_{i\in E} \pi_i f(i) = \mathbf{E}_\pi(f(X_0)).$$

Démonstration. Si X est transiente, chaque état j n'est visité qu'un nombre $\mathbf{P}_\lambda$-p.s. fini de fois et le temps τ_j d'atteinte de chaque état j charge $+\infty$ donc $\mathbf{E}_j(\tau_j) = +\infty$, on donc bien

$$\forall j \in E \quad \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{X_k = j} \xrightarrow[n\to+\infty]{\mathbf{P}_\lambda-\text{p.s.}} 0 = \frac{1}{\mathbf{E}_j(\tau_j)}$$

Supposons maintenant X récurrente. L'idée consiste à appliquer la loi forte des grands nombres à la suite de variables $\{\tau_j \circ \theta^{\tau_j^l}\}_{l\geq 1}$ qui sont indépendantes et de même loi pour la probabilité $\mathbf{P}_\lambda$. Rappelons que les instants de passage par j sont donnés par $\tau_j^1 = \tau_j$ et $\tau_j^{l+1} = \tau_j^l + \tau_j \circ \theta^{\tau_j^l}$.

Notons V_j^n la variable aléatoire nombre de visites à l'état j entre les instants 0 et n : $V_j^n = \sum_{k=1}^{n} \mathbf{1}_{X_k = j}$. Il est clair que $V_j^n \leq n$, ses valeurs sont données par :

$$V_j^n = 0 \quad \text{sur} \quad \{\tau_j > n\}$$
$$V_j^n = m \quad \text{sur} \quad \{\tau_j^m \leq n < \tau_j^{m+1}\} = \left\{\sum_{l=0}^{m-1} \tau_j \circ \theta^{\tau_j^l} \leq n < \sum_{l=0}^{m} \tau_j \circ \theta^{\tau_j^l}\right\} \quad (1 \leq m \leq n).$$

Sur l'ensemble $\{V_j^n \geq 1\} = \{\tau_j \leq n\}$ on a donc :

$$\sum_{l=0}^{V_j^n - 1} \tau_j \circ \theta^{\tau_j^l} \leq n < \sum_{l=0}^{V_j^n} \tau_j \circ \theta^{\tau_j^l}$$

11. On fait la convention $1/\infty = 0$ lorsque j est transient ou récurrent nul.
12. Le résultat peut se généraliser à f π-intégrable (voir [1]).

ou en divisant par V_j^n

$$\alpha_n(\omega) = \frac{1}{V_j^n} \sum_{l=0}^{V_j^n - 1} \tau_j \circ \theta^{\tau_j^l} \leq \frac{n}{V_j^n} < \frac{1}{V_j^n} \sum_{l=0}^{V_j^n} \tau_j \circ \theta^{\tau_j^l} = \beta_n(\omega). \tag{6}$$

Les variables $\{\tau_j \circ \theta^{\tau_j^l}\}_{l \geq 1}$ étant indépendantes et de même loi[13] pour la probabilité $\mathbf{P}_\lambda$, d'après la loi forte des grands nombres, pour $\mathbf{P}_\lambda$-presque tout ω

$$\gamma_n(\omega) = \frac{1}{n} \sum_{l=0}^{n} \tau_j \circ \theta^{\tau_j^l} \xrightarrow[n \to +\infty]{} \mathbf{E}_j(\tau_j).$$

Ceci vaut que τ_j soit intégrable ou non : si (Y_l) est une suite de variables positives, indépendantes, de même loi, non intégrables, $\frac{1}{n}\sum_{l=1}^{n} Y_l \xrightarrow[n\to+\infty]{\mathbf{P}-\text{p.s.}} +\infty$ (appliquer la loi forte des grands nombres aux variables tronquées $Y_l^k = \min(Y_l, k)$ et faire tendre k vers $+\infty$).

Lorsque n tend vers $+\infty$, $V_j^n(\omega)$ croît vers le nombre total $N_j(\omega)$ de visites à l'état j qui vaut $+\infty$ presque sûrement : la chaîne étant récurrente irréductible, pour tout i, $\mathbf{P}_i(N_j = +\infty) = 1$ (cf. proposition 3.8) donc $\mathbf{P}_\lambda(N_j = +\infty) = 1$. La suite $V_j^n(\omega)$ croît (au sens large) vers $+\infty$ $\mathbf{P}_\lambda$-p.s., quitte à sauter des termes (ceux d'indice n tels que $V_j^n = V_j^{n-1}$) les deux suites $\alpha_n(\omega)$ et $\beta_n(\omega)$ sont des sous-suites de $\gamma_n(\omega)$ et ont donc même limite que $\gamma_n(\omega)$ soit $\mathbf{E}_j(\tau_j)$.

Enfin, τ_j étant finie $\mathbf{P}_\lambda$-p.s., l'équation (6) vérifiée sur $\{\tau_j \leq n\}$, est vraie pour $\mathbf{P}_\lambda$-presque tout ω au moins pour n assez grand (dépendant de ω) et le quotient n/V_j^n, encadré par $\alpha_n(\omega)$ et $\beta_n(\omega)$ converge donc $\mathbf{P}_\lambda$-p.s. vers $\mathbf{E}_j(\tau_j)$. En passant aux inverses on a le résultat annoncé dans la première partie du théorème.

Étudions maintenant la limite de $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$ pour f bornée par M dans le cas où X admet une probabilité invariante π. Pour cela transformons la somme

$$\begin{aligned}
\frac{1}{n}\sum_{k=1}^{n} f(X_k) &= \frac{1}{n}\sum_{k=1}^{n} f(X_k) \sum_{i\in E} \mathbf{1}_{X_k = i} \\
&= \frac{1}{n}\sum_{k=1}^{n} \sum_{i\in E} f(i)\, \mathbf{1}_{X_k = i} \\
&= \frac{1}{n}\sum_{i\in E} f(i) \sum_{k=1}^{n} \mathbf{1}_{X_k = i} \\
&= \frac{1}{n}\sum_{i\in E} f(i) V_i^n = \sum_{i\in E} f(i) \frac{V_i^n}{n}
\end{aligned}$$

13. Le fait que pour $l = 0$, τ_j ait une loi différente (la chaîne ne part pas de j) n'affecte pas la limite puisque $\frac{1}{n}\tau_j \to 0$ lorsque $n \to +\infty$.

Majorons la différence

$$\Big|\frac{1}{n}\sum_{k=1}^{n} f(X_k) - \sum_{i\in E} f(i)\pi_i\Big| = \Big|\sum_{i\in E} f(i)\Big(\frac{V_i^n}{n} - \pi_i\Big)\Big| \leq M \sum_{i\in E}\Big|\frac{V_i^n}{n} - \pi_i\Big|$$

Chacun des termes de cette dernière somme tend vers 0 d'après la première partie du théorème ergodique, mais E n'est pas supposé fini, soit donc F une partie finie de E :

$$\begin{aligned}
\sum_{i\in E}\Big|\frac{V_i^n}{n} - \pi_i\Big| &\leq \sum_{i\in F}\Big|\frac{V_i^n}{n} - \pi_i\Big| + \sum_{i\notin F}\frac{V_i^n}{n} + \sum_{i\notin F}\pi_i \\
&\leq \sum_{i\in F}\Big|\frac{V_i^n}{n} - \pi_i\Big| + 1 - \sum_{i\in F}\frac{V_i^n}{n} + \sum_{i\notin F}\pi_i \quad \text{car } \sum_{i\in E}\frac{V_i^n}{n} = 1 \\
&\leq \sum_{i\in F}\Big|\frac{V_i^n}{n} - \pi_i\Big| + \sum_{i\notin F}\pi_i + \sum_{i\in F}\pi_i - \sum_{i\in F}\frac{V_i^n}{n} + \sum_{i\notin F}\pi_i \quad (\pi \text{ probabilité}) \\
&\leq 2\sum_{i\in F}\Big|\frac{V_i^n}{n} - \pi_i\Big| + 2\sum_{i\notin F}\pi_i
\end{aligned}$$

Pour majorer $\sum_{i\in E}|V_i^n/n - \pi_i|$ par ε, il suffit de choisir pour F une partie finie de E assez grande pour que $\sum_{i\notin F}\pi_i$ soit plus petite que $\varepsilon/4$, la première somme (finie) est alors majorée par $\varepsilon/4$ pour n assez grand. □

