Processus à temps discret

Léo Gayral
basées sur le cours de Christophe Sabot

2016/2017

20/09

1 Espérance conditionnelle dans le cas discret

— X une v.a. sur intégrable sur $(\Omega, \mathcal{F}, \mathbb{P})$ et $B \in \mathcal{F}$ tel que $\mathbb{P}(B) > 0$:
$\mathbb{E}[X|B] = \frac{\mathbb{E}[X \times \mathbb{1}_B]}{\mathbb{P}(B)}$
— Y une variable une variable à valeurs dans E dénombrable :
$E^* = \{y \in E, \mathbb{P}(Y = y) > 0\}$

Définition (Espérance conditionnelle de X selon Y)
$$\mathbb{E}[X|Y](\omega) = \begin{cases} \mathbb{E}[X|Y = y] & \text{si } Y(\omega) = y \in E^* \\ 0 & \text{si } Y(\omega) \notin E^* \end{cases}$$

Proposition
Soit Z une variable sur $\mathbb{R}^d$:
Z mesurable selon $\sigma(Y) \Leftrightarrow \exists C : (E, \mathcal{E}) \to \mathbb{R}^d, Z = C \circ Y$

Proposition (Caractérisation par la tribu $\sigma(Y)$)
$\mathbb{E}[\ |\ \mathbb{E}[X|Y]\ |\ ] \leq \mathbb{E}[\ |X|\ ]$
Pour toute variable Z $\sigma(Y)$-mesurable :
$\mathbb{E}[\ Z \times \mathbb{E}[X|Y]\ ] = \mathbb{E}[ZX]$
Cette relation caractérise l'espérance conditionnelle :
En particulier, si $\sigma(Y) = \sigma(Y')$ alors Y et Y' induisent la même espérance conditionnelle.

2 Définition générale et propriétés

Définition (Espérance conditionnelle de X selon $\mathcal{B}$)
Soient $\mathcal{B} \subset \mathcal{F}$ une sous-tribu sur $(\Omega, \mathcal{F}, \mathbb{P})$ et X une variable intégrable.
Il existe une unique v.a. $\mathcal{B}$-mesurable, notée $\mathbb{E}[X|\mathcal{B}]$, telle que :
$\forall B \in \mathcal{B}, \mathbb{E}[\ \mathbb{1}_B \times \mathbb{E}[X|\mathcal{B}]\ ] = \mathbb{E}[\mathbb{1}_B X]$
De façon équivalente on a :
$\forall Z \mathcal{B}\text{-mesurable, bornée}, \mathbb{E}[\ Z \times \mathbb{E}[X|\mathcal{B}]\ ] = \mathbb{E}[ZX]$

Définition (Cas positif non intégrable)
X une v.a. positive quelconque :
$\mathbb{E}[X|\mathcal{B}] = \lim\limits_{n\to\infty} \mathbb{E}[\min(X,n)|\mathcal{B}] \in [0;\infty]$
Cette espérance conditionnelle est caractérisée par la même relation que le cas intégrable.

Proposition (Propriétés élémentaires)
On a les points suivants :
- $\mathbb{E}[\bullet|\mathcal{B}]$ est linéaire
- $\mathbb{E}[\bullet|\mathcal{B}]$ est croissante (pour $X \geq_{\text{p.p.}} X'$)
- X $\mathcal{B}$-mesurable $\Rightarrow \mathbb{E}[X|\mathcal{B}] =_{\text{p.p.}} X$
- $\mathbb{E}[\ \mathbb{E}[X|\mathcal{B}]\ ] = \mathbb{E}[X]$
- $|\ \mathbb{E}[X|\mathcal{B}]\ | \ \leq_{\text{p.p.}} \mathbb{E}[\ |X|\ |\mathcal{B}]$

Définition (Espérance conditionnelle de X selon Y)
$\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]$

3 Règles de calcul

Proposition
Soient X et Y deux v.a. telles que X, Y et XY sont intégrables (ou positives).
Si Y est $\mathcal{B}$-mesurable, alors $\mathbb{E}[XY|\mathcal{B}] = \mathbb{E}[X|\mathcal{B}] \times Y$.

Proposition
Soient deux tribus $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{F}$:
$\mathbb{E}[\ \mathbb{E}[X|\mathcal{B}_1]\ |\mathcal{B}_2] = \mathbb{E}[\ \mathbb{E}[X|\mathcal{B}_2]\ |\mathcal{B}_1] = \mathbb{E}[X|\mathcal{B}_1]$

Proposition
Si X indépendant de $\mathcal{B}$, alors $\mathbb{E}[X|B] = \mathbb{E}[X]$.
Plus généralement : $\mathcal{B}_1$ indépendant de $\mathcal{B}_2 \Leftrightarrow \forall X$ intégrable, $\mathcal{B}_2$-mesurable, $\mathbb{E}[X|\mathcal{B}_1] = \mathbb{E}[X]$

27/09

4 Théorèmes usuels de l'espérance conditionnelle

Remarque
L'égalité caractéristique $\mathbb{E}[Z\mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[ZX]$ reste vrai dans un contexte assez général, tant que X et ZX sont intégrables (ou positifs).

Théorème (Lemme de Fatou)
Si $X_n \geq 0$ alors $\mathbb{E}[(\underline{\lim} X_n)|\mathcal{B}] \leq \underline{\lim}\, \mathbb{E}[X_n|\mathcal{B}]$.

Théorème (Convergence monotone)
Si $X_n \geq 0$ et $X_n \nearrow X$ alors $\lim\limits_{n\to\infty} \mathbb{E}[X_n|\mathcal{B}] = \mathbb{E}[X|\mathcal{B}]$

Théorème (Convergence dominée)
Si $|X_n| \leq Z$, Z intégrable et $X_n \to X$ alors $\mathbb{E}[X|\mathcal{B}] = \lim \mathbb{E}[X_n|\mathcal{B}]$ et $\mathbb{E}[\ |X|\ |\mathcal{B}] \leq \mathbb{E}[Z|\mathcal{B}]$

Théorème (Inégalité de Jensen)
Si X intégrable, $\phi : \mathbb{R} \to \mathbb{R}$ convexe et $\phi(X)$ intégrable alors $\phi(\mathbb{E}[X|\mathcal{B}]) \leq \mathbb{E}[\phi(X)|\mathcal{B}]$.

