Probabilités et statistiques

IV - Processus aléatoires

G. Chênevert

4 décembre 2023

JUNIA ISEN

Au menu aujourd’hui

Processus aléatoires

Chaînes de Markov

Processus continus

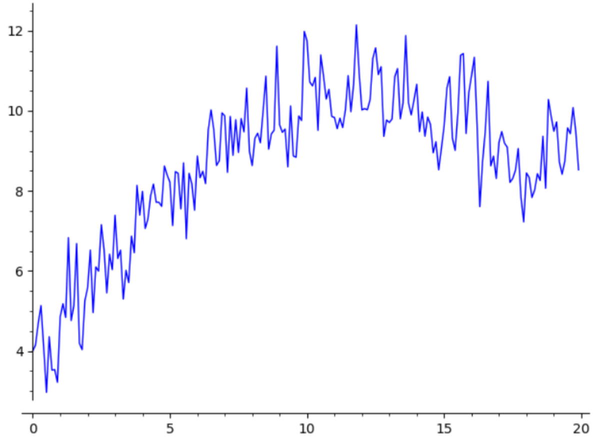
Notion de processus aléatoire

Modélisation d'un phénomène impliquant des v.a. correspondant à *plusieurs expériences aléatoires (dépendantes)* effectuées séquentiellement dans le temps

Exemples :

- cours de la bourse
- bruit sur un signal
- désintégration radioactive
- charge d'un serveur
- $\vdots$

Un exemple



Formalisme

Définition

Un **processus aléatoire** (ou **stochastique**, ou une **fonction aléatoire**) est la donnée d'une famille de variables aléatoires

$$X_t = X(t), \quad t \in \mathcal{T}.$$

Techniquement, puisqu'une variable aléatoire est une fonction $\Omega \to \mathbb{R}$, il s'agit donc d'une fonction de deux variables

$$\begin{aligned} X : \Omega \times \mathcal{T} &\longrightarrow \mathbb{R} \\ (\omega, t) &\mapsto X_t(\omega) = X(\omega, t) \end{aligned}$$

En d'autres termes

- Pour chaque $t \in \mathcal{T}$ fixé on a une variable aléatoire

$$X(t)$$

- En particulier, pour $\mathcal{T} = \{t_1, t_2, \dots\}$ discret, ce n'est qu'une suite de v.a.

$$X_{t_1}, X_{t_2}, \dots$$

- Chaque observation $\omega \in \Omega$ donne lieu à une **réalisation**

$$x(t) := X(\omega, t)$$

Au menu aujourd'hui

Processus aléatoires

Chaînes de Markov

Processus continus

Revenons aux pièces

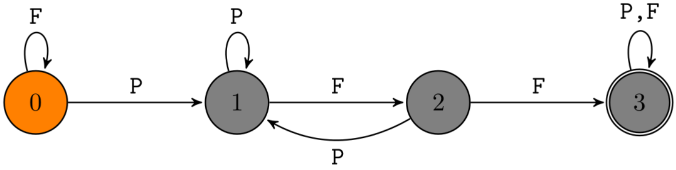
On peut se demander...

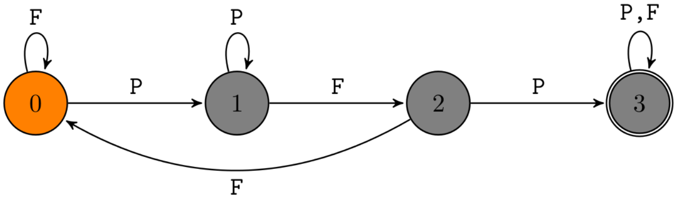
dans une suite de lancers d'une pièce équilibrée, quel est le nombre moyen de lancers nécessaires pour voir apparaître

$$\texttt{PFF}\,? \qquad \text{et} \qquad \texttt{PFP}\,?$$

Sur 3 lancers indépendants, chacun se produit avec probabilité $\frac{1}{8}$ mais...

Machines d’états





Une classe de processus aléatoires discrets

Définition

Une **chaîne de Markov** est un processus aléatoire à temps discret (disons $\mathcal{T} = \mathbb{N}$) à valeurs dans un ensemble fini (ou discret) $\mathcal{Q}$ d'**états**.

$$X_n = \text{numéro de l'état à la } n^{\text{e}} \text{ étape}$$

On suppose les probabilités de transition *constantes dans le temps* :

$$p_{ij} := \mathbb{P}[X_{n+1} = j \mid X_n = i]$$

(ne dépend pas de n).

