Fiche résumée du cours d'intégration et probabilités

1 Espaces mesurés

1.1 Ensembles mesurables

Définition 1.1.1 (Tribu) *Soit E un ensemble quelconque. Une* ***tribu*** *(ou σ-algèbre) sur E est une famille $\mathscr{E}$ de parties de E telle que :*

i) $E \in \mathscr{E}$

ii) $A \in \mathscr{E} \Rightarrow E - A \in \mathscr{E}$ *(noté A^C)*

iii) *si* $(A_n)_{n\in\mathbb{N}} \in \mathscr{E}$ *, alors* $\underset{n\in\mathbb{N}}{\cup} A_n \in \mathscr{E}$

On dit alors que $\mathscr{E}$ est l'ensemble des parties mesurables de E, ce qui fait de $(E,\mathscr{E})$ un éable. Si E est un espace topologique, on appelle ***tribu borélienne*** *la tribu engendrée par les ouverts, notée $\mathscr{B}(E)$.*

Définition 1.1.2 (Tribu engendrée par une partie) *Si $\mathscr{A} \in \mathscr{P}(E)$, on appelle* ***tribu engendrée par*** *$\mathscr{A}$ la plus petite tribu contenant $\mathscr{A}$, notée $\sigma(\mathscr{A})$.*

1.2 Mesures positives

Définition 1.2.1 (Mesure) *Soit $(E,\mathscr{E})$ un espace mesurable. Une* ***mesure*** *sur cet espace est une application $\mu : \mathscr{E} \to [0,+\infty]$, qui vérifie les propriétés suivantes :*

i) $\mu(\varnothing) = 0$

ii) Si $(A_n)_{n\in\mathbb{N}}$ est une suite de parties mesurables deux à deux disjointes, alors

$$\mu(\underset{n\in\mathbb{N}}{\sqcup} A_n) = \sum_n \mu(A_n)$$

Remarques : De façon générale, $\mu(\underset{n\in\mathbb{N}}{\cup} A_n) \leq \sum_n \mu(A_n)$. Par ailleurs, si $A \subset B$, tel que $\mu(A) < \infty$, alors, $\mu(B - A) = \mu(B) - \mu(A)$.

Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante, $\mu(\underset{n\in\mathbb{N}}{\cup} A_n) = \lim \uparrow \mu(A_n)$

Si $(B_n)_{n\in\mathbb{N}}$ est une suite décroissante,**telle que** $\mu(\mathbf{B_0}) < \infty$, alors $\mu(\underset{n\in\mathbb{N}}{\cap} B_n) = \lim \downarrow \mu(B_n)$

Définition 1.2.2 (Diverses) *i) On dire que μ est* ***finie*** *si $\mu(E) < \infty$.*

ii) On dire que μ est une ***probablilité*** *si $\mu(E) = 1$.*

iii) On dire que μ est ***σ-finie*** *s'il existe $(E_n)_{n\in\mathbb{N}}$ partition dénombrable de E tq, $\forall n \in \mathbb{N}$, $\mu(E_n)$ est finie.*

iv) On dira que $x \in E$ est un ***atome*** *de μ si $\{x\} \in \mathscr{E}$ et $\mu(\{x\}) > 0$.*

v) On dira que μ est ***diffuse****, ou* ***continue****, si elle n'admet pas d'atome.*

Définition 1.2.3 (Mesure de Dirac) *Soit $(E,\mathscr{E},\mu)$ un espace mesuré. On appelle* ***mesure de Dirac*** *la mesure $\delta(A) = \mathbf{1}_{\{0\in A\}}$*

Théorème 1.2.1 (Lemme de Borel-Cantelli) *Soit $(A_n)_{n\in\mathbb{N}}$ une suite de sous-ensembles d'un espace mesuré $(E,\mathscr{A},\mu)$. Alors, si $\sum\limits_{n\geq 1} \mu(A_n) < \infty$, on a $\mu(\limsup\limits_{n\to+\infty} A_n) = 0$, où l'on a posé $\limsup\limits_{n\to+\infty} A_n = \bigcap\limits_{n\geq 1} \bigcup\limits_{k\geq n} A_k$*

1.3 Fonctions mesurables

Définition 1.3.1 (Fonction mesurable) *Soient $(E,\mathscr{E})$, et $(F,\mathscr{F})$ deux ensembles mesurables. $f : E \to F$ est dite mesurable si, $\forall B \in \mathscr{F}$ $f^{-1}(B) \in \mathscr{E}$. f sera dite* ***borélienne*** *si elle est mesurable entre deux espaces munis de leur tribus borélienne.*

Définition 1.3.2 (Convergence en mesure) *Soit $(E, \mathscr{E}, \mu)$ un espace mesuré de mesure finie. Soient $(f_n)_{n\in\mathbb{N}}$ et f des fonctions $(E, \mathscr{E})$ dans $\mathbb{R}$. On dit que $(f_n)_{n\in\mathbb{N}}$ converge en mesure vers f si, $\forall\varepsilon$*

$$\mu(\{|f - f_n| > \varepsilon\}) \underset{n\to\infty}{\to} 0$$

Proposition 1.3.1 *Si $\mathscr{F} = \sigma(\mathscr{A})$ est une tribu sur F engendrée par une famille $\mathscr{A}$ de parties de F, $f : (E, \mathscr{E}) \to (F, \mathscr{F})$ est mesurable si et ssi $\forall A \in \mathscr{A}$, $f^{-1}(A) \in \mathscr{E}$. En conséquence, si E et F sont topologiques, toute fonction continue $f : E \to F$ est également borélienne.*

Opérations sur les fonctions mesurables : Le produit (ensembliste) de deux fonctions mesurables f_1 et f_2, $x \mapsto (f_1(x), f_2(x))$ est encore mesurable (sur l'espace produit muni de la tribu produit, engendrée par les pavés de $\mathscr{E} \times \mathscr{F}$). La somme, le sup et l'inf de deux fonctions mesurables le sont également.

NB : Toute fonction continue est à fortiori mesurable.

Définition 1.3.3 (limsup, liminf d'une suite de réels) *Soit $(a_n)_{n\in\mathbb{N}}$ une suite de réels. On définit la limite supérieure, (sup des valeurs d'adhérence de (a_n)) et la limite inférieure, (inf des valeurs d'adhérence de (a_n)) par :*

$$\limsup \quad a_n = \lim_{n\to+\infty} \downarrow \sup_{k\geq n}(a_n), \textit{ parfois notée } \overline{\lim_{n\to+\infty}}\, a_n$$

$$\liminf \quad a_n = \lim_{n\to+\infty} \uparrow \inf_{k\geq n}(a_n), \textit{ parfois notée } \underline{\lim_{n\to+\infty}}\, a_n$$

Par ailleurs, $\lim \quad a_n$ *existe si et ssi* $\limsup(a_n) = \liminf(a_n) = l$.

Proposition 1.3.2 *Soient f_n, $n \in \mathbb{N}$ des fonctions mesurables $E \to \mathbb{R}$. Alors, $\sup_n f_n : x \mapsto \sup_n f_n(x)$, $\inf_n f_n : x \mapsto \inf_n f_n(x)$, $\lim_{n\to+\infty} \sup(f_n)$ et $\lim_{n\to+\infty} \inf(f_n)$ sont mesurables. D'autre part, $\{x \in E \mid \exists \lim_{n\to+\infty} f_n(x) \in \mathbb{R}_+\}$ est mesurable.*

1.4 Théorème des classes monotones

Définition 1.4.1 (Classe monotone) *Une famille $\mathscr{M}$ de parties de E est dite **classe monotone** si :*

i) $E \in \mathscr{M}$

ii) $(A, B) \in \mathscr{M}^2$ *avec* $A \subset B$, *alors* $B - A \in \mathscr{M}$

iii) *si* $(A_n)_{n\in\mathbb{N}} \in \mathscr{M}$ *est une suite croissante de* $\mathscr{M}$, *i.e.* $A_n \subset A_{n+1} \forall n$, *alors* $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{M}$

Remarque : Pour qu'une classe monotone soit une tribu, il suffit qu'elle soit stable par union finie ou par intersection finie.

Théorème 1.4.1 (Des classes monotones) *Soit $\mathscr{C}$ une famille de parties de E. Si $(\mathscr{M}_i)$ est la famille de classes monotones contenant $\mathscr{C}$, alors $\cap\mathscr{M}_i$ est encore une classe monotone, c'est la plus petite classe monotone contenant $\mathscr{C}$, on la note $\mathscr{M}(\mathscr{C})$. Si $\mathscr{C}$ est stable par intersection finie, alors $\mathscr{M}(\mathscr{C}) = \sigma(\mathscr{C})$ est une tribu.*

Proposition 1.4.1 *Soient μ, ν deux mesures sur $(E, \mathscr{E})$, et $\mathscr{C}$ une classe de parties mesurables telles que $\mu(A) = \nu(A) \quad \forall A \in \mathscr{C}$. On suppose $\mathscr{C}$ stable par intersection finie.*

- *Si μ et ν sont finies, et que $\mu(E) = \nu(E)$ alors μ et ν coïncident sur $\sigma(\mathscr{C})$ (En particulier, $\nu = \mu$ si $\sigma(\mathscr{C}) = \mathscr{E}$).*
- *S'il existe $(E_n)_{n\in\mathbb{N}}$ une suite dans $\mathscr{E}$ avec $\mu(E_n) = \nu(E_n) < \infty$ et $\cup E_n = E$, alors μ et ν coïncident sur $\sigma(\mathscr{C})$*

2 Intégration

2.1 Intégrale de fonctions positives

Définition 2.1.1 (fonction étagée) *Une fonction $f : (E, \mathscr{E}) \to \mathbb{R}_+$, mesurable, est dite étagée si elle ne prend qu'un nombre fini de valeurs. f est étagée si ssi $f = \sum_{i=1}^{n} \alpha_i . \chi_{A_i}$, avec $0 \leq \alpha_i < ... < \alpha_n$ sont les différentes valeurs que peut prendre f et $A_i = \{x \in E, f(x) = \alpha_i\}$ On pose alors, pour toute mesure m, $\int f dm = \sum_{i=1}^{n} \alpha_i . m(A_i)$, avec la convention $0 * \infty = 0$ si $\alpha_i = 0$ et $m(A_i) = \infty$.*

Remarque : Propriétés élémentaires ; l'intégrale pour les fonctions étagées est linéaire et conserve l'ordre.

