

Chaînes de Markov

Introduction :

Dans le cours précédent, nous avons rencontré des graphes orientés. Ils vont servir dans ce chapitre à modéliser, en probabilité, des processus aléatoires qu'on appelle des chaînes de Markov.

Dans une première partie, nous définirons ce qu'est un graphe probabiliste et sa matrice de transition. Puis, en deuxième partie, nous étudierons l'évolution de certains processus aléatoires à l'aide des chaînes de Markov. Enfin, dans la troisième partie, nous nous intéresserons à la stabilisation de ces processus aléatoires.

1 Graphes probabilistes et matrices associées

a. Graphe probabiliste

Plaçons-nous dans la situation suivante.

Exemple

Dans un magasin, un produit de consommation courante est proposé suivant deux marques A et B . Une étude statistique, menée sur les clients qui achètent ce produit toutes les semaines, montre que :

- si un client achète la marque A une semaine, alors la probabilité qu'il achète la marque B la semaine suivante est 0,60 ;
- si un client achète la marque B une semaine, alors la probabilité qu'il rachète la même marque la semaine suivante est 0,30.

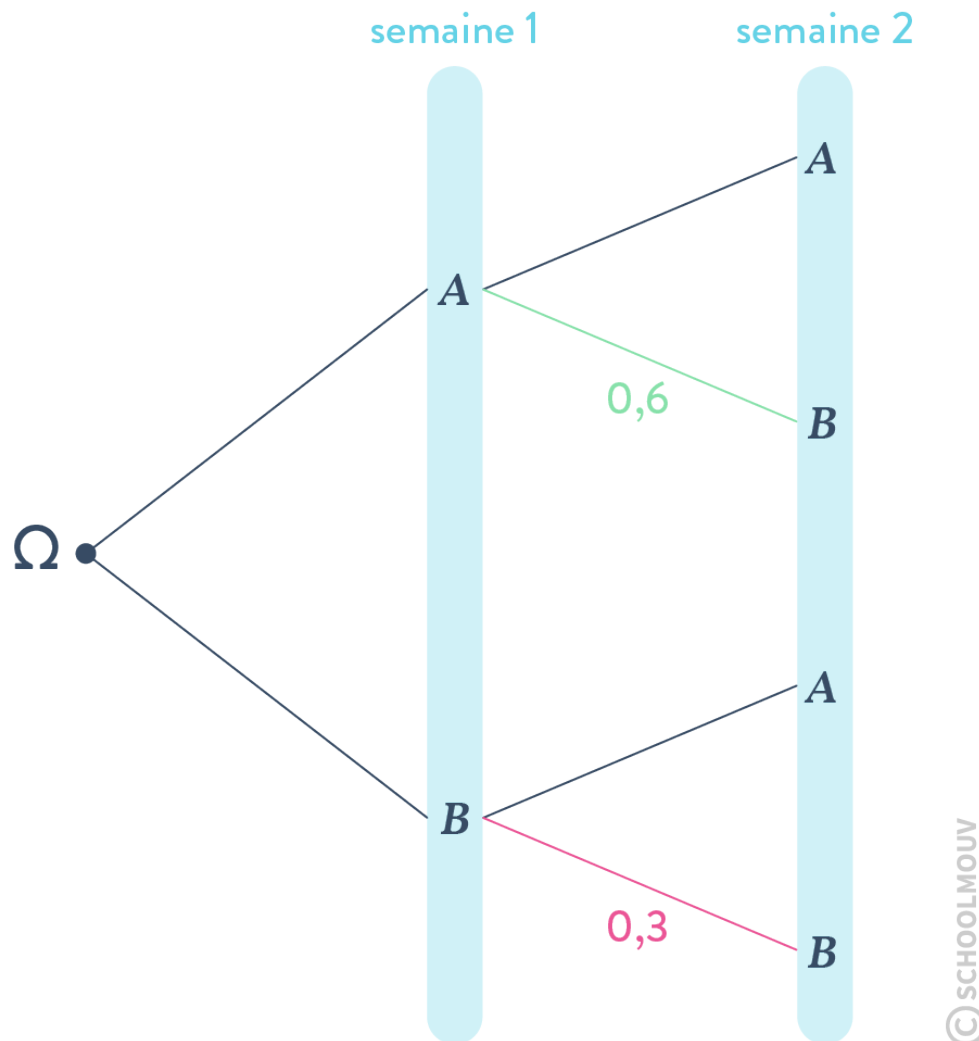
Un consommateur de ce produit est choisi au hasard et on l'interroge sur les marques qu'il choisit sur deux semaines.

→ Construisons l'arbre de probabilité représentant cette situation.

Rappelons-nous comment le construire.

Nous écrivons sur les branches de l'arbre les **probabilités conditionnelles** qui sont écrites dans l'énoncé :

- la probabilité que le client choisisse la marque B sachant qu'il a choisi la marque A la semaine précédente est $p_A(B) = 0,6$;
- la probabilité que le client choisisse la marque B sachant qu'il a choisi la marque B la semaine précédente est $p_B(B) = 0,3$.



Arbre pondéré avec les données de l'énoncé

Nous pouvons calculer les probabilités conditionnelles manquantes sur les branches de l'arbre en nous rappelant la propriété suivante.



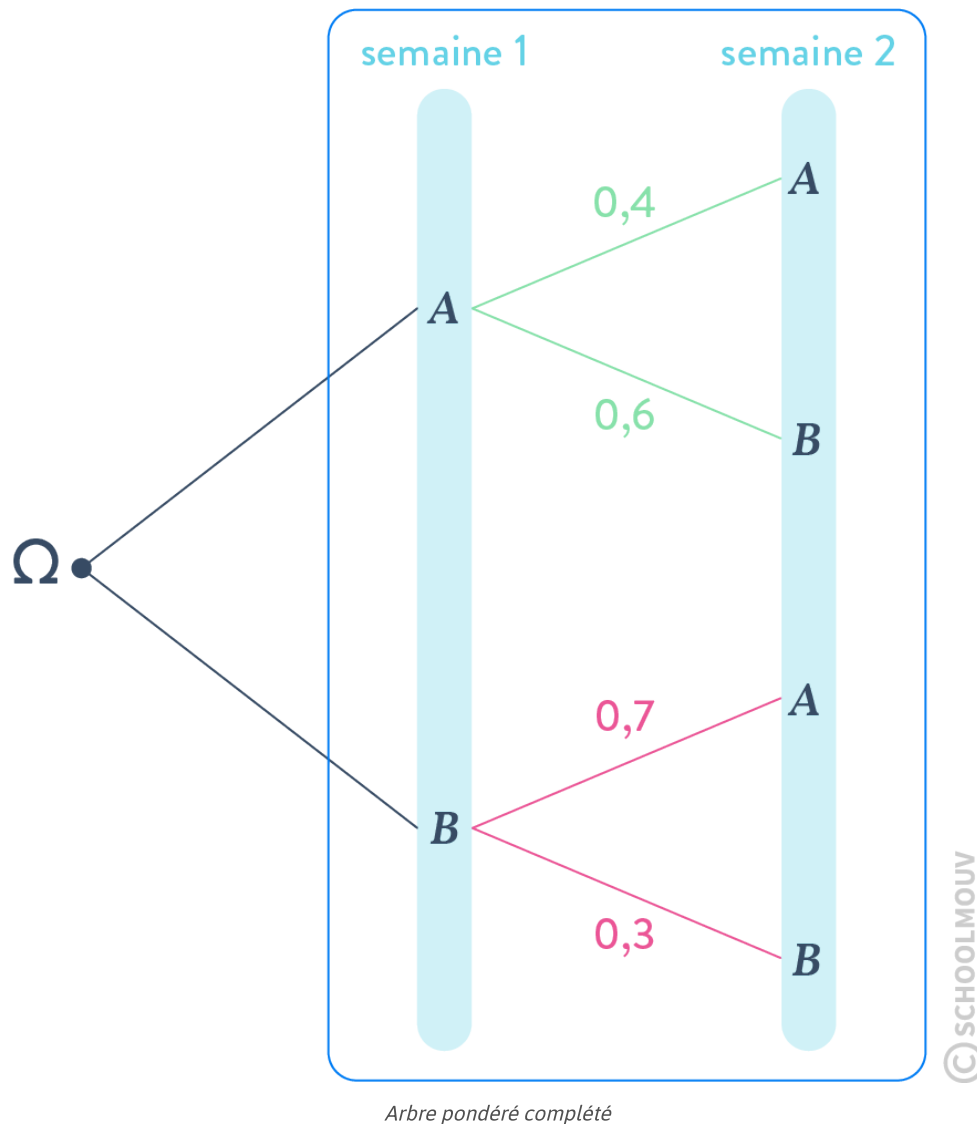
Propriété

Si A et B sont deux événements de l'univers, alors :

$$p_A(B) + p_A(\bar{B}) = 1$$

Autrement dit : quand on effectue la somme des branches issues d'un même nœud A de l'arbre, on doit trouver 1.

