MASTER 1 de MATHEMATIQUES **Université Paul Sabatier**

1M8M06M , 2006-2007 **U.F.R. M.I.G.**

36 h cours, 42 h TD.

INTRODUCTION

Ce cours comprend trois chapitres :

- les martingales discrètes (environ 10h30),

- les chaînes de Markov à états finis ou dénombrables (environ 24h),

- les processus de Poisson (environ 7h30).

Il s'agit de modèles mathématiques rendant compte de phénomènes concrets, physiques ou économiques par exemple. Ces sujets figurent également au programme de l'écrit de l'agrégation (analyse et probabilités) et à celui de l'option orale de mathématiques appliquées.

On se place sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et on définit ce que l'on appelle un **processus stochastique**, à savoir une application

$$X : (\Omega, F) \to \mathbb{R}$$

avec F ensemble des réels ou des entiers positifs, telle que pour tout $n \in F$, l'application, notée X_n, $X(., n) : (\Omega, \mathcal{A}) \to \mathbb{R}$ est une variable aléatoire réelle. Martingales discrètes et chaînes de Markov sont des processus stochastiques à temps discret ; les processus de Poisson sont des processus stochastiques à temps continu.

Définition 0.1 *On appelle* **filtration** *sur l'espace de probabilité* $(\Omega, \mathcal{A}, \mathbb{P})$ *une famille croissante de tribus* $\mathcal{F}_n$ *contenues dans* $\mathcal{A}$, $n \in F$, $F = \mathbb{R}^+$ *ou* $F = \mathbb{N}$..
On dit qu'un processus X *est* **adapté** *à la filtration* $(\mathcal{F}_t, t \in F)$ *si pour tout* $t \in F, X_t$ *est* $\mathcal{F}_t$*-mesurable. On dit qu'il est* **prévisible** *si pour tout* $t \in F, X_t$ *est* $\mathcal{F}_s$*-mesurable pour tout* $s \in F$, $s < t$.
On note $\mathcal{F}_\infty$ *la plus petite tribu contenant toutes les* $\mathcal{F}_n$: $\mathcal{F}_\infty = \vee_{t \in F} \mathcal{F}_t$.

Si un processus stochastique X est donné, on appelle **filtration naturelle** associée à X, notée $\mathcal{F}^X$ la filtration définie par la suite croissante de tribus $\mathcal{F}_n^X = \sigma(X_0, \cdots, X_n)$. Par définition, le processus X est adapté à sa filtration naturelle.

0.1 Plan

Martingales à temps discret :

Définitions, exemples, décompositions, temps d'arrêt, théorème d'arrêt, inégalités, théorèmes de convergence, martingales renversées et exemple de l'algorithme de Robbins Monro (cf. [1], chapitre 3).

Chaînes de Markov à temps discret sur un ensemble d'états fini ou dénombrable :

Définitions, exemples, fonctions excessives et invariantes, transience et récurrence, critères de transience, classification des états. Mesure invariante, classes de récurrence, périodes, convergence vers la mesure invariante, théorèmes ergodiques, temps d'atteinte, Exemples avec les files d'attente.

Processus de Poisson :

Définition par processus de comptage/temps de saut, propriété de Markov, propriété d'accroissements indépendants stationnaires.

1 Martingales à temps discret

Voir le livre de Jacques NEVEU [6] ou [1] chapitre 3.

La théorie des martingales a son origine dans l'étude des jeux : elle modélise d'une part le caractère aléatoire d'un phénomène mais aussi son évolution dans le temps. On étudie ici le temps discret.

1.1 Définitions et premiers exemples

Définition 1.1 *Soit un processus adapté X sur l'espace de probabilité filtré $(\Omega, \mathcal{A}, (\mathcal{F}_n, n \in \mathbb{N}), \mathbb{P})$ tel que pour tout entier n, X_n est intégrable. On dit que X est une* **martingale** *si pour tout entier $n, E[X_{n+1}/\mathcal{F}_n] = X_n$ presque sûrement.*
On dit que X est une **sous-martingale** *si pour tout entier $n, E[X_{n+1}/\mathcal{F}_n] \geq X_n$ presque sûrement.*
On dit que X est une **sur-martingale** *si pour tout entier $n, E[X_{n+1}/\mathcal{F}_n] \leq X_n$ presque sûrement.*

L'outil fondamental de ce chapitre, vu au premier semestre, est l'**espérance conditionnelle** ; il y a tout intérêt à se reporter au paragraphe 1.4 de [1] et au cours du premier semestre de monsieur BARTHE.

1.1.1 Exemples et exercices

(a) Montrer que si $H \in L^1(\Omega, \mathcal{A}), X_n = E[H/\mathcal{F}_n]$ est une martingale sur l'espace de probabilité filtré $(\Omega, \mathcal{A}, (\mathcal{F}_n, n \in \mathbb{N}), \mathbb{P})$.

(b) Soit $(Z_n, n \in \mathbb{N})$ une suite de variables aléatoires indépendantes et intégrables et $X_n = \sum_{i=1}^n Z_i$. Montrer que les tribus $\mathcal{F}^X$ et $\mathcal{F}^Z$ sont égales et
si pour tout entier $n, E(Z_n) = 0, X = (X_n)$ est une martingale pour sa filtration naturelle,
si pour tout entier $n, E(Z_n) \geq 0, X = (X_n)$ est une sous-martingale pour sa filtration naturelle,
si pour tout entier $n, E(Z_n) \leq 0, X = (X_n)$ est une sur-martingale pour sa filtration naturelle.

(c) L'exemple le plus simple de théorie des jeux est celui où les variables aléatoires Z_n sont des variables aléatoires de loi de Bernoulli de paramètre p avec valeurs $+1$ et -1. Alors, X_n représente la fortune d'un joueur après n paris. C'est une martingale pour sa filtration naturelle ; on l'appelle **marche aléatoire**.

Proposition 1.2 QC
Soit X une $\mathcal{F}$-martingale. Alors

$$(a) \forall n \geq 0, \forall k \geq 0, E[X_{n+k}/\mathcal{F}_n] = X_n \; ; \; E[X_n] = E[X_0].$$

(b) Si la martingale est de carré intégrable (soit pour tout n, X_n^2 est intégrable), les accroissements disjoints de X sont orthogonaux.
(c) Si X est une sur-martingale, $-X$ est une sous-martingale.
(d) L'espace des martingales est un espace vectoriel réel.
(e) Si X est une martingale et ϕ une application convexe mesurable telle que $\phi(X_n)$ est intégrable, alors le processus Y défini par $Y_n = \phi(X_n)$ pour tout n est une sous-martingale.

Preuve en exercice à chercher en amphi.
a) par récurrence
b) Pour $k \geq p \geq 0$, calculer $E[(X_{n+k} - X_{n+p})(X_{n+p} - X_n)]$ en utiliser à bon escient l'espérance conditionnelle sous l'espérance.
d) $L^1(\mathcal{F}_n)$ est un espace vectoriel et l'espérance conditionnelle est un opérateur linéaire.
e) ϕ mesurable montre que $\phi(X_n)$ est $\mathcal{F}_n$-mesurable, elle est intégrable par hypothèse et l'inégalité de Jensen appliquée à $E[\phi(X_{n+1})/\mathcal{F}_n]$ permet de conclure. •

Corollaire 1.3 *Si X est une sous-martingale, et ϕ croissante et convexe, $\phi(X)$ est une sous-martingale.*

1.2 Décompositions

Théorème 1.4 QC 2 *Décomposition de Doob : soit X une sous-martingale ; il existe une martingale M et un processus croissant prévisible A tels que pour tout entier n, $X_n = M_n + A_n$.*

Le processus A est appelé "compensateur" de X. En quelque sorte, A mesure le "défaut" de martingalité de X.
Preuve : On part de $A_0 = 0$ et $M_0 = X_0$. Puis on définit les deux suites par récurrence :

$$A_1 = E[X_1/\mathcal{F}_0] - X_0 \ ; \ M_1 = X_1 - A_1.$$

Puis :

$$A_n = E[X_n/\mathcal{F}_{n-1}] - M_{n-1} \ ; \ M_n = X_n - A_n.$$

Ainsi, par construction la suite M est adaptée et A_n est prévisible ; A_n et M_n sont intégrables par récurrence. Enfin, on calcule

$$E[M_n/\mathcal{F}_{n-1}] = E[X_n/\mathcal{F}_{n-1}] - E[A_n/\mathcal{F}_{n-1}] = E[X_n/\mathcal{F}_{n-1}] - A_n = M_{n-1}$$

par définition et parce que A est prévisible ($\mathcal{F}_{n-1}$ mesurable). •

Théorème 1.5 *Décomposition de Krickeberg : soit X une martingale bornée dans L^1 ; il existe deux martingales positives Y et Z telles que $X_n = Y_n - Z_n$ pour tout entier n.*

Preuve : L'hypothèse est qu'il existe une constante K telle que pour tout n, $E(|X_n|] \leq K$. Or, la suite $(|X_n|, n \in \mathbb{N})$ est une sous martingale : la suite des espérances $(E[|X_n|], n \in \mathbb{N})$ est donc croissante majorée par K et par la décomposition de Doob il existe une martingale M et un processus croissant prévisible A tels que pour tout entier $n, |X_n| = M_n + A_n$. La suite croissante A_n admet une limite presque sûre notée A_∞ : $|X_n| \leq M_n + A_\infty$. Or $A_n = |X_n| - M_n, E(A_n) \leq K - E[M_0]$ et la limite A_∞ est intégrable : $E(A_\infty) \leq K - E[M_0]$ et on peut donc en prendre l'espérance conditionnelle :

$$\forall n, |X_n| \leq M_n + E[A_\infty / \mathcal{F}_n] = Y_n.$$

La suite Y est une martingale positive, $Y - X$, notée Z, aussi puisque $Y \geq |X| \geq X$ et on a bien $X = Y - Z$. •

1.3 Arrêt

Définition 1.6 *On appelle* **temps d'arrêt** *relatif à une filtration* $(\mathcal{F}_n, n \in \mathbb{N})$ *une variable aléatoire* T *à valeurs dans* $\bar{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$ *telle que pour tout* $n \in \mathbb{N}$, *l'événement* $\{T = n\} \in \mathcal{F}_n$.
Remarquer que $\{T = n\} \in \mathcal{F}_n$ *équivaut à* $\{T \leq n\} \in \mathcal{F}_n$

Exemples :
. L'exemple le plus simple : $T = n_0$ constant est un $\mathcal{F}-$temps d'arrêt.
. Soit $(X_n, n \in \mathbb{N})$ une suite de variables aléatoires réelles, A un borélien de $\mathbb{R}$, on définit

$$T = \min\{n, X_n \in A\} \text{ si l'ensemble n'est pas vide, } + \infty \text{ sinon.}$$

Alors T est un temps d'arrêt pour la filtration naturelle de X. On l'appelle le **temps d'entrée** dans A.

Proposition 1.7 *Si* T *et* S *sont des* $\mathcal{F}$*-temps d'arrêt alors*

$$\sup(S,T) \textit{ noté } S \vee T, \ \inf(S,T) \textit{ noté } S \wedge T, \textit{ sont des } \mathcal{F} - \textit{temps d'arrêt.}$$

Preuve : en TD ; le principe général de ces preuves est le suivant : l'ensemble Ω se décompose selon la réunion disjointe

$$\Omega = \cup_{n \geq 0} \{T = n\} \cup \{T = +\infty\}$$

le dernier sous-ensemble étant vide si T est presque sûrement fini.

On note $\mathcal{F}_T := \{A \in \mathcal{F}_\infty; A \cap \{T \leq n\} \in \mathcal{F}_n, \forall n\}$. On l'appelle la **tribu des événements antérieurs** à T.

Proposition 1.8 *Une variable aléatoire X est $\mathcal{F}_T-$mesurable si et seulement si $\forall n$, $X\mathbf{1}_{\{T\leq n\}}$ est $\mathcal{F}_n-$mesurable. De plus, ceci est équivalent à $\forall n$, $X\mathbf{1}_{\{T=n\}}$ est $\mathcal{F}_n-$mesurable.*

Preuve :

Proposition 1.9 *a. Si T est un $\mathcal{F}$-temps d'arrêt, T est $\mathcal{F}_T$-mesurable.*

b. Si S est un temps d'arrêt tel que $T \leq S$, alors $\mathcal{F}_T \subset \mathcal{F}_S$.

Preuve : en exercice en amphi
a. $\{T \leq n\} \cap \{T = k\}$ est soit vide soit $\{T = k\}$ les deux appartenant à $\mathcal{F}_k$.
b. Soit $A \subset \mathcal{F}_T$, $A \cap \{S = k\} = \cap_{n\leq k} A \cup \{T = n\} \cap \{S = k\} \in \mathcal{F}_k$.

Proposition 1.10 *Soit une filtration $(\mathcal{F}_n, n \in \mathbb{N})$, $(X_n, n \in \mathbb{N})$ une suite de variables aléatoires réelles adaptées, et T un $\mathcal{F}$-temps d'arrêt. Alors X_T est $\mathcal{F}_T$-mesurable où*

$$X_T := \sum_{n\in\mathbb{N}} X_n \mathbf{1}_{\{T=n\}}.$$

Remarquer que X_T n'est définie que sur l'événement $\{T < \infty\}$. Hors de cet événement, X_T est nulle par convention, sauf si la limite presque sûre de la suite X_n existe auquel cas $X_T := \lim_{n\to\infty} X_n$ sur l'événement $\{T = \infty\}$.
Preuve : à chercher en amphi.
Il suffit de vérifier pour tout n que $X_T\mathbf{1}_{\{T\leq n\}} = \sum_{k=0}^n X_k\mathbf{1}_{\{T=k\}}$ est $\mathcal{F}_n-$mesurable. •

Lemme 1.11 *Pour toute variable aléatoire $g \in L^1$ ou positive, et T un $\mathcal{F}$-temps d'arrêt presque sûrement fini, $E[g/\mathcal{F}_T] = \sum_n E[g/\mathcal{F}_n]\mathbf{1}_{\{T=n\}}$.*

Preuve : (i) le "candidat" à être l'espérance conditionnelle est $\mathcal{F}_T$-mesurable par construction.
(ii) Soit $A \in \mathcal{F}_T$: $A \cap \{T = n\} \in \mathcal{F}_n$. L'espérance $E(g\mathbf{1}_A) = \sum_{\{n\geq 0\}} E(g\mathbf{1}_{A\cap\{T=n\}}) = \sum_{\{n\geq 0\}} E[E(g/\mathcal{F}_n)\mathbf{1}_{A\cap\{T=n\}}] = E[\sum_{\{n\geq 0\}} E(g/\mathcal{F}_n)\mathbf{1}_{\{T=n\}}\mathbf{1}_A]$. •

Théorème 1.12 *d'arrêt*
1. Soit $(X_n, n \in \mathbb{N})$ une $\mathcal{F}$-(sous, sur-)martingale et T un $\mathcal{F}$-temps d'arrêt.
Alors $(X_{n\wedge T}, n \in \mathbb{N})$ est une $\mathcal{F}$-(sous, sur)martingale.
2. Si T et S sont deux $\mathcal{F}$-temps d'arrêt bornés tels que $S \leq T$; alors
$E[X_T/\mathcal{F}_S](\geq, \ \leq) = X_S$.

Preuve : 1. On vérifie que pour tout $n, X_{T\wedge n}$ est $\mathcal{F}_n$ mesurable et intégrable :

$$X_{T\wedge n} = \sum_{k=0}^{n-1} X_k \mathbf{1}_{\{T=k\}} + X_n \mathbf{1}_{\{T\geq n\}},$$

c'est donc une somme finie d'éléments de $L^1(\Omega, \mathcal{F}_n, \mathbb{P})$.

On se sert de cette même décomposition pour montrer que $E[X_{T\wedge(n+1)}/\mathcal{F}_n] = \sum_{k=0}^{n} X_k \mathbf{1}_{\{T=k\}} + E[X_{n+1}/\mathcal{F}_n]\mathbf{1}_{\{T>n\}}$.

2. Soit N un entier qui domine les deux temps d'arrêt et Z une variable $\mathcal{F}_S$ mesurable et bornée.

$$E[X_N.Z] = \sum_{k=0}^{N} E[X_N.Z\mathbf{1}_{\{S=k\}}] = \sum_{k=0}^{N} E[X_k.Z\mathbf{1}_{\{S=k\}}]$$

puisque X est une martingale et $Z\mathbf{1}_{\{S=k\}}$ est $\mathcal{F}_k$ mesurable. La dernière expression n'est autre que l'espérance de $X_S.Z$ ce qui montre que $E[X_N/\mathcal{F}_S] = X_S$. De même, $E[X_N/\mathcal{F}_T] = X_T$ et en combinant les deux expressions et l'inclusion $\mathcal{F}_S \subset \mathcal{F}_T$, on obtient le résultat.

Pour une sous-martingale, on utilise la décomposition de Doob : $X_n = M_n + A_n, X_T = M_T + A_T, E[X_T/\mathcal{F}_S] = M_S + E[A_T/\mathcal{F}_S]$ et comme A est croissant, A_T est plus grand que A_S donc $E[A_T/\mathcal{F}_S] \geq A_S$. On peut utiliser l'équivalence : si X et Y sont $\mathcal{B}$-mesurables, alors $X \leq Y$ équivaut $\forall B \in \mathcal{B}$, $E(X\mathbf{1}_B) \leq E(Y\mathbf{1}_B)$. •

Corollaire 1.13 *Soit T_n une suite de temps d'arrêt bornés, croissante, alors $(X_{T_n}, n \in \mathbb{N})$ est une (sous,sur) martingale pour la filtration $(\mathcal{F}_{T_n}, n \in \mathbb{N})$.*

Preuve en exercice d'amphi
. X_{T_n} est $\mathcal{F}_{T_n}$ mesurable d'après la proposition 1.10 et intégrable comme somme finie de variables intégrables.
. la suite de tribus $\mathcal{F}_{T_n}$ est croissante (cf. proposition 1.10 b)
. puis le théorème d'arrêt.

Corollaire 1.14 *Soit $(X_n, n \in \mathbb{N})$ une $\mathcal{F}$-(sous)martingale et T un $\mathcal{F}$-temps d'arrêt presque sûrement fini. Alors, si l'une des deux hypothèses suivantes est réalisée :*

(i) T est un temps d'arrêt borné.

(ii) Il existe $Y \in L^1$ telle que pour tout $n, |X_{T\wedge n}| \leq Y$
alors $E[X_0](\leq) = E[X_T]$.

Preuve en exercice d'amphi : dans le premier cas, puisque T et 0 sont deux temps d'arrêt bornés et $T \geq 0$, le théorème d'arrêt montre que $E(X_T/\mathcal{F}_0)(\geq) = X_0$, et on prend l'espérance.

Dans le deuxième cas, on prend d'abord comme temps d'arrêt $T \wedge n$ qui est borné, et vérifie donc $E[X_{T\wedge n}](\geq) = E[X_0]$. L'hypothèse (ii) permet alors d'utiliser le théorème de Lebesgue de convergence dominée, puisque le fait que T est presque sûrement fini implique que $X_{T\wedge n}$ converge presque sûrement vers X_T. •

Corollaire 1.15 *Identité de Wald : soit* $(Y_n, n \in \mathbb{N})$ *une suite de variables aléatoires réelles adaptées, de même loi,d'espérance* m*, et pour tout* n*,* Y_n *est indépendante de* $\mathcal{F}_{n-1}$*. On pose* $S_n = Y_1 + \cdots + Y_n,\ n \geq 1$*. Alors, si* T *un* $\mathcal{F}$*-temps d'arrêt intégrable,* $\forall i,\ E[S_T] = E[Y_i]E(T)$*.*

Preuve :
. On remarque d'abord que T est presque sûrement fini puisqu'il est intégrable.
. Puis, on montre que le processus $(X_n = S_n - nm, n \geq 1)$ est une martingale : pour tout n, $X_n \in L^1(\mathcal{F}_n)$, puis $E[X_{n+1}/\mathcal{F}_n] = X_n + E[Y_{n+1} - m/\mathcal{F}_n] = X_n$ car $Y_{n+1} - m$ est centrée indépendante de $\mathcal{F}_n$.
. Soit $Y'_n = Y_n - m$ et $E(|Y'_n|) = \mu$. On majore pour tout n $|S_{n\wedge T}|$ par $Y = \sum_{1\leq k\leq T} |Y'_k|$.

$E[Y] = \sum_{n\geq 1} E[\sum_{k\leq n} |Y'_k|\mathbf{1}_{\{T=n\}}] = \sum_{1\leq k\leq n} E[|Y'_k|\mathbf{1}_{\{T=n\}}]$
$= \sum_{1\leq k} E[|Y'_k|\mathbf{1}_{\{T\geq k\}} = \sum_{1\leq k} E[|Y'_k|\mathbf{1}_{\{T\leq k-1\}^c}] = \mu\left(\sum_{1\leq k} E[\mathbf{1}_{\{T\geq k\}}]\right)$
car $\forall k,\ |Y'_k|$ est indépendante de $\{T \leq k-1\}$ ce qui fait finalement le produit $\mu E[T]$ (lemme classique : $E(T) = \sum_{1\leq k} \mathbb{P}\{T \geq k\}$).
On peut alors appliquer le (ii) du corollaire 1.14, $E[X_T] = E[X_1] = 0$ et il vient $E[S_T] = m.E[T]$. •

Application (exercice) : **la ruine du joueur**. *Dans l'exemple ci-dessus, on considère que* Y_n *est le résultat d'un jeu au temps* n*, variable aléatoire de Bernoulli valant* $+1$ *ou* -1 *avec probabilité* $p, 1-p$*. La fortune initiale du joueur est* a *et il espère atteindre la quantité* b*. Le jeu est terminé lorsque la fortune du joueur est soit nulle, auquel cas il a perdu, soit égale à* b*, auquel cas il a gagné. On note* T *l'instant de la fin du jeu. On demande*

- la probabilité que le joueur perde, ou gagne,

- l'espérance de T*.*

On remarque d'abord que la somme $S_n = \frac{Y_1+\cdots+Y_n+n}{2}$ suit une loi binomiale $B(n,p)$.

1. T est intégrable : soit $N > a + b > b$, donc $N > \frac{N+b}{2}$.

Soit $Z_n = \sum_{k=(n-1)N+1}^{nN} Y_k$. Les variables aléatoires $\frac{1}{2}(Z_n + N)$ forment une suite de variables aléatoires indépendantes de loi binomiale $B(N,p)$.
Si la fortune du joueur $X_k = a + Y_1 + \cdots + Y_k$ reste dans $]0, b[$ de $k = 1$ à nN, alors $Z_j \leq \frac{b+N}{2} < N$ pour tout j de 1 à n.
On remarque que $\{T \geq nN\} = \cap_{k=1,\ldots,nN}\{X_k \in]0,b[\} \subset \cap_{j=1,\ldots n}\{Z_j < b\}$.
Donc $\mathbb{P}\{T \geq nN\} \leq (B(N,p)]0, N-1])^n$ qui est une série sommable ;
or $E[T] = \sum_k \mathbb{P}\{T \geq k\} \leq N \sum_n \mathbb{P}\{T \geq nN\} < \infty$ et de plus T est presque sûrement fini.

2. Soit $r = \frac{1-p}{p}$. On montre que r^{X_n} est une martingale, par le théorème 1.12 on obtient pour tout n que $E[r^{X_{T\wedge n}}] = r^a$. Puisque par définition de $T, 0 \leq X_{T\wedge n} \leq b$, on peut appliquer le théorème de Lebesgue majoré et il vient $E[r^{X_T}] = r^a$.

3. Or, $X_T = 0$ ou $X_T = b$. D'une part $r^a = \mathbb{P}(ruine) + r^b \mathbb{P}(gain)$ et d'autre part $\mathbb{P}(ruine) + \mathbb{P}(gain) = 1$. Soit

$$\mathbb{P}(gain) = \frac{r^a - 1}{r^b - 1},\ \mathbb{P}(ruine) = \frac{r^b - r^a}{r^b - 1}.$$

4. Enfin, par l'identité de Wald, $E(U_T) = (2p-1)E(T)$, où $U_n = Y_1 + \cdots + Y_n$, $E(U_T) = -a\mathbb{P}(ruine) + (b-a)\mathbb{P}(gain)$ d'où l'on tire $E(T) = \frac{b(1-r^a)-a(1-r^b)}{(2p-1)(1-r^b)}$. •

Le théorème suivant permet de s'affranchir de l'hypothèse que les temps d'arrêt sont presque sûrement bornés.

Théorème 1.16 *de DOOB* **QC2**
Soit $(X_n, n \in \mathbb{N})$ une $\mathcal{F}$-martingale et S et T deux $\mathcal{F}$-temps d'arrêt tels que :

(i) $E[|X_S|]$ et $E[|X_T|]$ sont finies,

(ii) $\lim_{n\to+\infty} \int_{\{T>n\}} |X_n| d\mathbb{P} = \lim_{n\to+\infty} \int_{\{S>n\}} |X_n| d\mathbb{P} = 0$,

(iii) $S \leq T < \infty$ presque sûrement.

