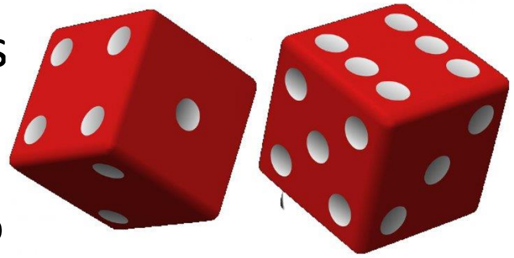
CHAPITRE II: CHAINES DE MARKOV

Centre Universitaire de AbdElhafid Boussouf, Mila
2ème Année Master Intelligence artificielle et ses applications
Année universitaire : 2021/2022
Matière: **Modélisation et simulation**
Responsable de la matière: DR. SADEK BENHAMMADA

Rappels: Probabilités – Variables aléatoires – Lois de probabilités

- **Univers:**
  - Un ensemble **Ω** de tous les résultats possibles d'une expérience aléatoire
  - Les éléments **ω** de l'ensemble **Ω** sont appelés des évènements élémentaires

  Exemple:
    - Le lancer d'un dé : $\Omega = \{1,2,3,4,5,6\}$
    - Le lancer de deux dés : $\Omega = \{(1,1), (1,2), \dots, (6,5), (6,6)\}$, ensemble des couples fo chiffres (avec ordre)



- **Evénement**
  - On appelle événement (au sens formel) tout sous-ensemble de Ω.
  - On dira qu'un événement *A* est réalisé lorsque l'évènement élémentaire ω effectivement réalisée est un élément de *A*, c'est-à-dire lorsque ω ∈ A.

  Exemple : lancer d'un dé à six faces
    - Ω={1,2,3,4,5,6}
    - Les évènements A, B, C:

A: Obtenir un six	A={6}
B: Obtenir un nombre pair	B={2,4,6}
C: Obtenir un nombre ≥ 4	C={4,5,6}

Terminologie des évènements aléatoires

Évènement = sous-ensemble de Ω, élément de $\mathcal{P}(\Omega)$.

Langage probabiliste	Notation	Langage ensembliste
issue ou résultat	$\omega \ (\in \Omega)$	élément de Ω
évènement A	$A \subseteq \Omega$	partie de Ω
A est réalisé	$\omega \in A$	appartenance
évènement contraire (non A)	$\overline{A} = \Omega \backslash A$	complémentaire
A et B	$A \cap B$	intersection
A ou B	$A \cup B$	union
évènements incompatibles	$A \cap B = \emptyset$	disjoints
A implique B	$A \subseteq B$	inclusion
évènement impossible	$\emptyset$	ensemble vide
évènement certain	Ω	partie maximale
système complet $A_1, \ldots, A_n$	$\Omega = \sqcup A_i$	partition

Espace probabilisé

- Un espace probabilisé ou espace de probabilité est un triplet $(\Omega, \mathcal{A}, \mathbb{P})$:
  - Ω: L'univers (L'ensemble de tous les résultats possibles)
  - $\mathcal{A}$: Un espace d'évènements, qui est un ensemble d' événements
  - Une application de probabilité $\mathbb{P}$:$\Omega \rightarrow [0,1]$, appelée probabilité sur Ω, tel que $\mathbb{P}(\Omega) = 1$

Variable aléatoire réelle

- Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé,
- Une variable aléatoire réelle X est une fonction définie sur un univers Ω et à valeur dans $\mathbb{R}$

$$X: \Omega \to \mathbb{R}$$

Exemple

- Soit l'expérience aléatoire : "On lance un dé à six faces et on regarde le résultat.

$$\Omega = \{1;\ 2;\ 3;\ 4;\ 5;\ 6\}$$

- On considère le jeu suivant :
  - Si le résultat est pair, on gagne 20 DA
  - Si le résultat est 1, on gagne 30 DA
  - Si le résultat est 3 ou 5, on perd 40 DA
- On a défini ainsi **une variable aléatoire X** sur Ω = {1; 2; 3; 4; 5; 6} qui peut prendre les valeurs 20, 30 ou -40:

X(1) = 30
X(2) = 20
X(3) = -40
X(4) = 20
X(5) = -40
X(6) = 20

$\boldsymbol{\omega \epsilon \Omega}$	**1**	**2**	**3**	**4**	**5**	**6**
$\boldsymbol{X(\omega)}$	30	20	-40	20	-40	20

Loi probabilisé

- Soit une variable aléatoire $\boldsymbol{X}$ définie sur un univers $\boldsymbol{\Omega}$ et prenant les valeurs $x_1, x_2, \dots, x_n$.
- La loi de probabilité de $\boldsymbol{X}$ associe à toute valeur x_i de $\boldsymbol{X}$ la probabilité $P(X = x_i)$.

Exemple

- Le lancer d'un dé : $\Omega = \{1,2,3,4,5,6\}$
- Chaque issue du lancer de dé est équiprobable et égale à $\frac{\mathbf{1}}{\mathbf{6}}$: $\forall \boldsymbol{\omega} \boldsymbol{\epsilon} \boldsymbol{\Omega}, \boldsymbol{P}(\boldsymbol{\omega}) = \frac{\mathbf{1}}{\mathbf{6}}$
- On considère la variable aléatoire X définie dans l'exemple précédent.

$\boldsymbol{P}(\boldsymbol{X} = \mathbf{20}) = \frac{\mathbf{1}}{\mathbf{6}} + \frac{\mathbf{1}}{\mathbf{6}} + \frac{\mathbf{1}}{\mathbf{6}} = \frac{1}{2}$

$\boldsymbol{P}(\boldsymbol{X} = \mathbf{30}) = \frac{\mathbf{1}}{\mathbf{6}}$

$\boldsymbol{P}(\boldsymbol{X} = -\mathbf{40}) = \frac{\mathbf{1}}{\mathbf{6}} + \frac{\mathbf{1}}{\mathbf{6}} = \frac{1}{3}$

- Tableau de la loi de probabilité de la variable aléatoire *X:*

x_i	**-40**	**20**	**30**
$P(X = x_i)$	$\frac{\mathbf{1}}{3}$	$\frac{\mathbf{1}}{2}$	$\frac{\mathbf{1}}{\mathbf{6}}$

Espérance, variance, écart-type

- Soit une variable aléatoire $\boldsymbol{X}$ définie sur un univers $\boldsymbol{\Omega}$ et prenant les valeurs $x_1, x_2, \ldots, x_n$.
- La loi de probabilité de $\boldsymbol{X}$ associe à toute valeur x_i de $\boldsymbol{X}$ la probabilité $P(X \ = x_i) = p_i$.
- L'espérance mathématique de la loi de probabilité de $\boldsymbol{X}$ est :

$$E(x) = \sum_{i=1}^{n} p_i x_i = p_1 x_1 + p_2 x_2 + \cdots + p_n x_n$$

- La variance de la loi de probabilité de X est :

$$\text{Var}(X) = \sum_{i=1}^{n} p_i \big(x_i - E(x)\big)^2 = p_1\big(x_1 - E(x)\big)^2 + p_2\big(x_2 - E(x)\big)^2 + \cdots + p_n\big(x_n - E(x)\big)^2$$

- \- L'écart-type de la loi de probabilité de $\boldsymbol{X}$ est :

$$\sigma(X) = \sqrt{\text{Var}(X)}$$

Exemple

x_i	**-40**	**20**	**30**
$P(X = x_i)$	$\frac{\mathbf{1}}{3}$	$\frac{\mathbf{1}}{2}$	$\frac{\mathbf{1}}{\mathbf{6}}$

$$E(x) = \sum_{i=1}^{n} p_i x_i = \frac{\mathbf{1}}{3} \times (-40) + \frac{\mathbf{1}}{2} \times (20) + \frac{\mathbf{1}}{6} \times (30) = \frac{\mathbf{10}}{6}$$

$$\mathrm{Var}(X) = \sum_{i=1}^{n} p_i \big(x_i - E(x)\big)^2 = \frac{\mathbf{1}}{3} \times \left(-40 - \frac{\mathbf{10}}{6}\right)^2 + \frac{\mathbf{1}}{2} \times \left(20 - \frac{\mathbf{10}}{6}\right)^2 + \frac{\mathbf{1}}{6} \times \left(30 - \frac{\mathbf{10}}{6}\right)^2 \approx 880{,}55$$

