

249. Suites de variables aléatoires de Bernoulli indépendantes.

Introduction : La loi de Bernoulli représente l'expérience très simple du jeu de pile ou face (2 issues possibles). Mais cette loi intervient pour la construction de nombreuses autres lois discrètes, ainsi qu'en modélisation.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilité.

1 Suites de variables de Bernoulli indépendantes

1.1 Indépendance

Définition 1. Deux évènements $A, B \in \mathcal{F}$ sont dits indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Définition 2. Une famille d'évènements $(A_i)_{i \in I} \in \mathcal{F}^I$ est dit mutuellement indépendantes si pour tous $i_1, \dots, i_k \in I$ distincts, $\mathbb{P}\left(\bigcap_{i=j}^k A_{i_j}\right) = \prod_{j=1}^k \mathbb{P}(A_{i_j})$.

Contre-exemple 3. L'indépendance deux à deux n'implique pas l'indépendance mutuelle. Ainsi, si on considère deux lancers de pièce, $A = \{\text{Face au premier lancer}\}$, $B = \{\text{Face au second lancer}\}$, $C = \{\text{Même résultat aux deux lancers}\}$, on a A, B, C deux à deux indépendants mais $\mathbb{P}(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

Définition 4. Une famille de variables aléatoires X_1, \dots, X_n à valeurs dans (E, \mathcal{B}) est dite (mutuellement) indépendante si pour tout $(A_1, \dots, A_n) \in \mathcal{B}^n$, la famille d'évènements $(X_i \in A_i)_{1 \leq i \leq n}$ est mutuellement indépendante.

Proposition 5. X_1, \dots, X_n est indépendante ssi pour toutes $\varphi_1, \dots, \varphi_n$ mesurables bornées, $\mathbb{E}\left[\prod_{i=1}^n \varphi(X_i)\right] = \prod_{i=1}^n \mathbb{E}[\varphi(X_i)]$.

1.2 Loi de Bernoulli

Définition 6. On dit qu'une variable aléatoire $X : \Omega \rightarrow \mathbb{R}$ suit une loi de Bernoulli de paramètre $p \in [0, 1]$ si $\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$.

Proposition 7. On a alors $\mathbb{E}[X] = p, \text{Var}(X) = p(1 - p)$.

Proposition 8. La fonction génératrice est donnée par $G_X(s) = 1 - p + ps$, et la fonction caractéristique par $\varphi_X(t) = 1 - p + pe^{it}$.

Définition 9. On dit que X suit une loi de Rademacher de paramètre $p \in [0, 1]$ si $\mathbb{P}(X = 1) = p, \mathbb{P}(X = -1) = 1 - p$.

Proposition 10 (Khintchine). Soient $\varepsilon_1, \dots, \varepsilon_n$ variables de Rademacher indépendantes, soient $a_1, \dots, a_n \in \mathbb{R}$, soit $f = \sum_{j=1}^n a_j \varepsilon_j$. On a $\sqrt{f} \|f\|_1 \geq \|f\|_2$.

1.3 Une construction explicite

On munit $\Omega = [0, 1[$ de sa tribu borélienne et de la mesure de Lebesgue.

Définition 11. Si $x \in \Omega$, on appelle développement dyadique de x toute suite $(D_k)_{k \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}}$ telle que $x = \sum_{k=0}^{+\infty} \frac{D_k}{2^k}$.

Exemple 12. $(0, 1, 0, 0, \dots)$ et $(0, 0, 1, 1, \dots)$ sont deux développements dyadiques de $\frac{1}{2}$.

Proposition 13. On pose, pour $x \in \Omega$, $R_0(x) = x$ et pour tout $n \in \mathbb{N}$, $D_{n+1}(x) = E[2R_n(x)], R_{n+1}(x) = 2R_n(x) - D_{n+1}(x)$. La suite $(D_n(x))_{n \in \mathbb{N}}$ est un développement dyadique de x .

Théorème 14. Sur $(\Omega, \mathcal{B}, \lambda)$, la suite $(D_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires de Bernoulli indépendantes.

2 Lois associées à des variables de Bernoulli

2.1 Loi binomiale

Il s'agit de la loi du nombre de pile lors de n lancers de pièce.

Définition 15. Soient $p \in [0, 1], n \in \mathbb{N}^*$. On dit que X suit une loi binomiale de paramètres n, p , et on note $X \sim \mathcal{B}(n, p)$, si $\forall k \in \llbracket 0, n \rrbracket, \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Proposition 16. On a alors $\mathbb{E}[X] = np, \text{Var}(X) = np(1 - p)$.

Proposition 17. Si X_1, \dots, X_n sont indépendantes et toutes de loi $\mathcal{B}(p)$, alors $\sum_{i=1}^n X_i$ suit une loi $\mathcal{B}(n, p)$.

2.2 Loi géométrique

Il s'agit de la loi du premier pile lors d'une suite de lancers de pièce.

Définition 18. Soit $p \in [0, 1]$. On dit que X suit une loi géométrique de paramètre p , et on note $X \sim \mathcal{G}(p)$, si pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = p(1 - p)^{k-1}$.

