Stochastic Process and Applications

Paul Lehaut

November 9, 2025

Contents

1 Chapitre 1: Rappels

1.1 Mesure

Soit $(\Omega, \mathcal{F})$ un espace mesurable.

Définition:

Une fonction $\mu : \mathcal{F} \longrightarrow [0, +\infty]$ est dite σ-additive si pour toute collection dénombrable $(A_i,\ i \in I)$ d'ensembles mesurables deux à deux disjoints, on a:

$$\mu(\bigcup_{i\in I} A_i) = \sum_{i\in I} \mu(A_i).$$

Une mesure μ sur $(\Omega, \mathcal{F})$ est σ-additive, à valeur dans $[0, +\infty]$, définie sur $\mathcal{F}$ telle que: $\mu(\varnothing) = 0$. On dit que $(\Omega, \mathcal{F}, \mu)$ est un espace mesuré, $A \in \mathcal{F}$ est de mesure nulle si $\mu(A) = 0$.
μ est dite σ-finie si il existe $(\Omega_n,\ n \in \mathbb{N})$ telle que:

$$\bigcup_{n\in\mathbb{N}} \Omega_n = \Omega \ \text{ et } \ \forall n \in \mathbb{N},\ \mu(\Omega_n) < +\infty.$$

Une mesure de probabilité $\mathbb{P}$ est une mesure telle que $\mathbb{P}(\Omega) = 1$.

Soit donc μ une mesure sur $(\Omega, \mathcal{F})$.

Proposition:

On a les propriétés suivantes:

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$$

$$A \subset B \implies \mu(A) \leq \mu(B)$$

Convergence monotone: pour $(A_n,\ n \in \mathbb{N})$ telle que $\forall n \in \mathbb{N},\ A_{n+1} \subset A_n,$ alors $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_n)$

Si $(A_i, i \in I)$ est une collection dénombrable d'ensembles mesurables, alors il vient: $\mu(\bigcup_{i\in I} A_i) \leq \sum_{i\in I} \mu(A_i)$.

Définition:

Les événements $(A_i, i \in I)$ sont indépendants si, pour tout sous-ensemble fini $J \subset I$, on a:

$$\mathbb{P}(\bigcap_{j\in J} A_j) = \prod_{j\in J} \mathbb{P}(A_j).$$

1.2 Fonctions Mesurables

Soient $(S, \mathcal{S})$ et $(E, \mathcal{E})$ deux espaces mesurables. Soit f une fonction de S dans E, alors:

$$\{f^{-1}(A),\ A \in \mathcal{E}\} = \sigma(f)$$

est une σ-algèbre.

Définition: Fonction mesurable

La fonction f est dite mesurable si $\sigma(f) \subset \mathcal{S}$.

Si μ est une mesure sur $(E, \mathcal{E})$ et que f est mesurable, alors $\mu_f = \mu \circ f^{-1}$ est une mesure sur $(S, \mathcal{S})$.
Une fonction continue définie sur un espace topologique et prenant ses valeurs dans un espace topologique est mesurable au sens de la σ-algèbre borélienne.
Pour f et g des fonctions mesurables à valeurs réelles définies sur le même espace mesurable, alors les fonction fg

et $\max(f,g)$ sont mesurables. Si par ailleurs ses fonctions ne prennent pas de valeurs infinies, alors la fonction $f+g$ est mesurable.
La composition de fonctions mesurables est également mesurable.

Proposition:
Soit (f_n) une suite de fonctions réelles mesurables, alors les fonctions $\liminf f_n$ et $\limsup f_n$ sont mesurables, en particulier, si (f_n) converge simplement, alors sa limite est mesurable.

Définition: Variable Aléatoire
Une variable aléatoire X définie de Ω dans E est une fonction mesurable définie sur $(\Omega, \mathcal{F})$ à valeurs dans $(E, \mathcal{E})$.
X est dite indépendante de la σ-algèbre $\mathcal{H}$ si, pour tout $(A,B) \in \mathcal{E} \times \mathcal{H}$, les événements $\{X \in A\}$ et B sont indépendants.

