Processus de Markov

Adapté de Yannis Korilis, Christian St-Jean, Dave DeBarr, Bob Carpenter, Jennifer Chu-Carroll et plusieurs autres

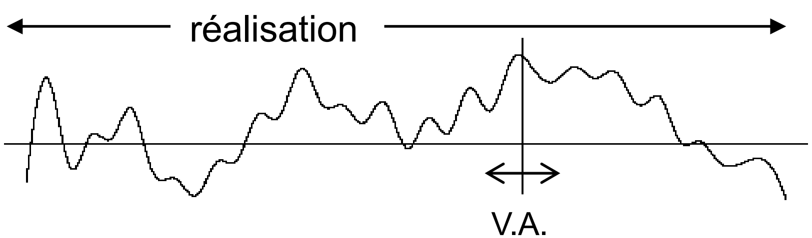
Processus stochastique

- Processus séquentiel dont chaque valeur est une variable aléatoire.
  - Caractérisé par ses différentes réalisations

Ex: Suite de lancers de dé (par ex., 1,3,2,5,3,6,2,4…

Signal radar,

Écho de voix



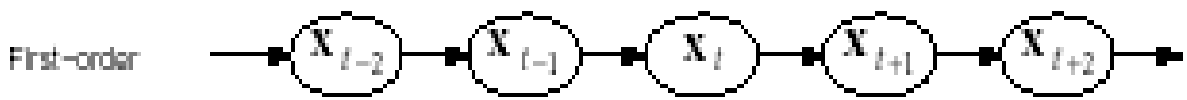
- Notation pour les processus chronologiques :
  - Temps continu : $x(t)$
  - Temps discret : $X(t_n)$; $X(nT_e)$; $X(n)$; X_n

Processus de Markov

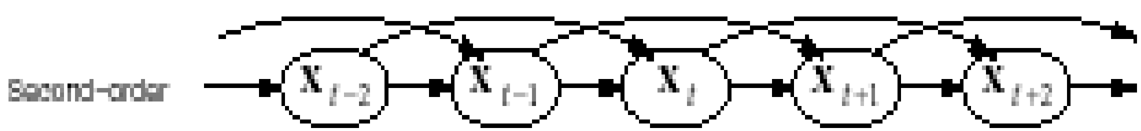
- L'évolution du processus dépend de son comportement passé (**processus à mémoire**)
  - 1^{er} ordre : Si on observe o_n à l 'instant t_n, alors

$$\boxed{P(X(t_n) \le o_n \,|\, X(t_{n-1}) \le o_{n-1}, \ldots, X(t_1) \le o_1) \;=\; P(X(t_n) \le o_n \,|\, X(t_{n-1}) \le o_{n-1})}$$

L'état courant du système suffit pour <u>prédire</u> son prochain état



  - 2^{nd} ordre : L'état courant du système et celui le précédant suffisent pour <u>prédire</u> le prochain état.



Andrei Markov (1856-1922) : statisticien russe spécialisé en modèles probabilistes temporels.

Chaîne de Markov

- Processus de Markov du 1er ordre à temps discret où les états X_n forment une suite d'indice {0,1,2,...}

Exemple: Alphabet : Σ = {'A','C','G','T'}

Séquence : $\{X_n\}$ = ACGCCTAGGCTAGCTTATCG

- Propriété de Markov :

$$\boxed{P\{X_{n+1}=j \mid X_n=i, X_{n-1}=i_{n-1},\ldots,X_0=i_0\} = P\{X_{n+1}=j \mid X_n=i\}}$$

  - Hypothèse de stationnarité : $P\{X_{n+1}=j \mid X_n=i\}=P_{ij}$
  - Probabilités de transition entre états : $P_{ij}\geq 0, \quad \sum_{j=0}^{\infty} P_{ij}=1$
  - Matrice des probabilités de transition $P=[P_{ij}]$
  - Probabilités initiales : $\pi_i \geq 0, \quad \sum_{i=0}^{m}\pi_i=1$

Représentation d'une chaîne de Markov

- Il faut définir:
  - L'espace des observations $\Sigma=\{o_1,\ldots,o_m\}$, où o_i est un élément d'un alphabet définissant la valeur possible de chaque état observé de $X(t)$
  - La matrice des probabilités de transitions $A=\{a_{i,j}\}$, où $a_{i,j} = P(o_j | o_i)$ est la probabilité d'observer o_j après avoir observé o_i à l'état précédent
  - Le vecteur des probabilités initiales $\Pi=\{\pi_i\}$, où $\pi_i=P(o_i)$ est la probabilité d'observer o_i à l'état initial