Proposition 19. On a alors $\mathbb{E}[X] = \frac{1}{p}$ et $\text{Var}(X) = \frac{1}{p^2}$.

Proposition 20. Si $(X_n)_{n \in \mathbb{N}^*}$ suite de v.a. indépendantes de loi $\mathcal{B}(p)$, alors la variable $T = \inf\{k \in \mathbb{N}^*, X_k = 1\}$ suit une loi $\mathcal{G}(p)$.

2.3 Loi de Poisson

Il s'agit d'une loi modélisant un nombre d'occurrences d'un événement sur une durée fixée.

Définition 21. Soit $\lambda > 0$. On dit que X suit une loi de Poisson de paramètre λ , et on note $X \sim \mathcal{P}(\lambda)$, si pour tout $k \in \mathbb{N}$, $\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

Proposition 22. On a alors $\mathbb{E}[X] = \lambda$ et $\text{Var}(X) = \lambda$.

Théorème 23 (événements rares). Soit $\lambda > 0$. Pour tout $n \in \mathbb{N}$, soit X_n v.a. de loi $\mathcal{B}(n, \frac{\lambda}{n})$. Alors $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une variable de loi $\mathcal{P}(\lambda)$.

3 Applications

3.1 Marche aléatoire

Soit $d \in \mathbb{N}^*$, soit (e_1, \dots, e_d) la base canonique de \mathbb{R}^d . Soit $(\theta_n)_{n \in \mathbb{N}}$ suite de variables indépendantes de loi uniforme dans $\{\pm e_1, \dots, \pm e_d\}$. On pose $X_0 = 0$ et pour tout $n \in \mathbb{N}$, $X_{n+1} = X_n + \theta_n$.

Théorème 24. On a :

- Si $d \in \{1, 2\}$, on a $\mathbb{P}(X_n = 0 \text{ infiniment souvent}) = 1$.
- Si $d \geq 3$, on a $\mathbb{P}(\|X_n\| \rightarrow +\infty) = 1$.

3.2 Théorème de Bernstein

Définition 25. Soit $f \in \mathcal{C}([0, 1])$.

On définit son *module de continuité* par $w_f(h) = \sup_{|x-y|<h} |f(x) - f(y)|$.

Théorème 26. Pour $p \in [0, 1]$, soient $X_1(p), \dots, X_n(p)$ iid suivant une loi de Bernoulli de paramètre p .

On pose, pour $p \in [0, 1]$, $S_n(p) = \sum_{i=1}^n X_i(p)$, et $p \in [0, 1]$, $B_n(p) = \mathbb{E} \left[f \left(\frac{S_n(p)}{n} \right) \right]$. $(B_n)_{n \in \mathbb{N}}$ est une suite de fonctions polynomiales qui converge uniformément vers f , on a $\|B_n - f\|_\infty \leq \frac{3}{2} w_f \left(\frac{1}{\sqrt{n}} \right)$ et cette inégalité est optimale.

3.3 Sondages

Proposition 27 (Loi des grands nombres). Soit $(X_n)_{n \in \mathbb{N}}$ suite de variables aléatoires iid dans $L^2(\mathbb{P})$. Alors $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathbb{E}[X_1]$ en probabilité.

Théorème 28 (central limite). Soit $(X_n)_{n \in \mathbb{N}^*}$ suite de variables aléatoires iid dans L^2 . Alors $\sqrt{\frac{n}{\text{Var}(X_1)}} \left(\frac{S_n}{n} - \mathbb{E}[X_1] \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$

Avant une élection opposant deux candidats A, B , on effectue un sondage sur un échantillon de n personnes. On note X_n la variable aléatoire valant 1 si le n -ème sondé annonce une intention de voter A et 0 sinon. On suppose que les $(X_n)_n$ sont indépendantes et suivent une loi de Bernoulli de paramètre $\theta \in [0, 1]$ inconnu.

Définition 29. On pose, pour $n \in \mathbb{N}^*$, $\widehat{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$, on appelle moyenne empirique cette variable aléatoire.

Proposition 30. $(\widehat{X}_n)_n$ est un estimateur de θ :

- Fortement consistant : $\widehat{X}_n \xrightarrow[n \rightarrow +\infty]{} \theta$ en probabilité et presque sûrement.
- Sans biais : $\forall n \in \mathbb{N}^*, \mathbb{E}[\widehat{X}_n] = \theta$.

Proposition 31. On a pour $n \in \mathbb{N}^*$, $\lim_{n \rightarrow \infty} \mathbb{P} \left(\theta \in \left[\overline{X}_n - \frac{q_\alpha}{\sqrt{2n}}, \overline{X}_n + \frac{q_\alpha}{\sqrt{2n}} \right] \right) \geq 1 - \alpha$.

Développements

- Théorème de Bernstein.
- Marche aléatoire sur \mathbb{Z} et \mathbb{Z}^2 .

Références

- [1] P. Barbé, M. Ledoux, *PROBABILITÉ*, EDP Sciences.
- [2] S. Méléard, *ALÉATOIRE*, École Polytechnique.
- [3] J.-Y. Ouvrard, *PROBABILITÉS - TOME 2*, Cassini.