Nous avons donc :

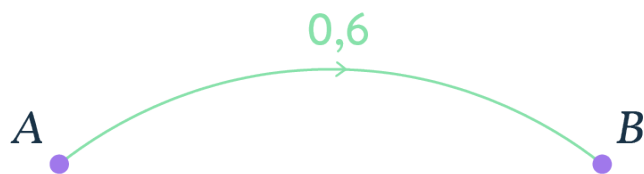


Nous allons transformer la partie encadrée de cet arbre en **graphe**. Elle représente le passage d'une semaine à la suivante.

- 1 Chaque semaine, le choix est A ou B , donc le graphe aura 2 sommets, que nous nommerons A et B également.

Ce graphe est orienté : l'arête qui part de A vers B représente la branche de l'arbre qui va de l'événement A (semaine 1) vers l'événement B (semaine 2). Nous écrivons près de l'arête le nombre $p_A(B) = 0,6$.

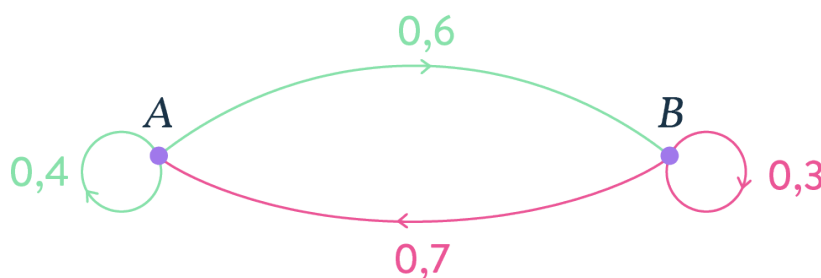
→ On dit qu'on a pondéré cette arête.



© SCHOOLMOUV

Pondération d'une arête dans un graphe orienté

- 2 Nous faisons de même avec chaque branche de l'arbre pour obtenir le graphe suivant :



© SCHOOLMOUV

Graphe orienté avec arêtes pondérées

- Nous constatons, comme dans l'arbre de probabilité, que la **somme des pondérations des arêtes qui partent de A vaut 1** ; de même, la somme des pondérations des arêtes qui partent de B vaut 1.
- Par ailleurs, il ne peut pas y avoir plus de 1 arête qui partent du sommet A vers le sommet B.

À partir de cet exemple, nous allons définir ce que sont un **graphe pondéré** et un **graphe probabiliste**.



Définition

Graphe pondéré :

On appelle graphe pondéré un graphe (orienté ou non) dont les arêtes sont affectées d'un nombre positif ou nul.

- Ces nombres s'appellent les **poids des arêtes** entre les sommets.



Définition

Graphe probabiliste :

G est un graphe d'ordre n orienté et pondéré. On numérote ses sommets par les nombres $1, 2, \dots, n$.

G est un graphe probabiliste lorsque, pour tous entiers i et j de $\{1, \dots, n\}$:

- il y a au plus 1 arête entre deux sommets i et j du graphe ;
- la somme des poids des arêtes partant d'un sommet i est égale à 1.



Astuce

Concrètement, le poids de chaque arête orientée de i vers j est la probabilité d'obtenir j sachant i .

→ C'est donc un nombre appartenant à $[0 ; 1]$.

Prenons un autre exemple avec 3 issues possibles, et sans passer par un arbre de probabilités que nous ne ferons plus désormais.



Exemple

Dans le même magasin, un autre produit est décliné selon 3 marques A , B et C . Le magasin étudie de la même manière l'achat de ce produit. Il résulte que :

- si un client achète la marque A une semaine, alors la probabilité que, la semaine suivante, il achète la marque B est 0,30, et la probabilité qu'il garde la marque A est 0,50,
- si un client achète la marque B une semaine, alors la probabilité qu'il rachète la même marque la semaine suivante est 0,45, et la probabilité qu'il achète la marque A est 0,5 ;
- si un client achète la marque C une semaine, alors la probabilité qu'il achète la marque A est 0,6, et il ne rachète pas la marque B .



À retenir

Méthodologie :

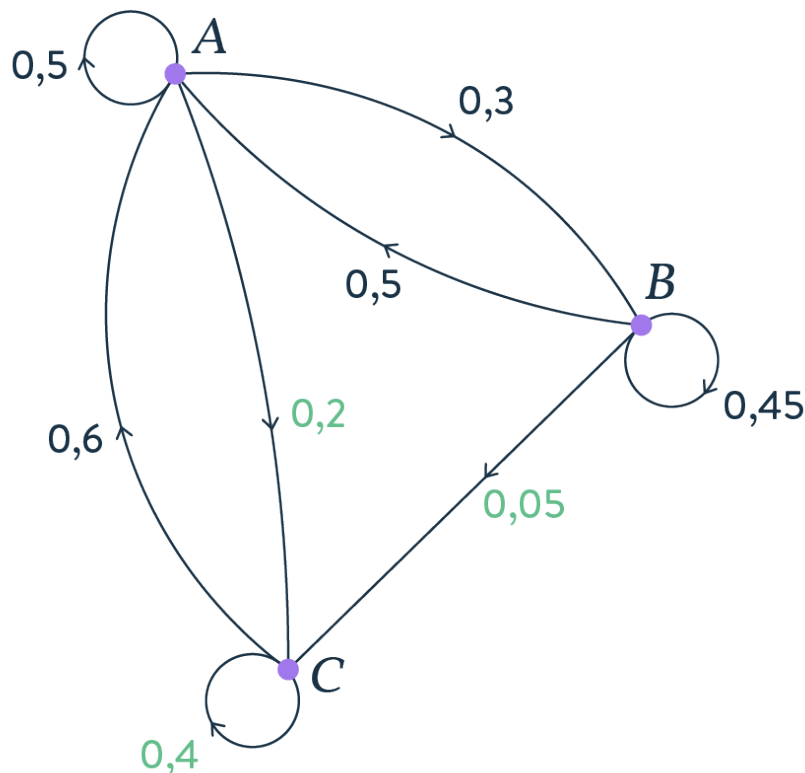
- On repère le nombre d'issues de l'univers.

→ C'est l'ordre du graphe.

- On représente chaque issue par un sommet i .
- Le poids de l'arête qui va du **sommet i vers j** est la probabilité d'**avoir j sachant i** .
- On complète le graphe probabiliste sachant que la somme des poids des arêtes partant d'un même sommet vaut 1.

Dans notre exemple, le graphe est d'ordre 3.

Débutons ce graphe avec les probabilités conditionnelles de l'énoncé que nous écrivons en noir, puis nous complétons les pondérations manquantes en vert.



Tracé du graphe probabiliste

Détaillons le cas du sommet A ; nous trouverons les autres pondérations avec la même méthode.

La somme des poids des arêtes qui **partent de A** vaut 1.