Définition 2.1.2 (Intégrale d'une fonction positive) *Soit $f : E \to \mathbb{R}_+$, mesurable. On définit $\int f d\mu = sup\{\int h d\mu$, pour h étagée , $0 \leq h \leq f\}$*

Remarque : Propriétés élémentaires ; l'intégrale pour les fonctions positives est linéaire et conserve l'ordre. (Cette propriété découle du fait que, pour toute fonction mesurable positive, on peut trouver une suite de fonctions étagées qui converge vers elle)

Théorème 2.1.1 (Beppo-Levi, théorème de convergence monotone) *Soit $(f_n)_{n \in \mathbb{N}}$, une suite croissante de fonctions mesurables $E \to \mathbb{R}_+$, avec $f_\infty = \lim \uparrow f_n$ qui est mesurable. Alors,*

$$\lim_{n \to +\infty} \uparrow \int f_n d\mu = \int \lim_{n \to +\infty} \uparrow f_n d\mu = \int f_\infty d\mu \, .$$

Proposition 2.1.1 *Soit f mesurable positive, $E \to \overline{\mathbb{R}}_+$*

i) (Inégalité de Markov) $\forall a > 0, \mu(\{x \in E, f(x) \geq a\}) \leq \frac{1}{a} \int f d\mu$

ii) Si $\int f d\mu < \infty$, alors $\mu(\{x \in E \, , \, f(x) = \infty\}) = 0$. On dit que f est finie μ-presque partout

iii) $\{x, f(x) > 0\}$ est de mesure nulle si et ssi $\int f d\mu = 0$ ($f = 0$ μ-presque partout)

Lemme 2.1.1 (Fatou) *Soit une suite de fonctions $f_n : E \to \mathbb{R}_+$ mesurables positives. On rapelle que $\liminf(f_n)$ est mesurable, et on a $\int \liminf(f_n) \leq \liminf \int f_n d\mu$*

2.2 Fonctions intégrables

Définition 2.2.1 *On dit que f est intégrable par rapport à une mesure μ si $\int |f| d\mu$. On notera*

$$\mathscr{L}_1(E, \mu) = \{f : E \to \mathbb{R}, \text{ mesurables telles que } \int |f| d\mu < \infty\}$$

Proposition 2.2.1 *L'intégrale de Lebesgue verifie l'inégalité triangulaire, est linéaire et conserve les inégalités. On peut également intégrer coordonnées par coordonnées des fonctions à valeurs dans $\mathbb{R}^d$*

Théorème 2.2.1 (Convergence dominée, Lebesgue) *Soit $f_n : E \to \mathbb{R}$ une suite de fonctions mesurables.On suppose :*

i) $f_n(x) \underset{n \to \infty}{\to} f(x)$ μ-presque partout

ii) $\exists g \in \mathscr{L}_1^+$ telle que $|f_n(x)| \leq g(x) \quad \forall n$ μ-presque partout

Alors, $f_n, f \in \mathscr{L}_1$, et $\int |f_n - f| d\mu \to 0$. En particulier, $\int f_n d\mu$ converge vers $\int f d\mu$.

Proposition 2.2.2 (Lemme de Scheffé) *Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions $E \to \mathbb{R}_+$, telles que $f_n \to f$ μ-presque partout. On suppose $\int f d\mu$ finie et que $\int f_n d\mu$ converge vers $\int f d\mu$. Alors, on a $\int |f_n - f| d\mu \to 0$*

Théorème 2.2.2 (Egoroff) *Soit $(E, \mathscr{E}, \mu)$ un espace mesuré, de mesure **finie**. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions telle que $f_n \to f$ $\mu(dx)$-presque partout. Alors, $\forall \varepsilon > 0$, $\exists A, \mu(A) < \varepsilon$, et tel que f_n converge uniformément vers f sur $E - A$*

2.3 Intégrales dépendant d'un paramètre

Théorème 2.3.1 *On se donne un espace métrique (U, d) et une application $f : U \times E \to \mathbb{R}$. On suppose que :*

i) $\forall u \in U$, $f(u, .) x \to f(u, x) \in \mathscr{L}_1(\mu)$

ii) $\mu(dx)$-presque partout, $f(., x) u \to f(u, x)$ est continue en un point u_0

iii) $\exists g \in \mathscr{L}_1^+(\mu)$ telle que $|f(u, x)| \leq g(x)$ $\mu(dx)$-presque partout, $\forall u \in U$

Alors, $I(u) = \int_E f(u, x) \mu(dx)$ est une application $U \to \mathbb{R}$, continue en u_0.

Théorème 2.3.2 *Soit I un intervalle ouvert de I, $u_0 \in I$, et une application $f : I \times E \to \mathbb{R}$. On suppose que :*

i) $\forall u \in I,\ f(u,.)x \to f(u,x) \in \mathscr{L}_1(\mu)$

ii) $\mu(dx)$*-presque partout,* $f(.,x)u \to f(u,x)$ *est dérivable en* u_0 *de dérivée* $\partial_u(u_0,x)$

iii) $\exists g \in \mathscr{L}_1^+(\mu)$ *telle que* $|f(u,x) - f(u_0,x)| \leq g(x)|u - u_0|$ $\mu(dx)$*-presque partout,* $\forall u \in U$

Alors, $F(u) = \int_E f(u,x)\mu(dx)$ *est une application* $I \to \mathbb{R}$*, dérivable en* u_0 *avec pour dérivée* $F'(u_0) = \int_E \partial_u f(u_0,x)\mu(dx)$.

3 Théorème de dérivation de Lebesgue

3.1 Dérivée des intégrales

Théorème 3.1.1 (Dérivation, Lebesgue) *Soit* $f \in \mathscr{L}^1(\mathbb{R})$*, formons*

$$F : x \mapsto \int_0^x f(t)dt$$

Alors, F *est dérivable* $\lambda-$*presque partout, et* $\partial F(x) = f(x)$ $\lambda(dx)$*-presque partout*

Définition 3.1.1 (Fonction maximale) *Soit* $f \in \mathscr{L}^1(\mathbb{R})$ *pour* $x \in \mathbb{R}$*, on définit*

$$Mf(x,r) = \frac{1}{\lambda(B(x,r))}\int_{B(x,r)} |f|(t)dt$$

et la fonction maximale

$$Mf(x) = \sup_{r>0} Mf(x,r)$$

Proposition 3.1.1 *La fonction maximale est borélienne.*

Définition 3.1.2 ($\mathscr{L}^1(\mathbb{R})$-faible) *Soit* $f : \mathbb{R} \to \mathbb{R}$*, borélienne, on pose*

$$\|f\|_{\mathscr{L}^1_W(\mathbb{R})} = \sup_{t>0}(\lambda\{x, |f(x)| > t\}.t)$$

Si $\|f\|_{\mathscr{L}^1_W} < \infty$*, on dit que* $f \in \mathscr{L}^1_W$

Théorème 3.1.2 (Hardy-Littlewood) *Soit* $f \in \mathscr{L}^1(\mathbb{R})$*, alors* $Mf \in \mathscr{L}^1_W$ *et* $\|Mf\|_{\mathscr{L}^1_W} \leq 3\|f\|_{\mathscr{L}^1}$

Lemme 3.1.1 (Vitali) *Soit* $\Omega = \bigcup_{i=1}^{n} B(x_i,r_i)$ *dans* $\mathbb{R}^r$*. Alors, il existe une suite extraite de boules disjointes, telle que* $\Omega \subset \bigcup_{k\geq 0} B(x_{i_k}, 3r_{i_k})$

3.2 Intégrales des dérivées

Proposition 3.2.1 *Supposons que l'on aie* F *dérivable* λ*-presque partout de dérivée* $\partial F \in \mathscr{L}^1$*, et* $F(x) = F(0) + \int_0^x \partial F(t)dt$*. Alors,* F *est uniformément continue, et* F *est à variations bornées.*

Définition 3.2.1 (Absolue continuité) $F : \mathbb{R} \to \mathbb{R}$ *est dite* ***absolument continue*** *si* $\forall \varepsilon > 0$, $\exists \eta > 0$ *pour tout* $]\alpha_1, \beta_1[...]\alpha_n, \beta_n[$ *disjoints, tels que* $\sum_{i=1}^{n} |F(\beta_i) - F(\alpha_i)| < \eta$*, on aie* $\sum_{i=1}^n |F(\alpha_i - \beta_i)| < \varepsilon$*. Ici, on a*

$$\sum_{i=1}^{n} |F(\alpha_i - \beta_i| \leq \int_D |\partial F(t)|dt \underset{D}{\to} 0$$

Où $D = \cup]\alpha_i, \beta_i[$.

