Chapitre 3. Introduction aux Processus stochastiques

LAKHEL El Hassan

Département: Informatiques et Réseaux-Télécom.
ENSA de Safi, A. U. 2022-2023

Page Web: https://sites.google.com/view/lakhelelhassan/ modélisation stochastique

Introduction :

Les processus stochastiques interviennent de façon très naturelle dans les applications pour l'analyse et la modélisation de systèmes complexes (biologie, physique, économie, ingénierie) ou en interaction avec un environnement mal connu (perturbations, bruits de mesures...)
La théorie des processus aléatoires concerne l'étude mathématique de phénomènes évoluant dans le temps, et dont l'évolution est de caractère aléatoire, c'est-à-dire non prévisible avec certitude.

Exemples

- *Jeux de hasard :* $(X_n)_{n\in\mathbb{N}}$ *représente la taille de la fortune d'un joueur qui participe à un jeu de hasard à l'étape n.*
- *Bourse :* $(X_t)_{t\in\mathbb{R}}$ *: cours d'une action le jour t, (une action : c'est un titre représentant une fraction de capitale dans certaines sociétés et donnant droit à une part de bénéfices.)*
- *Assurance :* $(X_n)_{n\in\mathbb{N}}$ *est le montant total des indemnités versées par une compagnie d'assurance pour les sinistres survenus le mois n.*
- *Epidémiologie :* $(X_n)_{n\in\mathbb{N}}$ *: nombre d'individus infectés par une maladie contagieuse au bout de n jours.*
- *Files d'attente :* $(X_t)_{t\in\mathbb{R}}$ *représente le nombre de clients à un guichet à l'instant t*

$\vdots$

La modélisation mathématique des phénomènes de ce genre est l'objet de la théorie des processus stochastique.

Pour définir un processus aléatoire, il faut :

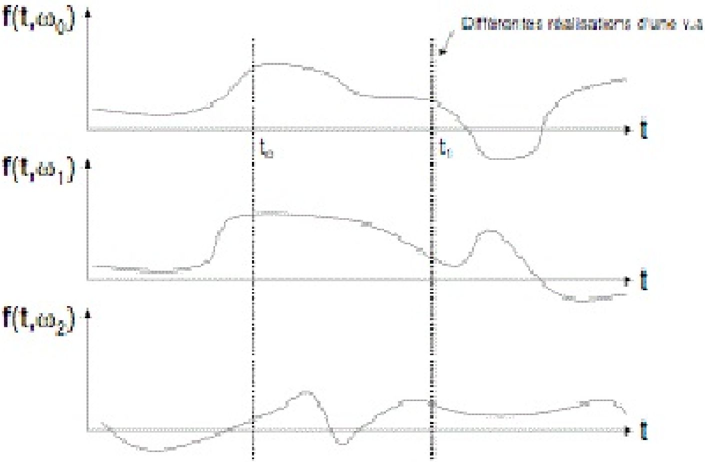
1. Un espace des temps $T \subseteq \mathbb{R}^+$: Les deux espaces des temps les plus utilisés sont :
   - Si $T = \mathbb{N}$ ou $\mathbb{N}^*$, on parle de processus stochastique en temps discret
   - Si $T = \mathbb{R}^+$, on parle de processus stochastique en temps continu.
2. Un espace des états : L'ensemble E peut être :
   - discret : une partie de $\mathbb{Z}$.
   - non discret : par exemple $E = \mathbb{R}$ ou une partie de $\mathbb{R}^2$ ...
3. une famille de variables ou vecteurs aléatoires indexée par un ensemble de paramètres $t \in T$ (le temps).

Autrement dit :

$$X : \begin{array}{ccc} T \times \Omega & \longrightarrow & E \\ (t, \omega) & \longrightarrow & X_t(\omega) \end{array}$$

Remarque

1. *$(X_t)_{t \in T}$ une famille de v. a. à valeurs dans $(E, \mathcal{E})$, définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.*
2. *$\omega \longrightarrow X_t(\omega)$ est une v. a. pour n'importe quelle valeur de $t \in T$ (t fixé).*
3. *Pour t et ω fixés $X_t(\omega)$ représente l'état du système à la date t pour une réalisation particulière ω. C'est un nombre.*
4. *Pour chaque $\omega \in \Omega$, $t \longrightarrow X_t(\omega)$ s'appelle une trajectoire du processus assouciée à ω.*


$f(t,\omega_0)$
$f(t,\omega_1)$
$f(t,\omega_2)$
t_0
t_1
t
t
t

Quelques relations de dépendance

Les éléments principaux qui différencient les processus aléatoires généraux sont l'espace des états E, l'espace des temps $\mathbb{T}$ et les relations de dépendance entre les X_t.

