Master I

Option : Probabilité et Statistiques

Département de Mathématiques, Faculté Des Sciences, UMMTO

$\boxed{\textbf{CHAÎNES DE MARKOV DISCRETES}}$

Y. Berkoun

1 Généralités sur les processus stochastiques

Soient $(\Omega,\mathcal{A},P)$ et $(E,\mathcal{F})$ respectivement un espace de probabilité et un espace probabilisable. Notons $\mathbb{T}$ un espace des temps pouvant être discret ($\mathbb{N}$ ou $\mathbb{Z}$) ou continu ($\mathbb{R}$).

Définition 1. *Un processus stochastique X est une application définie par*

$$\begin{aligned} X : \Omega \times \mathbb{T} &\longmapsto E \\ (w,t) &\longmapsto X(w,t) = X_t(w) \end{aligned}$$

$=$: Dans toute la suite, on notera $X = (X_t)_t$. E est dit espace des états pouvant être discret ou continu.
Si $\mathbb{T}$ et E sont discrets (respectivement continus), le processus X est dit processus à temps discret (continu) et à espace d'états discret (continu).
Ex: X peut représenter la valeur d'un indice boursier au fil du temps, l'évolution d'une population, etc.

Définition 2. *Pour ω fixé, l'application*

$$\begin{aligned} X : \mathbb{T} &\longmapsto E \\ t &\longmapsto X(w,t) = X_t(w) \end{aligned}$$

est dite trajectoire du processus associée l'aléa w.

Remarques 1
1- Pour t fixé et w variable, $X(w,t)$ est une variable aléatoire.
2- Le processus est dit du second ordre si $E(X_t^2) < \infty$, $\forall t$

Définition 3. *Bruit blanc*
Une suite $(\epsilon_t)_t$ de variables aléatoires identiquement distribuées est dite bruit blanc (au sens faible) si
a) $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma^2 < \infty \ \forall t$
b) $Cov(\epsilon_t,\epsilon_s) = 0 \ \forall t \neq s$

Remarque 2
On parlera de bruit blanc au sens fort si seulement la condition précédente a est vérifiée.

1.1 Stationnarité

Définition 4. *Stationnarité au sens faible*
Un processus $(X_t)_t$ est dit stationnaire au sens faible si
- $E(X_t) = constante < \infty$, $\forall t$
- $Cov(X_t, X_s)$ ne dépend que de la différence $|t - s|$

Définition 5. *Stationnarité au sens fort*
Un processus $(X_t)_t$ est dit stationnaire au sens fort ou strict si

$$\mathcal{L}(X_{t_1+h}, X_{t_2+h}, \dots X_{t_n+h}) = \mathcal{L}(X_{t_1}, X_{t_2}, \dots X_{t_n})$$

$$\forall t_1 < t_2 < \dots < t_n, t_i \in \mathbb{T}\ \forall i,\ et\ \forall h \in \mathbb{Z}$$

Où $\mathcal{L}$ désigne la loi de probabilité.
Ce qui signifie que $\mathcal{L}(X_{t+h}) = \mathcal{L}(X_t)$, $\forall t$, $\forall h$, i.e., l'invariance de la loi par translation.

Exemple
Soit $(X_t)_t$ un processus stationnaire au sens faible et notons $(Y_t)_t$ le processus défini par $Y_t = \sum_{i=1}^{\infty} a_i X_{t-i}$ où $(a_i)_i$ est une suite réelle vérifiant $\sum_{i\geq 1} |a_i| < \infty$. On a
- $E(Y_t) = m \sum_{i\geq 1} a_i < \infty$, $m = E(X_t)$
- $Cov(Y_t, Y_s) = \sum_{i\geq 1} \sum_{j\geq 1} a_i a_j Cov(X_t, X_s)$
Ce qui montre la stationnarité au sens faible de $(Y_t)_t$.

Remarques importantes 3
- Il est facile de montrer que la stationnarité au sens fort implique la sationnarité au sens faible (à condition que le moment d'ordre deux existe). Si $(\varepsilon_t)_t$ est une suite de variables aléatoires identiquement distribuées de même loi de Cauchy standard, alors $(\varepsilon_t)_t$ est strictement stationnaire, mais pas faiblement stationnaire du fait de la non-existence du moment d'ordre deux.
- L'inverse est généralement faux. Considérons une suite $(X_t)_t$ de variables aléatoires i.i.d de loi $N(0,1)$ et posons

$$Y_t = \begin{cases} X_t & \text{si t est pair} \\ \frac{1}{\sqrt{2}}(X_t^2 - 1) & \text{si t est impair} \end{cases}$$

$E(Y_t) = 0\ \forall t$, $Cov(Y_t, Y_s) = 0, \forall t, s$
Ce qui donne la stationnarité au sens faible de $(Y_t)_t$
$P(Y_t \leq 0) = P(X_t \leq 0) = \frac{1}{2}$ si t est pair
$P(Y_t \leq 0) = P(|X_t| \leq 1) \neq \frac{1}{2}$ si t est impair
Ce qui implique la non-stationnarité au sens strict de $(Y_t)_t$.

Définition 6. *Processus à accroissements indépendants*
Un processus $(X_t)_t$ *est dit à accroissements indépendants si*
$\forall t_1 < t_2 < \ldots < t_k, \forall k,$ *les variables aléatoires* $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_k} - X_{t_{k-1}}$ *sont indépendantes.*

Exemple Un bruit blanc est un processus à accroissements indépendants.

Définition 7. *Processus à accroissements stationnaires*
Un processus $(X_t)_t$ *est dit à accroissements stationnaires si*
$\forall t \in \mathbb{T}, \forall h > 0,$ *la* $\mathcal{L}(X_{t+h} - X_{t+h})$ *ne dépend pas de* t.