Pour trouver le poids de l'arête qui va vers C , nous effectuons simplement le calcul :

$$1 - (0,3 + 0,5) = 0,2$$

➔ Sachant qu'il a choisi la marque A la semaine précédente, la probabilité que le client choisisse la marque C est 0,2.

Comme dans le cours précédent, les graphes permettent de faire un traitement matriciel d'un problème donné. Nous allons ainsi définir dans ce cours un nouveau type de matrice associé aux graphes probabilistes.

b. Matrice de transition associée à un graphe probabiliste



Définition

Matrice de transition associée à un graphe probabiliste :

G est un graphe probabiliste d'ordre n , de sommets numérotés de 1 à n .

T est la matrice carrée d'ordre n définie par :

Pour tous i et j de $\{1, \dots, n\}$: $(T)_{i,j}$ est égal au poids de l'arête allant de i vers j .

→ T est appelée la matrice de transition du graphe probabiliste G .



Attention

La matrice de transition n'est pas la matrice d'adjacence, dont les coefficients donnent le nombre entier d'arêtes entre deux sommets.

→ Dans la matrice de transition d'un graphe probabiliste, tous les coefficients appartiennent à $[0 ; 1]$.



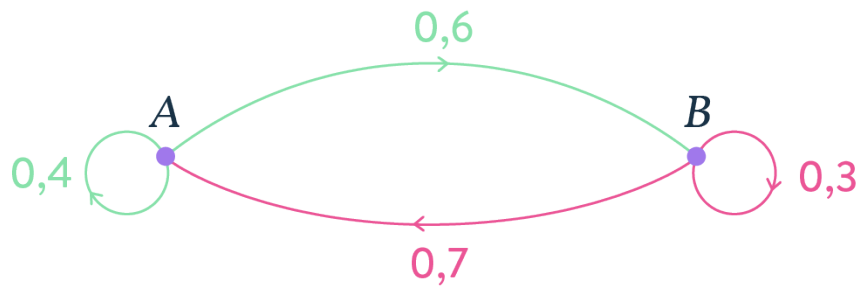
À retenir

Dans le cas d'un graphe probabiliste, le coefficient $(T)_{i,j}$ est égal à la probabilité d'obtenir l'événement représenté par le sommet j sachant l'événement représenté par i .

Pour illustrer ce qu'est une matrice de transition, reprenons les graphes du paragraphe précédent.



Exemple



© SCHOOLMOUV

Graphe orienté avec arêtes pondérées

Pour ce graphe probabiliste d'ordre 2, la matrice de transition T est carrée d'ordre 2.
Numérotons les sommets du graphe en suivant l'ordre alphabétique :

$$A \rightarrow 1 \text{ et } B \rightarrow 2$$

Appliquons la définition :

$$\begin{aligned}(T)_{1,1} &= \text{poids de l'arête allant de } A \text{ vers } A \\ &= p_A(A) \\ &= 0,4\end{aligned}$$

$$\begin{aligned}(T)_{1,2} &= \text{poids de l'arête allant de } A \text{ vers } B \\ &= p_A(B) \\ &= 0,6\end{aligned}$$

$$\begin{aligned}(T)_{2,1} &= \text{poids de l'arête allant de } B \text{ vers } A \\ &= p_B(A) \\ &= 0,7\end{aligned}$$

$$\begin{aligned}(T)_{2,2} &= \text{poids de l'arête allant de } B \text{ vers } B \\ &= p_B(B) \\ &= 0,3\end{aligned}$$

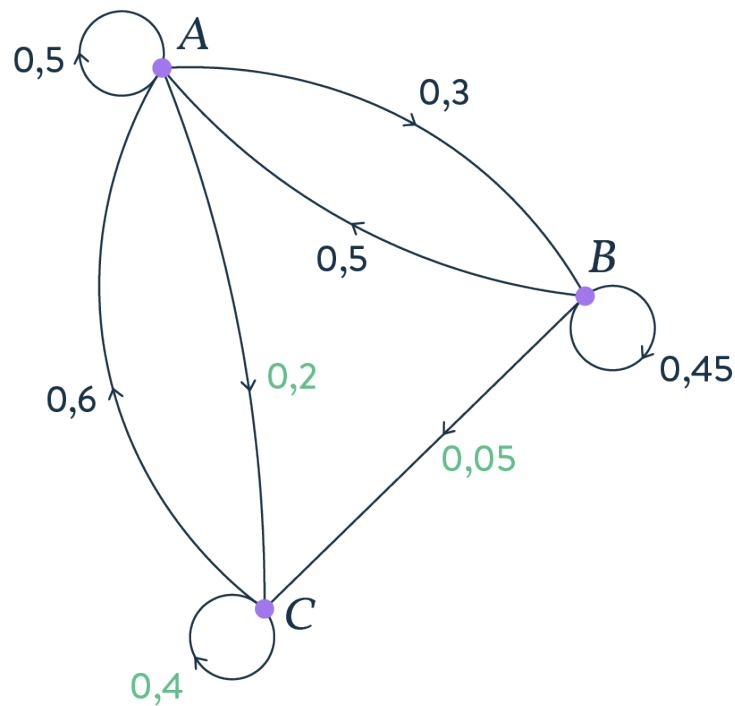
→ La matrice de transition obtenue est donc :

$$\begin{pmatrix} 0,4 & 0,6 \\ 0,7 & 0,3 \end{pmatrix}$$

Sur la première ligne de cette matrice sont écrits les poids des arêtes qui partent de A ; sur la deuxième ligne le poids des arêtes qui partent de B .

Exemple

Prenons maintenant le graphe d'ordre 3 :



Graphe probabiliste d'ordre 3

© SCHOOLMOUV

La matrice de transition est carrée d'ordre 3.

Numérotons les sommets dans l'ordre alphabétique.

→ La matrice de transition obtenue est :

$$\begin{pmatrix} 0,5 & 0,3 & 0,2 \\ 0,5 & 0,45 & 0,05 \\ 0,6 & 0 & 0,4 \end{pmatrix}$$

Sur la première ligne de cette matrice sont écrits les poids des arêtes qui partent de A ; sur la deuxième ligne, le poids des arêtes qui partent de B ; et sur la troisième ceux des arêtes qui partent de C .

De cette dernière remarque nous déduisons la propriété suivante.



Propriété

Soit G un graphe probabiliste et T la matrice de transition associée.

La somme des coefficients d'une ligne de la matrice vaut 1.

Nous avons découvert les graphes probabilistes et les matrices de transition associées, avec l'exemple concret d'une chaîne de magasins qui étudie la vente d'un produit.

Nous pouvons très bien nous figurer que ce magasin décide de prolonger l'étude pour déterminer la marque la plus achetée au bout d'un certain temps.

Les **chaînes de Markov** permettent de faire cette étude. Nous allons les définir dans la deuxième partie.

2 Chaînes de Markov

Le mathématicien russe Andreï Markov (1856-1922) a utilisé les graphes en théorie des probabilités pour étudier, dans certains cas, des processus aléatoires en fonction du temps.

→ Pour ce faire, il a utilisé des **variables aléatoires discrètes**.

Dans ce paragraphe, nous allons donc travailler avec les variables aléatoires que nous avons étudiées en première. Rappelons quelques définitions.

Rappel

Une **variable aléatoire** X sur l'univers Ω est une fonction, qui, à chaque issue de Ω , associe un nombre réel.

Soit $E \subset \mathbb{R}$.

X est une variable aléatoire sur l'univers Ω , à valeurs dans E .

Établir la **loi de la variable** X , c'est trouver, pour chaque valeur x de E , la probabilité que X soit égal à x , que l'on notera $p(X = x)$.

a. Chaîne de Markov

Dans tout ce paragraphe, nous considérons une expérience aléatoire à 2 ou 3 issues. Nous noterons l'univers Ω par $\{A, B\}$ ou $\{A, B, C\}$ suivant les cas.