Deux propriétés des processus stochastiques que l'on suppose vraies dans la définition des processus de Poisson sont :

Définition

- *Processus à accroissements indépendants : Un processus $(X_t)_{t\geq 0}$ est dit à accroissements indépendants si pour tout $n \in \mathbb{N}$, et pour tous $t_1 < t_2 < ... < t_n$, les accroissements $X_{t_1} - X_0$, $X_{t_2} - X_{t_1}$, ..., $X_{t_n} - X_{t_{n-1}}$ sont des variables aléatoires indépendantes.*

- *Si les v. a. $X_{t_2+s} - X_{t_1+s}$ et $X_{t_2} - X_{t_1}$ possède la même fct de répartition pour tout s ; on dit que le processus (X_t) est un processus à accroissements stationnaires.*
  Autrement dit les v. a. $X_{t_2+s} - X_{t_1+s}$ et $X_{t_2} - X_{t_1}$ sont identiquement distribuées pour tous $0 \leq t_1 \leq t_2$ et $s \geq 0$.

Premiers exemples : Processus de Bernoulli

- Processus de Bernoulli : $\{S_n, n \in \mathbb{N}\}$

$$S_n = X_1 + X_2 + ... + X_n,$$

avec les (X_i) sont i.i.d. de même loi de Bernoulli de paramètre p. Le processus $\{S_n, n \in \mathbb{N}\}$ s'appelle processus de Bernoulli.

Exercice

Considérons un automobiliste qui, sur le trajet qui l'amène au travail, rencontre 8 feux tricolores ; les réglages des feux sont tels que la probabilité pour chaque feu d'être vert est de $\frac{2}{3}$; on suppose pour simplifier que ces feux sont indépendants.

Définir un modèle probabiliste pour calculer les probabilités des événements suivants :

1. *lors d'un trajet, les 8 feux sont verts ;*
2. *lors d'un trajet, exactement 5 feux sur 8 sont verts.*

Solution :
Soit

$$X_i = \begin{cases} 1 & \textit{si le feu i est vert}; \\ 0 & \textit{sinon}. \end{cases}$$

On a

$$S_8 = X_1 + ... + X_8.$$

C'est la somme de huit v. a. indépendantes de bernoulli.
Donc

$$S_8 \sim B(8, p = \frac{2}{3}).$$

D'où

1. $\mathbb{P}(S_8 = 8) = (\frac{2}{3})^8 = ...$
2. $\mathbb{P}(S_8 = 5) = C_8^5(\frac{2}{3})^5(\frac{1}{3})^3 = ...$

Premiers exemples : Marche aléatoire

Définition

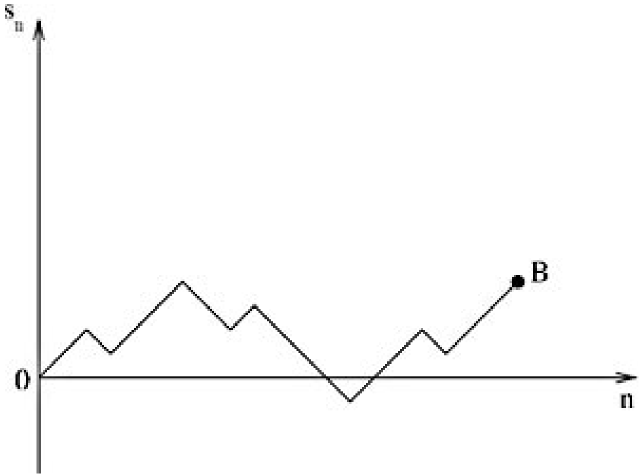
Soit X_n des v. a. indépendantes chacune prenant la valeur 1 avec une probabilité p et -1 avec une probabilité $1-p$. On note,

$$S_n = S_0 + \sum_{i=1}^{n} X_i, \quad n \in \mathbb{N}^*$$

La famille $S = \{S_n, n \in \mathbb{N}\}$ est un processus discret appelé marche aléatoire sur $\mathbb{R}$.

Remarque

Une particule qui, à l'instant 0, se trouve à l'origine. A chaque unité de temps, on lance une pièce de monnaie, si l'on obtient pile (resp. face) X_i prend $+1$ (resp. -1).
Ainsi, la v. a. (S_n) désigne la postition de la particule au bout de n lancers de la pièce.