1.3 Théorème de convergence pour l'intégration

Soit (f_n) une suite de fonctions mesurables à valeurs réelles. Cette suite converge presque partout si

$$\liminf f_n = \limsup f_n \text{ presque partout.}$$

On rappelle que la limite de cette suite de fonction est alors mesurable.

Théorème: Convergence monotone
Soit $(f_n,\ n \in \mathbb{N})$ une suite de fonctions mesurables à valeurs réelles telle que pour tout $n \in \mathbb{N},\ 0 \leq f_n \leq f_{n+1}$ presque partout, alors il vient:

$$\lim_{n\to\infty} \int f_n d\mu = \int \lim_{n\to\infty} f_n d\mu.$$

Le lemme de Fatou donne, pour (f_n, n) une suite de fonctions mesurables et positives presque partout, alors il vient:

$$\liminf \int f_n d\mu \geq \int \liminf f_n d\mu.$$

Théorème: Convergence dominée de Lebesgue
Soient $f,\ g$ deux fonctions à valeurs réelles mesurables, soit (f_n) une suite de fonctions à valeurs réelles mesurables. On suppose que, pout tout $n \in \mathbb{N}$, on a presque partout $|f_n| \leq g$, que f désigne la limite de (f_n) et que g est intégrable, alors:

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

1.4 Espaces L^p

On commence par rappeler les inégalités suivantes pour f et g des fonctions mesurables à valeurs réelles:

Inégalité de Hölder: Soient $p, g \in (1, +\infty)$ tels que $\frac{1}{p} + \frac{1}{q} = 1$, supposons que $|f|^p$ et $|g|^q$ soient intégrables, alors fg est intégrable et on a:

$$\int |fg| d\mu \leq (\int |f|^p d\mu)^{1/p} (\int |g|^q d\mu)^{1/q}.$$

Inégalité de Cauchy-Schwarz: Supposons que f et g soient de carré intégrable, alors fg est intégrable et on a:

$$\int |fg| d\mu \leq (\int f^2 d\mu)^{1/2} (\int g^2 d\mu)^{1/2}$$

on a égalité si et seulement si f et g sont propotionnelles presque partout.

Inégalité de Minkowski: Soit $p \in [1, +\infty)$, supposons que $|f|^p$ et $|g|^p$ soient intégrables, alors on a:

$$(\int |f+g|^p d\mu)^{1/p} \leq (\int |f|^p d\mu)^{1/p} + (\int |g|^p d\mu)^{1/p}.$$

<u>Proposition:</u>
Soit $p \in [1, +\infty)$, l'espace-vectoriel normé $(L^p, ||\cdot||_p)$ est complet.

<u>Théorème:</u> Fubini
Soient ν et μ deux mesures σ-finies respectivement sur $(E, \mathcal{E})$ et $(S, \mathcal{S})$, alors:
-il existe une unique mesure sur $(E \times S, \mathcal{E} \otimes \mathcal{S})$, notée $\nu \otimes \mu$, telle que:

$$\forall (A, B) \in \mathcal{E} \times \mathcal{S},\ \nu \otimes \mu(A \times B) = \nu(A)\mu(B)$$

c'est la mesure produit
-Soit f une fonction à valeurs réelles défnie sur $E \times S$, alors:

$$\int f(x,y)\nu \otimes \mu(dx, dy) = \int\int f(x,y)\mu(dy)\nu(dx) = \int\int f(x,y)\nu(dx)\mu(dy).$$

1.5 Espérance, Variance et Inégalités

Soit X une variable aléatoire, soit f une fonction à valeurs réelles, si $\mathbb{E}(f(X))$ est bien définie, alors on a:

$$\mathbb{E}(f(X)) = \int f(x)\mathbb{P}_X(dx).$$

Inégalité de Tchebychev: Soit X une VA$\mathbb{R}$, soit, $a > 0$, alors:

$$\mathbb{P}(|X| \geq a) \leq \frac{E(X^2)}{a^2}.$$

Inégalité de Jensen: Soit X une VA$\mathbb{R}^d$ intégrable, soit f une fonction à valeurs réelles convexe définie sur $\mathbb{R}^d$, alors $\mathbb{E}(f(X))$ est bien définie et:

$$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X)).$$

2 Espérance Conditionnelle

2.1 Espérance Conditionnelle

On s'intéresee à une VA$\mathbb{R}$ définie sur $(E, \mathcal{F})$ dont l'espérance est bien définie ainsi qu'à $\mathcal{H} \subset F$ une σ-algèbre.