Le théorème ergodique complète au moins partiellement l'information donnée par le théorème de convergence en loi. Comme les variables V_i^n/n sont comprises entre 0 et 1, on peut leur appliquer le théorème de Lebesgue et

$$\mathbf{E}_i\,\frac{V_j^n}{n} = \mathbf{E}_i\Big(\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{X_k=j}\Big) = \frac{1}{n}\sum_{k=1}^{n}\mathbf{P}_i(X_k = j) = \frac{1}{n}\sum_{k=1}^{n}P^k(i,j) \xrightarrow[n\to\infty]{} \frac{1}{\mathbf{E}_j(\tau_j)}.$$

pour toute chaîne de Markov homogène irréductible. Sous des hypothèses plus faibles (on ne suppose la chaîne ni récurrente ni apériodique) on obtient un résultat plus faible aussi : la convergence des moyennes de Césaro des $P^n(i,j)$ au lieu de celle de la suite des $P^n(i,j)$. On a établi le résultat suivant :

Proposition 4.19 *Les moyennes de Césaro des matrices* P^n *d'une chaîne de Markov homogène irréductible convergent :*

$$\forall (i,j)\in E\times E \quad \frac{1}{n}\sum_{k=1}^{n}P^k(i,j) \xrightarrow[n\to\infty]{} \frac{1}{\mathbf{E}_j(\tau_j)}$$

avec la convention $1/\infty = 0$ *pour les états* j *transients ou récurrents nuls.*

5 Propriétés algébriques des chaînes de Markov à espace d'états fini

Dans cette section nous supposons E fini et nous abordons l'étude de la convergence des matrices P^n sous un angle purement algébrique, celui des valeurs propres de P.

Définition 5.1 *On appelle matrice stochastique ou matrice de Markov, toute matrice à termes positifs ou nuls dont la somme de chaque ligne vaut* 1.

Proposition 5.2 *Soit* P *une matrice stochastique* $N \times N$.

a) P *admet la valeur propre* 1 *pour le vecteur propre colonne*[14] ${}^t(1, 1, \ldots, 1)$ *;*

b) les valeurs propres de P *sont toutes de module inférieur ou égal à* 1 *;*

c) plus précisément, elles sont toutes dans la réunion des disques de centre $P(i,i)$ *et de rayon* $1 - P(i,i)$ $(1 \le i \le N)$.

DÉMONSTRATION. a) traduit le fait que la somme des colonnes vaut 1.

b) Soit λ une valeur propre (complexe) de P et v le vecteur propre associé, de composantes $(v_1, v_2, \ldots, v_N)$. Considérons la composante de module maximum de v : $|v_i| = \max_{1 \le j \le N} |v_j|$. Le produit $Pv = \lambda v$ s'écrit pour cette composante :

$$\sum_{j=1}^{n} P(i,j) v_j = \lambda v_i \quad \text{d'où} \quad |\lambda|\,|v_i| \le \sum_{j=1}^{n} P(i,j)\,|v_j| \le |v_i| \sum_{j=1}^{n} P(i,j) = |v_i|$$

et $|\lambda| \le 1$ (v, vecteur propre, n'est pas nul donc $v_i \neq 0$).

c) On réécrit l'égalité $Pv = \lambda v$ pour la composante v_i de module maximum :

$$\lambda v_i = \sum_{j=1}^{n} P(i,j) v_j \Longleftrightarrow (\lambda - P(i,i)) v_i = \sum_{j \neq i} P(i,j) v_j$$

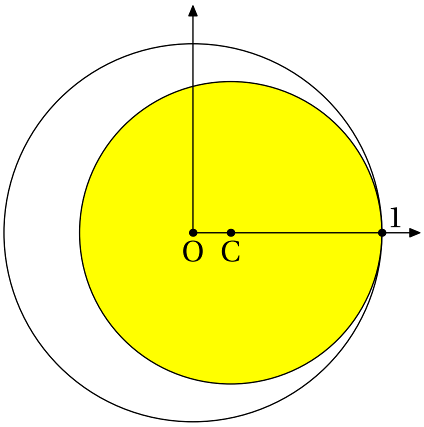
En passant aux modules on obtient :

$$|\lambda - P(i,i)|\,|v_i| \le \sum_{j \neq i} P(i,j)\,|v_j| \le |v_i| \sum_{j \neq i} P(i,j) = |v_i|\,(1 - P(i,i))$$

comme $|v_i| \neq 0$, il existe donc un i pour lequel $|\lambda - P(i,i)| \le 1 - P(i,i)$, ce qui établit le résultat annoncé. □

14. On notera ${}^t v$ et ${}^t P$ les transposés d'un vecteur v ou d'une matrice P.

Application :



Si tous les termes diagonaux de P sont strictement positifs, les valeurs propres de P sont toutes dans le disque de centre $\mathrm{C} = \min_i \mathrm{P}(i,i) > 0$ et de rayon $1 - \min_i \mathrm{P}(i,i)$; en particulier 1 est l'unique valeur propre de module 1 de P (éventuellement multiple).

Si P est diagonalisable, si 1 est sa seule valeur propre de module 1 et si 1 est valeur propre simple, il est facile de montrer que

$$\forall (i,j) \in \mathrm{E} \times \mathrm{E} \quad \mathrm{P}^n(i,j) \xrightarrow[n\to\infty]{} \pi_j \quad \text{indépendante de } i.$$

En effet dans ce cas la suite des matrices diagonales D^n converge vers une matrice D^∞ de rang 1, donc P^n converge[15] vers $\mathrm{P}^\infty = \mathrm{Q}\mathrm{D}^\infty\mathrm{Q}^{-1}$ où Q est une matrice inversible, donc P^∞ est également de rang 1 et comme toutes ses lignes ont même somme, toutes ses lignes doivent être égales (à π).

Il ne faudrait pas croire que toute matrice stochastique est diagonalisable, voici un exemple de matrice stochastique 3×3 triangulaire non diagonalisable :

$$\mathrm{P} = \begin{pmatrix} ½ & a & ½ - a \\ 0 & ½ & ½ \\ 0 & 0 & 1 \end{pmatrix}$$

Les valeurs propres sont ½, ½ et 1. Si $a \neq 0$, le sous-espace propre associé à ½ est de dimension 1 et P n'est donc pas diagonalisable.

Nous allons maintenant donner deux démonstrations simplifiées de l'existence d'une probabilité invariante dans le cas où l'espace d'états E est fini.

Proposition 5.3 *Toute chaîne de Markov homogène sur un espace d'états fini admet au moins une probabilité invariante.*

Voici une première démonstration algébrique :

Démonstration. une probabilité m est invariante si et seulement si $m\mathrm{P} = m$ soit en transposant ${}^t\mathrm{P}\,{}^tm = {}^tm$ c'est-à-dire ${}^tm$ est un vecteur propre de ${}^t\mathrm{P}$ pour la valeur propre 1. Or 1 est toujours valeur propre de P (cf. proposition précédente) et P et ${}^t\mathrm{P}$ ont même polynôme caractéristique donc mêmes valeurs propres ; ceci

15. Les passages à la limite ne posent pas de problème puisqu'il s'agit de combinaisons linéaires *finies*.

assure l'existence d'un vecteur ligne m tel que $m\mathrm{P} = m$, mais pas que les composantes de m soient toutes positives ! L'existence d'un vecteur m *positif* invariant découle du lemme suivant. Il est ensuite facile de le normer ($\sum_{i\in\mathrm{E}} m_i = 1$) pour en faire une probabilité. □

Lemme 5.4 (Perron-Frobenius) *Soit* P *une matrice stochastique* $\mathrm{N}\times\mathrm{N}$, λ *une valeur propre (réelle ou complexe) de module* 1 *de sa transposée* ${}^t\mathrm{P}$ *et* v *le vecteur propre (réel ou complexe) associé :* ${}^t\mathrm{P}v = \lambda v$. *Alors, le vecteur colonne* $w = {}^t(|v_1|,\dots,|v_\mathrm{N}|)$, *est vecteur propre de* ${}^t\mathrm{P}$ *pour la valeur propre* 1.

