

Chaînes de Markov (et applications)

Raphael Lachieze-Rey*

25 avril 2016

M1 Paris Descartes.

Table des matières

1	Chaînes de Markov homogènes	2
1.1	Exemples et définitions	2
1.2	Loi des X_n	2
2	Temps d'absorption	3
2.1	Temps d'arrêt	3
2.2	Probabilités et temps d'absorptions	4
3	Classification des états	5
3.1	Réurrence et transience	5
4	Distributions invariantes	6
4.1	Convergence à l'équilibre	8
4.2	Théorème ergodique	9
5	Chaînes de Markov et simulation	9
5.1	Algorithme Hit-and-run	9
5.2	Algorithme de Metropolis	10
6	Chaînes de Markov en temps continu, processus de Poisson	11
6.1	Lois sans mémoire	11
6.2	Processus de Poisson	11
6.3	Générateur infinitésimal	12
6.4	Processus de Poisson composé	12

(Ω, \mathbb{P}) est un espace probabilisé.

*lr.raphael@gmail.com

1 Chaînes de Markov homogènes

1.1 Exemples et définitions

Définition 1. Formellement, soit E un espace fini ou dénombrable. Ce sera l'**espace d'états**. Soit $X = \{X_n; n \geq 0\}$ une suite de variables aléatoires à valeurs dans E . On dit que X est une chaîne de Markov si, pour tout $x_1, \dots, x_{n+1} \in E$, on a

$$\mathbb{P}(\underbrace{X_{n+1} = x_{n+1}}_{\text{Le futur}} \mid \underbrace{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n}_{\text{Le passé (et le présent)}}) = \mathbb{P}(\underbrace{X_{n+1} = x_{n+1}}_{\text{Le futur}} \mid \underbrace{X_n = x_n}_{\text{Le présent}})$$

Cette propriété des chaînes de Markov est aussi connue comme **propriété de Markov**.

Définition 2. On dit qu'une matrice Q (éventuellement infinie) est **stochastique** ssi tous ses coefficients sont ≥ 0 et si la somme de chaque ligne fait 1 : $\forall x \in E$,

$$\sum_{y \in E} Q(x, y) = 1.$$

On dit aussi **matrice markovienne**.

Proposition 1. Si Q est la matrice de transition d'une chaîne de Markov, alors elle est stochastique.

Etant donné une matrice stochastique Q , il existe une chaîne de Markov de matrice de transition Q . Etant donné $x \in E$, la suite de variables aléatoires définie récursivement par $X_0 = x$,

$$\mathbb{P}(X_{n+1} = y \mid X_n = x, X_{n-1}, \dots, X_0) = Q(x, y)$$

est une chaîne de Markov de matrice de transition Q .

1.2 Loi des X_n

On appelle μ_0 la **loi initiale** de X , définie par

$$\mu_0(x) = \mathbb{P}(X_0 = x).$$

Connaissant μ_0 et Q , on peut calculer directement la loi de X_n .

Proposition 2. Pour toute suite $\{x_0, x_1, \dots, x_n\}$ dans E , on a

$$\begin{aligned} & \mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ &= \mu_0(x_0)Q(x_0, x_1)Q(x_1, x_2) \dots Q(x_{n-1}, x_n). \end{aligned}$$

Notation 1. Pour une mesure μ_0 et une matrice Q , on note la mesure

$$(\mu_0 Q)(y) = \sum_{x \in E} \mu_0(x) Q(x, y).$$

Cela revient à multiplier (matriciellement) la mesure μ_0 vue comme un vecteur $\mu_0 = (\mu_0(x_1), \mu_0(x_2), \dots)$ par la matrice Q .

Proposition 3. Si μ es la loi de X_0 , alors (μQ) est la loi de X_1 .

Pour une même chaîne X , on considère souvent plusieurs lois initiales différentes. Dans ce cas on précise la loi utilisée en notant

$$\mathbb{P} = \mathbb{P}_\mu$$

dans chaque calcul de probabilité, et l'espérance est alors notée \mathbb{E}_x . Si la loi est un "Dirac" $\mu = \delta_x$ pour un certain $x \in E$ (ce qui veut dire $X_0 = x$ p.s.), alors on note plus simplement $\mathbb{P}_{\delta_x} = \mathbb{P}_x, \mathbb{E}_{\delta_x} = \mathbb{E}_x$.