<u>Définition:</u> Espérance conditionnelle
On dit qu'une variable aléatoire Z, mesurable pour $\mathcal{H}$ telle que $\mathbb{E}(Z)$ soit bien définie, est l'espérance conditionnelle de X par rapport à $\mathcal{H}$ si:

$$\mathbb{E}(X1_A) = \mathbb{E}(Z1_A),\ \forall A \in \mathcal{H}.$$

Pour Z et Z' deux variables aléatoires $\mathcal{H}$ mesurables telles que leurs espérances soient bien définies et que:

$$\mathbb{E}(Z1_A) = \mathbb{E}(Z'1_A),\ \forall A \in \mathcal{H}$$

alors $Z = Z'$ presque partout.

Théorème: Radon-Nikodym

Soient μ et ν deux mesures σ-finie sur $(\Omega, \mathcal{H})$ telles que: $\nu(A) = 0 \implies \mu(A) = 0$, alors il existe une fonction mesurable f positive telle que:

$$\int_A f d\nu = \mu(A)$$

on note alors: $f = \frac{d\mu}{d\nu}$ et on l'appelle dérivée de Radon-Nikodym.

Proposition:

Pour X et Y des VA$\mathbb{R}$ de carré intégrable alors:

-Si X est positive presque partout, alors l'espérance conditionnelle de X est positive presque partout.

-On a presque partout: $\mathbb{E}(aX + bY|\mathcal{H}) = a\mathbb{E}(X|\mathcal{H}) + b\mathbb{E}(Y|\mathcal{H})$.

-Soit $(X_n,\ n \in \mathbb{N})$ une suite croissante de VA$\mathbb{R}$ positives de carré intégrable alors on a presque partout:

$$\lim \mathbb{E}(X_n|\mathcal{H}) = \mathbb{E}(\lim X_n|\mathcal{H}).$$

Proposition:

L'espérance conditionnelle de X par rapport à $\mathcal{H}$ existe toujours, par ailleurs on a:

$$\mathbb{E}(\mathbb{E}(X|\mathcal{H})) = \mathbb{E}(X)$$

donc l'intégrabilité de X entraîne celle de $\mathbb{E}(X|\mathcal{H})$.

Proposition:

-Si X est positive presque partoutalors son espérance conditionnelle l'est également.

-Si X et Y sont intégrables, alors, pour tout réels a et b, on a :

$$\mathbb{E}(aX + bY|\mathcal{H}) = a\mathbb{E}(X|\mathcal{H}) + b\mathbb{E}(Y|\mathcal{H})$$

et, si $X \leq Y$ presque partout, alors $\mathbb{E}(X|\mathcal{H}) \leq \mathbb{E}(Y|\mathcal{H})$ presque partout.

-Soit $(X_n,\ n \in \mathbb{N})$ une suite croissante de VA$\mathbb{R}$ positives presque partout, alors on a presque partout:

$$\lim \mathbb{E}(X_n|\mathcal{H}) = \mathbb{E}(\lim X_n|\mathcal{H}).$$

-Le lemme de Fatou s'écrit: Soit $(X_n,\ n \in \mathbb{N})$ une suite de VA$\mathbb{R}$ positives presque partout, alors on a presque partout:

$$\mathbb{E}(\liminf X_n|\mathcal{H}) \leq \liminf \mathbb{E}(X_n|\mathcal{H}).$$

-La convergence dominée de Lebesgue s'écrit: Soient $X, Y, (X_n)$ des VA$\mathbb{R}$ telles que (X_n) converge vers X presque partout et $|X_n| \leq Y$ avec Y intégrable, alors on a :

$$\lim \mathbb{E}(X_n|\mathcal{H}) = \mathbb{E}(X|\mathcal{H}).$$

Par ailleurs, les inégalités de Hölder, Cauchy-Schwarz, Minkowski et Jensen restent valables pour l'espérance conditionnelle.

