Leçon 249: Suites de variables de Bernoulli indépendantes

Adrien Fontaine

23 décembre 2012

Table des matières

Cadre : $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.

1 Constructions de suites de variables aléatoires de Bernoulli indépendantes

1.1 Définitions

Définition 1

On dit que deux évènements A et $B \in \mathcal{A}$ sont indépendants si :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Définition 2

Soient deux variables aléatoires X_1 et X_2 sur $\Omega, \mathcal{A}, \mathbb{P})$, à valeurs respectivement dans les espaces probabilisés $(\Omega_1, \mathcal{A}_1)$ et $(\Omega_2, \mathcal{A}_2)$. On dit que X_1 et X_2 sont indépendantes si pour tous $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$, on a :

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2 = \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2)$$

Remarque 1

Ces deux définitions se généralisent à un nombre quelconque fini d'évènements $(A_i)_{i \in I}$ (resp. de variables aléatoires $(X_i)_{i \in I}$). Pour toute partie finie J de I :

1. $$\mathbb{P}(\cap_{i \in J} A_i) = \prod_{i \in J} \mathbb{P}(A_i)$$
2. *(resp.)* $$\mathbb{P}(\cap_{i \in J} X_i \in A_i) = \prod_{i \in J} \mathbb{P}(X_i \in A_i)$$

Exemple 1

Lancé de dés indépendants (Ouvrard 1, p51)

Définition 3

Une variable aléatoire réelle X suit une loi de Bernoulli, de paramètre p, noté $X \sim \mathcal{B}(p)$, si $\mathbb{P}_X = p\delta_1 + (1-p)\delta_0$.

Exemple 2

- *Cette loi modélise un tirage pile ou face, ou aussi le tirage de boules blanches en proportion p et de boules noires en proportion $1-p$.*
- *Si $U \sim \mathcal{U}([0,1])$, alors $X = 1_{[0,p]}(U) \sim \mathcal{B}(p)$.*

1.2 Construction d'une suite finie

On peut construire un modèle probabiliste décrivant un jeu de pile ou face en n coups. Mathématiquement, cela revient à la construction d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et de n variables indépendantes $(X_i)_{i\in\{1,\dots,n\}}$ de même loi de Bernoulli. On peut prendre $\Omega = \{0,1\}^n$, muni de la probabilité uniforme

1.3 Construction d'une suite infinie

La généralisation dans le cas des suites infinies nous conduit à considérer $\Omega = \{0,1\}^{\mathbb{N}}$. Toutefois, l'existence d'une mesure de probabilité produit sur sur $\{0,1\}^{\mathbb{N}}$ n'est pas évidente (théorème de Kolmogorov, Ouvrard 2 p550)
Mais le problème posé possède une autre solution, où on prend pour Ω l'intervalle $[0,1[$ muni de la mesure de Lebesgue.

Définition 4

Soit $x \in [0,1[$. Définissons pour tout $x \in [0,1[$, les suites de terme général $D_n(x)$ et $R_n(x)$ par :

$$R_0(x) = x$$

et pour $n \in \mathbb{N}^*$,

$$D_n(x) = 2R_{n-1}(x) \text{ et } R_n(x) = 2R_{n-1}(x) - D_n(x)$$

On a alors

$$x = \sum_{j=1}^{+\infty} \frac{D_j(x)}{2^j}$$

$D_j(x)$ est un développement dyadique de x.

Démonstration : Ouvrard 2, p54 ■

Remarque 2

Le développement dyadique d'un réel n'est pas unique ; on a en effet

$$\forall n \in \mathbb{N}^*, \frac{1}{2^{n-1}} = \sum_{j=1}^{+\infty} \frac{1}{2^j}$$

mais cette définition de $D_n(x)$ donne un unique développement.

Proposition 1

Soit l'espace probabilisé $([0,1[, \mathcal{B}_{[0,1]}, \mathbb{P})$ où $\mathbb{P}$ est la mesure de Lebesgue sur $[0,1[$. Sur cet espace, la suite $(D_n)_{n\in\mathbb{N}^*}$ est une suite de variables indépendantes de même loi de Bernoulli $\mathcal{B}(\frac{1}{2})$.

Démonstration : Ouvrard 2, proposition 9.18p55 ■

Corollaire 1

Soit $(\mu_j)_{j\in\mathbb{N}^*}$ une suite de probabilités sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Il existe une suite de variables aléatoires réelles $(X_j)_{j\in\mathbb{N}^*}$ définies sur l'espace probabilisé $([0,1[, \mathcal{B}([0,1]), \mathbb{P})$ où $\mathbb{P}$ est la probabilité restriction de la mesure de Lebesgue à $[0,1[$, indépendantes, et telle que, pour tout $j \in \mathbb{N}^*$, X_j soit de loi μ_j.