Proposition 4. Pour tout n , la loi de X_n est μQ^n .

Remarque 1. TRES IMPORTANT!!

$$Q^k(x, y) \neq Q(x, y)^k.$$

membre de gauche : multiplication matricielle.

membre de droite : multiplication de réels (beaucoup plus facile).

Proposition 5. On a pour $n \geq 0, k \geq 0$

$$\mathbb{P}(X_{n+k} = y | X_n = x) = Q^k(x, y)$$

2 Temps d'absorption

2.1 Temps d'arrêt

Pour $x \in E$ on définit le temps aléatoire

$$T_x = \min\{n \geq 0 : X_n = x\},$$

premier moment où la chaîne atteint x .

Définition 3. Soit T une variable aléatoire à valeurs dans \mathbb{N} . T est un **temps d'arrêt** si pour tout n , l'évènement $(T = n)$ dépend uniquement du passé, c'est-à-dire si l'évènement $(T = n)$ est entièrement déterminé par les variables X_1, \dots, X_n (c'est-à-dire mesurable par rapport à $\sigma(X_1, \dots, X_n)$).

Exemple 1. Pour $x \in E$, le temps T_x est un temps d'arrêt : Pour $n \in \mathbb{N}$,

$$(T_x = n) = (X_1 \neq x, X_2 \neq x, \dots, X_{n-1} \neq x, X_n = x).$$

C'est bien un évènement qui est entièrement déterminé si on connaît les valeurs de X_1, \dots, X_n .

Proposition 6 (propriété de Markov forte). Soit $k \geq 1$, et T un temps d'arrêt. Pour $x, y \in E$,

$$\mathbb{P}(X_{T+k} = y | X_T = x) = \mathbb{P}(X_k = y | X_0 = x) = Q^k(x, y).$$

Proposition 7. Une autre manière de formuler la propriété de Markov est la suivante : Pour tout temps d'arrêt T , la chaîne

$$X' = (X'_0 = X_T, X'_1 = X_{T+1}, \dots)$$

est une chaîne de Markov dont la matrice de transition est Q et la loi initiale est X_T . De plus, la loi de X' est indépendante de (X_0, \dots, X_{T-1}) conditionnellement à X_T .

2.2 Probabilités et temps d'absorptions

Avec le langage introduit dans la section précédente, on s'intéresse pour $A \subset E$ aux quantités

$$\begin{aligned} h^A &= \mathbb{P}(T_A < \infty), \\ k^A &= \mathbb{E}(T_A). \end{aligned}$$

Remarquons que si $h_A \neq 1$, $k^A = \infty$, donc il faut calculer h^A en premier, et ensuite k^A si ça a du sens.

Si l'on conditionne par l'état de départ $x \in E$, on a

$$\begin{aligned} h_x^A &= \mathbb{P}_x(T_A < \infty) && \text{Probabilité d'arriver un jour en } A \text{ en partant de } x, \\ k_x^A &= \mathbb{E}_x(T_A) && \text{Temps moyen pour } y \text{ arriver.} \end{aligned}$$

Si $A = \{y\}$ est constitué d'un unique point, on note $h_x^{\{y\}} = h_x^y$, $k_x^{\{y\}} = k_x^y$.

Théorème 1. Si $x \notin A$, pour les calculer exactement il faut se persuader des deux faits suivants :

$$\begin{aligned} h_x^A &= \sum_{y \in E} Q(x, y) h_y^A, \text{ pour tout } x \in E \\ k_x^A &= 1 + \sum_{y \in E} Q(x, y) k_y^A, \text{ pour tout } x \in E. \end{aligned}$$

De plus, si ce système linéaire a plusieurs solutions, $(h_x^A)_x$ (resp. $(k_x^A)_x$) est la plus petite solution positive du système vérifiant $h_x^A = 1$ (resp. $k_x^A = 0$) pour $x \in A$.

