

Théorie de la mesure

Contexte et méthode

- Toujours se demander : dans quel ensemble je travaille, quelle est la tribu, quelle est la mesure ?
- Utiliser la régularité des mesures, les raisonnements par densité :
Indicatrices \rightarrow fonctions étagées \rightarrow positives (limite croissante) \rightarrow mesurables (parties positive/négative).
- User et abuser des \limsup et \liminf .

Motivations

Le problème de l'intégrale de Riemann est que la propriété de Riemann-intégrabilité n'est pas stable par limite simple, ce qui pose donc problème pour établir des théorèmes de limite intégrale généraux. De plus, même si l'on se restreint à une situation ne comportant que des fonctions continues (par morceaux), la preuve est très pénible. Une illustration concrète de ce phénomène est le résultat suivant :

THÉORÈME . — *Soit $f : [a, b] \mapsto \mathbb{R}$ une fonction bornée. Alors f est Riemann-intégrable si et seulement si l'ensemble de ses points de discontinuité est négligeable.*

où l'on peut définir la notion de négligeabilité sans avoir recours à la théorie de la mesure :

DÉFINITION 1. Une partie A de \mathbb{R} est négligeable si pour tout $\varepsilon > 0$, il existe une famille d'intervalles $(I_i)_{i \in \mathbb{N}}$ telle que $A \subset \bigcup_{i \in \mathbb{N}} I_i$ et $\sum_{i \in \mathbb{N}} l(I_i) \leq \varepsilon$, où $l(I_i)$ est la longueur de l'intervalle I_i .

Le prototype de fonction non intégrable au sens de Riemann est la fonction indicatrice des rationnels, qui bien que limite simple de fonctions en escalier, n'est pas Riemann-intégrable d'après le théorème précédent. Historiquement, ce furent plutôt les travaux de Fourier sur l'équation de la chaleur qui motivèrent le développement de la théorie de la mesure afin de rendre rigoureuses les manipulations du noyau de la chaleur.

Ainsi, il a fallu changer de méthode d'intégration, en intégrant par tranches rectangulaires horizontales, dont la base peut alors être un ensemble plus complexe qu'un simple intervalle. La limite simple de fonctions mesurables sera alors mesurable, mais on sera attentif aux subtilités induites par la convergence presque partout, qui sont discrètement évacuées grâce à la notion de complétion d'une tribu, qui permet justement de rendre mesurables les ensembles négligeables.

Définitions

Résultats principaux

Dans tout ce qui suit, on travaille dans un espace mesuré (E, \mathcal{A}, μ) , et on pourra utiliser un espace métrique (T, d) .

THÉORÈME 1 (Dynkin). — *Le σ -système engendré par un π -système est égal à la tribu engendrée par ce dernier.*

THÉORÈME 2 (Convergence monotone). — *Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables de E dans $[0, +\infty]$, telle que pour tout $n \geq 0$, $f_n \leq f_{n+1}$. Alors, en notant f la limite simple de cette suite, on a que f est mesurable, et :*

$$\int f = \lim_{n \rightarrow +\infty} \int f_n$$

THÉORÈME 3 (Convergence dominée). — *Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables qui converge simplement vers f . Si il existe g intégrable telle que pour tout $n \geq 0$, $|f_n| \leq g$, alors f est intégrable et :*

$$\int f = \lim_{n \rightarrow +\infty} \int f_n$$

REMARQUE. On peut seulement supposer les conditions ci-dessus vraies presque partout, mais dans ce cas il faut imposer la mesurabilité de la limite simple, ou bien travailler dans la tribu complétée, comme l'explique la remarque suivante.

REMARQUE. Si $\tilde{f} : E \mapsto F$ est presque partout égale à une fonction mesurable $f : (E, \mathcal{A}, \mu) \mapsto (F, \mathcal{B})$, alors \tilde{f} est mesurable pour la tribu complétée $\overline{\mathcal{A}}$ de μ sur E .

