Probabilités et Intégration

Licence 3, S5 Option Mathématiques Appliquées

et DUAS 1, S1

Table des matières

Chapitre 1

Espace probabilisé

1.1 Espace mesurable / probabilisable

Définition 1 *Soit E un ensemble. On note $\mathcal{P}(E)$ l'ensemble des parties de E. On dit que $\mathcal{E} \subset \mathcal{P}(E)$ est une tribu de parties de E si :*
i) L'ensemble E appartient à $\mathcal{E}$.
ii) Stabilité par passage au complémentaire : Si $A \in \mathcal{E}$ alors $A^C = E/A \in \mathcal{E}$.
iii) Stabilité par union dénombrable : Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{E}$ alors,

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{E}.$$

Le couple $(E, \mathcal{E})$ est appelé espace mesurable.

1.1.1 Tribu engendrée

Une méthode intéressante pour construire une tribu de parties de E est de la construire autour d'un ensemble $\mathcal{C} \subset \mathcal{P}(E)$ simple.

Définition 2 *Soit E un ensemble et soit $\mathcal{C} \subset \mathcal{P}(E)$. La tribu engendrée par $\mathcal{C}$ (notée $\sigma(\mathcal{C})$) est la plus petite tribu contenant $\mathcal{C}$.*

Tribu des boréliens

Pour définir cette tribu, rappelons qu'une topologie T sur un ensemble E est un ensemble d'ouverts.

Définition 3 *La tribu des boréliens de l'ensemble E (muni de la topologie T) est la tribu engendrée par les éléments (les ouverts) de T et on la note $\mathcal{B}(E)$.*

Dans le cas de l'espace $\mathbb{R}$ muni de sa topologie usuelle, les ouverts sont obtenus comme réunion dénombrable d'intervalles ouverts. Ainsi, $\mathcal{B}(\mathbb{R})$ est la tribu engendrée par les intervalles ouverts et plus généralement par les intervalles (ouverts ou non).

Définition 4 *La tribu des boréliens $\mathcal{B}(\mathbb{R})$ est la tribu engendrée par l'ensemble des intervalles $\{]a,b],\ -\infty \leq a < b \in \mathbb{R}\}$.*

1.1.2 Produit d'espaces mesurables

Définition 5 *Considérons n espaces mesurables $(E_1, \mathcal{E}_1), \ldots, (E_n, \mathcal{E}_n)$. L'espace produit de ces n espaces est l'espace $(E, \mathcal{E})$ avec :*
i) $E = E_1 \times \ldots \times E_n = \{(x_1, \ldots, x_n);\ x_1 \in E_1, \ldots, x_n \in E_n\}$,
ii) $\mathcal{E} = \mathcal{E}_1 \otimes \ldots \otimes \mathcal{E}_n = \sigma(\{A_1 \times \ldots \times A_n,\ A_i \in \mathcal{E}_i,\ i = 1, \ldots, n\})$. L'ensemble $\{A_1 \times \ldots \times A_n,\ A_i \in \mathcal{E}_i,\ i = 1, \ldots, n\}$ est appelé l'ensemble des pavés mesurables de $\mathcal{E}$.

1.1.3 Image d'une tribu par une application quelconque

Soit $f : E_1 \to E_2$ une application quelconque et soit $A \subset E_2$. On note $f^{-1}(A) := \{x \in E_1,\ f(x) \in A\}$ l'image réciproque par f de l'ensemble A.

Proposition 1 *Soit $f : E_1 \to E_2$ une application. Si $\mathcal{E}_2$ est une tribu de parties de E_2 alors $f^{-1}(\mathcal{E}_2) = \{f^{-1}(A),\ A \in \mathcal{E}_2\}$ est une tribu de parties de E_1. De plus, si $\mathcal{E}_1$ est une tribu de parties de E_1 alors l'ensemble $\{B \in \mathcal{P}(E_2) | f^{-1}(B) \in \mathcal{E}_1\}$ est une tribu de E_2.*

1.1.4 Espace probabilisable

Définition 6 *On appelle espace probabilisable le couple $(\Omega, \mathcal{F})$ où Ω est un ensemble quelconque et $\mathcal{F}$ une tribu des parties de Ω. L'ensemble Ω est appelé l'ensemble des possibles. Cet ensemble regroupe les résultats possibles d'une expérience aléatoire. Un élément A de $\mathcal{F}$ est un événement.*

Vocabulaire – Les éléments de Ω seront notés ω. Ainsi, ω est une réalisation possible de l'expérience aléatoire considérée. Notons A et B deux événements (*i.e.* $A \in \mathcal{F}$ et $B \in \mathcal{F}$).

- Si $\omega \in A$, on dit que A est réalisé.
- Si $\omega \in A \cup B$, on dit que A ou B sont réalisés.
- Si $\omega \in A \cap B$, on dit que A et B sont réalisés.

1.2 Espace mesuré / probabilisé

1.2.1 Mesure

Définition 7 *Soit $(E, \mathcal{E})$ un espace mesurable. Une mesure sur cet espace est une application $\mu : \mathcal{E} \to [0, \infty]$ vérifiant $\mu(\emptyset) = 0$ et σ-additive c'est-à-dire telle que pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments disjoints deux à deux de $\mathcal{E}$,*

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n).$$

Le triplet $(E, \mathcal{E}, \mu)$ *est appelé espace mesuré.*

Proposition 2 *On considère l'espace mesuré* $(E, \mathcal{E}, \mu)$*. Soient* A *et* B *deux éléments de* $\mathcal{E}$*.*
1) Si $\mu(E) < +\infty$, $\mu(A^C) = \mu(E) - \mu(A)$*.*
2) Si $A \subset B$ *alors* $\mu(A) \leq \mu(B)$*.*
3) Soit $(C_i)_{i \in I}$ *une collection dénombrable d'éléments de* $\mathcal{E}$ *disjoints deux à deux avec*

$$\bigcup_{i \in I} C_i = E,$$

alors,

$$\mu(A) = \sum_{i \in I} \mu(A \cap C_i).$$

4) Soit $(C_i)_{i=1,\ldots,p}$ *une collection finie d'éléments de* $\mathcal{E}$*. Si* $\mu(E) < +\infty$,

$$\mu\left(\bigcup_{i=1}^{p} C_i\right) = \sum_{i=1}^{p} \mu(C_i) + \sum_{j=2}^{p} (-1)^{j+1} \sum_{i_1 < \ldots < i_j} \mu(C_{i_1} \cap \ldots \cap C_{i_j}).$$

Cette formule est connue sous le nom de formule de Poincaré. En particulier, on a donc $\mu(C_1 \cup C_2) = \mu(C_1) + \mu(C_2) - \mu(C_1 \cap C_2)$*.*
5) Soit $(C_i)_{i=1,\ldots,p}$ *une collection finie d'éléments de* $\mathcal{E}$*.*

$$\mu\left(\bigcup_{i=1}^{p} C_i\right) \leq \sum_{i=1}^{p} \mu(C_i).$$

6) Si $(A_n)_{n \in \mathbb{N}}$ *est une suite croissante d'éléments de* $\mathcal{E}$ *(i.e pour tout* $n \in \mathbb{N}$, $A_n \subset A_{n+1}$*) alors*

$$\lim_{n \to \infty} \mu(A_n) = \mu\left(\bigcup_{n \in \mathbb{N}} A_n\right).$$