Définition 8. ***Fonctions d'autocovariance et d'autocorrélation***
Soit $(X_t)_t$ *un processus faiblement stationnaire. La fonction d'autocovariance* $\gamma_X(.)$ *est la fonction définie par*

$$\begin{aligned} \gamma_X : \mathbb{Z} &\longmapsto \mathbb{R} \\ h &\longmapsto \gamma_X(h) = Cov(X_t, X_{t+h}) \end{aligned}$$

la fonction d'autocorrélation notée $\varrho_X(.)$ *est définie par*

$$\begin{aligned} \varrho_X : \mathbb{Z} &\longmapsto [-1,1] \\ h &\longmapsto \varrho_X(h) = \frac{Cov(X_t, X_{t+h})}{\sqrt{V(X_t)}V(X_{t+h})} \end{aligned}$$

Proposition 1. *La fonction d'autocorrélation* ϱ_X *posséde les propriétès suivantes*
-$\varrho_X(0) = 1$
- $|\varrho_X(h)| \leq 1$
- $\varrho_X(h) = \varrho_X(-h)$

Définition 9. *Lois fini-dimensionnelles*
On appelle lois fini- dimensionnelles d'un processus $(X_t)_t$*, l'ensemble des lois*

$$\{\mathcal{L}(X_{t_1},X_{t_2},\ldots,X_{t_k}),\, t_1,\ldots,t_k \in \mathbb{T},\, k \in \mathbb{N}^*\}$$

Proposition 2. *(Kolmogorov) Loi d'un processus*
La loi $\mathbb{P}_X$ *d'un processus est entiérement déterminée par ses lois fini-dimensionnelles.*

Exercices

Exercice1
Dans la suite, $(\varepsilon_t)_t$ désignera un bruit blanc de variance σ^2.
a- Etudier la stationnarité du processus défini par $X_t = a + b\varepsilon_t + \varepsilon_{t-1}$, où $(a,b,c) \in \mathbb{R}^3$
b-A quelle condition le processus $X_t = at + b + \varepsilon_t$ est-il stationnaire ?

Exercice 2
Soit X une variable aléatoire non-symétrique telle que $E(X) = 0$, $V(X) = 1$. On pose $Y_t = (-1)^t X$, $t \in \mathcal{Z}$
Montrer que $(Y_t)_t$ est stationnaire au sens faible mais pas au sens fort.

Exercice 3
Soient X and Y deus processus faiblement stationnaires indépendants de fonction d'autocovariance ϱ_X et ϱ_Y. Montrer que le processus $X + Y$ est faiblement stationnaire de fonction d 'autocovariance $\varrho_{X+Y} = \varrho_X + \varrho_Y$.

Exercice 4
Soient A et B deux variables aléatoires de lois respectives $N(m_a,\sigma_a)$ et $N(m_b,\sigma_b)$ telles que $Cov(A,B) = \delta$. Notons $(Y_t)_t$ le processus défini par

$$Y_t = \begin{cases} A + \varepsilon_t & \text{si t est pair} \\ B + \varepsilon_t & \text{si t est impair} \end{cases}$$

où $(\varepsilon_t)_t$ est un bruit blanc de variance σ^2 indépendant de A et B.
a) $(Y_t)_t$ est-il sationnaire ? est-il du second ordre?
b) Etudier la stationnarité au sens fort ?

Exercice 5
Soit $(Y_t)_t$ un processus défini par $Y_t = \sum_{i=0}^{t}(-1)^i\varepsilon_i$, où $(\varepsilon_i)_i$ est un bruit blance de variance σ^2.
- $(Y_t)_t$ est-l du second ordre?
- Calculer $Cov(Y_t,Y_{t+h})$ pour $h \in \mathbb{Z}$. On distinguera le cas $h > 0$ et $h < 0$.
- $(Y_t)_t$ est-il stationnaire ?

Exercice 6
On considère le processus défini par

$X_t = \varepsilon_t - \theta\varepsilon_{t-1}$, $t \in \mathbb{Z}$ où $(\varepsilon)_t$ est un bruit blanc et $\theta \in]-1,+1[$.

a) Montrer que $(X_t)_t$ est stationnaire et calculer sa fonction d'auto-covariance.

2 Chaînes de Markov

Les chaînes de Markov ont été introduites par Andreï Andreïevitch Markov (1856-1922) qui était un éléve de Tchebychev. En lisant le roman *Eugène Onéguine* du russe Alexandre Pouchkine, Il examine les 20 000 premières lettres du texte et les considère comme 20 000 expériences aléatoires dont le résultat est consonne ou voyelle.Il montre alors, que l'occurrence d'une voyelle immédiatement après une consonne est un processus de Markov. Autrement dit, si on tire une lettre au hasard, c'est une voyelle ou une consonne selon des probabilités données par les statistiques (comme lorsqu'on joue à pile ou face), mais le statut d'une lettre dépend du statut de la précédente (événements non indépendants).

2.1 Chaînes de Markov discrètes

Considérons un processus stochastique $(X_t)_t$ à temps discret $\mathbb{T}$ et à espace d'états discret E.

Définition 10. *Chaîne de Markov*
$(X_t)_t$ est dite chaîne de Markov (cm) si

$$P(X_{n+1} = j/X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j/X_n = i) = P_{ij}(n)$$
$$i_0, i_1, \ldots, i_{n-1}, i, j \in E, n \in \mathbb{T}$$

Ce qui signifie que l'évoultion future du processus ne dépend que de son passé le plus récent, i.e., que le présent contient toute l'information passée du processus. En d'autres termes, le processus a tendance à oublier son passé; c'est ce qu'on appelle un processus sans mémoire. $P_{ij}(n)$ est la probabilité que le système démarrant de l'état initial i_0 atteigne l'état j au bout de $n+1$ transitions.

Exemple 1
Soit $(X_n)_n$ une suite positive de variables aléatoires indépendantes et équidistribuées de loi P donnée par

$$P(X_n = k) = q_k, \ \sum_{k \geq 0} q_k = 1$$

Soit $Y_n = \sum_{i=0}^{n} X_i$, alors $(Y_n)_n$ est une chaîne de Markov.
Remarquons dabord que $Y_{n+1} = Y_n + X_{n+1}$.

$$P(Y_{n+1} = j/Y_0 = i_0, \ldots, Y_{n-1} = i_{n-1}, Y_n = i) = P(X_{n+1} = j-i/X_0 = i_0, \, X_1 = i_1 - i_0 \ldots, \, X_n = i - i_{n-1}$$

Du fait que les variables aléatoires X_n sont indépendantes, le dernier terme est égal à $P(X_{n+1} = j - i) = P(Y_{n+1} = j/Y_n = i)$
$(Y_n)_n$ est donc une chaîne de Markov.

Pour l'instant, nous allons nous intéresser au cas des chaînes de Markov à valeurs dans un ensemble fini E, qui sont homogènes dans le temps. Leur loi est entièrement d'écrite par une matrice carrée (la matrice de transition de la chaîne), et leur étude se ramène alors essentiellement à des problèmes d'algèbre linéaire.