Formule des espérances

On s'intéresse souvent au temps de parcours pour atteindre un état absorbant.

Soit $Y_i(n)$ le nombre de transitions avant l'absorption en partant de l'état i après n étapes.

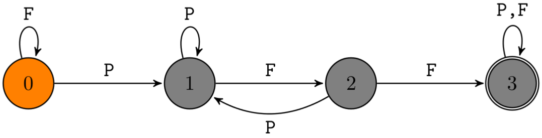
On a alors, pour tout i non absorbant :

$$Y_i(n) = 1 + \sum_j p_{ij} \, Y_j(n+1),$$

$$\implies \mathbb{E}[Y_i] = 1 + \sum_j p_{ij} \, \mathbb{E}[Y_j]$$

ce qui donne un système d'équations linéaires qu'on peut résoudre pour les $\mu_j := \mathbb{E}[Y_j]$.

Exemple : PFF



$$\begin{bmatrix}\mu_0\\ \mu_1\\ \mu_2\\ \mu_3\end{bmatrix} = \begin{bmatrix}1\\1\\1\\0\end{bmatrix} + \begin{bmatrix}\frac{1}{2} & \frac{1}{2} & 0 & 0\\ 0 & \frac{1}{2} & \frac{1}{2} & 0\\ 0 & \frac{1}{2} & 0 & \frac{1}{2}\\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\mu_0\\ \mu_1\\ \mu_2\\ \mu_3\end{bmatrix} \implies \begin{bmatrix}\mu_0\\ \mu_1\\ \mu_2\\ \mu_3\end{bmatrix} = \begin{bmatrix}8\\6\\4\\0\end{bmatrix}$$

Alors que pour PFP, on trouve $\mu_0 = 10$ (vérifiez !)

Point de vue matriciel

$$\mathbb{P}[X_{n+1} = i] = \sum_j p_{ji}\,\mathbb{P}[X_n = j]$$

i.e. si $\vec{v}_n$ désigne la distribution de probabilité à l'étape n,

$$\vec{v}_{n+1} = A\,\vec{v}_n \quad \text{avec} \quad A = [p_{ji}].$$

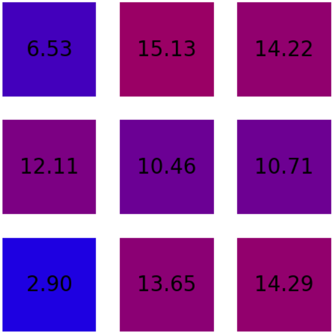
Sous certaines conditions il existe un unique point fixe

$$\vec{v}_\infty = A\,\vec{v}_\infty$$

(état stationnaire du système) vers lequel le processus converge.

Exemple : sauts de grenouille

Passage uniforme à une case voisine + 15 % de téléportation à un case quelconque



Ce type de modèle (marche aléatoire sur un graphe) a été appliqué avec un un certain succès à la définition d'un score de pertinence d'une page web en 1998

Marche aléatoire sur $\mathbb{Z}$

On considère un processus aléatoire (X_n) avec $X_0 = 0$,

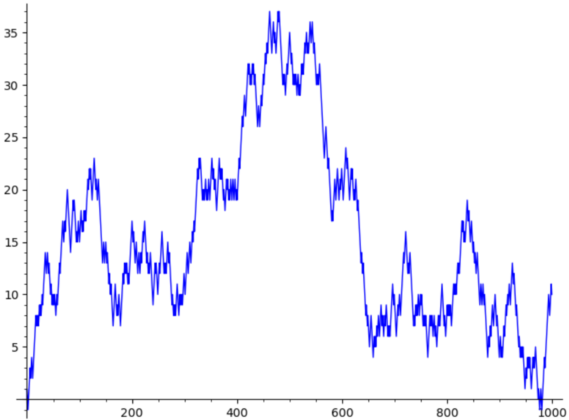
$$X_{n+1} = X_n + U_n, \qquad U_n \sim \mathcal{U}(\{-1, +1\})$$

