

MACS203 : Martingales

1 Théorie de la mesure

les théorèmes de convergence monotone, dominée, et le lemme de Fatou, - les inégalités de Markov, Chebysev, Cauchy-Schwarz, Hölder, Minkowsky, et de Jensen,

Espaces mesurables et mesures

Def. Soit $\mathcal{A} \subset \mathcal{P}(\Omega)$. On dit que

- (i) \mathcal{A} est une **algèbre** sur Ω si $\Omega \in \mathcal{A}$ et est stable par passage au complémentaire et par réunion,
- (ii) \mathcal{A} est une σ -algèbre si c'est une algèbre stable par union dénombrable. On dit alors que (Ω, \mathcal{A}) est un **espace mesurable**.

Def. Soit $\mathcal{I} \subset \mathcal{P}(\Omega)$. On dit que \mathcal{I} est un π -**système** s'il est stable par intersection finie.

Def. Soit \mathcal{A} une algèbre sur Ω et $\mu: \mathcal{A} \rightarrow \mathbf{R}_+$.

- (i) μ est dite **additive** si $\mu(\emptyset) = 0$ et $\forall A, B \in \mathcal{A}, A \cap B = \emptyset \implies \mu(A \cup B) = \mu(A) + \mu(B)$.
- (ii) μ est dite σ -**additive** si $\mu(\emptyset) = 0$ et $\forall (A_n)_{n \in \mathbf{N}} \subset \mathcal{A}$, si les A_n sont disjoints, $\mu(\cup_n A_n) = \sum_n \mu(A_n)$.
- (iii) Une fonction σ -additive μ sur un espace mesurable (Ω, \mathcal{A}) est appelée **mesure** et on dit que $(\Omega, \mathcal{A}, \mu)$ est un **espace mesuré**.
- (iv) Un espace mesuré $(\Omega, \mathcal{A}, \mu)$ est dit **fini** si $\mu(\Omega) < \infty$, et σ -**fini** s'il existe $(\Omega_n)_{n \in \mathbf{N}} \subset \mathcal{A}$ telle que $\mu(\Omega_n) < \infty$ et $\bigcup_{n \in \mathbf{N}} \Omega_n = \Omega$.

Prop. Soit \mathcal{I} un π -système, et μ, ν deux mesures finies sur $(\Omega, \sigma(\mathcal{I}))$. Si $\mu = \nu$ sur \mathcal{I} alors $\mu = \nu$ sur $\sigma(\mathcal{I})$.

Th (Extension de Carathéodory). Soit \mathcal{A}_0 une algèbre sur Ω et $\mu: \mathcal{A}_0 \rightarrow \mathbf{R}_+$ σ -additive. Alors il existe une mesure μ sur $\mathcal{A} := \sigma(\mathcal{A}_0)$ telle que $\mu = \mu_0$ sur \mathcal{A}_0 . Si de plus $\mu_0(\Omega) < \infty$ alors une telle extension est unique.

Def. (i) Sur un espace mesuré $(\Omega, \mathcal{A}, \mu)$, $N \in \mathcal{A}$ est dit **négligeable** si $\mu(N) = 0$.

- (ii) Soit $P(\omega)$ une propriété qui ne dépend que de $\omega \in \Omega$. On dit que P est vraie μ -presque partout si $\{\omega \in \Omega \mid P(\omega) \text{ n'est pas vraie}\}$ est inclus dans un ensemble négligeable.

Prop. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $(A_i)_{i \leq n} \subset \mathcal{A}$. Alors,

- (i) $\mu(\cup_{i \leq n} A_i) \leq \sum_{i=1}^n \mu(A_i)$,
- (ii) Si de plus $\mu(\Omega) < \infty$, on a $\mu(\cup_{i \leq n} A_i) = \sum_{k \leq n} (-1)^{k-1} \sum_{i_1 < \dots < i_k \leq n} \mu(A_{i_1} \cap \dots \cap A_{i_k})$.