Définition 11. *La chaîne de Markov $(X_n)_n$ est dite Homogène (dans le temps) ou à transitions stationnaires si*
$P(X_{n+1} = j/X_n = i) = P(X_n = j/X_{n-1} = i) \ldots, = P(X_1 = j/X_0 = i) = P_{ij}$

Ce qui signifie que la probabilité de passer de l'état i l'état j ne dépend pas du temps. P_{ij} est la probabilité de passage de l'état i à j en une étape.

Remarque 1
• $P_{ij} \geq 0, \forall i,j$ et $\sum_{j\in E} P_{ij} = 1\ \forall i$
Du fait que X_1 est discrète, $\Omega = \bigcup_{j\in E}(X_1 = j)$.

$$P(\Omega/X_0 = i) = 1 = P(\bigcup_{j\in E}(X_1 = j/X_0 = i) = \sum_{j\in E} P(X_1 = j/X_0 = i) = \sum_{j\in E} P_{ij}$$

Définition 12. *Matrice de Transition*
Soit $(X_n)_n$ une chaîne de Markov homogène à espace d'états E fini.On appelle matrice de transition P associée à $(X_n)_n$ la matrice $P = (P_{ij})_{i,j\in E}$.

Si $E = \{1,2,\ldots,n\}$, la matrice P s'écrit

$$P = \begin{pmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{pmatrix}$$

Définition 13. *Matrice Stochastique*
Une matrice $A = (a_{ij})_{i,j}$ est dite stochastique si
• $a_{ij} \geq 0,\ \forall i,j$
•• $\sum_j a_{ij} = 1,\ \forall i$

Proposition 3. *1-Si A et B sont deux matrices stochastiques, alors le produit $A \times B$ est aussi une matrice stochastique.*
2- Toute matrice stochastique admet la valeur $\lambda = 1$ comme valeur propre. C'est la seule laquelle on peut associer une distribution de probabilité.
3- Toute valeur propre λ est de module inférieur ou égal à 1.

Démonstration 1. *1- Soient deux matrices stochastiques $A = (a_{ik})_{i,k}$, $1 \leq i \leq n$, $1 \leq k \leq p$ et $B = (b_{kj})_{k,j}$, $1 \leq k \leq p$, $1 \leq j \leq q$. Notons $C = A \times B = (c_{ij})_{i,j}$.*

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \ \forall i,j$$

$$\sum_{j=1}^{q} c_{ij} = \sum_{j=1}^{q} \sum_{k=1}^{p} a_{ik} b_{kj} = \sum_{k=1}^{p} a_{ik} \sum_{j=1}^{q} b_{kj}$$

Du fait que A et B sont stochastiques $\sum_{j=1}^{q} b_{kj} = \sum_{k=1}^{p} a_{ik} = 1$.
Ce qui donne $\sum_{j=1}^{q} c_{ij} = 1, \forall i$.
2- la preuve est simple et elle est laissée au lecteur.
3- Soit λ une valeur propre d'une matrice stochastique P et notons π le vecteur propre associé.
On a $\lambda \pi_i = \sum_j P_{ij} \pi_j \ \forall i$.Ce qui implique que $|\lambda||\pi_i| \leq \sum_j P_{ij} |\pi_j| \ \forall i$
Notons $|\pi_k| = \max_i \{|\pi_i|\}$ alors $|\lambda||\pi_k| = \sum_j P_{ij} |\pi_k|$. Ce qui donne $|\lambda| \leq \sum_j P_{ij} = 1$ d'ou le résultat.

Exemple Soit P la matrice donnée par

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

Les valeurs propres sont données par le calcul du déterminant

$$det(P - \lambda I) = \begin{vmatrix} \frac{1}{3} - \lambda & 0 & \frac{2}{3} \\ \frac{1}{4} & \frac{1}{4} - \lambda & \frac{1}{2} \\ 0 & 0 & 1 - \lambda \end{vmatrix} = \begin{vmatrix} 1 - \lambda & 0 & \frac{2}{3} \\ 1 - \lambda & \frac{1}{4} - \lambda & \frac{1}{2} \\ 1 - \lambda & 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda) \begin{vmatrix} 1 & 0 & \frac{2}{3} \\ 1 & \frac{1}{4} - \lambda & \frac{1}{2} \\ 1 & 0 & 1 - \lambda \end{vmatrix} =$$
$$(1 - \lambda) \begin{vmatrix} 1 & 0 & \frac{2}{3} \\ 0 & \frac{1}{4} - \lambda & -\frac{1}{6} \\ 0 & 0 & \frac{1}{3} - \lambda \end{vmatrix} = (1 - \lambda)(\tfrac{1}{4} - \lambda)(\tfrac{1}{3} - \lambda).$$

Les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = \frac{1}{4}$, $\lambda_3 = \frac{1}{3}$
le vecteur propre associé à la valeur propre λ_1 est $\pi_1 = (0,0,1)$. On voit bien que les assertions de la proposition précédente sont vérifiées.

Remarques 2
1) Une matrice de transition P est une matrice stochastique et P admet la valeur $\lambda = 1$ comme valeur propre.

2) Il existe un vecteur propre à gauche π associé à la valeur propre $\lambda = 1$ donné par $\pi P = \pi$ qui définit une distribution de probabilité.

si $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, Pour déterminer π il suffit de résoudre l'équation $\pi P = \pi$ ce qui est équivalent à

$$\pi_i = \sum_{j=1}^{n} P_{ij}\pi_j \ \forall i, \ \pi_i \geq 0, \ \sum_{j=1}^{n} \pi_j = 1$$

3) La valeur propre $\lambda = 1$ est la seule valeur propre à laquelle on peut associer une distribution de probabilité(vecteur propre associé).

Définition 14. *Graphe de transitions*
Soit $(X_n)_n$ une chaîne de Markov homogène à espace d'états E fini de matrice de transition $P = (P_{ij})_{i,j\in E}$. On associe à P un graphe $G = \{(i,j), P_{i,j}\}$ dit graphe de transitions.

Le graphe d'une chaîne de Markov à temps discret est défini comme un graphe orienté dans lequel les noeuds représentent les états du processus et les arcs correspondant aux transitions, i.e., les probabilités de transition strictement positives.