Démonstration : Ouvrard 2, corollaire 9.19p58 ■

Remarque 3

Il est en fait, possible, à partir du modèle fondé sur $[0,1[$, de construire un modèle où l'espace fondamental est l'espace des suite $\{0,1\}^{\mathbb{N}}$, qui est le modèle naturel auquel nous faisions allusion en début de partie.

Démonstration : Ouvrard 2, complément p59 ■

Exemple 3

Dans une partie infinie de pile ou face, la probabilité de voir se réaliser une infinité de fois une suite $(\varepsilon_1, ..., \varepsilon_n)$ est égale à 1.

Démonstration : Ouvrard, exemple d'utilisation p60 ■

2 Construction de variables aléatoires à partir de variables aléatoires de Bernoulli

Proposition 2

Soit $(X_i)_{1\leq i\leq n}$ une suite finie de variables indépendantes de Bernoulli de paramètre p. Alors $S_n = \sum_{i=1}^n X_i$ suit une loi binomiale $\mathcal{B}(n,p)$, i.e

$$\forall k \in \{1,...,n\}\mathbb{P}(S_n = k) = \binom{n}{k} p^k(1-p)^{n-k}$$

Démonstration : Ouvrard 1, proposition 3.15p70 ■

Proposition 3

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables indépendantes de loi de Bernoulli $\mathcal{B}(p)$. On définit par récurrence (T_n) par :

$$T_1 = \inf\{k \geq 1, X_k = 1\} \text{ et } T_{n+1} = \inf\{k > T_n X_k = 1\}$$

Alors T_1, $T_2 - T_1$, ... , $T_n - T_{n-1}$, ..., sont i.i.d de loi géométrique $\mathcal{G}(p)$ (i.e $\mathbb{P}(T_1 = n) = p(1-p)^n, \forall n \in \mathbb{N}^*$).
De plus, $\forall n > 1$, $T_n \sim \mathcal{B}^-(n,p)$ loi binomiale négative i.e

$$\begin{cases} \mathbb{P}(T_n = l) = & \binom{l-1}{n-1} p^n(1-p)^{l-n} \text{ si } l \geq n \\ & 0 \text{ sinon} \end{cases}$$

Démonstration : Ouvrard 1, exercice 4.5p105 ■

Théorème 1 (*Théorème de Poisson*)

Soit (p_n) une suite de réels de l'intervalle $]0,1[$ telle que $\lim_{n\to+\infty} np_n = \lambda$ (où $\lambda > 0$). Considérons, pour chaque entier n, une variable aléatoire S_n de loi $\mathcal{B}(n,p_n)$. Alors pour tout entier

$k \in \mathbb{N}$, la suite de terme général $\mathbb{P}(S_n = k)$ est convergente et on a :

$$\lim_{n \to +\infty} \mathbb{P}(S_n = k) = \exp(-\lambda)\frac{\lambda^k}{k!}$$

Autrement dit, (S_n) converge en loi vers la loi de Poisson $\mathcal{P}(\lambda)$.

Démonstration : Ouvrard 1, théorème 7.1p226 ■

DÉVELOPPEMENT :

Théorème 2 (*Théorème des évènements rares de Poisson*)

Soit, pour tout $n \in \mathbb{N}^*$, une famille finie $\{A_{n,j}, 1 \leq j \leq M_n\}$ d'évènements indépendants définis sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On pose $\mathbb{P}(A_{n,j}) = p_{n,j}$, et on note

$$S_n = \sum_{j=1}^{M_n} 1_{A_{n,j}}$$

On suppose que la suite de terme général M_n tend en croissant vers $+\infty$, que

$$\max_{1 \leq j \leq M_n} p_{n,j} \to 0 \text{ et que } \sum_{j=1}^{M_n} p_{n,j} \to \lambda$$

où $\lambda > 0$. Alors la suite (S_n) converge en loi vers la loi de Poisson $\mathcal{P}(\lambda)$.

3 Applications

3.1 En statistiques

Définition 5

Une suite de variables aléatoires définies est dite convergente en probabilité vers la variable aléatoire X si pour tout $\varepsilon > 0$,

$$\lim_{n \to +\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$$

Théorème 3 (*Loi faible des grands nombres*)

Soit (X_n) une suite de variables aléatoires, indépendantes, de même loi, et admettant un moment d'ordre 2. Alors, la suite des variables aléatoires $\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ converge en probabilité vers $m = \mathbb{E}[X_1]$.

Démonstration : Ouvrard 1, théorème 7.9p234 ■

Corollaire 2 (*Théorème de Bernoulli*)

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'évènements indépendants et de même probabilité p. La suite des variables aléatoires $\frac{1}{n}\sum_{j=1} 1_{A_i}$ converge en probabilité vers p.