7) Si $(A_n)_{n \in \mathbb{N}}$ *est une suite décroissante d'événements (i.e pour tout* $n \in \mathbb{N}$, $A_n \subset A_{n-1}$*) et si* $\mu(A_0) < \infty$ *alors*

$$\lim_{n \to \infty} \mu(A_n) = \mu\left(\bigcap_{n \in \mathbb{N}} A_n\right).$$

Mesure de Lebesgue

Notons $\mathcal{C}$ l'ensemble formé des intervalles $]a, b]$, $-\infty \leq a \leq b < \infty$. Cet ensemble engendre la tribu borélienne $\mathcal{B}(\mathbb{R})$. On admettra que la fonction d'ensemble λ définie sur $\mathcal{C}$ par

$$\lambda\left(]a, b]\right) = b - a,$$

définit (par extension) une unique mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ appelée mesure de Lebesgue.

Sur l'espace produit $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ on définit la mesure de Lebesgue $\lambda^d = \lambda \otimes \ldots \otimes \lambda$ qui est entièrement définie par ses valeurs sur les pavés

$$\prod_{i=1}^{d}]a_i, b_i],$$

avec $a_i \in [-\infty, \infty[$ et $b_i \in \mathbb{R}$ pour tout $i = 1, \ldots, d$. La mesure λ^d est donnée par

$$\lambda^d \left(\prod_{i=1}^{d}]a_i, b_i] \right) = \prod_{i=1}^{d} \lambda(]a_i, b_i]) = \prod_{i=1}^{d} (b_i - a_i).$$

La mesure λ^d est également appelée mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^d)$. On notera parfois simplement λ au lieu de λ^d.

Mesure de probabilité

Définition 8 *Une mesure de probabilité $\mathbb{P}$ définie sur l'espace probabilisable $(\Omega, \mathcal{F})$ est une mesure vérifiant la propriété supplémentaire $\mathbb{P}(\Omega) = 1$. Le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ est appelé espace probabilisé.*

Définition 9 *Si l'ensemble Ω est fini, la probabilité uniforme $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ est définie pour tout $\omega \in \Omega$ par $\mathbb{P}(\{\omega\}) = 1/\mathrm{Card}(\Omega)$.*

Probabilité conditionnelle à un évènement

A partir d'une mesure de probabilité $\mathbb{P}$ définie sur l'espace $(\Omega, \mathcal{F})$ et d'un événement $B \in \mathcal{F}$ avec $\mathbb{P}(B) \neq 0$, on peut définir une autre mesure de probabilité appelée probabilité conditionnelle sachant l'événement B.

Proposition 3 *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et soit $B \in \mathcal{F}$ avec $\mathbb{P}(B) \neq 0$. La fonction d'ensemble $\mathbb{P}_B$ définie pour tout $A \in \mathcal{F}$ par $\mathbb{P}_B(A) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$ est une mesure de probabilité sur l'espace $(\Omega, \mathcal{F})$.*

On trouve souvent la notation $\mathbb{P}_B(A) = \mathbb{P}(A|B)$ qui se lit "probabilité de A sachant B".

1.2.2 Indépendance d'événements

Définition 10 *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et soient A et B deux événements. On dit que A et B sont indépendants (relativement à la mesure $\mathbb{P}$) si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

Définition 11 *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Des événements $A_1, \ldots, A_n$ sont indépendants si pour tout $k \in \{1, \ldots, n\}$ et tout ensemble d'indices $\{i_1 < i_2 \ldots < i_k\} \subset \{1, \ldots, n\}$,*

$$\mathbb{P}\left(\bigcap_{j=1}^{k} A_{i_j} \right) = \prod_{j=1}^{k} \mathbb{P}(A_{i_j}).$$

De plus, une suite $(A_n)_{n\geq 1}$ *d'événements est indépendante si pour tout* $n \in \mathbb{N}$, $A_1, \ldots, A_n$ *est un ensemble d'événements indépendants.*

Chapitre 2

Fonctions mesurables - Variables aléatoires

2.1 Définitions et résultats

Définition 12 *Soient* $(E_1, \mathcal{E}_1)$ *et* $(E_2, \mathcal{E}_2)$ *deux espaces mesurables. Une fonction* $f : E_1 \to E_2$ *est dite mesurable si pour tout* $A \in \mathcal{E}_2$, $f^{-1}(A) := \{x \in E_1 \mid f(x) \in A\} \in \mathcal{E}_1$.

Théorème 1 *Soient* $(E_1, \mathcal{E}_1)$, $(E_2, \mathcal{E}_2)$ *et* $(E_3, \mathcal{E}_3)$ *des espaces mesurables.*
1) Si $\mathcal{E}_2 = \sigma(\mathcal{C})$, $\mathcal{C} \subset \mathcal{P}(E_2)$. *La fonction* $f : E_1 \to E_2$ *est mesurable si et seulement si pour tout* $C \in \mathcal{C}$, $f^{-1}(C) \in \mathcal{E}_1$.
2) Soient $f : E_1 \to E_2$ *et* $g : E_2 \to E_3$ *deux fonctions mesurables. La fonction* $g \circ f$ *est également mesurable.*

Proposition 4 *Soit* $(f_n)_{n\geq 1}$ *une suite de fonctions mesurables de* $(E, \mathcal{E})$ *dans* $(\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$ *(où on rappelle que* $\bar{\mathbb{R}} = [-\infty, +\infty]$*). Alors les fonctions suivantes*

$$\inf_{n\geq 1} f_n, \ \sup_{n\geq 1} f_n, \ \liminf_{n\to\infty} f_n \ \text{et} \ \limsup_{n\to\infty} f_n,$$

sont mesurables. De plus, si f_n *admet une limite simple* f, *alors* f *est mesurable.*

Définition 13 *Soient* (E_1, T_1) *et* (E_2, T_2) *deux espaces topologiques. Une fonction* $f : (E_1, \mathcal{B}(E_1)) \to (E_2, \mathcal{B}(E_2))$ *mesurable est appelée fonction borélienne.*

On a le résultat important suivant.

Proposition 5 *Soient* (E_1, T_1) *et* (E_2, T_2) *deux espaces topologiques. Si une fonction* $f : E_1 \to E_2$ *est continue sur* E_1 *alors elle est borélienne.*

Dans le langage probabiliste, une application mesurable est appelée une variable aléatoire.

Définition 14 *Soient un espace probabilisable* $(\Omega, \mathcal{F})$ *et un espace mesurable* $(E, \mathcal{E})$. *Une variable aléatoire* X *est une fonction mesurable* $X : \Omega \to E$.

