Université Pierre et Marie Curie

Master 1 de Mathématiques

Processus de Sauts et Files d'Attente

Nicolas Fournier

d'après Philippe Bougerol

Chapitre 1

Marches aléatoires, processus de Bernoulli

Le but de ce chapitre est double. D'une part, faire quelques rappels et fixer les notations. D'autre part, introduire les marches aléatoires et le processus de Bernoulli, caricature à temps discret du très important processus de Poisson.

1.1 Rappels

On se donne un espace de probabilité $(\Omega, \mathcal{F}, \mathbf{P})$.

Soit $(E, \mathcal{E})$ un espace mesurable et X une v.a. à valeurs dans E. La loi de X est la mesure de probabilité μ_X sur $(E, \mathcal{E})$ définie par $\mu_X(A) = \mathbf{P}(X \in A)$, pour tout $A \in \mathcal{E}$.

Pour μ une probabilité sur $(E, \mathcal{E})$, on a $\mu_X = \mu$ si et seulement si $\mathbf{E}[\varphi(X)] = \int_E \varphi d\mu$ pour tout $\varphi \in \mathcal{C}$, où $\mathcal{C}$ est la classe des fonctions mesurables bornées de E dans $\mathbf{R}$. On peut aussi choisir pour $\mathcal{C}$ la classe des fonctions mesurables positives.

Définition 1.1.1 *Soit $T = \mathbf{N}$ ou $\mathbf{R}^+$ (ou éventuellement $\mathbf{Z}$ ou $\mathbf{R}$) et $(E, \mathcal{E})$ un espace mesurable.*

(i) Un processus à valeurs dans E est une famille $(X_t)_{t \in T}$ de v.a. à valeurs dans E.

(ii) On dit que deux processus $(X_t)_{t \in T}$ et $(Y_t)_{t \in T}$ à valeurs dans E ont la même loi si pour tout $k \geq 1$, pour tout $t_1 < t_2 < \ldots < t_k$ (avec $t_i \in T$), les v.a. $(X_{t_1}, \ldots, X_{t_k})$ et $(Y_{t_1}, \ldots, Y_{t_k})$ (à valeurs dans E^k, qui est muni de la tribu produit $\mathcal{E}^{\otimes k}$) ont la même loi.

L'ensemble T représente généralement le temps, qui peut être discret ou continu.

1.2 Marches aléatoires, processus de Bernoulli

Définition 1.2.1 *Soient μ une loi sur $\mathbf{R}^d$ et $X_1, X_2, \ldots$ des v.a. i.i.d. de loi μ. Le processus $(S_n)_{n\in\mathbf{N}}$ défini par $S_0 = 0$ et $S_n = X_1 + \ldots + X_n$ $(n \geq 1)$ est appelé marche aléatoire de "pas" de loi μ.*

Si $d = 1$ et si $\mu = (1-p)\delta_0 + p\delta_1$ pour un $p \in [0,1]$ (i.e. μ est la loi de Bernoulli de paramètre p), alors $(S_n)_{n\in\mathbf{N}}$ est appelé processus de Bernoulli. Bien sûr, pour n fixé, on a $S_n \sim$Bin(n,p) et donc $\mathbf{E}[S_n] = np$ et $\mathrm{Var}(S_n) = np(1-p)$. Par exemple, on joue indéfiniment au pile ou face avec une pièce qui tombe sur pile avec probabilité p : S_n représente le nombre de piles obtenus lors des n premiers lancers.

Si $d = 1$ et si $\mu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, alors le processus $(S_n)_{n\in\mathbf{N}}$ est appelé marche aléatoire simple symétrique.

Si $d = 2$ et si $\mu = \frac{1}{4}(\delta_{(-1,0)} + \delta_{(0,-1)} + \delta_{(1,0)} + \delta_{(0,1)})$, le processus $(S_n)_{n\in\mathbf{N}}$ est appelé marche aléatoire au plus proche voisin symétrique.

1.3 Propriété de Markov forte des marches aléatoires

La proposition suivante est évidente.

Proposition 1.3.1 *Soit $(S_n)_{n\in\mathbf{N}}$ une marche aléatoire de pas de loi μ. Pour $n_0 \in \mathbf{N}$ fixé, posons $S_n^{n_0} = S_{n_0+n} - S_{n_0}$. Alors $(S_n^{n_0})_{n\in\mathbf{N}}$ est une marche aléatoire de pas de loi μ et est de plus indépendante de $(S_0, \ldots, S_{n_0})$.*

Preuve: Il suffit d'observer que

(i) $S_0^{n_0} = 0$ et $S_{n_0+n} - S_{n_0} = X_{n_0+1} + \ldots + X_{n_0+n}$,

(ii) la suite $(X_{n_0+k})_{k\geq 1}$ est constituée de v.a. i.i.d. de loi commune μ,

(ii) la suite $(X_{n_0+k})_{k\geq 1}$ est indépendante du vecteur $(X_1, \ldots, X_{n_0})$ et donc aussi de $(S_0, S_1, ..., S_{n_0})$.

Les définitions qui suivent sont cruciales dans l'étude des processus et seront utilisées tout au long de ce cours.

Définition 1.3.2 *Soit $(S_n)_{n\in\mathbf{N}}$ un processus à temps discret. On introduit la filtration $(\mathcal{F}_n)_{n\in\mathbf{N}}$ (famille croissante de sous-tribus de $\mathcal{F}$) définie par*

$$\mathcal{F}_n = \sigma(S_0, \ldots, S_n).$$

(i) Une v.a. τ à valeurs dans $\mathbf{N} \cup \{\infty\}$ est un temps d'arrêt (pour le processus $(S_n)_{n\in\mathbf{N}}$) si pour tout $n \in \mathbf{N}$,

$$\{\tau \leq n\} \in \mathcal{F}_n.$$

(ii) On pose alors

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \ : \ A \cap \{\tau \leq n\} \in \mathcal{F}_n \ \forall\, n \in \mathbf{N}\}.$$

On vérifiera en exercice que $\mathcal{F}_\tau$ est une tribu (c'est une sous-tribu de $\mathcal{F}$). Quelques remarques : τ est un temps aléatoire, éventuellement infini.

Le point (i) peut se traduire ainsi : pour tout $n \in \mathbf{N}$, si l'on connaît les valeurs de $S_0, \dots, S_n$, alors on peut dire si $\tau \leq n$ ou non.

Pour (ii), un évènement A appartient à $\mathcal{F}_\tau$ si, lorsqu'on connaît les valeurs de $\tau, S_0, \dots, S_\tau$, on peut dire si A est réalisé ou non.

Exemple/Exercice : Soit $(S_n)_{n\in\mathbf{N}}$ un processus à valeurs dans E (muni d'une tribu $\mathcal{E}$) et $B \in \mathcal{E}$.

Alors $\tau = \inf\{n \geq 0 \ : \ S_n \in B\}$, avec la convention $\inf \emptyset = \infty$, est un temps d'arrêt. Les v.a. S_τ, $S_{\tau-1}$, $\sum_{k=0}^{\tau} \varphi(S_k)$ (pour une fonction mesurable $\varphi : E \mapsto \mathbf{R}$) sont toutes $\mathcal{F}_\tau$-mesurables. Par contre, $S_{\tau+1}$ n'est en général pas $\mathcal{F}_\tau$-mesurable.

C'est intuitivement clair : si on fixe $n \in \mathbf{N}$ et qu'on observe $S_0, ..., S_n$, alors on voit si le processus $(S_k)_{k\in\mathbf{N}}$ est passé dans B avant l'instant n, i.e. on voit si $\tau \leq n$ ou non. Si maintenant on observe $\tau, S_0, S_1, \dots, S_\tau$, i.e. on observe le processus jusqu'à ce qu'il arrive dans B, on peut déterminer les valeurs de S_τ, $S_{\tau-1}$, $\sum_{k=0}^{\tau} \varphi(S_k)$, mais pas la valeur de $S_{\tau+1}$.

Par contre, $\tau = \sup\{n \geq 0 \ : \ S_n \in B\}$ n'est en général pas (et presque jamais) un temps d'arrêt.

Pour n fixé, en observant $S_0, ..., S_n$, on ne peut généralement pas dire si $\tau \leq n$, puisqu'on ne sait pas si le processus repassera dans B ou non après n.

Théorème 1.3.3 Propriété de Markov forte pour les M.A. *Soit $(S_n)_{n\in\mathbf{N}}$ une marche aléatoire de pas de loi μ. Soit τ un temps d'arrêt (pour $(S_n)_{n\in\mathbf{N}}$) fini p.s. On pose $S_n^\tau = S_{\tau+n} - S_\tau$. Alors $(S_n^\tau)_{n\in\mathbf{N}}$ est une marche aléatoire de pas de loi μ indépendante de $\mathcal{F}_\tau$.*

Remarquons qu'il est crucial que τ soit un temps d'arrêt : si par exemple $(S_n)_{n\in\mathbf{N}}$ est un processus de Bernoulli de paramètre $1/2$ et si $\tau = \sup\{n \geq 0 \ : \ S_n = 3\}$ (qui n'est pas un temps d'arrêt), alors $S_1^\tau = S_{\tau+1} - S_\tau = 1$ p.s., donc $(S_n^\tau)_{n\in\mathbf{N}}$ n'est pas un processus de Bernoulli de paramètre $1/2$ (sinon, on aurait $\mathbf{P}(S_1^\tau = 1) = 1/2$).

Preuve: L'objectif est de montrer que pour tout $A \in \mathcal{F}_\tau$, pour tout $n \geq 0$, pour tous boréliens $B_1, \cdots, B_n$ de $\mathbf{R}^d$, on a $I = J$, où

$$I = \mathbf{P}(A, S_{\tau+1} - S_\tau \in B_1, \dots, S_{\tau+n} - S_{\tau+n-1} \in B_n)$$

et

$$J = \mathbf{P}(A)\mu(B_1)\dots\mu(B_n).$$

Mais, en se rappelant que $S_n = X_1 + \ldots + X_n$ où les X_i sont i.i.d. de loi μ, et en utilisant que τ est p.s. fini,

$$I = \sum_{k\geq 0} \mathbf{P}(A, \tau = k, S_{\tau+1} - S_\tau \in B_1, \dots, S_{\tau+n} - S_{\tau+n-1} \in B_n)$$

$$
\begin{aligned}
&= \sum_{k\geq 0} \mathbf{P}(A, \tau = k, S_{k+1} - S_k \in B_1, \ldots, S_{k+n} - S_{k+n-1} \in B_n) \\
&= \sum_{k\geq 0} \mathbf{P}(A, \tau = k, X_{k+1} \in B_1, \ldots, X_{k+n} \in B_n).
\end{aligned}
$$

Comme $A \in \mathcal{F}_\tau$, $A \cap \{\tau = k\} \in \mathcal{F}_k$. De plus, $(X_{k+1}, \ldots, X_{k+n})$ est indépendant de $\mathcal{F}_k$. Du coup,

$$
\begin{aligned}
I &= \sum_{k\geq 0} \mathbf{P}(A, \tau = k)\mathbf{P}(X_{k+1} \in B_1, \ldots, X_{k+n} \in B_n) \\
&= \sum_{k\geq 0} \mathbf{P}(A, \tau = k)\mu(B_1)\ldots\mu(B_n).
\end{aligned}
$$

Il n'y a plus qu'à remarquer que $\sum_{k\geq 0} \mathbf{P}(A, \tau = k) = \mathbf{P}(A)$ puisque τ est p.s. fini.

1.4 Loi géométrique

Considérons un processus de Bernoulli $(S_n)_{n\geq 0}$ de paramètre $p \in]0, 1]$ et

$$T = \inf\{n > 0; S_n = 1\}.$$

La loi de T s'appelle la loi géométrique sur $\mathbf{N}^*$ de paramètre p. On a donc

Définition 1.4.1 *Soit $p \in [0, 1]$. La loi géométrique de paramètre p est la probabilité μ sur $\mathbf{N}^*$ définie par $\mu(\{n\}) = (1-p)^{n-1}p$.*

Si X est une v.a. de loi géométrique, on a

$$\mathbf{P}(X > n) = (1-p)^n, \text{ pour tout } n \in \mathbf{N}.$$

Une expérience aléatoire se répète dans les mêmes conditions : le nombre d'essais nécessaires avant le premier succès a une loi géométrique. Vous jouez au loto: vous attendrez un temps géométrique avant de gagner, etc. Cette loi modélise tous les phénomènes d'attente à temps discret "sans mémoire", au sens où le fait d'attendre beaucoup ne change pas la loi du temps qu'il reste à attendre. En effet, on a

Proposition 1.4.2 *Une v.a. X à valeurs dans $\mathbf{N}^*$ suit une loi géométrique si et seulement si, pour tout $n, m > 0$,*

$$\mathbf{P}(X > m + n | X > n) = \mathbf{P}(X > m).$$

Preuve. Si X suit une loi géométrique (de paramètre p), on a alors trivialement que pour tout $n, m > 0$,

$$\mathbf{P}(X > m + n | X > n) = \frac{\mathbf{P}(X > m + n)}{\mathbf{P}(X > n)} = \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m = \mathbf{P}(X > m).$$

Supposons maintenant que pour tout $n, m > 0$, $\mathbf{P}(X > m + n | X > n) = \mathbf{P}(X > m)$ et posons $\lambda = \mathbf{P}(X > 1)$. Alors pour tout $n > 0$,

$$\mathbf{P}(X > n + 1)/\mathbf{P}(X > n) = \mathbf{P}(X > n + 1 | X > n) = \mathbf{P}(X > 1) = \lambda,$$

d'où

$$\mathbf{P}(X > n + 1) = \lambda \mathbf{P}(X > n).$$

On en déduit que $\mathbf{P}(X > n) = \lambda^n$ pour tout $n > 0$. C'est aussi vrai quand $n = 0$ car $\mathbf{P}(X > 0) = 1$ par hypothèse. Ainsi, pour tout $n > 0$,

$$\mathbf{P}(X = n) = \mathbf{P}(X > n - 1) - \mathbf{P}(X > n) = \lambda^{n-1} - \lambda^n = \lambda^{n-1}(1 - \lambda)$$

et X suit la loi géométrique de paramètre $p = 1 - \lambda$.

Le résultat suivant est intuitif. Il est facile à montrer de façon élémentaire. Mais il est plus instructif de le montrer en utilisant la propriété de Markov forte.

Proposition 1.4.3 *Soit $(S_n)_{n \in \mathbf{N}}$ un processus de Bernoulli de paramètre p et $T_k = \inf\{n \geq 0; S_n = k\}$, pour tout $k \geq 0$ (on a $T_0 = 0$ p.s.). Alors les variables aléatoires $\{T_{k+1} - T_k, k \geq 0\}$ sont i.i.d. de loi géométrique de paramètre p.*

Remarque 1.4.4 *On peut écrire $S_n = \sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq n\}}$ (exercice facile : pour tout $\ell \in \mathbf{N}$ fixé, $S_n = \ell$ ssi $T_\ell \leq n < T_{\ell+1}$ ssi $\sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq n\}} = \ell$). Par la proposition, on a donc $S_n = \sum_{k=1}^{\infty} \mathbf{1}_{\{Y_1 + \ldots + Y_k \leq n\}}$, pour une famille $(Y_k)_{k \geq 1}$ de v.a. i.i.d. de loi géométrique de paramètre p.*

Preuve: C'est un excellent exercice, dont nous ne donnons que les étapes.

(a) Montrer que T_1 suit la loi géométrique de paramètre p.

(b) Montrer que pour chaque $k \geq 1$, T_k est un temps d'arrêt (du processus $(S_n)_{n \in \mathbf{N}}$).

(c) Montrer que que pour chaque $k \geq 1$, $T_1, \ldots, T_k$ est $\mathcal{F}_{T_k}$-mesurable, et donc que $T_1, T_2 - T_1, \ldots, T_k - T_{k-1}$ est $\mathcal{F}_{T_k}$-mesurable.

(d) Fixons $k \geq 1$ et introduisons $\tilde{S}_n = S_n^{T_k} = S_{T_k + n} - S_{T_k}$. Par Markov forte, on sait que $(\tilde{S}_n)_{n \in \mathbf{N}}$ est un processus de Bernoulli de paramètre p indépendant de $\mathcal{F}_{T_k}$. Donc

$$\tilde{T}_1 = \inf\{n \geq 0 \ : \ \tilde{S}_n = 1\}$$

est indépendant de $\mathcal{F}_{T_k}$ et a la même loi que T_1 (la loi géométrique de paramètre p). Mais, comme $S_{T_k} = k$, on a

$$\tilde{T}_1 = \inf\{n \geq 0 \ : \ S_{T_k + n} - S_{T_k} = 1\} = \inf\{n \geq 0 \ : \ S_{T_k + n} = k + 1\} = T_{k+1} - T_k.$$

Ainsi, $T_{k+1} - T_k$ est indépendant de $\mathcal{F}_{T_k}$ et suit la loi géométrique de paramètre p.

(e) En exploitant les points (c) et (d), on voit que pour tout $k \geq 1$, les v.a. $T_1, T_2 - T_1, \ldots, T_{k+1} - T_k$ sont i.i.d. de loi géométrique de paramètre p. C'est ce que nous voulions démontrer.

1.5 Loi exponentielle

Les lois géométriques sont à valeurs entières. Nous allons considérer maintenant des lois ayant des propriétés analogues, mais à valeurs dans tout $\mathbf{R}^+$: les lois exponentielles. Ce sont, avec les lois de Gauss, les plus importantes du calcul des probabilités à temps continu.

Définition 1.5.1 *Soit $\lambda > 0$. On dit qu'une v.a. réelle T suit la loi exponentielle de paramètre λ (en bref, que $T \sim \mathcal{E}(\lambda)$) si $\mu_T(dt) = \lambda e^{-\lambda t}\mathbf{1}_{t>0}dt$.*

Si $T \sim \mathcal{E}(\lambda)$, alors $\mathbf{E}[T] = 1/\lambda$, $\mathrm{Var}(T) = 1/\lambda^2$. Aussi,

$$\mathbf{P}(T \geq t) = \mathbf{P}(T > t) = e^{-\lambda t} \text{ pour tout } t \geq 0.$$

Théorème 1.5.2 *Soit T une v.a. à valeurs dans $\mathbf{R}_*^+$. On a l'équivalence :*

(i) il existe $\lambda > 0$ tel que $T \sim \mathcal{E}(\lambda)$,

(ii) pour tout $s, t \geq 0$, $\mathbf{P}(T \geq t + s | T \geq t) = \mathbf{P}(T \geq s)$.

C'est très important. (ii) est une propriété qualitative (absence de mémoire). Donc toute v.a. sans mémoire suit nécessairement une loi exponentielle.

Le temps d'attente T, avant la prise d'un premier poisson, d'un pêcheur totalement inexpérimenté est une variable aléatoire sans mémoire (le temps d'attente résiduel ne dépend pas du temps d'attente écoulé). Donc *par nature*, T suit une loi exponentielle.

La durée de vie D d'un objet qui ne s'use pas est une variable aléatoire sans mémoire. Donc *par nature*, T suit une loi exponentielle. Etc.

Preuve: (i) implique (ii) est facile :

$$\mathbf{P}(T \geq t + s | T \geq t) = \frac{\mathbf{P}(T \geq t + s)}{\mathbf{P}(T \geq t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = \mathbf{P}(T \geq s).$$

Montrons maintenant que (ii) implique (i). Pour cela, introduisons $G(t) = \mathbf{P}(T \geq t)$. C'est une fonction décroissante sur $\mathbf{R}^+$, on a $G(0) = 1$ et $G(\infty) = 0$ par hypothèse. De plus (ii) donne directement que $G(t + s) = G(t)G(s)$ pour tout $s, t \geq 0$. On en déduit facilement que pour tout $p, q \in \mathbf{N}$, on a $G(p/q)^q = G(p)$ et que $G(p) = G(1)^p$, d'où $G(p/q) = G(1)^{p/q}$. En utilisant que $\mathbf{Q}$ est dense dans $\mathbf{R}^+$, que G est décroissante et que $t \mapsto G(1)^t$ est continue, on en déduit facilement que $G(t) = G(1)^t$ pour tout $t \in \mathbf{R}^+$.

De plus, $G(1) > 0$. Sinon, on aurait $G(t) = 0$ pour tout $t > 0$ et donc $T = 0$ p.s. (or on a supposé T à valeurs dans $\mathbf{R}_*^+$).

Aussi, $G(1) < 1$, sinon, on aurait $G(t) = 1$ pour tout $t > 0$, et donc $T = \infty$ p.s. (or on a supposé T à valeurs dans $\mathbf{R}_*^+$).

On peut donc poser $\lambda = -\ln G(1) > 0$ et conclure que $G(t) = e^{-\lambda t}$, i.e. que $T \sim \mathcal{E}(\lambda)$.

Lemme 1.5.3 (des 2 réveils) *Soient X et Y deux v.a. indépendantes de lois exponentielles de paramètres respectifs λ et μ. Soit $T = \min(X, Y)$ et $Z = \mathbf{1}_{(T=X)}$. Alors $T \sim \mathcal{E}(\lambda+\mu)$ et est indépendante de Z. De plus $\mathbf{P}(Z = 1) = \frac{\lambda}{\lambda+\mu}$ et $\mathbf{P}(Z = 0) = \frac{\mu}{\lambda+\mu}$.*

Par exemple, une mare contient des poissons rouges et des poissons gris. Un pêcheur totalement inexpérimenté y pêche. Soit X le premier temps de prise d'un poisson rouge et Y le premier temps de prise d'un poisson gris. Alors $T = \min\{X, Y\}$ est le temps de prise de son premier poisson, et $Z = 1$ s'il est rouge, 0 s'il est gris. *Par nature* (absence de mémoire), T suit forcément une loi exponentielle. Le lemme des deux réveils, très important, permet de calculer le paramètre de T et la loi de Z.

Preuve: Pour $t > 0$, on a

$$\begin{aligned}
\mathbf{P}(T > t, Z = 1) &= \mathbf{P}(X > t, X \le Y) \\
&= \int_0^\infty \int_0^\infty \mathbf{1}_{\{x>t, x\le y\}} \lambda\mu e^{-\lambda x} e^{-\mu y}\, dxdy \\
&= \int_0^\infty \mathbf{1}_{\{x>t\}} \lambda e^{-\lambda x} e^{-\mu x}\, dx \\
&= \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}.
\end{aligned}$$

De même,

$$\mathbf{P}(T > t, Z = 0) = \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t}.$$

En sommant, on conclut que $\mathbf{P}(T > t) = e^{-(\lambda+\mu)t}$, donc $T \sim \mathcal{E}(\lambda + \mu)$. En prenant $t = 0$, on voit que $\mathbf{P}(Z = 1) = \frac{\lambda}{\lambda+\mu}$ et $\mathbf{P}(Z = 0) = \frac{\mu}{\lambda+\mu}$. L'indépendance est aussi démontrée par les formules établies.

On peut généraliser ce lemme à plus de réveils.

Lemme 1.5.4 *Soit $\{X_k, k = 1, \ldots, n\}$ une famille de v.a. indépendantes de lois exponentielles de paramètres $\lambda_1, \ldots, \lambda_n$. Soit aussi $\lambda = \sum_{k=1}^n \lambda_k$. Soit $T = \min_{k=1,\ldots,n} X_k$ et soit $Z = \sum_{k=1}^n k\mathbf{1}_{\{T=k\}}$. Alors $T \sim \mathcal{E}(\lambda)$, T est indépendante de Z, et $\mathbf{P}(Z = k) = \lambda_k/\lambda$ pour tout $k = 1, \ldots, n$.*

Preuve: Pour $k \in \{1, \ldots, n\}$ fixé, on a $\mathbf{P}(X_i \ge x \ \forall\, i \ne k) = \exp(-\sum_{i\ne k} \lambda_i x)$, puis

$$\begin{aligned}
\mathbf{P}(T > t, Z = k) &= \mathbf{P}(X_k > t, X_i \ge X_k, \forall i \ne k) \\
&= \mathbf{E}\Big[\mathbf{1}_{\{X_k>t\}} \exp\Big(-\sum_{i\ne k} \lambda_i X_k\Big)\Big] \\
&= \int_0^\infty \mathbf{1}_{\{x>t\}} \exp\Big(-\sum_{i\ne k} \lambda_i x\Big) \lambda_k \exp(-\lambda_k x) dx \\
&= \int_t^\infty \lambda_k \exp\Big(-\lambda x\Big) dx \\
&= \frac{\lambda_k}{\lambda} e^{-\lambda t}.
\end{aligned}$$

En sommant sur k on voit que $\mathbf{P}(T > t) = \exp(-\lambda t)$, donc $T \sim \mathcal{E}(\lambda)$. En prenant $t = 0$, on voit que $\mathbf{P}(Z = k) = \lambda_k/\lambda$. L'indépendance est aussi prouvée.

Enfin, rappelons quelques faits sur les lois Gamma. Rappelons d'abord que la fonction Γ est définie sur $\mathbf{R}^+$ par la formule

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t}\, dt,\ x > 0.$$

On vérifie par récurrence, à l'aide d'une intégration par parties, que $\Gamma(n) = (n-1)!$ si $n \in \mathbf{N}^*$.

Définition 1.5.5 *On appelle loi Gamma de paramètre (λ, α) où $\lambda, \alpha \geq 0$, la probabilité sur $\mathbf{R}^+$ de densité*

$$f_{\lambda,\alpha}(x) = \frac{(\lambda x)^{\alpha-1} \lambda e^{-\lambda x}}{\Gamma(\alpha)} \mathbf{1}_{\{x>0\}}.$$

Pour $\alpha = 1$, on retrouve la loi exponentielle. La preuve de la proposition suivante est laissée au lecteur.

Proposition 1.5.6 *La somme de deux variables aléatoires indépendantes de lois Gamma de paramètres (λ, α_1) et (λ, α_2) est une variable aléatoire de loi Gamma de paramètre $(\lambda, \alpha_1 + \alpha_2)$.*

Corrolaire important :

Corollaire 1.5.7 *Si $X_1, X_2, \cdots, X_n$ sont des variables aléatoires indépendantes de loi exponentielle de paramètre λ, alors $X_1 + X_2 + \cdots + X_n$ a une loi Gamma de paramètre (λ, n).*

Chapitre 2

Processus de Poisson

2.1 P.A.I.S.

Nous voulons généraliser à $\mathbf{R}^+$ la construction de points aléatoires sur $\mathbf{N}$ obtenue par le processus de Bernoulli dans le chapitre précédent. Ceci nous conduit d'abord à la définition suivante, qui généralise au temps continu la notion de marche aléatoire.

Définition 2.1.1 *Un processus $(X_t)_{t\in\mathbf{R}^+}$ à valeurs dans $\mathbf{R}$ est un processus à accroissements indépendants et stationnaires (P.A.I.S.) si, en notant $\mathcal{F}_t = \sigma(X_r, r \leq t)$,*

1. *L'application $t \to X_t$ est p.s. continue à droite sur $\mathbf{R}^+$.*
2. *Pour tout $s, t \geq 0$, $X_{t+s} - X_s$ est indépendant de $\mathcal{F}_s$.*
3. *Pour tout $s, t \geq 0$, $X_{t+s} - X_s$ a la même loi que $X_t - X_0$.*
4. *$X_0 = 0$.*

Par exemple, le mouvement d'une particule de gaz est souvent modélisée par un PAIS. Plus concrètement, le nombre de poissons N_t pris par un pêcheur inexpérimenté durant $[0, t]$ est un PAIS (en supposant que la mare contient une infinité de poissons). En effet, les points 1. et 4. sont satisfaits. De plus, $N_{t+s} - N_t$, qui représente le nombre de poissons pêchés durant $]t, t+s]$ est indépendant de tout ce qui s'est passé jusqu'à l'instant t (point 2.) et a bien sûr même loi que le nombre de poissons $N_s = N_s - N_0$ pêchés durant $]0, s]$ (point 3.).

Nous allons généraliser la propriété de Markov forte des marches aléatoires.

Définition 2.1.2 *Soit $(X_t)_{t\in\mathbf{R}_+}$ un processus et, pour $t \geq 0$, $\mathcal{F}_t = \sigma(X_s, s \leq t)$.*

(i) Une v.a. τ à valeurs dans $\mathbf{R}^+ \cup \{\infty\}$ est un temps d'arrêt (pour $(X_t)_{t\in\mathbf{R}_+}$) si

$$\{\tau \leq t\} \in \mathcal{F}_t \qquad \textit{pour tout } t \geq 0.$$

(ii) On pose alors $\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty \ : \ A \cap \{\tau \leq t\} \in \mathcal{F}_t \ \forall\, t \geq 0\}$.

On montrera en exercice que $\mathcal{F}_\tau$ est une tribu et que si σ, τ sont deux temps d'arrêt tels que $\tau \leq \sigma$, alors $\mathcal{F}_\tau \subset \mathcal{F}_\sigma$ (observer que $A \cap \{\sigma \leq t\} = A \cap \{\tau \leq t\} \cap \{\sigma \leq t\}$).

Comme dans le cas discret, un temps aléatoire τ est un temps d'arrêt si pour tout $t \geq 0$, en observant les valeurs de X_s pour $s \in [0,t]$, on est capable de dire si $\tau \leq t$ ou non.

Aussi, pour un évenement $A \in \mathcal{F}$, on a $A \in \mathcal{F}_\tau$ si, lorsqu'on observe τ et X_s pour $s \in [0,\tau]$, on peut décider si A est réalisé ou non. De même, une v.a. Z est $\mathcal{F}_\tau$-mesurable si l'on peut déterminer sa valeur à partir de celles de τ et des X_s pour $s \in [0,\tau]$.

Lemme 2.1.3 *Tout temps d'arrêt τ est la limite d'une suite décroissante de temps d'arrêt τ_n, où τ_n est à valeurs dans l'ensemble dénombrable $\{k10^{-n}, k \in \mathbf{N}\} \cup \{+\infty\}$.*

Preuve: Il suffit de prendre
$$\tau_n = \frac{\lfloor 10^n \tau \rfloor + 1}{10^n}.$$
On a évidemment que τ_n décroit vers τ quand $n \to \infty$. Pour $n \geq 0$ fixé, montrons que τ_n est un temps d'arrêt : pour $t \geq 0$, on a
$$\{\tau_n \leq t\} = \{\lfloor 10^n \tau \rfloor \leq 10^n t - 1\} = \{\lfloor 10^n \tau \rfloor \leq \lfloor 10^n t - 1 \rfloor\},$$
soit encore
$$\{\tau_n \leq t\} = \{10^n \tau < \lfloor 10^n t - 1 \rfloor + 1\} = \{\tau < (\lfloor 10^n t - 1 \rfloor + 1)10^{-n}\},$$
qui est bien dans $\mathcal{F}_t$ car $(\lfloor 10^n t - 1 \rfloor + 1)10^{-n} \leq t$.

Les questions de mesurabilité concernant les temps d'arrêt en temps continu peuvent être délicates. A titre d'exemple montrons le lemme suivant.

Lemme 2.1.4 *Si $(X_t)_{t \in \mathbf{R}^+}$ est continu à droite, pour tout temps d'arrêt τ fini, X_τ est $\mathcal{F}_\tau$-mesurable.*

Preuve: Utilisons le lemme précédent pour écrire $\tau = \lim \tau_n$, où τ_n est une suite décroissante de temps d'arrêt à valeurs dans $\{k10^{-n}, k \in \mathbf{N}\}$. La variable aléatoire $X_{t \wedge \tau_n}$ est $\mathcal{F}_t$-mesurable car
$$X_{t \wedge \tau_n} = \sum_{\{k\,:\,k10^{-n} \leq t\}} X_{k10^{-n}} \mathbf{1}_{\{\tau_n = k10^{-n}\}} + X_t \mathbf{1}_{\{\tau_n > t\}}.$$
(Tous les termes de cette expression sont $\mathcal{F}_t$-mesurables : pour $s \leq t$, on a $\{\tau_n = s\} \in \mathcal{F}_s \subset \mathcal{F}_t$, X_s est $\mathcal{F}_s$-mesurable et donc $\mathcal{F}_t$-mesurable, enfin, $\{\tau_n > t\} = \{\tau_n \leq t\}^c \in \mathcal{F}_t$ puisque τ_n est un temps d'arrêt).

On en déduit que $X_{\tau \wedge t} = \lim_{n \to \infty} X_{t \wedge \tau_n}$ est aussi $\mathcal{F}_t$ mesurable (on utilise ici que X est continu à droite que que la suite τ_n décroit vers τ).