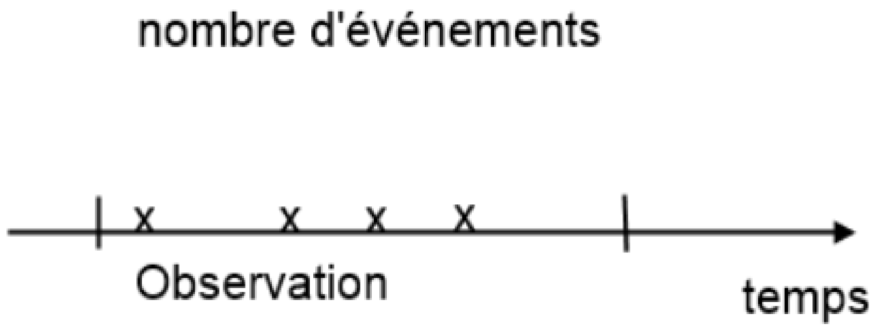
$$\sigma(X) = \sqrt{\mathrm{Var}(X)} = \sqrt{880{,}55} = 29{,}7$$

Rappels: Probabilités – Variables aléatoires – Lois de probabilités

- Les lois de probabilité permettent de décrire les variables aléatoires sous la forme d'une «expérience type»
- d'analyser cette expérience en détail pour pouvoir déduire les principales caractéristiques de toutes les expériences aléatoires qui sont du même type.
- On s'intéressera ici à quelques lois qui sont très fréquentes

Loi de Poisson

- La loi de Poisson est une loi de probabilité discrète qui permet la modélisation de l’observation d’un phénomène qui décrit le comportement du nombre d'événements se produisant dans un intervalle de temps fixé.
- Considérons *X* la variable aléatoire qui donne le nombre d’événements observés dans une unité de temps.
- On suppose
  - Un seul événement arrive à la fois
  - Le nombre d’événements se produisant ne dépend que du temps de l’observation
  - Les événements sont indépendants



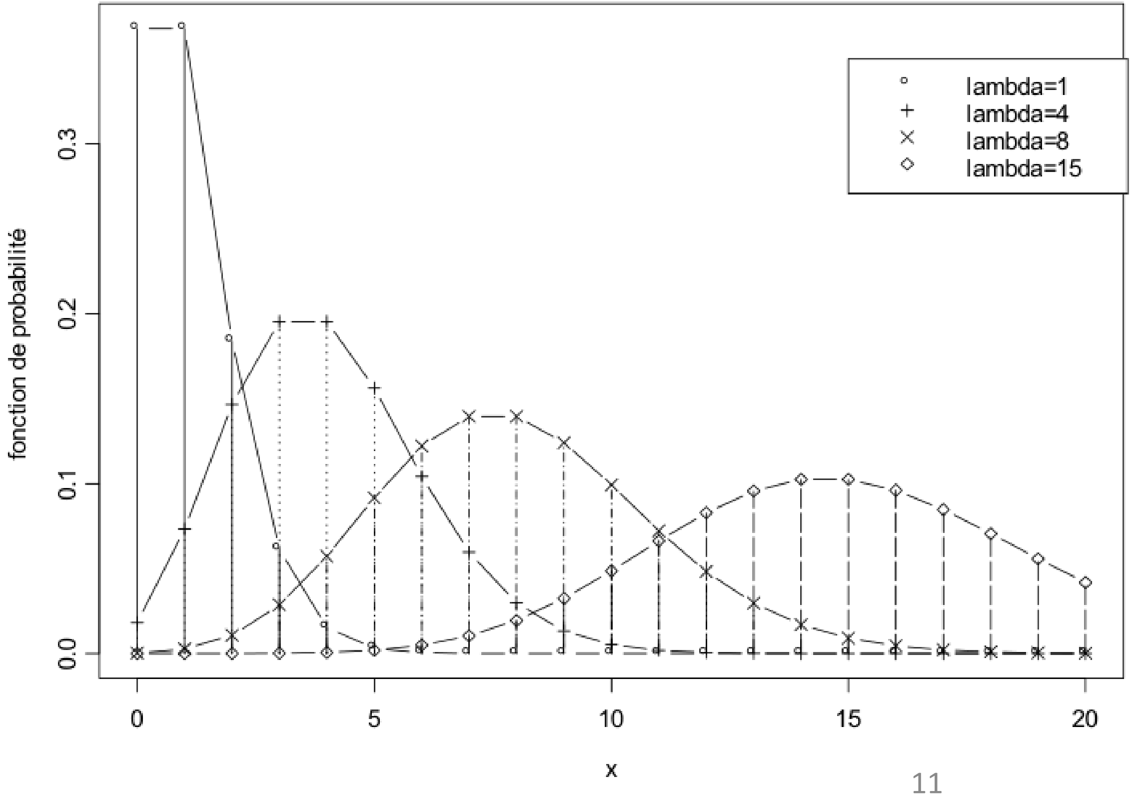
Loi de Poisson

- La variable aléatoire X qui donne le nombre d'événements par unité de temps suit une loi de Poisson, notée $X \sim P(\lambda)$, où $\boldsymbol{\lambda}$ est le nombre moyen d'événements par unité de temps.
- Les valeurs possibles de la variable aléatoire sont:
  - $S_X = \{0, 1, 2, \ldots\}$
  - La loi de probabilité est donnée par:

$$\boldsymbol{f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}} \quad \text{pour } x = 0,1, \ldots$$

- Espérance: $E(X) = \lambda$
- Variance: $Var(X) = \lambda$
- Ecart type: $E(X) = \sqrt{\lambda}$

Fonction de masse de loi de Poisson



Loi de Poisson

Exemple

- Dans une banque les clients arrivent à une fréquence moyenne **de 10 par heure**. Quelle est la probabilité qu'il y ait plus de **2 clients** en **10 min** ?
- Solution : Si on suppose que les clients arrivent indépendamment les uns des autres et que la moyenne est constante, la variable aléatoire ***X*** qui donne le nombre de clients en **10** min suit une loi $\boldsymbol{P(\lambda)}$, où $\boldsymbol{\lambda}$ est le nombre moyen de clients qui arrivent en **10 min**. $\boldsymbol{X \sim P(\lambda)}$

$$\boldsymbol{\lambda = 10 * \frac{10}{60} = \frac{10}{6}}$$

$$\boldsymbol{P(X > 2) = 1 - P(X \leq 2) = 1 - \big(P(X = 0) + P(X = 1) + P(X = 2)\big)}$$

$$\boldsymbol{= 1 - \big(f(0) + f(1) + f(2)\big)}$$

$$\boldsymbol{= 1 - \left(\frac{e^{-10/6}(10/6)^0}{0!} + \frac{e^{-10/6}(10/6)^1}{1!} + \frac{e^{-10/6}(10/6)^2}{2!}\right)}$$