2.2 Mesure image et loi d'une variable aléatoire

Définition 15 *Soient* $(E_1, \mathcal{E}_1)$ *et* $(E_2, \mathcal{E}_2)$ *deux espaces mesurables. Soit* $f : (E_1, \mathcal{E}_1) \to (E_2, \mathcal{E}_2)$ *une fonction mesurable. Si l'espace* $(E_1, \mathcal{E}_1)$ *est muni de la mesure* μ *alors la mesure image de* μ *par* f *est la mesure* ν *définie pour tout* $A \in \mathcal{E}_2$ *par* $\nu(A) := \mu(f^{-1}(A))$.

Définition 16 *Soit* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$ *une variable aléatoire. La loi de* X *est la mesure image* $\mathbb{P}_X$ *de* $\mathbb{P}$ *par* X *qui est définie pour tout* $A \in \mathcal{E}$ *par :*

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

1) Si $(E, \mathcal{E}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$, la loi de X est entièrement caractérisée par les valeurs $p_k := \mathbb{P}_X(\{k\})$ pour $k \in \mathbb{N}$.

2) Si $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, la loi de X est entièrement caractérisée par

$$\{F(x) := \mathbb{P}_X(]-\infty, x]) = \mathbb{P}(X \leq x),\ x \in \mathbb{R}\}.$$

Définition 17 *Soit* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *une variable aléatoire. La fonction* $F : \mathbb{R} \to [0, 1]$ *définie pour tout* $t \in \mathbb{R}$ *par*

$$F(t) = \mathbb{P}_X(]-\infty, t]) = \mathbb{P}(X \leq t)$$

est appelée la fonction de répartition de X.

2.3 Intégration d'une fonction mesurable

Soit $(E, \mathcal{E}, \mu)$ un espace mesuré et soit $f : (E, \mathcal{E}, \mu) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une fonction réelle mesurable. L'objectif de ce paragraphe est de donner un sens à l'expression

$$\int_E f d\mu,$$

que l'on appellera intégrale de Lebesgue de f par rapport à la mesure μ. Nous allons voir étape par étape la construction de cette intégrale.

2.3.1 Intégration d'une fonction étagée positive

Définition 18 *Une fonction* $g : (E, \mathcal{E}, \mu) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ *est dite étagée positive si elle s'écrit sous la forme*

$$g = \sum_{i=1}^{k} \alpha_i \mathbb{I}_{A_i},$$

où les α_i, $i = 1, \ldots, k$ *sont des réels positifs ou nuls et tous distincts et* $A_1, \ldots, A_k$ *sont des ensembles appartenant à* $\mathcal{E}$ *formant une partition de* E. *L'ensemble des fonctions étagées positives est noté* $\mathcal{H}^+$.

Définition 19 *L'intégrale par rapport à la mesure μ d'une fonction g étagée positive est la quantité :*

$$\int_E g d\mu = \left(\int_E g(x) d\mu(x) = \right) \sum_{i=1}^{k} \alpha_i \mu(A_i). \tag{2.1}$$

2.3.2 Intégration d'une fonction mesurable positive

Définition 20 *L'intégrale par rapport à la mesure μ d'une fonction f mesurable positive est la quantité :*

$$\int_E f d\mu = \sup \left\{ \int_E g d\mu,\ g \in \mathcal{H}^+,\ g \leq f \right\} \in [0, \infty].$$

Théorème 2 (Théorème de convergence monotone) *Soit $(f_n)_{n\in\mathbb{N}}$ une suite croissante de fonctions mesurables positives définies sur l'espace mesuré $(E, \mathcal{E}, \mu)$. Si f_n converge vers une fonction f μ-presque partout, alors f est une fonction mesurable positive et*

$$\lim_{n\to\infty} \int_E f_n d\mu = \int_E f d\mu.$$

Une conséquence importante de ce résultat est que si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions mesurables positives définies sur l'espace mesuré $(E, \mathcal{E}, \mu)$ alors,

$$\sum_{n\in\mathbb{N}} \int_E f_n d\mu = \int_E \sum_{n\in\mathbb{N}} f_n d\mu.$$

Théorème 3 (Lemme de Fatou) *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables positives définies sur l'espace mesuré $(E, \mathcal{E}, \mu)$. On a*

$$\int_E \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_E f_n d\mu.$$

Proposition 6 *Toute fonction f mesurable positive est limite d'une suite croissante $(g_n)_{n\geq 1}$ de fonctions étagées positives.*

Proposition 7 *Soit f une fonction mesurable positive.*

$$\int_E f d\mu = 0 \Leftrightarrow f = 0 \ \mu\text{-pp et } \int_E f d\mu < \infty \Rightarrow f < \infty \ \mu\text{-pp}.$$

2.3.3 Intégration d'une fonction mesurable

Définition 21 *Si $f : (E, \mathcal{E}, \mu) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une fonction mesurable réelle, on dit que f est intégrable par rapport à la mesure μ si*

$$\int_E |f| d\mu < \infty.$$

Dans ce cas, en posant $f_+ = \max(f, 0)$ et $f_- = -\min(f, 0)$, on définit l'intégrale de f par rapport à μ par :

$$\int_E f d\mu = \int_E f_+ d\mu - \int_E f_- d\mu.$$

Comparaison Riemann / Lebesgue

On se place dans le cas où $(E, \mathcal{E}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ avec λ la mesure de Lebesgue.

Proposition 8 *Si l'intégrale de Riemann d'une fonction f mesurable et bornée sur $[a, b]$ existe alors elle coïncide avec l'intégrale de Lebesgue et on a*

$$\int_a^b f(x)dx = \int_{[a,b]} f d\lambda.$$

Proposition 9 *Soient $a, b \in \bar{\mathbb{R}}$ avec $a < b$. Si $f :]a, b[\to \mathbb{R}$ est une fonction continue (et donc mesurable) sur $]a, b[$ et d'intégrale de Riemann généralisée absolument convergente c'est-à-dire telle que*

$$\int_a^b |f(x)|dx < \infty,$$

alors f est Lebesgue-intégrable sur $]a, b[$ et les deux intégrales coïncident :

$$\int_a^b f(x)dx = \int_{]a,b[} f d\lambda = \int_{[a,b]} f d\lambda.$$

Cas de la mesure de Dirac

Soit $(E, \mathcal{E})$ un espace mesurable et soit $a \in E$. La mesure de Dirac δ_a au point a est définie pour tout $A \in \mathcal{E}$ par $\delta_a(A) = 1$ si $a \in A$ et 0 sinon. Si $f : (E, \mathcal{E}, \delta_a) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une fonction mesurable et intégrable par rapport à la mesure δ_a, alors

$$\int_E f d\delta_a = f(a).$$

2.3.4 Quelques théorèmes importants de la théorie de l'intégration

Théorème 4 (admis) [Théorème de convergence dominée] *Soit $(E, \mathcal{E}, \mu)$ un espace mesuré et soit $(f_n)_{n\geq 1}$ une suite de fonctions réelles (i.e. à valeurs dans l'espace $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$) mesurables. Si f_n converge μ-presque-partout vers une fonction mesurable f et s'il existe une fonction mesurable h intégrable par rapport à μ telle que pour tout $n \geq 1$, $|f_n| \leq h$ μ-presque-partout alors f est intégrable par rapport à μ et*

