

10 Les processus de Markov à temps continu

Dans toute cette section, nous appellerons processus une famille $(X_t)_{t \geq 0}$ de variables aléatoires à valeurs dans un espace mesurable E quelconque, indexées par $t \in \mathbb{R}$. Si cet espace E est \mathbb{R} , ou \mathbb{N} , ou plus généralement un espace topologique muni de sa tribu borélienne, alors on dit que le processus est continu, continu à droite, continu à gauche, etc. si, pour (presque) tout ω , les fonctions $t \mapsto X_t(\omega)$ sont continues, continues à droite, etc. (La topologie que nous mettons sur \mathbb{N} est la topologie discrète, c'est à dire celle héritée de \mathbb{R} .)

Ce que nous appelons loi du processus est la collection de toutes les lois de n -uplets $(X_{t_1}, \dots, X_{t_n})$, c'est à dire en, en termes savants, la loi des marginales de rang fini.

10.1 Le processus de Poisson.

Le processus de Poisson est le plus simple des processus de Markov à temps continu. Plutôt que d'en donner une définition formelle, nous préférons le construire et donner ses propriétés élémentaires.

Nous commençons par quelques calculs de lois :

Définition 10.1. Soient (X_1, \dots, X_n) n variables aléatoires indépendantes de loi uniforme dans l'intervalle $[a, b]$. On note $(X_{(1)}, \dots, X_{(n)})$ leur réarrangement dans l'ordre croissant. La loi de $(X_{(1)}, \dots, X_{(n)})$ s'appelle la statistique d'ordre d'ordre n sur $[a, b]$, ou loi de Dirichlet d'ordre n sur $[a, b]$. On la note

$$D_n([a, b])(dx_1, \dots, dx_n).$$

Si on change l'intervalle $[a, b]$, ces lois se transforment par affinité.

Elle admet une densité par rapport à la mesure de Lebesgue sur \mathbb{R}^n qui vaut

$$\frac{n!}{(b-a)^n} \mathbf{1}_{a < \{x_1 < x_2 < \dots < x_n < b\}}.$$

Voici quelques propriétés élémentaires des lois de Dirichlet : la démonstration est laissée au lecteur à titre d'exercice.

Proposition 10.2. Soit (Y_1, \dots, Y_n) une variable aléatoire ayant la loi $D_n([a, b])$.

1. La loi de Y_n est $n \frac{(x-a)^{n-1}}{(b-a)^n} \mathbf{1}_{[a, b]}(x) dx$.

2. Si c est dans $[a, b]$, la loi de (Y_1, \dots, Y_{n-1}) sachant que $(Y_n = c)$ est $D_{n-1}([a, c])$.
3. De même, la loi de (Y_2, \dots, Y_n) sachant que $(Y_1 = c)$ est $D_{n-1}([c, b])$.
4. Pour $1 \leq k \leq p \leq n$, (Y_k, \dots, Y_p) est conditionnellement indépendante de $(Y_1, \dots, Y_k, Y_p, \dots, Y_n)$ sachant (Y_p, Y_k) , et la loi de $(Y_{k+1}, \dots, Y_{p-1})$ sachant que $\{(Y_k, Y_p) = (c, d)\}$ est $D_{p-k-1}([c, d])$.
5. La loi de (Y_1, \dots, Y_{n-1}) sachant que $\{Y_{n-1} \leq c \leq Y_n\}$ est $D_{n-1}([a, c])$.
6. La loi de (Y_2, \dots, Y_n) sachant que $\{Y_1 \leq c \leq Y_2\}$ est $D_{n-1}([c, b])$.

Dans ce qui va suivre, les lois exponentielles vont jouer un rôle particulier. Ce sont les seules lois "sans mémoire" sur \mathbb{R}_+ :

Proposition 10.3. *Soit T une variable aléatoire à valeurs dans \mathbb{R}_+ ayant la propriété suivante :*

$$\forall s, t > 0, \mathbf{P}(T > t + s/T > t) = \mathbf{P}(T > s).$$

Alors, si T n'est pas identiquement nulle, T suit une loi exponentielle. En d'autres termes, c'est la seule loi de variable T telle que la loi de $T - t$ sachant que $T > t$ est la même que la loi de T . On dit que T est sans mémoire.

Démonstration. — La preuve est directe. Si $F(t) = \mathbf{P}(T > t)$, F est bornée, continue à droite, et vérifie $F(t+s) = F(t)F(s)$. On vérifie alors immédiatement que pour tout entier k , $F(k) = F(1)^k$, puis que $F(k/p) = F(1)^{k/p}$, et on prolonge enfin par continuité à droite pour obtenir $F(t) = F(1)^t$. Ceci caractérise les variables exponentielles, sauf si $F(1) = 0$ auquel cas T est identiquement nulle. ■

Par définition, les lois de Dirichlet sont liées aux réarrangements de lois uniformes. Mais elles apparaissent aussi lorsqu'on conditionne des sommes de variables aléatoires exponentielles :

Proposition 10.4. *On considère une suite S_n de variables aléatoires indépendantes, à valeurs dans \mathbb{R}_+ , de loi exponentielle de paramètre λ . $\mathbf{P}(S_i \geq t) = \exp(-\lambda t)$. On pose $T_n = S_1 + \dots + S_n$. Alors, la loi de (T_1, \dots, T_n) sachant que $\{T_n \leq t \leq T_{n+1}\}$ est $D_{n-1}([0, t])$.*

De même, la loi de (T_1, \dots, T_{n-1}) sachant que $\{T_n = t\}$ est $D_{n-1}([0, t])$.

De plus, la loi de T_n est

$$\frac{t^{n-1}}{(n-1)!} \exp(-\lambda t) \mathbf{1}_{t>0} dt.$$

Démonstration. — Une fois de plus, ces propriétés sont élémentaires et laissées au lecteur. Tout repose sur le calcul de la loi de (T_1, \dots, T_n) qui est, par un changement de variables simple

$$\lambda^n \mathbf{1}\{0 < t_1 < \dots < t_n\} e^{-\lambda t_n} dt_1 \dots dt_n.$$

■

Nous pouvons alors définir le processus de Poisson :

Définition 10.5. Soit (S_1, \dots, S_n, \dots) une suite de variables aléatoires exponentielles indépendantes de paramètre λ . Posons $T_n = \sum_1^n S_i$ ($T_0 = 0$), et, pour tout t réel,

$$N_t = \sum_{n \geq 0} \mathbf{1}_{\{T_n \leq t\}}.$$

La famille de variables aléatoires (N_t) s'appelle le **processus de Poisson standard issu de 0 d'intensité (ou de paramètre) λ** .

Ce processus N_t compte le nombre de variables T_i qui sont dans $[0, t]$. Tous les processus ayant la même loi que celui-ci (en un sens que nous définirons plus bas) seront aussi des processus de Poisson.

Ce processus a les propriétés suivantes.

Proposition 10.6. 1. Pour tout t , la variable aléatoire N_t est finie presque sûrement.

2. Pour tout ω , la fonction $t \mapsto N_t$ est croissante, continue à droite, constante par morceaux et ne croît que par sauts de 1.

3. Si $N_{t-}(\omega)$ désigne la limite à gauche au point t de la fonction $N_t(\omega)$, alors pour tout t , $N_t = N_{t-}$ presque sûrement.

4. Pour tout t , la variable N_t suit une loi de Poisson de paramètre λt , c'est à dire

$$\mathbf{P}(N_t = k) = \frac{(\lambda t)^k}{k!} \exp(-\lambda t).$$

5. Pour tous $0 < t_1 < t_2 < \dots < t_n$, les variables $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ sont indépendantes.

6. Si $s < t$, la loi de $N_t - N_s$ est la même que celle de N_{t-s} .

Ces deux dernières propriétés disent que N est un **processus à accroissements indépendants homogène**.

Démonstration. — Le premier point est facile à voir : la suite T_n/n converge presque sûrement vers $\mathbf{E}(T_1) = 1/\lambda$, et donc T_n converge presque sûrement vers l'infini. Il n'y a donc qu'un nombre fini de termes dans la somme qui définit N_t .

De plus, pour tout ω , $N_t = \sum_n \mathbf{1}_{[S_n, \infty)}(t)$ et puisque cette somme est finie, c'est bien une fonction croissante continue à droite, ne croissant que par sauts de 1, et constante entre ses sauts.

L'aléa ω étant fixé, l'ensemble des points t de \mathbb{R} où $N_t \neq N_t^-$ est $A(\omega) = \{S_n(\omega), n \geq 1\}$. Or, t étant fixé, pour presque tout ω , et pour tout n , $\mathbf{P}(S_n = t) = 0$. Et donc, pour presque tout ω , $A \cap \{t\} = \emptyset$, et donc $N_t = N_t^-$. C'est un paradoxe apparent que la fonction $t \mapsto N_t(\omega)$ soit discontinue, mais que $N_t = N_t^-$ presque sûrement pour tout t : cela provient de ce que l'ensemble des nombres réels n'est pas dénombrable.

La suite est plus délicate. Commençons par la loi de N_t :

$$\mathbf{P}(N_t = k) = \mathbf{P}(T_k \leq t < T_{k+1}).$$

Nous avons vu plus haut la loi de (T_1, \dots, T_n) , et nous pouvons donc calculer

$$\mathbf{P}(N_t = k) = \int_{0 < t_1 < \dots < t_k < t < t_{k+1}} \lambda^{k+1} e^{-\lambda t_{k+1}} dt_1 \dots dt_{k+1}.$$

Dans la dernière intégrale, nous pouvons séparer l'intégrale en t_{k+1} des autres, et il reste

$$\lambda^k \exp(-\lambda t) \int_{0 < t_1 < \dots < t_k < t} dt_1 \dots dt_k = \lambda^k \exp(-\lambda t) \frac{t^k}{k!}.$$

La variable N_t a donc bien une loi de Poisson de paramètre λt .

Choisissons maintenant une suite $0 < t_1 < \dots < t_n$ et des entiers k_1, \dots, k_n . Cherchons

$$P = \mathbf{P}(N_{t_1} = r_1, N_{t_2} - N_{t_1} = k_1, \dots, N_{t_n} - N_{t_{n-1}} = k_n).$$

Posons $r_1 = k_1$, $r_2 = k_1 + k_2$, \dots , $r_n = k_1 + \dots + k_n$. Par un changement de notations immédiat, la probabilité cherchée est égale à

$$P = \mathbf{P}(N_{t_1} = r_1, N_{t_2} = r_2, \dots, N_{t_n} = r_n).$$

En écrivant la définition de N_t , on a

$$P = \mathbf{P}(T_{r_1} < t_1 < T_{r_1} + 1, T_{r_2} < t_2 < T_{r_2} + 1, \dots, T_{r_n} < t_n < T_{r_n} + 1).$$

Commençons par conditionner par rapport à l'événement $\{N_{t_n} = r_n\} = \{T_{r_n} < t_n < T_{r_n} + 1\}$: la loi conditionnelle de (T_1, \dots, T_{k_n}) sachant cet événement est une statistique d'ordre k_n sur l'intervalle $[0, t_n]$, c'est à dire la loi du réarrangement croissant de k_n variables uniformes indépendantes.