Théorème 3.2.1 (Lebesgue) *Soit* F *absolument continue. Alors* F *est dérivable presque partout de dérivée* $\partial F \in \mathscr{L}^1_{loc}(\mathbb{R})$*, et telle que* $F(x) = F(0) + \int_0^x \partial F(t)dt$.

4 Construction de mesures

4.1 Mesures extérieures et mesures

Définition 4.1.1 (Mesure extérieure) *Soit $\mathscr{E}$ un ensemble quelconque. Une **mesure extérieure** sur cet espace est une application $\mu^* : \mathscr{P}(E) \to [0, +\infty]$, qui vérifie les propriétés suivantes :*

i) $\mu^*(\varnothing) = 0$

ii) μ^* *croissante :* $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$

iii) ***(σ sous additivité)*** *Si* $(A_n)_{n\in\mathbb{N}} \in \mathscr{P}(E)^{\mathbb{N}}$

$$\mu^*(\underset{n\in\mathbb{N}}{\cup} A_n) \leq \sum_n \mu^*(A_n)$$

Définition 4.1.2 (Partie mesurable) *Une partie de E est dite **μ^*-mesurable**, et on note $B \in \mathscr{M}(\mu^*)$ si, $\forall A \in \mathscr{P}(E)$, $\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$*

Théorème 4.1.1 *i) $\mathscr{M}(\mu^*)$ est une tribu qui contient toutes les parties B telles que $\mu^*(B) = 0$*

ii) La restriction μ de μ^* à $\mathscr{M}(\mu^*)$ est une mesure.

4.2 Construction de la mesure de Lebesgue

Définition 4.2.1 (Mesure extérieure de Lebesgue) *Pour $A \subset \mathbb{R}$, on définit $\lambda^*(A) = inf\{\sum_{i=1}^{\infty}(b_i - a_i) \mid A \subset \overset{\infty}{\underset{i=1}{\cup}}]a_i, b_i[\}$*

Théorème 4.2.1 *i) λ^* est une mesure extérieure*

ii) $\mathscr{B}(\mathbb{R}) \subset \mathscr{M}(\lambda^*)$

iii) $\forall a \leq b$, *on a* $\lambda^*([a, b[) = b - a$

La restriction de λ^* à $\mathscr{B}(\mathbb{R})$ voire également à $\mathscr{M}(\lambda^*)$, s'apelle mesure de Lebesgue sur $\mathbb{R}$. C'est l'unique mesure sur $\mathscr{B}(\mathbb{R})$ qui vérifie iii)

Définition 4.2.2 (Mesure de Lebesgue sur $\mathbb{R}^d$) *On définit **la mesure de Lebesgue** sur $\mathbb{R}^d$ par :*

$$\lambda_d^*(A) = \inf\{\sum_{i=1}^{\infty}|P_i|,\ P_i \text{ pavés ouverts } \cup P_i \supset A\}$$

Où l'on a défini $|P| = \prod_{j=1}^{d}(b_j - a_j)$, pour $P = \prod_{j=1}^{d}]a_j, b_j[$

Théorème 4.2.2 (Complétion de mesure) *Soit $(E, \mathscr{E}, \mu)$ un espace mesuré. On dit qu'une partie $N \in \mathscr{P}(E)$ est μ négligeable, et on note $N \in \mathscr{N}$, s'il existe $A \in \mathscr{E}$, $N \subset A$ et $\mu(A) = 0$.On note $\overline{\mathscr{E}} = E \vee \mathscr{N} = \sigma(\mathscr{E}, \mathscr{N})$, la tribu engendrée par $\mathscr{E}$ et $\mathscr{N}$. $\overline{\mathscr{E}} = \{A \subset E \mid \exists B, B' \in \mathscr{E}$, avec $\mu(B' - B) = 0 \quad B \subset A \subset B'\} = \{A = B \cup N \mid (B, N) \in (\mathscr{E}, \mathscr{N})\}$*

Définition 4.2.3 (Tribu complétée) *$\overline{\mathscr{E}}$ est appellée tribu μ complétée.*

Proposition 4.2.1 *Si l'on pose $\overline{\mu}(A) = \mu(B)$, si $B \in \mathscr{E}$ et $B \subset A$ et $A - B \in \mathscr{N}$ ainsi définie est une mesure sur $\overline{\mathscr{E}}$, qui, restreinte à $\mathscr{E}$, coïncide avec μ.*

Proposition 4.2.2 *On a l'identité $\mathscr{M}(\lambda^*) = \overline{\mathscr{B}}(\mathbb{R}^d)$ (Tribu borélienne complète)*

Définition 4.2.4 (Mesure régulière) *La mesure de Lebesgue est dite **régulière** dans le sens suivant : $\forall A \in \overline{\mathscr{B}}(\mathbb{R}^d)$, on a*

$$\lambda(A) = \inf\{\lambda(\theta),\ \theta \text{ ouvert, } \theta \supset A\} = \sup\{\lambda(F),\ F \text{ fermé, } F \subset A\}$$

Proposition 4.2.3 *Soit (E, d) espace métrique, $\mathscr{B}(E)$ sa tribu borélienne, μ une mesure finie de E, alors μ est régulière dans le sens où, $\forall A \in \mathscr{B}(E)$, on a*

$$\mu(A) = \inf\{\mu(\theta),\ \theta \text{ ouvert, } \theta \supset A\} = \sup\{\mu(F),\ F \text{ fermé, } F \subset A\}$$

4.3 Comparaison avec l'intégrale de Riemann

Proposition 4.3.1 *Si g est Riemann-intégrable, Alors g est également mesurable par rapport à la tribu borélienne $\overline{\mathscr{B}}([a,b])$, et intégrable par rapport à la mesure de Lebesgue. Enfin,*

$$I(f) = \int g d\lambda$$

4.4 Mesure image

Définition 4.4.1 (Mesure image) *Soient $(E,\mathscr{E})$,$(F,\mathscr{F})$ deux espaces mesurables et $f : E \to F$, mesurable. Soit μ une mesure sur $(E,\mathscr{E})$. On appelle **mesure image** par f de μ la mesure ν, donnée par $\forall B \in \mathscr{F}$, $\nu(B) = \mu(f^{-1}(B))$*

Formule de changement de variable : Soit $\phi : F \to \mathbb{R}_+$, mesurable, alors $\int_F \phi d\nu = \int_E \phi \circ f d\mu$

Proposition 4.4.1 *i) $\forall x_0 \in \mathbb{R}^d$, l'image de λ_d par la translation x_0 est encore λ_d*

*ii) Réciproquement, si μ est une mesure **finie sur les parties bornées/compactes**, et si μ est invariante par translation sur $\mathbb{R}^d$, alors μ est proportionelle à la mesure de Lebesgue*

iii) Pour $c > 0$, l'image de λ_d par la dilatation $x \to cx$ est $c^{-d}\lambda_d$

4.5 Mesures de Stieltjes

Théorème 4.5.1 *Soit $F : \overline{\mathbb{R}} \to [0,1]$, croissante,avec $F(-\infty) = 0$ et $F(+\infty) = 1$.*

Définition 4.5.1 (Inverse continu à droite) *On introduit **l'inverse continu à droite, ou le quantile**, défini par :*

$$\begin{array}{cccccl} F^{-1} & : &]0,1[& \to & \mathbb{R} & \\ & & x & \mapsto & F(x) & = \inf\{y \in \mathbb{R} \mid F(y) > x\} \end{array}$$

Proposition 4.5.1 *Soit λ la mesure de Lebesgue sur $]0,1[$, la fonction F^{-1} est croissante, continue à droite, donc mesurable, et l'image de λ par F^{-1} est une mesure μ sur $\mathbb{R}$ dont la fonction de répartition est F. On appelle μ la mesure de Stieltjes associée à F, notée $\mu(dx) = dF(x)$ Enfin, si F est C^1, on a $dF(x) = F'(x)dx$.*

5 Les espaces L^p

5.1 Definitions et inégalités de Holder

Définition 5.1.1 ($\mathscr{L}^p$) *On se donne $(E,\mathscr{E},\mu)$ un espace mesuré, et $p \in [1,+\infty]$. $\mathscr{L}^p(E,\mu)$ est l'ensemble des fontions $f : E \to \mathbb{R}$, mesurables, et telles que :*

- *Si p est fini, $\int_{\mathbb{R}} |f|^p d\mu$ soit finie*
- *si $p = \infty$, f est essentiellement bornée, i.e $\exists c$, fini tq $\mu\{x, |f(x)| > c\} = 0$*

On introduit la relation d'équivalence $\sim$ sur $\mathscr{L}^p$ définie par

$$f \sim g \Leftrightarrow f = g \ \mu\text{-presque partout}$$

On note alors $L_p = \mathscr{L}^p / \sim$. On notera enfin $\|f\|_p = (\int |f|^p d\mu)^{\frac{1}{p}}$