Cette expérience est répétée : nous nous intéressons aux probabilités d'obtenir A ou B (ou éventuellement C) au bout d'un nombre n de répétitions.

- Dans le cas particulier $n = 0$, nous avons la première expérience.
- Pour $n = 1$, l'expérience est répétée 1 fois.
- *Et cætera*

Utilisons une variable aléatoire X_n qui associe, à chaque issue de Ω , son numéro dans la liste : $A \rightarrow 1$, $B \rightarrow 2$ ($C \rightarrow 3$), à la n -ième répétition.

Donnons deux exemples.

- L'événement $[X_3 = 2]$ est l'événement : « À la 3^e répétition de l'expérience, on obtient l'issue numérotée 2, c'est-à-dire B ».
→ Ainsi, $p(X_3 = 2)$ est la probabilité d'obtenir B à la 3^e répétition de l'expérience.
- L'événement $[X_{10} = 1]$ est l'événement : « À la 10^e répétition, on obtient l'issue numérotée 1, c'est-à-dire A ».
→ Ainsi, $p(X_{10} = 1)$ est la probabilité d'obtenir A à la 10^e répétition.

On construit ainsi une **suite de variables aléatoires** $(X_n)_{n \in \mathbb{N}}$ sur l'univers Ω .

→ Elles sont à valeurs dans $E = \{1, 2\}$ ou $E = \{1, 2, 3\}$.

Établissons maintenant le vocabulaire utile.



Définition

Espace des états :

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires définies sur un univers, et à valeurs dans un ensemble de nombres E .

→ Les valeurs de E s'appellent des **états**.

→ E porte le nom d'**espace des états**.

Pour tout $n \in \mathbb{N}$ et pour tout état i de E , l'événement $[X_n = i]$ se traduit par : « Le processus est dans l'état i à l'étape n ».

Prenons un exemple pour utiliser ce vocabulaire.



Exemple

Dans une cité scolaire, 3 activités sont proposées en association sportive : le handball ($\rightarrow 1$), le football ($\rightarrow 2$) et le volley-ball ($\rightarrow 3$).

→ Il s'agit des **états**, il y en a donc 3.

Les élèves inscrits le restent toute leur scolarité. Quand ils pratiquent une activité pendant un trimestre, ils peuvent demander une autre activité le trimestre suivant.

On interroge au hasard un élève inscrit à l'association durant sa scolarité et on lui demande quelle activité il pratique.

Soit $n \in \mathbb{N}$. On note $n + 1$ le numéro du trimestre.

Quand $n = 0$, l'élève est au 1^{er} trimestre de sa scolarité ; pour $n = 1$, il est au 2^e trimestre ; etc.

Considérons la variable aléatoire X_n qui, à chaque issue de Ω , associe le numéro du sport choisi à l'étape n . Ainsi :

- l'événement $[X_0 = 2]$ signifie que l'élève pratique le football pendant le 1^{er} trimestre ;
- l'événement $[X_5 = 3]$ signifie que l'élève a choisi le volley-ball pendant le $(5 + 1)$ -ième = 6^e trimestre.

→ On dit que $(X_n)_{n \in \mathbb{N}}$ forme une suite de variables aléatoires à valeurs dans l'**espace des états** $\{1, 2, 3\}$.

Allons un peu plus loin et considérons que, dans le processus de l'exemple précédent :

- le choix de l'élève au trimestre $(n + 1)$ dépend uniquement du choix qu'il a fait au trimestre n , et pas des choix précédant l'étape n ;
- le choix ne dépend pas de la valeur n .

Dans ces deux conditions, Markov a étudié l'évolution des variables au cours du temps, symbolisé ici par n .

→ C'est pourquoi ces suites de variables s'appellent des **chaînes de Markov**.



Définition

Chaîne de Markov :

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires à valeurs dans l'espace des états E .

$(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov sur E lorsque, pour tous i et j de E :

- l'événement $[X_{n+1} = i]$ dépend uniquement de l'état antérieur de X_n , et pas des états qui précèdent ;
- la probabilité de passer d'un état i à un état j ne dépend pas de n .



À retenir

Méthodologie :

Pour reconnaître une chaîne de Markov :

- on identifie l'espace des états de la variable aléatoire ;
- on justifie que l'état de X_{n+1} à l'étape $(n + 1)$ ne dépend que de l'état de X_n à l'étape n ; et pas des états précédents ;
- on justifie que la probabilité pour passer d'un état à un autre ne dépend pas du numéro de l'étape.

Nous concluons qu'on peut schématiser une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ sur un espace des états E par 2 étapes consécutives comme nous l'avons fait au premier paragraphe.

→ En particulier, nous allons associer un graphe probabiliste et une matrice de transition à une chaîne de Markov.

b. Graphe et matrice associés à une chaîne de Markov

Reprenons l'exemple ci-dessus.



Exemple

Les dirigeants de l'association sportive étudient les activités choisies pendant quatre ans.

Le 1^{er} trimestre (étape $n = 0$) :

- 30 % des élèves inscrits choisissent le hand,
- 60 % le foot,
- et 10 % le volley.

Puis, chaque trimestre, nous avons les éléments suivants.

- 75 % des élèves qui choisissent le handball un trimestre gardent cette activité le trimestre suivant ; 15 % d'entre eux changent pour le volley-ball ; les autres font ensuite du foot.
- 40 % des élèves qui choisissent le football un trimestre le gardent le trimestre suivant ; 15 % d'entre eux changent pour le handball, et les autres pour le volley.

- 50 % des élèves qui font du volley un trimestre changent pour le foot le trimestre suivant ;
20 % font du handball ; les autres gardent volley.

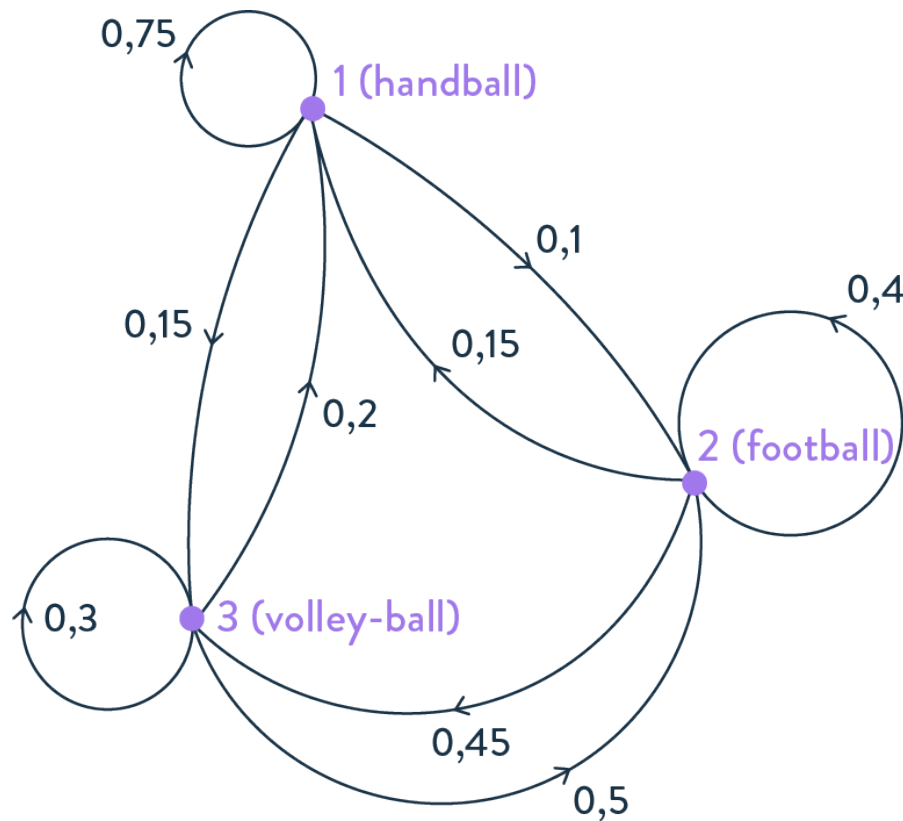
D'après l'énoncé :

- le choix d'une activité pour le trimestre $(n + 1)$ dépend uniquement de l'activité choisie le trimestre n précédent ;
- il ne dépend pas de la valeur n choisie puisque les fréquences restent stables chaque trimestre.