La probabilité conditionnelle que nous cherchons est donc la probabilité que, parmi k_n uniformes indépendantes, il y en ait k_1 dans $[0, t_1]$, k_2 dans $[t_1, t_2]$, \dots , k_n dans $[t_{n-1}, t_n]$. La probabilité pour chaque variable uniforme de tomber dans l'intervalle $[t_{i-1}, t_i]$ est $(t_i - t_{i-1})/t_n = p_i$, (en posant $t_0 = 0$ pour avoir des notations cohérentes), et donc la probabilité que nous cherchons est donnée par la loi multinômiale

$$r_n! \frac{p_1^{k_1}}{k_1!} \frac{p_2^{k_2}}{k_2!} \cdots \frac{p_n^{k_n}}{k_n!}.$$

Finalement, puisque nous connaissons la loi de N_{t_n} , qui est une loi de Poisson de paramètre λt_n , nous avons

$$P = r_n! \exp(-\lambda t_n) \frac{(\lambda t_n)^{r_n}}{r_n!} \prod_1^{n-1} \frac{p_i^{k_i}}{k_i!}.$$

En reprenant la valeur de $p_i = (t_i - t_{i-1})/t_n$, et si on se rappelle que $r_n = k_1 + \dots + k_n$, nous obtenons

$$P = \prod_1^{n-1} \frac{[\lambda(t_i - t_{i-1})]^{k_i} e^{-\lambda(t_i - t_{i-1})}}{k_i!}.$$

Ceci montre que les variables $N_{t_i} - N_{t_{i-1}}$ sont bien indépendantes, de loi de Poisson de paramètre $\lambda(t_i - t_{i-1})$. ■

Corollaire 10.7. *Si Y_1 et Y_2 sont des variables de Poisson indépendantes de paramètre a_1 et a_2 respectivement, alors $Y_1 + Y_2$ suit une loi de Poisson de paramètre $a_1 + a_2$*

Démonstration. — On n'a bien sûr pas besoin du résultat précédent pour démontrer ce résultat facile. (Utiliser la transformée de Fourier!) Mais il est amusant de le voir comme une conséquence de ce qui précède. En effet, si nous choisissons $\lambda = 1$, $a_2 + a_1 = t$, alors (Y_1, Y_2) a même loi que $(N_{a_1}, N_t - N_{a_1})$, et donc $Y_1 + Y_2$ a même loi que N_t . ■

La propriété d'accroissement indépendants permet de voir le comportement asymptotique de N_t :

Corollaire 10.8.

$$\lim_{t \rightarrow \infty} \frac{N_t}{t} = \lambda, \text{ presque sûrement.}$$

Démonstration. — On voit d’abord que N_n/n converge presque sûrement vers λ : en effet, N_n est la somme de n variables aléatoire indépendante de même loi (qui est la loi de Poisson de paramètre λ) et d’espérance λ . On peut donc appliquer la loi des grands nombres.

Ensuite, puisque N_t est croissant, sur l’intervalle $[n, n + 1]$, nous pouvons écrire un encadrement

$$\frac{N_n}{n} \frac{n}{n+1} \leq \frac{N_t}{t} \leq \frac{N_{n+1}}{n+1} \frac{n+1}{n},$$

ce qui permet de conclure à la convergence de $\frac{N_t}{t}$ vers λ . ■

Le calcul de la loi de N_t et des lois jointes de $(N_{t_1}, \dots, N_{t_n})$ semble un peu miraculeux. Nous allons essayer d’expliquer dans ce qui suit les raisons sous-jacentes à ce ”miracle”.

Tout d’abord, introduisons la filtration \mathcal{F}_t des événements antérieurs à t

$$\mathcal{F}_t = \sigma(N_u, u \leq t) = \sigma(N_t; N_u, u \in \mathbb{Q}, u \leq t).$$

(La dernière identité provenant de ce que la fonction $s \mapsto N_s$ est continue à droite, et donc que la connaissance de N_\bullet sur les rationnels de $[0, t[$ est équivalente à celle de N_\bullet sur le même intervalle.)

Nous avons

Proposition 10.9. *Pour toute fonction bornée f de \mathbb{N} dans \mathbb{R} , et tout $s < t$,*

$$\mathbf{E}(f(N_t) / \mathcal{F}_s) = Q_{t-s}(f)(N_s),$$

où

$$Q_t(f)(x) = \mathbf{E}(f(x + N_t)) = \exp(-\lambda t) \sum_n f(x + n) \frac{(\lambda t)^n}{n!}.$$

Cette propriété peut se voir comme la propriété de Markov du processus N_t : la loi de N_t sachant $(N_u, u \leq s)$ est la loi de N_t sachant N_s , et la loi de N_t sachant que $(N_s = k)$ est la loi de $k + N_{t-s}$. C’est un premier exemple de Processus de Markov homogène en temps et en espace.

Démonstration. — Par un argument de classes monotones, il suffit de montrer que, pour tout n -uplet $0 < t_1 < \dots < t_n = s$, et pour toute fonction bornée $F(x_1, \dots, x_n)$,

$$\mathbf{E}[F(N_{t_1}, N_{t_2}, \dots, N_{t_n})f(N_t)] = \mathbf{E}[F(N_{t_1}, N_{t_2}, \dots, N_{t_n})Q_{t-s}f(N_s)].$$

Commençons par faire un changement de variables, en appelant $Y_i = N_{t_{i+1}} - N_{t_i}$, et $Y = N_t - N_s$, et $F(N_{t_1}, N_{t_2}, \dots, N_{t_n}) = G(Y_1, \dots, Y_n)$. Il suffit donc de démontrer que

$$\mathbf{E}[G(Y_1, \dots, Y_n)f(Y_1 + \dots + Y_n + Y)] = \mathbf{E}[G(Y_1, \dots, Y_n)Q_{t-s}(f)(Y_1 + \dots + Y_n)].$$

Mais nous avons vu plus haut que les variables (Y_1, \dots, Y_n, Y) sont toutes indépendantes, et nous connaissons leurs lois.

Donc,

$$\mathbf{E}[f(Y_1 + \dots + Y_n + Y)/(Y_1, \dots, Y_n)] = H(Y_1 + \dots + Y_n),$$

avec $H(x) = \mathbf{E}(f(x + Y))$, par les propriétés élémentaires de l'espérance conditionnelle et des lois conditionnelles. Y étant une variable de Poisson de paramètre $\lambda(t - s)$, nous voyons donc que $H(x) = Q_{t-s}(f)(x)$. ■

Les opérateurs $f \mapsto Q_t(f)$ forment un semigroupe d'opérateurs.

Proposition 10.10. *Pour toute fonction $f : \mathbb{N} \mapsto \mathbb{R}$ bornée, on a $Q_t(Q_s(f)) = Q_{t+s}(f)$, et $\lim_{t \rightarrow 0} \frac{1}{t}[Q_t(f) - f] = \lambda D(f)$, où*

$$D(f)(x) = f(x + 1) - f(x).$$

De même, pour tout t ,

$$\frac{d}{dt}Q_t(f) = \lambda Q_t(D(f)) = \lambda D(Q_t(f)).$$

Démonstration. — Prenons deux variables de Poisson N et N' indépendantes de paramètres λt et λs respectivement. Alors

$$Q_t(Q_s(f))(x) = \mathbf{E}[Q_s(f)(x + N)] = \mathbf{E}[f(x + N + N')],$$

la dernière formule se calculant en conditionnant d'abord par N . Mais nous avons vu que $N + N'$ suit une loi de Poisson de paramètre $\lambda(t + s)$, d'où le résultat.

Ensuite, il suffit de voir que

$$Q_t(f)(x) = \exp(-\lambda t) \sum_k \frac{(\lambda t)^k}{k!} f(x+k).$$

La dérivation terme à terme en 0 dans cette série ne pose aucun problème tant que f est bornée, et nous obtenons le résultat.

Pour la dérivée au point t , il suffit d'écrire que

$$(10.6) \quad Q_{t+s}(f) - Q_t(f) = Q_t(Q_s(f) - f) = Q_s(Q_t)(f) - Q_t(f),$$

puis de diviser par s et de faire converger s vers 0, en utilisant la dérivée en 0. ■

La famille d'opérateurs linéaire $f \mapsto Q_t(f)$ est une famille vérifiant $Q_0 = Id$, $Q_t \circ Q_s = Q_{t+s}$. Formellement, ces opérateurs s'écrivent $Q_t = \exp(tA)$, où A est la dérivée en 0 de Q_t (on appelle A le générateur de Q_t). Ici, nous avons $A = \lambda(T - Id)$, où $T(f)(x) = f(x+1)$. Donc $\exp(t\lambda(T - Id)) = \exp(-\lambda t) \exp(\lambda t T)$. En fait, plus précisément, nous avons

Proposition 10.11. *Si $f : \mathbb{N} \mapsto \mathbb{R}$ est bornée, alors*

$$Q_t(f) = \exp(-\lambda t) \sum_{n=0}^{\infty} \frac{(\lambda t T)^n}{n!} (f).$$

Démonstration. — Immédiatement d'après la définition, on a $T^n(f)(x) = f(x+n)$, et donc la formule énoncée n'est rien d'autre qu'une réécriture de la formule 10.6. ■

10.2 Processus de Markov à temps continu.

Le processus de Poisson que nous venons de décrire semble résulter d'une construction très particulière. Nous allons voir qu'en fait il est caractérisé par sa propriété de Markov, plus le fait qu'il ne croît que par sauts de 1.

Pour cela, nous allons nous intéresser plus généralement aux processus de Markov à temps continu. Ici, nous ne nous intéresserons qu'à ceux qui sont à valeurs dans un ensemble fini ou dénombrable E . Un tel ensemble sera toujours muni de la topologie discrète.

Définition 10.12. Soit E un ensemble fini ou dénombrable, muni de la topologie discrète et de la tribu $\mathcal{P}(E)$ des parties de E .

Soit $(X_t)_{t \geq 0}$ un processus à valeurs dans E , à trajectoires continues à droite avec limites à gauche en tout point. On pose $\mathcal{F}_t^{(X)} = \sigma(X_u, u \leq t)$: c'est la filtration naturelle du processus X .

On dit que (X_t) est un processus de Markov à valeurs dans E si, pour tout $s < t$, et pour toute fonction borélienne bornée f de E dans \mathbb{R} ,

$$\mathbf{E}(f(X_t)/\mathcal{F}_s^{(X)}) = \mathbf{E}(f(X_t)/X_s).$$

De plus, si la loi de X_t sachant X_s ne dépend que de $t - s$, on dit que ce processus est **homogène**.

Dans ce cas, la loi de X_t sachant X_s est donnée par

$$\mathbf{P}(X_t = x / X_s = y) = Q_{t-s}(x, y),$$

où la famille $Q_t(x, y)$ est une famille de noyaux markoviens.

Insistons sur le fait que le fait pour un processus d'être un processus de Markov est une propriété de sa loi (à condition bien sûr qu'il soit continu à droite). En effet,

Lemme 10.13. Soit (X_t) un processus continu à droite avec limites à gauche, à valeurs dans E . Pour que (X_t) soit un processus de Markov, il suffit que, pour tous $0 < t_1 < \dots < t_n < t$, la loi de X_t sachant $(X_{t_1}, \dots, X_{t_n})$ soit celle de X_t sachant X_{t_n} .