Démonstration. Calculons la i-ème composante α_i du produit $({}^t\mathrm{P}-\mathrm{I})w$:

$$\alpha_i = \sum_{j\in\mathrm{E}} {}^t\mathrm{P}(i,j)w_j - w_i = \sum_{j\in\mathrm{E}} {}^t\mathrm{P}(i,j)|v_j| - |v_i| \geq \left|\sum_{j\in\mathrm{E}} {}^t\mathrm{P}(i,j)v_j\right| - |v_i| = 0$$

car $\sum_{j\in\mathrm{E}} {}^t\mathrm{P}(i,j)v_j = \lambda v_i$ et $|\lambda| = 1$. Il reste à remarquer que la somme des α_i est nulle :

$$\sum_{i\in\mathrm{E}} \alpha_i = \sum_{i\in\mathrm{E}}\Big(\sum_{j\in\mathrm{E}} {}^t\mathrm{P}(i,j)w_j - w_i\Big) = \sum_{j\in\mathrm{E}} w_j \sum_{i\in\mathrm{E}} {}^t\mathrm{P}(i,j) - \sum_{i\in\mathrm{E}} w_i = \sum_{j\in\mathrm{E}} w_j - \sum_{i\in\mathrm{E}} w_i = 0$$

donc les α_i sont tous nuls. □

Donnons maintenant une démonstration topologique de la proposition 5.3.

Démonstration. Une probabilité sur un ensemble E à N éléments est définie par les valeurs des probabilités de chaque singleton, c'est donc un vecteur de $\mathbf{R}^\mathrm{N}$ dont toutes les composantes sont comprises entre 0 et 1 et dont la somme vaut 1. Autrement dit, l'ensemble $\mathscr{M}_1(\mathrm{E})$ de ces probabilités est l'intersection de l'hypercube $[0,1]^\mathrm{N}$ et de l'hyperplan affine d'équation $\sum_{i=1}^{\mathrm{N}} x_i = 1$, c'est donc un compact de l'e.v. normé $\mathbf{R}^\mathrm{N}$.

On dit qu'une suite (μ_n) de probabilités sur E (fini) converge vers μ, si et seulement si, pour tout singleton $\{i\}$ de E, la suite numérique $\mu_n(\{i\})$ converge vers $\mu(\{i\})$ [16]. La topologie ainsi définie sur $\mathscr{M}_1(\mathrm{E})$ est équivalente à celle de $\mathbf{R}^\mathrm{N}$ (convergence des composantes), $\mathscr{M}_1(\mathrm{E})$ est donc compact.

Soit μ_0 une probabilité quelconque sur E, pour tout m de $\mathbf{N}$, $\mu_0\mathrm{P}^m$ est également une probabilité sur E, considérons la suite définie par :

$$\forall n \geq 1 \quad \mu_n = \frac{1}{n}(\mu_0 + \mu_0\mathrm{P} + \cdots + \mu_0\mathrm{P}^{n-1})$$

16. Noter que cette définition équivaut à la convergence faible (ou convergence en loi) de (μ_n) vers μ.

Les μ_n sont elles aussi des probabilités sur E et la compacité de $\mathscr{M}_1(E)$ permet d'extraire de la suite (μ_n) une sous-suite convergente (μ_{n_k}) ; pour tout k :

$$\mu_{n_k}P - \mu_{n_k} = \frac{1}{n_k}(\mu_0 P^{n_k} - \mu_0) \xrightarrow[k\to+\infty]{} 0$$

car pour tout singleton $\{i\}$ de E, $\mu_0 P^{n_k}(\{i\}) \leq 1$. La limite μ de la sous-suite (μ_{n_k}) vérifie donc $\mu P = \mu$, c'est une probabilité invariante. □

L'étude qui suit précise la nature des valeurs propres de module 1 de P.

Rappelons tout d'abord que P et ${}^tP$ ont mêmes valeurs propres et que pour tout λ les noyaux de $P - \lambda I$ et de sa transposée ${}^tP - \lambda I$ ont même dimension.

Lemme 5.5 *Soit* X *une chaîne de Markov homogène irréductible finie (donc récurrente positive), apériodique, de matrice de transition* P. P *n'admet aucune valeur propre de module* 1 *autre que* 1. *Le sous-espace propre pour* $\lambda = 1$ *est de dimension* 1.

Remarque : le résultat du lemme peut être légèrement amélioré ; sous les mêmes hypothèses, il résulte du théorème de Perron-Frobenius que la valeur propre $\lambda = 1$ est racine simple du polynôme caractéristique (voir [6] théorème 1.1 p. 3).

Démonstration. Soit π la probabilité invariante de la chaîne[17] ; si v est un vecteur propre (complexe) de ${}^tP$ pour la valeur propre $\lambda = e^{i\theta}$ de module 1, $m = {}^tv$ vérifie $mP = e^{i\theta}m$ et $mP^n = e^{ni\theta}m$, soit pour la composante j :

$$(mP^n)_j = \sum_{i\in E} m_i P^n(i,j) \xrightarrow[n\to+\infty]{} \sum_{i\in E} m_i \pi_j = \pi_j \sum_{i\in E} m_i. \tag{7}$$

mais m a au moins une composante non nulle et pour $m_j \neq 0$, $(mP^n)_j = e^{ni\theta}m_j$ n'a de limite que si $\theta = 0 \pmod{2\pi}$, ce qui impose $\lambda = 1$. De plus, pour $\lambda = 1$, $m_j = (mP^n)_j$ et (7) implique que m est un vecteur ligne proportionnel à π, le sous-espace propre de ${}^tP$ pour $\lambda = 1$ est donc de dimension 1, celui de P aussi. □

Le lemme suivant précise la structure d'une classe récurrente périodique.

Lemme 5.6 *Soit* X *une chaîne de Markov homogène irréductible finie de matrice de transition* P *et de période* $d \geq 2$. *Alors la chaîne de matrice* P^d *n'est pas irréductible : elle possède* d *classes* $C_0, C_1, \ldots, C_{d-1}$ *telles que si* $X_k \in C_r$, *alors* $X_{k+1} \in C_{r+1}$ *(avec la convention* $C_d = C_0$*).*

Exemple : la promenade aléatoire à barrière réfléchissante sur $E = \{0, 1, \ldots, N\}$, est de période $d = 2$ et $C_0 = \{i \in E \mid i \text{ pair}\}$, $C_1 = \{i \in E \mid i \text{ impair}\}$.

17. On utilise ici le résultat du corollaire 4.9 page 20.

DÉMONSTRATION. Soit i un état fixé de E. Remarquons que si n et n' sont deux entiers vérifiant $P^n(i,j) > 0$ et $P^{n'}(i,j) > 0$, alors $n \equiv n' \pmod d$: en effet, comme i et j communiquent, il existe un entier m tel que $P^m(j,i) > 0$ donc $P^{n+m}(i,i) \geq P^n(i,j)P^m(j,i) > 0$ et de même $P^{n'+m}(i,i) > 0$, donc d divise à la fois $n+m$ et $n'+m$ donc aussi leur différence $n-n'$.

Posons $C_r = \{j \in E \mid P^n(i,j) > 0 \Rightarrow n \equiv r \pmod d\}$ $(0 \leq r < d)$ et montrons que $C_0, C_1, \ldots, C_{d-1}$ sont les classes de la chaîne de matrice P^d.

Montrons d'abord que deux états d'une même classe C_r communiquent pour P^d. Soient j et k deux états de C_r ; ils communiquent avec i donc

$$\begin{array}{llllll} \exists n_j & P^{n_j}(i,j) > 0 & \text{et} & \exists n_k & P^{n_k}(i,k) > 0 & \text{avec} \quad n_j \equiv n_k \equiv r \pmod d \\ \exists m_j & P^{m_j}(j,i) > 0 & \text{et} & \exists m_k & P^{m_k}(k,i) > 0 & \end{array}$$

d'où $P^{n_j+m_j}(i,i) > 0$ et $P^{n_k+m_k}(i,i) > 0$ donc par définition de $d = d(i)$, $n_j + m_j \equiv 0 \pmod d$ et $n_k + m_k \equiv 0 \pmod d$ et comme $n_j \equiv n_k \equiv r \pmod d$, $m_j \equiv m_k \equiv d - r \pmod d$. Mais on a aussi $P^{m_j+n_k}(j,k) \geq P^{m_j}(j,i)P^{n_k}(i,k) > 0$ et $m_j + n_k = (d-r) + r \equiv 0 \pmod d$ donc $m_j + n_k$ est un multiple de d. Par symétrie il en va de même de $m_m + n_j$, donc j et k communiquent en un nombre d'étapes multiple de d.

Soient maintenant $j \in C_r$ et $k \in C_s$ avec $r \not\equiv s \pmod d$. Montrons par l'absurde que j et k ne communiquent pas pour P^d : supposons qu'il existe $n \geq 1$ et $m \geq 1$ tels que $P^{nd}(j,k) > 0$ et $P^{md}(k,j) > 0$. Comme i communique avec j, il existe l tel que $P^l(i,j) > 0$ et comme $j \in C_r$, on a $l \equiv r \pmod d$, mais alors $P^{l+nd}(i,k) \geq P^l(i,j)P^{nd}(j,k) > 0$ donc $l \equiv l + nd \equiv s \pmod d$ ce qui est incompatible avec $l \equiv r \pmod d$.