Loi exponentielle

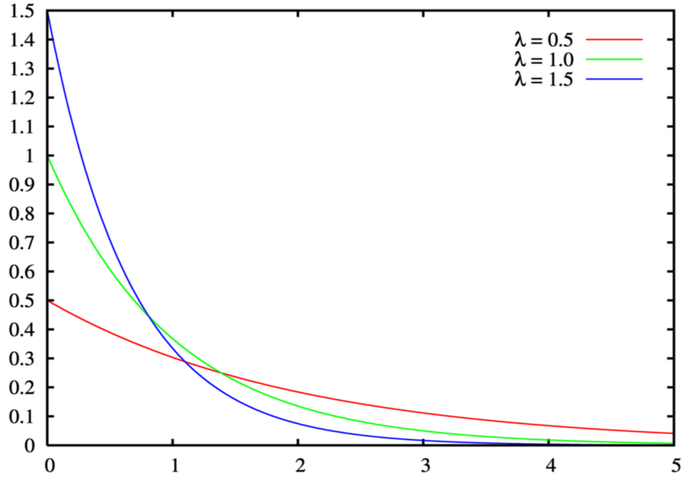
- La loi exponentielle donne le temps d'attente avant un événement lorsque le processus est régi par une loi de Poisson.
- Dans le cas de la loi de Poisson la variable aléatoire était le nombre d'événements tandis que dans la loi exponentielle c'est le temps d'attente avant le premier événement.
- Il est à noter que le nombre d'événements est une **variable aléatoire discrète** tandis que le temps d'attente est une **variable aléatoire continue**.
- La variable aléatoire **X** qui donne le temps d'attente avant la première apparition d'un phénomène de Poisson est **une loi exponentielle de paramètre** $\boldsymbol{\lambda = \frac{1}{E(X)}}$
- $\boldsymbol{E(X)}$ est l'espérance de **X** (temp moyen). Sa fonction de densité est donnée par:

$\boldsymbol{f(x) = \lambda e^{-\lambda x}}$ pour $0 \leq x \leq \infty$

$$\boldsymbol{P(a \leq X \leq b) = \int_a^b f(x)dx = e^{-\lambda a} - e^{-\lambda b}}$$

$$\boldsymbol{P(X \leq a) = P(0 \leq X \leq a) = e^{-\lambda \times 0} - e^{-\lambda a} = 1 - e^{-\lambda a}}$$
$$\boldsymbol{P(X \geq a) = 1 - P(X \leq x) = 1 - \left(1 - e^{-\lambda a}\right) = e^{-\lambda a}}$$

Fonction de densité de loi exponentielle



Loi exponentielle

Exemple 1: Dans une banque les clients arrivent à une fréquence moyenne **de 10 par heure**. Quelle est la probabilité qu'un client arrive dans 5 **minutes** ?

Solution:

Soit X la variable aléatoire qui donne la durée avant l'arrivée d'un client

X suit une loi exponentielle de paramètre $\boldsymbol{\lambda = \frac{10}{60} = 1/6}$

$$\boldsymbol{P(X \leq 5) = 1 - e^{-5\times\frac{1}{6}} \approx 0,56}$$

Exemple 2: Une montre digitale a une durée de vie moyenne de 100000 heures. Quelle est la probabilité qu'elle ne fonctionne plus après 5 ans ?

Solution :

Soit X la variable aléatoire qui donne la durée de vie en années d'une montre digitale.

- $\mathrm{E(X)=100000/365{,}25/24=}$ 11. 408 ans
- X suit une loi exponentielle de paramètre $\boldsymbol{\lambda = \frac{1}{11.\,408}}$
- $\boldsymbol{P(X \leq 5) = 1 - e^{-5\times\frac{1}{11.\,408}} \approx 0,355}$

Processus aléatoires : présentation

- La théorie des processus aléatoires concerne l’étude mathématique de phénomènes physiques, biologiques ou économiques évoluant dans le temps, et dont l’évolution est de caractère aléatoire, c’est-`a-dire non prévisible avec certitude.
- Pour définir un processus aléatoire, il faut :
    1. Un espace des temps T $(T \subset \mathbb{R}_+)$
    2. Un espace des états E.
    3. Une famille de variables aléatoires $(X_t)_{t \in T}$.

1. **Un espace des temps $\boldsymbol{T}$ ($\boldsymbol{T} \subset \mathbb{R}_+$)**

- Les deux espaces des temps les plus utilisés sont :
  - $\boldsymbol{T} = \mathbb{N}$: Le processus est dit discret ; on observe ce qu'il se passe à chaque unité de temps, ou bien on fait une suite d'opérations et on regarde ce qu'il se passe à chaque opération (ex : lancer une pièce).
  - $\boldsymbol{T} = \mathbb{R}_+$: le processus est dit continu : on observe un système qui évolue dans le temps à partir d'un instant t_0 que l'on prend pour origine des temps.

2. **Un espace des états $\boldsymbol{E}$**

- L'ensemble E peut être :
  - discret : c'est-à-dire fini ou dénombrable. Il sera, dans ce cas, souvent pratique d'identifier E avec une partie de $\mathbb{N}$ ou de $\mathbb{Z}$.
  - non discret : par exemple $\boldsymbol{E} = \mathbb{R}$ ou $\boldsymbol{E} \subset \mathbb{R}^{\boldsymbol{2}}$ (partie du plan) ou $\boldsymbol{E} \subset \mathbb{R}^{\boldsymbol{3}}$(partie de l'espace)

3. Une famille de variables aléatoires $(\boldsymbol{X_t})_{\boldsymbol{t \in T}}$

- Ces variables aléatoires sont toutes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$ et à valeurs dans l'espace des états E.
- Ainsi, à chaque instant $t \in T$, on associe, non pas une valeur déterministe mais une valeur aléatoire décrite par une variable aléatoire X_t à valeurs dans E.
- La variable aléatoire X_t peut représenter les résultats d'essais successifs comme par exemple, le jet d'une pièce à pile ou face, ou des observations successives sur une caractéristique d'une population.

3. Chaînes de Markov

- Les chaînes de Markov sont un outil mathématique permettant de modéliser l'évolution d'un système dont l'état au temps ***t+1*** ne dépend que de son état au temps ***t,*** et possédant un nombre fini d'états.
- À chaque étape, le système évolue en changeant d'état.
- Les probabilités de passer à chaque état au temps ***t+1*** à partir d'un état donné au temps ***t*** peuvent être regroupées sous forme d'une ***matrice carrée*** (matrice de transition)
- **La matrice de transition** dont les propriétés algébriques nous renseignent sur l'évolution du système.

3. Chaînes de Markov

3.1. Définition d'une chaîne de Markov en temps discret

- Soit $(X_t)_{t\in T}$ une famille de variables aléatoires à valeurs dans E un ensemble fini. Cette famille est une chaîne de Markov si elle vérifie la propriété de Markov : pour tout $n \in \mathbb{N}$, pour tout $x_0, x_1, \dots, x_{n+1} \in E$:

$$\boxed{P(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1} .., X_1 = x_1) = P(X_{n+1} = x_{n+1} | X_n = x_n)}$$

- Autrement dit la valeur de la variable X_{n+1} ne dépend que de la valeur de la variable X_n, et pas de tous ses états antécédents.

- **Chaîne de Markov homogène**
- Une chaîne de Markov est dite homogène si pour tout $n \in \mathbb{N}$ et tout i et j dans E, elle vérifie la propriété suivante :

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

Autrement dit une chaîne est homogène si la probabilité de transition d'un état à l'autre ne dépend pas de l'indice de la variable concernée mais uniquement de sa valeur : l'évolution du processus ne dépend pas de l'origine des temps.

