Probabilités

Méthode et contexte

- Une mesure de probabilité étant en particulier finie, on a dans ce cadre que les espaces L^p sont emboités, i.e : $L^\infty \subseteq \cdots \subseteq L^1$. Cela se traduit par le fait que si une variable aléatoire possède un moment d'ordre k, tous ses moments d'ordre inférieur sont également finis.

Définitions et propriétés élémentaires

Définition 1. Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé, et $(E, \mathcal{E})$ un espace mesurable.

1. Si $X : \Omega \to E$ est mesurable, alors X est appelée *variable aléatoire* (v.a.) à valeurs dans E.
2. Si X est une v.a. à valeurs dans E, on appelle loi de X la mesure image de P par X, notée P_X et vérifiant :
$$P_X(A) = P\left(X^{-1}(A)\right) = P\left(\{\omega \in \Omega \mid X(\omega) \in A\}\right) = P(X \in A).$$

Définition 2. Pour toute v.a.r X, on appelle *fonction de répartition* de X la donnée de $F_X : \mathbb{R} \to [0,1]$ définie par $F_X(t) = P(X \leq t) = P_X(]-\infty, t])$.

Remarque. F_X est continue à droite, limitée à gauche (càdlàg), croissante, tend vers 0 en $-\infty$, 1 en $+\infty$, et caractérise P_X.

Définition 3. Soit X une v.a. à valeurs dans $\mathbb{R}^d$. On appelle *fonction caractéristique* de X, notée Φ_X, la fonction de $\mathbb{R}^d$ dans $\mathbb{C}$ définie par

$$\Phi_X(\xi) := \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} \mathrm{d}P_X(x) = E\left(e^{i\langle X,\xi\rangle}\right).$$

Remarque. Φ_X est en fait la transformée de Fourier de la loi P_X. C'est une fonction uniformément continue, dont le module est borné par 1. Φ_X a autant de dérivées que X a de moments finis.

Remarque. Si $X \sim \mathcal{N}(\mu, \sigma^2)$, alors $\Phi_X(\xi) = exp(i\xi\mu - \frac{\xi^2\sigma^2}{2})$.

Définition 4. Si X est une v.a. à valeurs dans $\mathbb{N}$, on appelle *fonction génératrice* de X, la fonction $G_X : [0,1] \to \mathbb{R}^+$ définie par :

$$G_X(t) := E(t^X) = \sum_{n=0}^{+\infty} t^n P(X = n).$$

Remarque. G_X caractérise la loi de X, et détermine tous les moments de X comme l'explique la proposition suivante.

Proposition 1. — *Soit X une v.a. à valeurs dans $\mathbb{N}$, alors pour tout $k \geq 1$:*

$$E\left(\prod_{i=0}^{k-1}(X - i)\right) = \lim_{t \to 1^-} G_X^{(k)}(t).$$

Définition 5 (Indépendance).

- Des événements $(A_i)_{i\in I}$ sont dits indépendants si pour toute partie finie J de I, on a :

$$P\Big(\bigcap_{j\in J} A_j\Big) = \prod_{j\in J} A_j$$

- Des tribus $(\mathcal{A}_i)_{i\in I}$ sont dites indépendantes si pour toute famille $(A_i)_{i\in I}$ telle que $A_i \in \mathcal{A}_i$, les événements sont indépendants.
- Des variables aléatoires $(X_i)_{i\in I}$ à valeurs dans des espaces mesurables $(E_i, \mathcal{E}_i)$ sont dites indépendantes si la famille de tribus $(\sigma(X_i))_{i\in I}$ l'est.

Remarque. L'indépendance des $(X_i)_{i\in I}$ porte sur les tribus engendrées (sur Ω) et non sur les valeurs proprement dites de ces variables aléatoires. Par suite, si des $\Phi_i : (E_i, \mathcal{E}_i) \to (E_i', \mathcal{E}_i')$ sont mesurables, l'indépendance des $(X_i)_{i\in I}$ entraine celle des $(\Phi_i(X_i))_{i\in I}$.

Remarque. La vérification de l'indépendance des $(X_i)_{i\in I}$ se rammène à montrer que pour tout $J \subset I$ fini, pour toute famille $(A_j)_{j\in J}$ telle que $A_j \in \mathcal{E}_j$, on a $P\Big(\bigcap_{j\in J}(X_j \in A_j)\Big) = \prod_{j\in J}(X_j \in A_j)$.

