

Leçon 16

Existence de suites infinies de variables aléatoires indépendantes

1. Famille finie de variables aléatoires indépendantes
2. Construction d'un jeu de pile ou face infini
3. Suite quelconque de variables aléatoires réelles indépendantes

Exercices

L'objet de cette leçon est d'essayer de répondre à la question (naturelle) suivante : existe-t-il des suites infinies X_n , $n \in \mathbb{N}$, de variables aléatoires (réelles) sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, mutuellement indépendantes et de lois données. Cette question est au cœur de plusieurs énoncés classiques de la théorie des probabilités, comme par exemple la loi des grands nombres (Leçon 19).

1 Famille finie de variables aléatoires indépendantes

Afin de mettre la question initiale en perspective, et de comprendre les enjeux, il est bénéfique de rappeler pour commencer la situation d'une famille finie. Autrement dit, étant donné des lois de probabilités fixées $P_1, \dots, P_k$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (ou plus généralement sur un espace mesurable $(E, \mathcal{B})$, ou même des espaces différents $(E_1, \mathcal{B}_1), \dots, (E_k, \mathcal{B}_k)$), existe-t-il un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et des variables aléatoires $X_1, \dots, X_k$ sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, indépendantes et telles que X_n a pour loi P_n ($\mathbb{P}_{X_n} = P_n$), $n = 1, \dots, k$?

Si $k = 1$, le procédé est maintenant connu : prendre $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}(\mathbb{R})$, $\mathbb{P} = P_1$ et $X_1 : \mathbb{R} \rightarrow \mathbb{R}$ l'application identité ($X_1(\omega) = \omega$, $\omega \in \Omega = \mathbb{R}$). Alors la mesure image de $\mathbb{P} = P_1$ par l'application identité X_1 est P_1 , et donc X_1 a pour loi P_1 , puisque pour tout borélien B de $\mathbb{R}$,

$$\mathbb{P}(X_1 \in B) = \mathbb{P}(\{\omega \in \Omega; X_1(\omega) \in B\}) = \mathbb{P}(B) = P_1(B).$$

Si $k \geq 1$ est quelconque fixé, le procédé est en fait le même sur la base de la définition des mesures produit : prendre $\Omega = \mathbb{R} \times \dots \times \mathbb{R} = \mathbb{R}^k$, $\mathcal{A} = \mathcal{B}(\mathbb{R}^k)$, $\mathbb{P} = P_1 \otimes \dots \otimes P_k$ et $X_n : \mathbb{R}^k \rightarrow \mathbb{R}$, $n = 1, \dots, k$, les applications coordonnées,

$$X_n(\omega) = \omega_n, \quad \omega = (\omega_1, \dots, \omega_k) \in \Omega = \mathbb{R}^k.$$

Alors, les X_n sont indépendantes et chaque X_n a pour loi P_n , $n = 1, \dots, k$; en effet, pour tous boréliens $B_1, \dots, B_k$ de $\mathbb{R}$, par définition d'une mesure produit,

$$\begin{aligned}
\mathbb{P}((X_1, \dots, X_k) \in B_1 \times \dots \times B_k) &= \mathbb{P}(\{\omega \in \Omega; (X_1(\omega), \dots, X_k(\omega)) \in B_1 \times \dots \times B_k\}) \\
&= \mathbb{P}(\{\omega \in \Omega; (\omega_1, \dots, \omega_k) \in B_1 \times \dots \times B_k\}) \\
&= \mathbb{P}(B_1 \times \dots \times B_k) \\
&= P_1 \otimes \dots \otimes P_k(B_1 \times \dots \times B_k) \\
&= P_1(B_1) \dots P_k(B_k).
\end{aligned}$$

En choisissant $B_2 = \dots = B_k = \mathbb{R}$, ces égalités montrent en particulier que $\mathbb{P}(X_1 \in B_1) = P_1(B_1)$, donc que X_1 a pour loi P_1 , et de même pour tout $n = 1, \dots, k$, X_n a pour loi P_n . Et ainsi, pour donc tous boréliens $B_1, \dots, B_k$ de $\mathbb{R}$,

$$\mathbb{P}((X_1, \dots, X_k) \in B_1 \times \dots \times B_k) = \mathbb{P}(X_1 \in B_1) \dots \mathbb{P}(X_k \in B_k)$$

ce qui démontre l'indépendance mutuelle des $X_1, \dots, X_k$.

