

# Chapitre 1

## Modélisation markovienne.

### 1.1 Introduction

Un processus aléatoire est un phénomène dont une partie de l'évolution temporelle est aléatoire. On rencontre ces processus dans divers domaines de la physique, ou des sciences de l'ingénieur. Par exemple la répartition et l'évolution de la chaleur dans un corps, la turbulence atmosphérique; c'est aussi le cas des temps d'arrivée d'appels téléphoniques, les erreurs de mesures dues aux perturbations thermiques dans des capteurs électroniques de type radar ou sonar, ou encore l'évolution des cours de marchés boursiers.

La théorie de processus aléatoires est une théorie mathématique très riche, offrant de nombreuses interactions avec diverses branches des mathématiques, telles la théorie des graphes, l'analyse fonctionnelle, la théorie des opérateurs, ainsi que la théorie ergodique des systèmes dynamiques.

Formellement, un processus aléatoire est une succession de variables aléatoires  $(X_n)_{n \geq 0}$

$$X_0 \longrightarrow X_1 \longrightarrow \dots \longrightarrow X_n \longrightarrow X_{n+1} \longrightarrow \dots$$

L'indice  $n \in \mathbb{N}$  représente la paramètre temporel. Les variables  $X_n$  peuvent être à valeurs dans un espace discret  $E$ , dans des espaces euclidiens  $E = \mathbb{R}^d$ , ou dans des espaces plus complexes tels des espaces de chemins ou d'excursions.

Ces modèles trajectoriels sont utiles en biologie dans l'analyse d'évolutions ancestrales de population. On rencontre aussi des processus à valeurs dans des espaces de matrices en physique statistique, ou en analyse d'images. L'étude de tels modèles sort bien entendu du cadre de ce cours.

Les états  $X_n$  peuvent évoluer aléatoirement au cours du temps selon des mécanismes plus ou moins complexes. Dans certains cas, les épreuves  $X_n$  sont indépendantes; c'est le cas des séquences de lancers de dés, ou les successions de jets de pièces de monnaies.

Dans d'autres situations, le processus aléatoire est donné par une équation physique récursive de la forme suivante

$$X_n = F_n(X_{n-1}, U_n)$$

La v.a.  $X_0$  désigne la condition initiale du système,  $F_n$  des fonctions de dérive déterministes, et enfin  $U_n$  des variables aléatoires "indépendantes" des précédents états  $X_0, \dots, X_{n-1}$ . De telles séquences sont appelées des *chaînes de Markov*.

En traitement du signal, de tels systèmes peuvent représenter l'évolution temporelle d'un cible dans l'espace. Dans ce contexte, les variables aléatoires  $U_n$  ont une double dimension. Elles correspondent à la fois, aux erreurs de modélisation, ainsi qu'aux stratégies inconnues de guidage.

En mathématiques financières, l'évolution des prix d'actifs dans des marchés boursiers sont aussi modélisés par des chaînes de Markov. Dans ce contexte, les variables  $U_n$  représentent les fluctuations, et la volatilité stochastique du marché financier.

Nous reviendrons sur des exemples plus précis dans la suite du cours.

## 1.2 Chaînes de Markov discrètes

Les chaînes de Markov les plus élémentaires sont bien entendu celles dont les états aléatoires ne prennent qu'un nombre fini, ou au plus dénombrable de valeurs. Cette section est consacrée à l'étude de ces chaînes élémentaires. Nous insisterons sur les réalisations "canoniques", et dynamiques, de ces processus

aléatoires. La seconde partie de cette section concerne l'étude des semigroupes d'évolution des ces chaînes de Markov.

Afin de satisfaire la curiosité du lecteur, et souligner les points d'interaction avec d'autres domaines scientifiques, nous soulignerons les interactions entre la théorie des chaînes de Markov, l'analyse fonctionnelle, l'algèbre matricielle, et la théorie des graphes. Enfin, nous illustrerons ces nouvelles notions sur une variété d'exemples précis liés à des phénomènes aléatoires issus de la physique, de la biologie, et des sciences de l'ingénieur.

Dans la suite de cette section,  $X_n$  désigne une chaîne de Markov discrète à valeurs dans un espace  $E$ , au plus dénombrable. Ses transitions de probabilités seront données par une **collection**  $(M_n(x, \cdot))_{x \in E}$  de mesures de probabilités sur  $E$ .

Autrement dit, les probabilités de passage d'un état  $X_{n-1} = x$  à un nouvel état aléatoire  $X_n$  sont données par l'application suivante

$$y \in E \mapsto M_n(x, y) = \mathbb{P}(X_n = y \mid X_{n-1} = x) \in [0, 1]$$

On utilise parfois la notation

$$\mathbb{P}^{X_n | X_{n-1}}(y|x) = \mathbb{P}(X_n = y \mid X_{n-1} = x)$$

On désigne par la suite  $\eta_n = \mathbb{P}^{X_n}$ , la loi de l'état  $X_n$  de la chaîne, à chacun des instants  $n \in \mathbb{N}$ .

On remarquera que la loi de la trajectoire  $(X_0, \dots, X_n)$  de l'origine jusqu'à l'instant  $n$  est alors décrite par la formule multiplicative

$$\begin{aligned} \mathbb{P}^{(X_0, X_1, \dots, X_n)}(x_0, x_1, \dots, x_n) &= \mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) \\ &= \eta_0(x_0) M_1(x_0, x_1) \dots M_n(x_{n-1}, x_n) \end{aligned}$$

pour toute trajectoire  $(x_p)_{0 \leq p \leq n} \in E^{n+1}$ .

On peut clairement étendre la notion de chaîne de Markov précédente à des modèles markoviens  $X_n$  prenant leurs valeurs dans des espaces  $E_n$  liés au paramètre temporel! Dans ce contexte,  $M_{n+1}(x_n, x_{n+1})$  désigne la probabilité de passer d'un état  $x_n \in E_n$  vers un état  $x_{n+1} \in E_{n+1}$ . Plus formellement, nous avons à nouveau

$$M_{n+1}(x_n, x_{n+1}) = \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n)$$

et

$$\begin{aligned} \mathbb{P}^{(X_0, X_1, \dots, X_n)}(x_0, x_1, \dots, x_n) &= \mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) \\ &= \eta_0(x_0) M_1(x_0, x_1) \dots M_n(x_{n-1}, x_n) \end{aligned}$$

Bien que cette extension ne puisse paraître purement formelle, sans d'autres intérêts que mathématiques, cette notion apparaît naturellement dans la représentation de chaînes trajectorielles.