Théorème 5.1.1 (Inégalité de Holder) *Si $f \in L^p$,$g \in L^q$ sont deux fonctions mesurables $E \to \mathbb{R}$, et p, q conjugués dans $[1,+\infty]$ (ie $\frac{1}{p} + \frac{1}{q} = 1$), alors $fg \in L^1$ et*

$$\|fg\|_1 = \int |fg| d\mu \leq \|f\|_p . \|g\|_q$$

Corollaire 5.1.1 *Si μ est une mesure finie, et $1 \leq p \leq r \leq +\infty$, alors $L^r \subset L^p$. Si μ est une probabilité, on a même $\|f\|_p \leq .\|f\|_r$*

Proposition 5.1.1 (Inégalité de Jensen) *Soit μ une mesure de probabilité, $f \in L^1(\mu)$. Soit $\varphi : \mathbb{R} \to \mathbb{R}_+$, convexe (donc continue et mesurable), alors :*

$$\varphi(\int f d\mu) \leq \int \varphi(f) d\mu$$

5.2 Structure d'espace de Banach de L^p

Définition 5.2.1 (Espace de Banach) *Un espace vectoriel normé E est un **Espace de Banach** s'il est complet et que toute suite de Cauchy dans E converge dans E.*

Définition 5.2.2 (Convergence normale) *Soit (x_n) une suite dans un espace vectoriel normé E, on dit que la série (x_n) est normalement convergente si $\sum \|x_n\| < \infty$*

Lemme 5.2.1 *Pour qu'un e.v.n E soit un espace de Banach, il faut et il suffit que toute série de E normalement convergente converge dans E.*

Théorème 5.2.1 (Riesz) *$(L^p, \|.\|)$ est un espace vectoriel normé, et est un espace de Banach.*

Corollaire 5.2.1 *On munit L^2 de $< f, g >= \int fg d\mu$, pour $f, g \in L^2$. $< ., . >$ ainsi défini est un produit scalaire qui fait de L^2 un espace de Hilbert.*

Proposition 5.2.1 *i) Si (f_n) est une suite de L^p telle que $\sum \|f_n\|_p < \infty$, alors $\sum |f_i(x)| < \infty$ μ-presque partout*

ii) Si h_n est une suite de Cauchy dans L^p (elle converge alors dans L^p, au sens de $\|.\|_p$), on peut en extraire une sous suite qui converge μ-pp.

5.3 Approximation par des fonctions régulières

Définition 5.3.1 (Mesure extérieurement régulière) *Une mesure μ sur la tribu borélienne est **extérieurement régulière** si $\forall B \in \mathscr{B}$, on a $\mu(B) = Inf\{\mu(\theta), \theta$ ouvert contenant $B\}$*

Proposition 5.3.1 *Soit $p \in [1, +\infty[$. Soit (E, d) un espace métrique muni de la tribu Borélienne, et μ une mesure extérieurement régulière. Alors, $\forall f \in L^p, \exists (f_n)$ une suite de fonctions bornées et lipshitziennes telles que $\|f - f_n\|_p \to 0$*

Théorème 5.3.1 *Soit $p \in [1, +\infty[$, et $f \in L^p(\mathbb{R}_d, \lambda_d)$. Alors, $\exists (f_n)$ une suite de fonctions C^∞, à supports compacts, telles que $\|f - f_n\|_p \to 0$*

Théorème 5.3.2 *Considérons $\mathscr{F} : f \mapsto \begin{array}{cccc} \widehat{f} & : & \mathbb{R} & \to & \mathbb{C} \\ & & \xi & \mapsto & \int_{\mathbb{R}} e^{i\xi x} f(x) dx) \end{array}$ la transformée de fourrier. C'est un opérateur continu de $\mathscr{L}^1(\mathbb{R} \to \mathscr{C}_b(\mathbb{R}$ espace des fonctions continues bornées sur $\mathbb{R}$. Si $f \in \mathscr{L}^1(\mathbb{R})$, alors $\widehat{f} \in \mathscr{C}_0(\mathbb{R})$, ie la transformée de fourrier est continue et tend vers 0 en $+\infty$ et $-\infty$*

5.4 Le théorème de Radon Nikodym

Définition 5.4.1 (Mesures absolument continues, étrangères) *Soient μ et ν deux mesures sur (E, ξ). On dit que ν est **abslument continue** par rapport à μ, noté $\nu \ll \mu$ si toute partie μ-negligeable est ν-negligeable. Cette définition est également valide pour des mesures signées. On dit que μ et ν sont **étrangères** noté $\mu \perp \nu$, si il existe $N \subset E$ vérifiant $\mu(N) = \nu(N^C) = 0$*

Définition 5.4.2 (Densité de mesure) *Soient μ, ν deux mesures sur $(E,\mathscr{E})$. On dit qu'une fonction f mesurable positive est le mesure de densité de la mesure ν par rapport à μ si on peut écrire, $\forall A \in \mathscr{E}$,*

$$\nu(A) = \int_A f(x)\mu(dx)\ .$$

Théorème 5.4.1 (Radon-Nikodym) *Soient $(E,\mathscr{E})$ un espace mesurable, et μ et ν deux mesures **sigma finies**. Alors :*

i) Il existe deux mesures ν_a et ν_s, uniques, telles que :

$$\star \quad \nu = \nu_a + \nu_s \qquad \star \quad \nu_a \ll \mu \qquad \star \quad \nu_s \perp \mu$$

ii) $\exists f : E \rightarrow \mathbb{R}^+$, mesurable, et μ-essentiellement unique, telle que $d\nu_a = f d\mu$.

NB : Dans les applications, on utilise souvent $\nu \ll \mu \Rightarrow \exists f, d\nu = f d\mu$
Lemme : Soient μ et ν deux mesures finies sur E telles que $\nu \leq \mu$, ie $\forall A \in \mathscr{E}, \nu(A) \leq \mu(A)$ (cette notion est bcp plus forte que $\ll$). Alors, il existe une fonction mesurable $h : E \rightarrow [0;1]$ telle que $d\nu = h d\mu$.

6 Intégration sur un espace produit

6.1 Quelques préléminaires sur les tribus produit

Définition 6.1.1 (Produit tensoriel de tribus) *Soient $(E_i,\mathscr{E}_i))_{i\in\{1..k\}}$ des espaces mesurés. On munit alors $E = E_1 \times E_2 \times ... \times E_k$ de la tribu $\mathscr{E} = \mathscr{E}_1 \otimes \mathscr{E}_2 \otimes ... \otimes \mathscr{E}_k$, définie par $\mathscr{E}_1 \otimes \mathscr{E}_2 \otimes ... \otimes \mathscr{E}_k = \sigma(\mathscr{E}_1 \times \mathscr{E}_2 \times ... \times \mathscr{E}_k)$. De manière équivalente, la tribu produit est la plus petite tribu pour laquelle les projections sont mesurables.*

Proposition 6.1.1 *Soient (E_1,d_1), (E_2,d_2) deux espaces métriques séparables (c'est à dire à famille dénombrable dense). $E_1 \times E_2$ est muni de la topologie produit $(d((x,y),(x',y')) = sup(d_1(x,x'),d_2(y,y')))$. On a alors :*

$$\mathcal{B}(E_1 \times E_2) = \mathcal{B}(E_1) \otimes \mathcal{B}(E_2)$$

Définition 6.1.2 (Fibres) *Soit $C \subset E_1 \times E_2$, $x \in E_1$, et $y \in E_2$. On définit C_x, la **fibre au dessus de x**, par $C_x = \{y' \in E_2, (x,y') \in C\}$. De même, on définit C^y, la **fibre au dessus de y**, par $C^y = \{x' \in E_1, (x',y) \in C\}$.*

Proposition 6.1.2 *Soient (E_1,d_1), (E_2,d_2) espaces mesurables, et $C \in \mathscr{E}_1 \otimes \mathscr{E}_2$.*

i) Alors, $\forall x \in E_1$, $C_x \in \mathscr{E}_1$, et $\forall y \in E_1$, $C^y \in \mathscr{E}_2$

ii) Soit μ une mesure finie sur $(E_1,\mathscr{E}_1)$, et ν une mesure sur $(E_2,\mathscr{E}_2)$. Alors :

$$\begin{array}{ccc} E_1 & \rightarrow & \mathbb{R}^+ \\ x & \mapsto & \nu(C_x) \end{array} \quad \textit{est } \mathscr{E}_1\textit{-mesurable}$$

et

$$\begin{array}{ccc} E_2 & \rightarrow & \mathbb{R}^+ \\ y & \mapsto & \mu(C^y) \end{array} \quad \textit{est } \mathscr{E}_2\textit{-mesurable}$$

6.2 Mesure produit

Théorème 6.2.1 *Soit $(E,\mathscr{E})$, et $(F,\mathscr{F})$ deux espaces mesurables, munis de deux mesures σ-finies, μ et ν. Alors :*

i) Il existe une unique mesure m sur $(E \times F, \mathscr{E} \otimes \mathscr{F})$ telle que, pour tout pavé $A \times B \in \mathscr{E} \otimes \mathscr{F}$, $m(A \times B) = \mu(A)\nu(B)$. On appelle m la mesure produit notée $m = \mu \otimes \nu$, ou $m(dxdy) = \mu(dx)\nu(dy)$.

ii) Si $C \in \mathscr{E} \otimes \mathscr{F}$, on a

$$m(C) = \int_{E_1} \mu(dx)\nu(C_x) = \int_{E_2} \nu(dy)\mu(C^y)$$

Remarque : La mesure de Lebesgue sur $\mathbb{R}^d$ est la d-mesure produit des mesures de Lebesgue sur $\mathbb{R}$.