Alors $E[X_T/\mathcal{F}_S] = X_S$.

Preuve : Soit $A \in \mathcal{F}_S$ et n entier.
On calcule $E[X_{T\wedge n}\mathbf{1}_A] = E[X_{T\wedge n}\mathbf{1}_{A\cap S\leq n}] + E[X_{T\wedge n}\mathbf{1}_{A\cap S>n}]$.
L'événement $A \cap (S \leq n \in \mathcal{F}_S \cap \mathcal{F}_n = \mathcal{F}_{S\wedge n}$ et $S \wedge n \leq T \wedge n$ montrent que

$$E[X_{T\wedge n}/\mathcal{F}_{S\wedge n} = X_{S\wedge n}$$

parce qu'il s'agit d'une martingale (conséquence du téorème 1.12 1.).
Pour le second terme, $S > n$ implique que $T > n$, et donc ce terme est $E[X_n\mathbf{1}_{A\cap S>n}]$. Au total,

$$E[X_{T\wedge n}\mathbf{1}_A]E[X_S\mathbf{1}_{A\cap S\leq n}] + E[X_n\mathbf{1}_{A\cap S>n}] = E[X_{S\wedge n}\mathbf{1}_A].$$

Puis on fait tendre n vers l'infini dans cette égalité : pour le premier terme, on applique le théorème de convergence dominée avec l'hypothèse (i) puisque $|X_S\mathbf{1}_{A\cap S\leq n}| \leq |X_S|$, et on utilise le (ii) pour faire tendre vers 0 le second terme.
On fait de même pour

$$E[X_T\mathbf{1}_{A\cap T\leq n}] + E[X_n\mathbf{1}_{A\cap T>n}] = E[X_{T.\wedge n}\mathbf{1}_A].$$

•

Rappel : on dit qu'une famille de v.a. $(X_\alpha, \alpha \in I)$ est **uniformément intégrable** (ou équi-intégrable) si

$$\lim_{n\to\infty} \sup_\alpha \int_{\{|X_\alpha|\geq n\}} |X_\alpha| d\mathbb{P} = 0$$

ce qui équivaut à

$$sup_\alpha E[|X_\alpha|] < +\infty\ ;\ \forall\varepsilon > 0, \exists\delta > 0, \forall A \in \mathcal{A} : \mathbb{P}(A) \leq \delta, \int_A |X_\alpha| d\mathbb{P} < \varepsilon.$$

(cf. [2] page 126). Au premier semestre, on a vu : l'uniforme intégrabilité et la convergence en probabilité entraine la convergence dans L^1.

Exemple : Soit une martingale X telle que la famille de v.a. $(X_n, n \in \mathbb{N})$ est uniformément intégrable, alors les conditions ci-dessus (i) et (ii) sont satisfaites. En effet, on a alors par définition, $\sup_n E[|X_n|] < +\infty$ et $\sup_n \int_A |X_n| d\mathbb{P} \to 0$ quand $\mathbb{P}(A) \to_{n\infty} 0$. On en déduit

$$\int_{T>n} |X_n| d\mathbb{P} \leq \sup_m \int_{T>n} |X_m| d\mathbb{P} \to 0$$

puisque T est presque sûrement fini.

Par ailleurs $E[|X_T|] = \lim_{n\to\infty} \sum_{k=0}^n E[|X_T|\mathbf{1}_{T=k}]$. Or, $E[|X_T|\mathbf{1}_{T=k}] = E[|X_k|\mathbf{1}_{T=k}] \leq E[|X_n|\mathbf{1}_{T=k}]$ puisque $|X|$ est une sous martingale ; donc, $\sum_{k=0}^n E[|X_T|\mathbf{1}_{\{T=k\}}] \leq E[|X_n|\mathbf{1}_{\{T\leq n\}}] \leq E[|X_n|] \leq M$.

Donc, en particulier, **une martingale uniformément intégrable vérifie le théorème de Doob pour tout couple de temps d'arrêt S et $T, S \leq T$, presque sûrement finis.**

1.4 Inégalités.

Théorème 1.17 *Soit $(X_n, n \in \mathbb{N})$ une sous martingale positive et $\lambda \geq 0$. Alors*

$$\forall n \in \mathbb{N},\ \lambda \mathbb{P}\{\max_{k\leq n} X_k \geq \lambda\} \leq E[X_n \mathbf{1}_{\{\max_{k\leq n} X_k \geq \lambda\}}] \leq E[X_n].$$

Preuve : Si $T = \inf\{k \in \mathbb{N}, X_k \geq \lambda\}, \{T \leq n\} = \{\max_{k\leq n} X_k \geq \lambda\}$.
Soit les temps d'arrêt bornés $T \wedge n \leq n$, d'après le théorème 1.12,

$E[X_n/\mathcal{F}_{T\wedge n}] \geq X_{T\wedge n}$.
On intègre cette inégalité sur l'événement $\{T \leq n\} \in \mathcal{F}_{T\wedge n}$,

$$E(X_n) \geq E[X_n.\mathbf{1}_{T\leq n}] \geq E[X_{T\wedge n}.\mathbf{1}_{T\leq n}] = E(X_T \mathbf{1}_{T\leq n}) \geq \lambda \mathbb{P}(T \leq n).$$

Corollaire 1.18 *Si $(X_n, n \in \mathbb{N})$ est une martingale, $(|X_n|, n \in \mathbb{N})$ est une sous martingale positive et on a le résultat pour $|X_n|$:*

$$\forall n \in \mathbb{N},\ \lambda \mathbb{P}\{\max_{k\leq n} |X_k| \geq \lambda\} \leq E[|X_n|\mathbf{1}_{\max_{k\leq n} |X_k|\geq\lambda} \leq E[|X_n|].$$

Théorème 1.19 *Soit* $p > 1$ *et* $(X_n, n \in \mathbb{N})$ *une sous martingale positive dans* L^p. *Alors*

$$\| \max_{k \le n} X_k \|_p \le \frac{p}{p-1} \|X_n\|_p.$$

Corollaire 1.20 *Si* X *est une martingale, alors* $\|X_n^*\|_p \le \frac{p}{p-1} \|X_n\|_p$.

Preuve : on utilise la relation
$E(U^p) = \int \int_{\{0 \le t \le u\}} p t^{p-1} \mathbb{P}(du) dt = p \int_0^\infty t^{p-1} \mathbb{P}(U > t) dt$ (Tonelli et intégration par partie) pour toute variable aléatoire positive de L^p.
On pose $X_n^* = \max_{k \le n} X_k$ qui vérifie donc :

$$E[(X_n^*)^p] = p \int_0^\infty t^{p-1} \mathbb{P}(X_n^* > t) dt$$

et on applique le théorème précédent qui permet la majoration de $E[(X_n^*)^p]$ par

$$p \int_0^\infty t^{p-2} E[X_n \mathbf{1}_{\{X_n^* > t\}}] dt = pE[X_n \int_0^\infty t^{p-2} \mathbf{1}_{\{X_n^* > t\}} dt] = pE[X_n (X_n^*)^{p-1}/(p-1)]$$

que l'on majore par l'inégalité de Hölder qui donne le résultat.

•

1.5 Convergence des martingales

(Neveu page 62, BCD pages 48 et sq.)

Théorème 1.21 *Soit* $(X_n, n \in \mathbb{N})$ *une sous martingale positive bornée dans* L^2. *(c'est à dire* $\forall n,\ \exists C,\ E(X_n^2) \leq C$.*) Alors la suite* $(X_n, n \in \mathbb{N})$ *converge dans* L^2.

Preuve : L'application $x \mapsto x^2$ est convexe croissante sur $\mathbb{R}^+$ donc X_n^2 est encore une sous martingale, qui donc admet la décomposition $X_n^2 = M_n + A_n$. Soit $E(A_n) = E(X_n^2) - E(M_0) \leq C - E(M_0)$. On évalue alors la norme L^2 de $X_n - X_m$ pour $m > n$:

$$E[(X_n - X_m)^2] = E(X_n^2) - 2E(X_n X_m) + E(X_m^2) \leq E(X_n^2) - E(X_m^2)$$

et ceci est exactement $E(A_m) - E(A_n)$ qui est le reste de Cauchy d'une suite convergente (croissante majorée). •

Théorème 1.22 *Soit* $(X_n, n \in \mathbb{N})$ *une sous martingale telle que* $\sup_n E(X_n^+) < \infty$. *Alors la suite* $(X_n, n \in \mathbb{N})$ *converge presque sûrement vers une variable* X_∞ *intégrable qui vérifie de plus* $X_n \leq E[X_\infty / \mathcal{F}_n]$.

Corollaire 1.23 *Une sous martingale négative converge presque sûrement vers* X_∞ *élément de* L^1.

Preuve : (i) On suppose la sous-martingale positive bornée dans L^2.
D'après le théorème précédent, la suite $(X_n, n \in \mathbb{N})$ converge dans L^2 et si $X_n^2 = M_n + A_n$, la suite $(A_n, n \in \mathbb{N})$, suite croissante converge donc presque sûrement, vérifie $E(A_n) = E(X_n^2) - E(M_0) \leq C - E(M_0)$ donc converge dans L^1 et par conséquent, la suite $(M_n, n \in \mathbb{N})$ converge aussi dans L^1. A n fixé, on considère l'inégalité pour la martingale $(M_{n+k} - M_n, k \geq 0)$:

$$\mathbb{P}\{\max_{k \leq m} |M_{n+k} - M_n| \geq \lambda\} \leq \frac{1}{\lambda} E[|M_{n+m} - M_n|]$$

qui tend vers 0 quand n tend vers l'infini et ceci pour tout m (critère de Cauchy pour la convergence L^1 de M).

Par ailleurs, on a l'inclusion

$$\{\max_{0 \leq k,l \leq m} |M_{n+k} - M_{n+l}| \geq 2\lambda\} \subset \{\max_{0 \leq k \leq m} |M_{n+k} - M_n| \geq \lambda\}$$

et ainsi

$$\mathbb{P}\{\max_{0 \leq k,l \leq m} |M_{n+k} - M_{n+l}| \geq 2\lambda\} \leq \mathbb{P}\{\max_{0 \leq k \leq m} |M_{n+k} - M_n| \geq \lambda\} \leq \frac{1}{\lambda} E[|M_{n+m} - M_n|].$$

Comme M converge dans L^1, c'est une suite de Cauchy et $\forall \varepsilon > 0$, il existe n assez grand pour que $E[|M_{n+m} - M_n|] \leq \varepsilon$ pour tout m que l'on peut donc faire tendre vers l'infini :

$$\mathbb{P}\{\max_{k,l\geq 0} |M_{n+k} - M_{n+l}| \geq 2\lambda\} = \mathbb{P}\{\max_{i,j\geq n} |M_i - M_j| \geq 2\lambda\} \leq \frac{1}{\lambda} E[|M_\infty - M_n|] \leq \frac{\varepsilon}{\lambda},$$

ce qui montre que la suite indexée en n décroissante (positive donc convergente presque sûrement) $\max_{i,j\geq n} |M_i - M_j|$ converge en probabilité vers 0 lorsque n tend vers l'infini, donc presque sûrement vers 0.

On avait déjà la convergence presque sûre de A_n vers A_∞ : $(X_n^2, n \in \mathbb{N})$ converge presque sûrement vers X_∞^2 dont on savait déjà que c'est un élément de L^1.

(ii) On suppose que X est une martingale positive alors e^{-X} est une sous martingale positive bornée par 1, donc par (i) e^{-X_n} converge presque sûrement, son log aussi et donc X_n converge presque sûrement. Puis, par le lemme de Fatou

$$E[\liminf_n X_n] \leq \liminf_n E[X_n] \leq \limsup_n E[X_n] \leq E[\limsup_n X_n]$$

et donc puisque presque sûrement $\liminf_n X_n = \limsup_n X_n = X_\infty$, $E[X_\infty] = E[X_0] < \infty$.

(iii) On suppose X martingale bornée dans L^1. On utilise la décomposition de Krickeberg : $X_n = Y_n - Z_n$, Y et Z sont des martingales positives, par (ii) elles convergent presque sûrement vers Y_∞, $Z_\infty \in L^1$: donc X converge presque sûrement vers $X_\infty = Y_\infty - Z_\infty \in L^1$.

(iv) On suppose la sous-martingale bornée dans L^1. On utilise la décomposition de Doob : $X_n = M_n + A_n$, $E(A_n) = E(X_n) - E(M_0) \leq \sup_n E[|X_n|] - E(M_0)$. Donc A_∞ est intégrable et A converge presque sûrement et dans L^1.
Par ailleurs, $|M_n| \leq |X_n| + A_n$ donc M est une martingale bornée dans L^1 donc par (iii) converge presque sûrement vers $M_\infty \in L^1$, d'où le résultat pour X.

(v) Si X est une sous martingale telle que $\sup_n E[X_n^+] < \infty$, $|X| = -X + 2X^+$, $E[|X_n|] = 2E[X_n^+] - E[X_n] \leq 2\sup_n E[X_n^+] - E[X_0]$, et on est ramené à (iv). •

Corollaire 1.24 *Une sous-martingale majorée (sur-martingale minorée) converge presque sûrement vers une variable intégrable.*

Preuve : Il existe $C > 0$ tel que $\forall n,\ X_n \leq C$ donc $X_n^+ \leq C$ et on applique le théorème.

Théorème 1.25 QC3 *Soit M une martingale. Les conditions suivantes sont équivalentes :*

(i) M converge dans L^1.

(ii) La famille $(M_n, n \in \mathbb{N})$ est uniformément intégrable.

(iii) Il existe une variable aléatoire H intégrable telle que $M_n = E[H/\mathcal{F}_n]$.

Une telle martingale est appelée **martingale régulière.**
Preuve : $(i) \Rightarrow (iii)$: si M_n converge vers H dans L^1, alors M_n est bornée dans L^1 et le théorème 1.22 montre que M_n converge presque sûrement vers H.
Puis on considère k et $A \in \mathcal{F}_k$: pour tout $n \geq k$, $E(M_n \mathbf{1}_A) \to_{n\to\infty} E(H\mathbf{1}_A) = E[E(H/\mathcal{F}_k)\mathbf{1}_A]$. Or $E(M_n\mathbf{1}_A) = E(M_k\mathbf{1}_A)$ puisque $M_k = E(M_n/\mathcal{F}_k)$ et donc $E(M_k\mathbf{1}_A) = E[E(H/\mathcal{F}_k)\mathbf{1}_A]$ pour tout $A \in \mathcal{F}_k$ soit $M_k = E(H/\mathcal{F}_k)$.

$(iii) \Rightarrow (ii)$: $\forall n,\ \int_{\{|M_n|\geq a\}} |M_n| d\mathbb{P} \leq \int_{\{|M_n|\geq a\}} |H| d\mathbb{P}$ car $|M_n|$ et une sous martingale et $\{|M_n| \geq a\} \in \mathcal{F}_n$.
Par ailleurs, pour tout n,
$\int_{\{M_n|\geq a\}} |H| d\mathbb{P} \leq E(H\mathbf{1}_{\{|H|\geq\sqrt{a}\}}) + E(H\mathbf{1}_{\{M_n\geq a, |H|<\sqrt{a}\}}) \leq$
$E(H\mathbf{1}_{\{|H|\geq\sqrt{a}\}}) + \sqrt{a}\mathbb{P}\{|M_n| \geq a\} \leq E(H\mathbf{1}_{\{|H|\geq\sqrt{a}\}}) + \frac{1}{\sqrt{a}}E(|M_n|) \to 0$
quand a tend vers l'infini.

$(ii) \Rightarrow (i)$: uniforme intégrable implique borné dans L^1, donc il existe H dans L^1 limite presque sûre de M_n et convergence presque sûre et uniforme intégrable impliquent convergence dans L^1. •
Remarque : une martingale régulière converge presque sûrement et dans L^1 vers H telle que pour tout $n,\ M_n = E(H/\mathcal{F}_n)$.

Théorème 1.26 *Soit M une martingale régulière de la forme $(E(H/\mathcal{F}_n))$ et T un temps d'arrêt. Alors $M_T = E[H/\mathcal{F}_T]$ où, de fait, on peut choisir $H = \lim_n M_n$ au sens L^1 et presque sûr.*

Preuve : Soit $A \in \mathcal{F}_T$, pour tout $n,\ A \cap \{T = n\} \in \mathcal{F}_n$ et $H = M_\infty$.
1. Si la martingale est positive,

$$E(H\mathbf{1}_A) = \sum_n E(H\mathbf{1}_{A\cap\{T=n\}}) + E(H\mathbf{1}_{A\cap\{T=\infty\}}) = \sum_n E(M_n\mathbf{1}_{A\cap\{T=n\}}) + E(H\mathbf{1}_{A\cap\{T=\infty\}}) = E(M_T\mathbf{1}_A).$$

2. Sinon, M est la différence de deux martingales positives $M = X - Y$ et on montre que X qui vient de la décomposition de Krickeberg est régulière : $X_n = N_n + E(A_\infty/\mathcal{F}_n)$ où $N_n + A_n$ est la décomposition de Doob de $|M_n|$. La martingale N est régulière car majorée par $|M|$ qui est uniformément intégrable comme M.
Donc $Y = X - M$ est aussi régulière et le résultat est assuré. •

Corollaire 1.27 *Soit M une martingale régulière et S et T deux temps d'arrêt $S \leq T$, alors $M_S = E[M_T/\mathcal{F}_S]$.*

Preuve exercice immédiat.

Théorème 1.28 *(Neveu page 67-68) Une martingale (ou une sous martingale positive) bornée dans $L^p, p > 1$, converge dans L^p.*

Preuve : a) Si M est une sous-martingale positive, on utilise l'inégalité maximale :

$$\|X_n^*\|_p \leq \frac{p}{p-1}\|X_n\|_p$$

qui est borné par hypothèse. Par ailleurs, avec $p > 1$ X_n^p est une sous martingale, de plus positive et bornée dans L^1 donc elle converge presque sûrement vers $H^p \in L^1$. Donc aussi X_n converge presque sûrement vers $H \in L^p$. Enfin, X_n^* est une suite croissante donc converge presque sûrement, elle est bornée dans L^p, donc la limite aussi :

$$X_n^p \leq (X_n^*)^p \leq (X_\infty^*)^p \in L^1$$

et par le théorème de Lebesgue de convergence dominée, X_n converge vers H dans L^p.

b) Si X est une martingale bornée dans L^p, $|X|$ est une sous martingale positive bornée dans L^p. Par a), il existe $Y \in L^p$ telle que $|X_n|$ converge vers Y presque sûrement et dans L^p. De même X_n^+ est une sous martingale positive bornée dans L^p et il existe $Z \in L^p$ telle que X_n^+ converge vers Z presque sûrement et dans L^p. Globalement, $X_n = 2X_n^+ - |X_n|$ converge vers $2Z - Y$ presque sûrement et dans L^p. •

1.6 Martingales renversées

(Neveu pages 115-119)

Sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$, on considère une suite décroissante de tribus : $(\mathcal{F}_n, n \in \mathbb{N})$. On note $\mathcal{F}_\infty = \cap_n \mathcal{F}_n$.

Définition 1.29 *Une suite de variables aléatoires $(X_n, n \in \mathbb{N})$ sur $(\Omega, \mathcal{A}, \mathbb{P})$ est une* **(sous)martingale renversée** *si pour tout n, X_n est $\mathcal{F}_n$-mesurable, intégrable, et $E(X_n/\mathcal{F}_{n+1})(\geq) = X_{n+1}$.*

Ces processus ont des propriétés analogues à celles des martingales comme par exemple (*à faire en TD ou en DM*) :

une fonction convexe d'une martingale est une sous-martingale,
l'espérance d'une martingale est constante,
l'espérance d'une sous martingale est décroissante.

Montrons qu'il existe également des décompositions.

Théorème 1.30 *Décomposition de Doob : soit X une sous-martingale renversée (positive ou vérifiant $E[X_n]$ minorée) ; il existe une martingale renversée M et un processus décroissant positif A, A_n étant $\mathcal{F}_{n+1}-$mesurable, tels que pour tout entier n, $X_n = M_n + A_n$.*

Preuve : Soit $0 \leq B_n = \sum_{i=0}^{n-1}(E[X_i/\mathcal{F}_{i+1}] - X_{i+1})$ suite croissante dont on note B_∞ la limite presque sûre. On part de $A_0 = B_\infty$ et $M_0 = X_0 - A_0$. Puis on définit les deux suites :

$$A_n = A_0 - B_n = A_{n+1} + E[X_n/\mathcal{F}_{n+1}] - X_{n+1} \geq A_{n+1} \ ; \ M_n = X_n - A_n.$$

Ainsi, par construction la suite A est positive décroissante, A_n est $\mathcal{F}_{n+1}-$mesurable, M est adaptée et de proche en proche on montre que A_n et M_n sont intégrables. car $B\infty = A_0 \in L^1$, $E(B_\infty) = \lim_n E(X_0 - X_n)$.

Enfin, on calcule

$$E[M_n/\mathcal{F}_{n+1}] = E[X_n/\mathcal{F}_{n+1}] - E[A_{n+1} + E[X_n/\mathcal{F}_{n+1}] - X_{n+1}/\mathcal{F}_{n+1}] = -A_{n+1} + X_{n+1}$$

où l'on reconnait la définition de M_{n+1} et $E[M_n/\mathcal{F}_{n+1}] = M_{n+1}$.

Théorème 1.31 *Décomposition de Krickeberg : soit X une martingale renversée, bornée dans L^1 ; il existe deux martingales renversées positives Y et Z telles que $X_n = Y_n - Z_n$ pour tout entier n.*

Preuve : L'hypothèse est qu'il existe une constante M telle que pour tout n, $E(|X_n|] \leq M$. Or, la suite $(|X_n|, n \in \mathbb{N})$ est une sous martingale renversée positive donc par la décomposition de Doob il existe une martingale M et un processus décroissant A, A_n $\mathcal{F}_{n+1}-$mesurable, tels que pour tout entier $n, |X_n| = M_n + A_n$. Or $A_0 = \lim_n B_n$, $E(B_n) = E(X_0) - E(X_n) \leq 2M$ et la limite A_0 est intégrable et on peut donc en prendre l'espérance conditionnelle. On pose :

$$\forall n, M_n + E[A_0/\mathcal{F}_n] = Y_n \text{ et comme } |X_n| = M_n + A_n \leq Y_n,$$

la suite Y est une martingale positive renversée, $Y - X$ aussi puisque $Y_n - X_n \geq Y_n - |X_n| \geq 0$. •

Lemme 1.32 *Une $\mathcal{F}$-martingale renversée est bornée dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$ et c'est une famille uniformément intégrable.*

Preuve : en effet, dans ce cas $|X|$ est une sous martingale renversée donc

$$\forall n, |X_n| \leq E[|X_0|/\mathcal{F}_n].$$

•

Lemme 1.33 *Soit X une sous martingale positive renversée, alors*

$$\forall\lambda > 0,\ \lambda\mathbb{P}\{\sup_{k\geq n} X_k \geq \lambda\} \leq E(X_n).$$

Preuve : On pose $X_n^* = \sup_{k\geq n} X_k$, et $T = \sup\{k : X_k \geq \lambda\},\ +\infty$ si cet ensemble est vide. Les événements $\{T \geq n\}$ et $\{X_n^* \geq \lambda\}$ sont égaux et appartiennent à la tribu $\mathcal{F}_n$ puisqu'ils ne dépendent que du futur.
Par ailleurs,

$$\forall N,\ X_{T\wedge N}\mathbf{1}_{\{N\geq T\geq n\}} = \sum_{k=n}^{N} X_k\mathbf{1}_{\{T=k\}} \leq \sum_{k=n}^{N} E(X_n/\mathcal{F}_k)\mathbf{1}_{\{T=k\}}$$

$$\leq \sum_{k\leq n} E[X_n\mathbf{1}_{\{T=k\}}/\mathcal{F}_k]$$

puisque si $k \geq n$, $X_k \leq E(X_n/\mathcal{F}_k)$ et $\{T = k\} = \{X_k \geq \lambda, X_j < \lambda, \forall j > k\} \in \mathcal{F}_k$. De plus la dernière expression est d'espérance $E[X_n\mathbf{1}_{\{T\geq n\}}$.
Ainsi, la famille $(X_{T\wedge N}\mathbf{1}_{\{N\geq T\geq n\}}, N \geq n)$ est uniformément bornée dans L^1.
Si N tend vers l'infini, $X_{T\wedge N}\mathbf{1}_{\{N\geq T\geq n\}}$ converge presque sûrement vers $X_T\mathbf{1}_{\{T\geq n\}} \geq \lambda$, et le théorème de Lebesgue de convergence dominée montre alors que

$$\lambda\mathbb{P}(T \geq n) \leq E[X_T\mathbf{1}_{\{T\geq n\}}] \leq E[X_n\mathbf{1}_{\{T\geq n\}}]$$

d'où le résultat. •

On peut montrer la convergence suivante.