3.2. Matrice de transition

- **Probabilité de transition**
  - On appelle probabilité de transition pour aller de l'état i à l'état j la probabilité d'un état au suivant ainsi :

$$\boxed{P(X_{t+1} = j | X_t = i) = P_{i,j}}$$

- **Matrice de transition**
  - On définit une matrice carrée à coefficients réels à partir des probabilités de transition:

$$\mathcal{P} = \left(P_{x,y}\right)_{i,j \in E}$$

$$\mathcal{P} = \begin{pmatrix} p_{x_0,x_0} & p_{x_0,x_1} & p_{x_0,x_2} & \cdots \\ p_{x_1,x_0} & p_{x_1,x_1} & p_{x_1,x_2} & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix}$$

Toute matrice de transition vérifie les propriétés suivantes :

(1) pour tout couple $\forall\, j, i \in E$, $0 \leq P_{i,j} \leq 1$;

(2) La matrice $\mathcal{P}$ est stochastique ligne, c'est-à-dire : $\forall\, i \in E \ \sum_{j \in E} P_{i,j} = 1$

Représentation de la matrice de transition en tant que graphe

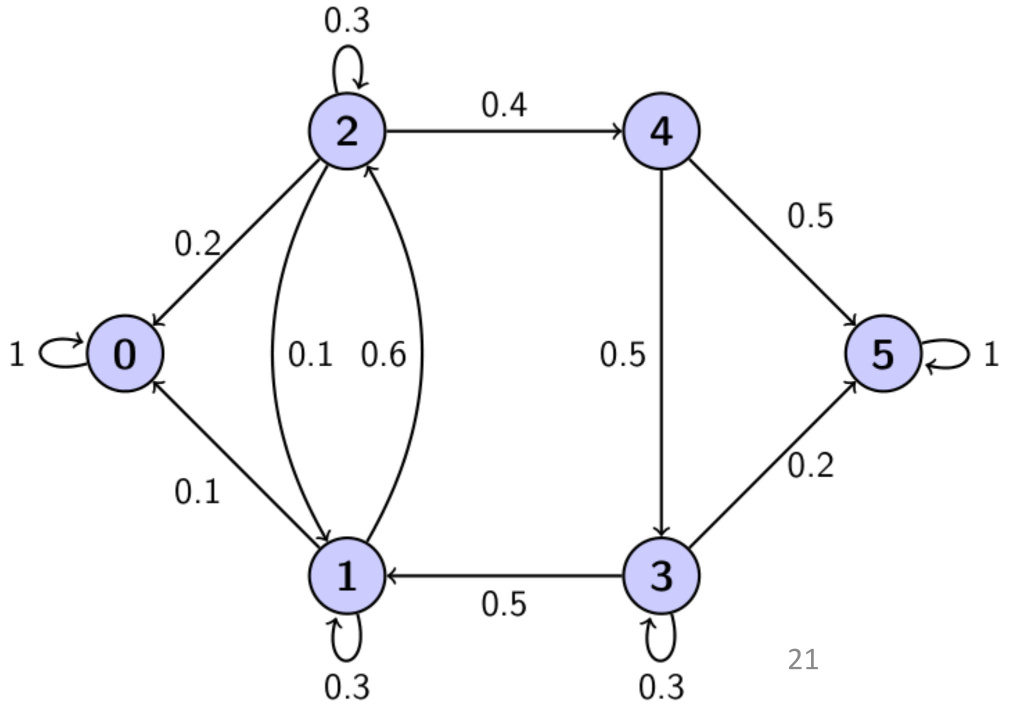
- Un chaine de Markov d'espace d'état E et de matrice de transition P peut être représenté par un graphe orienté étiqueté: $\mathrm{G} = (E, A)$, où les arêtes sont données par des transitions avec une probabilité non nulle:

$$\mathrm{A} = \{(i,j) | P_{i,j} > 0\}$$

- L'arête (i, j) est étiquetée par la probabilité $P_{i,j}$

Exemple:

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.1 & 0.3 & 0.6 & 0 & 0 & 0 \\ 0.2 & 0.1 & 0.3 & 0 & 0.4 & 0 \\ 0 & 0.5 & 0 & 0.3 & 0 & 0.2 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$



Probabilité de transition d'un état $\boldsymbol{i}$ à un état $\boldsymbol{j}$ en $\boldsymbol{n}$ étapes

Soit $(X_n)n \geq 0$ une chaîne de Markov de matrice de transition P

- On note $\boldsymbol{p_{i,j}^{n} = P(X_n = j | X_0 = i)}$ la probabilité de passer d'un état $\boldsymbol{i}$ à un état $\boldsymbol{j}$ en $\boldsymbol{n}$ étapes, pour $n \ \in \ N$.
- On note $P^{(n)}$ la matrice des coefficients $p_{i,j}^{n}$: $\boldsymbol{P^{(n)} = \left(p_{i,j}^{n}\right)_{i,j \in E}}$

Équations de Chapman-Kolmogorov

$$\forall \boldsymbol{n} \in \mathbb{N}, \boldsymbol{P^{(n)} = P^{n} = \left(p_{i,j}^{n}\right)_{i,j \in E}}$$

i.e, la probabilité de passer d'un état i à un état j en n étapes, 2ème $p_{i,j}^{n}$, est le coefficient situé à la $i^{\text{ème}}$ligne et la $j^{\text{ème}}$ colonne de la matrice P^{n}.

- En général: Les matrices de transitions d'ordre $n \ + \ m$ et n, m satisfont à la relation suivante:

$$\boldsymbol{P^{(n+m)} = P^{(n)} \ P^{(m)}}$$

Distribution stationnaire

- Soit (X_n) une chaine de Markov à espace d'états E de matrice de transition P**.**
- La distribution après $\boldsymbol{n}$ transitions notée $\boldsymbol{\pi_n} = (\boldsymbol{p_0},\ \boldsymbol{p_1}, \dots, ,\ \boldsymbol{p_m})$, est la loi de probabilité de la variable aléatoire $\boldsymbol{X_n}$
- On l'a note avec une matrice ligne $\boldsymbol{\pi_n} = (\boldsymbol{p_0}, \boldsymbol{p_1}, \dots, , \boldsymbol{p_m})$, le nombre de colonne correspond aux nombres d'états dans E.

Propriétés:

$$\forall\, n \in \mathbb{N}, \pi_{n+1} = \pi_n P$$
$$\forall\, n \in \mathbb{N}, \pi_n = \pi_0 P^n$$

Définition - Distribution stationnaire

Une distribution $\boldsymbol{\pi}$ est dite distribution stationnaire d'une chaîne de Markov de matrice de transition P, si c'est une probabilité qui vérifie:

$$\boldsymbol{\pi} = \boldsymbol{\pi} \times \boldsymbol{P}$$

Théorème (Perron-Frobenius)
P est la matrice de transition d'une chaine de Markov, on note π_0 la distribution initiale. S'il existe un entier $k \geq 1$, tel que la matrice P^k est une matrice strictement positive : tous ses éléments sont positifs non nuls (Donc P est une ***matrice primitive***), alors, la suite (π_n) des distributions converge vers une **unique distribution $\boldsymbol{\pi}$, invariante (stationnaire)** et indépendante de la distribution initiale π_0.

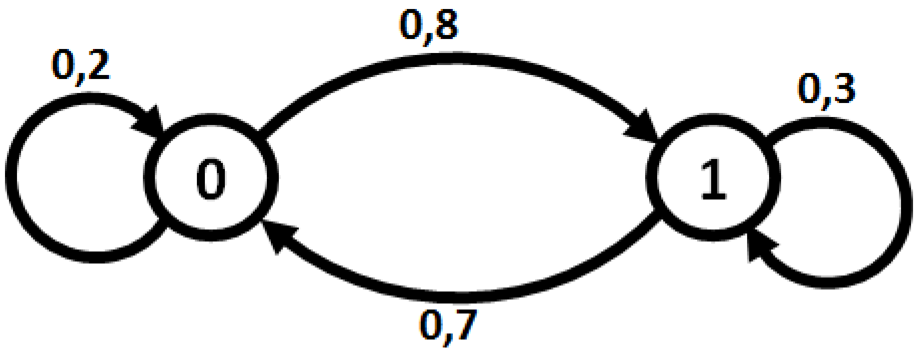
3. Chaînes de Markov

3.4. Transitions d’ordre *n* et loi à l’instant *n*

Exemple

Matrice de transitons

$$
\begin{array}{cc}
 & \begin{array}{cc} \mathbf{0} & \mathbf{1} \end{array} \\
P = & \begin{pmatrix} 0,2 & 0,8 \\ 0,7 & 0,3 \end{pmatrix} \begin{array}{c} \mathbf{0} \\ \mathbf{1} \end{array}
\end{array}
$$



- Initialement on est en 0, quelle est la probabilité d’ètre en 0 à la 3ème étape?