De plus, pour X et Y deux VA$\mathbb{R}$ telles que $\mathbb{E}(X)$ et $\mathbb{E}(XY)$ soient bien définies et que Y soit $\mathcal{H}$ mesurable, alors on a:

$$\mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(X|\mathcal{H})Y).$$

Proposition:
On suppose X intégrable, alors:

-Si X est $\mathcal{H}$ mesurable, alors: $\mathbb{E}(X|\mathcal{H}) = X$.

-Si X est indépendante de $\mathcal{H}$ alors: $\mathbb{E}(X|\mathcal{H}) = \mathbb{E}(X)$.

-Si Y est une VA $\mathcal{H}$ mesurable telle que $\mathbb{E}(XY)$ soit bien définie, alors: $\mathbb{E}(YX|\mathcal{H}) = Y\mathbb{E}(X|\mathcal{H})$.

-Si $\mathcal{G} \subset \mathcal{H}$ est une σ-algèbre, alors: $\mathbb{E}(\mathbb{E}(X|\mathcal{H})|\mathcal{G}) = \mathbb{E}(X|\mathcal{G})$.

-Si $A \in \mathcal{H}$, alors: $\mathbb{E}(X|A) = \mathbb{E}(X1_A)\mathbb{P}(A)$.

2.2 Espérance Conditionnée par une VA

Soit V une variable aléatoire définie sur $(E, \mathcal{E})$, on note $\mathbb{E}(X|V) = \mathbb{E}(X|\sigma(V))$, si $\mathbb{E}(X)$ est bien définie, alors il existe une fonction mesurable g définie sur E telle que: $\mathbb{E}(X|V) = g(V)$.

Si V est discrète, alors:

$$g(v) = \frac{\mathbb{E}(X1_{V=v})}{\mathbb{P}(V=v)} = \mathbb{E}(X|V=v) \ \text{ si } \mathbb{P}(V=v) > 0, \quad \text{et } g(v) = 0 \ \text{ sinon.}$$

3 Chaînes de Markov Discrètes

Une chaîne de Markov est une suite de variables aléatoires $X = (X_n)$ qui représente l'évolution dynamique d'un système stochastique. La propriété fondamentale des chaînes de Markov est que l'état du système à l'instant n ne dépend des états précédents qu'à travers X_n. C'est-à-dire que, conditionnellement à X_n, $(X_1, ..., X_n)$ et $(X_{n+k})_k$ sont indépendants.

Dans ce qui suit on considère l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et l'espace $(E, \mathcal{E})$ où E est au plus dénombrable, sans précisions supplémentairtes les VAs seront définies de $\mathcal{F}$ dans $\mathcal{E}$.

3.1 Définitions et Propriétés

Soit $X = (X_n)$ une chaîne de Markov, on définit l'ensemble des informations disponibles à l'instant n par la σ-algèbre $\mathcal{F}_n$

Définition: Filtre
Un filtre $\mathbb{F} = (\mathcal{F}_n)$ est une suite croissante de σ-algèbre incluses dans $\mathcal{F}$. Une suite de VAs (X_n) est adaptée pour $\mathbb{F}$ si X_n est $\mathcal{F}_n$ mesurable pour tout n.

Définition:
Le processus $X = (X_n)$ est une chaîne de Markov qui respecte le filtre $\mathbb{F}$ si elle est adaptée pour ce filtre et si elle vérifie la propriété de Markov: pour tout n, conditionnellement à X_n, $\mathcal{F}_n$ et $(X_{n+k})_{k\in\mathbb{N}}$ sont indépendants.

Proposition: Définitions équivalentes
On suppose que (X_n) est adaptée au filtre $\mathbb{F}$, alors on a les équivalences suivantes:

- (X_n) est une chaîne de Markov
- Pour tout n et $B \in \sigma(X_k)_{k\geq n}$ on a $\mathbb{P}(B|\mathcal{F}_n) = \mathbb{P}(B|X_n)$
- Pour tout n et $y \in E$ on a $\mathbb{P}(X_{n+1} = y|\mathcal{F}_n) = \mathbb{P}(X_{n+1} = y|X_n)$.

