Probabilité

M. KAFANDO
Ce cours a été redigé par **Denis Villemonais**,
enseignant à **IECL – Ecole des Mines de Nancy**

UO3S-LMI3

Semestre 6

Programme

1. Mesures et tribus
2. Fonctions mesurables et variables aléatoires Intégrale de Lebesgue et Espérance.
3. Théorèmes d'intégration
4. Fonction de répartition et fonction caractéristique
5. Théorème de Fubini
6. Indépendance de variables aléatoires
7. Variance, moments et espaces L^p
8. Suites de variables aléatoires
9. Vecteurs gaussiens

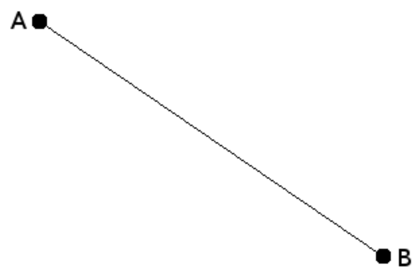
Probabilités
1. Mesures et tribus

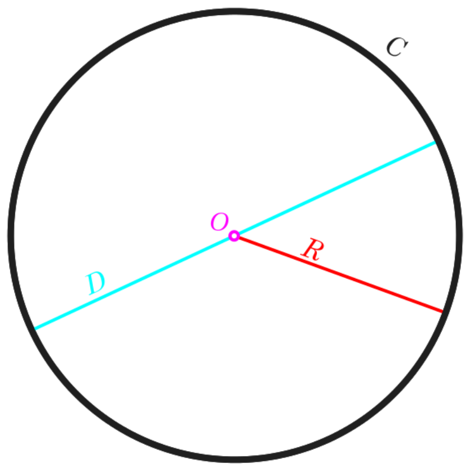
´

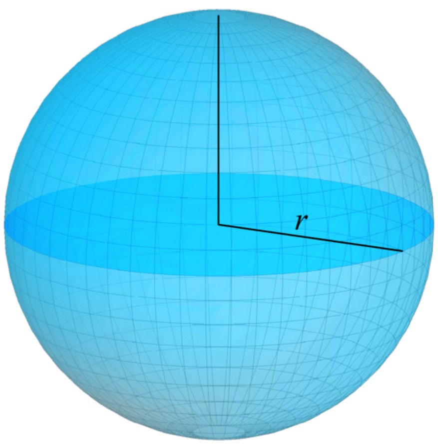
0. Motivations pour une théorie de la mesure

Nous savons déjà mesurer des ensembles simples...

- $\rightarrow$ Le segment $]a,b[$, de longueur $b-a$
- $\rightarrow$ Le disque de rayon r, de surface πr^2
- $\rightarrow$ La boule de rayon r, de volume $4\pi r^3/3$

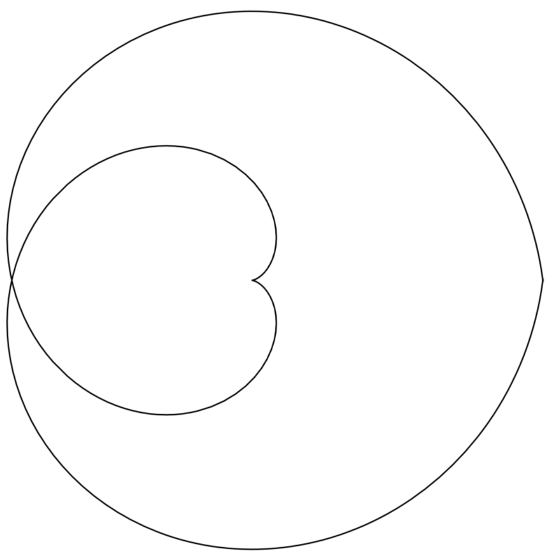




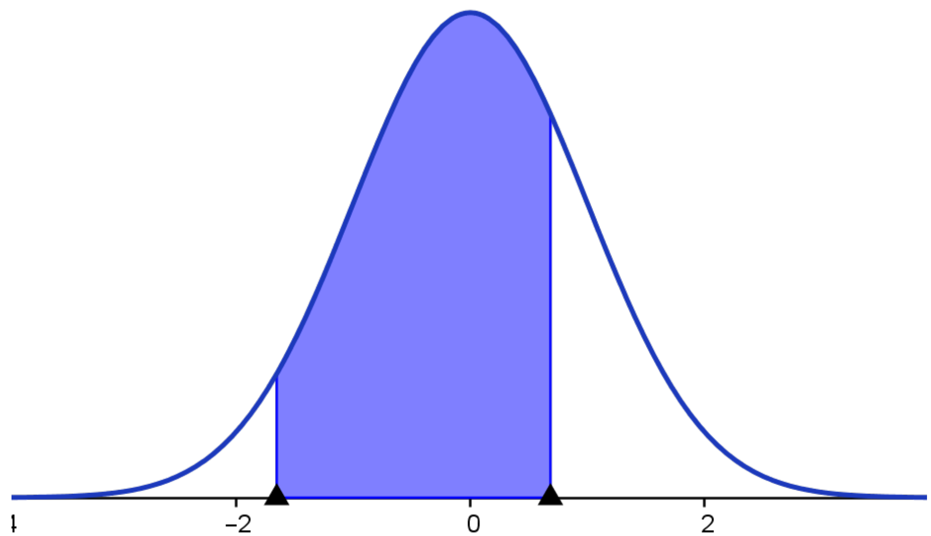


... et moins simples.

- $\rightarrow$ Les longueurs de courbes paramétrées régulières
- $\rightarrow$ Les surfaces délimitées par des courbes régulières



$$\left\{\begin{array}{l} x = \cos(t) + t\sin(t) \\ y = \sin(t) - t\sin(t) \end{array}\right.$$

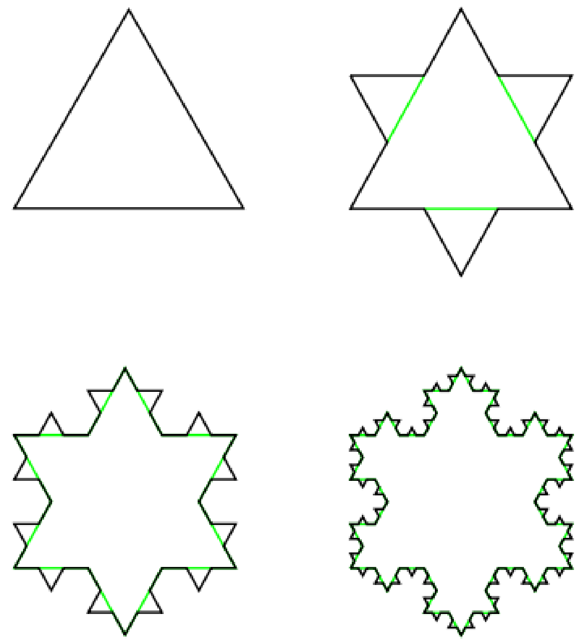


Courbe gaussienne

On peut étendre ce savoir à d'autres ensembles...

→ Des volumes extrudés

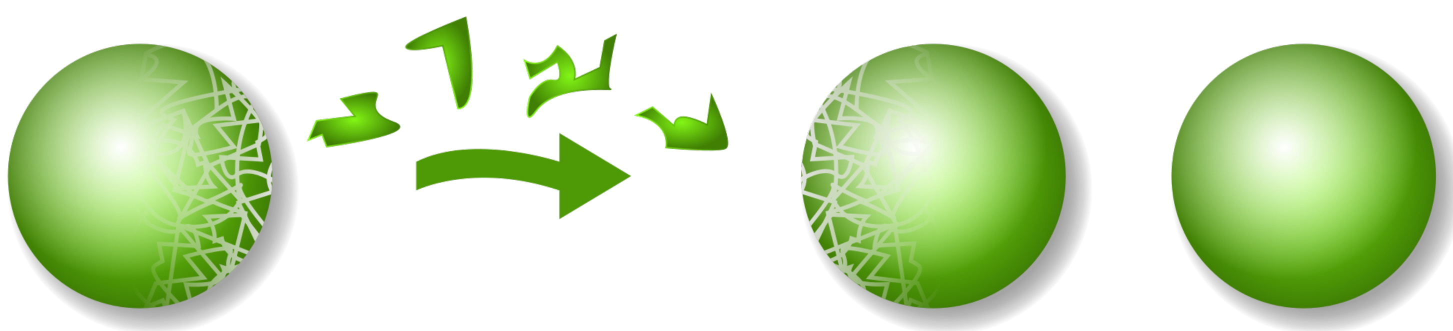
→ Des longueurs successives



… en prenant le risque de ne pas savoir ce que l'on fait.

→ Ensemble de Vitali : il contient [0,1] tout en étant de longueur nulle.

→ Paradoxe de Banach-Tarski : on peut découper une boule en un nombre fini de morceaux et reconstituer deux boules identiques disjointes (pas de conservation du volume !)



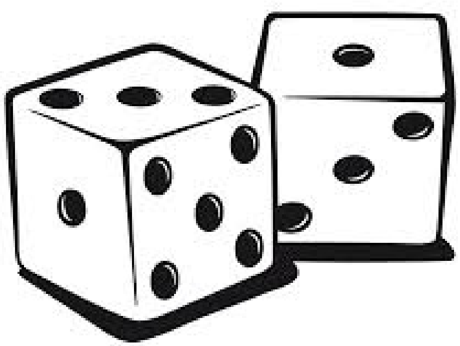
0. Motivations pour une théorie des probabilités

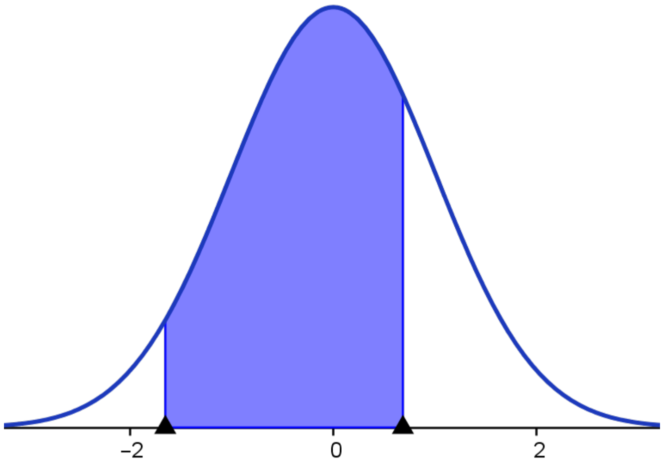
Nous savons déjà mener à bien des calculs de probabilités...

$\rightarrow$ Lancé de dés

$\rightarrow$ Tirage de boules dans une urne

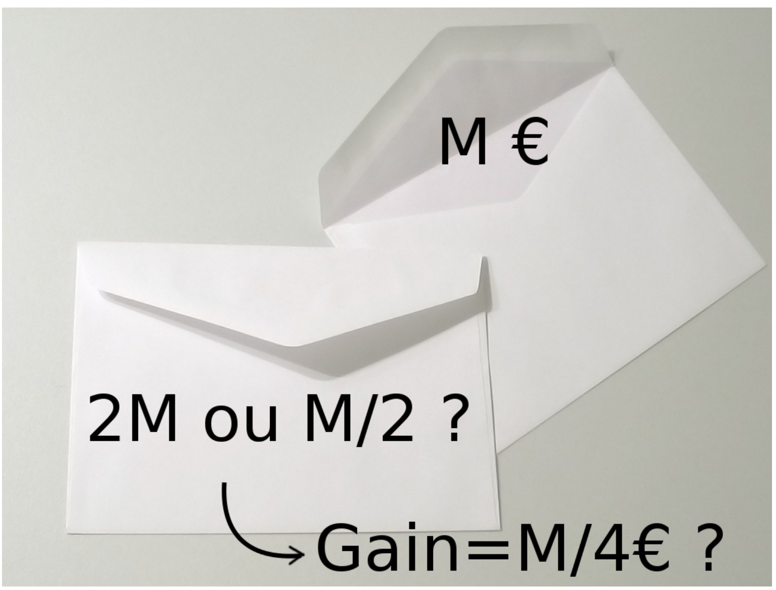
$\rightarrow$ Loi gaussienne : étude d'une variable aléatoire normale

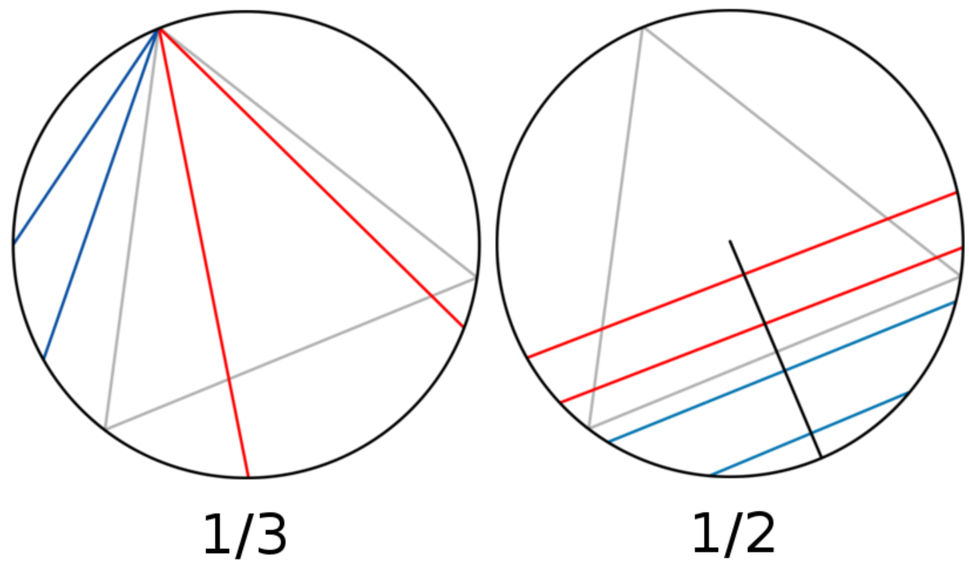




… mais ces notions vagues ne permettent pas de conduire une discussion probabiliste rigoureuse, même dans des situation simples…

- $\rightarrow$ Problème des deux enveloppes
- $\rightarrow$ Paradoxe de Bertrand : probab. qu'une corde $>\sqrt{3}/2$



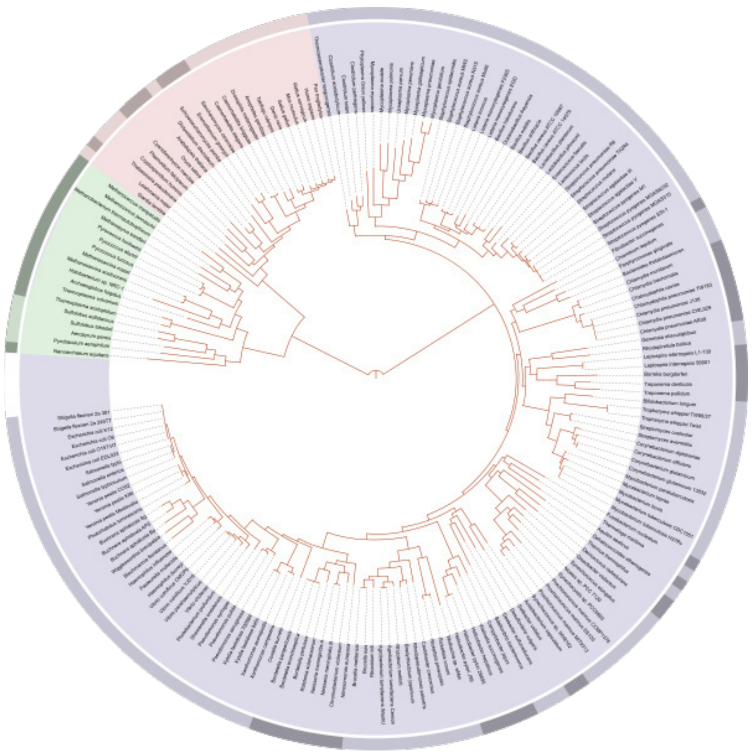


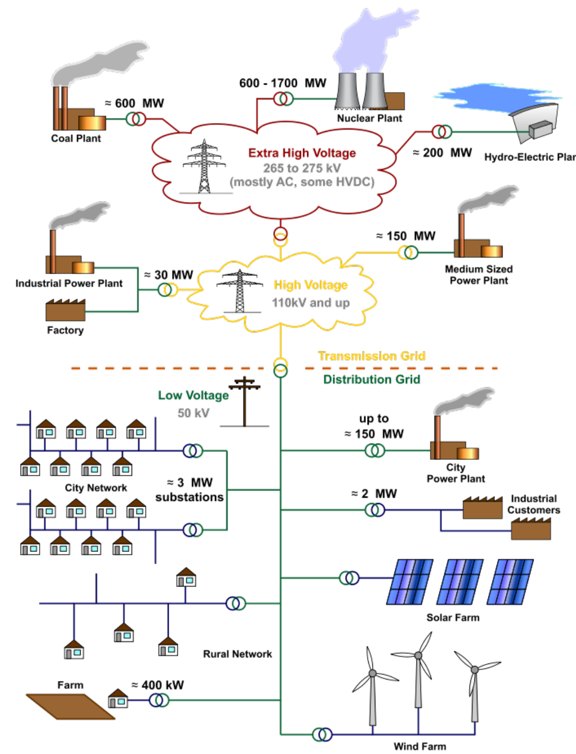
... sans compter que le réel est généralement bien plus complexe.

→ Arbre d'évolution des espèces

→ Réseau électrique

→ Big data...





1. Les tribus

Dans la suite, Ω est un ensemble non vide.

Tribu sur Ω

Un ensemble de parties $\mathcal{F} \subset \mathcal{P}(\Omega)$ est une *tribu* si

- $\Omega \in \mathcal{F}$ (ou $\emptyset \in \mathcal{F}$),
- $\mathcal{F}$ est stable par passage au complémentaire,
- $\mathcal{F}$ est stable par réunions (ou intersections) dénombrables.

Les éléments de $\mathcal{F}$ sont appelés les *ensembles mesurables* (ou les ensembles $\mathcal{F}$*-mesurables* en cas d'ambiguïté).

$\longrightarrow$ Exemples :

- la tribu grossière $\{\emptyset,\Omega\}$,
- la tribu totale $\mathcal{P}(\Omega)$,
- la plus petite tribu contenant $A \subset \Omega$ est égale à $\{\emptyset,A,A^c,\Omega\}$.