Maintenant il est clair que pour une suite infinie X_n , $n \geq 1$, la construction précédente va poser problème puisqu'elle va nécessiter de considérer un produit infini $P_1 \otimes \dots \otimes P_k \otimes \dots$ qui n'est pas nécessairement bien défini. La réponse, dans un cadre général, a été apportée par A. Kolmogorov. Il est présenté ici une construction dans le cas particulier où les lois P_k sont des lois de Bernoulli de paramètre $\frac{1}{2}$.

2 Construction d'un jeu de pile ou face infini

Rappeler pour commencer qu'un nombre $\omega \in [0, 1[$ admet un développement binaire

$$\omega = \sum_{n=1}^{\infty} \frac{1}{2^n} X_n(\omega)$$

avec $X_n(\omega) = 0$ ou 1 , unique en convenant de ne prendre pour les nombres dyadiques (de la forme $\frac{k}{2^\ell}$) que le développement qui n'a que des 0 à partir d'un certain rang (par exemple, le développement binaire de $\frac{3}{8}$ sera $0, 1, 1, 0, 0, 0 \dots$ plutôt que $0, 1, 0, 1, 1, 1, \dots$).

Prendre alors $\Omega = [0, 1[$, $\mathcal{A} = \mathcal{B}([0, 1[) = \mathcal{B}(\mathbb{R}) \cap [0, 1[$, $\mathbb{P} = \lambda$ la mesure de Lebesgue, et les X_n , $n \geq 1$, ainsi définies. Si $(\varepsilon_1, \dots, \varepsilon_k, \varepsilon_{k+1})$, $k \geq 1$, est une suite finie de 0 et de 1 fixée, l'événement

$$I = \{X_1 = \varepsilon_1, \dots, X_k = \varepsilon_k\}$$

correspond à l'intervalle dyadique de $[0, 1[$ des ω dont le développement commence par $\sum_{n=1}^k \frac{1}{2^n} \varepsilon_n$. Par exemple, si $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (0, 1, 1)$, $I = [\frac{3}{8}, \frac{1}{2}[$. Par définition du développement binaire, la variable X_{k+1} prend la valeur 0 sur la première moitié de cet intervalle, et 1 sur la seconde. Ainsi, puisque $\mathbb{P} = \lambda$,

$$\mathbb{P}(I \cap \{X_{k+1} = \varepsilon_{k+1}\}) = \frac{1}{2} \mathbb{P}(I).$$

Par itération,

$$\mathbb{P}(X_1 = \varepsilon_1, \dots, X_k = \varepsilon_k, X_{k+1} = \varepsilon_{k+1}) = \frac{1}{2^{k+1}}.$$

Ce résultat prouve deux choses. D'abord, les X_n , $n \geq 1$, sont toutes des variables de Bernoulli équilibrées sur $\{0, 1\}$. En effet, $\mathbb{P}(X_1 = \varepsilon_1) = \frac{1}{2}$,

$$\mathbb{P}(X_2 = \varepsilon_2) = \mathbb{P}(X_1 = 1, X_2 = \varepsilon_2) + \mathbb{P}(X_1 = 0, X_2 = \varepsilon_2) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

et ainsi de suite. Ensuite, elles sont mutuellement indépendantes puisque pour tout $k \geq 1$,

$$\mathbb{P}(X_1 = \varepsilon_1, \dots, X_k = \varepsilon_k) = \frac{1}{2^k} = \mathbb{P}(X_1 = \varepsilon_1) \cdots \mathbb{P}(X_k = \varepsilon_k).$$

Cette construction fournit donc bien une suite X_n , $n \geq 1$, de variables aléatoires indépendantes, sur l'espace probabilisé $([0, 1[, \mathcal{B}([0, 1[), \lambda)$, telle que chaque X_n est de loi de Bernoulli $\mathcal{B}(\frac{1}{2})$.