Exemple: X(t) : Temps qu'il fait sur 3 jours

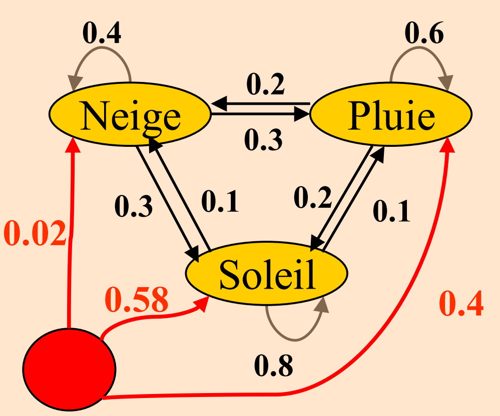
Σ={' neige ',' pluie ', 'soleil '}

$$A=\begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}$$

Π = {0.02 , 0.4, 0.58}

$$\sum_{j=1}^{m} a_{ij} = 1 \qquad \sum_{i=1}^{m} \pi_i = 1$$

Représentation Graphique



Probabilité d'une séquence d'états

Th. Bayes

$$P(o_n, o_{n-1}, \ldots, o_1) = P(o_n | o_{n-1}, \ldots, o_1) \times P(o_{n-1}, \ldots, o_1)$$

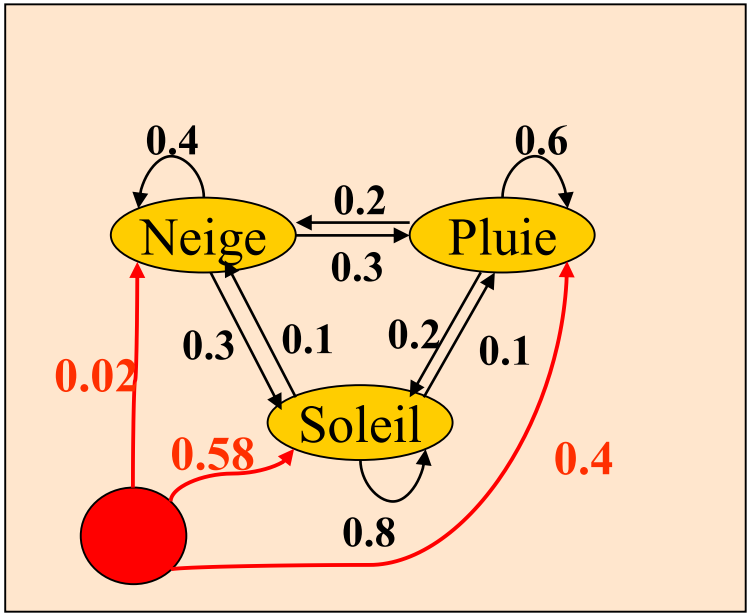
$$= P(o_n | o_{n-1}) \times P(o_{n-1} | o_{n-2}, \ldots, o_1) \times P(o_{n-2}, \ldots, o_1)$$

$$= P(o_n | o_{n-1}) \times P(o_{n-1} | o_{n-2}) \times \ldots \times P(o_2 | o_1) \times P(o_1)$$

$$= P(o_1) \times \prod_{t=2}^{n} P(o_t | o_{t-1})$$

Propriété de Markov

Représentation Graphique



Application:

Quelle est la probabilité qu'il fasse ' Soleil ', 5 jours de suite ?

P(Soleil, Soleil, Soleil, Soleil, Soleil)=

*0.58*0.8*0.8*0.8*0.8 =0.2375*

P(Neige, Neige, Pluie, Pluie, Soleil)=

*0.02*0.4*0.3*0.6*0.2=0,000288*

Et s'il y a plusieurs chemins ?