Démonstration. — Il suffit de recopier la démonstration faite plus haut pour le processus (N_t) . Remarquons que la donnée de la famille Q_t de noyaux markoviens sur E et de la loi de X_0 détermine entièrement les lois des n -uplets $(X_0, X_{t_1}, \dots, X_{t_n})$. ■

Remarques

1. La condition de continuité à droite est ici facile à comprendre : pour tout ω , la fonction $X_t(\omega)$ reste un certain temps positif à son point de départ x_0 , puis saute à un point x_1 , où elle reste à nouveau un certain temps positif avant de sauter à un point x_2 , etc. La condition de continuité à gauche est plus subtile : elle interdit au processus d'avoir une succession de sauts de plus en plus rapides, s'accumulant en temps fini. Cette hypothèse est inutile si l'espace est fini (elle sera automatiquement vérifiée).

2. La donnée fondamentale est celle de la famille $Q_t(x, y)$ de noyaux markoviens, qui représentent à la fois les probabilités conditionnelles $\mathbf{P}(X_t = y/X_0 = x)$, et $\mathbf{P}(X_{t+s} = y/X_s = x)$. Comme dans le cas du temps discret, on considère donc en fait des familles de processus markoviens issus de tous les points de départ possibles.

Beaucoup de propriétés des chaînes de Markov homogènes vont se répercuter aux processus à temps continu. En effet

Proposition 10.14. *Soit (X_t) un processus de Markov homogène à valeurs dans E . Pour tout entier n , posons $Y_k^{(n)} = X_{k/2^n}$. Alors la suite $(Y_k^{(n)})_{k \geq 0}$ est une chaîne de Markov homogène de noyau $P = Q_{1/2^n}$.*

Démonstration. — Il n’y a rien à démontrer, ou presque. Remarquons que la suite $Y_{\bullet}^{(n)}$ a la propriété de Markov par rapport à une filtration plus grosse que sa filtration naturelle. En effet, si $\mathcal{F}_t^{(X)}$ désigne la filtration naturelle de (X_t) et si $\mathcal{G}_k^{(n)} = \mathcal{F}_{k/2^n}^{(X)}$, alors, pour toute fonction bornée $f : E \mapsto \mathbb{R}$, on a

$$\mathbf{E}(f(Y_{k+1}^{(n)})/G_k^{(n)}) = \mathbf{E}(f(Y_{k+1}^{(n)})/Y_k^{(n)}).$$

■

Lemme 10.15. *Appelons Q_t l’opérateur linéaire*

$$Q_t(f)(x) = \mathbf{E}(f(X_t)/X_0 = x) = \sum_y Q_t(x, y)f(y),$$

qui agit sur les fonctions bornées $f : E \mapsto \mathbb{R}$. On a

$$Q_t \circ Q_s = Q_{t+s}.$$

Démonstration. — C’est la même propriété de semi-groupe que pour le processus de Poisson. On écrit

$$\begin{aligned} \mathbf{E}(f(X_{t+s+u})/\mathcal{F}_u) &= \mathbf{E}(f(X_{t+s+u})/\mathcal{F}_{s+u}/\mathcal{F}_u) \\ &= \mathbf{E}(Q_t(f)(X_{s+u})/\mathcal{F}_u) = Q_s(Q_t(f))(X_u), \end{aligned}$$

puis aussi que

$$\mathbf{E}(f(X_{t+s+u})/\mathcal{F}_u) = Q_{t+s}(f)(X_u).$$

■

Lemme 10.16. *Soit (X_t) un processus de Markov homogène en temps ; soit $X_{t-}(\omega)$ la limite à gauche de $s \mapsto X_s(\omega)$ au point t . Pour tout t ,*

$$\mathbf{P}(X_t = X_{t-}) = 1.$$

Démonstration. — Fixons (x, y) dans $E \times E$, et calculons $\mathbf{P}(X_{t-} = y, X_t = x)$. Fixons $t > 0$: puisque la topologie sur E est discrète par hypothèse, nous savons que $X_{t-1/n}(\omega) = X_{t-}(\omega)$ à partir d'un certain rang.

$$\text{Donc } \{X_t = x ; X_{t-} = y\} = \liminf_n \{X_t = x ; X_{t-1/n} = y\}.$$

Mais, par homogénéité,

$$\mathbf{P}(X_t = x, X_{t-1/n} = y / X_0 = z) = \mathbf{P}(X_{t+1/n} = x; X_t = y / X_{1/n} = z).$$

Donc, pour tout point z tel que $\mathbf{P}(X_0 = z) > 0$, on a

$$\mathbf{P}(X_t = x, X_{t-1/n} = y, X_0 = z) = \mathbf{P}(X_{t+1/n} = x; X_t = y, X_{1/n} = z) \frac{\mathbf{P}(X_{1/n} = z)}{\mathbf{P}(X_0 = z)}.$$

Par continuité à droite de X au point t et au point 0, ceci converge lorsque $n \rightarrow \infty$ vers $\mathbf{P}(X_t = x; X_t = y / X_0 = z)$, c'est à dire vers 0 si $y \neq x$. Et donc, en sommant en z , nous voyons que $\mathbf{P}(X_t = x; X_{t-} = y) = 0$ si $y \neq x$.

Ceci signifie que si $T(\omega)$ est un temps de saut de la fonction $X_t(\omega)$, alors la loi de T ne charge pas les points. ■

On déduit de cette propriété une propriété de régularité des noyaux Q_t (et c'est le seul point où on se sert de la continuité à gauche des trajectoires).

Lemme 10.17. *Pour toute fonction bornée f , et tout $x \in \mathbb{N}$, $t \mapsto Q_t(f)(x)$ est continue, et, si $s < t$,*

$$\mathbf{E}(f(X_t) / \mathcal{F}_s^{(X)}) = Q_{t-s}(f)(X_s).$$

Démonstration. — La fonction $t \mapsto X_t(\omega)$ étant continue à droite et f étant bornée, la continuité à droite de Q_t est une conséquence du théorème de convergence dominée. Pour la continuité à gauche, c'est une conséquence du lemme précédent : t étant fixé, $X_t = X_{t-}$ presque sûrement, et donc le même argument s'applique.

(En fait, si l'espace E est fini, nous verrons plus bas que pour une telle famille d'opérateurs, la continuité à droite en 0 suffit à assurer la continuité

des deux côtés, partout, et même le caractère analytique des fonctions $t \mapsto Q_t(x, y)$.)

Le seconde propriété n'est rien d'autre que la propriété de Markov. ■

Posons $F_t^{(X)} = \sigma(X_u, u \leq t)$: c'est une famille croissante de sous-tribus de \mathcal{A} . Nous allons d'abord montrer qu'aux ensembles de mesure nulle près, la famille croissante de tribus $\mathcal{F}_t^{(X)}$ est continue à droite. Plus précisément :

Proposition 10.18. *(Loi du 0 – 1.) Soit $\mathcal{F}_\infty = \vee_t \mathcal{F}_t^{(X)}$ et \mathcal{N} l'ensemble des négligeables de \mathcal{F}_∞ . (Un ensemble est dans \mathcal{N} s'il est inclus dans un événement B de \mathcal{F}_∞ de probabilité nulle.)*

Soit $\mathcal{F}_t = \sigma(F_t^{(X)}, \mathcal{N})$. Alors, pour tout $t \geq 0$,

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

Démonstration. — Appelons $\mathcal{F}_{t+}^{(X)} = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}^{(X)}$. Il suffit en fait de montrer qu'une variable bornée $\mathcal{F}_{t+}^{(X)}$ mesurable est égale presque sûrement à une variable $\mathcal{F}_t^{(X)}$ mesurable. En effet, un argument de classe monotones montre que, si \mathcal{A} est une tribu et \mathcal{N} une famille d'ensembles négligeables, alors toute variable bornée mesurable par rapport à la tribu $\sigma(\mathcal{A}, \mathcal{N})$ est presque sûrement égale à une variable \mathcal{A} -mesurable.

Donc, considérons une variable bornée Z \mathcal{F}_∞ -mesurable : il suffit de montrer que $\mathbf{E}(Z/\mathcal{F}_{t+}^{(X)}) = \mathbf{E}(Z/\mathcal{F}_t^{(X)})$. Par un argument de classes monotones, il suffit de le montrer pour une variable bornée Z de la forme $F_1(X_{t_1}) \dots F_n(X_{t_n})$. (On suppose les points t_i rangés dans l'ordre croissant.) Il ne coûte rien de supposer que t fait partie de la liste des t_i , soit $t = t_p$. Par la propriété de loi conditionnelle, on voit que, pour une telle variable Z , $\mathbf{E}(Z/\mathcal{F}_s^{(X)})$ est de la forme

$$F_1(X_{t_1}) \dots F_n(X_{t_r}) H(X_s),$$

où r est l'indice tel que $t_r \leq s < t_{r+1}$. (On conditionne successivement par rapport aux tribus $\mathcal{F}_{t_i}^{(X)}$ en partant du plus grand indice).

Donc, en conditionnant tout d'abord par rapport à $\mathcal{F}_{t_{p+1}}^{(X)}$, on se ramène au cas où

$$Z = F_1(X_{t_1}) \dots F_p(X_{t_p}) F(X_{t_{p+1}}).$$

Posons $Z_1 = F_1(X_{t_1}) \dots F_p(X_{t_p})$ et $t_{p+1} = s$, pour simplifier les notations.

Choisissons une suite s_n qui décroît vers t , de telle façon que $\cap_n \mathcal{F}_{s_n}^{(X)} = \mathcal{F}_{t+}^0$. Appelons $\mathcal{G}_n = \mathcal{F}_{s_n}^{(X)}$. Le théorème de convergence des martingales inverses nous dit que $\mathbf{E}(Z/\mathcal{G}_n)$ converge presque sûrement vers $\mathbf{E}(Z/\mathcal{F}_{t+}^{(X)})$.

Mais par ailleurs, nous pouvons explicitement calculer cette espérance conditionnelle pour la variable Z qui nous intéresse :

$$\mathbf{E}(Z/\mathcal{G}_n) = Z_1 Q_{s-s_n}(F)(X_{s_n}).$$

La suite $Q_{s-s_n}(F)(X_{s_n})$ converge presque sûrement vers $Q_{s-t}(X_t)$ lorsque $s_n \rightarrow t$, puisque $t \rightarrow Q_t$ est continue et que X_t est continue à droite, constante par morceaux.

D'autre part, $Z_1 Q_{s-t}(F)(X_t) = \mathbf{E}(Z/\mathcal{F}_t^0)$.

D'où le résultat. ■

Remarque. — Il faut faire attention ici que cette propriété ne provient pas du fait que la famille (X_t) est continue à droite. Voici un exemple très simple de processus X_t qui est continu et qui n'a pas cette propriété : si ε est une variable de Bernoulli, si on pose $X_t = t\varepsilon$; alors $\sigma(X_u, u \leq t) = \sigma(\varepsilon)$ pour tout $t > 0$, tandis que $X_0 = 0$ et $\mathcal{F}_0 = \{\Omega, \emptyset\}$ (exercice!).

Enfin, avant d'aller plus loin, nous étudions dans le cas fini la structure des semigroupes de noyaux Markoviens.

Théorème 10.19. *Soit E un ensemble fini et $(Q_t(x, y))_{t \geq 0}$ une famille de matrices markoviennes satisfaisant à la relation $Q_t Q_s = Q_{t+s}$, $Q_0 = Id$. Supposons qu'en plus Q soit continue en 0. Alors, il existe une matrice $A(x, y)$, satisfaisant*

$$\begin{cases} A(x, y) \geq 0 & \text{si } x \neq y, \\ \forall x \in E, \quad \sum_{y \in E} A(x, y) = 0, \end{cases}$$

telle que $Q_t = \exp(tA)$.

Réciproquement, si la matrice A vérifie ces deux conditions, alors pour tout $t \geq 0$, $\exp(tA)$ est une matrice markovienne.