S
n
0
n
B

Processus de comptage

Définition

Soit $(\tau_n)_{n\geq 1}$ *des v. a. positives.*
On note :

$$T_0 = 0 \quad et \quad T_n = \tau_n + T_{n-1}, \text{ pour } n \geq 1.$$

et

$$N_t = \begin{cases} 0 & si\ t < T_1 \\ n & si\ T_n \leq t < T_n + 1, \end{cases} \quad \text{on a aussi} \quad N_t = card\{n \geq 1, \quad T_n \leq t\} = \sum_{n\geq 1} 1_{\{T_n \leq t\}}$$

La suite $(T_n)_{n\in\mathbb{N}}$ *est appelée processus de renouvellement associé aux* (τ_n) *et* $\{N_t, t \in \mathbb{N}\}$ *s'appelle un processus de comptage.*

Remarque

- (T_n) *suite* $\uparrow$ *de v. a. réelles.*
- *Les* N_t *sont des v. a. entières.*

Exemples

1. *Si les* τ_n *modélisent les durées de vie d'une composante électrique,* N_t *représente le nombre de pièces changées avant l'instant* t*.*
2. *Si les* τ_n *modélisent les durées d'interarrivées des clients,* N_t *représente le nombre d'arrivées durant* $[0, t]$*.*

Processus de Poisson

Le processus de Poisson se rencontre fréquemment en pratique :

1. arrivées d'usagers à un arrêt d'autobus, à un guichet de poste ...
2. occurence de pannes de composants électriques ;
3. arrivées d'appels à un standard téléphonique, ...

On suppose que (τ_n) (les durées des intervalles interarrivées) est une suite de variables aléatoires i.i.d. de même loi exponentielle de paramètre λ.

On note T_1 l'instant d'arrivée du premier événement, T_2 le deuxième, T_k le k-ième et ainsi de suite.

On définit le processus $(N_t)_{t\geq 0}$ de la façon suivante :

$$N_t = \sum_{j\geq 1} \mathbf{1}_{\{T_j\leq t\}} = \sum_{j\geq 1} \mathbf{1}_{[0,t]}(T_j), \quad t \in \mathbb{R}^+.$$

Définition

Le processus $(N_t)_{t\geq 0}$ est un processus de Poisson d'intensité λ si :

(a) *Pour toutes suites $\nearrow$ de temps, les v. a. N_{t_0}, $N_{t_1} - N_{t_0}$,..., $N_{t_n} - N_{t_{n-1}}$ sont indépendantes.*

(b) *pour $0 \leq s < t$, la loi de $N_t - N_s$ ne dépend de t et s que par la différence $t - s$, plus précisément :*

$$N_t - N_s \sim \mathcal{P}oi(\lambda(t-s)).$$

Processus de Poisson (suite)

Remarque

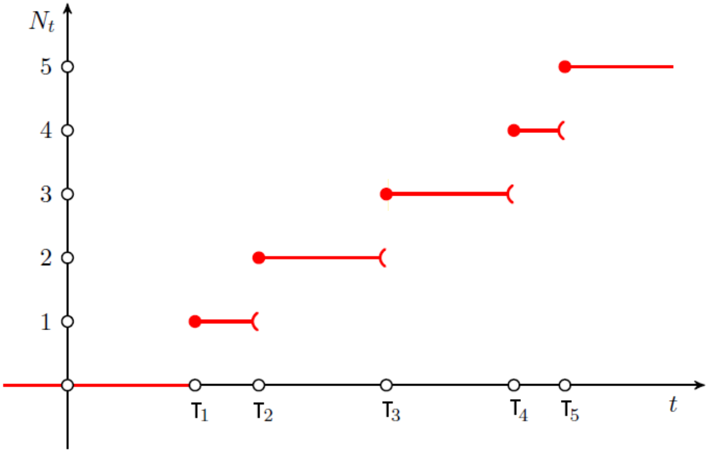
1. *La propriété (a) est appelée la propriété des accroissements indépendantes ; il signifie que les arrivées dans deux intervalles de temps disjoints sont indépendantes.*
2. *La propriété (b) est appelée "la stationnarité des accroisements" de (N_t).*
3. *La v. a. $N_t - N_s$ qui compte le nombre d'arrivées dans l'intervalle $]s, t]$ suit la loi de Poisson $\mathcal{P}oi(\lambda(t-s))$. Elle ne dépend que de la longueur de l'intervelle. De plus, la v. a. $N_t - N_s$ suit la même loi que N_{t-s}.*

Définition

Si les v. a. $(\tau_i)_{i\geq 1}$ sont i.i.d. et de même loi exponentielle de paramètre $\lambda > 0$, on dit que la suite $(T_k = \sum_{i=1}^{k} \tau_i)_{k\geq 1}$ est la suite des instants de sauts d'un processus de Poisson de paramètre λ.