5 Cas des variables de carré intégrable

Proposition
Soit $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. $\mathbb{E}[X|\mathcal{B}]$ est la projection orthogonale de X sur l'espace $L^2(\Omega, \mathcal{B}, \mathbb{P})$.

6 Variables à densité

$X \in \mathbb{R}^p$, $Y \in \mathbb{R}^q$ des variables aléatoires, (X, Y) de densité $f_{X,Y}$. On définit $f_Y(y) = \int f_{X,Y}(x, y)dx$ la loi marginale de Y.
La densité conditionnelle de X selon Y est $f_X^{Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$ (défini presque partout).
Soit $h : \mathbb{R}^p \to \mathbb{R}$ telle que $h(X)$ intégrable. On définit $\phi(y) = \int h(x) f_X^{Y=y}(x)dx$.
Alors $\mathbb{E}[h(X)|Y] = \phi(Y)$.

7 Vecteurs Gaussiens

Soit $(X, Y_1 ... Y_n)$ un vecteur gaussien centré.
Soit (λ_i) telle que $\forall 1 \leq i \leq n, \mathrm{Cov}(X - \sum \lambda_i Y_i, Y_i) = 0$. Alors $\mathbb{E}[X|Y_1 ... Y_n] = \sum \lambda_i Y_i$.
$\sum\limits_{i=}^{n} \lambda_i Y_i$ est en fait la projection orthogonale de X sur $\mathrm{Vect}(Y_1 ... Y_n)$. Une telle famille (λ_i) existe toujours. Si la matrice de covariance de $(X, Y_1 ... Y_n)$ est définie positive, alors les coefficients (λ_i) sont uniques.
Plus généralement, si $h(X)$ intégrable, alors $\mathbb{E}[h(X)|Y_1 ... Y_n] = \int h(x) d\mu(x)$ où $\mu \sim \mathcal{N}\left(\sum \lambda_i Y_i, \mathrm{Var}(X - \sum \lambda_i Y_i)\right)$.

8 Martingales

Définition (Filtration sur $(\Omega, \mathcal{F}, \mathbb{P})$)
Suite croissante de sous-tribus $(\mathcal{F}_n)$

Définition (Processus à temps discret)
Suite de variables aléatoires (X_n)

Définition (Processus (X_n) adapté à la filtration $(\mathcal{F}_n)$)
$\forall n \in \mathbb{N}, X_n$ est $\mathcal{F}_n$-mesurable.

Définition (Filtration canonique de (X_n))
$\mathcal{F}_n = \sigma(X_1 ... X_n)$

Définition (Martingale selon $(\mathcal{F}_n)$)
(M_n) un processus réel adapté à $(\mathcal{F}_n)$ tel que $\forall n \in \mathbb{N}, \mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$ et M_n intégrable.
Si $\mathbb{E}[M_{n+1}|\mathcal{F}_n] \leq M_n$, (M_n) une surmartingale.
Si $\mathbb{E}[M_{n+1}|\mathcal{F}_n] \geq M_n$, (M_n) une sous-martingale.

Proposition
(M_n) une $(\mathcal{F}_n)$-martingale $\Rightarrow$ martingale selon la filtration canonique (vrai pour les sur/sous-martingales).

Proposition
Pour une martingale, $\forall m \geq n, \mathbb{E}[M_n|\mathcal{F}_n] = M_n$ ($\leq$ pour une surmartingale, $\geq$ pour une sous-martingale).

Proposition
Soit f convexe telle que les $f(M_n)$ sont intégrables.
Si (M_n) est une (sous-)martingale, alors $(f(M_n))$ est une sous-martingale.

04/10

Définition
(H_n) est $(\mathcal{F}_n)$-prévisible si H_{n+1} est $\mathcal{F}_n$-mesurable.

Proposition
(M_n) martingale, (H_n) prévisible bornée. Alors $(H*M)_n = \sum\limits_{i=1}^{n} H_i(M_i - M_{i-1})$ définit une martingale.
Informellement, $M_n - M_{n-1}$ représente le gain du jeu au temps n et H_n la stratégie suivie par un joueur (la somme qu'il mise). Dans ce cas, le gain total du joueur $H * M$ est une martingale, aucune stratégie n'est gagnante.

9 Temps d'arrêt

Définition (Temps d'arrêt sur $(\mathcal{F}_n)$)
$T : \Omega \to \overline{\mathbb{N}}$ vérifie $\forall n \in \mathbb{N}$, "$T \leq n$" $\in \mathcal{F}_n$ (ou "$T = n$" $\in \mathcal{F}_n$ de façon équivalente).

Remarque (Temps d'arrêt classique)
(X_n) adapté à $(\mathcal{F}_n)$ à valeurs dans $(E, \mathcal{E})$ et $A \in \mathcal{E}$.
$T = \inf\{n \in \mathbb{N}, X_n \in A\}$ est un temps d'arrêt.

Définition
T un temps d'arrêt : $\mathcal{F}_T = \{A \in \mathcal{F}, \forall n \in \mathbb{N}, A \cap$ "$T \leq n$" $\in \mathcal{F}_n\}$ définit la sous-tribu des évènements antérieurs à T.
Si T strictement constant, égal à n sur Ω, alors $\mathcal{F}_T = \mathcal{F}_n$.

Proposition
Soient S, T deux temps d'arrêt sur $(\mathcal{F}_n)$.
Alors $\sup(S, T)$, $\inf(S, T)$ et $S + T$ des temps d'arrêt.
Si $S \leq T$ partout sur Ω, alors $\mathcal{F}_S \subset \mathcal{F}_T$.