Prop. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $(A_i)_n \subset \mathcal{A}$. Alors,

- (i) $A_n \uparrow A \implies \mu(A_n) \uparrow \mu(A)$,
- (ii) $(A_n \downarrow A \wedge (\exists k, \mu(A_k) < \infty)) \implies \mu(A_n) \downarrow \mu(A)$.

Lem (de Fatou pour les ensembles). Soit $(A_n)_n$ une suite dans \mathcal{A} . Alors, $\mu(\liminf A_n) \leq \liminf \mu(A_n)$.

Lem (inverse Fatou pour les ensembles). Supposons $(\Omega, \mathcal{A}, \mu)$ fini. Soit $(A_n)_n$ une suite dans \mathcal{A} . Alors, $\mu(\limsup A_n) \geq \limsup \mu(A_n)$.

Lem (de Borel-Cantelli). $\sum_n \mu(A_n) < \infty \implies \mu(\limsup A_n) = 0$.

L'intégrale de Lebesgue

Def. On dit qu'une fonction $f: (\Omega, \mathcal{A}) \rightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$ est **mesurable** si l'image réciproque de tout ensemble borélien est dans \mathcal{A} . On note $\mathcal{L}^0(\mathcal{A})$ l'ensemble des fonctions mesurables, $\mathcal{L}_+^0(\mathcal{A})$ si elles sont positives et $\mathcal{L}^\infty(\mathcal{A})$ si elles sont bornées.

Rem. Si $f: \Omega \rightarrow \mathbf{R}$ est continue avec Ω un espace topologique, alors f est $\mathcal{B}(\Omega)$ -mesurable et on dit qu'elle est **borélienne**.

Prop. (i) Pour $f, g \in \mathcal{L}^0(\mathcal{A})$, $h \in \mathcal{L}^0(\mathcal{B}(\mathbf{R}))$, $\lambda \in \mathbf{R}$, on a $f + g, \lambda f, fg, f \circ g \in \mathcal{L}^0(\mathcal{A})$.

- (ii) Pour une suite $(f_n)_n \subset \mathcal{L}_0(\mathcal{A})$, on a $\inf f_n, \liminf f_n, \sup f_n, \limsup f_n \in \mathcal{L}^0(\mathcal{A})$.

Th (des classes monotones). Soit \mathcal{H} une classe de fonctions réelles bornées sur Ω vérifiant

H1 \mathcal{H} est un espace vectoriel contenant la fonction constante 1,

H2 pour toute suite croissante $(f_n)_n \subset \mathcal{H}$ de fonctions positives dont la limite $f := \lim \uparrow f_n$ est bornée, on a $f \in \mathcal{H}$.

Soit \mathcal{I} un π -système tel que $\{1_A, A \in \mathcal{I}\} \subset \mathcal{H}$. Alors $\mathcal{L}^\infty(\sigma(\mathcal{I})) \subset \mathcal{H}$.

Not. L'intégrale $\int f d\mu$ sera aussi notée $\mu(f)$ par abus de notation.

Def. Pour $f \in \mathcal{L}_+^0(\mathcal{A})$, l'intégrale de f par rapport à μ est définie par $\mu(f) := \sup \{\mu(g) \mid g \in \mathcal{S}^+, g \leq f\}$ où \mathcal{S}^+ contient les fonctions de la forme $g = \sum_i a_i 1_{A_i}$, $a_i \in \mathbf{R}_+$ et $\mu(g) = \sum_i a_i \mu(A_i)$.

Lem. $\forall f_1, f_2 \in \mathcal{L}_0^+(\mathcal{A}, f_1 \leq f_2 \implies 0 \leq \mu(f_1) \leq \mu(f_2)$ et $\mu(f_1) = 0 \iff f_1 \stackrel{\mu-p.p.}{=} 0$.

Th (convergence monotone). Soit $(f_n)_n \subset \mathcal{L}_0^+(\mathcal{A})$ une suite croissante μ -p.p., i.e. $\forall n, f_n \stackrel{\mu-p.p.}{\leq} f_{n+1}$. Alors $\mu(\lim \uparrow f_n) = \lim \uparrow \mu(f_n)$.

Lem (Fatou). Soit $(f_n)_n \subset \mathcal{L}_0^+(\mathcal{A})$. On a $\mu(\liminf f_n) \leq \liminf \mu(f_n)$.