Les trajectoires de N sont croissantes et continues à droite.

Trajectoire d'un processus de Poisson



Exercice

Soit $(\tau_n)_{n\geq 1}$ *une suite de v. a. i.i.d. le loi* $\mathcal{E}(\lambda)$
Notons :

$$T_n = \tau_1 + \tau_2 + \ldots + \tau_n.$$

Montrer que T_n *suit la loi* $\gamma(n, \lambda)$, *i.e*

$$f_{T_n}(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} 1_{[0,\infty[}(x)$$

Remarque

La loi Gamma est habituellement surnommée la loi d'Erlang.

Solution

Par recurrence sur n.
Pour $n = 1$, il n'y a presque rien à montrer puisque T_1 suit la loi $\exp(\lambda)$ et celle-ci coincide avec $\gamma(1, \lambda)$.
Supposons que l'hypothèse est vrai pour n, et montrons qu'il est vrai pour $n + 1$.
En effet, on a :

$$\begin{aligned}
f_{T_{n+1}}(x) &= f_{T_n+\tau_{n+1}}(x) \\
&= \int_{-\infty}^{+\infty} f_{T_n}(u) f_{\tau_{n+1}}(x-u)du \\
&= \int_0^x \frac{\lambda^n}{(n-1)!} u^{n-1} e^{-\lambda u} \lambda e^{-\lambda(x-u)} du \\
&= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \int_0^x u^{n-1} du \\
&= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \left[\frac{u^n}{n}\right]_0^x \\
&= \frac{\lambda^{n+1}}{(n)!} x^n e^{-\lambda x}
\end{aligned}$$

Exercice

Un guichet ouvre au temps $T_0 = 0$. Soient T_1 et T_2 les instants d'arrivées respectifs des deux premiers clients. On suppose que les variables aléatoires T_1 et $T_2 - T_1$ sont indépendantes et ont pour loi $\mathcal{E}(\lambda)$ $\lambda > 0$. Enfin, on note N_t le nombre de clients qui se sont présentés au guichet entre l'instant 0 et l'instant t.

1. *Montrer que $P(N_t = 0) = P(T_1 > t)$ et calculer cette probabilité.*
2. *Calculer la fonction de répartition de T_2.*
3. *Justifier que*

$$\{N_t = 1\} = \{T_1 \leq t < T_2\} \quad \textit{et} \quad \{N_t = n\} = \{T_n \leq t < T_{n+1}\}$$

4. *En déduire que*

$$P(N_t = 1) = P(T_1 \leq t) - P(T_2 \leq t)$$

5. *Calculer $P(N_t = 1)$.*
6. *N_t suit une loi classique. Laquelle ? Que vaut $E(N_t)$?*

1. $\{N_t = 0\}$ signifie qu'aucun client ne s'est présenté au guichet jusqu'à l'instant t, c'est-à-dire le temps d'arrivée du premeir client est $> t$. Ainsi $P(N_t = 0) = P(T_1 > t)$. $T_1 \sim Exp(\lambda)$, donc
$P(N_t = 0) = P(T_1 > t) = 1 - F_{T_1}(t) = e^{-\lambda t}$.

2. **Méthode 1 :** on utilisera l'égalité $T_2 = T_1 + (T_2 - T_1)$, qui représente T_2 comme la somme de deux variables aléatoires indépendantes].
Si $x < 0$, on a $\mathbb{P}(T_2 \leq x) = 0$.
Si $x \geq 0$, les v. a. T_1 et $T_2 - T_1$ sont iid de même loi $Exp(\lambda)$, d'où
$f_{(T_1, T_2 - T_1)}(u, v) = \lambda^2 e^{-\lambda(u+v)} 1_{[0,\infty[\times[0,\infty[(u,v)}$.

$$\begin{aligned}
\mathbb{P}(T_2 \leq x) \quad & \mathbb{P}(T_1 + T_2 - T_1 \leq x) = \mathbb{E}(1_{\{(u,v)\in\mathbb{R}^2 \,|\, u+v\leq x\}}(T_1, T_2 - T_1)) \\
& = \int_{\{(u,v)\in\mathbb{R}^2 \,|\, u+v\leq x\}} f_{(T_1, T_2 - T_1)}(u, v)dudv \\
& = \int_0^x \lambda e^{-\lambda u} [\int_0^{x-u} \lambda e^{-\lambda v} dv]du = \int_0^x \lambda e^{-\lambda u}(1 - e^{-\lambda(x-u)})du \\
& = 1 - e^{-\lambda x} - \lambda x e^{-\lambda x}.
\end{aligned}$$

Méthode 2 : On utilisera la densité de T_2 donnée par l'exercice précédent.