3 Suite quelconque de variables aléatoires réelles indépendantes

En s'appuyant sur des exercices vus précédemment, il est possible d'étendre la construction précédente à des suites X_n , $n \in \mathbb{N}$, de variables réelles de lois données P_n , $n \in \mathbb{N}$, quelconques (par exemple les P_n sont des lois de Poisson, ou des lois normales, ou des mélanges des deux etc.).

L'étude débute par le cas où P_n est uniforme sur $[0, 1[$, $P_n = \mathcal{U}(0, 1)$, $n \in \mathbb{N}$. Il est présenté dans l'Exercice 1 que si X_n , $n \in \mathbb{N}$, est une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre $\frac{1}{2}$ sur $\{0, 1\}$, alors la série $\sum_{n \in \mathbb{N}} \frac{1}{2^n} X_n$ converge et est de loi $\mathcal{U}(0, 1)$ sur $[0, 1[$ (c'est en fait la réciproque de la construction du paragraphe précédent !). (Comme à l'habitude, il n'est pas distingué ici et plus loin si $\mathcal{U}$ est uniforme sur l'intervalle $(0, 1)$ ouvert, fermé, semi-ouvert, toutes ces mesures étant identiques puisque la mesure de Lebesgue ne charge pas les points.)

L'étape suivant considère une suite double $X_{m,n}$, $(m, n) \in \mathbb{N}^2$, de variables aléatoires indépendantes de loi de Bernoulli de paramètre $\frac{1}{2}$ sur $\{0, 1\}$ construite dans le paragraphe précédent (indexer $\mathbb{N} \times \mathbb{N}$ par $\mathbb{N}$). D'après ce qui précède, pour chaque $n \in \mathbb{N}$ fixé, la série $U_n = \sum_{m \in \mathbb{N}} \frac{1}{2^m} X_{m,n}$ définit une variable aléatoire uniforme sur $[0, 1[$. Les U_n , $n \in \mathbb{N}$, sont mutuellement indépendantes (en partitionnant la suite double $X_{m,n}$, $(m, n) \in \mathbb{N}^2$ suivant l'indice n) et de loi uniforme $\mathcal{U}(0, 1)$, et répondent donc à la construction attendue.

L'Exercice 2, Leçon 10, permet ensuite d'atteindre des lois arbitraires. Soit P_n , $n \in \mathbb{N}$, une famille quelconque de mesures de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$; soit, pour tout n , F_n la fonction de répartition de P_n , et $F_n^{(-1)}$ sa fonction inverse généralisée. Alors la suite $Z_n = F_n^{(-1)}(U_n)$, $n \in \mathbb{N}$, est constituée de variables aléatoires indépendantes (puisque les U_n le sont), et pour chaque $n \in \mathbb{N}$, $F_n^{(-1)}(U_n)$ a pour fonction de répartition F_n , et donc pour loi P_n . C'est bien le résultat attendu.

Exercices

(Une étoile * désignera une question de difficulté supérieure.)

Exercice 1. Soit X_n , $n \in \mathbb{N}$, une suite de variables aléatoires indépendantes de même loi de Bernoulli $\mathcal{B}(\frac{1}{2})$ sur $\{0, 1\}$; démontrer que la série $\sum_{n \in \mathbb{N}} \frac{1}{2^n} X_n$ converge presque sûrement, et montrer qu'elle suit la loi uniforme sur $[0, 1[$. (*Indication* : déterminer la mesure des ensembles dyadiques, ou décrire la fonction caractéristique de la somme partielle $S_k = \sum_{n=1}^k \frac{1}{2^n} X_n$.)