$$\lim_{n\to\infty} \int_E f_n d\mu = \int_E f d\mu.$$

Théorème 5 (admis) [Fubini] *Soient $(E_1, \mathcal{E}_1, \mu_1)$, $(E_2, \mathcal{E}_2, \mu_2)$ deux espaces mesurés où μ_1 et μ_2 sont des mesures σ-finies.*
1) Il existe une unique mesure ν sur l'espace mesuré produit $(E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$

telle que pour tout $A_1 \in \mathcal{E}_1$ et $A_2 \in \mathcal{E}_2$, $\nu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$. On note cette mesure $\nu = \mu_1 \otimes \mu_2$.
2) De plus, si la fonction mesurable $f : (E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est soit positive soit intégrable par rapport à ν alors

$$\begin{aligned}\int_{E_1 \times E_2} f d\nu &= \int_{E_1} \int_{E_2} f(x_1, x_2) d\mu_2(x_2) d\mu_1(x_1) \\ &= \int_{E_2} \int_{E_1} f(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2).\end{aligned}$$

On rappelle que la matrice Jacobienne d'une fonction $\psi : \mathbb{R}^d \to \mathbb{R}^d$ différentiable définie pour $x = (x_1, \ldots, x_d)$ par

$$\psi(x_1, \ldots, x_d) = (\psi_1(x_1, \ldots, x_d), \ldots, \psi_d(x_1, \ldots, x_d)),$$

est la matrice de dimension $d \times d$

$$J_\psi(x) = \begin{pmatrix} \frac{\partial \psi_1}{\partial x_1}(x) & \ldots & \frac{\partial \psi_1}{\partial x_d}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial \psi_d}{\partial x_1}(x) & \ldots & \frac{\partial \psi_d}{\partial x_d}(x) \end{pmatrix}$$

Soit $k \in \mathbb{N}^*$, on rappelle que ψ est un $\mathcal{C}^k$-difféomorphisme si ψ est bijective, k-fois dérivable de dérivée k-ième continue (*i.e.* de classe $\mathcal{C}^k$) et de bijection réciproque $\psi^{-1} \in \mathcal{C}^k$. Il est évident qu'un $\mathcal{C}^k$-difféomorphisme est une application mesurable de $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Théorème 6 (admis) [Changement de variable] *Soient U_1 et U_2 des ouverts de l'espace $\mathbb{R}^d$ que l'on munit de la tribu borélienne $\mathcal{B}(\mathbb{R}^d)$ et de la mesure de Lebesgue λ. Soit $\varphi : U_1 \to U_2$ un $\mathcal{C}^1$-difféomorphisme et soit $f : U_1 \to \mathbb{R}$ une fonction mesurable positive ou intégrable. Alors,*

$$\int_{U_1} f(x) dx = \int_{U_2} f(\varphi^{-1}(y)) \det(J_{\varphi^{-1}}(y)) dy,$$

où $J_{\varphi^{-1}}(y)$ est la matrice Jacobienne de φ^{-1} au point y.

2.3.5 Espaces L^p

Définition 22 *Soit un espace mesuré $(E, \mathcal{E}, \mu)$.*
i) Soit $p \in \mathbb{N}^*$. On note $\mathcal{L}^p(E, \mathcal{E}, \mu)$ l'ensemble des fonctions mesurables $f : (E, \mathcal{E}, \mu) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telles que

$$\int_E |f|^p d\mu < \infty.$$

ii) On note $L^p(E, \mathcal{E}, \mu)$ l'ensemble des classes d'équivalence pour l'égalité μ presque-partout des fonctions de $\mathcal{L}^p(E, \mathcal{E}, \mu)$.

On s'intéresse plus particulièrement à l'espace $L^2(E, \mathcal{E}, \mu)$.

Proposition 10 *L'ensemble $L^2(E, \mathcal{E}, \mu)$ est un espace vectoriel sur $\mathbb{R}$. De plus l'application $< \cdot, \cdot >$ définie par :*

$$\begin{array}{rcl} L^2(E, \mathcal{E}, \mu) \times L^2(E, \mathcal{E}, \mu) & \to & \mathbb{R} \\ (f, g) & \to & < f, g >:= \displaystyle\int_E f g d\mu, \end{array}$$

est une forme hermitienne définie positive c'est-à-dire un produit scalaire. L'espace $L^2(E, \mathcal{E}, \mu)$ est donc préhilbertien.

Proposition 11 *Soient f et g deux éléments de $L^2(E, \mathcal{E}, \mu)$. L'inégalité de Cauchy-Schwartz est donnée par*

$$\left| \int_E f g d\mu \right| \leq \sqrt{\int_E f^2 d\mu \int_E g^2 d\mu}.$$

Définition 23 *Soient $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions de $L^2(E, \mathcal{E}, \mu)$ et $f \in L^2(E, \mathcal{E}, \mu)$. On dit que f_n converge vers f dans $L^2(E, \mathcal{E}, \mu)$ si*

$$\lim_{n \to \infty} \|f_n - f\|_2 = 0.$$

On dira que f est la limite au sens de L^2 de la suite $(f_n)_{n \in \mathbb{N}}$.

Définition 24 *Soient f et g deux fonctions de $L^2(E, \mathcal{E}, \mu)$. On dit que f et g sont orthogonales si $< f, g >= 0$.*

Théorème 7 *Soit $(E, \mathcal{E}, \mu)$ un espace mesuré. L'espace $L^2(E, \mathcal{E}, \mu)$ est un espace de Hilbert.*

Théorème 8 (Théorème de la projection orthogonale) *Soit E un espace de Hilbert et F un sous espace fermé. Pour tout point $x \in E$, il existe un unique point $p_F(x) \in F$ tel que*

$$\|x - p_F(x)\|_2 = \inf\{\|x - y\|_2, \ y \in F\}.$$

En particulier, si F est un sous espace vectoriel fermé, le point $p_F(x)$ ainsi défini est l'unique point de F tel que

$$x - p_F(x) \in F^{\perp} := \{x \in E | < x, y >= 0 \ \forall \ y \in F\}.$$

2.3.6 Espérance d'une variable aléatoire

Dans le langage probabiliste, l'intégrale d'une variable aléatoire est appelée l'espérance.