La proposition précédente nous permet d'introduire le lemme suivant:

Lemme:

Soient un espace mesurable $(S, \mathcal{S})$ et (U_n) une suite de VAs indépendantes à valeurs dans S. Soient X_0 une variable aléatoire à valeurs dans E indépendante de (U_n) et f une fonction mesurable définie de $E \times S$ dans E, on peut alors définir le système dynamique stochastique suivant: $X_{n+1} = f(X_n, U_{n+1})$ qui est une chaîne de Markov.

Définition: Matrices de transition

Une chaîne de Markov X sur E a les matrices de transition $(P_n)_n$ si c'est une suite de matrices stochastiques sur E telle que: $\mathbb{P}(X_{n+1} = y | X_n) = P_{n+1}(X_n, y)$.

Une chaîne de Markov est homogène si (P_n) est constante.

Proposition:

La distribution d'une chaîne de Markov homogène est caractérisée par sa matrice de transition P et la distribution initiale de probabilité μ_0 pour X_0.

De plus, pour tout $n \in \mathbb{N}^*$, $x_1, ..., x_n \in E$, alors:

$$\mathbb{P}(X_0 = x_0, ..., X_n = x_n) = \mu_0(x_0) \prod_{k=1}^{n} P(x_{k-1}, x_k).$$

Proposition:

Soit (X_n) une chaîne de Markov de matrice de transition P, on note μ_n la densité de probabilité de X_n, soit f une fonction bornée ou positive, alors:

- $\mu_n = \mu_0 P^n$
- $\mathbb{E}(f(X_n)) = \mu_n f$ et $\mathbb{E}(f(X_n) | X_0 = x) = P^n f(x)$
- $\mathbb{E}(f(X_n) | \mathcal{F}_{n-1}) = P f(X_{n-1})$
- $\mathbb{E}(f(X_n) | X_0) = P^n f(X_0)$
- $\mathbb{P}(X_{n+k} = y | \mathcal{F}_n) = P^k(X_n, y)$

3.2 Probabilité Invariante et Réversible

On se donne dans cette partie une distribution de probabilité π et une matrice stochastique P.

Définition:

π est invariante pour P si $\pi = \pi P$. P est dite réversible par rapport à π si $\pi(x)P(x, y) = \pi(y)P(y, x)$.

On remarque alors que si P est réversible par rapport à π, alors π est invariante par rapport à P.

3.3 Classification d'Etats

On introduit les notation: $\mathbb{E}_x(X_n) = \mathbb{E}(X_n | X_0 = x)$ et $\mathbb{P}_x(X_n) = \mathbb{P}(X_n | X_0 = x)$.

Définition:

Un état y est accessible depuis un état x, noté $x \to y$, si $\mathbb{P}_x(X_n = y) > 0$.

x et y communiques $(x \leftrightarrow y)$ si y est accessible depuis x et inversement.

Proposition:

La relation $\leftrightarrow$ est une relation d'équivalence. On note alors C_x la classe d'équivalence d'un état x.

Définition: Irréductibilité

On dit qu'une chaîne de Markov est irréductible si E ne contient qu'une seule classe d'équivalence.

Définition: Classe fermée et état absorbant
Soit un état x, on dit que C_x est fermée si $C_x = \{y \in E | x \to y\}$. x est un état absorbant si $C_x = \{x\}$ C_x est fermée.

Si une chaîne de Markov possède au moins deux éléments et un état absorbant alors elle n'est pas irréductible.

Définition: Temps de retour, état transient et récurrent
On pose $T^x = \inf\{n \geq 1 | X_n = x\}$. x est dit transient si $\mathbb{P}(T^x = +\infty) > 0$, il est dit récurrent si $\mathbb{P}(T^x = +\infty) = 0$.

On peut alors noter $N^x = \sum_{n \in \mathbb{N}} 1_{X_n = x}$ le nombre de passage en x.