Remarque 4

Pour $A \in \mathcal{A}$, $\mathbb{P}(A)$ représente la probabilité d'occurrence de l'évènement A. Dans les premiers traités de théorie des probabilités, longtemps avant l'introduction de la théorie de la mesure, la probabilitée $\mathbb{P}(A)$ était définie de la manière suivante : on imagine qu'on répète l'expérience aléatoire un nombre N de fois, et on note N_A le nombre de répétitions pour lesquelles l'évènement A est réalisé ; alors, la proportion N_A/N converge quand $N \to +\infty$ vers la probabilité $\mathbb{P}(A)$.
Ce corollaire fait le lien entre notre approche axiomatique moderne et la définition historique de la probabilité comme fréquence d'apparition d'un évènement quand on répète un grand nombre de fois une expérience aléatoire.

Développement :
Théorème 4 (*Théorème de Bernstein*)

Soit $f : [0,1] \to \mathbb{C}$ une fonction continue, ω son module de continuité, i.e $\omega(h) = \sup\{|f(u) - f(v)|; |u-v| \leq h\}$. Pour $n \geq 1$, on considère le polynôme $B_n(f,x) = B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(\frac{k}{n})$, le n-ième polynôme de Bernstein de f. Alors :

1. *B_n converge uniformément vers f sur $[0,1]$.*
2. *Plus précisément, on a $\|f - B_n\|_\infty \leq C\omega(\frac{1}{\sqrt{n}}$, où C est une constante numérique.*
3. *L'estimation de 2) est optimale : il existe une fonction lipschitzienne f pour laquelle $\|f - B_n\|_\infty \geq \frac{\delta}{\sqrt{n}}$, où δ est une constante numérique.*

3.2 En théorie des martingales

Problème de la ruine du joueur :
Un joueur joue à pile ou face avec une pièce non nécessairement équilibrée, on note p la probabilité d'obtenir pile lors d'un jet. Il reçoit un euro de la banque s'il obtient pile et en donne un à la banque s'il obtient face. Sa fortune initiale est de $a \in \mathbb{N}^*$ euros et celle de la banque de $b \in \mathbb{N}^*$ euros. Le joueur joue jusqu'à sa ruine ou celle de la banque. On modélise ce jeu de la manière suivante : (Y_n) est une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, indépendantes de même loi $p\delta_1 + (1-p)\delta_{-1}$. On note S_n la fortune du joueur après n parties, pour un jeu qui ne s'arrêterait pas ; on pose

$$S_0 = a \text{ et } S_n = a + \sum_{j=1}^{n} Y_j$$

On note T le temps d'arrêt du jeu, c'est à dire

$$T = \inf\{n \in \mathbb{N}^*, S_n = 0 \text{ ou } a = b\}$$

Alors, on a :

$$\mathbb{E}(S_T) = (2p-1)\mathbb{E}(T) \text{ et } \mathbb{E}(\frac{1-p}{p} S_T) = 1$$

Par ailleurs,

1. Si $p = q = \frac{1}{2}$, alors

$$\mathbb{P}(S_T = a+b) = \frac{a}{a+b} \text{ et } \mathbb{E}(T) = ab$$

2. Si $p \neq \frac{1}{2}$, alors

$$\mathbb{P}(S_T = a+b) = \frac{1-(\frac{1-p}{p})^a}{1-(\frac{1-p}{p})^{a+b}} \text{ et } \mathbb{E}(T) = \frac{1}{2p-1}\left[\frac{1-(\frac{1-p}{p})^a}{1-(\frac{1-p}{p})^{a+b}} - a\right]$$

Démonstration : Ouvrard, exercice 15.2p394 ∎

3.3 En chaînes de Markov

Dans tout ce chapitre, E est un ensemble dénombrable et $P(E)$ est l'ensemble de ses parties.

Définition 6

On dit qu'une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$, à valeurs dans $(E, P(E))$, est une chaîne de Markov si, pour tout $(n+1)$-uplet $(i_0, ..., i_n)$ de points de E tel que $\mathbb{P}(\cap_{0\leq j\leq n-1}\{X_j = i_j\}) > 0$,

$$\mathbb{P}(X_n = i_n \mid \cap_{0\leq j\leq n-1}\{X_j = i_j\}) = \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1})$$

Autrement dit, la loi de X_n conditionnellement à $(X_0, ..., X_{n-1})$ et la loi de X_n conditionnellement à X_{n-1} sont identiques.

Exemple 4

Marche aléatoire sur $\mathbb{Z}$: soit $(Y_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de loi de Bernoulli sur $\{-1, 1\}$. Soit $X_n = \sum_{i=0}^{n} Y_i$. (X_n) est une chaîne de Markov à valeurs dans $\mathbb{Z}$.

Démonstration : Barbe-Ledoux, chapitre sur les chaînes de Markov ∎