### 1.2.1 Semigroupes de transitions

Les probabilités de transitions  $M_n$  d'une chaîne de Markov sur un espace  $E$  sont associées à deux opérateurs naturels :

1. Le premier agit à droite sur les fonctions bornées sur  $E$ . A chacune de ces fonctions  $f$ , on associe la fonction bornée  $M_n[f]$  définie par la formule suivante

$$M_n[f](x) =_{\text{déf.}} \mathbb{E}(f(X_n) \mid X_{n-1} = x) = \sum_{y \in E} M_n(x, y) f(y)$$

2. Le second agit à gauche sur les mesures de probabilités sur  $E$ . A chacune de ces mesures  $\eta$ , on associe la mesure de probabilité  $(\eta M_n)$  définie par

$$(\eta M_n)(y) = \sum_{x \in E} \eta(x) M_n(x, y) \in [0, 1]$$

Dans ce contexte, il est aussi très utile de voir une mesure de probabilité  $\eta$  comme un opérateur sur l'ensemble des fonctions  $f$  bornées sur l'espace des états  $E$

$$\eta[f] = \sum_{x \in E} \eta(x) f(x) \in \mathbb{R}$$

Avec ce système de notations, pour les fonctions indicatrices  $f = 1_B$  d'ensembles  $B \subset E$ , on retrouve la mesure  $\eta(B)$  des ensembles  $B$

$$\eta[1_B] = \sum_{x \in B} \eta(x) = \eta(B)$$

De même, on observe que

$$M_n[1_B](x) = \mathbb{E}(1_B(X_n) | X_{n-1} = x) = M_n(x, B) = \mathbb{P}(X_n \in B | X_{n-1} = x)$$

Lorsqu'il n'y a pas de confusions, il est coutume de noter  $\eta(f)$ , et  $M_n(f)$  plutôt que  $\eta[f]$ , et  $M_n[f]$ . Dans ce système de notations, nous avons

$$\mathbb{E}(f(X_n)) = \sum_{x \in E} f(x) \eta_n(x) = \eta_n(f)$$

et

$$\mathbb{E}(f(X_n) | X_{n-1} = x) = \sum_{y \in E} M_n(x, y) f(y) = M_n(f)(x)$$

En utilisant la formule des conditionnements emboîtés on montre que

$$\begin{aligned} \eta_n(f) &= \mathbb{E}(\mathbb{E}(f(X_n) | X_{n-1})) \\ &= \mathbb{E}(M_n(f)(X_{n-1})) \\ &= \sum_{x \in E} \eta_{n-1}(x) M_n(f)(x) = \eta_{n-1}(M_n(f)) \end{aligned}$$

Autrement dit, les lois  $\eta_n$  des différents états de la chaîne de Markov  $X_n$  peuvent "se calculer" récursivement. Ces dernières sont solution d'un système dynamique discret à valeurs dans l'espace des probabilités :

$$\eta_n = \eta_{n-1} M_n$$

Notre prochain objectif est de décrire plus précisément le semigroupe d'évolution du flot de mesures  $(\eta_n)_n$ . On introduit pour cela la probabilité conditionnelle de  $X_n$  en  $X_p$

$$M_{p,n}(x, y) = \mathbb{P}^{X_n|X_p}(y|x) = \mathbb{P}(X_n = y \mid X_p = x)$$

avec  $0 \leq p \leq n$ , Comme précédemment, on associe à toute fonction  $f$  sur  $E$ , la fonction  $M_{p,n}(f)$  sur  $E$  donnée par

$$M_{p,n}(f)(x) = \mathbb{E}(f(X_n)|X_p = x) = \sum_{y \in E} M_{p,n}(x, y) f(y) \quad (1.1)$$

En particulier pour  $p = n$ , on a  $M_{n,n}(f) = f$ ; et pour  $p = (n - 1)$ , nous avons

$$M_{n-1,n}(f)(x) = M_n(f)(x) = \mathbb{E}(f(X_n)|X_{n-1} = x)$$

D'après la formule de conditionnements emboîtés, nous avons la formule de récurrence

$$\begin{aligned} M_{p,n}(f)(x) &= \mathbb{E}(\mathbb{E}(f(X_n)|X_{n-1})|X_p = x) \\ &= \mathbb{E}(M_n(f)(X_{n-1})|X_p = x) = M_{p,n-1}(M_n(f))(x) \end{aligned}$$

Par conséquent, l'opérateur de transition  $M_{p,n}$  est donné par la formule (1.1) avec

$$\begin{aligned} M_{p,n}(x, y) &= \sum_{x_{p+1} \in E, \dots, x_{n-1} \in E} M_{p+1}(x, x_{p+1}) M_{p+1}(x_{p+1}, x_{p+2}) \dots M_n(x_{n-1}, y) \end{aligned}$$

On en déduit une formulation des transitions conditionnelles  $X_p \rightsquigarrow X_n$ , en terme de compositions d'opérateurs

$$M_{p,n}(f) = M_{p+1} M_{p+2} \dots M_n(f)$$

En utilisant la formule

$$\forall p \leq n \quad \mathbb{E}(f(X_n)) = \mathbb{E}(\mathbb{E}(f(X_n)|X_p)) = \mathbb{E}(M_{p,n}(f)(X_p))$$

on en conclut que

$$\begin{aligned} \eta_n(f) &= \eta_p(M_{p,n}f) = \sum_{x \in E} \eta_p(x) \left[ \sum_{y \in E} M_{p,n}(x, y) f(y) \right] \\ &= (\eta_p M_{p,n})(f) = \sum_{y \in E} \left( \sum_{x \in E} \eta_p(x) M_{p,n}(x, y) \right) f(y) \end{aligned}$$

L'ordre des sommes étant sans importance, on simplifie les notations, et on écrit tout simplement

$$\eta_n = \eta_p M_{p,n} \quad \text{avec} \quad M_{p,n} = M_{p+1} M_{p+2} \dots M_n \quad (1.2)$$

Nous avons donc montré que les opérateurs  $M_{p,n}$ ,  $\leq p \leq n$ , correspondent au semigroupe d'évolution du flot de mesures  $(\eta_n)_{n \geq 0}$ .