- $\pi_3 = \pi_0 P^3$
- $\pi_3 = (1 \quad 0)\begin{pmatrix} 0,2 & 0,8 \\ 0,3 & 0,7 \end{pmatrix}^3 = (1 \quad 0)\begin{pmatrix} 0,4 & 0,6 \\ 0,525 & 0,475 \end{pmatrix} = (0,4 \quad 0,6)$
  - La probabilité d’étre en 0 à la 3ème étape est =0,4
  - La probabilité d’étre en 1 à la 3ème étape est =0,6

Chaîne de Markov irréductible

- Soit x et y deux états. On dit que $\boldsymbol{x}$ **mène à** $\boldsymbol{y}$: $\boldsymbol{x} \rightarrow \boldsymbol{y}$, si:

$$\exists\ n \in \mathbb{N}, p_{x,y}^{n} = P(X_n = y | X_0 = x) > 0$$

- Cette relation signifie que partant de x nous avons une probabilité non nulle d'atteindre y à un certain temps n.
- On dit que $\boldsymbol{x}$ **et** $\boldsymbol{y}$ **communiquent** $(\boldsymbol{x} \leftrightarrow \boldsymbol{y})$ si x $(\boldsymbol{x} \rightarrow \boldsymbol{y}$)mène à y et y mène à x $(\boldsymbol{y} \rightarrow \boldsymbol{x})$.
- La relation de communication $\leftrightarrow$ est une **relation d'équivalence** (**Symétrique**, **réflexive**, et **transitive**)

> **Définition - Chaîne de Markov irréductible**
> Une chaîne de Markov est dite **irréductible** (**ergodique**) si:
> $$\forall\ x, y \in E, \boldsymbol{x} \leftrightarrow \boldsymbol{y}$$

- **Autrement dit:**

i.e, une chaîne de Markov est dite irréductible si:

$$\forall\ x, y \in E, \exists n \in \mathbb{N}, p_{x,y}^{n} = P(X_n = y | X_0 = x) > 0$$

- **Autrement dit:**

Une chaîne de Markov est dite irréductible si **son graphe est fortement connexe**, c'est-à-dire que pour tout couple d'états (x, y), il existe des chemins dirigés de x à y et de y à x

3.5. Classification des états

Classes d'équivalence

- Les états E de la chaîne peuvent être partitionnés en **classes d'équivalence** appelées classes irréductibles $\boldsymbol{C}$. $\forall\ \boldsymbol{x}, \boldsymbol{y}\ \in\ \boldsymbol{C}, \boldsymbol{x} \leftrightarrow \boldsymbol{y}$

Si E est réduit à une seule classe, la chaîne de Markov est irréductible.

Classe fermée

- Une classe d'équivalence $\boldsymbol{C}$ est dite fermée si pour tout $\boldsymbol{x}, \boldsymbol{y}$:

$$x \in\ C\ et\ x \rightarrow\ y\ \Rightarrow y\ \in\ C$$

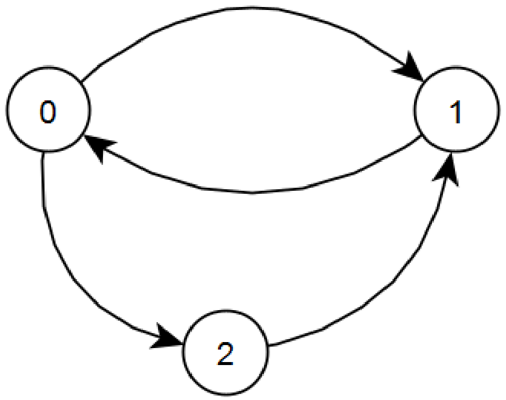
- autrement dit, une classe fermée C est une classe dont on ne peut pas en sortir.
- Sur le graphe d'un chaine de Markov, une classe est fermée si aucune transition n'en sort.

Etat absorbant

- Une classe fermée réduite à un état $C\ =\ \{x\}$ est appelée un état absorbant.
- Un état x est absorbant ssi $p_{x,x}=\ 1$.

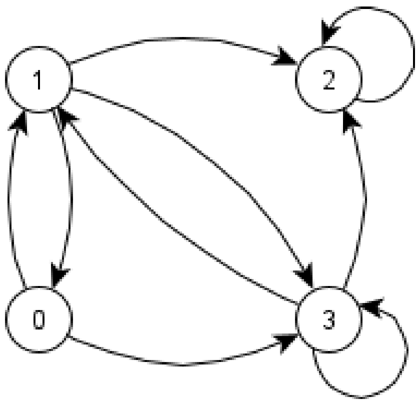
Exemples

Chaîne de Markov irréductible:

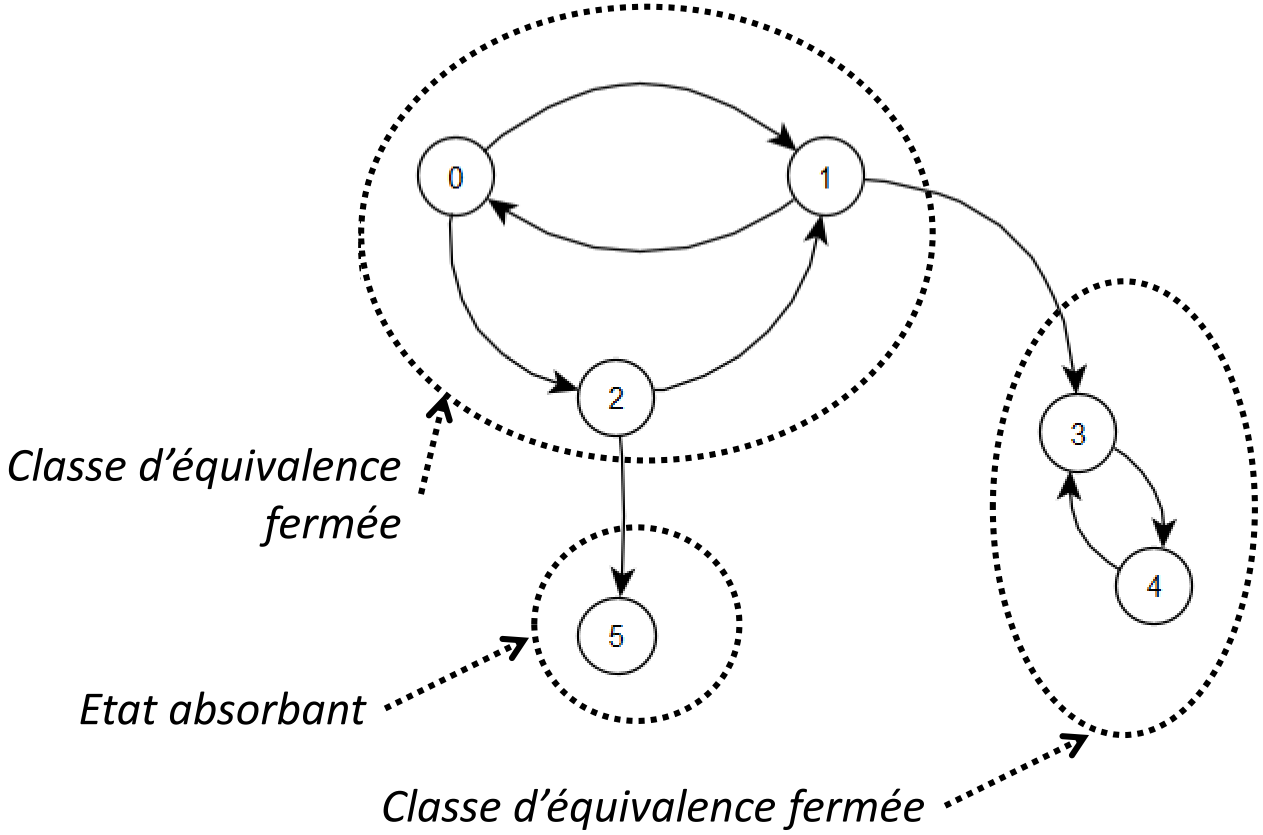


$$\begin{aligned} &\mathbf{0} \leftrightarrow \mathbf{1} \\ &\mathbf{1} \leftrightarrow \mathbf{0} \\ &\mathbf{0} \leftrightarrow \mathbf{2} \\ &\mathbf{2} \leftrightarrow \mathbf{0} \\ &\mathbf{1} \leftrightarrow \mathbf{2} \\ &\mathbf{2} \leftrightarrow \mathbf{1} \end{aligned}$$

Chaîne de Markov non irréductible, car aucun état ne peut être atteint à partir de 2.