3. Le temps d'arrivée du deuxème client est $T_2 > t$ et $t \geq T_1$.

4. On a
$$\mathbb{P}(T_1 \leq t < T_2) + \mathbb{P}(T_2 \leq t) = \mathbb{P}(T_1 \leq t < T_2 \text{ } ou \text{ } T_2 \leq t) = \mathbb{P}(T_1 \leq t).$$

5. Des questions précédentes, on tire
$$\mathbb{P}(N_t = 1) = \mathbb{P}(T_1 \leq t) - \mathbb{P}(T_2 \leq t) = \lambda t e^{-\lambda t}.$$

6. $N_t \sim \mathcal{P}(\lambda t)$ et on a $\mathbb{E}(N_t) = \lambda t$.

Proposition

Soit $N = (N_t)_{t\geqslant 0}$ un processus de Poisson d'intensité λ. Pour tout $t > 0$, la v.a.r. N_t est de loi de Poisson de paramètre λt, i.e.

$$P(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

Preuve. Puisque

$$T_n \sim Gamma(n, \lambda),$$

car la somme des v.a. iid de loi $exp(\lambda)$ suit la loi $gamma(n, \lambda)$ appelée aussi loi d'Erlang. Alors,

$$\mathbb{P}(N_t = n) = \mathbb{P}(T_n \leq t) - \mathbb{P}(T_{n+1} \leq t) = \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} dx - \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^n}{(n)!} dx.$$

Finalement,

$$\mathbb{P}(N_t = n) = \int_0^t \frac{d}{dx}\left(e^{-\lambda x} \frac{(\lambda x)^n}{(n)!}\right) dx = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Exercice

Les pannes d'une machine se produisent selon un processus de Poisson d'intensité $\lambda = 2$ par semaine, et qu'exactement deux pannes se sont produites dans l'intervalle $[0, 1]$. Soit $t_0 > 3$ une valeur quelconque de t.

1. *Calculer la probabilité que, à l'instant t_0, deux semaines sont écoulées depuis que la dernière panne s'est produite.*
2. *Quelle est la probabilité qu'il n'y ait aucune panne au cours des deux prochaines jours, à partir de t_0, si exactement une panne s'est produite (en tout) pendant les deux dernières semaines ?*

Remarque

Dans un problème sur le processus de Poisson, il faut d'abord déterminer la valeur du paramètre de la loi de Poisson pour chacune des questions.
Soit N_t le nb de pannes dans l'intervalle $[0, t]$, où t est en semaine.
On a :

$$N_t \sim \mathcal{P}(2t).$$

1. On a :

$$\mathbb{P}(N_{t_0} - N_{t_0-2} = 0) = \mathbb{P}(N_2 = 0) = e^{-4}.$$

2. On a aussi :

$$N_{t_0+\frac{2}{7}} - N_{t_0} \sim N_{\frac{2}{7}} \sim \mathcal{P}(\frac{4}{7}).$$

$$\mathbb{P}(N_{\frac{2}{7}} = 0) = e^{-\frac{4}{7}} = 0.56.$$

Exercice

Des clients arrivent dans un magasin selon un processus de Poisson d'intensité 30 *par heure. Quelle est la probabilité que le temps entre deux arrivées successives soit :*

1. *Plus de* 2 *minutes ?*
2. *Moins de 4 minutes ?*
3. *Entre 1 et 3 minutes ?*

Des clients arrivent dans un magasin selon un processus de Poisson d'intensité 30 par heure. Notons N_t le nombre de clients arrivés jusqu'a la date t (en minutes). La variable aléatoire N_t suit donc une loi de Poisson $\mathcal{P}(t/2)$. La durée T écoulée entre deux arrivées successives suit, elle, une loi exponentielle de paramètre $\lambda = 1/2$.