3 Classification des états

On dit qu'un état $x \in E$ mène à un état $y \in E$ si

$$\mathbb{P}_x(\exists n \geq 0, X_n = y) > 0.$$

La probabilité de passer par y après être passée par x est non-nulle. On note dans ce cas

$$x \rightsquigarrow y.$$

Si $x \rightsquigarrow y$ et $y \rightsquigarrow x$, on note

$$x \longleftrightarrow y$$

et on dit que x et y communiquent.

il est facile de voir que $x \rightsquigarrow y$ ssi il existe une suite d'états $x_0 = x, x_1, \dots, x_k = y$ qui "mène" de x à y et telle que $Q(x_m, x_{m+1}) > 0$. On appelle un tel chemin un chemin probable. $x \rightsquigarrow y$ ssi $\exists n \geq 0$ tel que $Q^n(x, y) > 0$.

Théorème 2. *La relation \longleftrightarrow est une relation d'équivalence et on peut partitionner E par l'ensemble des classes d'équivalences*

$$E = \cup_{x \in E} C_x$$

avec $C_x = C_y$ si $x \longleftrightarrow y$, et $C_x \cap C_y = \emptyset$ sinon.

3.1 Récurrence et transience

Définition 4. *On rappelle que T_x est le temps de 1er passage en x . Pour $r \geq 0$, on note*

$$T_x^{(r)} \text{ le temps de } r\text{-ème retour en } x,$$

défini par récurrence par

$$T_x^{(0)} = T_x; \quad T_x^{(r+1)} = \inf\{n > T_x^{(r)} : X_n = x\}.$$

Définition 5. *Un état x est dit récurrent si la probabilité de retour est 1, c'est-à-dire si*

$$\mathbb{P}_x(T_x^{(1)} < \infty) = 1.$$

Si cette propriété n'est pas vérifiée, on dit que l'état est transient.

Proposition 8. *Pour tous $x \in E, r \geq 0, T_x^{(r)}$ est un temps d'arrêt.*

On appelle r -ème excursion

$$S_x^{(r)} = T_x^{(r+1)} - T_x^{(r)}$$

le temps passé loin de x entre le r -ème et le $r + 1$ -ème passage.

Proposition 9. La loi de $S_x^{(r)}$ ne dépend pas de r : Pour tout $k \geq 1$

$$\mathbb{P}(S_x^{(r)} = k) = \mathbb{P}(S_x^{(1)} = k) = \mathbb{P}_x(T_x = k).$$

De plus, $S_x^{(r)}$ est indépendante de $(X_k; k \leq T_x^{(r)})$. Les $S_x^{(r)}, r \geq 0$, forment donc une suite de variables IID (indépendantes et identiquement distribuées).

Une autre manière de voir les choses est de considérer le nombre de visites en un point x après 0 sachant $X_0 = x$:

$$V_x = \#\{n \geq 1 : X_n = x\} = \sum_{n=1}^{\infty} \mathbf{1}_{X_n=x}.$$

Proposition 10. $x \in E$ est récurrent ssi pour toute loi initiale

$$V_x = \infty \quad p.s.$$

ssi

$$\sum_{n \geq 0} Q^n(x, x) = \infty.$$

Théorème 3 (Polya, 1921). La marche aléatoire sur \mathbb{Z}^d est récurrente (en tous points) si $d \leq 2$ et transiente (en tous points) si $d \geq 3$.

Proposition 11. Au sein d'une même classe d'équivalence, les états sont soit tous récurrents, soit tous transients. On parle alors de classe récurrente ou de classe transiente.

Remarque 2. Si E est fini, il y a toujours au moins une classe récurrente. (Il est impossible que tous les états n'aient été visités qu'un nombre fini de fois en un temps infini).

Il peut y avoir plusieurs classes récurrentes.

Définition 6. On dit qu'une chaîne de Markov est irréductible si il n'y a qu'une seule classe.

4 Distributions invariantes

Définition 7. Soit μ une mesure sur E . On dit que μ est invariante pour la chaîne de Markov X de matrice de transition Q si $\mu Q = \mu$.