Prenons maintenant j dans C_r et k dans C_s tels que $P(j,k) > 0$ et montrons que $s \equiv r+1 \pmod d$ ce qui terminera la démonstration. Comme j communique avec i, on a vu ci-dessus qu'il existe un entier l tel que $P^l(i,j) > 0$ et $l \equiv r \pmod d$. Mais $P^{l+1}(i,k) \geq P^l(i,j)P(j,k) > 0$ donc $l+1 \equiv s \pmod d$ et $s \equiv r+1 \pmod d$. □

Lemme 5.7 *Soit* X *une chaîne de Markov homogène irréductible finie, de matrice de transition* P *et de période* $d \geq 2$ *; les racines d-ièmes de l'unité sont valeurs propres de* P *et* P *n'a pas d'autre valeur propre de module* 1*. Le sous-espace propre pour chacune des racines d-ièmes de l'unité est de dimension* 1.

Remarque : le résultat du lemme peut être légèrement amélioré ; en fait chacune des racines d-ièmes de l'unité est racine simple du polynôme caractéristique (voir [6] théorème 1.15 p. 22).

DÉMONSTRATION. Chaque classe C_r étant irréductible pour P^d et finie, elle est récurrente positive et admet donc une unique probabilité $\pi^{(r)}$ invariante par P^d

(cf. corollaire 4.9). D'après le lemme précédent, $\pi^{(r)}\mathrm{P}$ est portée par la classe C_{r+1}, $\pi^{(r)}\mathrm{P}$ est également invariante par P^d et donc $\pi^{(r)}\mathrm{P} = \pi^{(r+1)}$ (unicité).

Posons $\lambda = \exp\frac{2i\pi}{d}$, en combinant les égalités $\pi^{(r)}\mathrm{P} = \pi^{(r+1)}$ on obtient

$$\Big(\sum_{r=0}^{d-1}\pi^{(r)}\Big)\mathrm{P} = \sum_{r=0}^{d-1}\pi^{(r)}\mathrm{P} = \sum_{r=0}^{d-1}\pi^{(r+1)} = \sum_{r=0}^{d-1}\pi^{(r)} \quad (\pi^d = \pi^0)$$

$$\forall k\ (1 \le k \le d-1),\ \Big(\sum_{r=0}^{d-1}\lambda^{kr}\pi^{(r)}\Big)\mathrm{P} = \sum_{r=0}^{d-1}\lambda^{kr}\pi^{(r)}\mathrm{P} = \sum_{r=0}^{d-1}\lambda^{kr}\pi^{(r+1)} = \lambda^{-k}\sum_{r=0}^{d-1}\lambda^{kr}\pi^{(r)}$$

ce qui montre que pour tout $0 \le k \le d-1$, le vecteur ligne $v_k = \sum_{r=0}^{d-1}\lambda^{kr}\pi^{(r)}$ est vecteur propre « à gauche » pour la valeur propre $\lambda^{-k} = \exp\frac{-2ik\pi}{d} = \exp\frac{-2i(d-k)\pi}{d}$ (son transposé est vecteur propre de ${}^t\mathrm{P}$). Les d racines d-ièmes de l'unité sont donc valeurs propres de ${}^t\mathrm{P}$, donc de P.

Réciproquement, soit λ une valeur propre de module 1 de P et v un vecteur propre de ${}^t\mathrm{P}$ associé à λ. Le vecteur ligne (complexe) $m = {}^tv$ vérifie $m\mathrm{P} = \lambda m$ et donc aussi $m\mathrm{P}^d = \lambda^d m$. Soit $m^{(r)}$ la trace de m sur la classe C_r : $m_i^{(r)} = m_i\,\mathbf{1}_{i\in\mathrm{C}_r}$. Comme $\mathrm{P}^d(i,j) = 0$ si i et j ne sont pas dans la même classe C_r (cf. prop. 3.9), $m^{(r)}$ vérifie $m^{(r)}\mathrm{P}^d = \lambda^d m^{(r)}$: pour tout état j de C_r

$$\lambda^d m_j^{(r)} = \lambda^d m_j = \big(m\mathrm{P}^d\big)_j = \sum_{i\in\mathrm{E}} m_i\mathrm{P}^d(i,j) = \sum_{i\in\mathrm{C}_r} m_i\mathrm{P}^d(i,j) = \sum_{i\in\mathrm{C}_r} m_i^{(r)}\mathrm{P}^d(i,j)$$

donc $m^{(r)}$ est un vecteur propre « à gauche » de P^d pour la valeur propre λ^d. Il est porté par la classe C_r qui est irréductible pour P^d d'après le lemme précédent et apériodique par définition de la période d. On applique le lemme 5.5 à la chaîne de matrice P^d sur classe C_r, 1 est la seule valeur propre de module 1, donc $\lambda^d = 1$ et λ est une racine d-ième de l'unité.

Enfin, d'après le lemme 5.7, $m^{(r)}$ vecteur propre « à gauche » de P^d pour la valeur propre 1, est proportionnel à $\pi^{(r)}$ unique probabilité invariante par P^d sur C_r. On en déduit que m s'écrit $m = \sum_{0\le r<d}\alpha_r\pi^{(r)}$ où les α_r sont des constantes complexes, mais l'équation $m\mathrm{P} = \lambda m$ et la relation $\pi^{(r)}\mathrm{P} = \pi^{(r+1)}$ imposent $\alpha_{r+1} = \lambda\alpha_r$ pour tout r, soit $\alpha_r = \lambda^r\alpha_0$, donc finalement pour tout λ racine d-ième de l'unité, il n'y a qu'une direction de vecteur propre « à gauche » : $m = \alpha_0\sum_{0\le r<d}\lambda^r\pi^{(r)}$. Pour chaque racine d-ième de l'unité le sous-espace propre associé pour ${}^t\mathrm{P}$ (ou P) est de dimension 1. □

Nous sommes en mesure d'énoncer le résultat final sur les valeurs propres de module 1 de P dans le cas général :

Proposition 5.8 *Soit* X *une chaîne de Markov homogène de matrice de transition* P *sur un espace d'états* E *fini.*

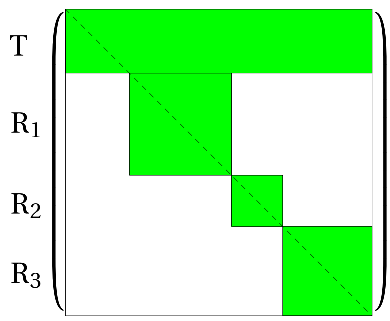
a) *La dimension du sous-espace propre associé à la valeur propre* 1 *pour* P *est égal au nombre de classes récurrentes de* X.

b) *Toute valeur propre de module* 1 *de* P *est racine* d*-ième de l'unité.*
Les racines d*-ièmes de l'unité sont valeurs propres de* P *si et seulement si* X *a au moins une classe récurrente de période* d*.*
La dimension du sous-espace propre associé à chaque racine d*-ième de l'unité (autre que* 1*) est précisément le nombre de classes récurrentes de période* d*.*

Remarque : compte tenu des remarques qui suivent les lemmes 5.5 et 5.7, l'énoncé ci-dessus est encore vrai si on remplace « dimension du sous-espace propre » par « ordre de la valeur propre ».

DÉMONSTRATION. Notons T l'ensemble des états transients et $R_1, \dots, R_k$ les classes récurrentes. T peut être vide, mais il y a toujours au moins une classe récurrente (prop. 3.10).

Il est toujours possible de renuméroter les états de façon à classer les états de T en premier, puis ceux de R_1 et ainsi de suite jusqu'à R_k. D'après la proposition 3.9, $P(i, j) = 0$ pour i récurrent et j n'appartenant pas à la classe de i, la matrice P présente donc des blocs de zéros représentés ci-contre dans le cas de 3 classes récurrentes : les zones coloriées représentent des termes positifs ou nuls, les zones claires ne contiennent que des termes nuls.



Sur chaque classe récurrente (finie), il existe une probabilité invariante (prop. 5.3), en la complétant par des 0 pour les états des autres classes on obtient encore une probabilité invariante par P (cf. la structure de la matrice P). Il est facile de voir que les vecteurs lignes ainsi construits sont linéairement indépendants, ce sont des vecteurs propres « à gauche » pour P et la valeur propre 1 (leurs transposés sont propres pour ${}^tP$ avec $\lambda = 1$), la dimension du sous-espace propre de ${}^tP$ pour la valeur propre 1 est donc au moins égale au nombre de classes récurrentes de X.