Chaîne de Markov non irréductible



Période

Définition.

- Étant donné un état $x \in E$, ***la période*** de l'état x, notée $d(x)$, est le plus grand commun diviseur des entiers n tels que $p_{x,x}^{n}$ est strictement positif :

$$d(x) = PGCD\{n \geq 1 \mid p_{x,x}^{n} > 0\}$$

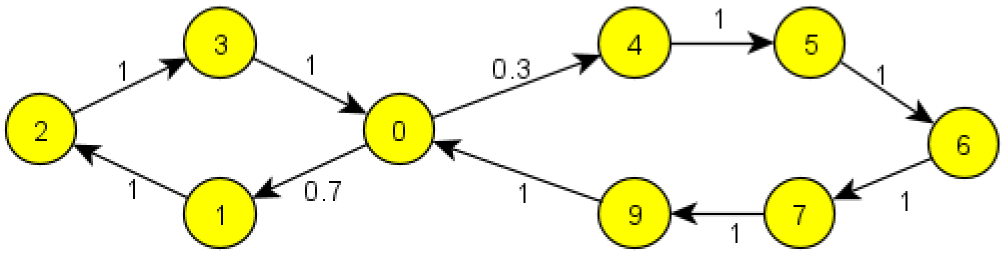
- Si $d(x) = 1$, alors on appelle l'état x **apériodique**.
- Une **chaîne de Markov est apériodique** si et seulement si tous ses états sont apériodiques.

Propriétés.

- Si deux états communiquent alors ils ont la même période.
- Tous les états d'une même classe irréductible ont la même période.
- La période d'une classe est la période de chacun de ses éléments.
- Une classe est dite apériodique si est seulement si sa période égale 1.

3.5. Classification des états

- **Exemple: Chaine périodique**



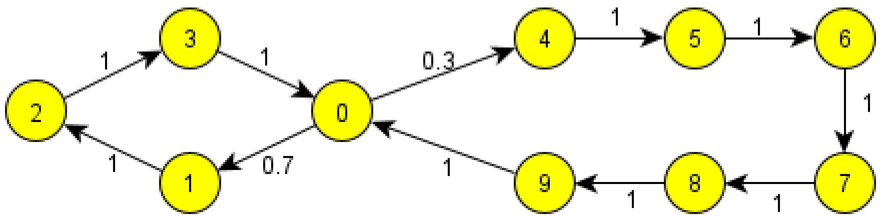
- Nous constatons qu'il y a deux cycles donc les états de chacun des deux cycles communiquent.
- De plus, ces deux cycles contiennent le même état 0, ce qui implique qu'ils communiquent, donc, **la chaîne est irréductible (une seule classe), par conséquent: tous les états ont la même période qui est la période de la chaîne de Markov**
- Partant de 0, nous sommes de nouveau en 0 au bout de 4 transitions en passant par le petit cycle et au bout de 6 transitions en passant par le grand cycle.

$p^4_{0,0}$ = 0.7 et $p^6_{0,0}$ = 0.3

- $\mathbf{d(0)} = \boldsymbol{PGCD\{n \geq 1 \mid p^n_{0,0} > 0\} = PGCD\{4 \times i + 6 \times j \mid (i,j) \neq (0,0)\} = PGCD\{4,6\} = 2}$.
- La chaîne est périodique de période 2.

3.5. Classification des états

- **Exemple: Chaine non périodique**



- **La chaîne est irréductible (une seule classe), par conséquent: tous les états ont la même période qui est la période de la chaîne de Markov**
- $\mathbf{d(0)} = \boldsymbol{PGCD}\{\boldsymbol{n} \geq \mathbf{1} \mid p_{0,0}^{n} > \mathbf{0}\} = \boldsymbol{PGCD}\{\mathbf{4} \times \boldsymbol{i} + \mathbf{7} \times \boldsymbol{j} \mid (\boldsymbol{i}, \boldsymbol{j}) \neq (\mathbf{0}, \mathbf{0})\} = \boldsymbol{PGCD}\{\mathbf{4}, \mathbf{7}\} = \mathbf{1}.$

- La chaîne est donc apériodique

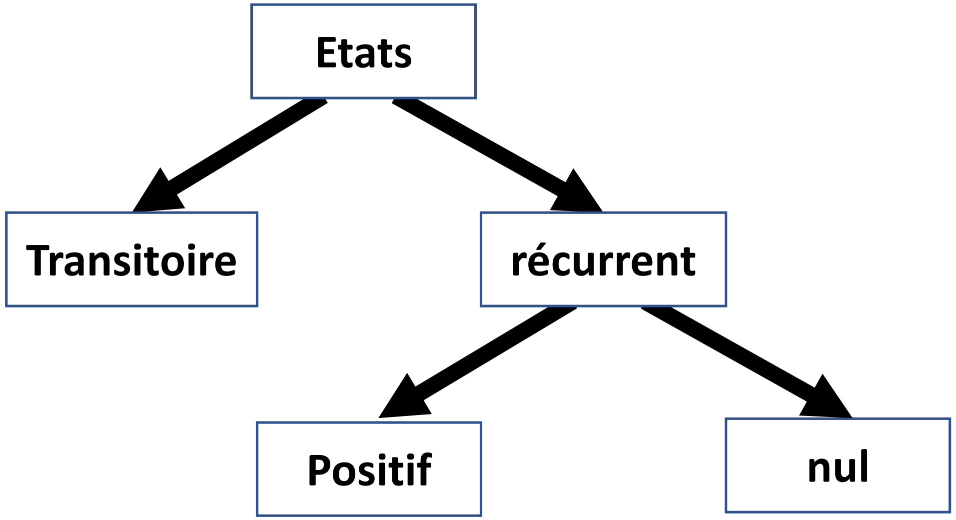
États récurrents / transitoires (transients)

- Un état est dit transitoire si, en partant de cet état au temps t, il existe une probabilité non nulle de ne jamais repasser par cet état. Un état non transitoire est dit récurrent.
- Soit $i \in E$
  - $f_{i,i}^{(n)}$: la probabilité que le premier retour en $x \in E$ ait lieu n étapes après l'avoir quitté,
  - $f_{x,x}$: La probabilité de repasser par x après l'avoir quitté:$f_{x,x} = \sum_{n\geq 1} f_{i,i}^{(n)}$
  - μ_i: Le temps moyen de retour en x: $\mu_i = \sum_{n\geq 1} n \times f_{i,i}^{(n)}$
- **Un** état **x est récurrent si et seulement si:**

$$f_{i,i} = \sum_{n\geq 1} f_{i,i}^{(n)} = 1$$

- μ_i représente le temps moyen de retour en i: il peut être infini, même si i est récurrent. On est donc conduit à une classification plus fine des états récurrents.
- Un état récurrent est dit **récurrent positif (non nul) si $\boldsymbol{\mu_i}$** $< +\infty$. Dans le cas contraire, il est dit **récurrent nul**.

3.5. Classification des états



Un état $i \in E$ récurrent nul ne peut exister que dans une chaine de Markov infinie ($\boldsymbol{E}$ est infini).

