

PROCESSUS STOCHASTIQUES

DAVID COUPIER

Table des matières

Introduction	4
Exercices	8
1 Chaînes de Markov	9
1.1 Qu'est-ce qu'une chaîne de Markov ?	9
1.1.1 Définitions et exemples	9
1.1.2 Loi d'une chaîne de Markov	12
1.2 Classification des états	16
1.2.1 Classes irréductibles	16
1.2.2 Récurrence et transience	17
1.2.3 Périodicité	21
1.2.4 Simulation	22
1.3 Théorèmes limites	23
1.3.1 Les mesures stationnaires	23
1.3.2 Les mesures réversibles	24
1.3.3 Lorsque l'espace d'états E est fini	24
1.3.4 Lorsque l'espace d'états E est infini	26
1.4 Exercices	28
2 Processus stochastiques markoviens en temps continu	31
2.1 Processus de Poisson	31
2.1.1 Trois propriétés de la loi exponentielle	31
2.1.2 Présentation du processus de Poisson	33
2.1.3 Simulation et modélisation	36
2.2 Processus markoviens de sauts	38
2.2.1 Définition et exemples	38
2.2.2 Dé-randomisation du temps et simulation	39
2.2.3 Probabilités de transition et générateur de Markov	40
2.2.4 Théorème limite	44
2.3 Application aux modèles de files d'attente	47
2.3.1 Présentation générale	47
2.3.2 Étude du cas $M/M/1$	47
2.4 Exercices	51

3 Méthodes MCMC	55
3.1 Méthodes de Monte Carlo	55
3.2 Méthodes MCMC	56
3.3 Exemple des mesures de Gibbs	59
Annexe : quelques formules	60
Bibliographie	61

Introduction

On peut définir un *processus stochastique* comme étant une famille $\{X_t\}_{t \in T}$ de variables aléatoires indexées par le temps t . Les mots *processus* et *stochastique* signifient respectivement *fonction* et *aléatoire*. Alors qu’une variable aléatoire X associe à chaque $\omega \in \Omega$ une réalisation $X(\omega)$, un processus stochastique $\{X_t\}_{t \in T}$ associe à chaque ω une fonction (ou trajectoire) $\{X_t(\omega)\}_{t \in T}$:

$$\begin{array}{ccc} T & \rightarrow & E \\ t & \mapsto & X_t(\omega) \end{array} ,$$

où E est l’espace d’arrivée des variables aléatoires X_t . Passer des variables aléatoires aux processus stochastiques revient à passer en analyse des points aux fonctions. À titre d’exemple, la trajectoire d’une mouche en fonction du temps peut être modélisée par un processus stochastique à valeurs dans $E = \mathbb{R}^3$. Lorsque l’ensemble des temps T est au plus dénombrable (par exemple $T = \mathbb{N}$), on parle de processus stochastiques à *temps discret*. Lorsqu’il est continu (i.e. $T = [0; t_0]$ ou $T = \mathbb{R}_+$), on parle de processus stochastiques à *temps continu*.

Dans tout ce cours, on abrège les expressions “variable aléatoire” en v.a. et “indépendantes et identiquement distribuées” en i.i.d.

Les situations réelles pouvant être modélisées par des processus stochastiques sont nombreuses. En voici quelques exemples :

EXEMPLES :

• **Problème de la ruine du joueur.** Considérons une suite de v.a. $(Y_n)_{n \geq 1}$ i.i.d. dont la loi commune est définie par

$$\mathbb{P}(Y_1 = 1) = p \text{ et } \mathbb{P}(Y_1 = -1) = 1 - p$$

et une quantité initiale (déterministe ou aléatoire) $Y_0 \in \mathbb{Z}$ indépendante des v.a. Y_n . On définit la *marche aléatoire simple* par

$$\begin{aligned} X_{n+1} &= X_n + Y_{n+1} \\ &= Y_0 + Y_1 + \dots + Y_n + Y_{n+1} , \end{aligned}$$

pour tout $n \in \mathbb{N}$. La suite $(X_n)_{n \geq 1}$ est un processus stochastique à temps discret $T = \mathbb{N}$ (ce sont les instants $n = 0, 1, 2, \dots$) et à valeurs dans $E = \mathbb{Z}$. La suite $(X_n)_{n \geq 1}$ représente l’évolution de la fortune d’un joueur (jouant à pile ou face, à la roulette...) gagnant un montant fixe (ici 1 euro) avec probabilité p et perdant le même montant avec probabilité $1 - p$. Les parties, dont les résultats sont les Y_n , sont supposées indépendantes. La fortune du joueur à l’issue de la n^e partie est X_n . La quantité Y_0 représente la fortune initiale du joueur.

Le joueur est ruiné (et donc arrête de jouer) dès que la suite $(X_n)_{n \geq 1}$ touche l’axe des abscisses, ce qui

arrivera avec probabilité 1 dès que $p \leq 1/2$. Dans tous les jeux d'argent, la probabilité p de gagner une partie est inférieure à $1/2$, et en général strictement. Stratégiquement, le joueur a intérêt à ajouter une autre condition d'arrêt du type sa fortune atteint un certain seuil $a > Y_0$ ou b tel que $Y_0 > b > 0$. Avec cette nouvelle stratégie, on aimerait connaître l'espérance de gain en fonction des paramètres Y_0 , a , b et p , ou encore la durée moyenne du jeu.

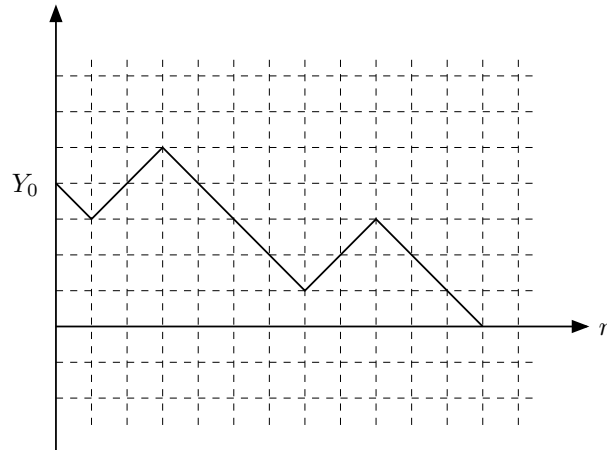


FIGURE 1 – Le joueur est ruiné à l'issue de la 12^e partie.

• **Problème de l'extinction d'une population.** Considérons une suite doublement indexée de v.a. $\{Y_{n,m}, n \in \mathbb{N}, m \in \mathbb{N}^*\}$ i.i.d. et à valeurs entières. La variable $Y_{n,m}$ représente le nombre de fils du m^e individu dans la n^e génération (s'il existe). Posons $X_0 = 1$; il y a initialement un seul individu (génération 0). Puis, pour tout n ,

$$X_{n+1} = \sum_{m=1}^{X_n} Y_{n,m}$$

représente le nombre d'individu dans la $(n+1)^e$ génération. La suite $(X_n)_{n \geq 1}$ est un processus stochastique à temps discret $T = \mathbb{N}$ (ce sont les générations) et à valeurs dans $E = \mathbb{N}$. Il est connu sous le nom de *processus de branchement* ou *arbre de Galton-Watson*.

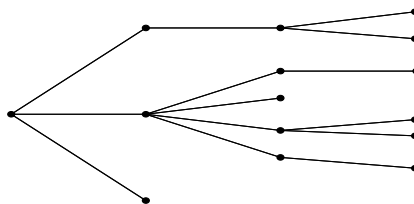


FIGURE 2 – Sont représentées les quatre premières générations d'un arbre de Galton-Watson. Le premier individu a $Y_{0,1} = 3$ fils. Ceux-ci auront respectivement $Y_{1,1} = 1$, $Y_{1,2} = 4$ et $Y_{1,3} = 0$ fils. La 2^e génération comprend donc 5 individus : $X_2 = 5$.

Historiquement, Galton et Watson ont introduit ce modèle pour étudier la perpétuation des lignées des Lords en Angleterre au 19^e siècle : les individus sont des Lords qui transmettent leur titre uniquement à

leurs fils. Il s'agit alors d'étudier l'évolution de la population au cours du temps, i.e. la quantité X_n quand n devient grand. Y aura-t'il extinction de la lignée de Lords ? Voici une première réponse pleine de bon sens : si le nombre moyen $\mathbb{E}[Y_{0,1}]$ de fils de chaque individu est élevé, la population devrait rapidement croître. À l'inverse, si $\mathbb{E}[Y_{0,1}]$ est proche de 0, la population devrait s'éteindre.

• **Séries temporelles ou chronologiques.** Les séries temporelles sont des processus stochastiques. Elles peuvent illustrer le nombre de morts suite à des accidents de la route dans un pays donné durant un intervalle de temps, le nombre de passagers dans les transports aériens ou encore les valeurs de clôtures journalières du CAC40.

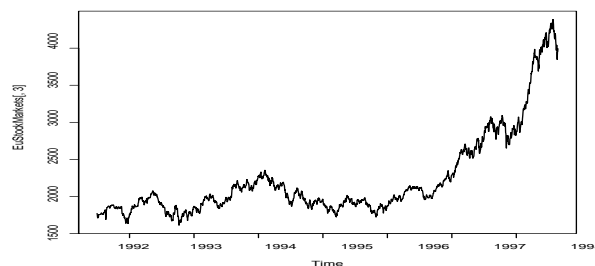


FIGURE 3 – Valeurs de clôtures journalières du CAC40 de 1991 à 1998.

Un des objectifs principaux de l'étude d'une série temporelle est la prévision des réalisations futures, très souvent pour des raisons économiques (prévoir l'évolution de la demande d'un produit pour ajuster au mieux les moyens de production, prévoir l'évolution d'un marché financier...).

Les séries temporelles feront l'objet de tout un cours en GIS5.

• **Files d'attente.** La salle de réservation d'une grande gare SNCF donne une bonne représentation d'une file d'attente (*queue* en anglais). Elle comprend un certain nombre de guichets et des clients qui sont soit en train d'être servis, soit en attente qu'un guichet se libère. Le nombre total de ces clients présents dans la salle de réservation au temps t est noté N_t . Le hasard intervient dans les arrivées des clients ainsi que dans la durée des services. La suite $(N_t)_{t \geq 0}$ est un processus stochastique à temps continu et à valeurs dans $E = \mathbb{N}$. L'objectif est d'étudier l'évolution de N_t au cours du temps afin d'optimiser le nombre de guichets nécessaires pour satisfaire en un temps raisonnable les clients.

On peut également modéliser par une file d'attente un système informatique dans lequel les nouvelles tâches se mettent en attente avant d'être traitées, ou encore un poste de travail dans une usine.

Les files d'attentes seront étudiées en détails dans ce cours.

En GIS3 ou en GIS4, vous avez déjà vu des processus stochastiques : n'importe quelle suite de v.a. i.i.d. en est un exemple ! L'indépendance entre les variables facilite les calculs et permet d'obtenir sans trop d'efforts des théorèmes limites intéressants et pratiques (Loi des Grands Nombres, Théorème Central Limite...). Malheureusement, l'indépendance n'est pas un phénomène courant dans la nature. Intégrer de la dépendance entre les variables permet de modéliser plus fidèlement la réalité. Il y a néanmoins un coût à payer ; les calculs sont plus difficiles à mener et les théorèmes plus chers.

Les processus stochastiques que nous étudierons dans ce cours prendront en compte une certaine dépendance entre les variables ; une dépendance de type *markovienne* (ou de *Markov*). Cela signifie que l'évolution future du processus ne dépend de son passé que par l'intermédiaire du présent. Par exemple pour

décider au mieux du 21^e coup à jouer dans une partie d'échecs, il suffit de connaître la configuration du jeu à l'issue du 20^e coup, le détail des 19 premiers coups n'ayant alors aucune importance.

Les exemples décrits précédemment sont markoviens. La fortune du joueur à l'issue de la $(n+1)^{\text{e}}$ partie ne dépend que de sa fortune à l'issue de la n^{e} et du résultat de la $(n+1)^{\text{e}}$ partie, mais pas de l'évolution totale de sa fortune depuis le début du jeu. Pour un processus de Galton-Watson, le nombre d'individus dans la génération à venir ne dépend que du nombre d'individus dans la génération actuelle et du nombre de fils qu'ils auront. Les files d'attente que nous étudierons seront aussi des processus stochastiques markoviens mais en temps continu. Par contre, dans le but de prédire plus fidèlement le futur, l'évolution dynamique des séries temporelles peut dépendre du passé. Elles ne rentrent donc pas dans le cadre de ce cours.

Exercices

Exercice 1 : Dénombrement de trajectoires d'une marche aléatoire

Partie 1. Considérons la marche aléatoire simple $(X_n)_{n \in \mathbb{N}}$ à valeurs dans \mathbb{Z} , définie par

$$\begin{cases} X_0 = y_0 \in \mathbb{Z} \\ X_n = y_0 + Y_1 + \dots + Y_n, \forall n \geq 1 \end{cases}$$

où $(Y_n)_{n \geq 1}$ est une suite de v.a. i.i.d., chacune valant 1 avec probabilité p et -1 avec probabilité $1 - p$, et $y_0 \in \mathbb{Z}$.

1. Quelles sont les valeurs que peut prendre X_{100} ? Notons E_{100} cet ensemble.
2. Soit $y \in E_{100}$. Combien y-a-t'il de trajectoires vérifiant $X_{100} = y$? Préciser ce nombre lorsque $y = y_0 + 100$, $y = y_0 - 100$ et $y = y_0$.
3. Soit $y \in E_{100}$. Montrer que toutes les trajectoires vérifiant $X_{100} = y$ ont la même probabilité. Quelle est cette probabilité ?
4. *Principe de réflexion.* Soient x, x', y, y' des entiers tels que $0 \leq x \leq x'$ et $yy' \geq 0$. Justifier heuristiquement qu'il y a autant de trajectoires de la marche aléatoire allant de (x, y) à (x', y') en touchant l'axe des abscisses, que de trajectoires allant de $(x, -y)$ à (x', y') .

Partie 2. Application au distributeur automatique de boissons.

Considérons un distributeur automatique de boissons, chacune valant 1 euro. Supposons que 60% des clients désirant acheter une boisson la paie avec une pièce de 1 euro, et le reste, avec une pièce de 2 euros. Dans ce dernier cas, le distributeur rend au consommateur sa monnaie, i.e. une pièce de 1 euro. À condition qu'il en ait... Il s'agit donc pour l'appareil de prévoir dans le distributeur, en début de journée, un stock suffisant de pièces de 1 euro. Mais pas trop pour ne pas bloquer inutilement de la trésorerie !

5. Supposons que dans une journée donnée, 100 clients se présentent et que exactement 60 d'entre eux paient avec une pièce de 1 euro. Quel stock initial permet d'assurer (disons à 95%) que chaque client récupère sa monnaie ?

Chapitre 1

Chaînes de Markov

Un modèle d'évolution dynamique en temps discret dans lequel on fait dépendre l'évolution future de l'état présent et du hasard est une chaîne de Markov. C'est un processus stochastique à temps discret. On en rencontre dans de nombreux domaines d'applications...

1.1 Qu'est-ce qu'une chaîne de Markov ?

1.1.1 Définitions et exemples

Définition 1.1.1 Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. à valeurs dans un espace E fini ou dénombrable, appelé *espace d'états*. On dit que $(X_n)_{n \in \mathbb{N}}$ est une *chaîne de Markov* si

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

pour tout entier $n \in \mathbb{N}$, pour tout état j et pour toute suite d'états $i_0, i_1, \dots, i_{n-1}, i$ pour lesquels la probabilité conditionnelle a un sens, i.e.

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

Si de plus la quantité $\mathbb{P}(X_{n+1} = j \mid X_n = i)$ ne dépend pas de n , i.e.

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_1 = j \mid X_0 = i)$$

alors la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ est dite *homogène*.

Il faut comprendre une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ comme une promenade dans l'espace d'états E , la variable X_n indiquant l'état dans lequel on est à l'instant n . La v.a. X_0 représente l'état initial duquel démarre la chaîne. Selon le contexte, X_0 pourra être aléatoire ou déterministe.

La propriété de Markov signifie que, connaissant le dernier état visité (disons à l'instant n), la loi du prochain état visité (i.e. la loi de X_{n+1}) ne dépend pas des états visités depuis l'instant 0 jusqu'à l'instant $n - 1$. Plus prosaïquement, on dit que

conditionnellement au présent, le futur ne dépend pas du passé.

Mais il dépend du présent : X_n et X_{n+1} n'ont aucune raison d'être indépendantes !

La propriété d'homogénéité d'une chaîne de Markov exprime quant à elle que la probabilité d'aller de i en

j reste la même au cours du temps. Elle permet de regrouper en une seule matrice (indépendante de n) les probabilités de transition entre deux états quelconques.

Définition 1.1.2 Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène à espace d'états E . Soient $i, j \in E$ deux états. La probabilité

$$p_{i,j} := \mathbb{P}(X_1 = j \mid X_0 = i)$$

est appelée *probabilité de transition de i à j* . La matrice $P := (p_{i,j})_{i,j \in E}$ est appelée *matrice de transition* de la chaîne.

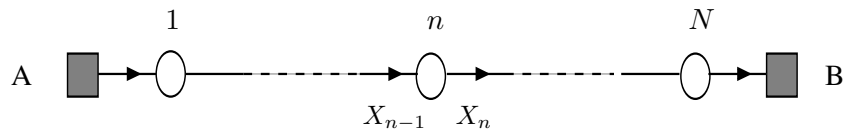
Lorsque l'espace d'états E est fini, la matrice P est carrée de taille $\text{Card}(E)$. Lorsqu'il est infini, elle admet un nombre infini de lignes et de colonnes. Les coefficients de la matrice P sont positifs ou nuls. Leur somme sur une même ligne vaut 1 : pour tout $i \in E$,

$$\sum_{j \in E} p_{i,j} = \sum_{j \in E} \mathbb{P}(X_1 = j \mid X_0 = i) = \mathbb{P}\left(\bigcup_{j \in E} \{X_1 = j\} \mid X_0 = i\right) = 1.$$

Comme nous le verrons dans les exemples suivants, il arrive fréquemment dans les applications que pour un état i donné, le nombre d'états j directement accessibles depuis i (i.e. tel que $p_{i,j} > 0$) soit faible. La matrice de transition est alors très creuse ; elle contient beaucoup de 0. Il est souvent plus économique (et plus pertinent) de résumer les probabilités de transition dans le *diagramme de transition*. C'est un graphe orienté et pondéré dont l'ensemble des sommets est E . Une arête de poids $p_{i,j}$ va de i à j si $p_{i,j} > 0$.

EXEMPLES :

• **Transmission d'un bit informatique.** Un bit informatique valant 0 ou 1 est transmis d'un poste A vers un poste B en passant par N intermédiaires. Chacun de ces intermédiaires transmet correctement le bit avec probabilité p et l'inverse avec probabilité $1 - p$, indépendamment les uns des autres. Le bit (aléatoire) d'entrée, disons X_0 , est supposé indépendant des intermédiaires. Pour $n = 1, \dots, N$, notons X_n le bit sortant du n^{e} intermédiaire.



La suite de v.a. $(X_n)_{0 \leq n \leq N}$ est à valeurs dans l'espace d'états $E = \{0, 1\}$. Vérifions que c'est une chaîne de Markov. Considérons pour ce faire, une suite i_0, \dots, i_n, i_{n+1} d'éléments de $\{0, 1\}$.

$$\begin{aligned} \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) &= \frac{\mathbb{P}(X_{n+1} = i_{n+1} \text{ et } X_n = i_n, \dots, X_0 = i_0)}{\mathbb{P}(X_n = i_n, \dots, X_0 = i_0)} \\ &= \frac{\mathbb{P}(|X_{n+1} - X_n| = |i_{n+1} - i_n| \text{ et } X_n = i_n, \dots, X_0 = i_0)}{\mathbb{P}(X_n = i_n, \dots, X_0 = i_0)} \\ &= \mathbb{P}(|X_{n+1} - X_n| = |i_{n+1} - i_n|), \end{aligned}$$

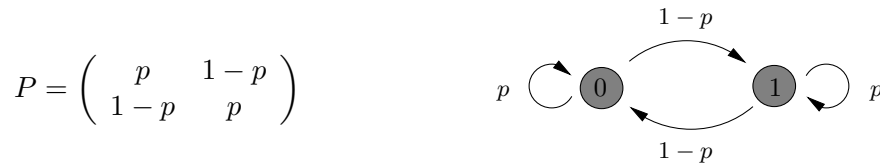
du fait de l'indépendance entre la v.a. $X_{n+1} - X_n$ (qui représente l'action du $(n+1)^e$ intermédiaire) et l'état du bit à la sortie des n premiers intermédiaires. Pour la même raison ;

$$\begin{aligned}\mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) &= \frac{\mathbb{P}(|X_{n+1} - X_n| = |i_{n+1} - i_n|) \mathbb{P}(X_n = i_n)}{\mathbb{P}(X_n = i_n)} \\ &= \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n) .\end{aligned}$$

Enfin, le caractère homogène de la chaîne $(X_n)_{0 \leq n \leq N}$ résulte du calcul :

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(|X_{n+1} - X_n| = |j - i|) = \begin{cases} p & \text{si } i = j \\ 1 - p & \text{sinon.} \end{cases} .$$

Voici la matrice et le graphe de transition de cette chaîne :



• **La marche aléatoire simple.** Soient $(Y_n)_{n \geq 1}$ une suite de v.a. i.i.d., chacune valant 1 avec probabilité p et -1 avec probabilité $1 - p$. Soit $Y_0 \in \mathbb{Z}$ une v.a. indépendante des Y_n , représentant le point de départ sur l'axe \mathbb{Z} de la chaîne. On définit la *marche aléatoire simple* par

$$\begin{aligned}X_{n+1} &= X_n + Y_{n+1} \\ &= Y_0 + Y_1 + \dots + Y_n + Y_{n+1} ,\end{aligned}$$

pour tout $n \in \mathbb{N}$. Le processus stochastique $(X_n)_{n \geq 1}$ est une chaîne de Markov homogène. Son espace d'états $E = \mathbb{Z}$ est cette fois infini. Comme dans l'exemple du bit informatique, le caractère markovien provient de l'indépendance des Y_n :

$$\begin{aligned}\mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) &= \frac{\mathbb{P}(X_{n+1} = i_{n+1} \text{ et } X_n = i_n, \dots, X_0 = i_0)}{\mathbb{P}(X_n = i_n, \dots, X_0 = i_0)} \\ &= \frac{\mathbb{P}(Y_{n+1} = i_{n+1} - i_n \text{ et } X_n = i_n, \dots, X_0 = i_0)}{\mathbb{P}(X_n = i_n, \dots, X_0 = i_0)} \\ &= \mathbb{P}(Y_{n+1} = i_{n+1} - i_n) \\ &= \frac{\mathbb{P}(Y_{n+1} = i_{n+1} - i_n) \mathbb{P}(X_n = i_n)}{\mathbb{P}(X_n = i_n)} \\ &= \frac{\mathbb{P}(X_{n+1} - X_n = i_{n+1} - i_n \text{ et } X_n = i_n)}{\mathbb{P}(X_n = i_n)} \\ &= \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n) ,\end{aligned}$$

où les i_0, \dots, i_n, i_{n+1} sont des états. Idem pour le caractère homogène :

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(Y_{n+1} = j - i) = \begin{cases} p & \text{si } j = i + 1 \\ 1 - p & \text{si } j = i - 1 \\ 0 & \text{sinon.} \end{cases} .$$

La matrice de transition de la chaîne $(X_n)_{n \geq 1}$ est de taille infinie. Sa i^{e} ligne est de la forme :

$$\cdots \quad 0 \quad 0 \quad 1-p \quad 0 \quad p \quad 0 \quad 0 \quad \cdots$$

où le “0” intercalé entre les coefficients $1-p$ et p est sur la i^{e} colonne. Son graphe de transition est donné par la Figure 1.1.