1. La probabilité que le temps écoulé entre deux arrivées soit plus de deux minutes est donc :

$$\mathbb{P}(T > 2) = e^{-2\lambda} = e^{-1}.$$

2. La probabilité que le temps écoulé entre deux arrivées soit moins de 4 minutes est donc :

$$\mathbb{P}(T \leq 4) = 1 - e^{-4\lambda} = 1 - e^{-2}.$$

3. La probabilité que le temps écoulé entre deux arrivées soit entre 1 et 3 minutes est donc :

$$\mathbb{P}(1 \leq T \leq 3) = P(T \leq 3) - \mathbb{P}(T \leq 1) = e^{-\frac{1}{2}} - e^{-\frac{3}{2}}.$$

Loi d'un processus

Un processus aléatoire est une généralisation d'un vecteur aléatoire. On parle de loi du processus stochastique pour décrire l'ensemble des lois fini-dimensionnelles.
Soit le processus stochastique $\{x_t | t \in T\}$ où T est continu ou discret.

- Fonction de répartition d'ordre n ($n \in \mathbb{N}^*$) : la fonction de répartition conjointe du v.a. $(X_{t_1}, ..., X_{t_n})$

$$F(x_1, ..., x_n; t_1, ..., t_n) = \mathbb{P}[X_{t_1} \leq x_1, ..., X_{t_n} \leq x_n]$$

- Fonction de densité d'ordre n :

$$f(x_1, ..., x_n; t_1, ..., t_n) = \frac{\partial^n F(x_1, ..., x_n; t_1, ..., t_n)}{\partial x_1 ... \partial x_n}$$

- Fonction de masse d'ordre n :

$$p(x_1, ..., x_n) = \mathbb{P}[X_1 = x_1, ..., X_n = x_n]$$

Caractéristiques

- Moyenne d'un processus stochastique à l'instant t :

$$m_X(t) = \mathbb{E}(X_t).$$

- Fonction d'autocorrélation en (t_1, t_2) :

$$R_X(t_1, t_2) = \mathbb{E}[X_{t_1} X_{t_2}]$$

- Fonction d'autocovariance en (t_1, t_2) :

$$C_X(t_1, t_2) = R_X(t_1, t_2) - m_X(t_1) m_X(t_2)$$

- Variance du processus à l'instant t :

$$V(X_t) = C_X(t, t) = \mathbb{E}[X_t^2] - [\mathbb{E}X_t]^2.$$

- Finalement, le coefficient de corrélation du p. s. en (t_1, t_2) :

$$\rho_X(t_1, t_2) = \frac{C_X(t_1, t_2)}{\sqrt{V_X(t_1) V_X(t_2)}}$$

Le signal télégraphique

Exercice

Un cas particulier du transformation du processus de Poisson est le signal télégraphique, défini par :

$$X(t) = (-1)^{N(t)} = \begin{cases} 1 & si \quad N(t) = 0, 2, 4, \dots \\ -1 & si \quad N(t) = 1, 3, \dots \end{cases}$$

1. *Déterminer la distribution de la v. a. $X(t)$.*
2. *Calculer $\mathbb{E}(X(t))$.*
3. *Calculer la fonction d'autocorrélation $R_X(t, t+s)$ de $X(t)$.*
4. *Déduire la fonction d'autocovariance $C_X(t, t+s)$ en $(t, t+s)$.*

1. On a

$$\mathbb{P}(X(t)=1)=\sum_{k=0}^{\infty}\mathbb{P}(N(t)=2k)=\sum_{k=0}^{\infty}e^{-\lambda t}\frac{(\lambda t)^{2k}}{(2k)!}=e^{-\lambda t}\sum_{k=0}^{\infty}\frac{(\lambda t)^{2k}}{(2k)!}=\frac{1+e^{-2\lambda t}}{2},$$

car

$$ch(\lambda t):=\frac{e^{\lambda t}+e^{-\lambda t}}{2}=\sum_{k=0}^{\infty}\frac{(\lambda t)^{2k}}{(2k)!}.$$

Ensuite, on a

$$\mathbb{P}(X(t)=-1)=1-\mathbb{P}(X(t)=1)=\frac{1-e^{-2\lambda t}}{2}.$$

2. On a

$$\mathbb{E}(X(t))=1.\frac{1+e^{-2\lambda t}}{2}+(-1).\frac{1-e^{-2\lambda t}}{2}=e^{-2\lambda t}.$$