Théorème (Théorème d'arrêt, version faible)
(M_n) une martingale, T un temps d'arrêt.
$M_n^T := M_{\min(n,T)}$ est une martingale.
Si T borné ($\leq_{p.p.} M \in \mathbb{N}$ fixé), alors $\mathbb{E}[M_T] = \mathbb{E}[M_0]$.
Si T fini (presque partout) et (M_n^T) uniformément L^1 ($|M_n^T| \leq Z \in L^1$) alors $\mathbb{E}[M_T] = \mathbb{E}[M_0]$.

Proposition
Si on a deux temps d'arrêt $S \leq_{p.p.} T \leq_{p.p.} C \in \mathbb{N}$ bornés, alors $\mathbb{E}[M_T | \mathcal{F}_S] =_{p.p.} M_S$.

10 Exemple : Ruine du joueur

(X_n) iid valent 1 avec probabilité p et -1 avec probabilité $(1 - p)$.
Le joueur dispose de $s_0 \in \mathbb{N}^*$ initialement, gagne ou perd 1 à chaque tour, et possède donc $S_n = s_0 + \sum\limits_{j=1}^{n} X_j$ au tour n.
Le jeu s'arrête au temps T (pour la filtration canonique) dès que (S_n) atteint 0 ou $N > s_0$ pour la première fois.
Si $p = 0.5$ alors (S_n) est une martingale et $\mathbb{P}(S_T = N) = \frac{s_0}{N}$.
Si $p \neq 0.5$ alors $M_n = \left(\frac{1-p}{p}\right)^{S_n}$ une martingale et $\mathbb{P}(S_T = N) = \frac{1-\left(\frac{1-p}{p}\right)^{s_0}}{1-\left(\frac{1-p}{p}\right)^{N}}$

11 Théorème de convergence presque sûre

Théorème
(X_n) une sous-martingale telle que $\sup\limits_{n\in\mathbb{N}} \mathbb{E}[X_n^+] < +\infty$. Alors $X_n \xrightarrow{\text{p.p.}} X_\infty \in L^1$.

Corollaire
On peut de façon considérer une surmartingale et les X_n^- dans le théorème précédent.
Il en découle que pour (X_n) une surmartingale positive, $X_n \xrightarrow{\text{p.p.}} X_\infty \in L^1$ et $X_n \geq_{p.p.} \mathbb{E}[X_\infty|\mathcal{F}_n]$.

Remarque
Pour une martingale (X_n), $\mathbb{E}[X_0] = \mathbb{E}[X_n] = \mathbb{E}[X_n^+] - \mathbb{E}[X_n^-]$.
Il en découle $\sup\limits_{n\in\mathbb{N}} \mathbb{E}[X_n^+] < +\infty \Leftrightarrow \sup\limits_{n\in\mathbb{N}} \mathbb{E}[X_n^-] < +\infty \Leftrightarrow \sup\limits_{n\in\mathbb{N}} \mathbb{E}[|X_n|] < +\infty$.

11/10

Corollaire
Si (X_n) une martingale converge vers X_∞ dans L^1, alors (X_n) converge presque sûrement vers X_∞.
L'implication réciproque est fausse en général.

12 Martingales fermées

Définition (Martingale fermée)
(X_n) telle que $\exists Z \in L^1,\ X_n = \mathbb{E}[Z|\mathcal{F}_n]$.

Proposition
$X_n \xrightarrow{L^1} X_\infty \Leftrightarrow \exists Z \in L^1,\ X_n = \mathbb{E}[Z|\mathcal{F}_n]$
Dans ce cas, $X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty],\ \mathcal{F}_\infty = \sigma(\mathcal{F}_n,\ n \in \mathbb{N})$ et donc $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$.

18/10

13 Convergence dans L^p

Théorème (Inégalité de Doob)
(X_n) une sous-martingale. $\forall a > 0$, on a :
$a \times \mathbb{P}(\max\limits_{i=1}^{n} X_i \geq a) \leq \mathbb{E}[X_n \times \chi_{\max\limits_{i=1}^{n} X_i \geq a}] \leq \mathbb{E}[X_n^+]$

Proposition
Z une variable positive, $p > 0$: $\mathbb{E}[Z^p] = \int\limits_0^\infty pt^{p-1}\mathbb{P}(Z \geq t)dt$.

Proposition
(X_n) sous-martingale positive, $X_n^* := \max\limits_{i=1}^{n} X_i$.
Alors $\forall p > 1,\ \mathbb{E}[(X_n^*)^p] \leq (\frac{p}{p-1})^p \times \mathbb{E}[(X_n)^p]$

Théorème (Convergence L^p)
Soit une martingale (X_n) bornée dans L^p pour $p > 1$ fixé.
Alors $X_n \overset{p.p.,\ L^p}{\longrightarrow} X_\infty \in L^p$ et $\mathbb{E}[|X_n|^p] \nearrow \mathbb{E}[|X_\infty|^p]$.
Avec $X_\infty^* = \sup\limits_{n\in\mathbb{N}} |X_n|$, on a $\mathbb{E}[(X_\infty^*)^p] \leq (\frac{p}{p-1})^p \times \mathbb{E}[|X_\infty|^p]$

14 Convergence dans L^1

Définition (Famille uniformément intégrable (u.i.))
$(X_i)_{i\in I}$ vérifie $\lim\limits_{a\to\infty} \left(\sup\limits_{i\in I}\ \mathbb{E}[\ |X_i| \times \chi_{|X_i|\geq a}\ ] \right) = 0$.

Remarque
(X_i) u.i. $\Rightarrow$ bornée dans L^1.
La réciproque est fausse en général.

Remarque
Si $X_1, ..., X_n \in L^1$, alors $(X_i)_{1\leq i\leq n}$ est u.i.

Proposition
$(X_i)_{i\in I}$ est u.i. $\Leftrightarrow \forall\epsilon > 0,\ \exists\delta > 0$ tel que si $\mathbb{P}(A) < \delta$ alors $\forall i \in I,\ \mathbb{E}[|X_i| \times \chi_A] \leq \epsilon$.

Théorème
$(X_n)_{n\in\mathbb{N}}$ une martingale :
(X_n) converge dans L^1 (donc presque sûrement) $\Leftrightarrow$ (X_n) fermée $\Leftrightarrow$ (X_n) u.i.