Théorème 1.34 *Soit $(X_n, n \in \mathbb{N})$ une $\mathcal{F}$-(sous)martingale renversée bornée dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$. Alors la suite X converge presque sûrement et dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$ vers $X_\infty \leq E(X_0/\mathcal{F}_\infty)$ pour une sous-martingale, avec l'égalité pour une martingale.*

Preuve : *à rédiger comme pb 1 ?*
La preuve de la convergence presque sûre est longue et compliquée comme celle du théorème 1.22, on l'admet dans un premier temps et alors l'uniforme intégrabilité montre la convergence L^1.

Pour la convergence presque sûre, reprenons les différentes étapes du théorème 1.22.
(i) X sous-martingale positive renversée bornée dans L^2. On se sert du lemme 1.33 et d ela décomposition de Doob.
(ii) X martingale positive
(iii) X martingale renversée bornée dans L^1.
(iv) X sous-martingale renversée bornée dans L^1.

Exemple d'application : la loi des grands nombres. *Soit une suite Z de variables aléatoires indépendantes intégrables de même loi et $S_n = \frac{Z_1+\cdots+Z_n}{n}$ et $\mathcal{F}_n =$*

$\sigma(S_n, S_{n+1}, ...)$.
1. Montrer d'abord que $\mathcal{F}_n = \sigma(S_n) \vee \sigma(Z_{n+1}, Z_{n+2}, ...)$
2. que S *est une* $\mathcal{F}$*-martingale renversée bornée dans* L^1.
3. En déduire que la suite S *converge presque sûrement et dans* $L^1(\Omega, \mathcal{A}, \mathbb{P})$ *vers* $E(Z_1)$.

1. Pour tout $k \geq n,\ kS_k = nS_n + Z_{n+1} + \cdots + Z_k$, d'où S_k est $\sigma(S_n) \vee \sigma(Z_{n+i}, i \geq 1)-$mesurable et $\mathcal{F}_n \subset \sigma(S_n) \vee \sigma(Z_{n+i}, i \geq 1)$.

Réciproquement, il est clair que $\sigma(S_n) \subset \mathcal{F}_n$ et comme $Z_k = kS_k - (k-1)S_{k-1}, \forall k \geq n+1$, Z_k est $\mathcal{F}_n$ mesurable et on a l'inclusion inverse.
2. S_n somme finie de variables intégrables $\mathcal{F}_n$ mesurables est aussi dans $L^1(\mathcal{F}_n)$.
$E[S_n/\mathcal{F}_{n+1}] = \frac{1}{n}\sum_{k=1}^{n} E[Z_k/\mathcal{F}_{n+1}]$; or $E[Z_k/F_{n+1}] = E[Z_k/\sigma(S_{n+1}) \vee \sigma(Z_{n+i}, i \geq 2)] = E[Z_k/\sigma(S_{n+1})]$ à cause de l'indépendance entre Z_k et $\sigma(Z_{n+i}, i \geq 2)$.
Les Z sont de même loi et S_n est symétrique en Z_i,

$$\forall k = 1, \cdots, n+1, E[Z_k/\sigma(S_{n+1})] = E[Z_1/\sigma(S_{n+1})] = E[S_n/\mathcal{F}_{n+1}].$$

Si l'on fait la moyenne à gauche, il vient :

$$E[S_{n+1}/S_{n+1}] = S_{n+1}$$

d'où l'égalité $S_{n+1} = E[S_n/\mathcal{F}_{n+1}]$.
3. On applique alors le théorème de convergence ci-dessus : (S_n) est bornée dans L^1 car $\forall n, E[|S_n|] \leq E(|Z_1|)$, donc S_n converge presque sûrement et dans L^1 vers S_∞.

$S_\infty = \lim_n E[S_1/\mathcal{F}_n]$ et $\mathcal{F}_n$ décroit : S_∞ est $\mathcal{F}_n$-mesurable pour tout n donc mesurable pour $\cap_n \mathcal{F}_n = \{\emptyset, \Omega\}$; c'est une constante. S_∞ est égale à son espérance égale à $E[S_1] = E[Z_1]$: on retrouve ainsi le résultat classique de la loi des grands nombres.

1.7 Algorithme de Robbins-Monro

Exemple : le dosage.
Il s'agit de déterminer le dosage optimal d'un produit chimique pour obtenir l'effet voulu ; on sait, à cet effet, effectuer des tests.
Au dosage $z \in \mathbb{R}$, on associe l'effet $F(z, \eta)$, où η est une variable aléatoire dont la réalisation dépend du test effectué (pas du dosage). L'effet moyen du dosage z est donné par $f(z) = E[F(z, \eta)]$; on suppose que f est strictement croissante et continue.
Soit $a \in \mathbb{R}$ l'effet désiré. On cherche à déterminer l'unique z^* tel que $f(z^*) = a$. Pour cela on construit une suite de variables aléatoires $(Z_n, n \in \mathbb{N})$ définie par récurrence :

$$Z_{n+1} = Z_n - \gamma_n[F(Z_n, \eta_{n+1}) - a]$$

où γ est une suite de réels positifs convergeant vers 0 et (η_n) est une suite de variables aléatoires indépendantes de même loi que η. On va donner des hypothèses pour que cette suite $(Z_n, n \in \mathbb{N})$ converge vers z^*.

Théorème 1.35 *de ROBBINS-MONRO : Soit* $h \in C(\mathbb{R}, \mathbb{R}), (\gamma_n)$ *une suite de réels positifs et* ε_n *des variables aléatoires de carré intégrable telles que si* $\mathcal{F}_n = \sigma(\varepsilon_1, \cdots, \varepsilon_n)$, $E(\varepsilon_{n+1}/\mathcal{F}_n) = 0$. *On pose*

$$Z_{n+1} = Z_n + \gamma_n[h(Z_n) + \varepsilon_{n+1}], Z_0 \text{ donnée constante,}$$

et on suppose
(i) $\sum_n \gamma_n = +\infty$; $\sum_n \gamma_n^2 < +\infty$,
(ii) $E(\varepsilon_{n+1}^2/\mathcal{F}_n) \leq c(1 + Z_n^2)$,
(iii) $|h(z)|^2 \leq c(1+z^2)$ *et il existe un unique* $z^* \in \mathbb{R} : h(z^*) = 0, (z - z^*)h(z) < 0, \ \forall z \neq z^*$.

Alors, la suite (Z_n) *converge presque sûrement vers* z^*.

Un exemple de $h : z \mapsto a - f(z)$ dans l'exemple di dosage ci-dessus.
Preuve en exercice
Première étape : trouver une majoration de la forme

$$E[(Z_{n+1} - z^*)^2/\mathcal{F}_n] \leq (Z_n - z^*)^2(1 + \alpha\gamma_n^2) + \gamma_n^2\beta.$$

En effet, d'une part les hypothèses montrent par récurrence que Z_n est de carré intégrable pour tout n, et d'autre part :

$$(Z_{n+1} - z^*)^2 = (Z_n - z^*)^2 + 2\gamma_n(Z_n - z^*)(h(Z_n) + \varepsilon_{n+1}) + \gamma_n^2(h(Z_n) + \varepsilon_{n+1})^2$$

que l'on conditionne par $\mathcal{F}_n$:

(1) $E[(Z_{n+1} - z^*)^2/\mathcal{F}_n] = (Z_n - z^*)^2 + 2\gamma_n(Z_n - z^*)h(Z_n) + \gamma_n^2[h(Z_n)^2 + E[\varepsilon_{n+1}^2/\mathcal{F}_n]]$.

Le deuxième terme est négatif par (iii) et le dernier se majore par (ii) et (iii) :

$$E[(Z_{n+1} - z^*)^2/\mathcal{F}_n] \leq (Z_n - z^*)^2 + \gamma_n^2 2c(1 + Z_n^2).$$

Soit une majoration de la forme $(Z_n - z^*)^2(1+\alpha\gamma_n^2)+\gamma_n^2\beta$ avec $\alpha = 4c$ et $\beta = 2c(1+2(z^*)^2)$.

Deuxième étape : on pose $u_n = \Pi_{k=1}^{n-1}(1 + \alpha\gamma_k^2)$, suite convergente puisque la série $\sum_n \gamma_n^2$ converge ; et $v_n = \sum_{k=1}^{n-1} \frac{\beta\gamma_k^2}{u_{k+1}}$ qui converge aussi puisque $u_{k+1} \geq 1$ et la série $\sum_n \gamma_n^2$ converge.

On définit la suite de variables aléatoires $X_n = \frac{(Z_n - z^*)^2}{u_n} - v_n$, soit $(Z_n - z^*)^2 = u_n(X_n + v_n)$. On vérifie que X est une surmartingale : X_n est $\mathcal{F}_n$-mesurable par définition et intégrable puisque Z_n est de carré intégrable ; et

$$E[X_{n+1}/\mathcal{F}_n] = \frac{1}{u_{n+1}}[E[(Z_{n+1}-z^*)^2/\mathcal{F}_n]-v_{n+1} \leq \frac{1}{u_{n+1}}[(Z_n-z^*)^2(1+\alpha\gamma_n^2)+\gamma_n^2\beta]-v_{n+1} = X_n.$$

De plus, elle est minorée par la suite convergente $(-v_n, n \in \mathbb{N})$ donc la suite X converge presque sûrement.

Troisième étape : ainsi, la suite $((Z_n - z^*)^2, n \in \mathbb{N})$ converge vers $(Z_\infty - z^*)^2$. On reprend l'équation (1) en gardant le deuxième terme négatif que l'on passe à gauche et il vient :

$$0 \leq 2\gamma_n(z^* - Z_n)h(Z_n) \leq (Z_n - z^*)^2(1 + \alpha\gamma_n^2) + \gamma_n^2\beta - E[(Z_{n+1} - z^*)^2/\mathcal{F}_n]$$

que l'on réexprime avec les suites (u_n) et (v_n) :

$$0 \leq 2\gamma_n(z^* - Z_n)h(Z_n) \leq u_n(X_n + v_n)\frac{u_{n+1}}{u_n} + \beta\gamma_n^2 - u_{n+1}(E[X_{n+1}/\mathcal{F}_n] + v_{n+1})$$

ce qui permet d'obtenir à droite la majoration $u_{n+1}(X_n - E[X_{n+1}/\mathcal{F}_n])$.
Ces termes sont tous positifs, on montre que la série associée est convergente, car le terme général de la série des espérances est $u_{n+1}(E[X_n] - E[X_{n+1}])$ qui est convergente puisque $E(X_n)$ est une suite décroissante et u_n est une suite convergente.

$$\sum_{n=0}^{N} u_{n+1}(E(X_n) - E(X_{n+1})) = \sum_{n=0}^{N}(u_{n+1} - u_n)E(X_n) + u_0 e(X_0) - u_{n+1}E(X_N).$$

Donc la série de terme général $\gamma_n(z^* - Z_n)h(Z_n)$ est aussi convergente dans L^1 (convergence monotone vers quelque chose d'intégrable). La divergence de la série de terme général γ_n et l'unicité du zéro de h permet alors de conclure.

•

Exercice : On applique ce théorème au dosage avec $h(z) = a - f(z)$,

$\varepsilon_{n+1} = f(Z_n) - F(Z_n, \eta_{n+1})$,
la suite γ_n du théorème et on suppose

(i) $|f(z)|^2 \leq c(1 + |z|^2$

(ii)$E(F(z,\eta)^2) \leq c(1 + |z|^2)$

Vérifier les hypothèses du théorème de R.M. en exercice et conclure.

2 Chaînes de Markov

2.1 Introduction

De nombreux phénomènes aléatoires ont la propriété suivante : la connaissance de l'état du phénomène à un instant donné apporte sur le futur autant d'information que la connaissance de tout le passé. Ce sont les processus de Markov, appelés "chaînes de Markov" si ce sont des processus à temps discret. Par exemple
. gestion de stocks
. prix d'une action au temps t : $P_{t+1} = P_t(1 + r_{t+1})$ où r est le taux au temps t.
. dynamique linéaire : $X_{k+1} = A_k X_k + B_k \varepsilon_{k+1}$.
. plus généralement, $X_{k+1} = F(X_n, \varepsilon_{k+1})$ où ε sont des variabbles aléatoires indépendantes.

On a ici des liens explicites, mais on peut avoir la donnée de la valeur d'une étape par sa loi conditionnelle sachant la valeur à l'étape précédente.

2.2 Exemples

2.2.1 Chaînes de Markov homogènes à espace d'états fini

Ici, on prend pour espace d'état $E = \{1, 2, \cdots, s\}$. La loi de passage d'une étape à l'autre est donnée par la matrice dite "de transition" Q avec
$(Q(i,j) = \mathbb{P}(X_{n+1} = j/X_n = i), 1 \leq i, j \leq s)$ dite aussi **matrice stochastique**, de taille $s \times s$. Ses termes sont positifs et représentent la probabilité que $X_n = j$ lorsque $X_{n-1} = i$. Donc, la somme $\sum_j Q(i,j) = 1$. On obtient également par itération la transition Q_n :

$$Q_n(i,j) = \sum_{i_1} \sum_{i_2} \cdots, \sum_{i_{n-1}} Q(i, i_1) Q(i_1, i_2) \cdots, Q(i_{n-1}, i_n)$$

c'est à dire que la matrice Q_n est simplement la puissance $n$ième de la matrice Q. Elle représente la loi de X_n sachant X_0 : $Q_n(i,j) = \mathbb{P}(X_n = j/X_0 = i)$. On peut ainsi faire une étude algébrique de ces chaînes. Voici quelques cas simples où l'on imagine ce qui se passe.
Exercices. Dans chaque exemple, écrire la matrice Q.

a) $Q(1,1) = \theta, Q(1,2) = 1-\theta, Q(2,1) = 1-\theta', Q(2,2) = \theta'$ où θ et $\theta' \in]0,1[$. Montrer que quelque soit le point de départ, on revient infiniment souvent en 1 et en 2.

$$\begin{aligned}\mathbb{P}(\exists n : X_n = 1/X_1 = 1) &= \sum_{n \geq 1} \mathbb{P}(X_2 = X_3 = \cdots = X_{n-1} = 2, X_n = 1/X_1 = 1) \\ &= \theta + \sum_{n \geq 2} (1-\theta)\theta'^{n-2}(1-\theta') = \theta + \frac{(1-\theta)(1-\theta')}{(1-\theta')} = 1.\end{aligned}$$

On peut terminer le calcul de la loi de X_n lorsque $\theta = \theta'$ en calculant la puissance $n$ème de la matrice Q.

b) $Q(1,1) = 0, Q(2,1) = 0, Q(1,2) = 1, Q(2,2) = 1$, écrire la matrice Q. Quoiqu'il se passe, on va en 2 : c'est un point "absorbant" : que l'on parte de 1 ou de 2, de toutes façons on va en 2 et on y reste....

c) Partant de 1, on va avec proba égale en 2 ou 3 ; et là on y reste.... écrire la matrice Q. Y a-t-il des points absorbants ?

d) On va sûrement de 1 à 2 et réciproquement.... écrire la matrice Q. Y a-t-il des points absorbants ?

2.2.2 Chaînes de Markov à espace d'états E dénombrable

Soit par exemple $E = \mathbb{N}$ et pour $n \in \mathbb{N}, Q(n, n+1) = \theta_n, Q(n, 0) = 1 - \theta_n$, avec $\theta_n \in]0, 1[$. On appelle T_0 le premier instant strictement positif où $X_n = 0$. On l'appelle parfois le "temps de retour".
Exercice : donner la loi du temps d'arrêt T_0 en partant de 0.

$$\mathbb{P}_0(T_0 = 0) = 0; \mathbb{P}_0(T_0 = 1) = 1 - \theta_0;$$

$$\forall n \geq 2, \mathbb{P}_0(T_0 = n) = \mathbb{P}_0(X_1 = 1, X_2 = 2, \cdots, X_{n-1} = n-1, X_n = 0) = \theta_0\theta_1 \cdots \theta_{n-2}(1-\theta_{n-1}).$$

Preuve : le temps d'arrêt T_0 est à valeurs dans $\mathbb{N}^*$.
$\mathbb{P}(T_0 = 1/X_0 = 0) = \mathbb{P}(X_1 = 0/X_0 = 0) = 1 - \theta_0$,
$\mathbb{P}(T_0 = n/X_0 = 0) = \mathbb{P}(X_n = 0, X_1, \cdots, X_{n-1} \neq 0/X_0 = 0) = \theta_0\theta_1 \cdots \theta_{n-2}(1 - \theta_{n-1})$. •

Dans le cas où la suite (θ) est constante, la loi de T_0 est une loi géométrique de paramètre $1 - \theta$ donc de moyenne $\frac{1}{1-\theta}$.
Alors en partant de 0, on y revient sûrement une fois : c'est à dire que T_0 est presque sûrement fini, la probabilité $\mathbb{P}_0\{T_0 < \infty\} = 1$, et la chaîne repassera presque sûrement en 0 un nombre infini de fois.... On dit que 0 est "récurrent".

C'est encore le cas si $\Pi_{n\in\mathbb{N}}\theta_n = 0$, car $\{T_0 = \infty\}$ est l'événement : aller de 0 à 1, de 1 à 2, etc.

$$\mathbb{P}_0(T_0 = \infty) = \Pi_{n\in\mathbb{N}}\theta_n \; ; \; \mathbb{P}_0(T_0 < \infty) = 1 - \Pi_{n\in\mathbb{N}}\theta_n = 1.$$

Mais si $\Pi_{n\in\mathbb{N}}\theta_n > 0$, partant de 0 la probabilité que la chaîne y revienne une fois est < 1 et la chaîne ne repassera presque sûrement en 0 qu'un nombre fini de fois et tend presque sûrement vers l'infini. On dit que 0 est "transient". En effet, sur $\{T_0 = n\}$, on définit la nouvelle chaîne $Y_k = X_{n+k}$ qui part de 0, on définit T_1 le temps de retour de Y en 0, qui vérifie aussi $\mathbb{P}(T_1 < \infty) < 1$ et sur $(T_1 = \infty)$ Y_k tend vers l'infini, donc X aussi.

On recherche alors $\mathbb{P}(X_n \to \infty)$.

$\mathbb{P}(T_1 = \infty) \cap (T_0 < \infty) = \mathbb{P}(T_1 = \infty)/T_0 < \infty)\mathbb{P}(T_0 < \infty) = \Pi_n\theta_k(1 - \Pi_n\theta_k)$
noté $\Pi(1 - \Pi)$.

On fait ensuite une récurrence sur les temps d'arrêt T_i temps du ième passage en 0 :
$\mathbb{P}(X_n \to \infty) = \Pi + \Pi(1 - \Pi) + \Pi(1 - \Pi)^2 + \cdots + \Pi(1 - \Pi)^n + \cdots = 1.$

2.3 Définitions

On donne dans ce paragraphe quelques définitions et on précise les notations.

Définition 2.1 *Soient* $(E, \mathcal{E})$ *et* $(F, \mathcal{F})$ *deux espaces mesurables. On appelle* **probabilité ou noyau de transition** *de* $(E, \mathcal{E})$ *dans* $(F, \mathcal{F})$ *une application* Q *de* $E \times \mathcal{F}$ *dans* $[0, 1]$ *telle que :*

. *pour tout* $B \in \mathcal{F}$, *l'application* $x \mapsto Q(x, B)$ *est* $\mathcal{E}$*-mesurable,*

. *pour tout* $x \in E$, *l'application* $B \mapsto Q(x, B)$ *est une probabilité sur* $(F, \mathcal{F})$.

Dans le cas où $E = F$ dénombrable, $\mathcal{E}$ est l'ensemble des parties de E et Q est définie sur tous les singletons avec $Q(x, \{y\}) = \mathbb{P}(X_{n+1} = y/X_n = x)$.
Notations : si Q est une probabilité de transition de $(E, \mathcal{E})$ dans $(F, \mathcal{F})$ et f une fonction mesurable positive (ou négative, ou bornée) sur $(F, \mathcal{F})$, on pose

$$Qf(x) = \int_F f(y) Q(x, dy).$$

Exercice : montrer que la fonction Qf *est mesurable sur* $(E, \mathcal{E})$.
Si $f = \mathbf{1}_B$, $(Q\mathbf{1}_B)(.) = Q(., B)$ est mesurable par définition,
donc c'est vrai pour toute fonction étagée,
l'ensemble de fonctions f qui vérifie la propriété est un espace vectoriel stable par limite croissante et contenant les indicatrices,
le théorème de classe monotone permet de conclure.

Exercices :
(i) μ *est une mesure sur* $(E, \mathcal{E})$, *on note*

$$\forall B \in \mathcal{F},\ \mu Q(B) = \int_F Q(x, B) d\mu(x).$$

Montrer que μQ *est une mesure sur* $(F, \mathcal{F})$.

(ii) Si Q *est une probabilité de transition de* $(E, \mathcal{E})$ *dans* $(F, \mathcal{F})$ *et si* R *est une probabilité de transition de* $(F, \mathcal{F})$ *dans* $(G, \mathcal{G})$, *on note pour tout* $C \in \mathcal{F} \otimes \mathcal{G}$ *:*

$$Q \otimes R(x, C) = \int_F Q(x, dy) \int_G R(y, dz) \mathbf{1}_C(y, z).$$

Montrer que $Q \otimes R$ *est une probabilité de transition de* $(E, \mathcal{E})$ *dans* $(F \times G, \mathcal{F} \otimes \mathcal{G})$.

- Soit $C \in \mathcal{F} \otimes \mathcal{G}$: $y \mapsto f(y) = \int_G R(y,dz)\mathbf{1}_C(y,z)$ est $\mathcal{F}$-mesurable positive, alors $Q \otimes R(x,C) = (Qf)(x)$ est $\mathcal{E}$ mesurable d'après le premier exercice.

- Soit x fixé dans E. $Q \otimes R(x, F \times G) = \int_F Q(x,dy) \int_G R(y,dz) = 1$.
Si $C_1 \cap C_2 = \emptyset$ dans $\mathcal{F} \otimes \mathcal{G}$, $\mathbf{1}_{C_1 \cup C_2} = \mathbf{1}_{C_1} + \mathbf{1}_{C_2}$ et les deux mesures successives sont additives.
Soit C_n suite croissante d'événements de $\mathcal{F}\otimes\mathcal{G}$, $\lim_n \mathbf{1}_{C_n} = \mathbf{1}_C, \forall y \in F,\ \int_G R(y,dz)\mathbf{1}_{C_n}(y,z)$ croit vers $\int_G R(y,dz)\mathbf{1}_C(y,z)$ où $C_n(y,.) = \{z \in G, (y,z) \in C_n\}$ par définition de la tribu produit.

2.4 Transitions et lois conditionnelles

Soient deux variables aléatoires X et Y sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, à valeurs respectivement dans $(E, \mathcal{E})$ et $(F, \mathcal{F})$. On note μ la loi de X et Q une probabilité de transition de $(E, \mathcal{E})$ dans $(F, \mathcal{F})$ telle que pour tout $B \in \mathcal{F}, Q(X, B)$ est une version de $\mathbb{P}\{Y \in B/X\}$. Ainsi, Q est une version de la loi conditionnelle de Y sachant X. (cf. cours de F. Barthe).
On a alors pour f fonction mesurable sur $(F, \mathcal{F})$:

$$E[f(Y)/X] = Qf(X).$$

Cette dernière égalité montre que la loi de Y est μQ. En effet, $E[f(Y)/X = x] = Qf(x)$ que l'on intègre sur E selon la loi μ soit $E[f(Y)] = \int_E Qf(x)\mu(dx)$ c'est à dire $E[f(Y)] = \mu Q(f)$.

En conclusion, un noyau de transition permet de définir une loi conditionnelle.

Bien entendu, si X et Y sont indépendantes, $Q(x, B)$ ne dépend pas de x et vaut $\mathbb{P}(Y \in B)$.

2.5 Chaînes de Markov

Définition 2.2 *On appelle* **chaînes de Markov** *sur l'espace de probabilité* $(\Omega, \mathcal{F}, \mathbb{P})$, *un processus* $(X_n, n \in \mathbb{N})$ *à valeurs dans* $(E, \mathcal{E})$ *tel que pour tout* $B \in \mathcal{F}$, *et tout* $n \in \mathbb{N}$,

$$\mathbb{P}(X_{n+1} \in B/X_0, \cdots, X_n) = \mathbb{P}(X_{n+1} \in B/X_n).$$

Si E est fini ou dénombrable,

$$\mathbb{P}\{X_{n+1} = j/X_0, \cdots, X_n\} = Q(X_n, j)\ ;\ \mathbb{P}\{X_{n+1} = j/X_0 = x_0, \cdots, X_n = x_n\} = Q(x_n, j)$$

avec Q la matrice de transition.