Enfin, pour montrer que X_τ est $\mathcal{F}_\tau$-mesurable, il faut vérifier que, pour tout borélien B, l'ensemble $\{X_\tau \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t$. C'est clair car
$$\{X_\tau \in B\} \cap \{\tau \leq t\} = \{X_{\tau \wedge t} \in B\} \cap \{\tau \leq t\}.$$

Théorème 2.1.5 Propriété de Markov forte des P.A.I.S. *Considérons un P.A.I.S.* $(X_t)_{t\in\mathbf{R}_+}$*. Pour tout temps d'arrêt* τ *presque sûrement fini, le processus* $(X_{t+\tau}-X_\tau)_{t\in\mathbf{R}^+}$ *est de même loi que* $(X_t)_{t\in\mathbf{R}^+}$ *et est indépendant de* $\mathcal{F}_\tau$*.*

Ce théorème est assez évident si $\tau = t_0$ est déterministe : c'est (en gros) la définition des P.A.I.S.

Preuve: Soit Z une v.a. $\mathcal{F}_\tau$-mesurable bornée, $\phi_1,\cdots,\phi_k$ des fonctions continues bornées sur $\mathbf{R}$ et $0\le t_1<\ldots<t_k$. L'objectif est de montrer que $I=J$, où

$$I=\mathbf{E}[Z\phi_1(X_{\tau+t_1}-X_\tau)\phi_2(X_{\tau+t_2}-X_\tau)\cdots\phi_k(X_{\tau+t_k}-X_\tau)]$$

et

$$J=\mathbf{E}[Z]\mathbf{E}[\phi_1(X_{t_1})\phi_2(X_{t_2})\cdots\phi_k(X_{t_k})].$$

Utilisons le lemme 2.1.3 pour écrire $\tau=\lim\downarrow\tau_n$ où chaque τ_n est à valeurs dans $\{\ell 10^{-n},\ \ell\in\mathbf{N}\}$ et introduisons

$$I_n=\mathbf{E}[Z\phi_1(X_{\tau_n+t_1}-X_{\tau_n})\phi_2(X_{\tau_n+t_2}-X_{\tau_n})\cdots\phi_k(X_{\tau_n+t_k}-X_{\tau_n})].$$

Nous allons montrer que $I_n=J$ pour tout n. Cela suffira, puisqu'on a $I=\lim_n I_n$ (car X est continu à droite, car τ_n décroit p.s. vers τ, car les fonctions ϕ_ℓ sont continues, et par convergence dominée, puisque tout est borné). Remarquons que Z est $\mathcal{F}_{\tau_n}$-mesurable (puisque $\tau\le\tau_n$) et écrivons

$$\begin{aligned} I_n &= \sum_{\ell\ge 0}\mathbf{E}[Z\mathbf{1}_{\{\tau_n=\ell 10^{-n}\}}\phi_1(X_{\tau_n+t_1}-X_{\tau_n})\cdots\phi_k(X_{\tau_n+t_k}-X_{\tau_n})] \\ &= \sum_{\ell\ge 0}\mathbf{E}[Z\mathbf{1}_{\{\tau_n=\ell 10^{-n}\}}\phi_1(X_{\ell 10^{-n}+t_1}-X_{\ell 10^{-n}})\cdots\phi_k(X_{\ell 10^{-n}+t_k}-X_{\ell 10^{-n}})]. \end{aligned}$$

Pour ℓ fixé, on voit que $Z\mathbf{1}_{\{\tau_n=\ell 10^{-n}\}}$ est $\mathcal{F}_{\ell 10^{-n}}$-mesurable (car τ_n est un temps d'arrêt et Z est $\mathcal{F}_{\tau_n}$-mesurable). De plus, par définition des P.A.I.S., le k-uplet

$$(X_{\ell 10^{-n}+t_1}-X_{\ell 10^{-n}},\cdots,X_{\ell 10^{-n}+t_k}-X_{\ell 10^{-n}})$$

est indépendant de $\mathcal{F}_{\ell 10^{-n}}$ et a même loi que $(X_{t_1},\cdots,X_{t_k})$. Ainsi,

$$I_n=\sum_{\ell\ge 0}\mathbf{E}[Z\mathbf{1}_{\{\tau_n=\ell 10^{-n}\}}]\mathbf{E}[\phi_1(X_{t_1})\phi_2(X_{t_2})\cdots\phi_k(X_{t_k})].$$

Cette dernière quantité n'est autre que J, puisque $\sum_{\ell\ge 0}\mathbf{1}_{\{\tau_n=\ell 10^{-n}\}}=1$.

2.2 Processus de Poisson

Définition 2.2.1 *Un processus ponctuel sur* $\mathbf{R}^+$ *est une suite strictement croissante de v.a.* $0\le T_1<T_2<\cdots$ *telle que* $T_n\to\infty$ *p.s. quand* $n\to\infty$*. Le processus de comptage associé à ce processus ponctuel est le processus* $(N_t)_{t\in\mathbf{R}_+}$ *défini par*

$$N_t=\sum_{k=1}^{+\infty}\mathbf{1}_{\{T_k\le t\}}=\#\{k\ge 1\ :\ T_k\le t\}.$$

Remarque 2.2.2 *Le processus de comptage* $(N_t)_{t\in\mathbf{R}_+}$ *de n'importe quel processus ponctuel sur* $\mathbf{R}^+$ *jouit des propriétés suivantes :* $t \to N_t$ *est continu à droite sur* $\mathbf{R}^+$*, croissant, à valeurs dans* $\mathbf{N}$*, on a* $N_0 = 0$ *et* $\lim_{t\to\infty} N_t = \infty$*. De plus, pour tout* $n \geq 1$*, on a*

$$T_n \leq t \Longleftrightarrow N_t \geq n$$

Enfin, pour tout $n \geq 1$*,*

$$T_n = \inf\{t \geq 0 \ : \ N_t = n\}$$

et T_n *est un temps d'arrêt (du processus* $(N_t)_{t\in\mathbf{R}^+}$*).*

Rappelons ce qu'on appelle **loi de Poisson** de paramètre $\alpha > 0$: une v.a. N à valeurs dans $\mathbf{N}$ suit la loi $\mathcal{P}(\alpha)$ si pour tout $n \in \mathbf{N}$,

$$\mathbf{P}(N = n) = e^{-\alpha}\frac{\alpha^n}{n!}.$$

Le résultat suivant est fondammental.

Théorème 2.2.3 *Soit* $0 \leq T_1 < T_2 < \cdots$ *un processus ponctuel et* $(N_t)_{t\geq 0}$ *le processus de comptage associé.*

(i) Si $(N_t)_{t\geq 0}$ *est un P.A.I.S., alors il existe* $\lambda > 0$ *tel que les v.a.* $(T_{n+1} - T_n)_{n\geq 0}$ *soient i.i.d. de loi* $\mathcal{E}(\lambda)$ *(avec* $T_0 = 0$*).*

(ii) Soit $\lambda > 0$*. Si les v.a.* $(T_{n+1} - T_n)_{n\geq 0}$ *sont i.i.d. de loi* $\mathcal{E}(\lambda)$ *(avec* $T_0 = 0$*), alors* $(N_t)_{t\geq 0}$ *est un P.A.I.S. De plus, pour tout* $0 \leq s < t$*,* $N_t - N_s \sim \mathcal{P}(\lambda(t-s))$*.*

On dit qu'un tel processus ponctuel $0 \leq T_1 < T_2 < \cdots$ *(ou, par abus de langage, son processus de comptage* $(N_t)_{t\geq 0}$*) est un processus de Poisson de paramètre* λ*.*

Par exemple, si N_t est le nombre de poissons pris par un pêcheur inexpérimenté durant $[0, t]$, le processus $(N_t)_{t\geq 0}$ est clairement un P.A.I.S. (on a déjà vu cela). Ainsi, on peut appliquer le théorème.

De même, dans certaines conditions, le nombre de clients N_t arrivant à un péage durant $[0, t]$ est un P.A.I.S. (en gros, les “conditions” reflètent la fluidité du trafic) et on peut appliquer le théorème.

Preuve:

Point (i), étape 1 Montrons déjà qu'il existe $\lambda > 0$ tel que $T_1 \sim \mathcal{E}(\lambda)$. Pour cela, il suffit de montrer que pour tout $s, t \geq 0$, $\mathbf{P}(T_1 > t + s) = \mathbf{P}(T_1 > t)\mathbf{P}(T_1 > s)$. Mais

$$\mathbf{P}(T_1 > t + s) = \mathbf{P}(N_{t+s} = 0) = \mathbf{P}(N_{t+s} - N_t = 0, N_t = 0).$$

Comme maintenant $(N_t)_{t\geq 0}$ est un P.A.I.S. par hypothèse, on trouve

$$\mathbf{P}(T_1 > t + s) = \mathbf{P}(N_{t+s} - N_t = 0)\mathbf{P}(N_t = 0) = \mathbf{P}(N_s = 0)\mathbf{P}(N_t = 0),$$

soit encore

$$\mathbf{P}(T_1 > t + s) = \mathbf{P}(T_1 > s)\mathbf{P}(T_1 > t).$$

Point (i), étape 2. Fixons un entier n. Par la propriété de Markov forte du P.A.I.S. $(N_t)_{t\geq 0}$ et comme T_n est un temps d'arrêt, on sait qu'en posant $\tilde{N}_t := N_{T_n+t} - N_{T_n}$, $(\tilde{N}_t)_{t\geq 0}$ a même loi que $(N_t)_{t\geq 0}$ et est indépendant de $\mathcal{F}_{T_n}$. Donc

$$\tilde{T}_1 = \inf\{t \geq 0 \ : \ \tilde{N}_t = 1\}$$

a même loi que T_1 (la loi $\mathcal{E}(\lambda)$) et est indépendant de $\mathcal{F}_{T_n}$. Mais, comme $N_{T_n} = n$, on a

$$\tilde{T}_1 = \inf\{t \geq 0 \ : \ N_{T_n+t} = n+1\} = T_{n+1} - T_n.$$

On a montré que pour tout $n \geq 1$, $T_{n+1} - T_n \sim \mathcal{E}(\lambda)$ et est indépendant de $\mathcal{F}_{T_n}$. Comme $T_1, \cdots, T_n$ sont clairement $\mathcal{F}_{T_n}$-mesurables, ainsi que $T_1, T_2 - T_1, \cdots, T_n - T_{n-1}$, on conclut aisément que les v.a. $(T_{n+1} - T_n)_{n\geq 0}$ sont i.i.d. de loi $\mathcal{E}(\lambda)$ (avec $T_0 = 0$).

Point (ii). On va vérifier que pour tout $k \geq 1$, pour tout $0 \leq t_1 < \cdots < t_k$, les v.a. $N_{t_i} - N_{t_{i-1}}$ sont indépendantes et de lois $\mathcal{P}(\lambda(t_i - t_{i-1}))$. Pour simplifier, on suppose que $k = 3$. On pose $X_1 = T_1$ et, pour $\ell \geq 2$, $X_\ell = T_\ell - T_{\ell-1}$. Par hypothèse, la famille $(X_\ell)_{\ell\geq 1}$ est i.i.d. de loi $\mathcal{E}(\lambda)$. Pour ℓ_1, ℓ_2, ℓ_3 dans $\mathbf{N}$, on a

$$\begin{aligned}
I &:= \mathbf{P}(N_{t_1} = \ell_1, N_{t_2} - N_{t_1} = \ell_2, N_{t_3} - N_{t_2} = \ell_3) \\
&= \mathbf{P}\Big(T_{\ell_1} \leq t_1 < T_{\ell_1+1}, T_{\ell_1+\ell_2} \leq t_2 < T_{\ell_1+\ell_2+1}, T_{\ell_1+\ell_2+\ell_3} \leq t_3 < T_{\ell_1+\ell_2+\ell_3+1}\Big) \\
&= \mathbf{P}\Big(X_1 + \cdots + X_{\ell_1} \leq t_1 < X_1 + \cdots + X_{\ell_1+1}, \\
&\qquad X_1 + \cdots + X_{\ell_1+\ell_2} \leq t_2 < X_1 + \cdots + X_{\ell_1+\ell_2+1}, \\
&\qquad X_1 + \cdots + X_{\ell_1+\ell_2+\ell_3} \leq t_3 < X_1 + \cdots + X_{\ell_1+\ell_2+\ell_3+1}\Big) \\
&= \lambda^{\ell_1+\ell_2+\ell_3+1} \int_{\mathbf{R}_+^{\ell_1+\ell_2+\ell_3+1}} e^{-\lambda(x_1+\cdots+x_{\ell_1+\ell_2+\ell_3+1})} \\
&\qquad \mathbf{1}_{\{x_1+\cdots+x_{\ell_1} \leq t_1 < x_1+\cdots+x_{\ell_1+1}\}} \mathbf{1}_{\{x_1+\cdots+x_{\ell_1+\ell_2} \leq t_2 < x_1+\cdots+x_{\ell_1+\ell_2+1}\}} \\
&\qquad \mathbf{1}_{\{x_1+\cdots+x_{\ell_1+\ell_2+\ell_3} \leq t_3 < x_1+\cdots+x_{\ell_1+\ell_2+\ell_3+1}\}} dx_1 \cdots dx_{\ell_1+\ell_2+\ell_3+1}.
\end{aligned}$$

On effectue alors le changement de variables triangulaire $s_1 = x_1$, $s_2 = x_1 + x_2$, ..., $s_{\ell_1+\ell_2+\ell_3+1} = x_1 + x_2 + \cdots + x_{\ell_1+\ell_2+\ell_3+1}$, dont le jacobien vaut 1, et on trouve

$$\begin{aligned}
I &= \lambda^{\ell_1+\ell_2+\ell_3+1} \int_{\mathbf{R}_+^{\ell_1+\ell_2+\ell_3+1}} e^{-\lambda s_{\ell_1+\ell_2+\ell_3+1}} \mathbf{1}_{\{s_1 < s_2 < \cdots < s_{\ell_1} \leq t_1\}} \\
&\qquad \mathbf{1}_{\{t_1 \leq s_{\ell_1+1} < \cdots < s_{\ell_1+\ell_2} \leq t_2\}} \mathbf{1}_{\{t_2 \leq s_{\ell_1+\ell_2+1} < \cdots < s_{\ell_1+\ell_2+\ell_3} \leq t_3\}} \\
&\qquad \mathbf{1}_{\{s_{\ell_1+\ell_2+\ell_3+1} \geq t_3\}} ds_1 \cdots ds_{\ell_1+\ell_2+\ell_3+1},
\end{aligned}$$

qui se "sépare" bien en $I = \lambda^{\ell_1+\ell_2+\ell_3+1} \times I_1 \times I_2 \times I_3 \times I_4$, où:

$$\begin{aligned}
I_1 &= \int_{(\mathbf{R}^+)^{\ell_1}} \mathbf{1}_{\{s_1 < s_2 < \cdots < s_{\ell_1} \leq t_1\}} ds_1 \cdots ds_{\ell_1}, \\
I_2 &= \int_{(\mathbf{R}^+)^{\ell_2}} \mathbf{1}_{\{t_1 \leq s_{\ell_1+1} < \cdots < s_{\ell_1+\ell_2} \leq t_2\}} ds_{\ell_1+1} \cdots ds_{\ell_1+\ell_2},
\end{aligned}$$

$$
\begin{aligned}
I_3 &= \int_{(\mathbf{R}^+)^{\ell_3}} \mathbf{1}_{\{t_2 \le s_{\ell_1+\ell_2+1} < \cdots < s_{\ell_1+\ell_2+\ell_3} \le t_3\}} ds_{\ell_1+\ell_2+1} \cdots ds_{\ell_1+\ell_2+\ell_3}, \\
I_4 &= \int_{\mathbf{R}^+} e^{-\lambda s_{\ell_1+\ell_2+\ell_3+1}} \mathbf{1}_{\{s_{\ell_1+\ell_2+\ell_3+1} \ge t_3\}} ds_{\ell_1+\ell_2+\ell_3+1}.
\end{aligned}
$$

Bien sûr, $I_4 = e^{-\lambda t_3}/\lambda$. On a de plus (exercice), pour tout $k \ge 1$, pour tout $a < b$,

$$
J_k(a,b) = \int_{\mathbf{R}^k} \mathbf{1}_{\{a < s_1 < \cdots < s_k < b\}} ds_1 \cdots ds_k = \frac{(b-a)^k}{k!}.
$$

Comme $I_1 = J_{\ell_1}(0, t_1)$, comme $I_2 = J_{\ell_2}(t_1, t_2)$ et comme $I_3 = J_{\ell_3}(t_2, t_3)$, on trouve

$$
I = \lambda^{\ell_1+\ell_2+\ell_3} e^{-\lambda t_3} \frac{t_1^{\ell_1}}{\ell_1!} \frac{(t_2-t_1)^{\ell_2}}{\ell_2!} \frac{(t_3-t_2)^{\ell_3}}{\ell_3!}.
$$

On conclut la preuve en remarquant que

$$
I = e^{-\lambda t_1} \frac{(\lambda t_1)^{\ell_1}}{\ell_1!} e^{-\lambda(t_2-t_1)} \frac{[\lambda(t_2-t_1)]^{\ell_2}}{\ell_2!} e^{-\lambda(t_3-t_2)} \frac{[\lambda(t_3-t_2)]^{\ell_3}}{\ell_3!}.
$$

Remarque 2.2.4 *(i) Soit $(N_t)_{t\ge 0}$ un processus de Poisson de paramètre λ. Alors pour tout $a > 0$, le processus $(N_{t+a} - N_a)_{t \ge 0}$ est un processus de Poisson de paramètre λ (car c'est bien sûr un P.A.I.S. de même loi que $(N_t)_{t\ge 0}$, cela découle directement de la définition des P.A.I.S.). Donc $T_1^a = \inf\{t \ge 0 \ : \ N_{t+a} - N_a = 1\}$ suit la loi $\mathcal{E}(\lambda)$. Or (exercice) $T_1^a = T_{N_a+1} - T_{N_a}$.*

(ii) Paradoxe des autobus. Des autobus arrivent à une station S suivant un processus $(N_t)_{t\ge 0}$ de Poisson de paramètre 4 (l'unité de temps étant l'heure). Il passe donc en moyenne 4 autobus par heure (par exemple, durant la première heure, N_1 autobus sont passés, et comme $N_1 \sim \mathcal{P}(4 \times 1)$, $\mathbf{E}[N_1] = 4$). J'arrive à l'instant a. Donc N_a autobus sont déjà passés, le suivant arrivera à l'instant T_{N_a+1}. La durée D de mon attente sera donc égale à $T_{N_a+1} - a$, qui suit une loi $\mathcal{E}(4)$, soit en moyenne $\mathbf{E}[D] = 1/4$ d'heure.

Comme on arrive "entre" deux autobus et qu'il passe en moyenne un autobus tous les 1/4 d'heure, on aurait pu penser que la durée moyenne de l'attente serait de 1/8 heure...

Le dernier résultat concernant les processus de Poisson sur $\mathbf{R}^+$ est tout aussi crucial, c'est une construction alternative à partir de v.a. de loi uniforme.

Proposition 2.2.5 *Soit $T > 0$ et $\lambda > 0$ fixés. Soit $Z \sim \mathcal{P}(\lambda T)$ et $(X_i)_{i \ge 1}$ des v.a. i.i.d. de loi uniforme sur $[0, T]$ (indépendantes de Z). Pour $t \in [0, T]$, soit*

$$
Y_t = \sum_{i=1}^{Z} \mathbf{1}_{\{X_i \le t\}}.
$$

Alors $(Y_t)_{t \in [0,T]}$ a même loi que $(N_t)_{t\in[0,T]}$, où $(N_t)_{t\ge 0}$ est un processus de Poisson de paramètre λ.

Preuve: Nous allons montrer, et cela suffit, que pour tout $k \geq 1$, pour tout $0 = t_0 < t_1 < \cdots < t_{k-1} < t_k = T$, pour tout $n_1 \in \mathbf{N}$, ..., $n_k \in \mathbf{N}$, $I = J$, où

$$I = \mathbf{P}(Y_{t_1} = n_1, Y_{t_2} - Y_{t_1} = n_2, \ldots, Y_{t_k} - Y_{t_{k-1}} = n_k)$$

et

$$J = \mathbf{P}(N_{t_1} = n_1, N_{t_2} - N_{t_1} = n_2, \ldots, N_{t_k} - N_{t_{k-1}} = n_k).$$

On observe que $\{Y_{t_1} = n_1, Y_{t_2} - Y_{t_1} = n_2, \ldots, Y_{t_k} - Y_{t_{k-1}} = n_k\}$ ssi $Z = n_1 + \ldots + n_k$ et parmi les $n_1 + ... + n_k$ v.a. X_i, exactement n_1 sont dans $[0, t_1]$, n_2 dans $]t_1, t_2]$, ..., exactement n_k sont dans $]t_{k-1}, t_k]$. Du coup,

$$\begin{aligned} I &= e^{-\lambda T} \frac{(\lambda T)^{n_1 + \ldots + n_k}}{(n_1 + \ldots + n_k)!} C_{n_1 + \ldots + n_k}^{n_1} \Big(\frac{t_1}{T}\Big)^{n_1} C_{n_2 + \ldots + n_k}^{n_2} \Big(\frac{t_2 - t_1}{T}\Big)^{n_2} \\ &\qquad \times \cdots \times C_{n_k}^{n_k} \Big(\frac{t_k - t_{k-1}}{T}\Big)^{n_k}. \end{aligned}$$

En développant tous les coefficients binomiaux, puis en simplifiant, on trouve

$$I = e^{-\lambda T} \lambda^{n_1 + \ldots + n_k} \frac{t_1^{n_1}}{n_1!} \frac{(t_2 - t_1)^{n_2}}{n_2!} \cdots \frac{(t_k - t_{k-1})^{n_k}}{n_k!}.$$

Enfin, comme $t_1 + (t_2 - t_1) + ... + (t_k - t_{k-1}) = t_k = T$,

$$I = e^{-\lambda t_1} \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda(t_2 - t_1)} \frac{[\lambda(t_2 - t_1)]^{n_2}}{n_2!} \times \cdots \times e^{-\lambda(t_k - t_{k-1})} \frac{[\lambda(t_k - t_{k-1})]^{n_k}}{n_k!}.$$

C'est bien sûr ce que vaut J, par définition du processus de Poisson.

2.3 Processus ponctuel de Poisson

Définition 2.3.1 *Soit $(E, \mathcal{E}, \mu)$ un espace mesuré muni d'une mesure σ-finie μ. Soit $(X_k)_{k \in K}$ un ensemble aléatoire fini ou dénombrable de points de E. Pour $A \in \mathcal{E}$, on pose*

$$N_A = \sum_{k \in K} \mathbf{1}_{\{X_k \in A\}}.$$

On dit que $(X_k)_{k \in K}$ est un processus ponctuel de Poisson (P.P.P.) sur E d'intensité μ (et que $(N_A)_{A \in \mathcal{E}}$ est la fonction de comptage associée) si pour tout $n \geq 1$, pour tout $A_1, \cdots, A_n \in \mathcal{E}$ deux à deux disjoints, les v.a. $N_{A_1}, \cdots, N_{A_n}$ sont indépendantes et de lois $\mathcal{P}(\mu(A_1)), \cdots, \mathcal{P}(\mu(A_n))$.

Par convention une variable aléatoire de loi de Poisson de paramètre $+\infty$ est une variable aléatoire identiquement infinie.

Remarque 2.3.2 *On admettra que la définition ci-dessus caractérise la loi du P.P.P. sur E d'intensité μ. On va montrer l'existence de ce processus quand $\mu(E) < \infty$ dans la Proposition 2.3.3 et quand $\mu(E) = \infty$ dans la Remarque 2.3.5. Ces énoncés pourront donc être considérés, a posteriori, comme des définitions.*

Remarquons que N_A n'est autre que le nombre de points (parmi $(X_k)_{k\in K}$) qui appartiennent à A.

Lorsque μ est finie (i.e. $\mu(E) < +\infty$), le nombre total N_E de points du P.P.P. est aléatoire mais fini presque sûrement (puisqu'il suit une loi de Poisson de paramètre $\mu(E)$). Lorsque $\mu(E) = \infty$, N_E est p.s. infini.

Notre premier but est de construire un processus ponctuel de Poisson. Cela prouvera leur existence. Cela permettra aussi de les simuler sur ordinateur. On commence par le cas fini.

Proposition 2.3.3 *Supposons que $\mu(E) < \infty$ et considérons la probabilité $\nu(\cdot) = \mu(\cdot)/\mu(E)$. Donnons-nous des v.a. i.i.d. $U_1, U_2, \cdots$ (à valeurs dans E) de loi ν et une v.a. $Z \sim \mathcal{P}(\mu(E))$ indépendante de la suite $(U_n)_{n\geq 1}$. Alors*

$$\{U_1, U_2, \cdots, U_Z\}$$

est un P.P.P. sur E d'intensité μ.

Preuve: Il suffit de montrer que pour $A_1, \cdots, A_n$ une partition de E en sous ensembles mesurables, en posant $N_A = \sum_{k=1}^{Z} \mathbf{1}_{\{U_k\in A\}}$, les v.a. $N_{A_1}, \cdots, N_{A_n}$ sont indépendantes de lois de Poisson de paramètres $\mu(A_1), \cdots, \mu(A_n)$. Pour $k_1, \cdots, k_n \in \mathbf{N}$,

$$\begin{aligned} I &:= \mathbf{P}\Big(N_{A_1} = k_1, N_{A_2} = k_2, \cdots, N_{A_n} = k_n\Big) \\ &= \mathbf{P}\Big(Z = k_1 + \cdots + k_n, N_{A_1} = k_1, N_{A_2} = k_2, \cdots, N_{A_n} = k_n\Big). \end{aligned}$$

Il faut donc que $Z = k_1 + \cdots + k_n$ et que parmi les $k_1 + \cdots + k_n$ points, k_1 tombent dans A_1, ..., k_n tombent dans A_n :

$$\begin{aligned} I &= e^{-\mu(E)} \frac{(\mu(E))^{k_1+\cdots+k_n}}{(k_1+\cdots+k_n)!} C^{k_1}_{k_1+\ldots+k_n} (\nu(A_1))^{k_1} C^{k_2}_{k_2+\ldots+k_n} (\nu(A_2))^{k_2} \\ &\qquad \times \cdots \times C^{k_n}_{k_n} (\nu(A_n))^{k_n}. \end{aligned}$$

En développant puis en simplifiant les coefficients binômiaux, on trouve

$$I = e^{-\mu(E)} (\mu(E))^{k_1+\cdots+k_n} \frac{(\nu(A_1))^{k_1}}{k_1!} \times \cdots \times \frac{(\nu(A_n))^{k_n}}{k_n!}.$$

En se rappelant que $\nu(A) = \mu(A)/\mu(E)$ et que $\mu(E) = \mu(A_1)+\cdots+\mu(A_n)$, on conclut que

$$I = e^{-\mu(A_1)} \frac{(\mu(A_1))^{k_1}}{k_1!} \times \cdots \times e^{-\mu(A_n)} \frac{(\mu(A_n))^{k_n}}{k_n!},$$

ce qu'on voulait.

Proposition 2.3.4 (Superposition) *Soit $(\mu_k)_{k\geq 1}$ une famille de mesures σ-finies sur $(E, \mathcal{E})$. Soit $\{(N^k_A)_{A\in\mathcal{E}}, k \in \mathbf{N}\}$, les fonctions de comptage d'une famille de P.P.P. sur E indépendants d'intensités $\{\mu_k, k \in \mathbf{N}\}$. Alors $N_A = \sum_{k=1}^{+\infty} N^k_A$ est la fonction de comptage d'un processus ponctuel de Poisson d'intensité $\mu = \sum_{k\geq 1} \mu_k$.*

En termes de processus ponctuels, on en déduit que l'union (dénombrable) de P.P.P. indépendants est un P.P.P. dont l'intensité est la somme des intensités (car sommer les fonctions de comptage revient à faire l'union des ensembles de points).

Remarque 2.3.5 *Ceci permet de construire des P.P.P. d'intensité μ σ-finie. En effet, on écrit $\mu = \sum_{k\geq 1} \mu_k$, où chaque mesure μ_k (sur E) est finie, pour chaque k, on construit un P.P.P. en utilisant la proposition 2.3.3, puis on fait l'union de ces P.P.P. (qu'on a choisis indépendants) en utilisant la propriété de superposition.*

Preuve: Il nous faut montrer que si $A_1, \cdots, A_n \in \mathcal{E}$ sont deux à deux disjoints, alors $N_{A_1}, \cdots, N_{A_n}$ sont indépendants (c'est évident en utilisant l'indépendance des $N^k_{A_1}, \cdots, N^k_{A_n}$ pour chaque k et l'indépendance des P.P.P.) et de lois de Poisson de paramètres $\mu(A_1), \cdots, \mu(A_n)$: ceci résulte facilement du fait que la somme de v.a. indépendantes de loi de Poisson de paramètres $(\lambda_k)_{k\geq 1}$ suit une loi de Poisson de paramètre $\sum_{k\geq 1} \lambda_k$ (y compris quand $\sum_{k\geq 1} \lambda_k = \infty$, exercice).

Remarque 2.3.6 *On note m (resp. m_+) la mesure de Lebesgue sur $\mathbf{R}$ (resp. $\mathbf{R}^+$).*

(i) Le processus de Poisson $0 < T_1 < T_2 < \cdots$ de paramètre $\lambda > 0$ est un P.P.P. sur $\mathbf{R}^+$ d'intensité λm_+. Et les notations se transposent ainsi : $N_t = N_{[0,t]}$, $N_t - N_s = N_{]s,t]}$.

(ii) Considérons deux processus de Poisson indépendants $0 < T_1 < T_2 < \cdots$ et $0 < T'_1 < T'_2 < \cdots$ de paramètre $\lambda > 0$. Alors $\{T_n, n \geq 1\} \cup \{-T'_n, n \geq 1\}$ est un P.P.P. sur $\mathbf{R}$ d'intensité λm.

Le (ii) se déduit du (i) par superposition (et un tout petit peu de travail pour voir que $\{-T'_n, n \geq 1\}$ est un P.P.P. sur $\mathbf{R}^-$ d'intensité λm_-. Quant au (i), il suffit de comparer la proposition 2.2.5 et la proposition 2.3.3.

On notera que dans (ii), le point 0 semble jouer un rôle particulier, puisqu'il est au milieu d'un intervalle dont la longueur est la somme de deux exponentielles. Mais tout point $t \in \mathbf{R}$ a cette propriété : longueur L_t de l'intervalle I_t contenant t a la même loi que $U+V$ où U et V sont indépendantes et de loi $\mathcal{E}(\lambda)$. J'entends par là que si $I_t = [A_t, B_t]$, où $B_t = \inf\{T_n : T_n \geq t, n \in \mathbf{Z}\}$ et $A_t = \sup\{T_n : T_n \leq t, n \in \mathbf{Z}\}$, alors $B_t - t$ et $t - A_t$ sont indépendantes et de loi $\mathcal{E}(\lambda)$. C'est lié au paradoxe des autobus.

Un outil souvent utile dans l'étude des processus ponctuels de Poisson est la fonctionnelle de Laplace.

Définition 2.3.7 *Soit $\{X_n, n \geq 1\}$ les points d'un processus ponctuel sur l'espace mesuré $(E, \mathcal{E})$. On appelle fonctionnelle de Laplace de ce processus l'application $\mathcal{L}$ qui à une fonction mesurable positive $f : E \to \mathbf{R}^+$ associe*

$$\mathcal{L}(f) = \mathbf{E}\Big[\exp\Big(-\sum_{n=1}^{+\infty} f(X_n)\Big)\Big].$$

Nous avons utilisé un petit abus de notation : si le processus ponctuel n'a qu'un nombre fini de points, alors la somme est bien sûr finie.

Théorème 2.3.8 *(i) La fonctionnelle de Laplace du processus ponctuel de Poisson sur $(E, \mathcal{E})$ d'intensité μ est donnée par (pour toute $f : E \mapsto \mathbf{R}^+$ mesurable)*

$$\mathcal{L}(f) = \exp\Big(- \int_E (1 - e^{-f(x)})\, \mu(dx) \Big).$$

(ii) Réciproquement, si $\{X_n, n \geq 1\}$ est un processus ponctuel dont la fonctionnelle de Laplace vérifie

$$\mathcal{L}(f) = \exp\Big(- \int_E (1 - e^{-f(x)})\, \mu(dx) \Big).$$

pour toute fonction $f : E \mapsto \mathbf{R}^+$ mesurable prenant un nombre fini de valeurs, alors $\{X_n, n \geq 1\}$ est un P.P.P. d'intensité μ.