Définition 25 *Soit* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *une variable aléatoire. On appelle espérance de* X *l'intégrale (si elle existe) de* X *par rapport à la probabilité* $\mathbb{P}$*. Ainsi si* X *est intégrable i.e. si*

$$\int_\Omega |X| d\mathbb{P} < \infty,$$

l'espérance de X *est donnée par :*

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

Définition 26 *Soit* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *une variable aléatoire et soit* $p \geq 1$*. On dira que* X *admet un moment d'ordre* p *si* $\mathbb{E}(|X|^p) < \infty$*. Le moment d'ordre* p *est alors donné par* $\mathbb{E}(X^p)$*.*

Théorème 9 (Théorème de Transfert) *Soit* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$ *une variable aléatoire de loi* $\mathbb{P}_X$ *et soit* $g : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *une fonction mesurable. Alors* $g(X)$ *est une variable aléatoire et si* $g(X)$ *est intégrable,*

$$\mathbb{E}(g(X)) = \int_\Omega g(X) d\mathbb{P} = \int_E g d\mathbb{P}_X \left(= \int_E g(x) d\mathbb{P}_X(x) \right). \tag{2.2}$$

Espérance conditionnelle à un événement

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Pour tout événenent $B \in \mathcal{F}$ avec $\mathbb{P}(B) \neq 0$, on définit l'espérance de X conditionnellement à l'événement B par

$$\mathbb{E}(X \mid B) := \int_\Omega X d\mathbb{P}_B,$$

avec, pour tout $A \in \mathcal{F}$, $\mathbb{P}_B(A) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$. On a le résultat suivant.

Proposition 12 *Pour tout événenent* $B \in \mathcal{F}$ *avec* $\mathbb{P}(B) \neq 0$*,*

$$\mathbb{E}(X \mid B) := \frac{1}{\mathbb{P}(B)} \int_B X d\mathbb{P} = \frac{\mathbb{E}(X \mathbb{I}_B)}{\mathbb{P}(B)}.$$

2.3.7 Quelques inégalités importantes

Proposition 13 (Inégalité de Markov) *Soit* $k > 0$ *et* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *une variable aléatoire telle que* $\mathbb{E}(|X|^k) < \infty$*. Pour tout* $\alpha > 0$*,*

$$\mathbb{P}(|X| \geq \alpha) \leq \frac{\mathbb{E}(|X|^k)}{\alpha^k}.$$

Proposition 14 (Inégalité de Bienaymé-Tchebychev) *Soit* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *une variable aléatoire admettant un moment d'ordre 2. On note* μ *son espérance et* σ^2 *sa variance. Alors, pour tout* $\epsilon > 0$*,*

$$\mathbb{P}(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

Proposition 15 (Inégalité de Jensen) *Soit $I \subset \mathbb{R}$, φ une fonction convexe sur I et $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (I, \mathcal{B}(I))$ une variable aléatoire intégrable et telle que $\mathbb{E}(|\varphi(X)|) < \infty$. Alors, $\varphi(\mathbb{E}(X)) \le \mathbb{E}(\varphi(X))$.*

Proposition 16 (Inégalité de Hölder) *Soient $p > 1$, $q > 1$ tels que $1/p + 1/q = 1$, $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ et $Y \in L^q(\Omega, \mathcal{F}, \mathbb{P})$ des variables aléatoires à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R})$. On a $\mathbb{E}(|XY|) \le \mathbb{E}^{1/p}(|X|^p)\mathbb{E}^{1/q}(|Y|^q)$.*

Proposition 17 (Inégalité de Lyapounov) *Soient $0 < \alpha \le \beta$ et $X \in L^\beta(\Omega, \mathcal{F}, \mathbb{P})$ une variable aléatoire à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R})$. On a*

$$\mathbb{E}^{1/\alpha}(|X|^\alpha) \le \mathbb{E}^{1/\beta}(|X|^\beta).$$

2.4 Indépendance de variables aléatoires

Définition 27 *Soient $X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (E_i, \mathcal{E}_i)$, $i = 1, \ldots, n$ des variables aléatoires. On dit qu'elles sont indépendantes si pour tout $A_1 \in \mathcal{E}_1, \ldots, A_n \in \mathcal{E}_n$,*

$$\mathbb{P}\left(\bigcap_{i=1}^n X_i^{-1}(A_i)\right) = \prod_{i=1}^n \mathbb{P}(X_i^{-1}(A_i)).$$

Proposition 18 *Soient $X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (E_i, \mathcal{E}_i)$, $i = 1, \ldots, n$ des variables aléatoires. Si pour tout $i = 1, \ldots, n$, on a $\mathcal{E}_i = \sigma(\mathcal{C}_i)$ où $\mathcal{C}_i \subset \mathcal{P}(E_i)$ est stable par intersections finies (i.e., si $A_1 \in \mathcal{C}_i$ et $A_2 \in \mathcal{C}_i$ alors $A_1 \cap A_2 \in \mathcal{C}_i$), les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes si et seulement si pour tout $A_1 \in \mathcal{C}_1, \ldots, A_n \in \mathcal{C}_n$,*

$$\mathbb{P}\left(\bigcap_{i=1}^n X_i^{-1}(A_i)\right) = \prod_{i=1}^n \mathbb{P}(X_i^{-1}(A_i)).$$

Corollaire 1 *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.*
i) Les variables aléatoire $X_1, \ldots, X_n$ définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ sont indépendantes si et seulement si pour tout $(t_1, \ldots, t_n) \in \mathbb{R}^n$,

$$\mathbb{P}(\{X_1 \le t_1\} \cap \ldots \cap \{X_n \le t_n\}) = \mathbb{P}(X_1 \le t_1) \times \ldots \times \mathbb{P}(X_n \le t_n).$$

ii) Soit $E_1, \ldots, E_n$ des ensembles au plus dénombrables. Les variables aléatoires discrètes $\{X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (E_i, \mathcal{P}(E_i)),\ i = 1, \ldots, n\}$ sont indépendantes si et seulement si pour tout $(t_1, \ldots, t_n) \in E_1 \times \ldots \times E_n$, $\mathbb{P}\{(X_1, \ldots, X_n) = (t_1, \ldots, t_n)\} = \mathbb{P}(X_1 = t_1) \ldots \mathbb{P}(X_n = t_n)$.

Proposition 19 *Soient $X_1 : (\Omega, \mathcal{F}, \mathbb{P}) \to (E_1, \mathcal{E}_1)$ et $X_2 : (\Omega, \mathcal{F}, \mathbb{P}) \to (E_2, \mathcal{E}_2)$ deux variables aléatoires et soient $f_1 : (E_1, \mathcal{E}_1) \to (E_3, \mathcal{E}_3)$ et $f_2 : (E_2, \mathcal{E}_2) \to (E_4, \mathcal{E}_4)$ deux fonctions mesurables. Si X_1 et X_2 sont indépendantes alors $f_1(X_1)$ et $f_2(X_2)$ sont indépendantes.*

Théorème 10 *Soient X_1 et X_2 deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Si X_1 et X_2 sont indépendantes alors $\mathbb{E}(X_1X_2) = \mathbb{E}(X_1)\mathbb{E}(X_2)$.*

2.5 Vecteurs aléatoires

Définition 28 *Un vecteur aléatoire $X = (X_1, \ldots, X_n)$ est une application mesurable de l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ vers l'espace produit $(E_1 \times \ldots \times E_n, \mathcal{E}_1 \otimes \ldots \otimes \mathcal{E}_n)$.*

Les composantes d'un vecteur aléatoire sont également des variables aléatoires.