3. Soit $s>0$, d'après la stationarité des accroissements, on a :

$$\begin{aligned} R_X(t,t+s) \quad &= \mathbb{E}[X(t)X(t+s)] = \mathbb{E}[(-1)^{N(t)}(-1)^{N(t+s)}] \\ &= \mathbb{E}[(-1)^{2N(t)}(-1)^{N(t+s)-N(t)}] = \mathbb{E}[(-1)^{N(t+s)-N(t)}] \\ &= \mathbb{E}[(-1)^{N(s)}] = \mathbb{E}(X(s)) = e^{-2\lambda s}. \end{aligned}$$

4. On a :

$$C_X(t,t+s)=R_X(t,t+s)-\mathbb{E}(X(t))\mathbb{E}(X(t+s))=e^{-2\lambda s}-e^{-2\lambda t}e^{-2\lambda(t+s)}.$$

Processus de sauts markoviens

Définition

Le processus stochastique $(X_t)_{t\in T}$ *est une chaîne de Markov en* *temps continu* *si :*
$\forall 0 \leq t_0 < t_1 < \ldots < t_n < t_{n+1}$

$$\mathbb{P}(X_{t_{n+1}} = x_{n+1} | X_{t_0} = x_0, \ldots, X_{t_n} = x_n) = \mathbb{P}(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n).$$

Cette propriété s'appelle propriété de Markov. Elle traduit le fait que le futur ne dépend du passé que par l'information la plus récente que l'on possède au sujet du processus.

Notation.

1. Ici, le temps t est continu, le plus souvent $t \in T = \mathbb{R}^+$. Dans le chapitre suivant, nous allons étudier le cas $t \in \mathbb{N}$.
2. Pour t fixé, $X_t \in E = \mathbb{N}$. La loi de la variable aléatoire X_t est

$$p(t) = [p_0(t), p_1(t), \ldots] = [\mathbb{P}(X_t = 0), \mathbb{P}(X_t = 1), \ldots]$$

   Les probabilités de transition $p(t, t')$ sont les matrices :

$$p_{ij}(t, t') = \mathbb{P}(X_{t'} = j | X_t = i).$$

3. Les chaîne de Markov en temps continu porte aussi le nom de **processus de sauts markoviens**.

Processus de Markov homogène

Dans de nombreux modèles, les transitions entre deux instants s et t ne dépendent pas des dates s et t, mais seulement du temps écoulé entre ces deux dates, c'est-à-dire de $(t - s)$. C'est ce qu'on appelle l'homogénéité (temporelle) de la chaîne :

Définition

La chaîne de Markov $(X_t)_{t \in \mathbb{R}+}$ est homogène si les probabilités de transition $p(t, s)$ ne dépendent de t et de s que par la différence $(t - s)$. On note alors :

$$p_{ij}(t) = p_{ij}(0, t) = \mathbb{P}(X_t = j | X_0 = i) = \mathbb{P}(X_{t+s} = j | X_s = i).$$

$\forall t > 0$ *la matrice $(p_{ij}(t))_{(i,j) \in E^2}$ est appelée matrice de transition dans le temps t.*

Convention : Dans toute la suite, on ne considérera que des processus homogènes, car l'emploi de processus Markoviens non homogènes est des plus rares.

En général la distribution initiale du processus $\mu = p(0)$ est donnée. La loi à tout instant est alors complètement déterminée par la matrice de transition $P(t)$. En effet, et si on note par $p_j(t)$ la loi marginale que le processus $(X_t)_{t\geq 0}$ soit à l'état j à l'instant t.

$$p_j(t) = \mathbb{P}(X_t = j), \qquad j \in \mathbb{N}.$$

D'après la formule des P.T., il est facile de voir que

$$p_j(t) = \sum_{i\in\mathbb{N}} \mathbb{P}(X_0 = i) P_{ij}(t) = \sum_{i\in\mathbb{N}} \mu_i P_{ij}(t).$$

Remarque

- *On a pour tout entier i et tout réel positif t :* $\sum_{j=0}^{\infty} p_{ij}(t) = 1.$
- *Nous allons ajouter l' hypothèse de régularité suivante :*

$$\forall (i,j) \in \mathbb{N}^2, \qquad p_{ij}(t) \longrightarrow_{t \longrightarrow 0} \delta_{ij} = \begin{cases} 1 & si \quad i = j \\ 0 & sinon, \end{cases}$$

c'est-à-dire :

$$\lim_{t \longrightarrow 0} P(t) = I,$$

où I est la matrice identité.
Ceci signifie qu'il n'y pas de saut instantané de l'état i à l'état j, mais pendant un intervalle de temps de durée > 0 *.*

Générateur infinitésimal

Nous allons définir une nouvelle matrice qui caractérise le processus :

Définition

La matrice A, définie par : $\forall (i,j) \in \mathbb{N}^2$,

$$A(i,j) = P'_{ij}(0) = \lim_{t \longrightarrow 0^+} \frac{P_{ij}(t) - I}{t} \qquad \left(i.e. \lim_{\Delta t \longrightarrow 0^+} \frac{P_{ij}(\Delta t) - P_{ij}(0)}{\Delta t}\right)$$

est appelée générateur infinitésimal, ou matrice génératrice, de la chaîne de Markov en temps continu.