Cette matrice A s'appelle le *générateur du processus*.

Une matrice A vérifiant ces deux propriétés s'appelle un *générateur markovien*.

Remarque. — Un exemple de générateur markovien est donné par $P - Id$,

où P est une matrice markovienne. On verra plus bas que ceci correspond à des processus de Markov très simples.

Démonstration. — La topologie que nous avons mise sur les matrices pour parler de continuité à droite en 0 est la topologie de \mathbb{R}^{n^2} . On aura intérêt à mettre une norme pour la définir qui soit une norme d'opérateurs, c'est à dire telle que $\|AB\| \leq \|A\|\|B\|$, par exemple

$$\|A\| = \sup_x \sum_y |A(x, y)|,$$

qui a l'avantage de donner la norme 1 aux matrices markoviennes.

D'autre part, nous identifions des matrices à des opérateurs linéaires agissant sur l'espace vectoriel F de dimension finie des fonctions de E dans \mathbb{R} , auquel cas, pour la norme décrite plus haut, $\|A\| = \sup_{\|f\|_\infty \leq 1} \|A(f)\|_\infty$.

Commençons par remarquer que la famille Q_t est partout continue à droite. En effet, $Q_{t+s} - Q_t = [Q_s - Id] \circ Q_t$, et donc la continuité à droite en 0 l'entraîne partout.

Mais il en va de même de la continuité à gauche car

$$\|Q_{t-s} - Q_t\| = \|Q_{t-s}(Id - Q_s)\| \leq \|Id - Q_s\| \rightarrow 0 \quad (s \rightarrow 0).$$

Pour exhiber la matrice A , nous allons voir que Q_t est dérivable partout, que sa dérivée à droite en 0 est A , et que $\frac{d}{dt}Q_t = AQ_t = Q_tA$.

Puisque la famille Q_t de matrices est continue, nous pouvons considérer pour $\lambda > 0$, les matrices

$$R_\lambda = \int_0^\infty \exp(-\lambda s) Q_s ds,$$

où l'intégrale peut être comprise coefficient par coefficient, chaque intégrale de coefficients étant convergente puisque tous les coefficients sont bornés par 1.

Nous avons

$$(\mu - \lambda)R_\lambda R_\mu = R_\lambda - R_\mu.$$

En effet, pour $\lambda \neq \mu$, nous écrivons

$$\begin{aligned}
 R_\lambda R_\mu &= \int_0^\infty \int_0^\infty \exp(-\lambda s) \exp(-\mu t) Q_s Q_t \, ds \, dt \\
 &= \int_0^\infty \int_0^\infty \exp(-\lambda s) \exp(-\mu t) Q_{s+t} \, ds \, dt \\
 &= \int_0^\infty \int_0^u \exp(-\lambda s) \exp(-\mu(u-s)) Q_u \, du \, ds \\
 &= \int_0^\infty \exp(-\mu u) Q_u \left[\int_0^u \exp(-(\lambda-\mu)s) \, ds \right] \, du \\
 &= \int_0^\infty \exp(-\mu u) Q_u \frac{1 - \exp(-(\lambda-\mu)u)}{\lambda - \mu} \, du \\
 &= \frac{R_\mu - R_\lambda}{\lambda - \mu}.
 \end{aligned}$$

Nous pouvons transformer cette équation pour voir que

$$R_\mu = R_\lambda [Id + (\mu - \lambda)R_\mu],$$

où l'on voit que l'image de R_μ est incluse dans celle de R_λ . Comme ceci reste vrai en échangeant λ et μ , nous voyons que l'image de R_λ est indépendante de λ .

Appelons F_0 le sous espace vectoriel image commune de toutes les matrices R_λ . Lorsque $\lambda \rightarrow \infty$, nous voyons que λR_λ converge vers l'identité. En effet

$$\lambda R_\lambda = \int_0^\infty \exp(-s) Q_{s/\lambda} \, ds.$$

Or, $Q_{s/\lambda}$ converge vers l'identité par continuité à droite de Q_t en 0, et il nous suffit d'appliquer le théorème de convergence dominée, tous les coefficients des matrices Q_s sont bornés par 1.

Donc, si f est un élément de F , $f = \lim_{\lambda \rightarrow \infty} \lambda R_\lambda(f)$, et donc f est limite d'éléments de F_0 . Ce sous-espace est donc dense dans F , et puisqu'en dimension finie les sous-espaces vectoriels sont fermés, $F_0 = F$ et toutes les applications R_λ sont surjectives.

Maintenant, si $f = R_1(g)$, alors

$$\begin{aligned} Q_t(f) &= \int_0^\infty \exp(-s)Q_{t+s}(g) ds \\ &= \int_t^\infty \exp(-(u-t))Q_u(g) du \\ &= \exp(t)[f - \int_0^t \exp(-u)Q_u(g) du] \end{aligned}$$

En particulier,

$$\frac{Q_t(f) - f}{t} = \frac{e^t - 1}{t} f - \frac{e^t}{t} \int_0^t e^{-u} Q_u(g) du.$$

Ceci converge vers $f - g$, lorsque t converge vers 0, et $Q_t(f)$ est donc dérivable en 0, de dérivée $f - g$.

Appelons $A(f) = \frac{d}{dt}Q_t(f)_{t=0}$.

Nous pouvons voir que Q_t est dérivable partout à droite en écrivant

$$\frac{Q_{t+s} - Q_t}{s}(f) = Q_t\left(\frac{Q_s(f) - f}{s}\right) \rightarrow Q_t A(f).$$

Par ailleurs, $Q_{t+s}(f) - Q_t(f)$ s'écrit aussi $Q_s(Q_t(f)) - Q_t(f)$ et ceci, divisé par s , converge vers $AQ_t(f)$, d'où l'on tire

$$\frac{d}{dt}Q_t = AQ_t = Q_t A.$$

La dérivée ici étant la dérivée à droite.

D'autre part, cette dérivée est aussi une dérivée à gauche puisque

$$\frac{1}{s}(Q_{t-s} - Q_t) = Q_{t-s}\left(\frac{Id - Q_s}{s}\right) \rightarrow -Q_t A \quad (s \rightarrow 0),$$

par continuité à gauche de Q_\bullet au point t .

Nous voulons en déduire que $Q_t = \exp(tA)$, où par définition

$$\exp(tA) = \sum_n \frac{t^n}{n!} A^n,$$

cette série étant convergente puisque $\|A^n\| \leq \|A\|^n$. Pour cela, posons $M_t = \exp(tA)$, qui vérifie aussi

$$\frac{d}{dt}M_t = AM_t = M_tA.$$

Écrivons

$$\begin{aligned} Q_t - M_t &= \int_0^t \frac{d}{ds}[Q_s M_{t-s}] ds \\ &= \int_0^t [Q'_s M_{t-s} - Q_s M'_{t-s}] ds \\ &= \int_0^t [Q_s A M_{t-s} - Q_s A M_{t-s}] ds = 0 \end{aligned}$$

D'autre part les propriétés de la matrice A sont faciles à vérifier.

La première provient de ce que, pour $x \neq y$, $Q_t(x, y) \geq 0$ et $Q_0(x, y) = 0$.

La seconde provient de ce que, pour tout t et tout x , $\sum_y Q_t(x, y) - Q_0(x, y) = 0$.

Réciproquement, si A est une matrice vérifiant les deux propriétés (un générateur markovien), alors $A(\mathbf{1}) = 0$, si $\mathbf{1}$ désigne le vecteur constant égal à 1, et donc $\exp(tA)(\mathbf{1}) = \mathbf{1}$.

D'autre part, il existe une constante k telle que $A+kId$ a tous ses coefficients positifs. Il en va de même de

$$\exp(t(A + kId)) = \sum_n \frac{[t(A + kId)]^n}{n!},$$

qui est la somme d'une série (convergente) de matrices à coefficients positifs, et qui est donc aussi à coefficients positifs.

Mais $\exp(t(A + kId)) = \exp(kt) \exp(tA)$, et nous voyons donc que $\exp(tA)$ est aussi une matrice à coefficients positifs. ■

Remarque. — La finitude du cardinal de E nous a servis ici pour s'assurer que la famille de matrices (Q_t) est dérivable en 0. Si le cardinal de E avait été infini, nous n'aurions obtenu cette dérivabilité dans les bons cas que sur un sous-espace vectoriel dense de fonctions. Les choses ont alors beaucoup plus compliquées.

10.3 Filtrations à temps continu.

A partir de maintenant, nous considérerons toujours cette famille \mathcal{F}_t de sous-tribus, croissante et continue à droite (c'est à dire que, pour tout t , $\bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t$). Une telle famille croissante de tribus est une **filtration** à temps continu. On peut développer pour ces filtrations les mêmes outils que ce qu'on avait développé en temps discret : *temps d'arrêt*, tribu \mathcal{F}_T , etc. La plupart des résultats du temps discret restent valables en temps continu, mais il y a quelques subtilités auxquelles il est bon de faire attention.

Définition 10.20. 1. On dit qu'un processus X_t est **adapté** si, pour tout $t \geq 0$, la variable X_t est \mathcal{F}_t mesurable.
 2. Soit $(\mathcal{F}_t)_{t \geq 0}$ une famille croissante de sous-tribus de \mathcal{A} , continue à droite. et soit $T : \Omega \mapsto \mathbb{R}_+ \cup \infty$ une variable aléatoire. On dit que T est un **temps d'arrêt** si et seulement,

$$\forall t \geq 0, \{T \leq t\} \in \mathcal{F}_t.$$

3. Si T est un temps d'arrêt, alors la tribu \mathcal{F}_T est définie par

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty, \forall t \in \mathbb{R}_+, A \cap \{T \leq t\} \in \mathcal{F}_t\}.$$

C'est la prolongation naturelle de la définition que nous avons adoptée pour les temps discrets.

Les théorèmes énoncés pour les filtrations à temps discret s'étendent avec quelques précautions aux processus à temps continu.

Tout d'abord, quelques propriétés élémentaires qui sont de simples extensions des propriétés équivalentes pour le temps discret, mais dont la démonstration demande un peu de précision.

Proposition 10.21. $T : \Omega \mapsto [0, \infty]$ est un temps d'arrêt si et seulement si, pour tout $t \in \mathbb{R}_+$, $\{T < t\} \in \mathcal{F}_t$.

Si S et T sont deux temps d'arrêt et si $A \in \mathcal{F}_T$, alors $A \cap \{T \leq S\} \in \mathcal{F}_S$.

En particulier, si $T \leq S$, alors $\mathcal{F}_T \subset \mathcal{F}_S$.

Démonstration. — Commençons par le premier point : $\{T < t\} = \bigcup_{r \in \mathbb{Q}, r < t} \{T \leq r\}$, et donc si T est un temps d'arrêt, $\{T < t\} \in \mathcal{F}_t$. Réciproquement, supposons que pour tout s , $\{T < s\} \in \mathcal{F}_s$; on a $\{T \leq t\} = \bigcap_n \{T < t + \frac{1}{n}\}$.

Les ensembles A_n décroissent, ils sont dans $\mathcal{F}_{t+1/n}$ par hypothèse, donc leur intersection est dans $\bigcap_n \mathcal{F}_{t+1/n} = \mathcal{F}_t$ par continuité à droite des tribus.