25/10

15 Théorème d'arrêt de Doob

Définition
Si $X_n \overset{p.s.}{\to} X_\infty$ et $T : \Omega \to \bar{\mathbb{N}}$, alors $X_T = \mathbb{1}_{T=\infty}X_\infty + \sum\limits_{n\in\mathbb{N}} \mathbb{1}_{T=n}X_n$ est bien défini.

Théorème (Théorème d'arrêt pour une martingale u.i.)
X_n une martingale u.i. de limite X_∞ et T un temps d'arrêt pour la même filtration.
Alors $X_T = \mathbb{E}[X_\infty | \mathcal{F}_T]$.
Si $S \leq T$, par inclusion des tribus, il en découle $\mathbb{E}[X_T | \mathcal{F}_S] = X_S$.

16 Martingales rétrogrades

Définition (Filtration rétrograde)
$(\mathcal{F}_n)_{n\leq 0}$ telle que $(\mathcal{F}_{-n})_{n\in\mathbb{N}}$ décroissante.
On note $\mathcal{F}_{-\infty} = \bigcap\limits_{n\leq 0} \mathcal{F}_n$ sa limite.

Définition (Martingale rétrograde)
$(X_n)_{n\leq 0}$ pour la filtration rétrograde $(\mathcal{F}_n)$:
$\forall m \leq n \leq 0,\ \mathbb{E}[X_n|\mathcal{F}_m] = X_m$ ($\geq$ pour les sous-martingales, $\leq$ pour les surmartingales).

Théorème (Convergence des sous-martingales rétrogrades)
Si (X_n) une sous-martingale rétrograde, bornée dans L^1, alors :
$\exists X_{-\infty},\ X_n \xrightarrow[n\to-\infty]{p.s.,L^1} X_{-\infty}$ et $X_{-\infty} \leq \mathbb{E}[X_n|\mathcal{F}_{-\infty}]$.

Corollaire (Convergence des martingales rétrogrades)
Si (X_n) une martingale rétrograde, bornée dans L^1, alors :
$\exists X_{-\infty},\ X_n \xrightarrow[n\to-\infty]{p.s.,L^1} X_{-\infty}$ et $X_{-\infty} = \mathbb{E}[X_n|\mathcal{F}_{-\infty}]$.

Corollaire
$Z \in L^1$, $(\mathcal{G}_n)_{n\in\mathbb{N}}$ décroissante, $\mathcal{G}_\infty = \bigcap\limits_{n\in\mathbb{N}} \mathcal{G}_n$.
Alors $\mathbb{E}[Z|\mathcal{G}_n] \xrightarrow[n\to\infty]{p.s.,L^1} \mathbb{E}[Z|\mathcal{G}_\infty]$.

Proposition
$\mathcal{G}, \mathcal{G}'$ deux tribus indépendantes. $Z \in L^1$ indépendante de $\mathcal{G}'$.
Alors $\mathbb{E}[Z|\mathcal{G}] = \mathbb{E}\left[Z|\sigma\left(\mathcal{G},\mathcal{G}'\right)\right]$.

17 Loi du 0-1 de Hewitt-Savage

Soient (X_n) iid, de $(\Omega,\mathcal{F})$ dans $(E,\mathcal{G})$, de loi μ.
On munit $E^{\mathbb{N}}$ de la tribu cylindrique $\mathcal{G}^{\otimes\mathbb{N}}$ et de la probabilité $\mathbb{P} = \mu^{\otimes\mathbb{N}}$.
Soit $\mathcal{E}_n = \{A \in \mathcal{G}^{\otimes\mathbb{N}}, (\omega_i)_{i\leq n}(\omega_i)_{i>n} \in A \Rightarrow \forall\sigma\in S_n, (\omega_{\sigma(i)})_{i\leq n}(\omega_i)_{i>n} \in A\}$.

$\mathcal{E} = \bigcap\limits_{n \geq 0} \mathcal{E}_n$ la tribu des évènements échangeables, invariants par toute permutation d'un nombre fini de coordonnées.
$\mathcal{E}$ est triviale : si $A \in \mathcal{E}$, alors $\mathbb{P}(A) \in \{0, 1\}$.

Remarque (Hewitt-Savage $\Rightarrow$ Kolmogoroff)
$\mathcal{H} = \sigma(X_n, X_{n+1}, ...)$ est une sous-tribu de $\mathcal{E}_n$, donc $\bigcap\limits_{n \leq 0} \mathcal{H}_n \subset \mathcal{E}$ est triviale.

18 Processus de Galton-Watson

$Z_0 > 0$ la population initiale fixée, Z_n la population au temps n.
$(\zeta_{n,k})$ iid la loi de descendance de l'individu k au tour n, égale à $\mu \in L^1$ une probabilité sur $\mathbb{N}$ de norme $m = \mathbb{E}[\mu]$.
Soit la loi d'évolution $Z_{n+1} = \sum\limits_{i=1}^{Z_n} \zeta_{n,i}$. On considère les tribus $\mathcal{F}_0 = \{\emptyset, \Omega\}$ et $\mathcal{F}_{n+1} = \sigma(\mathcal{F}_n, (\zeta_{n,i})_{i>0})$.
On définit $A = \bigcup Z_n^{-1}(0)$ l'évènement où la population s'éteint en temps fini.

Proposition
On a les résultats suivants :
- $(\frac{Z_n}{m^n})$ est une martingale pour la filtration $(\mathcal{F}_n)$.
- Dans le cas (sous-)critique où $m \leq 1$, $\mathbb{P}(A) = 1$.
- Dans le cas sur-critique où $m > 1$, $\mathbb{P}(A) = q < 1$ et $\phi(q) = q$ (avec $\phi(s) = \mathbb{E}[s^{Z_1}] = \mathbb{E}[s^{\mu}]$).
- Dans le cas sur-critique, si de plus $\mu \in L^2$, alors $\frac{Z_n}{m^n} \xrightarrow{L^2} M \in L^2$ et $\mathbb{P}(M = 0) = q$.