Lorsque la chaîne est homogène, ces opérateurs correspondent à la composition du même opérateur de transition  $M$ . Dans cette situation, on utilise souvent les notations synthétiques suivantes

$$M_{p,p+n} = M^n \quad \text{et} \quad \eta_n = \eta_0 M^n \quad \text{avec} \quad M^n = M^{n-1} M = M M^{n-1}$$

### 1.2.2 Processus historique

On considère une chaîne de Markov discrète  $X'_n$  à valeurs dans un espace au plus dénombrable  $E'$ . On note  $M'_n(x', y')$ , la probabilité de passage de  $X_{n-1} = x'$  vers  $X_n = y'$ . La séquence de trajectoires aléatoires

$$X_n = (X'_0, \dots, X'_n) \in E_n =_{\text{déf.}} (E')^{n+1}$$

forme à nouveau une chaîne de Markov à valeurs dans les espaces trajectoriels  $E_n$ . En effet, le passage de  $X_n$  à  $X_{n+1}$  s'effectue en deux temps. On conserve tout d'abord le segment de trajectoire  $X_n = (X'_0, \dots, X'_n)$ , puis on lui adjoint une extension élémentaire  $X'_{n+1} = x'$  de loi  $M'_{n+1}(X'_n, x')$ . Autrement dit, nous avons

$$\begin{aligned} X_n = (X'_0, \dots, X'_n) &\rightsquigarrow X_{n+1} = ((X'_0, \dots, X'_n), X'_{n+1}) \\ &= (\underbrace{X_n}_{(X'_0, \dots, X'_n)}, X'_{n+1}) \in E_{n+1} = (E_n \times E') \end{aligned}$$

On notera que pour toute fonction  $f_{n+1}$  bornée sur  $E_{n+1}$ , et pour tout segment de trajectoire  $x_n = (x'_0, \dots, x'_n) \in E_n$ , nous avons

$$\begin{aligned} & \mathbb{E}(f_{n+1}(X_{n+1}) \mid X_n = x_n) \\ &= \mathbb{E}(f_{n+1}([X'_0, \dots, X_n], X'_{n+1}) \mid (X'_0, \dots, X'_n) = (x'_0, \dots, x'_n)) \\ &= \sum_{x'_{n+1} \in E'} f_{n+1}([x'_0, \dots, x'_n], x'_{n+1}) M'_{n+1}(x'_n, x'_{n+1}) \end{aligned}$$

Cette équation s'exprime aussi sous la forme trajectorielle suivante

$$\begin{aligned} & \mathbb{E}(f_{n+1}(X_{n+1}) \mid X_n = x_n) \\ &= \sum_{(y'_0, \dots, y'_{n+1}) \in E_{n+1}} f_{n+1}(y'_0, \dots, y'_{n+1}) 1_{(x'_0, \dots, x'_n)}(y'_0, \dots, y'_n) M'_{n+1}(y'_n, y'_{n+1}) \end{aligned}$$

La dernière formule montre que  $X_n$  est une chaîne de Markov, de probabilités de transitions  $M_{n+1}$  de  $E_n$  vers  $E_{n+1}$ , données par

$$M_{n+1}((x'_0, \dots, x'_n), (y'_0, \dots, y'_{n+1})) = 1_{(x'_0, \dots, x'_n)}(y'_0, \dots, y'_n) M'_{n+1}(y'_n, y'_{n+1})$$

### 1.2.3 Interprétation matricielle

Lorsque l'espace d'état  $E$  est, soit fini, soit indexé de façon naturelle par  $\mathbb{N}$ , ou encore par  $\mathbb{Z}$ , les semigroupes de transitions  $M_{p,n}$  définis dans la section 1.2.1 correspondent à des compositions "élémentaires" de matrices. Pour préciser cette assertion, commençons par l'exemple le plus simple, où l'espace d'état est donné par un ensemble à deux points  $E = \{1, 2\}$ . Dans ce cas, la donnée des transitions de la chaîne

$$M_n(x, y) = \mathbb{P}(X_n = y \mid X_{n-1} = x)$$

est équivalente à la donnée des matrices  $(2 \times 2)$  suivantes

$$M_n = (M_n(x, y))_{x, y \in E} = \begin{pmatrix} M_n(1, 1) & M_n(1, 2) \\ M_n(2, 1) & M_n(2, 2) \end{pmatrix}$$



Par exemple, la matrice

$$M = \begin{pmatrix} 6/7 & 1/7 \\ 1/2 & 1/2 \end{pmatrix}$$

correspond au mouvement aléatoire entre deux états  $E = \{1, 2\}$ , synthétisé par le schéma ci-dessous

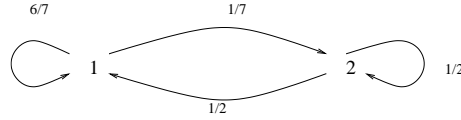


FIG. 1.1 –

Par définition de l'opérateur de transition  $M_n$ , nous avons, pour toute fonction  $f$  sur  $E = \{1, 2\}$ , la formule matricielle

$$\begin{aligned} \begin{bmatrix} M_n(f)(1) \\ M_n(f)(2) \end{bmatrix} &= \begin{bmatrix} M_n(1, 1) f(1) + M_n(1, 2) f(2) \\ M_n(2, 1) f(1) + M_n(2, 2) f(2) \end{bmatrix} \\ &= \begin{pmatrix} M_n(1, 1) & M_n(1, 2) \\ M_n(2, 1) & M_n(2, 2) \end{pmatrix} \begin{bmatrix} f(1) \\ f(2) \end{bmatrix} \end{aligned} \quad (1.3)$$

D'autre part, en utilisant la formule des conditionnements emboîtés

$$\mathbb{E}(f(X_n)) = \mathbb{E}(\mathbb{E}(f(X_n)|X_{n-1})) = \mathbb{E}(M_n(f)(X_{n-1}))$$

on obtient une nouvelle formule matricielle

$$\begin{aligned} [\eta_n(1), \eta_n(2)] &= \begin{bmatrix} \eta_{n-1}(1) M_n(1, 1) & \eta_{n-1}(2) M_n(2, 1) \\ \eta_{n-1}(1) M_n(1, 2) & \eta_{n-1}(2) M_n(2, 2) \end{bmatrix} \\ &= [\eta_{n-1}(1), \eta_{n-1}(2)] \begin{pmatrix} M_n(1, 1) & M_n(1, 2) \\ M_n(2, 1) & M_n(2, 2) \end{pmatrix} \end{aligned} \quad (1.4)$$