THÉORÈME 4 (Régularité des mesures boréliennes finies). — Si μ est une mesure borélienne finie sur un espace métrique (T, d) , alors μ est régulière, i.e pour tout $A \in \mathcal{B}(T)$:

$$\begin{aligned}\mu(A) &= \inf \{ \mu(U) \mid U \text{ ouvert, } A \subset U \} \\ &= \sup \{ \mu(F) \mid F \text{ fermé, } F \subset A \}\end{aligned}$$

REMARQUE. Dans un espace polonais (métrique séparable complet), on a de plus que la régularité intérieure est sur les compacts.

THÉORÈME 5 (Régularité des mesures de Radon). — Si μ est une mesure de Radon sur un espace métrique localement compact séparable (ex. \mathbb{R}^d), alors μ est régulière, et la régularité intérieure est sur les compacts.

THÉORÈME 6 (Représentation de Riesz-Markov). — Si E est un espace métrique localement compact séparable, et $\Lambda : C_c(E) \mapsto \mathbb{R}$ une forme linéaire positive, alors il existe une unique mesure de Radon μ sur E telle que, pour tout $f \in C_c(E)$:

$$\Lambda(f) = \int f d\mu$$

REMARQUE. Avec ce qui précède, on obtient de plus que la mesure μ est régulière.

THÉORÈME 7 (Inégalités d'Hölder et de Young). — Soient $1 \leq p, q, r \leq +\infty$, $f \in L^p$, $g \in L^q$, on a les inégalités :

- Si p et q sont des exposants conjugués, i.e : $\frac{1}{p} + \frac{1}{q} = 1$, alors $\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q}$.
- Si $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, alors $\|f * g\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^q}$.

THÉORÈME 8 (Riesz-Fischer). — Si $p \in [1; +\infty]$, alors $L^p(E, \mathcal{A}, \mu)$ est un espace de Banach.

REMARQUE. La preuve du théorème ci-dessus donne en particulier que la convergence en norme L^p implique l'existence d'une extraite qui converge presque partout. Le théorème de convergence dominée fournit une sorte de réciproque, i.e que la convergence presque partout entraîne la convergence L^p , sous réserve d'avoir une domination intégrable.

THÉORÈME 9 (Densité). — Soit $p \in [1; +\infty[$.

1. Les fonctions étagées sont denses dans L^p . De plus, si \mathcal{A} est généré par une famille dénombrable dense et si μ est σ -finie, alors $L^p(E, \mathcal{A}, \mu)$ est séparable.
2. (a) Si (E, d) est un espace métrique et μ une mesure borélienne finie, alors les fonctions lipschitziennes sont denses dans L^p .
- (b) Si de plus (E, d) est localement compact séparable et μ est une mesure de Radon, alors les fonctions lipschitziennes à support compact sont denses dans L^p .

REMARQUE. Si (E, d) est une métrique séparable, alors $\mathcal{B}(E)$ est dénombrablement généré.

REMARQUE. Pour le point 2.(a), il suffit en fait que la mesure soit extérieurement régulière sur les ouverts, et pour le point 2.(b), il suffit qu'elle soit de plus intérieurement régulière sur les compacts.

THÉORÈME 10 (Dérivée de Radon-Nikodym). — Soient μ et ν deux mesures σ -finies sur (E, \mathcal{A}) . Si ν est absolument continue par rapport à μ , alors il existe $f : E \mapsto \overline{\mathbb{R}}_+$ mesurable tel que $\nu = f\mu$.

REMARQUE. Si ν est une mesure signée (en particulier finie), et μ une mesure σ -finie, on a l'équivalence entre les propositions suivantes :

(i) $\nu \ll \mu$

(ii) Pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que $\forall A \in \mathcal{A}, \mu(A) \leq \delta \Rightarrow |\nu|(A) \leq \varepsilon$

L'hypothèse de finitude de ν est cruciale car elle donne l'intégrabilité de la dérivée de Radon-Nikodym par rapport à μ .

THÉORÈME 11 (Fubini-Tonelli). — Soient μ_1 et μ_2 deux mesures σ -finies sur (E_1, \mathcal{E}_1) et (E_2, \mathcal{E}_2) . Si $f : E_1 \times E_2 \mapsto [0, +\infty]$ est mesurable, alors :

1. $x \mapsto \int_{E_2} f(x, y) d\mu_2(y)$ est mesurable pour \mathcal{E}_1 , et $y \mapsto \int_{E_1} f(x, y) d\mu_1(x)$ est mesurable pour \mathcal{E}_2 .