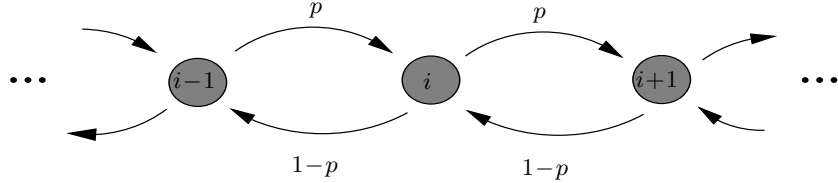


FIGURE 1.1 – Le graphe de transition de la marche aléatoire simple.

• **Le processus de Galton-Watson** décrit en introduction est également une chaîne de Markov homogène.

Dans la pratique, nous prouverons qu’une suite de v.a. est une chaîne de Markov en utilisant le résultat intuitif suivant :

Proposition 1.1.3 Une suite de v.a. $(X_n)_{n \geq 1}$ définie récursivement par $X_{n+1} = f(X_n, Y_{n+1})$ où

- $(Y_n)_{n \geq 1}$ est une suite de v.a. i.i.d. à valeurs dans un espace E' ,
- $X_0 \in E$ est une v.a. donnée et indépendante des $(Y_n)_{n \geq 1}$,
- $f : E \times E' \rightarrow E$ est une application déterministe,

est une chaîne de Markov homogène à espace d’états E .

Ce critère permet d’affirmer très rapidement que la marche aléatoire simple définie précédemment est une chaîne de Markov homogène. En effet, les incréments $Y_0, Y_1, Y_2 \dots$ sont des v.a. i.i.d. à valeurs dans l’espace $E' = \{-1; 1\}$ ($E' = \mathbb{Z}$ convenait également), $X_0 = Y_0$ et $X_{n+1} = f(X_n, Y_{n+1})$ avec comme application déterministe $f : f(x, y) = x + y$.

1.1.2 Loi d’une chaîne de Markov

Proposition 1.1.4 La loi d’une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ est entièrement déterminée par la donnée de sa matrice de transition P et de la loi de X_0 , appelée *loi initiale* et notée μ_0 :

$$\text{pour tout } i \in E, \mu_0(i) := \mathbb{P}(X_0 = i).$$

Plus précisément, pour tout entier n et toute suite d’états $i_0, i_1, \dots, i_{n-1}, i_n$ de E :

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = \mu_0(i_0) p_{i_0, i_1} \cdots p_{i_{n-1}, i_n}. \quad (1.1)$$

La formule (1.1) permet d’écrire la probabilité d’une intersection, i.e.

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0),$$

comme un produit de probabilités conditionnelles, les coefficients $p_{i,j}$.
En divisant par $\mu_0(i_0)$ dans (1.1), il s'ensuit que :

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1 \mid X_0 = i_0) = p_{i_0, i_1} \dots p_{i_{n-1}, i_n}.$$

Démonstration Le résultat repose sur la formule

$$\mathbb{P}\left(\bigcap_{k=0}^n A_k\right) = \mathbb{P}(A_0) \mathbb{P}(A_1 \mid A_0) \mathbb{P}(A_2 \mid A_1 \cap A_0) \dots \mathbb{P}(A_n \mid A_{n-1} \cap \dots \cap A_0)$$

à démontrer par récurrence sur $n \in \mathbb{N}$ et appliquée aux événements $A_k = \{X_k = i_k\}$. Il ne reste plus qu'à identifier les termes : $\mathbb{P}(A_0) = \mathbb{P}(X_0 = i_0)$ vaut $\mu(i_0)$, $\mathbb{P}(A_1 \mid A_0) = \mathbb{P}(X_1 = i_1 \mid X_0 = i_0)$ vaut par définition p_{i_0, i_1} et

$$\begin{aligned} \mathbb{P}(A_2 \mid A_1 \cap A_0) &= \mathbb{P}(X_2 = i_2 \mid X_1 = i_1, X_0 = i_0) \\ &= \mathbb{P}(X_2 = i_2 \mid X_1 = i_1) \\ &= p_{i_1, i_2}, \end{aligned}$$

par la propriété de Markov. C'est pareil pour les autres termes. ■

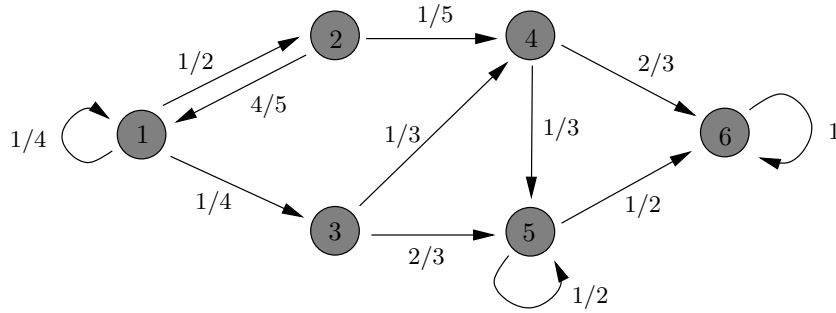


FIGURE 1.2 – Un graphe de transition sur l'espace d'états $E = \{1, 2, 3, 4, 5, 6\}$.

Partant de $X_0 = 1$, quel est le chemin le plus probable permettant $X_3 = 6$? On dénombre 3 chemins permettant de rejoindre l'état 6 depuis 1 en 3 coups : $1 \rightarrow 2 \rightarrow 4 \rightarrow 6$, $1 \rightarrow 3 \rightarrow 4 \rightarrow 6$ et $1 \rightarrow 3 \rightarrow 5 \rightarrow 6$. Leurs probabilités respectives sont :

$$\mathbb{P}(X_3 = 6, X_2 = 4, X_1 = 2 \mid X_0 = 1) = p_{1,2} p_{2,4} p_{4,6} = \frac{1}{2} \times \frac{1}{5} \times \frac{2}{3} = \frac{1}{15},$$

$$\mathbb{P}(X_3 = 6, X_2 = 4, X_1 = 3 \mid X_0 = 1) = p_{1,3} p_{3,4} p_{4,6} = \frac{1}{4} \times \frac{1}{3} \times \frac{2}{3} = \frac{1}{18},$$

$$\mathbb{P}(X_3 = 6, X_2 = 5, X_1 = 3 \mid X_0 = 1) = p_{1,3} p_{3,5} p_{5,6} = \frac{1}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{12}.$$

Le plus probable est donc $1 \rightarrow 3 \rightarrow 5 \rightarrow 6$.

Dans toute la suite, considérons une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ à valeurs dans E , de matrice de transition P et de loi initiale μ_0 .

Proposition 1.1.5 La loi de la chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ est invariante par translation dans le temps. Autrement dit, pour tous entiers n, m , toute suite d'états $i_0, i_1, \dots, i_n, i_{n+1}, \dots, i_{n+m}$ pour lesquels

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) > 0,$$

il vient :

$$\begin{aligned} \mathbb{P}(X_{n+m} = i_{n+m}, \dots, X_{n+1} = i_{n+1} \mid X_n = i_n, \dots, X_0 = i_0) = \\ \mathbb{P}(X_m = i_{n+m}, \dots, X_1 = i_{n+1} \mid X_0 = i_n). \end{aligned}$$

La probabilité conditionnelle d'une trajectoire donnée (i.e. les états i_{n+1}, \dots, i_{n+m}) reste la même au cours du temps, qu'elle ait lieu entre les instants $n+1$ et $n+m$ ou entre les instants 1 et m . Seul compte le dernier état visité, en l'occurrence i_n . Dans l'exemple donné dans la Figure 1.2, $1 \rightarrow 3 \rightarrow 5 \rightarrow 6$ reste le chemin le plus probable pour rejoindre l'état 6 depuis l'état 1 en 3 coups, que ce soit entre les instants 0 et 3, 47 et 50, ou encore 1234 et 1237 !

Une généralisation de ce résultat est connue sous le nom de *relation de Chapman-Kolmogorov* :

$$\mathbb{P}(X_{n+m} = j \mid X_0 = i) = \sum_{k \in E} \mathbb{P}(X_m = j \mid X_0 = k) \mathbb{P}(X_n = k \mid X_0 = i).$$

Il faut la lire comme suit : aller de i à j en $n+m$ pas, c'est aller de i à un certain état k en n pas puis de k à j en m pas.

Notons par μ_n la loi de X_n . C'est une mesure de probabilité sur E que l'on peut écrire sous la forme d'un vecteur ligne $(\mu_n(j))_{j \in E}$ (i.e. un élément de $\mathbb{R}^{\text{Card}(E)}$). L'objectif de la fin de cette section consiste à établir une relation matricielle liant μ_n à la loi initiale μ_0 et à la matrice P

Pour ce faire, notons par $(p_{i,j}^{(n)})_{i,j \in E}$ les coefficients de la matrice P^n , puissance n^e de P . L'expression brute de $p_{i,j}^{(n)}$ est

$$p_{i,j}^{(n)} = \sum_{i_1, \dots, i_{n-1} \in E} p_{i,i_1} p_{i_1,i_2} \dots p_{i_{n-1},j}.$$

Cette formule étant purement algébrique, voici une nouvelle expression du coefficient $p_{i,j}^{(n)}$ lui donnant davantage de sens :

Proposition 1.1.6 Soient n un entier et i, j des états.

- (1) $\mathbb{P}(X_n = j \mid X_0 = i) = p_{i,j}^{(n)}$.
- (2) $\mu_{n+1} = \mu_n P$.
- (3) $\mu_n = \mu_0 P^n$, ce qui signifie que $\mathbb{P}(X_n = j)$ est le j^e élément du vecteur ligne $\mu_0 P^n$.

En utilisant des vecteurs colonnes pour décrire les lois de X_n et X_{n+1} , le point (2) devient

$${}^t \mu_{n+1} = {}^t P {}^t \mu_n.$$

On peut donc voir l'évolution en loi de la suite $(X_n)_{n \in \mathbb{N}}$ comme un système itératif linéaire dont ${}^t P$ est la matrice d'évolution : on obtient la loi de X_{n+1} en multipliant (matriciellement et par la gauche) la loi de X_n par ${}^t P$.

Reprenons l'exemple donné dans la Figure 1.2. La probabilité de rejoindre l'état 6 depuis l'état 1, en 3 coups et quel que soit le chemin emprunté, se calcule :

$$\begin{aligned} p_{1,6}^{(3)} &= \sum_{i,j \in E} p_{1,i} p_{i,j} p_{j,6} \\ &= p_{1,2} p_{2,4} p_{4,6} + p_{1,3} p_{3,4} p_{4,6} + p_{1,3} p_{3,5} p_{5,6} \\ &= \frac{1}{15} + \frac{1}{18} + \frac{1}{12} \simeq 0.20 . \end{aligned}$$

En partant de l'état 1 avec probabilité 1, les lois de X_1 et X_2 se calculent de la manière suivante.

$$\mu_1 = \mu_0 P = (1 \ 0 \ 0 \ 0 \ 0 \ 0) \begin{pmatrix} 1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\ 4/5 & 0 & 0 & 1/5 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 2/3 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 2/3 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = (0.25 \ 0.5 \ 0.25 \ 0 \ 0 \ 0) .$$

$$\mu_2 = \mu_1 P = \mu_0 P^2 \simeq (0.46 \ 0.12 \ 0.07 \ 0.18 \ 0.17 \ 0) .$$

De même, $\mu_5 \simeq (0.15 \ 0.12 \ 0.06 \ 0.04 \ 0.16 \ 0.48)$, $\mu_{15} \simeq (0.01 \ 0.01 \ 0 \ 0 \ 0.01 \ 0.96)$.

Démonstration Seul le cas $n = 2$ sera traité pour donner l'intuition concernant le point (1). Le coefficient $p_{i,j}^{(2)}$ de la i^e ligne et de la j^e colonne de la matrice P^2 vaut

$$p_{i,j}^{(2)} = \sum_{k \in E} p_{i,k} p_{k,j} .$$

La formule des probabilités totales et la Proposition 1.1.4 permettent d'écrire les égalités :

$$\begin{aligned} \mathbb{P}(X_2 = j, X_0 = i) &= \sum_{k \in E} \mathbb{P}(X_2 = j, X_1 = k, X_0 = i) \\ &= \mu_0(i) \sum_{k \in E} p_{i,k} p_{k,j} , \end{aligned}$$

desquelles on en déduit le résultat ; $\mathbb{P}(X_2 = j | X_0 = i) = p_{i,j}^{(2)}$. Par ailleurs,

$$\begin{aligned} \mu_{n+1}(j) = \mathbb{P}(X_{n+1} = j) &= \sum_{i \in E} \mathbb{P}(X_{n+1} = j, X_n = i) \\ &= \sum_{i \in E} \mathbb{P}(X_{n+1} = j \mid X_n = i) \mathbb{P}(X_n = i) \\ &= \sum_{i \in E} p_{i,j} \mathbb{P}(X_n = i) . \end{aligned}$$

Cette égalité implique la relation matricielle $\mu_{n+1} = \mu_n P$ (i.e. le point (2)). Enfin, le point (3) s'obtient par une récurrence immédiate. ■

1.2 Classification des états

1.2.1 Classes irréductibles

Définition 1.2.1 Soient i et j deux états de E . On dit que l'état j est *accessible* depuis l'état i si

$$\exists n \in \mathbb{N}, p_{i,j}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = i) > 0.$$

On dit que les états i et j *communiquent* si chacun est accessible depuis l'autre. On note alors $i \leftrightarrow j$.

Proposition 1.2.2 La relation \leftrightarrow est une relation d'équivalence sur E .

Démonstration La réflexivité (i.e. $i \leftrightarrow i$) est immédiate : pour $n = 0$, $\mathbb{P}(X_0 = i \mid X_0 = i) = 1$. Il en va de même pour la symétrie ($i \leftrightarrow j$ implique $j \leftrightarrow i$). Enfin, la transitivité repose sur la relation de Chapman-Kolmogorov. Supposons que $i \leftrightarrow j$ et $j \leftrightarrow k$. En particulier, les états j et k sont respectivement accessibles depuis i et j . Il existe donc des entiers m et n tels que $\mathbb{P}(X_m = k \mid X_0 = j) > 0$ et $\mathbb{P}(X_n = j \mid X_0 = i) > 0$. Par conséquent,

$$\begin{aligned} \mathbb{P}(X_{n+m} = k \mid X_0 = i) &= \sum_{l \in E} \mathbb{P}(X_m = k \mid X_0 = l) \mathbb{P}(X_n = l \mid X_0 = i) \\ &\geq \mathbb{P}(X_m = k \mid X_0 = j) \mathbb{P}(X_n = j \mid X_0 = i) > 0, \end{aligned}$$

d'où k est accessible depuis l'état i . C'est la même chose dans l'autre sens. ■

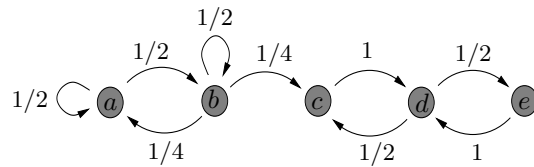
L'espace E peut donc être partitionné en classes d'équivalence pour la relation \leftrightarrow , appelées *classes irréductibles*. Nous insisterons dans les paragraphes suivants sur le fait que les états d'une même classe irréductible ont des propriétés équivalentes vis à vis de la chaîne (réurrence, transience et périodicité).

Lorsque l'espace E est réduit à une seule classe (i.e. tous les états communiquent), on dit que la chaîne est *irréductible*. En général, E se partitionne en états isolés dans lesquels on ne revient jamais une fois qu'on les a quittés, et en classes irréductibles disjointes.

Pour déterminer les classes irréductibles d'une chaîne de Markov, il est commode de travailler sur le graphe de transition plutôt que sur la matrice de transition P .

EXEMPLE : Considérons une chaîne de Markov à valeurs dans $E = \{a, b, c, d, e\}$ et dont la matrice et le graphe de transition sont données par :

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



La chaîne comporte deux classes irréductibles : $\{a, b\}$ et $\{c, d, e\}$.

1.2.2 Récurrence et transience

Définition 1.2.3 Soit $i \in E$. La v.a. T_i définie par

$$T_i = \inf\{n \geq 1, X_n = i\}$$

est appelée *temps d'atteinte de i* ou encore *temps de retour à i* lorsque la chaîne $(X_n)_{n \in \mathbb{N}}$ part de i . Par convention, lorsque pour tout $n \geq 1$, $X_n \neq i$, on pose $T_i = +\infty$.

Il faut bien comprendre que l'événement

$$\{T_i < +\infty\} = \{\exists n \geq 1, X_n = i\} = \bigcup_{n \geq 1} \{X_n = i\}$$

signifie que la chaîne $(X_n)_{n \in \mathbb{N}}$ repassera par l'état i (au moins une fois) alors que

$$\{T_i = +\infty\} = \{\forall n \geq 1, X_n \neq i\} = \bigcap_{n \geq 1} \{X_n \neq i\}$$

signifie que la chaîne $(X_n)_{n \in \mathbb{N}}$ ne repassera jamais par i .

Définition 1.2.4 Un état $i \in E$ est dit *récurrent* si, partant de i , on y revient presque sûrement en temps fini :

$$\mathbb{P}(T_i < +\infty \mid X_0 = i) = 1.$$

L'état i est dit *transient* dans le cas contraire, i.e. lorsque $\mathbb{P}(T_i = +\infty \mid X_0 = i) > 0$.

Autrement dit, un état est transient si avec probabilité strictement positive, on peut le quitter pour ne jamais y revenir. Comme cas particulier d'état transient, on retrouve les états pour lesquels $p_{i,i}^{(n)} = 0$, pour tout $n \geq 1$. Ce sont ceux que l'on quitte au premier pas pour ne jamais y revenir ;

$$\begin{aligned} \mathbb{P}(T_i < +\infty \mid X_0 = i) &= \mathbb{P}(\exists n \geq 1, X_n = i \mid X_0 = i) \\ &= \mathbb{P}\left(\bigcup_{n \geq 1} \{X_n = i\} \mid X_0 = i\right) \\ &\leq \sum_{n \geq 1} \mathbb{P}(X_n = i \mid X_0 = i) \\ &\leq \sum_{n \geq 1} p_{i,i}^{(n)} = 0. \end{aligned}$$

EXEMPLES :

• Reprenons le cas de la chaîne de Markov à valeurs dans $E = \{a, b, c, d, e\}$ définie dans le paragraphe précédent. L'état b est transient. En effet, l'état c est accessible depuis b mais pas l'inverse. Autrement dit, en allant en c depuis b , on est sûr de ne jamais y revenir :

$$\begin{aligned} \mathbb{P}(T_b = +\infty \mid X_0 = b) &= \mathbb{P}\left(\bigcap_{n \geq 1} \{X_n \neq b\} \mid X_0 = b\right) \\ &\geq \mathbb{P}(X_1 = c \mid X_0 = b) = 1/4. \end{aligned}$$

Ce qui donne $\mathbb{P}(T_b < +\infty \mid X_0 = b) < 1$. Il en va de même pour l'état a :

$$\mathbb{P}(T_a = +\infty \mid X_0 = a) \geq \mathbb{P}(X_2 = c, X_1 = b \mid X_0 = a) = \frac{1}{4} \times \frac{1}{2},$$

d'après la Proposition 1.1.4. Les états c , d et e seront quant à eux récurrents. Étudions le cas de c . En partant de c , l'unique solution pour ne jamais y revenir consiste à aller en d puis à transiter indéfiniment entre d et e , ce qui est de probabilité nulle. En effet, pour tout n ,

$$\begin{aligned} \mathbb{P}(T_c = +\infty \mid X_0 = c) &\leq \mathbb{P}(X_{2n+1} = d, X_{2n} = e, \dots, X_3 = d, X_2 = e, X_1 = d \mid X_0 = c) \\ &\leq p_{c,d}(p_{d,e}p_{e,d})^n, \end{aligned}$$

toujours d'après la Proposition 1.1.4. Avec $p_{d,e}p_{e,d} = 1/2$, on obtient la majoration :

$$\mathbb{P}(T_c = +\infty \mid X_0 = c) \leq \frac{1}{2^n}.$$

Il ne reste plus qu'à faire tendre n vers l'infini.

• En exercice, il sera montré que la marche aléatoire sur \mathbb{Z} est récurrente dans le cas symétrique, i.e. pour $p = 1/2$, et transiente sinon.

La v.a. N_i est le nombre de passages de la chaîne $(X_n)_{n \in \mathbb{N}}$ par l'état i après l'instant 0 :

$$N_i = \sum_{n \geq 1} \mathbb{1}_{X_n = i} = \text{Card}\{n \geq 1, X_n = i\}.$$

Partant de j , on accède en i en temps fini, puis on revient en i au moins n fois (i.e. $N_i \geq n$). Cela fait donc au moins $n + 1$ passages en i (partant de j).