Tribu engendrée par un ensemble de parties

Soit $\mathcal{S} \subset \mathcal{P}(E)$. La plus petite tribu (au sens de $\subset$) contenant $\mathcal{S}$ est appelée la *tribu engendrée par* $\mathcal{S}$. Elle est notée $\sigma(\mathcal{S})$.

$\rightarrow$ Remarques

- Si $\mathcal{F}$ est une tribu, alors $\sigma(\mathcal{F}) = \mathcal{F}$,
- Si $\mathcal{S}_1 \subset \sigma(\mathcal{S}_2)$, alors $\sigma(\mathcal{S}_1) \subset \sigma(\mathcal{S}_2)$,
- Si $\mathcal{S}_1 \subset \sigma(\mathcal{S}_2)$, alors $\sigma(\mathcal{S}_1) \subset \sigma(\mathcal{S}_2)$,
- Si $\mathcal{S} \subset \mathcal{F}$ tribu, alors $\sigma(\mathcal{S}) \subset \mathcal{F}$.

$\rightarrow$ Méthodes pour $\sigma(\mathcal{S}_1) = \sigma(\mathcal{S}_2)$ ou $\sigma(\mathcal{S}) = \mathcal{F}$

- $\mathcal{S}_1 \subset \sigma(\mathcal{S}_2)$ et $\mathcal{S}_2 \subset \sigma(\mathcal{S}_1) \Rightarrow \boxed{\sigma(\mathcal{S}_1) = \sigma(\mathcal{S}_2)}$,
- $\mathcal{F}$ est une tribu telle que $\mathcal{F} \subset \sigma(\mathcal{S})$ et $\mathcal{S} \subset \mathcal{F} \Rightarrow \boxed{\sigma(\mathcal{S}) = \mathcal{F}}$.

$\rightarrow$ Exercices

1. Montrer que $\sigma(\{A\}) = \{\emptyset, A, A^c, \Omega\}$,
2. Montrer que
$\sigma(\{\{a\},\{a,b\},\{a,b,c\}\}) = \sigma(\{\{a\},\{b\},\{c\}\}) = \mathcal{P}(\{a,b,c\})$.

Topologie induite : rappelons que $A \subset \Omega \subset \mathbb{R}^d$ est un ouvert de Ω $\Leftrightarrow \exists\ O$ ouvert de $\mathbb{R}^d$ tel que $A = O \cap \Omega$.

Tribu borélienne sur $\Omega \subset \mathbb{R}^d$

Soit $\Omega \subset \mathbb{R}^d$. On appelle *tribu borélienne sur* Ω la tribu engendrée par tous les ouverts de Ω. Elle est notée $\mathcal{B}(\Omega)$.

→ Remarques

- Si f est continue, alors $f^{-1}(A) \in \mathcal{B}(\mathbb{R}^d)$, $\forall A \in \mathcal{B}(\mathbb{R}^{d'})$,
- $\mathcal{B}(\Omega)$ est également engendré par les fermés de Ω.
- $\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R})$.

→ Exercices

1. (*) Montrer que

$$A := \{(x,y,z) \in \mathbb{R}^3,\ x^2 + yz^3 < -1\} \in \mathcal{B}(\mathbb{R}^3),$$

2. (*) Soit $A \in \mathcal{B}(\mathbb{R}^d)$. Montrer que $\mathcal{B}(A) \subset \mathcal{B}(\mathbb{R}^d)$,

2. Mesures

Soit Ω un ensemble non vide et $\mathcal{F}$ une tribu sur Ω.

Définition

$\mu : \mathcal{F} \to [0, +\infty]$ est appelée une *mesure* si

- $\mu(\emptyset) = 0$,
- $\forall A_0, A_1, \ldots, A_n, \ldots \in \mathcal{F}$ disjoints,

$$\mu\left(\bigsqcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n).$$

$\rightarrow$ Si $\mu(\Omega) = 1$, alors μ est appelée une *probabilité*.

$\rightarrow$ Exemples

- La mesure de Dirac $\delta_a : A \mapsto \mathbf{1}_{a \in A}$ est une probabilité,
- La mesure de comptage $A \mapsto \mathrm{Card}(A)$ est une mesure.

Soient $A, B, A_0, A_1, \ldots, A_n, \ldots \in \mathcal{F}$.

Propriétés des mesures

Soit μ une mesure sur $(\Omega, \mathcal{F})$, alors

- μ est *croissante* : si $A \subset B$, alors $\mu(A) \leq \mu(B)$
- μ est *sous-additive* : $\mu\left(\bigcup_{n\geq 0} A_n\right) \leq \sum_{n\geq 0} \mu(A_n)$
- μ est *continue* :

Si $(A_n) \nearrow$, alors $\mu\left(\bigcup_{n\geq 0} A_n\right) = \lim_{n\to\infty} \mu(A_n)$

Si $(A_n) \searrow$ et $\mu(A_0) < \infty$, alors $\mu\left(\bigcap_{n\geq 0} A_n\right) = \lim_{n\to\infty} \mu(A_n)$

$\rightarrow$ Exercice

1. Soit μ la mesure de comptage sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. On pose $A_n = [n, \infty[$ pour tout $n \geq 0$. Montrer que, sans l'hypothèse $\mu(A_0) < \infty$) le dernier point de la propriété ne s'applique pas.

Mesures discrètes sur $(\Omega,\mathcal{F})$

La mesure de Dirac en a est donnée par

$$\delta_a(A) = \mathbf{1}_{a\in A} = \begin{cases} 1 & \text{si } a \in A, \\ 0 & \text{sinon.} \end{cases}$$

Une mesure μ est dite *discrète* si il existe $K \subset \Omega$ au plus dénombrable et $(p_k)_{k\in K} \in \mathbb{R}_+^K$ tels que, $\forall A \in \mathcal{F}$,

$$\mu = \sum_{k\in K} p_k \delta_k.$$

→ Remarques

- Dans ce cas, $\mu(A) = \sum_{k\in K} p_k\delta_k(A) = \sum_{k\in A} p_k$,
- μ est une probabilité ssi $\sum_{k\in K} p_k = 1$.

→ Exemple

- $\mu = \frac{1}{2}\delta_1 + \frac{1}{3}\delta_2 + \frac{1}{6}\delta_3$ est une probabilité discrète sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$,

→ $\mu\left(\left[0,\frac{5}{2}\right]\right) = \frac{1}{2}\delta_1\left(\left[0,\frac{5}{2}\right]\right) + \frac{1}{3}\delta_2\left(\left[0,\frac{5}{2}\right]\right) + \frac{1}{6}\delta_3\left(\left[0,\frac{5}{2}\right]\right) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}.$

Mesure de Lebesgue sur $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$

La *mesure de Lebesgue* sur $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$, notée λ_d, est l'unique mesure positive telle que, pour tous $a_i < b_i \in \mathbb{R}$,

$$\lambda_d(]a_1,b_1[\times \cdots \times]a_d,b_d[) = \prod_{i=1}^{d}(b_i - a_i).$$

→ Remarques

- λ_1,λ_2 et λ_3 généralisent les notions de longueur, surface et volume déjà connues.
- $\lambda_d([0,1]^d) = 1$,
- λ_d est stable par rotation/translation/symétrie.

→ Exemple

- La mesure de Lebesgue λ_2 d'un disque de rayon r est πr^2.

→ Exercice (TD)

1. Montrer que $\lambda_1(\mathbb{Q}) = 0$ et $\lambda_d(\mathbb{Q} \times \cdots \times \mathbb{Q}) = 0$.

Mesures à densité continue par morceaux sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$

Soit f une fonction continue par morceaux positive sur $\mathbb{R}^d$. La mesure de densité f par rapport à la mesure de Lebesgue est l'unique mesure sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$, notée $f\,\lambda_1$, telle que, pour tous $a < b \in \mathbb{R}$,

$$f\,\lambda_1(]a,b[) = \int_a^b f(x)\,dx.$$

$\rightarrow$ Exemples

- La loi gaussienne est la mesure de densité $x \mapsto \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$.
- La mesure de densité $\mathbf{1}_{[0,2]}/2$ sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$, notée $\mathbf{1}_{[0,2]}/2\,\lambda_1$,

$$\frac{\mathbf{1}_{[0,2]}}{2}\,\lambda_1([1,3]) = \int_1^3 \frac{\mathbf{1}_{[0,2]}(x)}{2}\,dx = \int_1^2 \frac{1}{2}\,dx = \frac{1}{2}.$$

Ensemble négligeable d'un espace mesuré $(\Omega,\mathcal{F},\mu)$

Un ensemble $N \in \mathcal{P}(\Omega)$ est dit *négligeable* si il existe $A \in \mathcal{F}$ tel que $N \subset A$ et $\mu(A) = 0$.

$\rightarrow$ Remarques

- Un négligeable peut être non-mesurable ;
- Une propriété $P(\omega)$ est dite vraie *presque partout* (ou *presque sûrement* si μ est une probabilité) si l'ensemble

$$N = \{\omega \in \Omega,\ P(\omega) \text{ est faux}\}$$

  est négligeable.

$\rightarrow$ Exemples

- $\mathbb{Q}$ est un ensemble négligeable de $(\mathbb{R},\mathcal{B}(\mathbb{R}),\lambda_1)$,
- $P(x) =$"$x \notin \mathbb{Q}$" est vraie presque partout pour $(\mathbb{R},\mathcal{B}(\mathbb{R}),\lambda_1)$. En effet :

$$\lambda_1(\{x,\ P(x) \text{ est faux}\}) = \lambda_1(\{x,\ x \in \mathbb{Q}\}) = \lambda_1(\mathbb{Q}) = 0.$$

Espace mesurable complet

L'espace mesuré $(\Omega,\mathcal{F},\mu)$ est dit *complet* si $\mathcal{F}$ contient tous les ensembles négligeables.

→ Exemple

- La mesure de Lebesgue peut être étendue à la *tribu de Lebesgue*, définie par $\sigma(\mathcal{B}(\mathbb{R}) \cup \{N,\ N \text{ négligeable}\})$.

→ (*) Exercice (TD)

- On peut toujours "compléter" un espace en utilisant la tribu $\mathcal{F}' = \sigma(\mathcal{F} \cup \{N,\ N \text{ négligeable}\})$ et en étendant la mesure μ à la nouvelle tribu.

3. Indépendance entre des événements

Indépendance entre des évènements

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de proabilité.

Définition

Des événements $A_1,A_2,\dots,A_n,\dots \in \mathcal{F}$ sont dits **mutuellement indépendants** si, pour tous $i_1 < i_2 < \dots < i_k$,

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \prod_{l=1}^{k} \mathbb{P}(A_{i_l}).$$

$\rightarrow$ Remarques

- Il peut y avoir un nombre fini ou infini d'événements en jeu
- On dira parfois "indépendants" au lieu de "mutuellement indépendants"
- $\forall A \in \mathcal{F}$, les évenements A, $\emptyset$ et Ω sont mutuellement indépendants.

Exercice 1. On lance une pièce non truquée et on définit les évenements suivants :

$A = \{\text{Pile au 1er lancer}\}$ et $B = \{\text{même résultat aux deux lancers}\}$.

Est-ce que A et B sont indépendants ?

Solution. On a

$$\mathbb{P}(A) = \frac{1}{2} \text{ et } \mathbb{P}(B) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Or

$$\mathbb{P}(A \cap B) = \mathbb{P}(\text{Pile aux deux lancers}) = \frac{1}{4}.$$

Donc $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ donc

$$\boxed{A \text{ et } B \text{ sont indépendants.}}$$

Exercice 2. On lance une pièce non truquée et on définit les évenements suivants :

$$A = \{\text{Pile au 1er lancer}\} \text{ et } B = \{\text{Pile aux deux lancers}\}.$$

Est-ce que A et B sont indépendants ?

Solution. On a

$$\mathbb{P}(A) = \frac{1}{2} \text{ et } \mathbb{P}(B) = \frac{1}{4}.$$

Or

$$\mathbb{P}(A \cap B) = \mathbb{P}(\text{Pile aux deux lancers}) = \frac{1}{4}.$$

Donc $\mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B)$ donc

$$\boxed{A \text{ et } B \text{ ne sont pas indépendants.}}$$

4. Probabilités conditionnelles

Probabilités conditionnelles : Définition

Définition

Soit $A \in \mathcal{F}$ tel que $\mathbb{P}(A) > 0$. On définit alors pour tout $B \in \mathcal{F}$,

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}.$$

$\rightarrow$ Remarque

- Cette définition ne permet pas de construire les probabilités conditionnelles par rapport à des événements de probabilité nulle (mais ça existe).
- A et B indépendants ssi $\mathbb{P}(A) = 0$ ou $\mathbb{P}(B \mid A) = \mathbb{P}(B)$.

Formule de Bayes

Proposition

Soit $(A_n)_{n\in\mathbb{N}}$ des événements tels que $A_i \cap A_j = \emptyset$ pour $i \neq j$, et

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 1.$$

Alors, pour tout événements B, C,

$$\mathbb{P}(C \mid B) = \frac{\mathbb{P}(B \mid C)\mathbb{P}(C)}{\sum_{n\in\mathbb{N}} \mathbb{P}(B \mid A_n)\mathbb{P}(A_n)}.$$

Souvent

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B \mid A)\mathbb{P}(A) + \mathbb{P}(B \mid A^c)\mathbb{P}(A^c)}$$

Application 1

Un individu est tiré au hasard dans une population contenant une proportion de $\mathbf{10^{-4}}$ personnes atteintes d'une maladie rare.

Test : Vrai positif 99%, Faux positif 0.1%.

Quelle est la probabilité qu'un individu tiré au hasard et ayant un résultat positif soit effectivement porteur de cette maladie ?

Solution. Notons $A = \{\text{atteint}\}$ et $B = \{\text{test positif}\}$, ainsi

$$\mathbb{P}(A) = 10^{-4},\ \mathbb{P}(B|A) = 0{,}99 \text{ et } \mathbb{P}(B|A^c) = 10^{-3}.$$

Donc

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)} \simeq 0.09.$$

Application 2.

Le joueur attrape et lance une pièce au hasard parmi les trois suivantes :

$\rightarrow$ *PP* (deux côtés "pile"), *FF* (deux côtés "face") et *PF*.

En retombant, la pièce affiche un côté "pile". Quelle est la probabilité que l'autre côté de la pièce soit "face"?

Solution. On pose X le nom de la pièce choisie, X_1 le côté montré et X_2 le côté caché.

$\rightarrow \mathbb{P}(X_2 = face \mid X_1 = pile) = \mathbb{P}(X = PF \mid X_1 = pile) = ?$

Formule de Bayes :

$$\mathbb{P}(X = PF \mid X_1 = pile) = \frac{\mathbb{P}(X_1 = pile \mid X = PF)\mathbb{P}(X = PF)}{\mathbb{P}(X_1 = pile)}.$$

Or $\mathbb{P}(X_1 = pile \mid X = PF) = 1/2$, $\mathbb{P}(X = PF) = 1/3$ et $\mathbb{P}(X_1 = pile) = 1/2$, donc

$$\boxed{\text{Probabilité que le côté caché soit "face"} = 1/3}$$

5. Lemmes de Borel-Cantelli

Lemme (Borel-Cantelli pour des événements quelconques)

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{A}$ telle que

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < +\infty.$$

Alors,

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = \mathbb{P}\left(\bigcap_{n=0}^{+\infty}\bigcup_{k=n}^{+\infty} A_k\right) = 0.$$

c'est-à-dire que

$$\mathbb{P}\left(\{\omega \in \Omega \,/\, \exists \text{ une infinité d'indices } n \text{ tq } \omega \in A_n\}\right) = 0$$

ou encore

$$\text{"Card}\,\{n \in \mathbb{N} \,/\, \omega \in A_n\} < +\infty\text{" } \mathbb{P}(\omega)\text{-presque sûrement.}$$

Ex 1. On lance des dés non truqués, le $n^{\text{ème}}$ dé possédant n^2 faces. Quelle est la probabilité qu'un nombre infini de dés tombent sur la face 1 ?

Solution. Posons

$$A_n = \left\{ \text{le } n^{\text{ème}} \text{ dé tombe sur la face } 1 \right\}.$$

On a

$$\mathbb{P}(A_n) = \frac{1}{n^2}.$$

Donc

$$\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) < +\infty.$$

D'après le lemme de Borel-Cantelli,

$$\boxed{\mathbb{P}\left(\{\omega \in \Omega \,/\, \exists \text{ une infinité d'indices } n \text{ tq } \omega \in A_n\}\right) = 0}$$

soit

$$\boxed{\mathbb{P}\left(\text{un nombre infini de dés tombent sur la face } 1\right) = 0.}$$

Lemme (Borel-Cantelli pour des événements indépendants)

Soit $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements indépendants telle que

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty.$$

Alors,

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = \mathbb{P}\left(\bigcap_{n=0}^{+\infty} \bigcup_{k=n}^{+\infty} A_k\right) = 1.$$

c'est-à-dire que

$$\mathbb{P}\left(\{\omega \in \Omega \,/\, \exists \text{ une infinité d'indices } n \text{ tq } \omega \in A_n\}\right) = 1$$

ou encore

$$\text{"}\mathrm{Card}\,\{n \in \mathbb{N} \,/\, \omega \in A_n\} = +\infty\text{" } \mathbb{P}(\omega)\text{-presque sûrement.}$$

Probabilités
2. Fonctions mesurables et variables aléatoires
Intégrale de Lebesgue et Espérance

´

1. Fonctions mesurables

Soient E et F munis des tribus respectives $\mathcal{E}$ et $\mathcal{F}$.

Fonction mesurable

Une fonction $f : E \to F$ est dite mesurable par rapport à $\mathcal{E}$ et $\mathcal{F}$ si, pour tout $A \in \mathcal{F}$, $f^{-1}(A) \in \mathcal{E}$.

$\to$ Remarques

- Si $\mathcal{E}$ et $\mathcal{F}$ sont les tribus boréliennes, f est dite borélienne,
- Si $\mathcal{F} = \sigma(\mathcal{S})$, alors f est mesurable si et seulement si $f^{-1}(A) \in \mathcal{E}$ pour tout $A \in \mathcal{S}$.