Définition 2.3 *Soit Q une probabilité de transition de $(E, \mathcal{E})$ dans lui-même et ν une probabilité sur $(E, \mathcal{E})$. Une chaîne de Markov $(X_n, n \in \mathbb{N})$ est* **homogène** *de* **loi initiale** *ν et de* **transition** *Q si*

$$\forall B \in \mathcal{E},\ \mathbb{P}(X_0 \in B) = \nu(B)\ ;\ \forall n \in \mathbb{N}, \mathbb{P}(X_{n+1} \in B / X_0, \cdots, X_n) = Q(X_n, B).$$

Une chaîne **non homogène** est telle que $\mathbb{P}(X_{n+1} \in B / X_0, \cdots, X_n) = Q_n(X_n, B)$.

A partir des fonctions indicatrices de pavés, on démontre aisément le résultat suivant, par conditionnements successifs ou par récurrence sur p.

Proposition 2.4 QC 4
Soit une chaîne de Markov $(X_n, n \in \mathbb{N})$ de transition $Q, p \geq 1$ et f une fonction mesurable positive ou bornée sur $(E, \mathcal{E})^p$. alors

$$(*) \qquad E[f(X_{n+1}, \cdots, X_{n+p}) / X_0, \cdots, X_n] = E[f(X_{n+1}, \cdots, X_{n+p}) / X_n].$$

De plus cette dernière expression se récrit :

$$\int_{E^p} f(x_1, \cdots, x_p) Q(x_{p-1}, dx_p) Q(x_{p-2}, dx_{p-1}) ... Q(X_n, dx_1) =$$

$$= \int_{E^p} f(x_1, \cdots, x_p) Q^{\otimes p}(X_n; dx_1, \cdots, dx_p).$$

Preuve : La loi du p-uplet $(X_{n+1}, \cdots, X_{n+p})$ sachant le passé $\mathcal{F}_n$ est le produit successif des lois conditionnelles (théorème de Bayes)

$$\text{loi de } X_{n+p}/\mathcal{F}_{n+p-1} \otimes \cdots \otimes \text{loi de } X_{n+2}/\mathcal{F}_{n+1} \otimes \text{loi de } X_{n+1}/\mathcal{F}_n$$

On utilise la propriété de Markov et la loi du p-uplet $(X_{n+1}, \cdots, X_{n+p})$ sachant le passé $\mathcal{F}_n$ est le produit successif des lois conditionnelles

$$\text{loi de } X_{n+p}/X_{n+p-1} \otimes \cdots \text{loi de } X_{n+2}/X_{n+1} \otimes \text{loi de } X_{n+1}/X_n$$

D'où il vient :

$$E[f(X_{n+1}, \cdots, X_{n+p}) / X_0, \cdots, X_n) = \int_E (\int_E \cdots (\int_E f(x_1, \cdots, x_p) Q(x_{p-1}, dx_p)) \cdots) Q(x_1, dx_2)) Q(X_n, dx_1)$$

qui est en effet une fonction uniquement de X_n et pas de tout le passé. On la note $\Phi(X_n)$. La dernière écriture de cette fonction est conforme à la définition par récurrence du produit tensoriel des probabilités de transition.

Pour terminer la démonstration, on vérifie que $\Phi(X_n) \in L^1(\sigma(X_n))$ puis on considère g $\mathcal{E}$-mesurable bornée et on vérifie légalité de l'espérance des deux produits avec g borélienne bornée :

$$f(X_{n+1}, \cdots, X_{n+p}) g(X_n) \text{ et } \Phi(X_n) g(X_n)$$

en utilisant à gauche la loi du $p+1$-uplet $(X_n, \cdots, X_{n+p})$ et la loi de X_n à droite. •

Ainsi la loi conditionnelle de X_{n+1} sachant tout le passé est $Q(X_n, .)$ et plus généralement la loi conditionnelle de $(X_{n+1}, \cdots, X_{n+p})$ sachant tout le passé est $Q^{\otimes p}(X_n, .)$: c'est aussi sa loi conditionnelle sachant X_n.
Pour avoir maintenant la **loi conditionnelle de X_{n+p} sachant** $X_0, \cdots, X_n$, on applique le résultat précédent à une fonction uniquement de X_{n+p} :

$$E[f(X_{n+p})/X_0, \cdots, X_n)] = \int_{E^p} f(x_p)Q(x_{p-1}, dx_p) \cdots Q(x_1, dx_2)Q(X_n, dx_1)$$

que l'on note $Q^{\otimes p} f(X_n)$, c'est à dire que $Q^{\otimes p}$ est défini par

$$Q^{\otimes p}(x, B) = \int_E Q(x, dx_1) \int_E Q(x_1, dx_2) ... \int_B Q(x_{p-1}, dx_p).$$

Proposition 2.5 *Soit une chaîne de Markov $(X_n, n \in \mathbb{N})$ homogène de loi initiale ν et de transition Q. La loi de $(X_0, \cdots, X_p)$ est donnée $\forall p \geq 0$ par*

$$E[f(X_0, \cdots, X_p)] = \int_E \nu(dx_0) \int_E Q(x_0, dx_1) ... \int_E Q(x_{p-1}, dx_p) f(x_0, x_1, \cdots, x_p) =$$

$$= \int_{E^p} \nu(dx_0) Q^{\otimes p}(x_0; dx_1, \cdots, dx_p) f(x_0, x_1, \cdots, x_p).$$

Ceci s'écrit facilement dans le cas dénombrable :

$$E[f(X_0, \cdots, X_p)] = \sum_{x_i \in E, i=0,\cdots,p} f(x_0, \cdots, x_p) Q(x_{p-1}, x_p) \cdots Q(x_0, x_1) \nu(x_0).$$

Preuve : on applique la proposition 2.4 avec $n = 0$; on obtient

$$E[f(X_0, \cdots, X_p)/\mathcal{F}_0] = \int_E Q(X_0, dx_1) \int_E Q(x_1, dx_2) ... \int_E Q(x_{p-1}, dx_p) f(X_0, x_1, \cdots, x_p)$$

qui est une fonction de X_0 variable aléatoire de loi ν ; puis on intègre sur E en ν :

$$E[f(X_0, \cdots, X_p)] = \int_E \nu(dx_0) Q^{\otimes p}(x_0, dx_1 \cdots dx_p) f(x_0, x_1, \cdots, x_p).$$

•

Corollaire 2.6 *Lorsque f ne dépend que d'une variable, on obtient par intégration sous la loi ν*

$$E[f(X_p)] = \int_{E^p} f(x_p) Q^{\otimes p}(x_0, dx_1 \cdots dx_p) \nu(dx_0).$$

2.6 Chaine de Markov canonique

(cf. Mazliak-Priouret-Baldi pages 90-91).
Inversement, si l'on se donne l'espace d'états $(E, \mathcal{E})$, la probabilité ν et la probabilité de transition Q de $(E, \mathcal{E})$ dans lui-même, on peut se demander s'il existe toujours un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et un processus X qui serait une chaîne de Markov homogène de loi initiale ν et de transition Q.

Une réponse positive est apportée avec l'espace $\Omega = E^{\mathbb{N}}$, $X : (\omega, n) \mapsto \omega_n$, c'est à dire la $n$ième coordonnée, et $\mathcal{A}$ la plus petite tribu rendant mesurables tous les X_n, appelée la tribu cylindrique, $\mathcal{F}_n$ comme d'habitude est $\sigma(X_k, k \leq n)$ et $\mathcal{F}_\infty = \vee_n \mathcal{F}_n = \mathcal{A}$. Enfin, $\mathbb{P}_\nu$ est la probabilité définie par la propriété de la proposition 2.5. On admet le résultat :

Il existe une unique probabilité $\mathbb{P}_\nu$ ayant cette propriété ; elle est définie sur les cylindres par :

$$\mathbb{P}_\nu\{(X_0, \cdots, X_p) \in B_0 \times \cdots \times B_p\} = \int_{B_0} \nu(dx_0) \int_{B_1} Q(x_0, dx_1) ... \int_{B_p} Q(x_{p-1}, dx_p).$$

De fait, c'est une conséquence du théorème de Kolmogorov. On appelle $\mathcal{I}$ la classe des ensembles de la forme

$$\{\omega \in \Omega : (X_0, \cdots, X_n)(\omega) \in B_0 \times \cdots \times B_p\}.$$

Cette famille d'ensemble contient Ω, est stable par intersection finie et complémentaire, et par définition engendre la tribu $\mathcal{A}$, cf.le cours d'intégration de licence (théorème de Carathéodory).
Alors, on vérifie que, en effet, cette suite ainsi construite de $(\Omega, \mathcal{A}, \mathbb{P}_\nu)$ est bien une chaine de Markov de loi initiale ν et de noyau de transition Q.
C'est à dire qu'il faut vérifier que, pour tout n, la loi conditionnelle de X_{n+1} sachant $\mathcal{F}_n$ est la loi conditionnelle de X_{n+1} sachant X_n, et précisément $Q(X_n, dx)$.

Cet exemple de constuction est connu comme la **version canonique** de la chaîne de Markov homogène de loi initiale ν et de transition Q. L'avantage de cette version canonique est que les trajectoires sont intrinsèques : elles ne dépendent ni de ν ni de Q. Cela facilite l'étude simultanée de familles de processus correspondant à des familles de couples (ν, Q).

La plupart du temps, Q sera fixée, mais on fera varier ν, d'où la notation $\mathbb{P}_\nu$ pour indiquer la dépendance en ν. Lorsque ν est la mesure de Dirac au point $x \in E, \mathbb{P}_\nu$ se note $\mathbb{P}_x$. Remarquer que $\mathbb{P}_x(A)$ est aussi $\mathbb{P}_\nu(A/X_0 = x)$.

On utilisera systématiquement la filtration engendrée par le processus X notée $(\mathcal{F}_n, n \in \mathbb{N})$.

Lemme 2.7 *Pour tout $x \in E, B \in \mathcal{E}, n \geq 1$,*

$$\mathbb{P}_x\{X_n \in B\} = Q^{\otimes n}(x, E^{n-1} \times B).$$

Preuve : Comme dans la preuve de la proposition 2.5, on montre avec $f = \mathbf{1}_B$ et $p = 0$,

$$\mathbb{P}_x(X_n \in B) = \mathbb{P}_\nu(X_n \in B/X_0 = x) =$$

$$\int Q(x, dx_1) \int Q(x_1, dx_2) \cdots \int Q(x_{n-1}, dx_n) \mathbf{1}_B(x_n) = Q^{\otimes n}(x, E^{n-1} \times B).$$

•

Proposition 2.8 *Soit ν une loi initiale et $B \in \mathcal{A}$. Alors, $\mathbb{P}_\nu(B) = \int_E \mathbb{P}_x(B)\nu(dx)$. De même, pour toute Z $\mathcal{A}$-mesurable et positive ou $\mathbb{P}_\nu$-intégrable, $E_\nu(Z) = \int_E E_x(Z)\nu(dx)$.*

Preuve : en exercice
Il suffit de le montrer pour les événements $B \in \mathcal{I}$. Comme dans la preuve de la proposition 2.5, avec $f = \mathbf{1}_{\Pi_i B_i}$, on obtient

$$\mathbb{P}_x(B) = E[f(X_0, \cdots, X_n)/X_0 = x] = \mathbf{1}_{B_0}(x) Q^{\otimes n}(x, \Pi_{i=1}^n B_i),$$

que l'on intègre sur E par rapport à la mesure ν d'où $\mathbb{P}_\nu(B) = \int_E \mathbb{P}_x(B)\nu(dx)$.
On passe aux événements de la tribu cylindrique, puis à Z $\mathcal{A}$-mesurable et positive ou bornée par le théorème de la classe monotone :

. pour les événements, soit $B \in \mathcal{I} \subset \mathcal{A}$ donc de la forme

$$B = \{\omega : (X_0, \cdots, X_n)(\omega) \in \Pi_0^n B_i,\ B_i \in \mathcal{E}.$$

Soit $\mathcal{F}$ l'ensemble des $B \in \mathcal{A}$ tels que $\mathbb{P}_\nu(B) = \int_E \mathbb{P}_x(B)\nu(dx)$. On a l'inclusion $\mathcal{I} \subset \mathcal{F}$, cette classe $\mathcal{F}$ est stable par réunion croissante et par différence, $\Omega \in \mathcal{F}$, donc $\mathcal{F}$ contient $\sigma(\mathcal{I}) = \mathcal{A}$.

. pour les variables $\mathcal{A}$-mesurables bornées, soit $\mathcal{F}$ l'ensemble de celles telles que $E_\nu(Z) = \int_E E_x(Z)\nu(dx)$. Cet ensemble contient les indicatrices des événements de $\mathcal{A}$ d'après l'alinea précédent ; c'est un espace vectoriel stable par limite croissante, donc il contient toutes les variables $\mathcal{A}$-mesurables bornées.
Toute variable positive $Z = \lim_n Z \wedge n$, la propriété est vraie $\forall n$ pour $Z \wedge n$, et par limite monotone la propriété est vraie pour Z.
Enfin, on montre la propriété pour Z intégrable en décomposant Z en $Z^+ - Z^-$. •

2.7 Propriété de Markov

Soit donc $\Omega = E^{\mathbb{N}}$ sur lequel on définit l'opérateur de translation $\theta : \Omega \to \Omega, \omega = (\omega_n) \mapsto (\omega_{n+1})$. On a donc $X_n \circ \theta = X_{n+1}, \forall n \geq 0$. Si l'on compose p fois l'opérateur θ noté θ_p il vient :

$$X_n \circ \theta_p = X_{n+p}.$$

Si T est une variable aléatoire à valeurs entières, on définit de la même manière θ_T par :

$$X_n \circ \theta_T(\omega) = X_{n+T(\omega)}(\omega).$$

C'est à dire que si $\omega = (X_0(\omega), \cdots, X_n(\omega), ...)$ alors $\theta_T(\omega) = X_{T(\omega)}(\omega), X_{1+T(\omega)}(\omega).....$
A la filtration $(\mathcal{F}_n, n \in \mathbb{N})$, on peut ajouter la tribu $\mathcal{F}_\infty$, plus petite tribu contenant toutes les tribus $\mathcal{F}_n, n \in \mathbb{N}$. Si T est un $\mathcal{F}$-temps d'arrêt, on rappelle la notation

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty \ ; \ A \cap (T = n) \in \mathcal{F}_n, \forall n \in \mathbb{N}\},$$

vue dans le chapitre précédent : il s'agit là de la tribu des événements antérieurs à T et la variable aléatoire

(2) $$\mathbf{1}_{\{T<\infty\}} g(X_T) \text{ s'écrit } \sum_{n\geq 0} g(X_n)\mathbf{1}_{\{T=n\}},$$

elle est $\mathcal{F}_T$-mesurable pour toute fonction réelle g mesurable sur $(E, \mathcal{E})$.

On donne maintenant le théorème crucial de ce chapitre.

Théorème 2.9 Propriété de Markov forte : *Soit Y une variable aléatoire bornée ou positive sur $(\Omega, \mathcal{A})$ et une loi initiale ν.*
(i) Pour tout n, on a $\mathbb{P}_\nu$ presque sûrement

$$E_\nu[Y \circ \theta_n / \mathcal{F}_n] = E_{X_n}[Y].$$

(ii) Soit T un $\mathcal{F}$-temps d'arrêt ; on a $\mathbb{P}_\nu$ presque sûrement

$$E_\nu[Y \circ \theta_T \mathbf{1}_{\{T<\infty\}} / \mathcal{F}_T] = \mathbf{1}_{\{T<\infty\}} E_{X_T}[Y].$$

Preuve :

a) Si l'on pose $g(x) = E_x(Y)$, alors g est une fonction mesurable bornée (ou positive). En effet, si $Y = \mathbf{1}_B, \ B \in \mathcal{I}$ soit

$$Y = \mathbf{1}_{B_0}(X_0)\mathbf{1}_{B_1}(X_1) \cdots \mathbf{1}_{B_p}(X_p), B_i \in \mathcal{E},$$

alors par la proposition 2.4 :

$$E_x(Y) = \int_{B_0} \delta_x(x_0) \int_{B_1} Q(x_0, dx_1) \cdots \int_{B_p} Q(x_{p-1}, dx_p) = \mathbf{1}_{B_0}(x) Q^{\otimes p}(x, B_1 \times \cdots \times B_p)$$

est bien mesurable, et il en est de même par le théorème de classe monotone si Y est la fonction indicatrice d'un élément de la tribu $\mathcal{F}_\infty = \mathcal{A} = \sigma(\mathcal{I})$.
C'est donc encore vrai pour toute variable aléatoire $\mathcal{A}$-mesurable bornée (ou positive).

b) La candidate à être l'espérance conditionnelle, $g(X_n) = E_{X_n}(Y)$, est donc une variable aléatoire $\mathcal{F}_n$-mesurable comme fonction mesurable de X_n qui est $\mathcal{F}_n$-mesurable. Ensuite, considérons d'abord $Y = F(X_0, \cdots, X_p)$ avec F mesurable positive ou bornée sur E^p alors $Y \circ \theta_n = F(X_n, \cdots, X_{n+p})$. Dans ce cas, $g(x) = \int_{E^p} F(x, x_1, \cdots, x_p) Q^{\otimes p}(x, dx_1, \cdots, dx_p)$.
Soit Z $\mathcal{F}_n$-mesurable bornée, soit $Z = h(X_0, \cdots, X_n)$. On calcule

$$E_\nu[ZY] = \int_{E^{n+p+1}} \nu(dx_0) h(x_0, \cdots, x_n) F(x_n, \cdots, x_{n+p}) Q^{\otimes n}(x_0, dx_1, \cdots, dx_n) Q^{\otimes p}(x_n, dx_{n+1}, \cdots, dx_{n+p}) =$$

$$\int_E \nu(dx_0)\int_{E^n} h(x_0,\cdots,x_n)g(x_n)Q^{\otimes n}(x_0,dx_1,\cdots,dx_n) = E_\nu[Zg(X_n)]$$

ce qui montre bien que $E_\nu[Y\circ\theta_n/\mathcal{F}_n] = g(X_n) = E_\nu[F(X_n,\cdots,X_{n+p})/X_n]$, d'où le (i) est vrai pour les indicatrices d'événements de $\mathcal{I}$, classe telle que $\sigma(\mathcal{I}) = \mathcal{A}$.
L'espace vectoriel des variables Y qui vérifient (i) est stable par limite croissante et contient les indicatrices de $\mathcal{I}$, par le théorème des classes monotones, cet espace contient toutes les variables $\mathcal{A}$-mesurables positives ou bornées, ce qui montre (i).

c) De même si T est un $\mathcal{F}$-temps d'arrêt, $g(X_T)\mathbf{1}_{\{T<\infty\}}$ est une variable aléatoire $\mathcal{F}_T$-mesurable comme fonction mesurable de X_T qui est une variable aléatoire $\mathcal{F}_T$-mesurable (cf. (2)).
Il reste à montrer que pour toute loi initiale ν, elle est un représentant de l'espérance conditionnelle de $Y\circ\theta_T\mathbf{1}_{\{T<\infty\}}$ sachant la tribu $\mathcal{F}_T$.
Soit $A\in\mathcal{F}_T$, on veut montrer

$$E_\nu[Y\circ\theta_T\mathbf{1}_A] = E_\nu[E_{X_T}(Y)\mathbf{1}_A].$$

On utilise (i) pour montrer :

$$\int_A Y\circ\theta_T\mathbf{1}_{\{T<\infty\}}d\mathbb{P}_\nu = \sum_{n=0}^\infty\int_{A\cap\{T=n\}} Y\circ\theta_n d\mathbb{P}_\nu = \sum_{n=0}^\infty E_\nu[E_\nu[Y\circ\theta_n/\mathcal{F}_n]\mathbf{1}_{\{T=n\}\cap A}] =$$

$$= \sum_{n=0}^\infty E_\nu[E_{X_n}[Y]\mathbf{1}_{\{T=n\}\cap A}] = \sum_{n=0}^\infty\int_{A\cap\{T=n\}} g(X_n)d\mathbb{P}_\nu = \int_A g(X_T)\mathbf{1}_{\{T<\infty\}}d\mathbb{P}_\nu,$$

c'est à dire que

$$E_\nu[Y\circ\theta_T\mathbf{1}_{\{T<\infty\}}/\mathcal{F}_T] = \mathbf{1}_{\{T<\infty\}}E_{X_T}(Y).$$

•

2.8 Fonctions excessives et invariantes

Définition 2.10 *Une fonction réelle h mesurable positive sur l'espace d'états $(E,\mathcal{E})$ est dite* **invariante** *pour Q si $Qh = h$ et* **excessive** *pour Q si $Qh\le h$.*

Remarquons en préalable que pour toute fonction positive ou bornée $E_\nu[h(X_{n+1})/\mathcal{F}_n] = Qh(X_n)$. En effet en utilisant la définition et la question de cours 4 avec $p = 1$:

$$(3)\qquad E_\nu[h(X_{n+1})/\mathcal{F}_n] = E_\nu[h(X_{n+1})/X_n] = \int_E h(x)Q(X_n,dx) = Qh(X_n).$$

Exercice : si h est invariante, $h(X_n)$ est une $(\Omega,\mathcal{A},\mathcal{F},\mathbb{P}_\nu)$-martingale quelle que soit la loi initiale ν telle que h soit dans $L^1(E,\mathcal{E},\nu)$.
. Par définition de la filtration, $h(X_n)$ est $\mathcal{F}_n-$mesurable,

. Puisque $h \in L^1(E, \mathcal{E}, \nu), E_\nu[h(X_0)] = \int h(x)\nu(dx) < \infty$. On suppose par récurrence que $h(X_n)$ est aussi dans $L^1(\Omega, \mathcal{A}, \mathbb{P}_\nu)$, alors, comme h est positive $E_\nu[h(X_{n+1})/\mathcal{F}_n] \in \overline{\mathbb{R}}^+$, et puisqu'elle est invariante, utilisant (3) il vient :

$$E_\nu[h(X_{n+1})/\mathcal{F}_n] = Qh(X_n) = h(X_n)$$

qui est intégrable donc aussi $h(X_{n+1})$ l'est aussi.
. La relation ci-dessus montre enfin que l'on a bien une martingale.
. Enfin, c'est une martingale positive, d'après le corollaire 1.24, elle converge $\mathbb{P}_\nu$ presque sûrement vers une variable intégrable. •

Exercice : si h est excessive, $h(X_n)$ est une $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P}_\nu)$-surmartingale positive quelle que soit la loi initiale ν telle que h soit dans $L^1(E, \mathcal{E}, \nu)$. De plus, elle converge donc $\mathbb{P}_\nu$-presque sûrement quand n tend vers l'infini.
Par définition de la filtration et la mesurabilité de h, $h(X_n)$ est $\mathcal{F}_n$-mesurable.
Comme ci-dessus, $h(X_0) \in L^1(\Omega, \mathcal{A}, \mathbb{P}_\nu)$. On suppose par récurrence que $h(X_n) \in L^1(\Omega, \mathcal{A}, \mathbb{P}_\nu)$.
Puisque h est excessive et positive,

$$0 \leq E_\nu[h(X_{n+1})/\mathcal{F}_n] = Qh(X_n) \leq h(X_n)$$

donc $E_\nu[h(X_{n+1})/\mathcal{F}_n]$ est intégrable, $h(X_{n+1})$ aussi, et il s'agit d'une $\mathcal{F}$-surmartingale.
Enfin, c'est une sur-martingale positive, donc minorée ; d'après le corollaire 1.24, elle converge presque sûrement vers une variable intégrable. •

Définition 2.11 *Soit $B \in \mathcal{E}$, on pose $T_B = \inf\{n > 0 : X_n \in B\}$, ($+\infty$ si l'ensemble est vide) et*

$$R_B = \overline{lim}\{X_n \in B\} = \{\omega : \sum_{n \in \mathbb{N}} \mathbf{1}_{\{X_n \in B\}}(\omega) = +\infty\}.$$

On appelle T_B le temps d'entrée de la chaîne dans B, et R_B l'ensemble de récurrence en B.

Il est clair que T_B est un $\mathcal{F}$-temps d'arrêt et que $R_B \in \mathcal{A} = \mathcal{F}_\infty$. Pour $p \in \mathbb{N}$ sur l'événement $\{T_B > p\}$, $T_B - p = \inf\{n > 0 : X_{n+p} \in B\}$, soit
$T_B = p + \inf\{n > 0 : X_{n+p} \in B\} = p + \inf\{n > 0 : X_n \circ \theta_p \in B\} = p + T_B \circ \theta_p$.
Donc, sur $\{T_B > p\}$, on a $T_B = p + T_B \circ \theta_p$.

Ceci est encore vrai pour tout temps d'arrêt quelconque T : $T_B = T + T_B \circ \theta_T$ sur $\{T_B > T\}$. En effet,

$$\{T_B > T\} = \cup_p(\{T_B > p\} \cap \{T = p\}) \cup \{T_B = \infty\} \cap \{T = \infty\}.$$

Sur $\{T_B > p\} \cap \{T = p\}, T_B = p + T_B \circ \theta_p = T + T_B \circ \theta_T$, et sur $\{T = \infty\}, \infty = \infty$...