On note X_n la variable aléatoire qui associe à une issue de l'univers {handball, football, volley-ball} le numéro de l'activité choisie pendant le trimestre $(n + 1)$.

→ $(X_n)_{n \in \mathbb{N}}$ est bien une chaîne de Markov sur $E = \{1, 2, 3\}$.

Nous pouvons schématiser cette chaîne par le graphe probabiliste suivant qui représente $(X_n)_{n \in \mathbb{N}}$ sur deux étapes consécutives.



Représentation de la chaîne de Markov sur deux étapes consécutives

© SCHOOLMOUV

La matrice de transition associée à ce graphe est :

$$T = \begin{pmatrix} 0,75 & 0,1 & 0,15 \\ 0,15 & 0,4 & 0,45 \\ 0,2 & 0,5 & 0,3 \end{pmatrix}$$

→ On dit que c'est aussi la **matrice de transition de la chaîne de Markov**.



Définition

Matrice de transition d'une chaîne de Markov :

$(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov sur E à 2 (resp. 3) états.

La matrice de transition de cette chaîne de Markov est la matrice carrée T d'ordre 2 (resp. 3), telle que, pour tous i et j de E :

- $(T)_{i,j}$ vaut la probabilité de passer de l'état i à l'état j ;
- $(T)_{i,j}$ est le poids de l'arête allant de i vers j sur le graphe probabiliste associé.



Attention

Si cette arête n'existe pas, cette probabilité et le coefficient de la matrice associé valent 0.

Comme dans une suite de nombres, nous allons nous intéresser à présent à l'évolution des états de la chaîne de Markov suivant la valeur de n .

c. Distributions dans une chaîne de Markov

Reprenons les données de l'exemple précédent.

Le 1^{er} trimestre (étape $n = 0$), nous avons : 30 % des élèves inscrits en hand, 60 % en foot et 10 % en volley.

Ainsi :

- $p(X_0 = 1) = 0,30$;
- $p(X_0 = 2) = 0,60$;
- $p(X_0 = 3) = 0,10$.

→ Ces trois valeurs forment la loi de probabilité de X_0 ; elles portent aussi le nom de **distribution initiale**.



Définition

Distribution initiale d'une chaîne de Markov :

$(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov sur E à 2 ou 3 états.

La loi de probabilité de la variable aléatoire X_0 est appelée la distribution initiale de la chaîne $(X_n)_{n \in \mathbb{N}}$.

→ On la représente sous la forme d'une **matrice ligne** à 2 ou 3 colonnes et on la note π_0 :

- s'il y a 2 états :

$$\pi_0 = (p(X_0 = 1) \quad p(X_0 = 2))$$

- s'il y a 3 états :

$$\pi_0 = (p(X_0 = 1) \quad p(X_0 = 2) \quad p(X_0 = 3))$$

Comme nous nous intéressons à l'évolution du processus, nous posons une autre définition.



Définition

Distribution après n transitions :

$(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov sur E à 2 ou 3 états. La loi de probabilité de la variable aléatoire X_n est appelée la distribution après n transitions de la chaîne $(X_n)_{n \in \mathbb{N}}$.

→ On la représente sous la forme d'une **matrice ligne** à 2 ou 3 colonnes et on la note π_n :

- s'il y a 2 états :

$$\pi_n = (p(X_n = 1) \quad p(X_n = 2))$$

- s'il y a 3 états :

$$\pi_n = (p(X_n = 1) \quad p(X_n = 2) \quad p(X_n = 3))$$



À retenir

π_0 est la matrice ligne des probabilités d'avoir A ou B (ou éventuellement C) à la première expérience aléatoire.

De même, π_n est la matrice ligne des probabilités d'avoir A ou B (ou éventuellement C) à la n -ième étape.

Appliquons ces définitions à l'exemple de l'association sportive et déterminons les matrices π_0 et π_1 .

① Nous avons déjà la loi de probabilité de X_0 :

$$\pi_0 = (0,3 \quad 0,6 \quad 0,1)$$

② Déterminons π_1 .

Commençons par calculer $p(X_1 = 1)$.

➔ Nous cherchons la probabilité que l'élève interrogé ait choisi l'activité 1 au 2^e trimestre.

Utilisons la loi des probabilités totales :

$$\begin{aligned} p(X_1 = 1) &= p((X_0 = 1) \cap (X_1 = 1)) \\ &\quad + p((X_0 = 2) \cap (X_1 = 1)) \\ &\quad + p((X_0 = 3) \cap (X_1 = 1)) \end{aligned}$$

En effet, rappelons-nous que si l'élève choisit le hand au 2^e trimestre, il est dans l'une des trois configurations suivantes :

- il a pratiqué le hand au 1^{er} trimestre et il garde le hand au 2^e :

$$(X_0 = 1) \cap (X_1 = 1)$$

- il a pratiqué le foot au 1^{er} trimestre et il choisit le hand au 2^e :

$$(X_0 = 2) \cap (X_1 = 1)$$

- il a pratiqué le volley au 1^{er} trimestre et il choisit le hand au 2^e :

$$(X_0 = 3) \cap (X_1 = 1)$$

Pour trouver ces probabilités, nous appliquons la formule sur les probabilités conditionnelles :

$$p(A \cap B) = p(B) \times p_A(B)$$

➔ Nous avons donc :

$$\begin{aligned} p(X_1 = 1) &= p(X_0 = 1) \times p_{X_0=1}(X_1 = 1) \\ &\quad + p(X_0 = 2) \times p_{X_0=2}(X_1 = 1) \\ &\quad + p(X_0 = 3) \times p_{X_0=3}(X_1 = 1) \end{aligned}$$

Nous connaissons π_0 , donc les valeurs : $p(X_0 = 1) = 0,3$, $p(X_0 = 2) = 0,6$ et $p(X_0 = 3) = 0,1$.

Nous savons aussi que les probabilités conditionnelles se trouvent soit sur le graphe, soit dans la matrice de transition T de la chaîne de Markov. Utilisons cette dernière :

$$T = \begin{pmatrix} 0,75 & 0,1 & 0,15 \\ 0,15 & 0,4 & 0,45 \\ 0,2 & 0,5 & 0,3 \end{pmatrix}$$

- $p_{X_0=1}(X_1 = 1)$ est la probabilité de passer de l'état 1 à l'état 1.
→ C'est donc le coefficient $(T)_{1,1} = 0,75$.
- $p_{X_0=2}(X_1 = 1)$ est la probabilité de passer de l'état 2 à l'état 1.
→ C'est donc le coefficient $(T)_{2,1} = 0,15$.
- Finalement, $p_{X_0=3}(X_1 = 1)$ est la probabilité de passer de l'état 3 à l'état 1.
→ C'est donc le coefficient $(T)_{3,1} = 0,2$.

Calculons donc :

$$\begin{aligned} p(X_1 = 1) &= 0,3 \times 0,75 + 0,6 \times 0,15 + 0,1 \times 0,2 \\ &= \boxed{0,335} \end{aligned}$$



Astuce

Nous remarquons qu'en faisant le produit de la matrice π_0 par la première colonne de T , nous aurions obtenu exactement ce calcul :

$$\begin{aligned} \pi_0 \times \begin{pmatrix} 0,75 \\ 0,15 \\ 0,2 \end{pmatrix} &= (0,3 \quad 0,6 \quad 0,1) \times \begin{pmatrix} 0,75 \\ 0,15 \\ 0,2 \end{pmatrix} \\ &= 0,3 \times 0,75 + 0,6 \times 0,15 + 0,1 \times 0,2 \end{aligned}$$

Ce qui s'explique par le fait que la 1^{re} colonne de la matrice T sont les probabilités conditionnelles de passer de l'état i à l'état 1, celles dont nous avons justement besoin !