$\to$ Exemples

- Une fonction constante est toujours mesurable,
- Si $\mathcal{E} = \{\emptyset, E\}$ et $\mathcal{F} = \mathcal{P}(F)$, alors les seules fonctions mesurables sont les fonctions constantes,
- Si $\mathcal{E} = \mathcal{P}(E)$, alors toutes les fonctions sont mesurables,
- Si $E = \{0,1,2\}$, $\mathcal{E} = \sigma(\{0\},\{1,2\})$ et $\mathcal{F} = \mathcal{P}(F)$, alors les fonctions mesurables vérifient $f(1) = f(2)$.

Propriétés des fonctions boréliennes

On suppose que $E \subset \mathbb{R}^d$, $\mathcal{E} = \mathcal{B}(E)$, $F = \mathbb{R}^{d'}$ et $\mathcal{F} = \mathcal{B}(\mathbb{R}^{d'})$.

- Les fonctions indicatrices de boréliens ($x \mapsto \mathbf{1}_A(x)$, avec $A \subset E$ borélien) sont boréliennes ;
- Les fonctions continues sont boréliennes ;
- Les sommes, produits et composées de fonctions boréliennes sont boréliennes ;
- Les limites ponctuelles de fonctions boréliennes sont également boréliennes ;
- Les sup, inf, lim sup et lim inf de suites de fonctions boréliennes sont boréliennes.

$\rightarrow$ Remarque

- Les fonctions étagées (c'est-à-dire les sommes pondérées de fonctions indicatrices) sont boréliennes ;
- Les fonctions dérivées sont boréliennes (voir TD), ainsi que les primitives de fonctions mesurables.

$\rightarrow$ Exemples importants.

- Soient $A \in \mathcal{B}(\mathbb{R})$ et $f : A \to \mathbb{R}$ borélienne. Alors

  1. La fonction $g : \mathbb{R} \to \mathbb{R}$, définie par $g(x) = \mathbf{1}_A(x) f(x)$, c'est-à-dire
  $$g(x) = \begin{cases} f(x) \text{ si } x \in A, \\ 0 \text{ sinon}, \end{cases}$$
  est borélienne en tant que produit de fonctions boréliennes.

  2. Application. La fonction $h : \mathbb{R} \to \mathbb{R}$ définie par
  $$h(x) = \begin{cases} 1/x \text{ si } x > 0, \\ 0 \text{ si } x = 0, \\ 1/(1-x) \text{ si } x < 0 \end{cases}$$
  est borélienne.

2. Construction de l'intégrale de Lebesgue

Soit $(\Omega,\mathcal{A},\mu)$ un espace mesuré.

Nous allons construire, pour des fonction mesurables

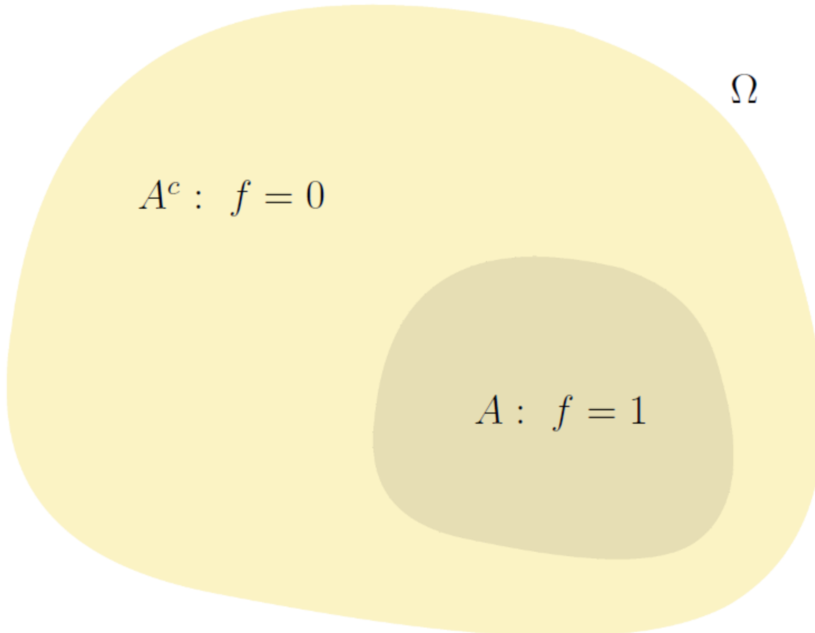
$$f : \Omega \to \mathbb{R},$$

une intégrale de f par rapport à μ.

Cette intégrale sera notée indifféremment

$$\boxed{\int_\Omega f\, d\mu = \int_\Omega f(x)\, d\mu(x) = \int_\Omega f(x)\, \mu(dx).}$$

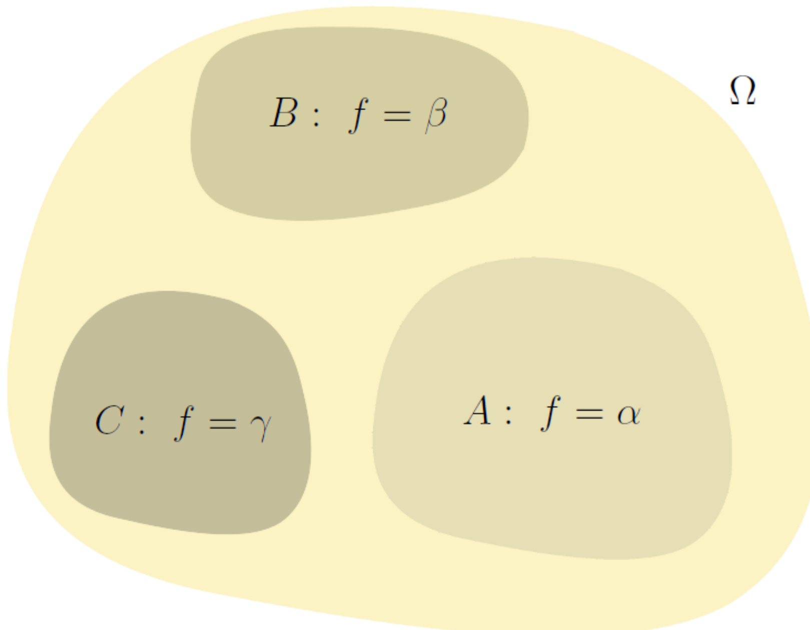
Soient $A \in \mathcal{A}$ et $f = \mathbf{1}_A$.



Dans cette situation, on pose

$$\boxed{\int_\Omega f \, d\mu := \mu(A)}$$

Soient $A, B, C \in \mathcal{A}$ et $f = \alpha\,\mathbf{1}_A + \beta\,\mathbf{1}_B + \gamma\,\mathbf{1}_C$.



Dans cette situation, on pose

$$\boxed{\int_\Omega f\,d\mu := \alpha\,\mu(A) + \beta\,\mu(B) + \gamma\,\mu(C)}$$

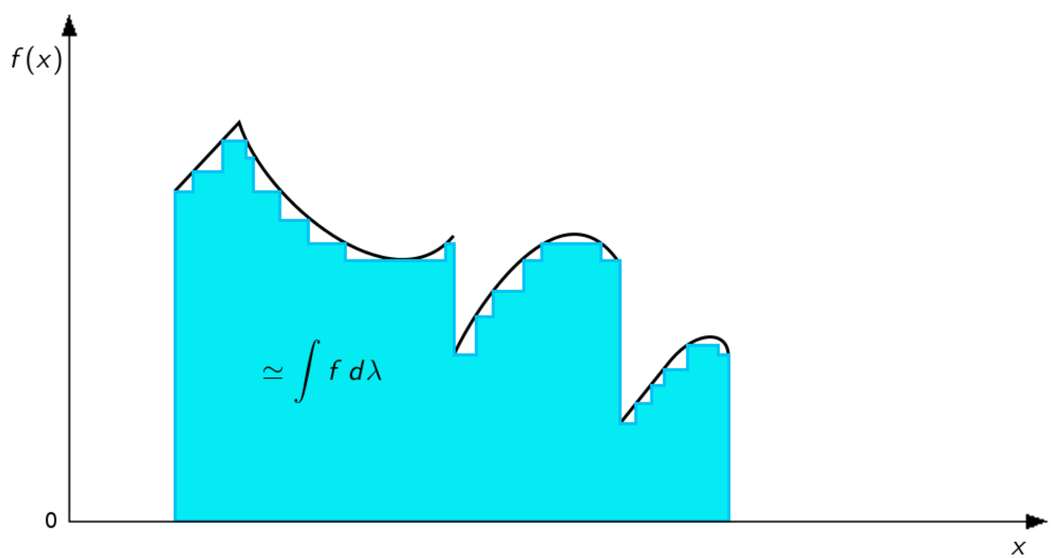
$\rightarrow$ De même pour toutes les fonctions étagées positives.

Soit $f : \Omega \to \overline{\mathbb{R}}_+$ Lebesgue mesurable et **positive**. Pour toute suite croissante de fonctions étagées positives (f_n) telles que

$$f_n(x) \xrightarrow[n\to\infty]{} f(x), \text{ pour } \mu\text{-presque tout } x,$$

on pose

$$\boxed{\int_\Omega f\,d\mu := \lim_{n\to\infty} \int_\Omega f_n d\mu.}$$



$\rightarrow$ Remarques

- La limite ne dépend pas de la suite (f_n) choisie,
- On peut toujours trouver une suite croissante de fonctions étagées positives (f_n) telles que

$$f_n(x) \xrightarrow[n \to \infty]{} f(x), \text{ pour } \mu\text{-presque tout } x,$$

- La définition suivante est équivalente

$$\int_\Omega f \, d\mu := \sup \left\{ \int_\Omega f_0 d\mu, \text{ où } f_0 \text{ étagée telle que } 0 \leq f_0 \leq f \right\},$$

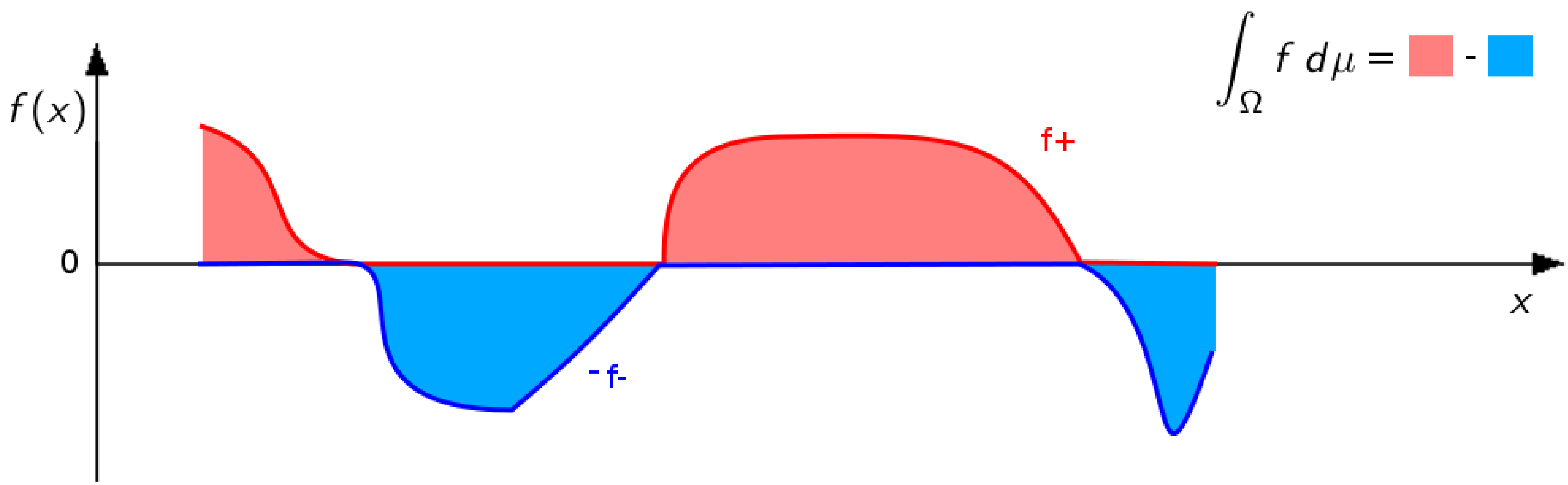
- Si f est mesurable positive, alors l'intégrale est toujours bien définie, avec éventuellement $\int_\Omega f \, d\mu = \infty$.

Soit $f : (\Omega,\mathcal{A}) \to (\overline{\mathbb{R}},\mathcal{B}(\overline{\mathbb{R}}))$ mesurable. On note f_+ sa partie positive et f_- sa partie négative. Si $\int_\Omega |f| d\mu < \infty$, c'est-à-dire si

$$\int_\Omega f_+ \, d\mu < \infty \text{ et } \int_\Omega f_- \, d\mu < \infty,$$

alors f est dite *intégrable* et on pose

$$\boxed{\int_\Omega f \, d\mu := \int_\Omega f_+ \, d\mu - \int_\Omega f_- \, d\mu.}$$



Récapitulatif

$\rightarrow$ Construction de l'intégrale de Lebesgue pour les fonctions

1. indicatrices,
2. étagées positives,
3. mesurables positives,
4. mesurables intégrables.

Critère d'intégrabilité immédiat

Une fonction mesurable f est intégrable ssi $|f|$ est intégrable. Dans ce cas,

$$\left|\int_\Omega f\, d\mu\right| \leq \int_\Omega |f|\, d\mu.$$

Si f et g sont égaux presque partout, alors f intégrable $\Leftrightarrow$ g intégrable. Le cas échéant, $\int_\Omega f d\mu = \int_\Omega g d\mu$.

Soient f,g des fonctions mesurables.

Linéarité pour les fonctions positives (resp. intégrables)

Si f et g sont positives (resp. intégrables), alors, pour tous $\alpha,\beta > 0$ (resp. $\alpha,\beta \in \mathbb{R}$),

$$\int_\Omega \alpha f + \beta g \, d\mu = \alpha \int_\Omega f \, d\mu + \beta \int_\Omega g \, d\mu.$$

Croissance de l'intégrale

Supposons $f \leq g$ (positives ou intégrables) μ-presque sûrement. Alors

$$\int_\Omega f \, d\mu \leq \int_\Omega g \, d\mu.$$

De plus, $\int_\Omega f \, d\mu = \int_\Omega g \, d\mu \Leftrightarrow f = g$ μ-presque sûrement.

En pratique dans ce cours...

$\rightarrow$ Si $\mu = \lambda_d$ sur $\mathbb{R}^d$ et f continue par morceaux $\rightarrow$ Riemann :

$$\int_{\mathbb{R}^d} f(x)\,\lambda_d(dx) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, \ldots, x_d)\,dx_1 \cdots dx_d.$$

$\rightarrow$ Si $\mu = h\,\lambda_d$ avec $h : \mathbb{R}^d \rightarrow \mathbb{R}_+$ borélienne,

$$\int_\Omega f(x)\,\mu(dx) = \int_\Omega f(x)\,h(x)\,\lambda_d(dx).$$

$\rightarrow$ Si $\mu = \sum_{n\in\mathbb{N}} p_n \delta_{a_n}$ est discrète, alors

$$\int_\Omega f(x)\,\mu(dx) = \sum_{n\in\mathbb{N}} p_n f(a_n).$$

3. Les variables aléatoires

Ensemble des possibilités

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité.

- Ω est l'*ensemble des possibles*, il permet de décrire l'ensemble des résultats possibles de l'expérience aléatoire
- $\mathbb{P}$ est une mesure de probabilité sur $(\Omega,\mathcal{F})$

$\rightarrow$ Exemples

- Lancé de 2 dés. $\Omega = \{1,2,\ldots,6\}^2$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P} = 1/36 \cdot \mathrm{Card}$.
- Lancé de n dés. $\Omega = \{1,2,\ldots,6\}^n$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P} = 1/6^n \cdot \mathrm{Card}$.
- Jeux des deux enveloppes. Soient a et $2a$ les contenus des deux enveloppes (a est une inconnu). Une possibilité est décrite par le contenu de l'enveloppe choisie puis de l'autre enveloppe, soit $\Omega = \{(a,2a),(2a,a)\}$, $\mathcal{F} = \mathcal{P}(E)$ puis $\mathbb{P} = 1/2 \cdot \mathrm{Card}$.
- Tirage uniforme sur $\Omega = [0,1]^2$, $\mathcal{F} = \mathcal{B}([0,1]^2)$, $\mathbb{P} = \lambda_1$.

Variable aléatoire

Soit $(E,\mathcal{E})$ un espace mesurable. Toute application mesurable

$$X : (\Omega,\mathcal{F}) \to (E,\mathcal{E})$$

est appelée une *variable aléatoire* (*v.a.*).

- → Si X est une v.a. à valeurs dans $\mathbb{R}^d$, alors, sauf mention contraire, $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$ ou $\mathcal{E}$ =tribu de Lebesgue.
- → Exemples (mêmes tribus que précédemment)
  - $X_1 : (d_1,d_2) \in \{1,2,\ldots,6\}^2 \mapsto d_1 + d_2$.
  - $X_2 : (d_1,d_2) \in \{1,2,\ldots,6\}^2 \mapsto d_1 - d_2$.
  - $X_3 : (d_1,\ldots,d_n) \in \{1,2,\ldots,6\}^n \mapsto \sum_{i=1}^n d_i$.
  - Gain: $(e_1,e_2) \in \{(a,2a),(2a,a)\} \mapsto e_2 - e_1$.
  - $X_4 : (x,y) \in [0,1]^2 \mapsto \min(x,y)$.

Soient $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité et $X : \Omega \to E$ une v.a.

Loi d'une variable aléatoire

La *loi de* X est la mesure de probabilité $\mathbb{P}_X$ sur $(E,\mathcal{E})$ définie par

$$\begin{aligned}\mathbb{P}_X(A) = \mathbb{P}(X \in A) &:= \mathbb{P}(\{\omega \in \Omega,\ X(\omega) \in A\})\\ &= \mathbb{P}(X^{-1}(A)),\ \forall A \in \mathcal{E}.\end{aligned}$$

- $\to$ $\mathbb{P}_X$ est la mesure image de $\mathbb{P}$ par X.
- $\to$ Exercice
  1. Soient $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$, $\mathbb{P} = \lambda_1([0,1])$ et $p \in [0,1]$.
     Déterminer la loi de $X : x \mapsto \mathbf{1}_{x<p}$.
     $\to$ Solution : déterminons $\mathbb{P}_X(0)$ et $\mathbb{P}_X(1)$.
     $\mathbb{P}_X(0) = \mathbb{P}(X^{-1}(0)) = \mathbb{P}([p,1]) = \lambda_1([p,1]) = 1 - p$,
     $\mathbb{P}_X(1) = \mathbb{P}(X^{-1}(1)) = \mathbb{P}([0,p[) = \lambda_1([0,p[) = p$,
     donc $\mathbb{P}_X$ est une probabilité discrète et $\mathbb{P}_X = (1-p)\delta_0 + p\delta_1$.