Preuve: On commence par (i) dans le cas où f ne prend qu'un nombre fini de valeurs (distinctes) $\alpha_1, \cdots, \alpha_k$. Si $A_i = \{f = \alpha_i\}$ on a $f = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}$. De plus, les A_i sont deux à deux disjoints. Les v.a. $N_{A_i} = \sum_n \mathbf{1}_{\{X_n \in A_i\}}$ sont indépendantes et de loi de Poisson de paramètre $\mu(A_i)$. En se rappelant que si $N \sim \mathcal{P}(\lambda)$, alors $\mathbf{E}[\exp(-\alpha N)] = \exp(-\lambda(1 - e^{-\alpha}))$, on trouve donc

$$\begin{aligned}
\mathcal{L}(f) &= \mathbf{E}\Big[\exp\Big(- \sum_{n=1}^{+\infty} \sum_{i=1}^{k} \alpha_i \mathbf{1}_{\{X_n \in A_i\}} \Big) \Big] \\
&= \mathbf{E}\Big[\exp\Big(- \sum_{i=1}^{k} \sum_{n=1}^{+\infty} \alpha_i \mathbf{1}_{\{X_n \in A_i\}} \Big) \Big] \\
&= \mathbf{E}\Big[\exp\Big(- \sum_{i=1}^{k} \alpha_i N_{A_i} \Big) \Big] \\
&= \prod_{i=1}^{k} e^{-\mu(A_i)(1 - e^{-\alpha_i})} \\
&= \exp\Big(- \sum_{i=1}^{k} \mu(A_i)(1 - e^{-\alpha_i}) \Big) \\
&= \exp\Big(- \int_E (1 - e^{-f(x)})\, \mu(dx) \Big).
\end{aligned}$$

On montre maintenant (ii) : On considère $A_1, \ldots, A_k \in \mathcal{E}$ deux à deux disjoints, et on veut montrer que les v.a. $N_{A_1}, \cdots, N_{A_k}$ sont indépendantes, de lois de Poisson de paramètres $\mu(A_1), \cdots, \mu(A_k)$. Pour cela, il suffit de vérifier que pour tout $\alpha_1, \cdots, \alpha_k \geq 0$, on a

$$\mathbf{E}\Big[\exp(- \sum_1^k \alpha_i N_{A_i}) \Big] = \prod_1^k \exp\Big(- \mu(A_i)(1 - e^{-\alpha_i}) \Big).$$

C'est précisément ce que dit l'hypothèse appliquée à $f = \sum_{i=1}^{k} \alpha_i \mathbf{1}_{A_i}$, qui donne

$$\begin{aligned}\mathbf{E}\Big[\exp(-\sum_1^k \alpha_i N_{A_i})\Big] &= \mathcal{L}(f)\\ &= \exp\Big(-\int_E (1-e^{-f(x)})\,\mu(dx)\Big) = \prod_1^k \exp\Big(-\mu(A_i)(1-e^{-\alpha_i})\Big).\end{aligned}$$

On termine maintenant (i) : soit donc f mesurable positive quelconque, et f_k une suite de fonctions mesurables positives ne prenant qu'un nombre fini de valeurs croissant vers f (toutes les fonctions vont de E dans $\mathbf{R}$). Nous savons déjà que pour tout k,

$$\mathcal{L}(f_k) = \exp\left(-\int_E (1-e^{-f_k(x)})\,\mu(dx)\right).$$

Il est très facile de conclure en faisant tendre k vers l'infini, par convergences monotone et dominée.

Le théorème suivant est très important.

Théorème 2.3.9 *Soit $\{X_n, n \geq 1\}$ un P.P.P. sur E, d'intensité μ et $\{Y_n, n \geq 1\}$ une suite de v.a. i.i.d. à valeurs dans un espace mesurable $(F, \mathcal{F})$, de loi ν, indépendante de $\{X_n, n \geq 1\}$. Alors $\{(X_n, Y_n), n \geq 1\}$ est un P.P.P. sur $E \times F$ d'intensité $\mu \otimes \nu$.*

Le résultat suivant est un corollaire immédiat (choisir $E = \mathbf{R}^+$ et $\mu = \lambda m_+$).

Théorème 2.3.10 (Processus de Poisson marqué) *Soient $0 < T_1 < T_2 < \cdots$ les points d'un processus de Poisson sur $\mathbf{R}^+$ de paramètre λ et soit $Y_1, Y_2, \cdots$ une suite de variables aléatoires à valeurs dans un ensemble F, i.i.d. de loi ν, et indépendante de $\{T_n, n \geq 1\}$. Alors, $\{(T_n, Y_n), n \geq 1\}$ forme un P.P.P. sur $\mathbf{R}^+ \times F$ d'intensité $\lambda m_+ \otimes \nu$.*

Preuve du théorème 2.3.9. Calculons la fonctionnelle de Laplace du processus $\{(X_n, Y_n), n \geq 1\}$. Soit $f : E \times F \to \mathbf{R}$ une fonction mesurable positive. Posons

$$g(x) = \int_F \exp(-f(x,y))\,\nu(dy).$$

En utilisant l'indépendance des processus $\{X_n, n \geq 1\}$ et $\{Y_n, n \geq 1\}$, on trouve

$$\begin{aligned}\mathcal{L}(f) &= \mathbf{E}\Big[\exp\Big(-\sum_{n=1}^{+\infty} f(X_n, Y_n)\Big)\Big]\\ &= \mathbf{E}\Big(\mathbf{E}\Big[\exp\Big(-\sum_{n=1}^{+\infty} f(X_n, Y_n)\Big)\Big| X_k, k \in \mathbf{N}\Big]\Big)\\ &= \mathbf{E}\Big(\prod_{n=1}^{+\infty} \mathbf{E}\Big[\exp(-f(X_n, Y_n))\Big| X_k, k \in \mathbf{N}\Big]\Big)\\ &= \mathbf{E}\Big(\prod_{n=1}^{+\infty} g(X_n)\Big).\end{aligned}$$

Remarquons alors que, puisque $\{X_n, n \geq 1\}$ est un P.P.P. d'intensité μ,

$$\mathbf{E}\Big(\prod_{n=1}^{+\infty} g(X_n)\Big) = \mathbf{E}\Big(\sum_{n=1}^{+\infty} \exp(-\ln g(X_n))\Big) = \exp\Big(-\int_E (1-e^{\ln g(x)})\,\mu(dx)\Big),$$

soit encore

$$\mathcal{L}(f) = \exp\Big(\int_E (g(x)-1)\,\mu(dx)\Big)$$

En remplaçant g par sa définition, on obtient donc que

$$\begin{aligned}\mathcal{L}(f) &= \exp\Big(\int_E \Big(\int_F e^{-f(x,y)}\nu(dy)-1\Big)\,\mu(dx)\Big)\\ &= \exp\Big(\int_{E\times F}(e^{-f(x,y)}-1)\,(\mu\otimes\nu)(dx,dy)\Big)\end{aligned}$$

comme désiré.

Exemple. Un pêcheur attrape des poissons suivant un processus de Poisson $0 < T_1 < T_2 < \cdots$ de paramètre $\lambda > 0$. L'unité de temps est l'heure (il pêche donc en moyenne λ poissons par heure). Notons Y_i la masse du i-ème poisson attrapé. On suppose que les Y_i sont i.i.d. de loi ν (sur $\mathbf{R}^+$) et indépendants du processus de Poisson. On appelle X_t la masse totale de poissons pêchés durant $[0,t]$.

Alors $\{(T_n, Y_n), n \geq 1\}$ est un P.P.P. sur $\mathbf{R}^+ \times \mathbf{R}^+$ d'intensité $\lambda m_+ \otimes \nu$.

De plus, $X_t = \sum_{n\geq 1} Y_n \mathbf{1}_{\{T_n \leq t\}} = \sum_{n\geq 1} f(T_n, Y_n)$, où $f(s,y) = y\mathbf{1}_{\{s\leq t\}}$. On peut donc calculer, pour tout $\alpha > 0$,

$$\mathbf{E}[e^{-\alpha X_t}] = \exp\Big(-\int_0^\infty\int_0^\infty (1-e^{-\alpha f(s,y)})\lambda ds\nu(dy)\Big) = \exp\Big(-\lambda t\int_0^\infty (1-e^{\alpha y})\nu(dy)\Big).$$

Un petit calcul montre enfin que

$$\mathbf{E}[X_t] = -\frac{d}{d\alpha}\mathbf{E}[e^{-\alpha X_t}]|_{\alpha=0} = \lambda t\int_0^\infty y\nu(dy).$$

C'est logique : durant $[0,t]$, on attrape en moyenne λt poissons, et la masse moyenne des poissons vaut $\int_0^\infty y\nu(dy)$.

Etudions ce qu'on appelle parfois les processus de Poisson effacés.

Corollaire 2.3.11 *Soit $T_1 < T_2 < \cdots$, les points d'un processus de Poisson sur $\mathbf{R}^+$ d'intensité λ. Soit $(Y_n)_{n\geq 1}$ une famille i.i.d. de v.a. de Bernoulli de paramètre p (indépendante du processus de Poisson). Alors*

$$N_t^0 = \sum_{n\geq 1}(1-Y_n)\mathbf{1}_{\{T_n\leq t\}} \quad \textit{et} \quad N_t^1 = \sum_{n\geq 1} Y_n\mathbf{1}_{\{T_n\leq t\}}$$

sont les processus de comptage de 2 processus de Poisson indépendants de paramètres $\lambda(1-p)$ et λp.

Preuve: On sait que $\{(T_n, Y_n), n \geq 1\}$ est un P.P.P. sur $\mathbf{R}^+ \times \{0,1\}$ d'intensité $\xi = \lambda m_+ \otimes ((1-p)\delta_0 + p\delta_1)$. Soit N sa fonction de comptage. On a alors, pour tout $0 \leq s < t$ (noter que $1 - Y_n = \mathbf{1}_{\{Y_n=0\}}$ et $Y_n = \mathbf{1}_{\{Y_n=1\}}$)

$$N_t^0 - N_s^0 = N(]s,t] \times \{0\}) \quad et \quad N_t^1 - N_s^2 = N(]s,t] \times \{1\}).$$

L'indépendance de N^0 et N^1 s'ensuit facilement : pour tout $0 < s < t$ et $0 < u < v$, les ensembles $]s,t] \times \{0\}$ et $]u,v] \times \{1\}$ sont disjoints. Pour montrer que N^1 (par exemple) est un processus de Poisson de paramètre λp, il suffit de montrer que pour tout $0 = t_0 < t_1 < \cdots < t_k$, les v.a. $N_{t_i}^1 - N_{t_{i-1}}^1, i = 1, \ldots, k$ sont indépendantes de de lois de Poisson de paramètre $\lambda p(t_i - t_{i-1})$. L'indépendance découle du fait que les ensembles $]t_{i-1}, t_i] \times \{1\}$ sont disjoints. De plus, on sait que $N_{t_i}^1 - N_{t_{i-1}}^1 = N(]t_{i-1}, t_i] \times \{1\})$ suit une loi de Poisson de paramètre $\xi(]t_{i-1}, t_i] \times \{1\})$, qui vaut bien $\lambda p(t_i - t_{i-1})$.

Exemple. Un pêcheur attrape des poissons suivant un processus de Poisson $0 < T_1 < T_2 < \cdots$ de paramètre $\lambda > 0$. Chaque poisson est rouge avec probabilité p et gris avec probabilité $1-p$. Si N_t^G (resp. N_t^R) est le nombre de poissons gris (resp. rouges) pêchés durant $[0,t]$, les processus N^G et N^R sont deux processus de Poisson indépendants, de paramètres $\lambda(1-p)$ et λp.

2.4 La file d'attente $M/G/\infty$

Des clients arrivent aux instants $0 < T_1 < T_2 < \cdots$ formant un processus de Poisson de paramètre λ. Le temps de service du client n est une variable aléatoire Y_n de loi η (sur $\mathbf{R}^+$). On suppose que la famille $(Y_n)_{n\geq 1}$ est i.i.d. et indépendante de $(T_n)_{n\geq 1}$. On considère le cas où il n'a pas d'attente: tout client est immédiatement servi. Ceci revient à supposer qu'il y a une infinité de serveurs.

On emploie la notation $M/G/\infty$ pour cette situation: le M indique que les arrivées sont poissoniennes (M est pour Markov, nous verons plus tard pourquoi). Le G indique que la loi des services est arbitraire (G pour Général). Enfin ∞ indique qu'il y a une infinité de serveurs.

Théorème 2.4.1 *Le nombre X_t de clients présents dans le système à l'instant t suit une loi de Poisson de paramètre $\lambda \int_0^\infty (y \wedge t)\, \eta(dy)$.*

Preuve: On introduit le processus $\{(T_n, Y_n), n \geq 1\}$. C'est un processus ponctuel de Poisson sur $\mathbf{R}^+ \times \mathbf{R}^+$ d'intensité $\lambda m_+ \otimes \eta$, où m_+ est la mesure de Lebesgue sur $\mathbf{R}^+$. Remarquons ensuite que

$$X_t = \sum_{n=1}^{+\infty} \mathbf{1}_{\{T_n \leq t < T_n + Y_n\}} = \sum_{n=1}^{+\infty} \mathbf{1}_{\{(T_n, Y_n) \in I_t\}},$$

où $I_t = \{(x,y); 0 \leq x \leq t \leq x + y\}$. Autrement dit, en notant N la fonction de comptage du P.P.P., on a $X_t = N_{I_t}$. Par définition d'un processus ponctuel de Poisson,

X_t a une loi de Poisson de paramètre $(\lambda m_+ \otimes \eta)(I_t)$. Or

$$\begin{aligned}
(\lambda m_+ \otimes \eta)(I_t) &= \lambda \int_0^\infty \int_0^\infty \mathbf{1}_{I_t}(x,y)\, dx\, \eta(dy) \\
&= \lambda \int_0^\infty \Big(\int_0^\infty \mathbf{1}_{0<x<t<x+y}\, dx \Big) \eta(dy) \\
&= \lambda \int_0^\infty (y \wedge t)\, \eta(dy).
\end{aligned}$$

On déduit facilement de ce théorème que :

(i) si le temps de service moyen $\int_0^\infty y\eta(dy)$ est fini, alors X_t converge en loi, quand $t \to \infty$, vers la loi de Poisson de paramètre $\lambda \int_0^\infty y\eta(dy)$. Le système est donc stable (n'explose pas).

(ii) si le temps de service moyen $\int_0^\infty y\eta(dy)$ est infini, alors X_t tend vers l'infini (en probabilité) quand $t \to \infty$.

Chapitre 3

Processus régénératifs

On trouve dans la littérature de nombreuses définitions des processus régénératifs. Nous avons choisi celle qui est la plus adaptée aux théorèmes ergodiques. Dans la suite $\mathbf{T}$ désigne soit l'ensemble $\mathbf{N}$ des entiers soit l'ensemble $\mathbf{R}^+$. Nous supposerons toujours que les processus $\{X_t, t \in \mathbf{R}^+\}$ à valeurs dans un espace topologique E, sont continus à droite.

Définition 3.1.1 *On dit qu'un processus $(X_t)_{t\in\mathbf{T}}$, est régénératif si il existe une suite croissante $\{\tau_n, n \in \mathbf{N}\}$ de variables aléatoires à valeurs dans $\mathbf{T}$ telle que, pour toute fonction mesurable $f : E \to \mathbf{R}^+$, les variables aléatoires*

$$Z_n = \int_{\tau_n}^{\tau_{n+1}} f(X_s)\, ds \ \text{si}\ \mathbf{T} = \mathbf{R}^+, \qquad Z_n = \sum_{k=\tau_n}^{\tau_{n+1}-1} f(X_k)\ \text{si}\ \mathbf{T} = \mathbf{N},$$

sont i.i.d.

Très souvent $\tau_0 = 0$, mais ce n'est pas une nécéssité. En prenant pour f la fonction identiquement égale à 1, on voit que les variables aléatoires $\{\tau_{n+1} - \tau_n, n \in \mathbf{N}\}$, sont positives, indépendantes et de même loi. On voit donc que τ_n est une marche aléatoire sur $\mathbf{R}$ à valeurs positives. On appelle parfois un tel processus un processus de renouvellement.

Exemple. Considérons une file d'attente où les délais d'arrivée $X_1, X_2, \ldots$ entre les clients sont i.i.d. (de loi commune μ sur $\mathbf{R}^+$). Autrement dit, les temps d'arrivée des clients sont $T_1 = X_1$, $T_2 = X_1+X_2$, $T_3 = X_1+X_2+X_3$, etc. Soit $N_t = \sum_{n\geq 1} \mathbf{1}_{\{T_n\leq t\}}$ le nombre de clients arrivés avant l'instant $t \geq 0$. Alors les processus $X_t = t-T_{N_t}$ (temps écoulé, à l'instant t, depuis la dernière arrivée) et $Y_t = T_{N_t+1} - t$ (temps résiduel d'attente, à l'instant t, avant la prochaine arrivée), sont des processus régénératifs, et on peut choisir $\tau_n = T_n$.

En effet, considérons f mesurable positive sur $\mathbf{R}^+$, alors les v.a.

$$Z_n = \int_{T_n}^{T_{n+1}} f(X_s)ds = \int_{T_n}^{T_{n+1}} f(s - T_n)ds = \int_0^{T_{n+1}-T_n} f(u)du = \int_0^{X_{n+1}} f(u)du$$

sont bien entendu i.i.d., donc $(X_t)_{t\geq 0}$ est régénératif. On a utilisé le changement de variables $u = s - T_n$.

De même, les v.a.

$$Z_n = \int_{T_n}^{T_{n+1}} f(Y_s)ds = \int_{T_n}^{T_{n+1}} f(T_{n+1} - s)ds = \int_0^{T_{n+1}-T_n} f(u)du = \int_0^{X_{n+1}} f(u)du$$

sont bien entendu i.i.d., donc $(Y_t)_{t\geq 0}$ est régénératif. On a utilisé le changement de variables $u = T_{n+1} - s$.

Nous utiliserons la version suivante de la loi des grands nombres.

Théorème 3.1.2 *Soit $(Y_k)_{k\geq 1}$ des v.a. i.i.d. à valeurs positives. Presque sûrement,*

$$\lim_{n\to+\infty} \frac{Y_1 + Y_2 + \cdots + Y_n}{n} = \mathbf{E}(Y_1) \leq \infty.$$

Preuve: Lorsque $\mathbf{E}(Y_1) < +\infty$, c'est la loi des grands nombres classique. Si $\mathbf{E}(Y_1) = +\infty$, on remarque que pour tout $a > 0$, puisque $\min(Y_1, a)$ est d'espérance finie,

$$\lim_{n\to+\infty} \frac{\sum_{k=1}^n Y_k}{n} \geq \lim_{n\to+\infty} \frac{\sum_{k=1}^n \min(Y_k, a)}{n} = \mathbf{E}(\min(Y_1, a))$$

Ceci étant vrai pour tout a, par application du théorème de convergence monotone,

$$\lim_{n\to+\infty} \frac{\sum_{k=1}^n Y_k}{n} \geq \lim_{a\to+\infty} \mathbf{E}(\min(Y_1, a)) = \infty.$$

Théorème 3.1.3 Théorème ergodique des processus régénératifs, $\mathbf{T} = \mathbf{R}^+$. *Considérons un processus régénératif $(X_t)_{t\in\mathbf{R}^+}$ à valeurs dans E. Pour toutes fonctions mesurables $f, g : E \to \mathbf{R}^+$, presque sûrement,*

$$\lim_{t\to+\infty} \frac{\int_0^t f(X_s)\, ds}{\int_0^t g(X_s)\, ds} = \frac{\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s)\, ds]}{\mathbf{E}[\int_{\tau_0}^{\tau_1} g(X_s)\, ds]},$$

si $0 < \mathbf{E}[\int_{\tau_0}^{\tau_1} g(X_s)\, ds] < +\infty$. En particulier, si $0 < \mathbf{E}[\tau_1 - \tau_0] < +\infty$,

$$\lim_{t\to+\infty} \frac{1}{t} \int_0^t f(X_s)\, ds = \int_E f(x)\, \pi(dx),$$

où π est la probabilité sur E définie par (pour $A \in \mathcal{E}$)

$$\pi(A) = \frac{1}{\mathbf{E}[\tau_1 - \tau_0]} \mathbf{E}\Big[\int_{\tau_0}^{\tau_1} \mathbf{1}_{\{X_s\in A\}}\, ds\Big].$$

Preuve:

Etape 1. Posons $N_t = \sum_{k=1}^{+\infty} \mathbf{1}_{\{\tau_k\leq t\}}$. Alors N_t tend p.s. vers l'infini quand $t \to \infty$ (car N_t est p.s. croissant et car $N_{\tau_k} = k$). De plus, pour tout $t \geq \tau_0$, par définition de N_t,

$$\tau_{N_t} \leq t \leq \tau_{N_t+1}.$$

On a donc, puisque f est à valeurs positives,

$$\int_0^{\tau_{N_t}} f(X_s)\,ds \le \int_0^t f(X_s)\,ds \le \int_0^{\tau_{N_t+1}} f(X_s)\,ds.$$

Etape 2. On écrit

$$\frac{1}{N_t}\int_0^{\tau_{N_t}} f(X_s)\,ds = \frac{1}{N_t}\int_0^{\tau_0} f(X_s)\,ds + \frac{1}{N_t}\sum_{k=0}^{N_t-1}\int_{\tau_k}^{\tau_{k+1}} f(X_s)\,ds.$$

Par hypothèse, les v.a. $Z_k = \int_{\tau_{k-1}}^{\tau_k} f(X_s)\,ds$ sont i.i.d. (et positives), donc par la LGN, $\frac{1}{n}\sum_{k=0}^{n-1} Z_k$ converge p.s. vers $\mathbf{E}(Z_1) = \mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s)\,ds]$. Quand $t \to +\infty$, puisque $N_t \to +\infty$ p.s., on en déduit que $\frac{1}{N_t}\sum_{k=0}^{N_t-1} Z_k$ converge p.s. vers la même limite $\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s)\,ds]$. Comme enfin $\frac{1}{N_t}\int_0^{\tau_0} f(X_s)\,ds$ tend vers 0 p.s. (puisque $N_t \to \infty$), on conclut que p.s.,

$$\lim_{t\to\infty}\frac{1}{N_t}\int_0^{\tau_{N_t}} f(X_s)\,ds = \mathbf{E}\Big[\int_{\tau_0}^{\tau_1} f(X_s)\,ds\Big].$$

Etape 3. On montre de même (utiliser aussi que $(N_t+1)/N_t \to 1$) que p.s., quand $t \to \infty$

$$\lim_{t\to\infty}\frac{1}{N_t}\int_0^{\tau_{N_t+1}} f(X_s)\,ds = \lim_{t\to\infty}\frac{N_t+1}{N_t}\frac{1}{N_t+1}\int_0^{\tau_{N_t+1}} f(X_s)\,ds = \mathbf{E}\Big[\int_{\tau_0}^{\tau_1} f(X_s)\,ds\Big].$$

Etape 4. Les étapes 1,2,3 montrent que pour f positive, on a p.s.

$$\lim_{t\to\infty}\frac{1}{N_t}\int_0^t f(X_s)ds = \mathbf{E}\Big[\int_{\tau_0}^{\tau_1} f(X_s)\,ds\Big].$$

Supposons maintenant qu'on a 2 fonctions positives f, g et que $0 < \mathbf{E}[\int_{\tau_0}^{\tau_1} g(X_s)\,ds] < \infty$. On écrit alors

$$\lim_{t\to\infty}\frac{\int_0^t f(X_s)ds}{\int_0^t g(X_s)ds} = \lim_{t\to\infty}\frac{N_t^{-1}\int_0^t f(X_s)ds}{N_t^{-1}\int_0^t g(X_s)ds} = \frac{\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s)\,ds]}{\mathbf{E}[\int_{\tau_0}^{\tau_1} g(X_s)\,ds]}.$$

Finalement, si on peut appliquer ce résultat avec $g = 1$, i.e. si $0 < \mathbf{E}[\tau_1 - \tau_0] < \infty$, alors on trouve que pour toute f mesurable positive,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f(X_s)ds = \frac{\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s)\,ds]}{\mathbf{E}[\tau_1 - \tau_0]}.$$

Le membre de droite n'est autre que $\int_E f\,d\pi$.

Exemple : processus de renouvellement alterné. Modélisons le comportement d'une machine qui est successivement en état de marche et en panne comme suit. La machine tombe en panne, après la n-ième réparation, au bout d'un temps X_{n+1}. Se fait réparer, après la n-ième panne, pendant une durée Y_n. Autrement dit, la machine marche du temps $T_0 = 0$ au temps $T_1 = X_1$, elle est en réparation du temps T_1 au temps $T_2 = X_1 + Y_1$, puis elle marche du temps T_2 au temps $T_3 = T_2 + X_2$, puis elle est en réparation du temps T_3 au temps $T_4 = T_3 + Y_2$, etc.

Soit $A_t = 1$ si la machine marche à l'instant t, $A_t = 0$ sinon. On suppose que les v.a. X_i sont i.i.d. (positives), indépendantes des Y_i (positives) aussi i.i.d. On suppose aussi que $0 < E[X_1] < \infty$ et que $0 < \mathbf{E}[Y_1] < \infty$.

Le processus $(A_t)_{t\geq 0}$ est régénératif et on peut prendre $\tau_n = T_{2n}$: en effet, pour f mesurable positive, les v.a.

$$\begin{aligned} Z_n &= \int_{T_{2n}}^{T_{2n+2}} f(A_s)ds \\ &= \int_{T_{2n}}^{T_{2n}+X_{n+1}} f(1)ds + \int_{T_{2n}+X_{n+1}}^{T_{2n}+X_{n+1}+Y_{n+1}} f(0)ds \\ &= f(1)X_{n+1} + f(0)Y_{n+1} \end{aligned}$$

sont bien sûr i.i.d. Donc on déduit du théorème que p.s.,

$$\lim_{t\to\infty} \frac{A_t}{t} = \frac{\mathbf{E}[\int_{T_0}^{T_2} A_s ds]}{\mathbf{E}[T_2 - T_0]} = \frac{\mathbf{E}[X_1]}{\mathbf{E}[X_1 + Y_1]}.$$

Autrement dit, la proportion du temps (sur un intervalle de temps infini) pendant laquelle la machine marche vaut $\mathbf{E}[X_1]/\mathbf{E}[X_1 + Y_1]$, ce qui est intuitivement clair.

On montre de la même façon:

Théorème 3.1.4 Théorème ergodique des processus régénératifs, $T = \mathbf{N}$. *Considérons un processus régénératif* $(X_n)_{n\in\mathbf{N}}$ *à valeurs dans* E. *Pour toutes fonctions mesurables* $f, g : E \to \mathbf{R}^+$, *presque sûrement,*

$$\lim_{n\to+\infty} \frac{\sum_{k=0}^{n} f(X_k)}{\sum_{k=0}^{n} g(X_k)} = \frac{\mathbf{E}[\sum_{k=\tau_0}^{\tau_1-1} f(X_k)]}{\mathbf{E}[\sum_{k=\tau_0}^{\tau_1-1} g(X_k)]},$$

si $0 < \mathbf{E}[\sum_{k=\tau_0}^{\tau_1-1} g(X_k)] < +\infty$. *Lorsque* $0 < \mathbf{E}[\tau_1 - \tau_0] < +\infty$,

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n} f(X_k) = \int_E f(x)\, \pi(dx),$$

où π *est la probabilité sur* E *définie par (pour* $A \in \mathcal{E}$*)*

$$\pi(A) = \frac{1}{\mathbf{E}[\tau_1 - \tau_0]} \mathbf{E}\Big[\sum_{k=\tau_0}^{\tau_1-1} \mathbf{1}_{\{X_k\in A\}}\Big].$$

Chapitre 4

Chaînes de Markov

Ce chapitre est une introduction générale aux chaînes de Markov. Dans tout le chapitre, E est un espace fini ou dénombrable.

4.1 Matrices de transition

Définition 4.1.1 *On appelle matrice (ou noyau, ou probabilité) de transition sur E, une famille $\{P(i,j), i,j \in E\}$ de réels telle que*

i) $P(i,j) \geq 0$, pour tout $i,j \in E$.

ii) Pour tout $i \in E$, $\sum_{j \in E} P(i,j) = 1$.

Quelques propriétés/notations. Si P, Q sont deux matrices de transitions sur E, si f est une fonction de E dans $\mathbf{R}$ (on considère alors $f = (f(i))_{i \in E}$ comme un vecteur colonne) et si μ est une probabilité sur E (on considère alors $\mu = (\mu(i))_{i \in E}$ comme un vecteur ligne),

• PQ est la matrice de transition définie par $PQ(i,j) = \sum_{k \in E} P(i,k)Q(k,j)$ (vérifier que c'est une matrice de transition).

• P^n est la matrice de transition définie par $P^0(i,j) = \mathbf{1}_{\{i=j\}}$ et $P^{n+1} = P^n P$.

• Pf est la fonction de E dans $\mathbf{R}$ définie par $Pf(i) = \sum_{j \in E} P(i,j)f(j)$.

• μP est la probabilité sur E définie par $\mu P(j) = \sum_{i \in E} \mu(i)P(i,j)$ (vérifier que c'est une probabilité).

• μf est le nombre $\sum_{i \in E} \mu(i)f(i)$. C'est $\mathbf{E}[f(X)]$, si X est une v.a. à valeurs dans E de loi μ.

Bref, tous les produits *matriciels* usuels sont utilisables dans ce contexte, en identifiant les fonctions à des vecteurs colonnes et les mesures à des vecteurs lignes.

Vérifier que $\mu(Pf) = (\mu P)f$.

4.2 Chaîne de Markov

Définition 4.2.1 *Soit P une matrice de transition sur E. Un processus $(X_n)_{n\in\mathbf{N}}$ à valeurs dans E est une chaîne de Markov de transition P si pour tout $n \geq 0$, tous $i_0, \cdots, i_n \in E$, tout $j \in E$,*

$$\mathbf{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = P(i_n, j).$$

La loi de X_0 s'appelle la loi initiale de la chaîne.

Exercice. On a alors $\mathbf{P}(X_{n+1} = j | X_n = i) = P(i,j)$ pour tout $n \geq 0$, tout $i, j \in E$.

Une chaîne de Markov vérifie donc deux propriétés importantes :

(i) X_{n+1} ne dépend de tout le passé jusqu'à l'instant n qu'à travers la valeur de X_n (propriété de Markov),

(ii) la loi de X_{n+1} sachant X_n ne dépend pas du temps n (homogénéité).

Pour montrer qu'un processus est une chaîne de Markov, on utilise très souvent le lemme (très facile) suivant.

Lemme 4.2.2 *Soit E un ensemble dénombrable, $(\varepsilon_n)_{n\geq 1}$ une suite de v.a. i.i.d. à valeurs dans un ensemble F. Soit $\phi : E \times F \to E$ une application mesurable. Pour tout X_0 indépendant de $(\varepsilon_n)_{n\geq 1}$, la suite définie par récurrence par*

$$X_{n+1} = \phi(X_n, \varepsilon_{n+1})$$

est une chaîne de Markov à valeurs dans E de transition $P(i,j) = \mathbf{P}(\phi(i, \varepsilon_1) = j)$.

Preuve: Ceci résulte du fait que

$$\begin{aligned}
\mathbf{P}(X_{n+1} = j | X_0 = i_0, \cdots, X_n = i_n) &= \mathbf{P}(\phi(X_n, \varepsilon_{n+1}) = j | X_0 = i_0, \cdots, X_n = i_n) \\
&= \mathbf{P}(\phi(i_n, \varepsilon_{n+1}) = j | X_0 = i_0, \cdots, X_n = i_n) \\
&= \mathbf{P}(\phi(i_n, \varepsilon_{n+1}) = j) \\
&= \mathbf{P}(\phi(i_n, \varepsilon_1) = j) = P(i_n, j).
\end{aligned}$$

On a utilisé que ε_{n+1} est indépendant de $X_0, \cdots, X_n$, puis que ε_{n+1} a la même loi que ε_1.

Exemples. (i) Rappelons qu'une marche aléatoire de pas de loi μ à valeurs dans $\mathbf{Z}^d$ est définie par $X_0 = 0$ et, pour $n \geq 1$, $X_n = Y_1 + \cdots + Y_n$, où les Y_k sont i.i.d. de loi μ. En appliquant le lemme avec $E = F = \mathbf{Z}^d$, avec $\varepsilon_n = Y_n$, et avec $\phi(x,y) = x + y$ (on a bien $X_{n+1} = X_n + Y_{n+1} = \phi(X_n, Y_{n+1})$, on trouve que $(X_n)_{n\geq 0}$ est une chaîne de Markov à valeurs dans $\mathbf{Z}^d$ de transition $P(i,j) = \mathbf{P}(\phi(i, Y_1) = j) = \mu(\{j - i\})$ (c'est logique : pour passer de i à j, il faut sauter de $j - i$).

(ii) Considérons une famille de v.a. i.i.d. $(Y_n)_{n\geq 1}$ de loi $(\delta_{-1} + \delta_1)/2$ et posons $X_0 = 0$ puis $X_n = Y_1 + \cdots + Y_n$ pour $n \geq 1$. Alors $(X_n)_{n\geq 0}$ est la marche aléatoire simple symétrique sur $\mathbf{Z}$, c'est donc une chaîne de Markov. Par contre, $S_n = \sup_{k=0,\dots,n} X_k$ n'est pas une chaîne de Markov : on peut se convaincre, par exemple, que

$$\mathbf{P}(S_4 = 2 | S_0 = 0, S_1 = 1, S_2 = 1, S_3 = 1) \neq \mathbf{P}(S_4 = 2 | S_0 = 0, S_1 = 0, S_2 = 0, S_3 = 1)$$

(la probabilité de gauche vaut 1/4, celle de droite 1/2).