6.3 les théorèmes de Fubini

Théorème 6.3.1 (Fubini-Tonelli, cas positif) *Soient* $(E, \mathscr{E}, \mu)$, *et* $(F, \mathscr{F}, \nu)$ *deux espaces mesurés, avec* μ *et* ν σ*-finies. On considère* $f : E \times F \to \mathbb{R}_+$. *Alors, les applications*

$$x \mapsto \int_F f(x,y)\nu(dy)$$

$$y \mapsto \int_E f(x,y)\mu(dx)$$

sont respectivement $\mathscr{E}$*-mesurables et* $\mathscr{F}$*-mesurables. Enfin*

$$\int\limits_{E\times F} fd(\mu\otimes\nu) = \int\limits_E \mu(dx)\left(\int\limits_F \nu(dy)f(x,y)\right) = \int\limits_E \mu(dx)\left(\int\limits_F \nu(dy)f(x,y)\right)$$

Théorème 6.3.2 (Fubini-Lebesgue) *Soient* $(E, \mathscr{E}, \mu)$, *et* $(F, \mathscr{F}, \nu)$ *deux espaces mesurés, avec* μ *et* ν σ*-finies. On considère* $f : E \times F \to \mathbb{R}$. ***On suppose que*** $\mathbf{f \in L^1(E \times F}, \mu\otimes\nu)$. *On a alors :*

i) $\mu(dx)$*-presque partout,* $y \mapsto f(x,y) \in L^1(F,\nu)$

ii) nu(dy)-presque partout, $x \mapsto f(x,y) \in L^1(E,\mu)$

iii) $x \mapsto \int_F f(x,y)\nu(dy)$ *est* $\mathscr{E}$*-mesurable, et dans* $L^1(E,\mu)$

iv) $y \mapsto \int_E f(x,y)\mu(dx)$ *est* $\mathscr{F}$*-mesurable, et dans* $L^1(F,\nu)$

v) $\int\limits_{E\times F} fd(\mu\otimes\nu) = \int\limits_E \mu(dx)\left(\int\limits_F \nu(dy)f(x,y)\right) = \int\limits_E \mu(dx)\left(\int\limits_F \nu(dy)f(x,y)\right)$

NB : Pour appliquer ce théorème (vérifier que $\mathbf{f \in L^1(E\times F}, \mu\otimes\nu))$**, on applique Fubini-Tonelli à** $|\mathbf{f}|$

6.4 Quelques applications

Définition 6.4.1 (Produit de convolution) *On définit le* ***produit de convolution*** *de deux fontions mesurables* f *et* g *par* $f * g : x \mapsto \int\limits_{\mathbb{R}^d} f(x-y)g(y)dy$. *Le produit de convolution est commutatif.*

Proposition 6.4.1 *Si* $f, g \in L^1(\mathbb{R}^d)$, *alors* $f * g(x)$ *est bien défini dx-pp. Plus précisément,* $f * g \in L^1(\mathbb{R}^d)$, *et* $\|f * g\|_1 \leq \|f\|_1\|g\|_1$

Proposition 6.4.2 *Soit* $p \in]1; +\infty[$, q *son conjugué. Si* $f \in L^p(\mathbb{R}^d)$, $g \in L^q(\mathbb{R}^d)$, *alors* $f * g(x)$ *existe* $\forall x \in \mathbb{R}^d$ *et* $f * g$ *est continue, avec limite 0 en l'infini.*

Lemme 6.4.1 *Si* $p \in [1; +\infty[$, $f \in L^p(\mathbb{R}^d, \lambda_d)$, *et* $f_\varepsilon(x) = f(x+\varepsilon)$, *alors* f_ε *converge dans* L^p *vers* f *quand* $\varepsilon \to 0$.

6.4.1 Définition ***(Approximation de l'unité)***

Considérons une famille $(\varphi_n)_{n\in\mathbb{N}}$ de fonctions $\mathbb{R}^d \to \mathbb{R}_+$. Si, $\forall n \in \mathbb{N}$, φ_n est continue, à support compact, et telle que $\int\limits_{\mathbb{R}^d} \varphi_n d\lambda_d = 1$, et si $\forall\varepsilon > 0$, $\lim\limits_{n\to\infty} \int \mathbf{1}_{\{\|x\|>\varepsilon\}}\varphi_n(x)dx = 0$, on dit alors que $(\varphi_n)_{n\in\mathbb{N}}$ est une **approximation de l'unité**.

6.4.2 Proposition

i) Si $f \in L^p$ et $(\varphi_n)_{n\in\mathbb{N}}$ est une approximation de l'unité, alors, $f*\varphi_n \in L^p$ et $\|f * \varphi_n\| \leq \|f\|_p$

ii) Si $f \in \mathscr{C}_c$ (ensemble des fonctions continues à support compact), alors $f * \varphi_n(x) \underset{n\to+\infty}{\to} f(x) \forall x$, et est dans L^p

iii) Si $f \in L^p$, $f * \varphi_n \to f$ dans L^p

7 Mesures signées

7.1 Mesures signées et variation totale

Définition 7.1.1 (Mesure signée) *Soit* $(E, \mathscr{E})$ *un espace mesurable, et* $\mu : \mathscr{E} \to \mathbb{R}$*. On dit que* μ *est une mesure signée si :*

i) $\mu(\emptyset) = 0$

ii) $\forall\ (A_n)_{n\in\mathbb{N}}$ *suite de* $\mathscr{E}$*, telle que* $A_i \cap A_j = \emptyset \quad \forall i \neq j$*,* $\sum \mu(A_n)$ *est absolument convergente, et* $\mu(\sqcup A_n) = \sum \mu(A_n)$

Théorème 7.1.1 *Si* μ *est une mesure signée, on note,* $\forall B \in \mathscr{E}$*,* $|\mu|(B) = sup \sum |\mu(A_n)|$ *pour* $(A_n)_{n\in\mathbb{N}} \in \mathscr{E}^{\mathbb{N}}$*,* $\sqcup A_n = B$*. Alors,* $|\mu|$ *est une mesure (positive) finie sur* $(E, \mathscr{E})$*. on l'appelle* ***variation totale de*** μ*.*

Lemme 7.1.1 *Supposons que* A *vérifie* $|\mu|(A) = \infty$*. Alors, il existe une partition* $A = B \sqcup C$ *telle que* $|\mu(B)| > 1$*,* $|\mu(C)| > 1$

7.2 Decomposition de Jordan, théorème de Radon-Nikodym

Définition 7.2.1 (Décomposition d'une mesure signée) *Soit* μ *une mesure signée, on pose*

$$\mu^+(A) = \frac{1}{2}(|\mu|(A) + \mu(A)) \text{ , et } \mu^-(A) = \frac{1}{2}(|\mu|(A) - \mu(A))$$

μ^+ *et* μ^- *ainsi définies sont des mesures positives, et on a la décomposition de Jordan* $|\mu| = \mu^+ + \mu^-$*, et* $\mu = \mu^+ - \mu^-$

Théorème 7.2.1 *Soit* μ *une mesure signée, alors il existe* $B \in \mathscr{E}$ *telle que*

$$\forall A \in \mathscr{E} \quad \mu^+(A) = \mu(A \cap B), \text{ et } \mu^-(A) = \mu(A \cap B^c)$$

En particulier, les mesures μ^+ *et* μ^- *sont orthogonales. Enfin, la décomposition* $\mu = \mu^+ - \mu^-$ *en deux mesures positives orthogonales est unique.*

Théorème 7.2.2 *Si* $f_n \in L^1(|\mu|) \forall n$*, et que* $f_n \to f$ $|\mu|$*-presque partout, s'il existe* $g \in L^1(|\mu|)$*, telle que* $|f_n| < g$ $|\mu|$*-presque partout* $\forall n$*, alors* $f \in L^1(|\mu|)$ *et* $\int f_n d\mu$ *converge vers* $\int f d\mu$

Théorème 7.2.3 (Radon-Nikodym) *Soient* $(E, \mathscr{E})$ *un espace mesurable,* ν *une mesures positive* ***sigma finie****, et* μ *une mesure signée. Alors , les assertions suivantes sont équivalentes :*

i) $\mu \ll \nu$

ii) (Uniforme absolue continuité) $\forall \varepsilon > 0, \exists \eta > 0$ *tel que* $\nu(A) < \eta \Rightarrow |\mu|(A) < \varepsilon$

iii) $\exists g \in L^1(\nu)$*, mesurable, et* μ*-essentiellement unique, telle que* $d\mu = g d\nu$*, et* $|\mu| = |g|\nu$

8 Dualité $L^p - L^q$

Définition 8.0.2 (Isométrie entre L^q et L^p*) *Soit* $(E, \mathscr{E}, \nu)$ *un espace mesuré, et soient* p, q *deux réels conjugués. Soit* $g \in L^q$ *posons*

$$\Phi_g(f) = \int_E |fg| d\nu \forall f \in L^p$$

L'application

$$\begin{array}{cccc} \Phi & : & L^q & \to & L^p* \\ & & g & \mapsto & \Phi_g \end{array}$$

est une isométrie, linéaire de L^q *dans* L^p*, qui est bijective si* $p < \infty$