8/11

19 Chaînes de Markov

Définition (Matrice stochastique)
E un ensemble dénombrable. $Q : E \times E \to \mathbb{R}$ une matrice stochastique :
$\forall x \in E,\ \sum\limits_{y \in E} Q(x, y) = 1,\ \forall y \in E, Q(x, y) \geq 0$.
Autrement dit, $\forall x \in E,\ Q(x, \bullet)$ une probabilité sur E.
De façon équivalente, on considère l'opérateur $Q \in \mathcal{L}(\mathbb{R}^E)$ tel que $Qf(x) = \sum\limits_{y \in E} Q(x, y)f(y)$.

Remarque
Soient Q et Q' stochastiques. $(Q \times Q')(x, y) := \sum\limits_{z \in E} Q(x, z)Q'(z, y)$ définit une matrice stochastique.

Définition (Chaîne de Markov)
Soit Q une matrice de transition.
(X_n) un processus sur E vérifie $\forall y \in E,\ \mathbb{P}(X_{n+1} = y \mid \sigma(X_0 ... X_n)\ ) = \mathbb{P}(X_{n+1} = y \mid X_n\ )$
De façon équivalente, $\forall x_0 ... x_n, y \in E,\ \mathbb{P}(X_0 = x_0 ... X_n = x_n) > 0 \Rightarrow \mathbb{P}(X_{n+1} = y | X_1 = x_1 ... X_n = x_n) = Q(x_n, y)$.

Proposition
(X_n) une chaîne de Markov $\Leftrightarrow \forall x_0 ... x_n \in E,\ \mathbb{P}(X_0 = x_0 ... X_n = x_n) = \mathbb{P}(X_0 = x_0) \times \prod\limits_{i=0}^{n-1} Q(x_i, x_{i+1})$
Alors $\mathbb{P}(X_0 = x) > 0 \Rightarrow \mathbb{P}(X_n = y | X_0 = x) = Q^n(x, y)$.

Remarque
On nomme $\mu(x) = \mathbb{P}(X_0 = x)$ la loi initiale de la chaîne.

Proposition
$Qf(x) = \mathbb{E}[f(X_1)|X_0 = x]$ et $Q^n f(x) = \mathbb{E}[f(X_n)|X_0 = x]$.

Théorème
$\forall Q$ stochastique, $\forall \mu$ une probabilité sur E, $\exists (\Omega, \mathcal{F}, \mathbb{P})$, $\exists (X_n) \in E$ un processus tels que :
(X_n) une chaîne de Markov sur cet espace de loi de transition Q et de distribution initiale μ.

20 Espace canonique

Définition (Espace canonique)
(X_n) une chaîne de Markov : $\left(\Omega, \mathcal{F}\right) = \left(E^{\mathbb{N}}, (\mathcal{P}(E))^{\otimes \mathbb{N}}\right)$ son espace canonique.
Le processus canonique sur un tel espace désigne les projection $X_n\left((\omega_i) \in E^\infty\right) = \omega_n$.
$\mathcal{F}_n = \sigma(X_1 ... X_n) = \{A \subset E^\infty,\ \exists x_1 ... x_n \in E,\ \forall \omega \in A,\ \omega_1 = x_1 ... \omega_n = x_n\}$ et $\mathcal{F} = \sigma(\mathcal{F}_n, n \in \mathbb{N})$.

Proposition
$\forall Q$ stochastique, $\forall \mu$ une loi sur E, $\exists ! \mathbb{P}_\mu$ sur l'espace canonique telle que (X_n) une chaîne de Markov pour (Q, μ).
En particulier, on note $\mathbb{P}_x = \mathbb{P}_\mu$ pour la loi $\mu = \delta_x$: $\forall \mu,\ \mathbb{P}_\mu = \sum\limits_{x \in E} \mu(x) \times \mathbb{P}_x$.

Remarque
Soit X une variable de loi μ : $\mathbb{E}_X[\bullet] = \mathbb{E}_\mu[\bullet] = \int \bullet d\mathbb{P}_\mu(x)$

21 Propriétés de Markov

Définition (Opérateur de décalage)
$\Theta : (\omega_0, \omega_1, ...) \mapsto (\omega_1, ...)$ et de même, $\Theta_n(\omega_0, ...) = (\omega_n, ...)$.

Proposition (Markov faible)
$(\Omega, \mathcal{F})$ l'espace canonique sur (X_n) la chaîne de Markov (Q, μ).
Soient F et G mesurables bornées (ou positives).
Si F est $\mathcal{F}_n$-mesurable, $\mathbb{E}_\mu[F \times (G \circ \Theta_n)] = \mathbb{E}_\mu[F \times \mathbb{E}_{X_n}[G]]$.
Autrement dit, $\mathbb{E}_\mu[G \circ \Theta_n | \mathcal{F}_n] =_{p.s.} \mathbb{E}_{X_n}[G]$.

Proposition (Markov fort)
Avec les notations précédentes, soit T un $(\mathcal{F}_n)$-temps d'arrêt.
Si F est $\mathcal{F}_T$-mesurable, $\mathbb{E}_\mu[F \times \mathbb{1}_{T<\infty} \times (G \circ \Theta_T)] = \mathbb{E}_\mu[F \times \mathbb{1}_{T<\infty} \times \mathbb{E}_{X_T}[G]]$.
Autrement dit, $\mathbb{E}_\mu[\mathbb{1}_{T<\infty} G \circ \Theta_T | \mathcal{F}_T] =_{p.s.} \mathbb{1}_{T<\infty} \mathbb{E}_{X_T}[G]$.

16/11

22 Fonction de Green

Définition (Fonction de Green)
Soient (X_n) une chaîne de Markov Q, μ sur E et $x, y \in E$:

$N(x) := \#\{n \in \mathbb{N}, X_n = x\} = \sum_{n\in\mathbb{N}} \mathbb{1}_{X_n=x}$ une variable aléatoire.
On définit la fonction de Green $G(x, y) = \mathbb{E}_x[N(y)] \in \mathbb{R}^+ \cup \{\infty\}$.