Notation importante On considèrera souvent plusieurs lois initiales pour la même matrice de transition. Il nous faudra donc souvent préciser avec quelle loi initiale on travaille. On notera en indice cette loi initiale dans les espérances et les probabilités, $\mathbf{E}_\nu$, $\mathbf{P}_\nu$. Par exemple, si on travaille avec une chaîne de Markov $(X_n)_{n\geq 0}$ de transition P, on note $\mathbf{E}_\nu(X_k)$ (resp. $\mathbf{P}_\nu(X_k \in A)$) l'espérance de X_k (resp. la probabilité que $X_k \in A$) si $(X_n)_{n\geq 0}$ est la chaîne de transition P et de loi initiale ν.

Dans le cas où $\nu = \delta_x$ pour un $x \in E$ (i.e. quand $X_0 = x$ p.s.), on simplifie la notation et on écrit $\mathbf{P}_x, \mathbf{E}_x$ pour $\mathbf{P}_{\delta_x}, \mathbf{E}_{\delta_x}$.

Lemme 4.2.3 *Soit $(X_n)_{n\in\mathbf{N}}$ une chaîne de transition P. Pour toute probabilité ν sur E,*

$$\mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) = \nu(x_0)P(x_0, x_1)\cdots P(x_{n-1}, x_n).$$

En particulier (avec $\nu = \delta_x$),

$$\mathbf{P}_x(X_1 = x_1, \cdots, X_n = x_n) = P(x, x_1)\cdots P(x_{n-1}, x_n).$$

Preuve: Par récurrence sur n : si $n = 0$, c'est évident, puisque $\mathbf{P}_\nu(X_0 = x_0) = \nu(x_0)$ par définition de $\mathbf{P}_\nu$. Si la formule est vraie avec n, alors on écrit

$$\begin{aligned}
& \mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n, X_{n+1} = x_{n+1}) \\
= \; & \mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) \\
& \qquad \times \mathbf{P}_\nu(X_{n+1} = x_{n+1} | X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) \\
= \; & \nu(x_0)P(x_0, x_1)\cdots P(x_{n-1}, x_n) \times P(x_n, x_{n+1}).
\end{aligned}$$

Corollaire 4.2.4 *La loi d'une chaîne de Markov est entièrement caractérisée par sa matrice de transition et sa loi initiale.*

Proposition 4.2.5 *Soit $(X_n)_{n\in\mathbf{N}}$ une chaîne de Markov de transition P et $n \geq 0$ fixé.*

(i) La loi de X_n (si $X_0 \sim \nu$) est νP^n, i.e. $\mathbf{P}_\nu(X_n = y) = \nu P^n(y)$ pour tout $y \in E$. En particulier, $\mathbf{P}_x(X_n = y) = P^n(x, y)$.

(ii) Pour $f : E \to \mathbf{R}^+$, on a $\mathbf{E}_\nu(f(X_n)) = \nu P^n f$. En particulier, $\mathbf{E}_x(f(X_n)) = P^n f(x)$.

On peut donc dire que $P^n(x, y)$ est la probabilité que la chaîne passe de x à y en n coups.

Preuve: (ii) découle de (i). Pour (i), fixons $y \in E$, et écrivons

$$\begin{aligned}
\mathbf{P}_\nu(X_n = y) &= \sum_{x_0,\cdots,x_{n-1}\in E} \mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1}, X_n = y) \\
&= \sum_{x_0,\cdots,x_{n-1}\in E} \nu(x_0, x_1)P(x_1, x_2)\cdots P(x_{n-2}, x_{n-1})P(x_{n-1}, y).
\end{aligned}$$

En utilisant les produits matriciels, ceci donne $\nu P^n(y)$ (rappelons que ν est un vecteur ligne et P une matrice, donc νP^n un vecteur ligne). Avec le choix particulier $\nu = \delta_x$, on trouve

$$\mathbf{P}_x(X_n = y) = \delta_x P^n(y) = \sum_{z\in E} \delta_x(z)P^n(z, y) = P^n(x, y).$$

4.3 Propriété de Markov

Soit $(X_n)_{n\geq 0}$ une chaîne de Markov sur E. On note $\mathcal{F}_n$ la tribu engendrée par $(X_0, X_1, \cdots, X_n)$.

Théorème 4.3.1 (Propriété de Markov simple) *Pour toute fonction mesurable $f : E^{\mathbf{N}} \to \mathbf{R}$, par exemple bornée ou positive, pour toute probabilité ν, presque sûrement,*

$$\mathbf{E}_\nu[f(X_n, X_{n+1}, \cdots)|\mathcal{F}_n] = \mathbf{E}_{X_n}[f(X_0, X_1, \cdots)].$$

Ceci signifie que pour tout $x_0, x_1, \cdots, x_n \in E$, on a

$$\mathbf{E}_\nu[f(X_n, X_{n+1}, \cdots)|X_0 = x_0, \cdots, X_n = x_n] = \mathbf{E}_{x_n}[f(X_0, X_1, \cdots)].$$

Preuve: Il suffit de vérifier ce résultat lorsque f est une fonction indicatrice $f = \mathbf{1}_A$ où A est de la forme

$$A = \{a_0\} \times \{a_1\} \times \cdots \times \{a_d\} \times E \times E \times \cdots.$$

On a alors $f(X_n, X_{n+1}, \cdots) = \mathbf{1}_{\{X_n=a_0, X_{n+1}=a_1,\cdots,X_{n+d}=a_d\}}$. L'objectif est donc de montrer que pour tout $a_0, \cdots, a_d, x_0, \cdots, x_n \in E$, on a $I = J$, où

$$I = \mathbf{P}_\nu(X_n = a_0, X_{n+1} = a_1, \cdots, X_{n+d} = a_d|X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n)$$

et

$$J = \mathbf{P}_{x_n}(X_0 = a_0, X_1 = a_1, \cdots, X_d = a_d).$$

Si $a_0 \neq x_n$, c'est évident, puisque $I = 0 = J$. Si maintenant $a_0 = x_n$, on a d'une part

$$J = P(a_0, a_1)\cdots P(a_{d-1}, a_d)$$

et d'autre part

$$\begin{aligned}
I &= \frac{\mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n, X_{n+1} = a_1, \cdots, X_{n+d} = a_d)}{\mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n)} \\
&= \frac{\nu(x_0)P(x_0, x_1)\cdots P(x_{n-1}, x_n)P(x_n, a_1)\cdots P(a_{d-1}, a_d)}{\nu(x_0)P(x_0, x_1)\cdots P(x_{n-1}, x_n)} \\
&= P(x_n, a_1)\cdots P(a_{d-1}, a_d).
\end{aligned}$$

Comme $x_n = a_0$, on a bien $I = J$.

Rappelons qu'un temps d'arrêt du processus $(X_n)_{n\geq 0}$ est une v.a. à valeurs dans $\mathbf{N} \cup \{\infty\}$ telle que pour tout $n \in \mathbf{N}$, l'événement $\{T = n\}$ est dans $\mathcal{F}_n = \sigma(X_0, \cdots, X_n)$. La tribu $\mathcal{F}_T$ est formée des événements $A \in \mathcal{F}_\infty$ tels que pour tout $n \in \mathbf{N}$, $A \cap \{T = n\} \in \mathcal{F}_n$. En particulier $T, \mathbf{1}_{\{T<\infty\}}X_T$ et $X_{T\wedge n}$ sont des variables aléatoires $\mathcal{F}_T$-mesurables.

Lemme 4.3.2 *Soit T un temps d'arrêt. Pour toute variable aléatoire Z bornée (ou positive), pour tout $n \geq 0$,*

$$\mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_T) = \mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_n).$$

Preuve: Déjà, $\{T = n\} \in \mathcal{F}_T$. En effet, il faut vérifier que pour tout $k \geq 0$, on a $\{T = n\} \cap \{T = k\} \in \mathcal{F}_k$. C'est vrai si $k \neq n$ (car alors $\{T = n\} \cap \{T = k\} = \emptyset \in \mathcal{F}_k$) et si $k = n$ (car alors $\{T = n\} \cap \{T = k\} = \{T = k\} \in \mathcal{F}_k$ puisque T est un temps d'arrêt).

Du coup, $\mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_T) = \mathbf{E}(\mathbf{1}_{\{T=n\}}Z|\mathcal{F}_T)$. Il faut donc montrer que

$$\mathbf{E}(\mathbf{1}_{\{T=n\}}Z|\mathcal{F}_T) = \mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_n),$$

ce qui signifie que (i) $\mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_n)$ est $\mathcal{F}_T$-mesurable et (ii) pour tout $A \in \mathcal{F}_T$, $\mathbf{E}[\mathbf{1}_A\mathbf{1}_{\{T=n\}}Z] = \mathbf{E}[\mathbf{1}_A\mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_n)]$.

Pour (i), il faut vérifier que pour tout $B \in \mathcal{B}(\mathbf{R})$, $\{\mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_n) \in B\} \in \mathcal{F}_T$, soit encore que pour tout $k \geq 0$, $\{\mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_n) \in B\} \cap \{T = k\} \in \mathcal{F}_k$: c'est juste si $k = n$, puisque $\{T = n\} \in \mathcal{F}_n$ et puisque $\mathbf{E}(Z|\mathcal{F}_n)$ est $\mathcal{F}_n$-mesurable ; c'est aussi juste si $k \neq n$, puisqu'alors $\{\mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_n) \in B\} \cap \{T = k\} = \{0 \in B\} \cap \{T = k\} \in \mathcal{F}_k$ (noter que $\{0 \in B\}$ vaut soit $\emptyset$ soit Ω).

Pour montrer (ii), on utilise que $\mathbf{1}_A\mathbf{1}_{\{T=n\}}$ est $\mathcal{F}_n$ mesurable (car $A\cap\{T = n\} \in \mathcal{F}_n$ car $A \in \mathcal{F}_T$ par hypothèse), d'où

$$\mathbf{E}[\mathbf{1}_A\mathbf{1}_{\{T=n\}}\mathbf{E}(Z|\mathcal{F}_n)] = \mathbf{E}[\mathbf{E}[\mathbf{1}_A\mathbf{1}_{\{T=n\}}Z|\mathcal{F}_n]] = \mathbf{E}[\mathbf{1}_A\mathbf{1}_{\{T=n\}}Z].$$

On déduit Markov forte des deux résultats précédents.

Théorème 4.3.3 (Propriété de Markov forte) *Soit T un temps d'arrêt de la chaîne de Markov $(X_n)_{n\geq 0}$. Pour toute fonction mesurable $f : E^{\mathbf{N}} \to \mathbf{R}$ bornée ou positive, pour toute probabilité ν sur E, p.s.*

$$\mathbf{1}_{\{T<\infty\}}\mathbf{E}_\nu[f(X_T, X_{T+1}, \cdots)|\mathcal{F}_T] = \mathbf{1}_{\{T<\infty\}}\mathbf{E}_{X_T}[f(X_0, X_1, \cdots)].$$

Preuve: Par le lemme, on a, pour tout $n \geq 0$,

$$\begin{aligned}
\mathbf{1}_{\{T=n\}}\mathbf{E}_\nu[f(X_T, X_{T+1}, \cdots)|\mathcal{F}_T] &= \mathbf{1}_{\{T=n\}}\mathbf{E}_\nu[f(X_T, X_{T+1}, \cdots)|\mathcal{F}_n] \\
&= \mathbf{1}_{\{T=n\}}\mathbf{E}_\nu[f(X_n, X_{n+1}, \cdots)|\mathcal{F}_n] \\
&= \mathbf{1}_{\{T=n\}}\mathbf{E}_{X_n}[f(X_0, X_1, \cdots)] \\
&= \mathbf{1}_{\{T=n\}}\mathbf{E}_{X_T}[f(X_0, X_1, \cdots)].
\end{aligned}$$

On a utilisé Markov simple à l'avant-dernière ligne. En sommant sur n et en utilisant que $\sum_{n\geq 0}\mathbf{1}_{\{T=n\}}=\mathbf{1}_{\{T<\infty\}}$, on trouve

$$\mathbf{1}_{\{T<\infty\}}\mathbf{E}_\nu[f(X_T,X_{T+1},\cdots)|\mathcal{F}_T]=\mathbf{1}_{\{T<\infty\}}\mathbf{E}_{X_T}[f(X_0,X_1,\cdots)].$$

Notons la conséquence suivante de la propriété de Markov forte.

Corollaire 4.3.4 *Soit T un temps d'arrêt de la chaîne de Markov $(X_n)_{n\geq 0}$, fini presque sûrement et tel que $X_T=x$ p.s. (pour un certain $x\in E$). Alors, $(X_{T+n})_{n\geq 0}$ est une chaîne de Markov de transition P (la même que celle de $(X_n)_{n\geq 0}$), de loi initiale δ_x, qui est de plus indépendante de $\mathcal{F}_T$.*

Preuve: Il suffit de montrer que, pour tout $A\in\mathcal{F}_T$ et tous $x_1,\cdots,x_k\in E$, on a $I=J$, où

$$I=\mathbf{P}_\nu(A,X_T=x,X_{T+1}=x_1,\cdots,X_{T+k}=x_k)$$

et

$$J=\mathbf{P}_\nu(A)P(x,x_1)\cdots P(x_{k-1},x_k).$$

Mais, comme $A\in\mathcal{F}_T$,

$$\begin{aligned}I&=\mathbf{E}_\nu[\mathbf{1}_A\mathbf{E}[\mathbf{1}_{\{X_T=x,X_{T+1}=x_1,\cdots,X_{T+k}=x_k\}}|\mathcal{F}_T]]\\&=\mathbf{E}_\nu[\mathbf{1}_A\mathbf{E}_{X_T}[\mathbf{1}_{\{X_0=x,X_1=x_1,\cdots,X_k=x_k\}}]]\end{aligned}$$

par Markov forte. Puis, comme $X_T=x$,

$$\begin{aligned}I&=\mathbf{E}_\nu[\mathbf{1}_A\mathbf{E}_x[\mathbf{1}_{\{X_0=x,X_1=x_1,\cdots,X_k=x_k\}}]]\\&=\mathbf{E}_\nu[\mathbf{1}_AP(x,x_1)\cdots P(x_{k-1},x_k)],\end{aligned}$$

qui n'est autre que J.

4.4 Propriétés de récurrence

Nous allons étudier le comportement des trajectoires d'une chaîne de Markov X_n lorsque $n\to+\infty$. On va distinguer deux types de comportement suivant que la chaîne "tend vers l'infini" ou non. Introduisons les notations suivantes : pour $i\in E$,

$$\tau_i=\inf\{k>0;X_k=i\}.$$

Puis $\tau_i^{(0)}=0$ et, pour tout entier $\ell\geq 0$,

$$\tau_i^{(\ell+1)}=\inf\{k>\tau_i^{(\ell)};X_k=i\}.$$

On voit que $\tau^{(1)}=\tau_i$ est le temps d'atteinte de l'état i (en excluant si nécessaire le temps $n=0$), puis $\tau_i^{(\ell)}$ est l'instant de ℓ-ième passage en i (en excluant si nécessaire le temps $n=0$). Ce sont des temps d'arrêt. Pour tout $i,j\in\mathbf{E}$, on introduit

$$G(i,j)=\sum_{n=0}^{+\infty}P^n(i,j)=\sum_{n=0}^{+\infty}\mathbf{E}_i(\mathbf{1}_{\{X_n=j\}})=\mathbf{E}_i\Big[\sum_{n=0}^{+\infty}\mathbf{1}_{\{X_n=j\}}\Big].$$

C'est l'espérance du nombre de passages par j en partant de i.

Définition 4.4.1 *Un état i est récurrent si $\mathbf{P}_i(\tau_i < +\infty) = 1$.*

Théorème 4.4.2 *Les assertions suivantes sont équivalentes:*

1. *i est récurrent.*
2. *$\sum_{n=0}^{+\infty} \mathbf{1}_{\{X_n=j\}} = +\infty$, $\mathbf{P}_i$-p.s.*
3. *$G(i,i) = +\infty$.*

On a donc, pour un état i donné, l'équivalence : 1. partant de i, on est sûr de revenir en i ; 2. partant de i, on est sûr de revenir une infinité de fois en i ; 3. le nombre moyen de passages en i, partant de i, est infini.

L'équivalence entre 1 est 2 est claire intuitivement, et bien sûr 2 implique 3. Par contre, il est assez surprenant que le nombre de passages par i soit infini p.s. dès que son espérance est infinie (c'est ce que dit 3 implique 2).

Preuve:

Etape 1. Soit $N = \sum_{n=0}^{+\infty} \mathbf{1}_{\{X_n=j\}}$ le nombre de passages par i. On va montrer ici que pour tout $n \geq 0$, $\mathbf{P}_i(N > n) = \alpha^n$, où $\alpha = \mathbf{P}_i(\tau_i < +\infty)$.

C'est clair si $n = 0$ (car comme on part de i, on a $N \geq 1$ $\mathbf{P}_i$-p.s.). Il suffit donc de montrer que pour tout $n \geq 0$, $\mathbf{P}_i(N > n+1) = \alpha \mathbf{P}_i(N > n)$.

On écrit donc

$$\begin{aligned}
\mathbf{P}_i(N > n+1) &= \mathbf{P}_i(\tau_i^{(n+1)} < \infty) \\
&= \mathbf{E}_i[\mathbf{P}_i(\tau_i^{(n+1)} < \infty | \mathcal{F}_{\tau_i^{(n)}})] \\
&= \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{P}_i(\tau_i^{(n+1)} - \tau_i^{(n)} < \infty | \mathcal{F}_{\tau_i^{(n)}})\Big]. \qquad (*)
\end{aligned}$$

La dernière égalité vient du fait que $\tau_i^{(n+1)} < \infty$ ssi $\tau_i^{(n)} < \infty$ et $\tau_i^{(n+1)} - \tau_i^{(n)} < \infty$ et du fait que $\tau_i^{(n)}$ est $\mathcal{F}_{\tau_i^{(n)}}$-mesurable.

Considérons l'application $\Phi : E^{\mathbf{N}} \mapsto \mathbf{N} \cup \{\infty\}$ définie par

$$\Phi((x_k)_{k\geq 0}) = \inf\{k \geq 1 \ : \ x_k = i\},$$

et $\Psi : E^{\mathbf{N}} \mapsto \{0,1\}$ par $\Psi((x_k)_{k\geq 0}) = \mathbf{1}_{\{\Phi((x_k)_{k\geq 0}) < \infty\}}$. On observe alors que

$$\tau_i = \Phi((X_k)_{k\geq 0}) \qquad et \qquad \tau_i^{(n+1)} - \tau_i^{(n)} = \Phi((X_{\tau_i^{(n)}+k})_{k\geq 0})$$

donc

$$\mathbf{1}_{\{\tau_i < \infty\}} = \Psi((X_k)_{k\geq 0}) \qquad et \qquad \mathbf{1}_{\{\tau_i^{(n+1)} - \tau_i^{(n)} < \infty\}} = \Psi((X_{\tau_i^{(n)}+k})_{k\geq 0}).$$

Ainsi,

$$\begin{aligned}\mathbf{1}_{\{\tau_i^{(n)}<\infty\}}\mathbf{P}_i\Big(\tau_i^{(n+1)}-\tau_i^{(n)}<\infty|\mathcal{F}_{\tau_i^{(n)}}\Big) &= \mathbf{1}_{\{\tau_i^{(n)}<\infty\}}\mathbf{E}_i\Big[\Psi((X_{\tau_i^{(n)}+k})_{k\geq 0})|\mathcal{F}_{\tau_i^{(n)}}\Big] \\ &= \mathbf{1}_{\{\tau_i^{(n)}<\infty\}}\mathbf{E}_{X_{\tau_i^{(n)}}}\Big[\Psi((X_k)_{k\geq 0})\Big]\end{aligned}$$

par Markov forte, soit encore, comme $X_{\tau_i^{(n)}} = i$ (quand $\tau_i^{(n)} < \infty$),

$$\mathbf{1}_{\{\tau_i^{(n)}<\infty\}}\mathbf{P}_i\Big(\tau_i^{(n+1)}-\tau_i^{(n)}<\infty|\mathcal{F}_{\tau_i^{(n)}}\Big) = \mathbf{1}_{\{\tau_i^{(n)}<\infty\}}\mathbf{E}_i\Big[\mathbf{1}_{\{\tau_i<\infty\}}\Big] = \alpha\mathbf{1}_{\{\tau_i^{(n)}<\infty\}}.$$

En revenant à $(*)$, on trouve donc

$$\mathbf{P}_i(N>n+1) = \alpha\mathbf{P}_i(\tau_i^{(n)}<\infty) = \alpha\mathbf{P}_i(N>n).$$

Etape 2. On conclut.

1 implique 2 : Si i est récurrent, alors $\alpha = 1$ et donc par l'étape 1, on a

$$\mathbf{P}_i(N=\infty) = \lim_{n\to\infty}\mathbf{P}_i(N>n) = 1,$$

ce qui prouve que $N = \infty$ p.s.

2 implique 3 est évident, puisque $G(i,i) = \mathbf{E}_i(N)$.

3 implique 1 : Si $G(i,i) = \mathbf{E}_i(N) = \infty$, alors forcément $\alpha = 1$, donc i est récurrent (rappelons que $\alpha = \mathbf{P}_i(\tau_i < \infty)$). En effet, si $\alpha < 1$, alors

$$\mathbf{E}_i(N) = \sum_{n\geq 0}\mathbf{P}_i(N>n) = \sum_{n\geq 0}\alpha^n < \infty.$$

Théorème 4.4.3 *Soit i un état récurrent (s'il en existe un) de la chaîne de Markov $(X_n)_{n\geq 0}$. Alors, sous $\mathbf{P}_i$, la suite des temps $(\tau_i^{(n)})_{n\geq 0}$ fait de $(X_n)_{n\geq 0}$ un processus régénératif.*

Preuve: Il s'agit de montrer que pour tout $f : E \to \mathbf{R}^+$, les variables aléatoires

$$Z_n = \sum_{k=\tau_i^{(n)}}^{\tau_i^{(n+1)}-1} f(X_k)$$

sont indépendantes et de même loi.

Fixons $n \geq 0$. Comme i est récurrent, on sait que $\tau_i^{(n)} < \infty$ $\mathbf{P}_i$-p.s. et on a bien sûr $X_{\tau_i^{(n)}} = i$. Par le corollaire 4.3.4 (et comme $\tau_i^{(n)} < \infty$ p.s.), on sait que $(\tilde{X}_k)_{k\geq 0}$ définie par $\tilde{X}_k = X_{\tau_i^{(n)}+k}$ est une chaîne de même transition P et de loi initiale δ_i, indépendante de $\mathcal{F}_{\tau_i^{(n)}}$. De plus, on voit facilement que $\tilde{\tau}_i = \inf\{k \geq 1 \ : \ \tilde{X}_k = i\}$ n'est autre que $\tau_i^{(n+1)} - \tau_i^{(n)}$. Ainsi, on peut écrire

$$Z_n = \sum_{k=0}^{\tau_i^{(n+1)}-\tau_i^{(n)}-1} f(X_{\tau_i^{(n)}+k}) = \sum_{k=0}^{\tilde{\tau}_i-1} f(\tilde{X}_k)$$

et il est clair que Z_n a même loi que $Z_1 = \sum_{k=0}^{\tau_i-1} f(X_k)$.

Enfin, comme $Z_1, \cdots, Z_{n-1}$ sont $\mathcal{F}_{\tau_i^{(n)}}$-mesurables, on voit que Z_n est indépendant de $(Z_1, \cdots, Z_{n-1})$.

On déduit donc directement du théorème 3.1.4:

Corollaire 4.4.4 *Soit i un état récurrent (s'il en existe un) de la chaîne de Markov $(X_n)_{n\geq 0}$. Pour toutes fonctions $f, g : E \mapsto \mathbf{R}^+$, $\mathbf{P}_i$-presque sûrement,*

$$\lim_{n\to+\infty} \frac{\sum_{k=0}^n f(X_k)}{\sum_{k=0}^n g(X_k)} = \frac{\int_E f \ dm_i}{\int_E g \ dm_i}$$

si $0 < \int_E g \ dm_i < \infty$, où m_i est la mesure sur E définie par

$$m_i(j) = \mathbf{E}_i\Big[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=j\}}\Big] \quad \forall\, j \in E.$$

4.5 Mesures et probabilités invariantes

Dans la théorie des chaînes et des processus de Markov, la notion de probabilité invariante est sans doute la plus importante. Elle généralise celle d'état d'équilibre en physique.

Définition 4.5.1 *On dit qu'une mesure m sur E est une mesure invariante de la chaîne de Markov de noyau de transition P si $mP = m$, i.e. si pour tout $j \in E$,*

$$m(j) = \sum_{i\in E} m(i)P(i,j).$$

On remarque que si X_0 a pour loi m, alors X_n a pour loi $mP^n = mP(P^{n-1}) = mP^{n-1} = \cdots = mP = m$. Donc la loi de la chaîne est *invariante*, c'est la même pour tous les temps.

Les mesures introduites au corollaire 4.4.4 vont avoir un rôle important.

Lemme 4.5.2 *Soit i un état récurrent de la chaîne de Markov $(X_n)_{n\geq 0}$ (s'il en existe un). Alors, la formule*

$$m_i(j) = \mathbf{E}_i\Big[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=j\}}\Big] \quad \forall\, j \in E$$

définit une mesure invariante.

Preuve: Sous $\mathbf{P}_i$, $X_0 = X_{\tau_i} = i$, on peut donc écrire (distinguer les cas $j = i$ et $j \neq i$)

$$m_i(j) = \mathbf{E}_i\Big[\sum_{k=0}^{\tau_i - 1} \mathbf{1}_{\{X_{k+1}=j\}}\Big] = \sum_{k\geq 0} \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_i > k\}}\mathbf{1}_{\{X_{k+1}=j\}}\Big].$$

Comme $\{\tau_i > k\} \in \mathcal{F}_k$, par Markov simple,

$$m_i(j) \quad = \quad \sum_{k\geq 0} \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_i>k\}}\mathbf{P}_i(X_{k+1} = j|\mathcal{F}_k)\Big] = \sum_{k\geq 0} \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_i>k\}}\mathbf{P}_{X_k}(X_1 = j)\Big].$$

Mais $P_\ell(X_1 = j) = P(\ell, j)$, donc

$$m_i(j) \quad = \quad \sum_{k\geq 0} \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_i>k\}}P(X_k, j)\Big] = \sum_{\ell\in E} P(\ell, j) \sum_{k\geq 0} \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_i>k\}}\mathbf{1}_{\{X_k=\ell\}}\Big].$$

Mais $\sum_{k\geq 0} \mathbf{E}_i[\mathbf{1}_{\{\tau_i>k\}}\mathbf{1}_{\{X_k=\ell\}}] = \mathbf{E}_i[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=\ell\}}] = m_i(\ell)$. On a donc bien

$$m_i(j) = \sum_{\ell\in E} P(\ell, j)m_i(\ell).$$

Définition 4.5.3 *Une chaîne de Markov de transition P est dite irréductible si pour tout $i, j \in E$, il existe $n \in \mathbf{N}$ tel que $P^n(i, j) > 0$.*

Autrement dit, pour toute paire d'états $i, j \in E$, il est possible d'aller de i à j

Lemme 4.5.4 *Si m est une mesure invariante d'une chaîne irréductible, alors on a $m(j) > 0$ pour tout $j \in E$ (ou $m(j) = 0$ pour tout $j \in E$).*

Preuve: Supposons donc qu'il existe i tel que $m(i) > 0$. Soit ensuite $j \in E$. On sait qu'il existe n tel que $P^n(i, j) > 0$. Du coup,

$$m(j) = (mP^n)(j) = \sum_{k\in E} m(k)P^n(k, j) \geq m(i)P^n(i, j) > 0.$$

Lemme 4.5.5 *Si une chaîne de Markov irréductible possède une probabilité invariante π, toute autre mesure invariante est proportionnelle à π. En particulier, π est la seule probabilité invariante.*

Remarquons que l'hypothèse d'irréductibilité est cruciale : par exemple, supposons que $E = \{1, 2, 3\}$ et que $P(1, 2) = P(2, 1) = P(3, 3) = 1$ (toutes les autres probabilités de transition étant nulles). Cette chaîne n'est pas irréductible (exercice) et on a plusieurs probabilités invariantes, par exemple $m = (1/2\ 1/2\ 0)$ (notation matricielle) et $m = (0\ 0\ 1)$.

Preuve: Soit m une mesure invariante et i un état fixé de E. Posons $\lambda = \frac{\pi(i)}{m(i)}$ et $m' = \lambda m$ et montrons que $m' = \pi$.

Soit $\mu(k) = \min(m'(k), \pi(k))$. Alors μ est invariante, car $\mu P(k) \leq m'P(k) = m'(k)$ et $\mu P(k) \leq \pi P(k) = \pi(k)$, donc $\mu P(k) \leq \mu(k)$ (pour tout $k \in E$). Comme de plus $\sum_{k\in E} \mu P(k) = \sum_{k\in E} \mu(k)$ (car P est une matrice de transition), on conclut que $\mu P(k) = \mu(k)$ pour tout $k \in E$.

Donc $\pi - \mu$ est une mesure (positive), invariante, et on a $(\pi - \mu)(i) = 0$ (par définition, $m'(i) = \pi(i)$ puis $\mu(i) = \pi(i)$) donc, par le lemme $(\pi - \mu)(j) = 0$ pour tout $j \in E$. Autrement dit, $\pi = \mu$.

De même, $m' - \mu$ est une mesure (positive), invariante, et on a $(m' - \mu)(i) = 0$. Donc, par le lemme $(m' - \mu)(j) = 0$ pour tout $j \in E$. Autrement dit, $m' = \mu$.

On a donc bien $m' = \pi$.

Théorème 4.5.6 *Considérons une chaîne de Markov $(X_n)_{n\geq 0}$ irréductible. Les conditions suivantes sont équivalentes.*

(i) Il existe un état $i \in E$ tel que $\mathbf{E}_i(\tau_i) < +\infty$.

(ii) Pour tout état $i \in E$, $\mathbf{E}_i(\tau_i) < +\infty$.

(iii) La chaîne de Markov possède une probabilité invariante π.

Sous ces conditions la chaîne est dite **récurrente positive**. *Sa probabilité invariante π est unique. Pour tout $k \in E$,*

$$\pi(k) = \frac{1}{\mathbf{E}_k(\tau_k)} \qquad (*)$$

et pout tout $i, k \in E$

$$\pi(k) = \frac{1}{\mathbf{E}_i(\tau_i)} \mathbf{E}_i\Big[\sum_{n=0}^{\tau_i - 1} \mathbf{1}_{\{X_n = k\}}\Big]. \qquad (**)$$

Intuitivement, si π est la probabilité invariante d'une chaîne, $\pi(k)$ représente la proportion du temps passé en j, i.e. $\pi(j) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \mathbf{1}_{\{X_k = j\}}$ (on verra cela plus loin). Donc $(*)$ n'est pas déraisonnable : en gros, quand on est en j, on met un temps $\mathbf{E}_j[\tau_j]$ à revenir en j, donc on passe en j "tous les" $\mathbf{E}_j[\tau_j]$, donc la proportion du temps passé en j vaut $1/\mathbf{E}_j[\tau_j]$.

Preuve:

Bien sûr, (ii) implique (i) Montrons que (i) implique (iii). Soit $i \in E$ tel que $\mathbf{E}_i(\tau_i) < +\infty$. Alors i est récurrent. Soit donc m_i la mesure invariante associée à i, i.e. $m_i(k) = \mathbf{E}_i[\sum_{n=0}^{\tau_i - 1} \mathbf{1}_{\{X_n = k\}}]$, voir le lemme 4.5.2. On a $\sum_{k\in E} m_i(k) = \mathbf{E}_i(\tau_i)$. Donc

$$\pi(k) = \frac{1}{\mathbf{E}_i(\tau_i)} \mathbf{E}_i\Big[\sum_{n=0}^{\tau_i - 1} \mathbf{1}_{\{X_n = k\}}\Big]$$

est une probabilité invariante.

