Probabilités

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.
- Des variables indépendantes sont de covariance nulle, mais la réciproque est très fausse ! Par exemple en considérant une loi gaussienne et son produit par une v.a. de Rademacher, leur covariance est nulle mais elles ne sont pas indépendantes, sinon leurs valeurs absolues le seraient, et donc la gaussienne serait indépendante d'elle même, i.e. constante.
- À l'inverse, des variables a priori corrélées peuvent êtres indépendantes : si U est une loi exponentielle et V une loi uniforme sur $[0,1]$, alors $\sqrt{U}\cos(2\pi V)$ et $\sqrt{U}\sin(2\pi V)$ sont indépendantes et suivent chacune la loi $\mathcal{N}(0,1/2)$.

Méthode

- Utiliser les outils adaptés : pour étudier une somme de v.a. indépendantes on utilise la transformée de Fourier, pour étudier leur min on utilise la fonction de répartition, etc ...
- Pour calculer la loi d'un couple (X,Y) de v.a., on prend f mesurable positive et on essaye d'écrire $E\left(f(X,Y)\right) = \displaystyle\int f(x,y)\mathrm{d}\mu(x,y)$, et alors le couple est de loi μ.

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 $F_X(t) = P(X \le t) = P_X(]-\infty, t])$.

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 $(X_i)_{1\leq i\leq n}$ une famille de variables aléatoires réelles, et $X := (X_1, \ldots, X_n)$.*

1. *Si les $(X_i)_{1\leq i\leq n}$ sont indépendants, et ont $(f_{X_i})_{1\leq i\leq n}$ comme densités respectives par rapport à la mesure de Lebesgue, alors $P_X \ll \lambda_n$ et a pour densité $f_X(x_1, \ldots, x_n) = f_{X_1}(x_1)\ldots f_{X_n}(x_n)$.*
2. *Réciproquement, si $P_X \ll \lambda_n$, de densité s'écrivant $f_X(x_1, \ldots, x_n) = f_{X_1}(x_1)\ldots f_{X_n}(x_n)$ où les $(f_i)_{1\leq i\leq n}$ sont des densités de probabilité (i.e. positives, d'intégrale valant 1), alors les $(X_i)_{1\leq i\leq n}$ sont indépendants, de densités respectives $(f_{X_i})_{1\leq i\leq n}$.*

Corollaire 3. — *Soit $(X_i)_{1\leq i\leq n}$ des variables aléatoires réelles, les propriétés suivantes sont équivalentes :*

1. *$(X_i)_{1\leq i\leq n}$ est une famille de variables aléatoires indépendantes ;*
2. *$P_X = \otimes_{i=1}^n P_{X_i}$*
3. *Pour toute famille $(f_i)_{1\leq i\leq n}$ de fonctions boréliennes positives, $E\left(\prod_{i=1}^n f_i(X_i)\right) = \prod_{i=1}^n E\left(f_i(X_i)\right)$*
4. *$\Phi_X = \otimes_{i=1}^n \Phi_{X_i}$*

Résultats principaux

Théorème 4 (Inégalité de Markov). — *Soit X une variable aléatoire réelle presque surement positive, alors pour $\alpha > 0$:*

$$P(X \geq \alpha) \leq \frac{E(X)}{\alpha}$$

Théorème 5 (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 $\Phi(E(X)) \leq E(\Phi(X))$. Si Φ est de plus strictement convexe, alors il y a égalité si et seulement si X est p.s constante.*

Théorème 6 (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 7 (Coalitions). — *Soit* $(\mathcal{A}_i)_{i\in I}$ *une famille de tribus engendrées par les* π*-systèmes* $(C_i)_{i\in I}$*. Alors ces tribus sont indépendantes si et seulement si les* π*-systèmes générateurs le sont. En particulier, si* $(X_i)_{i\in I}$ *est une famille de variables aléatoires indépendantes, et* $(I_k)_{k\in K}$ *une partition de* I*, alors les tribus* $\big(\sigma\left(X_i,\ i\in I_k\right)\big)_{k\in K}$ *sont indépendantes.*

Théorème 8 (Loi faible des grands nombres). — *Soit* $(X_i)_{i\in\mathbb{N}^*}$ *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 :*

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

Lemme 9 (Borel-Cantelli). — *Soit* $(A_n)_{n\geq 0}$ *une suite d'évènements. On note* $\limsup A_n := \bigcap\limits_{n\geq 0}\bigcup\limits_{k\geq n} A_k$*, qui correspond à l'évènement "être dans une infinité de* A_n*". On a alors la dichotomie suivante :*

1. *Si* $\displaystyle\sum_n P(A_n)<\infty$*, alors* $P(\limsup A_n)=0$*, i.e.* $\sum_n \mathbb{1}_{A_n}<+\infty$ *presque sûrement.*
2. *Si* $\displaystyle\sum_n P(A_n)=\infty$ *et que les* $(A_n)_{n\geq 0}$ *sont indépendants, alors* $P(\limsup A_n)=1$*.*

Outils importants

Lemme 10 (Fekete). — *Si* $(u_n)_{n\in\mathbb{N}}$ *est une suite sous additive, i.e.* $\forall n,m\in\mathbb{N},\ u_{n+m}\leq u_n+u_m$*, alors* $(\dfrac{u_n}{n})_{n\in\mathbb{N}}$ *converge, et on a l'égalité :* $\lim\limits_{n\to\infty}\dfrac{u_n}{n}=\inf\limits_{n\geq 1}\dfrac{u_n}{n}\in\mathbb{R}\cup\{-\infty\}$*.*

Proposition 11 (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* $f\geq 0$ *p.p. ou* $E\big(|f(X)|\big)<\infty$*, alors :*

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

Corollaire 12 (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 13 (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 14. — *Soit* I *un intervalle de* $\mathbb{R}$*. Si* $\Phi:I\to\mathbb{R}$ *est une fonction convexe, alors pour tout* $x\in\mathring{I}$ *:*

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

Lemme 15. — *Pour une v.a. à valeurs dans* $\mathbb{R}^+$*,* $E(X)=\displaystyle\int_0^\infty P(X\geq x)\mathrm{d}x$*, car* $X=\displaystyle\int_0^\infty 1_{x\leq X}\mathrm{d}x$