Lemme 1.2.5 Pour tout entier n et tous états i et j ,

$$\mathbb{P}(N_i \geq n + 1 \mid X_0 = j) = \mathbb{P}(T_i < +\infty \mid X_0 = j) \mathbb{P}(N_i \geq n \mid X_0 = i).$$

Les états transients sont ceux dans lesquels on ne passe qu'un nombre fini de fois. Par opposition, on revient une infinité de fois dans un état récurrent.

Proposition 1.2.6 Les conditions suivantes sont équivalentes :

- (1) l'état i est récurrent : $\mathbb{P}(T_i < +\infty \mid X_0 = i) = 1$;
- (2) conditionnellement à $X_0 = i$, la chaîne de markov $(X_n)_{n \in \mathbb{N}}$ revient presque sûrement une infinité de fois en i :

$$\mathbb{P}(N_i = +\infty \mid X_0 = i) = 1 ;$$

- (3) la série $\sum p_{i,i}^{(n)}$ diverge.

Proposition 1.2.7 Les conditions suivantes sont équivalentes :

- (4) l'état i est transient : $\mathbb{P}(T_i < +\infty \mid X_0 = i) < 1$;
- (5) conditionnellement à $X_0 = i$, la v.a. $N_i + 1$ suit la loi géométrique de paramètre $1 - \alpha_i$, avec $\alpha_i = \mathbb{P}(T_i < +\infty \mid X_0 = i)$:

$$\forall n \in \mathbb{N}, \mathbb{P}(N_i = n \mid X_0 = i) = \alpha_i^n (1 - \alpha_i) ;$$

(en particulier N_i est presque sûrement finie) ;

(6) conditionnellement à $X_0 = i$, la v.a. N_i est intégrable : $\mathbb{E}[N_i | X_0 = i] = \sum p_{i,i}^{(n)} < +\infty$.

Remarquons que l'identité entre l'espérance conditionnelle $\mathbb{E}[N_i | X_0 = i]$ et la série $\sum p_{i,i}^{(n)}$ est valable quelle que soit la nature de cette dernière, en la considérant comme un élément de $[0; +\infty]$. Elle repose sur le théorème de Fubini.

$$\begin{aligned} \mathbb{E}[N_i | X_0 = i] &= \mathbb{E} \left[\sum_{n \geq 1} \mathbb{1}_{X_n = i} \mid X_0 = i \right] \\ &= \sum_{n \geq 1} \mathbb{E}[\mathbb{1}_{X_n = i} | X_0 = i] \\ &= \sum_{n \geq 1} \mathbb{P}(X_n = i | X_0 = i) \\ &= \sum_{n \geq 1} p_{i,i}^{(n)}. \end{aligned}$$

Reprenons l'exemple de la Section 1.2.1. On a déjà établi que

$$\mathbb{P}(T_b = +\infty | X_0 = b) \geq 1/4.$$

On peut montrer qu'il y a en fait égalité (en effet, partir de b pour aller en a et ne jamais en revenir est un événement non vide, certes, mais de probabilité nulle). Ainsi la Proposition 1.2.7 affirme que le nombre de passages en l'état b , partant de b , suit la loi géométrique de paramètre $\frac{3}{4}$. On en déduit son espérance :

$$\mathbb{E}[N_b | X_0 = b] = \mathbb{E}[N_b + 1 | X_0 = b] - 1 = \frac{1}{1 - 3/4} - 1 = 3.$$

Les deux propositions ci-dessus se démontrent simultanément.

Démonstration Prouvons $(1) \Rightarrow (2)$. En combinant, $\mathbb{P}(T_i < +\infty | X_0 = i) = 1$ et le Lemme 1.2.5, il vient :

$$\mathbb{P}(N_i \geq n + 1 | X_0 = i) = \mathbb{P}(N_i \geq n | X_0 = i) = \dots = \mathbb{P}(N_i \geq 1 | X_0 = i),$$

en itérant. Par ailleurs, $\{N_i \geq 1\} = \{T_i < +\infty\}$. On en déduit que pour tout n , $\mathbb{P}(N_i \geq n | X_0 = i) = 1$. Il ne reste plus qu'à faire tendre n vers l'infini pour obtenir (2) :

$$\begin{aligned} \mathbb{P}(N_i = +\infty | X_0 = i) &= \mathbb{P} \left(\bigcap_{n \geq 1} \{N_i \geq n\} \mid X_0 = i \right) \\ &= \lim_{n \rightarrow +\infty} \mathbb{P}(N_i \geq n | X_0 = i) \\ &= 1. \end{aligned}$$

L'implication $(2) \Rightarrow (3)$ repose sur :

$$\sum_{n \geq 1} p_{i,i}^{(n)} = \mathbb{E}[N_i | X_0 = i] \geq (+\infty) \mathbb{P}(N_i = +\infty | X_0 = i) = +\infty.$$

Supposons (4) : $\alpha_i = \mathbb{P}(T_i < +\infty | X_0 = i) < 1$. Le Lemme 1.2.5 donne :

$$\mathbb{P}(N_i \geq n+1 | X_0 = i) = \alpha_i \mathbb{P}(N_i \geq n | X_0 = i) = \dots = \alpha_i^n \mathbb{P}(N_i \geq 1 | X_0 = i) .$$

Ainsi,

$$\mathbb{P}(N_i \geq n+1 | X_0 = i) = \alpha_i^n \mathbb{P}(T_i < +\infty | X_0 = i) = \alpha_i^{n+1} .$$

On en déduit d'une part que N_i est presque sûrement finie

$$\mathbb{P}(N_i = +\infty | X_0 = i) = \lim_{n \rightarrow +\infty} \mathbb{P}(N_i \geq n | X_0 = i) = 0$$

(car $\alpha_i < 1$) et d'autre part que sa loi est géométrique de paramètre α_i . Ce qui prouve (5).

$$\mathbb{P}(N_i = n | X_0 = i) = \mathbb{P}(N_i \geq n | X_0 = i) - \mathbb{P}(N_i \geq n+1 | X_0 = i) = \alpha_i^n (1 - \alpha_i) .$$

L'implication (5) \Rightarrow (6) est immédiate :

$$\mathbb{E}[N_i | X_0 = i] = \sum_{n \geq 1} \mathbb{P}(N_i \geq n | X_0 = i) = \frac{\alpha_i}{1 - \alpha_i} < +\infty .$$

Enfin, (3) \Rightarrow (1) et (6) \Rightarrow (4) sont respectivement les contraposées de (4) \Rightarrow (6) et (1) \Rightarrow (3). Les boucles sont bouclées ! ■

La récurrence et la transience sont des propriétés de classe irréductible :

Proposition 1.2.8 Si les états i et j communiquent alors ils sont tous deux récurrents ou tous deux transients.

Ainsi, tous les états appartenant à la même classe irréductible qu'un état transient (resp. récurrent), le seront également. La classe sera alors qualifiée de transiente (resp. récurrente).

Démonstration Si les états i et j communiquent, il existe deux instants k et l tels que $p_{i,j}^{(k)} > 0$ et $p_{j,i}^{(l)} > 0$. Soit $m \geq k + l$. Partant de i , une façon d'y revenir en m pas consiste à d'abord aller en j en l pas, puis d'y revenir en $m - k - l$ pas pour enfin aller en i en k pas :

$$p_{i,i}^{(m)} \geq p_{i,j}^{(l)} p_{j,j}^{(m-k-l)} p_{j,i}^{(k)} .$$

De la même manière,

$$p_{j,j}^{(m)} \geq p_{j,i}^{(k)} p_{i,i}^{(m-k-l)} p_{i,j}^{(l)} .$$

Les deux séries $\sum_m p_{i,i}^{(m)}$ et $\sum_m p_{j,j}^{(m)}$ sont donc de même nature. Les Propositions 1.2.6 et 1.2.7 permettent alors de conclure. ■

Proposition 1.2.9 La probabilité de sortir d'une classe irréductible récurrente est nulle. Plus précisément, si i est un état récurrent et $C(i)$ sa classe alors

$$\forall j \notin C(i), \forall n \in \mathbb{N}, \mathbb{P}(X_n = j | X_0 = i) = p_{i,j}^{(n)} = 0 .$$

Démonstration Soit $j \notin C(i)$ et supposons par l'absurde qu'il existe un entier n tel que $p_{i,j}^{(n)} > 0$. Il existe alors un entier m tel que $p_{j,i}^{(m)} > 0$. Dans le cas contraire, i.e. $\forall m, p_{j,i}^{(m)} = 0$, l'état i ne pourrait être récurrent puisqu'il existerait une probabilité non nulle de quitter i et de ne jamais y revenir. Dès lors, les états i et j communiquent, ce qui contredit $j \notin C(i)$. ■

Proposition 1.2.10 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, étant donné un état i transient, le nombre moyen de passage par l'état i en partant d'un état j , i.e. $\mathbb{E}[N_i | X_0 = j]$, est finie. Cette affirmation repose sur le Lemme 1.2.5, la Proposition 1.2.7, le fait que $\mathbb{E}[Z] = \sum_{n \geq 1} \mathbb{P}(Z \geq n)$ pour une v.a. positive Z et le calcul suivant :

$$\begin{aligned} \mathbb{E}[N_i | X_0 = j] &= \sum_{n \in \mathbb{N}} \mathbb{P}(N_i \geq n + 1 | X_0 = j) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(T_i < +\infty | X_0 = j) \mathbb{P}(N_i \geq n | X_0 = i) \\ &= \mathbb{P}(T_i < +\infty | X_0 = j) \sum_{n \in \mathbb{N}} \mathbb{P}(N_i \geq n | X_0 = i) \\ &= \mathbb{P}(T_i < +\infty | X_0 = j) (1 + \mathbb{E}[N_i | X_0 = i]) \\ &< +\infty . \end{aligned}$$

Donc, presque sûrement, chaque état de E est visité un nombre fini de fois. C'est impossible : E étant fini, notre chaîne fréquente au moins l'un des états de E une infinité de fois. ■

1.2.3 Périodicité

Définition 1.2.11 La *période* d'un état i est l'entier $d(i)$ défini par

$$d(i) = \text{PGCD}\{n \geq 1, p_{i,i}^{(n)} > 0\} .$$

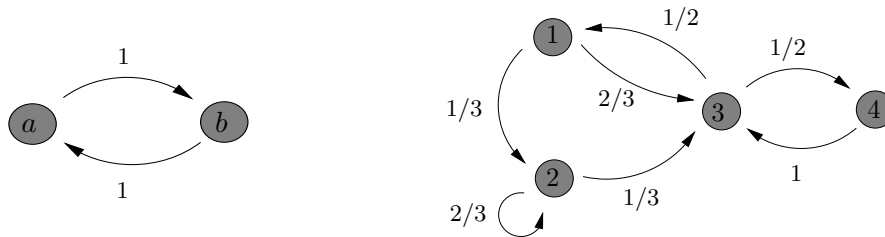
Lorsque $d(i) = 1$, l'état i est qualifié de *apériodique*.

Un état en lequel on peut rester avec probabilité non nulle, i.e. $p_{i,i} > 0$, est automatiquement apériodique.

Voici deux exemples. À gauche, l'état a est de période 2 (idem pour b). En effet, chaque chemin de probabilité non nulle partant de a et y revenant comprend un nombre pair de pas. À droite, tous les états sont apériodiques. Étudions en particulier le cas de l'état 1. Depuis cet état, il est possible d'y revenir en 3 pas (en faisant $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$) :

$$p_{1,1}^{(3)} = \sum_{i,j \in E} p_{1,i} p_{i,j} p_{j,1} \geq p_{1,2} p_{2,3} p_{3,1} > 0 .$$

Il est également possible d'y revenir en 5 pas : $p_{1,1}^{(5)} > 0$ (en faisant $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 3 \rightarrow 1$). Puisque 3 et 5 sont premiers entre eux, l'état 1 est apériodique.



La périodicité est une propriété de classe irréductible :

Proposition 1.2.12 Si les états i et j communiquent alors ils ont la même période.

Tous les états d'une même classe irréductible ont donc la même période. Si celle-ci vaut 1, la classe est alors qualifiée de apériodique.

Démonstration Soient i et j deux états qui communiquent. Il suffit de montrer que $d(j)$ divise $d(i)$. En effet, par symétrie, on aura également $d(i)$ divise $d(j)$, et donc $d(j) = d(i)$. Comme i et j communiquent, il existe deux entiers l et m tels que $p_{i,j}^{(l)} > 0$ et $p_{j,i}^{(m)} > 0$. Considérons maintenant un entier n tel que $p_{i,i}^{(n)} > 0$. Les inégalités

$$p_{j,j}^{(m+n+l)} \geq p_{j,i}^{(m)} p_{i,i}^{(n)} p_{i,j}^{(l)} > 0$$

et

$$p_{j,j}^{(m+l)} \geq p_{j,i}^{(m)} p_{i,j}^{(l)} > 0$$

impliquent respectivement que $d(j)$ divise les entiers $m + n + l$ et $m + l$: il divise donc la différence, i.e. n . Autrement dit, $d(j)$ divise tous les entiers n tels que $p_{i,i}^{(n)} > 0$. Il divise donc leur PGCD $d(i)$. ■

1.2.4 Simulation

Simuler une chaîne de Markov est très facile. Voici le code **R** de la fonction `iteration` qui prend en entrée un état $x \in E = \{1, 2, 3, 4\}$ et renvoie l'état visité juste après conformément à la dynamique décrite par le graphe de transition ci-dessus à droite.

```
iteration <- function(X)
{ U <- runif(1,min=0,max=1)
  if (X==a) { if (U < 1/3) {Y=b} else {Y=c}}
  if (X==b) { if (U < 2/3) {Y=b} else {Y=c}}
  if (X==c) { if (U < 1/2) {Y=a} else {Y=d}}
  if (X==d) {Y=c}
  Y
}
```

Il suffit alors de répéter n fois cette procédure pour obtenir une réalisation de X_n sachant $X_0 = x$. Une boucle `for` fera parfaitement l'affaire.

1.3 Théorèmes limites

Étant donné une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$, l'objectif de cette dernière partie est d'approximer la loi de la v.a. X_n lorsque n tend vers l'infini.

1.3.1 Les mesures stationnaires

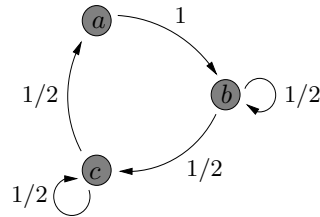
Définition 1.3.1 Une *mesure stationnaire* (ou *invariante*) d'une chaîne de Markov de matrice de transition P est une loi de probabilité sur E , disons $\pi = (\pi(j))_{j \in E}$, vérifiant $\pi = \pi P$.

Soit π une mesure stationnaire pour la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ de matrice de transition P . Rappelons que le vecteur ligne $(\mu_n(j))_{j \in E}$ désigne la loi de la v.a. X_n . La formule $\mu_{n+1} = \mu_n P$ (Proposition 1.1.6) implique que si la loi de X_n est π (i.e. $\mu_n = \pi$) alors il en est de même pour la loi de X_{n+1} (i.e. $\mu_{n+1} = \pi$). Par conséquent, si la loi initiale μ_0 (celle de X_0) est π alors toutes les v.a. X_n seront distribuées selon π . C'est ce qui justifie le qualificatif de stationnaire. Cela signifie que la probabilité de se trouver dans un état donné reste constante au cours du temps, bien que la chaîne saute constamment d'état en état. Une mesure stationnaire doit donc être comprise comme un équilibre dynamique en loi.

EXEMPLE :

Considérons une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ à espace d'états $E = \{a, b, c\}$ dont la matrice et le graphe de transition sont donnés par :

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



Réolvons le système linéaire $\pi = \pi P$ où les coordonnées du vecteur π sont notées (π_a, π_b, π_c) . On obtient une droite vectorielle de solutions, engendrée par le vecteur $(1/2, 1, 1)$. Par ailleurs, la mesure stationnaire π est une mesure de probabilité. Elle doit donc satisfaire la condition $\pi_a + \pi_b + \pi_c = 1$. Ceci détermine π de manière unique ; $\pi = (1/5, 2/5, 2/5)$.

Attention, contrairement à l'exemple ci-dessus, une chaîne de Markov peut ne pas admettre de mesure stationnaire, ou même en admettre plusieurs...

Considérons enfin une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ à espace d'états E fini et de matrice de transition P . Supposons de plus que la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une mesure de probabilité sur E , notée ν :

$$\text{pour tout } j \in E, \quad \lim_{n \rightarrow +\infty} \mu_n(j) = \nu(j) .$$

Puisque E est fini, l'application linéaire tP est continue donc ${}^tP^t \mu_n$ converge (coordonnée par coordonnée) vers ${}^tP^t \nu$, lorsque n tend vers l'infini. D'autre part, ${}^tP^t \mu_n = {}^t\mu_{n+1}$ converge également (coordonnée par coordonnée) vers la loi ${}^t\nu$. Celle-ci vérifie donc ${}^tP^t \nu = {}^t\nu$, ou encore $\nu = \nu P$. En conclusion, si la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une mesure de probabilité alors cette loi limite est nécessairement une mesure stationnaire pour la chaîne de Markov correspondante.

1.3.2 Les mesures réversibles

Définition 1.3.2 Une *mesure réversible* d'une chaîne de Markov de matrice de transition P est une loi de probabilité sur E , disons $\pi = (\pi(i))_{i \in E}$, vérifiant

$$\forall i, j \in E, \pi(i)p_{i,j} = \pi(j)p_{j,i} . \quad (1.2)$$

Proposition 1.3.3 Toute mesure réversible d'une chaîne de Markov est stationnaire pour cette même chaîne.

Démonstration Soient $j \in E$ et $\pi = (\pi(i))_{i \in E}$ une mesure réversible d'une chaîne de Markov de matrice de transition P . Dès lors,

$$\begin{aligned} \sum_{i \in E} \pi(i)p_{i,j} &= \sum_{i \in E} \pi(j)p_{j,i} \\ &= \pi(j) \sum_{i \in E} p_{j,i} \\ &= \pi(j) \end{aligned}$$

car la somme des coefficients d'une même ligne de P vaut 1. Il vient $\pi P = \pi$; π est donc stationnaire pour la chaîne de Markov de matrice de transition P . ■

Attention la réciproque est fautive : la mesure stationnaire $\pi = (1/5, 2/5, 2/5)$ identifiée dans la section précédente n'est pas réversible. En effet,

$$\pi_a p_{a,b} = \frac{1}{5} \times 1 \text{ alors que } \pi_b p_{b,a} = 0 .$$

Cependant, en pratique (et notamment pour les projets), pour obtenir une mesure stationnaire, il est recommandé de commencer par chercher une mesure réversible, plus facile à identifier quand elle existe.

1.3.3 Lorsque l'espace d'états E est fini

Considérons une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ à espace d'états fini E et de matrice de transition P . Comme la somme des coefficients sur chaque ligne de P vaut 1, tout vecteur à coordonnées constantes est vecteur propre de P associé à la valeur propre 1. Une matrice carrée et sa transposée ayant les mêmes valeurs propres, 1 est donc également valeur propre de tP . Notons E_1 l'espace propre associé à la valeur propre 1 pour tP :

$$E_1 = \{v \in \mathbb{R}^{\text{Card}(E)}, {}^tPv = v\} .$$

On vient de montrer que $\dim E_1 \geq 1$. Dès lors, tout vecteur (colonne) ${}^t\pi$ appartenant à E_1 et dont la somme des coordonnées vaut 1, produit un vecteur ligne π , loi de probabilité sur E et vérifiant $\pi = \pi P$. C'est donc une mesure stationnaire pour la chaîne de Markov de matrice de transition P . En conclusion, l'existence d'au moins une mesure stationnaire est automatique lorsque l'espace d'états E est fini.

Lorsque la chaîne est irréductible (dans ce cas, tous les états sont récurrents), l'espace propre E_1 est de dimension 1, ce qui implique l'unicité de la mesure stationnaire.

Proposition 1.3.4 Considérons une chaîne de Markov homogène à espace d'états fini E et de matrice de transition P . Elle admet au moins une mesure stationnaire. Si la chaîne est irréductible alors il y a unicité de la mesure stationnaire (notons-la π). Dans ce cas, elle charge tous les états

$$\pi(i) > 0, \text{ pour tout } i \in E$$

et le temps moyen de retour en i est égal à l'inverse de $\pi(i)$:

$$\mathbb{E}[T(i) \mid X_0 = i] = \frac{1}{\pi(i)}, \text{ pour tout } i \in E.$$

Cependant, l'existence et l'unicité de la mesure stationnaire π n'assure pas la convergence en loi de la suite $(X_n)_{n \in \mathbb{N}}$ vers π . Cette convergence repose en fait sur l'apériodicité de la chaîne. Le résultat suivant est non trivial. Il ne sera pas démontré.

Théorème 1.3.5 Considérons une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ à espace d'états fini E et de matrice de transition P . Supposons de plus que la chaîne est irréductible et apériodique. Notons π sa mesure stationnaire.

- (1) La matrice P^n converge quand n tend vers l'infini vers une matrice dont toutes les lignes sont égales à π .
- (2) Quelle que soit la loi de X_0 , la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi quand n tend vers l'infini vers π : pour tout état i , $\mathbb{P}(X_n = i) \rightarrow \pi(i)$.

Les points (1) et (2) s'interprètent par de l'indépendance asymptotique : en deux instants éloignés l'un de l'autre, la chaîne se comporte de manière presque indépendante. En effet, les quantités $\mathbb{P}(X_n = j \mid X_0 = i)$ et $\mathbb{P}(X_n = j)$ tendent toutes les deux vers la même limite $\pi(j)$. Elles sont donc proches asymptotiquement, ce qui donne

$$\mathbb{P}(X_n = j, X_0 = i) \simeq \mathbb{P}(X_n = j) \mathbb{P}(X_0 = i).$$

En utilisant l'invariance par translation dans le temps, la relation ci-dessus se généralise en

$$\mathbb{P}(X_{n+m} = j, X_m = i) \simeq \mathbb{P}(X_{n+m} = j) \mathbb{P}(X_m = i).$$

Considérons une nouvelle fois l'exemple du bit informatique dont la matrice de transition est donnée par

$$P = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix}.$$

La chaîne de Markov correspondante est irréductible si et seulement si $p < 1$. Dans ce cas, l'unique mesure stationnaire est le vecteur ligne $\pi = (1/2, 1/2)$. Si de plus $p > 0$, alors la chaîne est apériodique. Le théorème limite s'applique :

$$\forall i \in \{0, 1\}, \lim_{n \rightarrow +\infty} \mathbb{P}(X_n = i) = \frac{1}{2}.$$

Autrement dit, quelque soit le bit initial X_0 et quelle que soit la valeur de $0 < p < 1$, au bout d'un certain temps, le bit sortant suit une loi uniforme sur $\{0, 1\}$.