Proposition 2 (Caractérisations de l'indépendance). — *Soit $(X_i)_{1\le i\le n}$ une famille de variables aléatoires avec $X_i : (\Omega, \mathcal{F}) \to (E_i, \mathcal{E}_i)$. Soit $X := (X_1, \dots, X_n) : (\Omega, \mathcal{F}) \to (\prod_{i=1}^n E_i, \otimes_{i=1}^n \mathcal{E}_i)$.*

1. *Les $(X_i)_{1\le i\le n}$ sont indépendants si et seulement si $P_X = \otimes_{i=1}^n P_{X_i}$, ce qui équivaut à $\Phi_X = \otimes_{i=1}^n \Phi_{X_i}$.*
2. (a) *Si les $(X_i)_{1\le i\le n}$ sont indépendants, et ont $(f_{X_i})_{1\le i\le n}$ comme densités respectives par rapport à la mesure de Lebesgue, alors $P_X \ll \lambda_n$ et a pour densité $f_X(x_1, \dots, x_n) = f_{X_1}(x_1)\dots f_{X_n}(x_n)$.*
   (b) *Réciproquement, si $P_X \ll \lambda_n$, de densité s'écrivant $f_X(x_1, \dots, x_n) = f_{X_1}(x_1)\dots f_{X_n}(x_n)$, alors les $(X_i)_{1\le i\le n}$ sont indépendants, de densités respectives $(f_{X_i})_{1\le i\le n}$.*

Résultats principaux

Théorème 3 (Inégalité de Markov). — *Soit X une variable aléatoire réelle presque surement positive, alors pour $\alpha > 0$:*

$$P(X \ge \alpha) \le \frac{E(X)}{\alpha}$$

Théorème 4 (Inégalité de Jensen). — *Soient $X \in L^1$, et Φ une fonction convexe sur un intervalle I tel que $P(X \in I) = 1$ et $E(|\Phi(X)|) < \infty$. Alors $\Phi(E(X)) \le E(\Phi(X))$. Si Φ est de plus strictement convexe, alors il y a égalité si et seulement si X est p.s constante.*

Théorème 5 (Injectivité de la transformée de Fourier). — *Soient X_1 et X_2 des variables aléatoires à valeurs dans $\mathbb{R}^d$. Si $\Phi_{X_1} = \Phi_{X_2}$, alors $P_{X_1} = P_{X_2}$.*

Théorème 6 (Coalitions). — *Soient $(X_i)_{i\in I}$ une famille de variables aléatoires indépendantes, et $(I_k)_{k\in K}$ une partition de I. Alors les tribus $\big(\sigma(X_i,\ i \in I_k)\big)_{k\in K}$ sont indépendantes.*

Théorème 7 (Loi faible des grands nombres). — *Soient $(X_i)_{i\in\mathbb{N}^*}$ une famille de variables aléatoires indépendantes de L^2, telles que $\lim\limits_{n\to+\infty} \frac{1}{n}\sum_{i=1}^n E(X_i) = \mu$ et $\sup_i V(X_i) = \sigma^2$. Alors :*

1. *La moyenne empirique $\overline{X}_n := \dfrac{1}{n}\displaystyle\sum_{i=1}^n X_i$ converge dans L^2 vers la moyenne théorique μ.*
2. *Pour $\varepsilon > 0$, on a l'estimation suivante :*

$$P(|\overline{X}_n - E(\overline{X}_n)| \ge \varepsilon) \le \frac{\sigma^2}{n\varepsilon^2}.$$

Outils importants

Lemme 8 (Fekete). — *Si $(u_n)_{n\in\mathbb{N}}$ est une suite sous additive, i.e. $\forall n, m \in \mathbb{N},\ u_{n+m} \le u_n + u_m$, alors $(\dfrac{u_n}{n})_{n\in\mathbb{N}}$ converge, et on a l'égalité : $\lim\limits_{n\to\infty} \dfrac{u_n}{n} = \inf\limits_{n\ge 1} \dfrac{u_n}{n} \in \mathbb{R} \cup \{-\infty\}$.*

Proposition 9 (Changement de variable). — *Soit X une v.a. à valeurs dans $(E, \mathcal{E})$, et $f : E \to \overline{\mathbb{R}}$ une fonction mesurable telle que $f \ge 0$ p.p. ou $E\big(|f(X)|\big) < \infty$, alors :*

$$E(f(X)) = \int_E f(x)\mathrm{d}P_X(x).$$

Corollaire 10 (Inégalité de Bienaymé-Tchebychef). — *Si* $X \in L^2$ *est une v.a.r, alors pour tout* $\varepsilon > 0$ *:*

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

Proposition 11 (Inégalité de Hoeffding). — *Soient* $(X_i)_{i\in\mathbb{N}^*}$ *une famille de v.a. indépendantes à valeurs dans* $[a, b]$. *Alors pour tout* $\epsilon > 0$ *:*

$$P(|\overline{X}_n - E(\overline{X}_n)| \geq \epsilon) \leq 2\exp(-2\frac{n\varepsilon^2}{(b-a)^2})$$

Autres résultats

Lemme 12. — *Soit* I *un intervalle de* $\mathbb{R}$. *Si* $\Phi : I \to \mathbb{R}$ *est une fonction convexe, alors pour tout* $x \in \mathring{I}$ *:*

$$\Phi(x) = \sup_{a,b \ | \ l_{a,b}\leq\Phi} l_{a,b}(x)$$