Chaînes de Markov :

définitions et exemples

infos pratiques :

mail : pierre-loic.meliot@u-psud.fr

page web avec notes de cours, feuilles d'exercice, etc.
↳ Google « meliot maths », en haut de la rubrique Teachings

horaires : cours 9h-10h, TD 10h15-12h15
mercredi 9 sept. ⟶ mercredi 14 octobre.

Les premiers résultats de la théorie des probabilités concernent les suites de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ iid (indépendantes et identiquement distribuées).

__objectif__ : étendre ces résultats au cas où :

" X_{n+1} ne dépend du passé $(X_0, \dots, X_n)$ qu'au travers de X_n de façon homogène en temps ".

$\rightsquigarrow$ notion de chaîne de Markov.

1. Matrices stochastiques et chaînes de Markov

$\mathcal{X}$ = ensemble fini ou dénombrable = espace des états.

Une *matrice stochastique* sur $\mathcal{X}$ est une famille $(P(x,y))_{x,y\in\mathcal{X}}$ de nombres positifs telle que

$$\sum_{y\in\mathcal{X}} P(x,y) = 1 \quad \text{pour tout } x\in\mathcal{X}.$$

<u>exemple</u> : $\mathcal{X} = \{1,2,3\}$, $P = \begin{pmatrix} 0 & 1/3 & 2/3 \\ 1/2 & 1/6 & 1/3 \\ 1/4 & 0 & 3/4 \end{pmatrix}$.

On représentera une probabilité $\pi = (\pi(x))_{x\in\mathcal{X}}$ sur $\mathcal{X}$ par un vecteur ligne $(\pi(x_1), \pi(x_2), \dots)$

P matrice stochastique $\Longrightarrow$ chaque ligne est une probabilité.

Si π est une probabilité $\left(\sum_{x\in\mathcal{X}} \pi(x) = 1\right)$, alors πP est une probabilité.
P une matrice stochastique sur $\mathcal{X}$

$$\sum_{x\in\mathcal{X}} (\pi P)(x) = \sum_{x,y} \pi(y) P(y,x) \underset{\substack{\uparrow \\ P \text{ stochastique}}}{=} \sum_{y} \pi(y) \underset{\substack{\uparrow \\ \pi \text{ probabilité}}}{=} 1$$

Fixons $\left| \begin{array}{l} P \text{ matrice stochastique sur } \mathcal{X}. \\ \pi_0 \text{ mesure de probabilités sur } \mathcal{X} \end{array} \right.$

<u>Définition</u> : Une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ de matrice P et de loi initiale π_0 est un élément aléatoire de $\mathcal{X}^{\mathbb{N}}$ (suite ou trajectoire aléatoire) telle que :

$$\mathbb{P}[X_0 = x_0, X_1 = x_1, \dots X_n = x_n] \qquad (*)$$
$$= \pi_0(x_0) P(x_0, x_1) P(x_1, x_2) \dots P(x_{n-1}, x_n)$$

pour tout $n \geq 0$, pour tous $x_0, x_1, \dots, x_n \in \mathcal{X}$.

<u>Point subtil</u> (de théorie de la mesure) :

cette formule permet de calculer toutes les probabilités de cylindres $\{X_0 = x_0, X_1 = x_1, \dots X_n = x_n\}$, mais aussi d'événements plus généraux $E \subset \mathcal{X}^{\mathbb{N}}$.

ex : $E = \{(X_n)_{n \in \mathbb{N}} \text{ est stationnaire à } x\}$

$$= \bigcup_{n=0}^{+\infty} \bigcup_{x_0, \dots x_{n-1} \in \mathcal{X}} \bigcap_{m \geq 0} C(x_0, x_1, \dots x_{n-1}, \underbrace{x, \dots, x}_{m \text{ termes}})$$

Donc $\mathbb{P}[E] = \lim_{n \to +\infty}^{\nearrow} \left(\sum_{x_0 \dots x_{n-1} \in \mathcal{X}} \lim_{m \geq 0}^{\searrow} \mathbb{P}[C(x_0, x_1 \dots x_{n-1}, x, \dots x)] \right)$

Il existe pour $\left\{ \begin{array}{l} \text{toute loi initiale } \pi_0 \\ \text{toute matrice stochastique } P \end{array} \right.$ une unique probabilité $\mathbb{P} = \mathbb{P}_{(\pi_0, P)} = \mathbb{P}_{\pi_0}$ sur $\mathcal{X}^{\mathbb{N}}$ dont les probabilités sur les cylindres sont données par $(*)$

<u>Proposition</u> : Soit $(X_n)_{n \in \mathbb{N}}$ une CM de loi $\mathbb{P}_{(\pi_0, P)}$.