4. Lois discrètes

Définition et propriété

La loi d'une variable aléatoire X est dite *discrète* si il existe un ensemble K fini ou dénombrable tel que

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

Dans ce cas,

$$\mathbb{P}_X = \sum_{n \in K} \mathbb{P}(X = n)\delta_n.$$

$\rightarrow$ Remarque

- Calculer la loi d'une v.a. discrète X revient donc à calculer

$$\mathbb{P}(X = n), \forall n \in K.$$

Quelques lois discrètes classiques

On dit qu'une variable aléatoire X suit une...

- Loi de Dirac en $x \in E$ si et seulement si $\mathbb{P}(X = x) = 1$.
  $\rightarrow$ pour modéliser un phénomène déterministe.
- Loi uniforme discrète sur un ensemble fini K si et seulement si

$$\mathbb{P}(X = x) = \frac{1}{Card(K)}, \forall x \in K.$$

  $\rightarrow$ pour modéliser un phénomène dont les résultats sont équiprobables (lancer de dés non biaisés).
- Loi de Bernouilli de paramètre $p \in [0,1]$ si et seulement si

$$\mathbb{P}(X = 0) = 1 - p \text{ et } \mathbb{P}(X = 1) = p.$$

  $\rightarrow$ pour modéliser une expérience à deux issues possibles (succès/échec, pile ou face biaisé).

- Loi binomiale de paramètre $n \in \mathbb{N}$ et $p \in [0,1]$ si et seulement si
$$\mathbb{P}(X = k) = C_n^k (1-p)^{n-k} p^k, \forall k \in \{0,1,\ldots,n\}.$$
$\rightarrow$ pour modéliser le nombre de succès obtenus à l'issue de n épreuves de Bernoulli indépendantes de paramètre p (nombre de tirs atteignant une cible après n essais).
- Loi géométrique de paramètre $p \in]0,1]$ si et seulement si
$$\mathbb{P}(X = k) = (1-p)^{k-1} p, \forall k \in \mathbb{N}.$$
$\rightarrow$ pour modéliser le premier temps de succès dans une suite de tirage de variable aléatoire de Bernouilli indépendantes (premier tir atteignant une cible).
- Loi de Poisson de paramètre $\lambda > 0$ si et seulement si
$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \forall k \in \mathbb{N}.$$
$\rightarrow$ pour modéliser le nombre d'événements se produisant dans un intervalle de temps fixé (nombre de connexion à un site pendant une heure), λ étant leur moyenne.

5. Lois absolument continues

Définition

La loi d'une variable aléatoire X est dite *absolument continue* si il existe $f_X : \mathbb{R}^d \to [0, +\infty]$ borélienne, appelée la *densité* de X, telle que

$$\forall A \in \mathcal{B}(\mathbb{R}^d),\ \mathbb{P}_X(A) = \int_A f_X(u)\, d\lambda_d(u).$$

- Si f_X et g_X densité de X, alors $f_X = g_X$ λ_d-presque partout.
- On a $\int_{\mathbb{R}^d} f_X(u) d\lambda_d(u) = \mathbb{P}(X \in \mathbb{R}^d) = 1$.
- Si $f_X = f_Y$ presque partout, alors $\mathbb{P}_X = \mathbb{P}_Y$.
- **Notations.** $d\mathbb{P}_X = f_X d\lambda_d$ ou $\mathbb{P}_X(du) = f_X(u)\lambda_d(du)$.

Quelques lois absolument continues classiques

La variable aléatoire X suit une...

→ Loi uniforme (1d) sur l'intervalle $[a,b]$ ($a < b$) si et seulement si sa densité est définie par

$$f_X(u) = \frac{1}{b-a}\mathbf{1}_{[a,b]}(u) \ \forall u \in \mathbb{R}.$$

→ Loi uniforme (multi-d) sur le pavé $[a_1,b_1] \times \ldots \times [a_d,b_d]$ si et seulement si sa densité est définie par

$$f_X(u) = \frac{1}{\prod_{i=1}^{d}(b_i - a_i)}\mathbf{1}_{[a_1,b_1]\times\ldots\times[a_d,b_d]}(u) \ \forall u \in \mathbb{R}^d.$$

→ Loi gaussienne $\mathcal{N}(m,\sigma^2)$ sur $\mathbb{R}$ si et seulement si sa densité est définie par

$$f_X(u) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(u-m)^2}{\sigma^2}\right).$$

6. Espérance et calculs de loi

Soient $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité et $X : \Omega \to \mathbb{R}^d$ une v.a.

Espérance d'une variable aléatoire

La variable aléatoire X est dite *intégrable* si $\int_\Omega |X(\omega)|\, d\mathbb{P}(\omega) < \infty$.
Si X est intégrable ou positive, on appelle *espérance de* X la quantité

$$\boxed{\mathbb{E}(X) := \int_\Omega X(\omega)\, d\mathbb{P}(\omega).}$$

- $\to$ Si $X \geq 0$ presque sûrement, alors $\mathbb{E}(X)$ est bien défini.
- $\to$ Si $\exists a \geq 0$ tel que $\mathbb{P}(|X| < a) = 1$, alors X est intégrable.
- $\to$ Exercices
  1. Calculer l'espérance de Gain (jeu des 2 enveloppes).
  2. Calculer l'espérance de X_1, de loi λ_1 sur $[0,1]$.

Théorème du transport

Soit $\varphi : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ borélienne.

- Les propriétés suivantes sont équivalentes
  a. $\varphi(X)$ est une variable aléatoire intégrable,
  b. φ est $\mathbb{P}_X$-intégrable.
- Dans ce cas (ou si φ positive), alors

$$\boxed{\mathbb{E}\left(\varphi(X)\right) = \int_{\mathbb{R}^d} \varphi(u)\, \mathbb{P}_X(du).}$$

$\to$ Calcul de $\mathbb{E}(\varphi(X))$ directement à partir de $\mathbb{P}_X$ (la loi de X).

$\to$ Si X est absolument continue, $\mathbb{P}_X(du) = f_X(u)\lambda_d(du)$:

$$\mathbb{E}\left(\varphi(X)\right) = \int_{\mathbb{R}^d} \varphi(u)\, f_X(u)\lambda_d(du).$$

Exercice 1. Soit X une v.a. de loi binomiale de paramètre (n,p), i.e.

$$\mathbb{P}_X = \sum_{k=0}^{n} C_n^k (1-p)^{n-k} p^k \delta_k.$$

La variable aléatoire $\sinh(X)$ est-elle intégrable ? Le cas échéant, exprimer son intégrale.

Solution. D'après le théorème du transport,

$$\mathbb{E}(|\sinh(X)|) = \int_{\mathbb{R}} |\sinh(x)| \mathbb{P}_X(dx) = \sum_{k=0}^{n} C_n^k (1-p)^{n-k} p^k |\sinh(k)|$$

donc $\mathbb{E}(|\sinh(X)|) < \infty$ et $\boxed{\sinh(X) \text{ est donc intégrable}}$. De plus,

$$\boxed{\mathbb{E}(\sinh(X)) = \int_{\mathbb{R}} \sinh(x) \mathbb{P}_X(dx) = \sum_{k=0}^{n} C_n^k (1-p)^{n-k} p^k \sinh(k).}$$

Exercice 2. Soit X une v.a. de loi gaussienne centrée $\mathcal{N}(0,1)$, i.e.

$$\mathbb{P}_X(du) = \frac{e^{-u^2/2}}{\sqrt{2\pi}}\,\lambda_1(du).$$

La variable aléatoire $\sinh(X)$ est-elle intégrable ? Le cas échéant, exprimer son intégrale.

Solution. D'après le théorème du transport,

$$\mathbb{E}(|\sinh(X)|) \leq \mathbb{E}(|e^{|X|}|) = \int_{\mathbb{R}} |e^{|u|}|\mathbb{P}_X(du) = \int_{\mathbb{R}} e^{|u|}\,\frac{e^{-u^2/2}}{\sqrt{2\pi}}\,\lambda_1(du)$$

donc $\mathbb{E}(|\sinh(X)|) < \infty$ et $\boxed{\sinh(X) \text{ est donc intégrable}}$. De plus,

$$\boxed{\mathbb{E}(\sinh(X)) = \int_{\mathbb{R}} \sinh(u)\mathbb{P}_X(du) = \int_{\mathbb{R}} \sinh(u)\frac{e^{-u^2/2}}{\sqrt{2\pi}}\,\lambda_1(du).}$$

Caractérisation d'une loi

Deux variables aléatoires X, Y à valeurs dans $\mathbb{R}^d$ ont même loi si et seulement si

$$\mathbb{E}(\varphi(X)) = \mathbb{E}(\varphi(Y))$$

pour toute fonction $\varphi : \mathbb{R}^d \to \mathbb{R}$ <u>borélienne</u>

ou

pour toute fonction $\varphi : \mathbb{R}^d \to \mathbb{R}$ <u>borélienne bornée</u>

ou

pour toute fonction $\varphi : \mathbb{R}^d \to \mathbb{R}$ <u>borélienne $\mathbb{P}_X$ et $\mathbb{P}_Y$ intégrable</u>

ou

pour toute fonction $\varphi : \mathbb{R}^d \to \mathbb{R}$ <u>continue bornée</u>.

Exercice 3. Soient $X \sim \mathcal{N}(0,1)$ et $Y \sim \mathcal{N}(0,\sigma^2)$. Montrer que σX et Y ont la même loi.

Solution. On a $f_X(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$ et $f_Y(u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2}$.

Pour toute fonction φ mesurable bornée, $\varphi(\sigma X)$ et $\varphi(Y)$ sont intégrables car bornées p.s et (cdv $v = \sigma u$)

$$\begin{aligned}
\mathbb{E}(\varphi(\sigma X)) &= \int_{\mathbb{R}} \varphi(\sigma u) \frac{1}{\sqrt{2\pi}} e^{-u^2/2} d\lambda_1(u) \\
&= \int_{\mathbb{R}} \varphi(v) \frac{1}{\sigma\sqrt{2\pi}} e^{-v^2/2\sigma^2} d\lambda_1(v) \\
&= \mathbb{E}(\varphi(Y)).
\end{aligned}$$

Ceci étant vrai pour toute fonction φ mesurable bornée, on en déduit que les lois de X et Y sont identiques.

Exercice 4. Soit X une v.a. de loi exponentielle de paramètre 1, i.e.

$$\mathbb{P}_X(du) = e^{-u}\mathbf{1}_{u>0}\,\lambda_1(du), \quad \text{soit } f_X(u) = e^{-u}\mathbf{1}_{u>0}.$$

Calculer la loi de $Y = \exp(X)$.

Solution. Pour toute fonction φ mesurable bornée, $\varphi(Y)$ est intégrable car bornée p.s et (cdv $v = \exp(u)$)

$$\begin{aligned}
\mathbb{E}(\varphi(Y)) = \mathbb{E}(\varphi(\exp(X))) &= \int_{\mathbb{R}} \varphi(\exp(u))e^{-u}\mathbf{1}_{u>0}\lambda_1(du) \\
&= \int_{]0,+\infty[} \varphi(\exp(u))e^{-u}\lambda_1(du) \\
&= \int_{]1,+\infty[} \varphi(v)\frac{1}{v^2}\,\lambda_1(dv).
\end{aligned}$$

Cette relation étant vraie pour toute fonction mesurable positive bornée φ, on en déduit que

$$\boxed{\mathbb{P}_Y(dv) = \frac{1}{v^2}\mathbf{1}_{v>1}\lambda_1(dv), \quad \text{soit } f_Y(v) = \frac{1}{v^2}\mathbf{1}_{v>1}.}$$

Exercice 5. Soit X une v.a. de loi exponentielle de paramètre 1, i.e.

$$\mathbb{P}_X(du) = e^{-u}\mathbf{1}_{u>0}\,\lambda_1(du), \quad \text{soit } f_X(u) = e^{-u}\mathbf{1}_{u>0}.$$

Calculer la loi de Z =partie entière de $\exp(X)$.

Solution. Z est à valeurs dans $\mathbb{N}^*$ p.s. (donc discrète) et, $\forall n \in \mathbb{N}^*$,

$$\begin{aligned}\mathbb{P}(Z = n) &= \mathbb{P}(Y \in [n,n+1[) = \mathbb{E}(\mathbf{1}_{Y\in[n,n+1[})\\ &= \int_{\mathbb{R}} \mathbf{1}_{u\in[n,n+1[}\, f_Y(u)\,\lambda_1(du)\\ &= \int_{[n,n+1[} \frac{1}{v^2}\,\lambda_1(du) = \frac{1}{n(n+1)}.\end{aligned}$$

Donc $\boxed{\mathbb{P}_Z = \sum_{n=1}^{+\infty} \frac{1}{n(n+1)}\,\delta_n.}$

Probabilités
3. Théorèmes d'intégration

Rappels de la séance précédente

Rappel I. Loi d'une variable aléatoire

Soient $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité et $X : \Omega \to \mathbb{R}^d$ une variable aléatoire.
On appelle loi de X la mesure de probabilité $\mathbb{P}_X$ sur $\mathbb{R}^d$ définie par

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A).$$

Par exemple :

a. Les lois discrètes : $\mathbb{P}_X(dx) = \sum_{n \in K} p_n \, \delta_n(dx)$
b. Les lois absolument continues : $\mathbb{P}_X(dx) = f_X(x) \, \lambda_d(dx)$
c. Les lois mixtes : $\mathbb{P}_X(dx) = \sum_{n \in K} p_n \, \delta_n(dx) + f_X(x) \, \lambda_d(dx)$.

Attention ! Il existe des lois qui ne rentrent pas dans ce cadre.

Rappel II. Intégrabilité

Soient $(\Omega,\mathcal{F},\mu)$ un espace mesuré et $f : \Omega \to \mathbb{R}$ mesurable.

1. Si f est positive, alors $\boxed{\displaystyle\int_\Omega f\, d\mu \in [0,+\infty] \text{ est bien définie}}$
2. Si $\int_\Omega |f|\, d\mu < \infty$, alors

$$\boxed{f \text{ est dite } \textbf{intégrable} \text{ et } \int_\Omega f\, d\mu \in]-\infty,\infty[\text{ est bien définie}}$$

Soient $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité et $Y : \Omega \to \mathbb{R}$ une variable aléatoire à valeurs réelles.

1. Si Y est positive p.s, alors $\boxed{\mathbb{E}(Y) \in [0,+\infty] \text{ est bien définie}}$
2. Si $\mathbb{E}(|Y|) < \infty$, alors

$$\boxed{Y \text{ est dite } \textbf{intégrable} \text{ et } \mathbb{E}(Y) \in]-\infty,\infty[\text{ est bien définie}}$$

Rappel III. Calcul d'espérances

Soient $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité et $X : \Omega \to E$ une variable aléatoire de loi $\mathbb{P}_X$.

Théorème du transport

Pour toute fonction $\varphi : E \to \mathbb{R}$ telle que $\varphi(X)$ est positive p.s ou intégrable,

$$\mathbb{E}(\varphi(X)) = \int_E \varphi(x)\,\mathbb{P}_X(dx)$$

Remarque. Pour tout $A \subset \mathbb{R}^d$ mesurable,

$$\mathbb{P}(X \in A) = \mathbb{P}_X(A) = \int_E \mathbf{1}_A(x)\,\mathbb{P}_X(dx) = \mathbb{E}(\mathbf{1}_A(X)).$$

a. Variables aléatoires de loi discrète

Soit X une variable aléatoire de loi discrète

$$\mathbb{P}_X = \sum_{n \in K} p_n \, \delta_n.$$

Alors, pour tout $A \in \mathcal{B}(\mathbb{R}^d)$, on a par définition

$$\mathbb{P}(X \in A) = \mathbb{P}_X(A) = \sum_{n \in K} p_n \mathbf{1}_{n \in A}.$$

De plus, pour toute fonction φ mesurable positive,

$$\boxed{\mathbb{E}(\varphi(X)) = \int_E \varphi(x) \, \mathbb{P}_X(dx) = \sum_{n \in K} p_n \, \varphi(n).}$$

b. Variables aléatoires de loi absolument continue

Soit X une variable aléatoire de loi abs. continue dans $E = \mathbb{R}^d$

$$\mathbb{P}_X(dx) = f_X(x)\,\lambda_d(dx).$$

Alors, pour tout $A \in \mathcal{B}(\mathbb{R}^d)$, on a par définition

$$\mathbb{P}(X \in A) = \mathbb{P}_X(A) = \int_{\mathbb{R}^d} \mathbf{1}_A(x) f_X(x)\,\lambda_d(dx).$$

De plus, pour toute fonction φ mesurable positive,

$$\boxed{\mathbb{E}(\varphi(X)) = \int_{\mathbb{R}^d} \varphi(x)\,\mathbb{P}_X(dx) = \int_{\mathbb{R}^d} \varphi(x) f_X(x)\,\lambda_d(dx).}$$

c. Variable aléatoire de loi mixte

Soit X une variable aléatoire de loi mixte dans $E = \mathbb{R}^d$

$$\mathbb{P}_X(dx) = \sum_{n \in K} p_n \, \delta_n(dx) + f_X(x) \, \lambda_d(dx).$$

Alors, pour tout $A \in \mathcal{B}(\mathbb{R}^d)$, on a

$$\mathbb{P}(X \in A) = \mathbb{P}_X(A) = \sum_{n \in K} p_n \mathbf{1}_{n \in A} + \int_{\mathbb{R}^d} \mathbf{1}_A \, f_X(x) \, \lambda_d(dx).$$

De plus, pour toute fonction φ mesurable positive,

$$\boxed{\mathbb{E}(\varphi(X)) = \int_{\mathbb{R}^d} \varphi(x) \, \mathbb{P}_X(dx) = \sum_{n \in K} p_n \varphi(n) + \int_{\mathbb{R}^d} \varphi(x) f_X(x) \, \lambda_d(dx)}$$

Si φ n'est pas positive...