Def. $f \in \mathcal{L}_0(\mathcal{A})$ est dite μ -intégrable si $\mu(|f|) < \infty$ et son intégrale est définie par $\mu(f) := \mu(f^+) - \mu(f^-)$. On note $\mathcal{L}^1(\mathcal{A}, \mu)$ leur ensemble.

Th (convergence dominée). Soit $(f_n)_n \subset \mathcal{L}_0(\mathcal{A})$ une suite telle que $f_n \stackrel{\mu-p.p.}{\rightarrow} f \in \mathcal{L}_0(\mathcal{A})$. Si $\sup_n |f_n| \in \mathcal{L}^1(\mathcal{A}, \mu)$, alors $f_n \rightarrow f$ dans $\mathcal{L}^1(\mathcal{A}, \mu)$, i.e. $\mu(|f_n - f|) \rightarrow 0$. En particulier, $\mu(f_n) \rightarrow \mu(f)$.

Lem (Scheffé). Soit $(f_n)_n \subset \mathcal{L}^1(\mathcal{A}, \mu)$ telle que $f_n \stackrel{\mu-p.p.}{\rightarrow} f \in \mathcal{L}^1(\mathcal{A}, \mu)$. Alors $f_n \rightarrow f$ dans $\mathcal{L}^1(\mathcal{A}, \mu)$ ssi $\mu(|f_n|) \rightarrow \mu(|f|)$.

Transformées de mesures

Def. Soit $(\Omega_1, \mathcal{A}_1, \mu_1)$ un espace mesuré, $(\Omega_2, \mathcal{A}_2)$ un espace mesurable et $f: \Omega_1 \rightarrow \Omega_2$ une fonction mesurable. Alors $\mu_2 = \mu_1 \circ f^{-1}$, notée $\mu_1 f^{-1}$, définit une mesure appelée **mesure image**.

Th (transfert). Soit $\mu_2 = \mu_1 f^{-1}$ et $h \in \mathcal{L}^0(\mathcal{A}_2)$. Alors $h \in \mathcal{L}^1(\mathcal{A}_2, \mu_2) \iff h \circ f \in \mathcal{L}^1(\mathcal{A}_1, \mu_1)$. Dans ces conditions on a $\int_{\Omega_2} h d\mu_2 = \int_{\Omega_1} h \circ f d\mu_1$.

Def. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $f \in \mathcal{L}_0^+(\mathcal{A})$. On définit $\forall A \in \mathcal{A}, \nu(A) := \mu(f \mathbf{1}_A) = \int_A f d\mu$.

- (i) $\nu = f \cdot \mu$ est une mesure appelée mesure de **densité** f par rapport à μ .
- (ii) Soit μ_1, μ_2 deux mesures sur un espace mesurable (Ω, \mathcal{A}) . On dit que μ_2 est **absolument continue** par rapport à μ_1 , $\mu_2 \prec \mu_1$, si $\forall A \in \mathcal{A}, \mu_2(A) = 0 \implies \mu_1(A) = 0$. Sinon on dit que μ_2 est **étrangère** à μ_1 .
- (iii) Si $\mu_2 \prec \mu_1$ et $\mu_1 \prec \mu_2$, on dit que μ_1 et μ_2 sont **équivalentes**, $\mu_1 \sim \mu_2$. Si $\mu_2 \not\prec \mu_1$ et $\mu_1 \not\prec \mu_2$, on dit que μ_1 et μ_2 sont **singulières**.

Th. (i) Pour $g: \Omega \rightarrow \bar{\mathbf{R}}_+$ \mathcal{A} mesurable, on a $(f \cdot \mu)(g) = \mu(fg)$.

(ii) Pour $g \in \mathcal{L}_0^+(\mathcal{A})$, on a $g \in \mathcal{L}^1(\mathcal{A}, f \cdot \mu)$ ssi $fg \in \mathcal{L}^1(\mathcal{A}, \mu)$ et alors $(f \cdot \mu)(g) = \mu(fg)$.