Probabilité d'un retour à 0 en $2n$ étapes :

$$\binom{2n}{n} \frac{1}{2^{2n}} \approx \frac{1}{\sqrt{\pi n}}$$

On montre qu'on repasse presque sûrement par 0 un jour

Voyez plutôt



Marche aléatoire sur $\mathbb{Z}^2$

On considère un couple (X_n, Y_n) avec

$$(X_{n+1}, Y_{n+1}) = (X_n, Y_n) + \vec{U}_n, \quad \vec{U}_n \sim \mathcal{U}\big(\{(\pm 1, 0), (0, \pm 1)\}\big).$$

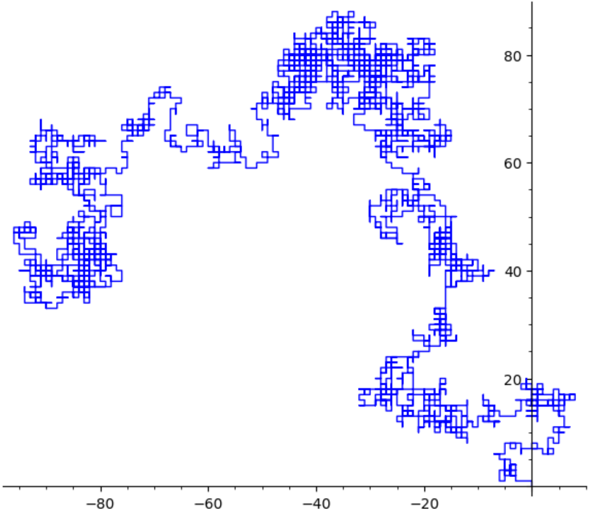
On est encore presque sûr de revenir à l'origine

(même si la durée espérée du parcours est infinie)

Fait : en 3D,

$$\mathbb{P}[\text{retour}] \approx 28{,}22\,\% \text{ (!)}$$

Marche aléatoire 2D



Au menu aujourd'hui

Processus aléatoires

Chaînes de Markov

Processus continus

Processus continus

La description d'un processus aléatoire $X(t)$ est en général un peu plus délicate.

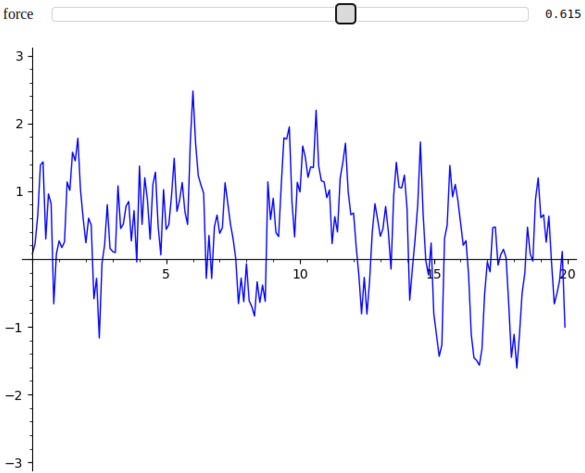
Première étape : décrire la loi de *chaque* v.a. $X(t)$ (t fixé)

Restera alors étudier la *dépendance* entre les $X(t)$ à différents instants...

On fait typiquement des hypothèses simplificatrices pour éviter d'avoir à travailler avec une densité de probabilité sur un espace de dimension infinie

$$f(x_t)_{t \in \mathcal{T}} \qquad \dots$$

$X(t) \sim \mathcal{N}(0,1)$ plus ou moins dépendants



Statistiques d'ordre 2

On étudie la distribution des couples aléatoires

$$\big(X(t_1),\, X(t_2)\big).$$

Si le processus est **stationnaire** (invariant par translation temporelle), cela se ramène à l'étude de

$$\big(X(s),\, X(s+t)\big)$$

qui ne dépend que de $t \in \mathbb{R}$.

Autocovariance

Pour mesurer à quel point les valeurs de $X(t)$ à différents instants sont corrélées, on définit pour un processus stationnaire

$$r_X(t) := \mathrm{Cov}(X(s), X(s+t)) = \mathbb{E}[X(s)\,X(s+t)] - \mu^2,$$

voire

$$r_X(t) := \mathrm{Cov}(\overline{X(s)}, X(s+t)) = \mathbb{E}\left[\overline{X(s)}\,X(s+t)\right] - |\mu|^2$$

pour un signal stationnaire à valeurs complexes.