Soient X une variable aléatoire de loi $\mathbb{P}_X$ et $\varphi(X)$ de signe quelconque. On souhaite calculer $\mathbb{E}(\varphi(X))$.

1. La variable aléatoire $|\varphi(X)|$ est positive donc
$$\mathbb{E}(|\varphi(X)|) = \int_E |\varphi(x)| \, \mathbb{P}_X(dx).$$

2. Deux situations :
   a. Si $\mathbb{E}(|\varphi(X)|) = \infty$, on s'arrête là car $\mathbb{E}(\varphi(X))$ indéfinie
   b. Si $\mathbb{E}(|\varphi(X)|) < \infty$, alors
$$\boxed{\mathbb{E}(\varphi(X)) = \int_E \varphi(x) \, \mathbb{P}_X(dx)}$$

   avec les mêmes formules que précédemment.

Rappel IV. Calcul de loi

Soit X une variable aléatoire dont on ne connait pas la loi, mais dont on sait exprimer $\mathbb{E}(\varphi(X))$, pour toute fonction $\varphi : \mathbb{R}^d \to \mathbb{R}$ mesurable positive bornée, sous la forme

$$\mathbb{E}(\varphi(X)) = \int_E \varphi(x)\,\nu(dx),$$

où ν est une mesure $\to$ ν est alors la loi de X.

Typiquement : $X = f(Z)$, avec Z de loi $\mathbb{P}_Z$ **connue**. On écrit alors

$$\mathbb{E}(\varphi(X)) = \mathbb{E}(\varphi(f(Z))) = \int_E \varphi(f(z))\,\mathbb{P}_Z(dz) \qquad \textit{(Th. du transport).}$$

Après cdv, factorisation, identification... on se ramène à

$$\mathbb{E}(\varphi(X)) = \int_E \varphi(x)\,\nu(dx),$$

pour toute fonction φ positive mesurable bornée. On en déduit que

$$\boxed{\mathbb{P}_X = \nu.}$$

Exemple. Soit Z une variable aléatoire de loi

$$\mathbb{P}_Z(dz) = \frac{1}{2}\delta_9(dz) + \mathbf{1}_{z\in[1,+\infty[}\frac{1}{2z^2}\,\lambda_1(dz).$$

On pose $X = \sqrt{Z}$. Quelle est la loi de X ?

Réponse. Pour toute fonction φ mesurable positive bornée,

$$\begin{aligned}\mathbb{E}(\varphi(X)) &= \mathbb{E}(\varphi(\sqrt{Z}))\\ &\overset{\text{Th. transp.}}{=} \frac{1}{2}\varphi(\sqrt{9}) + \int_{\mathbb{R}} \varphi(\sqrt{z})\mathbf{1}_{z\in[1,+\infty[}\frac{1}{2z^2}\,d\lambda_1(z).\end{aligned}$$

Donc, par changement de variable,

$$\mathbb{E}(\varphi(X)) = \frac{1}{2}\varphi(3) + \int_{]1,+\infty[} \varphi(x)\frac{1}{x^3}\,d\lambda_1(x) = \int_{\mathbb{R}} \varphi(x)\,\mathbb{P}_X(dx),$$

avec

$$\boxed{\mathbb{P}_X(dx) = \frac{1}{2}\delta_3(dx) + \mathbf{1}_{x\in]1,+\infty[}\frac{1}{x^3}\,\lambda_1(dx)}$$

Chapitre 3. Théorèmes d'intégration

1. Exposé du problème

Soient $(\Omega,\mathcal{F},\mu)$ un espace mesuré, f une fonction mesurable et $(f_n)_n$ une suite de fonctions mesurables à valeurs dans $\bar{\mathbb{R}}$ telles que

$$f(x) = \lim_{n\to\infty} f_n(x), \text{ pour } \mu\text{-presque tout } x \in \Omega.$$

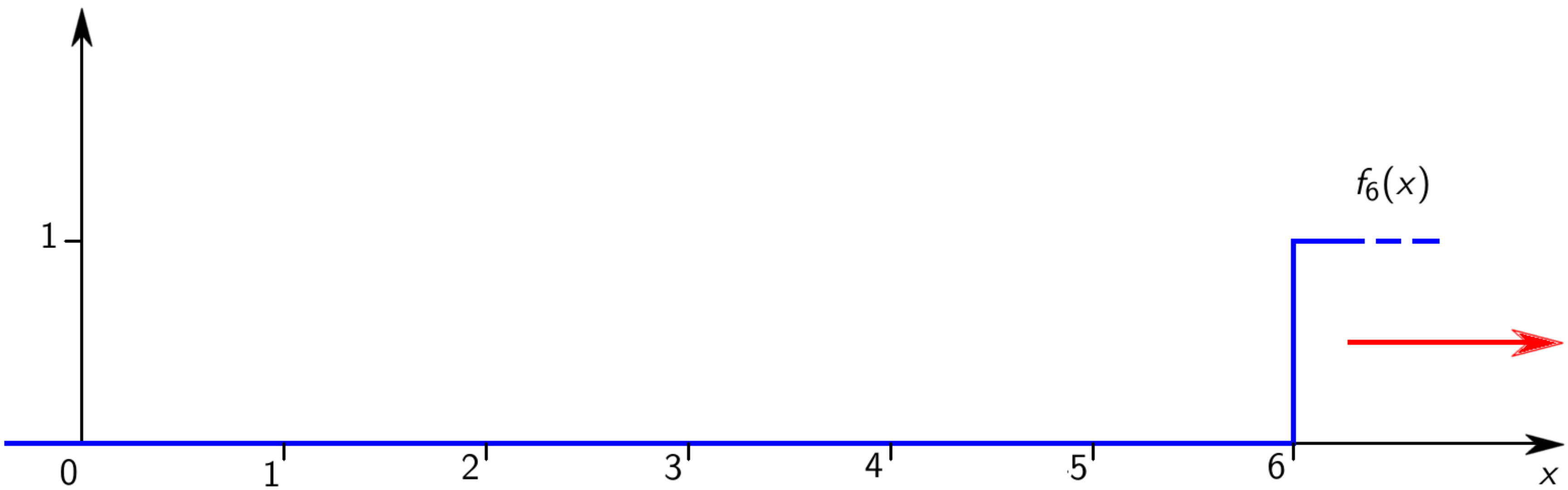
Question. Peut-on écrire

$$\int_\Omega f(x)\,\mu(dx) = \boxed{\int_\Omega \lim_{n\to\infty} f_n(x)\,\mu(dx) = \lim_{n\to\infty} \int_\Omega f_n(x)\,\mu(dx) \ ?}$$

Autrement dit : Peut-on intervertir les signes $\lim$ et $\int_\Omega$?
Peut-on intervertir les signes $\lim$ et $\mathbb{E}$?

Ex 1. Contre-exemple sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. On pose, pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$,

$$f_n(x) = \mathbf{1}_{x \in [n, n+1]} = \begin{cases} 1 \text{ si } x \in [n, n+1], \\ 0 \text{ sinon.} \end{cases}$$



On obtient, pour tout $x \in \mathbb{R}$,

$$f(x) := \lim_{n \to \infty} f_n(x) = 0$$

Rappelons que, pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$,

$$f_n(x) = \mathbf{1}_{x \in [n,n+1]}.$$

Ainsi, pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$,

$$\boxed{\int_{\mathbb{R}} \lim_{n\to\infty} f_n(x)\,\lambda(dx) = \int_{\mathbb{R}} f(x)\,\lambda(dx) = \int_{\mathbb{R}} 0\,\lambda(dx) = 0}$$

et

$$\boxed{\int_{\mathbb{R}} f_n(x)\,\lambda(dx) = \int_{\mathbb{R}} \mathbf{1}_{x\in[n,n+1]}\,\lambda(dx) = \lambda([n,n+1]) = 1}$$

En définitive,

$$\boxed{\int_{\mathbb{R}} \lim_{n\to\infty} f_n(x)\,\lambda(dx) = 0 \neq \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\,\lambda(dx) = 1.}$$

2. Théorème de convergence monotone

Soient $(\Omega,\mathcal{F},\mu)$ un espace mesuré, f une fonction mesurable et $(f_n)_n$ une suite de fonctions mesurables de $(\Omega,\mathcal{F})$ dans $(\overline{\mathbb{R}},\mathcal{B}(\overline{\mathbb{R}}))$.

Théorème de convergence monotone (version $\int$)

Si, pour μ-presque tout x, $(f_n(x))_n$ est une suite positive et croissante, alors

$$\int_\Omega \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_\Omega f_n \, d\mu$$

$\rightarrow$ Remarques

- Aucune hypothèse d'intégrabilité sur les f_n.
- D'après les hypothèses, $\lim_{n\to\infty} f_n(x)$ est bien définie pour μ-presque tout x.

Soient $(\Omega,\mathcal{F},\mathbb{P})$ un espace probabilisé, X une variable aléatoire et $(X_n)_n$ une suite de variables aléatoires à valeurs réelles.

Théorème de convergence monotone (version $\mathbb{E}$)

Si $(X_n)_n$ est positive, croissante presque sûrement,alors

$$\mathbb{E}\left(\lim_{n\to\infty} X_n\right) = \lim_{n\to\infty} \mathbb{E}(X_n).$$

$\to$ Remarques

- D'après les hypothèses, $\lim_{n\to\infty} X_n$ est une variable aléatoire bien définie presque sûrement.
- Preuve : on se ramène à une intégrale :

$$\begin{aligned}
\mathbb{E}\left(\lim_{n\to\infty} X_n\right) &= \int_\Omega \lim_{n\to\infty} X_n(\omega)\,\mathbb{P}(d\omega) \\
&= \lim_{n\to\infty} \int_\Omega X_n(\omega)\,\mathbb{P}(d\omega) \\
&= \lim_{n\to\infty} \mathbb{E}(X_n) \quad \to \square
\end{aligned}$$

Ex 2. Calcul de la limite d'une espérance. Soit X une variable aléatoire à valeurs strictement positives. Montrer que

$$\lim_{n\to\infty} \mathbb{E}\left(\frac{1}{\sqrt{X+\frac{1}{n}}}\right) = \mathbb{E}\left(\frac{1}{\sqrt{X}}\right).$$

Solution. La suite de variables aléatoires $X_n = 1/\sqrt{X+1/n}$ est <u>positive et croissante</u> presque-sûrement, donc, d'après le théorème de convergence monotone,

$$\lim_{n\to\infty} \mathbb{E}\left(\frac{1}{\sqrt{X+\frac{1}{n}}}\right) = \mathbb{E}\left(\lim_{n\to\infty}\frac{1}{\sqrt{X+\frac{1}{n}}}\right).$$

Or, presque sûrement,

$$\lim_{n\to\infty} \frac{1}{\sqrt{X+\frac{1}{n}}} = \frac{1}{\sqrt{X}},$$ ce qui permet de conclure.

3. Théorème de convergence dominée

Théorème de convergence dominée (Version $\int$)

Si il existe $g : \Omega \to \mathbb{R}$ intégrable telle que

1. $|f_n(x)| \leq |g(x)|$, $\quad \mu(x)$-presque partout, $\forall n \geq 0$,
2. $f_n(x) \xrightarrow[n\to\infty]{} f(x)$, $\quad \mu(x)$-presque partout,

alors f est intégrable et

$$\int_\Omega f \, d\mu = \int_\Omega \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_\Omega f_n \, d\mu$$

$\rightarrow$ Remarques

- On a de plus $\int_\Omega |f| d\mu \leq \int_\Omega |g| d\mu$, donc f intégrable.
- Autre cas de convergence : lemme de Scheffé (TD 3).

Théorème de convergence dominée (Version $\mathbb{E}$)

Si il existe une variable aléatoire Y intégrable telle que

1. $|X_n| \leq |Y|$, presque sûrement, $\forall n \geq 0$,,
2. $X_n \xrightarrow[n\to\infty]{} X$, presque sûrement,

alors X est intégrable et

$$\mathbb{E}(X) = \mathbb{E}\left(\lim_{n\to\infty} X_n\right) = \lim_{n\to\infty} \mathbb{E}(X_n).$$

$\rightarrow$ Remarque

- On a de plus $\mathbb{E}(|X|) \leq \mathbb{E}(|Y|)$,
- Cas de convergence des densités : voir le lemme de Scheffé.

Ex 3. Convergence en loi. Soit X une v.a. à valeurs réelles. Montrer que, pour toute fonction continue bornée $h : \mathbb{R} \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}\left(h\left(X + \frac{1}{n}\right)\right) = \mathbb{E}(h(X)).$$

Solution. La fonction h est uniformément bornée, donc

$\left|h\left(X + \frac{1}{n}\right)\right| \leq Y := \|h\|_\infty < \infty$, où Y est intégrable car borné.

De plus, h est continue, donc

$$h\left(X + \frac{1}{n}\right) \xrightarrow[n\to\infty]{} h(X) \quad \text{presque sûrement.}$$

Le théorème de convergence dominée permet de conclure.

4. Lemme de Fatou

Lemme de Fatou (version $\int$ et $\mathbb{E}$)

Si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions <u>positives</u> μ presque partout, alors

$$\int_\Omega \liminf_{n\to\infty} f_n(x)\,\mu(dx) \leq \liminf_{n\to\infty} \int_\Omega f_n(x)\,\mu(dx)$$

Si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires <u>à valeurs positives</u> presque sûrement, alors

$$\mathbb{E}\left(\liminf_{n\to\infty} X_n\right) \leq \liminf_{n\to\infty} \mathbb{E}(X_n).$$

$\rightarrow$ Remarque

- On ne suppose pas que (f_n) ou (X_n) converge.
- Exemple d'inégalité stricte vu en début de séance.
- Exemple d'application : voir TD 1 question 2.

5. Continuité et dérivabilité sous le signe somme

Soient E métrique (par ex. E ouvert de $\mathbb{R}^d$) et $(\Omega,\mathcal{F},\mu)$ mesuré.

Continuité

Soit $f : E \times \Omega \to \mathbb{R}$ telle que

- $\omega \mapsto f(x,\omega)$ est mesurable pour tout $x \in E$,
- $x \mapsto f(x,\omega)$ est continue pour μ presque tout $\omega \in \Omega$,
- $\exists g : \Omega \to \mathbb{R}$ intégrable telle que

$$|f(x,\omega)| \leq |g(\omega)|,\ \mu(\omega) - pp, \forall x \in E.$$

Alors la fonction

$$x \in E \mapsto \int_\Omega f(x,\omega)\,\mu(d\omega) \text{ est continue.}$$

Soient E ouvert de $\mathbb{R}$ et $(\Omega,\mathcal{F},\mu)$ mesuré.

Dérivabilité

Soit $f : E \times \Omega \to \mathbb{R}$ telle que

- $\omega \mapsto f(x,\omega)$ est mesurable et μ-intégrable pour tout $x \in E$,
- $x \mapsto f(x,\omega)$ est dérivable pour μ presque tout $\omega \in \Omega$,
- $\exists g : \Omega \to \mathbb{R}$ intégrable telle que

$$\left|\frac{\partial f}{\partial x}(x,\omega)\right| \leq |g(\omega)|,\ \mu\text{-pp}, \forall x \in E.$$

Alors $F : x \in E \mapsto \int_\Omega f(x,\omega)\mu(d\omega)$ est dérivable en x et

$$F'(x) = \int_\Omega \frac{\partial f}{\partial x}(x,\omega)\mu(d\omega).$$

$\rightarrow$ Remarque

- On peut considérer $E \subset \mathbb{R}^d$ et $\partial f/\partial x_i$.

Ex 4. Dérivation de la transformée de Laplace. Soit X une variable aléatoire à valeurs positives telle que $\mathbb{E}(Xe^X) < \infty$. Montrer que l'application

$$\varphi(t) : t \in]-\infty, 1[\mapsto \mathbb{E}(e^{tX}) \text{ est dérivable.}$$

Solution.

- Pour tout $t \in]-\infty, 1[$, e^{tX} est une variable aléatoire.
- Presque sûrement, la fonction $t \mapsto e^{tX}$ est dérivable.
- On a, pour tout $t \in]-\infty, 1[$,

$$\left| \frac{\partial e^{tX}}{\partial t} \right| = |Xe^{tX}| \leq Xe^X \text{ ps, où } Xe^X \text{ est intégrable par hyp.}$$

D'après le théorème de dérivation sous le signe somme, φ est dérivable sur $]-\infty, 1[$ et

$$\boxed{\frac{\partial \varphi(t)}{\partial t} = \mathbb{E}(Xe^{tX}).}$$

Probabilités
4. Fonction de répartition et fonction caractéristique

´

Continuité/Dérivabilité sous le signe $\mathbb{E}$

Soient E métrique (par ex. E ouvert de $\mathbb{R}^d$) et X une v.a. à valeurs dans un espace mesurable $(F,\mathcal{F})$.

Continuité

Soit $f : E \times F \to \mathbb{R}$ telle que

- $f(x,X)$ est une variable aléatoire pour tout $x \in E$,
- $x \mapsto f(x,X)$ est continue presque sûrement,
- $\exists Y$ une variable aléatoire intégrable telle que

$$|f(x,X)| \leq |Y|, \text{ p.s}, \forall x \in E.$$

Alors la fonction

$$x \in E \mapsto \mathbb{E}(f(x,X)) \text{ est continue.}$$

Soient E ouvert de $\mathbb{R}$ et X une v.a. à valeurs dans un espace mesurable $(F,\mathcal{F})$.

Dérivabilité

Soit $f : E \times F \to \mathbb{R}$ telle que

- $f(x,X)$ est une variable aléatoire intégrable pour tout $x \in E$,
- $x \mapsto f(x,X)$ est dérivable presque sûrement,
- $\exists Y$ une variable aléatoire intégrable telle que

$$\left|\frac{\partial f}{\partial x}(x,X)\right| \leq |Y|, \text{ p.s}, \forall x \in E.$$

Alors $F : x \in E \mapsto \mathbb{E}(f(x,X))$ est dérivable en x et

$$F'(x) = \mathbb{E}\left(\frac{\partial f}{\partial x}(x,X)\right).$$

$\rightarrow$ Preuve : appliquer le théorème avec intégrales à

- $F(x) = \int_\Omega f(x,X(\omega))\,\mathbb{P}(d\omega)$ ou $F(x) = \int_F f(x,u)\,\mathbb{P}_X(du)$.