Remarque

On a

$$\begin{cases} a_{ij} & = \lim_{t \longrightarrow 0^+} \frac{p_{ij}(t)}{t} \quad si \quad i \neq j \\ a_{ii} & = \lim_{t \longrightarrow 0^+} \frac{p_{ii}(t) - 1}{t} \quad sinon \end{cases}$$

Rappel : Le théorème de Taylor assure qu'une fonction f dérivable au point 0 admet un DL_1 en ce point :

$$f(t) = f(0) + f'(0)t + \circ(t).$$

Les développements limités en 0 à l'ordre 1 donnent :

$$p_{ij}(t) = a_{ij}t + o(t).$$

Si la chaîne est dans l'état i initialement, la probabilité qu'elle soit dans l'état j à l'instant t est environ $a_{ij}t$, avec t petit. Le nombre a_{ij} est alors appelé taux de départ infinitesimal de l'état i vers l'état j. De même, on a :

$$1 - p_{ii}(t) = -a_{ii}t + o(t).$$

Si le processus est dans l'état i initialement, la probabilité qu'il l'ait quitté à l'instant t est environ $-a_{ii}t$. Le coefficient positif $-a_{ii}$ est le taux de départ infinitesimal de l'état i.
Par ailleurs, pour tout i fixé et pour tout $t \geq 0$, la somme des $p_{ij}(t)$ fait 1. Avec les développements limités ci-dessus, on a donc formellement :

$$1 = \sum_{j} p_{ij}(t) = 1 + (a_{ii} + \sum_{j \neq i} a_{ij})t + o(t).$$

D'où :

$$\forall i \in \mathbb{N} \qquad -a_{ii} = \sum_{j \neq i} a_{ij} \qquad (*)$$

Remarque

1. *La relation $(*)$ veut dire que la somme des taux de transition de l'état i vers l'ensemble des autres états j est égale au taux de départ de l'état i.*
2. *D'après la relation $(*)$, la somme de chaque ligne de A est nulle. Résultat à ne pas confondre avec celui concernant les matrices de transition $P(t)$, pour lesquelles la somme sur chaque ligne vaut 1.*

Processus de Markov particulier : Processus de Poisson

Soit $(N_t)_{t\in T}$ un processus de Poisson de paramètre λ. Nous avons vu que le processus de comptage $(N_t)_{t\geq 0}$ est un processus aléatoire à valeurs dans $\mathbb{N}$, qui vaut 0 au temps 0 et qui saute de 1 à chacun des instants aléatoires T_i qui caractérisent le modèle. De plus

* La v. a. N_t suit une loi de Poisson $\mathcal{P}(\lambda t)$:

$$\mathbb{P}(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

* Quand le processus est dans l'état i , il saute en $i+1$ au bout d'un temps aléatoire de loi exponentielle de paramètre λ.

Soit $x \in \mathbb{N}$, posons

$$X_t = x + N_t.$$

1. Déterminer la matrice de transition de ce processus.
2. Déterminer le générateur infinitésimal de ce processus.

Solution :

1. Un calcul explicite donne la matrice de transition P_t dont les coefficients sont :
   - Pour $i \leq j$, on a :

$$P_{ij}(t) = \mathbb{P}(N_t + x = j | x = i) = \mathbb{P}(N_t = j - i) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!} I_{i \leq j}.$$

2. On en déduit que

$$A(i, i+1) = P'_{i,i+1}(0) = \lambda, \qquad A(i,i) = P'_{i,i}(0) = -\lambda$$

et que les autres $A(i,j)$ sont nuls.
On trouve le générateur :

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \ldots \\ 0 & -\lambda & \lambda & 0 & \ldots \\ 0 & 0 & -\lambda & \lambda & \ldots \\ & & & & \ddots \\ 0 & 0 & 0 & -\lambda & \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

Si le processus se trouve initialement en l'état i, il va y rester pour une durée aléatoire de loi $\mathcal{E}(\lambda)$, pour entrer alors nécessairement dans l'état $i+1$, etc.