Montrons que (iii) implique (ii). Soit donc π une probabilité invariante. Pour tout $i \in E$, par le principe du maximum (cf. lemme 4.5.7),

$$\sum_{j \in E} \pi(j) G(j,i) \leq \sum_{j \in E} \pi(j) G(i,i) = G(i,i).$$

Mais

$$\sum_{j \in E} \pi(j) G(j,i) = \sum_{j \in E} \sum_{n \geq 0} \pi(j) P^n(j,i) = \sum_{n \geq 0} \pi P^n(i) = \sum_{n \geq 0} \pi(i) = \infty.$$

Donc $G(i,i) = \infty$ et i est récurrent. Puis, par le lemme 4.5.2, la mesure $m_i(k) = \mathbf{E}_i[\sum_{n=0}^{\tau_i - 1} \mathbf{1}_{\{X_n = k\}}]$ est invariante, elle est donc proportionnelle à π. En particulier, $\mathbf{E}_i(\tau_i) = \sum_{k \in E} m_i(k) < \infty$.

Enfin, on a montré $(**)$ dans (i) implique (iii). Et $(*)$ n'est autre que $(**)$ quand $k = i$.

Lemme 4.5.7 (Principe du maximum) *Pour tous $i, j \in E$, $G(i,j) \leq G(j,j)$.*

Preuve: On suppose $i \neq j$ et on écrit

$$\begin{aligned} G(i,j) &= \mathbf{E}_i\Big[\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k = j\}}\Big] = \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_j < \infty\}} \sum_{k=0}^{\infty} \mathbf{1}_{\{X_{k+\tau_j} = j\}}\Big] \\ &= \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_j < \infty\}} \mathbf{E}_i\Big(\sum_{k=0}^{\infty} \mathbf{1}_{\{X_{k+\tau_j} = j\}} \Big| \mathcal{F}_{\tau_j}\Big)\Big] \\ &= \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_j < \infty\}} \mathbf{E}_i\Big(\Phi((X_{k+\tau_j})_{k \geq 0}) \Big| \mathcal{F}_{\tau_j}\Big)\Big], \end{aligned}$$

où $\Phi : E^{\mathbf{N}} \mapsto \mathbf{R}^+$ est définie par $\Phi((x_k)_{k \geq 0}) = \sum_{k=0}^{\infty} \mathbf{1}_{\{x_k = j\}}$. Par Markov forte,

$$\begin{aligned} G(i,j) &= \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_j < \infty\}} \mathbf{E}_{X_{\tau_j}}\Big(\Phi((X_k)_{k \geq 0})\Big)\Big] \\ &= \mathbf{E}_i\Big[\mathbf{1}_{\{\tau_j < \infty\}} \mathbf{E}_j\Big(\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k = j\}}\Big)\Big] = \mathbf{P}_i(\tau_j < +\infty) G(j,j). \end{aligned}$$

Théorème 4.5.8 Thm ergodique des CdM récurrentes positives. *Si $(X_n)_{n \geq 0}$ est une chaîne de Markov récurrente positive de probabilité invariante π, pour toute loi initiale ν, et pour toute fonction $f : E \to \mathbf{R}^+$, $\mathbf{P}_\nu$-presque sûrement,*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \int_E f \, d\pi. \qquad (*)$$

(on a $\int_E f \, d\pi = \sum_{i \in E} f(i) \pi(i)$) et pour tout $i, j \in E$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k(i,j) = \pi(j).$$

En choisissant $f = \mathbf{1}_{\{j\}}$, on voit qu'on a bien, pour n'importe quelle loi initiale, $\pi(j) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \mathbf{1}_{\{X_k=j\}}$: c'est la proportion du temps passé en j.

Preuve: On sait déjà que pour tout $i \in E$ (si i est un état récurrent, ce qui est le cas), pour tout $f, g : E \mapsto \mathbf{R}^+$, $\mathbf{P}_i$-p.s.,

$$\lim_{n\to+\infty} \frac{\sum_{k=0}^{n} f(X_k)}{\sum_{k=0}^{n} g(X_k)} = \frac{\int_E f\, dm_i}{\int_E g\, dm_i}$$

où $m_i(j) = \mathbf{E}_i[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=j\}}]$. Donc, avec $g = 1$, on trouve

$$\lim_{n\to+\infty} \frac{\sum_{k=0}^{n} f(X_k)}{n} = \frac{1}{\mathbf{E}_i[\tau_i]} \mathbf{E}_i\Big[\sum_{0}^{\tau_i-1} f(X_k)\Big] = \int_E f d\pi$$

$\mathbf{P}_i$-p.s. pour tout $i \in E$ (et donc $\mathbf{P}_\nu$-p.s. pour toute loi initiale ν).

Avec $f = \mathbf{1}_{\{j\}}$, on obtient $\lim_n \frac{1}{n} \sum_0^n \mathbf{1}_{\{X_k=j\}} = \pi(j)$ $\mathbf{P}_i$-p.s. Donc par convergence dominée, $\lim_n \mathbf{E}_i[\frac{1}{n} \sum_0^n \mathbf{1}_{\{X_k=j\}}] = \pi(j)$, soit encore $\lim_n \frac{1}{n} \sum_0^n P^n(i,j) = \pi(j)$.

Ce théorème donne deux moyens pratiques d'approcher π si, comme c'est souvent le cas, on ne sait pas la calculer explicitement. La première façon est la méthode de Monte Carlo, qui consiste à simuler sur ordinateur une longue trajectoire X_n de la chaîne, et d'utiliser que $\pi(j) \simeq \frac{1}{n} \sum_0^n \mathbf{1}_{\{X_k=j\}}$. L'autre façon est de calculer itérativement P^n, par exemple dans le cas où E est fini. Puis de faire la moyenne des $P^n(i,j)$ pour approcher $\pi(j)$.

Définition 4.5.9 *On dit qu'une suite $(x_n)_{n\geq 0}$ à valeurs dans E tend vers l'infini si pour toute partie finie F de E, il existe $n_F \geq 1$ t.q. pour tout $n \geq n_F$, $x_n \notin F$.*

Théorème 4.5.10 *Pour une chaîne de Markov irréductible, on a l'alternative :*

Cas récurrent: *Tous les états sont récurrents et partant de tout point, la chaîne visite une infinité de fois tous les autres, p.s.*

Cas transitoire: *Partant de tout point, $(X_n)_{n\geq 0}$ tend vers l'infini, p.s.*

Preuve: S'il existe un état récurrent i, considérons la mesure invariante $m_i(j) = \mathbf{E}_i[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=j\}}]$ qui lui est associée. Soit $j \in E$ fixé. La chaîne étant irréductible, $m_i(j) > 0$. Or, $\mathbf{P}_i$-p.s.,

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=j\}}}{\sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=i\}}} = \frac{m_i(j)}{m_i(i)} > 0.$$

Puisque i est récurrent, $\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k=i\}} = +\infty$. La limite précédente assure donc que $\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k=j\}} = +\infty$ $\mathbf{P}_i$-p.s. En prenant l'espérance, on voit que $G(i,j) = \infty$, d'où $G(j,j) = +\infty$ par le principe du maximum : l'état j est récurrent.

Supposons qu'il n'y a pas d'état récurrent. Donc pour tout $j \in E$, $G(j,j) < \infty$. Donc si F est une partie finie de E,

$$\mathbf{E}_i\Big(\sum_{k=0}^{\infty} \mathbf{1}_F(X_k)\Big) = \sum_{j\in F} G(i,j) \quad \leq \quad \sum_{j\in F} G(j,j) < +\infty$$

Donc $\sum_{k=0}^{\infty} \mathbf{1}_F(X_k)$ est une somme finie, $\mathbf{P}_i$-p.s., ce qui n'est possible que si X_n quitte F définitivement après un certain temps.

Corollaire 4.5.11 *Si E est fini, toute chaîne irréductible est récurrente positive.*

Exemple. Avec $E = \{1,2,3\}$ et $P(2,1) = P(2,3) = 1/2$ et $P(1,2) = P(3,2) = 1$. Faire un dessin. Comprendre qu'intuitivement, $\pi = (1/4\ 1/2\ 1/4)$. Calculer π en résolvant $\pi = \pi P$. Calculer $\mathbf{E}_i[\tau_i]$ (à la main) pour $i = 1,2,3$ et vérifier qu'on a bien $\pi(i) = 1/\mathbf{E}_i(\tau_i)$.

Proposition 4.5.12 *Une chaîne irréductible récurrente admet une et une seule (à constante près) mesure invariante.*

Preuve: L'existence d'une mesure invariante m a été vue au lemme 4.5.2. Pour l'unicité, on procède comme au lemme 4.5.5 : on considère deux mesures invariantes m et ℓ, toutes deux non identiquement nulles, donc strictement positives partout (voir le lemme 4.5.4). On fixe $i \in E$ et on pose $m' = \lambda \ell$, où $\lambda = m(i)/\ell(i)$.

On vérifie que $\mu(j) = \min\{m(j), m'(j)\}$ est invariante. En effet, on a bien sûr $\mu P(j) \leq \mu(j)$ (comme au lemme 4.5.5). On considère ensuite la mesure positive $\nu(j) = \mu(j) - \mu P(j)$, et on écrit, pour $n \geq 1$,

$$\sum_{k=0}^{n} \nu P^k = \sum_{k=0}^{n} \mu P^k - \mu P^{k+1} = \mu - \mu P^{n+1} \leq \mu.$$

Donc $\nu G = \sum_{k=0}^{\infty} \nu P^k \leq \mu$, i.e. pour tout $j \in E$, $\sum_{i\in E} \nu(i) G(i,j) \leq \mu(j)$. C'est impossible si ν n'est pas identiquement nulle, puisque $G(i,j) = \infty$ pour tout $i, j \in E$ (rappelons que la chaîne est supposée récurrente). Nous avons montré que $\mu P = \mu$, qui est donc invariante.

Donc $m - \mu$ est une mesure (positive), invariante, nulle en i (par construction), donc identiquement nulle (cf lemme 4.5.4). Donc $m = \mu$.

De même, $m' = \mu$.

Donc $m = m'$, i.e. m et ℓ sont proportionnelles.

Remarque 4.5.13 *Il est possible qu'une chaîne irréductible admette une mesure invariante sans qu'elle soit récurrente.*

Exemple/Exercice. Considérons la marche aléatoire simple sur $\mathbf{Z}$, i.e. $P(i, i+1) = p$, $P(i, i-1) = 1 - p$, avec $p \in]0,1[$. Montrer qu'elle est irréductible. Montrer que $m(i) = 1$ pour tout i est invariante (quelque soit la valeur de p).

A l'aide de la loi des grands nombres (écrire X_n comme une somme de v.a. i.i.d.), montrer que $(X_n)_{n\geq 0}$ est transitoire si $p \neq 1/2$.

Si $p = 1/2$, montrer que $\mathbf{P}_0(X_{2n} = 0) = C_{2n}^n 2^{-n} \sim \sqrt{2\pi n}(n/e)^n$. Montrer que $G(0,0) = \infty$. Donc $(X_n)_{n\geq 0}$ est récurrente. Mais non récurrente positive (sinon, elle aurait une probabilité invariante, $m \equiv 1$ lui serait proportionnelle, ce qui est absurde). Remarquer aussi que, comme $m(i) = \mathbf{E}_0[\sum_{k=0}^{\tau_0 - 1} \mathbf{1}_{\{X_k = i\}}]$ est invariante (et vérifie $m(0) = 1$), on a donc $m(i) = 1$ pour tout $i \in \mathbf{Z}$. C'est très surprenant : le nombre moyen de passages par $i \in \mathbf{Z}$ entre deux passages en 0 vaut toujours 1, que $i = 1$ ou que $i = 10000$!!

Résumé.

• Si une chaîne irréductible a une probabilité invariante, alors elle est récurrente positive (et la probabilité invariante est unique).

• Il est possible qu'une chaîne irréductible admette une (ou plusieurs) mesures invariantes sans qu'elle soit récurrente.

• Si une chaîne est récurrente, alors elle admet une unique (à constante près) mesure invariante.

4.6 Stationnarité et réversibilité

On dit qu'un processus $(X_n)_{n \in \mathbf{N}}$ est stationnaire (au sens strict) si pour tous $m, n \geq 0$ la loi de $(X_m, X_{m+1}, \cdots, X_{m+n})$ est la même que celle de $(X_0, X_1, \cdots, X_n)$.

Théorème 4.6.1 *Une chaîne de Markov de loi initiale ν est un processus stationnaire si et seulement si ν est une probabilité invariante.*

Preuve: Si (X_n) est stationnaire, alors X_0 et X_1 ont la même loi, et donc $\nu = \nu P$, ν est bien invariante.

Si ν est invariante, alors $(X_m, X_{m+1}, \cdots)$ est une chaîne de transition P (la même) et de loi initiale $\nu P^m = \nu$. Donc bien sûr, $(X_m, X_{m+1}, \cdots, X_{m+n})$ a la même loi que $(X_0, X_1, \cdots, X_n)$.

Proposition 4.6.2 *Soit $(X_n)_{n \geq 0}$ une chaîne de Markov récurrente positive de probabilité invariante ν. Sous $\mathbf{P}_\nu$, le processus $(X_{N-n})_{n \in \{0, \ldots, N\}}$ (avec N grand fixé) est une chaîne de Markov de transition.*

$$Q(i,j) = \frac{\nu(j)}{\nu(i)} P(j,i)$$

Preuve: Pour tout $n \in \{0, \ldots, N\}$ et $a_0, a_1, \cdots, a_n \in E$, on a:

$$\begin{aligned}
&\mathbf{P}_\nu(X_N = a_0, X_{N-1} = a_1, \cdots, X_{N-n} = a_n) \\
&\quad = \mathbf{P}_\nu(X_{N-n} = a_n, X_{N-n+1} = a_{n-1}, \cdots, X_N = a_0) \\
&\quad = \nu P^{N-n}(a_n) P(a_n, a_{n-1}) \cdots P(a_1, a_0) \\
&\quad = \nu(a_n) P(a_n, a_{n-1}) \cdots P(a_1, a_0) \\
&\quad = \frac{\nu(a_n)}{\nu(a_{n-1})} P(a_n, a_{n-1}) \frac{\nu(a_{n-1})}{\nu(a_{n-2})} P(a_{n-1}, a_{n-2}) \cdots \frac{\nu(a_1)}{\nu(a_0)} P(a_1, a_0) \nu(a_0) \\
&\quad = Q(a_n, a_{n-1}) Q(a_{n-1}, a_{n-2}) \cdots Q(a_1, a_0) \nu(a_0)
\end{aligned}$$

ce qui montre la proposition. On a utilisé que la loi de X_{N-n} est νP^{N-n} et que $\nu P^{N-n} = \nu$ car ν est invariante.

Définition 4.6.3 *Une probabilité ν sur E est dite réversible (ou P-réversible) si pour tout $i, j \in \mathbf{E}$*

$$\nu(i)P(i,j) = \nu(j)P(j,i).$$

Une probabilité réversible ν est invariante : il suffit de sommer sur j l'égalité précédente.

4.7 Processus de naissance et mort à temps discret

Dans cette section, on étudie un cas particulier de chaînes, mais les méthodes utilisées (pour montrer la transience ou la récurrence) sont assez générales.

Définition 4.7.1 *Une chaîne de Markov à temps discret à valeurs dans $E = \mathbf{N}$ est appelée processus de naissance et mort (PNM) à temps discret si $P(i,j) = 0$ dès que $|i-j| \neq 1$. On pose alors $p_i = P(i, i+1)$ (pour $i \geq 0$) et $q_i = P(i, i-1)$ (pour $i \geq 1$). On a forcément $p_0 = 1$ et $p_i + q_i = 1$ pour $i \geq 1$.*

Remarque 4.7.2 *Un PNM à temps discret est irréductible si et seulement si $p_i > 0$ et $q_i > 0$ pour tout $i \geq 1$.*

Théorème 4.7.3 *Un PNM à temps discret irréductible est*

(i) récurrent si et seulement si

$$\sum_{i \geq 1} \frac{q_1 q_2 \dots q_i}{p_1 p_2 \dots p_i} = \infty;$$

(ii) récurrent positif si et seulement si

$$S = 1 + \sum_{i \geq 1} \frac{p_1 p_2 \dots p_{i-1}}{q_1 q_2 \dots q_i} < \infty.$$

Son unique probabilité invariante est alors donnée par $\pi_0 = 1/S$ et, pour tout $i \geq 1$, $\pi_i = (p_1 p_2 \dots p_{i-1})/(S q_1 q_2 \dots q_i)$.

Preuve: Commençons par (ii). Une mesure m est invariante si et seulement si $mP = m$, ce qui se réécrit $m_0 = m_1 q_1$ et $m_i = m_{i+1} q_{i+1} + m_{i-1} p_{i-1}$ pour $i \geq 1$. Cette récurrence d'ordre 2 se résout ainsi :

$$m_i = m_0 \frac{p_1 p_2 \dots p_{i-1}}{q_1 q_2 \dots q_i}, \quad i \geq 1.$$

Le point (ii) s'ensuit aisément (rappelons qu'une chaîne est récurrente positive si et seulement si elle admet une probabilité invariante).

Pour le point (i), remarquons que la chaîne est récurrente si et seulement si $\mathbf{P}_1(T_0 < \infty) = 1$, où $T_0 = \inf\{n \geq 0 \ : \ X_n = 0\}$, puisque par Markov simple et comme $X_1 = 1$ sous $\mathbf{P}_0$,

$$\mathbf{P}_0(\tau_0 < \infty) = \mathbf{P}_0(X_1 = 1, \tau_0 < \infty) = \mathbf{P}_1(\tau_0 < \infty) = \mathbf{P}_1(T_0 < \infty) = 1.$$

On fixe $a \in \mathbf{N}_*$ (grand), et on introduit $T_a = \inf\{n \geq 0 \ : \ X_n = a\}$ et on pose $u(i) = \mathbf{P}_i(T_0 < T_a)$ pour $i = 0, ..., a$. On a bien sûr $u(0) = 1$ et $u(a) = 0$ et montre facilement, par Markov simple, que

$$u(i) = q_i u(i-1) + p_i u(i+1) \quad \text{pour } i = 1, \dots, a-1.$$

Ainsi, $p_i(u(i+1) - u(i)) = q_i(u(i) - u(i-1))$, et donc, toujours pour $i = 1, \dots, a-1$, $u(i+1) - u(i) = (q_i/p_i) \dots (q_1/p_1)[u(1) - u(0)]$, puis

$$u(i) = u(0) + \sum_{k=0}^{i-1} (u(k+1) - u(k)) = u(0) + [u(1) - u(0)]\Big[1 + \sum_{k=1}^{i-1} \frac{q_1 \dots q_k}{p_1 \dots p_k}\Big].$$

En choisissant $i = a$, comme $u(a) = 0$ et $u(0) = 1$, on trouve

$$u(1) = 1 - \frac{1}{1 + \sum_{k=1}^{a-1} \frac{q_1 \dots q_k}{p_1 \dots p_k}}.$$

On a donc

$$\mathbf{P}_1(T_0 < T_a) = 1 - \frac{1}{1 + \sum_{k=1}^{a-1} \frac{q_1 \dots q_k}{p_1 \dots p_k}}.$$

Mais $T_a < \infty$ $\mathbf{P}_1$-p.s. pour tout $a \in \mathbf{N}_*$ (car c'est évident si la chaîne est récurrente, et c'est vrai aussi si la chaîne est transitoire, puisqu'alors $X_n \to \infty$ et est donc forcée de passer par a en partant de 1). De plus $T_a \geq a - 1 \to \infty$ $\mathbf{P}_1$-p.s. quand $a \to \infty$. Donc, à un évenement $\mathbf{P}_1$-négligeable près,

$$\{T_0 < \infty\} = \cup_{a \geq 1} \{T_0 < T_a\}$$

puis $\mathbf{P}_1(T_0 < \infty) = \lim_{a \to \infty} \mathbf{P}_1(T_0 < T_a)$. Donc en effet, $\mathbf{P}_1(T_0 < \infty) = 1$ si et seulement si $\sum_{k=1}^{\infty} \frac{q_1 \dots q_k}{p_1 \dots p_k} = \infty$.

4.8 Problèmes d'absorption

Considérons une chaîne de Markov $(X_n)_{n \geq 0}$ d'espace d'état fini E, de transition P.

On suppose que $E = F \cup G$, avec

(a) pour tout $i \in G$, $P(i, i) = 1$,

(b) pour tout $i \in F$, il existe $j \in G$ et $n \geq 1$ t.q. $P^n(i, j) > 0$.

Noter que la matrice s'écrit donc sous la forme

$$P = \begin{pmatrix} Q & A \\ 0 & I \end{pmatrix}.$$

Ces matrices sont toutes de natures différentes : $(Q(i,j))_{i\in F, j\in F}$, $(A(i,j))_{i\in F, j\in G}$ et $(I(i,j))_{i\in G, j\in G}$.

Noter qu'une telle chaîne n'est pas du tout irréductible. Les éléments de G sont appelés états absorbants (une fois qu'on y est, on y reste). L'énoncé suivant permet de calculer, partant de $i \in F$, la probabilité de finir absorbé en $j \in G$, ainsi que d'autres quantités tout à fait intéressantes.

Proposition 4.8.1 *(a) La matrice $N = \sum_{k=0}^{+\infty} Q^k$ existe et vérifie $N = (I-Q)^{-1}$.*

(b) On a $\lim_{n\to+\infty} P^n = \begin{pmatrix} 0 & NA \\ 0 & I \end{pmatrix}$.

(c) Soit $T_G = \inf\{n \geq 0 \ : \ X_n \in G\}$. Pour tout $i \in F$ et $j \in G$, on a

$$\mathbf{P}_i(X_{T_G} = j) = (NA)(i,j) \quad et \quad \mathbf{E}_i[T_G] = \sum_{j\in F} N(i,j).$$

Preuve: Remarquons d'abord (récurrence) que

$$P^n = \begin{pmatrix} Q^n & (I + Q + Q^2 + \cdots + Q^{n-1})A \\ 0 & I \end{pmatrix}.$$

Pour tout $i \in F$, i est transitoire. En effet, il suffit de montrer que $\mathbf{P}_i(\tau_i < \infty) < 1$. Mais, par (b), on sait qu'il existe $j \in G$ et $n \geq 1$ t.q. $P^n(i,j) > 0$, ce qui implique qu'on peut trouver $i_1, \ldots, i_{n-1} \in F$ t.q. $P(i,i_1)P(i_1,i_2)\cdots P(i_{n-1},j) > 0$. Quitte à raccourcir ce chemin, on peut supposer que tous les i_k sont différents de i. Du coup,

$$\begin{aligned} \mathbf{P}_i(\tau_i = \infty) &\geq \mathbf{P}_i(X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = j) \\ &= P(i,i_1)P(i_1,i_2)\cdots P(i_{n-1},j) > 0. \end{aligned}$$

Donc pour tout $i \in F$, $G(i,i) < +\infty$. Il résulte donc du principe du maximum que $G(i,j) < +\infty$ pour tout $i,j \in F$. Autrement dit la matrice N est finie (observer que $P^n(i,j) = Q^n(i,j)$ pour tout $i,j \in F$). En particulier Q^n tend vers 0. Comme $(I-Q)(I+Q+\cdots+Q^n) = I - Q^{n+1}$ on en déduit que $(I-Q)N = I$, soit encore $N = (I-Q)^{-1}$. On conclut aussi que

$$\lim_{n\to+\infty} P^n = \begin{pmatrix} 0 & NA \\ 0 & I \end{pmatrix}.$$

Notons que $T_G = \sum_{n\geq 0}\sum_{j\in F} \mathbf{1}_{\{X_n = j\}}$. Ainsi, pour tout $i \in F$,

$$\mathbf{E}_i[T_G] = \sum_{n\geq 0}\sum_{j\in F} P^n(i,j) = \sum_{n\geq 0}\sum_{j\in F} Q^n(i,j) = \sum_{j\in F} N(i,j),$$

cette somme (finie) étant bien sûr finie.

Enfin, la chaîne une fois dans G ne bouge plus, et donc $X_{T_G} = \lim_{n\to+\infty} X_n$. Ainsi, pour $i \in F$ et $j \in G$,

$$\mathbf{P}_i(X_{T_G} = j) = \lim_{n\to+\infty} \Pr_i(X_n = j) = \lim_{n\to+\infty} P^n(i,j) = (NA)(i,j).$$

Exemple. J'ai 1 euro et j'ai besoin de 3 euros. Je joue à pile ou face avec un ami. Je mise toujours 1 (j'en récupère 0 avec probabilité $1/2$ et 2 avec probabilité $1/2$. J'arrête quand j'ai 0 ou 3 euros.

Ma fortune X_n à l'instant n est une chaîne de Markov, de loi initiale δ_1, sur $E = \{0, 1, 2, 3\}$. On a $P(0,0) = 1$, $P(1,0) = P(1,2) = 1/2$, $P(2,1) = P(2,3) = 1/2$ et $P(3,3) = 1$. On peut écrire $E = F \cup G$ avec $F = \{1,2\}$ et $G = \{0,3\}$, et on a

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 0 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 0 & 3 \end{array} \\ \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

qui nous donne Q et A. On peut facilement calculer

$$N = (I - Q)^{-1} = \begin{pmatrix} 4/3 & 2/3 \\ 2/3 & 4/3 \end{pmatrix} \qquad \text{et} \qquad NA = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}.$$

(Attention aux natures de ces matrices). On trouve que $NA(1,3) = 1/3$: la probabilité, partant d'une fortune de 1 euro, de terminer avec 3 euros, vaut, comme il se doit, $1/3$. L'espérance du temps de jeu est $\mathbf{E}_1(T_G) = N(1,1) + N(1,2) = 4/3 + 2/3 = 2$.

Chapitre 5

Processus markoviens de sauts

Dans toute la suite, E est un ensemble dénombrable, éventuellement fini, qu'on munit de la distance discrète : $d(i,j) = \mathbf{1}_{\{i \neq j\}}$. Nous allons étudier les "processus de Markov à temps continu" à valeurs dans E. On les appelle aussi "chaînes de Markov à temps continu" ou encore "processus markoviens de saut".

5.1 Premières propriétés

Définition 5.1.1 *Soit $(P_t)_{t \in \mathbf{R}^+}$ une famille de matrices de transition sur E, i.e. $P_t(i,j) \geq 0$ et $\sum_{j \in E} P_t(i,j) = 1$.*

Un processus $(X_t)_{t \geq 0}$ à valeurs dans E est appelé PMS (processus markovien de saut) de semigroupe $(P_t)_{t \geq 0}$ si, pour tout $j \in E$, tout $0 \leq s \leq t+s$,

$$\mathbf{P}(X_{s+t} = j | \mathcal{F}_s) = P_t(X_s, j),$$

et si $t \mapsto X_t$ est p.s. continu à droite sur $\mathbf{R}^+$. On a noté $\mathcal{F}_s = \sigma(X_u, u \leq s)$.

On peut montrer (à l'aide du théorème de classe monotone) le résultat suivant.

Remarque 5.1.2 *Pour montrer qu'un processus p.s. continu à droite $(X_t)_{t \geq 0}$ à valeurs dans E est un PMS de semigroupe $(P_t)_{t \geq 0}$, il suffit de montrer que pour tout $n \in \mathbf{N}$, pour tous $0 \leq t_1 \leq \cdots \leq t_{n+1}$, pour tout $i_1, \cdots, i_{n+1} \in E$,*

$$\mathbf{P}(X_{t_{n+1}} = i_{n+1} | X_{t_1} = i_1, \cdots, X_{t_n} = i_n) = P_{t_{n+1} - t_n}(i_n, i_{n+1}).$$

Donnons un premier exemple de tel processus, qui s'avérera relativement général.

Proposition 5.1.3 *Soit $(Z_n)_{n \in \mathbf{N}}$ une chaîne de Markov sur E de probabilité de transition Q indépendante d'un processus de Poisson $(N_t)_{t \in \mathbf{R}^+}$ de paramètre λ. Alors*

$$(X_t = Z_{N_t},\ t \in \mathbf{R}^+)$$

est un PMS de semigroupe

$$P_t(i,j) = \sum_{n=0}^{+\infty} e^{-\lambda t} \frac{(\lambda t)^n Q^n(i,j)}{n!}.$$

Preuve: Notons $0 < T_1 < T_2 ...$ les instants de saut du processus de Poisson. On a alors $X_t = Z_0$ pour $t \in [0, T_1[$, puis $X_t = Z_1$ pour $t \in [T_1, T_2[$, etc. $(X_t)_{t \geq 0}$ est donc bien continu à droite p.s. On écrit ensuite (notons ν la loi de Z_0)

$$\begin{aligned}
&\mathbf{P}(X_0 = i_0, X_{t_1} = i_1, \cdots, X_{t_r} = i_r) \\
&= \sum_{k_1, \cdots, k_r \geq 0} \mathbf{P}\Big(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \cdots, N_{t_r} - N_{t_{r-1}} = k_r, \\
&\qquad\qquad Z_0 = i_0, Z_{k_1} = i_1, Z_{k_1+k_2} \cdots, Z_{k_1+\cdots+k_r} = i_r\Big) \\
&= \sum_{k_1, \cdots, k_r \geq 0} \frac{(\lambda t_1)^{k_1}}{k_1!} e^{-\lambda t_1} \times \cdots \times \frac{(\lambda(t_r - t_{r-1}))^{k_r}}{k_r!} e^{-\lambda(t_r - t_{r-1})} \\
&\qquad\qquad \times \nu(i_0) Q^{k_1}(i_0, i_1) \cdots Q^{k_r}(i_{r-1}, i_r) \\
&= \nu(i_0) P_{t_1}(i_0, i_1) P_{t_2 - t_1}(i_1, i_2) \times \cdots \times P_{t_r - t_{r-1}}(i_{r-1}, i_r).
\end{aligned}$$

On conclut bien sûr que

$$\begin{aligned}
&\mathbf{P}(X_{t_r} = i_r | X_0 = i_0, X_{t_1} = i_1, \cdots, X_{t_{r-1}} = i_{r-1}) \\
&= \frac{\nu(i_0) P_{t_1}(i_0, i_1) P_{t_2 - t_1}(i_1, i_2) \times \cdots \times P_{t_r - t_{r-1}}(i_{r-1}, i_r)}{\nu(i_0) P_{t_1}(i_0, i_1) P_{t_2 - t_1}(i_1, i_2) \times \cdots \times P_{t_{r-1} - t_{r-2}}(i_{r-2}, i_{r-1})} \\
&= P_{t_r - t_{r-1}}(i_{r-1}, i_r).
\end{aligned}$$

Proposition 5.1.4 Equation de Chapman Kolmogorov. *Supposons que E soit le “vrai” espace d'états du PMS, c'est à dire que pour tout $i \in E$, il existe $t_i \geq 0$ (déterministe) tel que $\mathbf{P}(X_{t_i} = i) > 0$. Alors $(P_t)_{t \in \mathbf{R}^+}$, d'un PMS a la propriété de semigroupe suivante: pour tous $t, s \geq 0$, $P_{t+s} = P_t P_s$.*

Preuve: On écrit

$$\begin{aligned}
P_{t+s}(i,j) &= \mathbf{P}(X_{t_i+t+s} = j | X_{t_i} = i) \\
&= \sum_{k \in E} \mathbf{P}(X_{t_i+t+s} = j | X_{t_i+t} = k, X_{t_i} = i) \mathbf{P}(X_{t_i+t} = k | X_{t_i} = i) \\
&= \sum_{k \in E} P_s(k,j) P_t(i,k) = P_t P_s(i,j).
\end{aligned}$$

Comme dans le cas des chaînes, on note $\mathbf{P}_i, \mathbf{E}_i$ (resp. $\mathbf{P}_\nu, \mathbf{E}_\nu$) la loi du processus lorsque $X_0 = i$ (resp. $X_0 \sim \nu$).

Rappelons que temps d'arrêt T du processus $(X_t)_{t \geq 0}$ est une v.a. à valeurs dans $\mathbf{R}^+ \cup \{\infty\}$ telle que pour tout $t \geq 0$, $\{T \leq t\} \in \mathcal{F}_t$. On introduit alors la tribu $\mathcal{F}_T = \{A \in \mathcal{F} \ : \ A \cap \{T \leq t\} \in \mathcal{F}_t \ \forall \ t \geq 0\})$. Comme d'habitude, on a posé préalablement $\mathcal{F}_t = \sigma(\{X_s, s \leq t\})$.