Ensuite, nous écrivons

$$A \cap \{T \leq S\} \cap \{S \leq t\} = A \cap \{T \leq t\} \cap [\{T \leq S\} \cap \{S \leq t\}].$$

Le premier ensemble est dans \mathcal{F}_t par définition de \mathcal{F}_T , et pour le second, nous écrivons

$$\{T \leq S\} \cap \{S \leq t\} = \bigcap_n [\{T < S + \frac{1}{n}\} \cap \{S \leq t\}].$$

Posons $A_n = \{T < S + \frac{1}{n}\} \cap \{S \leq t\}$. C'est une suite décroissante d'ensembles. Nous avons

$$\{T < S + \frac{1}{n}\} \cap \{S \leq t\} = \bigcup_{r \in \mathbb{Q}} \{T \leq r\} \cap [\{r < S + \frac{1}{n}\} \cap \{S \leq t\}].$$

Or, pour tout r rationnel, l'événement $\{T \leq r\} \cap [\{r < S + \frac{1}{n}\} \cap \{S \leq t\}]$ est vide si $r > t + 1/n$, et $\mathcal{F}_r \cap \mathcal{F}_t$ mesurable sinon. C'est donc dans tous les cas un événement $\mathcal{F}_{t+1/n}$ mesurable. Donc A_n est $\mathcal{F}_{t+1/n}$ mesurable.

L'intersection décroissante des A_n est donc mesurable par rapport à l'intersection des tribus $\mathcal{F}_{t+1/n}$, qui est \mathcal{F}_t par continuité à droite de la filtration.

La second point est immédiat. ■

Donnons ensuite un petit résultat technique qui étend la continuité à droite des tribus à des *temps d'arrêt*.

Lemme 10.22. *Soit T_n une suite décroissante de temps d'arrêt, et soit T leur limite. Alors T est un temps d'arrêt et $\mathcal{F}_T = \bigcap_n \mathcal{F}_{T_n}$.*

Démonstration. — Tout d'abord, T est un *temps d'arrêt*, puisque $\{T < t\} = \bigcap_n \{T_n < t\}$. (Ceci ne serait pas vrai avec les inégalités $\{T \leq t\}$).

Ensuite, nous savons que $\mathcal{F}_T \subset \mathcal{F}_{T_n}$ si $T \leq T_n$, et donc $\mathcal{F}_T \subset \bigcap_n \mathcal{F}_{T_n}$.

Soit $A \in \bigcap_n \mathcal{F}_{T_n}$ et soit $t \in \mathbb{R}_+$. Nous voulons montrer que $A \cap \{T \leq t\} \in \mathcal{F}_t$. Posons $\varepsilon > 0$.

$$A \cap \{T < t + \varepsilon\} = \bigcup_n A \cap \{T_n < t + \varepsilon\} \in \mathcal{F}_{t+\varepsilon},$$

la dernière assertion provenant de ce que chaque T_n est un *temps d'arrêt* et que $A \cap \{T_n < t\} = A \cap \{T_n \leq t\} \cap \{T_n < t\}$. Alors, en posant $\varepsilon = \frac{1}{m}$, et

$A_m = A \cap \{T < t + \frac{1}{m}\}$, nous voyons que la suite A_m est décroissante, son intersection est donc élément de $\cap_m \mathcal{F}_{t+1/m} = \mathcal{F}_t$ et $\cap_m A_m = A \cap \{T \leq t\}$.

C'est ce que nous voulions démontrer. ■

Nous pouvons maintenant démontrer les analogues des propriétés obtenues en temps discret :

Proposition 10.23. *Soit X_t un processus adapté, à valeurs dans un espace topologique E , continu à droite (dans la pratique, E sera \mathbb{R} ou bien un ensemble fini ou dénombrable muni de la topologie discrète).*

Soit A un ouvert de \mathbb{R} . Alors, $T = \inf\{s \geq 0, X_s \in A\}$ est un temps d'arrêt.

Si T est un temps d'arrêt, $X_T(\omega) = X_{T(\omega)}(\omega)$ est \mathcal{F}_T mesurable.

Démonstration. — Commençons par montrer que le temps T est un temps d'arrêt. (Comme d'habitude, $\inf\{\emptyset\} = \infty$.) Puisque X_t est continu à droite et que A est ouvert, alors

$$\{T \leq t\} = \{X_t \in A\} \cup \{\exists s < t, s \in \mathbb{Q}, X_s \in A\}.$$

On voit directement sur cette décomposition que cet événement est dans \mathcal{F}_t . Remarquons que la continuité à droite nous sert à nous ramener à une famille dénombrable d'instant s .

Remarquons aussi que la propriété pour T d'être un temps d'arrêt est équivalente au fait que le processus continu à droite $\mathbf{1}_{[T, \infty[}(t) = \mathbf{1}_{T \leq t}$ est adapté.

Pour montrer que X_T est \mathcal{F}_T -mesurable, nous allons utiliser une astuce qui consiste à se ramener au temps discret.

Soit T un temps d'arrêt, et Δ_n l'ensemble des dyadiques d'ordre n sur \mathbb{R}_+ , c'est à dire l'ensemble des nombres de la forme $k/2^n$, $k \in \mathbb{N}$. Posons $T_n = \inf\{s \in \Delta_n, s \geq T\}$. C'est un temps d'arrêt : $\{T_n \leq t\} = \cup_{k/2^n \leq t} \{T \leq k/2^n\}$. Cette suite T_n est donc une suite de temps d'arrêt qui converge vers T en décroissant (c'est l'approximation dyadique par excès de T).

Mais, $S_n = 2^n T_n$ est à valeurs dans $\mathbb{N} \cup \{\infty\}$, et c'est un temps d'arrêt par rapport à la filtration (à temps discret) $\mathcal{G}_k^{(n)} = \mathcal{F}_{k/2^n}$. Le processus $Y_k^{(n)} = X_{k/2^n}$ est adapté pour cette filtration $\mathcal{G}_k^{(n)}$. Donc, Y_{S_n} est mesurable par rapport à la tribu $\mathcal{G}_{S_n}^{(n)}$ dont on voit immédiatement qu'elle est égale à \mathcal{F}_{T_n} .

Par continuité à droite de (X_t) , nous voyons que la suite Y_{S_n} converge vers X_T , qui est donc \mathcal{F}_{T_n} mesurable pour tout n , et par suite \mathcal{F}_T mesurable.

■

10.4 La propriété de Markov forte.

Comme dans le cas fini, nous pouvons nous placer sur l'espace canonique Ω des fonctions continues à droite avec limites à gauche de \mathbb{R}_+ dans E . On munit cet espace de la plus petite tribu qui rend mesurables les applications coordonnées $X_t(\omega) = \omega(t)$. Etant donné une famille de noyaux Markoviens $Q_t(x, y)$, nous pouvons considérer une probabilité \mathbf{P}_x sur cet espace mesurable qui fait du processus des coordonnées $X_t(\omega) = \omega(t)$ un processus de Markov de noyau Q_t et de loi initiale δ_x .

Il faut bien sûr des conditions sur Q_t pour pouvoir construire de telles mesures, en particulier il faut la condition de semigroupe $Q_{t+s} = Q_t \circ Q_s$, mais ce n'est pas la seule.

Dans ce qui suit, nous supposons toujours qu'on s'est donné un processus de Markov à trajectoires à valeurs continues à droites et avec limites à gauche et de noyaux Q_t , et la mesure \mathbf{P}_x n'est rien d'autre que la loi de $(X_t)_{t \geq 0}$ sachant que $X_0 = x$. Cette construction est toujours possible si l'espace E est fini et que $Q_t = \exp(tA)$, où A est un générateur markovien.

D'autre part, sur cet espace canonique, nous pouvons considérer l'opérateur de translation θ_t défini par

$$\theta_t(\omega)_s = \omega_{t+s},$$

que nous étendons comme dans le cas du temps discret en $\theta_t(Z)(\omega) = Z(\theta_t(\omega))$, ainsi qu'aux *temps d'arrêt* finis presque sûrement (ou aux *temps d'arrêt* infinis sur $\{T < \infty\}$, par $\theta_T Z(\omega) = Z(\theta_{T(\omega)}(\omega))$.

Nous pouvons alors énoncer le théorème d'arrêt

Théorème 10.24. *Soit (X_t) un processus de Markov sur E fini ou dénombrable, soit T un temps d'arrêt et soit Z une variable aléatoire bornée. Alors*

$$\mathbf{E}(\theta_T(Z)\mathbf{1}_{T < \infty} / \mathcal{F}_T) = \mathbf{E}_{X_T}(Z)\mathbf{1}_{T < \infty}.$$

Démonstration. — Comme d'habitude, par un argument de classes monotones, on peut se ramener au cas où $Z = f_1(X_{t_1}) \dots f_k(X_{t_k})$.

On peut même faire un peu mieux ici : on se ramène au cas où les t_i sont des nombres dyadiques : en effet, par continuité à droite du processus (toutes les fonctions f_i étant continues pour la topologie discrète), si dans une fonction

Z de la forme précédente, on remplace les temps t_i par leurs approximations dyadiques d'ordre n par excès, on obtient une variable Z_n qui converge vers Z , presque sûrement. Si la formule est démontrée pour Z_n , alors

$$\mathbf{E}(Z/\mathcal{F}_T) = \lim_n \mathbf{E}(Z_n/\mathcal{F}_T), \quad \mathbf{E}_{X_T}(Z) = \lim_n \mathbf{E}_{X_T}(Z_n),$$

les deux convergences étant des conséquences du théorème de convergence dominée.

Ensuite, tous les temps t_i intervenant dans la définition de Z étant dyadiques d'ordre k , appelons T_n l'approximation dyadique d'ordre n par excès de T , avec $k \geq n$. Nous avons déjà vu que le processus X considéré aux temps dyadiques est une chaîne de Markov, et que le temps T_n est un *temps d'arrêt* pour cette chaîne. Nous pouvons donc appliquer la propriété de Markov forte des chaînes, et nous obtenons

$$\mathbf{E}(Z\mathbf{1}_{T_n < \infty}/\mathcal{F}_{T_n}) = \mathbf{E}_{X_{T_n}}(Z)\mathbf{1}_{T_n < \infty}.$$

Il reste à passer à la limite. Remarquons que $\{T_n < \infty\} = \{T < \infty\}$: tout d'abord, nous avons vu que $\cap_n \mathcal{F}_{T_n} = \mathcal{F}_T$, et le membre de gauche converge vers $\mathbf{E}(Z/\mathcal{F}_T)$ par le théorème de convergence des martingales inverses. Quand au membre de droite, X_{T_n} étant égal à X_T à partir d'un certain rang, il converge vers $\mathbf{E}_{X_T}(Z)$.

■

Théorème 10.25. *Soit X_t un processus de Markov à valeurs dans E fini ou dénombrable. Soit T un temps d'arrêt presque sûrement fini de la filtration \mathcal{F}_t . Alors, le processus X_{T+s} est un processus de Markov de même noyau de transition que X_t , de loi initiale μ_T , où μ_T est la loi de X_T , et conditionnellement indépendant de \mathcal{F}_T sachant X_T .*

Démonstration. — On fait exactement comme dans la démonstration du théorème équivalent en temps discret : la propriété de Markov forte montre que le processus $(Y_t) = (X_{T+t})$ est un processus de Markov de mêmes noyaux que (X_t) .