Pour établir l'inégalité inverse on considère un vecteur propre v quelconque de ${}^tP$ pour la valeur propre 1 ; le vecteur ligne $m = {}^tv$ est invariant par P (ce n'est pas une *mesure* invariante car ses composantes ne sont pas nécessairement positives).

Tout vecteur propre v de ${}^tP$ associé à une valeur propre λ de module 1 vérifie pour tout j de T $\quad v_j = 0$, en effet : posons $m = {}^tv$, si j est transient, pour tout i, $P^n(i, j) \to 0$ lorsque $n \to +\infty$ (cf. prop. 4.1), donc $\lambda^n m_j = (mP^n)_j = \sum_{i \in E} m_i P^n(i, j)$ tend vers 0, ce qui implique $m_j = 0$ et donc $v_j = 0$.

Soit R une classe récurrente fixée, il existe une unique probabilité $\pi^{(R)}$ sur R invariante par (la restriction de) P. Notons $m^{(R)}$ la trace de m sur R : $m_i^{(R)} = m_i \mathbf{1}_{i\in R}$. Comme $P^n(i,j) = 0$ si i est récurrent et j n'appartient pas à la classe de i (cf. prop. 3.9), on a

$$\begin{aligned}
\forall j \in R, \quad m_j^{(R)} = m_j = (mP^n)_j &= \sum_{i\in E} m_i P^n(i,j) \\
&= \sum_{i\in T} m_i P^n(i,j) + \sum_{i\in R} m_i P^n(i,j) \\
&= \sum_{i\in R} m_i P^n(i,j) \quad (m \text{ est nulle sur T}) \\
&\xrightarrow[n\to+\infty]{} \sum_{i\in R} m_i \pi_j^{(R)} \quad \text{si R est apériodique.}
\end{aligned}$$

Si R est une classe périodique, il suffit de considérer les moyennes de Césaro pour pouvoir passer à la limite (cf. prop. 4.19) :

$$\forall j \in R, \quad m_j^{(R)} = m_j = \frac{1}{N}\sum_{n=1}^{N} (mP^n)_j = \sum_{i\in R} m_i \frac{1}{N}\sum_{n=1}^{N} P^n(i,j) \xrightarrow[N\to+\infty]{} \sum_{i\in R} m_i \pi_j^{(R)}.$$

Donc, dans tous les cas, $m^{(R)}$ et $\pi^{(R)}$ sont proportionnels :

$$\forall j \in R, \quad m_j^{(R)} = \sum_{i\in R} m_i \pi_j^{(R)} = \pi_j^{(R)} \sum_{i\in R} m_i.$$

Comme m est nulle sur T, m est la somme sur toutes les classes récurrentes de vecteurs de la forme $\alpha_R \pi^{(R)}$, la dimension du sous-espace propre associé à la valeur propre 1 pour ${}^tP$ est donc au plus égale au nombre de classes récurrentes de X. Le point a) est établi.

Compte tenu de la forme de la matrice P (voir dessin ci-dessus) le polynôme caractéristique de P peut se calculer par blocs : c'est le produit des déterminants des blocs diagonaux correspondants à $T \times T$, $R_1 \times R_1$, ..., $R_k \times R_k$. Le bloc $T \times T$ ne fournit aucune valeur propre de module 1 (on a vu ci-dessus que tout vecteur propre v pour λ de module 1 est nul sur T), le résultat b) s'obtient en appliquant le lemme 5.7 aux différentes classes récurrentes. □

Exercices

Exercice 1. Ruine du joueur sur N états

La fortune initiale du joueur A est k $(0 \leq k \leq \mathrm{N})$, celle de son adversaire est $\mathrm{N}-k$. À chaque partie le joueur A prend un € à son adversaire avec une probabilité p ou lui donne un € avec une probabilité q, la probabilité d'une partie nulle étant r $(p+q+r=1)$. Le jeu s'arrête dès que l'un des joueurs est ruiné.

La fortune X_n du joueur A après la n-ième partie est une chaîne de Markov de matrice

$$\mathrm{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ q & r & p & 0 & \cdots & 0 \\ 0 & q & r & p & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & q & r & p \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}$$

1) Montrer que la probabilité $u_\mathrm{N}(k)$ de ruine du joueur A vérifie :

$$u_\mathrm{N}(0) = 1 \quad u_\mathrm{N}(\mathrm{N}) = 0$$

$$\forall k \in \{1,2,\ldots,\mathrm{N}-1\} \quad (p+q)u_\mathrm{N}(k) = p u_\mathrm{N}(k+1) + q u_\mathrm{N}(k-1)$$

et résoudre ce système[18].

2) Quelle est la probabilité $v_\mathrm{N}(k)$ que le joueur A ruine son adversaire ? Montrer que la durée du jeu est presque sûrement finie.

Éléments de réponse : 1) $u_\mathrm{N}(k) = 1 - k/\mathrm{N}$ si $p=q$,
$u_\mathrm{N}(k) = \frac{a^k - a^\mathrm{N}}{1-a^\mathrm{N}}$ avec $a = q/p$ si $p \neq q$.

Exercice 2.

Un joueur possède 10 € et veut essayer d'en gagner 10 de plus en jouant à la roulette. Il envisage deux stratégies :

a) Il mise ses 10 € en une seule fois sur rouge ou sur noir.

b) Il joue 1 € à la fois soit sur rouge soit sur noir et persévère jusqu'à ce que sa fortune atteigne 20 € (à moins qu'il ne soit ruiné avant ...).

18. On pourra, soit considérer la suite auxiliaire $u'_\mathrm{N}(k) = u_\mathrm{N}(k) - u_\mathrm{N}(k-1)$, soit utiliser les techniques classiques de calcul des suites récurrentes $w_{k+1} = a w_k + b w_{k-1}$ (a, b constantes données).

Calculer pour ces deux stratégies la probabilité qu'a le joueur de sortir du jeu avec 20 € en poche, sachant qu'à la roulette la probabilité que le rouge (ou le noir) sorte est $\frac{18}{37}$ et que le joueur double sa mise s'il a joué la bonne couleur et qu'il la perd sinon. Commenter le résultat.

Éléments de réponse : a) $\mathrm{P}_a \approx 0.486$. b) $\mathrm{P}_b \approx 0.368$.

Exercice 3. Promenade aléatoire sur **Z**

On considère la promenade aléatoire $\mathrm{S}_n = \mathrm{S}_0 + \sum_{i=1}^n \mathrm{Y}_i$ où les Y_i sont des variables aléatoires indépendantes valant $+1$, -1, ou 0 avec les probabilités respectives p, q, r, telles que $p+q+r=1$ et S_0 une variable aléatoire entière indépendante des $(\mathrm{Y}_i)_{i\geq 1}$. $(\mathrm{S}_n)_{n\in\mathbf{N}}$ est une chaîne de Markov sur **Z** dont la matrice de transition est donnée par

$$\forall i \in \mathbf{Z}, \quad \mathrm{P}(i, i+1) = p \quad \mathrm{P}(i, i-1) = q \quad \mathrm{P}(i,i) = r$$

Pour tout k de **Z** on note $u(k)$ la probabilité de passage par l'état 0 à partir de l'état k :

$$u(k) = \mathbf{P}_k(\exists n \geq 0, \mathrm{S}_n = 0) = \mathbf{P}(\exists n \geq 0, \mathrm{S}_n = 0 \mid \mathrm{S}_0 = k).$$

1) Montrer que la suite $u(k)$ vérifie :

$$\forall k \in \mathbf{Z}^*, \quad (p+q)u(k) = pu(k+1) + qu(k-1) \quad \text{et} \quad u(0) = 1$$

et que

a) si $p = q \quad \forall k \neq 0 \quad u(k) = 1$,

b) si $p > q \quad \forall k < 0 \quad u(k) = 1$,

c) si $p > q \quad \forall k > 0 \quad u(k) = (q/p)^k$: comparer les événements $\mathrm{A}(k)$, passage par l'état 0 de la promenade aléatoire qui part de l'état k et $\mathrm{A}_\mathrm{N}(k)$ ruine du joueur de l'exercice 1 et remarquer que $\mathrm{A}(k)$ est la limite croissante en N des $\mathrm{A}_\mathrm{N}(k)$.

2) En déduire que la probabilité partant de 0 d'y revenir, $\mathbf{P}_0(\exists n \geq 1, \mathrm{S}_n = 0)$, vaut dans tous les cas $1 - |p-q|$. Autrement dit, la promenade aléatoire S_n est récurrente si $p = q$ et transiente si $p \neq q$.