Théorème 5.1.5 Markov forte. *Soit $(X_t)_{t \geq 0}$ un PMS à valeurs dans E et T un temps d'arrêt. Pour toute fonction mesurable $f : E^{\mathbf{R}^+} \to \mathbf{R}$, bornée ou positive,*

$$\mathbf{1}_{\{T<\infty\}} \mathbf{E}\Big[f((X_{T+t})_{t \geq 0}) \Big| \mathcal{F}_T\Big] = \mathbf{1}_{\{T<\infty\}} \mathbf{E}_{X_T}\Big[f((X_t)_{t \geq 0})\Big].$$

Remarque 5.1.6 *On dit qu'on utilise Markov simple quand on applique Markov forte avec un temps déterministe (tout temps déterministe est un temps d'arrêt).*

Preuve: Il suffit de montrer le théorème quand $f((x_t)_{t\geq 0}) = \phi(x_{t_1}, \ldots, x_{t_n})$ avec $\phi : E^n \mapsto \mathbf{R}^+$ et $0 \leq t_1 < t_2 \cdots < t_n$. Et on peut supposer que $\phi(x_1, \ldots, x_n) = \mathbf{1}_{\{x_1=i_1,\cdots,x_n=i_n\}}$, car toute fonction ϕ sur E^n est une combinaison linéaire de telles fonctions. L'objectif est donc de montrer que

$$\mathbf{1}_{\{T<\infty\}}\mathbf{P}_\nu(X_{T+t_1} = i_1, \cdots, X_{T+t_n} = i_n|\mathcal{F}_T) = \mathbf{1}_{\{T<\infty\}}\mathbf{P}_{X_T}(X_{t_1} = i_1, \cdots, X_{t_n} = i_n).$$

Etape 1. On établit ici la formule lorsque T est à valeurs dans l'ensemble dénombrable $\{k2^{-\ell} : k \in \mathbf{N}\} \cup \{\infty\}$. Pour cela, on écrit

$$\begin{aligned} I &:= \mathbf{1}_{\{T<\infty\}}\mathbf{P}_\nu(X_{T+t_1} = i_1, \cdots, X_{T+t_n} = i_n|\mathcal{F}_T) \\ &= \sum_{k\geq 0} \mathbf{1}_{\{T=k2^{-\ell}\}}\mathbf{P}_\nu(X_{k2^{-\ell}+t_1} = i_1, \cdots, X_{k2^{-\ell}+t_n} = i_n|\mathcal{F}_T) \\ &= \sum_{k\geq 0} \mathbf{1}_{\{T=k2^{-\ell}\}}\mathbf{P}_\nu(X_{k2^{-\ell}+t_1} = i_1, \cdots, X_{k2^{-\ell}+t_n} = i_n|\mathcal{F}_{k2^{-\ell}}). \end{aligned}$$

Pour la dernière ligne, on a utilisé un argument du type du lemme 4.3.2. Ceci donne

$$\begin{aligned} I &= \sum_{k\geq 0} \mathbf{1}_{\{T=k2^{-\ell}\}}P_{t_1}(X_{k2^{-\ell}}, i_1)P_{t_2-t_1}(i_1, i_2)\cdots P_{t_n-t_{n-1}}(i_{n-1}, i_n) \\ &= \sum_{k\geq 0} \mathbf{1}_{\{T=k2^{-\ell}\}}\mathbf{P}_{X_{k2^{-\ell}}}(X_{t_1} = i_1, \cdots, X_{t_n} = i_n) \\ &= \mathbf{1}_{\{T<\infty\}}\mathbf{P}_{X_T}(X_{t_1} = i_1, \cdots, X_{t_n} = i_n). \end{aligned}$$

Etape 2. Si T est un temps d'arrêt quelconque, on l'approche par une suite de temps d'arrêt T_ℓ de sorte que $\lim_\ell \downarrow T_\ell = T$, que pour chaque ℓ, T_ℓ soit à valeurs dans $\{k2^{-\ell} : k \in \mathbf{N}\} \cup \{\infty\}$, et que $\{T_\ell < \infty\} = \{T < \infty\}$. Par l'étape 1, on a donc, pour tout ℓ,

$$\mathbf{1}_{\{T<\infty\}}\mathbf{P}_\nu(X_{T_\ell+t_1} = i_1, \cdots, X_{T_\ell+t_n} = i_n|\mathcal{F}_{T_\ell}) = \mathbf{1}_{\{T<\infty\}}\mathbf{P}_{X_T}(X_{t_1} = i_1, \cdots, X_{t_n} = i_n).$$

Comme X est continu à droite et comme T_ℓ décroit vers T, on a $\lim_\ell X_{T_\ell+t} = X_{T+t}$ pour tout t. Donc, en utilisant que $\mathcal{F}_T \subset \mathcal{F}_{T_\ell}$ (et la convergence dominée),

$$\begin{aligned} &\mathbf{1}_{\{T<\infty\}}\mathbf{P}_\nu(X_{T+t_1} = i_1, \cdots, X_{T+t_n} = i_n|\mathcal{F}_T) \\ &= \lim_{\ell\to\infty} \mathbf{1}_{\{T<\infty\}}\mathbf{P}_\nu(X_{T_\ell+t_1} = i_1, \cdots, X_{T_\ell+t_n} = i_n|\mathcal{F}_T) \\ &= \lim_{\ell\to\infty} \mathbf{1}_{\{T<\infty\}}\mathbf{E}_\nu(\mathbf{P}_\nu(X_{T_\ell+t_1} = i_1, \cdots, X_{T_\ell+t_n} = i_n|\mathcal{F}_{T_\ell})|\mathcal{F}_T) \\ &= \lim_{\ell\to\infty} \mathbf{E}_\nu(\mathbf{1}_{\{T<\infty\}}\mathbf{P}_\nu(X_{T_\ell+t_1} = i_1, \cdots, X_{T_\ell+t_n} = i_n|\mathcal{F}_{T_\ell})|\mathcal{F}_T) \\ &= \lim_{\ell\to\infty} \mathbf{E}_\nu(\mathbf{1}_{\{T<\infty\}}\mathbf{P}_{X_{T_\ell}}(X_{t_1} = i_1, \cdots, X_{t_n} = i_n)|\mathcal{F}_T) \\ &= \mathbf{E}_\nu(\mathbf{1}_{\{T<\infty\}}\mathbf{P}_{X_T}(X_{t_1} = i_1, \cdots, X_{t_n} = i_n)|\mathcal{F}_T) \\ &= \mathbf{1}_{\{T<\infty\}}\mathbf{P}_{X_T}(X_{t_1} = i_1, \cdots, X_{t_n} = i_n). \end{aligned}$$

5.2 Description dynamique

Considérons un PMS $(X_t)_{t\geq 0}$ et posons $T_0 = 0$ puis, par récurrence, pour $n \geq 0$,

$$T_{n+1} = \inf\{t > T_n \ : \ X_t \neq X_{T_n}\}.$$

Ce sont des temps d'arrêt.

Remarque 5.2.1 *Si on suppose que pour tout $i \in E$, $\mathbf{P}_i(T_1 < \infty) = 1$, alors p.s., pour tout $n \geq 1$, $T_n < \infty$: cela se montre par récurrence, en utilisant Markov forte.*

On rappelle que si $U \sim \mathcal{E}(1)$, alors $U/\lambda \sim \mathcal{E}(\lambda)$.

Théorème 5.2.2 *On suppose que pour tout $i \in E$, $\mathbf{P}_i(T_1 < \infty) = 1$, et on pose $Q(i,j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i(T_1)$, pour $i, j \in E$.*

- *$(X_{T_n})_{n\in\mathbf{N}}$ est une chaîne de Markov de transition Q.*
- *Sachant $(X_{T_n})_{n\in\mathbf{N}}$, les v.a. $T_1, T_2 - T_1, \cdots$ sont indépendantes de lois $\mathcal{E}(\lambda(X_{T_0}))$, $\mathcal{E}(\lambda(X_{T_1}))$, ...*

Les $(\lambda(i))_{i\in E}$ sont les “taux de saut” du PMS et $(Q(i,j))_{i,j\in E}$ est sa “matrice de transition”. Voici une autre façon de dire la même chose.

Corollaire 5.2.3 *Avec les mêmes hypothèses et notations que dans le théorème, il existe une chaîne de Markov $(\xi_n)_{n\geq 0}$ de transition Q, indépendante d'une suite $(U_n)_{n\geq 1}$ de v.a. indépendantes de loi exponentielle de paramètre 1 telles que, pour tout $t \in \mathbf{R}^+$,*

$$X_t = \sum_{n=0}^{+\infty} \xi_n \mathbf{1}_{\{t\in[T_n, T_{n+1}[\}}$$

où $T_0 = 0$ et $T_{n+1} = T_n + U_{n+1}/\lambda(\xi_n)$.

Preuve: Il suffit de poser $\xi_n = X_{T_n}$ et $U_n = \lambda(\xi_{n-1})(T_n - T_{n-1})$ et d'appliquer le théorème.

Preuve du Théorème.

Etape 1. Montrons déjà que pour tout $i \in E$, $\mathbf{P}_i(T_1 \geq t) = e^{-\lambda(i)t}$ pour tout $t \geq 0$. On écrit

$$\begin{aligned}
\mathbf{P}_i(T_1 > t+s) &= \mathbf{P}_i(\forall\, u \in [0, t+s],\ X_u = i) \\
&= \mathbf{E}_i\Big[\mathbf{1}_{\{\forall\, u\in[0,t],\ X_u = i\}} \mathbf{P}_i(\forall\, u \in [t, t+s],\ X_u = i | \mathcal{F}_t)\Big].
\end{aligned}$$

Soit $\Phi : E^{\mathbf{R}^+} \mapsto \mathbf{R}^+$ définie par $\Phi((x_u)_{u\geq 0}) = \mathbf{1}_{\{\forall\, u\in[0,s],\ x_u = i\}}$. Par Markov simple,

$$\begin{aligned}
\mathbf{P}_i(\forall\, u \in [t, t+s],\ X_u = i | \mathcal{F}_t) &= \mathbf{E}_i\Big[\Phi((X_{t+u})_{u\geq 0})\Big|\mathcal{F}_t\Big] \\
&= \mathbf{E}_{X_t}\Big[\Phi((X_u)_{u\geq 0})\Big] \\
&= \mathbf{P}_{X_t}(\forall\, u \in [0, s],\ X_u = i).
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\mathbf{P}_i(T_1 > t+s) &= \mathbf{E}_i\Big[\mathbf{1}_{\{\forall\, u\in[0,t],\ X_u=i\}}\mathbf{P}_i(\forall\, u\in[0,s],\ X_u=i)\Big] \\
&= \mathbf{P}_i(\forall\, u\in[0,t],\ X_u=i)\mathbf{P}_i(\forall\, u\in[0,s],\ X_u=i) \\
&= \mathbf{P}_i(T_1>t)\mathbf{P}_i(T_1>s).
\end{aligned}$$

Donc T_1 suit une loi exponentielle sous $\mathbf{P}_i$. Son paramètre est bien sûr $\lambda(i) = 1/\mathbf{E}_i[T_1]$.

Etape 2. Remarquons ensuite que pour tout $i, j \in E$, et $t \geq 0$,

$$\mathbf{P}_i(T_1 > t, X_{T_1} = j) = e^{-\lambda(i)t}Q(i,j).$$

En effet,

$$\mathbf{P}_i(T_1 > t, X_{T_1} = j) = \mathbf{E}_i\Big[\mathbf{1}_{\{T_1>t\}}\mathbf{P}_i(X_{T_1}=j|\mathcal{F}_t)\Big] = \mathbf{E}_i\Big[\mathbf{1}_{\{T_1>t\}}\mathbf{E}_i[\Phi((X_{t+s})_{s\geq 0})|\mathcal{F}_t]\Big]$$

où $\Phi : E^{\mathbf{R}^+} \mapsto \mathbf{R}^+$ est définie par $\Phi((x_s)_{s\geq 0}) = \mathbf{1}_{\{x_{t_1((x_s)_{s\geq 0})}=j\}}$, où $t_1((x_s)_{s\geq 0}) = \inf\{s \geq 0 \ : \ x_s \neq x_0\}$. On a utilisé que $T_1 = t_1((X_s)_{s\geq 0}) > t$ implique que $\Phi((X_{t+s})_{s\geq 0}) = \Phi((X_t)_{s\geq 0})$. Donc par Markov simple,

$$\mathbf{P}_i(T_1 > t, X_{T_1} = j) = \mathbf{E}_i\Big[\mathbf{1}_{\{T_1>t\}}\mathbf{E}_{X_t}(\Phi((X_s)_{s\geq 0}))\Big] = \mathbf{E}_i\Big[\mathbf{1}_{\{T_1>t\}}\mathbf{P}_{X_t}(X_{T_1}=j)\Big],$$

puis (comme $T > t_1$ implique que $X_t = i$),

$$\mathbf{P}_i(T_1 > t, X_{T_1} = j) = \mathbf{E}_i\Big[\mathbf{1}_{\{T_1>t\}}\mathbf{P}_i(X_{T_1}=j)\Big] = \mathbf{E}_i\Big[\mathbf{1}_{\{T_1>t\}}Q(i,j)\Big]$$

par définition de Q, et on conclut à l'aide de l'étape 1.

Etape 3. Pour $n \geq 1$, $u_1, \ldots, u_n > 0$ et $i_1, \ldots, i_n \in E$,

$$\begin{aligned}
&\mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \cdots, T_n - T_{n-1} > u_n, X_{T_n} = i_n) \\
&= \mathbf{E}_i[\mathbf{1}_{\{T_1>u_1, X_{T_1}=i_1,\cdots,T_{n-1}-T_{n-2}>u_n, X_{T_{n-1}}=i_{n-1}\}}\mathbf{E}_i(\mathbf{1}_{\{T_n-T_{n-1}>u_n, X_{T_n}=i_n\}}|\mathcal{F}_{T_{n-1}})] \\
&= \mathbf{E}_i[\mathbf{1}_{\{T_1>u_1, X_{T_1}=i_1,\cdots,T_{n-1}-T_{n-2}>u_n, X_{T_{n-1}}=i_{n-1}\}}\mathbf{E}_i(\Phi((X_{T_{n-1}+s})_{s\geq 0})|\mathcal{F}_{T_{n-1}})],
\end{aligned}$$

où $\Phi((x_s)_{s\geq 0}) = \mathbf{1}_{\{t_1((x_s)_{s\geq 0})=i_n, x_{t_1((x_s)_{s\geq 0})}=i_n\}}$. (On a utilisé que $t_1(X_{T_{n-1}+s})_{s\geq 0}) = T_n - T_{n-1}$). Par Markov forte,

$$\begin{aligned}
&\mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \cdots, T_n - T_{n-1} > u_n, X_{T_n} = i_n) \\
&= \mathbf{E}_i[\mathbf{1}_{\{T_1>u_1, X_{T_1}=i_1,\cdots,T_{n-1}-T_{n-2}>u_n, X_{T_{n-1}}=i_{n-1}\}}\mathbf{E}_{X_{T_{n-1}}}(\Phi((X_s)_{s\geq 0}))] \\
&= \mathbf{E}_i[\mathbf{1}_{\{T_1>u_1, X_{T_1}=i_1,\cdots,T_{n-1}-T_{n-2}>u_{n-1}, X_{T_{n-1}}=i_{n-1}\}}\mathbf{E}_{X_{T_{n-1}}}(\mathbf{1}_{\{T_1>u_n, X_{T_1}=i_n\}})].
\end{aligned}$$

Mais $\mathbf{E}_{X_{T_{n-1}}}(\mathbf{1}_{\{T_1>u_n, X_{T_1}=i_n\}}) = Q(X_{T_{n-1}}, i_n)e^{-\lambda(X_{T_{n-1}})u_n}$ par l'étape 2. De plus, nous pouvons remplacer $X_{T_{n-1}}$ par i_{n-1}, donc

$$\begin{aligned}
&\mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \cdots, T_n - T_{n-1} > u_n, X_{T_n} = i_n) \\
&= \mathbf{E}_i[\mathbf{1}_{\{T_1>u_1, X_{T_1}=i_1,\cdots,T_{n-1}-T_{n-2}>u_{n-1}, X_{T_{n-1}}=i_{n-1}\}}]Q(i_{n-1}, i_n)e^{-\lambda(i_{n-1})u_n}.
\end{aligned}$$

On montre donc, en répétant ce procédé, que

$$\begin{aligned}&\mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \cdots, T_n - T_{n-1} > u_n, X_{T_n} = i_n)\\ &\quad = Q(i, i_1)e^{-\lambda(i)u_1}Q(i_1, i_2)e^{-\lambda(i_1)u_2} \cdots Q(i_{n-1}, i_n)e^{-\lambda(i_{n-1})u_n}.\end{aligned}$$

En choisissant $u_1 = \ldots = u_n = 0$, on conclut aisément que $(X_{T_n})_{n\geq 0}$ est une chaîne de Markov de transition Q, puisque

$$\mathbf{P}_i(X_{T_1} = i_1, \cdots, X_{T_n} = i_n) = Q(i, i_1)Q(i_1, i_2) \cdots Q(i_{n-1}, i_n).$$

Enfin, on voit que

$$\begin{aligned}&\mathbf{P}_i(T_1 > u_1, \cdots, T_n - T_{n-1} > u_n | X_{T_1} = i_1, \cdots, X_{T_n} = i_n)\\ &\quad = \frac{\mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \cdots, T_n - T_{n-1} > u_n, X_{T_n} = i_n)}{\mathbf{P}_i(X_{T_1} = i_1, \cdots, X_{T_n} = i_n)}\\ &\quad = e^{-\lambda(i)u_1}e^{-\lambda(i_1)u_2} \cdots e^{-\lambda(i_{n-1})u_n}.\end{aligned}$$

Ceci achève la preuve.

On a montré aussi

Proposition 5.2.4 *Pour tout $i, j \in E$,*

$$\mathbf{P}_i(X_{T_1} = j, T_1 > t) = e^{-\lambda(i)t}Q(i, j).$$

Simulation d'un PMS avec $(Q(i,j))_{i,j\in E}$ et $(\lambda(i))_{i\in E}$ donnés.

On part d'un état i (donné ou simulé). On simule $U_1 \sim \mathcal{E}(1)$ et on pose $T_1 = U_1/\lambda(i)$. Pour $t \in [0, T_1[$ on pose $X_t = i$.

On simule $\xi_1 \sim Q(i, \cdot)$ et on pose $X_{T_1} = \xi_1$. On simule $U_2 \sim \mathcal{E}(1)$ et on pose $T_2 = T_1 + U_2/\lambda(\xi_1)$.

On simule $\xi_2 \sim Q(\xi_1, \cdot)$ et on pose $X_{T_2} = \xi_2$. On simule $U_3 \sim \mathcal{E}(1)$ et on pose $T_3 = T_2 + U_3/\lambda(\xi_2)$.

Etc.

Enfin, une remarque qui peut s'avérer pratique.

Proposition 5.2.5 *Soit $(Q(i, j))_{i,j\in E}$ une matrice de transition t.q. $Q(i, i) = 0$ pour tout $i \in E$ et soit $(\lambda(i))_{i\in E}$ une famille de réels positifs tels que $\sup_{i\in E} \lambda(i) \leq \lambda < \infty$. Soit $(Z_n)_{n\geq 0}$ une chaîne de Markov de transition P définie par*

$$P(i, j) = 1 - \frac{\lambda(i)}{\lambda} \textit{ si } i = j \quad \textit{et} \quad P(i, j) = \frac{\lambda(i)Q(i, j)}{\lambda} \textit{ si } i \neq j.$$

Soit aussi $(N_t)_{t\geq 0}$ un processus de Poisson de paramètre λ, indépendant de $(Z_n)_{n\geq 0}$. Alors $(Z_{N_t})_{t\geq 0}$ est un PMS de taux de sauts $(\lambda(i))_{i\in E}$ et de matrice de transition $(Q(i, j))_{i,j\in E}$.

Preuve: On sait déjà que $(Z_{N_t})_{t\geq 0}$ est un PMS (voir le lemme 5.1.3). Notons T_1 l'instant du premier saut de $(Z_{N_t})_{t\geq 0}$. On doit montrer que $\mathbf{E}_i[T_1] = 1/\lambda(i)$ et que $\mathbf{P}_i[Z_{N_{T_1}} = j] = Q(i,j)$.

$$\mathbf{P}_i[T_1 > t] = \sum_{k\geq 0} \mathbf{P}_i(N_t = k, Z_1 = \cdots = Z_k = i) = \sum_{k\geq 0} e^{-\lambda t}\frac{(\lambda t)^k}{k!}(P(i,i))^k,$$

d'où $\mathbf{P}_i[T_1 > t] = \exp(-\lambda(1-P(i,i))t) = \exp(-\lambda(i)t)$. Donc $\mathbf{E}_i[T_1] = 1/\lambda(i)$. Ensuite, pour $j \neq i$, (si $i = j$, on a $Q(i,i) = 0 = \mathbf{P}_i[Z_{N_{T_1}} = i]$)

$$\mathbf{P}_i[Z_{N_{T_1}} = j] = \sum_{\ell\geq 1} \mathbf{P}_i(Z_1 = \cdots = Z_{\ell-1} = i, Z_\ell = j) = \sum_{\ell\geq 1}(P(i,i))^{\ell-1}P(i,j),$$

donc $\mathbf{P}_i[Z_{N_{T_1}} = j] = P(i,j)/(1-P(i,i)) = Q(i,j)$.

5.3 Explosion

Il est a priori possible que, avec une probabilité non nulle, $\lim_{n\to+\infty} T_n < +\infty$, on dit alors qu'il y a **explosion**. Afin de donner un critère de non explosion, montrons d'abord:

Lemme 5.3.1 *Soit $U_n, n \geq 1$, une suite de variables aléatoires indépendantes de loi exponentielle de paramètre $\lambda_1, \lambda_2, \cdots$.*

(i) Si $\sum_{k=1}^\infty 1/\lambda_k < \infty$, alors $\sum_{k=1}^\infty U_k < \infty$ p.s.

(ii) Si $\sum_{k=1}^\infty 1/\lambda_k = \infty$, alors $\sum_{k=1}^\infty U_k = \infty$ p.s.

Preuve: Comme $\mathbf{E}[\sum_{k=1}^\infty U_k] = \sum_{k=1}^{+\infty} 1/\lambda_k$, le point (i) est trivial. Pour (ii), on écrit

$$\mathbf{E}\Big[\exp\Big(-\sum_{k=1}^\infty U_k\Big)\Big] = \prod_{k=1}^\infty \frac{\lambda_k}{1+\lambda_k} = \exp\Big(-\sum_{k\geq 1}\log(1+1/\lambda_k)\Big) = 0.$$

Donc $\sum_{k=1}^\infty U_k = \infty$ p.s.

Proposition 5.3.2 *Considérons un PMS $(X_t)_{t\geq 0}$. Il n'y a pas explosion, c'est à dire que $\lim_{n\to+\infty} T_n = +\infty$ p.s., si et seulement si $\sum_{k=1}^{+\infty} \frac{1}{\lambda(X_{T_k})} = +\infty$ p.s.*

Preuve: En utilisant le lemme ci-dessus et le fait que sachant $(X_{T_n})_{n\geq 0}$, les $T_{n+1} - T_n$ sont indépendants et $\mathcal{E}(\lambda(X_{T_n}))$, on peut écrire que

$$\begin{aligned}
\mathbf{P}\Big(\lim_{n\to+\infty} T_n < \infty\Big) &= \mathbf{E}\Big[\mathbf{P}\Big(\lim_{n\to+\infty} T_n < \infty \Big| (X_{T_n})_{n\geq 0}\Big)\Big] \\
&= \mathbf{E}\Big(\mathbf{1}_{\{\sum_{k=1}^\infty \frac{1}{\lambda(X_{T_k})} < \infty\}}\Big) \\
&= \mathbf{P}\Big(\sum_{k=1}^\infty \frac{1}{\lambda(X_{T_k})} < \infty\Big).
\end{aligned}$$

Corollaire 5.3.3 *Si (i)* $\sup_{i\in E}\lambda(i)<\infty$ *ou (ii) la chaîne* $(X_{T_n})_{n\geq 0}$ *est récurrente, alors il n'y a pas d'explosion.*

Preuve: Il résulte de la proposition précédente que s'il y a explosion alors $\lambda(X_{T_n})\to +\infty$ lorsque $n\to+\infty$. C'est impossible dans les deux cas (i) et (ii).

L'explosion entraîne de nombreux phénomènes désagréables. Pour les éviter **nous supposerons toujours désormais qu'il n'y a pas d'explosion**.

5.4 Le générateur

Quand on modélise un phénomène, on part toujours des taux d'évènement $(\lambda(i))_{i\in E}$ et de la matrice de transition $(Q(i,j))_{i,j\in E}$. S'il est presque toujours impossible de décrire le semigroupe $(P_t)_{t\geq 0}$, on peut souvent l'étudier, en utilisant les équations de Kolmogorov. Commençons par définir le générateur, qui n'est pour l'instant qu'une notation.

Définition 5.4.1 *On appelle générateur d'un PMS la "matrice"* $A(i,j), i,j\in E$, *définie par*

$$A(i,j)=\begin{cases}\lambda(i)Q(i,j), & si\ \ i\neq j,\\ -\lambda(i), & si\ \ i=j.\end{cases}$$

Remarquons que $Q(i,i)=0$ donc $\sum_{j\in E}A(i,j)=0$. Pour décrire le lien entre le générateur et le semigroupe, montrons le lemme important suivant qui dit qu'en un temps petit, le processus n'a pas le temps de sauter deux fois. De façon plus précise, on a le lemme suivant. On note $o(t)$ une fonction telle que $o(t)/t\to 0$ quand $t\to 0$.

Lemme 5.4.2 *Pour tout* $i\in E$, $\mathbf{P}_i(T_2\leq t)=o(t)$.

Preuve: En utilisant le théorème de la section précédente, on peut écrire:

$$\mathbf{P}_i(T_2\leq t)=\sum_{j\in E}\mathbf{P}_i(X_{T_1}=j,T_2\leq t)=\sum_{j\in E}\mathbf{P}_i(X_{T_1}=j)\,\mathbf{P}_i(T_2\leq t|X_{T_1}=j).$$

Mais

$$\begin{aligned}\mathbf{P}_i(T_2\leq t|X_{T_1}=j) &\leq \mathbf{P}_i(T_1\leq t,T_2-T_1\leq t|X_{T_1}=j)\\ &= (1-e^{-\lambda(i)t})(1-e^{-\lambda(j)t})\\ &\leq \lambda(i)t(1-e^{-\lambda(j)t}).\end{aligned}$$

Donc

$$\mathbf{P}_i(T_2\leq t)\leq\lambda(i)t\sum_{j\in E}\mathbf{P}_i(X_{T_1}=j)(1-e^{-t\lambda(j)t}).$$

C'est un $o(t)$ car $\lim_{t\to 0}\sum_{j\in E}\mathbf{P}_i(X_{T_1}=j)(1-e^{-t\lambda(j)t})=0$ par convergence dominée.

Théorème 5.4.3 *Le semigroupe* $(P_t)_{t\geq 0}$ *vérifie* $P_0=I$ *et, pour tout* $t\geq 0$,

- $\dfrac{dP_t}{dt} = AP_t$ *(équation de Kolmogorov backward);*
- *Si il n'y a pas explosion,* $\dfrac{dP_t}{dt} = P_t A$ *(équation de Kolmogorov forward).*

Remarque 5.4.4 *L'équation de Kolmogorov backward se réécrit : pour tout* $i, j \in E$,

$$P_t'(i,j) = -\lambda(i)P_t(i,j) + \sum_{k \neq i} \lambda(i)Q(i,k)P_t(k,j).$$

L'équation de Kolmogorov forward se réécrit : pour tout $i, j \in E$,

$$P_t'(i,j) = -P_t(i,j)\lambda(j) + \sum_{k \neq j} P_t(i,k)\lambda(k)Q(k,j).$$

Nous allons montrer proprement l'équation backward et "rapidement" l'équation forward.

Preuve rapide de l'équation forward (Cette preuve est correcte si E est fini). Pour $0 \le t < t+h$, on écrit

$$P_{t+h}(i,j) = \mathbf{P}_i(X_{t+h} = j) = I + J + K,$$

où

$$\begin{aligned}
I &= \mathbf{P}_i(X_t = j \text{ et pas de saut entre } t \text{ et } t+h),\\
J &= \sum_{k \neq j} \mathbf{P}_i(X_t = k \text{ et exactement un de saut de } k \text{ à } j \text{ entre } t \text{ et } t+h),\\
K &= \mathbf{P}_i(X_{t+h} = j \text{ et au moins deux sauts entre } t \text{ et } t+h).
\end{aligned}$$

Par le lemme, on sait que $K = o(h)$. De plus, par Markov simple, on a

$$I = \mathbf{P}_i(X_t = j)\mathbf{P}_j(T_1 > h) = P_t(i,j)e^{-\lambda(j)h} = P_t(i,j)(1 - \lambda(j)h) + o(h).$$

Enfin, toujours par Markov simple (et par le lemme),

$$\begin{aligned}
J &= \sum_{k \neq j} \mathbf{P}_i(X_t = k)\mathbf{P}_k(T_1 < h < T_2, X_{T_1} = j)\\
&= \sum_{k \neq j} \mathbf{P}_i(X_t = k)\mathbf{P}_k(T_1 < h, X_{T_1} = j) + o(h)\\
&= \sum_{k \neq j} P_t(i,k)(1 - e^{-\lambda(k)h})Q(k,j) + o(h)\\
&= \sum_{k \neq j} P_t(i,k)\lambda(k)Q(k,j)h + o(h).
\end{aligned}$$

Ainsi,

$$P_{t+h}(i,j) = P_t(i,j)(1 - \lambda(j)h) + \sum_{k \neq j} P_t(i,k)\lambda(k)Q(k,j)h + o(h),$$

donc

$$\frac{P_{t+h}(i,j) - P_t(i,j)}{h} = -P_t(i,j)\lambda(j) + \sum_{k\neq j} P_t(i,k)\lambda(k)Q(k,j) + \frac{o(h)}{h},$$

et il n'y a plus qu'à faire tendre h vers 0.

Preuve de l'équation backward.

Etape 1. On montre ici que pour tout $i, j \in E$, tout $t \geq 0$,

$$P_t(i,j) = e^{-\lambda(i)t}\mathbf{1}_{\{i=j\}} + \sum_{k\neq i} \int_0^t e^{-\lambda(i)(t-s)} A(i,k) P_s(k,j)\, ds$$

Pour cela, on écrit

$$\begin{aligned} P_t(i,j) &= \mathbf{P}_i(X_t = j) = \mathbf{P}_i(X_t = j, T_1 > t) + \mathbf{P}_i(X_t = j, T_1 \leq t) \\ &= \mathbf{1}_{\{i=j\}} e^{-\lambda(i)t} + \sum_{k\neq i} \mathbf{P}_i(X_{T_1} = k, X_t = j, T_1 \leq t). \end{aligned}$$

Mais

$$\begin{aligned} \mathbf{P}_i(X_{T_1} = k, X_t = j, T_1 \leq t) &= \mathbf{E}_i[\mathbf{1}_{\{X_{T_1}=k, T_1\leq t\}} \mathbf{P}_i(X_t = j | \mathcal{F}_{T_1})] \\ &= \mathbf{E}_i[\mathbf{1}_{\{X_{T_1}=k, T_1\leq t\}} P_{t-T_1}(X_{T_1}, j)]. \end{aligned}$$

Du coup,

$$\begin{aligned} \mathbf{P}_i(X_{T_1} = k, X_t = j, T_1 \leq t) &= \mathbf{E}_i(\mathbf{1}_{\{X_{T_1}=k, T_1\leq t\}} P_{t-T_1}(k, j)) \\ &= \int_0^t e^{-\lambda(i)s} P_{t-s}(k,j)\lambda(i)Q(i,k)\, ds \\ &= \int_0^t e^{-\lambda(i)(t-s)} A(i,k) P_s(k,j)\, ds. \end{aligned}$$

Ligne 2, on a utilisé que T_1 et X_{T_1} sont indépendants, que $T_1 \sim \mathcal{E}(\lambda(i))$ (sous $\mathbf{P}_i$) et que $\mathbf{P}_i(X_{T_1} = k) = Q(i,k)$. Ligne 3, on a utilisé le changement de variables $s \to t - s$. L'étape est terminée.

Etape 2. On a donc

$$P_t(i,j) = e^{-\lambda(i)t}\Big[\mathbf{1}_{\{i=j\}} + \sum_{k\neq i} \int_0^t e^{\lambda(i)s} A(i,k) P_s(k,j)\, ds\Big].$$

En utilisant que P_s est borné par 1 et comme $\sum_{k\neq i} A(i,k) = \lambda(i)$, on déduit de cette formule que $P_t(i,j)$ est continu, puis dans un second temps qu'il est dérivable et que

$$\begin{aligned} P_t'(i,j) &= -\lambda(i)P_t(i,j) + e^{-\lambda(i)t}\Big[\sum_{k\neq i} e^{\lambda(i)t} A(i,k) P_t(k,j)\Big] \\ &= -\lambda(i)P_t(i,j) + \sum_{k\neq i} A(i,k) P_t(k,j) \end{aligned}$$

comme désiré.

5.5 Mesures invariantes

Définition 5.5.1 *(i) On dit qu'un PMS $(X_t)_{t\geq 0}$ est :*
- *irréductible si la chaîne $(X_{T_n})_{n\geq 0}$ de transition Q est irréductible,*
- *récurrente si la chaîne $(X_{T_n})_{n\geq 0}$ de transition Q est récurrente,*
- *transitoire si la chaîne $(X_{T_n})_{n\geq 0}$ de transition Q est transitoire.*

(ii) On dit que m est une mesure invariante si pour tout $t \geq 0$, $mP_t = m$.