Proposition 2.12 *La fonction $x \mapsto \mathbb{P}_x(T_B < \infty)$ est excessive pour Q, et la fonction $x \mapsto \mathbb{P}_x(R_B)$ est invariante pour Q.*

Preuve : Rappel : Pour toute f positive ou bornée $Qf(x) = E_x[f(X_1)]$ et pour tout $A \in \mathcal{E}$, $\mathbb{P}_x(A) = Q(x, A)$.

(i) Posons $f(x) = \mathbb{P}_x(T_B < \infty)$. C'est une fonction mesurable et positive. On calcule alors en utilisant la propriété de Markov forte :

$$Qf(x) = E_x[\mathbb{P}_{X_1}(T_B < \infty)] = E_x[E_{X_1}(\mathbf{1}_{T_B<\infty})] =$$

$$E_x[E_x(\mathbf{1}_{\{T_B<\infty\}} \circ \theta_1/\mathcal{F}_1)] = E_x[\mathbf{1}_{\{T_B\circ\theta_1<\infty\}}],$$

c'est à dire la probabilité partant de x qu'il existe $n > 1$ tel que $X_n \in B$, ce qui est plus petit que $f(x)$, et f est bien excessive.

(ii) Posons $g(x) = \mathbb{P}_x(R_B) = Q(x, R_B)$ (encore mesurable et positive). On opère le même type de calculs :

$$Qg(x) = E_x[\mathbb{P}_{X_1}(R_B)] = E_x[E_x(\mathbf{1}_{R_B} \circ \theta_1)/\mathcal{F}_1)] =$$

$$E_x[\mathbf{1}_{R_B} \circ \theta_1] = \mathbb{P}_x(\sum_{n\geq 1} \mathbf{1}_{\{X_n\in B\}} = +\infty) = \mathbb{P}_x(\sum_{n\geq 0} \mathbf{1}_{\{X_n\in B\}} = +\infty) = g(x).$$

•

Remarquons que dans R_B, la chaîne X passe infiment souvent dans B, c'est pour cela que l'on parle de récurrence. Donc, alors $T_B < \infty$ et on a l'inclusion (cf. BMD page 95 et sq.)
$R_B \subset \{T_B < \infty\}$.

2.9 Etats récurrents ou transitoires

On suppose dans ce paragraphe que E est dénombrable muni de la tribu $\mathcal{E}$ engendrée par les singletons et on note

$$Q(y, x) = Q(y, \{x\})$$

comme on l'a fait dans l'exemple où E était un espace fini.

2.9.1 Passages successifs en un point

(cf. temps de retour, BCD 5.4 page 129)
Soit $B \in \mathcal{E}$ et T_B^p le p-ième passage de la chaîne dans B, c'est à dire :

$$T_B^0 = 0 \; ; \; T_B^1 = \inf\{n > 0 : X_n \in B\} = T_B \; ; \; \cdots ; T_B^p = \inf\{n > T_B^{p-1} : X_n \in B\}.$$

Sur l'événement $\{T_B^p < \infty\}, T_B^{p+1} = T_B^p + \inf\{n > 0 : X_n \circ \theta_{T_B^p} \in B\} = T_B^p + T_B \circ \theta_{T_B^p}$.

On a $R_B = \overline{lim}\{X_n \in B\} = \cap_p\{T_B^p < \infty\}$.
Pour tout $x \in E$, on note T_x^p pour $T_{\{x\}}^p$.

Lemme 2.13 $\mathbb{P}_x\{T_x^p < \infty\} = (\mathbb{P}_x\{T_x < \infty\})^p$.

Preuve : en utilisant la propriété de Markov forte et le fait que sur l'événement $\{T_x^p < \infty\}, X_{T_x^p}$ vaut x.

La propriété est évidente pour $p = 1$. On fait une preuve par récurrence. Donc on calcule

$$\mathbb{P}_x\{T_x^{p+1} < \infty\} = \mathbb{P}_x\{T_x^p + T_x \circ \theta_{T_x^p} < \infty\} = \mathbb{P}_x(\{T_x^p < \infty\} \cap \{T_x \circ \theta_{T_x^p} < \infty\})$$

et l'on passe aux intégrales d'indicatrices pour utiliser la propriété de Markov forte :

$$E_x[\mathbf{1}_{T_x^p<\infty}\mathbf{1}_{T_x<\infty} \circ \theta_{T_x^p}] = E_x[\mathbf{1}_{T_x^p<\infty} E_x[\mathbf{1}_{T_x<\infty} \circ \theta_{T_x^p}/\mathcal{F}_{T_x^p}]] =$$

$$E_x[\mathbf{1}_{T_x^p<\infty} E_{X_{T_x^p}}[\mathbf{1}_{T_x<\infty}]] = E_x[\mathbf{1}_{T_x^p<\infty} E_x[\mathbf{1}_{T_x<\infty}]] =$$

$$= (\mathbb{P}_x\{T_x < \infty\})^p \mathbb{P}_x\{T_x < \infty\}$$

ce qu'il fallait démontrer. •

Définition 2.14 *On appelle* **noyau potentiel** *de la chaîne de Markov l'application de $E \times \mathcal{E}$ dans $\overline{\mathbb{R}^+}$ définie par :*

$$U : (x, B) \mapsto \sum_{n\geq 0} Q^n(x, B) = \sum_{n\geq 0} \mathbb{P}_x\{X_n \in B\}$$

où $Q_0(x, B) = \delta_x(B)$.

Proposition 2.15 $P_x(R_x) = \lim_{p\to\infty} \mathbb{P}_x(T_x^p < \infty) = \lim_{p\to\infty}(\mathbb{P}_x(T_x < \infty))^p = 0$ *ou* 1.

Le potentiel $U(x, x) = \sum_{p\in\mathbb{N}}(\mathbb{P}_x(T_x < \infty))^p$.

Preuve : (i) L'événement $R_x = \cap_{p\geq 1}\{T_x^p < \infty\}$ or la suite $\{T_x^p < \infty\}$ est décroissante donc $\mathbb{P}(R_x) = \lim_p \mathbb{P}\{T_x^p < \infty\} = \lim_p(\mathbb{P}\{T_x < \infty\})^p$ d'après le lemme.
(ii) Le potentiel $U(x, x) = \sum_{n\geq 0} \mathbb{P}_x\{X_n = x\} = E_x[\sum_{n\geq 0} \mathbf{1}_x(X_n)]$ nombre moyen de passages en x. En particulier $\sum_{n\geq 1} \mathbf{1}_x(X_n) = \sum_{p\geq 1} \mathbf{1}_{\{T_x^p<\infty\}}$ d'espérance

$$\sum_{p\geq 1}(\mathbb{P}_x\{T_x < \infty\})^p = +\infty \text{ ou } \frac{\mathbb{P}_x\{T_x < \infty\}}{1 - \mathbb{P}_x\{T_x < \infty\}}$$

selon qu'il s'agit d'une série géométrique divergente ou convergente
(selon que $\mathbb{P}_x\{T_x < \infty\} = 1$ ou < 1).

•

Théorème 2.16 *Soit $x \in E$. Il n'y a que deux possibilités :*

a) $\mathbb{P}_x(T_x < \infty) = 1$: on dit que x est **récurrent** *et dans ce cas $\mathbb{P}_x(R_x) = 1, U(x,x) = +\infty$.*

b) $\mathbb{P}_x(T_x < \infty) < 1$: on dit que x est **transitoire** *ou transient et dans ce cas $\mathbb{P}_x(R_x) = 0, U(x,x) < +\infty$.*

Preuve immédiate : précisément $U(x,x) = \frac{1}{1-\mathbb{P}_x(T_x<\infty)}$. •

Proposition 2.17 *Soit x un point récurrent de E et $B \in \mathcal{E}$. L'un des deux cas est vérifié :*

a) $\mathbb{P}_x(R_B) = \mathbb{P}_x(T_B < \infty) = 1$,

b) $\mathbb{P}_x(R_B) = \mathbb{P}_x(T_B < \infty) = 0$.

Preuve : Rappel : $R_B \subset \{T_B < \infty\}$.
. Si x est récurrent et si $T_B < \infty$, alors il existe n tel que $X_n \in B$, et il existe p tel que $T_x^p \leq n < T_x^{p+1}$.
. Pour tout p, on pose

$$A_p = \cup_{n\geq 0}[\{X_n \in B\} \cap \{T_x^p \leq n < T_x^{p+1}\}].$$

Lemme 2.18 *Si x est récurrent, alors $\mathbb{P}_x$ presque sûrement : $\{T_B < \infty\} = \cup_p A_p$ et $R_B = \overline{\lim} A_p$.*

Preuve :
1. Si $T_B < \infty$, alors il existe n tel que $X_n \in B$ et p tel que $T_x^p \leq n < T_x^{p+1}$,

et réciproquement si $\omega \in A_p, \exists n / X_n \in B$ c'est à dire que $T_B(\omega) < \infty$.
2. Par définition, $\omega \in \overline{\lim} A_p$ signifie que cet élément appartient à une infinité de A_p, et pour chacun de ces p, il existe n tel que $T_x^p \leq n < T_x^{p+1}$ ce qui est dire qu'il y a une infinité de n avec $X_n \in B$ donc une infinité de passage en B soit R_B. •

Suite de la preuve de la proposition :
1. $A_p \in \mathcal{F}_{T_x^{p+1}}$ car $\forall k$ l'intersection
$A_p \cap \{T_x^{p+1} = k\} = \cup_{n=0}^{k-1}[\{X_n \in B\} \cap \{T_x^p \leq n\} \cap \{T_x^{p+1} = k\}]$.
2. Par ailleurs, si x est récurrent, $T_x^p < \infty$ et sur A_p $\theta_{T_x^p} \in A_0$ car $T_x^{x+1} = T_x^p + T_x \circ \theta_{T_x^p}$:

$$A_p \subset \{T_x^p < \infty \text{ et } \theta_{T_x^p} \in A_0\}$$

et réciproquement, si $T_x^p < \infty$ et $\theta_{T_x^p} \in A_0$, on est dans A_p, d'où l'égalité et

$$\mathbf{1}_{A_p} = \mathbf{1}_{\{T_x^p<\infty\}} \mathbf{1}_{A_0} \circ \theta_{T_x^p}.$$

3. On remarque que $X_{T_x^p} = x$

4. Ainsi $\mathbb{P}_x(A_p/\mathcal{F}_{T_x^p}) = \mathbf{1}_{\{T_x^p<\infty\}}\mathbb{P}_x(A_0)$.

Comme x est récurrent, pour tout p, $\mathbb{P}_x\{T_x^p < \infty\} = 1$ et $\mathbb{P}_x(A_p/\mathcal{F}_{T_x^p}) = \mathbb{P}_x(A_0)$ et A_p est indépendant de $\mathcal{F}_{T_x^p}$ donc de A_{p-1} et de plus tous les A_p ont la même probabilité. On applique le lemme de Borel-Cantelli à la suite d'événements A_p :

$$\{T_B < \infty\} = \cup_p A_p \text{ et } R_B = \limsup A_p.$$

Viennent les deux cas possibles annoncés :

. Si $\mathbb{P}_x(A_0) = 0$, alors $\mathbb{P}_x\{T_B < \infty\} = 0$ et $R_B \subset \{T_B < \infty\}$ donc $\mathbb{P}_x(R_B) = 0$.

.. Si $\mathbb{P}_x(A_0) \neq 0$, alors $R_B = \limsup A_p$ est de probabilité 1 donc aussi $\{T_B < \infty\}$ puisque $\sum_p \mathbb{P}_x(A_p) = \infty$. •

2.9.2 Communication

Définition 2.19 *On dit que x* **conduit à** *B si $\mathbb{P}_x(T_B < \infty) > 0$, à y si $\mathbb{P}_x(T_y < \infty) > 0$, et on admet que x conduit à x.*

Proposition 2.20 *La relation x conduit à y est transitive.*

Preuve : soient $x, y, z \in E$, (rappel : sur $\{T_z > T_y\}$, $T_z = T_y + T_z \circ \theta_{T_y}$)

$(T_z < \infty) = (T_z < \infty) \cap (T_z > T_y) \cup (T_z < \infty) \cap (T_z < T_y) \supset (T_y < \infty) \cap (T_z \circ \theta_{T_y} < \infty)$.

d'où

$$\mathbb{P}_x(T_z < \infty) \geq \mathbb{P}_x[(T_y < \infty) \cap (T_z \circ \theta_{T_y} < \infty)] =$$
$$E_x[\mathbf{1}_{T_y<\infty} E_x(\mathbf{1}_{T_z<\infty} \circ \theta_{T_y}/\mathcal{F}_{T_y})] = \mathbb{P}_x(T_y < \infty)\mathbb{P}_y(T_z < \infty).$$

•

Définition 2.21 *Soient x et y deux éléments de E. On dit que x et y* **communiquent** *si on peut atteindre y à partir de x avec une probabilité strictement positive (x conduit à y) et réciproquement.*

exercice : c'est une relation d'équivalence (réflexive, symétrique, transitive).

Proposition 2.22 QC5
Soit x un point récurrent et y tel que x conduit à y ; alors y conduit à x et y est aussi récurrent.

Preuve : (i) Si x est récurrent, on est $\mathbb{P}_x$ presque sûrement dans R_x. Si on est de plus sur $\{T_y < \infty\}$, il existe k tel que $T_x^{k-1} < T_y < T_x^k$. Alors, sur cet ensemble, on a $T_x^k = T_y + T_x \circ \theta_{T_y}$ c'est à dire le premier instant après T_y où on passe en x.

Ceci montre l'inclusion $\mathbb{P}_x$ presque sûre :

$$\{T_y < \infty\} \cap R_x \subset \{T_y < \infty\} \cap \{T_x \circ \theta_{T_y} < \infty\}$$

On intègre cette inclusion sous $\mathbb{P}_x$ et on utilise la propriété de Markov :

$$0 < \mathbb{P}_x\{T_y < \infty\} = \mathbb{P}_x(\{T_y < \infty\} \cap R_x) \leq \mathbb{P}_x(\{T_y < \infty\} \cap \{T_x \circ \theta_{T_y} < \infty\})$$
$$= \mathbb{P}_x\{T_y < \infty\}\mathbb{P}_y\{T_x < \infty\}$$

donc $\mathbb{P}_y(T_x < \infty) = 1$: y conduit à x.

(ii) Par ailleurs, x récurrent et $\mathbb{P}_x(T_y < \infty) > 0$ implique $\mathbb{P}_x(T_y < \infty) = 1$ grâce à la proposition 2.17.

Enfin, on déduit de l'inclusion

$$\{T_x < \infty\} \cap \{T_y \circ \theta_{T_x} < \infty\} \subset \{T_y < \infty\}$$

toujours de la même façon en intégrant cette inclusion sous $\mathbb{P}_y$

$$\mathbb{P}_y(\{T_x < \infty\}) \times \mathbb{P}_x(\{T_y < \infty\}) \leq \mathbb{P}_y(\{T_y < \infty\}),$$

donc que y est récurrent puisque le produit à gauche est 1. •

Corollaire 2.23 *Si x est récurrent et si h est une fonction invariante (ou excessive) alors $h(X_n) = h(x), \mathbb{P}_x$ presque sûrement.*

Preuve : On a vu que $(h(X_n), n \in \mathbb{N})$ est une $\mathbb{P}_x$-(sur)martingale, positive, donc elle converge presque sûrement.

Par ailleurs, $\mathbb{P}_x(\overline{\lim}\{X_n = x\}) = 1$, puisque x est récurrent : la chaine passe infiniment souvent en x. Donc, il existe une suite extraite telle que $X_{n_i} = x$, donc une suite extraite de $(h(X_n))$ est constante et égale à $h(x)$ et la limite presque sûre de $(h(X_n))$ est nécessairement $h(x)$.
Soit y une valeur prise par la suite (X_n) partant de x, c'est à dire que x conduit à y. D'après la proposition 2.22 y est alors aussi récurrent et d'après la proposition 2.17, puisque $\mathbb{P}_x(T_y < \infty) > 0$, $\mathbb{P}_x(T_y < \infty) = 1$. Puisque y est récurrent, il existe une suite extraite telle que $X_{n_i} = y$, donc une suite extraite de $(h(X_n))$ est constante et égale à $h(y)$ et la limite presque sûre de $(h(X_n))$ devrait être aussi $h(y)$. A cause de l'unicité de la limite, on a $h(x) = h(y)$ pour tout y auquel conduit x, donc $h(X_n)$ est constante et égale à $h(x)$ lorsque l'on part de x récurrent. •

Soit E un ensemble dénombrable et a un point récurrent de E. Soit C l'ensemble des points auxquels conduit a.
D'après la proposition précédente, tous les points de C sont aussi récurrents et ne conduisent qu'à des éléments de C.

Définition 2.24 *On appelle C, classe d'équivalence de a,* **la classe récurrente** *de a.*

On peut restreindre l'espace d'état à C ; la chaîne est alors dite **récurrente** : tous ses états sont récurrents et communiquent entre eux.

Rappel : le noyau potentiel pour $x \in E$ et $B \in \mathcal{E}$,

$$U(x, B) = \sum_{n\geq 0} Q^n(x, B) = \sum_{n\geq 0} \mathbb{P}_x(X_n \in B).$$

Proposition 2.25 *Pour tout y différent de x : 1. $U(y, x) = \mathbb{P}_y(T_x < \infty)U(x, x)$.*
2. Si x est transient, $U(y, x) < \infty$.
3. Si x est récurrent et que y conduit à $x, U(y, x) = \infty$.

Preuve :

$$U(y, x) = E_y[\sum_{n\geq T_x} \mathbf{1}_{\{X_n=x\}}] = E_y[\sum_{k\geq 0} \mathbf{1}_{\{X_k=x\}} \circ \theta_{T_x} \mathbf{1}_{\{T_x<\infty\}}] =$$

$$E_y[\mathbf{1}_{\{T_x<\infty\}} E_{X_{T_x}}(\sum_{k\geq 0} \mathbf{1}_{\{X_k=x\}})] = \mathbb{P}_y(T_x < \infty)U(x, x).$$

Si x est transient, $U(x, x)$ est fini et donc aussi $U(y, x)$.
Si x est récurrent, $U(x, x)$ est infini ; si y conduit à x, $\mathbb{P}_y(T_x < \infty) > 0$ et $U(y, x) = \infty$. •

Définition 2.26 *Un ensemble $B \in \mathcal{E}$ est appelé* **absorbant** *si aucun x de B ne conduit à B^c, c'est à dire $\forall x \in B$, $\mathbb{P}_x(T_{B^c} < \infty) = 0$.*

Ceci veut dire que pour $x \in B, Q(x, .)$ est concentrée sur B. On peut prendre la restriction de Q à l'espace d'états $(B, \mathcal{E} \cap B)$, cela définit une nouvelle chaîne.

Exemple avec $E = \{a, b, c, d, e, f\}$ et une matrice Q bien choisie pour avoir deux ensembles absorbants ou l'on voit les classes irréductibles.

Exemples : soit $A \in \mathcal{E}$.

a) L'ensemble B des points qui ne conduisent pas à A est **absorbant** *:*
$B = \{x, \mathbb{P}_x\{T_A < \infty\} = 0\}$
En effet, si $T_{B^c} < \infty$ et $T_A \circ \theta_{T_{B^c}} < \infty$, alors $T_A = T_{B^c} + T_A \circ \theta_{T_{B^c}} < \infty$.
Soit $x \in B$:

$$\mathbb{P}_x(T_A < \infty) = 0 \geq \mathbb{P}_x(T_{B^c} < \infty, T_A \circ \theta_{T_{B^c}} < \infty),$$

d'où

$$E_x(\mathbf{1}_{\{T_{B^c}<\infty\}} E_{X_{T_{B^c}}}(\mathbf{1}_{\{T_A<\infty\}})) = 0.$$

Or, $y = X_{T_{B^c}} \in B^c$, et donc $E_{X_{T_{B^c}}}(\mathbf{1}_{\{T_A<\infty\}}) > 0$. Dans l'espérance ci-dessus, il faut alors que l'autre facteur soit nul sous $\mathbb{P}_x$, soit $\mathbb{P}_x(T_{B^c} < \infty) = 0$, c'est à dire que B est aborbant. Notons de plus que $B \subset A^c$ car $\forall x \in A, x$ conduit à lui-même et ne peut être dans B.

b) L'ensemble $D = \{y : \mathbb{P}_y(R_A) = 1\}$ est **absorbant**.

En effet, soit $x \in D$, sur $\{T_{D^c} < \infty\}$, revenir un nombre infini de fois en A signifie revenir un nombre infini de fois en A après T_{D^c} :
$\{T_{D^c} < \infty\} \cap R_A = \{T_{D^c} < \infty\} \cap R_A \circ \theta_{T_{D^c}}$

$$E_x[\mathbf{1}_{\{T_{D^c}<\infty\}}\mathbf{1}_{R_A}/\mathcal{F}_{T_{D^c}}] = E_x[\mathbf{1}_{\{T_{D^c}<\infty\}}E_x(\mathbf{1}_{R_A} \circ \theta_{T_{D^c}}/\mathcal{F}_{T_{D^c}})] =$$
$$= \mathbf{1}_{\{T_{D^c}<\infty\}}E_{X_{T_{D^c}}}(\mathbf{1}_{R_A}),$$

par la propriété de Markov. Puisque $X_{T_{D^c}} \in D^c, E_{X_{T_{D^c}}}(\mathbf{1}_{R_A}) < 1$, la variable aléatoire $\mathbf{1}_{\{T_{D^c}<\infty\}}(1 - E_{X_{T_{D^c}}}(\mathbf{1}_{R_A})) > 0$ et on en prend l'espérance sous $\mathbb{P}_x$:

$$E_x[\mathbf{1}_{\{T_{D^c}<\infty\}}(1 - E_{X_{T_{D^c}}}(\mathbf{1}_{R_A}))] > 0.$$

Mais par ailleurs

$$\mathbb{P}_x(R_A \cap \{T_{D^c} < \infty\}) = \mathbb{P}_x(\{T_{D^c} < \infty\}) = E_x[\mathbf{1}_{\{T_{D^c}<\infty\}}E_{X_{T_{D^c}}}(\mathbf{1}_{R_A})] < \mathbb{P}_x\{T_{D^c} < \infty\},$$

ce qui nécessite $\mathbb{P}_x(\{T_{D^c} < \infty\} = 0$. •

Certains auteurs utilisent une autre terminologie pour la notion d'ensemble absorbant :

Définition 2.27 *Une partie F de E est* **close** *si pour tout $x \in F, \mathbb{P}_x(X_n \in F,\ \forall n) = 1$.*

Une partie close est de fait un ensemble absorbant : pour tout $x \in F, \mathbb{P}_x(T_{F^c} < \infty) = 0$.

Remarquons que si E est dénombrable et si a est un état récurrent, alors la classe récurrente C de a est contenue dans $D = \{x : \mathbb{P}_x(R_a) = 1\}$ (si a est récurrent et conduit à x, alors $\mathbb{P}_x\{T_a < \infty\} = 1$ soit $\mathbb{P}_x(R_a = 1)$).
L'inclusion est parfois stricte ; il suffit par exemple de prendre
$E = (a, b), Q(a, a) = Q(b, a) = 1.\ D = E,\ C = \{a\}$.

2.9.3 Classification des états

Rappel : $x \neq y \to U(y, x) = \mathbb{P}_y(T_x < \infty)U(x, x),\ U(x, y) = \mathbb{P}_x(T_y < \infty)U(y, y)$.

Proposition 2.28 *([5] prop. 4.18).*
Il existe une partition $(C_i, i \in I)$ des états récurrents de E telle que

(i) si $x, y \in C_i,\ \mathbb{P}_x(R_y) = 1,\ U(x, y) = +\infty$.

(ii) si $x \in C_i, y \in C_j, i \neq j, U(x, y) = 0$ et $\mathbb{P}_x(\sum_{n\geq 0} \mathbf{1}_{\{y\}}(X_n) = 0) = 1$.

Preuve en exercice

(i) x et y sont récurrents et équivalents : $\mathbb{P}_x(T_y < \infty) = 1$, donc $\mathbb{P}_x(R_y) = 1$ d'après la proposition 2.17.
Et la proposition 2.25 montre que $U(x, y) = \mathbb{P}_x(T_y < \infty)U(y, y) = +\infty$ puisque y est récurrent.

(ii) x et y sont récurrents et dans deux classes différentes.
$\mathbb{P}_x(T_y < \infty) = 0, \forall n > 0, X_n \neq y, U(x, y) = 0$.

Définition 2.29 *Si E est la seule partie close (ou absorbante), la chaîne est dite* **irréductible**. *Une matrice de transition (matrice stochastique) est dite irréductible si la chaîne associée est irréductible.*

Lemme 2.30 *Une chaîne est irréductible si et seulement si*

(H) pour tout x de $E, U(x,y) > 0, \forall y \in E$,
c'est à dire que tous les états communiquent entre eux deux à deux.