2. On a :

$$\int f(x, y) d(\mu_1 \otimes \mu_2)(x, y) = \int_{E_1} \left(\int_{E_2} f(x, y) d\mu_2(y) \right) d\mu_1(x) = \int_{E_2} \left(\int_{E_1} f(x, y) d\mu_1(x) \right) d\mu_2(y)$$

REMARQUE. On dispose aussi de la version Fubini-Lebesgue dans le cas intégrable, qui aboutit aux mêmes conclusions, avec en plus que, pour $|f|$, les deux fonctions de 1. sont finies presque-partout.

REMARQUE. Les égalités des théorèmes de Fubini reposent entièrement sur le théorème de la classe monotone, d'où l'importance de l'hypothèse de σ -finitude.

COROLLAIRE 12 (Intégration par parties). — Soient f et g sont deux fonctions mesurables, localement intégrables de \mathbb{R} dans \mathbb{R} . On note, pour $x \in \mathbb{R}$:

$$F(x) := \int_0^x f(t) dt, \quad \text{et} \quad G(x) := \int_0^x g(t) dt.$$

Alors, F et G sont absolument continues, dérivables presque partout, de dérivées presque partout égales respectivement à f et g , et de plus, pour tout $a < b$:

$$\int_a^b f(t) G(t) dt = \left(F(b)G(b) - F(a)G(a) \right) - \int_a^b F(t) g(t) dt$$

THÉORÈME 13 (Changement de variable). — Soit φ un C^1 -difféomorphisme entre deux ouverts U et V de \mathbb{R}^n . On note $\varphi'(x)$ sa différentielle et $J_\varphi(x) := \det(\varphi'(x))$ le jacobien de φ au point $x \in U$. Alors :

1. Pour toute fonction $f \geq 0$ borélienne et tout borélien B de V , on a :

$$\int_B f(y) dy = \int_{\varphi^{-1}(B)} f(\varphi(x)) |J_\varphi(x)| dx$$

2. Pour toute fonction borélienne f sur V , la condition que f soit intégrable sur V est identique à la condition que la fonction $|J_\varphi(x)| f(\varphi(x))$ soit intégrable sur U , et dans ce cas, la formule du changement de variable ci-dessus est encore valide.

Outils importants

COROLLAIRE 14 (Unicité des mesures). — Soient μ et ν deux mesures sur (E, \mathcal{A}) qui coïncident sur un π -système \mathcal{C} tel que $\mathcal{A} = \sigma(\mathcal{C})$. Alors :

1. Si $\mu(E) = \nu(E) < +\infty$, alors $\mu = \nu$.

2. Si il existe une suite croissante $(A_n)_{n \in \mathbb{N}}$ d'éléments de \mathcal{C} tels que $\bigcup_n A_n = E$, et que pour tout $n \in \mathbb{N}$,

$$\mu(A_n) = \nu(A_n) < +\infty, \text{ alors } \mu = \nu.$$

LEMME 15 (Fatou). — Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables positives. Alors :

$$\int \liminf_n f_n \leq \liminf_n \int f_n$$

PROPOSITION 16 (Continuité des intégrales). — Soit $t_0 \in T$ et $f : T \times E \mapsto \overline{\mathbb{R}}$. Si :

- Pour tout $t \in T$, $x \mapsto f(t, x)$ est mesurable.
- Pour presque tout $x \in E$, $t \mapsto f(t, x)$ est continue en t_0 .
- Il existe g intégrable telle que pour tout $t \in T$, pour presque tout $x \in E$, on a $|f(t, x)| \leq g(x)$

Alors, $t \mapsto \int f(t, x) d\mu(x)$ est continue en t_0 .