Théorème 8.0.4 *Supposons que ν est une mesure σ-finie sur $(E,\mathscr{E})$, et $p \in [1,+\infty[$. Alors, si $\Phi : L^p \to \mathbb{R}$ est une forme linéaire continue sur L^p, alors il existe une unique $g \in L^q$ (p, q conjugués) telle que $\Phi(f) = \int_E fgd\nu$. En particulier, $\|\Phi\| = \|g\|_q$*

8.1 Complément culturel

Théorème 8.1.1 (Riesz) *Soit E un espace séparable métrique localement compact. Soit $\mathbb{C}_0 = \{f, \text{continue tendant vers 0 à l'infini}\}$. On munit $\mathbb{C}_0$ de la norme uniforme. Si $\Phi : \mathbb{C}_0 \to \mathbb{R}$ est une forme linéaire continue, alors il existe une unique mesure signée μ sur E muni de la tribu borélienne telle que $\Phi(f) = \int fd\mu$. On peut remarquer que μ est une mesure positive si et seulement si Φ est positive.*

9 Formule de changement de variable

Proposition 9.0.1 *Soit $b \in \mathbb{R}^d$, et M matrice $d \times d$, et $f : \begin{array}{ccc} \mathbb{R}^d & \to & \mathbb{R}^d \\ x & \mapsto & Mx+b \end{array}$. Pour tout A borélien de $\mathbb{R}^d$, on a*

$$\lambda(f(A)) = |detM|\lambda(A)$$

Théorème 9.0.2 *Pour toute fonction borélienne $f : D \to \mathbb{R}_+$ on a*

$$\int_D f(x)dx = \int_U f(\varphi(y))|j_\varphi(y)|dy$$

où j_φ est le jacobien, déterminant de la matrice jacobienne $J_\varphi(y) = \left(\frac{\partial \varphi_i}{\partial y_j}\right)_{i,j}$

Proposition 9.0.2 *Soit $K \subset U$ un compact ,$\varepsilon > 0$, on peut alors choisir $\delta > 0$ assez petit de sorte que, pour tout cube C de centre $u_0 \in K$ et de longueur $2l < \delta$, alors*

$$(1-\varepsilon)|j_\varphi(u_0)|\lambda(C) \leq \lambda(\varphi(C)) \leq (1+\varepsilon)|j_\varphi(u_0)|\lambda(C)$$

Théorème 9.0.3 *Soit A borélien de S^{d-1} (sphère en dimension d). Posons $\Gamma(A) = \{rx, r \in [0;1], x \in A\}$, et $\omega_d(A) = d.\lambda(\Gamma(A))$. ω_d est une mesure finie sur S_{d-1}, appellée mesure de Lebesgue sur S_{d-1} . $\forall f : \mathbb{R}^n \to \mathbb{R}_+$ borélienne,*

$$\int_{\mathbb{R}^n} f(x)dx = \int_0^\infty ds.s^{n-1} \int_{S^{n-1}} \omega_d(dz)f(sz)$$

10 les bases de la théorie des probabilités

10.1 Axiomatique de Kolmogorov

Définition 10.1.1 (Espace de probabilités, variable aléatoire) *Un **espace de probabilités** est un triplet $(\Omega,\mathscr{F},\mathbb{P})$, où $\mathscr{F}$ est une tribu sur un ensemble Ω, et $\mathbb{P}$ une mesure positive de probabilité $\mathbb{P}(\Omega) = 1$.*

- *L'élément $\omega \in \Omega$ s'appelle **l'aléa***
- *Une partie $\Lambda \in \mathscr{F}$ est un **évènement***
- *Une **variable aléatoire** est une application mesurable $X : \Omega \to E$ (E mesurable). X est une fonction mesurable qui dépend de l'aléa. La mesure image de $\mathbb{P}$ par X s'appelle **la loi de X**.*
- *Si X est une variable aléatoire positive, on apelle **espérance** de X, ou **valeur moyenne** de X la quantité*

$$E(X) = \int_\Omega X(\omega)d\mathbb{P}(\omega)$$

Définition 10.1.2 (Loi d'une variable aléatoire) *La **loi** de X, aussi apellée **distribution** de X, notée $\mathbb{P}_X$ est la mesure de probabilités définie par*
$$\mathbb{P}_X(B) = \mathbb{P}(x \in B) = \mathbb{P}(\omega \in \Omega, X(\omega) \in B\mathbb{P}(X^{-1}(B))$$

Lemme 10.1.1 *Si X est une variable aléatoire de loi $\mathbb{P}_X$ sur E, et si $f : E \to \mathbb{R}_+$ est mesurable alors $f(X)$ est une variable aléatoire (positive), et*
$$E(f(X)) = \int_E f(x)\mathbb{P}_X(dx)$$

i) **Lois dicrètes, à valeurs dans $\mathbb{Z}_+$**

a) Masse de dirac en k, δ_k, loi de la variable aléatoire constante égale à k

b) Loi de Bernouilli, sur $\{0;1\}$: Soit $p \in]0;1[$, β loi de Bernouilli avec paramètre p : $\beta(1) = p$, et $\beta(0) = 1 - p$

c) Loi Binomiale , sur $\{0, 1..n\}$: Soit $n \in \mathbb{N}$, $p \in]0;1[$. La loi binomiale de paramètres (n, p) est la mesure ρ sur $\{0, 1..n\}$, telle que
$$\rho(k) = \binom{k}{n} p^k (1-p)^{n-k}$$

d) Loi géométrique, sur $\mathbb{N}$: pour $p \in]0;1[$ elle est donnée par
$$\gamma(k) = p^{k-1}(1-p)$$

e) Loi de Poisson : on se donne $c \in]0, \infty[$. La loi de poisson de paramètre c est la mesure ν sur $\mathbb{Z}_+$ donnée par
$$\nu(k) = e^{-c}\frac{c^k}{k!}$$

ii) **Lois continues, à valeurs dans $\mathbb{R}$. On regarde principalement les mesures absolument continues par rapport à $\Lambda_\mathbf{d}$.**

a) Loi uniforme sur $[0;1]$: c'est la mesure de Lebesgue sur $[0;1]$. la loi uniforme sur $[a;b]$ est la loi de densité
$$\frac{1}{b-a}\mathbf{1}_{[a;b]}$$

b) Loi exponentielle de paramètre $c > 0$. C'est la loi de densité
$$ce^{-cx}$$
.

c) Loi de cauchy, avec pour densité
$$\frac{1}{\pi(1+x^2)}$$

d) Loi gaussienne, ou loi normale centrée réduite, avec pour densité
$$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$
Si $m \in \mathbb{R}$, $\sigma^2 > 0$, on note $\mathcal{N}(m, \sigma^2)$ la loi gaussienne de moyenne m et de variance σ^2, ie la loi de densité
$$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-m)^2}{2\sigma^2}}$$
Si ξ a pour loi $\mathcal{N}(0,1)$, alors $\xi' = \sigma\xi + m$ suit une loi $\mathcal{N}(m, \sigma^2)$

Définition 10.1.3 (Tribu engendrée par une variable aléatoire) *Soit $X : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ une variable aléatoire sur $(\Omega, \mathscr{F}, \mathbb{P})$ espace de probabilités, on appelle **tribu engendrée par X**, et on note $\sigma(X) = \{X^{-1}(B), B \in \mathscr{E}\}$ la plus petite tribu rendant X mesurable.*

Proposition 10.1.1 *Soit X une variable aléatoire à valeurs dans $(E, \mathscr{E})$, et Y une variable aléatoire réelle, $\sigma(X)$-mesurable, alors, il existe $f : E \to \mathbb{R}$, mesurable, telle que $Y = f(X)$*

10.2 Moments d'une variable aléatoire

Définition 10.2.1 (Moments entiers/positifs) *Soit X une variable aléatoire à valeurs dans $\mathbb{R}$. $\forall p > 0$, on dit que X a un moment absolu d'ordre p, et on note $X \in L^p$ si $E(|X|) < \infty$. Si $p \in \mathbb{N}$ et $X \in L^p$, on appelle moment d'ordre p de X la quantité $E(X^p) = \int x^p \mathbb{P}_X(dx)$*