Par conséquent, si l'on représente une fonction numérique  $f$ , et une mesure de probabilité  $\eta$  sur  $E$  par les vecteurs colonnes et lignes

$$f = \begin{bmatrix} f(1) \\ f(2) \end{bmatrix} \quad \text{et} \quad \eta = [\eta(1), \eta(2)]$$

alors, les équations (1.3) et (1.4) s'expriment sous la forme d'un **semigroupe matriciel**

$$M(f) = M f \quad \text{et} \quad \eta_n = \eta_{n-1} M_n = \eta_0 M_1 \dots M_n$$

Ces interprétations matricielles restent valables pour des chaînes de Markov à valeurs dans des espaces d'états finis abstraits  $E = \{x_1, \dots, x_d\}$ . Dans ce contexte, les matrices de transitions sont données par

$$M_n = \begin{pmatrix} M_n(x_1, x_1) & \dots & M_n(x_1, x_d) \\ \vdots & \vdots & \vdots \\ M_n(x_d, x_1) & \dots & M_n(x_d, x_d) \end{pmatrix}$$

Les mesures de probabilités  $\eta$ , et les fonctions  $f$  sur  $E$  sont associées au vecteurs lignes et colonnes suivants

$$\eta = [\eta(x_1), \dots, \eta(x_d)] \quad \text{et} \quad f = \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_d) \end{bmatrix}$$

**Exemple 1.2.1** *La figure suivante*

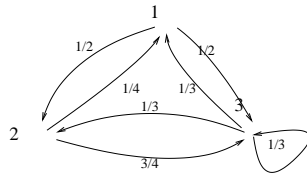


FIG. 1.2 – Chaîne à 3 points

présente un schéma d'évolution de chaîne de Markov sur un espace à trois points  $E = \{1, 2, 3\}$ , et de matrice de transition homogène donnée par

$$M = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 0 & 3/4 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Lorsque l'espace d'états est dénombrable, et indexé par les entiers positifs  $E = \{x_i, i \in \mathbb{N}\}$ , les matrices de transition sont infini-dimensionnelles

$$M_n = (M_n(x, y))_{x, y \in E} = \begin{pmatrix} M_n(x_0, x_0) & M(x_0, x_1) & M(x_0, x_2) & \dots \\ M_n(x_1, x_0) & M(x_1, x_1) & M(x_1, x_2) & \dots \\ M_n(x_2, x_0) & M(x_2, x_1) & M(x_2, x_2) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Dans ce cas, les probabilités  $\eta$ , et les fonctions  $f$  sur  $E$  sont associés au vecteurs

$$\eta = [\eta(x_0), \eta(x_1), \eta(x_2), \dots] \quad \text{et} \quad f = \begin{bmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \vdots \end{bmatrix}$$

**Exemple 1.2.2** Une marche aléatoire homogène sur  $\mathbb{N}$ , absorbée en 0, peut être représentée par le schéma suivant

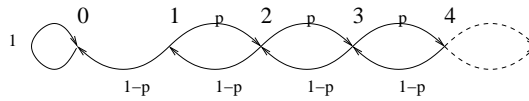


FIG. 1.3 – Chaîne absorbée en 0

La matrice de transition associée à cette chaîne est donnée par

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ (1-p) & 0 & p & 0 & \dots \\ 0 & (1-p) & 0 & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Lorsque le point 0 est réfléchissant, la matrice de transition s'écrit sous la forme

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ (1-p) & 0 & p & 0 & \dots \\ 0 & (1-p) & 0 & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Cette situation correspond au schéma suivant

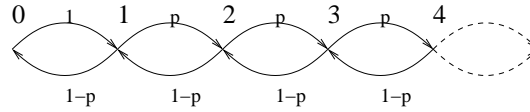


FIG. 1.4 – Chaîne réfléchie en 0

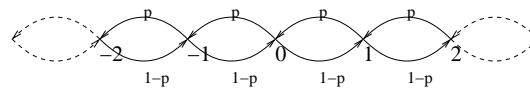
Dans le dernier cas, où l'espace d'états est dénombrable, et indexé par les entiers relatifs  $E = \{x_i, i \in \mathbb{Z}\}$ , les matrices de transition sont données par

$$M_n = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & M(x_{-1}, x_{-1}) & M_n(x_{-1}, x_0) & M(x_{-1}, x_1) & \dots \\ \dots & M(x_0, x_{-1}) & M_n(x_0, x_0) & M(x_0, x_1) & \dots \\ \dots & M(x_1, x_{-1}) & M_n(x_1, x_0) & M(x_1, x_1) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

et les probabilités  $\eta$ , et les fonctions  $f$  sur  $E$  sont associés au vecteurs

$$\eta = [\dots, \eta(x_{-1}), \eta(x_0), \eta(x_1), \dots] \quad \text{et} \quad f = \begin{bmatrix} \vdots \\ f(x_{-1}) \\ f(x_0) \\ f(x_1) \\ \vdots \end{bmatrix}$$

**Exemple 1.2.3** Une marche aléatoire simple sur  $\mathbb{Z}$  peut être représenté par le schéma suivant

FIG. 1.5 – Marche simple sur  $\mathbb{Z}$

La matrice de transition de cette évolution est donnée par

$$M = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & 0 & (1-p) & 0 & \dots \\ \dots & p & 0 & (1-p) & \dots \\ \dots & & p & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

## 1.3 Exemples

### 1.3.1 Files d'attentes

Soit  $(U_n)_{n \geq 1}$  une suite de v.a. positives et indépendantes de lois respectives  $(\mu_n)_{n \geq 1}$ . On considère la chaîne de Markov définie de façon récursive par l'équation suivante

$$\begin{cases} X_{n+1} = (X_n - 1)^+ + U_{n+1} \\ X_0 = 0 \end{cases}$$

Dans la formule précédente  $a^+ = \max(a, 0)$  désigne le maximum entre un nombre réel  $a \in \mathbb{R}$  et 0.