Si nous voulons continuer à déterminer π_1 , calculons maintenant $p(X_1 = 2)$.

D'après la loi des probabilités totales :

$$\begin{aligned}
p(X_1 = 2) &= p((X_0 = 1) \cap (X_1 = 2)) \\
&\quad + p((X_0 = 2) \cap (X_1 = 2)) \\
&\quad + p((X_0 = 3) \cap (X_1 = 2)) \\
&= p(X_0 = 1) \times p_{X_0=1}(X_1 = 2) \\
&\quad + p(X_0 = 2) \times p_{X_0=2}(X_1 = 2) \\
&\quad + p(X_0 = 3) \times p_{X_0=3}(X_1 = 2)
\end{aligned}$$

On retrouve les coefficients de π_0 et les coefficients placés sur la 2^e colonne de T .

→ Nous avons ainsi :

$$\begin{aligned}
\pi_0 \times \begin{pmatrix} 0,1 \\ 0,4 \\ 0,5 \end{pmatrix} &= (0,3 \quad 0,6 \quad 0,1) \times \begin{pmatrix} 0,1 \\ 0,4 \\ 0,5 \end{pmatrix} \\
&= 0,3 \times 0,1 + 0,6 \times 0,4 + 0,1 \times 0,5 \\
&= \boxed{0,32}
\end{aligned}$$

$p(X_1 = 3)$ s'obtient de même en calculant :

$$\pi_0 \times \begin{pmatrix} 0,15 \\ 0,45 \\ 0,3 \end{pmatrix} = \boxed{0,345}$$

C En conclusion, nous obtenons :

$$\pi_1 = (0,335 \quad 0,32 \quad 0,345)$$

Les calculs pour obtenir π_2 à partir de π_1 sont analogues.

Pour éviter le très long développement que pourrait occasionner le calcul de π_n pour n donné, nous allons utiliser la propriété suivante.



Propriété

$(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov sur E à 2 (ou 3) états.

Soit π_0 la matrice ligne représentant sa distribution initiale ; π_n la matrice ligne représentant sa distribution après n transitions ; et T la matrice de transition de la chaîne X_n .

- Pour tout entier naturel n : $\pi_{n+1} = \pi_n \times T$.
- Pour tout entier naturel n : $\pi_n = \pi_0 \times T^n$.

Démontrons ces propriétés en considérant un espace à 2 états, et commençons par la première. La démonstration suit le raisonnement que nous avons appliqué dans l'exemple précédent.



Démonstration

Soit $n \in \mathbb{N}$: nous cherchons à exprimer la matrice π_{n+1} en fonction de la matrice π_n .

Par définition, $\pi_{n+1} = (p(X_{n+1} = 1) \quad p(X_{n+1} = 2))$, puisque la chaîne de Markov est à valeurs dans un espace à 2 états.

→ Calculons les coefficients de π_{n+1} .

① D'après la loi des probabilités totales :

$$\begin{aligned} p(X_{n+1} = 1) &= p((X_n = 1) \cap (X_{n+1} = 1)) \\ &\quad + p((X_n = 2) \cap (X_{n+1} = 1)) \\ &= p(X_n = 1) \times p_{X_n=1}(X_{n+1} = 1) \\ &\quad + p(X_n = 2) \times p_{X_n=2}(X_{n+1} = 1) \end{aligned}$$

Nous reconnaissons les coefficients de $\pi_n = (p(X_n = 1) \quad p(X_n = 2))$ et les probabilités conditionnelles d'obtenir $X_{n+1} = 1$ sachant que $X_n = 1$ ou $X_n = 2$.

② De même, la loi des probabilités totales donne :

$$\begin{aligned} p(X_{n+1} = 2) &= p((X_n = 1) \cap (X_{n+1} = 2)) \\ &\quad + p((X_n = 2) \cap (X_{n+1} = 2)) \\ &= p(X_n = 1) \times p_{X_n=1}(X_{n+1} = 2) \\ &\quad + p(X_n = 2) \times p_{X_n=2}(X_{n+1} = 2) \end{aligned}$$

Ici encore, nous trouvons les coefficients de π_n et les probabilités conditionnelles d'obtenir $X_{n+1} = 2$ sachant que $X_n = 1$ ou $X_n = 2$.

③ Or nous savons que la matrice de transition de la chaîne de Markov est la matrice dont les coefficients sont, pour tous i et j de $\{1, 2\}$, $(T)_{i,j}$, soit la probabilité de passer de l'état i à l'état j .

$$\text{Ainsi : } p(X_{n+1} = 1) = p(X_n = 1) \times (T)_{1,1} + p(X_n = 2) \times (T)_{2,1}$$

$$\text{Et : } p(X_{n+1} = 2) = p(X_n = 1) \times (T)_{1,2} + p(X_n = 2) \times (T)_{2,2}$$

→ Nous avons donc :

$$\begin{aligned} \pi_{n+1} &= (p(X_{n+1} = 1) \quad p(X_{n+1} = 2)) \\ &= (p(X_n = 1) \quad p(X_n = 2)) \times \begin{pmatrix} (T)_{1,1} & (T)_{1,2} \\ (T)_{2,1} & (T)_{2,2} \end{pmatrix} \end{aligned}$$

Ⓒ En conclusion, nous avons donc, pour tout $n \in \mathbb{N}$:

$$\pi_{n+1} = \pi_n \times T$$

Cette relation ressemble beaucoup à celle que nous avons pour les suites réelles géométriques.

Nous allons montrer la deuxième propriété par **récurrence** sur n .

Démonstration

Pour tout entier naturel n , nous voulons montrer la propriété P_n : « $\pi_n = \pi_0 \times T^n$ ».

1 Initialisation :

$$\begin{aligned}\pi_0 \times T^0 &= \pi_0 \times I_2 \text{ [où } I_2 \text{ est la matrice identité d'ordre 2]} \\ &= \pi_0\end{aligned}$$

→ P_0 est vérifiée.

2 Hérité :

Supposons que P_k : $\pi_k = \pi_0 \times T^k$ est vraie pour un entier naturel k quelconque.

→ Nous cherchons à démontrer que, alors, P_{k+1} est vraie, c'est-à-dire :

$$\pi_{k+1} = \pi_0 \times T^{k+1}$$

D'après la propriété précédente, nous savons que :

$$\pi_{k+1} = \pi_k \times T$$

Utilisons l'hypothèse de récurrence :

$$\begin{aligned}\pi_{k+1} &= \pi_0 \times T^k \times T \\ &= \pi_0 \times T^{k+1}\end{aligned}$$

→ Ainsi, nous avons montré que, si P_k est vraie, alors P_{k+1} est vraie.

C Conclusion :

Nous avons démontré que la proposition était vraie pour $n = 0$ et que, à partir de là, elle était héréditaire.

→ Nous pouvons conclure que la proposition est vraie pour tout $n \in \mathbb{N}$:

$$\pi_n = \pi_0 \times T^n$$



Attention

La matrice $\pi_n = \pi_0 \times T^n$ se calcule en faisant le produit de π_0 par T^n dans cet ordre.

Rappelons-nous que le produit de deux matrices n'est pas commutatif, surtout quand elles n'ont pas la même dimension. Ici, on ne peut pas calculer $T^n \times \pi_0$.

Cette dernière propriété amène également la propriété suivante.



Propriété

$(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov sur E à 2 (ou 3) états.

Soit T sa matrice de transition et $n \in \mathbb{N}^*$.

Le coefficient $(T^n)_{i,j}$ situé à la i -ème ligne et la j -ième colonne de T^n est la probabilité de passer de l'état i à l'état j en n transitions.

Reprenons maintenant notre exemple de l'association sportive pour montrer l'efficacité de ces propriétés.



Exemple

Cherchons la distribution de la chaîne après 4 transitions.