Définition 29 *La fonction de répartition du vecteur X est définie par :*

$$F(t_1, \ldots, t_n) = \mathbb{P}(X_1 \leq t_1 \cap \ldots \cap X_n \leq t_n).$$

Définition 30 *L'espérance du vecteur $X = (X_1, \ldots, X_n)$ est le vecteur donné par : $\mathbb{E}(X) = (\mathbb{E}(X_1), \ldots, \mathbb{E}(X_n))^\top$. Si pour tout $i = 1, \ldots, n$ la variable aléatoire X_i admet un moment d'ordre 2, la matrice de variance du vecteur X est donnée par $\Sigma = \mathbb{E}\{(X-\mu)(X-\mu)^\top\} \in \mathbb{R}^{n \times n}$. On a donc $\Sigma_{i,j} = \mathrm{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$.*

Proposition 20 *Soit $X = (X_1, \ldots, X_n) : (\Omega, \mathcal{F}, \mathbb{P}) \to (E_1 \times \ldots \times E_n, \mathcal{E}_1 \otimes \ldots \otimes \mathcal{E}_n)$ un vecteur aléatoire. Les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes si et seulement si la loi $\mathbb{P}_X$ du vecteur X est égale à la mesure produit $\mathbb{P}_{X_1} \otimes \ldots \otimes \mathbb{P}_{X_n}$.*

2.6 Variables aléatoires particulières

2.6.1 Variables aléatoires discrètes

Définition 31 *Une variable aléatoire $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$ est dite discrète si l'ensemble E est fini ou dénombrable. Dans ce cas, on prendra souvent pour $\mathcal{E}$ l'ensemble $\mathcal{P}(E)$ des parties de E.*

Dans toute la suite, on notera $\{e_i, \ i \in I\}$ les éléments de l'ensemble E avec I un ensemble au plus dénombrable.

Proposition 21 *Considérons l'application $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{P}(E))$. Si pour tout $k \in I$, $X^{-1}(\{e_k\}) \in \mathcal{F}$ alors X est une variable aléatoire.*

Proposition 22 *Soit la variable aléatoire discrète $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{P}(E))$. La loi de X est la mesure de probabilité*

$$\mathbb{P}_X = \sum_{k \in I} p_k \delta_{e_k},$$

où pour tout $k \in I$, $p_k = \mathbb{P}[X^{-1}(\{e_k\})]$.

Proposition 23 *Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{P}(E))$ une variable aléatoire discrète et $g : (E, \mathcal{P}(E)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une fonction mesurable. Si*

$$\sum_{k \in I} p_k |g(e_k)| < \infty$$

où $p_k := \mathbb{P}_X(\{e_k\})$*, alors,*

$$\mathbb{E}[g(X)] = \sum_{k\in I} p_k g(e_k).$$

Définition 32 *La fonction génératrice d'une variable aléatoire discrète* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ *est la fonction* $G_X : [0,1] \to [0,1]$ *définie par*

$$G_X(s) = \mathbb{E}(s^X) = \sum_{k=1}^{\infty} \mathbb{P}_X(\{k\})s^k.$$

Proposition 24 *La fonction génératrice d'une variable aléatoire discrète* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ *caractérise entièrement la loi* $\mathbb{P}_X$ *de* X *puisque pour tout* $k \in \mathbb{N}$*,*

$$\mathbb{P}_X(\{k\}) = \frac{G_X^{(k)}(0)}{k!},$$

où $G_X^{(k)}$ *dénote la dérivée* k*-ième de* G_X*.*

Principales lois discrètes à connaître

1) **Loi de Bernoulli** – Une variable aléatoire X suit une loi de Bernoulli de paramètre $p \in]0,1[$ si elle prend ses valeurs dans $\{0,1\}$ avec $\mathbb{P}_X(\{0\}) = 1-p$ et $\mathbb{P}_X(\{1\}) = p$. Pour cette loi, il est facile de montrer que $\mathbb{E}(X) = p$, $\mathrm{Var}(X) = 1-p$ et $G_X(s) = 1 + p(s-1)$.
2) **Loi binomiale** – Une variable aléatoire X suit une loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in]0,1[$ ($X \sim \mathcal{B}(n,p)$) si elle prend ses valeurs dans l'ensemble $\{0,1,\ldots,n\}$ avec

$$\mathbb{P}_X(\{k\}) = C_n^k p^k (1-p)^{n-k}.$$

Pour la loi binomiale $\mathcal{B}(n,p)$, $\mathbb{E}(X) = np$, $\mathrm{Var}(X) = np(1-p)$ et $G_X(s) = (1+p(s-1))^n$.
3) **Loi Géométrique** – Une variable aléatoire X suit une loi géométrique de paramètre $p \in]0,1[$ si elle prend ses valeurs dans $\mathbb{N}^*$ avec $\mathbb{P}_X(\{k\}) = p(1-p)^{k-1}$. On a $\mathbb{E}(X) = 1/p$, $\mathrm{Var}(X) = (1-p)/p^2$ et $G_X(s) = sp/(1-s(1-p))$.
4) **Loi de Poisson** – Une variable aléatoire X suit une loi de Poisson de paramètre $\lambda > 0$ si elle prend ses valeurs dans $\mathbb{N}$ avec

$$\mathbb{P}_X(\{k\}) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}.$$

On a $\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$ et $G_X(s) = \exp(\lambda(s-1))$.

2.6.2 Variables aléatoires de loi absolument continue

Proposition 25 *Soit* $(E, \mathcal{E}, \mu)$ *un espace mesuré. Soit* $f : (E, \mathcal{E}) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ *une fonction mesurable positive.*

1) La fonction d'ensemble ν définie par :

$$\nu : A \in \mathcal{E} \to \int_A f d\mu \in [0, \infty]$$

est une mesure sur l'espace $(E, \mathcal{E})$.
2) Soit $g : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une fonction mesurable et intégrable par rapport à ν,

$$\int_E g d\nu = \int_E g f d\mu.$$

Définition 33 *Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$ une variable aléatoire de loi de probabilité $\mathbb{P}_X$. On dit que $\mathbb{P}_X$ est une loi absolument continue par rapport à une mesure μ sur $(E, \mathcal{E})$ s'il existe une fonction mesurable $f : (E, \mathcal{E}) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ telle que pour tout $A \in \mathcal{E}$,*

$$\mathbb{P}_X(A) = \int_A f d\mu.$$

La fonction f est appelée la densité de X et son intégrale sur E par rapport à μ est égale à 1.

Variable aléatoire à densité Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une variable aléatoire. On dit que X est à densité (ou que la loi de X est à densité) si $\mathbb{P}_X$ est absolument continue par rapport à la mesure de Lebesgue.

Proposition 26 *Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une variable aléatoire.*
i) Soit $g : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une fonction mesurable. Si la loi de X admet une densité f (par rapport à la mesure de Lebesgue) et si $g(X)$ est intégrable alors,

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g f d\lambda = \int_{\mathbb{R}} g(x) f(x) d\lambda(x).$$

ii) Réciproquement, si pour toute fonction g mesurable et bornée, il existe une fonction positive et mesurable f telle que

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) f(x) d\lambda(x).$$

alors f est la densité de $\mathbb{P}_X$ par rapport à la mesure de Lebesgue.