Pour tout n, $\pi_n(x) = \mathbb{P}[X_n = x] = (\pi_0 P^n)(x)$.

$\pi_n = \pi_0 P^n$. (formule pour les lois marginales)

En effet :

$$\pi_n(x) = \mathbb{P}_{\pi_0}[X_n = x] = \sum_{x_0, x_1 \dots x_{n-1}} \mathbb{P}_{\pi_0}[X_0 = x_0, X_1 = x_1, \dots X_{n-1} = x_{n-1}, X_n = x]$$

$$= \sum_{x_0 \dots x_{n-1}} \pi_0(x_0)\, P(x_0, x_1)\, P(x_1, x_2) \dots P(x_{n-1}, x)$$

produits matriciels

$$= (\pi_0 P^n)(x).$$

<u>Proposition</u> : $(X_n)_{n\in\mathbb{N}}$ est une CM de loi $\mathbb{P}_{(\pi_0, P)}$ ssi $\mathbb{P}[X_0 = x_0] = \pi_0(x_0)$ et, $\forall x_0, \dots, x_{n-1}, x_n \in \mathcal{X}$,

$$\mathbb{P}[X_n = x_n \mid X_0 = x_0, X_1 = x_1, \dots X_{n-1} = x_{n-1}]$$

$$= P(x_{n-1}, x_n).$$

En effet, pour le sens direct,

$$\mathbb{P}[X_n = x_n \mid X_0 = x_0, \dots, X_{n-1} = x_{n-1}]$$

$$= \frac{\pi_0(x_0) P(x_0, x_1) \dots P(x_{n-1}, x_n)}{\pi_0(x_0) P(x_0, x_1) \dots P(x_{n-2}, x_{n-1})} = P(x_{n-1}, x_n)$$ et la réciproque procède du même calcul.

<u>Proposition</u> : Soit $f : \mathcal{X} \to \mathbb{R}$ une fonction représentée par un vecteur colonne $\begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \end{pmatrix}$. Pour tout $n \in \mathbb{N}$,

$$\mathbb{E}_{\pi_0}[f(X_n)] = \sum_{x \in \mathcal{X}} \pi_n(x) f(x) = \pi_n \times f$$

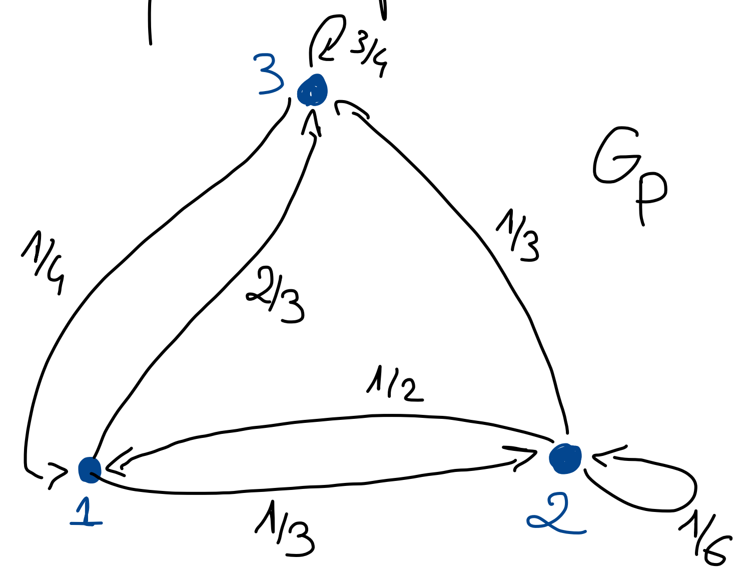
$$= \pi_0 P^n f.$$

$\Longrightarrow$ calculs de probabilités = calcul matriciel.