Chapitre 4. Fonction de répartition et fonction caractéristique

Fonction de répartition : définition et généralités

Définition

Si $X : \Omega \to \mathbb{R}$ est une variable aléatoire réelle alors, sa **fonction de répartition** est la fonction F_X définie par

$$\forall t \in \mathbb{R}, \boxed{F_X(t) = \mathbb{P}(X \leq t)} = \mathbb{P}_X(]-\infty, t]) = \int_{]-\infty,t]} 1 \, \mathbb{P}_X(dx).$$

Théorème de caractérisation

Deux variables aléatoires X et Y ont la même loi si et seulement si elles ont la même fonction de répartition, i.e. si et seulement si $F_X(t) = F_Y(t)$, $\forall t \in \mathbb{R}$.

Objectif : trouver la loi de X à partir de sa fonction de répartition.

Exemple. Soit X une variable aléatoire de loi uniforme sur $[0,1]$. Déterminer la loi de

$$Y := -\ln(1-X).$$

Solution. Nous avons $\mathbb{P}_X(dx) = \mathbf{1}_{[0,1]}\,\lambda_1(dx)$, donc, $\forall t \in \mathbb{R}$,

$$\begin{aligned} F_Y(t) &= \mathbb{P}(Y \leq t) = \mathbb{P}(X \leq 1 - e^{-t}) \\ &= \int_{]-\infty,1-e^{-t}]} \mathbf{1}_{[0,1]}\,\lambda_1(dx) = (1-e^{-t})\mathbf{1}_{t>0}. \end{aligned}$$

Or la fonction caractéristique d'une variable aléatoire Z de loi exponentielle de paramètre 1 est

$$F_Z(t) = \int_{]-\infty,t]} e^{-x}\mathbf{1}_{x>0}\,\lambda_1(dx) = (1-e^{-t})\mathbf{1}_{t>0}.$$

On en déduit que

$$\boxed{Y \text{ est de loi exponentielle de paramètre } 1.}$$

La fonction de répartition $F_X(t) = \mathbb{P}_X(]-\infty,t])$ est

- **croissante**, car les ensembles $]-\infty,t]$ sont croissants en t,
- **continue à droite**, car, par continuité décroissante de $\mathbb{P}_X$,

$$F_X(t_+) = \lim_{h\to 0_+} \mathbb{P}_X(]-\infty,t+h]) = \mathbb{P}_X\left(\cap_{h>0}\,]-\infty,t+h]\right)$$
$$= \mathbb{P}_X(]-\infty,t]) = F_X(t)$$

- telle que, par continuité décroissante de $\mathbb{P}_X$,

$$\lim_{t\to-\infty} F_X(t) = \lim_{t\to-\infty} \mathbb{P}_X(]-\infty,t]) = \mathbb{P}_X(\cap_{t\in\mathbb{R}}\,]-\infty,t])$$
$$= \mathbb{P}_X(\emptyset) = 0$$

- telle que, par continuité croissante de $\mathbb{P}_X$,

$$\lim_{t\to+\infty} F_X(t) = \lim_{t\to+\infty} \mathbb{P}_X(]-\infty,t]) = \mathbb{P}_X(\cup_{t\in\mathbb{R}}\,]-\infty,t])$$
$$= \mathbb{P}_X(]-\infty,+\infty[) = 1$$

- F_X admet une **limite à gauche** et, par continuité de $\mathbb{P}_X$,

$$\begin{aligned} F_X(t_-) &= \lim_{h\to 0_+} \mathbb{P}_X(]-\infty,t-h]) = \mathbb{P}_X\left(\cup_{h>0}\,]-\infty,t-h]\right) \\ &= \mathbb{P}_X(]-\infty,t[) = \mathbb{P}(X \in]-\infty,t[). \end{aligned}$$

- On a ainsi, pour tout $-\infty \leq a \leq b \leq +\infty$,

$$\begin{aligned} \mathbb{P}(X \in]a,b[) &= \mathbb{P}(X \in]-\infty,b[) - \mathbb{P}(X \in]-\infty,a]) \\ &= F_X(b_-) - F_X(a), \\ \mathbb{P}(X \in]a,b]) &= F_X(b) - F_X(a), \\ \mathbb{P}(X \in [a,b[) &= F_X(b_-) - F_X(a_-), \\ \mathbb{P}(X \in [a,b]) &= F_X(b) - F_X(a_-). \end{aligned}$$

- En particulier, $\mathbb{P}\left(X = x\right) = F_X(x) - F_X(x_-)$, donc

$$\mathbb{P}(X = x) > 0 \iff F_X \text{ est discontinue au point } x.$$

Calculs de fonction de répartition

Variables aléatoires discrètes

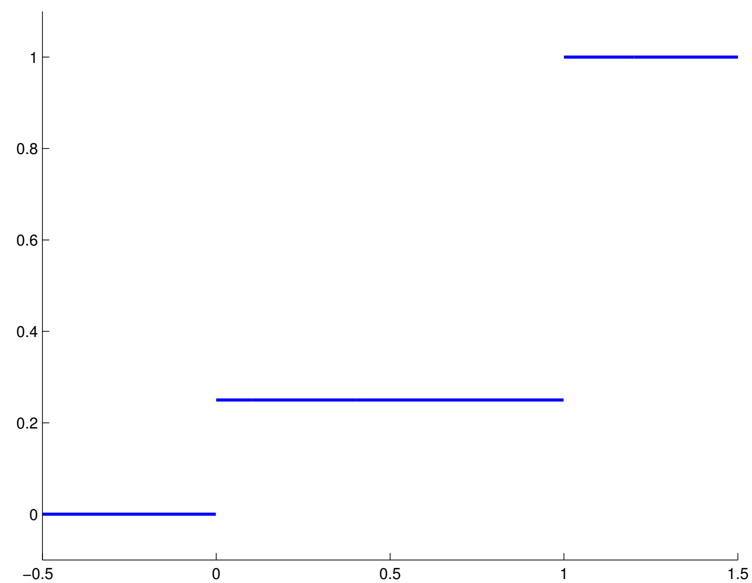
Soit X une v.a. discrète à valeurs dans $K = \{k_1 < k_2 < \dots\} \subset \mathbb{R}$ presque sûrement. Alors F_X est constante par morceaux et

$$F_X(t) = \sum_{k \in K,\, k \le t} \mathbb{P}(X = k), \quad \forall t \in \mathbb{R},$$

donc $F_X(t) - F_X(t_-) = \mathbb{P}(X = t)$, pour tout $t \in \mathbb{R}$.

Exemple. X de loi $\mathbb{P}_X = \frac{1}{4}\delta_0 + \frac{3}{4}\delta_1$,

$$F_X(t) = \begin{cases} 0 & \text{si } t < 0, \\ 1/4 & \text{si } 0 \le t < 1, \\ 1 & \text{si } 1 \le t. \end{cases}$$



Variables aléatoires absolument continues

Soit X une v.a. à valeurs dans $\mathbb{R}$, de loi $\mathbb{P}_X = f_X \, d\lambda_1$. Alors F_X est continue, dérivable λ_1-presque partout et $F_X' = f_X$ λ_1-presque partout, soit

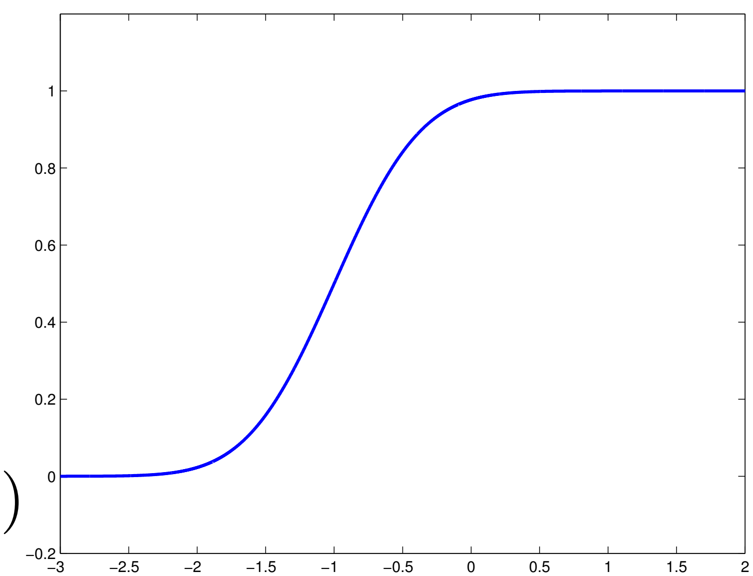
$$F_X(t) = \int_{]-\infty,t[} f_X(x)\,\lambda_1(dx), \quad \forall t \in \mathbb{R}.$$

Exemple. X de loi $\mathcal{N}(-1,1)$

$$\mathbb{P}_X(x) = \frac{\exp(-(x+1)^2/2)}{\sqrt{2\pi}}\,\lambda_1(dx),$$

donc

$$F_X(t) = \int_{]-\infty,t]} \frac{\exp(-(x+1)^2/2)}{\sqrt{2\pi}}\,\lambda_1(dx)$$



Ex 2 (loi mixte). Calculer F_X, pour X une v.a. réelle de loi

$$\mathbb{P}_X(dx) = \frac{1}{2}\delta_{-1}(dx) + \frac{1}{4}\delta_2(dx) + \frac{\mathbf{1}_{x>0}}{4}e^{-x}\,\lambda_1(dx).$$

Solution. Pour tout $t \in \mathbb{R}$, $F_X(t) = \int_{]-\infty,t]} 1\,\mathbb{P}_X(dx)$, donc

$$F_X(t) = \frac{1}{2}\mathbf{1}_{-1\leq t} + \frac{1}{4}\mathbf{1}_{2\leq t} + \int_{]-\infty,t]} \frac{\mathbf{1}_{x>0}}{4}e^{-x}\,\lambda_1(dx).$$

Par conséquent,

$$\boxed{F_X(t) = \begin{cases} 0 & \text{si } t < -1, \\ \frac{1}{2} & \text{si } -1 \leq t < 0, \\ \frac{1}{2} + \frac{1}{4}(1 - e^{-t}) & \text{si } 0 \leq t < 2 \\ \frac{3}{4} + \frac{1}{4}(1 - e^{-t}) & \text{si } 2 \leq t. \end{cases}}$$

Calcul d'une loi à partir d'une fonction de répartition

Soit X une variable aléatoire de fonction de répartition F_X.

Propriété

Supposons que F_X est de classe C^1 sur $\mathbb{R} \setminus K$, où K est au plus dénombrable. Alors X est de loi mixte et sa loi s'écrit :

$$\mathbb{P}_X(dx) = F'_X(x)\,\mathbf{1}_{x \notin K}\,\lambda_1(dx) + \sum_{k \in K} p_k \delta_k(dx)$$

avec $p_k = \mathbb{P}(X = k) = F_X(k) - F_X(k_-)$.

$\rightarrow$ Remarques

- F'_X définie partout sauf aux points $k \in K$,
- Si F_X est continue, alors X de loi absolument continue, de densité $f_X = F'$
- Si F_X constante par morceaux ($F' = 0$ λ_1-p.p.), alors X de loi discrète à valeurs dans K.

Souvent, on se donnera une fonction $F : \mathbb{R} \to \mathbb{R}$ et il faudra dans un premier temps déterminer l'existence d'une variable aléatoire X telle que $F_X = F$, avant de déterminer la loi de X.

Existence d'une v.a. de fonction de répartition donnée

Une fonction $F : \mathbb{R} \to \mathbb{R}$ est la fonction de répartition d'une variable aléatoire si et seulement si elle est

- croissante,
- continue à droite
- et telle que

$$\lim_{t \to -\infty} F(t) = 0 \quad \text{et} \quad \lim_{t \to +\infty} F(t) = 1.$$

Exemple. Montrer qu'il existe une v.a. X de fonction de répartition

$$F(t) = \begin{cases} 0 & \text{si } t < 3 \\ 1/3 & \text{si } 3 \leq t < 6 \\ 1 & \text{si } 6 \leq t. \end{cases}$$

Puis déterminer la loi de X.

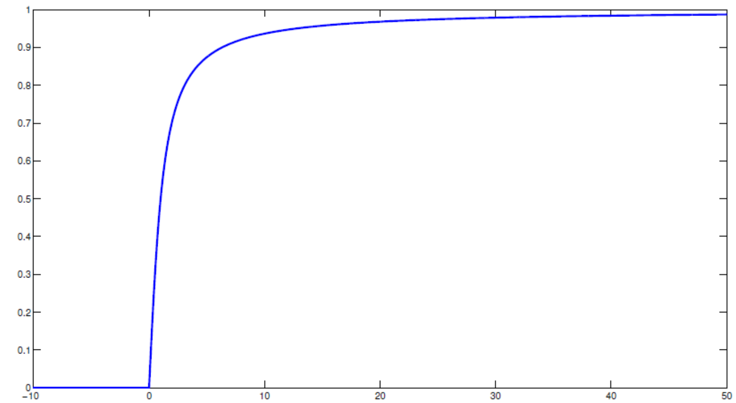
Solution. F est croissante et continue à droite, $\lim_{t \to -\infty} F(t) = 0$ et $\lim_{t \to +\infty} F(t) = 1$, donc il existe X telle que $F_X = F$.
De plus, F_X est constante par morceaux, donc X est discrète et

$$F_X(t) - F_X(t_-) = \begin{cases} 1/3 \text{ si } t = 3, \\ 2/3 \text{ si } t = 6, \\ 0 \text{ sinon.} \end{cases}$$

donc la loi de X est $\boxed{\mathbb{P}_X = \frac{1}{3}\delta_3 + \frac{2}{3}\delta_6}$

Exemple. Montrer qu'il existe une v.a. X de fonction de répartition

$$F(t) = \begin{cases} 0 & \text{si } t < 0, \\ \frac{2}{\pi}\arctan(t) & \text{si } t \geq 0. \end{cases}$$



Solution. F est croissante et continue à droite, $\lim_{t\to-\infty} F(t) = 0$ et $\lim_{t\to+\infty} F(t) = 1$, donc il existe X telle que $F_X = F$.
De plus, F_X est C^1 par morceaux et continue, donc X est absolument continue de densité $f_X : \mathbb{R} \to \mathbb{R}$ telle que

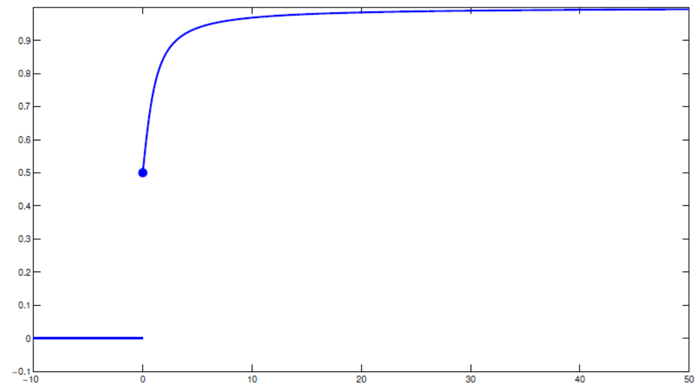
$$f_X(x) = F'(x)\,\mathbf{1}_{x\neq 0} = \begin{cases} 0 & \text{si } x \leq 0, \\ \frac{2}{\pi(1+x^2)} & \text{si } x > 0. \end{cases}$$

En définitive,

$$\boxed{\mathbb{P}_X(dx) = \frac{2}{\pi(1+x^2)}\mathbf{1}_{x>0}\,\lambda_1(dx)}$$

Exemple. Montrer qu'il existe une v.a. X de fonction de répartition

$$F(t) = \begin{cases} 0 & \text{si } t < 0, \\ \frac{1}{\pi}\left(\arctan(t) + \frac{\pi}{2}\right) & \text{si } t \geq 0. \end{cases}$$



Solution. F est croissante et continue à droite, $\lim_{t\to-\infty} F(t) = 0$ et $\lim_{t\to+\infty} F(t) = 1$, donc il existe X telle que $F_X = F$.
De plus, F est C^1 en tout $x \in \mathbb{R}^*$ et discontinue en $x = 0$. Par suite, il s'agit de la fonction de répartition d'une v.a. X de loi mixte donnée par

$$\mathbb{P}_X(dx) = (F(0) - F(0_-))\delta_0(dx) + F'(x)\mathbf{1}_{x\neq 0}\,\lambda_1(dx),$$

soit

$$\boxed{\mathbb{P}_X(dx) = \frac{1}{2}\delta_0(dx) + \frac{1}{\pi(1+x^2)}\mathbf{1}_{x>0}\,\lambda_1(dx).}$$

Fonction caractéristique : Définition

Définition

Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$. La fonction

$$\begin{array}{rccl} \varphi_X : & \mathbb{R}^d & \longrightarrow & \mathbb{C} \\ & u & \longmapsto & \mathbb{E}\left(e^{i\langle u,X\rangle}\right) = \int_{\mathbb{R}^d} e^{i\langle u,x\rangle}\mathbb{P}_X(dx) \end{array},$$

où $\langle u,X\rangle = u_1X_1 + \cdots + u_dX_d$, est bien définie et est appelée la **fonction caractéristique** de X.

Caractérisation de la loi

Deux v.a. X et Y à valeurs dans $\mathbb{R}^d$ ont même loi si et seulement si elles ont même fonction caractéristique, i.e.

$$\varphi_X(u) = \varphi_Y(u), \ \forall u \in \mathbb{R}^d.$$

Exemple. Soit X une variable aléatoire à valeurs réelles de loi

$$\mathbb{P}_X(dx) = \frac{1}{2}\delta_{-1}(dx) + \frac{1}{4}\delta_2(dx) + \frac{\mathbf{1}_{x>0}}{4}\,e^{-x}\,\lambda_1(dx).$$

Déterminer la fonction caractéristique de X.

Solution. D'après le théorème du transport, $\forall u \in \mathbb{R}$,

$$\begin{aligned}\varphi_X(u) &= \int_{\mathbb{R}} e^{iux}\,\mathbb{P}_X(dx)\\ &= \frac{1}{2}e^{-iu} + \frac{1}{4}e^{2iu} + \int_{\mathbb{R}} e^{iux}\,\frac{\mathbf{1}_{x>0}}{4}\,e^{-x}\,\lambda_1(dx),\end{aligned}$$

soit, $\forall u \in \mathbb{R}$,

$$\boxed{\varphi_X(u) = \frac{1}{2}e^{-iu} + \frac{1}{4}e^{2iu} + \frac{1}{4(1-iu)}}$$

Théorème (Existence d'une densité)

Soit X une v.a. à valeurs dans $\mathbb{R}^d$. Si φ_X est intégrable par rapport à $\lambda_d(du)$, alors X suit une loi absolument continue de densité f_X donnée par

$$f_X(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi_X(u)\, e^{-i\langle u,x\rangle}\, \lambda_d(du).$$

Remarques.

