Chaînes de Markov, Processus • Formulaire

Définitions et propriétés

Espérance mathématique :

Cas discret
$$E[X]=\sum_{x\in\aleph} xP(X=x)$$
Cas continu
$$E[X]=\int_{\mathbb{R}} xf(x)dx$$

Variance : Pour une variable aléatoire X, la variance est :
$$\mathrm{Var}(X)=E\left[\left(X-E[X]\right)^2\right]=E\left[X^2\right]-E[X]^2$$

Covariance : La covariance entre X et Y est la grandeur :
$$\mathrm{Cov}(X,Y)=E\left[\left(X-E[X]\right)\left(Y-E[Y]\right)\right]=E[XY]-E[X]E[Y]$$

Coefficient de corrélation linéaire :
$$\rho(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

Moment : Pour une variable aléatoire X, le moment d'ordre k est :
$$E\left[X^k\right]$$
Le moment centré d'ordre k est $E\left[(X-\mu)^k\right],\mu=E[X]$.

Fonction génératrice des moments : La f.g.m. d'une variable X est :
$$\phi(t)=E\left[e^{tX}\right]$$

Cas discret
$$\phi(t)=\sum_x e^{tx}p(x)$$
Cas continu
$$\phi(t)=\int_{-\infty}^{+\infty} e^{tx}f(x)dx$$

Propriétés : La fonction génératrice des moments caractérise entièrement la variable aléatoire. La f.g.m. de la somme de deux variables est le produit de leurs f.g.m.

Espérance conditionnelle : C'est l'espérance de la loi conditionnelle :

Cas discret, Cas continu
$$E[X\mid Y=y]=\sum_x x.p_{X|Y}(x\mid y) \qquad \int_{-\infty}^{+\infty} x.f_{X|Y}(x\mid y)dx$$

Théorème : Espérance conditionnelle de X sachant Y :
$$E[X]=E\left[E[X\mid Y]\right]=E\left[r(Y)\right]$$

Variance conditionnelle : C'est la variance de la loi conditionnelle :
$$\mathrm{Var}(X\mid Y=y)=E\left[X^2\mid Y=y\right]-E[X\mid Y=y]^2$$
$$\mathrm{Var}(X)=E\left[\mathrm{Var}(X\mid Y)\right]+\mathrm{Var}\left(E[X\mid Y]\right)$$

Processus de Branchement

Définition : Un processus stochastique est une collection de variables aléatoires $X_t(\omega)$, $t\in T$ et $\omega\in\Omega$.

Nombre de descendants : $X_n=\sum_{i=1}^{X_{n-1}} Z_i$

Espérance
$$E[X_n]=\mu^n,\mu=E[Z]$$
Variance
$$\mathrm{Var}(X_n)=\sigma^2\mu^{n-1}+\mu^2\,\mathrm{Var}(X_{n-1})$$

Extinction ou explosion ?

	$\mu<1$	$\mu=1$	$\mu>1$
$E[X]$	$\searrow_0$	$=1$	$\nearrow^{+\infty}$
$\mathrm{Var}(X)$	$\searrow_0$	$=n\sigma^2$	$\nearrow^{+\infty}$
	S'éteint toujours	S'éteint toujours	Ne s'éteint pas toujours

Probabilité que le processus s'éteigne :
$$\pi_0=\lim_{n\to\infty}P_0(n) \text{ et } P_0(n)=\phi\left(P_0^{(n-1)}\right)$$
π_0 est tel que $\pi_0=\phi(\pi_0)$ avec $\phi(0)=P_0$ et $\phi(1)=1$

Marche aléatoire

Formalisation : La position entre 0 et N au temps n est donnée par X_n. Le point de départ i est X_0. Le gain, ou variable aléatoire du déplacement à l'étape n est I_n.
$$P(I_n=1)=p,P(I_n=-1)=q\Rightarrow X_n=X_0+\sum_{k=1}^n I_k$$

Notations : P_i est la probabilité d'arriver en N étant parti de i. Q_i est la probabilité d'arriver en 0, étant parti de i.
$$P_i=P_{i+1}p+P_{i-1}q\Rightarrow P_i=C.r^i\Rightarrow P_i=\begin{cases}\dfrac{1-(q/p)^i}{1-(q/p)^N},q\neq p\\[2mm] \dfrac{i}{N},q=p\end{cases}$$

Conditions limites : Lorsque $N\to\infty$:
$$p>q\Rightarrow P_i\to 1-(q/p)^i \text{ ou } p<q\Rightarrow P_i\to\frac{(q/p)^i}{(q/p)^N}\to 0$$

Chaînes de Markov

Définition : Une chaîne de Markov est un processus discret et à temps discret, noté X_n, $n\in\mathbf{N}$, tel que :
$$P\left(X_{n+1}=j\mid X_n=i,X_{n-1}=...,X_1=...\right)=P\left(X_{n+1}=j\mid X_n=i\right)$$