2. Graphe d'une chaîne et représentation avec des aléas i.i.d.

Le graphe d'une matrice stochastique P a pour sommets les états $x \in \mathcal{X}$, et une arête orientée étiquetée

$$x \xrightarrow[P(x,y)]{} y \quad \text{si} \quad P(x,y) > 0$$



Construction intuitive de $(X_n)_{n\in\mathbb{N}}$ chaîne de Markov de loi $\mathbb{P}_{(\pi_0, P)}$:

- on tire au hasard X_0 de loi π_0 (point de départ)

- au temps n, si $X_n = x$, on fait un saut / une transition vers $y = X_{n+1}$ avec probabilité $P(x,y)$, indépendamment des précédentes transitions.

Alors, $\mathbb{P}[X_0 = x_0, X_1 = x_1, \dots X_n = x_n]$

$$= \underbrace{\pi_0(x_0)}_{\text{l'état de départ est } x_0} \quad \underbrace{P(x_0, x_1)}_{\text{le premier saut va de } x_0 \text{ à } x_1} \quad \dots \quad \underbrace{P(x_{n-1}, x_n)}_{\text{le n-ième saut va de } x_{n-1} \text{ à } x_n}$$

l'état de départ est x_0

le premier saut va de x_0 à x_1

le n-ième saut va de x_{n-1} à x_n.

Théorème (représentation des chaînes de Markov)

- Soit $(\xi_n)_{n\geq 0}$ des v.a. iiD à valeurs dans un espace E

  X_0 indépendante de $(\xi_n)_{n\geq 0}$, de loi π_0 sur $\mathcal{X}$

  $f : \mathcal{X} \times E \longrightarrow \mathcal{X}$.

La suite définie par récurrence par $X_{n+1} = f(X_n, \xi_n)$ est une CM de loi initiale π_0 et matrice

$$P(x,y) = \mathbb{P}[f(x,\xi) = y].$$

- Réciproquement, étant donnés π_0, P, il existe

E, $(\xi_n)_{n\in\mathbb{N}}$ et $f : \mathcal{X} \times E \to \mathcal{X}$ telle que la construction ci-dessus fournisse une CM $(X_n)_{n\in\mathbb{N}}$ de loi $\mathbb{P}_{(\pi_0, P)}$.

Preuve du sens direct :

$$\mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n]$$

$$= \mathbb{P}[X_0 = x_0, f(X_0, \xi_0) = x_1, \dots, f(X_{n-1}, \xi_{n-1}) = x_n]$$

$$= \mathbb{P}[X_0 = x_0, f(x_0, \xi_0) = x_1, \dots, f(x_{n-1}, \xi_{n-1}) = x_n]$$

$$= \mathbb{P}[X_0 = x_0]\ \mathbb{P}[f(x_0, \xi_0) = x_1] \dots \mathbb{P}[f(x_{n-1}, \xi_{n-1}) = x_n]$$

indépendance des ξ_n et de X_0

$$= \pi_0(x_0)\ P(x_0, x_1) \dots P(x_{n-1}, x_n)$$

car les ξ_n sont toutes de même loi.

3. Exemples importants de chaînes de Markov.

- suite $(X_n)_{n\in\mathbb{N}}$ de variables i i D à valeurs dans $\mathcal{X}$

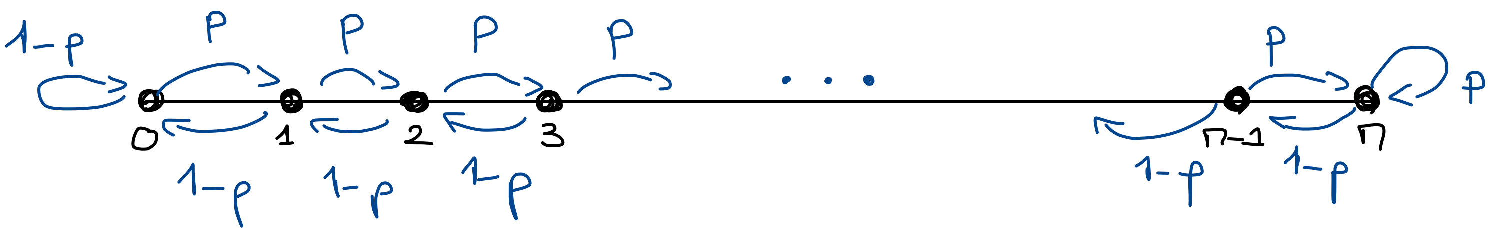
$$\mathbb{P}[X_0 = x_0, \dots X_n = x_n] = \pi(x_0)\,\pi(x_1)\dots\pi(x_n)$$