Lorsque $p = 0$, la matrice de transition devient

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

La chaîne est alors 2-périodique. Puisque P^2 est égale à la matrice identité, il en va de même pour toutes les puissances paires de P , tandis que les puissances impaires sont égales à P . Par conséquent, le coefficient $p_{0,0}^{(n)}$ prend alternativement les valeurs 0 et 1 : la suite $(p_{0,0}^{(n)})_{n \in \mathbb{N}}$ ne peut converger.

Terminons par une comparaison. La Théorème 1.3.5 stipule que, sous les bonnes hypothèses, $p_{i,j}^{(n)}$ tend vers une limite strictement positive (Proposition 1.3.4). Ceci concerne le cas où j est récurrent. Lorsque j est transient, $p_{i,j}^{(n)}$ tend vers 0. En effet, l'espérance conditionnelle $\mathbb{E}[N_j | X_0 = i]$ est finie (voir la démonstration de la Proposition 1.2.10). Par conséquent, la série

$$\begin{aligned} \sum_{n \geq 1} p_{i,j}^{(n)} &= \sum_{n \geq 1} \mathbb{P}(X_n = j \mid X_0 = i) \\ &= \sum_{n \geq 1} \mathbb{E}[\mathbb{1}_{X_n=j} \mid X_0 = i] \\ &= \mathbb{E}[N_j \mid X_0 = i] \end{aligned}$$

est convergente. Son terme général, i.e. $p_{i,j}^{(n)}$, tend donc vers 0 quand n tend vers l'infini.

1.3.4 Lorsque l'espace d'états E est infini

L'étude théorique de la convergence en loi d'une chaîne de Markov devient plus délicate lorsque l'espace d'états E est infini. À plus forte raison que dans la partie précédente, les résultats qui suivent ne seront pas démontrés.

Considérons une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ irréductible et récurrente, à espace d'états infini E et de matrice de transition P . On peut alors montrer qu'il existe un vecteur colonne $v \in \mathbb{R}^E$, à coordonnées strictement positives, vérifiant ${}^t P v = v$. Si la somme des coordonnées de v est finie

$$M := \sum_{j \in E} v(j) < +\infty, \quad (1.3)$$

alors il suffit de renormaliser le vecteur v par sa masse totale M pour obtenir une mesure stationnaire (qui sera unique). Cependant la condition (1.3) n'est pas toujours vérifiée. Une hypothèse supplémentaire sur la chaîne doit être faite pour assurer l'existence (et l'unicité) d'une mesure stationnaire.

Définition 1.3.6 Considérons une chaîne de Markov homogène irréductible et récurrente à espace d'états infini E . Soit $i \in E$. Le temps de retour $T(i)$ à i est fini presque sûrement (car i est récurrent). Si, de plus,

$$\mathbb{E}[T(i) | X_0 = i] < +\infty,$$

l'état i est qualifié de *récurrent positif*. Dans le cas contraire, i est qualifié de *récurrent nul*.

Être récurrent nul signifie que l'on revient presque sûrement en temps fini en cet état mais que le temps de retour est plutôt long. Si long que son espérance est infinie.

La marche aléatoire simple et symétrique (i.e. avec $p = 1/2$) sur \mathbb{Z} est un exemple de chaîne de Markov pour laquelle tous les états sont récurrents nuls.

Le théorème suivant établit que la récurrence positive est une propriété de classe (même chose pour la récurrence nulle) : c'est une condition nécessaire et suffisante pour l'existence d'une mesure stationnaire. Si c'est le cas, il y a unicité de la mesure stationnaire.

Théorème 1.3.7 Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène irréductible, à espace d'états infini E . Sont équivalentes :

- (i) tous les états sont récurrents positifs ;
- (ii) il existe au moins un état récurrent positif ;
- (iii) la chaîne $(X_n)_{n \in \mathbb{N}}$ admet (au moins) une mesure stationnaire.

Lorsque ces conditions sont vérifiées, il y a unicité de la mesure stationnaire. Notons-la π . Elle est définie par :

$$\forall i \in E, \frac{1}{\pi(i)} = \mathbb{E}[T(i) \mid X_0 = i] .$$

Si de plus, la chaîne $(X_n)_{n \in \mathbb{N}}$ est apériodique alors :

- (1) la matrice P^n converge quand n tend vers l'infini vers une matrice dont toutes les lignes sont égales à π ;
- (2) quelle que soit la loi de X_0 , la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi quand n tend vers l'infini vers π : pour tout état i , $\mathbb{P}(X_n = i) \rightarrow \pi(i)$.

1.4 Exercices

Exercice 1 : Transmission d'un bit informatique

Un bit informatique valant 0 ou 1 est transmis d'un poste A vers un poste B en passant par N intermédiaires. Chacun de ces intermédiaires transmet correctement le bit avec probabilité p et l'inverse avec probabilité $1 - p$, indépendamment les uns des autres. Le bit (aléatoire) d'entrée, disons X_0 , est supposé indépendant des intermédiaires. Pour $n = 1, \dots, N$, notons X_n le bit sortant du n^{e} intermédiaire.

1. Il a été vérifié en cours que la suite de v.a. $(X_n)_{0 \leq n \leq N}$ est une chaîne de Markov homogène. Rappeler sa matrice et son graphe de transition.

2. Calculer la probabilité que le bit arrive correct en B. Que vaut cette probabilité lorsque N tend vers l'infini ? En déduire la loi de X_N lorsque N tend vers l'infini.

Commentaires : le théorème limite 1.3.5 permet de retrouver rapidement ces résultats.

Exercice 2 : Jeu de l'oie

Soit le petit jeu de l'oie suivant. Les cases sont 0, 1, 2, 3 et forment un cycle. Sur la case 3, il est écrit "au prochain tour, reculez de deux cases". On part de la case 0 et on gagne dès que l'on y retourne. À chaque étape, on lance une pièce symétrique : on avance d'une case si c'est face et de deux si c'est pile. On note X_n la suite des états (le joueur continue de jouer après avoir gagné).

1. Montrer que la suite $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov homogène issue de 0 et donner sa matrice de transition.

2. Étudier la convergence en loi de la chaîne.

3. En déduire le nombre moyen de coups joués pour gagner.

Exercice 3 : Gestion de stock

Dans un magasin, on souhaite suivre l'évolution au cours du temps du stock d'un article donné, i.e. le nombre d'exemplaires de cet article en réserve dans le magasin. Notons X_n ce nombre au matin du n^{e} jour. Ce jour-là, D_{n+1} exemplaires de cet article sont demandés. Le stock évolue comme suit. Si $X_n - D_{n+1} \geq s$ alors on ne réapprovisionne pas. Dans le cas contraire, on réapprovisionne à hauteur de S . Les entiers s et S sont les niveaux minimal et maximal du stock. La suite de v.a. $(D_n)_{n \geq 1}$ est i.i.d. et le niveau initial du stock X_0 peut être supposé indépendant des $(D_n)_{n \geq 1}$.

1. Montrer que la suite $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov homogène et donner son graphe de transition (du moins, les arêtes issues d'un état $s \leq x \leq S$).

2. Supposons désormais que $s = 1$, $S = 2$ et que la loi commune des $(D_n)_{n \geq 1}$ est uniforme sur $\{0, 1, 2\}$. Donner la matrice et le graphe de transition. Quelle est la probabilité d'être en rupture de stock à long terme ?

Exercice 4 (Examen GIS 2009) :

Considérons la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ à valeurs dans $E = \{a, b, c, d\}$ et dont la matrice de transition est donnée par :

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

1. Représenter le graphe de transition de la chaîne $(X_n)_{n \in \mathbb{N}}$ et montrer qu'elle est irréductible.
2. Calculer le coefficient $p_{d,d}^{(n)}$ pour $n = 1, 2, 3, 4$.
3. La série $\sum p_{d,d}^{(n)}$ est-elle convergente ou divergente ?
4. Étudier la convergence en loi de la chaîne. En déduire le temps moyen de retour en l'état d .

Exercice 5 (Examen GIS 2013) :

Le jeu vidéo *Dario Cros* contient 4 niveaux. À partir du niveau x , pour $x = 1, 2, 3$, on accède au niveau $x + 1$ avec probabilité $\frac{1}{2}$, sinon on retourne au niveau 1. On réussit le niveau 4 avec probabilité $\frac{1}{2}$: c'est la victoire finale de *Dario Cros* ! Sinon on retourne au niveau 1 et tout est à recommencer...

1. Modéliser ce jeu vidéo par une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ à valeurs dans un espace E à 5 états (les 4 niveaux et l'état VICTOIRE). Pour cette modélisation, nous supposons que l'état VICTOIRE mène systématiquement au niveau 1. Préciser l'unité de temps. Donner le graphe de transition et la matrice de transition P .

Dans la suite, on s'intéresse au temps T , le temps d'atteinte de l'état VICTOIRE :

$$T := \inf\{n \geq 1, X_n = \text{VICTOIRE}\}.$$

2. Calculer P^2 . En déduire la loi de X_2 sachant $X_0 = 1$, puis sachant $X_0 = 3$.
3. Déterminer les classes irréductibles de la chaîne $(X_n)_{n \in \mathbb{N}}$. En partant du niveau 1, est-il possible de jouer indéfiniment sans jamais gagner (i.e. sans jamais parvenir à l'état VICTOIRE) ? Vous pouvez utiliser la formule

$$\mathbb{P}(T < +\infty \mid X_0 = 1) = \mathbb{P}(T < +\infty \mid X_0 = \text{VICTOIRE})$$

(qui rapporte 1.5 points en bonus si vous la démontrez).

4. Déterminer le comportement asymptotique de la chaîne $(X_n)_{n \in \mathbb{N}}$. Lorsque n est très grand, a-t-on plus de chance de trouver X_n dans l'état 1 ou dans $E \setminus \{1\}$? Est-ce surprenant ?
5. Sachant $X_0 = \text{VICTOIRE}$, quelle est l'espérance de T ? En déduire le temps moyen qu'il faudra à *Dario Cros*, partant du niveau 1, pour parvenir à la victoire finale.

Exercice 6 : La marche aléatoire simple

Soient $(Y_n)_{n \geq 1}$ une suite de v.a. i.i.d., chacune valant 1 avec probabilité p et -1 avec probabilité $1 - p$. Soit $Y_0 \in \mathbb{Z}$ une v.a. indépendante des Y_n , représentant le point de départ sur l'axe \mathbb{Z} de la chaîne. On définit la marche aléatoire simple par

$$\begin{aligned} X_{n+1} &= X_n + Y_{n+1} \\ &= Y_0 + Y_1 + \dots + Y_n + Y_{n+1}, \end{aligned}$$

pour tout $n \in \mathbb{N}$. Il a été vérifié que le processus stochastique $(X_n)_{n \geq 1}$ est une chaîne de Markov homogène sur $E = \mathbb{Z}$.

1. Montrer que, pour tout n , $p_{0,0}^{(2n+1)} = 0$ et $p_{0,0}^{(2n)} = C_{2n}^n p^n q^n$. Puis, donner un équivalent de $p_{0,0}^{(2n)}$ en utilisant la formule de Stirling :

$$\exists A > 0, n! \sim A\sqrt{n} \left(\frac{n}{e}\right)^n.$$

2. Conclure quant à la récurrence ou la transience de la chaîne en fonction du paramètre p . Interpréter.

Exercice 7 : Existence de plusieurs mesures stationnaires

Considérons la chaîne de Markov $(X_n)_{n \geq 0}$ à valeurs dans $E = \{a, b, c, d\}$ et dont la matrice de transition est donnée par :

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

1. Déterminer les classes irréductibles. Lesquelles sont transientes, récurrentes ?
2. Montrer que la matrice P^n converge (coefficient par coefficient) vers une matrice P^∞ que l'on déterminera.
3. En déduire l'existence de plusieurs mesures stationnaires (associées à P).

Exercice 8 (Examen 2015 GIS2A) :

On suppose qu'un trait est gouverné par deux gènes, pouvant être chacun de deux types, G et g . Les trois paires de gènes possibles sont donc GG , Gg et gg . Le type G est dominant par rapport au type g , qui sera qualifié de récessif. Cela signifie que dans la paire Gg , c'est le gène G qui s'exprime. Les paires GG , Gg et gg sont respectivement dites dominante, hybride et récessive.

Hypothèse d'appariement. Soient $A, B, C, D \in \{G, g\}$. L'appariement de la paire AB avec CD produit avec probabilité $1/4$ l'une des paires AC , AD , BC ou BD (par un tableau de croisement).

1. Un éleveur adopte la stratégie suivante : à chaque fois, il apparie l'individu de la n^e génération avec une paire hybride.

(a) Modéliser la situation par une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ où X_n représente la paire de gènes de l'individu de la n^e génération. Donner son graphe de transition et sa matrice de transition P .

(b) Déterminer les classes irréductibles de la chaîne, la nature (transiente ou récurrente) et la période de chacun des états.

(c) Déterminer le comportement asymptotique de la chaîne $(X_n)_{n \in \mathbb{N}}$.

2. Effectuer la même étude lorsque l'éleveur adopte une seconde stratégie ; à chaque fois, il apparie l'individu de la n^e génération avec une paire dominante. Comparer qualitativement les résultats des deux stratégies au bout d'un temps long.

Chapitre 2

Processus stochastiques markoviens en temps continu

Un modèle d'évolution dynamique *en temps continu* dans lequel on fait dépendre l'évolution future de l'état présent et du hasard est un processus de Markov. On en rencontre dans de nombreux domaines d'applications, comme par exemple l'étude des files d'attente.

2.1 Processus de Poisson

2.1.1 Trois propriétés de la loi exponentielle

La loi exponentielle est l'ingrédient de base pour modéliser des temps d'attente d'événements “imprévisibles”. Cette partie contient quelques-unes de ses propriétés élémentaires.

Soit X une v.a. suivant la loi exponentielle de paramètre λ , notée $\mathcal{E}(\lambda)$. Sa densité est :

$$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{x \geq 0} .$$

Elle est sans atome, i.e. $\forall x, \mathbb{P}(X = x) = 0$. Sa fonction de répartition vaut

$$F_X(x) = \mathbb{P}(X \leq x) = (1 - e^{-\lambda x}) \mathbb{1}_{x \geq 0} , \quad (2.1)$$

son espérance et sa variance valent respectivement $1/\lambda$ et $1/\lambda^2$. De l'expression (2.1), on déduit qu'à partir d'une v.a. X de loi exponentielle de paramètre λ , on obtient une v.a. Y de loi exponentielle de paramètre 1 en posant simplement $Y = \lambda X$.

En pratique, une v.a. de loi exponentielle représente une durée, typiquement le temps d'attente d'un événement ou une durée de vie. La propriété importante des lois exponentielles est d'être “sans mémoire”. Dans le cas particulier d'un composant électronique dont la durée de vie serait modélisée par une loi exponentielle, cela signifie que la probabilité pour que le composant vive un temps t est la même, qu'il soit neuf ou qu'il ait déjà vécu un temps s . Cette absence de mémoire est caractéristique des lois exponentielles.

Proposition 2.1.1 Une v.a. X à valeurs dans \mathbb{R}_+ et de fonction de répartition continue suit une loi exponentielle si et seulement si pour tous réels $s, t \geq 0$,

$$\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t) . \quad (2.2)$$

Démonstration Si X suit la loi $\mathcal{E}(\lambda)$ alors, d'après (2.1),

$$\begin{aligned}\mathbb{P}(X > s+t \mid X > s) &= \frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} = \mathbb{P}(X > t) .\end{aligned}$$

Réciproquement, en notant $g(t) = 1 - F_X(t)$, l'équation (2.2) s'écrit $g(s+t) = g(s)g(t)$, pour tous réels $s, t \geq 0$. Un petit travail permet de montrer que les solutions continues de cette équation sont de la forme $g(t) = e^{at}$, i.e. $F_X(t) = 1 - e^{at}$. Enfin, le fait que $F_X(t)$ tende vers 1 quand t tend vers l'infini force le réel a à être strictement négatif. La v.a. X suit donc la loi exponentielle de paramètre $-a$. ■

Proposition 2.1.2 Considérons n v.a. indépendantes X_1, \dots, X_n de lois respectives $\mathcal{E}(\lambda_1), \dots, \mathcal{E}(\lambda_n)$. Posons $Y = \min\{X_1, \dots, X_n\}$. Alors Y suit la loi exponentielle de paramètre $\lambda_1 + \dots + \lambda_n$ et pour tout indice $i = 1, \dots, n$,

$$\mathbb{P}(Y = X_i) = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n} .$$

Démonstration Pour identifier la loi de Y , le plus simple est de calculer sa fonction de répartition :

$$\begin{aligned}\forall t \geq 0, \mathbb{P}(Y > t) &= \mathbb{P}(\forall i = 1, \dots, n, X_i > t) \\ &= \prod_{i=1}^n \mathbb{P}(X_i > t) \\ &= e^{-(\lambda_1 + \dots + \lambda_n)t} .\end{aligned}$$

On conclut en utilisant le fait qu'une fonction de répartition caractérise la loi.

L'événement $Y = X_i$ équivaut à $X_i \leq X_j, \forall j$. De plus, l'indépendance entre les v.a. X_1, \dots, X_n nous permet d'écrire :

$$\begin{aligned}\mathbb{P}(Y = X_i) &= \int_{\{(x_1, \dots, x_n), x_i \leq x_j \forall j\}} \left(\prod_{1 \leq j \leq n} \lambda_j e^{-\lambda_j x_j} \mathbb{1}_{x_j \geq 0} \right) dx_1 \dots dx_n \\ &= \int_0^\infty \prod_{j \neq i} \left(\int_{x_i}^\infty \lambda_j e^{-\lambda_j x_j} dx_j \right) \lambda_i e^{-\lambda_i x_i} dx_i \\ &= \int_0^\infty \lambda_i e^{-(\lambda_1 + \dots + \lambda_n)x_i} dx_i \\ &= \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n} .\end{aligned}$$

■

Proposition 2.1.3 La somme de n v.a. indépendantes de loi exponentielle de paramètre λ suit une loi, appelée loi Gamma de paramètres n et λ et notée $\Gamma(n, \lambda)$, dont la densité est donnée par :

$$f(x) = \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} \mathbb{1}_{x \geq 0}.$$

2.1.2 Présentation du processus de Poisson

Définition 2.1.4 Soit $(X_n)_{n \geq 1}$ une suite de v.a. indépendantes et de même loi, exponentielle de paramètre λ . Posons $S_0 = 0$ et pour tout entier $n \geq 1$, $S_n = X_1 + \dots + X_n$. Pour tout réel $t \geq 0$, définissons la v.a. N_t , à valeurs entières, par :

$$N_t = n \Leftrightarrow S_n \leq t < S_{n+1}.$$

Le processus stochastique $\{N_t\}_{t \geq 0}$ est appelé *processus de Poisson d'intensité λ* .

Le processus de Poisson est un modèle de comptage d'événements aléatoires isolés dans le temps, comme des "tops" d'horloge séparés par des durées aléatoires. Dans ce modèle :

- X_n est la durée séparant le $(n-1)^e$ top du n^e ;
- S_n est la date à laquelle survient le n^e top. D'après la Proposition 2.1.3, la v.a. S_n suit la loi Gamma de paramètres n et λ ;
- N_t est le nombre de tops comptés entre l'instant 0 et l'instant t :

$$N_t = \sum_{n \geq 1} \mathbb{1}_{S_n \leq t}.$$

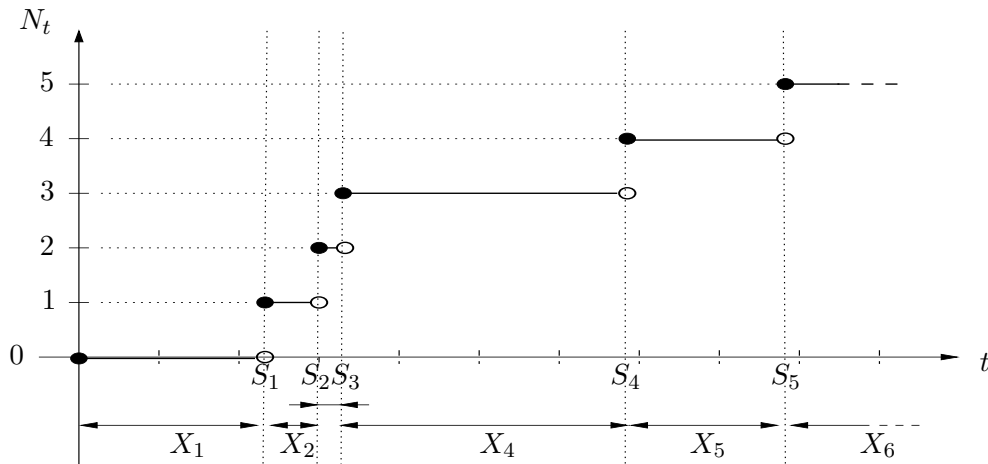


FIGURE 2.1 – Ci-dessus est représentée la trajectoire simulée d'un processus de Poisson d'intensité $\lambda = 1/2$ sur l'intervalle $[0; 10]$. On constate que d'assez longues périodes sans aucun top alternent avec des rafales de tops rapprochés : il n'y a pas d'autre explication à la fameuse "loi des séries" chère aux journalistes. Si des événements (comme par exemple des accidents d'avion) arrivent rarement, de manière imprévisible et indépendante, on ne peut pas imaginer qu'ils surviennent à des intervalles de temps réguliers. Il faut plutôt s'attendre à les voir survenir parfois de manière rapprochée, par "séries".

Soit $0 \leq s \leq t$ des réels. La différence $N_t - N_s$ est appelée *accroissement*. Elle est égale au nombre de tops entre les instants s et t .