Remarquons que $U(x,x) = +\infty$ dans le cas récurrent et $\frac{1}{1-\mathbb{P}_x(T_x<\infty)} > 0$ dans les deux cas.
Preuve : (H) implique que la chaîne est irréductible : supposons qu'il existe une partie F de E, stricte, et close.
Pour tout $x \in F, \mathbb{P}_x(T_{F^c} < \infty) = 0$.
Puisque F est un sous ensemble strict, F^c n'est pas vide : il existe $y \in F^c$ qu'on ne peut atteindre à partir de $x : U(x,y) = 0$, ce qui contredit l'hypothèse (H).

Réciproquement, supposons la chaine irréductible et que (H) n'est pas vérifiée : il existe $y \in E$ et $x \neq y, U(x,y) = 0, (U(y,y) \geq 1$ ne peut être nul), donc en particulier, $\mathbb{P}_x\{T_y < \infty\} = 0$: x ne conduit pas à y.
Alors, d'après l'exemple a) ci-dessus, l'ensemble
$F = \{x' \in E$ qui ne conduisent pas à $y\}$ est absorbant.
Par ailleurs, F n'est pas vide, il contient x, et est strictement contenu dans E puisque $y \in F^c$. Et on a obtenu l'existence d'une partie close stricte.

Proposition 2.31 QC6
Une partie F est close (ou absorbante) si et seulement si $\forall x \in F, Q(x,F) = 1$ ou si et seulement si $\forall x \in F, \forall y \in F^c, U(x,y) = 0$.

Preuve :

$1 \Rightarrow 2$. si F est close, pour tout $x \in F, \mathbb{P}_x\{X_n \in F, \ \forall n\} = 1$. En particulier, $Q(x,F) = \mathbb{P}_x\{X_1 \in F\} = 1$.

$2 \Rightarrow 3$. $Q(x,F) = 1, \forall x \in F$ implique que $Q(x,F^c) = 0$, et ceci pour tout $x \in F$. Donc, $\mathbb{P}_x$ presque sûrement, par récurrence, $X_1 \in F, X_2 \in F, \cdots, X_n \in F, \forall n$. C'est à dire que $Q^n(x,F) = 1, Q^n(x,F^c) = 0$, et par conséquent, pour tout $y \in F^c, U(x,y) = 0$.

$3 \Rightarrow 1$. Si pour tout $x \in F$ et tout $y \in F^c, U(x,y) = 0$, cela signifie que pour tout $n, \mathbb{P}_x(X_n = y) = 0$, soit $\mathbb{P}_x\{X_n \in F^c\} = 0$. Plus précisément, $\mathbb{P}_x\{T_{F^c} < \infty\} = 0, \forall x \in F$, c'est à dire que F est close ou absorbante. •

Définition 2.32 *Lorsqu'une chaîne est irréductible :*
si tout $x \in E$ est récurrent, on dit qu'elle est **récurrente** *;*
sinon, tous les états sont transients, et on dit qu'elle est **transiente**.

On peut remarquer que dans le cas d'une chaîne irréductible, tous les états communiquent, ils sont donc tous transients ou tous récurrents.

$$U(y,x) = \mathbb{P}_y(T_x < \infty)U(x,x)$$

sont finis ou infinis en même temps puisque $\mathbb{P}_y(T_x < \infty) > 0$.

2.10 Mesure invariante

Définition 2.33 *On appelle* **mesure invariante** *une mesure non nulle μ sur $(E,\mathcal{E})$ telle que $\mu Q = \mu$, c'est à dire $\forall x \in E, \sum_{y\in E} \mu(y)Q(y,x) = \mu(x)$.*

Hors programme : plus généralement, si E n'est pas discret, μ est invariante si elle est positive non nulle $\sigma-$finie et $\int \mu(dy)Q(y,dx) = \mu(dx)$.

Lemme 2.34 *Si X est une chaîne de Markov irréductible qui possède une probabilité invariante, alors X est récurrente.*

Preuve : si μ est invariante, alors pour tout $n, \mu = \mu Q^n$, soit :

$$\forall n,\ \forall x \in E, \mu(x) = \sum_{y\in E} \mu(y)Q^n(y,x).$$

Supposons la chaîne transiente : il existe x avec $U(x,x) < \infty$. Alors pour tout $y, U(y,x) < \infty$ donc $Q^n(y,x) \to 0$ quand n tend vers l'infini. Comme $Q^n(y,x) \leq 1$, on peut appliquer le théorème de convergence dominée et passer à la limite ci-dessus, soit $\mu(x) = 0,\ \forall x$, ce qui n'est pas possible (par définition, μ est non nulle). Donc la chaîne n'est pas transiente : il existe au moins un état récurrent ; comme X est irréductible, tous les états communiquent et sont donc tous récurrents. •

Proposition 2.35 *Si X est une chaîne de Markov irréductible et récurrente et si f est une fonction excessive, f est constante.*

Preuve : on a vu que si f est excessive, $(f(X_n) = f(x), n \geq 0)$ $\mathbb{P}_x$ presque sûrement.

La chaîne est récurrente : soit $y \neq x$, ce sont des points récurrents, x conduit à y donc il existe n tel que $X_n = y$ donc $f(X_n) = f(y)$ $\mathbb{P}_x$ presque sûrement, ceci nécessite que $f(x) = f(y)$. •

Voici la contraposée.

Corollaire 2.36 *Si une chaîne irréductible admet une fonction excessive non constante, elle est transiente.*

se rappeler l'exercice : si X est de Markov, x récurrent et h invariante, alors la suite $h(X_n)$ est constante $\mathbb{P}_x$-presque sûrement.

Exemple de chaîne irréductible transiente : $E = \mathbb{Z}$, $Q(n, n-1) = p$, $Q(n, n+1) = 1 - p$, $0p \neq \frac{1}{2}$ sinon.
En effet, une suite récurrente permet de trouver au moins une fonction invariante non constante et positive en partant d'un couple $(f(0), f(1))$ vérifiant $f(1) < f(0) < \frac{p}{1-p} f(1)$.

Exercices
1. Soit μ une probabilité invariante pour une chaîne X. Alors, sous la loi $\mathbb{P}_\mu$, la chaîne est stationnaire, c'est à dire que la loi du vecteur aléatoire $(X_r, \cdots, X_{r+n})$ ne dépend pas de r.

Rappelons que la loi de X_r sous $\mathbb{P}_\mu$ est μQ^r. Puis notons que l'invariance de μ montre que pour tout $r, \mu Q^r = \mu$, donc sous $\mathbb{P}_\mu$, la loi de X_r est μ. Soit maintenant f borélienne bornée sur $\mathbb{R}^n, r \geq 0$.

$$E_\mu[f(X_r, \cdots, X_{r+n})] = E_\mu[f(X_0, \cdots, X_n) \circ \theta_r] = E_\mu[E_\mu[f(X_0, \cdots, X_n) \circ \theta_r / \mathcal{F}_r]] =$$

$$= E_\mu[E_{X_r}[f(X_0, \cdots, X_n)]] = \int_E \mu(dx) E_x[f(X_0, \cdots, X_n)] = E_\mu[f(X_0, \cdots, X_n)]$$

puisque X_r est de loi μ sous $\mathbb{P}_\mu$.

2. Chaîne duale : si la mesure invariante μ vérifie $\forall x \in E, \mu(x) \in \mathbb{R}_*^+$ on pose $\hat{Q}(x, y) = \frac{\mu(y)}{\mu(x)} Q(y, x)$. Montrer que $\hat{Q}$ est un noyau de transition et que si Q est irréductible récurrente, il en est de même de $\hat{Q}$.

On suppose que pour tout $x, \mu(x) > 0$, donc $\hat{Q}$ est bien définie. Montrons que c'est un noyau de transition :

- $\hat{Q}(x, .)$ est une probabilité : c'est une application positive sur E et $\sum_y \hat{Q}(x, y) = \sum_y \frac{\mu(y)}{\mu(x)} Q(y, x) = \frac{1}{\mu(x)} \sum_y \mu(y) Q(y, x) = 1$ à cause de l'invariance de μ.

- l'application $x \mapsto \hat{Q}(x, y)$ est mesurable pour tout y, car définie sur un ensemble dénombrable muni de la tribu des parties.

Remarquons par récurrence que $\hat{Q}^n(x, y) = \frac{\mu(y)}{\mu(x)} Q^n(y, x)$:

$$\hat{Q}^n(x, y) = \sum_z \hat{Q}(x, z) \hat{Q}_{n-1}(z, y) = \sum_z \frac{\mu(z)}{\mu(x)} \frac{\mu(y)}{\mu(z)} Q(z, x) Q_{n-1}(y, z) = \frac{\mu(y)}{\mu(x)} Q^n(x, y).$$

Donc $\hat{U}(x, y) = \frac{\mu(y)}{\mu(x)} U(y, x)$ et cette relation montre que Q et $\hat{Q}$ sont irréductibles et récurrentes en même temps. •

Lemme 2.37 *Soit une chaîne irréductible récurrente de transition Q et une mesure μ telle que $\mu \geq \mu Q$. Alors*

$\exists x$ tel que $\mu(x) = 0$ implique $\mu(y) = 0,\ \forall y$.
$\exists x$ tel que $\mu(x) < \infty$ implique $\mu(y) < \infty,\ \forall y$.
$\exists x$ tel que $\mu(x) > 0$ implique $\mu(y) > 0,\ \forall y$.

Preuve : Puisque la chaîne est irréductible, $U(y,x) > 0$ pour $x \neq y$: $\exists n, Q^n(y,x) > 0$. Comme $\mu \geq \mu Q^n, \mu(x) \geq \sum_z \mu(z) Q^n(z,x) \geq \mu(y) Q^{\otimes n}(y,x)$.
Ainsi, $\mu(x) = 0 \Rightarrow \mu(y) = 0, \mu(x) < \infty \Rightarrow \mu(y) < \infty$.
S'il existe x avec $\mu(x) > 0$, supposons qu'il existe $y,\ \mu(y) = 0$. Alors tous les points sont de mesure nulle, d'où la contradiction : ou ils sont tous de mesure > 0 ou tous de mesure nulle. •

Théorème 2.38 QC7
Soit X une chaîne irréductible récurrente. Alors X possède une mesure invariante unique à un coefficient multiplicatif près.
Cette mesure vérifie $\forall y \in E, \mu(y) \in \mathbb{R}_*^+$ et pour $x \in E$, la mesure unique μ_x telle que $\mu_x(x) = 1$ est donnée par

$$\mu_x(y) = E_x[\sum_{k=0}^{T_x - 1} \mathbf{1}_{\{y\}}(X_k)].$$

De plus, μ_x est telle que $\mu_x(E) = E_x(T_x) \in \overline{\mathbb{R}}^+$.

Preuve : (i) On définit une mesure $\nu_x(y) = E_x[\sum_{k=1}^{T_x} \mathbf{1}_{\{y\}}(X_k)]$ et pour f mesurable positive, $\mu_x(f) = E_x[\sum_{k=0}^{T_x-1} f(X_k)]$ et $\nu_x(f) = E_x[\sum_{k=1}^{T_x} f(X_k)]$. Par ailleurs, $\mu_x(x) = \nu_x(x) = 1, \mu_x(E) = E_x(T_x) = \nu_x(E)$. De fait, $\mu_x = \nu_x$.

. μ_x est invariante

$$\mu_x Q(f) = \mu_x(Qf) = E_x[\sum_{k=0}^{T_x-1} Qf(X_k)] = \sum_{k \geq 0} E_x[\mathbf{1}_{k<T_x} E_{X_k}(f(X_1))] =$$

$$= \sum_{k \geq 0} E_x[\mathbf{1}_{k<T_x} f(X_{k+1})] = \sum_{j \geq 1} E_x[\mathbf{1}_{j<T_x+1} f(X_j)] = \nu_x(f) = \mu_x(f),$$

c'est dire que μ_x est invariante.
(ii) Le lemme 2.37, avec μ_x invariante et $\mu_x(x) = 1$, montre que μ_x est à valeurs dans $\mathbb{R}_*^+$.
(iii) Unicité : soit λ une mesure invariante $0 < \lambda(y) < \infty,\ \forall y$. On définit une chaine duale $\hat{Q}$ avec μ_x comme dans l'exercice 2. Pour $\hat{Q}$, la fonction $f : y \mapsto \frac{\lambda(y)}{\mu_x(y)}$ est excessive ; en effet :

$$\hat{Q} f(y) = \sum_z \hat{Q}(y,z) \frac{\lambda(z)}{\mu_x(z)} = \sum_z Q(z,y) \frac{\lambda(z)}{\mu_x(y)} \leq \frac{\lambda(y)}{\mu_x(y)} = f(y)$$

puisque λ est excessive. Comme la chaîne est irréductible récurrente, f est constante, et donc λ est proportionnelle à μ. •

Pour conclure, on est en présence de la dichotomie suivante : ou bien $\mu(E) = +\infty$ pour toute mesure invariante μ, ou bien $\mu(E) < \infty$.

Théorème 2.39 *(cf. le théorème 4.23 page 104 de [5])*
Soit X une chaîne irréductible récurrente de mesure invariante μ. Alors deux cas sont possibles :

(i) $\mu(E) = +\infty$ auquel cas pour tout $x \in E, E_x(T_x) = +\infty$. la chaîne est dite **récurrente nulle**.

(ii) $\mu(E) < \infty$ auquel cas pour tout $x \in E, E_x(T_x) < +\infty$. la chaîne est dite **récurrente positive**. *Dans ce cas, l'unique probabilité invariante est donnée par*

$$\pi(y) = \frac{E_x[\sum_{k=0}^{T_x-1} \mathbf{1}_{\{y\}}(X_k)]}{E_x(T_x)}.$$

En particulier, $\forall x, \pi(x) = \frac{1}{E_x(T_x)}$.

Preuve : μ est unique à un coefficient multiplicatif près. Par le théorème précédent, pour x fixé, $\mu_x(E) = E_x(T_x)$.

(i) Si $E_x(T_x) = +\infty, \mu_x(E) = \mu(E) = +\infty$ pour tout autre mesure invariante.

(ii) Si $E_x(T_x) < +\infty, \mu_x(E)$ et $\mu(E) < \infty$ pour tout autre mesure invariante. Dans ce cas, on obtient une probabilité invariante en divisant μ_x par $\mu_x(E) = E_x(T_x)$, d'où le résultat. •

Plus généralement, on a toujours le résultat suivant concernant les mesures invariantes, même pour un espace E non dénombrable. Mais alors, il faut montrer de plus la σ-finitude.

Théorème 2.40 *Soit une chaine pas nécessairement irréductibe. Soit $a \in E$ un état récurrent et la mesure définie pour tout $A \in \mathcal{E}$ par*

$$\mu(A) = E_a[\sum_{k=0}^{T_a-1} \mathbf{1}_A(X_k)]$$

est invariante pour Q ; de plus $\mu(A)$ est l'espérance du temps de séjour en A entre deux passages en a. Cette mesure est concentrée sur l'ensemble absorbant (partie close) $D = \{x \in E, \mathbb{P}_x(R_a) = 1\}$.

Preuve : (i) Puisque a est récurrent, $a \in D$, et puisque D est absorbant, on a $E_a[\mathbf{1}_{X_n \in D^c}] = 0$ pour tout n donc $\mu(D^c) = 0$ et μ est à support dans D.

(ii) μ est $\sigma-$finie (facultatif).

(iii) μ est invariante (cf. la preuve de la QC7) : pour $x \in E$,

$$\mu Q(x) = \sum_y \mu(y)Q(y,x) = E_a[\sum_{k=0}^{T_a-1} Q(X_k, x)] =$$

$$\sum_{k\geq 0} E_a[\mathbf{1}_{\{k<T_a\}}Q(X_k,A)] = \sum_{k\geq 0} E_a[\mathbf{1}_{\{k<T_a\}}E_{X_k}[\mathbf{1}_A(X_1)]] =$$

$$\sum_{k\geq 0} E_a[\mathbf{1}_{\{k<T_a\}}E_a[\mathbf{1}_A(X_{1+k})/\mathcal{F}_k]] = \sum_{k\geq 0} E_a[(\mathbf{1}_{\{k+1<T_a\}} + \mathbf{1}_{\{k+1=T_a\}})\mathbf{1}_A(X_{1+k})] =$$

$$\sum_{k\geq 1} E_a[\mathbf{1}_{\{k<T_a\}}\mathbf{1}_A(X_k)] + \mathbf{1}_A(a) = \mu(A) - \mathbb{P}_a(X_0\in A) + \mathbf{1}_A(a) = \mu(A)$$

et μ est invariante.

Il est enfin clair dans la définition que $\mu(A)$ est, sous $\mathbb{P}_a$, l'espérance du temps de passage entre 0 et T_a, soit entre deux temps de passage en a. passage entre 0 •

Définition 2.41 *Un état a est dit* **récurrent positif** *si la mesure précédente est finie, c'est à dire si* $E_a(T_a) < +\infty$*. Il est dit* **récurrent nul** *si la mesure précédente est infinie, c'est à dire si* $E_a(T_a) = +\infty$*.*

En effet, il vient de la définition que $\mu(E) = E_a(T_a)$.

On verra des exemples de tels états dans les marches aléatoires et les files d'attente.

Quelques remarques supplémentaires à traiter en exercice :

Si E est fini, il existe au moins un point récurrent.

$$\sum_{y\in E} U(x,y) = \sum_{y\in E}\sum_{n\geq 0}\mathbb{P}_x\{X_n = y\} = \sum_{n\geq 0}\mathbb{P}_x\{X_n\in E\} = +\infty$$

Donc, si E est fini, cela veut dire qu'il existe au moins un $y\in E$ tel que $U(x,y)=+\infty$. On utilise alors la proposition 2.25: pour tout y différent de $x, U(x,y) = \mathbb{P}_x(T_y<\infty)U(y,y)$. Donc $U(y,y)$ est infini et y est récurrent. •

Si E est fini et que la chaine est irréductible, elle est récurrente positive.
En conséquence, si E est fini et que la chaine est irréductible, elle est admet un point récurrent, elle est donc elle admet une mesure invariante μ. Dans ce cas, la chaine est nécessairement récurrente et récurrente positive car E fini montre que μ est finie. •

Si π est une probabilité invariante et que x est transient, alors $\pi(x) = 0$.
$U(x,x) = \sum_n \mathbb{P}_x(X_n = x) < \infty$. Or $U_N(y,x) = \sum_{n\leq N} Q^n(y,x)$ converge vers $U(y,x)$ et $U(y,x)\leq U(x,x)<\infty$. (toujours proposition 2.25 qui donne $U(y,x) = U(x,x)\mathbb{P}_y(T_x < \infty)$).
Donc, d'une part pour tout $n, \pi(x) = \sum_{y\in E}\pi(y)Q^n(y,x) = \frac{1}{N+1}\sum_{y\in E}\pi(y)U_N(y,x)$. Puis on fait tendre N vers l'infini puisque U_N est une suite croissante convergente vers U, donc $\frac{U_N}{N+1}$ tend vers 0 et $\pi(x) = 0$. •

2.11 Théorèmes ergodiques

Il s'agit de l'étude de limites de quantité du type $\frac{1}{n}\sum_{k=0}^{n-1} f(X_k)$ sans avoir l'indépendance, c'est à dire sans la loi des grands nombres...

Théorème 2.42 *Soit X une chaîne irréductible récurrente de mesure invariante μ et de loi initiale λ. Soient f et g dans $L^1(\mu)$ et $\mu(g) \neq 0$. Alors, $\mathbb{P}_\lambda$ presque sûrement,*

$$\lim \frac{\sum_{k=0}^{n-1} f(X_k)}{\sum_{k=0}^{n-1} g(X_k)} = \frac{\mu(f)}{\mu(g)}.$$

Preuve : On fixe x dans E et μ la mesure invariante μ_x. On note T_x^p le $p$ième temps de passage en x, et pour $f \in L^1(\mu)$, on pose

$$Z_0 = \sum_{k=0}^{T_x - 1} f(X_k) \ ; \ Z_p = Z_0 \circ \theta_{T_x^p} = \sum_{k=T_x^p}^{T_x^{p+1}-1} f(X_k).$$

Par définition de $\mu = \mu_x, E_x(Z_0) = \mu_x(f)$. Soit λ une loi initiale pour la chaîne. Pour tout p et g mesurable et bornée,

$$E_\lambda[g(Z_p)] = E_\lambda[g(Z_0) \circ \theta_{T_p}] = E_\lambda[E_\lambda[g(Z_0) \circ \theta_{T_p} / \mathcal{F}_{T_p}]] =$$

$$E_\lambda[E_{X_{T_p}} g(Z_0)]] = E_\lambda[E_x(g(Z_0))] = E_x[g(Z_0)]$$

donc les Z_p sont toutes de même loi.

On peut montrer qu'elles sont également indépendantes : intuitivement, la loi de Z_p ne dépend que de la position de la chaîne en T_p, c'est à dire x, et pas de ce qui s'est passé strictement avant T_p, c'est à dire que Z_p est indépendante des $Z_i, i < p$. Plus précisément, soit un borélien A et on calcule

$$\mathbb{P}_\lambda[Z_p \in A/\mathcal{F}_{T_{p-1}}] = E_\lambda[\mathbb{P}_\lambda[Z_0 \circ \theta_p \in A/\mathcal{F}_{T_p}]/\mathcal{F}_{T_{p-1}}] = E_\lambda[\mathbb{P}_x[Z_0 \in A]/\mathcal{F}_{T_p - 1}] = \mathbb{P}_x[Z_0 \in A]$$

ce qui donne en même temps l'indépendance et l'équidistribution.

On peut donc appliquer à cette suite Z la loi des grands nombres : $\mathbb{P}_\lambda$ presque sûrement

$$\frac{1}{n} \sum_{k=0}^{T_x^n - 1} f(X_k) = \frac{1}{n} \sum_{k=0}^{n-1} Z_k \to \mu(f) = E_x(Z_0).$$

Donc, avec f et g il vient le quotient limite :

$$u_n = \frac{\sum_{k=0}^{T_n - 1} f(X_k)}{\sum_{k=0}^{T_n - 1} g(X_k)} \to \frac{\mu(f)}{\mu(g)}$$

puisque $\mu(g) \neq 0$.

On passe ensuite à n au lieu de T_n : soit n entier ; il existe p tel que $T_p < n \leq T_{p+1}$ car la récurrence montre que la suite de ces temps d'arrêt est croissante et il vient :

$$\frac{\sum_{k=0}^{n-1} f(X_k)}{\sum_{k=0}^{n-1} g(X_k)} = \frac{\sum_{k=0}^{T_p-1} f(X_k) + \sum_{T_p}^{n-1} f(X_k)}{\sum_{k=0}^{T_p-1} g(X_k) + \sum_{T_p}^{n-1} g(X_k)} = u_p.k_{n,p}$$

avec

$$k_{n,p} = \frac{1 + \frac{\sum_{T_p}^{n-1} f(X_k)}{\sum_{k=0}^{T_p-1} f(X_k)}}{1 + \frac{\sum_{T_p}^{n-1} g(X_k)}{\sum_{k=0}^{T_p-1} g(X_k)}}$$

dont la limite est 1 : il suffit de voir par exemple que le numérateur

$$1 + \frac{\frac{1}{p}\sum_{T_p}^{n-1} f(X_k)}{\frac{1}{p}\sum_{k=0}^{T_p-1} f(X_k)}$$

converge presque sûrement vers 1. Or ce numérateur moins 1 est un quotient de dénominateur convergeant presque sûrement vers $\mu(f)$, et le numérateur est majoré en valeur absolue par $\frac{1}{p}\sum_{T_p}^{T_{p+1}-1} |f|(X_k) = \frac{1}{p}Z_p^{|f|}$ qui tend presque sûrement vers 0. •

Corollaire 2.43 *Soit X une chaîne irréductible récurrente de mesure invariante μ et de loi initiale λ. Le quotient du nombre de visites en y avant n sur le nombre de visites en x avant n converge $\mathbb{P}_\lambda$ presque sûrement vers $\frac{\mu(y)}{\mu(x)}$.*

Il suffit de prendre $f = \mathbf{1}_y$ et $g = \mathbf{1}_x$.

Théorème 2.44 *Soit X une chaîne irréductible récurrente positive de probabilité invariante π. Soit $f \in L^1(\pi)$. Alors presque sûrement*

$$\lim \frac{1}{n}\sum_{k=0}^{n-1} f(X_k) = \pi(f). \tag{4}$$

Preuve : Si la chaîne est irréductible récurrente positive, $\mu(E) < \infty$, et la probabilité π existe. Il suffit d'appliquer le théorème précédent avec $g = 1$. •

Proposition 2.45 *(i) Soit X une chaîne irréductible récurrente positive de probabilité invariante π. Alors pour tout $x \in E$, le nombre de visites en x avant n converge $\mathbb{P}_\lambda$ presque sûrement*

$$\lim \frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}_{\{x\}}(X_k) = \pi(x).$$

(ii) Si X une chaîne irréductible récurrente nulle. Alors pour tout $x \in E$, presque sûrement :

$$\lim \frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}_{\{x\}}(X_k) = 0.$$

Preuve : (i) C'est le théorème 2.44 avec $f = \mathbf{1}_x$.