Remarque 3
Si $(X_n)_n$ est une chaîne de Markov et A est un événement engendré par les variables $(X_1, X_2, \ldots, X_n)$, alors

$$(P(X_{n+1} = j/\ A, X_n = i) = P(X_{n+1} = j/X_n = i)$$

Définition 15. *Probabilité de transition en n-étapes*
On appelle probabilité de transition en n-étapes notée $P_{ij}^{(n)} = P(X_n = j/X_0 = i)$, la probabilité de passer de l'état i à l'état j en n étapes, avec $P_{ij}^{(0)} = \delta_{ij} = \begin{cases} 1 & si\ i = j \\ 0 & sinon. \end{cases}$

La matrice de transitions en n- étapes est la matrice.
Il est intéressant de voir s'il y a un lien entre les matrices P et P_n. Le résultat suivant nous apporte une réponse.

Théorème 1. *Soit $P_n = (P_{ij}^{(n)})_{i,j}$ la matrice de transitions en n-étapes associée à une chaîne de Markov homogène finie de matrice de transition P, alors $P_n = P^n$.*

Démonstration 2. - *Si $n=1$ évident du fait que $P_1 = P^1 = P$*
- *Si $n=2$,*

$$P_{ij}^{(2)} = P(X_2 = j/X_0 = i) = P(X_2 = j, \bigcup_{k\in E}(X_1 = k)/X_0 = i) = \sum_{k\in E} P(X_2 = j,(X_1 = k)/X_0 = i)$$

$$= \sum_{k\in E} P(X_1 = k/X_0 = i)P(X_2 = j/X_0 = i,X_1 = k) = \sum_{k\in E} P(X_1 = k/X_0 = i)P(X_2 = j/X_1 = k)$$

$= \sum_{k\in E} P_{ik}P_{kj} = PP = P^2$

Supposons que la propriétè est vraie à l'ordre n, montrons qu'elle est vraie à l'ordre $n+1$.

$$\begin{aligned} P_{ij}^{(n+1} &= P(X_{n+1} = j/X_0 = i) = P(X_{n+1} = j, \bigcup_{k\in E}(X_n = k)/X_0 = i) \\ &= \sum_{k\in E} P(X_{n+1} = k/X_n = k,X_0 = i)P(X_n = k/X_0 = i) \\ &= \sum_{k\in E} P(X_{n+1} = k/X_n = k)P(X_n = k/X_0 = i) = \sum_{k\in E} P_{ik}^{(n)}P_{kj} \end{aligned}$$

Ce qui donne que $P_n = P^nP = P^{n+1}$, d'où le résultat.

Théorème 2. *Equation de Kolmogorov-Chapman*

$$\forall m,n \in \mathbb{N}^\star,\ P_{ij}^{(n+m)} = \sum_{k\in E} P_{ik}^{(m)}P_{kj}^{(n)},\ \forall i,j \in E$$

Ce qui signifie que $P_{n+m} = P^nP^m = P^{n+m}$

Démonstration 3.

$$\begin{aligned} P_{ij}^{(n+m)} &= P(X_{n+m} = j/X_0 = i) = \sum_{k\in E} P(X_{n+m} = j,X_n = k/X_0 = i) \\ &= \sum_{k\in E} P(X_{n+m} = j/X_n = k)P(X_n = k/X_0 = i) = \sum_{k\in E} P_{ik}^{(n)}P_{kj}^{(m)} \end{aligned}$$

Ce qui donne $P_{n+m} = P^nP^m = P^{n+m}$

Dans tout ce qui précède, on a utiliser que des probabilités conditionnelles. Si nous voulons calculer la probabilité d'un état i à l'instant n, il est nécessaire de connaître la loi initiale processus.

Loi d'une chaîne de Markov

Notons $\Pi_k(n) = P(X_n = k)$, $n \geq 0$, $k \geq 1$.
Nous pouvons écrire la distributionde X_n sous forme d'un vecteur ligne

$$\Pi(n) = (\Pi_1(n), \Pi_2(n), \ldots,) \text{ avec} \sum_{j\geq 1} \Pi_j(n) = 1$$

Pour déterminer $\Pi(n)$, il est nécessaire de connaître la loi initiale du processus, soit sa distribution
$\Pi(0) = (\Pi_1(0), \Pi_0(n), \ldots,)$ avec $\Pi_k(0) = P(X_0 = k)$

$$\begin{aligned}\Pi_k(n) = P(X_n = k) = P(X_n = k, \bigcup_{j\in E}(X_0 = j)) &= \sum_{j\in E} P(X_n = k, X_0 = j)\\ = P(X_0 = j)P(X_n = k/X_0 = j) &= \sum_{j\in E} \Pi_j(0)P_{jk}^{(n)}\end{aligned}$$

Ce qui donne $P(X_n = k) = \sum_{j\in E} \Pi_j(0)P_{jk}^{(n)}$ et donc $\Pi(n) = \Pi(0)P^n$, $\forall n \geq 1$.

Ce qui signifie que la loi de X_n est entièrement déterminée par sa loi initiale et sa matrice de transition.

Remarque 4

$\Pi(n+1) = \Pi(0)P^{n+1} = \Pi(0)P^nP = \Pi(n)P$, $\forall n \geq 1$
Supposons que $\lim_{n\to+\infty} \Pi(n) = \Pi$, alors $\lim_{n\to+\infty} \Pi(n+1) = \lim_{n\to+\infty} \Pi(n)P \to \Pi = \Pi P$
Si une distribution limite Π existe, elle doit vérifier $\Pi = \Pi P$.

Exemple 2
Soit la matrice P défine par $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et posons $\pi(0) = (\alpha, 1-\alpha)$, $0 \leq \alpha \leq 1$, une distribution initiale.

$P^{2n} = I_2$, $P^{2n+1} = P$ où I_2 est la matrice identité d'ordre 2.

$$\lim_{n\to+\infty} \Pi(n) = \Pi(0) \lim_{n\to+\infty} P^n = \begin{cases} \pi(0) & \text{si } n \text{ est pair} \\ (1-\alpha, \alpha) & \text{sinon} \end{cases}$$

Si $\alpha = \frac{1}{2}$, $\pi(0) = (\frac{1}{2}, \frac{1}{2})$, $\pi(0) = \lim_{n\to+\infty} \Pi(n)$
Si $\alpha \neq \frac{1}{2}$, $\lim_{n\to+\infty} \Pi(n) \neq \pi(0)$.
La chaîne ne converge pas pour $\alpha = \frac{1}{2}$.