- Probabilité de faire une transition en *n* sauts

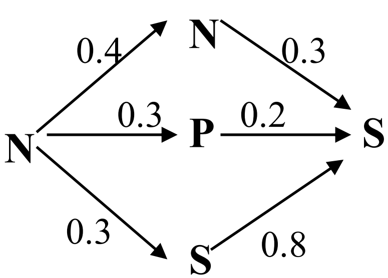
$$P_{ij}^{n} = P\{X_{n+m} = j \mid X_m = i\}, \quad n,m \geq 0,\ i,j \geq 0$$

Note : P_{ij}^{n} est l'élément (*i, j*) de la matrice P^n ; $P_{ij}^{n} = \left[\boldsymbol{P}^n\right]_{ij}$

- Les équations de Chapman-Kolmogorov permettent des « *pauses* »

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^{n} P_{kj}^{m}, \quad n,m \geq 0,\ i,j \geq 0$$

- *Quelle est la probabilité qu'il fasse ' Soleil' dans 2 jours sachant qu'il neige aujourd'hui ?*



$$P(X_3 = S \mid X_1 = N) =$$

$$P(S|N) \times P(N|N) + P(S|P) \times P(P|N) + P(S|S) \times P(S|N)$$

$$= 0.3 * 0.4 + 0.2 * 0.3 + 0.8 * 0.3 = 0.42$$

ou $P(X_3 = S \mid X_1 = N) = \left[\boldsymbol{P}^2\right]_{NS}$

système d’évènements complet

- On s’intéresse uniquement au résultat, en suivant les chemins possibles sans se soucier de l’état de départ :

$$P(X_t = j) = \sum_{i=1}^{m} \left[P(X_t = j \mid X_{t-1} = i) \times P(X_{t-1} = i) \right]$$

- *Quelle est la probabilité qu’il fasse ‘ Soleil ’ dans 3 jours ?*

$$P(X(3)=S) = P(S|N) \times P(X(2)=N) + P(S|P) \times P(X(2)=P) + \cdots$$

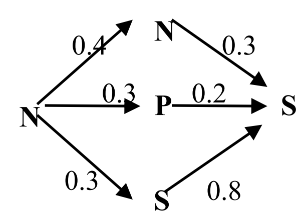
$$P(S|S) \times P(X(2)=S)$$

$\ldots$

$$P(X(3)=S) = 0.3*(0.4*0.02+0.2*0.4+0.1*0.58)$$

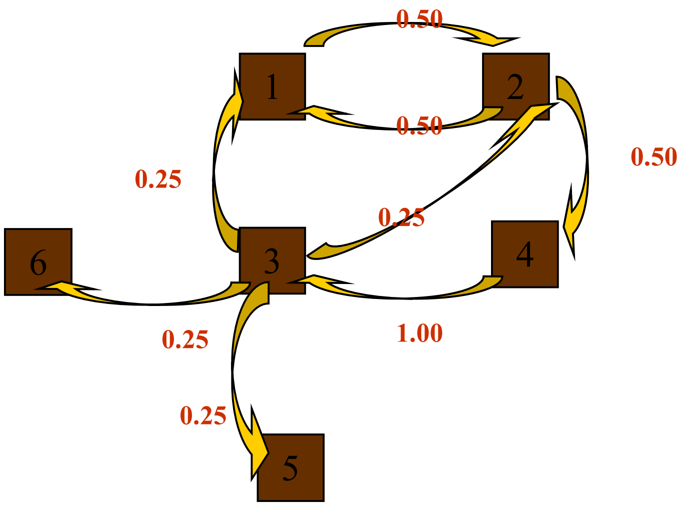
$$+0.2*(0.3*0.02+0.6*0.4+0.1*0.58)$$

$$+0.8*(0.3*0.02+0.2*0.4+0.8*0.58) = 0.5446$$



Exemple

- Un programme informatique comprend 5 modules indépendants, sp1, .., sp5, et un module de sortie sp6
  - De sp1 on peut :
    - aller à sp2 avec probabilité de ½
    - boucler avec probabilité de ½
  - De sp2 on peut :
    - aller à sp1 avec probabilité de ½
    - aller à sp4 avec probabilité de ½
  - De sp3 on peut :
    - aller à sp1 avec probabilité de ¼
    - aller à sp2 avec probabilité de ¼
    - aller à sp5 avec probabilité de ¼
    - à sp6 avec probabilité de ¼
  - De sp4 on va à sp3
  - De sp5, on boucle
  - De sp6, on boucle



- Partant de sp2, quelle probabilité d'y être à nouveau au temps « 4 » ?