Ce processus aléatoire peut s'interpréter comme la longueur d'**une file d'attente**, ou encore le **temps d'attente d'un client** arrivant à un guichet, servant une personne par unité de temps. Dans ce contexte, la v.a.  $U_{n+1}$  représente le nombre de clients arrivant dans la file d'attente au temps  $(n + 1)$ .

On peut aussi interpréter  $X_n$  comme le nombre de paquets (symboles binaires représentant de l'information : voix, vidéo, données,...) en attente dans la **mémoire d'un canal de communication**, transmettant un paquet par unité de temps. Dans cette situation, la v.a.  $U_{n+1}$  représente le nombre de paquets arrivant dans le canal à l'instant  $(n + 1)$ .

On notera que les transitions de cette chaîne sont données pour tout  $i \geq 1$ , et pour tout  $j \geq 0$ , par la formule suivante

$$\mathbb{P}(X_{n+1} = (i-1) + j \mid X_n = i) = \mu_{n+1}(j) = \mathbb{P}(X_{n+1} = j \mid X_n = 0)$$

### 1.3.2 Modèle d'urne

On considère une urne contenant initialement  $B_0$  boules blanches, et  $N_0$  boules noires. A chaque instant  $n$ , on choisit au hasard une boule, puis on remet cette boule dans l'urne accompagnée d'une nouvelle boule de la même couleur. On note  $(B_n, N_n)$  le nombre de boules blanches, et noires, dans l'urne au temps  $n$ . Par construction, le couple  $X_n = (B_n, N_n)$  est une chaîne de Markov à valeurs dans  $E = \mathbb{N}^2$ , et de probabilités de transitions

$$\mathbb{P}^{X_n | X_{n-1}}(\cdot | (b, m)) = \frac{b}{m+b} \delta_{(b+1, m)} + \frac{m}{b+m} \delta_{(b, m+1)}$$

### 1.3.3 Marche aléatoire sur $\mathbb{Z}$

La marche aléatoire simple sur  $\mathbb{Z}$  correspond à un mouvement aléatoire d'une particule sur les entiers relatifs, se déplaçant soit d'un pas vers la droite, avec une probabilité  $p$ , soit d'un pas vers la gauche, avec une probabilité  $(1-p)$ , avec  $p \in (0, 1)$ . Ce mouvement peut être représenté schématiquement par la figure suivante :

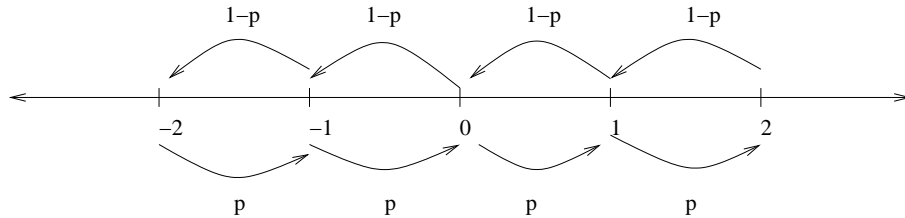


FIG. 1.6 – Marche aléatoire simple sur  $\mathbb{Z}$

Plus formellement, ce mouvement aléatoire est défini par une chaîne de Markov  $X = (X_n)_{n \geq 0}$  définie sur un espace de probabilités  $(\Omega, \mathcal{F}, \mathbb{P})$ , d'origine  $X_0 = 0$

et de probabilités de transitions homogènes

$$\begin{aligned} M(x, y) &=_{\text{déf.}} \mathbb{P}(X_n = y \mid X_{n-1} = x) = \mathbb{P}^{X_n|X_{n-1}}(y|x) \\ &= p \mathbb{1}_{x+1}(y) + (1-p) \mathbb{1}_{x-1}(y) \end{aligned}$$

Pour décrire une réalisation dynamique de cette chaîne, on se donne une suite de v.a. indépendantes  $U = (U_n)_{n \geq 1}$  distribuées sur  $\{-1, +1\}$  selon la même loi de Bernoulli

$$\mathbb{P}(U_n = +1) = 1 - \mathbb{P}(U_n = -1) = p$$

On suppose, comme d'habitude, que cette suite est définie sur un espace de probabilités  $(\Omega, \mathcal{F}, \mathbb{P})$ . On associe à  $U$ , le système dynamique aléatoire donné par

$$\begin{cases} X_n = X_{n-1} + U_n \\ X_0 = 0 \end{cases} \quad (1.5)$$

Cette interprétation dynamique offre un certain nombre d'avantages. Par exemple, elle permet une représentation "explicite" de l'état  $X_n$  en terme de v.a. indépendantes, et simplifie l'analyse des transitions de probabilités.

**Exercice 1.3.1** *Cet exercice a pour objectif d'analyser plus en profondeur la marche aléatoire sur  $\mathbb{Z}$  décrite dans l'exemple 1.3.3.*

1. *Vérifier que l'interprétation dynamique introduite en (1.5) correspond bien à la donnée d'une marche aléatoire simple sur  $\mathbb{Z}$ .*
2. *Montrer que la position moyenne de la particule au temps  $n$  est donnée par la formule  $\mathbb{E}(X_n) = n \times (2p - 1)$ . En conclure que*

$$\begin{cases} \lim_{n \rightarrow \infty} \mathbb{E}(X_n) = -\infty & \text{si } p \in [0, 1/2) \\ \mathbb{E}(X_n) = 0 & \text{si } p = 1/2 \\ \lim_{n \rightarrow \infty} \mathbb{E}(X_n) = +\infty & \text{si } p \in (1/2, 1] \end{cases}$$

3. *Vérifier que les transitions de la chaîne entre deux instants,  $m$  et  $(m+n)$ , sont données par la formule*

$$\mathbb{P}(X_{m+n} = x + [k - (n - k)] \mid X_m = x) = C_n^k p^k (1-p)^{n-k}$$

*pour tous les  $k \in \{0, \dots, n\}$ , et*

$$\mathbb{P}(X_{m+n} \notin \{2k - n : k = 0, \dots, n\} \mid X_m = x) = 0$$

4. En déduire que

$$\mathbb{P}(X_{m+2k} = 0 \mid X_m = 0) = \frac{(2k)!}{k!k!} (p(1-p))^k$$

En utilisant la formule de Stirling ( $k! \simeq \sqrt{2\pi k} k^k e^{-k}$ ), montrer que

$$\mathbb{P}(X_{m+2k} = 0 \mid X_m = 0) \simeq \frac{(4p(1-p))^k}{\sqrt{\pi k}} \quad (= 1/\sqrt{\pi k} \text{ si } p = 1/2)$$