(iii) On dit que le PMS est récurrent positif s'il admet une probabilité invariante.

Il est possible que $(X_t)_{t\geq 0}$ soit récurrent positif sans que $(X_{T_n})_{n\geq 0}$ le soit (l'idée est d'accélerer le temps, nous ferons cela plus loin).

Lemme 5.5.2 *Si le PMS $(X_t)_{t\geq 0}$ est (irréductible et) récurrent, la chaîne de Markov de transition*

$$\Pi(i,j) = \int_0^\infty e^{-t} P_t(i,j)\, dt$$

est aussi récurrente.

Preuve: Par définition, comme $(X_t)_{t\geq 0}$ est récurrent, on a $\sum_{n\geq 0} Q^n(i,j) = \infty$ pour tout $i, j \in E$. Nous allons montrer que

$$\sum_{n\geq 1} \Pi^n(i,j) = (\lambda(j))^{-1} \sum_{n\geq 0} Q^n(i,j), \qquad (*)$$

ce qui permettra d'affirmer qu'en effet, la chaîne de transition Π est récurrente.

On montre par récurrence que $\Pi^n(i,j) = \int_0^\infty e^{-t} \frac{t^{n-1}}{(n-1)!} P_t(i,j)\, dt$. En effet, c'est vrai si $n = 1$, et si c'est vrai avec n, alors, comme $P_t P_s = P_{t+s}$,

$$\begin{aligned}
\Pi^{n+1} &= \Big(\int_0^\infty e^{-t} \frac{t^{n-1}}{(n-1)!} P_t\, dt\Big)\Big(\int_0^\infty e^{-t} P_s\, ds\Big) \\
&= \int_0^\infty \int_0^\infty e^{-t-s} \frac{t^{n-1}}{(n-1)!} P_{t+s} dt ds \\
&= \int_0^\infty e^{-u} P_u \int_0^u \frac{t^{n-1}}{(n-1)!} dt du \\
&= \int_0^\infty e^{-u} P_u \frac{u^n}{n!} du.
\end{aligned}$$

Par suite,

$$\sum_{n\geq 1} \Pi^n(i,j) = \sum_{n\geq 1} \int_0^\infty e^{-t} \frac{t^{n-1}}{(n-1)!} P_t(i,j)\, dt = \int_0^\infty P_t(i,j) dt.$$

Remarquons ensuite que

$$\int_0^\infty P_t(i,j)\, dt = \mathbf{E}_i \left(\int_0^\infty \mathbf{1}_{\{X_t=j\}}\, dt\right) = \sum_{n=0}^\infty \mathbf{E}_i \left(\mathbf{1}_{\{X_{T_n}=j\}} (T_{n+1} - T_n)\right)$$

car $X_t = \sum_{n\geq 0} X_{T_n} \mathbf{1}_{\{t\in[T_n,T_{n+1}[\}}$. Donc

$$\int_0^\infty P_t(i,j)\,dt = \mathbf{E}_i\left(\int_0^\infty \mathbf{1}_{\{X_t=j\}}\,dt\right) = \sum_{n=0}^\infty \mathbf{E}_i\left(\mathbf{1}_{\{X_{T_n}=j\}}\mathbf{E}_i[T_{n+1}-T_n|(X_{T_k})_{k\geq 0}]\right).$$

Mais sachant $(X_{T_k})_{k\geq 0}$, $T_{n+1}-T_n \sim \mathcal{E}(\lambda(X_{T_n}))$. Ainsi,

$$\int_0^\infty P_t(i,j)\,dt = \sum_{n=0}^\infty \mathbf{E}_i\left(\mathbf{1}_{\{X_{T_n}=j\}}(\lambda(X_{T_n}))^{-1}\right) = (\lambda(j))^{-1}\sum_{n=0}^\infty Q^n(i,j).$$

Nous avons montré $(*)$.

Proposition 5.5.3 *Si X_t est un PMS récurrent positif, alors il est récurrent.*

Preuve: On note π sa probabilité invariante (telle que $\pi P_t = \pi$ pour tout $t \geq 0$). On rappelle, voir quelques lignes plus haut, que $\int_0^\infty P_t(i,j)\,dt = (\lambda(j))^{-1}G(i,j)$, où $G(i,j) = \sum_{n=0}^\infty Q^n(i,j)$. On rappelle que $G(i,j) \leq G(j,j)$ et que notre objectif est de vérifier que $G(j,j) = \infty$ (pour n'importe quel $j \in E$ fixé).

On multiplie l'égalité précédente par $\pi(i)$ et on somme sur $i \in E$, on en déduit que $\int_0^\infty (\pi P_t)(j)\,dt = (\lambda(j))^{-1}\sum_{i\in E}\pi(i)G(i,j) \leq (\lambda(j))^{-1}G(j,j)$. Ainsi,

$$G(j,j) \geq \lambda(j)\int_0^\infty (\pi P_t)(j)\,dt = \lambda(j)\int_0^\infty \pi(j)\,dt = \infty.$$

Théorème 5.5.4 *Soit $(X_t)_{t\geq 0}$ un PMS* **récurrent**. *Il admet une mesure invariante m, unique à une constante près. De plus, on a*

1. Pour un état i fixé, si $S_i = \inf\{t > T_1; X_t = i\}$, $m_i(j) = \mathbf{E}_i[\int_0^{S_i} \mathbf{1}_{\{X_s=j\}}\,ds]$ est une mesure invariante.

2. Si m est une mesure telle que ν définie par $\nu(j) = \lambda(j)m(j)$ vérifie $\nu Q = \nu$, alors m est invariante.

3. Si $mA = 0$ (i.e. pour tout $i \in E$, $m(i)\lambda(i) = \sum_{j\neq i} m(j)\lambda(j)Q(j,i)$), alors m est invariante.

Preuve: noter qu'informellement, par Kolmogorov rétrograde, si $mA = 0$, alors pour tout $t \geq 0$, $(mP_t)' = mP_t' = mAP_t = 0$, donc $mP_t = mP_0 = m$. Mais nous avons échangé une somme et une dérivation, donc ce n'est rigoureux que dans le cas où E est fini.

Unicité. Si m est une mesure invariante, $mP_t = m$ pour tout $t \geq 0$. En particulier, $m\Pi = m$, où Π est la transition introduite dans le lemme. Donc m est une mesure invariante de la chaîne de noyau Π. Or on sait qu'il n'existe qu'une seule telle mesure, à une constante près, car cette chaîne est récurrente.

Existence et point 1. On fixe $i \in E$ et on montre que la mesure définie au point 1 est invariante. On fixe $t > 0$ et $k \in E$, et on montre que $m_iP_t(k) = m_i(k)$. On se rappelle que $m_iP_t(k) = \sum_{j\in E} m_i(j)P_t(j,k)$, et on utilise la définition de m_i :

$$m_iP_t(k) = \sum_{j\in E}\mathbf{E}_i\Big[\int_0^{S_i}\mathbf{1}_{\{X_s=j\}}\mathbf{P}_j(X_t=k)ds\Big] = \mathbf{E}_i\Big[\int_0^{S_i}\mathbf{P}_{X_s}(X_t=k)ds\Big].$$

Mais par Markov simple, $\mathbf{P}_{X_s}(X_t = k) = \mathbf{P}_i(X_{t+s} = k|\mathcal{F}_s)$ puis, comme S_i est un temps d'arrêt, $\mathbf{1}_{\{S_i \geq s\}}\mathbf{P}_{X_s}(X_t = k) = \mathbf{E}_i(\mathbf{1}_{\{S_i \geq s, X_{t+s}=k\}}|\mathcal{F}_s)$. Donc

$$\begin{aligned} m_i P_t(k) &= \mathbf{E}_i\Big[\int_0^\infty \mathbf{E}_i(\mathbf{1}_{\{S_i \geq s, X_{t+s}=k\}}|\mathcal{F}_s)ds\Big] = \mathbf{E}_i\Big[\int_0^\infty \mathbf{1}_{\{S_i \geq s, X_{t+s}=k\}}ds\Big] \\ &= \mathbf{E}_i\Big[\int_t^{t+S_i} \mathbf{1}_{\{X_u=k\}}du\Big] = A + B, \end{aligned}$$

où $A = \mathbf{E}_i[\int_0^{S_i} \mathbf{1}_{\{X_u=k\}}du]$ et $B = \mathbf{E}_i[\int_{S_i}^{S_i+t} \mathbf{1}_{\{X_u=k\}}du] = \mathbf{E}_i[\int_0^t \mathbf{1}_{\{X_{S_i+u}=k\}}du]$. En utilisant Markov forte,

$$B = \mathbf{E}_i\Big[\int_0^t \mathbf{E}_i[\mathbf{1}_{\{X_{S_i+u}=k\}}|\mathcal{F}_{S_i}]du\Big] = \mathbf{E}_i\Big[\int_0^t \mathbf{E}_{X_{S_i}}[\mathbf{1}_{\{X_u=k\}}]du\Big].$$

Comme $X_{S_i} = i$, on conclut que $B = \mathbf{E}_i[\int_0^t \mathbf{E}_i[\mathbf{1}_{\{X_u=k\}}]du] = \mathbf{E}_i[\int_0^t \mathbf{1}_{\{X_u=k\}}du]$ puis, enfin, que

$$m_i P_t(k) = A + B = \mathbf{E}_i\Big[\int_0^{S_i} \mathbf{1}_{\{X_u=k\}}du\Big] = m_i(k)$$

par définition de m_i.

Point 2. On fixe $i \in E$, on considère la mesure m_i construite au point 1, on pose $\nu(j) = \lambda(j)m_i(j)$, pour tout $j \in E$, et on montre que $\nu Q = \nu$. (Par unicité de la mesure invariante de Q, si $\tilde{\nu}$ est une autre mesure vérifiant $\tilde{\nu}Q = \tilde{\nu}$, $\tilde{\nu}$ est proportionnelle à ν, donc $\tilde{m}$ définie par $\tilde{m}(j) = \tilde{\nu}(j)/\lambda(j)$ est proportionnelle à m_i, donc $\tilde{m}$ est invariante pour le PMS).

En posant $\tau_i = \inf\{n \geq 1 \ : \ X_{T_n} = i\}$, on a (car $X_t = \sum_{n\geq 0} X_{T_n}\mathbf{1}_{\{t\in[T_n,T_{n+1}[\}}$),

$$\forall\, j \in E, \quad m_i(j) = \mathbf{E}_i\Big[\sum_{n=0}^{\tau_i-1}(T_{n+1} - T_n)\mathbf{1}_{\{X_{T_n}=j\}}\Big] = \frac{1}{\lambda(j)}\mathbf{E}_i\Big[\sum_{n=0}^{\tau_i-1}\mathbf{1}_{\{X_{T_n}=j\}}\Big].$$

La seconde égalité utilise que sachant $X_{T_n} = j$, $(T_{n+1} - T_n) \sim \mathcal{E}(\lambda(j))$. Du coup,

$$\nu(j) = \mathbf{E}_i\Big[\sum_{n=0}^{\tau_i-1}\mathbf{1}_{\{X_{T_n}=j\}}\Big],$$

qui est invariante pour la chaîne $(X_{T_n})_{n\geq 0}$ (cf Lemme 4.5.2) et donc vérifie $\nu = \nu Q$.

Point 3. Il suffit de voir que si m est une mesure vérifiant $mA = 0$, alors ν définie par $\nu(j) = \lambda(j)m(j)$ satisfait $\nu = \nu Q$, ce qui est immédiat, puis d'utiliser le point 2.

Proposition 5.5.5 *Si X_t est un PMS récurrent positif, de probabilité invariante π, alors pour toute $f : E \mapsto \mathbf{R}$ (bornée ou positive), pour tout $i \in E$, $\mathbf{P}_i$-p.s.,*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(X_s)\, ds = \int_E f\, d\pi.$$

Preuve: Soit $i \in E$ fixé et $S_i = \inf\{t > T_1 \ : \ X_t = i\}$. On peut montrer que $(X_t)_{t\geq 0}$ est régénératif, avec $\tau_0 = 0$, $\tau_1 = S_i$, etc. (poser $\tau_{n+1} = \inf\{t > \tau_n \ : X_t = i$ et $\exists\ s \in [\tau_n, t]$ t.q. $X_s \neq i\}$). Il résulte donc du théorème 3.1.3 que $\mathbf{P}_i$-p.s., $\lim_{t\to\infty} t^{-1} \int_0^t f(X_s)ds = \int_E f d\pi$, où

$$\pi(j) = \frac{1}{\mathbf{E}_i[S_i]}\mathbf{E}_i\Big[\int_0^{S_i} \mathbf{1}_{\{X_s=j\}}ds\Big].$$

Par le théorème précédent, π est bien l'unique probabilité invariante (elle est invariante, et bien sûr, $\sum_{j\in E}\pi(j) = \mathbf{E}_i[S_i]/\mathbf{E}_i[S_i] = 1$).

Résumé important.

• Si $(Q(i,j))_{i,j\in E}$ est irréductible récurrente, alors le PMS associé à $(Q(i,j))_{i,j\in E}$ et $(\lambda(i))_{i\in E}$ admet une unique mesure invariante (à constante près) m, caractérisée par $mA = 0$, i.e. $\lambda(i)m(i) = \sum_{j\neq i} m(j)\lambda(j)Q(j,i)$ pour tout $i \in E$.

• Il est possible, si $(Q(i,j))_{i,j\in E}$ n'est PAS récurente, de trouver une probabilité m telle que $mA = m$ sans que le PMS (ni Q) soit récurrent (nous donnerons un exemple dans quelques pages).

• S'il existe une probabilité m telle que $mA = 0$ ET $\sum_{i\in E} m(i)\lambda(i) < \infty$, alors m est une probabilité invariante du PMS, qui est donc récurrent positif. (En effet, on a alors que $\nu(i) = \lambda(i)m(i)$ est invariante pour Q, comme c'est une mesure finie, on peut trouver une probabilité invariante pour Q, la chaîne associée est donc récurrente positive, donc le PMS est récurrent et on peut appliquer le théorème: m est invariante).

• Exemple où la chaîne est récurrente nulle et le PMS est récurrent positif : on considère $E = \mathbf{Z}$, $Q(i, i-1) = Q(i, i+1) = 1/2$. On a vu (c'est la marche aléatoire simple symétrique) que la chaîne de transition Q est récurrente nulle, de mesure invariante $\nu(i) = 1$ pour tout $i \in \mathbf{Z}$. Si $\lambda(i) = (1 + i^2)$, alors $m(i) = \nu(i)/\lambda(i) = 1/(1 + i^2)$ vérifie $mA = 0$ (pour le PMS de caractéristiques Q et λ), c'est donc une mesure invariante finie du PMS, et $\pi(i) = a/(1 + i^2)$ (où $a = 1/\sum_{i\in\mathbf{Z}}(1 + i^2)^{-1}$) est une probabilité invariante du PMS, qui est donc récurrent positif.

5.6 Processus de naissance et mort

Les processus de naissance et mort représentent la classe la plus simple (et la plus importante) de PMS à valeurs dans $E = \mathbf{N}$.

Définition 5.6.1 *On appelle PNM (processus de naissance et mort) tout PMS à valeurs dans* $\mathbf{N}$ *tel que* $Q(i,j) = 0$ *si* $j \notin \{i-1, i+1\}$.

On a donc aussi $A(i,j) = 0$ *si* $j \notin \{i-1, i+1\}$. *On pose* $\alpha_i = A(i, i-1)$ *pour tout* $i \geq 1$, $\beta_i = A(i, i+1)$ *pour tout* $i \geq 0$. *On a alors*

(i) $\lambda(0) = \beta_0$ *et* $Q(0,1) = 1$ *;*

(ii) pour $i \geq 1$, $\lambda(i) = \alpha_i + \beta_i$, $Q(i, i-1) = \frac{\alpha_i}{\alpha_i+\beta_i}$, $Q(i, i+1) = \frac{\beta_i}{\alpha_i+\beta_i}$.

L'idée (on verra cela plus ou moins rigoureusement au prochain chapitre) est que le PMS $(X_t)_{t\geq 0}$ modélise une population : X_t est le nombre d'individus vivants à l'instant t. Pour tout $i \geq 1$, α_i est le *taux de mort* quand i individus sont vivants, et β_i est le *taux de naissance* quand i individus sont vivants. Donc (c'est lié au lemme des deux réveils) quand on est dans l'état i, on saute avec taux $\lambda(i) = \beta_i + \alpha_i$, et c'est une naissance (donc on saute de 1) avec probabilité $\frac{\beta_i}{\alpha_i+\beta_i}$ ou c'est une mort (donc on saute de -1) avec probabilité $\frac{\alpha_i}{\alpha_i+\beta_i}$.

La remarque suivante est très facile (dessiner le graphe associé à la matrice Q).

Remarque 5.6.2 *Un PNM est irréductible si et seulement si $\alpha_i > 0$ pour tout $i \geq 1$ et $\beta_i > 0$ pour tout $i \geq 0$.*

Théorème 5.6.3 *Un PNM irréductible est*

(i) récurrent si et seulement si

$$\sum_{i\geq 1} \frac{\alpha_1\alpha_2\cdots\alpha_i}{\beta_1\beta_2\cdots\beta_i} = \infty.$$

(ii) récurrent positif si et seulement si

$$\sum_{i\geq 1} \frac{\alpha_1\alpha_2\cdots\alpha_i}{\beta_1\beta_2\cdots\beta_i} = \infty \quad et \quad S = 1 + \sum_{i\geq 1} \frac{\beta_0\beta_1\cdots\beta_{i-1}}{\alpha_1\alpha_2\cdots\alpha_i} < \infty.$$

Sa probabilité invariante est alors $\pi(0) = 1/S$ et $\pi(i) = (\beta_0\beta_1\cdots\beta_{i-1})/(S\alpha_1\alpha_2\cdots\alpha_i)$ pour $i \geq 1$.

Preuve: (i) Par définition, le PMS est récurrent ssi la chaîne de transition Q l'est. Mais cette chaîne est un PNM à temps discret. On peut donc appliquer le théorème 4.7.3, et on voit que notre PNM est récurrent ssi

$$\sum_{i\geq 1} \frac{q_1q_2\cdots q_i}{p_1p_2\cdots p_i} = \infty,$$

où $q_i = Q(i, i-1)$ et $p_i = Q(i, i+1)$. La conclusion s'ensuit immédiatement, puisque $q_i/p_i = \alpha_i/\beta_i$.

(ii) Supposons $S < \infty$. Comme le processus est récurrent, il suffit de vérifier que π, qui est bien sûr une probabilité, vérifie $\pi A = 0$, i.e., pour tout $i \in \mathbf{N}$, $\pi(i)\lambda(i) = \sum_{j\neq i} \pi(j)\lambda(j)Q(j,i)$.

Si $i = 0$, on veut

$$\pi(0)\beta_0 = \pi(1)(\alpha_1 + \beta_1)\frac{\alpha_1}{\alpha_1 + \beta_1},$$

i.e. $\pi(0)\beta_0 = \pi(1)\alpha_1$, ce qui est vrai.

Si $i \geq 1$, il faut

$$\pi(i)(\alpha_i + \beta_i) = \pi(i-1)(\alpha_{i-1} + \beta_{i-1})\frac{\beta_{i-1}}{\alpha_{i-1} + \beta_{i-1}} + \pi(i+1)(\alpha_{i+1} + \beta_{i+1})\frac{\alpha_{i+1}}{\alpha_{i+1} + \beta_{i+1}},$$

soit encore $\pi(i)(\alpha_i + \beta_i) = \pi(i-1)\beta_{i-1} + \pi(i+1)\alpha_{i+1}$, c'est vrai.

Si $S = \infty$, on vérifie que $m(0) = 1$ et $m(i) = (\beta_0\beta_1 \cdots \beta_{i-1})/(\alpha_1\alpha_2 \cdots \alpha_i)$ pour $i \geq 1$ est invariante, c'est exactement la même preuve. Comme m n'est pas sommable, le PMS n'est pas récurrent positif.

Un contre-exemple: Prenons $\alpha_n = 3^n, \beta_n = 2 \cdot 3^n$. Alors $\sum_{n=1}^{\infty} \frac{\alpha_1 \cdots \alpha_n}{\beta_1 \cdots \beta_n} = \sum \frac{1}{2^n} < +\infty$, donc le processus n'est pas récurrent. Pourtant, $S < \infty$ et la probabilité π du point (ii) vérifie bien $\pi A = 0$. Elle n'est pas invariante, puisque sinon, le processus serait récurrent positif.

Chapitre 6

Files d'attente et autres

Nous allons étudier quelques modèles de files d'attentes, et quelques autres modèles. L'objectif principal est de comprendre comment trouver le générateur, i.e. les taux d'évènements $(\lambda(i))_{i\in E}$ et la matrice $(Q(i,j))_{i,j\in E}$ dans des situations concrètes.

Une file d'attente est constituée de **clients** qui arrivent de l'extérieur pour rejoindre cette file, de **guichets** où les clients vont se faire servir par des **serveurs**. Dans certains cas les clients attendent dans une salle d'attente de capacité limitée. Un client servi disparaît. Les instants d'arrivée des clients et les temps de service sont aléatoires.

Une file d'attente est décrite par la loi d'interarrivée des clients, la loi des temps de service, le nombre de serveurs, la longueur maximale de la file (égale à la taille de la salle d'attente éventuelle). Nous supposerons toujours ici que les interarrivées sont des variables aléatoires indépendantes et de même loi, indépendantes des temps de service, eux mêmes indépendants et de même loi. Pour les files simples, on utilise les notations de Kendall:

Loi d'interarrivée / Loi de service / Nombre de serveurs (/ Longueur maximale).

Les lois sont notées symboliquement: M lorsqu'elles sont exponentielles (M pour Markov), G (G pour Général) sinon. On ne spécifie pas la longueur maximale de la file lorsqu'elle est infinie. Par exemple une file $M/M/s$ est une file d'attente à s guichets, telle que le flot d'arrivée des clients est poissonien et les temps de service exponentiels, sans restriction sur la taille de la file d'attente. Nous nous limiterons à l'étude de files markoviennes.

La question essentielle est de savoir si la taille de la file va exploser ou au contraire s'équilibrer. Dans ce dernier cas, il peut être intéressant de calculer la taille moyenne de la file, la loi du temps d'attente d'un client, etc...

6.1 La file $M/M/1$.

Il y a 1 serveur. Les interarrivées sont des v.a. exponentielles de paramètre α et les temps de service des v.a. exponentielles de paramètre μ, toutes indépendantes. On note

$$X_t = \text{nombre de clients dans file+caisse à l'instant } t.$$

Théorème 6.1.1 *Le processus $(X_t)_{t\geq 0}$ est un PMS à valeurs dans $\mathbf{N}$, et on a $\lambda(0) = \alpha$, $Q(0,1) = 1$, $\lambda(i) = \alpha + \mu$ et $Q(i, i-1) = \mu/(\alpha+\mu)$ et $Q(i, i+1) = \alpha/(\alpha+\mu)$ si $i \geq 1$. Dans tous les autres cas, $Q(i,j) = 0$.*

Preuve: Il est possible mais pénible d'écrire une preuve complète de ce théorème. Contentons nous de comprendre pourquoi il est vrai.

Etape 1. C'est un PMS car les lois en jeu sont exponentielles (sans mémoire). Si on connaît $(X_s)_{s\in[0,t]}$, alors ce qui se passe après t ne dépend que de X_t. En effet,

• Si $X_t = 0$, il n'y a aucun client dans la file ni à la caisse. On est en train d'attendre le prochain client, et on sait depuis combien de temps on l'attend. Mais comme la loi interarrivée est exponentielle, la loi du temps qui reste à l'attendre est encore $\mathcal{E}(\alpha)$. Donc le futur après t ne dépend de $(X_s)_{s\in[0,t]}$ que par le fait que $X_t = 0$.

• Si $X_t = i \geq 1$, alors il y a un client au service et $i-1$ clients en train d'attendre. Un client est en train d'arriver, et on sait depuis combien de temps on l'attend. Mais, comme la loi interarrivée est exponentielle, la loi du temps qui reste à l'attendre est encore $\mathcal{E}(\alpha)$. Un client est en train de se faire servir, et on sait depuis combien de temps. Mais, comme la loi du temps de service est exponentielle, la loi du temps qu'il restera au service après t est encore $\mathcal{E}(\mu)$. Ainsi, le futur après t ne dépend de $(X_s)_{s\in[0,t]}$ que par le fait que $X_t = i$.

Noter que tout cela est faux si les lois ne sont pas exponentielles.

Etape 2. On détermine le générateur. On pose donc $T_1 = \inf\{t \geq 0 \ : X_t \neq X_0\}$, et on cherche $Q(i,j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0 = i$.

Si $i = 0$, soit T le temps d'arrivée du prochain client. Par hypothèse, $T \sim \mathcal{E}(\alpha)$. On a forcément $T_1 = T$ et $X_{T_1} = 1$. Du coup, on a bien $Q(0,1) = 1$ et $\lambda(0) = \alpha$.

Si $i \geq 1$, alors on a $i-1$ clients dans la file et 1 au service, soit S son temps de service, on a $S \sim \mathcal{E}(\mu)$. Soit aussi T le temps d'arrivée du prochain client, on a $T \sim \mathcal{E}(\alpha)$. Ainsi, $T_1 = \min\{S,T\}$, et $X_{T_1} = i+1$ si $T < S$ et $X_{T_1} = i-1$ si $S < T$. Par le lemme des deux réveils, on voit que $T_1 \sim \mathcal{E}(\alpha+\mu)$, donc $\lambda(i) = 1/\mathbf{E}_i[T_1] = \alpha+\mu$ et que $Q(i,i+1) = \mathbf{P}_i(X_{T_1} = i+1) = \mathbf{P}(T < S) = \alpha/(\alpha+\mu)$ et que $Q(i,i-1) = \mathbf{P}_i(X_{T_1} = i-1) = \mathbf{P}(S < T) = \mu/(\alpha+\mu)$.

Proposition 6.1.2 *Le PMS $(X_t)_{t\geq 0}$ est récurrent ssi $\alpha \leq \mu$. Il est récurrent positif ssi $\alpha < \mu$ et sa probabilité invariante est alors donnée par $\pi(i) = (1-\alpha/\mu)(\alpha/\mu)^i$ pour $i \geq 0$.*

C'est assez logique : le processus est récurrent s'il repasse par 0 infiniment souvent. Autrement dit, la file n'explose pas, elle est vide régulièrement : il faut pour cela que le taux d'arrivée α (i.e. le nombre de clients qui arrivent par unité de temps) soit inférieur au taux de départ μ. Et quand il y a égalité, c'est critique : le PMS est récurrent, mais non récurrent positif.

Notons que dans le cas récurrent positif, le nombre de clients dans file+caisse *à l'équilibre* vaut $\sum_{i\geq 0} i\pi(i) = \alpha/(\mu-\alpha)$, qui est d'autant plus petit que μ est grand (logique).

Preuve: On a affaire à un PNM, avec $\alpha_i = \lambda(i)Q(i,i-1) = \mu$ (pour $i \geq 1$) et $\beta_i = \lambda(i)Q(i,i+1) = \alpha$ (pour $i \geq 0$). Donc le processus est récurrent ssi

$$\sum_{i\geq 1} \frac{\alpha_1\cdots\alpha_i}{\beta_1\cdots\beta_i} = \sum_{i\geq 1}\left(\frac{\mu}{\alpha}\right)^i = \infty,$$

i.e. si $\mu \geq \alpha$. Il est alors récurrent positif si

$$S = 1 + \sum_{i\geq 1} \frac{\beta_0\cdots\beta_{i-1}}{\alpha_1\cdots\alpha_i} = 1 + \sum_{i\geq 1}\left(\frac{\alpha}{\mu}\right)^i < \infty,$$

i.e. si $\mu < \alpha$. Alors $S = \mu/(\mu-\alpha)$ et la probabilité invariante est donnée par $\pi(0) = 1/S = (1-\alpha/\mu)$ et $\pi(i) = (\beta_0\cdots\beta_{i-1})/(S\alpha_1\cdots\alpha_i) = (1-\alpha/\mu)(\alpha/\mu)^i$.

Kolmogorov forward. Vérifier que KF s'écrit, pour ce processus, pour tout $i \geq 0$,

$$\begin{cases} P_t'(i,0) = -\alpha P_t(i,0) + \mu P_t(i,1), \\ P_t'(i,j) = -(\alpha+\mu)P_t(i,j) + \alpha P_t(i,j-1) + \mu P_t(i,j+1), \quad j \geq 1. \end{cases}$$

C'est "limpide" : par exemple si $j \geq 1$,

- le terme $-(\alpha+\mu)P_t(i,j)$ exprime qu'on part de j avec taux $\alpha+\mu$;
- le terme $+\alpha P_t(i,j-1)$ exprime qu'on passe de $j-1$ à j à taux α;
- le terme $+\mu P_t(i,j+1)$ exprime qu'on passe de $j+1$ à j à taux μ.

On peut s'amuser à sommer (sans trop se poser de questions) les égalités ci-dessus. Par exemple, en utilisant que $\mathbf{E}_i[X_t] = \sum_{j\geq 1} jP_t(i,j)$, on peut écrire

$$\frac{d}{dt}\mathbf{E}_i[X_t] = \alpha - \mu\mathbf{P}_i[X_t \geq 1]. \qquad (*)$$

La formule obtenue n'est pas très intéressante, mais : le terme $+\alpha$ exprime que les clients arrivent à taux α, le terme $-\mu\mathbf{P}_i[X_t \geq 1]$ exprime qu'ils partent à taux μ POURVU qu'il y ait au moins un client présent. Voici comment on établit $(*)$.

$$\begin{aligned}
\frac{d}{dt}\mathbf{E}_i[X_t] &= \sum_{j\geq 1} jP_t'(i,j) \\
&= -(\alpha+\mu)\sum_{j\geq 1} jP_t(i,j) + \alpha\sum_{j\geq 1} jP_t(i,j-1) + \mu\sum_{j\geq 1} jP_t(i,j+1) \\
&= -(\alpha+\mu)\mathbf{E}_i[X_t] + \alpha\sum_{\ell\geq 0}(\ell+1)P_t(i,\ell) + \mu\sum_{\ell\geq 2}(\ell-1)P_t(i,\ell) \\
&= -(\alpha+\mu)\mathbf{E}_i[X_t] + \alpha\mathbf{E}_i[X_t+1] + \mu\mathbf{E}_i[(X_t-1)\mathbf{1}_{\{X_t\geq 2\}}] \\
&= \alpha - \mu\mathbf{E}_i[X_t - (X_t-1)\mathbf{1}_{\{X_t\geq 2\}}].
\end{aligned}$$

Il n'y a plus qu'à vérifier que pour tout $n \in \mathbf{N}$, on a bien $n-(n-1)\mathbf{1}_{\{n\geq 2\}} = \mathbf{1}_{\{n\geq 1\}}$. C'est vrai si $n=0$, si $n=1$, et si $n \geq 2$. Ainsi, $\mathbf{E}_i[X_t - (X_t-1)\mathbf{1}_{\{X_t\geq 2\}}] = \mathbf{E}_i[\mathbf{1}_{\{X_t\geq 1\}}] = \mathbf{P}_i(X_t \geq 1)$.

6.2 La file $M/M/s$.

Il y a $s \geq 1$ serveurs. Les interarrivées sont des v.a. exponentielles de paramètre α et les temps de service des v.a. exponentielles de paramètre μ, toutes indépendantes. On note

$$X_t = \text{nombre de clients dans file+caisse à l'instant } t.$$

Théorème 6.2.1 *Le processus $(X_t)_{t\geq 0}$ est un PMS à valeurs dans $\mathbf{N}$. De plus, on a pour tout $i \geq 0$, $\lambda(i) = \alpha + \min\{i, s\}\mu$ et*

$$Q(i, i-1) = \frac{\min\{i, s\}\mu}{\alpha + \min\{i, s\}\mu} \quad \text{et} \quad Q(i, i+1) = \frac{\alpha}{\alpha + \min\{i, s\}\mu}.$$

Noter que $Q(0, -1)$ (qui n'existe pas vraiment puisque $E = \mathbf{N}$) vaut bien 0. Dans tous les autres cas, $Q(i, j) = 0$.

Preuve:

Etape 1. C'est un PMS car les lois en jeu sont exponentielles (sans mémoire). Si on connaît $(X_s)_{s\in[0,t]}$, alors ce qui se passe après t ne dépend que de X_t. En effet,

si $X_t = i \geq 0$, il y a précisément $k = \min\{s, i\}$ clients en train de se faire servir et $i - k$ (qui est nul si $i \leq s$) en train d'attendre. Pour chaque client au service, on sait depuis quand il est arrivé, quand il a commencé à se faire servir, mais cela n'apporte aucune information : comme la loi de service $\mathcal{E}(\mu)$ est sans mémoire, son temps de service résiduel est encore de loi $\mathcal{E}(\mu)$. Un client est en train d'arriver, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\mathcal{E}(\alpha)$ étant sans mémoire, notre temps résiduel d'attente sera encore de loi $\mathcal{E}(\alpha)$. Le futur après t ne dépend donc de $(X_s)_{s\in[0,t]}$ que par le fait que $X_t = i$.