<u>Première solution</u> : graphique

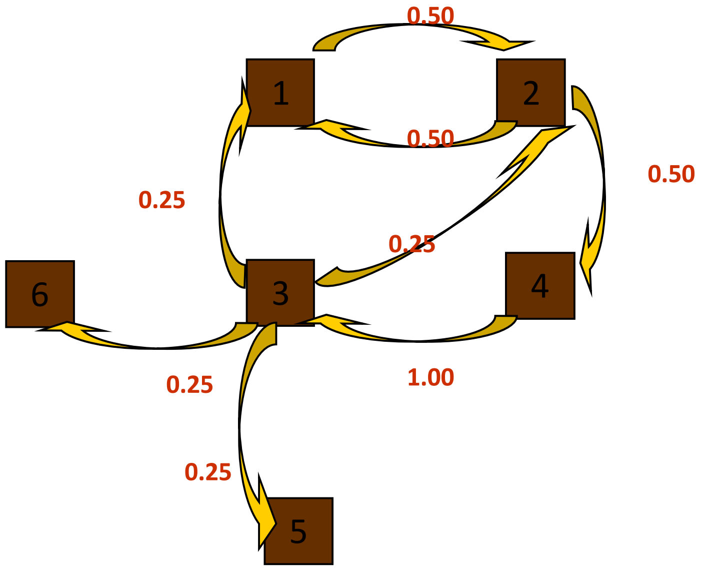
Il existe 3 chemins pour aller de 2 à 2 en quatre temps :

2→4→3→1→2 avec une proba : 0.50x1x0.25x0.50=1/16

2→1→2→1→2 avec une proba : 0.50x0.50x0.50x0.50=1/16

2→1→1→1→2 avc une proba : 0.50x0.50x0.50x0.50=1/16

Soit une proba de 3/16=0.187

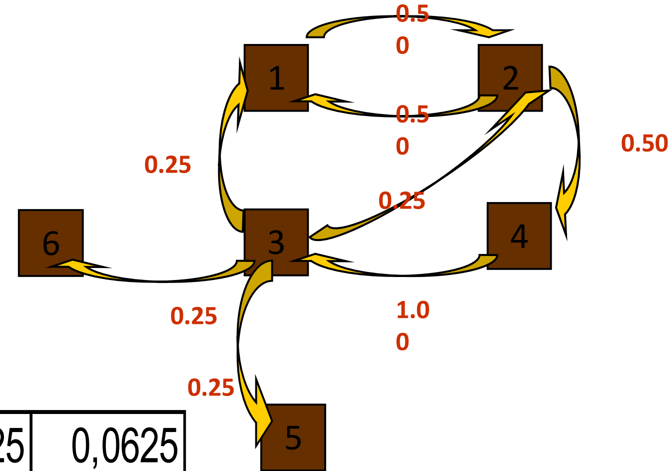


Deuxième résolution : par matrice

Si on pose les probabilités p_{ij} sous forme de matrice P, on a $$P(X_4 = 2 / X_0 = 2) = p_{22}^4$$

Matrice initiale P

0,50	0,5	0	0	0	0
0,5	0	0	0,5	0	0
0,25	0,25	0	0	0,25	0,25
0	0	1	0	0	0
0	0	0	0	1	0
0	0	0	0	0	1



P^4

0,375	0,25	0,125	0,125	0,0625	0,0625
0,3125	0,1875	0,125	0,125	0,125	0,125
0,1875	0,125	0,0625	0,0625	0,28125	0,28125
0,1875	0,125	0,125	0,0625	0,25	0,25
0	0	0	0	1	0
0	0	0	0	0	1

Distribution limite des états

- Probabilités des états en fonction du temps:

$$\pi_j^n = P\{X_n = j\}, \qquad \boldsymbol{\pi}^n = (\pi_0^n, \pi_1^n, \ldots)$$

$$P\{X_n = j\} = \sum_{i=0}^{\infty} P\{X_{n-1} = i\} P\{X_n = j \mid X_{n-1} = i\} \Rightarrow \pi_j^n = \sum_{i=0}^{\infty} \pi_i^{n-1} P_{ij}$$

Sous forme matricielle:

$$\pi^n = \pi^{n-1} P = \pi^{n-2} P^2 = \ldots = \pi^0 P^n$$

- Si π^n tend vers une valeur stationnaire π, on a :

$$\pi = \lim_{n \to \infty} \pi^n$$

et

$$\pi = \pi P$$ (En partant de $\pi^n = \pi^{n-1} P$)