Proposition 10.2.1 *On peut énoncer à nouveau les principaux resultats d'intégration.*

i) *Limite monotone : Si $X_n \geq 0$, $X_n \nearrow X$, alors$E(X_n) \nearrow E(X)$*

ii) *Fatou : Si $X_n \geq 0$, $E(\liminf X_n) \leq \liminf(E(X_n))$*

iii) *Lebesgue : $(X_n)_n$ une suite de variable aléatoire qui converge vers X, et si il existe $Y \geq 0$, $E(Y) < \infty$, et $|X_n| \leq Y$, alors*

$$E(|X_n - X|) \to 0$$

iv) *Holder : Soient p et q conjugués, $X \in L^p$, et $Y \in L^q$, alors $Y \in L^1$, et*

$$E(|XY|) \leq E(|X|^p)^{\frac{1}{p}}.E(|Y|^q)^{\frac{1}{q}}$$

v) *Jensen : Si $X \in L^1$, et $\varphi : \mathbb{R} \to \mathbb{R}_+$, convexe,*

$$E(\varphi(X)) \geq \varphi(E(X))$$

Définition 10.2.2 (Variance, covariance, et ecart type) *Soit X une variable aléatoire de L^2. On note que $L^2 \subset L^1$. On appelle **variance** de X, si on note $m = E(X)$, la quantité*

$$Var(X) = E(X^2) - E(X)^2 = E((X-m)^2)\ .$$

*En particulier, on pose $Var(X) = \sigma_X^2$, et on appelle σ_X **l'ecart type** de X. Si X, Y sont deux variables aléatoires, on définit la **covariance** de X et Y*

$$cov(X,Y) = E(XY) - E(X)E(Y) = E((X - E(X))(Y - E(Y)))\ .$$

Proposition 10.2.2 *On a également :*

i) *Inegalité de Markov : Soit $X \in L^1$, et $a > 0$, on a*

$$P(|X| > a) < \frac{1}{a}E(|X|)$$

ii) *Inegalité de Bienaymé-Tchebytchev : Soit $X \in L^2$, et $a > 0$, on a*

$$P(|X - E(X)| > a) < \frac{1}{a^2}Var(X)$$

Définition 10.2.3 (Fonction caractéristique) *On appelle fonction caractéristique de X la transformation de fourrier de sa loi*

$$\begin{array}{cccc} \Phi_X & : & \mathbb{R} & \to & \mathbb{C} \\ & & \xi & \mapsto & \int_{\mathbb{R}} e^{i\xi x}\mathbb{P}_X(dx) = E(e^{i\xi x}) \end{array}$$

Dans le cas d-dimensionnel, $\Phi_X(\xi) = E(e^{i\xi.x})$, avec $\xi.x = \sum \xi_i x_i$

Lemme 10.2.1 *Si X a pour loi $\mathcal{N}(0,\sigma^2)$, alors $\Phi_X(\xi) = e^{-\frac{\sigma^2\xi^2}{2}}$*

Proposition 10.2.3 *On a toujours $\Phi_X : \mathbb{R}^d \to \mathbb{C}$ est continue, bornée, avec $\Phi_X(0) = 1$. Par ailleurs, si la loi de X admet une densité par rapport à la mesure de Ldebesgue, alors $\Phi_X \underset{\infty}{=} 0$. Si X est réelle, $X \in L^k(\mathbb{P})$, $k \in \mathbb{N}$, alors Φ_X est C^k et, plus précisément, $\Phi^{(k)}(\xi) = E((ix)^k e^{i\xi X})$. Enfin, si $X \in \mathbb{R}^d$, $E(|X|^2) < \infty$, alors on a le développement limité*

$$\Phi_X(\xi) = 1 + i\sum_{k=1}^{d}\xi_k E(X_k) - \frac{1}{2}\sum_{i=1}^{d}\sum_{i=1}^{d}\xi_i\xi_n E(X_i X_n) + o(|\xi|^2)$$

Ou $X = (X_1 ... X_d)$

Théorème 10.2.1 *La fonction caractéristique d'une variable aléatoire X caractérise la loi de cette variable. Si X et X' sont deux variable aléatoire dans $\mathbb{R}^d$ telles que $\Phi_X = \Phi_{X'}$, alors X et X' ont la même loi. On a une formule d'inversion, dite **Formule de Plancherel**; si $\varphi : \mathbb{R}^d \to \mathbb{R}$, de transformée de fourier intégrable (ex , φC^2 à support compact), alors*

$$E(\varphi(X)) = \int \varphi(x)P_X(dx) = 2\pi^{-d}\int \widehat{\varphi}(\xi)\Phi_X(-\xi)d\xi$$

Lemme 10.2.2 *$f_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\int e^{i\xi x} g_{\frac{1}{\sigma}}(\xi)\widehat{\mu}(-\xi)d\xi$. En particulier, μ_σ s'exprime en fonction de $\widehat{\mu}$*

Lemme 10.2.3 *Si φ est borélienne bornée, alors $\int \varphi d\mu = \lim_{\sigma \to 0} \int \varphi d\mu_\sigma$*

11 Indépendance

11.1 Notion d'indépendance

Définition 11.1.1 (Evènements indépendants) *Soient $\lambda, \lambda' \in A$, on dit qu'ils sont indépendants, noté $\lambda \perp \lambda'$, si $\mathbb{P}(\lambda \wedge \lambda') = \mathbb{P}(\lambda)\mathbb{P}(\lambda')$*

Remarque : l'indépendance est stable par passage au complémentaire.

Définition 11.1.2 (Tribus indépendantes) *Si $\mathscr{F}_1 ... \mathscr{F}_n$ sont des sous tribus de $\mathscr{F}$, on dit qu'elles sont indépendantes si $\forall \lambda_i \in \mathscr{F}_i, i \in \{1...n\}$ on a $\mathbb{P}(\lambda_1 \wedge \lambda_2 \wedge ... \wedge \lambda_n) = \prod_{i=1}^{n} \mathbb{P}(\lambda_i)$*

Rappel : Si $(\mathscr{F}_i)_{i \in I}$ une famille de sous tribus , on note $\bigvee_{i \in I} \mathscr{F}_i = \sigma(\bigcup_{i \in I} \mathscr{F}_i)$ la plus petite tribu contenant les $\mathscr{F}_i$

Définition 11.1.3 (variable aléatoire indépendantes) *Des variables aléatoires $X_1 ... X_n$ sont dites indépendantes si les tribus qu'elles engendrent le sont, i.e. si*

$$\mathbb{P}(X_1 \in B_1, ..., Xn \in B_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in B_i)$$

On remarque que des évènement sont indépendants si les fonctions indicatrices associées sont indépendantes. Finalement, $X_1 ... X_n$ sont indépendantes si et seulement si pour toute fonctions boréliennes $f_1 ... f_n$ bornées, on a

$$E(f(X_1) ... f(X_n)) = E(f(X_1)) ... E(f(X_n))$$

Théorème 11.1.1 *Soient $X_1, ... X_n$ des variables aléatoires. Elles sont indépendantes si et seulement si la loi de $\underline{X} = (X_1 ... X_n)$ est donnée par $P_{\underline{X}} = P_{X_1} \otimes ... \otimes P_{X_n}$, ou encore si pour toute famille $(f_i)_i$ de fonctions mesurables, on a*

$$E(\prod_{i=1}^{n} f_i(X_i)) = \prod_{i=1}^{n} E(f_i(X_i))$$

Dans le cas de variables à densité, on a indépendance si la densité de $\underline{X}$ peut s'écrire comme fonction des $(x_i)_i$

Proposition 11.1.1 (Regroupement par paquets) *Soit $(\mathscr{B}_n)_{n \in \mathbb{N}}$ une suite de sous tribus indépendantes, et $I \subset \mathbb{N}$ quelconque. On pose $\mathscr{F}_I = \bigvee_{i \in I} \mathscr{B}_i$, $\mathscr{F}_{I^c} = \bigvee_{j \in I^c} \mathscr{B}_j$, alors $\mathscr{F}_I \perp \mathscr{F} I^c$. Plus généralement, si $I_1 ... I_k$ est une partition de $\mathbb{N}$ alors les $\mathscr{F}_{I_k}$ sont indépendantes.*

Théorème 11.1.2 (Loi du tout ou rien de Kolmogorov) *Soit $(\mathscr{B}_n)_{n \in \mathbb{N}}$ une suite de sous tribus indépendantes, posons $G_n = \bigvee_{i \geq n} \mathscr{B}_i$, et $G_\infty = \bigcap_n$ (Tribu queue). Alors, G_∞ est triviale, ie si $\Lambda \in G_\infty$, alors $\mathbb{P}(\Lambda) = 0$ ou 1*

11.2 Lemme de Borel-Cantelli

Lemme 11.2.1 (Borel Cantelli, partie directe) *Soient $(\lambda_n)_{n \in \mathbb{N}}$ une suite d'évènements. Si $\sum \mathbb{P}(\lambda_i) < \infty$. Alors*

$$\mathbb{P}(\limsup \lambda_n) = 0 \ .$$

Réciproque : *On suppose de plus que les λ_n sont indépendants. Alors, si $\sum \lambda_i = \infty$, alors $\mathbb{P}(\limsup \lambda_n) = 1$*

11.3 Somme de variables aléatoires

Définition 11.3.1 (Convolution de mesures) *Soient μ et ν deux mesures sur $\mathbb{R}^d$ on note $\mu * \nu$ l'unique mesure sur $\mathbb{R}^d$ définie par :*

$$\int_{\mathbb{R}^d} f(x) \mu * \nu(dx) = \int \int f(x+y) \mu(dx) \nu(dy)$$

Ou $f : \mathbb{R}^d \to \mathbb{R}^d$ mesurable.