PROPOSITION 17 (Dérivation des intégrales). — Soit I un intervalle ouvert de \mathbb{R} , $t_0 \in I$, et $f : I \times E \mapsto \overline{\mathbb{R}}$. Si :

- Pour tout $t \in I$, $x \mapsto f(t, x)$ est mesurable et intégrable.
- Pour presque tout $x \in E$, $t \mapsto f(t, x)$ est dérivable en t_0 .
- Il existe g intégrable telle que pour tout $t \in I$, pour presque tout $x \in E$, on a

$$|f(t, x) - f(t_0, x)| \leq g(x) |t - t_0|$$

Alors, $t \mapsto \int f(t, x) d\mu(x)$ est dérivable en t_0 , de dérivée $\int \frac{\partial f}{\partial t}(t_0, x) d\mu(x)$.

REMARQUE. On peut remplacer les deux dernières conditions par les suivantes, plus fortes, qui assurent alors la dérivabilité en tout point de l'intervalle :

- Pour presque tout $x \in E$, $t \mapsto f(t, x)$ est dérivable sur tout I .
- Il existe g intégrable tel que pour presque tout $x \in E$, pour tout $t \in I$, on a $\left| \frac{\partial f}{\partial t}(t, x) \right| \leq g(x)$.

PROPOSITION 18 (Mesure de Stieltjes). — Il y a une correspondance entre les fonctions croissantes continues à droite, et les mesures de Radon. Plus précisément :

- Si μ est une mesure de Radon sur \mathbb{R} , alors la fonction F_μ suivante est croissante, continue à droite.

$$F_\mu(x) := \begin{cases} \mu([0; x]) & \text{si } x \geq 0 \\ -\mu([x; 0]) & \text{si } x < 0 \end{cases}$$

- Si $F : \mathbb{R} \mapsto \mathbb{R}$ est croissante, continue à droite, alors il existe une unique mesure de Radon μ telle que, pour tout $a < b \in \mathbb{R}$, $\mu([a; b]) = F(b) - F(a)$.

COROLLAIRE 19 (Décomposition de Lebesgue). — Si μ et ν sont deux mesures σ -finies sur (E, \mathcal{A}) , alors il existe une unique décomposition $\nu = \nu_a + \nu_s$ en deux mesures telles que ν_a est absolument continue par rapport à μ , et ν_s est étrangère à μ .

PROPOSITION 20 (Points de Lebesgue). — Si $f \in L^1(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$, alors presque tous les points de \mathbb{R}^d sont des points de Lebesgue pour f .

COROLLAIRE 21. — Si μ est une mesure de Radon absolument continue par rapport à la mesure de Lebesgue, alors sa dérivée de Radon-Nikodym f est localement intégrable, et pour presque tout $x \in \mathbb{R}^d$:

$$\lim_{r \rightarrow 0} \frac{\mu(B(x, r))}{\lambda_d(B(x, r))} = f(x)$$

REMARQUE. On peut en fait montrer que pour une mesure étrangère à λ_d , le taux d'accroissement ci-dessus tend vers 0. On en déduit que pour une mesure borélienne σ -finie quelconque, la limite de ce taux est presque-partout égal à la dérivée de Radon-Nikodym de la partie absolument continue de la décomposition de Lebesgue de cette mesure.

COROLLAIRE 22. — L'intégrale d'une fonction $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ contre la mesure de Lebesgue est presque partout dérivable, de dérivée f .

REMARQUE. La réciproque est fautive : une fonction borélienne dérivable presque partout n'est pas forcément égale à l'intégrale de sa dérivée. Par exemple l'escalier de Cantor est croissant continu sur $[0, 1]$, de dérivée presque partout nulle, mais $F(1) - F(0) = 1$.

En fait, si f est croissante continue, par théorème de Stieltjes elle définit une mesure de Radon, et si la partie singulière dans sa décomposition de Lebesgue est nulle, i.e la mesure est absolument continue par rapport à la mesure de Lebesgue, alors on dit que f est absolument continue. Sinon, il manque la contribution de la partie singulière.

Autres résultats

REMARQUE. Il ya des raisons profondes au fait que l'on ne peut pas prendre l'ensemble des parties comme tribu pour faire de l'analyse. On peut par exemple montrer qu'il n'existe pas de mesure non triviale (non atomique) sur $(0, 1)$ muni de la tribu des parties.