Proposition:
Soit (X_n) une chaîne de Markov de matrice de transition P, alors:

- Si x est récurrent, alors $\mathbb{P}(N^x = +\infty) = 1$ et $\sum_n P^n(x,x) = +\infty$
- S'il est transient, alors $\mathbb{P}(N^x < +\infty) = 1$ et $\sum_n P^n(x,x) < +\infty$
- Les états de C_x sont ou tous transients ou tous récurrents
- Si C_x est ouvert, alors ses états sont transients.

3.4 Théorèmes Asymptotiques

On note, pour un état x, $\pi(x) = 1/\mathbb{E}_x(T^x)$.

Définition:
Un état récurrent x est: récurrent nul si $\pi(x) = 0$, récurrent positif si $\pi(x) > 0$.

Théorème:
Soit (X_n) une chaîne de Markov irréductible, alors elle est ou bien récurrente nulle, ou bien récurrente positive, ou bien transiente.
Si elle est récurrente nulle ou transiente, alors il n'y a pas de probabilité invariante.
On a finalement le résultat suivant:

$$\forall x \in E, \frac{1}{n} \sum_{k=1}^{n} 1_{X_k = x} \longrightarrow_{ps} \pi(x).$$

4 Points A Retenir

4.1 Fonctions Mesurables

On rappelle, pour f de $(S, \mathcal{S})$ dans $(E, \mathcal{E})$: $\sigma(f) = \{f^{-1}(A),\ A \in \mathcal{E}\}$.
La fonction f est alors dite mesurable si $\sigma(f) \in \mathcal{S}$.
En général le produit, le max et la somme de fonction mesurable est mesurable.
La limite simple d'une suite de fonctions mesurables est mesurable.
Le théorème de convergence monotone donne, pour (f_n) une suite de fonctions mesurables à valeurs réelles telle que $0 \leq f_n \leq f_{n+1}$ presque partout, alors il vient:

$$\lim \int f_n d\mu = \int \lim f_n d\mu.$$

4.2 Espaces L^p

Inégalités:

- Hölder: pour $p, q \in (1, +\infty)$ tels que $\frac{1}{p} + \frac{1}{q} = 1$, supposons que $f \in L^p$ et $g \in L^q$, alors fg est intégrable et:

$$\int |fg| d\mu \leq (\int |f|^p d\mu)^{\frac{1}{p}} (\int |g|^q d\mu)^{\frac{1}{q}}.$$

- CS: supposons désormais que f et g soient de carré intégrable, alors:

$$\int |fg| d\mu \leq (\int |f|^2 d\mu)^{\frac{1}{2}} (\int |g|^2 d\mu)^{\frac{1}{2}}.$$

- Minkowski: soit $p \in [1, +\infty)$, on suppose que $f, g \in L^p$, alors:

$$(\int |f + g|^p d\mu)^{\frac{1}{p}} \leq (\int |f|^p d\mu)^{\frac{1}{p}} + (\int |g|^p d\mu)^{\frac{1}{p}}.$$

Pour $p \in [1, +\infty)$, l'espace-vectoriel normé $(L^p,\ || \cdot ||_p)$ est complet.

4.3 Espérance

Inégalité de Tchebychev: pour $a > 0$, on a: $\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(X^2)}{a^2}$. Inégalité de Jensen: pour f convexe: $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$.

On rappelle la définition: une variable aléatoire Z, mesurable pour $\mathcal{H}$ telle que son espérance soit bien définie, est l'espérance conditionnelle de X par rapport à $\mathcal{H}$ si:

$$\mathbb{E}(X1_A) = \mathbb{E}(Z1_A),\ \forall A \in \mathcal{H}.$$

L'espérance conditionnelle de X par rapport à $\mathcal{H}$ existe toujours et, par ailleurs: $\mathbb{E}(\mathbb{E}(X|\mathcal{H})) = \mathbb{E}(X)$.
La convergence dominée de Lebesgue s'écrit ici: soient $X, Y, (X_n)$ des VA$\mathbb{R}$ telles que (X_n) converge vers X presque partout et $|X_n| \leq Y$ avec Y intégrable, alors on a :

$$\lim \mathbb{E}(X_n|\mathcal{H}) = \mathbb{E}(X|\mathcal{H}).$$

De plus, pour X et Y deux VA$\mathbb{R}$ telles que $\mathbb{E}(X)$ et $\mathbb{E}(XY)$ soient bien définies et que Y soit $\mathcal{H}$ mesurable, alors il vient;

$$\mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(X|\mathcal{H})Y).$$

Propriétés importantes:

- Si X est $\mathcal{H}$ mesurable, alors: $\mathbb{E}(X|\mathcal{H}) = X$
- Si X est indépendante de $\mathcal{H}$, alors: $\mathbb{E}(X|\mathcal{H}) = \mathbb{E}(X)$
- Si Y est une VA $\mathcal{H}$ mesurable telle que $\mathbb{E}(XY)$ soit bien définie, alors: $\mathbb{E}(YX|\mathcal{H}) = Y\mathbb{E}(X|\mathcal{H})$
- Si $\mathcal{G} \subset \mathcal{H}$ est une σ-algèbre, alors: $\mathbb{E}(\mathbb{E}(X|\mathcal{H})|\mathcal{G}) = \mathbb{E}(X|\mathcal{G})$
- Si $A \in \mathcal{H}$, alors: $\mathbb{E}(X|A) = \mathbb{E}(X1_A)\mathbb{P}(A)$.

4.4 Chaînes de Markov discrètes

Propriété fondamentale: conditionnellement à X_n, les événements $X_1, ..., X_{n-1}$ et $(X_{n+k})_{k\geq 0}$ sont indépendants.

Le processus (X_n) est une chaîne de Markov qui respecte le filtre $\mathbb{F}$ si elle est adaptée pour ce filtre et si elle vérifie: pour tout n, conditionnellement à X_n, les événements $\mathcal{F}_n$ et (X_{n+k}) sont indépendants.

On a les équivalneces pour (X_n) adaptée à $\mathbb{F}$: (X_n) est une chaîne de Markov ssi pour tout $B \in \sigma((X_k)_{k\geq n})$, $\mathbb{P}(B|\mathcal{F}_n) = \mathbb{P}(B|X_n)$ pour tout n ssi $\mathbb{P}(X_{n+1} = y|\mathcal{F}_n) = \mathbb{P}(X_{n+1} = y|X_n)$ pour tout n et $y \in E$.

Le résultat suivant est un résultat fondamental sur la construction d'une chaîne de Markov: Soient un espace mesurable $(S, \mathcal{S})$ et (U_n) une suite de VAs indépendantes à valeurs dans S. Soient X_0 une VA dans E indépendante de (U_n) et f mesurable de $E \times S$ dans E, alors

$$\underline{X_{n+1} = f(X_n, U_{n+1}) \text{ est une chaîne de Markov.}}$$

Pour une chaîne de Markov homogène, la distribution initiale de probabilité μ_0 pour X_0 et la matrice de transition P, alors:

$$\mathbb{P}(X_0 = x_0, ..., X_n = x_n) = \mu_0(x_0) \prod_{j=1}^{n} P(x_{j-1}, x_j).$$

On a les propositions suivantes, pour une chaîne de Markov homogène:

- $\mu_n = \mu_0 P^n$
- $\mathbb{E}(f(X_n)) = \mu_0 P^n f$
- $\mathbb{E}(f(X_n)|\mathcal{F}_{n-1}) = Pf(X_{n-1})$
- $\mathbb{E}(f(X_{n+k}|X_n)) = P^k f(X_n)$
- $\mathbb{E}(X_{n+k} = y|X_n) = P^k(X_n, y)$
- Si x est récurrent: $P(N^x = +\infty) = 1$ et $\sum_n P^n(x,x) = +\infty$
- S'il est transient: $P(N^x < +\infty) = 1$ et $\sum_n P^n(x,x) < +\infty$
- Les états d'une même classe d'équivalence sont ou tous transients ou tous récurrents, si la classe est ouverte, ils sont tous transients

Une chaîne de Markov irréductible est ou bien récurrente nulle/positive ou transiente.

Si elle est récurrente nulle ou transiente, alors il n'y a pas de probabilité invariante.

On a $\frac{1}{n}\sum_{k=1}^{n} 1_{X_k=x} \longrightarrow_{ps} 1/\mathbb{E}_x(T^x)$.