### 1.3.4 Marche aléatoire sur $\mathbb{Z}^d$

On note  $|\cdot|$  la distance  $l^1$  sur  $E = \mathbb{Z}^d$  définies par

$$|x| = \sum_{i=1}^d |x^i|$$

pour tout  $x = (x^i)_{1 \leq i \leq d} \in \mathbb{Z}^d$ . On associe à une mesure de probabilité  $p$  sur l'ensemble des  $2d$  vecteurs unitaires directionnels

$$\mathcal{U} = \{u \in \mathbb{Z}^d : |u| = 1\}$$

la transition homogène

$$M(x, y) = \sum_{u \in \mathcal{U}} p(u) 1_{x+u}(y)$$

L'évolution aléatoire de la chaîne  $X_n$  associée à  $M$  est claire. A chaque étape  $n$ , la particule choisit aléatoirement un vecteur  $u \in \mathcal{U}$  avec la probabilité  $p(u)$ , et se déplace dans cette direction. Autrement dit, si  $(U_n)_{n \geq 1}$  désigne une suite de v.a. indépendantes de même loi  $p$  sur  $\mathcal{U}$ , on a une représentation dynamique de l'évolution

$$X_n = X_{n-1} + U_n = X_0 + \sum_{i=1}^n U_i$$

La figure suivante présente une réalisation d'une trajectoire aléatoire de la chaîne  $X_n$  sur  $\mathbb{Z}^2$ , d'origine  $X_0 = 0 \in \mathbb{Z}^2$ .







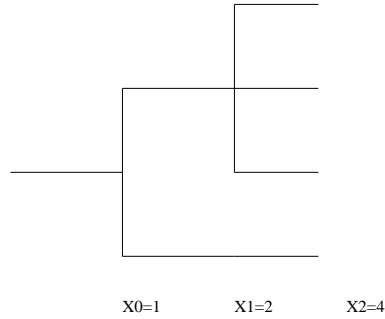


FIG. 1.9 – Processus de branchement

Examinons la situation où chaque individu se dédouble avec une probabilité  $p$ , ou disparaît avec la probabilité  $(1 - p)$ . Ce modèle correspond au choix d'une v.a.  $N$  de Bernoulli loi

$$\mathbb{P}^N = p \delta_2 + (1 - p) \delta_0$$

Dans ce contexte, on notera que  $\mathbb{E}(N) = 2p$ . De plus, on montre les transitions de cette chaîne sont données par la formule binomiale

$$\mathbb{P}(X_{n+1} = 2y \mid X_n = x) = C_x^y p^y (1 - p)^{x-y}$$

pour tout  $y \in \{0, \dots, x\}$ , et  $\mathbb{P}(X_{n+1} \in 2\mathbb{N} + 1 \mid X_n = x) = 0$ , pour tout  $x \in \mathbb{N}$ .

## 1.4 Chaînes de Markov abstraites

Comme nous l'avons souligné dans l'introduction, la notion de chaîne de Markov, est loin d'être restreinte à des phénomènes aléatoires à valeurs dans des espaces discrets. Les fluctuations de température d'un liquide par unité de temps correspondent à des chaînes de Markov à valeurs réelles. L'évolution d'une cible spatiale peut être modélisée par une chaîne de Markov à valeurs dans l'espace euclidien  $\mathbb{R}^3$ , ou bien dans  $\mathbb{R}^9$  si l'on tient compte des coordonnées de position, vitesse et accélération. Enfin, si l'on tient compte des évolutions de cette même cible entre certains paliers, on obtient une chaîne de Markov à valeurs dans des espaces de chemins ou d'excursions.

Tous ces modèles physiques, s'inscrivent dans la théorie abstraite des chaînes de Markov à valeurs dans des espaces mesurables. Leur analyse s'exprime de façon naturelle dans la théorie de l'intégration de Lebesgue. Afin d'éviter un langage

trop technique, j'ai volontairement choisi de ne pas inscrire ce cours d'ingénierie stochastique dans ce cadre trop mathématique. Nous ignorerons donc les notions d'ensembles et de fonctions mesurables, les théorèmes de Fubini, et d'autres propriétés fondamentales de la théorie de l'intégration de Lebesgue. La nature est suffisamment stable pour polir les erreurs et les confusions dues à de tels écarts. J'ai plutôt essayé de coller au plus près à la réalité scientifique et technique, tout en aiguisant la curiosité mathématique du lecteur.

Dans la suite du cours, on utilisera donc la terminologie "chaîne de Markov abstraite" pour désigner une chaîne de Markov  $X_n$  à valeurs dans des ensembles suffisamment réguliers  $E_n$ , de transitions de probabilités

$$M_n(x_{n-1}, dx_n) =_{\text{déf.}} \mathbb{P}^{X_n|X_{n-1}}(dx_n|x_{n-1}) = \mathbb{P}(X_n \in dx_n \mid X_{n-1} = x_{n-1})$$

et de loi initiale  $\eta_0(dx_0) = \mathbb{P}^{X_0}(dx_0)$  sur  $E_0$ . On note alternativement

$$\eta_n(dx_n) = \mathbb{P}^{X_n}(dx_n) = \mathbb{P}(X_n \in dx_n)$$

la loi de la v.a.  $X_n$  sur  $E_n$ , donnée pour toute fonction  $f_n$  bornée sur  $E_n$  par la formule

$$\mathbb{E}(f_n(X_n)) = \int f_n(x_n) \eta_n(dx_n) =_{\text{déf.}} \eta_n(f_n)$$

Dans ce système de notations abusives, une mesure donnée sur  $E_n$  s'interprète, soit comme un opérateur intégral, soit comme une fonction ensembliste. Ainsi, pour tout sous ensemble suffisamment régulier  $B_n \subset E_n$ , nous avons les représentations équivalentes

$$\begin{aligned} \eta_n(B_n) &= \mathbb{P}^{X_n}(B_n) = \eta_n[1_{B_n}] \\ M_n(x_{n-1}, B_n) &= \mathbb{P}^{X_n|X_{n-1}}(B_n|x_{n-1}) \\ &= \mathbb{P}(X_n \in B_n \mid X_{n-1} = x_{n-1}) = M_n[1_{B_n}](x_{n-1}) \end{aligned}$$