THÉORÈME 23 (Egorov). — *Si μ est une mesure finie, et $(f_n)_{n \in \mathbb{N}}$ est une suite de fonctions mesurables réelles convergeant μ -presque-partout vers f mesurable, alors pour tout $\varepsilon > 0$, il existe $A \in \mathcal{A}$ tel que $(f_n)_{n \in \mathbb{N}}$ converge uniformément vers f sur A , et $\mu(E \setminus A) \leq \varepsilon$.*

THÉORÈME 24 (Lusin). — *Soit $f : [a, b] \mapsto \mathbb{R}$ borélienne. Pour tout $\varepsilon > 0$, il existe un compact $K \subset [a, b]$ avec $\lambda([a, b] \setminus K) \leq \varepsilon$, tel que la restriction de f à K est continue.*

REMARQUE. Le résultat ci dessus est en fait une équivalence, i.e une fonction est mesurable si et seulement si elle est continue sur un ensemble de mesure arbitrairement grande.

THÉORÈME 25 (Décomposition de Jordan). —

REMARQUE. Version signée/complexe de représentation de Riesz.

THÉORÈME 26 (Dualité). — *Si p et q sont deux exposants conjugués finis, alors L^q est le dual topologique de L^p .*

Probabilités

Contexte

- Une mesure de probabilité étant en particulier finie, on a dans ce cadre que les espaces L^p sont emboîtés, i.e : $L^\infty \subseteq \dots \subseteq L^1$. 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 être 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 $E(f(X, Y)) = \int f(x, y) d\mu(x, y)$, et alors le couple est de loi μ .

Définitions et propriétés élémentaires

DÉFINITION 2. Soit (Ω, \mathcal{F}, P) un espace probabilisé, et (E, \mathcal{E}) un espace mesurable.

1. Si $X : \Omega \rightarrow E$ 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 :

$$P_X(A) = P(X^{-1}(A)) = P(\{\omega \in \Omega \mid X(\omega) \in A\}) = P(X \in A).$$

DÉFINITION 3. Pour toute v.a.r X , on appelle *fonction de répartition* de X la donnée de $F_X : \mathbb{R} \rightarrow [0, 1]$ 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 4. 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

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

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 5. Si X est une v.a. à valeurs dans \mathbb{N} , on appelle *fonction génératrice* de X , la fonction $G_X : [0, 1] \rightarrow \mathbb{R}^+$ 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 27. — Soit X une v.a. à valeurs dans \mathbb{N} , alors pour tout $k \geq 1$:

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

DÉFINITION 6 (Indépendance).

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

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

- 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 $\Phi_i : (E_i, \mathcal{E}_i) \rightarrow (E'_i, \mathcal{E}'_i)$ sont mesurables, l'indépendance des $(X_i)_{i \in I}$ entraîne celle des $(\Phi_i(X_i))_{i \in I}$.

REMARQUE. La vérification de l'indépendance des $(X_i)_{i \in I}$ se ramène 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 $P \left(\bigcap_{j \in J} (X_j \in A_j) \right) = \prod_{j \in J} P(X_j \in A_j)$.

PROPOSITION 28 (Caractérisations de l'indépendance). — Soit $(X_i)_{1 \leq i \leq n}$ une famille de variables aléatoires réelles, et $X := (X_1, \dots, X_n)$.

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 $f_X(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$.
2. Réciproquement, si $P_X \ll \lambda_n$, de densité s'écrivant $f_X(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$ 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 29. — 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, $E \left(\prod_{i=1}^n f_i(X_i) \right) = \prod_{i=1}^n E(f_i(X_i))$
4. $\Phi_X = \otimes_{i=1}^n \Phi_{X_i}$

Résultats principaux

THÉORÈME 30 (Inégalité de Markov). — Soit X une variable aléatoire réelle presque sûrement positive, alors pour $\alpha > 0$:

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

THÉORÈME 31 (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 32 (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 $(\sigma(X_i, i \in I_k))_{k \in K}$ sont indépendantes.