(ii) On reprend encore, au lieu de π, une mesure invariante μ. Soit F fini, $F \in \mathcal{E}$, $f = \mathbf{1}_x$ et $g = \mathbf{1}_F$. On obtient pour tout F

$$\lim_n \frac{\sum_{k=0}^{n-1} \mathbf{1}_x(X_k)}{\sum_{k=0}^{n-1} \mathbf{1}_F(X_k)} = \frac{\mu(x)}{\mu(F)}.$$

Soit alors une suite croissante d'ensembles finis F_j de réunion E, pour tout j :

$$\frac{\sum_{k=0}^{n-1} \mathbf{1}_x(X_k)}{\sum_{k=0}^{n-1} \mathbf{1}_{F_j}(X_k)} = \frac{\mu(x)}{\mu_x(F_j)}$$

pour tout F_j si grand soit il. Si $\mu_x(E)$ est infini, cette limite est donc nulle. •

Exercices
1. Si X une chaîne irréductible récurrente positive de probabilité invariante π. Alors

$$\lim \frac{1}{n} \sum_{k=0}^{n-1} Q_k(y,x) = \pi(x).$$

2. Si X une chaîne irréductible récurrente nulle. Alors

$$\lim \frac{1}{n} \sum_{k=0}^{n-1} Q_k(y,x) = 0.$$

2.12 Marches aléatoires

Soit ν une probabilité sur $\mathbb{R}^k$, $(Y_n, n \geq 1)$ une suite de variables aléatoires indépendantes de loi ν, définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. On pose $X_n = \sum_{i=1}^n Y_i$, alors

Exercice : X est une chaîne de Markov de condition initiale 0, de transition Q définie par $Q(x,B) = \nu(B - x)$.
La suite $(X_n + x, n \geq 1)$ est pour tout $x \in \mathbb{R}^k$, une version de la chaîne de Markov de condition initiale x et de transition Q. Une telle chaîne est appelée **marche aléatoire**.

Proposition 2.46

$$\mathbb{P}(X_n \in B) = Q^n(0,B) = \nu^{*n}(B),$$

$$U(0,B) = U(x, x+B), \forall x \in E,$$

en particulier $U(x,x) = U(0,0)$.

Preuve : Notons que $Q^n(0,.)$ est la loi de X_n sachant $X_0 = 0$, c'est à dire que $Q^n(0,B) = \mathbb{P}(Y_1+\cdots+Y_n \in B) = \nu^{*n}(B)$, c'est la convolution de n variables aléatoires indépendantes.

Puis $U(0,B) = \sum_n Q^n(0,B) = \sum_n \nu^{*n}(B) = \sum_n Q^n(x,x+B)$ et ceci pour tout x de E.

•

Quelques conséquences :
. Si 0 est récurrent, tout point x est récurrent puisque pour tout $x \in E, U(0,0) = U(x,x)$.
. Si ν a une moyenne m, par la loi des grands nombres, X_n/n tend $\mathbb{P}$ presque sûrement vers m quand n tend vers l'infini. Il est clair que pour m non nul, X_n ne revient une infinité de fois dans aucun ensemble borné de $\mathbb{R}^k$; donc 0 n'est pas récurrent ici.

Etudions le cas où ν est une probabilité discrète concentrée sur Z : c'est à dire que l'on choisit $E = \mathbb{Z}$ comme espace d'états. Posons $S = \{x \in \mathbb{Z}^* : \nu(x) > 0\}$ le support de ν éventuellement privé de 0 et supposons S non vide. Le cas serait le même avec E dénombrable.

Définition 2.47 *Posons d le PGCD de S : on l'appelle la* **période** *de la marche aléatoire. Si* $d = 1$, *on dit que la chaîne est* **apériodique**.

Proposition 2.48 *Soit ν une loi sur $\mathbb{Z}$, de période d.*
Le nombre d est le plus grand entier $d \geq 1$ tel que ν est concentrée sur $d\mathbb{Z}$.
Soit $I = \{x \in \mathbb{Z} : \sum_{n>0} \nu^{*n}(x) > 0\}$. a) Si ν est concentrée sur $d\mathbb{N}, I$ contient $n_0 d + d\mathbb{N}$ pour un certain entier n_0 ; b) si ν n'est concentrée ni sur $d\mathbb{N}$ ni sur $-d\mathbb{N}, I = d\mathbb{Z}$.

Remarquons que si 0 est récurrent, I est la classe de récurrence de 0 : c'est l'ensemble des x auxquels 0 conduit.
De plus, $d = PGCD(S)$ dit que tout $x \in S$ est divisible par d. A priori $S \subset d\mathbb{Z}^*$.

Preuve : Il existe par définition du PGCD d'après l'identité de Bezout, un entier $p \geq 1$, des entiers $a_1, \cdots, a_p$ dans S, et des entiers relatifs $u_1, \cdots, u_p$ tels que $a_1u_1 + \cdots + a_pu_p = d$. Si l'on pose pour i entre 1 et $p, u_i^+ = max(u_i, 0), u_i^- = -min(u_i, 0)$, on vérifie que $a_1u_1^+ + \cdots + a_pu_p^+ = d + a_1u_1^- + \cdots + a_pu_p^-$.
On considère alors l'ensemble $I = \{x \in \mathbb{Z} : \sum_{n>0} \nu^{*n}(x) > 0\} = \{x \in \mathbb{Z} : 0 \to x\}$.

Le fait que pour tout $y, Q^n(y, y+x) = \nu^{*n}(x),\ U(y, y+x) = \sum_{n>0} Q^n(y, y+x) = \sum_{n>0} \nu^{*n}(x)$, montre que si $x \in I, \sum_{n>0} \nu^{*n}(x) > 0$, donc tout y de $\mathbb{Z}$ conduit à $x+y$. En particulier, si $y \in I$, 0 conduit à y et si $x \in I$, 0 conduit à $x+y$ par transitivité et $x+y \in I$.

De plus, il est clair que $S \subset I$, donc $a_j \in I$ pour tout j et $a_1u_1^+ + \cdots + a_pu_p^+ \in I, a_1u_1^- + \cdots + a_pu_p^- \in I$. Il existe donc dans I deux nombres a et b qui, comme les a_i, sont nécessairement multiples de d, tels que $a + d = b, a = kd, b = (k+1)d$.

. Si ν est à support dans $\mathbb{N}^*$, a et b sont positifs. Pour tout $n, na, (n-1)a+b, \cdots, nb \in I$ et on remarque que deux de ces entiers consécutifs diffèrent de d :

$$(n-l)a + lb - [(n-l+1)a + (l-1)b)] = b - a = d.$$

Pour $n > k, (n+1)a = (n+1)kd < n(k+1)d = nb$, et deux séries consécutives se chevauchent :
pour $n = k+1$, on obtient les éléments de I $(k+1)kd, ((k+1)k+1)d, \cdot, (k+1)^2 d$,
pour $n = k+2$, on obtient $(k^2+2k)d, (k^2+2k+1)d, \cdots, (k+1)(k+2)d \in I$, avec $k^2+2k < (k+1)^2$,
et coetera et ainsi I contient tous les entiers multiple de d à partir de $n_0 = (k+1)k$.

Même chose si a et b sont négatifs (le support de μ dans les entiers négatifs).

. Sinon, il y a plusieurs cas : si a et b ne sont pas de même signe, comme ils sont multiples de d et différents de d, par exemple si $a < 0$ et $b \geq 0, a = -kd, b = (-k+1)d$ avec $d \geq kd$, soit $k = 1$, donc $a = -d$, et $b = 0 \in I$, ce qui prouve que I contient tous les entiers jd pour $j \leq 0$. Mais, comme μ n'est pas concentrée sur les entiers négatifs, il existe $n_1 d > 0$ dans S, multiple de d son PGCD. Donc, $n_1 d + a = (n_1 - 1)d \in I$, et par récurrence $d \in I$. Alors I contient tout $\mathbb{Z}d$.

. Si a et b sont positifs, comme plus haut, on sait que I contient tous les entiers nd pour $n \geq n_0$. Mais S contient nécessairement un entier négatif, soit $-n_1 d \in S \subset I$. Par addition, on a $\forall k \geq n_0,\ \forall j \geq 1\ (k - jn_1)d \in I$. Or, la famille d'entiers $(k - jn_1)$ lorsque $k \geq n_0, j \geq 1$ contient -1 ; donc $-d \in I$ et alors on est ramené au cas précédent.
(Soit i tel que $n_0 - in_1 < 0$: on ajoute 1 jusqu'à dépasser -1, ce qui est possible puisque $\mathbb{Z}$ est un corps archimédien). •

Proposition 2.49 *Soit μ une loi sur $\mathbb{Z}$ de moyenne m. Alors la marche aléatoire qui lui est associée est une chaîne de Markov récurrente pour $m = 0$, transiente sinon.*

Dans le cas récurrent, la classe récurrente de 0 est $d\mathbb{Z}$ où d est la période de μ.

Preuve : (i) Si $m \neq 0, \lim_n \frac{X_n}{n} = m$ et nécessairement X_n devient infini et ne peut revenir en 0.

(ii) Si $m = 0, \lim_n \frac{X_n}{n} = 0$ presque sûrement donc en probabilité et $\mathbb{P}(|X_n| \leq \varepsilon n)$ tend vers 1 quand n tend vers l'infini. C'est à dire que $\lim_n \mu^{*n}[-\varepsilon n, +\varepsilon n] = 1$.

Soit $x \in \mathbb{Z} : U(0,x) = \mathbb{P}_0(T_x < \infty)U(x,x) \leq U(x,x) = U(0,0)$, la première égalité est conséquence de la proposition 2.25. On somme cette inégalité pour $x \in [-\varepsilon n, +\varepsilon n]$:

$$U(0, [-\varepsilon n, +\varepsilon n]) \leq (2\varepsilon n + 1)U(0,0).$$

Par ailleurs,

$$\frac{1}{n}\sum_{p=1}^{n} \mu^{*p}[-\varepsilon p, +\varepsilon p] = \frac{1}{n}\sum_{p=1}^{n} \mathbb{P}(|X_p| \leq \varepsilon p) \leq \frac{1}{n}\sum_{p=1}^{n} \mathbb{P}(|X_p| \leq \varepsilon n) \leq \frac{1}{n} U(0, [-\varepsilon n, +\varepsilon n]).$$

Le membre de gauche tend par la loi de Césarot vers 1, et la limite de celui de droite est majoré par $2\varepsilon U(0,0), \forall \varepsilon$. Donc $U(0,0) = +\infty, 0$ est récurrent, comme tout autre point x puisque $U(0,0) = U(x,x)$.

(iii) I est la classe récurrente de $0, \mu$ est de support des deux côtés de 0 puisque centrée, et le b) de la proposition 2.48 permet de conclure.

•

Remarquons que si E est dénombrable et sous ensemble de $\mathbb{R}$ (soit stable par addition), la mesure de comptage sur $E, \mu(n) = 1, \forall n \in E$, est invariante pour toute marche aléatoire à valeurs sur E. Si E est infini, $\mu(E) = +\infty$, on a donc dans le cas récurrent une récurrence nulle.

3 Processus de Poisson

(cf. BCD, proposition 10.6 page 154)

3.1 Introduction

Définition 3.1 *Une première définition est la suivante : un processus de Poisson est un processus à valeurs entières noté N vérifiant la propriété suivante :*

$$
\begin{aligned}
&\forall n \geq 1, \forall (t_1, \cdots, t_n) \in (\mathbb{R}^+)^n, t_1 \leq t_2 \leq \cdots \leq t_n, \\
(5) \qquad &(N_{t_1}, \cdots, N_{t_n} - N_{t_{n-1}}) \textit{ sont des variables aléatoires indépendantes,} \\
&\textit{pour tout } i, , N_{t_i} - N_{t_{i-1}} \textit{ suit une loi de Poisson de paramètre } \lambda(t_i - t_{i-1}).
\end{aligned}
$$

*Ce paramètre λ s'appelle l'***intensité** *du processus.*

C'est dire que la loi de tout $n-$uplet $(N_{t_1}, \cdots, N_{t_n})$ est l'image de la loi $\otimes_{i=1}^n \mathcal{P}(\lambda(t_i - t_{i-1}))$ par l'application de $\mathbb{R}^n$ dans lui-même : $x \mapsto x_1, x_1 + x_2, \cdots, \sum x_i$. Comme on l'a fait pour les chaînes de Markov, l'existence d'un tel processus est assurée par une construction canonique et le théorème de Kolmogorov. On choisit pour espace $\Omega = \mathbb{N}^{\mathbb{R}^+}$, muni de la tribu cylindrique $\mathcal{A}$, et on sait qu'il existe sur l'espace mesurable $(\Omega, \mathcal{A})$ une probabilité $\mathbb{P}$ telle que sous $\mathbb{P}$ le processus canonique $(X_t(\omega) = \omega_t)$ est un processus de Poisson au sens ci-dessus.

De fait, avec des hypothèses minimales que l'on va prendre comme définition ci-dessous, on va voir que les propriétés énoncées ci-dessus seront vérifiées. On aura de plus quelques propriétés des "temps de saut", c'est à dire des instants où N_t s'incrémente de 1.

Définition 3.2 *On appelle "processus de Poisson" une famille de variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P}), (N_t, t \in \mathbb{R}^+)$ à valeurs entières telle que*

(i) $\{\omega : t \mapsto N_t(\omega)$ càd croissante $N_t - N_{t^-} = 0$ ou $1\}$ est de probabilité 1.

(ii) $\forall t \geq 0, \forall s \geq 0, N_{t+s} - N_t$ est indépendante de $N_u, \forall u \leq t$.

(iii) La loi de $N_{t+s} - N_s$ ne dépend pas de s, on la note μ_t.

Exercice : Les propriétés (ii) et (iii) sont vérifiées par un processus tel celui défini en (5).

Les paragraphes suivants vont permettre de montrer la réciproque, c'est à dire que la définition 3.2 implique (5) en ajoutant l'hypothèse

H : $\mathbb{P}\{\forall t \in \mathbb{R}^+, N_t = 0\} < 1$, c'est à dire que l'on peut décoller de 0.

Dans les paragraphes suivants, on suppose maintenant que le processus N vérifie la définition 3.2 et l'hypothèse **H**.

3.2 Premières propriétés

1. *Exercice* : Si N est un processus de Poisson, $t \mapsto M_t = N_t - N_0$ est aussi un processus de Poisson issu de 0 : $M_0 = 0$. en effet, on peut vérifier que les trois propriétés sont préservées car $M_{t+s} - M_t = N_{t+s} - N_t$.

A partir de maintenant, on prendra en général $N_0 = 0$*, c'est à dire un processus de Poisson issu de* 0.

2. *Exercice* : $\forall n, \forall t_1 \leq, \cdots, t_n$, la loi de $(N_{t_1}, \cdots, N_{t_n})$ est l'image de $\mu_{t_1} \otimes \cdots \otimes \mu_{t_n - t_{n-1}}$ par l'application $f : x \mapsto (x_1, x_1 + x_2, \cdots, x_1 + \cdots + x_n)$.
C'est une conséquence du fait que $(N_{t_1}, \cdots, N_{t_1}) = f(N_{t_i - t_{i-1}})$ où les $N_{t_i - t_{i-1}}$ sont des variables aléatoires indépendantes de loi $\otimes_i \mu_{t_i - t_{i-1}}$.

3.

Lemme 3.3 QC 8
Soit N un processus de Poisson (au sens de la définition 3.2) issu de 0 *vérifiant l'hypothèse* **H** *et soit $G_t(u) = E(u^{N_t})$ la fonction génératrice de N_t. Alors*

1. $\forall u \in]0,1[, t \mapsto G_t(u)$ est càd et il existe une fonction C strictement positive sur $]0,1[$ telle que $G_t(u) = e^{-C(u).t}$.

2. Si $N_0 = 0$, il existe $\lambda > 0 : \mathbb{P}(N_t - N_0 = 0) = e^{-\lambda t}$.

Preuve : (i) l'application $t \mapsto N_t$ est continue à droite presque sûrement ; puis $0 < u < 1$ implique $0 < u^{N_t} < 1$ et $t \mapsto u^{N_t}$ est continue à droite presque sûrement et bornée ; le théorème de Lebesgue montre alors que $t \mapsto G_t(u)$ est continue à droite.

(ii) $G_{t+s}(u) = E[u^{N_{t+s} - N_s} u^{N_s} = G_t(u) G_s(u)$ par (ii) et (iii) de la définition 3.2. De plus N est issu de 0 montre que $G_0(u) = 1$.

(iii) Par un théorème classique d'analyse, les trois faits : continuité à doite, $G_0(u) = 1$ et $G_{t+s} = G_t G_s$ montre l'existence d'une constante positive $C(u)$ telle que $g_t(u) = e^{-C(u)t}$.

L'application $f : t \mapsto \mathbb{P}(N_t = 0)$ vérifie $f(0) = 1$ et $f(s+t) = f(s) f(t)$ par le même argument que pour 1. De plus elle est continue à droite : en effet, puisque N est croissant, si une suite s_n décroit vers t, la suite $f(s_n)$ est croissante vers $\mathbb{P}(\cup_n \{N_{s_n}) = 0\})$. Mais N est à valeurs entière et càd : qd s_n tend vers t, sur lévénement $(N_t = 0)$, il existe n à partir duquel $N_{s_n} = 0$, et donc $\{N_t = 0\} = \cup_n \{N_{s_n}) = 0\}$, d'où la continuité à droite. •

Corollaire 3.4 H *implique* $\mathbb{P}(\overline{\lim}(N_t = 0)) = 0$.

Preuve : $\forall k \geq 0, \mathbb{P}(N_k = 0) = e^{-\lambda k}, \sum_k e^{-\lambda k} < \infty$ et le lemme de Borel-Cantelli permettent de conclure. •

Avant de donner la loi de N_t, on va s'occuper des temps de saut et des temps d'arrêt.

3.3 Temps de saut, temps d'arrêt

Rappel : $N_t - N_{t-} = 0$ *ou* 1.
On note T_i les temps de saut successifs du processus. Comme $N_0 = 0, T_1(\omega) = \inf\{t > 0, N_t(\omega) = 1\}$; $T_n(\omega) = \inf\{t > 0, N_t(\omega) = n\}$. C'est une suite croissante stricte (car $N_.$ est cà d). On note $S_n = T_n - T_{n-1}, S_1 = T_1$ et on pose $S_0 = 0$.

Remarque 3.5 $N_t = \sum_{n>0} \mathbf{1}_{\{T_n \leq t\}}$ *: c'est le nombre de sauts avant t.*

Comme on l'a vu pour les martingales à temps discret, on va introduire la notion de filtration mais cette fois-ci à temps continu.

Définition 3.6 *On appelle "filtration naturelle" du processus N la suite croissante de tribus $\mathcal{F}_t = \sigma\{N_s, s \leq t\}$, de fait on complète chaque tribu par l'ensemble des négligeables $\mathcal{N}$ et on peut montrer que la suite est continue à droite (cf[7]) :*

$$\mathcal{F}_t = \cap_\varepsilon(\sigma\{N_s, s \leq t + \varepsilon\} \vee \mathcal{N}.$$

On dit qu'une filtration $\mathcal{F}$ est cà d dès que

$$\cap_{\varepsilon>0}\mathcal{F}_{t+\varepsilon} = \mathcal{F}_t.$$

C'est le cas ici par construction, simplement.

Définition 3.7 *On appelle $\mathcal{F}$-temps d'arrêt une variable aléatoire T à valeurs positives telle que pour tout $t \geq 0, \{T \leq t\} \in \mathcal{F}_t$.*

Par exemple :

Proposition 3.8 *Les temps de sauts de N sont des $\mathcal{F}$-temps d'arrêt.*

Preuve : *en exercice* par récurrence et en utilisant la remarque

$$\{T_n > t\} = \{T_{n-1} > t\} \cup \{T_{n-1} \leq t < T_n\} = \{T_{n-1} > t\} \cup \{N_t = n - 1\}.$$

D'abord $\{T_1 > t\} = \{N_t = 0\} \in \mathcal{F}_t$ par définition de la filtration.
Supposons par récurrence que T_{n-1} est un $\mathcal{F}$-temps d'arrêt ; soit $t > 0$,
$\{T_n > t\} = \{T_{n-1} > t\} \cup \{T_{n-1} \leq t < T_n\} = \{T_{n-1} > t\} \cup \{N_t = n - 1\} \in \mathcal{F}_t$. •

Proposition 3.9 *Sous l'hypothèse* **H** *ces temps d'arrêt sont presque sûrement finis.*

Preuve : par récurrence *en exercice.*
$\mathbb{P}\{T_1 > t\} = \mathbb{P}\{N_t = 0\} = e^{-\lambda t}$ ce qui montre que T_1 est de loi exponentielle de paramètre λ donc presque sûrement fini.
Supposons par récurrence que T_{n-1} est presque sûrement fini. Ainsi, $\cap_{k\geq 1}(T_{n-1} > k) = \emptyset$. Alors :

$$\cap_{k\geq 1}(T_n > k) = \cap_{k\geq 1}(\{T_{n-1} > k\} \cup \{N_k = n-1\})$$

qui est une limite décroissante. Comme $\cap_{k\geq 1}(T_{n-1} > k) = \emptyset$, cette limite est $\cap_{k\geq 1}\{N_k = n-1\} = \{N_k = n-1 \forall k \geq 1\}$ ce qui contredit **H** en prenant le processus de Poisson $N'_k = N_k - (n-1)$. •

La définition suivante est analogue à celle déjà vue en temps discret.

Définition 3.10 *Pour tout $\mathcal{F}$-temps d'arrêt T, on définit la tribu des événements antérieurs à T :*

$$\mathcal{F}_T = \{A \in \mathcal{A}, \forall t \geq 0, A \cap (T \leq t) \in \mathcal{F}_t\}.$$

Proposition 3.11 *Si T est un $\mathcal{F}$-temps d'arrêt, $\mathcal{F}_T$ est une tribu et T est $\mathcal{F}_T$-mesurable.*

Preuve en exercice (tout à fait analogue au cas discret).

Lemme 3.12 *d'approximation fondamental : Soit T un $\mathcal{F}$-temps d'arrêt et*

$$T_n = \sum_{k=0}^{n2^n-1} \frac{k+1}{2^n} \mathbf{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n}[}(T) + n\mathbf{1}_{[n,\infty[}(T), n \geq 1.$$

Alors,

(i) $(T_n, n \geq 1)$ *est une suite de temps d'arrêt tels que* $\forall n, T_n \leq n$,

(ii) $\forall \omega \in \Omega, \lim_n T_n(\omega) = T(\omega)$.

(iii) $\forall \omega \in \{T < \infty\}, \exists N(\omega)$ *tel que après* $N(\omega)$, *la suite* $(T_n(\omega))$ *est décroissante.*

Preuve : par construction, $T_n \leq n$; par ailleurs si $T \leq n$, $T_n \geq T$, et remarquons que $0 \leq T_n - T \leq 2^{-n}$.

(i) Soit $t \geq n, \{T_n \leq t\} = \{T_n \leq n\} = \Omega \in \mathcal{F}_t$. Et si $t < n$, et $k = [t2^n], k2^{-n} \leq t < (k+1)2^{-n}$. Alors, $\{T_n \leq t\} = \{T_n \leq k2^{-n}\} = \{T < k2^{-n}\} \in \mathcal{F}_{k2^{-n}} \subset \mathcal{F}_t$.

(ii) Sur l'événement $(T = \infty), T_n = n$ tend vers l'infini.
On fixe $\omega \in \{T < \infty\}$, pour n assez grand tel que $T(\omega) < n$, $(T_n - T)(\omega) \leq 2^{-n}$ qui tend vers 0.

(iii) Soit donc pour $T(\omega) < \infty$, un entier $N(\omega) > T(\omega)$, à partir de cet entier, la suite est décroissante : en effet, soit $k_n(\omega) = [T(\omega)2^n]$: $T_n(\omega) = (k_n+1)2^{-n}$. Ainsi $\frac{2k_n}{2^{n+1}} \leq T(\omega) < \frac{2(k_n+1)}{2^{n+1}}$ implique que $T_{n+1}(\omega) = (2k_n+1)2^{-(n+1)} \leq T_n(\omega)$.