Proposition
$G(x, y) = \sum_{n\in\mathbb{N}} Q^n(x, y)$

Remarque
Si $I - Q$ inversible, alors $G = (I - Q)^{-1}$.

Définition
Soit $x \in E$:
$\tau_x = \min\{n \geq 0, X_n = x\}$ et $\tau_x^+ = \min\{n > 0, X_n = x\}$

Lemme
On a les résultats suivants :
- Sous $\mathbb{P}_x$, $N(x) \sim \mathcal{G}_{eom}(\mathbb{P}_x[\tau_x^+ < \infty])$
- Avec la convention $\frac{1}{0} = \infty$, $G(x, y) = \frac{\mathbb{P}_x[\tau_y<\infty]}{1-\mathbb{P}_y[\tau_y^+<\infty]}$.
- En particulier, $G(x, x) = \frac{1}{1-\mathbb{P}_x[\tau_x^+<\infty]}$

Proposition (États récurrents et transitoires)
On dit que $x \in E$ est récurrent s'il vérifie, de façon équivalente :
- $\mathbb{P}_x(\tau_x^+ < \infty) = 1$
- $N(x) = \infty$ presque sûrement pour la loi $\mathbb{P}_x$
- $G(x, x) = \infty$

A contrario, on dit que x est transient s'il vérifie, de façon équivalente :
- $\mathbb{P}_x(\tau_x^+ < \infty) < 1$
- $N(x) < \infty$ presque sûrement pour la loi $\mathbb{P}_x$
- $G(x, x) < \infty$

23 Classification des états

Définition (Communication)
x conduit à y ($x \rightsquigarrow y$) lorsque $\mathbb{P}_x[\tau_y < \infty] > 0$.
x et y communiquent ($x \sim y$) si $x \rightsquigarrow y$ et $y \rightsquigarrow x$.

Proposition
$x \sim y \Leftrightarrow G(x, y) > 0 \Leftrightarrow \exists n \in \mathbb{N}, Q^n(x, y) > 0$

Remarque
$\rightsquigarrow$ est une relation transitive.
$\sim$ est une relation d'équivalence sur E.

Lemme
Si x récurrent et $x \rightsquigarrow y$, alors y est récurrent et $\mathbb{P}_x[\tau_y < \infty] = \mathbb{P}_y[\tau_x < \infty] = 1$, d'où $x \sim y$.
En conséquence, si $x \sim y$, alors x récurrent $\Leftrightarrow$ y récurrent.

Théorème (Classification des états)
Soit R l'ensemble des états récurrents de E, Q :
Si $x \in R$ et $x \rightsquigarrow y$ alors $y \in R$, donc $\sim$ se restreint correctement à R.
Soient $(R_i)_{i \in I}$ une énumération de $R/\sim$ et $T = \inf\{n \in \mathbb{N}, X_n \in R\}$:
- Si $x \in R_i$ et $y \in R_i$ alors $\mathbb{P}_x$ presque sûrement, $N(y) = \infty$.
- Si $x \in R_i$ et $y \in E \backslash R_i$ alors $\mathbb{P}_x$ presque sûrement, $N(y) = 0$.
- Si $x \in E \backslash R$ et $y \in R_i$ alors $\mathbb{P}_x$ presque sûrement :
  - $T < \infty$ et $X_T \in R_i$, alors $N(y) = \infty$.
  - $T < \infty$ et $X_T \in R_j, j \neq i$, alors $N(y) = 0$.
  - $T = \infty$, alors $N(y) = 0$.
- Si $x \in E \backslash R$ et $y \in E \backslash R$ alors $\mathbb{P}_x$ presque sûrement $N(y) < \infty$.

Définition
Q irréductible $\Leftrightarrow \forall x, y \in E$, $x \sim y$.

Proposition
Soit Q irréductible. Avec la classification ci-dessus, on a nécessairement :
- $R = E$, alors la chaîne est dite irréductible récurrente.
- $R = \emptyset$, alors la chaîne est dite irréductible transiente.

Définition
$x \in E$ est absorbant lorsque $Q(x, x) = 1$.

29/11

24 Marche aléatoire sur $\mathbb{Z}$

$\xi_i \sim \mu$ des variables iid sur $\mathbb{Z}$, X_0 la loi initiale et $X_{n+1} = X_n + \xi_{n+1}$.
(X_n) une chaîne de Markov pour $Q(x, y) = \mu(y - x)$.

Théorème
On suppose $\mu \in L^1$ et $m = \mathbb{E}[\mu]$.
- Tout état $x \in \mathbb{Z}$ est récurrent $\Leftrightarrow m = 0$
- La chaîne Q est irréductible $\Leftrightarrow$ supp(μ) engendre $(\mathbb{Z}, +)$ (en tant que monoïde)

25 Marche aléatoire simple sur $\mathbb{Z}^d$

Avec le même formalisme que précédemment adapté à $\mathbb{Z}^d$, $Q(x,y) = \mu(y - x) = \frac{1}{2d} \times \delta_{||y-x||_1=1}$.
Cette chaîne de Markov est irréductible.

Théorème (Théorème de Polya)
Si $d \in \{1,2\}$, alors la MAS est récurrente.
Sinon, si $d \geq 3$, la MAS est transiente.

26 05/12

27 Chaîne de Markov tuée

$\rho \in [0,1]$, (E,Q) une chaîne de Markov et $\Delta \notin E$ le cimetière.
On a la chaîne Q' qui reste dans (E,Q) avec probabilité $1-\rho$ et meurt en Δ avec probabilité ρ, définie par :
Si $x,y \in E$, $Q'(x,y) = (1-\rho)Q(x,y)$. Si $x \in E$, $Q'(x,\Delta) = \rho$, et enfin $Q'(\Delta,y) = \delta_{y=\Delta}$.
En particulier, si $x,y \in E$, alors $(Q')^n(x,y) = (1-\rho)^n Q^n(x,y)$.
Il en découle $G'(x,y) = \sum\limits_{n \in \mathbb{N}} (Q')^n(x,y) < \infty$, $G'|_{E\times E} \cong (I-(1-\rho)Q)^{-1}$.