➤ L'existence de π dépend de la structure de la chaîne

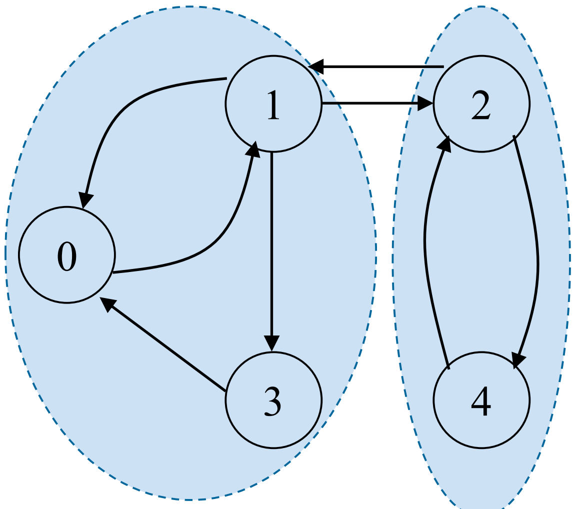
- La distribution stationnaire existe et est indépendante de l'état initial si la chaîne est irréductible et apériodique

Irréductible :

- Deux états i et j communiquent si :

$$\exists n,m: \ P_{ij}^{n} > 0, P_{ji}^{m} > 0$$

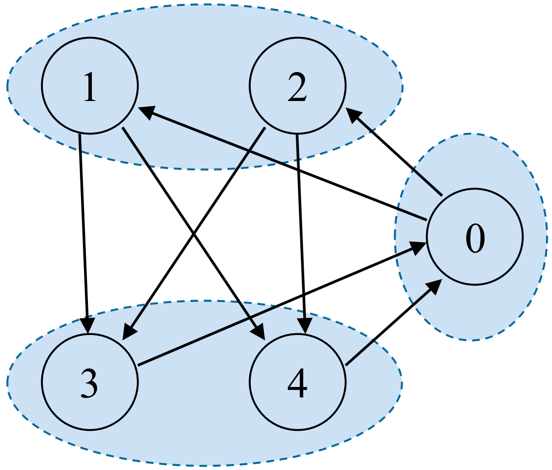
- La chaîne est irréductible si tous les états communiquent



Apériodique :

- Une chaîne est apériodique si aucun état n'est périodique
- Un état i est périodique si :

$$\exists\, d > 1: \ P_{ii}^{n} > 0 \Rightarrow n = \alpha d$$



Détermination de la distribution stationnaire

A. Nombre d'états fini

Deux possibilités :

- Résoudre le système d'équations :

$$\pi_j = \sum_{i=0}^{m} \pi_i P_{ij}, \quad j = 0, 1, \ldots, m$$

$$\sum_{i=0}^{m} \pi_i = 1$$

- Dériver π de $\lim_{n\to\infty} P^n$, sachant que

$$\pi_j = P\{X_n = j | X_0 = i\} = P_{ij}^n$$

et donc que P^n converge vers une matrice dont les lignes valent π

➡ Valable pour un petit nombre d'états

B. Nombre d'états infini

- Méthodes précédentes non applicables
- On doit deviner la solution, ou la trouver par itération:

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad j = 0, 1, \ldots,$$

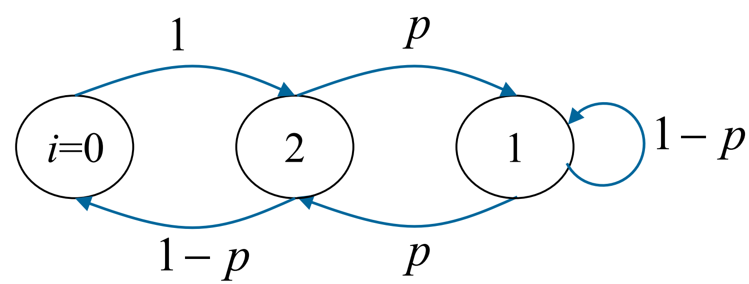
$$\sum_{i=0}^{\infty} \pi_i = 1$$

Exemple

- Une professeure possède deux parapluies. S’il pleut et l’un d’eux est disponible, elle le prend; sinon, elle sort sans parapluie.

Si p est la probabilité qu’il pleuve à chaque fois qu’elle se déplace, quelle est la probabilité qu’elle prenne une douche involontaire alors?