Dans la grande majorité des cas, l'intégrale de Lebesgue coïncide avec celle de Riemann et donc

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) f(x) dx.$$

Proposition 27 *Si la variable aléatoire X est à densité f alors la fonction de répartition F est continue et dérivable λ-presque-partout. De plus, $f(t) = F'(t)$.*

Proposition 28 *Si la fonction de répartition F d'une variable aléatoire X est dérivable λ-presque-partout et telle que $\int_{\mathbb{R}} F' d\lambda = 1$ alors la loi $\mathbb{P}_X$ est à densité de densité F'.*

Vecteurs aléatoires à densité

Proposition 29 *Si le vecteur aléatoire* $X = (X_1, \ldots, X_n)^\top$ *est à densité* $f : \mathbb{R}^n \to \mathbb{R}^+$ *par rapport à la mesure de Lebesgue* λ^n *alors*

$$f(t_1, \ldots, t_n) = \frac{\partial^n F}{\partial t_1 \ldots \partial t_n}(t_1, \ldots, t_n) \ \lambda^n \ p.p.$$

Proposition 30 *Si la loi de* $X = (X_1, \ldots, X_n)^\top$ *admet une densité* f *alors, pour tout* $k \in \{1, \ldots, n\}$ *et pour tout* $1 \leq i_1 < \ldots < i_k \leq n$*, le vecteur* $(X_{i_1}, \ldots, X_{i_k})$ *est également à densité de fonction de densité :*

$$h(x_{i_1}, \ldots, x_{i_k}) = \int_{\mathbb{R}} \ldots \int_{\mathbb{R}} f(x_1, \ldots, x_n) dx_{j_1} \ldots dx_{j_{n-k}},$$

où $\{j_1 < \ldots < j_{n-k}\} = \{i_1 < \ldots < i_k\}^C$.

Principales lois de probabilité à densité

1) **Loi uniforme** – Une variable aléatoire X suit une loi uniforme de paramètres $-\infty < a < b < +\infty$ si $\mathbb{P}_X$ admet pour densité la fonction

$$f(x) = \frac{1}{b-a} \mathbb{I}_{\{x \in [a,b]\}}.$$

On montre que $\mathbb{E}(X) = (a+b)/2$, $\operatorname{Var}(X) = (a-b)^2/12$ et

$$\Phi_X(t) = \frac{\exp(itb) - \exp(ita)}{it(b-a)}.$$

2) **Loi normale** – Une variable aléatoire X suit une loi normale de paramètres $\mu \in \mathbb{R}$ et $\sigma > 0$ si $\mathbb{P}_X$ admet pour densité la fonction

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right\}.$$

On a $\mathbb{E}(X) = \mu$, $\operatorname{Var}(X) = \sigma^2$ et $\Phi_X(t) = \exp(it\mu - \sigma^2 t^2/2)$.
3) **Loi gamma** – Une variable aléatoire X suit une loi gamma de paramètres $a > 0$ et $\lambda > 0$ si $\mathbb{P}_X$ admet pour densité la fonction

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} \exp(-\lambda x) \mathbb{I}_{\{x \geq 0\}}.$$

On a alors, $\mathbb{E}(X) = a/\lambda$, $\operatorname{Var}(X) = a/\lambda^2$ et $\Phi_X(t) = (1 - it/\lambda)^{-a}$. La loi Gamma de paramètres $a = 1$ et $\lambda > 0$ est appelée la loi exponentielle de paramètre λ.
4) **Loi de Cauchy** – Une variable aléatoire suit une loi de Cauchy (standard) si $\mathbb{P}_X$ admet pour densité la fonction

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}.$$

Cette loi n'admet pas d'espérance et $\Phi_X(t) = \exp(-|t|)$. Notez que la loi de Cauchy est un cas particulier de la loi de student que nous verrons plus tard.

2.6.3 Lois de mélange

Il existe des lois qui ne sont ni discrètes ni absolument continues mais construites comme un mélange de ces deux types de lois.

Définition 34 *Soit* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *une variable aléatoire. On dit que la loi de* X *est une loi de mélange s'il existe deux mesures* $\mathbb{P}_X^{(1)}$ *et* $\mathbb{P}_X^{(2)}$ *avec* $\mathbb{P}_X^{(1)}$ *loi discrète et* $\mathbb{P}_X^{(2)}$ *loi admettant une densité par rapport à la mesure de Lebesgue et un paramètre* $\alpha \in]0, 1[$ *tels que* $\mathbb{P}_X = \alpha \mathbb{P}_X^{(1)} + (1 - \alpha)\mathbb{P}_X^{(2)}$.

Annexe A – Rappel des notions à connaître

Théorie des ensembles

Un ensemble E est la réunion de plusieurs éléments. Par exemple, $\mathbb{N} = \{0, 1, 2, \ldots\}$ est l'ensemble des entiers naturels, $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ est l'ensemble des entiers relatifs, etc .... L'ensemble vide $\emptyset$ est l'ensemble ne contenant pas d'éléments.

Définition 35 *Un ensemble E est fini s'il existe $n \in \mathbb{N}^* = \{1, 2, \ldots\}$ et une bijection $f : E \to \{1, \ldots, n\}$. Le nombre d'éléments de E est alors n que l'on appelle cardinal de E. On note* $\text{Card}(E) = n$. *Un ensemble E est dénombrable s'il existe une bijection $f : E \to \mathbb{N}$.*

Par exemple, l'ensemble $\mathbb{N}$ et l'ensemble des rationnels $\mathbb{Q}$ sont dénombrables. Ce n'est évidemment pas le cas des ensembles $\mathbb{R}$ et $\mathbb{C}$.

Définition 36 *On dit que A est un sous-ensemble de E si tout élément de A est aussi dans E. On note $A \subset E$.*

L'ensemble vide $\emptyset$ est un sous-ensemble de n'importe quel ensemble. Il ne faut pas confondre un él'ement x de E avec le singleton $\{x\}$ qui est le sous-ensemble de E ne contenant que l'élément x.

Définition 37 *Soit E un ensemble. Les sous-ensembles de E forment un ensemble appelé "ensemble des parties de E" et noté $\mathcal{P}(E)$. L'ensemble $\mathcal{P}(E)$ est donc un ensemble d'ensembles. Si* $\text{Card}(E) = n$ *alors* $\text{Card}(\mathcal{P}(E)) = 2^n$.

Par exemple, l'ensemble des parties de $E = \{0, 1\}$ est donné par $\mathcal{P}(E) = \{\{0, 1\},\ \{0\},\ \{1\}, \emptyset\}$.