28 Mesure invariante et convergence

Définition (Mesure invariante)
$\mu \in (\mathbb{R}^+)^E$ non nulle telle que $\forall x \in E, \mu(x) = \sum_{z \in E} \mu(z) Q(z,x)$.
Autrement dit, $\mu = \mu Q$ est un vecteur propre à gauche.

Définition (Probabilité invariante)
μ une mesure invariante telle que $\mu(E) = 1$, μ une probabilité.
Plus généralement, $\mu(E) < \infty$ suffit, à normalisation près.

Proposition
Si $X_0 \sim \mu$ où μ une proba invariante de Q, alors $\forall n \in \mathbb{N},\ X_n \sim \mu$.

Remarque
Plus généralement, on a toujours $X_{n+1} \sim X_n \times Q$.

Remarque
Si E fini et Q irréductible (nécessairement récurrente), par Perron-Frobenius, $\exists! \mu$ proba invariante.
Si E fini et Q non irréductible, alors il existe une infinité de probas invariantes,
combinaisons convexes des probas sur les composantes irréductibles.

Proposition
Soit $x \in E$ récurrent.
Alors $\mu_x(y) := \mathbb{E}_x[\sum_{k=0}^{\tau_x^+ - 1} \mathbb{1}_{X_k = y}]$ (nombre moyen de passages en y entre deux passages en x) une mesure invariante.
En particulier, $\mu_x(x) = 1$.

Définition (Mesure réversible)
$\mu \geq 0$ non nulle telle que $\forall x, y \in E,\ \mu(x) Q(x,y) = \mu(y) Q(y,x)$.

Proposition
Si μ réversible, alors μ invariante.

Théorème
Soit Q récurrente irréductible : $\exists! \mu$ invariante (à constante multiplicative près).

Définition (États récurrents positifs ou nuls)
Soit $x \in (E, Q)$ un état récurrent.
- Si $\mathbb{E}_x[\tau_x^+] < \infty$, x est récurrent positif.
- Si $\mathbb{E}_x[\tau_x^+] = \infty$, x est récurrent nul.

Proposition (Chaîne récurrente positive ou nulle)
Soit Q récurrente irréductible. Alors de façon mutuellement exclusive, on a nécessairement :

- $\forall x \in E$, x récurrent positif : $\mu(x) = \frac{1}{\mathbb{E}_x[\tau_x^+]}$ l'unique proba invariante.
- $\forall x \in E$, x récurrent nul : il n'existe pas de proba invariante ($\Leftrightarrow$ de mesure invariante finie).

Remarque
Si (E, Q) récurrente irréductible et E fini, alors la chaîne est positive.

Proposition
Si μ proba invariante sur Q irréductible, alors nécessairement Q récurrente positive.

Remarque
Ceci est faux si on n'a qu'une mesure invariante ($\mathbb{Z}^3$ transiente a une mesure invariante).

Théorème (Théorème Ergodique)
Soit μ mesure invariante sur (E, Q) récurrente irréductible.
Soient $f, g \in L^1(E \to \mathbb{R}, \mu)$ avec $g > 0$. Alors en partant de tout point, $\forall \mathbb{P}_x$:
$\frac{\sum\limits_{k=0}^{n-1} f(X_k)}{\sum\limits_{k=0}^{n-1} g(X_k)} \xrightarrow[n\to\infty]{ps} \frac{\int f d\mu}{\int g d\mu}$

Remarque
$\int g d\mu \neq 0$ suffit pour le résultat, mais alors les sommes finies ne sont pas toujours définies.

06/12

Corollaire
Si la mesure $\mu(E) < \infty$ est finie (Q récurrente positive), avec $g = 1$, on a :
$\frac{1}{n} \sum\limits_{k=0}^{n-1} f(X_k) \xrightarrow[n\to\infty]{\mathbb{P}_x\ ps} \int f \frac{d\mu}{\mu(E)}$
En particulier, avec $f = \delta_y$, le temps moyen passé en y converge vers la proba invariante $\mu(y)$.

Corollaire
Si Q récurrente nulle, $\forall f \in L^1(\mu)$, on a :
$\frac{1}{n} \sum\limits_{k=0}^{n-1} f(X_k) \xrightarrow{ps} 0$

29 Convergence en loi vers la proba invariante

Définition (Période de x récurrent)
Soit $x \in E$ un état récurrent : $L(x) := \{n \in \mathbb{N}, Q^n(x,x) > 0\}$ additif.
$d(x) := \gcd(L(x))$ et on a $d(x)\mathbb{Z} = L(x) - L(x)$.

Proposition
Si (E, Q) irréductible, tous les états ont la même période notée $d(Q)$.
Si $d(Q) = 1$, on dit que Q est apériodique.
Si Q irréductible apériodique, alors $\forall x, y \in E,\ \exists n_0 \in \mathbb{N}, \forall n > n_0, Q^n(x,y) > 0$.

Définition (Distance en variation totale)
Soient μ et ν des probas sur E.
$d_{TV}(\mu,\nu) := \frac{1}{2} \sup\limits_{x \in E} |\mu(x) - \nu(x)|$ est une distance majorée par 1 sur les probas de E.

Théorème
Soit μ proba invariante sur (E, Q) récurrente irréductible, positive apériodique.
$\forall X_0$ la loi initiale, on a $d_{TV}(X_n, \mu) \underset{n\to\infty}{\longrightarrow} 0$.
Comme $X_n \sim X_0 Q^n$, de façon équivalente, $d_{TV}(Q^n(X_0, \bullet), \mu) \to 0$.

Corollaire
En général, $\mu_n \overset{d_{TV}}{\to} \mu \Rightarrow \mu_n \overset{\text{loi}}{\to} \mu$.
En particulier, ici, $X_n \overset{\text{loi}}{\to} \mu$ la proba invariante.