μ est une mesure invariante ssi pour tout $x \in E$

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

Proposition 12. Si μ_0 la distribution initiale est invariante, alors μ_0 est également la distribution de X_1, X_2, \dots , c'est-à-dire pour tout $x \in E$

$$\mathbb{P}_{\mu_0}(X_1 = x) = \mathbb{P}_{\mu_0}(X_2 = x) = \dots = \mu_0(x).$$

Proposition 13. On suppose E fini. On sait que pour $x \in E$, pour tout $n \geq 0$, $\mu_{x,n} = (Q^n(x, y))_{y \in E} = (\mathbb{P}_x(X_n = y))_{y \in E}$ est une mesure de probabilité.

Si il existe $x \in E$ et une mesure de probabilité π tel que pour chaque y

$$\mu_{x,n}(y) \rightarrow \pi(y),$$

alors π est une distribution invariante.

Théorème 4 (Admis). Toute chaîne de Markov irréductible récurrente admet au plus une mesure invariante à une constante multiplicative près (et donc au plus une probabilité invariante).

Proposition 14. Soit X une chaîne de Markov IR, et $x \in E$. On appelle μ_x la mesure définie par

$$\mu_x(y) = \mathbb{E}_x \left(\sum_{n=1}^{T_x^{(1)}} \mathbf{1}_{X_n=y} \right), y \in E$$

le nombre moyen de visites en y entre 2 passages en x . Alors pour tout x , μ_x est une mesure invariante.

De plus, $0 < \mu_x(y) < \infty$ pour tout $x, y \in E$.

Comme μ_x est une mesure invariante, si sa masse est finie, alors

$$\pi(y) = \frac{\mu_x(y)}{\mu_x(E)}, y \in E.$$

est une distribution invariante (et ne dépend pas de x).

On a

$$\mu_x(E) = \sum_{y \in E} \mathbb{E}_x \sum_{n=1}^{T_x^{(1)}} \mathbf{1}_{X_n=y} = \mathbb{E}_x \sum_{n=1}^{T_x^{(1)}} \sum_y \mathbf{1}_{X_n=y} = \mathbb{E}_x \sum_{n=1}^{T_x^{(1)}} 1 = \mathbb{E}_x T_x^{(1)}.$$

On en déduit

$$\pi(y) = \frac{\mu_x(y)}{\mathbb{E}_x T_x^{(1)}}$$

Avec $x = y$, ça nous donne notamment une relation entre la valeur de la distribution invariante en x et le temps de retour moyen :

$$\pi(x) = \frac{1}{\mathbb{E}_x T_x^{(1)}}$$

Même si c'est théoriquement intéressant, $\mathbb{E}_x T_x^{(1)}$ (ou $\pi_x(y)$) est dur à calculer en pratique : Il faut résoudre un systèmes de $|E|$ équations a $|E|$ inconnues.

Voici un outil plus pratique :

Proposition 15. On dit qu'une distribution μ est réversible si

$$\mu(x)Q(x, y) = \mu(y)Q(y, x), x, y \in E.$$

Toute distribution réversible est aussi invariante

Théorème 5. Soit X une chaîne de Markov irréductible. Alors on a les équivalences suivantes :

(i) X admet une distribution invariante unique, définie par

$$\pi(x) = \frac{1}{\mathbb{E}_x(T_x^{(1)})}$$

(ii) Tout état x vérifie

$$\mathbb{E}_x(T_x^{(1)}) < \infty$$

(iii) Un état x le vérifie.

Si X vérifie la condition (i) (ou de manière équivalente (ii) et (iii)), on dit que X est **récurrente positive**, sinon on dit qu'elle est **récurrente nulle**.

Théorème 6. Soit X une chaîne de Markov IR sur un espace d'états fini. Alors X est IRP.

4.1 Convergence à l'équilibre

Définition 8. Une chaîne irréductible est dite apériodique ssi pour tout $x \in E$ $Q^n(x, x) > 0$ pour n suffisamment grand.

A l'inverse, pour une chaîne périodique, il existe un entier $p > 1$ minimal, appelée période, tel que pour un certain $x \in E$ et un certain entier k , pour tout $n \geq 1$,

$$Q^{k+np}(x, x) = 0.$$

(admis)

Théorème 7. Une chaîne est apériodique si pour un $x \in E$ le pgcd de tous les temps n tels que $Q^n(x, x) > 0$ est 1 .