On dit qu'une chaîne de Markov **réelle**  $X = (X_n)_{n \geq 0}$  est **absolument continue**, lorsque la v.a. initiale  $X_0$  est absolument continue de densité  $p_0$  sur  $\mathbb{R}$ , et lorsque ses transitions de probabilités sont données en terme d'une famille de densités de probabilités  $\{p_n(x, \cdot), x \in \mathbb{R}\}$  sur  $\mathbb{R}$

$$M_n(x_{n-1}, dx_n) = \mathbb{P}^{X_n|X_{n-1}}(dx_n|x_{n-1}) = p_n(x_{n-1}, x_n) dx_n \quad (1.6)$$

Dans la formule précédente  $dx_n$  désigne la mesure de Lebesgue sur  $\mathbb{R}$ . A titre illustratif, on pourra considérer les familles de densités gaussiennes, ou

exponentielles données par la formule

$$p_n(x, y) = \frac{e^{-\frac{1}{2}(y-x)^2}}{\sqrt{2\pi}} \quad \text{ou} \quad p_n(x, y) = 1_{[0, \infty)}(y) (|x| + 1) e^{-(|x|+1)y}$$

Dans ce contexte, la trajectoire  $(X_0, \dots, X_n)$  de l'origine jusqu'à l'instant  $n$  est une v.a. absolument continue, de loi donnée par la formule

$$\mathbb{P}^{(X_0, \dots, X_n)}(d(x_0, \dots, x_n)) = p_0(x_0)p_1(x_0, x_1) \dots p_n(x_{n-1}, x_n) dx_0 dx_1 \dots dx_n$$

### 1.4.1 Semigroupe des transitions

Comme dans le cas des chaînes discrètes, il existe un système de notations naturel permettant de décrire les semigroupes d'évolution des lois des états  $X_n$ .

Les probabilités de transitions  $M_n(x_{n-1}, dx_n)$  permettent de définir deux opérateurs intégraux naturels :

1. Le premier agit à droite sur les fonctions bornées. A chacune de ces fonctions  $f$ , on associe la fonction bornée  $M_n[f]$  définie par la formule suivante

$$M_n[f] : x \in \mathbb{R} \mapsto M_n[f](x) =_{\text{déf.}} \mathbb{E}(f(X_n) | X_{n-1} = x) \in \mathbb{R}$$

2. Le second agit à gauche sur les mesures de probabilités sur  $\mathbb{R}$ . A chacune de ces mesures  $\eta$ , on associe la mesure de probabilité  $(\eta M_n)$  définie par

$$(\eta M_n) : B \subset \mathbb{R} \mapsto (\eta M_n)(B) = \int \eta(dx) M_n(x, B) \in [0, 1]$$

Dans ce contexte, une mesure de probabilité  $\eta$  correspond à un opérateur intégral sur l'ensemble des fonctions  $f$  mesurables et bornées

$$\eta[f] = \int \eta(dx) f(x) \in \mathbb{R}$$

Avec ce système de notations, pour les fonctions indicatrices  $f = 1_B$  d'ensembles  $B \subset \mathbb{R}$ , nous avons  $\eta(B) = \eta[1_B]$ . Lorsqu'il n'y a pas de confusions, il est coutume de noter  $\eta(f)$  et  $M_n(f)$  plutôt que  $\eta[f]$  et  $M_n[f]$ . Cet abus de notation évident est parfois poussé à l'extrême. Certains auteurs notent parfois tout simplement  $\eta f$ , pour insister sur le fait que cette opération intégrale n'est autre qu'une extension naturelle du produit matriciel.

À la différence des chaînes discrètes, les lois  $\eta_n(dx)$  des états  $X_n$  de chaînes abstraites sont données par des équations d'évolution intégrales. Il est bien entendu hors de question d'être tenté de calculer, ou d'estimer ces formules de transport. Il convient néanmoins de souligner qu'il existe des stratégies numériques et probabilistes pour le faire! Ces techniques sont connues sous le nom de *méthodes de Monte-Carlo*, en référence au fait qu'elles sont basées sur des simulations concrètes de trajectoires aléatoires.

En utilisant la propriété de Markov, on notera que pour toute fonction  $f_{n+1}$ , mesurable et bornée sur  $E_{n+1}$ , on a

$$\begin{aligned} \mathbb{E}(f_{n+1}(X_{n+1})|X_{n-1}) &= \mathbb{E}(\mathbb{E}(f_{n+1}(X_{n+1})|X_{n-1}, X_n) |X_{n-1}) \\ &= \mathbb{E}(\mathbb{E}(f_{n+1}(X_{n+1})|X_n) |X_{n-1}) \\ &= \mathbb{E}(M_{n+1}(f_{n+1})(X_n) |X_{n-1}) \\ &= M_n[M_{n+1}(f_{n+1})](X_{n-1}) \end{aligned}$$

On note  $(M_n M_{n+1})(x_{n-1}, dx_{n+1})$  la collection de mesures sur  $E_{n+1}$ , indexée par les  $x_{n-1} \in E_{n-1}$  et définies pour tout  $B_{n+1} \subset E_{n+1}$  par la formule de composition intégrale

$$(M_n M_{n+1})(x_{n-1}, B_{n+1}) = \int_{E_n} M_n(x_{n-1}, dx_n) M_{n+1}(x_n, B_{n+1})$$

En terme d'indicatrices cette équation s'exprime sous la forme suivante

$$(M_n M_{n+1})(1_{B_{n+1}})(x_{n-1}) = \int_{E_n} M_n(x_{n-1}, dx_n) M_{n+1}(1_{B_{n+1}})(x_n)$$