Définition 16. *Distribution stationnaire*
On dit qu'une distribution π est stationnaire par rapport à une matrice de transition P si $\pi P = P$. Pour déterminer $\pi = (\pi_1,\pi_2,\dots)$, on résoud

$$\star \begin{cases} \pi P = \pi \\ \sum_{k\in E} \pi_k = 1 \end{cases}$$

$\star$ est dite équation de balance.

Exemple 3

$$\text{Soit } P = \begin{pmatrix} 1/2 & 01/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix} \text{ et notons } \pi = (\pi_1,\pi_2,\pi_3) \text{ une distribution stationnaire.}$$

$$\pi P = \pi \longrightarrow \begin{cases} \pi_1 + \pi_2 = 2\pi_1 \\ \frac{\pi_1}{2} + \pi_3 = \pi_2 \\ \frac{\pi_2}{2} = \pi_3 \end{cases}$$

$$\longrightarrow \pi = (a,a,\tfrac{a}{2}) \text{ avec } a + a + \tfrac{a}{2} = 1 \longrightarrow a = \tfrac{2}{5} \longrightarrow \pi = (\tfrac{2}{5},\tfrac{2}{5},\tfrac{1}{5})$$

Probabilité d'un chemin
Supposons que l'on veuille calculer la probabilité du chemin $(X_0 = i_0,X_1 = i_1,\dots,X_n = i_n)$, à savoir

$$\begin{aligned} P(X_0 = i_0,X_1 = i_1,\dots,X_n = i_n) &= P(X_0 = i_0)P(X_1 = i_1/X_0 = i_0)\dots,P(X_n = i_n/X_{n-1} = i_{n-1}) \\ &= P(X_0 = i_0)P_{i_0i_1}P_{i_1i_2}\dots P_{i_{n-1}i_n} \end{aligned}$$

Le calcul de la dernière quantité nécessite la connaissance de la loi initiale de X_0.

2.1.1 Comportement asymptotique d'une chaîne de Markov

- **Régime permanent et régime stationnaire**

Dans la remarque 4, on a étudié la probabilité $P(X_n = k)$ en fonction du nombre de transitions n, ce qui revient à étudier le régime transitoire. Il est évident que la loi de X_n varie en fonction du temps et de Π_0. Il est intéressant de voir si Π_n converge vers une distribution limite qui ne dépend pas de Π_0. Si c'est le cas, on dit qu'on a un régime permanent.

Définition 17. *On dit qu'une chaîne de Markov converge vers Π (ou admet une distribution limite) si*
$\lim\limits_{n\to+\infty} \Pi(n) = \Pi$ *indépendemment de $\Pi(0)$. La chaîne est dite stationnaire ou ergodique.*

Remarques 5
a) Π est une distribution de probabilité.
b) $\Pi(n+1) = \Pi(n)P$ implique que $\lim_{n\to+\infty} \Pi(n+1) = \lim_{n\to+\infty} \Pi(n) = \Pi$
Or $\Pi(n) = \Pi(0)P^n$, donc si $\lim_{n\to+\infty} \Pi(n)$ existe, forcément $\lim_{n\to+\infty} P^n$ existe aussi.
Si $P^\star = \lim_{n\to+\infty} P^n$ alors $\Pi = \Pi_0 P^\star$. La convergence d'une chaîne de Markov ne dépend que de sa matrice de transition.

c) Il ne faut pas confondre entre une distribution limite et une distribution stationnaire. Une distribution limite est forcément stationnaire, mais l'inverse est généralement faux.
Si $P^\star = \lim_{n\to+\infty} P^n$ alors $\Pi = \Pi_0 P^\star$.
La convergence d'une chaîne de Markov ne dépend que de sa matrice de transition.

Proposition 4. *si Π est la distribution limite d'une chaîne ergodique, alors elle est l'unique distribution stationnaire*

Théorème 3. *Si $(X_n)_n$ est une chaîne de Markov homogène finie, alors il existe au moins une distribution stationnaire(ce qui n'est pas le cas pour une chaîne à espace d'états infini).*

Exemple 4
Soit la matrice de transition

$$P = \begin{pmatrix} 0 & 1/2 & 01/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

les distributions stationnaires sont données par $\pi = (0{,}1-2\alpha{,}\alpha{,}\alpha)$, $0 \leq \alpha \leq \frac{1}{2}$.
On voit bien qu'il y a une infinité de distributions stationnaires.

Théorème 4. *Soit $(X_n)_n$ est une chaîne de Markov homogène finie de matrice de transition P. Si $\exists n_0 \in \mathbb{N}^\star$ tel que P^{n_0} a tous ses éléments qui sont strictement positifs, alors $\lim_{n\to+\infty} \Pi(n) = \Pi$ indépendement de Π_0 et $\lim_{n\to+\infty} P^n = P^\star$.*
Π est un vecteur de probabilité de composantes strictement positives et $P^\star$ est une matrice dont toutes les lignes sont identiques à Π et $\Pi P^\star = \Pi$

Exemple 5
Soit la matrice de transition

$$P = \begin{pmatrix} 1/2 & 01/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

La matrice P^2 a tous ses élements qui sont strictement positifs. La distribution stationnaire est $\Pi = (2/7{,}3/7{,}2/7)$ et $\lim\limits_{n \to +\infty} P^n = P^\star$.
$P^\star$ a toutes ses lignes qui sont identiques à Π.

Théorème 5. *Conditions suffisantes*
Si une matrice de transition P n'a pas d'éléments nuls sur sa diagonale principale et sur ses deux diagonales obliques qui l'encadrent alors, la chaîne admet une distribution limite.

Théorème 6. *Conditions nécessaires*
Soit $(X_n)_n$ est une chaîne de Markov homogène finie de matrice de transition P. Si la valeur propre $\lambda = 1$ est simple et que toutes les autres propres sont de module strictement inférieur 1, alors la chaîne admet une distribution limite.

Exemple 7
Soit $P = \begin{pmatrix} 1/4 & 01/4 & 1/2 \\ 0 & 1/4 & 3/4 \\ 0 & 3/4 & 1/4 \end{pmatrix}$

$$det(P - \lambda I) = \frac{1 - 4\lambda}{8}(\lambda - 1)(2\lambda + 1)$$

$\lambda_1 = 1$ est simple et $|\lambda_2| = \frac{1}{4} < 1$, $|\lambda_3| = |-\frac{1}{2}| < 1$
la chaîne associée P admet une distribution limite.