- Le membre de droite est une fonction continue en x
- Si on voit la fonction caractéristique comme une transformée de Fourier, alors le théorème ci-dessus donne la transformée de Fourier inverse.

Calcul de Moments

Théorème

Soit X une variable aléatoire à valeurs réelles. Soit $k \in \mathbb{N}$ tel que X^k est intégrable, c'est-à-dire tel que

$$\mathbb{E}(|X|^k) < +\infty.$$

Alors φ_X est k fois dérivable et $\varphi_X^{(k)}(0) = i^k \mathbb{E}(X^k)$.

De plus, au voisinage de 0,

$$\varphi_X(t) = \sum_{n=0}^{k} \frac{(it)^n}{n!} \mathbb{E}(X^n) + o(|t|^k).$$

Une réciproque partielle

Soit X une variable aléatoire à valeurs réelles. Si φ_X est $2n$ fois dérivable, avec $n \in \mathbb{N}^*$, alors

$$\mathbb{E}\left(|X|^k\right) < +\infty, \forall k \in \{0,1\ldots,2n\}$$

et

$$\mathbb{E}\left(X^k\right) = (-i)^k \varphi^{(k)}(0).$$

Probabilités
5. Théorème de Fubini

´

Mesurer sur un espace produit

Soient $(\Omega_1,\mathcal{F}_1,\mu_1)$ et $(\Omega_2,\mathcal{F}_2,\mu_2)$ deux espaces mesurés.
But : transférer les structures d'espaces mesurés à $\Omega_1 \times \Omega_2$.

Espace mesuré produit

On définit l'espace mesuré produit sur $\Omega_1 \times \Omega_2$ par

- la tribu produit $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(A_1 \times A_2,\ A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2)$,
- la mesure produit $\mu_1 \otimes \mu_2$, seule mesure sur $\mathcal{F}_1 \otimes \mathcal{F}_2$ telle que

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1) \times \mu_2(A_2),\ \forall A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2.$$

$\rightarrow$ Remarques

- Les éléments de $\mathcal{F}_1 \otimes \mathcal{F}_2$ ne sont pas tous de la forme $A_1 \times A_2$.
- On généralise aisément à $\Omega_1 \times \cdots \times \Omega_d$.

Exemples de tribus et de mesures produits

- Tribus boréliennes.

$$\begin{aligned}
\mathcal{B}(\mathbb{R}^2) &= \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \stackrel{not.}{=} \mathcal{B}(\mathbb{R})^{\otimes 2}, \\
\mathcal{B}(\mathbb{R}^3) &= \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}^2) \otimes \mathcal{B}(\mathbb{R}) \\
&= \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \stackrel{not.}{=} \mathcal{B}(\mathbb{R})^{\otimes 3}, \\
\mathcal{B}(\mathbb{R}^d) &= \mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R}) \stackrel{not.}{=} \mathcal{B}(\mathbb{R})^{\otimes d}.
\end{aligned}$$

- La mesure de Lebesgue.

$$\begin{aligned}
\lambda_2 &= \lambda_1 \otimes \lambda_1 \stackrel{not.}{=} \lambda_1^{\otimes 2}, \\
\lambda_3 &= \lambda_1 \otimes \lambda_2 \stackrel{not.}{=} \lambda_2 \otimes \lambda_1 = \lambda_1 \otimes \lambda_1 \otimes \lambda_1 \stackrel{not.}{=} \lambda_1^{\otimes 3}, \\
\lambda_d &= \lambda_i \otimes \lambda_{d-i} = \lambda_1 \otimes \cdots \otimes \lambda_1 \stackrel{not.}{=} \lambda_1^{\otimes d}.
\end{aligned}$$

- Produits de deux mesures à densité. Soient $f_1 : \mathbb{R}^{d_1} \to \mathbb{R}$ et $f_2 : \mathbb{R}^{d_2} \to \mathbb{R}$ deux densités, c'est-à-dire deux fonctions positives presque sûrement et telles que

$$\int_{\mathbb{R}^{d_1}} f_1(x)\,\lambda_{d_1}(dx) = \int_{\mathbb{R}^{d_2}} f_2(y)\,\lambda_{d_2}(dy) = 1.$$

Alors

$$\begin{aligned}(f_1(x)\lambda_{d_1}(dx)) \otimes (f_2(y)\lambda_{d_2}(dy)) &= f_1(x)f_2(y)\lambda_{d_1}(dx) \otimes \lambda_{d_2}(dy) \\ &= f_1(x)f_2(y)\lambda_{d_1+d_2}(dx,dy).\end{aligned}$$

- Produits de n mesures à densité. Soient $f_1, \ldots, f_n$ des densités sur $\mathbb{R}^{d_1}, \ldots, \mathbb{R}^{d_n}$. Alors

$$(f_1\lambda_{d_1}) \otimes \ldots \otimes (f_n\lambda_n) = f_1 \otimes \cdots \otimes f_n\,\lambda_{d_1+\cdots+d_n},$$

où $f_1 \otimes \cdots \otimes f_n(x_1, \ldots, x_n) = f(x_1)\cdots f(x_n)$.

- Produits de mesures discrètes. On a

$$\left(\sum_{m\in K_1} p_m\delta_m\right)\otimes\left(\sum_{n\in K_2} q_n\delta_n\right)=\sum_{(m,n)\in K_1\times K_2} p_m q_n\delta_{(m,n)}.$$

- Produits de mesures discrètes et à densité. On a

$$\left(\sum_{m\in K} p_m\delta_m\right)\otimes(f(x)\lambda_d(dx))=\sum_{m\in K} p_m\,f(x)\,\delta_m\otimes\lambda_d(dx).$$

Remarque : tous ces résultats sont obtenus par une forme de bilinéarité... par exemple,

$$\begin{aligned}\left(\sum_{m\in K} p_m\delta_m\right)\otimes(f(x)\lambda_d(dx)) &= \sum_{m\in K} p_m\left[\delta_m\otimes(f(x)\lambda_d(dx))\right]\\ &= \sum_{m\in K} p_m\,f(x)\,\left[\delta_m\otimes\lambda_d(dx)\right].\end{aligned}$$

Calcul d'intégrale sur un espace produit

Théorème de Fubini–Tonelli

Soit $f : (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2) \to \overline{R}$ mesurable positive ou $\mu_1 \otimes \mu_2$-intégrable, alors les trois intégrales suivantes sont bien définies et égales

$$I_1 = \int_{\Omega_1 \times \Omega_2} f(\omega_1 \times \omega_2)\, \mu_1 \otimes \mu_2(d\omega_1, d\omega_2),$$
$$I_2 = \int_{\Omega_1} \int_{\Omega_2} f(\omega_1, \omega_2)\, \mu_2(d\omega_2)\, \mu_1(d\omega_1),$$
$$I_3 = \int_{\Omega_2} \int_{\Omega_1} f(\omega_1, \omega_2)\, \mu_1(d\omega_1)\, \mu_2(d\omega_2).$$

Donc on peut inverser les intégrales sur Ω_1 et Ω_2

- si f est positive,
- ou si f est $\mu_1 \otimes \mu_2$-intégrable, *i.e.* $\int_{\Omega_1 \times \Omega_2} |f|\, d\mu_1 \otimes \mu_2 < \infty$.

Situation usuelle. On souhaite calculer une intégrale double en inversant l'ordre des intégrales, soit

$$\int_{\Omega_1}\int_{\Omega_2} f(\omega_1,\omega_2)\,\mu_2(d\omega_2)\,\mu_1(d\omega_1) = \int_{\Omega_2}\int_{\Omega_1} f(\omega_1,\omega_2)\,\mu_1(d\omega_1)\,\mu_2(d\omega_2).$$

- Si f est positive, ce changement d'intégrale est licite.
- Si f est de signe qcq, on doit avoir $\int_{\Omega_1\times\Omega_2} |f|\,d\mu_1\otimes\mu_2 < \infty$, ce qui, d'après le théorème de Fubini, est équivalent à

$$\begin{aligned} &\int_{\Omega_1}\int_{\Omega_2} |f(\omega_1,\omega_2)|\,\mu_2(d\omega_2)\,\mu_1(d\omega_1) < \infty \\ \text{ou } &\int_{\Omega_2}\int_{\Omega_1} |f(\omega_1,\omega_2)|\,\mu_1(d\omega_1)\,\mu_2(d\omega_2) < \infty. \end{aligned}$$

  Si cela est vérifié, alors le changement d'intégrale est licite.

Dans le dernier cas, $f(\cdot,y)$ est μ_1-intégrable pour μ_2-presque tout $y\in\Omega_2$ et $f(x,\cdot)$ est μ_2-intégrable pour μ_1-presque tout $x\in\Omega_1$.

Cas 1. Intégrale par rapport à la mesure de Lebesgue

Soit $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ une fonction mesurable. Pour étudier son intégrale par rapport à λ_2, on procède comme suit

1. Si f est positive, alors, d'après le théorème de Fubini-Tonelli,

$$\begin{aligned}\int_{\mathbb{R}^2} f(x,y)\, d\lambda_2(x,y) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\, d\lambda_1(x) \right) d\lambda_1(y) \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\, d\lambda_1(y) \right) d\lambda_1(x).\end{aligned}$$

2. Si f est de signe qcq, alors on applique l'étape 1 à $|f|$.
   - Si $\int_{\mathbb{R}^2} |f|\, d\lambda_2 = \infty$, alors f n'est pas intégrable.
   - Si $\int_{\mathbb{R}^2} |f|\, d\lambda_2 < \infty$, alors f est intégrable et on peut calculer l'intégrale de f en appliquant les égalités de l'étape 1 à f directement.

Ex 1. Calculer, lorsqu'elle est bien définie, l'intégrale de

$$f : (x,y) \in]0,1]^2 \mapsto x/(xy)^\alpha$$

par rapport à λ_2.

Solution. f est positive donc, d'après le théorème de Fubini-Tonelli,

$$\begin{aligned}
\int_{]0,1]^2} \frac{x}{(xy)^\alpha} \lambda_2(dx,dy) &= \int_{]0,1]} \int_{]0,1]} \frac{x}{(xy)^\alpha} \lambda_1(dy)\lambda_1(dx) \\
&= \begin{cases} \int_{]0,1]} \infty \, \lambda_1(dx) \text{ si } \alpha \geq 1, \\ \int_{]0,1]} \frac{x}{(1-\alpha)x^\alpha} \lambda_1(dx) \text{ si } \alpha < 1, \end{cases} \\
&= \begin{cases} \infty \text{ si } \alpha \geq 1, \\ \frac{1}{(1-\alpha)(2-\alpha)} \text{ si } \alpha < 1. \end{cases}
\end{aligned}$$

Conclusion. f est intégrable ssi $\alpha < 1$. Dans ce cas,

$$\int_{]0,1]^2} f \, d\lambda_2 = \frac{1}{(1-\alpha)(2-\alpha)}.$$

Ex 2. Calculer, lorsqu'elle est bien définie, l'intégrale de

$$f : (x,y) \in]0,1]^2 \mapsto (x+y)\sin(2\pi x)/(xy)$$

par rapport à λ_2.

Solution. f étant de signe quelconque, intéressons nous à l'intégrale de $|f|$. Cette fonction est positive donc, d'après le théorème de Fubini-Tonelli,

$$\int_{]0,1]^2} \frac{|x\sin(2\pi x)|}{|xy|}\lambda_2(dx,dy) = \int_{]0,1]}\int_{]0,1]} \frac{|(x+y)\sin(2\pi x)|}{|xy|}\lambda_1(dy)\lambda_1(dx)$$

$$= \int_{]0,1]} \frac{|\sin(2\pi x)|}{|x|}\int_{]0,1]} \frac{|x+y|}{|y|}\lambda_1(dy)\lambda_1(dx)$$

$$= \int_{]0,1]} \frac{|\sin(2\pi x)|}{|x|}\infty \, d\lambda_1(x) = \infty.$$

Conclusion. L'intégrale n'est pas bien définie.

Cas 2. Échange des signes $\sum$ et $\int$

Pour tout $n \geq 0$, soit $u_n : \mathbb{R}^d \mapsto \overline{\mathbb{R}}$ mesurable. On souhaite étudier

$$\int_{\mathbb{R}^d} \sum_{n=0}^{\infty} u_n(x)\, \lambda_d(dx) \qquad \left(= \int_{\mathbb{R}^d \times \mathbb{N}} u_n(x)\, \lambda_d(dx) \otimes \sum_{n=0}^{\infty} \delta_n \right).$$

1. Si les u_n sont positives, alors, d'après le théorème de Fubini-Tonelli (appliqué à $\lambda_d \otimes$ Card),

$$\int_{\mathbb{R}^d} \sum_{n=0}^{\infty} u_n(x)\, \lambda_d(dx) = \sum_{n=0}^{\infty} \int_{\mathbb{R}^d} u_n(x)\, \lambda_d(dx).$$

2. Si u_n est de signe qcq, alors on applique l'étape 1 à $|u_n|$.
   - Si $\int_{\mathbb{R}^d} \sum_{n=0}^{\infty} |u_n(x)|\, \lambda_d(dx) = \infty$, alors expression indéfinie.
   - Si $\int_{\mathbb{R}^d} \sum_{n=0}^{\infty} |u_n(x)|\, \lambda_d(dx) < \infty$, alors on peut intervertir l'intégrale et la somme, comme à l'étape 1.

Ex 3. Exprimer l'intégrale de $f : x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \mapsto \sum_{n=1}^{\infty} \frac{\cos(nx)}{n^2}$ par rapport à λ_1 sous la forme d'une série.

Solution. D'après le théorème de Fubini-Tonelli,

$$\int_{[-\frac{\pi}{2},\frac{\pi}{2}]} \sum_{n=1}^{\infty} \left| \frac{\cos(nx)}{n^2} \right| d\lambda_1(x) = \sum_{n=1}^{\infty} \int_{[-\frac{\pi}{2},\frac{\pi}{2}]} \left| \frac{\cos(nx)}{n^2} \right| d\lambda_1(x) \\ \leq \sum_{n=1}^{\infty} \frac{\pi}{n^2} < \infty.$$

Donc, en appliquant à nouveau le théorème de Fubini-Tonelli,

$$\int_{[-\frac{\pi}{2},\frac{\pi}{2}]} \sum_{n=1}^{\infty} \frac{\cos(nx)}{n^2} d\lambda_1(x) = \sum_{n=1}^{\infty} \int_{[-\frac{\pi}{2},\frac{\pi}{2}]} \frac{\cos(nx)}{n^2} d\lambda_1(x) \\ = \sum_{n=1}^{\infty} \left[\frac{\sin(nx)}{n^3} \right]_{-\pi/2}^{\pi/2} = \sum_{n=1}^{\infty} \frac{2(-1)^n}{(2n+1)^3}.$$

Ex 4. Soit (X,Y) une v.a. absolument continue de densité

$$f_{(X,Y)}(x,y) = \frac{1}{\sqrt{2\pi}} \mathbf{1}_{x\in[0,1]}(x) e^{-\frac{y^2}{2}}.$$

Quelle est la loi de $Z = -\ln(1-X)$?

Solution. On a $\mathbb{P}(Z \leq 0) = 0$ et, pour tout $t > 0$,

$$\begin{aligned}
F_Z(t) &= \mathbb{P}(Z \leq t) = \mathbb{P}(X \leq 1 - e^{-t}) = \mathbb{P}((X,Y) \in [0,1-e^{-t}] \times \mathbb{R}) \\
&= \int_{[0,1-e^{-t}]\times\mathbb{R}} \frac{1}{\sqrt{2\pi}} \mathbf{1}_{x\in[0,1]}(x) e^{-\frac{y^2}{2}} \lambda_2(dx,dy) \\
&= \int_{[0,1-e^{-t}]} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \lambda_1(dy)\lambda_1(dx) \text{ par Fubini} \\
&= \int_{[0,1-e^{-t}]} d\lambda_1(x) = 1 - e^{-t}.
\end{aligned}$$

Donc Z est une variable aléatoire de loi exponentielle de param 1.

Ex 5. Soient f_X et f_Y des densités de v.a. sur $\mathbb{R}$. Montrer que

1. $g : (x,y) \mapsto f_X(x)f_Y(y)$ est la densité d'une v.a. Z sur $\mathbb{R}^2$;
2. Montrer que $\varphi_Z(x,y) = \varphi_X(x)\varphi_Y(y)$, $\forall (x,y) \in \mathbb{R}^2$.

Solution. 1. La fonction g est mesurable positive ; de plus

$$\begin{aligned}\int_{\mathbb{R}^2} g(x,y)\lambda_2(dx,dy) &= \int_{\mathbb{R}}\int_{\mathbb{R}} f_X(x)f_Y(y)\lambda_1(dx)\lambda_1(dy)\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f_X(x)\lambda_1(dx)f_Y(y)\lambda_1(dy)\\ &= \int_{\mathbb{R}} f_X(x)\lambda_1(dx)\int_{\mathbb{R}} f_Y(y)\lambda_1(dy) = 1.\end{aligned}$$

2. Nous avons, d'après le théorème de Fubini-Tonelli,

$$\begin{aligned}\varphi_Z(x,y) &= \mathbb{E}\left(e^{i\langle(x,y),Z\rangle}\right) = \int_{\mathbb{R}^2} e^{i(xu+yv)}g(u,v)\,\lambda_2(du,dv)\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{ixu}e^{iyv}\,\lambda_1(du)\lambda_1(dv) = \int_{\mathbb{R}} e^{ixu}\lambda_1(du)\int_{\mathbb{R}} e^{iyv}\lambda_1(dv)\\ &= \mathbb{E}\left(e^{ixX}\right)\mathbb{E}\left(e^{iyY}\right) = \varphi_X(x)\varphi_Y(y).\end{aligned}$$

Probabilités
6. Indépendance de variables aléatoires

´

Définition

Soient X_1,X_2 deux variables aléatoires définies sur un même espace de probabilité $(\Omega,\mathcal{F},\mathbb{P})$.

- X_1 à valeurs dans $(E_1,\mathcal{E}_1)$
- X_2 à valeurs dans $(E_2,\mathcal{E}_2)$.