Théorème 2.1.5 Considérons un processus de Poisson $\{N_t\}_{t \geq 0}$ d'intensité λ . Il vérifie les propriétés suivantes :

- (P1) $\{N_t\}_{t \geq 0}$ est un *processus de comptage* ; il est à valeurs entières, vérifie $N_0 = 0$ p.s. et pour tous réels $0 \leq s \leq t$, $N_s \leq N_t$.
- (P2) $\{N_t\}_{t \geq 0}$ est un *processus à accroissements indépendants* ; pour tout entier k et pour toute suite d'instants $0 \leq t_1 < t_2 < \dots < t_k$, les accroissements $N_{t_2} - N_{t_1}, \dots, N_{t_k} - N_{t_{k-1}}$ sont des v.a. indépendantes.
- (P3) Les accroissements du processus $\{N_t\}_{t \geq 0}$ sont poissonniens ; pour tous réels $0 \leq s < t$, la v.a. $N_t - N_s$ suit la loi de Poisson de paramètre $\lambda(t - s)$;
- (P4) $\{N_t\}_{t \geq 0}$ est un processus *homogène* ou à *accroissements stationnaires* ; pour tous instants $0 \leq t_1 < t_2$ et $s \geq 0$, la v.a. $N_{t_2+s} - N_{t_1+s}$ suit la même loi que $N_{t_2} - N_{t_1}$;
- (P5) $\{N_t\}_{t \geq 0}$ est un *processus d'événements rares* ; $\mathbb{P}(N_{t+h} - N_t \geq 2) = o(h)$.

Nous ne démontrerons pas ces résultats et nous nous contenterons de quelques commentaires.

Par définition, un processus de Poisson est une fonction du hasard et du temps, à valeurs entières, qui est croissante dans le temps. De plus,

$$\mathbb{P}(N_0 = 0) = 1 - \mathbb{P}(N_0 \neq 0) \geq 1 - \mathbb{P}(X_1 = 0) = 1 .$$

C'est donc un processus de comptage.

La démonstration de la propriété (P2) est liée à l'absence de mémoire de la loi exponentielle (Proposition 2.1.1). Elle signifie que les nombres de tops comptés dans des intervalles de temps disjoints sont indépendants. En particulier, les nombres de tops dans les intervalles $[0; t]$ et $]t; t + s]$ sont indépendants. C'est la propriété (P3) qui est à l'origine du nom du processus de Poisson. La probabilité pour qu'il y ait n tops entre les instants s et t (avec $s < t$) est donnée par :

$$\mathbb{P}(N_t - N_s = n) = \frac{(\lambda(t - s))^n}{n!} e^{-\lambda(t-s)} .$$

En particulier,

$$\mathbb{E}[N_t - N_s] = \lambda(t - s) .$$

Le nombre moyen de tops comptés dans un intervalle de temps est donc proportionnel à la longueur de cet intervalle et le coefficient de proportionnalité n'est autre que λ . Par conséquent, le nombre moyen de tops par unité de temps est égal à λ , fort justement appelé *intensité* du processus. La durée moyenne entre deux tops étant de $1/\lambda$ (ce sont des exponentielles de paramètre λ), il est normal de compter λ tops par unité de temps en moyenne.

Avec la condition $N_0 = 0$ p.s., le nombre de tops jusqu'à l'instant t est distribué selon la loi de Poisson de paramètre λt :

$$N_t \sim \mathcal{P}(\lambda t) .$$

La propriété (P4) est une conséquence immédiate de (P3) : les v.a. $N_{t_2+s} - N_{t_1+s}$ et $N_{t_2} - N_{t_1}$ suivent toutes les deux la loi de Poisson de paramètre $\lambda((t_2 + s) - (t_1 + s)) = \lambda(t_2 - t_1)$.

La propriété (P5) signifie que la probabilité de compter plus d'un top dans un petit intervalle de temps est négligeable. Elle découle également de la propriété (P3). Rappelons tout d'abord que $f(h) = o(h)$ signifie que le rapport $f(h)/h$ tend vers 0 quand h tend vers 0. La notation $o(h)$ a un caractère infinitésimal (et en aucun cas global). Puisque

$$\mathbb{P}(N_{t+h} - N_t = 0) = e^{-\lambda h} = 1 - \lambda h + o(h)$$

et

$$\mathbb{P}(N_{t+h} - N_t = 1) = (\lambda h)e^{-\lambda h} = \lambda h + o(h) ,$$

on en déduit que

$$\mathbb{P}(N_{t+h} - N_t \geq 2) = 1 - \mathbb{P}(N_{t+h} - N_t = 0) - \mathbb{P}(N_{t+h} - N_t = 1) = o(h) .$$

Réciproquement, un processus stochastique vérifiant les propriétés **(P1)**, **(P2)**, **(P4)** et **(P5)** est un processus de Poisson. Il vérifie donc également la propriété **(P3)**.

Les instants de saut d'un processus de Poisson vérifient la propriété intéressante suivante : lorsqu'on sait qu'exactly n sauts ont eu lieu dans l'intervalle $[0; t]$, la distribution de ces sauts dans $[0; t]$ est celle de n points tirés uniformément au hasard dans cet intervalle.

Proposition 2.1.6 Conditionnellement en $\{N_t = n\}$, la loi du vecteur (S_1, \dots, S_n) des n premiers instants de saut d'un processus de Poisson est la même que celle du réordonnement par ordre croissant d'une suite de n v.a. indépendantes et de même loi uniforme sur $[0; t]$. Autrement dit, si on tire des v.a. U_1, \dots, U_n indépendantes de loi uniforme sur $[0; t]$ et si on les réordonne en posant

$$U_{(1)} = \min_{1 \leq i \leq n} U_i < U_{(2)} < \dots < U_{(n)} = \max_{1 \leq i \leq n} U_i ,$$

alors la loi conditionnelle de (S_1, \dots, S_n) sachant $N_t = n$ est la même que celle de $(U_{(1)}, \dots, U_{(n)})$.

Démonstration Nous démontrons uniquement le cas $n = 1$. Étant donnés des réels $0 \leq u \leq t$, on peut écrire :

$$\begin{aligned} \mathbb{P}(S_1 \leq u, N_t = 1) &= \mathbb{P}(N_u = 1, N_t = 1) \\ &= \mathbb{P}(N_u = 1, N_t - N_u = 0) \\ &= \mathbb{P}(N_u = 1) \mathbb{P}(N_t - N_u = 0) , \end{aligned}$$

du fait de l'indépendance des accroissements **(P2)**. Par la propriété **(P3)**, les v.a. N_u et $N_t - N_u$ suivent des lois de Poisson de paramètres respectifs λu et $\lambda(t - u)$ (où λ est l'intensité du processus). Ainsi,

$$\begin{aligned} \mathbb{P}(S_1 \leq u, N_t = 1) &= \mathbb{P}(N_u = 1) \mathbb{P}(N_t - N_u = 0) \\ &= \lambda u e^{-\lambda u} \times e^{-\lambda(t-u)} \\ &= \lambda u e^{-\lambda t} . \end{aligned}$$

Finalement,

$$\mathbb{P}(S_1 \leq u \mid N_t = 1) = \frac{u}{t} .$$

Autrement dit, la fonction de répartition de la v.a. S_1 conditionnellement à $\{N_t = 1\}$ est celle d'une loi uniforme sur l'intervalle $[0; t]$. ■

Terminons en donnant deux résultats sur la superposition de processus de Poisson indépendants et sur l'opération inverse de décomposition.

Proposition 2.1.7

1. Soient $\{N_t^1\}_{t \geq 0}$ et $\{N_t^2\}_{t \geq 0}$ deux processus de Poisson indépendants (i.e. tout événement lié à l'un est indépendant de tout événement lié à l'autre), d'intensités respectives λ_1 et λ_2 . On appelle *superposition* de ces deux processus le processus somme :

$$\forall t \geq 0, N_t = N_t^1 + N_t^2.$$

Alors $\{N_t\}_{t \geq 0}$ est encore un processus de Poisson d'intensité $\lambda_1 + \lambda_2$.

2. Inversement, si $\{N_t\}_{t \geq 0}$ est un processus de Poisson d'intensité λ , on associe aux instants de saut $(S_n)_{n \in \mathbb{N}}$ de ce processus une suite de v.a. $(Y_n)_{n \in \mathbb{N}}$ indépendantes, indépendantes du processus $\{N_t\}_{t \geq 0}$, et de même loi de Bernoulli :

$$\forall n \in \mathbb{N}, \mathbf{P}(Y_n = 1) = p \text{ et } \mathbf{P}(Y_n = 0) = 1 - p.$$

Soit $\{N'_t\}_{t \geq 0}$ le processus dont les instants de saut sont les $(S_n)_{n \in \mathbb{N}}$ tels que $Y_n = 1$. Alors, $\{N'_t\}_{t \geq 0}$ est encore un processus de Poisson d'intensité $\lambda' = p\lambda$.

Voici deux exemples d'application :

- Sur une voie à double sens, si les flux de véhicules venant de droite et de gauche peuvent être considérés comme deux processus de Poisson indépendants, d'intensité respectives λ_1 et λ_2 , alors le flux global de véhicules est encore un processus de Poisson d'intensité $\lambda_1 + \lambda_2$.
- Si le flux global de véhicules est un processus de Poisson d'intensité λ et que 5% des véhicules sont rouges alors le flux des véhicules rouges est encore un processus de Poisson d'intensité $\lambda' = 0,05\lambda$.

2.1.3 Simulation et modélisation

De nombreuses situations sont modélisées par les processus de Poisson : arrivées de clients dans une file d'attente, appels à un central téléphonique, désintégration de particules radioactives, pannes de composants électroniques, passages de voitures sur une route à grande circulation non embouteillée. . .

Pour se faire une idée concrète de ce qu'est un processus de Poisson, rien ne vaut la simulation. On peut bien sûr programmer en suivant la définition, ce qui implique des tirages successifs de variables exponentielles. On peut aussi appliquer la propriété (P3) et la Proposition 2.1.6 : pour simuler des événements poissonniens sur l'intervalle de temps $[0; t]$, on commence par choisir la valeur n prise par N_t suivant la loi $\mathcal{P}(\lambda t)$, puis on simule les dates d'arrivées de ces n événements en les tirant au hasard suivant la loi uniforme sur $[0; t]$. En **R**, cela donne :

```
n = rpois(1, λt)
v = runif(n, 0, t)
w = sort(v)
```

Le vecteur w contient les instants de sauts S_1, \dots, S_n du processus de Poisson d'intensité λt sur l'intervalle $[0; t]$.

Il est important de comprendre quelles hypothèses de modélisation l'utilisation d'un processus de Poisson sous-entend. Dire que des événements surviennent selon un processus de Poisson, c'est supposer qu'à chaque instant, l'arrivée du prochain événement est parfaitement imprévisible, i.e. indépendante de ce qui a précédé. On suppose de plus que l'intensité (nombre moyen d'événements par unité de temps) reste constante. Supposons que l'on souhaite modéliser des arrivées de clients dans un magasin. Les questions sont les suivantes.

- Les clients arrivent-ils un par un (événements rares) ?

- Le temps qui s'est écoulé depuis l'arrivée du dernier client est-il sans incidence sur les chances d'en voir arriver un bientôt (accroissements indépendants) ?
- Dans une seconde fixée, les chances de voir arriver un client sont-elles constantes (homogénéité en temps) ?

Même si les réponses à ces questions sont généralement plutôt négatives, il est raisonnable de commencer par modéliser cette situation par un processus de Poisson. Lorsque cette première modélisation conduit à des résultats aberrants, on a recours à des modèles plus sophistiqués qui sont des généralisations du processus de Poisson. Par exemple, les *processus de renouvellement* sont des processus de comptage d'événements rares séparés par des durées indépendantes, mais de loi autre qu'exponentielle. Les *processus de Poisson non homogènes* sont des processus de comptage, à accroissements indépendants, d'événements rares, mais la loi de $N_{t+h} - N_t$ dépend de t . L'intérêt des *processus de Poisson composés* réside dans la possibilité d'ajouter, à chaque top, une v.a. entière plutôt que d'augmenter systématiquement de 1. Ceci permet en particulier de tenir compte des arrivées groupées dans les modèles d'attente.

Signalons enfin que le processus de Poisson peut être étendu à des dimensions supérieures à 1. Les tops sont alors interprétés comme des points de l'espace, pouvant par exemple modéliser les individus d'une espèce protégée dans un parc naturel. Dans la fenêtre $[0; 20]^2$, le nombre de points d'un processus de Poisson (bi-dimensionnel) d'intensité λ suit la loi de Poisson de paramètre $\lambda \times 20^2$. Puis, en ces points, nous centrons des disques dont les rayons sont des v.a. i.i.d. L'union de ces disques forme un sous-ensemble aléatoire Γ , appelé *Modèle Booléen* ; voir la Figure 2.2. Le Modèle Booléen permet de modéliser de nombreux phénomènes physiques ou biologiques : la porosité d'un milieu ou d'un matériau, la propagation d'une maladie sur un territoire, etc.

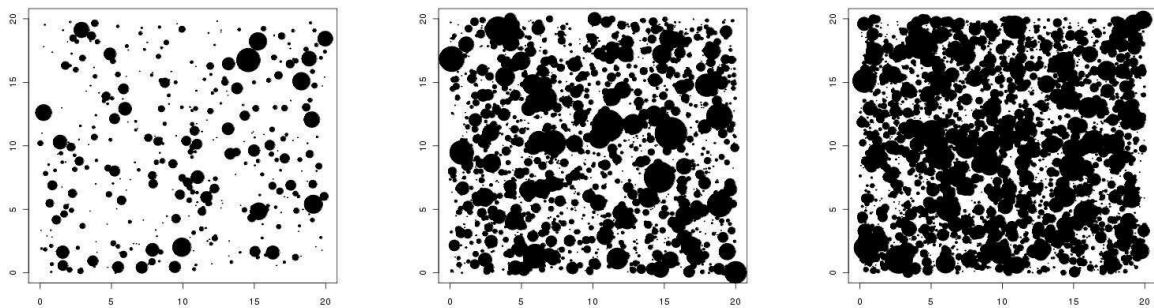


FIGURE 2.2 – Trois simulations du Modèle Booléen. Dans les trois cas, les rayons sont distribués selon la loi exponentielle de paramètre 1. L'intensité du processus de Poisson sous-jacent vaut, de gauche à droite, $\lambda = 1, 4, 7$. Il n'est donc pas étonnant que l'ensemble Γ (en noir) remplisse de plus en plus la fenêtre $[0; 20]^2$. Dans la dernière simulation, est-il possible de traverser la fenêtre $[0; 20]^2$, disons de gauche à droite, tout en restant dans Γ ?

Voici le code en **R** produisant ces simulations :

```
n = rpois(1,  $\lambda L^2$ )
X = runif(n, 0, L)
Y = runif(n, 0, L)
R = rexp(n, 1)
```

```
plot(0:L, 0:L, type="n", xlab="", ylab="")
for (i in 1:n) points(X[i],Y[i],cex=R[i],pch=16)
```

2.2 Processus markoviens de sauts

Les Processus markoviens de sauts sont la généralisation des chaînes de Markov au temps continu. Le passage du temps discret au temps continu se fait en remplaçant le pas de temps fixe d'une chaîne de Markov par des intervalles de temps aléatoires indépendants de loi exponentielle. Le processus de Poisson sera notre premier exemple de processus markoviens de sauts. Un second sera lié aux files d'attente et sera abordé dans la dernière partie de ce cours.

2.2.1 Définition et exemples

Considérons un ensemble E fini ou dénombrable et une suite croissante de v.a. $(S_n)_{n \in \mathbb{N}}$ à valeurs dans \mathbb{R}_+ . Dans la suite $S_0 = 0$. Un *processus de sauts* $\{X_t\}_{t \geq 0}$ à espace d'états E et d'instant de sauts $(S_n)_{n \in \mathbb{N}}$ est un processus stochastique dont la valeur ne peut changer qu'en ses instants de sauts :

$$\forall n \in \mathbb{N}, \exists ! x \in E \text{ tel que } \forall t \in [S_n; S_{n+1}[, X_t = x .$$

Nous nous intéressons exclusivement à une classe particulière de processus de sauts, appelés *processus markoviens de sauts*. Ces processus évoluent de la manière suivante. Supposons que le processus se trouve à l'état x à l'issu du saut intervenant à l'instant S_n .

1. Le *temps de séjour* dans l'état x , à savoir la v.a. $S_{n+1} - S_n$, suit la loi exponentielle de paramètre $\lambda(x)$. Le paramètre de cette loi peut donc dépendre de l'état x où le processus se trouve. Mais à part cette dépendance en l'état x , la v.a. $S_{n+1} - S_n$ est indépendante du passé du processus.
2. À l'instant S_{n+1} , le processus saute de l'état x vers l'état y (avec $y \neq x$) avec une probabilité $q_{x,y}$, quantité indépendante de $S_{n+1} - S_n$ et du passé.

L'évolution du processus est donc déterminée par la suite $(\lambda(x))_{x \in E}$ et par la matrice $Q := (q_{x,y})_{x,y \in E}$. Voici une définition plus formelle :

Définition 2.2.1 Un processus de sauts $\{X_t\}_{t \geq 0}$ à espace d'états E et d'instant de sauts $(S_n)_{n \in \mathbb{N}}$ est *markovien* s'il existe :

- une suite bornée $(\lambda(x))_{x \in E}$ de réels strictement positifs,
- une matrice $Q = (q_{x,y})_{x,y \in E}$ de réels positifs vérifiant

$$\forall x \in E, q_{x,x} = 0 \text{ et } \sum_{y \in E} q_{x,y} = 1 ,$$

telles que, pour tout entier n , pour tous états $x_0, x_1, \dots, x_n, x_{n+1}$ et pour tous réels positifs t_1, \dots, t_n, t_{n+1} :

$$\begin{aligned} \mathbb{P} \left(\begin{array}{c} X_{S_{n+1}} = x_{n+1}, \\ S_{n+1} - S_n > t_{n+1} \end{array} \middle| \begin{array}{c} X_{S_n} = x_n, S_n - S_{n-1} > t_n, \dots, \\ X_{S_1} = x_1, S_1 > t_1, X_0 = x_0 \end{array} \right) \\ = \mathbb{P} \left(\begin{array}{c} X_{S_{n+1}} = x_{n+1}, \\ S_{n+1} - S_n > t_{n+1} \end{array} \middle| X_{S_n} = x_n \right) \end{aligned} \quad (2.3)$$

$$= e^{-\lambda(x_n)t_{n+1}} q_{x_n, x_{n+1}} . \quad (2.4)$$

Le passé du processus $\{X_t\}_{t \geq 0}$, à savoir les états visités et les temps de séjours dans chacun de ces états, est décrit par l'événement

$$X_{S_n} = x_n, S_n - S_{n-1} > t_n, \dots, X_{S_1} = x_1, S_1 - S_0 > t_1, X_{S_0} = x_0.$$

L'égalité (2.3) traduit le caractère markovien du processus : conditionnellement à l'état présent, l'évolution future ne dépend pas du passé. Dans l'égalité (2.4), le terme $e^{-\lambda(x_n)t_{n+1}}$ représente la probabilité pour un temps de séjour de loi exponentielle de paramètre $\lambda(x_n)$ d'être supérieur à t_{n+1} et le terme $q_{x_n, x_{n+1}}$ la probabilité de sauter de x_n vers x_{n+1} . Le produit des termes $e^{-\lambda(x_n)t_{n+1}}$ et $q_{x_n, x_{n+1}}$ traduit l'indépendance entre le temps de séjour et le futur état visité.

Le coefficient $\lambda(x)$ doit être compris comme le *taux de saut* à partir de l'état x . Son inverse $1/\lambda(x)$ est le temps moyen avant le prochain saut lorsque le processus est dans l'état x . Le fait que la suite $(\lambda(x))_{x \in E}$ soit bornée est une hypothèse technique ; elle assure que la suite $(S_n)_{n \in \mathbb{N}}$ des instants de sauts du processus tend presque sûrement vers l'infini, et donc que la v.a. X_t est bien définie pour tout $t \geq 0$. Remarquons que cette condition est automatiquement remplie quand l'espace d'états E est fini.

Enfin, l'égalité $q_{x,x} = 0$, pour tout x , interdit au processus de sauter depuis un état dans lui-même.

Voici deux exemples de processus markoviens de sauts. Un troisième exemple concernant le modèle de file d'attente à un serveur sera abordé dans la dernière partie de ce cours.

EXEMPLES :

• **Processus de Poisson.** Le processus de Poisson $\{N_t\}_{t \geq 0}$ (d'intensité λ) est un processus de sauts à espace d'états $E = \mathbb{N}$: il vaut n sur tout l'intervalle de temps $[S_n; S_{n+1}[$. De plus, tous les temps de séjour suivent la loi exponentielle de paramètre λ et, depuis l'état n (qu'il atteigne à l'issue du n^e saut), il saute en $n+1$.

$$\forall n \in \mathbb{N}, \lambda(n) = \lambda \text{ et } q_{n,m} = \begin{cases} 1 & \text{si } m = n+1 \\ 0 & \text{sinon} \end{cases}.$$

Le processus de Poisson remplit aisément la définition d'un processus markovien de sauts. L'équation (2.3) repose sur l'indépendance des temps de séjour $S_1, S_2 - S_1, \dots, S_{n+1} - S_n$ et sur le fait que N_{S_n} soit toujours égal à n . Ce n'est pas plus difficile pour (2.4) :

$$\begin{aligned} \mathbb{P}(N_{S_{n+1}} = n+1, S_{n+1} - S_n > t_{n+1} \mid N_{S_n} = n) &= \mathbb{P}(S_{n+1} - S_n > t_{n+1}) \\ &= e^{-\lambda t_{n+1}} \\ &= e^{-\lambda t_{n+1}} q_{n,n+1}. \end{aligned}$$

• **Automate binaire.** Considérons une machine qui peut être soit en état de marche, soit en panne. Notons X_t la v.a. égale à 0 si la machine est en panne à l'instant t et égale à 1 si elle est en état de marche à l'instant t . Nous faisons l'hypothèse que les temps de fonctionnement de cette machine sont des v.a. de loi exponentielle de paramètre λ et que les temps de réparation sont des v.a. de loi exponentielle de paramètre μ . Nous supposons également toutes ces v.a. indépendantes.

Alors, le processus $\{X_t\}_{t \geq 0}$ est un processus markovien de sauts à valeurs dans $E = \{0, 1\}$. Les paramètres qui interviennent sont $\lambda(0) = \mu$, $\lambda(1) = \lambda$ et $q_{0,1} = q_{1,0} = 1$.