2.1.2 Classification des états d'une chaîne de Markov

Définition 18. *Accessibilté*
Soient i, $j \in E$. On dit dit que l'état i est accessible partir de l'état j si $\exists n_0 \in \mathbb{N}$ $P_{ij}^{(n_0)}) > 0$. *On note* $i \curvearrowright j$.

Définition 19. *Communication*
On dit que les états i et j sont communicants si $\exists$ $m,n \in \mathbb{N}^\star / P_{ij}^{(n)} > 0$ *et* $P_{ji}^{(m)} > 0$. *On note* $i \leftrightarrow j$

On définit sur E une relation R donnée par $iRj \Leftrightarrow i \leftrightarrow j$

Proposition 5. *R est une classe d'équivalence*

Démonstration 4. *- R est reflexive car* $\forall i \in E$, iRi *du fait que* $P_{ii}^{(0)} = 1$.
- R est symétrique (par définition)
- R transitive
Si iRj et jRk, montrons que iRk.
$iRj \Leftrightarrow \exists\ n_1,m_1 \in \mathbb{N}^\star / P_{ij}^{(n_1)} > 0$ *et* $P_{ji}^{(m_1)} > 0$

$jRk \Leftrightarrow \exists\ n_2,m_2 \in \mathbb{N}^\star / P_{jk}^{(n_2)} > 0$ *et* $P_{kj}^{(m_2)} > 0$

$$P_{ik}^{(n_1+n_2)} = P(X_{n_1+n_2} = k/X_0 = i) = \sum_{l \in E} P(X_{n_1} = l/X_0 = i)P(X_{n_1+n_2} = k/X_{n_1} = l)$$

$$= \sum_{l \in E} P_{il}^{(n_1)} P_{lk}^{(n_2)} = P_{ij}^{(n_1)} P_{jk}^{(n_2)} + \sum_{l \in E,\, l \neq j} P_{il}^{(n_1)} P_{lk}^{(n_2)} > 0$$

De même $P_{ki}^{(n_2+m_2)} > 0$ *et donc* iRk

Remarque 6
La relation d'équivalence R va partitionner l'espace des états E en classes d'équivalence $C_1,C_2,\ldots,C_r$ telles que $E = \bigcup_{i=1}^{r} C_i$, $C_i \bigcap C_j \neq \emptyset$ $i \neq j$. La communication est une propriétè de classe, i.e., que les états d'une même classe sont tous communicants.

Définition 20. *Chaîne irréductible*
Une chaîne de markov homogène finie est dite irréductible si la relation d'équivalence R n'induit qu'une seule classe d'équivalence, i.e.; que tous les états sont communicants.

Exemple Tous les états de l'exemple 7 sont communicants.

Définition 21. *Classe récurrente*
Une classe est dite récurrente, s'il est impossssible de la quitter.

Définition 22. *Classe transitoire*
Une classe est dite transitoire, s'il est posssible d'en sortir mais jamais de revenir.

Exemple 8

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

Les classes transitoires sont (1,2) et (3). (4) est une classe récurrente.

Définition 23. *Classe fermée*
Un ensemble d'états C est dit fermé si :$\forall i \in C, \forall j \notin C \ P_{ij} = 0$

Proposition 6. *Si C est un ensemble d'états fermé, alors $\sum_{j \in C} P_{ij}^{(n)} = 1$ (Au bout de n transitions , on est toujours dans C).*

Démonstration 5. *- Si $n = 0$, $P_{ii}^{(0)} = \sum_{j \in C} P_{ij}^{(n)} = 1$*
- Soit $i \in C$,

$$\sum_{j \in C} P_{ij}^{(n)} = \sum_{j \in C} P(X_n = j/X_0 = i) = \sum_{j \in E} P(X_n = j, \bigcup_{k \in E} (X_{n-1} = k)/X_0 = i)$$

$$= \sum_{j \in C} \sum_{k \in E} P(X_{n-1} = k/X_0 = i) P(X_n = j)/X_{n-1} = k)$$

$$= \sum_{j\in C}\sum_{k\in E} P_{ik}^{(n-1)} P_{kj} = \sum_{j\in C}\sum_{k\in C} P_{ik}^{(n-1)} P_{kj} + \underbrace{\sum_{j\in C}\sum_{k\notin C} P_{ik}^{(n-1)} P_{kj}}_{0}$$

$$= \sum_{j\in C}\sum_{k\in C} P_{ik}^{(n-1)} P_{kj} = \sum_{k\in C} P_{ik}^{(n-1)} \underbrace{\sum_{j\in C} P_{kj}}_{1} = \sum_{k\in C} P_{ik}^{(n-1)} = 1$$

Celà vient du fait que $\sum_{k\in E} P_{ij}^{(n-1)} = \sum_{k\in C} P_{ij}^{(n-1)} + \underbrace{\sum_{k\notin C} P_{ij}^{(n-1)}}_{0} = 1$

Définition 24. *Etat absorbant*
Si un ensemble C fermé est réduit à un seul élément i, alors cet état est dit absorbant, i.e.,$\forall n \geq 0$ $P_{ii}^{(n)} = 1$ (quand on est dans l'état i, on ne sort plus).

Définition 25. *Périodicité d'un état*
Un état i est dit périodique de période $d(i)$ si

$$d(i) = pgcd\{n \geq 1,\, P_{ii}^{(n)} > 0\} > 1$$

- Si $d(i) = 1$, on dit que l'état est apériodique.
- Si $n \neq kd(i)$, $k \in \mathbb{N}^\star$ alors $P_{ii}^n = 0$.(l'état i ne peut être revisité qu'au bout d'un nombre de transitions multiple de $d(i)$).

Exemple 9

• Soit $P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}$

$$d(1) = pgcd\{n \geq 1, P_{11}^{(n)} > 0\} = pgcd\{4,6,8,\ \ldots,\} = 2 = d(2) = d(3) = d(4)$$

Tous les états sont communicants, la chaîne est périodique de période 2.

•• Soit $P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 1/3 \\ 1/4 & 0 & 3/4 \end{pmatrix}$

$d(1) = pgcd\{2,3,4,\ \ldots,\} = 1$. La chaîne est irréductible, donc tous les états sont de même période (voir exercice) et donc la chaîne est apériodique.