D'après la propriété, nous avons $\pi_4 = \pi_0 \times T^4$ et nous connaissons :

$$\pi_0 = (0,3 \quad 0,6 \quad 0,1)$$
$$\text{et : } T = \begin{pmatrix} 0,75 & 0,1 & 0,15 \\ 0,15 & 0,4 & 0,45 \\ 0,2 & 0,5 & 0,3 \end{pmatrix}$$

La calculatrice nous permet d'obtenir, en arrondissant à 10^{-4} près :

$$T^4 = \begin{pmatrix} 0,4766 & 0,2664 & 0,2570 \\ 0,3607 & 0,3348 & 0,3046 \\ 0,3686 & 0,3298 & 0,3016 \end{pmatrix}$$

$$\begin{aligned} \text{et : } \pi_4 &= \pi_0 \times T^4 \\ &= (0,3 \quad 0,6 \quad 0,1) \times \begin{pmatrix} 0,4766 & 0,2664 & 0,2570 \\ 0,3607 & 0,3348 & 0,3046 \\ 0,3686 & 0,3298 & 0,3016 \end{pmatrix} \\ &= (0,3962 \quad 0,3137 \quad 0,2900) \end{aligned}$$

Concrètement, après 4 transitions, donc au 5^e trimestre de la scolarité des inscrits :

- la probabilité qu'un élève soit inscrit en handball sera d'environ 0,3962 ;
- celle qu'il soit inscrit en football sera d'environ 0,3137 ;
- et celle qu'il soit inscrit en volley-ball d'environ 0,29.

La matrice T^4 permet également de trouver les probabilités de passer d'un état à un autre en 4 transitions.

Prenons par exemple le coefficient $(T^4)_{2,2}$. C'est la probabilité que l'élève, initialement inscrit en football, soit inscrit en football en 4 transitions.

→ Lisons le coefficient : cette probabilité vaut 0,3348.

Dans cette partie, nous avons rencontré les chaînes de Markov qui permettent de calculer, dans certaines conditions, l'état probabiliste d'un processus.

Nous pouvons nous interroger : existe-t-il, comme pour les suites, une limite de ces chaînes, c'est-à-dire : le processus peut-il se stabiliser ? C'est ce que nous allons voir dans le prochain point.

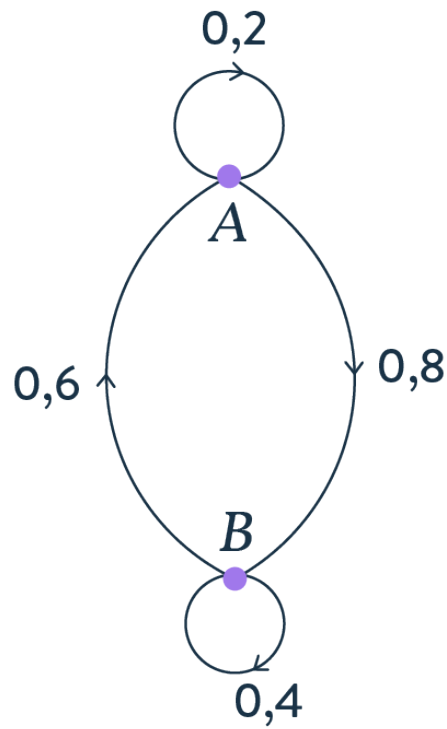
3 Distributions invariantes d'une chaîne de Markov à 2 ou 3 états

a. Distribution invariante d'une chaîne de Markov

Prenons l'exemple de la chaîne de Markov suivante.

Exemple

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à 2 états dont on donne le graphe probabiliste et la matrice de transition T ci-dessous :



© SCHOOLMOUV

Chaîne de Markov à deux états

$$T = \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix}$$

Supposons que la distribution initiale est :

$$\pi_0 = \begin{pmatrix} \frac{3}{7} & \frac{4}{7} \end{pmatrix}$$

Nous savons que $\pi_1 = \pi_0 \times T$ (nous notons les coefficients de T sous forme de fraction) :

$$\begin{aligned} \pi_1 &= \pi_0 \times T \\ &= \begin{pmatrix} \frac{3}{7} & \frac{4}{7} \end{pmatrix} \times \begin{pmatrix} \frac{1}{5} & \frac{4}{5} \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix} \\ &= \begin{pmatrix} \frac{3}{7} & \frac{4}{7} \end{pmatrix} \end{aligned}$$

→ Nous avons donc $\pi_1 = \pi_0$.

Or, pour tout entier naturel n , $\pi_{n+1} = \pi_n \times T$.

Ainsi, $\pi_2 = \pi_1 \times T = \pi_0 \times T = \pi_0$.

Et nous pouvons écrire la même chose pour tous les π_n qui suivent. Pour cette distribution initiale-là, la distribution après n transitions est toujours égale à la distribution initiale.

→ On parle de **distribution invariante**.



Définition

Distribution invariante d'une chaîne de Markov :

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à 2 (ou 3) états et T sa matrice de transition. Supposons qu'il existe π une matrice ligne à 2 (ou 3) colonnes telle que :

- $\pi = \pi \times T$;
- et la somme des coefficients de π est égale à 1.

Cette matrice est alors appelée distribution invariante de la chaîne de Markov.



À retenir

Pour prouver qu'une matrice ligne π est une distribution invariante d'une chaîne de Markov, il suffit de calculer $\pi \times T$ et de prouver que cette matrice est π .

Prenons maintenant un exemple où nous devons chercher une distribution invariante.



Exemple

Considérons une chaîne de Markov dont la matrice de transition est :

$$T = \begin{pmatrix} 0,7 & 0,2 & 0,1 \\ 0,15 & 0,7 & 0,15 \\ 0,2 & 0,5 & 0,3 \end{pmatrix}$$

Nous cherchons s'il existe une distribution invariante pour cette chaîne de Markov.



À retenir

Méthodologie :

On cherche une distribution invariante pour une chaîne de Markov : il faut résoudre l'équation matricielle (E) : $\pi = \pi \times T$:

- ① On nomme les coefficients de $\pi = (a \quad b \quad c)$ ($\pi = (a \quad b)$ s'il y a 2 états).
- ② On calcule la matrice ligne $\pi \times T$.

- ③ L'équation (E) devient un système de 3 équations à 3 inconnues (resp. de 2 équations à 2 inconnues), que l'on résout.
- Ⓒ On n'oublie pas de vérifier que la somme des coefficients de π est égale à 1.

Appliquons cette méthode à notre exemple.

- ① Notons une matrice ligne $\pi = (a \quad b \quad c)$.

➔ Nous voulons résoudre $\pi = \pi \times T$.

- ② Calculons $\pi \times T$:

$$\begin{aligned}\pi \times T &= (a \quad b \quad c) \times \begin{pmatrix} 0,7 & 0,2 & 0,1 \\ 0,15 & 0,7 & 0,15 \\ 0,2 & 0,5 & 0,3 \end{pmatrix} \\ &= (0,7a + 0,15b + 0,2c \quad 0,2a + 0,7b + 0,5c \quad 0,1a + 0,15b + 0,3c)\end{aligned}$$

- ③ $\pi = \pi \times T$, donc leurs coefficients sont deux à deux égaux.