Définition 38 *Soit $(A_i)_{i \geq 1}$ une suite de sous-ensembles de E. On dit que $(A_i)_{i \geq 1}$ est une partition de E si pour tout $i \neq i'$, $A_i \cap A_{i'} = \emptyset$ et si $A_1 \cup A_2 \cup \ldots = E$.*

Définition 39 *Soit A un sous-ensemble de E. Le complémentaire de A dans E est le sous-ensemble A^C qui contient tous les éléments de E qui ne sont pas dans A. Si E est fini,* $\text{Card}(A^C) = \text{Card}(E) - \text{Card}(A)$.

Proposition 31 *Soit $(A_i)_{i \geq 1}$ une suite de sous-ensembles de E et B un sous-ensemble de E. On a les propriétés suivantes :*

$$\left(\bigcup_{i \geq 1} A_i \right)^C = \bigcap_{i \geq 1} A_i^C, \quad \left(\bigcap_{i \geq 1} A_i \right)^C = \bigcup_{i \geq 1} A_i^C, \quad B \cap \bigcup_{i \geq 1} A_i = \bigcup_{i \geq 1} (B \cap A_i)$$

et

$$B \cup \bigcap_{i \geq 1} A_i = \bigcap_{i \geq 1} (B \cup A_i).$$

Définition 40 *Soient E et F deux ensembles. Le produit cartésien de E et F est l'ensemble $E \times F = \{(x, y) \mid x \in E,\ y \in F\}$.*

Dénombrement

Le dénombrement sert à répondre à la question suivante : combien y a t'il de façons de construire un ensemble de p éléments pris dans un ensemble de n éléments. Ce nombre est différent selon que l'on autorise ou non les répétitions (prendre plusieurs fois le même élément) et que l'on tienne compte ou pas de l'ordre des éléments ($\{e_1, e_2\} = \{e_2, e_1\}$ ou $\{e_1, e_2\} \neq \{e_2, e_1\}$). Le tableau ci dessous regroupe les formules utiles :

	avec ordre	sans ordre
avec répétition	n^p	C_{n+p-1}^p
sans répétition	A_n^p	C_n^p

Remarquons que si l'on n'autorise pas les répétitions, alors forcément $n \geq p$. On rappelle que

$$C_n^p = \frac{n!}{p!(n-p)!} \text{ et } A_n^p = p!C_n^p = \frac{n!}{(n-p)!}.$$

Enfin, soient E et F deux ensembles dénombrables de cardinal respectif p et n (ici on peut avoir $n < p$). La valeur n^p correspond au nombre d'applications possibles de E dans F.
Par exemple, si $E = \{1, 2, 3\}$ et $F = \{a, b\}$, il y a $8 = 2^3$ applications possibles : $\{1, 2, 3\} \to \{a, a, a\}, \{1, 2, 3\} \to \{a, a, b\}, \ldots$ Ainsi, A_n^p au nombre possibles d'applications injectives de E dans F (si $n \geq p$). Donc, si $n = p$, $A_n^p = n!$ correspond au nombre de bijection de E dans F ou autrement dit au nombre de permutations d'un ensemble de n éléments.

Construction de l'intégrale de Riemann

Nous rappelons ici les étapes menant à la construction de l'intégrale de Riemann d'une fonction réelle bornée f définie sur un intervalle $[a, b]$. La première étape consiste à définir l'intégrale de fonctions positives en escalier.

Définition 41 *Soient $-\infty < a < b < \infty$. Une fonction g en escalier définie sur un intervalle $[a, b]$ s'écrit sous la forme :*

$$g = \sum_{i=1}^{k} \alpha_i \mathbb{I}_{\{]x_{i-1}, x_i]\}},$$

où $\alpha_i \in \mathbb{R}$ pour tout $i = 1, \ldots, k$ et $a = x_0 \leq x_1 \leq \ldots \leq x_k = b$ est une subdivision de l'intervalle $[a, b]$. On notera $\mathcal{G}([a, b])$ l'ensemble de ces fonctions.

L'intégrale de Riemann d'une fonction $g \in \mathcal{G}([a, b])$ est donnée par :

$$\int_a^b g(x)dx := \sum_{i=1}^{k} \alpha_i (x_i - x_{i-1}).$$

Elle correspond à l'aire sous la fonction en escalier g. Nous passons à présent à la définition d'une fonction Riemann intégrable. On note

$$\begin{aligned} I^-(f) &= \sup\left\{\int_a^b g(x)dx \mid g \in \mathcal{G}([a,b]) \text{ et } g \leq f\right\} \\ \text{et } I^+(f) &= \inf\left\{\int_a^b g(x)dx \mid g \in \mathcal{G}([a,b]) \text{ et } g \geq f\right\}. \end{aligned}$$

Définition 42 *Soit f une fonction bornée définie sur un intervalle $[a,b]$. La fonction f est dite Riemann intégrable si $I^-(f) = I^+(f)$. On a alors :*

$$I^+(f) = I^-(f) = \int_a^b f(x)dx.$$

Il est facile de montrer que l'ensemble des fonctions Riemann intégrables sur $[a,b]$ forme un sous espace vectoriel de l'ensemble des fonctions bornées sur $[a,b]$.

Théorème 11 *Si f est continue sur $[a,b]$ alors f est Riemann intégrable et la fonction*

$$F(x) := \int_a^x f(t)dt,$$

est dérivable sur $[a,b]$ et telle que $F'(x) = f(x)$. La fonction F est une primitive de f.

Intégrale impropre

Définition 43 *Soient $-\infty \leq a < b \leq \infty$ et soit $f :]a,b[\to \mathbb{R}$ une fonction localement intégrable (c'est-à-dire Riemann intégrable sur tout intervalle $[c,d] \subset]a,b[$).*

i) On dit que l'intégrale de f sur $]a,b[$ converge s'il existe $c \in]a,b[$ tel que les limites

$$\lim_{x \uparrow b} \int_c^x f(t)dt \text{ et } \lim_{x \downarrow a} \int_x^c f(t)dt,$$

existent et sont finies. On note alors

$$\int_a^b f(t)dt = \lim_{x \downarrow a} \int_x^c f(t)dt + \lim_{x \uparrow b} \int_c^x f(t)dt,$$

l'intégrale de Riemann impropre de f sur $]a,b[$.

ii) Une fonction $f :]a,b[\to \mathbb{R}$ d'intégrale convergente est dite intégrable au sens de Riemann généralisé si l'intégrale de f sur $]a,b[$ est absolument convergente c'est-à-dire s'il existe $c \in]a,b[$ tel que

$$\lim_{x \uparrow b} \int_c^x |f(t)|dt < \infty \text{ et } \lim_{x \downarrow a} \int_x^c |f(t)|dt < \infty.$$

Un exemple classique de fonction d'intégrale convergente mais pas absolument convergence est la fonction $t \in]0, \infty[\mapsto \sin(t)/t$. En effet, on a

$$\lim_{x\to\infty} \int_0^x \frac{\sin(t)}{t} dt = \frac{\pi}{2} \text{ et } \lim_{x\to\infty} \int_0^x \frac{|\sin(t)|}{t} dt = \infty.$$

La fonction $\sin(t)/t$ n'est donc pas intégrable au sens de Riemann généralisé.