2.1.3 Classification probabiliste des Chaînes de Markov

Soit $(X_n)_n$ une chaîne de Markov homogène et finie.

a-Probabilité de premier passage

Notons $f_{ij}(n) = P(X_n = j,\, X_{n-1} \neq j,\, \ldots; X_1 \neq j / X_0 = i)$
$f_{ij}(n)$ est la probabilité pour que le système soit pour la première fois dans l'état j sachant qu'il a démarré de i.
$f_{ij}(1) = P(X_1 = j/X_0 = i) = P_{ij}$, $f_{ii}(0) = 1$.

De même, notons $F_{ij}(n) = \sum_{k=1}^{n} f_{ij}(k)$ = probabilité que le système atteigne j en moins de $(n+1)$ transitions sachant qu'il a démarré de i.

Soit aussi $F_{ij}(+\infty) = \sum_{k=1}^{+\infty} f_{ij}(k)$= probabilité que le système atteigne j en démarrant de i.

$F_{ij}(+\infty) > 0$ signifie que j est accessible à partir de i.

$F_{ii}(+\infty)$: est la probabilité que le système revienne pour la première fois dans i sachant qu'il a démarré de i.

b- Instant de premier passage

Soit T_{ij} la variable aléatoire définie par $T_{ij} = Inf_k\{k \geq 1,\, X_k = j/X_0 = i\}$
T_{ij} est l'instant de premier passage en j et prend ses valeurs dans $\mathbb{N}^\star \bigcup (+\infty)$.

$T_{ij} = +\infty$, si $\forall k\ X_k \neq j$

Notons que $f_{ij}(n) = P(T_{ij} = n) = P(X_n = j,\, X_{n-1} \neq j,\, \ldots; X_1 \neq j/X_0 = i)$

Définition 26. *Etat récurrent et transitoire*
Un état i est dit récurrent si $F_{ii}(+\infty) = 1$
Il sera dit transitoire si $F_{ii}(\infty) < 1$.

$F_{ii}(+\infty) = 1$ signifie que le système partant de i repassera à coup sûr dans i.
$F_{ii}(+\infty) < 1$ signifie que le système démarrant de i peut ne pas repasser par i.

Classification des états récurrents
On distingue deux types d'états récurrents:

- Récurrent nul si $\mu_i = \sum_{n\geq 1} f_{ii}(n) = +\infty = \sum_{n\geq 1} P(T_{ii} = n) = E(T_{ii})$
- Récurrent non-nul si $\mu_i = \sum_{n\geq 1} f_{ii}(n) < +\infty$

μ_i est le nombre moyen de retours en i.

Exemple 9

$$\text{Soit } P = \begin{pmatrix} 1/4 & 3/4 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1/3 & 2/3 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(5) et (4) sont transitoires. 1,2 et (3) sont récurrents.
$F_{11}(+\infty) = 1$?

$$f_{11}(1) = P(X_1 = 1/X_0 = 1) = 1/4$$

$$f_{11}(2) = 3/4 \times 1/2,\ f_{11}(3) = 3/4(1/2)^2, \ldots,\ f_{11}(n) = 3/4(1/2)^{n-1}$$

$F_{11}(+\infty) = 1/4 + 3/4 \sum_{n\geq 2} (1/2)^{n-1} = 14 + 3/4 = 1$

$\mu_1 = \sum_{n\geq 1} f_{ii}(n) = 1/4 + 3/4 \sum_{n\geq 2} n(1/2)^{n-1} = \sum_{n\geq 1} a_n$

$\lim\limits_{n\to+\infty} \dfrac{a_{n+1}}{a_n} < 1$, ce qui implique que $\sum_{n\geq 1} a_n < +\infty$ et donc 1 est récurrent non nul.
(3) est absorbant et donc récurrent.

Théorème 7. *On a*

$$P_{ij}^{(n)} = \sum_{m=0}^{n} f_{ij}(m) P_{jj}^{(n-m)}$$

et

$$f_{ij}(n) = \sum_{k\in E,\ k\neq j} P_{ik} f_{kj}(n-1)$$

Démonstration 6. *1-*

$$P_{ij}^{n} = P(X_n = j/X_0 = i) = P(X_n = j, \bigcup_{m=0}^{n} (T_{ij} = m)/X_0 = i)$$

Car $\Omega = [\bigcup_{m\geq 0} (T_{ij} = m)] \bigcup (T_{ij} = +\infty)$

$$(X_n = j \bigcap \Omega = \bigcup_{m\geq 0}(X_n = j,\, T_{ij} = m) = \bigcup_{m=0}^{n}(X_n = j,\, T_{ij} = m)$$

$$P_{ij}^{(n)} = P(X_n = j,\, T_{ij} = m/X_0 = i) = \sum_{m=0}^{n} P(T_{ij} = m/X_0 = i)P(X_n = j/T_{ij} = m,\, X_0 = i)$$

$$= \sum_{m=0}^{n} f_{ij}(m)P(X_m = j/X_{m-1} \neq j, \dots, X_1 \neq j,\, X_0 = i) = \sum_{m=0}^{n} f_{ij}(m)P_{jj}^{(n-m)}$$

2-

$$f_{ij}(n) = P(T_{ij} = n) = P(X_n = j,\, X_{n-1} \neq j, \dots, X_1 \neq j/X_0 = i)$$

Or $\Omega = (X_1 < +\infty) = \bigcup_{k\in E}(X_n = k)$*, ce qui implique que*

$$f_{ij}(n) = P(X_1 < +\infty, X_1 \neq j,\ \dots,\ X_{n-1} \neq j, X_n = j/X_0 = i)$$

$$= P(\bigcup_{k\in E}(X_n = k),\, X_1 \neq j,\ \dots,\ X_{n-1} \neq j, X_n = j/X_0 = i)$$

$$= \sum_{k\in E} P(X_1 = k,\, X_1 \neq j,\ \dots,\ X_{n-1} \neq j, X_n = j/X_0 = i)$$

Si $k = j$, $(X_1 = k)\bigcap(X_1 \neq j) = \emptyset$

$$f_{ij}(n) = \sum_{\substack{k\in E\\ k\neq j}} \sum = P(X_1 = k,\, X_2 \neq j,\ \dots,\ X_{n-1} \neq j, X_n = j/X_0 = i)$$

$$= \sum_{\substack{k\in E\\ k\neq j}} \sum P_{ik}P(X_n = j,\, X_{n-1} \neq j,\ \dots,\ ,X_2 \neq j/X_1 = k) = \sum_{\substack{k\in E\\ k\neq j}} \sum P_{ik}f_{kj}(n-1)$$