Plus généralement, on a la formule de conditionnement

$$\mathbb{E}(f_{n+p}(X_{n+p})|X_{n-1}) = M_n M_{n+1} \dots M_{n+p}(f_{n+p})(X_{n-1})$$

pour toute fonction  $f_{n+p}$  bornée sur  $E_{n+p}$ , et pour tout décalage d'indice temporel  $p \geq 1$ . Les opérateurs intégraux

$$M_{n,n+p} =_{\text{déf.}} M_n M_{n+1} \dots M_{n+p}$$

avec  $n \geq 1$  et  $p \geq 1$ , forment *un semigroupe d'opérateurs intégraux*, en ce sens où

$$\forall n \leq m \leq n + p \quad M_{n,n+p} = M_{n,m}M_{m,n+p}$$

### 1.4.2 Équations de Chapman-Kolmogorov

D'après la formule des conditionnements emboîtés, nous avons

$$\begin{aligned} \eta_n(f_n) &= \mathbb{E}(\mathbb{E}(f_n(X_n)|X_{n-1})) = \mathbb{E}(M_n(f_n)(X_{n-1})) \\ &= \int_{E_{n-1}} \eta_{n-1}(dx_{n-1}) M_n(f_n)(x_{n-1}) = \eta_{n-1}(M_n(f_n)) \end{aligned} \quad (1.7)$$

Soit  $(\eta_{n-1}M_n)$  la mesure sur  $E_n$  définie, pour tout  $B_n \subset E_n$ , par la formule

$$(\eta_{n-1}M_n)(1_{B_n}) = \int_{E_{n-1}} \eta_{n-1}(dx_{n-1}) M_n(1_{B_n})(x_{n-1})$$

Par construction, nous avons les représentations équivalentes suivantes

$$\begin{aligned} \eta_n(f_n) &= \eta_{n-1}(M_n(f_n)) \\ &= \int_{E_{n-1}} \eta_{n-1}(dx_{n-1}) \left[ \int_{E_n} M_n(x_{n-1}, dx_n) f_n(x_n) \right] \\ &= \int_{E_n} \left[ \int_{E_{n-1}} \eta_{n-1}(dx_{n-1}) M_n(x_{n-1}, dx_n) \right] f_n(x_n) \\ &= (\eta_{n-1}M_n)(f_n) =_{\text{déf.}} \eta_{n-1}M_n(f_n) \end{aligned}$$

En utilisant (1.7), nous obtenons la formule de transport intégral des lois des états de la chaîne

$$\eta_n = \eta_{n-1}M_n = \eta_0M_1M_2 \dots M_n \quad (1.8)$$

Cette équation intégrale, appelée **la formule de Chapman-Kolmogorov**, permet de voir les lois  $\eta_n$  comme solution d'un système dynamique (déterministe) intégral (et donc linéaire) sur les espaces de mesures de probabilités.

### 1.4.3 Processus historique

Comme nous l'avons vu pour les chaînes discrètes dans la section 1.2.2, le cadre non homogène est utile pour représenter des modèles trajectoriels, tel le processus historique associé à une chaîne de Markov élémentaire. La construction abstraite de ces modèles trajectoriels est analogue à celle présentée à l'exemple 1.2.2. Ainsi, si  $X'_n$  est une chaîne de Markov de transitions  $M'_n$  sur des ensembles  $E'_n$ , les séquences de trajectoires

$$X_n = (X'_0, \dots, X'_n)$$

forment un processus de Markov sur les espaces produits

$$E_n = (E'_0 \times \dots \times E'_n)$$

Le passage de  $X_n$  à  $X_{n+1}$  s'effectue en deux temps. On conserve tout d'abord le segment de trajectoire  $X_n = (X'_0, \dots, X'_n)$ , puis on lui adjoint une extension élémentaire aléatoire  $X'_{n+1}$  de loi  $M'_{n+1}(X'_n, dx')$ . Plus formellement, nous avons

$$\begin{aligned} X_n = (X'_0, \dots, X'_n) &\rightsquigarrow X_{n+1} = ((X'_0, \dots, X'_n), X'_{n+1}) \\ &= (\underbrace{X_n}_{(X'_0, \dots, X'_n)}, X'_{n+1}) \in E_{n+1} = (E_n \times E') \end{aligned}$$

On notera que pour toute fonction  $f_{n+1}$  bornée sur  $E_{n+1}$ , et pour tout segment de trajectoire  $x_n = (x'_0, \dots, x'_n) \in E_n$ , nous avons

$$\begin{aligned} \mathbb{E}(f_{n+1}(X_{n+1}) \mid X_n = x_n) \\ &= \mathbb{E}(f_{n+1}([X'_0, \dots, X'_n], X'_{n+1}) \mid (X'_0, \dots, X'_n) = (x'_0, \dots, x'_n)) \\ &= \int_{x'_{n+1} \in E'} f_{n+1}([x'_0, \dots, x'_n], x'_{n+1}) M'_{n+1}(x'_n, dx'_{n+1}) \end{aligned}$$

Cette équation s'exprime aussi sous la forme trajectorielle suivante

$$\begin{aligned} \mathbb{E}(f_{n+1}(X_{n+1}) \mid X_n = x_n) \\ &= \int_{(y'_0, \dots, y'_{n+1}) \in E_{n+1}} f_{n+1}(y'_0, \dots, y'_{n+1}) \delta_{(x'_0, \dots, x'_n)}(d(y'_0, \dots, y'_n)) M'_{n+1}(y'_n, dy'_{n+1}) \end{aligned}$$



La dernière formule montre que  $X_n$  est une chaîne de Markov, de probabilités de transitions  $M_{n+1}$  de  $E_n$  vers  $E_{n+1}$ , données par la formule

$$\begin{aligned} & M_{n+1}((x'_0, \dots, x'_n), d(y'_0, \dots, y'_{n+1})) \\ &= \delta_{(x'_0, \dots, x'_n)}(d(y'_0, \dots, y'_n)) M'_{n+1}(y'_n, dy'_{n+1}) \end{aligned}$$

Ces processus historiques interviennent de façon naturelle dans divers problèmes issus de la physique, ou de la biologie. Ils offrent un cadre markovien naturel pour modéliser et analyser des évolutions aléatoires complexes, liées le plus souvent à des effets de dépendance trajectorielles. Ainsi, dans la section 2.8, ces processus historiques nous permettront de définir des modèles d'arbres généalogiques en terme d'algorithmes génétiques trajectoriels. Dans la section 2.9, nous utiliserons à nouveau ces modèles pour représenter des explorations évolutives basées sur des mécanismes de mémoire renforçant les probabilités de retours vers des sites qui ont déjà été visités.