2.2.2 Dé-randomisation du temps et simulation

Considérons un processus markovien de sauts $\{X_t\}_{t \geq 0}$ à espace d'états E et d'instant de sauts $(S_n)_{n \in \mathbb{N}}$ associé à la suite $(\lambda(x))_{x \in E}$ et à la matrice $Q = (q_{x,y})_{x,y \in E}$ (voir Définition 2.2.1). On s'intéresse ici à la

suite $(Z_n)_{n \in \mathbb{N}}$ définie par :

$$Z_n = X_{S_n} . \quad (2.5)$$

Autrement dit, Z_n est l'état visité après le n^{e} saut, et la suite $(Z_n)_{n \in \mathbb{N}}$ correspond à la suite des états visités par le processus $\{X_t\}_{t \geq 0}$ sans tenir compte du temps passé dans chacun de ses états.

L'opération consistant à passer du processus $\{X_t\}_{t \geq 0}$ à la suite $(Z_n)_{n \in \mathbb{N}}$ définie comme en (2.5) est qualifiée de *dé-randomisation du temps*. En effet, on a remplacé la suite aléatoire des instants de sauts $(S_n)_{n \in \mathbb{N}}$ par la grille rigide \mathbb{N} . De plus :

Proposition 2.2.2 La suite $(Z_n)_{n \in \mathbb{N}}$ obtenue précédemment est une chaîne de Markov homogène à valeurs dans E et de matrice de transition Q . Elle a la particularité de vérifier :

$$\forall n \in \mathbb{N}, Z_n \neq Z_{n+1} .$$

Puisque l'incrémentement du temps est la seule différence entre le processus à temps continu $\{X_t\}_{t \geq 0}$ et sa version dé-randomisée en temps $(Z_n)_{n \in \mathbb{N}}$, l'algorithme pour le simuler sera très proche de celui de la chaîne de Markov $(Z_n)_{n \in \mathbb{N}}$. Cette routine, nommée `markov`, renvoie l'état dans lequel se trouve le processus markovien de sauts $\{X_t\}_{t \geq 0}$ au temps $T > 0$. Elle prend en entrée l'état initial x^0 du processus, le réel T , ainsi que le vecteur $\lambda = (\lambda(x))_{x \in E}$ et la matrice $Q = (q_{x,y})_{x,y \in E}$ déterminant la dynamique du processus. Les éléments de E sont numérotés et notés comme suit : x_1, x_2, \dots

```
markov = fonction( $x^0, T, \lambda, Q$ )
{
  X =  $x^0$ 
  T = T - rexp(1,  $\lambda[x^0]$ )
  while (T > 0)
  {
    U = runif(1, 0, 1)
    R =  $q_{X, x_1}$ 
    i = 1
    while (R < U)
    {
      i = i + 1
      R = R +  $q_{X, x_i}$ 
    }
    X =  $x_i$ 
    T = T - rexp(1,  $\lambda[x_i]$ )
  }
  X
}
```

Le cœur de cette routine consiste à choisir le prochain état visité par le processus selon le vecteur de probabilité $(q_{X, x_i})_{i \geq 1}$, l'état occupé étant X . Cette simulation peut se simplifier considérablement suivant le contexte (processus de Poisson, automate binaire).

2.2.3 Probabilités de transition et générateur de Markov

Le résultat suivant est l'analogue de la propriété de Markov et de l'homogénéité en temps qui définissent habituellement les chaînes de Markov. Comme dans le cas discret, cela conduit à une caractérisation simple de la loi du processus.

Proposition 2.2.3 Un processus markovien de sauts $\{X_t\}_{t \geq 0}$ à espace d'états E satisfait les deux égalités suivantes : pour tout entier n , pour tous états $x_0, x_1, \dots, x_n, x_{n+1}$ et pour toute suite croissante de réels positifs $0 < t_1 < \dots < t_n < t_{n+1}$,

$$\begin{aligned} \mathbb{P}(X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n, \dots, X_{t_1} = x_1, X_0 = x_0) &= \mathbb{P}(X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n) \\ &= \mathbb{P}(X_{t_{n+1}-t_n} = x_{n+1} \mid X_0 = x_n) . \end{aligned}$$

Dès lors, la probabilité conditionnelle $\mathbb{P}(X_t = y \mid X_s = x)$ ne dépend que des états x, y et de l'accroissement $t - s$, ce qui justifie la notation suivante. Pour tout réel positif t et pour tous états x, y , la *probabilité de transition de x vers y sur un intervalle (de temps) de longueur t* est définie par :

$$p_{x,y}^{(t)} := \mathbb{P}(X_t = y \mid X_0 = x) .$$

Attention, la notation $p_{x,y}^{(t)}$ ne précise pas le nombre de sauts pour aller de l'état x à l'instant 0 à l'état y à l'instant t .

La matrice $P^{(t)} := (p_{x,y}^{(t)})_{x,y \in E}$ est appelée *matrice de transition sur un intervalle (de temps) de longueur t* . Comme dans le cas discret, la taille de la matrice $P^{(t)}$ dépend du cardinal de E (éventuellement infini), ses coefficients sont des réels positifs et la somme de ces coefficients sur toute une ligne vaut 1.

Comme dans le cas discret, la loi μ_t de la v.a. X_t est obtenue par produit matriciel entre la loi initiale μ_0 (i.e. celle de X_0) et la matrice de transition $P^{(t)}$:

$$\mu_t = \mu_0 P^{(t)} .$$

Les Propositions 2.2.4 et 2.2.5 fournissent deux autres points communs entre processus markoviens à temps discret et à temps continu. La donnée de la famille de matrices $\{P^{(t)}\}_{t \geq 0}$ et de la loi de X_0 suffit à caractériser la loi d'un processus markovien de sauts :

Proposition 2.2.4 Soit $\{X_t\}_{t \geq 0}$ un processus markovien de sauts de matrices de transition $\{P^{(t)}\}_{t \geq 0}$. Pour tout entier n , pour tous états x_0, x_1, \dots, x_n et pour toute suite croissante de réels positifs $0 < t_1 < \dots < t_n$, on a :

$$\mathbb{P}(X_{t_n} = x_n, \dots, X_{t_1} = x_1, X_0 = x_0) = \mathbb{P}(X_0 = x_0) p_{x_0, x_1}^{(t_1)} p_{x_1, x_2}^{(t_2 - t_1)} \dots p_{x_{n-1}, x_n}^{(t_n - t_{n-1})} .$$

La démonstration de cette dernière égalité est la même que dans le cas discret. Elle utilise exclusivement les caractères markovien et homogène du processus $\{X_t\}_{t \geq 0}$ (Proposition 2.2.3).

Le résultat suivant s'interprète comme suit : "aller de x vers y en un temps $t + s$, c'est aller de x vers un état z en un temps t puis de z vers y en un temps s .

Proposition 2.2.5 (*Relation de Chapman-Kolmogorov*) Soit $\{X_t\}_{t \geq 0}$ un processus markovien de sauts de matrices de transition $\{P^{(t)}\}_{t \geq 0}$. Soient $s, t \geq 0$. Alors

$$P^{(t+s)} = P^{(t)} P^{(s)} .$$

Ou encore, les coefficients de ces matrices satisfont les relations : pour tous états x, y ,

$$p_{x,y}^{(t+s)} = \sum_{z \in E} p_{x,z}^{(t)} p_{z,y}^{(s)} .$$

Démonstration Écrivons pour commencer

$$\begin{aligned}
p_{x,y}^{(t+s)} &= \mathbb{P}(X_{t+s} = y \mid X_0 = x) \\
&= \sum_{z \in E} \mathbb{P}(X_{t+s} = y, X_t = z \mid X_0 = x) \\
&= \sum_{z \in E} \mathbb{P}(X_{t+s} = y \mid X_t = z, X_0 = x) \mathbb{P}(X_t = z \mid X_0 = x)
\end{aligned}$$

en multipliant au numérateur et au dénominateur par la facteur $\mathbb{P}(X_t = z, X_0 = x)$. La démonstration s'achève en utilisant la propriété de Markov et l'homogénéité :

$$\begin{aligned}
\mathbb{P}(X_{t+s} = y \mid X_t = z, X_0 = x) &= \mathbb{P}(X_{t+s} = y \mid X_t = z) \\
&= \mathbb{P}(X_s = y \mid X_0 = z) \\
&= p_{z,y}^{(s)}.
\end{aligned}$$

■

Les similitudes entre les cas discret et continu s'arrêtent ici. En effet, les probabilités de transition $p_{x,y}^{(t)}$ sont en général très difficile à calculer, du fait de leur dépendance en le temps. Cependant, la relation $P^{(t+s)} = P^{(t)}P^{(s)}$ implique que toute la dynamique de la chaîne est contenue dans $P^{(\epsilon)}$ avec $\epsilon > 0$ petit. En effet, pour tout n , il vient $P^{(n\epsilon)} = (P^{(\epsilon)})^n$. Sont alors connues les matrices de transition $P^{(t)}$ pour $t = \epsilon, 2\epsilon, 3\epsilon, \dots$. Autrement dit, il s'agit de comprendre comment varie l'application $t \mapsto P^{(t)}$ sur l'intervalle de temps infinitésimal $[0; \epsilon]$. Cette idée motive l'introduction du *générateur* du processus. Auparavant remarquons que, puisque le processus markovien de sauts $\{X_t\}_{t \geq 0}$ issu de l'état x , y demeure jusqu'au premier saut S_1 , il vient

$$p_{x,y}^{(0)} := \lim_{t \rightarrow 0^+} p_{x,y}^{(t)} = \begin{cases} 0 & \text{si } x \neq y \\ 1 & \text{si } x = y \end{cases}.$$

Théorème 2.2.6 Soit $\{X_t\}_{t \geq 0}$ un processus markovien de sauts associé à la suite $(\lambda(x))_{x \in E}$ et à la matrice $Q = (q_{x,y})_{x,y \in E}$. Soient $x, y \in E$. L'application $t \mapsto p_{x,y}^{(t)}$ est dérivable à droite en $t = 0$ et cette dérivée vaut :

$$a_{x,y} := \lim_{t \rightarrow 0^+} \frac{p_{x,y}^{(t)} - p_{x,y}^{(0)}}{t} = \begin{cases} \lambda(x)q_{x,y} & \text{si } x \neq y \\ -\lambda(x) & \text{si } x = y \end{cases}. \quad (2.6)$$

Lorsque les états x et y sont distincts, cette dérivée est appelée *taux de transition de x vers y* . La matrice $A = (a_{x,y})_{x,y \in E}$ est appelée *générateur de Markov* du processus $\{X_t\}_{t \geq 0}$.

Soient x et y des états distincts. La probabilité $q_{x,y}$ d'aller en y lorsqu'on quitte x doit être pondérée par l'intensité $\lambda(x)$ avec laquelle on quitte x , pour obtenir le taux de transition $a_{x,y}$; $a_{x,y} = \lambda(x)q_{x,y}$. Par exemple, si le temps passé en x est en moyenne plutôt long (disons $\lambda(x)^{-1} = 100$) alors le taux de transition de x vers y sera faible, i.e. $a_{x,y} \leq 0,01$, et ce même si lorsqu'on quitte x c'est pour aller systématiquement en y ($q_{x,y} = 1$).

Démonstration Seul le cas où $x = y$ est traité : cela suffit à comprendre la trame du Théorème 2.2.6. Étant en x au temps 0, supposons que le processus y soit de nouveau au temps t . Il y a deux possibilités.

Soit le processus n'a pas quitté l'état x , ce qui signifie que le premier instant de saut S_1 survient après t . Soit le processus a quitté l'état x et y est revenu, ce qui implique au moins deux sauts avant l'instant t .

$$\mathbb{P}(X_t = x \mid X_0 = x) = \mathbb{P}(S_1 > t \mid X_0 = x) + \mathbb{P}(S_1 \leq t, S_2 \leq t, X_t = x \mid X_0 = x). \quad (2.7)$$

Sachant $X_0 = x$, l'instant du premier saut S_1 suit la loi exponentielle de paramètre $\lambda(x)$. Le premier terme du membre de droite de (2.7) devient :

$$\mathbb{P}(S_1 > t \mid X_0 = x) = e^{-\lambda(x)t} = 1 - \lambda(x)t + o(t),$$

lorsque t tend vers 0. Si le second terme du membre de droite de (2.7) est un $o(t)$ alors $\mathbb{P}(X_t = x \mid X_0 = x)$ vaut également $1 - \lambda(x)t + o(t)$. D'où la limite recherchée :

$$\frac{\mathbb{P}(X_t = x \mid X_0 = x) - 1}{t} = -\lambda(x) + o(1).$$

Nous savons ce qu'il reste à faire. Le fait que les deux premiers sauts du processus $\{X_t\}_{t \geq 0}$ surviennent à des instants très proches de 0 est improbable. C'est la raison pour laquelle $\mathbb{P}(S_1 \leq t, S_2 \leq t, X_t = x \mid X_0 = x)$ est un $o(t)$. Formellement,

$$\begin{aligned} \mathbb{P}(S_1 \leq t, S_2 \leq t, X_t = x \mid X_0 = x) &\leq \mathbb{P}(S_1 \leq t, S_2 \leq t \mid X_0 = x) \\ &\leq \mathbb{P}(S_1 \leq t, S_2 - S_1 \leq t \mid X_0 = x) \\ &\leq \sum_{z \in E} \mathbb{P}(S_1 \leq t, X_{S_1} = z, S_2 - S_1 \leq t \mid X_0 = x). \end{aligned}$$

Soit $z \in E$. Par définition du processus markovien de sauts, il vient :

$$\begin{aligned} \mathbb{P}(S_1 \leq t, X_{S_1} = z, S_2 - S_1 \leq t \mid X_0 = x) &= \mathbb{P}(X_{S_1} = z, S_2 - S_1 \leq t \mid S_1 \leq t, X_0 = x) \\ &\quad \times \mathbb{P}(S_1 \leq t \mid X_0 = x) \\ &= \mathbb{P}(X_{S_1} = z, S_2 - S_1 \leq t \mid S_1 \leq t)(1 - e^{-\lambda(x)t}) \\ &= (1 - e^{-\lambda(z)t})q_{x,z}(1 - e^{-\lambda(x)t}). \end{aligned}$$

En utilisant l'inégalité $e^{-u} \geq 1 - u$ et le fait que la suite $(\lambda(z))_{z \in E}$ soit bornée, disons par une constante $M > 0$, il vient :

$$\begin{aligned} \mathbb{P}(S_1 \leq t, S_2 \leq t, X_t = x \mid X_0 = x) &\leq \sum_{z \in E} \mathbb{P}(S_1 \leq t, X_{S_1} = z, S_2 - S_1 \leq t \mid X_0 = x) \\ &\leq \sum_{z \in E} (1 - e^{-\lambda(z)t})q_{x,z}(1 - e^{-\lambda(x)t}) \\ &\leq \sum_{z \in E} \lambda(z)q_{x,z}\lambda(x)t^2 \\ &\leq M\lambda(x)t^2 \sum_{z \in E} q_{x,z} \\ &\leq M\lambda(x)t^2 = o(t). \end{aligned}$$

■

Le générateur de Markov $A = (a_{x,y})_{x,y \in E}$ est entièrement déterminé par la suite $(\lambda(x))_{x \in E}$ et la matrice $Q = (q_{x,y})_{x,y \in E}$ définissant le processus markovien de sauts. C'est une matrice carrée à coefficients réels positifs excepté sur la diagonale. Le terme d'ordre x de la diagonale vérifie

$$a_{x,x} = -\lambda(x) = -\lambda(x) \sum_{y, y \neq x} q_{x,y} = - \sum_{y, y \neq x} a_{x,y} .$$

La somme des coefficients d'une même ligne de la matrice A vaut donc 0.

Dans les applications, un modèle markovien continu est défini par ses taux de transition $a_{x,y}$ qui ont en général une signification concrète (nombres moyens d'arrivées, de services, de pannes ou de réparations par unité de temps). De plus, l'intérêt du générateur par rapport aux matrices de transition $\{P^{(t)}\}_{t \geq 0}$ est que celui-ci ne dépend plus du temps : c'est une dérivée (à droite) en $t = 0$. Ainsi, on résume souvent l'information qu'il contient par un *graphe de transition*. C'est un graphe orienté et pondéré dont l'ensemble des sommets est E . Une arête de poids $a_{x,y}$ va de x vers $y \neq x$ si $a_{x,y} > 0$.

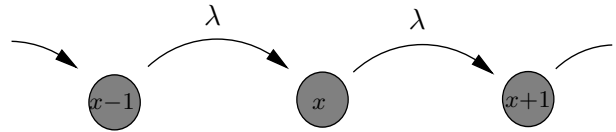
EXEMPLES :

• **Processus de Poisson.** Pour tout entier x , $\lambda(x)$ est égal à λ et $q_{x,x+1}$ égal à 1. Dès lors, le taux de transition de x vers $x + 1$ vaut

$$a_{x,x+1} = \lambda(x) q_{x,x+1} = \lambda .$$

Il est nul vers tout autre entier y : pour tout $y \in \mathbb{N} \setminus \{x, x + 1\}$, $a_{x,y} = 0$. Enfin, le coefficient $a_{x,x}$ vaut $-\lambda$.

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \ddots \\ 0 & 0 & -\lambda & \lambda & \ddots \\ 0 & 0 & 0 & -\lambda & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$



• **Automate binaire.** La v.a. X_t est égale à 0 ou 1 selon que la machine soit en état de panne ou de marche à l'instant t . Rappelons que les temps de fonctionnement de cette machine sont des v.a. de loi exponentielle de paramètre λ et que les temps de réparation sont des v.a. de loi exponentielle de paramètre μ , toutes ces v.a. étant indépendantes. D'après la Proposition 2.2.6, le taux de transition de l'état de panne vers l'état de fonctionnement est

$$a_{0,1} = \lambda(0) q_{0,1} = \mu$$

tandis le taux de transition de l'état de fonctionnement vers l'état de panne est

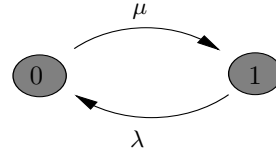
$$a_{1,0} = \lambda(1) q_{1,0} = \lambda .$$

Enfin, les coefficients $a_{0,0}$ et $a_{1,1}$ valent respectivement $-\mu$ et $-\lambda$.

2.2.4 Théorème limite

Comme dans le cas discret, l'étude du comportement asymptotique d'un processus markovien de sauts passe par l'identification d'une mesure stationnaire sur l'espace d'états E .

$$A = \begin{pmatrix} -\mu & \mu \\ \lambda & -\lambda \end{pmatrix}$$



Définition 2.2.7 Une *mesure stationnaire* (ou *invariante*) d'un processus markovien de sauts de matrices de transition $\{P^{(t)}\}_{t \geq 0}$ est une loi de probabilité sur E , disons $\pi = (\pi(x))_{x \in E}$, vérifiant pour tout t la relation $\pi = \pi P^{(t)}$.

Rappelons que la loi μ_t de la v.a. X_t vérifie $\mu_t = \mu_0 P^{(t)}$. Dès lors, la relation $\pi = \pi P^{(t)}$ s'interprète comme suit : si la v.a. initiale X_0 a pour loi la mesure stationnaire π alors, pour tout temps t , la loi de X_t est encore π .

La relation matricielle $\pi = \pi P^{(t)}$ est équivalente au système linéaire

$$\forall y \in E, \sum_{x \in E} \pi(x) p_{x,y}^{(t)} = \pi(y)$$

(de taille égale au cardinal de E).

Les matrices de transition $\{P^{(t)}\}_{t \geq 0}$ étant en général inaccessibles, nous privilégierons la caractérisation des mesures stationnaires en termes de générateur.

Proposition 2.2.8 Une loi de probabilité π sur E est une mesure stationnaire d'un processus markovien de sauts de générateur A si et seulement si $\pi A = 0$, ce qui s'écrit :

$$\forall y \in E, \sum_{x \in E} \pi(x) a_{x,y} = 0. \quad (2.8)$$

EXEMPLES :

- **Processus de Poisson.** Pour tout entier x , le taux de transition $a_{x,y}$ vaut λ si et seulement $y = x + 1$. Il est nul vers tout autre entier y . Les équations de stationnarité (2.8) donnent :

$$\begin{cases} -\lambda\pi(x) = 0 \\ \lambda\pi(x) - \lambda\pi(x+1) = 0, \forall x \in \mathbb{N} \end{cases}$$

ou encore $\pi(x) = 0$ pour tout x . Le processus de Poisson en tant que processus markovien de sauts n'admet donc pas de mesure stationnaire. C'est relativement intuitif ; il ne peut pas exister d'équilibre en loi pour un processus qui tend p.s. vers l'infini.

- **Automate binaire.** Avec le générateur

$$A = \begin{pmatrix} -\mu & \mu \\ \lambda & -\lambda \end{pmatrix},$$

les équations de stationnarité donnent $\mu\pi(0) = \lambda\pi(1)$. Puisque $\pi(0) + \pi(1) = 1$, on trouve comme unique mesure stationnaire

$$\pi = \left(\frac{\lambda}{\lambda + \mu}, \frac{\mu}{\lambda + \mu} \right).$$

La notion de mesure réversible existe aussi pour les processus markoviens de sauts, exactement comme pour les chaînes de Markov.

Définition 2.2.9 Une *mesure réversible* d'un processus markovien de sauts de générateur A est une loi de probabilité sur E , disons $\pi = (\pi(x))_{x \in E}$, vérifiant

$$\forall x, y \in E, \pi(x)a_{x,y} = \pi(y)a_{y,x} . \quad (2.9)$$

Proposition 2.2.10 Toute mesure réversible pour un processus est stationnaire pour ce processus.

En pratique (et notamment pour les projets), pour obtenir une mesure stationnaire, il est recommandé de commencer par chercher une mesure réversible, plus facile à identifier quand elle existe. D'ailleurs, la mesure stationnaire

$$\pi = \left(\frac{\lambda}{\lambda + \mu}, \frac{\mu}{\lambda + \mu} \right)$$

de l'automate binaire est également réversible. En effet, elle satisfait l'équation

$$\pi(0)a_{0,1} = \pi(1)a_{1,0} .$$

Définition 2.2.11 Un processus markovien de sauts de générateur A est dit *irréductible sur E* si pour tous états $x, y \in E$ distincts il existe des états $x_1, \dots, x_n \in E$ tous différents tels que :

$$a_{x,x_1}a_{x_1,x_2} \dots a_{x_{n-1},x_n}a_{x_n,y} > 0 .$$

Comme dans le cas discret, être irréductible signifie que l'on peut passer (en plusieurs étapes si nécessaire) de n'importe quel état x à n'importe quel état y avec une probabilité strictement positive. Le processus markovien de sauts correspondant à l'automate binaire est irréductible. Ce n'est pas le cas du processus de Poisson.