Inégalités remarquables

Th. Soit f une fonction \mathcal{A} -mesurable, et $g: \mathbf{R} \rightarrow \mathbf{R}_+$ une fonction borélienne croissante positive.

- (i) $g \circ f$ est mesurable et $\forall c \in \mathbf{R}, \mu(g \circ f) \geq g(c)\mu(\{f \geq c\})$ (**Inégalité de Markov**),
- (ii) Si $f^2 \in \mathcal{L}^1(\mathcal{A}, \mu)$, $\forall c > 0, c^2 \mu(\{|f| \geq c\}) \leq \mu(f^2)$ (**inégalité de Tchebyshev**).

Espaces produits

Th (Fubini). L'application $\mu: A \mapsto \int (\int \mathbf{1}_A d\mu_1) d\mu_2 = \int (\int \mathbf{1}_A d\mu_2) d\mu_1$ sur $\mathcal{A}_1 \otimes \mathcal{A}_2$ est une mesure sur $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$, appelée **mesure produit** de μ_1 et μ_2 , et notée $\mu_1 \otimes \mu_2$. C'est l'unique mesure sur $\Omega_1 \times \Omega_2$ vérifiant $\forall (A_1, A_2) \in \mathcal{A}_1 \times \mathcal{A}_2, \mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$.

De plus, pour tout f dans $\mathcal{L}_0^+(\mathcal{A}_1 \times \mathcal{A}_2)$ ou $\mathcal{L}^1(\mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \otimes \mu_2)$, $\int f d\mu = \int (\int f d\mu_1) d\mu_2 = \int (\int f d\mu_2) d\mu_1 \in \bar{\mathbf{R}}_+$.

Soit $g: \Omega_1 \rightarrow \Omega_2$ avec Ω_1 et Ω_2 des ouverts de \mathbf{R}^n . Si g est différentiable en x , on note $Dg(x) := \left(\frac{\partial g_i}{\partial x_j} \right)_{1 \leq i, j \leq n}$ sa matrice jacobienne en x .

g est un \mathcal{C}^1 -difféomorphisme si g est une bijection telle que g et g^{-1} sont de classe \mathcal{C}^1 . Dans ce cas $\det[Dg^{-1}(y)] = \frac{1}{\det[Dg \circ g^{-1}(y)]}$.

Th. Soit μ_1 une mesure sur $(\Omega_1, \mathcal{B}(\Omega_1))$ de densité par rapport à la mesure de Lebesgue $f_1 \in \mathcal{L}_0^+(\mathcal{B}(\Omega_1))$, i.e. $\mu_1(dx) = \mathbf{1}_{\Omega_1} f_1(x) \cdot dx$. Soit g un \mathcal{C}^1 -difféomorphisme. La mesure image $\mu_2 = \mu_1 g^{-1}$ est absolument continue par rapport à la mesure de Lebesgue, de densité $f_2: y \mapsto \mathbf{1}_{\Omega_2}(y) f_1(g^{-1}(y)) |\det[Dg^{-1}(y)]|$ et pour toute fonction $h: \Omega_2 \rightarrow \mathbf{R}$ positive ou μ_2 -intégrable, $\int_{\Omega_1} h \circ g(x) f_1(x) dx = \int_{\Omega_2} h(y) f_2(y) dy$.

2 Préliminaires de la théorie des probabilités

Variables aléatoires

Def. Soit \mathbf{T} un ensemble et $\{X_\tau, \tau \in \mathbf{T}\}$ une famille quelconque de v.a. La σ -algèbre \mathcal{X} engendrée par cette famille est la plus petite σ -algèbre sur Ω telle que X_τ est \mathcal{X} -mesurable pour tout $\tau \in \mathbf{T}$, i.e.

$$\mathcal{X} = \sigma(X_\tau, \tau \in \mathbf{T}) = \sigma(\{X_\tau^{-1}(A) \mid \tau \in \mathbf{T}, A \in \mathcal{B}(\mathbf{R})\}).$$

Lem. Soit X et Y deux v.a. sur $(\Omega, \mathcal{A}, \mathbf{P})$ à valeurs respectivement dans \mathbf{R} et \mathbf{R}^n . Alors X est $\sigma(Y)$ -mesurable ssi $\exists f: \mathbf{R}^n \rightarrow \mathbf{R}, X = f(Y)$.