Théorème 8. *Soit* $(X_n)_n$ *une chaîne de Markov homogène finie, alors*
i *récurrent* $\Leftrightarrow \sum_{n\geq 0} P_{ii}^{(n)} = +\infty$
i *transitoire* $\Leftrightarrow \sum_{n\geq 0} P_{ii}^{(n)} < +\infty$

Démonstration 7. *On a montré que* $P_{ij}^{(n)} = \sum_{m=0}^{n} f_{ij}(m)P_{jj}^{(n-m)} \longrightarrow \sum_{n=1}^{N} P_{ij}^{(n)} = \sum_{n=1}^{N}\sum_{m=0}^{n} f_{ij}(m)P_{jj}^{(n-m)}$

En inversant les sommes et en posant $n - m = k$*, on obtient*

$$\sum_{n=1}^{N} P_{ij}^{(n)} = \sum_{m=1}^{N} f_{ij}(m) \sum_{k=0}^{N-m} P_{jj}^{(k)} \leq \sum_{m=1}^{N} f_{ij}(m) \sum_{k=0}^{N} P_{jj}^{(k)} \tag{1}$$

$$\sum_{m=1}^{N} f_{ij}(m) \sum_{k=0}^{N-m} P_{jj}^{(k)} \geq \sum_{m=1}^{N} f_{ij}(m) \sum_{k=0}^{N-N_1} P_{jj}^{(k)}, \text{ avec } N_1 < N \tag{2}$$

On peut prendre $N_1 = [\frac{N}{2}] + 1$
De (1), on en déduit que

$$\sum_{m=1}^{N} f_{ij}(m) \geq \frac{\sum_{n=1}^{N} P_{ij}^{(n)}}{\sum_{k=0}^{N} P_{jj}^{(k)}} = \frac{\sum_{n=1}^{N} P_{ij}^{(n)}}{1 + \sum_{k=1}^{N} P_{jj}^{(k)}}$$

.

De (2), on en déduit que

$$\sum_{m=1}^{N} f_{ij}(m) \leq \frac{\sum_{n=1}^{N} P_{ij}^{(n)}}{\sum_{k=0}^{N-N_1} P_{jj}^{(k)}} = \frac{\sum_{n=1}^{N} P_{ij}^{(n)}}{1 + \sum_{k=1}^{N-N_1} P_{jj}^{(k)}}$$

En posant $i = j$ *dans l'équation précedente et en faisant tendre* N *vers l'infini, on obtient*

$$\sum_{n=1}^{+\infty} f_{ii}(n) = F_{ii}(+\infty) \leq \frac{\sum_{n=1}^{+\infty} P_{ii}^{(n)}}{1 + \sum_{n=1}^{+\infty} P_{ii}^{(n)}}$$

Dans (1) en faisant tendre N *vers l'infini, on obtient*

$$\sum_{n=1}^{+\infty} f_{ii}(n) = F_{ii}(+\infty) \geq \frac{\sum_{n=1}^{+\infty} P_{ii}^{(n)}}{1 + \sum_{n=1}^{+\infty} P_{ii}^{(n)}}$$

Ce qui donne $\sum_{n=1}^{+\infty} f_{ii}(n) = F_{ii}(+\infty) = \dfrac{\sum_{n=1}^{+\infty} P_{ii}^{(n)}}{1 + \sum_{n=1}^{+\infty} P_{ii}^{(n)}}$

Si $F_{ii}(+\infty) = 1 \longrightarrow \sum_{n\geq 1} P_{ii}^{(n)} = +\infty$

Si $F_{ii}(+\infty) < 1 \longrightarrow \sum_{n\geq 1} P_{ii}^{(n)} < \infty$, *d'où le résultat.*

Remarques 7
a) Si i est transitoire aors $\lim\limits_{n\to+\infty} P_{ii}^{(n)} = 0$ car la série $\sum_{n\geq 1} P_{ii}^{(n)} < \infty$.
b) Soit A_n la variable aléatoire définie par

$$A_n = \begin{cases} 1 & \text{si } X_n = i \\ 0 & \text{sinon.} \end{cases}$$

Posons $N_i = \sum_{n\geq 0} A_n$ (nombre de fois où le processus visite l'état i).

$$E(N_i/X_0 = i) = \sum_{n\geq 0} E(A_n/X_0 = i) = \sum_{n\geq 0} P(X_n = i/X_0 = i) = \sum_{n\geq 0} P_{ii}^{(n)}$$

Théorème 9. *Si i est récurrent et $i \longleftrightarrow j$ alors j est récurrent.*

Démonstration 8. *$i \leftrightarrow j$ implique que $\exists\, m,k \in \mathbb{N}^\star / P_{ij}^{(n)} > 0$ et $P_{ji}^{(k)} > 0$.*

$$P_{jj}^{(m+n+k)} = P(X_{m+n+k} = j / X_0 = i) = P(X_{m+n+k} = j,\ \bigcup_{u\in E}(X_{n+k} = u),\ \bigcup_{v\in E}(X_k = v)/X_0 = j)$$

$$= \sum_{u\in E}\sum_{v\in E} P(X_{m+n+k} = j,\ X_{n+k} = u, X_k = v / X_0 = j)$$

$$= \sum_{u\in E}\sum_{v\in E} P(X_k = v / X_0 = j) P(X_{n+k} = u / X_k = v) P(X_{m+n+k} = j / X_{n+k} = u)$$

$$= \sum_{u\in E}\sum_{v\in E} P_{jv}^{(k)} P_{vu}^{(n)} P_{uj}^{(m)} \geq P_{ji}^{(k)} P_{ii}^{(n)} P_{ij}^{(m)}$$

En sommant sur n, on obtient

$$= \sum_{n\geq 0} P_{jj}^{(m+n+k)} \geq (\sum_{n\geq 0} P_{ii}^{(n)}) P_{ij}^{(m)} P_{ji}^{(k)},\ \text{ce qui implique que } \sum_{n\geq 0} P_{ii}^{(n)} = +\infty.$$

Remarque 8

Dans une chaîne de Markov homogène finie et irréductible, tous les états sont de même nature:

- Tous transistoires
- Tous récurrents nuls ou non-nuls.
- Périodiques de même période.