Nous admettons le théorème limite suivant.

Théorème 2.2.12 Considérons un processus markovien de sauts $\{X_t\}_{t \geq 0}$ de matrices de transition $\{P^{(t)}\}_{t \geq 0}$, irréductible sur E et admettant une mesure stationnaire π . Alors :

- (1) π est l'unique mesure stationnaire du processus $\{X_t\}_{t \geq 0}$;
- (2) la matrice $P^{(t)}$ converge quand t tend vers l'infini vers une matrice dont toutes les lignes sont égales à π :

$$\forall x, y \in E, \lim_{t \rightarrow +\infty} p_{x,y}^{(t)} = \pi(y) ;$$

- (3) quelle que soit la loi de X_0 , la loi de X_t converge quand t tend vers l'infini vers π :

$$\forall x \in E, \lim_{t \rightarrow +\infty} \mathbb{P}(X_t = x) = \pi(x) ;$$

Le comportement asymptotique d'un processus markovien de sauts irréductible est donc décrit par l'unique mesure stationnaire quand elle existe. Dans ce cas, le Théorème 2.2.12 exprime qu'au bout d'un certain temps, le système se stabilise dans un régime d'équilibre appelé *régime stationnaire*. Une fois ce régime atteint, la probabilité d'être dans l'état x est donnée par $\pi(x)$.

Remarquons enfin que la notion de périodicité n'a pas de sens en temps continu, ce qui simplifie d'autant la discussion. Par contre, celles de récurrence et de transience sont conservées : ce sont toujours des propriétés de classes irréductibles.

2.3 Application aux modèles de files d'attente

2.3.1 Présentation générale

La salle de réservation dans une grande gare SNCF donne une bonne représentation d'une file d'attente. Elle comprend un certain nombre de guichets et des clients qui sont, soit en train d'être servis, soit en attente. La salle de réservation forme le système. Elle peut être de capacité totale finie ou infinie. Les clients arrivent de manière aléatoire selon un flot régulier. Les temps de services sont également aléatoires. L'objectif est de savoir si la longueur de la file d'attente admet un comportement stationnaire et dans ce cas de calculer sa loi. Cela permet d'optimiser le nombre de guichets nécessaires pour satisfaire les clients.

Une file d'attente est décrite par plusieurs éléments :

1. Les *instants d'arrivée des clients* sont en général aléatoires. Certaines hypothèses sont faites sur leurs lois. Tout d'abord, il n'arrive qu'un client à la fois. La deuxième hypothèse est l'homogénéité dans le temps. Cela se traduit par le fait que les temps d'interarrivée des clients sont des v.a. de même loi. Ils sont également supposés indépendants. Cette hypothèse simplifie notablement l'étude mais elle est moins claire à justifier en pratique. Enfin, la loi des temps d'interarrivée est supposée connue. Le cas le plus courant est celui où cette loi est exponentielle. Dans ce cas, le modèle des temps d'arrivée des clients est un processus de Poisson. Évidemment d'autres cas peuvent se présenter : temps d'interarrivés constants, de loi uniforme, lognormale ou encore gamma.
2. Les *durées de service* sont supposées être des v.a. indépendantes, de même loi et indépendante du processus des arrivées.
3. Le *nombre de guichets ou serveurs* est évidemment un paramètre important du modèle.
4. La *longueur maximum de la file d'attente* est également un paramètre du modèle. Il est raisonnable de supposer dans certains cas que la file d'attente puisse être aussi longue que l'on veut. Cependant dans d'autres cas, la longueur est limitée et lorsque cette limite est atteinte, un client arrivant ne peut entrer dans la file d'attente. Il repart.
5. Le plus souvent les clients sont servis dans leur ordre d'arrivée. Cette *discipline de service* est appelée *FIFO* (First In First Out). Mais d'autres disciplines pourraient être utilisées comme par exemple, servir en priorité certains types de clients, ceux demandant un service de courte durée.

Pour résumer les caractéristiques d'une file d'attente on utilise classiquement la notation de Kendall :

loi d'interarrivée / loi de service / nombre de serveurs / longueur maximum de la file

Les lois d'interarrivées et les lois de services sont notées symboliquement M pour une loi exponentielle (M pour Markov, on verra pourquoi plus tard), D pour une loi déterministe (v.a. constante), U pour une loi uniforme, G pour une loi quelconque (G pour générale). Par exemple, une file M/M/s/∞ signifie que le flot d'arrivée des clients est poissonnien, la loi des services est exponentielle, il y a s serveurs et la capacité de la salle d'attente est illimitée. Lorsqu'on ne spécifie pas le dernier paramètre celui-ci est infini. Sauf avis contraire la discipline de service est FIFO.

Dans la suite, nous n'étudierons que le modèle de file d'attente M/M/1.

2.3.2 Étude du cas M/M/1

Considérons un guichet ou un serveur. Les arrivées des clients à ce guichet suivent un processus de Poisson d'intensité λ . Le temps de service pour chaque client est une v.a. de loi exponentielle de paramètre

μ . Toutes les v.a. qui interviennent sont supposées indépendantes. Les clients se mettent en file d'attente et sont servis selon leur ordre d'arrivée (discipline FIFO). La capacité de la file d'attente est illimitée.

Considérons le processus $\{X_t\}_{t \geq 0}$ représentant le nombre de clients en attente (y compris le client en train d'être servi) au temps t . C'est un processus de sauts à valeurs dans \mathbb{N} . Quand un client arrive, le processus saute de $+1$ et quand un client s'en va à la fin de son service, le processus saute de -1 .

Décrivons plus précisément cette dynamique. Lorsque le processus saute en $x > 0$, il y reste un temps aléatoire qui vaut $\min(Y_1, Y_2)$ où Y_1 est le temps avant l'arrivée du prochain client et Y_2 est le temps avant la fin du service en cours. Ces deux v.a. sont indépendantes de loi exponentielle de paramètres respectifs λ et μ (d'après la propriété de perte de mémoire de la loi exponentielle : Proposition 2.1.1). D'après la Proposition 2.1.2, le temps de séjour dans l'état x , i.e. la v.a. $\min(Y_1, Y_2)$, suit la loi exponentielle de paramètre $\lambda + \mu$. Le saut suivant est de $+1$ si $\min(Y_1, Y_2) = Y_1$ ce qui se produit avec probabilité $\lambda/(\lambda + \mu)$. Il est de -1 si $\min(Y_1, Y_2) = Y_2$ ce qui se produit avec probabilité $\mu/(\lambda + \mu)$. Le fait que Y_1 et Y_2 soient p.s. différents signifie que deux clients ne peuvent entrer et sortir au même instant.

Le cas $x = 0$ (il n'y a plus personne dans la file d'attente) est un peu différent. Une fois en 0, le processus y reste un temps aléatoire Y_1 qui est le temps avant l'arrivée du prochain client. C'est une v.a. de loi exponentielle de paramètre λ . Le saut à l'issue de ce temps est nécessairement de $+1$.

En résumé, la description de la dynamique du processus $\{X_t\}_{t \geq 0}$ est celle d'un processus markovien de sauts :

- le temps de séjour en l'état x suit la loi exponentielle de paramètre $\lambda(x) = \lambda + \mu$ si $x \geq 1$ et $\lambda(0) = \lambda$;
- les coefficients de la matrice $Q = (q_{x,y})_{x,y \in E}$ vérifient

$$q_{x,y} = \begin{cases} \frac{\lambda}{\lambda + \mu} & \text{si } y = x + 1 \\ \frac{\mu}{\lambda + \mu} & \text{si } y = x - 1 \\ 0 & \text{sinon} \end{cases}$$

lorsque $x \geq 1$, et $q_{0,1} = 1$ lorsque $x = 0$.

Enfin, le générateur de Markov du processus se calcule au moyen de la Proposition 2.2.6.

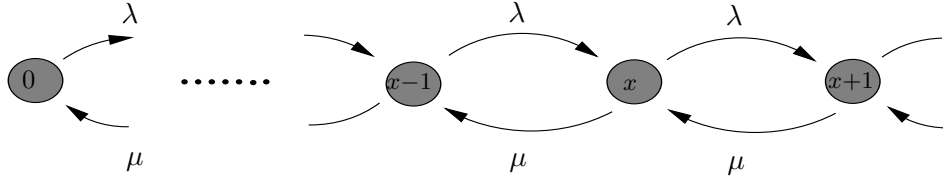
Proposition 2.3.1 Le processus stochastique $\{X_t\}_{t \geq 0}$ représentant le nombre de clients dans la file d'attente M/M/1 est un processus markovien de sauts à espace d'états \mathbb{N} et de générateur de Markov

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ \mu & -(\lambda + \mu) & \lambda & 0 & 0 & \ddots \\ 0 & \mu & -(\lambda + \mu) & \lambda & 0 & \ddots \\ 0 & 0 & \mu & -(\lambda + \mu) & \lambda & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Voici le graphe de transition de la file d'attente M/M/1 :

Le processus markovien de sauts $\{X_t\}_{t \geq 0}$ représentant le nombre de clients dans la file d'attente M/M/1 est irréductible. D'après le Théorème 2.2.12, s'il admet une mesure stationnaire alors celle-ci est unique. Le plus simple est de chercher une mesure réversible. Les équations de réversibilité (2.9) donnent, pour tout entier $x \in \mathbb{N}$

$$\pi(x)\lambda = \pi(x+1)\mu,$$



ce qui, après un petit raisonnement par récurrence, aboutit à

$$\forall x \in \mathbb{N}, \pi(x) = \pi(0) \left(\frac{\lambda}{\mu} \right)^x.$$

Il existe donc une mesure réversible si et seulement si la série de terme général $\pi(x)$ est convergente, i.e. si $\lambda < \mu$. Cette mesure réversible s'écrit alors :

$$\forall x \in \mathbb{N}, \pi(x) = \left(1 - \frac{\lambda}{\mu} \right) \left(\frac{\lambda}{\mu} \right)^x$$

($\pi(0)$ est choisi afin de satisfaire la condition $\sum \pi(x) = 1$).

Proposition 2.3.2 Le processus stochastique $\{X_t\}_{t \geq 0}$ représentant le nombre de clients dans la file d'attente M/M/1 admet une mesure stationnaire si et seulement si le paramètre λ du processus de Poisson des arrivées est strictement inférieur au paramètre μ de la loi exponentielle des services. Dans ce cas, elle est unique et apparait comme limite lorsque t tend vers l'infini de la loi du nombre de clients dans la file au temps t :

$$\forall x \in \mathbb{N}, \lim_{t \rightarrow +\infty} \mathbb{P}(X_t = x) = \left(1 - \frac{\lambda}{\mu} \right) \left(\frac{\lambda}{\mu} \right)^x$$

(cette loi limite est celle de $G - 1$ où G suit la loi géométrique de paramètre λ/μ).

Ce résultat est tout à fait intuitif. En effet, les paramètres λ et μ représentent respectivement le nombre moyen d'arrivée de clients par unité de temps et μ le nombre moyen de clients servis par unité de temps. Ainsi, la file d'attente se stabilise s'il n'arrive pas trop de clients pour saturer l'offre de service, i.e. si λ est plus petit que μ . En régime stationnaire (lorsque $\lambda < \mu$), la longueur de la file d'attente a pour moyenne et variance :

$$\frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}} \text{ et } \frac{\frac{\lambda}{\mu}}{\left(1 - \frac{\lambda}{\mu}\right)^2}.$$

Pour compléter ces résultats, signalons que lorsque $\lambda > \mu$, la longueur de la file d'attente tend p.s. vers l'infini.

Intéressons-nous enfin au temps total R passé dans le système par un client donné. Supposons qu'à son arrivée dans la file d'attente, il y a n clients. Le temps total qu'il passera dans le système correspond donc :

- à la fin du service du client au guichet lors de son arrivée : c'est une exponentielle de paramètre μ d'après la propriété de perte de mémoire ;
- au $n - 1$ services des clients présents devant lui dans la file et non encore servis : ce sont des exponentielles de paramètre μ ;
- à son propre temps de service : encore une exponentielle de paramètre μ .

Toutes ces v.a. étant indépendantes, ce temps total suit donc une loi Gamma de paramètres $n + 1$ et μ . Autrement dit,

$$\mathbb{P}(R \leq t \mid X_0 = n) = \int_0^t \mu^{n+1} \frac{x^n}{n!} e^{-\mu x} dx . \quad (2.10)$$

En régime stationnaire, la loi de X_0 est connue (voir Proposition 2.3.2). La loi du temps total R passé dans le système ne dépend plus du client choisi. Il est appelé *temps de réponse* du système.

Proposition 2.3.3 Considérons la file d'attente M/M/1 dans le cas $\lambda < \mu$. En régime stationnaire, le temps de réponse R est distribué selon la loi exponentielle de paramètre $\mu - \lambda$.

Démonstration Soit $t \geq 0$. D'après (2.10),

$$\begin{aligned} \mathbb{P}(R \leq t) &= \sum_{n=0}^{\infty} \mathbb{P}(X_0 = n) \mathbb{P}(R \leq t \mid X_0 = n) \\ &= \sum_{n=0}^{\infty} \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^n \int_0^t \mu^{n+1} \frac{x^n}{n!} e^{-\mu x} dx \\ &= \int_0^t (\mu - \lambda) e^{-\mu x} \left(\sum_{n=0}^{\infty} \frac{\lambda^n x^n}{n!}\right) dx \\ &= \int_0^t (\mu - \lambda) e^{-(\mu - \lambda)x} dx \\ &= 1 - e^{-(\mu - \lambda)t} . \end{aligned}$$

On reconnaît la fonction de répartition de la loi exponentielle de paramètre $\mu - \lambda$. Puisque les fonctions de répartition caractérisent les lois, le résultat est démontré. ■

2.4 Exercices

Exercice 1 : Bit informatique poissonnien

Considérons un processus de Poisson $\{N_t\}_{t \geq 0}$ d'intensité λ et une v.a. X_0 , indépendante du processus $\{N_t\}_{t \geq 0}$, valant -1 ou 1 avec probabilité $1/2$. Soit $\{X_t\}_{t \geq 0}$ le processus à valeurs dans $\{-1, 1\}$, changeant de signe à chaque saut du processus $\{N_t\}_{t \geq 0}$ et issu de X_0 .

1. Après avoir donné une expression de X_t en fonction de N_t et de X_0 , déterminer son espérance et sa variance.

2. Calculer la probabilité pour que la valeur du bit au temps t_0 soit la même qu'initialement. Puis la limite de cette probabilité quand $t_0 \rightarrow +\infty$. Comparer avec le bit informatique à temps discret. On pourra utiliser la formule

$$\sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!} = \frac{1}{2} (e^z + e^{-z}) .$$

Exercice 2 : Temps d'attente de voyageurs

Supposons que des voyageurs en partance pour Londres arrivent à la gare Lille Europe selon un processus de Poisson d'intensité λ . Le départ de leur train est prévu au temps t_0 (nous conviendrons que la gare ouvre ses portes au temps $t = 0$).

1. Calculer le nombre moyen de voyageurs prenant ce train.
2. Calculer la moyenne de la somme des temps d'attente de ces voyageurs.

Exercice 3 : Loi binomiale et processus de Poisson

Comme il fait beau ce matin, je prends mon petit-déjeuner dans le jardin. Des merles et des moineaux viennent chiper des miettes de pain que j'éparpille volontairement sur la table. Les passages des merles et des moineaux peuvent être modélisés par deux processus de Poisson indépendants d'intensités respectives λ et λ' .

1. Calculer la probabilité pour que le premier oiseau venant sur la table soit un merle.
2. Calculer la probabilité pour que les deux premiers oiseaux venant sur la table soient des merles.
Indication : utiliser la propriété de perte de mémoire de la loi exponentielle.
3. Quelle est la loi du nombre d'oiseaux venant sur la table jusqu'au temps t ? Sachant que ce nombre vaut n , quelle est la loi du nombre de merles venant sur la table jusqu'au temps t ?

Exercice 4 : Radioactivité

Une source radioactive émet des particules selon un processus de Poisson d'intensité λ . Chaque particule est émise suivant une direction aléatoire, de loi uniforme sur $[0; 2\pi]$, et indépendante des autres particules. Un compteur Geiger, placé près de la source, enregistre les particules émises de direction comprise entre $-\pi/4$ et $\pi/4$. Quelle est la loi du nombre de particules enregistrées jusqu'au temps t ?

Exercice 5 (Examen GIS 2009) :

Une fois la nuit tombée, Pierre et Paul se retrouvent à l'observatoire afin de contempler des étoiles filantes. Le passage des étoiles filantes dans le ciel est modélisé par un processus de Poisson d'intensité 0.1 (l'unité de temps étant la minute), noté $\{N_t\}_{t \geq 0}$. Soit X_n le temps aléatoire séparant le passage de la $(n-1)^e$ étoile filante du passage de la n^e .

1. Quelle est la probabilité pour Pierre et Paul de ne voir aucune étoile filante durant la première heure d'observation ? Quel est le nombre moyen d'étoiles filantes qu'ils devraient voir durant cette première heure d'observation ? Et pendant la seconde heure d'observation ?

2. Paul est de nature impatient (seulement pour cette question) : s'il ne voit aucune étoile filante durant 20 minutes (consécutives) alors il s'en va ! Quelle est la probabilité qu'il en voit exactement 3 ? Plus généralement, quelle est la loi du nombre d'étoiles filantes que Paul verra ? En déduire le nombre moyen d'étoiles filantes que l'impatient Paul verra.

3. On suppose enfin que Paul est arrivé en retard au rendez-vous (disons au temps $t_0 > 0$ alors que Pierre est arrivé au temps $t = 0$).

(a) Sachant qu'aucune étoile filante n'est passée avant l'arrivée de Paul, quelle est la loi du temps qu'il devra patienter pour en observer une ?

(b) Supposons dans cette question que $t_0 = 20$. Sachant qu'une seule étoile filante est passée avant l'arrivée de Paul, quelle aurait été la probabilité pour Paul de la voir s'il était arrivé 5 minutes plus tôt ?

Exercice 6 : Loi géométrique et processus de Poisson

Je souhaite traverser une route à grande circulation dont le flux de véhicules peut être modélisé par un processus de Poisson d'intensité $\lambda = 1$ (l'unité étant la minute). Depuis l'ablation de mes deux jambes, il me faut tout de même 3 minutes pour traverser une telle route. Avant de me lancer, je préfère réfléchir un peu. . .

1. Quelle est la loi du nombre de véhicules que je dois laisser passer avant de pouvoir traverser ?

2. Quel est le temps moyen d'attente avant de traverser ?

Exercice 7 (Examen GIS 2011) : Problèmes de simulations

Précisons que les générateurs mentionnés ci-dessous génèrent des v.a. indépendantes entre elles, mais aussi de tout le reste !

1. Comment simuler une loi exponentielle de paramètre 2 lorsqu'on dispose uniquement d'un générateur de loi exponentielle de paramètre 1 ?

2. Comment simuler un processus de Poisson d'intensité 1 sur l'intervalle $[0; 100]$ sans utiliser de lois exponentielles (quel que soit le paramètre) ?

3. À partir de la simulation précédente et d'un générateur de loi de Bernoulli de paramètre $\frac{1}{2}$, comment simuler un processus de Poisson d'intensité $\frac{1}{2}$ sur l'intervalle $[0; 100]$?

Exercice 8 (Examen GIS 2013) :

Un avion de guerre survole une ville selon une trajectoire rectiligne mesurant 10 km. Nous appellerons x -ième arrondissement la portion de ville survolée entre les kilomètres $x - 1$ et x . Supposons que les points d'impact des bombes larguées par l'avion suivent un processus de Poisson $(N_t)_{t \in [0,10]}$ d'intensité $\frac{1}{2}$. Une seule bombe larguée entre les kilomètres $x - 1$ et x suffit à détruire totalement l'arrondissement correspondant.

1. Combien de bombes en moyenne sont larguées lors du passage de l'avion ? Quelle est la probabilité pour que le premier arrondissement soit détruit ?

2. Supposons qu'une seule bombe soit larguée durant toute la traversée. Quel arrondissement a-t'il le plus de chance d'être détruit ?

3. Que représente la variable aléatoire

$$X = \sum_{x=0}^9 \mathbb{1}_{N_{x+1}-N_x \geq 1} ?$$

Calculer son espérance. Comparer avec le nombre moyen de bombes larguées (question 1).

4. Calculer la probabilité pour que tous les arrondissements soient détruits.

5. Supposons enfin que l'avion effectue une seconde traversée, indépendante de la première et dont le largage est encore modélisé par un processus de Poisson d'intensité $\frac{1}{2}$. Quelle est dans ce cas le nombre moyen d'arrondissements détruits (s'inspirer de la question 3.) ?

Exercice 9 : Deux automates binaires dépendants

Considérons un système à deux composants identiques tel que, en cas de panne de l'un des deux composants, l'autre puisse aussi tomber en panne. Les hypothèses de modélisation sont les suivantes.

- Le temps de fonctionnement de chaque composant suit la loi $\mathcal{E}(\lambda)$.
- Le temps de réparation d'un composant seul (pendant le fonctionnement de l'autre) suit la loi $\mathcal{E}(\mu)$.
- Le temps de réparation des deux composants simultanément suit la loi $\mathcal{E}(\nu)$.
- En cas de panne d'un des composants, l'autre continue à fonctionner avec probabilité c .
- Toutes les v.a. du modèle sont indépendantes.

Soit Z_t l'état du système au temps t : $Z_t \in E = \{00, 10, 01, 11\}$ où, par exemple, l'état 01 signifie que le premier composant est en panne tandis que le second est en fonctionnement.

1. Nous admettons que $\{Z_t\}_{t \geq 0}$ est un processus markovien de sauts sur E . Préciser cependant la suite $(\lambda(x))_{x \in E}$ et la matrice $(q_{x,y})_{x,y \in E}$ le définissant. Déterminer son générateur de Markov et dessiner son graphe de transition.