Corollaire
Sous les mêmes hypothèses, si de plus E fini, alors la convergence est exponentielle :
$\forall X_0, \exists C < \infty,\ \exists \epsilon \in [0,1[,\ d_{TV}(X_n, \mu) \leq C \times \epsilon^n$

13/12

30 Processus de Poisson

Remarque (Rappels sur la loi exponentielle)
$Z \sim \mathcal{E}(\lambda) \Leftrightarrow \mathbb{P}(Z = t) = \lambda e^{-\lambda t} dt$
$\exists \lambda > 0,\ Z \sim \mathcal{E}(\lambda) \Leftrightarrow \forall t, s > 0,\ \mathbb{P}(Z > t + s | Z > s) = \mathbb{P}(Z > t)$
Si $(Z_i)_{i \leq n}$ indépendantes, exponentielles de paramètres $(\lambda_i)_{i \leq n}$, alors $\inf\limits_{i \leq n}(Z_i) \sim \mathcal{E}(\sum\limits_{i \leq n} \lambda_i)$.

Définition (Processus ponctuel sur $\mathbb{R}^+$)
(T_i) des variables telles que presque sûrement, $(T_n) \nearrow \infty$.

Définition (Processus de Comptage)
$(N_t)_{t \in \mathbb{R}^+}$ définie par $N_t := \#\{i \in \mathbb{N}^*, T_i \leq t\} = \sum\limits_{i=1}^{\infty} \mathbb{1}_{T_i \leq t}$.
On a $N_t - N_s = \#\{i > 0, s < T_i \leq t\}$ et $t \mapsto N_t$ (presque sûrement) continue à droite, convergente à gauche.

Remarque
La donnée des (T_i) équivaut à celle des (N_t).

Définition (Processus de Poisson)
Un processus de Poisson est un processus ponctuel sur $\mathbb{R}^+$ qui vérifie l'accroissement stationnaire indépendant :
$\forall s > 0, (N_t - N_s)_{t \geq s} \perp\!\!\!\perp \sigma\,(N_u,\ u < s)$
$\forall t \geq s,\ N_t - N_s \sim N_{t-s}$

Proposition
Soit (N_t) un processus de Poisson. $\exists! \lambda > 0$ l'intensité du processus, tel que : $\forall t \geq s,\ (N_t - N_s) \sim \mathcal{P}(\lambda \times (t - s))$ (avec $\mu \sim \mathcal{P}(\lambda)$ une loi de Poisson, $\mu(n \in \mathbb{N}) = \frac{\lambda^n}{n!} \times e^{-\lambda}$)

Théorème
Soit (N_t) d'intensité λ : avec $T_0 = 0$ fixé, on a $\tau_i = T_{i+1} - Ti \sim \mathcal{E}(\lambda)$ des variables iid.
Réciproquement, une telle famille $(\tau_i)_{i \geq 0}$ induit naturellement un processus de Poisson d'intensité λ.

Définition (Temps d'arrêt)
T un temps d'arrêt pour la filtration $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$:
$\forall t \geq 0,\ \{\omega \in \Omega,\ T(\omega) \leq t\} \in \mathcal{F}_t$.
On définit $\mathcal{F}_T = \{A \in \sigma(\mathcal{F}_t, t \in \mathbb{R}^+),\ \forall t \in \mathbb{R}^+,\ A \cap T^{-1}([0, t]) \in \mathcal{F}_t\}$ la tribu des évènements antérieurs à T.

Proposition (Markov Fort pour un processus de Poisson d'intensité λ)
Soit T un temps d'arrêt selon la filtration canonique $\mathcal{F}_t^N = \sigma(N_u, u \leq t)$.
On pose $N_t^T := N_{T+t} - N_T$, bien défini lorsque $T(\omega) < \infty$.
Conditionnellement à $T^{-1}(0, \infty[)$, $(N_t^T)_{t \geq 0}$ est un processus de Poisson d'intensité λ, indépendant de la tribu $\mathcal{F}_T^N$.

Théorème (LGN et TCL pour un processus de Poisson)
On retrouve ici deux résultats classiques des probabilités :
- Loi des Grands Nombres : $\frac{N_t}{t} \xrightarrow[t \to \infty]{ps} \lambda$
- Théorème Central Limite : $\frac{N_t - \lambda t}{\sqrt{\lambda \times t}} \xrightarrow[t \to \infty]{\text{loi}} \mathcal{N}(0, 1)$

31 Introduction aux chaînes de Markov à temps continu

On se place toujours sur E dénombrable (ou fini). (X_t) homogène en temps vérifie :

— $\forall t \geq 0,\ \exists Q_t : E \times E \to [0,1]$ de sorte que $\forall t \geq 0,\ \forall x \in E,\ Q_t(x, \bullet)$ une probabilité sur E.

— $\forall t, s \geq 0, \mathbb{P}(X_{t+s} = y | X_u, u \leq s) = \mathbb{P}(X_{t+s} = y | X_s) = Q_t(X_s, y)$.

(Q_t) est le groupe de transition de la chaîne, et vérifie $\forall t, s \geq 0,\ Q_{t+s} = Q_t \times Q_s$ (produit matriciel).
Autrement dit, $\forall t, s \geq 0,\ \forall x, y \in E,\ Q_{t+s}(x,y) = \sum\limits_{z \in E} Q_t(x,z) Q_s(z,y)$.
(X_t) est à trajectoires continues à droite, limitées à gauche.
$\exists L : E \times E \to \mathbb{R}$ le générateur de Markov, ou terme de saut, tel que :

— $\forall x \neq y \in E,\ L_{x,y} \geq 0,\ L_x := \sum\limits_{y \neq x} L_{x,y} < \infty$ et $L_{x,x} = -L_x$.

— Sachant que $X_t = x$, on va attendre un temps aléatoire de loi $\mathcal{E}(L_x)$ puis faire un saut en $y \in E,\ y \neq x$ avec probabilité $\frac{L_{x,y}}{L_x}$.

Remarque
Un processus de Poisson (N_t) est une chaîne de Markov à temps continu particulier,
où on n'effectue que des sauts de n à $(n+1)$ sur $E = \mathbb{N}$.
On a alors l'intensité qui décrit $L_{n,n+1} = L_n = \lambda$ le temps de saut.