3) En utilisant les résultats de l'exercice 1, montrer que la promenade aléatoire S_n symétrique $(p = q)$ sort $\mathbf{P}_0$-presque sûrement de tout compact :

$$\forall \mathrm{K} \in \mathbf{N}, \quad \mathbf{P}_0(\exists n \geq 1, |\mathrm{S}_n| \geq \mathrm{K}) = 1.$$

Exercice 4. Principe du miroir

On considère la promenade aléatoire symétrique définie par $S_0 = 0$ et $S_n = \sum_{i=1}^{n} Y_i$ pour $n \geq 1$, où les Y_i sont des variables aléatoires indépendantes de loi :

$$\mathbf{P}(Y_i = +1) = \frac{1}{2}, \quad \mathbf{P}(Y_i = -1) = \frac{1}{2}.$$

1) On se propose de calculer la loi de la v.a. ν_{-5}, premier instant où $S_n = -5$.
   - Montrer que cette loi ne charge que les entiers impairs supérieurs ou égaux à 5 et calculer $\mathbf{P}(\nu_{-5} = 5)$.
   - Pour évaluer $\mathbf{P}(\nu_{-5} = 2k+1)$ avec $k \geq 3$ on remarque d'abord que sur l'événement $\{\nu_{-5} = 2k+1\}$, $S_{2k} = -4$. On calcule alors le nombre total N_k de trajectoires vérifiant $S_0 = 0$ et $S_{2k} = -4$ en considérant les nombres M et D de « montées » et « descentes » de S_n, c'est à dire les nombres de fois où $Y_i = +1$ (S_n « monte ») et où $Y_i = -1$ (S_n « descend »).
   - On calcule ensuite le nombre N'_k de trajectoires vérifiant $S_0 = 0$, $S_{2k} = -4$ et ayant atteint l'état -5 avant l'instant $2k$ en appliquant le principe du miroir au premier instant n où $S_n = -5$.
   - En déduire la loi de ν_{-5} :

$$\forall k \geq 2, \quad \mathbf{P}(\nu_{-5} = 2k+1) = \frac{5(2k)!}{(k-2)!(k+3)!}\left(\frac{1}{2}\right)^{2k+1}$$

2) Calculer de manière analogue la loi de la v.a. ν_{10}, premier instant où $S_n = 10$:

$$\forall k \geq 5, \quad \mathbf{P}(\nu_{10} = 2k) = \frac{10(2k-1)!}{(k+5)!(k-5)!}\left(\frac{1}{2}\right)^{2k}$$

3) Calculer la loi de la v.a. $\nu = \inf\{n > 0 \mid S_n = 10 \text{ ou } S_n = -5\}$, premier instant où la promenade atteint soit 10 soit -5, c'est-à-dire la loi de la durée du jeu lorsque les fortunes initiales des joueurs sont 5 et 10.

Exercice 5. Promenades aléatoires symétriques sur $\mathbf{Z}^2$ et $\mathbf{Z}^3$

1) Une particule part de l'origine O du plan $\mathbf{Z}^2$ et se déplace de la façon suivante : partant à l'instant n du point (x, y), la particule saute à l'instant $n+1$ à l'un des quatre points « voisins » $(x+1, y+1)$, $(x+1, y-1)$, $(x-1, y+1)$, $(x-1, y-1)$ avec la probabilité $\frac{1}{4}$.

   M_n désignant la position de la particule à l'instant n, on cherche à calculer les probabilités $\mathbf{P}_O(M_n = O) = \mathbf{P}(M_n = O \mid M_0 = O)$ pour tout $n \geq 0$. Ramener ce problème à l'étude de deux promenades aléatoires symétriques sur $\mathbf{Z}$ et montrer que

$$\forall n > 0, \quad \mathbf{P}_O(M_{2n} = O) = \left(C_{2n}^{n}\frac{1}{2^{2n}}\right)^2$$

   En déduire que la série $\sum_{n \geq 0} \mathbf{P}_O(M_n = O)$ diverge.

2) On reprend le problème dans $\mathbf{Z}^3$: la particule part de l'origine et si à un instant donné elle est au point (x, y, z), à l'instant suivant elle saute à l'un des huit points « voisins » (x', y', z') où $x' = x \pm 1$, $y' = y \pm 1$, $z' = z \pm 1$ avec la probabilité $\frac{1}{8}$.
Montrer que

$$\forall n > 0, \quad \mathbf{P}_{\mathrm{O}}(\mathrm{M}_{2n} = \mathrm{O}) = \left(\mathrm{C}_{2n}^{n} \frac{1}{2^{2n}}\right)^3$$

En déduire que la série $\sum_{n \geq 0} \mathbf{P}_{\mathrm{O}}(\mathrm{M}_n = \mathrm{O})$ converge.
Conclusion : Les promenades aléatoires symétriques sur $\mathbf{Z}$ et $\mathbf{Z}^2$ sont récurrentes, la promenade aléatoire sur $\mathbf{Z}^3$ est transiente.

Exercice 6. Séries de succès

Soit $(p_k)_{k \in \mathbf{N}}$ une suite de réels de $]0,1[$ et soit (X_n) la chaîne de Markov sur $\mathbf{N}$ de matrice de transition

$$\mathrm{P} = \begin{pmatrix} q_0 & p_0 & 0 & 0 & \cdots & 0 & \cdots \\ q_1 & 0 & p_1 & 0 & \cdots & 0 & \cdots \\ q_2 & 0 & 0 & p_2 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 & \cdots \\ q_n & 0 & 0 & 0 & \cdots & p_n & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

1) Montrer que (X_n) est une chaîne irréductible apériodique.
2) Soit τ_j le temps d'atteinte de l'état j. Montrer que

$$\forall i \in \mathbf{N},\ \forall k \geq 1, \quad \mathbf{P}_i(\tau_0 > k) = \prod_{j=0}^{k-1} p_{i+j}.$$

En déduire la loi de τ_0 pour la probabilité $\mathbf{P}_i$ et une condition nécessaire et suffisante portant sur $\prod_{j=0}^{\infty} p_j$ pour que la chaîne (X_n) soit récurrente.
Dans toute la suite du problème on notera $\forall k \geq 1$, $\beta_k = \prod_{n=0}^{k-1} p_n$, $\beta_\infty = \prod_{n=0}^{\infty} p_n$ et $\beta_0 = 1$.
3) Existe-t-il des mesures réversibles non nulles ? Donner une condition nécessaire et suffisante portant sur les (β_k) pour qu'il existe une probabilité invariante et la calculer.
4) Étude de la loi du temps d'atteinte de l'état j partant de l'état i, pour $0 \leq j \leq i$. En remarquant que l'état j ne peut être atteint à partir de i qu'en passant par 0, montrer que $\mathbf{P}_i(\tau_j = k)$ est nulle si $k \leq j$ et que

$$\forall k > j, \quad \mathbf{P}_i(\tau_j = k) = \sum_{n=1}^{k} \mathbf{P}_i(\tau_0 = n)\ \mathbf{P}_0(\tau_j = k - n).$$

5) Étude de la loi du temps d'atteinte de i partant de 0. Montrer que si $i > 0$

$$\mathbf{P}_0(\tau_i = i) = \beta_i \quad \text{et} \quad \forall k > i, \quad \mathbf{P}_0(\tau_i = k) = \sum_{n=1}^{i} \mathbf{P}_0(\tau_0 = n)\, \mathbf{P}_0(\tau_i = k - n).$$

6) Dans cette question on suppose la chaîne (X_n) transiente et on se propose de calculer les probabilités $f_{ij} = \mathbf{P}_i(\tau_j < \infty)$.

a) Montrer que si $i < j$, $f_{ij} = 1$: pour cela considérer les v.a. $N_k = \sum_{n=0}^{\infty} 1_{X_n = k}$ (nombre de visites à l'état k) et montrer que $\sum_{k=0}^{j-1} N_k = +\infty$ $\mathbf{P}_i$-p.s. sur $\{\tau_j = +\infty\}$; conclure en utilisant la transience de la chaîne.

b) Cas $i \geq j > 0$: montrer en utilisant 4) que si $i \geq j > 0$, $f_{ij} = f_{i0}\, f_{0j} = 1 - \frac{\beta_\infty}{\beta_i}$.

c) En déduire les valeurs de $U(i,j) = \sum_{n=0}^{\infty} P^n(i,j)$.