Etape 2. On détermine le générateur. On pose donc $T_1 = \inf\{t \geq 0 \ : X_t \neq X_0\}$, et on cherche $Q(i, j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0 = i$.

Si $i = 0$, soit T le temps d'arrivée du prochain client. Par hypothèse, $T \sim \mathcal{E}(\alpha)$. On a forcément $T_1 = T$ et $X_{T_1} = 1$. Du coup, on a bien $\lambda(0) = \alpha = \alpha + \min\{0, s\}\mu$ et $Q(0, 1) = 1$.

Si $i \geq 1$, il y a $k = \min\{s, i\}$ clients au service, notons $S_1, \ldots, S_k$ leurs temps de service, qui sont i.i.d. $\mathcal{E}(\mu)$. Il y a un client en train d'arriver, notons $T \sim \mathcal{E}(\alpha)$ son temps d'arrivée. Notons $S = \min\{S_1, \ldots, S_k\}$, dont la loi est $\mathcal{E}(k\mu)$ (par le lemme des k réveils). Alors $T_1 = \min\{S, T\}$ et on a $X_{T_1} = i + 1$ si $T < S$ et $X_{T_1} = i - 1$ si $S < T$. Par le lemme des deux réveils, on voit que $T_1 \sim \mathcal{E}(\alpha + k\mu)$, donc $\lambda(i) = 1/\mathbf{E}_i[T_1] = \alpha + k\mu$ et que $Q(i, i+1) = \mathbf{P}_i(X_{T_1} = i+1) = \mathbf{P}(T < S) = \alpha/(\alpha + k\mu)$ et que $Q(i, i-1) = \mathbf{P}_i(X_{T_1} = i-1) = \mathbf{P}(S < T) = k\mu/(\alpha + k\mu)$.

En utilisant les résultats généraux sur les PNM et qu'on a ici $\alpha_i = \lambda(i)Q(i, i-1) = \min\{i, s\}\mu$ (pour $i \geq 1$) et $\beta_i = \lambda(i)Q(i, i+1) = \alpha$ (pour $i \geq 0$), on peut montrer la proposition suivante.

Proposition 6.2.2 *Le PMS $(X_t)_{t\geq 0}$ est récurrent ssi $\alpha \leq s\mu$. Il est récurrent positif ssi $\alpha < s\mu$ et sa probabilité invariante est moche mais calculable explicitement.*

Le processus est récurrent si la file n'explose pas : il faut pour cela que le taux d'arrivée α soit inférieur au taux de départ maximal μs (taux de départ par serveur multiplié par le nombre de serveurs).

Kolmogorov forward. Vérifier que KF s'écrit, pour ce processus, pour tout $i \geq 0$,

$$\begin{cases} P'_t(i,0) = -\alpha P_t(i,0) + \mu P_t(i,1), \\ P'_t(i,j) = -(\alpha + \mu \min\{j,s\})P_t(i,j) + \alpha P_t(i,j-1) + \mu \min\{j+1,s\}P_t(i,j+1), \end{cases}$$

la seconde ligne étant valide pour tout $j \geq 1$. C'est "limpide" : par exemple si $j \geq 1$,

- le terme $-(\alpha + \mu \min\{j,s\})P_t(i,j)$ exprime qu'on part de j avec taux $\alpha + \mu \min\{j,s\}$;
- le terme $+\alpha P_t(i,j-1)$ exprime qu'on passe de $j-1$ à j à taux α;
- le terme $+\mu \min\{j+1,s\}P_t(i,j+1)$ exprime qu'on passe de $j+1$ à j à taux $\mu \min\{j+1,s\}$.

On peut s'amuser à sommer (sans trop se poser de questions) les égalités ci-dessus. Par exemple, en utilisant que $\mathbf{E}_i[X_t] = \sum_{j\geq 1} jP_t(i,j)$, on peut écrire

$$\frac{d}{dt}\mathbf{E}_i[X_t] = \sum_{j\geq 1} jP'_t(i,j) = \cdots = \alpha - \mu\mathbf{E}_i[\min\{X_t,s\}].$$

Le calcul et l'interprétation sont laissés au lecteur.

6.3 La file $M/M/\infty$.

Il y a un nombre infini de serveurs. Tout client est immédiatement servi et personne n'attend. Les interarrivées sont des v.a. exponentielles de paramètre α et les temps de service des v.a. exponentielles de paramètre μ, toutes indépendantes. On note

$$X_t = \text{nombre de clients dans file+caisse à l'instant } t.$$

Théorème 6.3.1 *Le processus $(X_t)_{t\geq 0}$ est un PMS à valeurs dans $\mathbf{N}$, et on a, pour tout $i \geq 0$, $\lambda(i) = \alpha + i\mu$ et*

$$Q(i,i-1) = \frac{i\mu}{\alpha + i\mu} \quad \textit{et} \quad Q(i,i+1) = \frac{\alpha}{\alpha + i\mu}.$$

Noter que $Q(0,-1)$ (qui n'existe pas vraiment puisque $E = \mathbf{N}$) vaut bien 0. Dans tous les autres cas, $Q(i,j) = 0$.

Preuve:

Etape 1. C'est un PMS car les lois en jeu sont exponentielles (sans mémoire). Si on connaît $(X_s)_{s\in[0,t]}$, alors ce qui se passe après t ne dépend que de X_t. En effet,

Si $X_t = i$, alors il y a i clients au service. Pour chaque client au service, on sait depuis quand il est arrivé, quand il a commencé à se faire servir, mais cela n'apporte aucune information : comme la loi de service $\mathcal{E}(\mu)$ est sans mémoire, son temps de service résiduel est encore de loi $\mathcal{E}(\mu)$. Un client est en train d'arriver, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\mathcal{E}(\alpha)$ étant sans mémoire, notre temps résiduel d'attente sera encore de loi $\mathcal{E}(\alpha)$. Le futur après t ne dépend donc de $(X_s)_{s\in[0,t]}$ que par le fait que $X_t = i$.

Etape 2. On détermine le générateur. On pose donc $T_1 = \inf\{t \geq 0 \ : X_t \neq X_0\}$, et on cherche $Q(i,j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0 = i$.

Si $i = 0$, soit T le temps d'arrivée du prochain client. Par hypothèse, $T \sim \mathcal{E}(\alpha)$. On a forcément $T_1 = T$ et $X_{T_1} = 1$. Du coup, on a bien $\lambda(0) = \alpha = \alpha + i\mu$ et $Q(0,1) = 1$.

Si $i \geq 1$, il y a i clients au service, notons $S_1, \dots, S_i$ leurs temps de service, qui sont i.i.d. $\mathcal{E}(\mu)$. Il y a un client en train d'arriver, notons $T \sim \mathcal{E}(\alpha)$ son temps d'arrivée. Notons $S = \min\{S_1, \dots, S_i\}$, dont la loi est $\mathcal{E}(i\mu)$ (par le lemme des i réveils). Alors $T_1 = \min\{S, T\}$ et on a $X_{T_1} = i+1$ si $T < S$ et $X_{T_1} = i-1$ si $S < T$. Par le lemme des deux réveils, on voit que $T_1 \sim \mathcal{E}(\alpha + i\mu)$, donc $\lambda(i) = 1/\mathbf{E}_i[T_1] = \alpha + i\mu$ et, enfin, que $Q(i, i+1) = \mathbf{P}_i(X_{T_1} = i+1) = \mathbf{P}(T < S) = \alpha/(\alpha + i\mu)$ et que $Q(i, i-1) = \mathbf{P}_i(X_{T_1} = i-1) = \mathbf{P}(S < T) = i\mu/(\alpha + i\mu)$.

Proposition 6.3.2 *Le PMS $(X_t)_{t\geq 0}$ est récurrent positif, et sa probabilité invariante est donnée par la loi de Poisson $\mathcal{P}(\alpha/\mu)$.*

Preuve: On a affaire à un PNM, avec $\alpha_i = \lambda(i)Q(i,i-1) = i\mu$ (pour $i \geq 1$) et $\beta_i = \lambda(i)Q(i,i+1) = \alpha$ (pour $i \geq 0$). Donc le processus est récurrent ssi

$$\sum_{i\geq 1} \frac{\alpha_1 \cdots \alpha_i}{\beta_1 \cdots \beta_i} = \sum_{i\geq 1} \frac{\mu^i i!}{\alpha^i} = \infty,$$

ce qui est toujours le cas. Il est alors récurrent positif si

$$S = 1 + \sum_{i\geq 1} \frac{\beta_0 \cdots \beta_{i-1}}{\alpha_1 \cdots \alpha_i} = 1 + \sum_{i\geq 1} \frac{\alpha^i}{\mu^i i!} < \infty,$$

ce qui est toujours le cas, et on a $S = e^{\alpha/\mu}$. La probabilité invariante est donnée par $\pi(0) = 1/S = e^{-\alpha/\mu}$ et $\pi(i) = (\beta_0 \cdots \beta_{i-1})/(S\alpha_1 \cdots \alpha_i) = e^{-\alpha/\mu}\frac{\alpha^i}{\mu^i i!}$.

Kolmogorov forward. Vérifier que KF s'écrit, pour ce processus, pour tout $i \geq 0$,

$$\begin{cases} P_t'(i,0) = -\alpha P_t(i,0) + \mu P_t(i,1), \\ P_t'(i,j) = -(\alpha + \mu j)P_t(i,j) + \alpha P_t(i,j-1) + \mu(j+1)P_t(i,j+1), \quad j \geq 1. \end{cases}$$

En en utilisant que $\mathbf{E}_i[X_t] = \sum_{j\geq 1} jP_t(i,j)$, on peut écrire

$$\frac{d}{dt}\mathbf{E}_i[X_t] = \alpha - \mu\mathbf{E}_i[X_t]. \qquad (*)$$

En effet,

$$\begin{aligned}
\frac{d}{dt}\mathbf{E}_i[X_t] &= \sum_{j\geq 1} jP_t'(i,j) \\
&= -\sum_{j\geq 1}(\alpha+\mu j)jP_t(i,j) + \alpha\sum_{j\geq 1} jP_t(i,j-1) + \mu\sum_{j\geq 1} j(j+1)P_t(i,j+1) \\
&= -\mathbf{E}_i[(\alpha+\mu X_t)X_t] + \alpha\sum_{\ell\geq 0}(\ell+1)P_t(i,\ell) + \mu\sum_{\ell\geq 2}(\ell-1)\ell P_t(i,\ell) \\
&= -\mathbf{E}_i[(\alpha+\mu X_t)X_t] + \alpha\mathbf{E}_i[X_t+1] + \mu\mathbf{E}_i[(X_t-1)X_t\mathbf{1}_{\{X_t\geq 2\}}] \\
&= \alpha - \mu\mathbf{E}_i[X_t^2 - (X_t-1)X_t\mathbf{1}_{\{X_t\geq 2\}}] \\
&= \alpha - \mu\mathbf{E}_i[X_t].
\end{aligned}$$

La dernière ligne utilise que pour tout $n \in \mathbf{N}$, on a $n^2 - (n-1)n\mathbf{1}_{\{n\geq 2\}} = n$. En utilisant que $\mathbf{E}_i[X_0] = i$, et en résolvant $(*)$, on trouve que

$$\mathbf{E}_i[X_t] = \frac{\alpha}{\mu} + \Big(i - \frac{\alpha}{\mu}\Big)e^{-\mu t}.$$

On pourra comparer $\lim_{t\to\infty}\mathbf{E}_i[X_t]$ avec la moyenne de la probabilité invariante du processus.

6.4 La file $M/M/1/k$.

Il y a 1 serveur. Les interarrivées sont des v.a. exponentielles de paramètre α et les temps de service des v.a. exponentielles de paramètre μ, toutes indépendantes. Par contre, la salle d'attente est limitée : si un client qui arrive trouve k personnes dans la salle d'attente, il s'en va immédiatement. On note

$$X_t = \text{nombre de clients dans file+caisse à l'instant } t.$$

Théorème 6.4.1 *Le processus $(X_t)_{t\geq 0}$ est un PMS à valeurs dans $\{0,\cdots,k+1\}$, et on a*

$$\lambda(i) = \begin{cases} \alpha & si \quad i = 0 \\ \alpha+\mu & si \quad i \in \{1,\cdots,k\} \\ \mu & si \quad i = k+1 \end{cases}$$

et

$$Q(i,j) = \begin{cases} 1 & si \quad i=0,\ j=1 \\ \dfrac{\alpha}{\alpha+\mu} & si \quad i\in\{1,\cdots,k\},\ j=i+1 \\ \dfrac{\mu}{\alpha+\mu} & si \quad i\in\{1,\cdots,k\},\ j=i-1 \\ 1 & si \quad i=k+1,\ j=k. \end{cases}$$

Dans tous les autres cas, $Q(i,j) = 0$.

Preuve:

Etape 1. C'est un PMS car les lois en jeu sont exponentielles (sans mémoire). Si on connaît $(X_s)_{s\in[0,t]}$, ce qui se passe après t ne dépend que de X_t. L'argumentaire est laissé au lecteur. Distinguer les 3 cas $X_t = 0$, $X_t \in \{1, \dots, k\}$, et $X_t = k + 1$.

Etape 2. On détermine le générateur. On pose donc $T_1 = \inf\{t \geq 0\ : X_t \neq X_0\}$, et on cherche $Q(i,j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0 = i$.

Si $i = 0$, soit T le temps d'arrivée du prochain client. Par hypothèse, $T \sim \mathcal{E}(\alpha)$. On a forcément $T_1 = T$ et $X_{T_1} = 1$. Du coup, on a bien $Q(0,1) = 1$ et $\lambda(0) = \alpha$.

Si $i = k + 1$. Aucun client ne peut arriver (sans être rejeté). Il y a un client au service, soit S son temps de service, qui suit la loi $\mathcal{E}(\mu)$. On a bien sûr $S \sim \mathcal{E}(\mu)$. Forcément, $T_1 = S$ et $X_{T_1} = k$. Du coup, on a bien $Q(k+1,k) = 1$ et $\lambda(k+1) = \mu$.

Si $i \in \{1, \cdots, k\}$, alors on a $i-1$ clients dans la file et 1 au service, soit S son temps de service, on a $S \sim \mathcal{E}(\mu)$. Soit aussi T le temps d'arrivée du prochain client, on a $T \sim \mathcal{E}(\alpha)$. Ainsi, $T_1 = \min\{S, T\}$, et $X_{T_1} = i + 1$ si $T < S$ et $X_{T_1} = i - 1$ si $S < T$. Par le lemme des deux réveils, on voit que $T_1 \sim \mathcal{E}(\alpha + \mu)$, donc $\lambda(i) = 1/\mathbf{E}_i[T_1] = \alpha + \mu$ et que $Q(i, i+1) = \mathbf{P}_i(X_{T_1} = i+1) = \mathbf{P}(T < S) = \alpha/(\alpha+\mu)$ et que $Q(i, i-1) = \mathbf{P}_i(X_{T_1} = i-1) = \mathbf{P}(S < T) = \mu/(\alpha+\mu)$.

6.5 Une file un peu plus compliquée

Des clients arrivent dans un magasin suivant un processus de Poisson de paramètre $\alpha > 0$ (autrement dit, les interarrivées sont $\mathcal{E}(\alpha)$). Ce magasin dispose de 3 employés : un vendeur, un caissier et un manutentionnaire. Quand un client arrive, il doit se faire servir par le vendeur (temps de service de loi exponentielle de paramètre $\mu_1 > 0$), puis par le caissier (temps de service de loi exponentielle de paramètre $\mu_2 > 0$), puis par le manutentionnaire (temps de service de loi exponentielle de paramètre $\mu_3 > 0$). Les objets aléatoires ci-dessus sont bien sûr mutuellement indépendants.

Soit $X_t = (X_t^1, X_t^2, X_t^3)$ où X_t^1 (respectivement X_t^2, X_t^3) est le nombre de clients en train d'attendre ou de se faire servir par le vendeur (respectivement le caissier, le manutentionnaire) à l'instant $t \geq 0$.

Théorème 6.5.1 *Le processus $(X_t)_{t\geq 0}$ est un PMS à valeurs dans $E = \mathbf{N}^3$. De plus, on a, pour tout $i = (i_1, i_2, i_3)$,*

$$\lambda(i) = \alpha + \min\{i_1, 1\}\mu_1 + \min\{i_2, 1\}\mu_2 + \min\{i_3, 1\}\mu_3,$$

et, pour $i=(i_1,i_2,i_3)\in\mathbf{N}^3$ *et* $j=(j_1,j_2,j_3)\in\mathbf{N}^3$,

$$Q(i,j)=\begin{cases}\dfrac{\alpha}{\lambda(i)} & si\quad j=(i_1+1,i_2,i_3),\\ \dfrac{\min\{i_1,1\}\mu_1}{\lambda(i)} & si\quad j=(i_1-1,i_2+1,i_3),\\ \dfrac{\min\{i_2,1\}\mu_2}{\lambda(i)} & si\quad j=(i_1,i_2-1,i_3+1),\\ \dfrac{\min\{i_3,1\}\mu_3}{\lambda(i)} & si\quad j=(i_1,i_2,i_3-1).\end{cases}$$

Et $Q(i,j)=0$ *dans tous les autres cas.*

Preuve:

Etape 1. C'est un PMS car les lois en jeu sont toutes exponentielles (sans mémoire). Si on connaît $(X_s)_{s\in[0,t]}$, alors ce qui se passe après t ne dépend que de X_t. En effet, supposons que $X_t=i=(i_1,i_2,i_3)$. Alors on sait qu'il y a un client en train d'arriver, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\mathcal{E}(\alpha)$ étant sans mémoire, notre temps résiduel d'attente sera encore de loi $\mathcal{E}(\alpha)$. D'autre part, il y a $\min\{i_1,1\}$ client en train de se faire servir par le vendeur, on sait depuis combien de temps, etc. Mais, son temps de service étant $\mathcal{E}(\mu_1)$, son temps résiduel de service, à l'instant t, est encore de loi $\mathcal{E}(\mu_1)$. Il y a aussi $\min\{i_2,1\}$ client en train de se faire servir par le caissier, on sait depuis combien de temps, etc. Mais, son temps de service étant $\mathcal{E}(\mu_2)$, son temps résiduel de service, à l'instant t, est encore de loi $\mathcal{E}(\mu_2)$. Enfin, il y a $\min\{i_3,1\}$ client en train de se faire servir par le manutentionnaire, on sait depuis combien de temps, etc. Mais, son temps de service étant $\mathcal{E}(\mu_3)$, son temps résiduel de service, à l'instant t, est encore de loi $\mathcal{E}(\mu_3)$. Ainsi, la connaissance de $(X_s)_{s\in[0,t]}$ n'apporte rien de plus que celle de X_t pour le comprtement à venir.

Etape 2. On détermine le générateur. On pose donc $T_1=\inf\{t\geq 0\ :X_t\neq X_0\}$, et on cherche $Q(i,j)=\mathbf{P}_i(X_{T_1}=j)$ et $\lambda(i)=1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0=i=(i_1,i_2,i_3)$.

• Soit $T\sim\mathcal{E}(\alpha)$ le temps d'arrivée du prochain client.

• Si $i_1\geq 1$, soit $S_1\sim\mathcal{E}(\mu_1)$ le temps de service du client en train de se faire servir par le vendeur. Sinon, on pose $S_1=\infty$. Donc dans tous les cas, $S_1\sim\mathcal{E}(\min\{i_1,1\}\mu_1)$. (On rappelle qu'une v.a. de loi $\mathcal{E}(0)$ est une v.a. infinie p.s.).

• Si $i_2\geq 1$, soit $S_2\sim\mathcal{E}(\mu_2)$ le temps de service du client en train de se faire servir par le caissier. Sinon, on pose $S_2=\infty$. Donc dans tous les cas, $S_2\sim\mathcal{E}(\min\{i_2,1\}\mu_2)$.

• Si $i_3\geq 1$, soit $S_3\sim\mathcal{E}(\mu_3)$ le temps de service du client en train de se faire servir par le manutentionnaire. Sinon, on pose $S_3=\infty$. Donc dans tous les cas, $S_3\sim\mathcal{E}(\min\{i_3,1\}\mu_3)$.

On a donc $T_1=\min\{T,S_1,S_2,S_3\}$ et $X_{T_1}=(i_1+1,i_2,i_3)$ si $T_1=T$, $X_{T_1}=(i_1-1,i_2+1,i_3)$ si $T_1=S_1$, $X_{T_1}=(i_1,i_2-1,i_3+1)$ si $T_1=S_2$, $X_{T_1}=(i_1,i_2,i_3-1)$

si $T_1 = S_3$ (ouf !). Par le lemme des 4 réveils, on conclut qu'en effet, on a $T_1 \sim \mathcal{E}(\alpha + \min\{i_1, 1\}\mu_1 + \min\{i_2, 1\}\mu_2 + \min\{i_3, 1\}\mu_3)$, donc

$$\lambda(i) = \alpha + \min\{i_1, 1\}\mu_1 + \min\{i_2, 1\}\mu_2 + \min\{i_3, 1\}\mu_3,$$

et que

$$Q(i, (i_1 + 1, i_2, i_3)) = \mathbf{P}_i(X_{T_1} = (i_1 + 1, i_2, i_3)) = \mathbf{P}(T_1 = T) = \frac{\alpha}{\lambda(i)},$$

$$Q(i, (i_1 - 1, i_2 + 1, i_3)) = \mathbf{P}_i(X_{T_1} = (i_1 - 1, i_2 + 1, i_3)) = \mathbf{P}(T_1 = S_1) = \frac{\min\{i_1, 1\}\mu_1}{\lambda(i)},$$

etc.

6.6 Un modèle simple de populations

Considérons une population d'individus assexués. La durée de vie de chaque individu suit une loi exponentielle de paramètre μ. D'autre part, chaque individu (vivant) produit des portées suivant un processus de Poisson de paramètre α. La loi du nombre d'individus issus d'une portée est donnée par une probabilité ν sur $\mathbf{N}^*$. Enfin, des individus immigrent suivant un processus de poisson de paramètre κ.

Soit X_t le nombre d'individus vivant à l'instant $t \geq 0$.

Théorème 6.6.1 *Le processus $(X_t)_{t\geq 0}$ est un PMS à valeurs dans $E = \mathbf{N}$. De plus, on a, pour tout $i \in \mathbf{N}$,*

$$\lambda(i) = \kappa + i\mu + i\alpha$$

et, pour $i = \mathbf{N}$ et $j \in \mathbf{N}$,

$$Q(i,j) = \begin{cases} \dfrac{i\mu}{\lambda(i)} & si \quad j = i - 1, \\ \dfrac{\kappa + \alpha\nu(1)}{\lambda(i)} & si \quad j = i + 1, \\ \dfrac{i\alpha\nu(k)}{\lambda(i)} & si \quad j = i + k \ (avec\ k \geq 2), \end{cases}$$

et $Q(i,j) = 0$ dans tous les autres cas.

Preuve:

Etape 1. C'est un PMS car les lois en jeu sont toutes exponentielles (sans mémoire). Si on connaît $(X_s)_{s\in[0,t]}$, alors ce qui se passe après t ne dépend que de X_t. En effet, supposons que $X_t = i$. Alors on sait qu'un immigré arrive, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\mathcal{E}(\kappa)$ étant sans mémoire, notre temps résiduel d'attente sera encore de loi $\mathcal{E}(\kappa)$. D'autre part, il y a i individus en train de mourir, et on sait quand ils sont nés, mais, comme leur durée de vie suit une loi exponentielle, leur durée de vie résiduelle à l'instant t est encore $\mathcal{E}(\mu)$.

Enfin, i individus vont produire (potentiellement) une portée. Pour chacun, on sait son âge et la date de sa dernière portée éventuelle, mais, les portées arrivant suivant des processus de Poisson de paramètre α, sa prochaine portée se produira dans un temps de loi $\mathcal{E}(\alpha)$. Enfin, la taille de cette portée ne dépend de rien.

Ainsi, la connaissance de $(X_s)_{s\in[0,t]}$ n'apporte rien de plus que celle de X_t pour le comportement à venir.

Etape 2. On détermine le générateur. On pose donc $T_1 = \inf\{t \geq 0 \ : X_t \neq X_0\}$, et on cherche $Q(i,j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0 = i$.

• Soit $T \sim \mathcal{E}(\kappa)$ le temps d'arrivée du prochain immigré.

• Soient $M_1, \dots, M_i$ les dates de mort des i individus, on sait qu'elles sont i.i.d. de loi $\mathcal{E}(\mu)$.

• Soient $P_1, \dots, P_i$ les dates des premières portées des i individus, on sait qu'elles sont i.i.d. de loi $\mathcal{E}(\alpha)$. Enfin, soient $Z_1, \dots, Z_i$ les tailles de ces portées, qui sont i.i.d. de loi ν.

On a donc $T_1 = \min\{T, M_1, \dots, M_i, P_1, \dots, P_i\}$ et par le lemme des $2i+1$ réveils, on a donc $T_1 \sim \mathcal{E}(\kappa + i\mu + i\alpha)$, soit $\lambda(i) = \kappa + i\mu + i\alpha$.

De plus, $X_{T_1} = i+1$ si $T_1 = T$, $X_{T_1} = i-1$ si $T_1 \in \{M_1, \dots, M_i\}$, et enfin, pour tout $j = 1, \dots, i$, $X_{T_1} = i + Z_j$ si $T_1 = P_j$.

Donc $X_{T_1} = i-1$ si $T_1 \in \{M_1, \dots, M_i\}$, et par le lemme des réveils,

$$Q(i, i-1) = \sum_{j=1}^{i} \mathbf{P}(T_1 = M_j) = \sum_{j=1}^{i} \frac{\mu}{\lambda(i)} = \frac{i\mu}{\lambda(i)}.$$

Aussi, $X_{T_1} = i+1$ si $T_1 = T$ ou si, pour un $j \in \{1, \dots, i\}$, $T_1 = P_j$ et $Z_j = 1$, donc

$$Q(i, i+1) = \mathbf{P}(T_1 = T) + \sum_{j=1}^{i} \mathbf{P}(T_1 = P_j, Z_j = 1) = \frac{\kappa}{\lambda(i)} + \sum_{j=1}^{i} \frac{\alpha}{\lambda(i)}\nu(1) = \frac{\kappa + i\alpha\nu(1)}{\lambda(i)}.$$

Enfin, pour $k \geq 2$, $X_{T_1} = i+k$ si, pour un $j \in \{1, \dots, i\}$, $T_1 = P_j$ et $Z_j = k$, donc

$$Q(i, i+k) = \sum_{j=1}^{i} \mathbf{P}(T_1 = P_j, Z_j = k) = \sum_{j=1}^{i} \frac{\alpha}{\lambda(i)}\nu(k) = \frac{i\alpha\nu(k)}{\lambda(i)}.$$

Kolmogorov forward. Vérifier que KF s'écrit, pour ce processus, pour tout $i \geq 0$, tout $j \geq 1$ (inutile d'écrire $P_t'(i,0)$)

$$\begin{aligned} P_t'(i,j) &= -(\kappa + j\mu + j\alpha)P_t(i,j) + \mu(j+1)P_t(i,j+1) + \kappa P_t(i,j-1) \\ &\quad + \sum_{k=1}^{j} \alpha(j-k)\nu(k)P_t(i,j-k). \end{aligned}$$

En en utilisant que $\mathbf{E}_i[X_t] = \sum_{j\geq 1} jP_t(i,j)$, on trouve

$$\frac{d}{dt}\mathbf{E}_i[X_t] = \kappa + (\alpha\rho - \mu)\mathbf{E}_i[X_t], \qquad (*)$$

où $\rho = \sum_{k\geq 1} k\nu(k)$. C'est logique (essayer d'interpréter). En effet,

$$\frac{d}{dt}\mathbf{E}_i[X_t] = \sum_{j\geq 1} jP_t'(i,j) = -A + B + C + D,$$

où

$$A = \sum_{j\geq 1}(\kappa + j\mu + j\alpha)jP_t(i,j) = \mathbf{E}_i[(\kappa + \mu X_t + \alpha X_t)X_t],$$

$$\begin{aligned} B &= \mu\sum_{j\geq 1}(j+1)jP_t(i,j+1) = \mu\sum_{\ell\geq 2}\ell(\ell-1)P_t(i,\ell) = \mu\mathbf{E}_i[X_t(X_t-1)\mathbf{1}_{\{X_t\geq 2\}}] \\ &= \mu\mathbf{E}_i[X_t(X_t-1)] \end{aligned}$$

car $n(n-1)\mathbf{1}_{\{n\geq 2\}} = n(n-1)$ pour tout $n \in \mathbf{N}$,

$$C = \kappa\sum_{j\geq 1} jP_t(i,j-1) = \kappa\sum_{\ell\geq 0}(\ell+1)P_t(i,\ell) = \kappa\mathbf{E}_i[X_t+1],$$

et enfin, le plus difficile,

$$\begin{aligned} D &= \alpha\sum_{j\geq 1} j\sum_{k=1}^{j}(j-k)\nu(k)P_t(i,j-k) \\ &= \alpha\sum_{k\geq 1}\sum_{j\geq k} j(j-k)\nu(k)P_t(i,j-k) \\ &= \alpha\sum_{k\geq 1}\sum_{\ell\geq 0}(\ell+k)\ell\nu(k)P_t(i,\ell) \\ &= \alpha\sum_{k\geq 1}\mathbf{E}_i[(X_t+k)X_t\nu(k)] \\ &= \alpha\mathbf{E}_i[X_t^2] + \alpha\rho\mathbf{E}_i[X_t]. \end{aligned}$$

Au final, on a donc

$$\begin{aligned} \frac{d}{dt}\mathbf{E}_i[X_t] &= -\mathbf{E}_i[(\kappa + \mu X_t + \alpha X_t)X_t] + \mu\mathbf{E}_i[X_t(X_t-1)] \\ &\quad + \kappa\mathbf{E}_i[X_t+1] + \alpha\mathbf{E}_i[X_t^2] + \alpha\rho\mathbf{E}_i[X_t] \\ &= \kappa + (\alpha\rho - \mu)\mathbf{E}_i[X_t]. \end{aligned}$$

En utilisant que $\mathbf{E}_i[X_0] = i$, et en résolvant $(*)$, on trouve, si $\alpha\rho \neq \mu$ (sinon, c'est une autre formule explicite)

$$\mathbf{E}_i[X_t] = -\frac{\kappa}{\alpha\rho - \mu} + \Big(i + \frac{\kappa}{\alpha\rho-\mu}\Big)e^{(\alpha\rho-\mu)t}.$$

On trouve trouve donc deux comportements bien différents :

- si $\alpha\rho < \mu$, alors $\lim_{t\to\infty}\mathbf{E}_i[X_t] = \kappa/(\mu - \alpha\rho)$, la population n'explose pas,
- si $\alpha\rho > \mu$, alors $\lim_{t\to\infty}\mathbf{E}_i[X_t] = \infty$.

C'est logique, car $\alpha\rho$ est le taux de naissance des portées multiplié par le nombre moyen de bébés par portée, qu'on compare au taux μ de décès.

Table des matières

1 Marches aléatoires, processus de Bernoulli **3**
1.1 Rappels 3
1.2 Marches aléatoires, processus de Bernoulli 4
1.3 Propriété de Markov forte des marches aléatoires 4
1.4 Loi géométrique 6
1.5 Loi exponentielle 8

2 Processus de Poisson **11**
2.1 P.A.I.S. 11
2.2 Processus de Poisson 13
2.3 Processus ponctuel de Poisson 17
2.4 La file d'attente $M/G/\infty$ 23

3 Processus régénératifs **25**

4 Chaînes de Markov **29**
4.1 Matrices de transition 29
4.2 Chaîne de Markov 30
4.3 Propriété de Markov 32
4.4 Propriétés de récurrence 34
4.5 Mesures et probabilités invariantes 37
4.6 Stationnarité et réversibilité 43
4.7 Processus de naissance et mort à temps discret 44
4.8 Problèmes d'absorption 45

5 Processus markoviens de sauts **49**
5.1 Premières propriétés 49
5.2 Description dynamique 52
5.3 Explosion 55
5.4 Le générateur 56
5.5 Mesures invariantes 59
5.6 Processus de naissance et mort 62

6 Files d'attente et autres **65**
6.1 La file $M/M/1$. . . . 65
6.2 La file $M/M/s$. . . . 68
6.3 La file $M/M/\infty$. . . . 69
6.4 La file $M/M/1/k$. . . . 71
6.5 Une file un peu plus compliquée 72
6.6 Un modèle simple de populations 74