THÉORÈME 33 (Loi faible des grands nombres). — Soit $(X_i)_{i \in \mathbb{N}^*}$ une famille de variables aléatoires indépendantes de L^2 , telles que $\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu$ et $\sup_i V(X_i) = \sigma^2$. Alors :

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

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

LEMME 34 (Borel-Cantelli). — Soit $(A_n)_{n \geq 0}$ une suite d'évènements. On note $\limsup A_n := \bigcap_{n \geq 0} \bigcup_{k \geq n} A_k$, qui correspond à l'évènement "être dans une infinité de A_n ". On a alors la dichotomie suivante :

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

THÉORÈME 35 (Loi du 0-1). — Si $(X_n)_{n \in \mathbb{N}}$ est une famille de variables aléatoires indépendantes, on note $\mathcal{F}_n^+ := \sigma(X_k, k \geq n)$ et $\mathcal{F}_\infty^+ := \bigcap_{n=1}^\infty \mathcal{F}_n^+$ la tribu asymptotique. Alors pour tout $A \in \mathcal{F}_\infty^+$, $P(A) \in \{0, 1\}$.

REMARQUE. On en déduit qu'une variable aléatoire \mathcal{F}_∞^+ -mesurable est presque sûrement constante en considérant sa fonction de répartition. En particulier, les \liminf et \limsup d'une suite de v.a. indépendantes sont constantes presque sûrement, donc soit cette suite converge p.s. vers une constante, soit la limite n'existe p.s. jamais.

THÉORÈME 36 (Loi forte des grands nombres). — Si $(X_n)_{n \in \mathbb{N}}$ est une suite de v.a. i.i.d. telle que $E(|X_1|) < \infty$, alors P p.s. on a :

$$\frac{X_1 + \cdots + X_n}{n} \rightarrow E(X_1)$$

REMARQUE. Avec des résultats d'uniforme intégrabilité, on peut aussi obtenir la convergence au sens L^1 dans le théorème ci-dessus.

THÉORÈME 37 (Théorème central limite). — Si $(X_n)_{n \in \mathbb{N}}$ est une suite de v.a. i.i.d. telle que $E(X_1^2) < \infty$, alors en notant $m := E(X_1)$ et $\sigma^2 := V(X_1)$, on a :

$$\sqrt{n}(\bar{X}_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

Outils importants

LEMME 38 (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 $(\frac{u_n}{n})_{n \in \mathbb{N}}$ converge, et on a l'égalité : $\lim_{n \rightarrow \infty} \frac{u_n}{n} = \inf_{n \geq 1} \frac{u_n}{n} \in \mathbb{R} \cup \{-\infty\}$.

PROPOSITION 39 (Formule de transfert). — Soit X une v.a. à valeurs dans (E, \mathcal{E}) , et $f : E \rightarrow \bar{\mathbb{R}}$ une fonction mesurable telle que $f \geq 0$ p.p. ou $E(|f(X)|) < \infty$, alors :

$$E(f(X)) = \int_E f(x) dP_X(x).$$

THÉORÈME 40 (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.

COROLLAIRE 41 (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 42 (Inégalité de Hoeffding). — Soient $(X_i)_{i \in \mathbb{N}^*}$ une famille de v.a. indépendantes à valeurs dans $[a, b]$. Alors pour tout $\epsilon > 0$:

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

THÉORÈME 43 (Lévy). — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a. à valeurs dans \mathbb{R}^d , alors : $X_n \xrightarrow{\mathcal{L}} X_\infty$ si et seulement si $\Phi_{X_n} \rightarrow \Phi_{X_\infty}$ simplement.

Autres résultats

LEMME 44. — Soit I un intervalle de \mathbb{R} . Si $\Phi : I \rightarrow \mathbb{R}$ est une fonction convexe, alors pour tout $x \in \overset{\circ}{I}$:

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

LEMME 45. — Pour une v.a. à valeurs dans \mathbb{R}^+ , $E(X) = \int_0^\infty P(X \geq x) dx$, car $X = \int_0^\infty 1_{x \leq X} dx$

Martingales et processus aléatoires

Systemes dynamiques