2. Où dans ce modèle apparaît la dépendance entre les deux composants ?

3. Nous supposons désormais que le temps moyen de fonctionnement d'un composant est 1 jour, le temps moyen de réparation d'un composant seul est 6 heures et celui des deux composants simultanément est 8 heures.

- (a) Comment traduire ces hypothèses sur les paramètres λ , μ et ν (avec le jour comme unité) ?
- (b) Déterminer la mesure stationnaire du processus $\{Z_t\}_{t \geq 0}$. Est-elle réversible ?
- (c) Supposons que le système global fonctionne dès qu'au moins un des deux composants fonctionne. Quelle condition sur le paramètre c assure le fonctionnement du système dans 90% du temps ?

Exercice 10 : Étude de la file M/M/2

Soit $\{X_t\}_{t \geq 0}$ le processus stochastique représentant le nombre de clients dans la file d'attente M/M/2 pour laquelle λ est l'intensité du processus de Poisson régissant l'arrivée des clients et μ est le paramètre de la loi exponentielle modélisant les durées de service. Rappelons que toutes les v.a. intervenant dans ce modèle sont indépendantes, y compris les durées de service des 2 guichets.

1. Étude du comportement asymptotique.

(a) Déterminer les taux de transition du processus markovien de sauts $\{X_t\}_{t \geq 0}$. Représenter son graphe de transition.

(b) Déterminer le comportement asymptotique du processus $\{X_t\}_{t \geq 0}$.

2. Longueur moyenne de la file d'attente en régime stationnaire.

(a) Calculer en régime stationnaire le nombre moyen de clients dans la file d'attente. Rappelons à cet effet que :

$$\sum_{n \geq 1} nr^{n-1} = \frac{1}{(1-r)^2} \text{ dès que } 0 < r < 1.$$

(b) Comparer cette espérance au nombre moyen de clients dans la file M/M/1 pour laquelle l'intensité du processus de Poisson des arrivées est encore λ et le paramètre de la loi exponentielle des durées de service est cette fois 2μ . Comment expliquez-vous cette différence ?

Exercice 11 : Étude de la file M/M/ ∞ (Examen GIS 2015)

On considère un système composé d'un nombre infini de guichets identiques et indépendants les uns des autres. Dès qu'un client arrive, il entre donc instantanément en service (car il y a toujours au moins un guichet libre). On suppose que le processus d'arrivée des clients est un processus de Poisson d'intensité $\lambda > 0$ et que les temps de services sont des exponentielles de paramètre $\mu > 0$ (pour tous les guichets). On note X_t le nombre de clients dans le système au temps t . Nous admettrons que $(X_t)_{t \geq 0}$ est un processus markovien de sauts à temps continu.

1. Soit $n \in \mathbb{N}$.

(a) Montrer que le paramètre $\lambda(n)$ du temps de séjour du processus $(X_t)_{t \geq 0}$ en l'état n vaut $\lambda(n) = n\mu + \lambda$.

(b) Lorsque le processus $(X_t)_{t \geq 0}$ quitte l'état n , déterminer la probabilité $q_{n,m}$ avec laquelle il va en $m \in \mathbb{N}$.

(c) En déduire les taux de transition $(a_{n,m})_{n,m \in \mathbb{N}}$ du processus $(X_t)_{t \geq 0}$ et donner son graphe de transition.

2.(a) Déterminer une mesure réversible π pour le processus $(X_t)_{t \geq 0}$. Quelle loi reconnaissez-vous ?

(b) Est-ce la seule mesure invariante du processus $(X_t)_{t \geq 0}$?

(c) Sous quelles conditions sur λ et μ , la variable aléatoire X_t converge quand $t \rightarrow \infty$ vers sa mesure invariante π ? Comparer avec la file d'attente M/M/1.

(d) Déterminer le nombre moyen de clients dans le système en régime stationnaire.

(e) Supposons $\lambda = 1$ et $\mu = 1/2$. Donner une valeur approchée de la probabilité $\mathbb{P}(X_t > 5)$ quand t est grand.

Chapitre 3

Méthodes MCMC

3.1 Méthodes de Monte Carlo

On appelle méthode de Monte Carlo toute méthode algorithmique visant à calculer une valeur numérique approchée d'une certaine quantité (souvent inconnue) par un procédé aléatoire. Les méthodes de Monte Carlo sont notamment utilisées pour obtenir des valeurs approchées d'intégrales, comme des aires ou des volumes.

Voici un exemple. Supposons que nous souhaitons obtenir une approximation de l'aire de l'ensemble

$$D = \{(x, y) \in [0; \infty[; x^2 + y^2 \leq 1\}$$

qui est simplement un quart de disque de rayon 1. Son aire est théoriquement égale à $\frac{\pi}{4}$, dont la valeur approchée à 10^{-3} près est 0.785. Quitte à multiplier par 4, obtenir une valeur numérique approchée de l'aire de D revient à obtenir une valeur numérique approchée de π .

Pour ce faire, considérons une suite $(X_n)_{n \geq 1}$ de vecteurs aléatoires i.i.d. dont la loi commune est la loi uniforme sur le carré $[0; 1]^2$. Rappelons que la densité de cette loi est donnée par :

$$f(x, y) = \mathbb{1}_{[0; 1]^2}(x, y) .$$

Posons pour tout $n \geq 1$, $Y_n = \mathbb{1}_{\{X_n \in D\}}$. La v.a. Y_n vaut 1 lorsque X_n appartient au quart de disque D et 0 sinon. Les v.a. Y_n , $n \geq 1$, sont encore indépendantes et de même loi, à savoir la loi de Bernoulli de paramètre

$$p := \mathbb{E}[Y_1] = \mathbb{P}(X_1 \in D) = \frac{\text{aire}(D)}{\text{aire}([0; 1]^2)} = \frac{\pi}{4} .$$

La Loi des Grands Nombres (LGN) s'applique : avec probabilité 1, la moyenne empirique

$$\frac{1}{n} \sum_{i=1}^n Y_i \tag{3.1}$$

converge vers l'espérance $p = \mathbb{E}[Y_1] = \frac{\pi}{4}$. La quantité (3.1) représente la proportion des X_i , $1 \leq i \leq n$, appartenant à D . Terminons cet exemple par une simulation numérique. Simulons un n -échantillon (X_1, \dots, X_n) de loi la loi uniforme sur le carré $[0; 1]^2$ et vérifions si chaque réalisation X_i appartient ou non à D . Pour $n = 1000$ et $n = 100000$, la moyenne arithmétique (3.1) vaut respectivement 0.824 et 0.78708. Plus n (i.e. la taille de l'échantillon) est grand et plus l'approximation sera précise. Dans ce contexte, la LGN fournit un procédé aléatoire produisant une valeur numérique approchée de l'aire de D .

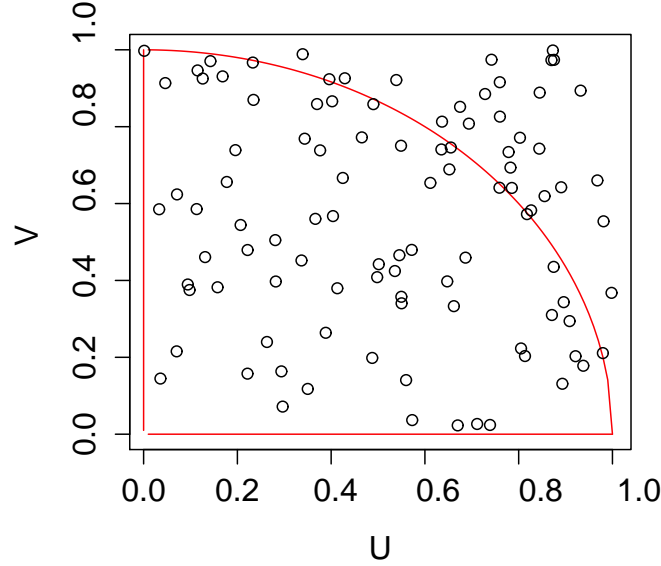


FIGURE 3.1 – Le quart de disque D est délimité en rouge. Les points représentent des réalisations de X_1, \dots, X_{100} . Certains d’entre eux sont à l’intérieur de D , d’autres non.

3.2 Méthodes MCMC

L’acronyme MCMC signifie Monte Carlo Markov Chain, i.e. méthodes de Monte Carlo par chaînes de Markov. On parle également d’exploration markovienne ce qui – vous le verrez – a plus de sens.

Considérons une loi de probabilité $\pi = (\pi(i), i \in E)$ définie sur un ensemble fini E que l’on ne sait pas simuler directement ; soit parce qu’elle est inconnue, soit parce qu’on ne dispose pas de méthode directe ou algorithmiquement raisonnable pour le faire. L’idée (géniale !) des méthodes MCMC consiste à construire une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ sur E admettant π comme mesure stationnaire et convergeant vers π (quel que soit le point de départ $X_0 = x_0$ de la chaîne). Supposons avoir à disposition une telle chaîne de Markov $(X_n)_{n \in \mathbb{N}}$. Alors, en partant d’un élément quelconque $x_0 \in E$ et au bout d’un temps n très grand, le loi de la v.a. X_n est proche de π . Autrement dit, la valeur X_n qu’il est facile d’obtenir puisque les chaînes de Markov se simulent en général aisément, est une *réalisation approchée* de la loi π .

Il s’agit donc de construire une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$, i.e. une matrice de transition P , ayant de bonnes propriétés. Pour cela, nous utiliserons le résultat précieux ci-dessous. Quitte à restreindre l’espace d’états E , nous supposons dorénavant que la mesure de probabilité ciblée π charge chaque état : $\forall i \in E, \pi(i) > 0$.

Proposition 3.2.1 Soit $Q = (q_{i,j})_{i,j \in E}$ une matrice de transition sur un ensemble fini E vérifiant : $q_{i,j} > 0$ si et seulement si $q_{j,i} > 0$. Définissons une matrice carrée $P = (p_{i,j})_{i,j \in E}$ comme suit. Si $i \neq j$,

$$p_{i,j} = \begin{cases} q_{i,j} \min \left\{ \frac{\pi(j)q_{j,i}}{\pi(i)q_{i,j}}, 1 \right\} & \text{si } q_{i,j} \neq 0 \\ 0 & \text{sinon.} \end{cases} \quad (3.2)$$

et, pour tout i , le coefficient $p_{i,i}$ est choisi de telle sorte que

$$\sum_{j \in E} p_{i,j} = 1 .$$

Alors P est une matrice de transition et toute chaîne de Markov sur E ayant P pour matrice de transition admet π comme mesure réversible.

Démonstration Remarquons tout d'abord que par construction, pour $i \neq j$, $p_{i,j} \leq q_{i,j}$, ce qui implique pour tout $i \in E$ fixé :

$$\sum_{j, j \neq i} p_{i,j} \leq \sum_{j, j \neq i} q_{i,j} \leq 1 ,$$

car Q est une matrice de transition. Il est donc possible de choisir le coefficient diagonal $p_{i,i} \geq 0$ de telle sorte que $\sum_j p_{i,j} = 1$.

Vérifions maintenant que, pour tous indices $i \neq j$, on a $\pi(i)p_{i,j} = \pi(j)p_{j,i}$. Si $q_{i,j} = 0$ alors il en est de même pour $q_{j,i}$ par hypothèse sur Q , et les coefficients $p_{i,j}$ et $p_{j,i}$ sont tous deux nuls d'après (3.2). Dans ce cas, $\pi(i)p_{i,j} = 0 = \pi(j)p_{j,i}$. Supposons désormais que $q_{i,j} > 0$ et que

$$\frac{\pi(j)q_{j,i}}{\pi(i)q_{i,j}} \leq 1$$

(l'autre alternative se traitant exactement de la même manière). D'après (3.2), $p_{i,j}$ vaut

$$q_{i,j} \min \left\{ \frac{\pi(j)q_{j,i}}{\pi(i)q_{i,j}}, 1 \right\} = \frac{\pi(j)q_{j,i}}{\pi(i)} .$$

Par hypothèse, le rapport $\frac{\pi(i)q_{i,j}}{\pi(j)q_{j,i}}$ est plus grand que 1, ce qui implique $p_{j,i} = q_{j,i}$ d'après (3.2). On déduit de ce qui précède l'identité recherchée :

$$p_{i,j} = \frac{\pi(j)q_{j,i}}{\pi(i)} = \frac{\pi(j)p_{j,i}}{\pi(i)} .$$

■

La Proposition 3.2.1 permet donc de modifier une matrice de transition Q en une nouvelle matrice de transition P qui cette fois admet π comme mesure réversible. Ce résultat sera appliqué au cas important des mesures de Gibbs. Voir Section 3.3.

Supposons avoir construit grâce à la Proposition 3.2.1 une chaîne de Markov admettant π comme mesure réversible, donc stationnaire. Il reste toutefois à s'assurer de la convergence de cette chaîne de Markov vers π . L'espace d'états E étant fini, il s'agit de prouver l'irréductibilité et l'apériodicité de la chaîne (ou de la matrice de transition P). Remarquons que ces deux propriétés ne dépendent que de la structure de graphe du graphe de transition et non des poids.

• **Irréductibilité de P .** D'après (3.2), $q_{i,j} > 0$ implique $p_{i,j} > 0$ (pour $i \neq j$). Ainsi, la matrice P hérite de l'irréductibilité de la matrice de transition Q (si celle-ci l'est).

• **Apériodicité de P .** En combinant le fait que $q_{i,j} > 0$ implique $p_{i,j} > 0$ (pour $i \neq j$) et $p_{i,i} \geq q_{i,i}$, on obtient que l'apériodicité de Q (si tel est le cas) se transmet à P . Vérifions que

$$\begin{aligned}
 p_{i,i} &= 1 - \sum_{j,j \neq i} p_{i,j} = 1 - \sum_{j,j \neq i} p_{i,j} \mathbb{1}_{\{p_{i,j} > 0\}} = 1 - \sum_{j,j \neq i} q_{i,j} \min \left\{ \frac{\pi(j)q_{j,i}}{\pi(i)q_{i,j}}, 1 \right\} \mathbb{1}_{\{p_{i,j} > 0\}} \\
 &\geq 1 - \sum_{j,j \neq i} q_{i,j} \mathbb{1}_{\{q_{i,j} > 0\}} \\
 &= 1 - \sum_{j,j \neq i} q_{i,j} \\
 &= q_{i,i}.
 \end{aligned} \tag{3.3}$$

Mais il se peut qu'en pratique la matrice P soit apériodique sans que Q le soit (et ce sera le cas dans la Section 3.3). Supposons qu'il existe des indices i et j tels que le rapport

$$\frac{\pi(j)q_{j,i}}{\pi(i)q_{i,j}}$$

soit strictement plus petit que 1. Dès lors, l'inégalité (3.3) devient stricte : $p_{i,i} > q_{i,i} \geq 0$. L'état i est donc apériodique, tout comme la matrice P si celle-ci est irréductible.

Ainsi, l'irréductibilité et l'apériodicité de la chaîne $(X_n)_{n \in \mathbb{N}}$ dépendent du choix de la matrice Q . En général, plusieurs bons choix sont possibles. Un exemple sera donné dans la prochaine section.

Terminons cette section avec l'algorithme, dit de Metropolis, permettant de simuler la v.a. X_N en suivant la méthode proposée dans la Proposition 3.2.1.

```

Initialiser  $X = x_0$ 
 $n = 0$ 
Répéter jusqu'à  $n = N$ 
{
   $i = X$ 
  Choisir  $j$  avec probabilité  $q_{i,j}$ 
   $\rho = \frac{\pi(j)q_{j,i}}{\pi(i)q_{i,j}}$ 
  If  $(\rho \geq 1)$  then {  $X = j$  }
  else {  $U = \text{runif}(1,0,1)$ 
        If  $(U < \rho)$  then {  $X = j$  } }
   $n = n+1$ 
}
X

```

Vérifions que les transitions données par cet algorithme correspondent bien à la matrice de transition P . Lorsque $\rho \geq 1$, la chaîne saute de i vers $j \neq i$ avec probabilité $q_{i,j}$, qui vaut justement $p_{i,j}$ dans ce cas. Lorsque $\rho < 1$;

$$\begin{aligned}
 \mathbb{P}(\text{aller en } j \mid X = i) &= \mathbb{P}(U < \rho, \text{ choisir } j \mid X = i) \\
 &= \mathbb{P}(U < \rho \mid \text{choisir } j, X = i) \mathbb{P}(\text{choisir } j \mid X = i) \\
 &= \rho \times q_{i,j},
 \end{aligned}$$

qui vaut encore $p_{i,j}$ dans ce cas. Il découle de ce qui précède que la probabilité de rester en i vaut $1 - \sum_{j,j \neq i} p_{i,j}$, i.e. par définition $p_{i,i}$.

3.3 Exemple des mesures de Gibbs

Considérons un espace d'état E fini dont un élément générique σ sera appelé *configuration*. À chaque configuration σ est associée une énergie positive $H(\sigma)$. Une énergie $H(\sigma) \simeq 0$ (faible) correspond à une configuration σ qu'on s'attend à observer (σ est dite stable). Tandis qu'une énergie $H(\sigma) \gg 1$ (très grande) correspond à une configuration σ rare (σ est dite instable). Puis on souhaite définir la probabilité de la configuration σ par la formule $\mathbb{P}(\sigma) \simeq e^{-H(\sigma)}$. Ainsi $\mathbb{P}(\sigma)$ est proche de 0 quand σ est instable et proche de 1 lorsqu'elle est stable.

Cependant, pour définir une mesure de probabilité (avec une masse totale égale à 1), il faut diviser par la constante de normalisation

$$Z := \sum_{\sigma \in E} e^{-H(\sigma)} .$$

On obtient alors une mesure de probabilité sur E , appelée *mesure de Gibbs*, et définie comme suit : pour tout $\sigma \in E$,

$$\mathbb{P}(\sigma) = \frac{1}{Z} e^{-H(\sigma)} .$$

Dans de nombreuses situations intéressantes comme le modèle d'Ising ou le problème du voyageur de commerce (voir les projets correspondants), le cardinal de l'ensemble E est très grand (disons de l'ordre de 10^{12}), rendant ainsi le calcul explicite de la constante de normalisation Z algorithmiquement impossible. On est alors incapable de donner la valeur de $\mathbb{P}(\sigma)$ – alors que l'énergie $H(\sigma)$ est connue – ou de simuler la mesure de Gibbs \mathbb{P} . L'algorithme de Metropolis vient alors à notre secours et nous permet d'obtenir une réalisation approchée de la loi cible $\pi = \mathbb{P}$. En effet, pour le mettre en oeuvre, il nous suffit de connaître la loi \mathbb{P} uniquement à travers les rapports

$$\frac{\mathbb{P}(\sigma)}{\mathbb{P}(\sigma')} = e^{-(H(\sigma) - H(\sigma'))} ,$$

pour toutes configurations σ, σ' , dans lesquels la constante inconnue Z a miraculeusement disparue !

Afin d'appliquer la Proposition 3.2.1, il reste à choisir une matrice de transition Q convenable. Voici une façon de procéder. Munissons l'espace d'états E d'une structure de graphe non orientée, propre au contexte et dans laquelle E représente l'ensemble des sommets. Nous noterons alors $\sigma \sim \sigma'$ lorsque $\{\sigma, \sigma'\}$ est une arête du graphe. Nous dirons alors que les configurations σ et σ' sont *voisines*. La probabilité de transition d'une configuration σ vers une autre configuration σ' sera donnée par la formule :

$$q_{\sigma, \sigma'} = \frac{1}{\text{Card}\{\eta \in E, \eta \sim \sigma\}}$$

si $\sigma \sim \sigma'$ et $q_{\sigma, \sigma'} = 0$ sinon. La dynamique de la matrice de transition Q peut être décrite comme suit : venant de σ , on choisit la configuration suivante uniformément parmi les voisines de σ . Ce choix garantit la condition de la Proposition 3.2.1, à savoir $q_{\sigma, \sigma'} > 0$ si et seulement si $q_{\sigma', \sigma} > 0$. Remarquons enfin que la matrice Q est irréductible dès que la structure de graphe sur E définit un graphe connexe.

Annexe : quelques formules

Voici quelques formules ultra-classiques du calcul des probabilités dont vous ne pourrez pas vous passer durant ce cours.

On se donne un triplé $(\Omega, \mathcal{F}, \mathbb{P})$ où :

- Ω est un ensemble ;
- \mathcal{F} est une tribu sur Ω , i.e. une très grande famille de sous-ensembles de Ω ;
- \mathbb{P} est une mesure de probabilité sur Ω : les seuls sous-ensembles de Ω pouvant être mesurés par \mathbb{P} sont ceux de la tribu \mathcal{F} .

Dans la suite, A , B et $(A_n)_{n \in \mathbb{N}}$ sont dans la tribu \mathcal{F} .

[A1] On a toujours

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$$

avec égalité dès que les événements A_n , $n \in \mathbb{N}$, sont deux à deux disjoints.

[A2] Supposons que la famille $(A_n)_{n \in \mathbb{N}}$ forme une partition de Ω , i.e. les A_n , $n \in \mathbb{N}$, sont deux à deux disjoints et leur union remplit tout Ω . Alors ;

$$\mathbb{P}(B) = \sum_{n \in \mathbb{N}} \mathbb{P}(B \cap A_n) = \sum_{n \in \mathbb{N}} \mathbb{P}(B | A_n) \mathbb{P}(A_n) .$$

[A3] Les événements A et B sont indépendants si et seulement si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$, ce qui s'écrit encore

$$\mathbb{P}(A | B) = \mathbb{P}(A) \text{ dès que } \mathbb{P}(B) > 0.$$

[A4] Si la suite $(A_n)_{n \in \mathbb{N}}$ est croissante, i.e. $A_n \subset A_{n+1}$, alors

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) .$$

[A5] Si la suite $(A_n)_{n \in \mathbb{N}}$ est décroissante, i.e. $A_n \supset A_{n+1}$, alors

$$\mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) .$$

[A6] La probabilité d'un événement est égale à l'espérance de son indicatrice :

$$\mathbb{P}(A) = \mathbb{E}[\mathbb{1}_A] .$$

Bibliographie

Ce cours s'inspire fortement des cours de Daniel Flipo

<http://daniel.flipo.free.fr/cours/markov.html>

<http://daniel.flipo.free.fr/cours/ffa.html>

et du livre de Bernard Ycart :

Modèles et algorithmes markoviens, Springer, 2002.