Espérance de variables aléatoires

Th. Soit $X \in \mathcal{L}^1(\mathcal{A}, \mathbf{P})$ et $g: \mathbf{R}^d \rightarrow \bar{\mathbf{R}}$ une fonction convexe telle que $\mathbf{E}(|g(X)|) < \infty$. Alors $\mathbf{E}(g(X)) \geq g(\mathbf{E}(X))$.

Def. Soit X une v.a. à valeurs dans \mathbf{R}^d . Sa fonction caractéristique est $\Phi_X: \begin{matrix} \mathbf{R}^d & \rightarrow & \mathbf{C} \\ u & \mapsto & \mathbf{E}[e^{i\langle u|X \rangle}] \end{matrix}$.

Lem. $\Phi_X(0) = 1$ et Φ_X est continue bornée (par 1) sur \mathbf{R}^d .

Prop. Soit $X \sim \mathcal{N}(b, V)$. On a $\Phi_X(u) = e^{i\langle u|b \rangle - \frac{1}{2}\langle u|Vu \rangle}$.

Prop. Soit X réelle avec $\mathbf{E}(|X|^p) < \infty$ pour un certain $p \in \mathbf{N}^*$. Alors Φ_X est p fois dérivable et $\forall k \in \llbracket 1; p \rrbracket$, $\Phi_X^{(k)}(0) = i^k \mathbf{E}(X^k)$.

Espaces \mathcal{L}^p et convergences fonctionnelles des v.a.

La corrélation entre deux v.a. X et Y est $\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\|X\|_2 \|Y\|_2}$.

Le théorème de Pythagore s'écrit

$$\mathbf{E}(XY) = 0 \implies \mathbf{E}[(X+Y)^2] = \mathbf{E}[X^2] + \mathbf{E}[Y^2] \quad \text{ou} \quad \text{Cov}(X, Y) = 0 \implies \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y).$$

et la loi du parallélogramme s'écrit $\|X+Y\|_2^2 + \|X-Y\|_2^2 = 2\|X\|_2^2 + 2\|Y\|_2^2$.

Def. Soit $(X_n)_n$ et X dans \mathcal{L}^0 . On dit que $(X_n)_n$ **converge en probabilité** vers X si $\forall \varepsilon > 0$, $\lim_{n \rightarrow \infty} \mathbf{P}[|X_n - X| \geq \varepsilon] = 0$.

Lem. La convergence p.s. ou la convergence en norme dans L^p impliquent la convergence en probabilité.

Lem. La convergence en probabilité est équivalente à la convergence au sens de la distance $D: (X, Y) \mapsto \mathbf{E}(|X - Y| \wedge 1)$

Th. (L^0, D) est un espace métrique complet.

Th. Soit $(X_n)_n$ et X des v.a. dans \mathcal{L}^0 .

(i) $X_n \rightarrow X$ p.s. ssi $\sup_{m \geq n} |X_m - X| \rightarrow 0$ en probabilité.

(ii) $X_n \rightarrow X$ en probabilité ssi de toute suite croissantes $(n_k)_k \subset \mathbf{N}$, on peut extraire une sous-suite $(n_{k_j})_j$ telle que $X_{n_{k_j}} \rightarrow X$ p.s.

Cor (Slutsky). Soit ϕ continue. Si $X_n \rightarrow X$ en probabilités, alors $\phi(X_n) \rightarrow \phi(X)$ en probabilité.

Def. Une famille C de v.a. est dite **uniformément intégrable** (U.I.) si $\lim_{c \rightarrow \infty} \sup_{X \in C} \mathbf{E}[|X| \mathbf{1}_{|X| \geq c}] = 0$.

Th. Soit $(X_n)_n$ et X des v.a. dans \mathcal{L}^1 . Alors $X_n \rightarrow X$ dans L^1 si et seulement si $X_n \rightarrow X$ en probabilité et $(X_n)_n$ est U.I.

Convergence en loi