Théorème 8 (Convergence à l'équilibre). Soit X une chaîne de Markov IRPA. Supposons que π soit une distribution invariante pour X . Alors pour tout $x \in E$,

$$\mathbb{P}_x(X_n = y) \rightarrow \pi(y).$$

4.2 Théorème ergodique

Théorème 9 (Théorème ergodique). *Soit X une chaîne de Markov irréductible de distribution initiale une probabilité μ_0 . Pour $n \geq 1$, on note $V_x(n)$ le nombre de visites en x avant le temps n*

$$V_x(n) = \sum_{k=1}^n \mathbf{1}_{X_k=x}.$$

Alors pour tout état $x \in E$ p.s.

$$\frac{1}{n} V_x(n) = \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{X_k=x} \rightarrow \frac{1}{\mathbb{E}_x(T_x^{(1)})}.$$

Remarque 3. *Comme on a pas fait d'hypothèse de récurrence, cette dernière quantité peut être nulle.*

5 Chaines de Markov et simulation

5.1 Algorithme Hit-and-run

Exercice 1 (Simulation de la loi uniforme, Hit-and-Run Algorithm). Soit A un sous-ensemble mesurable de \mathbb{R}^d tel que $\lambda(A) > 0$, où λ est la mesure de Lebesgue d -dimensionnelle. On rappelle qu'une variable uniforme sur A a la loi

$$\mu_A(dx) = \frac{\lambda(dx)}{\lambda(A)}.$$

Pour certains ensembles A , les méthodes accept-or-reject sont très peu efficaces, typiquement lorsque A est très "mince", ou que la dimension d est très grande. Les chaînes de Markov peuvent fournir une alternative.

Pour simplifier et parce qu'on travaille dans un cadre discret, on suppose ici que A est un sous-ensemble fini de \mathbb{Z}^2 . On suppose de plus que A est connexe, où un point est relié à un autre ssi ils sont reliés par une arête de \mathbb{Z}^2 .

On considère la suite de variables aléatoires au comportement suivant :

- $X_0 \in A$.
- A chaque temps n , on choisit avec probabilité 1/2 la ligne sur laquelle se situe X_n ou la colonne sur laquelle se situe X_n .
- On tire uniformément $X_{n+1} \in A$ sur cette ligne ou cette colonne.

1. Identifier la matrice de transition Q et les propriétés de la chaîne de Markov .
2. Montrer que la distribution uniforme sur A est invariante.
3. En déduire une manière de simuler approximativement une variable uniforme sur A . Quelle convergence a-t-on ?
4. Comment pourrait-on généraliser cette méthode dans le cadre continu (informel) ?

5.2 Algorithme de Metropolis

Etant donné une mesure de probabilité π sur un espace E , le but de l'algorithme de Metropolis est de construire une chaîne de Markov $X = (X_n)$ telle que la loi de X_n converge vers π ,

$$X_n \xrightarrow{\mathcal{L}} \pi.$$

On suppose sans perte de généralité que $\pi(x) > 0$ sur E (autrement il suffit d'ôter de E les points où π s'annule). Une manière pour approximer π de cette manière est de trouver une matrice stochastique $Q(x, y)$ telle que la chaîne de Markov correspondante soit IRPA et π est invariante pour Q . L'algorithme de Metropolis consiste en les étapes suivantes :

- Construire matrice de transition $P(x, y)$ quelconque telle que la chaîne de Markov correspondante qui vit dans le bon espace d'états E soit irréductible aperiodique. Il faut de plus que P soit symétrique : $P(x, y) = P(y, x)$. Pour le bon fonctionnement de l'algorithme de simulation, il faut que la chaîne de Markov correspondante soit facile à simuler, c'est-à-dire que la loi $P(x, \cdot)$ doit être facile à calculer.
- Tirer X_1 suivant une loi quelconque μ (typiquement $\mu = \delta_x$ pour une certaine configuration $x \in E$)
- Pour chaque n , tirer Y_{n+1} suivant la loi $P(X_n, \cdot)$ (c'est-à-dire tirer Y_{n+1} comme si (X_n, Y_{n+1}, \dots) était une chaîne de Markov de matrice de transition $P(x, y)$)
- Tirer U_n une variable de loi uniforme sur $[0, 1]$ indépendamment de (X_n) et (Y_n) .
- Si $\pi(Y_{n+1})/\pi(X_n) \geq U_n$, poser $X_{n+1} = Y_{n+1}$
- Sinon, garder $X_{n+1} = X_n$.