$$\Pi_0 = \pi, \quad P(x,y) = \pi(y)$$

- ruine du joueur. $\mathcal{X} = [\![0, N]\!]$

$$P(k,k+1) = p \quad \text{si } k < N \qquad\qquad P(0,0) = 1-p$$

$$P(k,k-1) = 1-p \quad \text{si } k > 0 \qquad\qquad P(N,N) = p$$



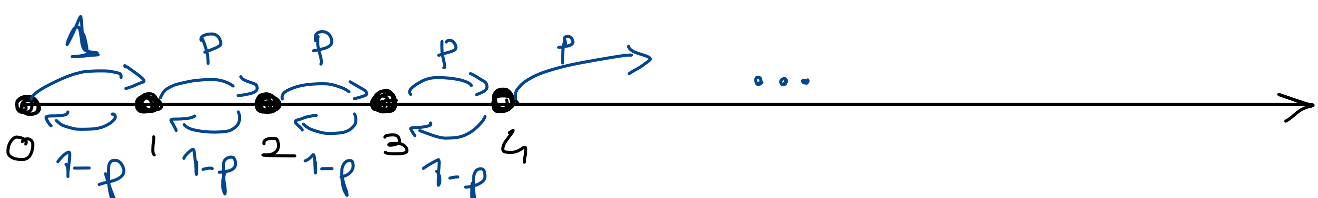
Cette chaîne modélise la quantité d'argent d'un joueur qui mise une unité à chaque tour, avec probabilité p de gagner $1-p$ de perdre.

On modifie parfois $P(0,0)$ et $P(0,1)$
$P(n,n)$ et $P(n,n-1)$

(par exemple on peut prendre $P(0,0) = 1$
$P(n,n) = 1$).

<u>question importante</u> : quelle est la probabilité d'atteindre 0 avant n ?

- file d'attente. $\mathcal{X} = \mathbb{N}$



modélise le nombre de personnes dans une file d'attente.
(guichet à la poste, serveur informatique, etc.)

<u>questions</u> : – combien de personnes en moyenne dans la file sur un long intervalle de temps ?

– peut-on avoir $X_n \xrightarrow[n \to +\infty]{} +\infty$? (explosion de la file)

- marche aléatoire sur $\mathbb{Z}$. $\mathcal{X} = \mathbb{Z}$

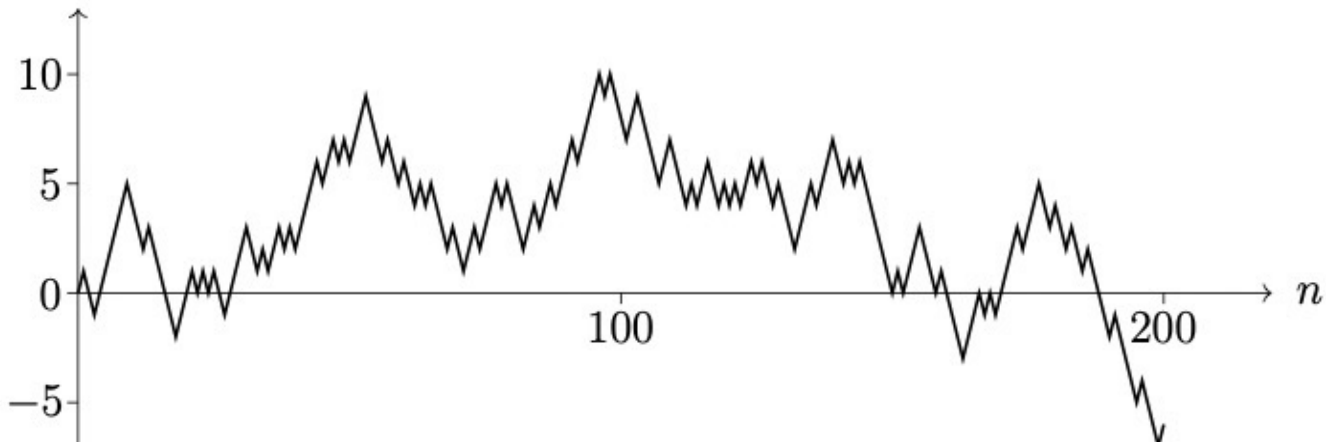
$P(k, k+1) = p$, $P(k, k-1) = 1-p$ $\forall k \in \mathbb{Z}$.