- Modèle markovien
  - Avec *i* le nombre de parapluies disponibles à l’endroit présent, on obtient le diagramme de transition

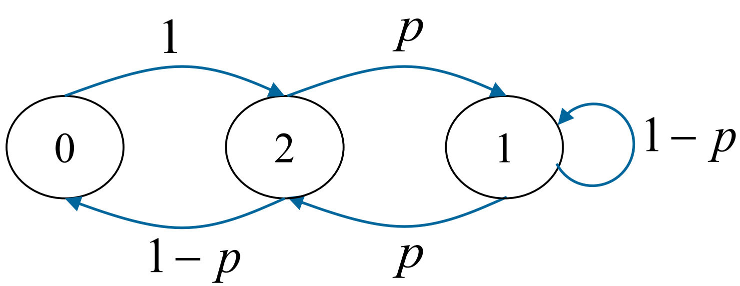


  - La mMatrice des probabilités de transitions entre états

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1-p & p \\ 1-p & p & 0 \end{bmatrix}$$

$$P\{\textit{Se mouiller}\} = \pi_0 p$$

Ex.: chaîne Markov avec un nombre fini d'états



$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1-p & p \\ 1-p & p & 0 \end{bmatrix}$$

$$\begin{cases} \pi = \pi P \\ \sum_i \pi_i = 1 \end{cases} \Leftrightarrow \begin{cases} \pi_0 = (1-p)\pi_2 \\ \pi_1 = (1-p)\pi_1 + p\pi_2 \\ \pi_2 = \pi_0 + p\pi_1 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases} \Leftrightarrow \pi_0 = \frac{1-p}{3-p}, \pi_1 = \frac{1}{3-p}, \pi_2 = \frac{1}{3-p}$$

$$P\{\textit{Se mouiller}\} = \pi_0 p = p\frac{1-p}{3-p}$$

Ex.: chaîne Markov avec un nombre fini d'états

- Si on prend $p = 0.1$:

$$\pi = \left(\frac{1-p}{3-p}, \frac{1}{3-p}, \frac{1}{3-p}\right) = (0.310, 0.345, 0.345)$$

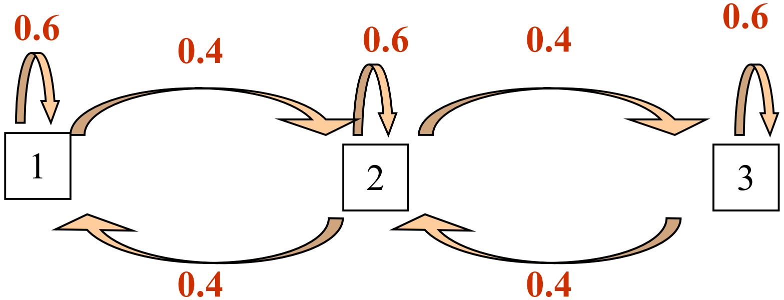
$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0.9 & 0.1 \\ 0.9 & 0.1 & 0 \end{bmatrix} \qquad P\{\textbf{\textit{Se mouiller}}\} = \pi_0 p = p\frac{1-p}{3-p}$$

- On peut aussi calculer π en évaluant $\lim_{n\to\infty} P^n$ numériquement

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 0.310 & 0.345 & 0.345 \\ 0.310 & 0.345 & 0.345 \\ 0.310 & 0.345 & 0.345 \end{bmatrix} \qquad (n \approx 150)$$

  - L'efficacité du calcul dépend de la structure de P

Exemple montrant l'indépendance de π des états initiaux



À l'équilibre, on a :

$\pi = [1/4 \quad 1/2 \quad 1/4]$

0,6	0,4	0
0,2	0,6	0,2
0	0,4	0,6

Ainsi, si on part de trois états différents :

$$\pi_a^0 = [1,0,0] \qquad \pi_b^0 = [0,1,0] \qquad \pi_c^0 = [0,0,1]$$

On obtient par simulation :

$$\pi^1 \quad [0.6 \quad 0.4 \quad 0.0] \qquad [0.2 \quad 0.6 \quad 0.2] \qquad [0.0 \quad 0.4 \quad 0.6]$$

$$\pi^{20} \quad \begin{matrix} [0.25 & 0.50 & 0.25] \\ [0.25 & 0.50 & 0.25] \\ [0.25 & 0.50 & 0.25] \end{matrix}$$

Autres processus markoviens

- Chaînes de Markov généralisées (ou Semi-Markov) :
  - Le prochain état dépend de l'état présent, *plus* la durée de temps passée dans ce dernier.
- Chaînes de Markov à temps continu :
  - Cas spécial des processus semi-markoviens.
- Champs de Markov
  - Modèle de Markov s'appliquant à une distribution spatiale de variables aléatoires
- Chaînes de Markov cachées :
  - Elles sont accessibles uniquement à travers des observations indirectes