En d'autres termes, on fait évoluer $X = (X_n)$ comme une chaîne de Markov normale de matrice de transition P , à la différence qu'à chaque itération on ne garde la nouvelle valeur X_{n+1} que si le nouveau ratio $\pi(X_{n+1})/\pi(X_n)$ est suffisamment élevé, autrement on laisse l'ancienne valeur $X_{n+1} = X_n$.

Exercice 2. Pourquoi (X_n) est une chaîne de Markov (homogène)? Quelle est sa matrice de transition? Montrer qu'elle est irréductible et réversible. Qu'en déduisez-vous sur la limite de X_n ? Par quel type plus général de condition peut-on remplacer

$$U_n \leq \frac{\pi(Y_{n+1})}{\pi(X_n)}?$$

Barker a proposé la condition

$$U_n \leq \frac{\pi(Y_{n+1})}{\pi(X_n) + \pi(Y_{n+1})}$$

6 Chaînes de Markov en temps continu, processus de Poisson

6.1 Lois sans mémoire

Pour $t, s \geq 0$,

$$\mathbb{P}(T \in [t, t + s] | T \geq t) = \mathbb{P}(T \in [0, s]).$$

Cette propriété caractérise ce qu'on appelle les lois **sans mémoire**, et il y a peu de solutions.

Exercice 3. Les seules variables aléatoires sans mémoire sont les variables exponentielles, c'est-à-dire avec fonction de distribution

$$\mathbb{P}(T \geq t) = \exp(-\lambda t), t \geq 0,$$

où $\lambda > 0$ est le paramètre de la loi.

6.2 Processus de Poisson

Le processus décrit précédemment est appelé le processus de Poisson. Formellement, on le définit ainsi :

- Soit $\lambda > 0$.
- Soit $T_i, i \geq 1$, une famille de variables aléatoires IID exponentielles de paramètre λ .
- Soit $S_n = \sum_{k=1}^n T_k, n \geq 0$ (avec $S_0 = 0$),
- Pour $t \geq 0$, on pose

$$X(t) = \max\{n : S_n \leq t\} = \min\{n : S_{n+1} > t\} = \sum_{k=1}^{\infty} \mathbf{1}_{\{S_k \leq t\}}.$$

Ce processus possède la propriété de Markov continue :

$$\forall n \geq 1, 0 < t_1 < t_2 < \dots < t_n, x_1 \in \mathbb{N}, \dots, x_n \in \mathbb{N},$$

$$\mathbb{P}(X_{t_n} = m_n | X_{t_1} = x_1, \dots, X_{t_{n-1}} = x_{n-1}) = \mathbb{P}(X_{t_n} = m_n | X_{t_{n-1}} = x_{n-1})$$

Au lieu de matrice de transition, on parle pour le processus X_t de noyau de transition $Q_t(x, y), t > 0, x, y \in \mathbb{N}$,

$$Q_t(x, y) = \mathbb{P}(X_t = y | X_0 = x).$$

Proposition 16. Pour $t > 0$, soit N_t une variable de Poisson de paramètre λt . Pour le processus de Poisson,

$$Q_t(x, y) = \mathbb{P}(N_t = y - x), t > 0, x \leq y \in \mathbb{N}.$$

Proposition 17. On déduit de l'exercice précédent le résultat suivant : Le processus de Poisson est à accroissements stationnaires, c'est-à-dire

$$\forall t_1, t_2, s > 0, X_{t_1+s} - X_{t_1} \stackrel{(d)}{=} X_{t_2+s} - X_{t_2}.$$

6.3 Générateur infinitésimal

Soit f une fonction bornée et dérivable. On pose

$$Lf(x) = \lim_{\varepsilon \rightarrow 0} \frac{\mathbb{E}[f(X_{t+\varepsilon}) - f(X_t)]}{\varepsilon}$$

en supposant que la limite existe. Remarquons que si le processus est homogène, la limite ne dépend pas de t .