Propriétés

Soit $(X_n)_{n\geq 0}$ une chaine de Markov à espace d'états E

- Si x est récurrent $et\ x \to y$, alors y est récurrent.
- Les états d'une classe d'équivalence C sont tous récurrents (C est une classe récurrente), ou tous transients (C est une classe transiente).
- Une classe C qui n'est pas fermée, est transiente.
- Une classe C fermée et sur un espace fini est récurrente.
- Une chaîne Markov irréductible sur un espace d'états fini est récurrente.

Théorème

- Soit $(X_n)_{n\geq 0}$ une chaine de Markov à espace d'états fini E. Alors elle possède au moins une loi de probabilité stationnaire.
- Si (X_n) est **irréductible et récurrente**, alors la loi de probabilité stationnaire est **unique**.

  On note π la loi de probabilité stationnaire **unique**; elle vérifie pour tout $i \in E$, $\pi(i) = \frac{1}{\mu_i}$ où μ_i est le temps moyen de retour en i.
- Si (X_n) est **irréductible, récurrente** et **apériodique**, alors P^m converge vers la matrice dont toutes les lignes sont constantes égales à π.

$$P^m = \begin{pmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{pmatrix}$$

Propriété

- Si π est une distribution stationnaire d'une chaîne irréductible, alors $\pi_x > 0$ pour tout $x \in E$.

3. Chaînes de Markov à temps continu (CMTC)

- Soit $(X_t)_{t \in T}$ une famille de variables aléatoires à valeurs dans l'espace d'état E, et espace du temps $\boldsymbol{T} \subset \mathbb{R}_+$ Cette famille est une chaîne de Markov si elle vérifie la propriété de Markov : pour tout $x_0, x_1, \dots, x_{n+1} \in E$ et $x_0 < x_1 < \cdots < x_{n+1}$:

- $P\left(X_{t_{n+1}} = x_{n+1} \middle| X_{t_n} = x_n, X_{t_{n-1}} = x_{n-1} .., X_{t_0} = x_0\right) = P(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n)$

- Autrement dit la valeur de la variable X_{n+1} ne dépend que de la valeur de la variable X_n, et pas de tous ses états antécédents.

- **Chaîne de Markov à temps continu homogène**
- Une chaîne de Markov **à temps continu** est dite homogène si pour tout $n \in \mathbb{N}$ et tout i et j dans E, elle vérifie la propriété suivante :

$$\boxed{P(X_{t+h} = i | X_t = j) = P(X_h = i | X_0 = j)}$$

Autrement dit une chaîne à temps continu est homogène si la probabilité de transition d'un état à l'autre ne dépend des instants t_{n+h} et t_n, mais de la durée $(t_{n+1} - t_n)$.

Matrice de taux de transition (Générateur infinitésimal) d'une CMTC

- Dans une chaine de Markov à temps continu:
- Le temps d'attente pour la transition d'un état i à un état j est une variable aléatoire exponentielle de paramétré λ_{ij}.
- $\lambda_{i,j}$ est le nombre moyen de transition de l'état i à l'état j par unité du temps.
- Le temps passé dans un état i est une variable aléatoire exponentielle de taux λ_i.
- Plus simplement, sachant que la chaîne X est dans l'état i, elle y reste un temps exponentiel de paramètre q_i puis saute à un nouvel état, en choisissant l'état j avec probabilité $\pi_{i,j}$.

$$\lambda_{ij} = q_i \pi_{ij}$$

Une CMTC sur une espace d'état dénombrable E est défini par une matrice de taux de transitions (ou Générateur infinitésimal) $Q = \left(q_{ij}\right)_{i,j\in E}$:

$$q_{ij} = \begin{cases} \lambda_{ij} & Si\ i \neq j \\ -q_i = -\sum_{k\neq i} \lambda_{ik} & Si\ i = j \end{cases}$$

Remarque: Pour toute $i \in E, \sum_{j\in E} q_{ij} = 0$ (La somme des éléments d'une ligne de Q est nulle)

Matrice de probabilités de sauts d'une CMTC

- A un générateur $Q = \left(q_{ij}\right)_{i,j\in E}$, on associe la matrice de saut $\Pi = \left(\pi_{ij}\right)_{i,j\in E}$ donnée par:

$$\pi_{ij} = \begin{cases} q_{ij}/q_i & \text{si } j \neq i \text{ et } q_i \neq 0, \\ 0 & \text{si } j \neq i \text{ et } q_i = 0, \end{cases}$$

$$\pi_{ii} = \begin{cases} 0 & \text{si } q_i \neq 0, \\ 1 & \text{si } q_i = 0. \end{cases}$$

- **Remarque:** La matrice Π est stochastique : ces coefficients sont positifs et leur somme sur chaque ligne vaut 1.

4. Chaînes de Markov à temps continu (CMTC)

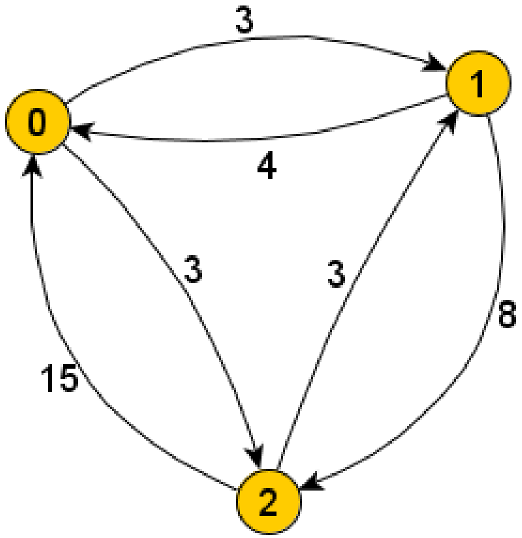
Stationnarité d'une chaine de Markov à temps continu

- Une chaine CMTC irréductible à espace d'états fini E est stationnaire.
- Dans ce cas, elle possède une distribution stationnaire unique $\pi = (\pi_i)_{i \in E}$ définit par:

$$\begin{cases} \pi Q = 0 \\ \sum_{i \in E} \pi_i = 1 \end{cases}$$

Exemple

Considérons une chaine de Markov à temps continu décrit par le graphe de taux de transitions suivant:

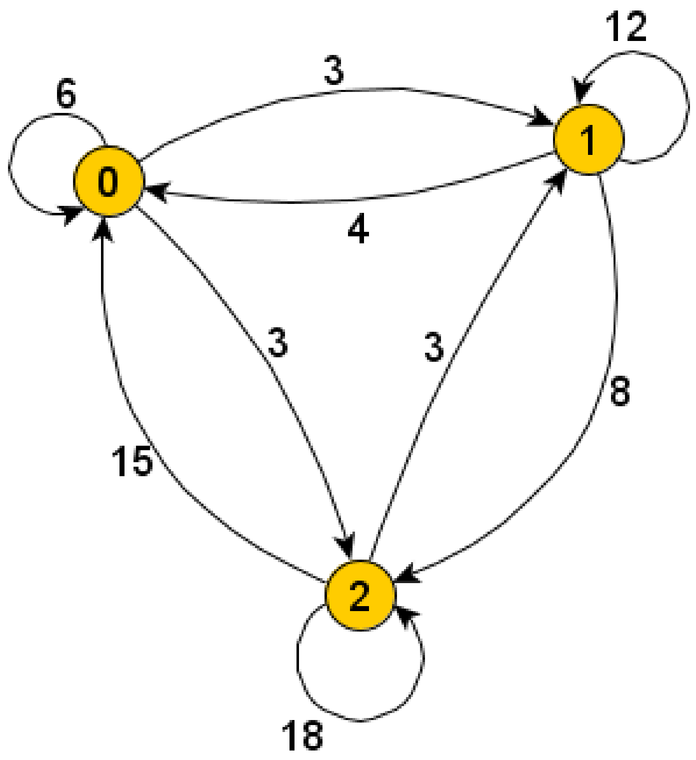


1. Donnez la matrice de taux de transitions (ou Générateur infinitésimal) et la matrice de saut de cette chaine de Markov
2. Démontrer que cette chaine est ergodique, et calculer sa distribution stationnaire.