•

Proposition 3.13 *Soit sur un espace de probabilité filtré* $(\Omega, \mathcal{F}_t; t \in \mathbb{R}^+, \mathbb{P})$ *un processus stochastique càd* $\mathcal{F}$*-adapté* $(X_t, t \in \mathbb{R}^+)$ *à valeurs dans* E *sous ensemble discret de réels ;* T *un* $\mathcal{F}^X$*-temps d'arrêt presque sûrement fini. Alors* $\omega \mapsto X_{T(\omega)}(\omega)$ *est* $\mathcal{F}_T$*-mesurable.*

Preuve : On utilise le lemme d'approximation précédent. Comme X est càd, et que T_n décroit vers T, la suite X_{T_n} tend presque sûrement vers X_T. Or, X est à valeurs discrètes :

$$\forall \omega \in \Omega, \exists N(\omega), \forall n \geq N(\omega), X_{T_n}(\omega) = X_T(\omega).$$

Le but est démontrer que pour tout x de E et tout réel positif t,
$\{X_T = x\} \cap \{T \leq t\} \in \mathcal{F}_t$. Or :

. $\{T \leq t\} = \{T < t\} \cup \{T = t\}$; $\{T = t\} = \{T \leq t\} \cap_{n>0} \{T > t - 1/n\} \in \mathcal{F}_t$.

. $\{X_T = x\} \cap \{T = t\} = \{X_t = x\} \cap \{T = t\} \in \mathcal{F}_t$.

.$\{X_T = x\} \cap \{T = t\} = \cup_{k\geq 0} \cap_{p\geq k} \{X_{T_p} = x\} \cap \{T_p = t\}$ car T_p décroit vers T et pour p assez grand, $X_{T_p} = X_T$.

a) si $p < t$, $T_p \leq p < t$ et
$\{X_{T_p} = x\} = \cup_{k=0}^{p2^p-1} \{X_{2^{-p}(k+1)} = x\} \cap \{2^{-p}k \leq T < 2^{-p}(k+1)\} \cup \{X_p = x\} \cap \{T \geq p\} \in \mathcal{F}_t$.

b) si $p \geq t$,
$\{X_{T_p} = x\} \cap \{T < t\} = \cup_{t>(k+1)2^{-p}} \{X_{2^{-p}(k+1)} = x\} \cap \{2^{-p}k \leq T < 2^{-p}(k+1)\} \in \mathcal{F}_t$
car aucun indice ne dépasse t. Ainsi X_T est-il $\mathcal{F}_T$ mesurable. •

3.4 Processus de Markov, processus à accroissements indépendants

Définition 3.14 *Un* **processus à accroissements indépendants** *(PAI) est un processus stochastique qui vérifie la propriété (ii) de la définition 3.2.*
Un **processus à accroissements stationnaires** *(PAS) est un processus stochastique qui vérifie la propriété (iii) de la définition 3.2.*
On note PAIS un processus qui vérifie les deux propriétés.

On appelle **processus de Markov** *sur l'espace de probabilité filitré* $(\Omega, \mathcal{F}_t, \mathbb{P})$, *un processus* $(X_t, t \in \mathbb{R}^+)$ *à valeurs réelles tel que pour tout borélien* B, *et tout couple de réels positifs* (t, s),

$$\mathbb{P}(X_{t+s} \in B/\mathcal{F}_t) = \mathbb{P}(X_{t+s} \in B/X_t) \text{ noté } \mathbb{P}_{X_t}(X_s \in B).$$

On dit qu'il est **homogène** *si de plus cette loi conditionnelle ne dépend pas de* t.

Exemple : un processus de Poisson est un PAIS.

Proposition 3.15 *(i) Un PAI est un processus de Markov.*

(ii) Un PAIS est de plus homogène.

Preuve : (i) Soit un couple de réels positifs $(t, s), A \in \mathcal{F}_s^X, f$ une fonction borélienne bornée :

$$E[f(X_{s+t})\mathbf{1}_A] = E[f(X_{s+t} - X_s + X_s)\mathbf{1}_A] =$$

$$\int_A \int f(x + X_s) d\mu_{s,t}(x) d\mathbb{P}$$

où $\mu_{s,t}$ est la loi de l'accroissement $X_{s+t} - X_s$, donc $E[f(X_{s+t})/\mathcal{F}_s^X] = \int f(x + X_s) d\mu_{s,t}(x)$ qui est $\sigma(X_s)$-mesurable, donc c'est aussi $E[f(X_{s+t})/X_s]$.

(ii) Si de plus le processus est stationnaire, $\mu_{s,t}$ ne dépend que de t, et donc aussi la loi conditionnelle. •

Remarque : La loi conditionnelle est la convolée $\tau_x * (\mu_t)$ sur $\{X_s = x\}$ où τ_x est la translation de x sur $\mathbb{R}$.

Corollaire 3.16 *Un processus de Poisson est un processus de Markov homogène.*

Définition 3.17 *On note θ_s l'opérateur de translation sur les trajectoires de l'espace canonique :*

$$\theta_s : E^{\mathbb{R}^+} \to E^{\mathbb{R}^+}, \omega \mapsto (t \mapsto \omega(t + s)).$$

Si T est un temps d'arrêt, on a encore l'opérateur θ_T.

Définition 3.18 *Un processus de Markov X est* **fort** *si pour tout temps d'arrêt T presque sûrement fini et quelle que soit Y mesurable bornée,*

$$(*) \qquad [Y \circ \theta_T / \mathcal{F}_T^X] = E[Y \circ \theta_T / X_T].$$

Si de plus le processus est homogène, $E[Y \circ \theta_T / X_T] = g(X_T)$ où $g(x) = E[Y/X_0 = x]$.

Théorème 3.19 *Un processus de Markov X à valeurs dans E discret, càd et homogène est un processus de Markov fort.*

Preuve : il suffit de montrer le théorème pour des variables aléatoires cylindriques, soit $Y = \phi(X_{t_1}, \cdots, X_{t_n})$ avec ϕ borélienne bornée, et d'utiliser le théorème des classes monotones, d'où le plan de la preuve en notant $\mathcal{L}$ l'ensemble de Y vérifiant (*).
(i) les constantes sont mesurables bornées et vérifient (*) ; (*) est une relation linéaire, donc $\mathcal{L}$ est un espace vectoriel contenant les constantes.
(ii) Soit une suite strictement croissante $0 < t_1 < \cdots < t_p$ et des fonctions f_i mesurables bornées. On pose $Y = \Pi_i f_i(X_{t_i})$. Soit $A \in \mathcal{F}_T^X : E[Y \circ \theta_T \mathbf{1}_A] = E[\Pi_i f_i(X_{t_i+T})\mathbf{1}_A]$.
On introduit $T_n = \frac{l+1}{2^n}$ si $l = [T2^n]$ qui est une suite de temps d'arrêt décroissant vers T. Le théorème de convergence dominée et le fait que X soit à valeurs discrètes montrent que

$$E[Y \circ \theta_T \mathbf{1}_A] = \lim_n E[\Pi_i f_i(X_{t_i+T_n})\mathbf{1}_{A\cap(T<n)}].$$

(c'est le pus long)
(iii)$\mathcal{L}$ est stable par limite croissante (Tonelli)
Conclusion : $\mathcal{L}$ est l'ensemble de toutes les variables aléatoires boréliennes bornées et X est un Markov fort. $\bullet$

Notation Soit $T : \Omega \to \mathbb{R}^+; \omega \mapsto \inf\{t > 0, X_t(\omega) \neq X_0(\omega)\}$.

Proposition 3.20 *Soit* $(X_t, t \geq 0)$ *un processus stochastique à valeurs dans* E*, càd,* T *est un* $\mathcal{F}^X$*-temps d'arrêt.*

Preuve : On remarque que $\{\omega : T(\omega) > t\} = \{\omega \ : \ X_u(\omega) = X_0(\omega) \ \forall u \leq t\}$. Comme X est cà d, cet événement est $\cap_{t_n \leq t, t_n \in Q}\{X_{t_n})X_0\}$ qui est bien dans $\mathcal{F}_t$. $\bullet$

Théorème 3.21 *Soit* $(X_t, t \geq 0)$ *un processus stochastique càd, à valeurs dans* E *discret contenu dans* $\mathbb{R}$*, de Markov, homogène et tel que* $X_t - X_0$ *est indépendant de* X_0*. Alors*

(i)) $T = \inf\{t > 0, X_t \neq X_0\}$ *suit une loi exponentielle.*

(ii) T *et* $X_T - X_0$ *sont deux variables aléatoires indépendantes.*

Preuve :les hypothèses donnent que T est indépendant de X_0
(i) Soit $F : t \mapsto \mathbb{P}(T \geq t)$, on a $F(0) = 1$, X càd entraine que F l'est aussi. De plus $F(s+t) = \mathbb{P}(T > s+t) = \mathbb{P}(T > s \text{ et } T \circ \theta_s > t)$ car sur $(T > s), X_s = X_0$ et $T \circ \theta_s > t = \inf\{u > 0, \ X_{s+u} \neq X_s = X_0\}$ et réciproquement.

Donc par la propriété de Markov

$$F(t+s) = E[\mathbf{1}_{T>s}\mathbf{1}_{T>t} \circ \theta_s] = F(s)F(t)$$

puisque $\mathbb{P}(T > t / X_0 = x) = \mathbb{P}(T > t)$ par l'indépendance. Et on peut conclure (i).
(ii) $\mathbb{P}(T > t, X_T - X_0 = x) = E[\mathbf{1}_{T>t} E[\mathbf{1}_{X_T - X_0 = x} / \mathcal{F}_t]]$ qui par la propriété de Markov est $E[\mathbf{1}_{T>t} \mathbb{P}_{X_t}\{X_T - X_0 = x\}]$. Or sur $(T > t)$, on a $X_t = X_0$ et donc $\mathbb{P}(T > t, X_T - X_0 = x) = F(t)\mathbb{P}_{X_0}\{X_T - X_0 = x\}$

$\bullet$

3.5 Propriétés fondamentales du processus de Poisson

On a maintenant tous les outils pour montrer les autres propriétés du processus de Poisson. En particulier, le théorème 3.19 dit qu'un processus de Poisson est un Markov fort.

Théorème 3.22 *Soit $(N_t, t \geq 0)$ un processus de Poisson et T un $\mathcal{F}^N$-temps d'arrêt preque sûrement fini. Alors, le processus $X : t \mapsto N_{t+T} - N_T$ est un processus de Poisson issu de 0 indépendant de $\mathcal{F}_T^N$.*

Preuve : (i) est vraie car X est aussi càd croissant, a les mêmes sauts que N.

(ii) et (iii) Pour tout s et tout t, $X_{s+t} - X_t = N_{s+t+T} - N_{t+T}$; la tribu $\mathcal{F}_s^X$ est engendrée par $(N_{u+T} - N_T, u \leq s)$ donc contenue dans $\mathcal{F}_{s+T}^N$. Il vient

$$E[f(N_{s+t+T} - N_{s+T})/\mathcal{F}_s^X] = E[E[f(N_{s+t+T} - N_{s+T})/\mathcal{F}_{s+T}^N]/\mathcal{F}_s^X].$$

On pose $Y = f(X_t - X_0)$; alors $f(N_{s+t+T} - N_{s+T}) = f(N_t - N) \circ \theta_{s+T} = Y \circ \theta_{s+T}$, et par la propriété de Markov $E[f(N_{s+t+T} - N_{s+T})/\mathcal{F}_{s+T}^N] = E[Y \circ \theta_{s+T}/\mathcal{F}_{s+T}^N] = E[Y \circ \theta_{s+T}/N_{s+T}]$ et par l'homogénéité $= E[Y/N_0] = E[Y]$ car $N_t - N_0$, donc Y, est indépendant de N_0 pour $t > 0$.

En conclusion, on obtient que $E[f(N_{s+t+T} - N_{s+T})/\mathcal{F}_s^X] = E[f(N_t - N_0)]$ d'où l'indépendance des accroissements, et le fait qu'ils sont de même loi que ceux de N. •

Théorème 3.23 *On pose $T_0 = 0$ et on suppose $N_0 = 0$. Les variables aléatoires $(S_n = T_n - T_{n-1})$ sont indépendantes et de loi exponentielle (λ).*

Preuve :

$$\mathbb{P}(T_n - T_{n-1} > t_{n-1}, \cdots, T_1 > t_1) = E[\Pi_{i=1}^{n-1} \mathbf{1}_{\{T_i - T_{i-1} > t_{i-1}\}} \mathbb{P}(T_n - T_{n-1} > t_{n-1}/\mathcal{F}_{T_{n-1}}^N)$$

On utilise le théorème 3.22 à savoir que $N_{t+T_{n-1}} - N_{T_{n-1}}$ est un processus de Poisson indépendant de $\mathcal{F}_{T_{n-1}}^N$:

$$\mathbb{P}(T_n - T_{n-1} > t_{n-1}/\mathcal{F}_{T_{n-1}}^N) =$$

$$\mathbb{P}[\{N_{t_{n-1}+T_{n-1}} - N_{T_{n-1}} = 0\}/\mathcal{F}_{T_{n-1}}^N] = \mathbb{P}\{N_{t_{n-1}} - N_0 = 0\} = \mathbb{P}\{T_1 > t_{n-1}\} = e^{-\lambda t_{n-1}}.$$

Puis on effectue une récurrence.

•

Le corollaire suivant achève de montrer que la définition 3.2 implique la définition 5.

Rappelons que l'on peut supposer $N_0 = 0$.

Corollaire 3.24 *La variable N_t suit une loi de Poisson de paramètre λt.*

Première **Preuve** avec une succession de trois lemmes .

Lemme 3.25 *La variable N_t est intégrable pour tout $t \in \mathbb{R}^+$.*

Preuve : $N_t = \sum_{n\in\mathbb{N}^*} \mathbf{1}_{\{T_n \le t\}}$, nombre de sauts avant t.

Donc en utilisant la loi exponentielle des inter-temps de saut, $E[N_t] = \sum_{n\ge 1} \int_{[0,t]} \frac{u^{n-1}}{(n-1)!} e^{-\lambda u} \lambda^n du$. Soit en permutant somme et intégrale, $E[N_t] = \lambda \int_{[0,t]} e^{\lambda u} e^{-\lambda u} du = \lambda t$. $\bullet$

Lemme 3.26 *On a $\mathbb{P}(N_t = 1) = \lambda t e^{-\lambda t}$, et donc le quotient $\frac{\mathbb{P}\{N_t=1\}}{t}$ converge vers λ lorsque t tend vers 0.*

Preuve : L'événement $(N_t = 1) = (T_1 \le t < T_2) = (S_1 \le t < S_1 + S_2)$. On peut donc calculer cette probabilité avec la loi du couple indépendant (S_1, S_2) de variables exponentielles.

Le calcul donne $\mathbb{P}(N_t = 1) = \lambda t e^{-\lambda t}$, d'où le résultat. $\bullet$

Lemme 3.27 *Le quotient $\frac{\mathbb{P}(N_t \ge 2)}{t}$ converge vers 0 lorsque t tend vers 0.*

Preuve : L'événement $(N_t \ge 2) = (T_2 \le t) = (S_1 + S_2 \le t)$. On peut donc calculer cette probabilité avec la loi de la somme des deux exponentielles, soit une loi $\Gamma(2, \lambda)$.
Ou, plus simplement, si l'on se rappelle que N_t est à valeurs entières, $\mathbb{P}(N_t = 0) = \mathbb{P}(T > t) = e^{-\lambda t}, \mathbb{P}(N_t \ge 2) = 1 - \mathbb{P}(N_t = 0) - \mathbb{P}(N_t = 1)$.
Le calcul donne $\mathbb{P}(N_t \ge 2) = 1 - e^{-\lambda t} - \lambda t e^{-\lambda t}$, d'où le résultat. $\bullet$

Preuve du corollaire : on rappelle que la fonction génératrice $G_t(u) = E[u^{N_t}]$ est de la forme $e^{-C(u)t}$. Donc,

$$C(u) = \lim_{t\to 0} \frac{1 - G_t(u)}{t}.$$

Par ailleurs, par définition, $G_t(u) = \sum_{n\ge 0} u^n \mathbb{P}(N_t = n)$, d'où

$$1 - G_t(u) = 1 - \mathbb{P}(N_t = 0) - u\mathbb{P}(N_t = 1) - \sum_{n\ge 2} u^n \mathbb{P}(N_t = n).$$

Le lemme 3.27 dit que la limite est nulle quand t tend vers 0. Le lemme 3.26 donne que la limite est $C(u) - (\lambda - \lambda u) = 0$. Ainsi on a identifié $C(u) = \lambda(1 - u)$; la fonction génératrice est $G_t(u) = e^{-\lambda(1-u)t}$, c'est à dire celle d'une loi de Poisson de paramètre λt. $\bullet$

On a une deuxième preuve plus rapide de ce corollaire

Corollaire 3.28 *La variable N_t suit une loi de Poisson de paramètre λt.*

Preuve : L'événement $(N_t \geq k) = (T_k \leq t)$. Or T_k est une somme de k variables indépendantes de loi exponentielles de paramètre λ, donc T_k suit une loi $\Gamma(k, \lambda)$. En déduire en **exercice** la loi de N_k.

$\bullet$

3.6 Construction du processus de Poisson

On revient enfin sur le lien entre la définition (5) et la définition (3.2) : on montre l'implication (3.2) implique (5). Sous la première définition, avec le paramètre $\lambda > 0$, et la loi μ_t loi de Poisson de paramètre λt, on construit par le théorème de Kolmogorov la probabilité $\mathbb{P}$ sur l'espace canonique $(\mathbb{N}^{\mathbb{R}^+}, \mathcal{A})$ tribu cylindrique, telle que pour tout n, pour tout n-uple de réels positifs ordonné $(t_1, \cdots, t_n), N_0 = 0$, la loi de $(N_{t_1}, N_{t_2} - N_{t_1}, \cdots, N_{t_n} - N_{t_{n-1}})$ où N_t est la t-ième projection canonique, est le produit tensoriel $\otimes_{i=1,n}\mu_{t_i - t_{i-1}}$. On a l'existence et l'unicité de $\mathbb{P}$. On a vu brièvement au début du chapitre que ceci est bien un processus de Poisson selon la définition (3.2). On le montre ici plus précisément bien que pas entièrement.

1. On admet qu'il existe une version de l'application

$$\Omega = \mathbb{N}^{\mathbb{R}^+} \rightarrow \mathbb{N} \; ; \; \omega \mapsto \omega(t)$$

telle que $t \mapsto \omega(t)$ soit càd (théorie générale des processus).

2. Par construction de $\mathbb{P}, N_{t+s} - N_s$ est indépendante de N_u pour tout $u \leq s$, et donc $N_{t+s} - N_s$ est indépendante de $\mathcal{F}_s^N$.

3. Par construction, $N_{t+s} - N_s$ est de loi μ_t qui ne dépend pas de s.

4. Il reste à prouver que les sauts sont de taille 1 ce que l'on a éludé au début du chapitre....

Tel que ce processus est construit, c'est un PAIS, donc un processus de MARKOV homogène, à valeurs dans les entiers. On peut appliquer le paragraphe 3.4 : si T est le premier temps de saut, T est de loi exponentielle de paramètre λ, et T et N_T sont indépendantes ($N_0 = 0$.) :

$$\mathbb{P}\{N_T \geq 2\} \cap \{T < t\}) = \mathbb{P}(N_T \geq 2)\mathbb{P}(T < t)$$

Puisque N est croissant, $\mathbb{P}(\{N_T \geq 2\} \cap \{T < t\}) \leq \mathbb{P}(N_t \geq 2)$.
Comme $\mathbb{P}(T < t) \neq 0, \mathbb{P}(N_T \geq 2) \leq \frac{\mathbb{P}(N_t \geq 2)}{\mathbb{P}(T<t)}$. Or, par construction dans ce paragraphe, la loi de N_t est connue : c'est μ_t, loi de Poisson de paramètre λ, d'où $\frac{\mathbb{P}(N_t \geq 2)}{\mathbb{P}(T<t)} = \frac{1-e^{-\lambda t}-\lambda t e^{-\lambda t}}{1-e^{-\lambda t}}$ de limite nulle quand t tend vers 0. Donc le premier saut est de taille 1.

On itère ensuite le raisonnement en utilisant le nouveau processus de Poisson $X_t = N_{t+T} - N_T$: on peut vérifier que $N'_t = N_{t+T} - N_T$ satisfait à la définition 3.1 à son tour ; on se sert de la même preuve que pour le théorème 3.22 qui utilise seulement que le processus que l'on décalle du temps T est un processus de Markov fort homogène.

3.7 Deuxième construction du processus de Poisson

On se donne ici $\lambda > 0$, une famille de variables indépendantes et de loi exponentielle de paramètre $\lambda, (S_n, n \geq 1)$. C'est à dire que l'on se donne la loi des durées inter-sauts. On pose $T_n = \sum_{i=1,n} S_i$ et

$$X_t = \sum_{n \geq 1} \mathbf{1}_{\{T_n \leq t\}}.$$

Remarquons que dans cette construction encore on l'égalité des événements $\{X_t = n\} = \{T_n \leq t < T_{n+1}\}$. Ce processus a pour valeurs le nombre de sauts avant l'instant t. Par construction, c'est un processus continu à droite, de sauts tous égaux à 1, et la propriété (i) est vraie. On déduit (ii) et (iii) du théorème suivant.

Théorème 3.29 *Soit* $(0 < t_1 < \cdots < t_n)$*; les variables aléatoires* $(X_{t_i} - X_{t_{i-1}}), i = 1, \cdots, n)$ *sont indépendantes et de loi de Poisson de paramètres* $\lambda(t_i - t_{i-1})$.

Preuve : On calcule la loi du n-uplet par une succession de lemmes.

Lemme 3.30 *Soit* $\Delta_n = \{t \in \mathbb{R}^n : 0 < t_1 < \cdots < t_n\}$. *Pour tout* n, *la loi de* $(T_1, \cdots, T_n)$ *est de densité* $f_n(t) = \mathbf{1}_{\Delta_n}(t)\lambda^n e^{-\lambda t_n}$.

Preuve :

•

Lemme 3.31

$$E[f(T_1,\cdots,T_n)\mathbf{1}_{\{X_t=n\}}] = \lambda^n e^{-\lambda t}\int_{\Delta_n} f(t_1,\cdots,t_n)\mathbf{1}_{[0,t]}(t_n)dt_1\cdots dt_n.$$

Preuve

•

Corollaire 3.32 *X_t est de loi de Poisson de paramètre λt.*

Preuve :

•

Corollaire 3.33 *La loi de $(T_1,\cdots,T_n)$ sachant $\{X_t=n\}$ est de densité $\mathbf{1}_{\Delta_n}(t)\frac{n!}{t^n}\mathbf{1}_{[0,t]}(t_n)$ (loi de Dirichlet d'ordre n).*

Preuve :

•

Lemme 3.34 *Soit* $(U_1, \cdots, U_n)$ *n variables aléatoires indépendantes uniformes sur* $[0, t]$ *;* $0 = t_0 < t_1 < t_2 < \cdots < t_p = t$ *;* $M_k = card\{i \leq n : U_i \in]t_{k-1}, t_k[\}$. *Alors, la loi de* $(M_1, \cdots, M_p)$ *est multinomiale* (p, q) *où* $q_k = \frac{t_k - t_{k-1}}{t}, k = 1, \cdots, p$.

Preuve

•

Lemme 3.35 *Soit* σ *variable aléatoire à valeurs dans l'espace des permutations* $\mathcal{S}_n$ *de loi uniforme et indépendante de* $(T_k, k \geq 1)$ *; on pose* $Z_k = T_{\sigma(k)}$. *Alors, la loi de* $(Z_1, \cdots, Z_n)$ *sachant* $\{X_t = n\}$ *est de densité* $\Pi_{j=1,n} \frac{1}{t} \mathbf{1}_{[0,t]}(t_j)$.

Preuve

•

Ces deux derniers lemmes permettent de montrer :

Corollaire 3.36 *La loi de* $(X_{t_1}, \cdots, X_{t_p} - X_{t_{p-1}})$ *sachant* $\{X_t = n\}$ *est multinomiale* (p, q) *où* $q_k = \frac{t_k - t_{k-1}}{t}$.

Preuve :

•

Preuve du théorème

•

References

[1] D. BAKRY, L. COUTIN, T. DELMOTTE : “Martingales et chaînes de Markov”, http://www.lsp.ups-tlse.fr.

[2] Ph. BARBE et M. LEDOUX : “Probabilités, de la licence à l'agrégation”, Belin, Paris, 1998.*en 14 exemplaires à la BU...*

[3] R. DURETT : “Probability : Theory and Examples”, Wadsworth and Brooks, Pacific Grove, california, 1991.

[4] W. FELLER : “An Introduction to Probability Theory and its Applications” I and II, John Wiley, New York, 1966.

[5] L. MAZLIAK, P. PRIOURET et P. BALDI : “Martingales et chaînes de Markov”, Hermann, Paris, 1998. *en 5 exemplaires à la BU...*

[6] J. NEVEU : “Martingales à temps discret”, Masson, Paris, 1972.

[7] P. PROTTER : “Stochastic Integration and Differential Equations”, Springer, Berlin, 1990.

[8] G. SAPORTA : “Probabilités, analyse des données et statistique”, ed. Technip, Paris, 1990.*en ? exemplaires à la BU...*

[9] P. S. TOULOUSE, Thèmes de probabilités et statistique, agrégation de mathématiques, Dunod, Paris, 1999.