Nous obtenons ainsi :

$$\begin{aligned}\pi = \pi \times T &\Leftrightarrow \begin{cases} a = 0,7a + 0,15b + 0,2c \\ b = 0,2a + 0,7b + 0,5c \\ c = 0,1a + 0,15b + 0,3c \end{cases} \\ &\Leftrightarrow \begin{cases} 0,3a - 0,15b - 0,2c = 0 & L_1 \\ -0,2a + 0,3b - 0,5c = 0 & L_2 \\ -0,1a - 0,15b + 0,7c = 0 & L_3 \end{cases}\end{aligned}$$

Nous allons résoudre ce système par combinaison : nous allons éliminer le a de L_1 et L_2 en utilisant L_3 , que nous conservons (nous indiquons en gris les opérations sur les lignes que nous effectuons) :

$$\begin{aligned}\pi = \pi \times T &\Leftrightarrow \begin{cases} -0,15b - 0,2c - 0,45b + 2,1c = 0 & L_1 \leftarrow L_1 + 3L_3 \\ 0,3b - 0,5c + 0,3b - 1,4c = 0 & L_2 \leftarrow L_2 - 2L_3 \\ -0,1a - 0,15b + 0,7c = 0 & L_3 \leftarrow L_3 \end{cases} \\ &\Leftrightarrow \begin{cases} -0,6b + 1,9c = 0 \\ 0,6b - 1,9c = 0 \\ -0,1a - 0,15b + 0,7c = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} 0,6b = 1,9c \\ 0,6b = 1,9c \\ -0,1a - 0,15b + 0,7c = 0 \end{cases}\end{aligned}$$

En remplaçant b par son expression en fonction de c dans L_3 , nous obtenons :

$$\begin{aligned}\pi = \pi \times T &\Leftrightarrow \begin{cases} b = \frac{19}{6}c \\ -0,1a - 0,15 \times \frac{19}{6}c + 0,7c = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} b = \frac{19}{6}c \\ 0,1a = \frac{9}{40}c \end{cases} \\ &\Leftrightarrow \begin{cases} b = \frac{19}{6}c \\ a = \frac{9}{4}c \end{cases}\end{aligned}$$

Nous n'avons plus que 2 équations pour 3 inconnues. Rappelons-nous alors que $a + b + c = 1$. Nous avons donc :

$$\begin{aligned}\frac{9}{4}c + \frac{19}{6}c + c &= 1 \\ \text{Soit : } \frac{77}{12}c &= 1 \\ \text{Donc : } c &= \frac{12}{77}\end{aligned}$$

➔ Nous en déduisons les autres coefficients de la matrice ligne π :

$$\begin{aligned}a &= \frac{9}{4} \times \frac{12}{77} \\ &= \frac{27}{77} \\ \text{et : } b &= \frac{19}{6} \times \frac{12}{77} \\ &= \frac{38}{77}\end{aligned}$$

Ⓒ En conclusion, la chaîne de Markov associée à la matrice T admet bien une distribution invariante et celle-ci est :

$$\pi = \left(\frac{27}{77} \quad \frac{38}{77} \quad \frac{12}{77} \right)$$



Attention

Exerçons notre esprit critique en vérifiant que les coefficients de la matrice π sont bien des nombres de $[0 ; 1]$.

Nous avons vu comment chercher une distribution invariante, et la trouver lorsqu'elle existe. Pour prouver l'existence d'une telle distribution, nous pourrions utiliser le résultat suivant.

b. Existence d'une distribution invariante pour une chaîne de Markov

Reprenons la chaîne de Markov du paragraphe précédent.

Et supposons maintenant que la distribution initiale est :

$$\pi_0 = (0,3 \quad 0,7)$$

Déterminons, à l'aide de la calculatrice π_1 , π_5 , π_{10} , π_{20} et π_{30} en prenant des valeurs arrondies à 10^{-5} près.

→ Nous obtenons :

$$\begin{aligned}\pi_1 &= \pi_0 \times T \\ &= (0,3 \quad 0,7) \times \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix} \\ &= (0,48 \quad 0,52) \\ &\neq \pi_0 \\ &[\pi_0 \text{ n'est donc pas une distribution invariante}]\end{aligned}$$

$$\begin{aligned}\pi_5 &= \pi_0 \times T^5 \\ &= (0,3 \quad 0,7) \times \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix}^5 \\ &= (0,42989 \quad 0,57011)\end{aligned}$$

$$\begin{aligned}\pi_{10} &= \pi_0 \times T^{10} \\ &= (0,3 \quad 0,7) \times \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix}^{10} \\ &= (0,42856 \quad 0,57144)\end{aligned}$$

$$\begin{aligned}\pi_{20} &= \pi_0 \times T^{20} \\ &= (0,3 \quad 0,7) \times \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix}^{20} \\ &= (0,42857 \quad 0,57143)\end{aligned}$$

$$\begin{aligned}\pi_{30} &= \pi_0 \times T^{30} \\ &= (0,3 \quad 0,7) \times \begin{pmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{pmatrix}^{30} \\ &= (0,42857 \quad 0,57143)\end{aligned}$$

Nous constatons que les coefficients de la matrice ligne π_n semblent tendre respectivement vers 2 valeurs.

Nous constatons également que : $\frac{3}{7} \approx 0,42857$ et $\frac{4}{7} \approx 0,57143$.

Tout se passe comme si la suite de matrices (π_n) convergeait vers la distribution invariante.

→ Dans ce cas, nous dirons que le **processus se stabilise** autour d'un état probabiliste au bout d'un certain temps.

Nous admettons la propriété suivante.



Propriété

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à 2 (ou 3) états et T sa matrice de transition.

Si la matrice T n'a pas de coefficient égal à 0, alors la suite des matrices lignes $(\pi_n)_{n \in \mathbb{N}}$ converge vers une distribution invariante.

→ De plus, cette distribution est unique.

Cette propriété nous permet de trouver les états stables des chaînes de Markov.

Dans notre exemple, la distribution converge vers la distribution invariante :

$$\begin{pmatrix} \frac{3}{7} & \frac{4}{7} \end{pmatrix}$$

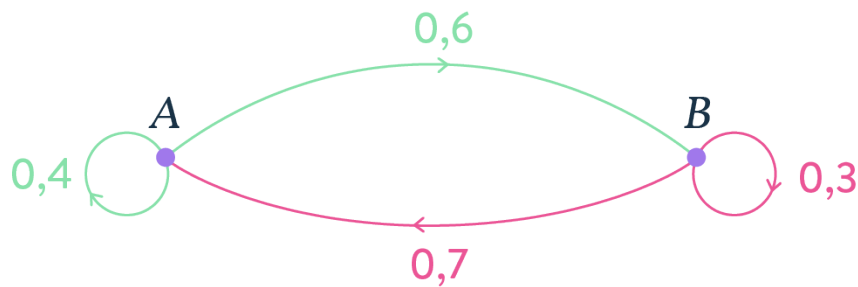
Revenons maintenant sur notre tout premier exemple.



Exemple

Nous avons étudié l'achat de deux marques A et B .

Reprenons le graphe probabiliste :



© SCHOOLMOUV

Graphe probabiliste à deux sommets

Nous supposons que c'est une chaîne de Markov qui est représentée : à savoir que les clients n'achètent un produit qu'en fonction de celui qu'ils ont acheté la semaine précédente, et que cet achat n'est pas conditionné par la semaine choisie.

→ La matrice de transition associée à cette chaîne de Markov est :

$$T = \begin{pmatrix} 0,4 & 0,6 \\ 0,7 & 0,3 \end{pmatrix}$$

D'après la propriété ci-dessus, comme les coefficients de T sont tous non nuls, les distributions de la chaîne de Markov vont tendre vers une distribution particulière qui est la distribution invariante.

→ Avec la même méthode développée ci-dessus, nous trouvons que la distribution invariante est :

$$\pi = \begin{pmatrix} \frac{7}{13} & \frac{6}{13} \end{pmatrix}$$

Concrètement, au bout d'un certain nombre de semaines, les habitudes des clients achetant ce produit se stabilisent :

- $\frac{7}{13}$ des clients achètent la marque A ;
- $\frac{6}{13}$ achètent la marque B .

Conclusion :

Dans ce cours, nous avons étudié ce qu'est une chaîne de Markov. Cet objet mathématique permet, en alliant les graphes, les matrices et les probabilités, d'étudier des processus aléatoires en fonction du temps, et de prédire éventuellement leur stabilisation.

C'est un outil qui est très utilisé dans plusieurs disciplines, que ce soit pour prédire des résultats d'élection, étudier un génome...