Homogénéité : On ne considérera que des chaînes homogènes :
$$\forall n,P\left(X_{n+1}=j\mid X_n=i\right)=P_{ij}$$

Matrice de transition : Une chaîne de Markov est complètement caractérisée par les P_{ij}, qui forment la matrice de transition :
$$\overline{\overline{P}}=\begin{bmatrix}P_{11}&P_{12}&\cdots&P_{1N}\\P_{21}&P_{22}&\cdots&P_{2N}\\\vdots&\vdots&\ddots&\vdots\\P_{N1}&P_{N2}&\cdots&P_{NN}\end{bmatrix}$$
Cette matrice est **stochastique**, elle a les propriétés suivantes :
$$0\le P_{ij}\le 1 \text{ et } \forall i,\sum_j P_{ij}=1$$

Probabilités simultanées :
$$\pi_0(i)=P(X_0=i),i=1\ldots N \text{ et } P_{ij}^{(n)}=P\left(X_{n+m}=j\mid X_m=i\right)$$

Théorème de Kolmogorov :
$$\overline{\overline{P}}^{(n+m)}=\overline{\overline{P}}^{(n)}\overline{\overline{P}}^{(m)}$$
Corollaire : La matrice de transition à n étages est la matrice de transition élevée à la puissance n.

Classification des états :
- Une **classe** est l'ensemble des états qui communiquent entre eux.
- Une chaîne composée d'une seule classe est **irréductible**.
- Un état est **absorbant** si P_{ii}=1. On ne peut pas en sortir.
- Un état est **récurrent** si on repasse par cet état une infinité de fois.
- Un état est **transitoire** dans le cas contraire.

Période des états : un état i est dit périodique de période d si :
$$\begin{cases}P_{ii}^{(n)}=0 \text{ si } n\neq kd\\P_{ii}^{(n)}>0 \text{ si } n=kd\end{cases}$$
Une chaîne **ergodique** est à la fois irréductible et apériodique.

Probabilités stationnaires : Si une chaîne est ergodique, on a :
$$\text{si } \lim_{n\to\infty}P_{ij}^{(n)}=\pi_j,\pi^t=\pi^t\overline{\overline{P}} \text{ avec } \pi_n\underset{n\to\infty}{\to}\pi$$

Propriété d'ergodicité : La proportion d'apparition de l'état i sur le long terme est donnée par :
$$\pi_i=\lim_{n\to\infty}P(X_n=i)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\Im(x_k=i)$$

Condition de réversibilité : Une chaîne est réversible dans le temps si son déroulement dans le temps ou inversement au temps est identique :
$$Q_{ij}=P_{ij}\Leftrightarrow\pi_iP_{ij}=\pi_jP_{ji}$$

Processus de Poisson

Absence de mémoire : Une variable aléatoire X est dite sans mémoire si et seulement si :

$$\forall(s,t), P(X > s+t \mid X > t) = P(X > s)$$

Une propriété caractéristique de la fonction d'une telle variable est :

$$\overline{F}(s+t) = \overline{F}(s)\overline{F}(t)$$

La seule variable aléatoire continue ayant la propriété d'absence de mémoire est la variable aléatoire **exponentielle.**

Définitions :

- Un **processus de comptage** $\{N(t),t\geq 0\}$ est le nombre d'évènements aléatoires entre 0 et t.
- Un processus de comptage est dit à **accroissement indépendant** si pour deux intervalles A et B disjoints, $N(A)$ et $N(B)$ sont indépendants.
- Un processus de comptage est à **accroissement stationnaire** si $N(]s,t[)=N(t-s)$; il dépend de l'amplitude de l'intervalle, non de sa position.

Déf. 1/2 : Un processus de comptage est un processus de Poisson ssi :
(i) $N(0)=0$
(ii) Il est à accroissement indépendant et stationnaire
(iii) $N(t) \sim \mathrm{Poiss}(\lambda t)$ ou bien $\begin{cases} P(N(t)=1) = \lambda(t) + o(t) \\ P(N(t)=2) = o(t) \end{cases}$

Déf. 3 : Soient les T_i des variables aléatoires $Exp(\lambda)$ indépendantes, et $S_n = \sum_{i=1}^{n} T_i$. Alors $N(t) = \max\{n : S_n \leq t\}$ est un processus de Poisson.

Processus de Poisson inhomogène : Un processus de comptage est un processus de Poisson inhomogène d'intensité $\lambda(t)$ ssi :
(i) $N(0)=0$
(ii) $N(t)$, $t\geq 0$ a des incréments indépendants
(iii) $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$
(iv) $P(N(t+h) - N(t) = 2) = +o(h)$

$$N(t+s) - N(s) \sim \mathrm{Poiss}\left(\int_s^{s+t} \lambda(u)\,du\right)$$

Processus de Poisson composé : C'est un processus tel que :

$$X(t) = \sum_{i=1}^{N(t)} Y_i$$

$$E[X(t)] = \lambda t E[Y] \text{ et } Var(X(t)) = \lambda t E[Y^2]$$

Files d'attente

Système M/M/1 : L'arrivée des clients se fait selon un processus de Poisson, le temps de service est exponentiel, et il n'y a qu'un serveur.