■

Comme application, nous pouvons voir que les temps de saut du processus suivent des lois exponentielles :

Théorème 10.26. *Soit $T = \inf\{t \mid X_t \neq X_0\}$. Alors T suit sous \mathbf{P}_x une loi exponentielle de paramètre $\lambda(x)$.*

Démonstration. — Remarquons que T est presque sûrement non nul, puisque les trajectoires sont continues à droite pour la topologie discrète. Appelons $F(t) = \mathbf{P}_x(T > t)$; c'est une fonction continue à droite. Soit A_t l'événement $\{T > t\}$. Nous avons

$$A_{t+s} = A_t \cap \theta_t(A_s).$$

En appliquant la propriété de Markov, nous avons

$$F(t+s) = \mathbf{E}_x[\mathbf{1}_{A_t} \mathbf{E}_{X_t}(A_s)].$$

Mais sur A_t , $X_t = x$, et donc $F(t+s) = F(t)F(s)$. Nous avons déjà vu plus haut que ceci caractérise les lois exponentielles, c'est à dire que $F(t) = \exp(-\lambda(x)t)$.

■

Remarque. — Il se peut que $\lambda(x)$ soit nul, auquel cas T est presque sûrement infini : le point x est alors une trappe. On peut aussi voir que le temps T est soit presque sûrement infini, soit presque sûrement fini.

De même, les temps de passages successifs dans les différents points sont exponentiels, tant qu'ils sont finis :

Proposition 10.27. *Soit $T_n = \inf\{t > T_{n-1} \mid X_t \neq X_{T_{n-1}}\}$, avec $T_n = \infty$ si $T_{n-1} = \infty$. Alors, $T_n - T_{n-1}$ est conditionnellement indépendante de $\mathcal{F}_{T_{n-1}}$ sachant X_{T_n} , et la loi de $T_n - T_{n-1}$ sachant que $X_{T_{n-1}} = x$ est une loi exponentielle de paramètre $\lambda(x)$, .*

Démonstration. —

La condition d'indépendance conditionnelle, s'écrit : pour toute fonction F borélienne bornée de $[0, \infty]$ dans \mathbb{R} ,

$$\mathbf{E}(F(T_n - T_{n-1})/\mathcal{F}_{T_{n-1}}) = \mathbf{E}(F(T_n - T_{n-1})/X_{T_{n-1}}).$$

Il suffit d'appliquer la propriété de Markov forte au temps T_{n-1} , tant qu'il est fini : $T_n - T_{n-1} = \theta_{T_{n-1}}(T_1)$: il est soit fini presque sûrement, soit presque sûrement infini. La démonstration est immédiate. ■

Avec ces résultats, nous pouvons maintenant donner une caractérisation simple du processus de Poisson :

Théorème 10.28. *Soit $(S_t)_{t \geq 0}$ un processus de Markov, à valeurs dans \mathbb{N} , ayant les propriétés suivantes :*

1. $S_0 = 0$.

2. $\forall t > 0, \mathbf{P}(S_t = 0) < 1$.
3. Pour tout ω , les fonctions $t \mapsto S_t(\omega)$ sont continues à droite, croissant par sauts de 1.
4. Pour tout $s < t$, la loi de S_t sachant que $X_s = k$ est la loi de $k + X_{t-s}$.

Alors, le processus S_t est un processus de Poisson. Plus exactement, pour un certain $\lambda > 0$, pour le processus de Poisson N_t d'intensité λ , et $\forall t_1 < t_2 \dots < t_n$, les lois jointes de $(N_{t_1}, \dots, N_{t_n})$ et de $(S_{t_1}, \dots, S_{t_n})$ sont les mêmes.

Démonstration. — Remarquons tout d'abord qu'un tel processus S_t a automatiquement des limites à gauche. Tout d'abord, la loi du temps

$$T_1 = \inf\{t \mid S_t > 0\}$$

est une loi exponentielle de paramètre $\lambda > 0$, puisque le processus ne reste pas en 0.

Puisque le processus ne croît que par sauts de 1, $X_{T_1} = 1$, et le processus $S'_t = S_{T_1+t} - 1$ a même loi que le processus S_t . En effet, la loi de $(\hat{S}_t) = (S_{T_1+t})$ est un processus de Markov de mêmes noyaux que S_t et de loi initiale δ_1 , et la loi de S_t sachant que $S_0 = 1$ est la même que la loi de $S_t - 1$ sachant que $S_0 = 0$.

Posons $T_2 = \inf\{n \mid S_t > 1\}$. Nous voyons que $T_2 - T_1$ suit la même loi que T_1 , et c'est donc une exponentielle de même paramètre λ . De plus, $T_2 - T_1$ est conditionnellement indépendant de \mathcal{F}_{T_1} sachant X_{T_1} . Mais X_{T_1} est constant, et donc $T_2 - T_1$ est indépendant de \mathcal{F}_{T_1} . De plus, $X_{T_2} = 2$, puisque S_\bullet ne croît que par sauts de 1.

On peut donc recommencer le raisonnement, et montrer que le troisième temps de saut T_3 est tel que $T_3 - T_2$ est un temps exponentiel de paramètre λ , indépendant de \mathcal{F}_{T_2} , et ainsi de suite pour les temps de sauts successifs.

Nous voyons ainsi que les temps de sauts T_n de S_t sont des sommes de variables aléatoires exponentielles de même paramètre λ , indépendantes. Puisque le processus ne fait que croître par sauts de 1, alors sa valeur au temps t est le nombre de sauts qui précèdent t , c'est à dire que

$$S_t = \sum_n \mathbf{1}_{\{T_n \leq t\}}.$$

On reconnaît là la définition du processus de Poisson.

■

Ces résultats nous permettent d'obtenir une propriété remarquable du processus de Poisson

Proposition 10.29. *Soit N_t un processus de Poisson d'intensité λ , et soit T un temps d'arrêt presque sûrement fini. Alors, $N_{T+t} - N_T$ est encore un processus de Poisson d'intensité λ .*

Démonstration. — Le processus construit ainsi est un processus de Markov homogène à valeurs dans \mathbb{N} , ne croissant que par sauts de 1. C'est donc un processus de Poisson. On obtient son intensité en regardant la loi de son premier temps de saut. ■

10.5 Construction des processus dans le cas fini.

Nous nous donnons dans ce paragraphe un générateur Markovien A sur un ensemble fini, et nous nous proposons de construire le processus à partir de A . Nous posons $Q_t = \exp(tA)$.

Nous commençons par décrire ce qu'est la matrice A pour le processus X_t .

Proposition 10.30. *Soit A un générateur markovien. Posons $A(x, x) = -m(x)$, et, supposons que pour tout x , $m(x) \neq 0$. Posons $P(x, y) = A(x, y)/m(x)$, pour $x \neq y$, ainsi que $P(x, x) = 0$.*

Alors P est une matrice markovienne.

Démonstration. — C'est immédiat à partir de la définition des générateurs markoviens. ■

Théorème 10.31. *Soit (X_t) un processus de Markov sur un ensemble E fini, de générateur A , avec $A(x, x) \neq x$, $\forall x$. Définissons les temps de saut successifs de X par $T_0 = 0$, $T_n = \inf\{t > T_{n-1}, X_t \neq X_{T_{n-1}}\}$. Alors, $Y_n = X_{T_n}$ est une chaîne de Markov de matrice P .*

Si $S_{n+1} = T_{n+1} - T_n$, alors la loi de (S_1, \dots, S_n) sachant que $(Y_0, \dots, Y_{n-1}) = (x_0, \dots, x_{n-1})$ est la loi d'une somme de variables aléatoires indépendantes, exponentielles de paramètres $(m(x_0), \dots, m(x_{n-1}))$.

En particulier, T_1 suit une loi exponentielle de paramètre $-A(x, x)$, et $A(x, y)/(-A(x, x))$ est la probabilité que $X_{T_1} = y$ sachant que $X_0 = x$.

Nous commençons par un lemme qui nous sera fort utile par la suite :

Lemme 10.32. *Soit T_2 le second temps de saut de X_t . Alors, pour tout $x \in E$, $\lim_{t \rightarrow 0} \frac{1}{t} \mathbf{P}_x(T_2 < t) = 0$.*

Démonstration. — Nous écrivons, en nous rappelant que $\theta_{T_1}(T_1) = T_2 - T_1$,

$$\begin{aligned} \mathbf{P}_x(T_2 \leq t) &= \mathbf{P}_x(T_1 \leq t, \theta_{T_1}(T_2 - T_1) \leq t - T_1) \\ &= \sum_{y \neq x} \mathbf{P}_x(T_1 \leq t, X_{T_1} = y, \theta_{T_1}(T_1) \leq t - T_1). \end{aligned}$$

Nous savons que T_1 suit sous \mathbf{P}_y une loi exponentielle de paramètre $\lambda(y)$ et donc que $\mathbf{P}_y(T_1 < s) = 1 - e^{-\lambda(y)s}$.

Appliquons la propriété de Markov forte au temps T_1 :

$$\mathbf{P}_x(T_2 \leq t) = \sum_y \mathbf{E}_x(\mathbf{1}_{\{T_1 < t\}} \mathbf{1}_{\{X_{T_1} = y\}} (1 - e^{-\lambda(y)(t-T_1)}).$$

Posons $\|\lambda\| = \sup_y(\lambda(y))$, et écrivons, sur $\{T_1 \leq t\}$,

$$1 - e^{-\lambda(y)(t-T_1)} \leq \lambda(y)(t - T_1) \leq \|\lambda\|(t - T_1);$$

il vient

$$\mathbf{P}_x(T_2 \leq t) \leq \|\lambda\| \mathbf{E}_x[\mathbf{1}_{\{T_1 \leq t\}}(t - T_1)].$$

Si T_1 suit sous \mathbf{P}_x une loi exponentielle de paramètre $\lambda(x)$, nous obtenons

$$\mathbf{E}_x[\mathbf{1}_{\{T_1 \leq t\}}(t - T_1)] = \int_0^t (t - s)\lambda(x)e^{-\lambda(x)s} ds = \frac{\lambda(x)t - 1 - e^{-\lambda(x)t}}{\lambda(x)}.$$

Cette quantité est équivalente à $\lambda(x)t^2/2$ lorsque $t \rightarrow 0$. ■

Démonstration. — (Du théorème 10.31.) Nous commençons par démontrer que, sous \mathbf{P}_x , la loi de T_1 est une loi exponentielle de paramètre $\lambda(x)$. Nous savons déjà que T_1 suit une loi exponentielle. Par hypothèse, $Q_t(x, y) = \mathbf{P}_x(X_t = y)$, et $\lambda(x) = \lim_{t \rightarrow 0} (Q_t(x, x) - 1)/t$. Appelons T_2 le deuxième temps de saut de X_t . Alors,

$$\begin{aligned} Q_t(x, x) &= \mathbf{E}_x(\mathbf{1}_{\{X_t = x\}}) \\ &= \mathbf{E}_x(\mathbf{1}_{\{X_t = x\}} \mathbf{1}_{\{T_1 > t\}}) + \mathbf{E}_x(\mathbf{1}_{\{X_t = x\}} \mathbf{1}_{\{T_2 \leq t\}}), \end{aligned}$$

la dernière identité est que, sous \mathbf{P}_x , on ne peut avoir $X_t = x$ que si $t \notin [T_1, T_2[$. D'après le lemme précédent, le second terme est un $o(t)$. Quant au premier, nous avons

$$\frac{\mathbf{E}_x(\mathbf{1}_{\{X_t=x\}}\mathbf{1}_{\{T_1>t\}}) - 1}{t} = \frac{\mathbf{P}_x(T_1 > t) - 1}{t} = \frac{e^{-\lambda(x)t} - 1}{t} \rightarrow -\lambda(x) \quad (t \rightarrow 0).$$

Nous voyons donc ainsi que la dérivée de $Q_t(x, x)$ en 0 est $-\lambda(x)$.