Solution

Générateur infinitésimal $Q = \left(q_{ij}\right)_{i,j \in E}$

$$q_{ij} = \begin{cases} \lambda_{ij} & Si\ i \neq j \\ -q_i = -\sum_{k \neq i} \lambda_{ik} & Si\ i = j \end{cases}$$



$$q_{00} = -(3+3) = -6$$
$$q_{01} = 3$$
$$q_{02} = 3$$
$$q_{10} = 4$$
$$q_{11} = -(4+8) = -12$$
$$q_{12} = 8$$
$$q_{20} = 15$$
$$q_{21} = 3$$
$$q_{22} = -(3+15) = -18$$

$$Q = \begin{pmatrix} -6 & 3 & 3 \\ 4 & -12 & 8 \\ 15 & 3 & -18 \end{pmatrix}$$

4. Chaînes de Markov à temps continu (CMTC)

Solution

Matrice de sauts $\boldsymbol{\Pi}$:

$$\pi_{ij} = \begin{cases} q_{ij}/q_i & \text{si } j \neq i \text{ et } q_i \neq 0, \\ 0 & \text{si } j \neq i \text{ et } q_i = 0, \end{cases}$$

$$\pi_{ii} = \begin{cases} 0 & \text{si } q_i \neq 0, \\ 1 & \text{si } q_i = 0. \end{cases}$$

$$Q = \begin{pmatrix} -6 & 3 & 3 \\ 4 & -12 & 8 \\ 15 & 3 & -18 \end{pmatrix}$$

$\pi_{00} = 0$

$\pi_{01} = \frac{3}{6} = \frac{1}{2}$

$\pi_{02} = \frac{3}{6} = \frac{1}{2}$

$\pi_{10} = \frac{4}{12} = \frac{1}{4}$

$\pi_{11} = 0$

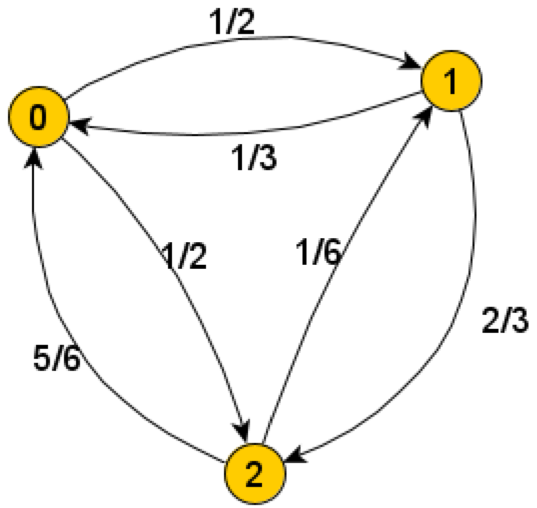
$\pi_{12} = \frac{8}{12} = \frac{2}{3}$

$\pi_{20} = \frac{15}{18} = \frac{5}{6}$

$\pi_{21} = \frac{3}{18} = \frac{1}{6}$

$\pi_{22} = 0$

$$\boldsymbol{\Pi} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 5/6 & 1/6 & 0 \end{pmatrix}$$



- Une CMTC est ergodique si:
  - son espace d'état est fini
  - Irréductible

$E = \{0,1,2\}$

$|E| = 3$, donc E est fini

$$\left.\begin{array}{l} 0 \longrightarrow 1 \; et \; 1 \longrightarrow 0 : 0 \leftrightarrow 1 \\ 0 \longrightarrow 2 \; et \; 2 \longrightarrow 0 : 0 \leftrightarrow 2 \\ 1 \longrightarrow 2 \; et \; 2 \longrightarrow 1 : 1 \leftrightarrow 2 \end{array}\right\} \text{La chaine est irréductible}$$

La chaine est finie et irréductible, donc, elle est ergodique,

Solution

Distribution stationnaire:

$$\begin{cases} \pi Q = 0 \\ \sum_{i \in E} \pi_i = 1 \end{cases}$$

$$Q = \begin{pmatrix} -6 & 3 & 3 \\ 4 & -12 & 8 \\ 15 & 3 & -18 \end{pmatrix}$$

$$\begin{cases} (\pi_0 \quad \pi_1 \quad \pi_2) \begin{pmatrix} -6 & 3 & 3 \\ 4 & -12 & 8 \\ 15 & 3 & -18 \end{pmatrix} = 0 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases}$$

$$\begin{cases} -6\pi_0 + 4\pi_1 + 15\pi_2 = 0 \\ 3\pi_0 - 12\pi_1 + 3\pi_2 = 0 \\ 3\pi_0 + 8\pi_1 - 18\pi_2 = 0 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases}$$

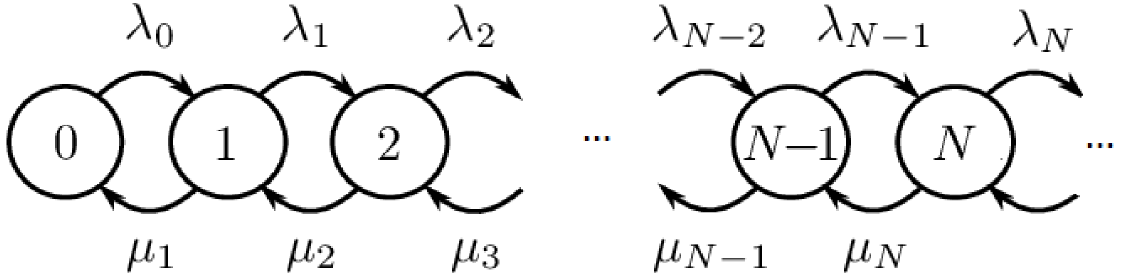
$$\begin{cases} \pi_0 = 64/105 \\ \pi_1 = \frac{1}{5} \\ \pi_2 = 4/21 \end{cases}$$

$$\boxed{\boldsymbol{\pi} = (\mathbf{64/105} \quad \mathbf{1/5} \quad \mathbf{4/21})}$$

4. Chaînes de Markov à temps continu (CMTC)

Processus naissance-mort

- Le processus de naissance-mort (ou processus de naissance et de mort) est une classe de chaines de Markov à temps continu et à espace d'états $E = \{0,1,2, \dots\}$ où les transitions d'état sont de deux types :
  - Les "naissances", où l'état passe de n à $n+1$,
  - Les "décès, où l'état passe de n à $n-1$
- Représentation du processus naissance-mort par un graphe orienté:



- Matrice de taux de transitions (Générateur infinitésimal) du processus naissance-mort:

$$\boldsymbol{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \dots & \\ \mu_1 & -\lambda_1-\mu_1 & \lambda_1 & 0 & \dots \\ 0 & \mu_2 & -\lambda_2-\mu_2 & \lambda_2 & \\ \vdots & \ddots & \ddots & \ddots & \end{pmatrix}.$$

Processus naissance-mort

- Un processus de naissance-mort est ergodique si et seulement si:

$$\sum_{i=1}^{\infty} \prod_{n=1}^{i} \frac{\mu_n}{\lambda_n} = \infty \text{ et } \sum_{i=1}^{\infty} \prod_{n=1}^{i} \frac{\lambda_{n-1}}{\mu_n} < \infty$$

- Si un processus de naissance-mort est ergodique, alors, il existe une distribution stationnaire $\pi = (\pi_i)_{i \in E}$ définit par:

$$\pi_k = \pi_0 \prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i}, k = 1,2, \ldots$$

$$\pi_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i}}$$

Processus naissance-mort

- Les processus naissance-mort ont de nombreuses applications:
  - En démographie ,
  - En théorie des files d'attente ,
  - En épidémiologie ,
  - En biologie,
  - etc.