Remarque : souvent appelée abusivement « fonction d'autocorrélation » même lorsqu'elle n'est pas normalisée en divisant par σ^2

Interprétation fréquentielle ($\mu = 0$)

Écrivons notre signal à la Fourier (par rapport à t) :

$$X(t) = \int_{-\infty}^{+\infty} \widehat{X}(f)\, e^{+2\pi \mathrm{i} f t}\, \mathrm{d}f$$

d'où

$$\overline{X}(s) = \int_{-\infty}^{+\infty} \overline{\widehat{X}(f)}\, e^{-2\pi \mathrm{i} f s}\, \mathrm{d}f$$

$$X(s+t) = \int_{-\infty}^{+\infty} \widehat{X}(f)\, e^{+2\pi \mathrm{i} f (s+t)}\, \mathrm{d}f$$

et alors

$$\overline{X}(s)\,X(s+t) = \int_{-\infty}^{+\infty} |\widehat{X}(f)|^2\, e^{+2\pi \mathrm{i} f t}\,\mathrm{d}f$$

d'où

$$r_X(t) = \mathbb{E}\big[\overline{X}(s)\,X(s+t)\big] \;= \int_{-\infty}^{+\infty} \mathbb{E}\Big[|\widehat{X}(f)|^2\Big]\; e^{+2\pi \mathrm{i} f t}\,\mathrm{d}f$$

i.e. moralement (quelques subtilités quand même !)

Théorème (Wiener-Khintchine)

$$\widehat{r_X}(f) = \mathbb{E}\Big[|\widehat{X}(f)|^2\Big]\,.$$

Densité spectrale de puissance

Définition

$$s_X(f) := \mathbb{E}\left[|\widehat{X}(f)|^2\right]$$

Le théorème de Wiener-Khintchine peut donc se reformuler :

$$\boxed{\widehat{r_X}(f) = s_X(f)}$$

la densité spectrale de puissance est la transformée de la fonction d'autocovariance.

Interprétation

Énergie totale d'un signal déterministe $x(t)$:

$$E = \int_{-\infty}^{+\infty} |x(t)|^2 \, \mathrm{d}t \quad \Longrightarrow \quad |x(t)|^2 \text{ puissance instantanée.}$$

Mais aussi (théorème de Parseval) :

$$E = \int_{-\infty}^{+\infty} |\widehat{x}(f)|^2 \, \mathrm{d}f \quad \Longrightarrow \quad |\widehat{x}(f)|^2 \text{ densité spectrale de puissance.}$$

Pour les signaux aléatoires, on travaille avec l'espérance de cette densité.

Exemple : bruit blanc

On dit que $X(t)$ est **blanc** si

$$s_X(f) = s_0 \quad \text{constante.}$$

On a alors $r_X(t) = s_0\,\delta(t)$... variance infinie !

Préférons-lui un **bruit blanc échantillonné** à la fréquence f_e :

$$s_X(f) = s_0\,\Pi_{f_e}(f)$$

$$r_X(t) = f_e s_0\,\mathrm{sinc}(\pi f_e t)$$

Bruit blanc échantillonné gaussien : lorsque

$$X_n := X(nt_e) \sim \mathcal{N}(0,\sigma), \quad \sigma^2 = f_e s_0.$$

Par exemple si les X_n sont indépendantes (mais pas que).

Condition d'ergodicité

En général, la condition souhaitée est que la loi des grands nombres s'applique :

avec

$$\mathrm{moy}_t(X) := \frac{1}{t}\int_0^t X(u)\,\mathrm{d}u$$

on veut

$$\mathrm{moy}_t(X) \underset{t\to\infty}{\longrightarrow} \mu \qquad \text{presque sûrement}$$

i.e. que les moyennes temporelles convergent vers l'espérance distributionnelle

On parle alors de **processus ergodique**

Filtrage

Passons ce signal $X(t)$ dans un filtre $\mathcal{F}$ avec réponse impulsionnelle $h(t)$.

On obtient comme sortie la convolution $Y(t) = h(t) * X(t)$.

Fréquentiellement $\widehat{Y}(f) = \widehat{h}(f) \cdot \widehat{X}(f)$ donc en prenant $\mathbb{E}\left[|\cdot|^2\right]$:

$$\boxed{s_Y(f) = |\widehat{h}(f)|^2 \cdot s_X(f).}$$

$\Longrightarrow$ goto Traitement de signal !


That's all Folks!