Si de plus le processus est à accroissements stationnaires, comme c'est le cas pour le processus de Poisson, cet opérateur ne dépend pas de t : En effet, $X_{t+\varepsilon} - X_\varepsilon \stackrel{(d)}{=} X_\varepsilon - X_0$ pour tous $t, \varepsilon > 0$.

Dans ce cas, l'opérateur L transforme une fonction en une autre fonction, qui dénote la manière dont $f(X_t)$ varie au voisinage de 0 si $X_0 = x$.

Dans le cas Poissonien, $X_\varepsilon - X_0$ est une variable de Poisson de paramètre $\lambda\varepsilon$.

$$\begin{aligned} \mathbb{E}[f(X_\varepsilon) - f(X_0) | X_0 = x] &= (f(x+0) - f(x)) * \exp(-\lambda\varepsilon) + (f(x+1) - f(x))\lambda\varepsilon \exp(-\lambda\varepsilon) + 2\|f\|_\infty o(\varepsilon) \\ Lf(x) &= \lambda(f(x+1) - f(x)) \end{aligned}$$

- On dit que X_t est un *processus de sauts*, car il ne peut varier que par discontinuités.
- Dans ce cas, X_t est un processus de saut à taux constant λ : λ est l'intensité avec laquelle le processus *saute*, sous-entendu la répartition des sauts est Poissonnienne, c'est-à-dire que les sauts sont séparés par des variables exponentielles IID.

6.4 Processus de Poisson composé

On introduit une loi de probabilité μ sur \mathbb{R} , qui représente la variable aléatoire de la quantité d'argent déposée par un client.

Pour modéliser la quantité d'argent déposée à un instant $t \geq 0$, on introduit une suite de variables iid $Y_k, k \geq 1$, indépendants de μ_t , avec comme loi commune μ , et on suppose que le i -ème client a apporté une quantité d'argent $Y_k \in \mathbb{R}$.

En appelant X_t le processus de Poisson défini au chapitre précédent (avec paramètre d'intensité $\lambda > 0$), la quantité d'argent déposée à l'instant t est donc

$$Z_t = \sum_{k=1}^{\infty} Y_k \mathbf{1}_{\{X_t \geq k\}},$$

en gros, on comptabilise les sommes de tous les clients qui sont effectivement déjà passés à l'instant t .

En utilisant le fait que X_t est une variable de Poisson de paramètre λt , et que les Y_i sont IID de loi μ , on peut déterminer la loi de Z_t , via la fonction caractéristique.

Théorème 10. Avec les notations précédentes, pour $t > 0$, la fonction caractéristique de Z_t est

$$\psi_{Z_t}(\theta) = \exp(\lambda t (\psi_{Y_1}(\theta) - 1)), \theta \in \mathbb{R}.$$

Cela nous permet par exemple de déterminer les premiers moments de Z_t :

$$\mathbb{E}Z_t = -i \frac{d}{d\theta} \Big|_{\theta=0} \psi_{Z_t}(\theta) = -i \lambda t \psi'_{Y_1}(0) \exp\left(\lambda t \underbrace{(\psi_{Y_1}(0) - 1)}_{=1}\right) = \lambda t \mathbb{E}[Y_1]$$

$$\mathbb{E}Z_t^2 = -\frac{d^2}{dt^2} \Big|_{t=0} \psi_{Z_t}(\theta) = -\exp(\dots) [\lambda t \psi''_{Y_1}(0) + (\lambda t \psi'_{Y_1}(0))^2] = (\lambda t \mathbb{E}Y_1^2 + \lambda^2 t^2 (\mathbb{E}Y_1)^2)$$

Du coup, la variance est

$$\text{Var}(Z_t) = \mathbb{E}[Z_t^2] - (\mathbb{E}Z_t)^2 = \lambda t \mathbb{E}Y_1^2.$$

Le GI de Z_t est

$$Lf(x) = \lambda \mathbb{E}[f(x + Y_1) - f(x)].$$