- Plusieurs serveurs : **M/M/k**
- Temps de service générique : **M/G/k**
- Arrivée des clients générique : **G/G/k**

Grandeurs et notations :

- Nombre moyen de clients dans le système : $L = \lim_{t\to\infty} \frac{1}{t}\int_0^t X(t)\,dt$
- Nombre moyen de clients dans la file : $L_q = \lim_{t\to\infty} \frac{1}{t}\int_0^t X_q(t)\,dt$
- Temps moyen passé par un client dans le système : $W = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} W_i$

Equations de coût : Le taux moyen de gain du système est, λ_a est le taux d'arrivée des requêtes et k le nombre de serveurs :

$$\lambda_a E[S] \leq k$$

Equations d'équilibre : a_n est la proportion de clients qui trouvent n clients dans le système en arrivant, et d_n, est la proportion de clients qui laissent n clients dans le système en partant. Il se trouve que :

$$a_n = d_n$$

Caractéristiques du serveur M/M/1 : $P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$

$$L = \frac{\lambda}{\mu - \lambda} \quad W = \frac{1}{\mu - \lambda} \quad W_q = \frac{\lambda}{\mu}\frac{1}{\mu - \lambda} \quad L_q = \frac{\lambda^2}{\mu}\frac{1}{\mu - \lambda}$$

Processus de naissance et de mort : Il est paramétré par un ensemble de couples (v_n,p_n) :

$$\lambda_n + \mu_n = v_n \text{ et } \frac{\lambda_n}{\lambda_n + \mu_n} = \rho_n$$

$$P_0 = \left[1 + \sum_{n=1}^{+\infty} \frac{\lambda_0\lambda_1 \ldots \lambda_{n-1}}{\mu_1\mu_2 \ldots \mu_n}\right]^{-1}$$

$$L = \sum_{n=1}^{+\infty} nP_0 \frac{\lambda_0\lambda_1 \ldots \lambda_{n-1}}{\mu_1\mu_2 \ldots \mu_n} \qquad \overline{\lambda} = \sum_{n=0}^{+\infty} \lambda_n P_n \qquad W = \frac{\sum_{n=0}^{+\infty} nP_n}{\sum_{n=0}^{+\infty} \lambda_n P_n}$$

Variables aléatoires

Modélisation probabiliste discrète :

- <u>Variable aléatoire uniforme discrète</u>

$$P(X = x) = \frac{1}{n+1}, k \leq x \leq k+n$$

- <u>Variable aléatoire de Bernoulli</u>

$$P(X = 1) = P(\text{"}succès\text{"})$$
$$E[X] = p, \mathrm{Var}(X) = p(1-p)$$

- <u>Variable aléatoire géométrique</u>

$$P(X = k) = (1-p)^{k-1} p, k = 1,2,\ldots$$
$$E[X] = \frac{1}{p}, \mathrm{Var}(X) = \frac{1-p}{p^2}$$

- <u>Variable aléatoire binomiale</u>

$$P(X = k) = p^k (1-p)^{n-k} C_n^k, k = 0,1,\ldots,n$$
$$E[X] = np, \mathrm{Var}(X) = np(1-p) \qquad \phi(t) = \left(pe^t + 1 - p\right)^n$$

- <u>Variable aléatoire de Poisson</u>

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, k = 0,1,\ldots, \lambda > 0$$
$$E[X] = \lambda, \mathrm{Var}(X) = \lambda \qquad \phi(t) = e^{\lambda\left(e^t - 1\right)}$$

Modélisation probabiliste continue :

- <u>Variable aléatoire uniforme</u>

$$a < x < b, f(x) = c = \frac{1}{b-a} \qquad F(x) = \frac{x-a}{b-a}$$
$$E[X] = \frac{a+b}{2}, \mathrm{Var}(X) = \frac{(a-b)^2}{12}$$

- <u>Variable aléatoire exponentielle</u>

$$x > 0, f(x) = \lambda e^{-\lambda x} \qquad F(x) = -e^{-\lambda x}$$
$$E[X] = \frac{1}{\lambda}, \mathrm{Var}(X) = \frac{1}{\lambda^2} \qquad \phi(t) = \frac{\lambda}{\lambda - t}$$

- <u>Variable aléatoire gaussienne</u>

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$
$$E[X] = \mu, \mathrm{Var}(X) = \sigma^2 \qquad \phi(t) = \exp\left(\frac{\sigma^2 t^2}{2} + \mu t\right)$$

- <u>Variable aléatoire Gamma</u>

$$x \geq 0, f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \text{ où } \Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1}\,dx$$
$$E[X] = \frac{\alpha}{\lambda}, \mathrm{Var}(X) = \frac{\alpha}{\lambda^2} \qquad \phi(t) = \frac{\lambda^\alpha}{(\lambda - t)^\alpha}$$