7) Dans cette question on suppose la chaîne (X_n) récurrente et on se propose de calculer les $m(i,j) = \mathbf{E}_i(\tau_j)$.

a) Montrer en utilisant 2) que pour $i \geq 0$, $m(i,0) = \sum_{n\in\mathbf{N}} \frac{\beta_{i+n}}{\beta_i}$.

b) Montrer en utilisant 5) que pour $i > 0$

$$m(0,i) = i\,\beta_i + \sum_{n=1}^{i} n\,\beta_{n-1}\, q_{n-1} + m(0,i) \sum_{n=1}^{i} \beta_{n-1}\, q_{n-1}$$

Montrer que $\sum_{n=1}^{i} \beta_{n-1}\, q_{n-1} = 1 - \beta_i$ et que $\sum_{n=1}^{i} n\,\beta_{n-1}\, q_{n-1} = \sum_{n=0}^{i-1} \beta_n - i\,\beta_i$ et en déduire que $m(0,i) = \frac{1}{\beta_i} \sum_{n=0}^{i-1} \beta_n$.

c) Justifier les égalités :

$$\forall (i,j)\ i \geq j > 0, \quad m(i,j) = m(i,0) + m(0,j)$$

$$\forall (i,j)\ 0 < i < j, \quad m(0,j) = m(0,i) + m(i,j)$$

d) Calculer $m(i,j)$. Retrouver $m(i,i)$ à partir de la probabilité invariante.

Exercice 7. Probabilités d'absorption (capture par les classes récurrentes)

Soit $(X_n)_{n\geq 0}$ une chaîne de Markov sur un espace d'états E fini, admettant un ensemble non vide T d'états transients et un ensemble non vide R d'états récurrents (formés d'une ou plusieurs classes). On note ν le temps d'atteinte de l'ensemble R par la chaîne X_n :

$$\nu(\omega) = \inf\{n \geq 1 \mid X_n(\omega) \in R\} \quad \text{ou} \quad +\infty$$

1) Montrer que pour tout état i transient, la variable ν est $\mathbf{P}_i$-presque sûrement finie et même $\mathbf{P}_i$-intégrable.

2) On pose $m_i = \mathbf{E}_i(\nu)$. Montrer que les m_i sont solutions du système linéaire :

$$\forall i \in \mathrm{T}, \quad m_i = 1 + \sum_{j \in \mathrm{T}} \mathrm{P}(i,j) m_j$$

Indication : $m_i = \mathbf{E}_i(\nu) = \sum_{n \in \mathbf{N}} \mathbf{P}_i(\nu \geq n+1)$.

3) On suppose dans cette question que l'ensemble R est constitué de r classes de récurrence $(\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_r)$, avec $r \geq 2$. On note μ_{ik} la probabilité d'atteinte de la classe C_k à partir de l'état i :

$$\mu_{ik} = \mathbf{P}_i(\exists n \geq 1 \;\; \mathrm{X}_n \in \mathrm{C}_k) = \mathbf{P}_i(\mathrm{X}_\nu \in \mathrm{C}_k)$$

Montrer que les μ_{ik} sont solutions du système linéaire :

$$\forall i \in \mathrm{T}, \quad \mu_{ik} = \sum_{j \in \mathrm{C}_k} \mathrm{P}(i,j) + \sum_{j \in \mathrm{T}} \mathrm{P}(i,j) \mu_{jk}$$

4) Que peut-on dire de la limite lorsque n tend vers l'infini de $\mathrm{P}^n(i,j)$ où i est transient et j appartient à la classe récurrente C_k ?
Indication : remarquer que $\lim_n \mathrm{P}^n(i,j) = \lim_n \mathbf{P}_i(\nu < \infty \cap \mathrm{X}_\nu \in \mathrm{C}_k \cap \mathrm{X}_{\nu+n} = j)$.

5) Retrouver les résultats de l'exercice 1 sur la ruine du joueur et préciser la durée moyenne du jeu.

Éléments de réponse : 5)
$$\mathbf{E}_i(\nu) = \frac{i\,(\mathrm{N}-i)}{p+q} \quad \text{si } p = q.$$
$$\mathbf{E}_i(\nu) = \frac{1}{q-p}\left(i - \mathrm{N}\,\frac{1-(q/p)^i}{1-(q/p)^{\mathrm{N}}}\right) \quad \text{si } p \neq q.$$

Exercice 8. Délais et probabilités d'atteinte

Soit $(\mathrm{X}_n)_{n \geq 0}$ une chaîne de Markov irréductible sur un espace d'états fini. On s'intéresse aux quantités
– m_{ij} = temps moyen mis par la chaîne pour passer de l'état i à l'état j et
– π_{ij}^k = probabilité d'atteindre l'état j *avant* l'état k à partir de l'état i.

1) Montrer qu'on peut résoudre ce problème grâce aux résultats de l'exercice précédent, à condition de modifier convenablement la matrice de transition de la chaîne.

2) Application à la promenade du scarabée
Un scarabée de déplace le long des arêtes d'un tétraèdre régulier, en changeant de sommet à chaque unité de temps et en choisissant sa destination au hasard parmi l'un des trois autres sommets (avec la même probabilité $\frac{1}{3}$).

   a) Calculer le temps moyen mis par le scarabée pour atteindre le sommet S_2 à partir de S_1.

b) Quelle est la probabilité d'atteindre S_2 avant S_4 en partant de S_1 ?

c) Calculer la loi du temps d'atteinte de S_2 à partir de S_1 et retrouver le résultat du a).

Éléments de réponse : 2) a) Temps moyen = 3. 2) b) Proba = ½.
2) c) Loi géométrique : $\forall k \geq 1, \mathbf{P}(\nu = k) = (2/3)^{k-1}(1/3)$, on retrouve $\mathbf{E}(\nu) = 3$.

Exercice 9. Étude d'une file d'attente

On considère une station service disposant d'une seule pompe et de deux places de parking pour véhicules en attente de service. On discrétise le problème en faisant les hypothèses suivantes :

– les clients n'arrivent qu'aux instants entiers, en nombre Y_n à l'instant n et les variables $(Y_n)_{n\geq 0}$ sont indépendantes et de même loi, donnée par :

$$\mathbf{P}(Y_n = 0) = 0{,}4 \quad \mathbf{P}(Y_n = 1) = 0{,}4 \quad \mathbf{P}(Y_n = 2) = 0{,}2$$

– les services commencent aux instants entiers et durent une unité de temps.
– tout véhicule trouvant à son arrivée les deux places de parking occupées renonce à attendre et cherche une autre station.

1) Montrer que le nombre X_n de véhicules présents, en attente ou en service, juste après l'instant n est une chaîne de Markov et préciser sa matrice de transition.

2) Quelle est la probabilité en régime stationnaire de l'événement « aucun véhicule n'est présent à la station » ?

3) Si aucun client n'est présent à l'instant 0, quel laps de temps moyen faudra-t-il attendre jusqu'à la saturation de la station (premier instant où les deux places de stationnement sont occupées) ?

Éléments de réponse 1) $X_{n+1} = \min\big((X_n - 1)^+ + Y_{n+1}, 3\big)$. 2) $\pi(0) = 8/35$.
3) $m_0 = 20$ (cf. exercices 7 et 8).

Exercice 10. Définition et propriétés des fonctions génératrices

Soit X une variable aléatoire à valeurs dans $\mathbf{N}$. On pose pour $s \in \mathbf{C}$

$$g_X(s) = E\big(s^X\big) = \sum_{k\in\mathbf{N}} s^k\, \mathbf{P}(X = k).$$

1) Quel est le domaine de définition de g_X ? On appelle fonction génératrice de X la restriction de g_X à l'intervalle $[0, 1]$.

2) Montrer que g_X est indéfiniment dérivable sur $[0,1[$, que toutes ses dérivées sont croissantes sur $[0,1[$ et que

$$\forall k \in \mathbf{N}, \quad g_X^{(k)}(0) = k!\, \mathbf{P}(X = k)$$

En déduire que la fonction génératrice d'une variable aléatoire entière caractérise entièrement sa loi, autrement dit l'application $\mathbf{P}_X \longrightarrow g_X$ est injective.