Proposition 11.3.1 *Soit X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{R}^d$ de lois respectives μ et ν.*

i) $X+Y$ *a pour loi* $\mu * \nu$

ii) la fonction caractéristique Φ_{X+Y} *est notée par le produit* $\Phi_X \Phi_Y$

iii) Si $X, Y \in L^2(\mathbb{P})$ *alors* $Var(X,Y) = Var(X) + Var(Y)$

Remarque : *La somme de variables aléatoires indépendantes gaussiennes est une variable aléatoire gaussienne de moyenne la somme des moyennes, et de variance la somme des variances.*

12 Convergence de variables aléatoires

12.1 Notion de convergence forte

Définition 12.1.1 (Convergence presque-sûre) *On dit que $(X_n)_{n\in\mathbb{N}}$* ***converge presque surement*** *vers X, et on note $X_n \to X$ ps s'il existe un évènement λ de probabilité 1 tel que $\forall \omega \in \lambda$, $\lim_{n\to\infty} X_n(\omega) = X(\omega)$*

Définition 12.1.2 (Convergence L^p) *On dit, pour $p \geq 1$, que $(X_n)_{n\in\mathbb{N}}$* ***converge dans L^p vers X****, et on note $X_n \to X$ dans L^p , si $E(|X_n - X|) \underset{n\to\infty}{\to} 0$*

Définition 12.1.3 (Convergence en probabilités) *On dit, pour $p \geq 1$, que $(X_n)_{n\in\mathbb{N}}$* ***converge en probabilités vers X****, et on note $X_n \to X$ en proba , si $\forall \varepsilon > 0$ $\mathbb{P}(\{|X_n - X| > \varepsilon\}) \underset{n\to\infty}{\to} 0$*

Proposition 12.1.1 *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires.*

i) $\lim X_n = X$ *en proba* $\Leftrightarrow \lim_{n\to\infty} E(|X_n - X|) \wedge 1) = 0$

ii) $X_n \to X$ *presque surement ou dans* $L^p \Rightarrow X_n \to X$ *en proba*

iii) Si $X_n \to X$ en proba, alors on peut en extraire une sous suite qui converge presque surement.

iv) Si $X_n \to X$ en proba, et si $Y \in L^p$ tel que $|X_n| < Y$, alors $X_n \to X$ dans L^p

Proposition 12.1.2 *Notons $L^0(\Omega, \mathbb{P})$ l'espace des classes d'équivalence des variables aléatoires à valeurs dans $\mathbb{R}^d$, avec $X \sim X'$ si et ssi*

$$\mathbb{P}(X = X') = 1 .$$

On pose $dist(X,Y) = E(|X - Y| \wedge 1)$, distance sur L^0 qui métrise la convergence en probabilités. Par ailleurs, L^0 est complet pour cette distance.

12.2 Loi des grands nombres

Théorème 12.2.1 (Loi des grands nombres) *On considère $(X_n)_{n\in\mathbb{N}}$ des variables aléatoires à valeurs dans $\mathbb{R}^d$, iid. On suppose $X_1 \in L^1$, et on note $m = E(X_1)$. Alors :*

$$\text{(loi faible)} \quad \frac{X_1 + ... + X_n}{n} \to m \text{ dans } L^1$$

$$\text{(loi forte)} \quad \frac{X_1 + ... + X_n}{n} \to m \text{ presque surement}$$

12.3 Convergence en loi (convergence faible)

Définition 12.3.1 (Convergence étroite/faible) *Soient $\mu, \mu_1 ...$ des mesures de probabilités sur $\mathbb{R}^d$, on dit que $(\mu_n)_{n\in\mathbb{N}}$ converge étroitement , ou faiblement vers μ, et on notera $\mu_n \Rightarrow \mu$, (ou $\mu_n \overset{(e)}{\Rightarrow} \mu$) si*

$$\forall f \in \mathscr{C}_b(\mathbb{R}^d), \int f d\mu_n \to \int f d\mu .$$

Si $X, X_1 ... X_n ...$ sont des variables aléatoires à valeurs dans $\mathbb{R}^d$, on dit que X_n converge en loi vers X si la loi des X_n converge étroitement vers celle de X, ie. $E(f(X_n)) \to E(f(X)) \forall f \in \mathscr{C}_b(\mathbb{R}^d)$, noté $X_n \Rightarrow X$

Proposition 12.3.1 *Si $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X, alors $X_n \Rightarrow X$*

Théorème 12.3.1 (Du porte manteau) *Si μ, $(\mu_n)_{n\in\mathbb{N}}$ sont des mesures de probabilités sur $\mathbb{R}^d$, les propriétés suivantes sont équivalentes :*

i) $\mu_n \Rightarrow \mu$

ii) $\forall G \subset \mathbb{R}^d$ *ouvert,* $\liminf \mu_n(G) \geq \mu(G)$

iii) $\forall F \subset \mathbb{R}^d$ *fermé,* $\limsup \mu_n(F) \leq \mu(F)$

iv) $\forall B$ *borélien, si* $\partial B = \overline{B} - \overset{\circ}{B}$ *est tel que* $\mu(\partial B) = 0$, *alors* $\mu_n(B) \to \mu(B)$

Proposition 12.3.2 *Soit M un sous espace de $\mathscr{C}_b$ (muni de $\|.\|_\infty$) tel que $\mathscr{C}_c \subset \overline{M}$. Alors, pour $\mu, \mu_1...$ des mesures de probabilités, les propriétés suivantes sont équivalentes :*

i) $\mu_n \Rightarrow \mu$

ii) $\int f d\mu_n \to \int f d\mu \ \forall f \in \mathscr{C}_c$

iii) $\int f d\mu_n \to \int f d\mu \ \forall f \in M$

Théorème 12.3.2 *Si $\mu, \mu_1...$ sont des mesures de probabilités sur $\mathbb{R}^d$ alors $\mu_n \to \mu$ si et seulement si $\widehat{\mu}_n(\xi) \underset{n\to\infty}{\to} \widehat{\mu}(\xi) \ \forall \xi \in \mathbb{R}^d$*

Théorème 12.3.3 (Central limite) *Soit $X_1, X_2...$ une suite de variables aléatoires iid à valeurs dans $\mathbb{R}$, $X_1 \in {}^2(\mathbb{P})$. On pose $m = E(X_1)$, $\sigma^2 = var(X_1)$. Alors :*

$$\frac{X_1 + ... + X_n - nm}{\sqrt{n}} \Rightarrow N(0, \sigma^2)$$

Définition 12.3.2 (Vecteur gaussien) *Un vecteur aléatoire $\underline{X} = (X_1...X_d)$ $\Omega \to \mathbb{R}^d$ est dit gaussien si $\forall \underline{\lambda} = (\lambda_1...\lambda_d) \in \mathbb{R}^d$, $\underline{\lambda}.\underline{X}$ est une variable aléatoire gaussienne.*

Définition 12.3.3 (Matrice de dispersion) *Si $\underline{X} = (X_1...X_d)$ est un vecteur gaussien , alors $\underline{m} = (E(X_1...E(X_d))$ s'appelle le vecteur de la moyenne, et $D_{ij} = cov(X_i, X_j) = E(X_iX_j) - E(X_i)E(X_j)$. $[D_{ij}]_{1\leq i,j\leq d}$ s'appelle la matrice de dispersion/covariance.*

Corollaire 12.3.1 *Un vecteur gaussien a une loi déterminée par son vecteur moyenne et sa matrice de covariance. Plus précisément, sa fonction caractéristique est donnée par la formule suivante :*

$$\Phi_{\underline{X}}(\underline{\lambda}) = E(e^{i<\underline{\lambda},\underline{X}>}) = e^{i<\underline{\lambda},\underline{m}> - \frac{1}{2}<\underline{\lambda},D\underline{\lambda}>}$$

Lemme 12.3.1 *Soit une matrice symétrique positive D, soit C sa racine carrée, $\underline{m}$ un vecteur. Soient enfin $N_1...N_d \sim N(0,1)$, alors $\underline{X} = C.\underline{N} + \underline{m}$ est un vecteur gaussien de moyenne $\underline{m}$ et de covariance D*

Corollaire 12.3.2 (densité d'un vecteur gaussien $\underline{m} \in \mathbb{R}^d$) *Soit D une matrice $d \times d$ symétrique positive. alors un vecteur gaussien $\underline{X}$ de moyenne $\underline{m}$ et de matrice de dispersion D va admettre une densité par rapport à Lebesgue si et seulement si D est inversible. Dans ce cas, la densité est donnée par*

$$\mathbb{P}_{\underline{X}}(d\underline{x}) = (2\pi)^{-\frac{d}{2}}(det(D)^{-\frac{1}{2}}e^{-\frac{1}{2}<\underline{x}-\underline{m},D^{-1}(\underline{x}-\underline{m})>}\lambda_d(d\underline{x})$$

Corollaire 12.3.3 *Les coordonnées $X_1...X_d$ d'un vecteur gaussien sont indépendantes si et seulement si la matrice de dispersion D est diagonale.*

Théorème 12.3.4 (Central limite d-dimensionnel) *Soit $X_1, X_2...$ une suite de variables aléatoires iid à valeurs dans $\mathbb{R}^d$, $X_1 \in {}^2(\mathbb{P})$. On pose $\underline{m} = E(X_1)$, $D_{ij} = cov(X_1^{(i)}X_1^{(j)})$, où $X = (Y^{(1)}...Y^{(d)})$ (la matrice de covariance est toujours symétrique positive, il existe donc un vecteur gaussien $N(\underline{m}, D)$) . Alors :*

$$\frac{X_1 + ... + X_n - n\underline{m}}{\sqrt{n}} \Rightarrow N(0, D)$$