représentation si $X_0 = 0$ p.s. :

$(\xi_n)_{n \in \mathbb{N}}$ suite de v.A. iiD avec
$$\mathbb{P}[\xi_n = 1] = p$$
$$\mathbb{P}[\xi_n = -1] = 1-p$$

$X_n = \xi_1 + \xi_2 + \cdots + \xi_n$.

loi initiale : $\pi_0 = \delta_0$ (Dirac en 0). On note la loi de $(X_n)_{n\in\mathbb{N}}$ dans $\mathcal{X}^{\mathbb{N}}$ $\mathbb{P}_0$ au lieu de $\mathbb{P}_{\delta_0}$.



trajectoire $(X_n)_{0\le n\le 200}$ si $p = 1/2$.

questions : - a-t-on $X_n \underset{n \to +\infty}{\longrightarrow} +\infty$ ou $-\infty$? avec quelle probabilité ?

- si ce n'est pas le cas, quels sont les états visités ? combien de fois ?

• marche aléatoire sur un graphe.

P matrice stochastique $\longrightarrow$ G_P graphe dirigé étiqueté.

G graphe non dirigé $\longrightarrow$ matrice de transition P_G

$= (V$ sommets, E arêtes$)$

x $\qquad\qquad$ $\{x, y\}$

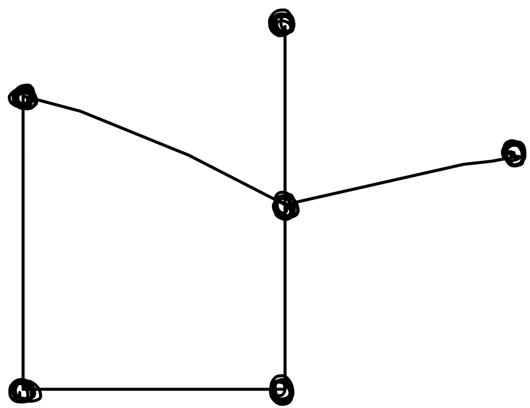
Le degré d'un sommet d'un graphe est $\deg(x) = \operatorname{card}\{y \in V \mid \{x,y\} \in E\}$

Si $\deg x \geq 1 \ \forall x \in V$:

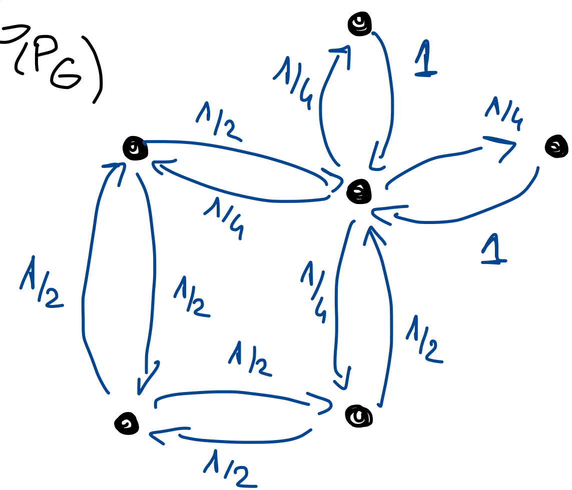
$$P_G(x,y) = \begin{cases} \dfrac{1}{\deg x} & \text{si } y \sim x \\ 0 & \text{sinon} \end{cases}$$

<u>exemple</u> :

$G =$



$G_{(P_G)}$



exemple : marche aléatoire sur le réseau $\mathbb{Z}^2$

$$V = \mathbb{Z}^2 \quad ; \quad E = \{ (x,y) \mid \|x-y\|_1 = 1 \}$$