3) Montrer que toute série entière $S(x) = \sum_{k=0}^{\infty} a_k x^k$ de rayon de convergence 1, dont tous les a_k sont *positifs* vérifie :
$S(x) \xrightarrow[x \to 1_-]{} \sum_{k=0}^{\infty} a_k$ si $\sum_{k=0}^{\infty} a_k < \infty$,
$S(x) \xrightarrow[x \to 1_-]{} +\infty$ si $\sum_{k=0}^{\infty} a_k = \infty$.
En déduire que g est continue à gauche au point 1 et que

$$\mathbf{E}(X) = \lim_{s \nearrow 1} \uparrow g'_X(s) \quad \text{et} \quad \mathbf{E}(X(X-1)) = \lim_{s \nearrow 1} \uparrow g''_X(s) \quad (\text{égalités dans } \overline{\mathbf{R}}).$$

Application : calculer la fonction génératrice d'une variable aléatoire de Poisson de paramètre λ et en déduire son espérance et sa variance. Même question pour une variable aléatoire X binomiale négative de paramètres r et a (rappel : une telle variable aléatoire est à valeurs dans $\{n \in \mathbf{N} \mid n \geq r\}$ et sa loi est donnée par $\mathbf{P}(X = r + k) = C_{r+k-1}^{r-1} a^r (1-a)^k$ pour tout $k \geq 0$; pour $r = 1$ on retrouve la loi géométrique).

4) Calculer la fonction génératrice $g_{X_1+X_2}$ de la somme de deux variables aléatoires indépendantes X_1 et X_2 à valeurs dans $\mathbf{N}$ en fonction de g_{X_1} et g_{X_2}.
Application : Quelle est la loi d'une somme de deux variables aléatoires indépendantes de loi de Poisson de paramètres λ_1 et λ_2. Quelle est la loi d'une somme de n variables aléatoires indépendantes de Poisson de même paramètre λ ?

5) Soit $(X_i)_{i \in \mathbf{N}^*}$ une suite de variables aléatoires indépendantes de même loi à valeurs dans $\mathbf{N}$ et ν une variable aléatoire à valeurs dans $\mathbf{N}^*$ indépendante des (X_i). Montrer que l'expression

$$S(\omega) = \begin{cases} \sum_{i=1}^{\nu(\omega)} X_i(\omega) & \text{si } \nu(\omega) \geq 1 \\ 0 & \text{si } \nu(\omega) = 0 \end{cases}$$

définit une variable aléatoire entière. On note g la fonction génératrice commune des X_i, g_ν celle de ν et g_S celle de S.
Montrer que $\forall s \in [0,1] \quad g_S(s) = g_\nu \circ g\,(s)$.
En déduire l'espérance et la variance de S en fonction de celles de ν et X_i.

Applications :

a) Déterminer la loi de S lorsque ν est une variable aléatoire de Poisson de paramètre λ et les X_i sont de Bernoulli : $\mathbf{P}(X_i = 1) = p \quad \mathbf{P}(X_i = 0) = 1 - p$.

b) Le nombre d'accidents de la route se produisant par semaine est une variable aléatoire de moyenne M et d'écart-type Σ. Les nombres de blessés lors de chaque accident ont des distributions indépendantes, chacune de moyenne m et d'écart-type σ. Calculer la moyenne et l'écart-type du nombre de blessés de la route par semaine.

Éléments de réponse

1) Le domaine de définition contient au moins le disque fermé de rayon 1, mais peut être $\mathbf{C}$ tout entier en particulier si X est bornée : dans ce cas g_X est un polynôme.

3) Si $X \sim \mathscr{P}(\lambda)$, $g_X(s) = \exp\lambda(s-1)$, $\mathbf{E}(X) = \lambda$, $\mathbf{Var}(X) = \lambda$.
Si X binomiale négative, $g_X(s) = \left(\frac{sa}{1-s(1-a)}\right)^r$, $\mathbf{E}(X) = \frac{r}{a}$, $\mathbf{Var}(X) = r\,\frac{1-a}{a^2}$.

4) $g_{X_1+X_2}(s) = g_{X_1}(s)\,g_{X_2}(s)$; $X_1 + X_2 \sim \mathscr{P}(\lambda_1 + \lambda_2)$.

5) $\mathbf{E}(S) = \mathbf{E}(X_1)\,\mathbf{E}(\nu)$; $\mathbf{Var}(S) = (\mathbf{E}(X_1))^2\,\mathbf{Var}(\nu) + \mathbf{E}(\nu)\,\mathbf{Var}(X_1)$.

a) S suit une loi de Poisson de paramètre λp.

Exercice 11. Chaînes de Galton-Watson

Soit $(Y_{n,k})_{(n,k)\in\mathbf{N}^2}$ une suite de v.a. à valeurs dans $\mathbf{N}$, indépendantes et de même loi : $Y_{n,k}$ représente le nombre de descendants du k-ième individu de la génération n.

La taille des générations successives est déterminée par celle de la population initiale (variable entière positive X_0) et la relation de récurrence :

$$\forall n \geq 0, \quad X_{n+1} = \sum_{k=1}^{X_n} Y_{n,k}.$$

On pose $Y = Y_{1,1}$, $m = \mathbf{E}(Y)$ et note $g(s) = \sum_{k=0}^{\infty} s^k\,\mathbf{P}(Y = k)$ la fonction génératrice commune des $Y_{n,k}$ et g_n celle de X_n.

1) En utilisant les résultats de l'exercice précédent,

a) donner une expression de g_{n+1} en fonction de g_n et g,

b) montrer que $\forall n \geq 1, \quad \mathbf{E}(X_n) = m^n\,\mathbf{E}(X_0)$ et en déduire que si $m < 1$ la chaîne est presque sûrement absorbée par l'état 0.

2) Soit $u_n = \mathbf{P}_1(X_n = 0) = g_n(0)$ la probabilité que la chaîne se trouve à l'état 0 à l'instant n. On suppose dans cette question $X_0 = 1$ et $\mathbf{P}(Y = 0) > 0$.

   Montrer que $\forall n \geq 1,\ \forall s \in [0,1], \quad g_{n+1}(s) = g \circ g_n(s)$ et en déduire que u_n vérifie l'équation $u_{n+1} = g(u_n)$. Montrer que g est continue, croissante et convexe sur $[0,1]$ et étudier la convergence de la suite u_n dans les deux cas $m \leq 1$ et $m > 1$.

   Soit ν l'instant de premier passage par 0 de la chaîne (X_n). En remarquant que $\{X_n = 0\} = \{\nu \leq n\}$, montrer que
   - si $m \leq 1, \quad \mathbf{P}_1(\nu < +\infty) = 1$,
   - si $m > 1, \quad \mathbf{P}_1(\nu < +\infty) = \alpha$ $(0 < \alpha < 1)$ et $\mathbf{P}_1(X_n \to +\infty) = 1 - \alpha$ (remarquer que tous les états autres que 0 sont transients).

3) Montrer que la suite $\left(\frac{X_n}{m^n}\right)$ est une martingale qui converge lorsque $n \to +\infty$.

4) *Applications* :

   a) Des particules se désintègrent en donnant naissance à 0, 1, ou 2 particules identiques avec les probabilités $p_0 > 0, p_1 > 0, p_2 > 0$ $(p_0 + p_1 + p_2 = 1)$. Exprimer en fonction de p_0, p_1, p_2 la probabilité que les « descendants » d'une même particule disparaissent tous.

   b) Cent jeunes couples aux noms tous différents colonisent une île déserte. Chaque couple procrée jusqu'à avoir au moins un garçon et une fille ou au plus trois enfants. Les probabilités de naissance d'un garçon ou d'une fille sont supposées égales. Quelle probabilité ont chacun des noms de famille initiaux de disparaître ?

Éléments de réponse

4) *Applications* :

a) Si $p_2 \leq p_0$ (cas $m \leq 1$), la probabilité d'extinction est 1.
   Si $p_2 > p_0$ (cas $m > 1$), $\alpha = \frac{p_0}{p_2}$.

b) L'espérance de la loi du nombre de garçons mis au monde par un couple est $m = \frac{5}{4}$ et $\alpha = \sqrt{2} - 1$.

Références

[1] Ph. Barbe & M. Ledoux, *Probabilité*, Belin (1998), ISBN 2-7011-2131-0.

[2] K. L. Chung, *Markov Chains with stationary transition probabilities*, Springer (1967).

[3] S. Karlin, *Initiation aux processus stochastiques*, Dunod (1968).

[4] J. Neveu, *Initiation aux processus stochastiques*, Polycopié de cours de maîtrise/DEA (1984).

[5] J. R. Norris, *Markov Chains*, Cambridge Academic Press (1997).

[6] E. Seneta, *Non-negative Matrices and Markov Chains*, Springer Verlag (1981).

[7] P. S. Toulouse, *Thèmes de probabilités et statistique*, Dunod (1999) ISBN 2-10-0004308-0.