Considérons maintenant le cas où $x \neq y$. Nous avons

$$Q_t(x, y) = \mathbf{P}_x(X_t = y) = \mathbf{P}_x(X_{T_1} = y; T_1 \leq t < T_2) + \mathbf{P}_x(X_t = y; T_2 \leq t),$$

cette identité venant de ce que, si $x \neq y$ et $X_t = y$, alors, sous \mathbf{P}_x , $T_1 \leq t$, et que si $t < T_2$, alors $X_t = X_{T_1}$.

Comme plus haut, le second terme est un $o(t)$, d'après le lemme, et quant au premier, nous l'écrivons

$$\mathbf{P}_x(X_{T_1} = y; T_1 \leq t) - \mathbf{P}_x(X_{T_1} = y; T_2 \leq t).$$

le second terme est à nouveau un $o(t)$, et il nous reste à étudier

$$\lim_{t \rightarrow 0} \frac{\mathbf{P}_x(X_{T_1} = y; T_1 \leq t)}{t}.$$

Écrivons

$$\{X_{T_1} = y; t < T_1\} = \{t < T_1\} \cap \theta_t\{X_{T_1} = y\}.$$

Il vient, en appliquant la propriété de Markov au temps t ,

$$\mathbf{P}_x(X_{T_1} = y; t < T_1) = \mathbf{E}_x[\mathbf{1}_{\{t < T_1\}}\mathbf{P}_x(X_{T_1} = y)] = \mathbf{P}_x(X_{T_1} = y) \exp(-\lambda(x)t).$$

Appelons $B(x, y)$ la quantité $\mathbf{P}_x(X_{T_1} = y)$.

Nous obtenons

$$\begin{aligned} \mathbf{P}_x(X_{T_1} = y; T_1 \leq t) &= \mathbf{P}_x(X_{T_1} = y) - \mathbf{P}_x(X_{T_1} = y; t < T_1) \\ &= B(x, y)[1 - \exp(-\lambda(x)t)], \end{aligned}$$

d'où nous déduisons que

$$\lim_{t \rightarrow 0} \frac{\mathbf{P}_x(X_{T_1} = y; T_1 \leq t)}{t} = B(x, y)\lambda(x).$$

Nous voyons donc que $B(x, y) = -A(x, y)/\lambda(x)$.

Maintenant, la propriété de Markov appliquée aux temps T_n montre que $(Y_n) = (X_{T_n})$ est une chaîne de Markov, puisque $Y_{n+1} = \theta_{T_n}(Y_1)$, et nous venons de calculer sa matrice de transition.

Ensuite, puisque $T_{n+1} - T_n = \theta_{T_n}(T_1)$, nous voyons que, pour toute fonction F bornée,

$$\mathbf{E}[F(T_{n+1} - T_n)/\mathcal{F}_{T_n}] = \mathbf{E}_{X_{T_n}}(F(T_1)),$$

ce qui montre que $T_{n+1} - T_n$ est conditionnellement indépendante de \mathcal{F}_{T_n} sachant X_{T_n} , et que sa loi sachant que $X_{T_n} = y$ est une loi exponentielle de paramètre $\lambda(y)$. On en déduit aisément par récurrence la loi conditionnelle de $(T_1, T_2 - T_1, \dots, T_{n+1} - T_n)$ sachant (Y_0, Y_1, \dots, Y_n) . ■

Remarque. — La condition $A(x, x) \neq 0$ n'est là que pour nous simplifier la tâche. Si $A(x, x) = 0$, alors x est un point trappe, et le processus arrivé en x n'en repart plus. Les temps de saut successifs T_n ne sont plus alors nécessairement finis. Cela dépend de la probabilité pour la chaîne d'arriver en x , et donc du statut du point de départ (récurrent ou transitoire), et de x lui même.

Réciproquement, nous pouvons construire un processus de Markov de générateur A en considérant d'abord une chaîne de Markov de matrice P , en se donnant indépendamment pour chaque $x \in E$ une suite $(S_n(x))$ de temps exponentiels indépendants de paramètre $\lambda(x)$, puis en posant

$$T_n = S_1(Y_0) + S_2(Y_1) + \dots + S_n(Y_{n-1}).$$

Alors, le processus qui saute aux temps T_n en suivant la suite Y_n est un processus de Markov de matrice A .

Plus précisément,

$$X_t = \sum_n Y_n \mathbf{1}_{[T_n, T_{n+1}[}(t)$$

est un processus de Markov, à trajectoires continues à droites et avec limites à gauche, qui est de générateur A , où A est définie à partir de P et des $\lambda(x)$ comme dans l'énoncé.

Ce n'est pas si facile de vérifier la propriété de Markov directement sur cette définition. Le problème vient de la non-indépendance des sauts et des positions de la chaîne (Y_n) . Ce qui est clair d'après ce qui précède, c'est que s'il existe un processus de Markov de générateur A , il a la même loi que celui que nous venons de construire, et donc que celui que nous venons de construire est bien un processus de Markov de générateur A . Encore faut-il qu'il en existe au

moins un. C'est ce que nous allons faire ci-dessous, en partant d'une remarque simple.

Si les valeurs de $\lambda(x)$ sont indépendantes du point x (soit λ leur valeur commune), alors la suite (T_n) est indépendante de la suite (Y_n) , et dans ce cas, si $N_t = \sum_n n \mathbf{1}_{[T_n, T_{n+1}[}(t)$ est un processus de Poisson de paramètre λ . Alors dans ce cas, $X_t = Y_{N_t}$.

Pour construire notre processus de Markov, nous allons plutôt utiliser cette construction :

Proposition 10.33. *Soit (Y_n) une chaîne de Markov homogène de matrice P . Soit N_t un processus de Poisson indépendant de paramètre λ . Alors, le processus $X_t = Y_{N_t}$ est un processus de Markov de générateur*

$$A = \lambda(P - Id).$$

Démonstration. — Ce processus est par construction continu à droite et pourvu de limites à gauche.

Vérifions d'abord que ce processus est Markovien. Cela vient des propriétés de Markov combinées de la chaîne Y et du processus N_t . Pour vérifier le caractère markovien, il suffit de montrer que, pour tous $(t_1 < \dots < t_n < t)$, la loi de X_t sachant $(X_{t_1}, \dots, X_{t_n})$ est la loi de X_t sachant X_{t_n} , et il faut en plus calculer cette loi.

Posons $Q_t = \exp(t\lambda(P - Id)) = \exp(-\lambda t) \exp(\lambda t P)$. Considérons deux fonctions $f(x)$ et $F(x_1, \dots, x_n)$ à valeurs dans \mathbb{R} , définies respectivement sur E et E^n . Nous écrivons

$$\begin{aligned}
& \mathbf{E}[f(X_t)F(X_{t_1}, \dots, X_{t_n})] \\
= & \sum_{k_1 \leq \dots \leq k_n \leq k} \mathbf{E}[f(Y_k)F(Y_{k_1}, \dots, Y_{k_n})\mathbf{1}_{\{N_t=k, N_{t_1}=k_1, \dots, N_{t_n}=k_n\}}] \\
= & \sum_{k_1 \leq \dots \leq k_n \leq k} \mathbf{E}[f(Y_k)F(Y_{k_1}, \dots, Y_{k_n})]\mathbf{P}(N_t = k, N_{t_1} = k_1, \dots, N_{t_n} = k_n) \\
& \text{(par indépendance de } N \text{ et } Y) \\
= & \sum_{k_1 \leq \dots \leq k_n \leq k} \mathbf{E}[P^{k-k_n}(f)(Y_{k_n})F(Y_{k_1}, \dots, Y_{k_n})]\mathbf{P}(N_{t-t_n} = k - k_n) \times \\
& \mathbf{P}(N_{t_1} = k_1, \dots, N_{t_n} = k_n) \text{ (propriété de Markov de } Y \text{ et } N) \\
= & \sum_{k_1 \leq \dots \leq k_n} \sum_{k'} \mathbf{E}[P^{k'}(f)(Y_{k_n})F(Y_{k_1}, \dots, Y_{k_n})]\mathbf{P}(N_{t-t_n} = k') \times \\
& \mathbf{P}(N_{t_1} = k_1, \dots, N_{t_n} = k_n) \\
= & e^{-\lambda(t-t_n)} \sum_{k_1 \leq \dots \leq k_n} \sum_{k'} \frac{[\lambda(t-t_n)]^{k'}}{(k')!} \mathbf{E}[P^{k'}(f)(Y_{k_n})F(Y_{k_1}, \dots, Y_{k_n})] \times \\
& \mathbf{P}(N_{t_1} = k_1, \dots, N_{t_n} = k_n) \text{ (loi de } N_{t-t_n}) \\
= & e^{-\lambda(t-t_n)} \sum_{k_1 \leq \dots \leq k_n} \mathbf{E}\left(\sum_{k'} \frac{(\lambda(t-t_n)P)^{k'}}{k'!} F(Y_{k_1}, \dots, Y_{k_n}) \times \right. \\
& \left. \mathbf{P}(N_{t_1} = k_1, \dots, N_{t_n} = k_n)\right) \\
= & \sum_{k_1 \leq \dots \leq k_n} \mathbf{E}[Q_{t-t_n}(f)(Y_{k_n})F(Y_{k_1}, \dots, Y_{k_n})]\mathbf{P}(N_{t_1} = k_1, \dots, N_{t_n} = k_n) \\
& \text{(définition de } Q_t) \\
= & \sum_{k_1 \leq \dots \leq k_n} \mathbf{E}[Q_{t-t_n}(f)(Y_{k_n})F(Y_{k_1}, \dots, Y_{k_n})\mathbf{1}_{\{N_{t_1}=k_1, \dots, N_{t_n}=k_n\}}] \\
& \text{(indépendance de } Y \text{ et } N) \\
= & \mathbf{E}[Q_{t-t_n}(f)(X_{t_n})F(X_{t_1}, \dots, X_{t_n})].
\end{aligned}$$

Nous obtenons finalement à la fois la propriété de Markov et le noyau de X_t , qui est $Q_t = \exp(\lambda(P - Id))$. C'est bien ce que nous voulions démontrer. ■

Remarquons que dans ce qui précède, nous ne demandons pas à la matrice P de vérifier $P(x, x) = 0$. Les temps de sauts du processus ne sont pas nécessairement les temps de saut de N_t : il se peut que N_t saute sans que le processus bouge ; il s'agit alors de sauts fictifs.

Finalement, nous pouvons donner une construction simple d'un processus

de Markov de générateur donné A .

Corollaire 10.34. *Soit A un générateur markovien sur l'ensemble fini E , et soit λ un majorant de $-\sup_x A(x, x)(x)$.*

On pose

$$P(x, x) = 1 + \frac{A(x, x)}{\lambda}; \quad P(x, y) = \frac{A(x, y)}{\lambda} \text{ pour } x \neq y.$$

Alors, si Y_n est une chaîne de Markov de matrice P et si N_t est un processus de Poisson d'intensité λ , alors Y_{N_t} est un processus de Markov de générateur A .

Démonstration. — Il suffit de vérifier qu'avec les définitions données, on a bien $A = \lambda(P - Id)$: c'est immédiat. ■

10.6 Mesures Invariantes.

Comme dans le cas des chaînes de Markov, une mesure de probabilité μ sur E sera dite invariante lorsque, chaque fois que X_0 a la loi μ , X_t garde la loi μ pour tout temps $t > 0$. ceci se traduit par l'équation $\mu Q_t = \mu$.

En dérivant en $t = 0$, on obtient $\mu A = 0$, où A est le générateur. Réciproquement, l'équation $\mu A = 0$ donne $\mu Q_t = \mu$, puisque $\partial_t Q_t = A Q_t$.

Nous pouvons récrire cette équation en faisant apparaître la matrice $P(x, y)$ de la chaîne sous-jacente, ainsi que les paramètres $\lambda(x)$ des temps de saut. On obtient

$$\sum_{y \neq x} \mu(y) \lambda(y) P(y, x) = \mu(x) \lambda(x).$$

En d'autres termes, la mesure $\nu(x) = \mu(x) \lambda(x)$ est une mesure invariante pour la chaîne de matrice P .

En éliminant le cas trivial où l'un des $\lambda(x)$ est nul, (auquel cas x est une trappe et la mesure de Dirac en x est une mesure invariante), on voit que la recherche des mesures invariantes pour X_t revient à la recherche des mesures invariantes pour la chaîne sous-jacente. On a alors la proposition suivante, qui découle immédiatement des considérations qui précèdent

Proposition 10.35. *Si tous les $\lambda(x)$ sont non nuls et si P désigne la matrice de la chaîne sous-jacente ($P(x, y) = A(x, y)/\lambda(x)$, $x \neq y$, $P(x, x) = 0$), alors*

la mesure invariante est unique si et seulement si il n'y a qu'une seule classe de récurrence pour P . Dans ce cas, si ν est l'unique mesure invariante pour P , alors l'unique mesure invariante pour X_t est proportionnelle à $\nu(x)/\lambda(x) = \nu(x)\mathbf{E}_x(T_x)$.

En d'autres termes, il faut multiplier la mesure invariante de la chaîne sous-jacente par le temps de séjour moyen au point x .

La convergence vers la mesure invariante est plus simple que pour les chaînes discrètes. En effet, si la chaîne est récurrente irréductible, on a toujours, pour toute mesure initiale μ_0 ,

$$\lim_{t \rightarrow \infty} \mu_0 Q_t = \mu,$$

où μ est la mesure invariante (sans conditions d'apériodicité). Cela se voit assez facilement sur la représentation $Q_t = \exp(tA)$ et $A = \lambda(P - Id)$, où la matrice Markovienne P peut être choisie apériodique.

De plus, le théorème ergodique s'énonce alors

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(X_s) ds = \int_E f d\mu (p.s.)$$

Ce théorème se démontre plus ou moins selon les mêmes lignes que le théorème correspondant dans le cas du temps discret.

10.7 Mesures de Poisson.

Si on observe une trajectoire d'un processus de Poisson sur \mathbb{N} , c'est la fonction de répartition d'une mesure positive (aléatoire) sur \mathbb{N} : cette mesure est la mesure $\nu(\omega, dx) = \sum_{n \geq 1} \delta_{T_n}$, où les temps T_n sont les temps de saut successifs du processus. C'est une mesure ponctuelle (elle n'est formée que de sommes de masses de Dirac affectées de coefficients 1).

Cette mesure a les propriétés suivantes

Proposition 10.36. *Pour tout borélien A de \mathbb{R}_+ de mesure de Lebesgue finie, notons*

$$N(A) = \int_A \nu(\omega, dx) = \sum_{n \geq 1} \mathbf{1}_A(T_n).$$

Alors, pour tout A , $N(A)$ suit une loi de Poisson de paramètre $\lambda|A|$, où $|A|$ désigne la mesure de Lebesgue de A .

Plus généralement, si A_1, \dots, A_n sont des boréliens disjoints de mesure de Lebesgue finie, alors $N(A_1), \dots, N(A_n)$ sont des variables de Poisson indépendantes de paramètres respectifs $\lambda|A_i|$.

Démonstration. — Lorsque $A = [0, t]$, $N(A) = N_t$. De même, lorsque $A =]s, t]$, $N(A) = N_t - N_s$. Ainsi, lorsque $A_i =]t_i, t_{i+1}]$, on retrouve ici la propriété d'accroissements indépendants du processus de Poisson.

Nous allons montrer directement le second point, en commençant par le cas où tous les boréliens A_i sont inclus dans $[0, t]$, pour un réel t fixé. Remarquons que, quitte à rajouter à la liste des A_i de complémentaire de $\cup_i A_i$ dans $[0, t]$, nous pouvons toujours supposer que les A_i forment une partition de $[0, t]$. Nous allons alors faire le même raisonnement que ce qui nous a amenés à la propriété d'accroissements indépendants de N_\bullet .

Tout d'abord, $\sum_{i=1}^n N(A_i) = N([0, t]) = N_t$, et cette variable suit une loi de Poisson de paramètre λt .

Ensuite, conditionnons la loi de $(N(A_1), \dots, N(A_n))$ par $\{N_t = k\}$. Conditionnellement à $\{N_t = k\}$, les temps (T_1, \dots, T_k) ont la même loi que le réarrangement croissant de k variables uniformes indépendantes dans $[0, t]$. La probabilité qu'il y en ait k_1 dans A_1 , k_2 dans A_2 , ..., k_n dans A_n suit une loi multinômiale de paramètres $p_i = \frac{|A_i|}{t}$. (Ici, $k = k_1 + \dots + k_n$.)

On obtient ainsi

$$\mathbf{P}(N(A_1) = k_1, \dots, N(A_n) = k_n / N_t = k) = k! \frac{p_1^{k_1}}{k_1!} \dots \frac{p_n^{k_n}}{k_n!},$$

et on conclut exactement comme dans le cas déjà traité où $A_i =]t_i, t_{i+1}]$.

Pour passer au cas général où les A_i ne sont pas tous dans un intervalle borné, on pose $A_i^K = A_i \cap [0, K]$. Les variables $N(A_i^K)$ convergent toutes en croissant vers $N(A_i)$ (convergence monotone pour les mesures $\nu(\omega, dx)$, si l'on veut). Ces variables $N(A_i^K)$ ont comme espérance $\lambda|A_i^K|$ (l'espérance d'une variable de Poisson de paramètre μ est égale à μ). Par convergence monotone, on voit donc que les variables $N(A_i)$ sont intégrables d'espérance $\lambda|A_i|$. Elles sont de plus indépendantes comme limites presque sûres de variables indépendantes.

Il reste à voir qu'elles suivent une loi de Poisson. Ce sont des limites presque sûres de variables de Poisson. La convergence presque sûre entraîne la convergence en loi. Mais si $\mu_n \rightarrow \mu$, la loi de Poisson de paramètre μ_n converge

étroitement vers la loi de Poisson de paramètre μ , comme on le voit immédiatement sur la transformée de Fourier

$$\Phi_{\mu_n}(u) = \exp(\mu(e^{iu} - 1)).$$

■

Cette mesure aléatoire sur \mathbb{R} formée de sommes de masses de Dirac est un moyen commode de représenter un "nuage de points" aléatoires. À un ensemble aléatoire de points (X_i) , on fait correspondre la mesure (aléatoire) $\sum_i \delta_{X_i}$. Ces mesures ont la propriété particulière de prendre leurs valeurs dans $\mathbb{N} \cup \infty$: on les appelle des mesures ponctuelles. La mesure d'un ensemble A est le nombre de points X_i qui sont dans A .

Dire que cette mesure est presque sûrement de Radon, par exemple, signifie que presque sûrement il n'y a qu'un nombre fini de points dans chaque compact.

Pour fixer les idées :

Définition 10.37. Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace de probabilité. Une mesure aléatoire sur (E, \mathcal{E}) est un noyau de transition $M(\omega, dx)$ mais qui n'est pas nécessairement une probabilité. C'est à dire

1. $\forall B \in \mathcal{E}, \omega \mapsto M(\omega, B)$ est une variable aléatoire à valeurs dans $[0, \infty]$.
2. $\forall \omega \in \Omega, B \mapsto M(\omega, B)$ est une mesure positive σ -finie sur E .

On peut se poser la question de savoir si la construction que nous avons faite sur \mathbb{R} à partir du processus de Poisson (ou de variables exponentielles) peut se faire pour d'autres mesures. C'est ce que nous allons faire ci dessous :

Définition 10.38. Soit (E, \mathcal{E}, μ) un espace mesuré, muni d'une mesure μ σ -finie. On dit qu'une mesure aléatoire $M(\omega, dx)$ est une mesure de Poisson d'intensité μ sur A si, pour tout n -uplet (A_1, \dots, A_n) d'éléments disjoints de \mathcal{E} de mesures $\mu(A_i)$ finies, les variables $M(A_1), \dots, M(A_n)$ sont indépendantes, et suivent des lois de Poisson de paramètres respectifs $\mu(A_1), \dots, \mu(A_n)$.

Ainsi, le processus de Poisson construit sur \mathbb{R}_+ est une mesure de Poisson dont l'intensité est λdx , où dx est la mesure de Lebesgue. Nous allons voir qu'on peut toujours construire une telle mesure de Poisson. Commençons par le cas où μ est une mesure finie.

Proposition 10.39. Soit μ_1 une probabilité sur E , et X_n une suite de variables indépendantes à valeurs dans E , de loi μ_1 . Soit N une variable aléatoire

indépendante de toute la suite (X_n) , suivant une loi de Poisson de paramètre λ . Alors

$$\mu(\omega, dx) = \sum_{i=1}^{N(\omega)} \delta_{X_i}$$

est une mesure de Poisson d'intensité $\lambda\mu_1$.

Si μ est finie, en posant $\mu_1 = \mu/\mu(E)$ et $\lambda = \mu(E)$, on voit ainsi comment construire une mesure de Poisson : on prend un nombre Poissonien de variables aléatoires indépendantes de loi μ_1 .

Démonstration. — Soit (A_1, \dots, A_n) une partition de E . Alors, $\sum \mu(A_i) = N$ suit une loi de Poisson de paramètre λ . Le raisonnement déjà fait plus haut pour le cas de \mathbb{R}_+ montre que la loi conditionnelle de $(N(A_1), \dots, N(A_n))$ sachant que $N = k$ est une loi multinômiale de paramètres $p_i = \mu_1(A_i)$.

On a donc, pour $k_1 + \dots + k_n = k$,

$$\mathbf{P}(N(A_1) = k_1, \dots, N(A_n) = k_n) = k! \prod_1^n \frac{p_i^{k_i}}{k_i!},$$

et donc

$$\mathbf{P}(N(A_1) = k_1, \dots, N(A_n) = k_n) = e^{-\mu} \prod_1^n \frac{(\lambda p_i)^{k_i}}{k_i!},$$

ce qui est bien la formule attendue. ■

Dans le cas général où μ est seulement σ -finie, E est par définition réunion dénombrable d'ensembles A_n disjoints de masse finie. Nous construisons indépendamment la mesure de Poisson sur chaque A_i par le procédé décrit plus haut, et il suffit de recoller les morceaux, en utilisant le résultat suivant :

Proposition 10.40. *Soit M_n une suite de mesures de Poisson indépendantes d'intensité μ_n portée par des ensembles disjoints A_n . Alors la mesure $M(\omega, dx) = \sum_n M_n(\omega, dx)$ est une mesure de Poisson d'intensité $\mu = \sum_n \mu_n$.*

Démonstration. — Exercice! ■