Définition

X_1 et X_2 indépendantes si, pour tous ensembles $B_1 \in \mathcal{E}_1$, $B_2 \in \mathcal{E}_2$,

$$\mathbb{P}(X_1 \in B_1,X_2 \in B_2) = \mathbb{P}(X_1 \in B_1)\,\mathbb{P}(X_2 \in B_2).$$

ce qui est équivalent à

$$\mathbb{P}_{(X_1,X_2)} = \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2}.$$

Généralisation : Indépendance de n v.a.

Définition

Les v.a. $X_1, \ldots, X_n$ sont mutuellement indépendantes si

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

pour tous ensembles mesurables $B_1, \ldots, B_n$.
C'est-à-dire si

$$\mathbb{P}_{(X_1, \ldots, X_n)} = \mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n}.$$

$\rightarrow$ Remarque

- $X_1, \ldots, X_n$ indépendantes $\Rightarrow$ X_i, X_j indépendantes $\forall i \neq j$.
- X_i, X_j indépendantes $\forall i \neq j$ $\not\Rightarrow$ $X_1, \ldots, X_n$ indépendantes

Si (X_1,X_2) est discret. (on peut supposer E_1 et E_2 discrets)

Propriété

X_1 et X_2 sont indépendantes si et seulement si, $\forall x_1 \in E_1, x_2 \in E_2$,

$$\mathbb{P}(X_1 = x_1 \text{ et } X_2 = x_2) = \mathbb{P}(X_1 = x_1)\,\mathbb{P}(X_2 = x_2).$$

Exemple 1. Loi d'un couple (X_1,X_2) donné par

$X_1 \setminus X_2$	0	1	2	$\mathbb{P}(X_1 = i)$
0	$8/27$	$1/27$	0	$1/3$
1	$4/27$	$11/27$	$3/27$	$2/3$
$\mathbb{P}(X_2 = j)$	$4/9$	$4/9$	$1/9$	

$$\mathbb{P}(X_1 = 0 \text{ et } X_2 = 0) \neq \mathbb{P}(X_1 = 0)\,\mathbb{P}(X_2 = 0).$$

donc $\boxed{X_1 \text{ et } X_2 \text{ ne sont pas indépendants.}}$

Ex 2. X_i, X_j indépendantes $\forall i \neq j \;\; \not\Rightarrow \;\; X_1, \ldots, X_n$ indépendantes

Soient X_1 et X_2 deux v.a. indépendantes de loi $\frac{1}{2}(\delta_{-1} + \delta_1)$.

$\to X_1$, X_2 et $X_3 = X_1 X_2$ sont 2 à 2 indépendantes de même loi.

Mais

$$\begin{aligned} \mathbb{P}(X_1 = 1, X_2 = 1, X_3 = 1) = \mathbb{P}(X_1 = 1, X_2 = 1) &= \frac{1}{4} \\ \neq \prod_{i=1}^{3} \mathbb{P}(X_i = 1) &= \frac{1}{8}. \end{aligned}$$

Ainsi,

$$\boxed{X_1, X_2 \text{ et } X_3 \text{ ne sont pas indépendants}}$$

Caractérisation

Espérance d'un produit de variables indépendantes

Les v.a. $X_1, \ldots, X_n$ sont mutuellement indépendantes ssi

$$\mathbb{E}\left(\prod_{i=1}^{n} h_i(X_i)\right) = \prod_{i=1}^{n} \mathbb{E}\left(h_i(X_i)\right)$$

pour toutes fonctions $h_i : E_i \to \overline{\mathbb{R}}$ mesurables positives, $1 \leq i \leq n$.

Preuve. Si $h_i = \mathbf{1}_{A_i}$ pour tout $i \in \{1, \ldots, n\}$, alors

$$\mathbb{E}\left(\prod_{i=1}^{n} h_i(X_i)\right) = \mathbb{E}\left(\mathbf{1}_{X_1 \in A_1, \ldots, X_n \in A_n}\right) = \mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n)$$

$$= \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i) = \prod_{i=1}^{n} \mathbb{E}\left(h_i(X_i)\right).$$

Cas général : passage à la limite par convergence monotone.

Corollaire

Si les v.a. $X_1, \ldots, X_n$ sont mutuellement indépendantes, alors, pour toutes fonctions mesurables $f_i : E_i \to F_i$, $1 \leq i \leq n$, les variables aléatoires

$$f_1(X_1), \ldots, f_n(X_n)$$

sont mutuellement indépendantes.

Preuve. Pour toutes fonctions $h_i : F_i \to \overline{\mathbb{R}}$ mesurables positives, $h_i \circ f_i : E_i \to \overline{\mathbb{R}}$ est une fonction mesurable positive, donc, d'après le théorème précédent (implication $\Rightarrow$),

$$\mathbb{E}\left(\prod_{i=1}^{n} h_i(f_i(X_i))\right) = \prod_{i=1}^{n} \mathbb{E}\left(h_i(f_i(X_i))\right).$$

Donc, d'après le théorème précédent (implication $Leftarrow$),

$$\boxed{f_1(X_1), \ldots, f_n(X_n) \text{ sont mutuellement indépendantes}}$$

Fonctions de répartition

Les variables aléatoires réelles $X_1,\ldots,X_n$ sont mutuellement indépendantes si et seulement si

$$\mathbb{P}(X_1 \leq t_1,\ldots,X_n \leq t_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq t_i),\ \forall (t_1,\ldots,t_n) \in \mathbb{R}^n$$

Fonctions caractéristiques

Les variables aléatoires $X_1,\ldots,X_n$ à valeurs dans $\mathbb{R}^d$ sont mutuellement indépendantes si et seulement si

$$\mathbb{E}\left(e^{i(u_1X_1+\cdot+u_nX_n)}\right) = \prod_{i=1}^{n} \mathbb{E}\left(e^{iu_iX_i}\right),\ \forall (u_1,\ldots,u_n) \in \mathbb{R}^n$$

Si (X_1,X_2) est de loi absolument continue. $E_1 = \mathbb{R}^m$ et $E_2 = \mathbb{R}^n$.

Propriété (Cas (X_1,X_2) absolument continu)

X_1 et X_2 sont indépendantes si et seulement si il existe deux fonctions mesurables positives g et h telles que

$$\boxed{f_{(X_1,X_2)}(x_1,x_2) = g(x_1)h(x_2)} \quad \lambda_{m+n}(x_1,x_2)\text{-presque partout,}$$

Propriété (Cas X_1 et X_2 absolument continues)

Soient X_1 et X_2 absolument continues. X_1 et X_2 sont indépendantes si et seulement si (X_1,X_2) est absolument continu et

$$\boxed{f_{(X_1,X_2)}(x_1,x_2) = f_{X_1}(x_1)\,f_{X_2}(x_2)} \quad \lambda_{m+n}(x_1,x_2)\text{-presque partout.}$$

Remarque :

- Si g et h sont des densités, alors $f_{X_1} = g$ et $f_{X_2} = h$ p.p.

Exemple 3. Supposons $E_1 = E_2 = \mathbb{R}$ et (X_1,X_2) de densité

$$f_{(X_1,X_2)}(x_1,x_2) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2-x_2}\, \mathbf{1}_{[0,+\infty[}(x_2).$$

Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?

Solution. En définissant les fonctions positives mesurables

$$g(x_1) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \text{ et } h(x_2) = e^{-x_2}\, \mathbf{1}_{[0,+\infty[}(x_2),$$

on a

$$f_{(X_1,X_2)}(x_1,x_2) = g(x_1)h(x_2) \text{ pour tout } x_1,x_2.$$

On en déduit que

$$\boxed{X_1 \text{ et } X_2 \text{ sont indépendants.}}$$

De plus $f_{X_1} = g$ et $f_{X_2} = h$, car g et h sont des densités.

Exemple 4. Supposons $E_1 = E_2 = \mathbb{R}$ et (X_1,X_2) de densité

$$f_{(X_1,X_2)}(x_1,x_2) = \frac{1}{\pi} e^{-x_1^2 - x_2^2} \, \mathbf{1}_{x_1 \neq x_2}.$$

Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?

Solution. En définissant les fonctions positives mesurables

$$g(x_1) = \frac{1}{\sqrt{\pi}} e^{-x_1^2} \text{ et } h(x_2) = \frac{1}{\sqrt{\pi}} e^{-x_2^2},$$

on a

$$f_{(X_1,X_2)}(x_1,x_2) = g(x_1)h(x_2) \text{ pour tout } (x_1,x_2) \notin \mathcal{D},$$

où $\mathcal{D} = \{(x_1,x_2) \text{ tel que } x_1 = x_2\}$ est de mesure λ_2 nulle. On en déduit que

$$\boxed{X_1 \text{ et } X_2 \text{ sont indépendants.}}$$

De plus $f_{X_1} = g$ et $f_{X_2} = h$ (densité d'une loi $\mathcal{N}(0,\sqrt{1/2})$).

Exemple 5. Supposons $E_1 = E_2 = \mathbb{R}$ et (X_1,X_2) de densité

$$f_{(X_1,X_2)}(x_1,x_2) = x_1\, e^{-x_1(1+x_2)}\, \mathbf{1}_{[0,+\infty[}(x_1)\, \mathbf{1}_{[0,+\infty[}(x_2).$$

Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?

Solution. On ne voit pas de factorisation évidente. On calcule donc f_{X_1} et f_{X_2} et on vérifie si $f_{(X_1,X_2)}$ est égale ou non à $f_{X_1} f_{X_2}$.

On a

$$f_{X_1}(x_1) = e^{-x_1}\, \mathbf{1}_{[0,+\infty[}(x_1) \text{ et } f_{X_2}(x_2) = \frac{1}{(1+x_2)^2}\, \mathbf{1}_{[0,+\infty[}(x_2).$$

Question : A-t-on $f_{(X_1,X_2)} = f_{X_1} f_{X_2}$ presque partout ?
Les ingrédients :

- $f_{(X_1,X_2)}(0,1) \neq f_{X_1}(0) f_{X_2}(1)$,
- $f_{(X_1,X_2)}$ et $f_{X_1} f_{X_2}$ sont continues sur $[0, +\infty[^2$.

Question : A-t-on $f_{(X_1,X_2)} = f_{X_1} f_{X_2}$ presque partout ?

Les ingrédients :

- $f_{(X_1,X_2)}(0,1) \neq f_{X_1}(0) f_{X_2}(1)$,
- $f_{(X_1,X_2)}$ et $f_{X_1} f_{X_2}$ sont continues sur $[0,+\infty[^2$.

On en déduit qu'il existe $\varepsilon > 0$ tel que

$$f_{(X_1,X_2)}(x_1,x_2) \neq f_{X_1}(x_1) f_{X_2}(x_2),\ \forall (x_1,x_2) \in [0,\varepsilon] \times [1,1+\varepsilon].$$

Or $\lambda_2([0,\varepsilon] \times [1,1+\varepsilon]) = \varepsilon^2$, donc

"$f_{(X_1,X_2)} = f_{X_1} f_{X_2}$ presque partout" est FAUX.

En particulier,

$$\boxed{X_1 \text{ et } X_2 \text{ ne sont pas indépendants.}}$$

Indépendance et corrélation

Soient X_1 et X_2 deux variables aléatoires admettant un moment d'ordre 2, c'est-à-dire telles que $\mathbb{E}(X_i^2) < \infty$.

Proposition

X_1 et X_2 indépendantes $\Longrightarrow \mathrm{Cov}(X_1,X_2) = 0$.

Preuve. On a

$$\begin{aligned}\mathrm{Cov}(X_1,X_2) &:= \mathbb{E}\left((X_1 - \mathbb{E}(X_1))(X_2 - \mathbb{E}(X_2))\right) \\ &= \mathbb{E}(X_1 - \mathbb{E}(X_1)) \times \mathbb{E}(X_2 - \mathbb{E}(X_2)) \\ &= (\mathbb{E}(X_1) - \mathbb{E}(X_1)) \times (\mathbb{E}(X_2) - \mathbb{E}(X_2)) = 0.\end{aligned}$$

ATTENTION : Réciproque FAUSSE

Contre exemple 6. Supposons

$$\mathbb{P}_{X_1} = \frac{1}{4}(\delta_{-1} + \delta_1) + \frac{1}{2}\delta_0 \text{ et } X_2 = \mathbf{1}_{\{0\}}(X_1).$$

Alors,

$$\mathbb{P}_{X_2} = \frac{1}{2}(\delta_0 + \delta_1)$$

Or $X_1 X_2 = 0$ presque sûrement, donc

$$\mathbb{E}(X_1 X_2) = 0,\ E(X_2) = 1/2 \text{ et } \mathbb{E}(X_1) = 0.$$

ainsi $\mathrm{Cov}(X_1, X_2) := \mathbb{E}(X_1 X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2) = 0$.

Mais $\mathbb{P}(X_1 = 1, X_2 = 1) = 0 \neq \mathbb{P}(X_1 = 1)\,\mathbb{P}(X_2 = 1) = 1/8$, donc

$$\boxed{X_1 \text{ et } X_2 \text{ ne sont pas indépendants.}}$$

Somme de variables aléatoires

Théorème du transport pour la somme de v.a indépendantes

Soient X et Y deux variables aléatoires indépendantes à valeurs dans E. Alors, pour toute fonction mesurable positive $f : E \mapsto \mathbb{R}$,

$$\begin{aligned}\mathbb{E}(f(X+Y)) &= \int_E f(x+y)\, d\mathbb{P}_{(X,Y)}(x,y) \\ &= \int_E \int_E f(x+y)\, d\mathbb{P}_X(x)\, d\mathbb{P}_Y(y)\end{aligned}$$

$\rightarrow$ Si, après calcul, on arrive à une écriture du type

$$\mathbb{E}(f(X+Y)) = \int_E f(z)\, d\mu(z),$$

alors μ est la loi de $X+Y$.

Question : Comment faire en pratique ?

Proposition (Variables aléatoires discrètes)

Soient X_1 et X_2 deux variables aléatoires de lois discrètes et indépendantes. Alors $X_1 + X_2$ est de loi discrète et, pour tout $k \in E$,

$$\mathbb{P}(X_1 + X_2 = k) = \sum_{m \in E} \sum_{n \in E} \mathbf{1}_{m+n=k} \mathbb{P}(X_1 = m)\mathbb{P}(X_2 = n)$$
$$= \sum_{j \in E} \mathbb{P}(X_1 = k - j)\mathbb{P}(X_2 = j).$$

Exemple 7. Si X_1,X_2 sont de loi uniforme sur $\{1, \dots, 6\}$, alors, $\forall k \in \mathbb{N}$,

$$\mathbb{P}(X_1 + X_2 = k) = \sum_{j=1}^{6} \mathbb{P}(X_1 = k - j)\mathbb{P}(X_2 = j) = \sum_{j=1}^{6} \frac{\mathbf{1}_{k-j \text{ et } j \in \{1,\dots,6\}}}{36}$$
$$= \frac{1}{36} \begin{cases} k - 1 & \text{si } k \in \{2, \dots, 7\} \\ 14 - k - 1 & \text{si } k \in \{8, \dots, 12\} \\ 0 & \text{sinon.} \end{cases}$$

Somme de variables aléatoires indépendantes

Proposition (Variables aléatoires abs. continues)

Soient X_1 et X_2 de loi absolument continues, de densités respectives f_{X_1} et f_{X_2}. Si X_1 et X_2 sont indépendantes, alors $X_1 + X_2$ est de loi absolument continue de densité

$$f_{X_1+X_2}(x) = f_{X_1} * f_{X_2}(x) \overset{\text{déf}}{=} \int_{\mathbb{R}} f_{X_1}(x-t) f_{X_2}(t)\, d\lambda_1(t).$$

Autrement dit,

$$\mathbb{P}_{X_1+X_2} = f_{X_1} * f_{X_2}\, d\lambda_1.$$

Autre méthode. La fonction caractéristique. Si $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes, alors, en posant $Z = X_1 + \cdots + X_n$, nous avons, pour tout $u \in \mathbb{R}$,

$$\begin{aligned}\varphi_Z(u) &= \mathbb{E}\left(e^{iu(X_1+\cdots+X_n)}\right) = \mathbb{E}\left(e^{iuX_1}\cdots e^{iuX_n}\right) \\ &= \mathbb{E}\left(e^{iuX_1}\right)\cdots\mathbb{E}\left(e^{iuX_n}\right) = \varphi_{X_1}(u)\cdots\varphi_{X_n}(u).\end{aligned}$$

Si on a des expressions explicites pour $\varphi_{X_1}(u), \ldots, \varphi_{X_n}(u)$, alors on peut calculer la fonction caractéristique φ_Z de Z, ce qui permet (parfois) de retrouver sa loi.

Exemples

- $X_1 \sim \mathcal{B}(n,p)$, $X_2 \sim \mathcal{B}(m,p)$ ind. $\Rightarrow X_1 + X_2 \sim \mathcal{B}(n+m,p)$,
- $X_1 \sim \mathcal{P}(a_1)$, $X_2 \sim \mathcal{P}(a_2)$ ind. $\Rightarrow X_1 + X_2 \sim \mathcal{P}(a_1 + a_2)$,
- $X_1 \sim \mathcal{N}(m_1,\sigma_1^2)$, $X_2 \sim \mathcal{N}(m_2,\sigma_2^2)$ ind.

$$\Longrightarrow X_1 + X_2 \sim \mathcal{N}(m_1 + m_2,\sigma_1^2 + \sigma_2^2)$$

Il faut travailler ces exemples chez vous !

Probabilités

7. Variance, moments et espaces L^p

´

1. Variance et covariance

Définition/proposition

Soient X, Y telles que $\mathbb{E}(X^2) < \infty$ et $\mathbb{E}(Y^2) < \infty$.

- La variance de X est bien définie, donnée par

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] \geq 0. \end{aligned}$$

- La covariance de X et Y est bien définie, donnée par

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right] \end{aligned}$$

- La fonction Cov est bilinéaire symétrique.
- X, Y indépendants $\Rightarrow \text{Cov}(X, Y) = 0$.

Exemple 1. Soit X de loi $\frac{1}{\lambda}e^{-x/\lambda}\mathbf{1}_{x>0}\lambda_1(dx)$ et $Y = X + 1$.

1. Calculer la variance de X.
2. Calculer la covariance de X et Y.

Solution.

1. Nous avons $\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 \frac{1}{\lambda} e^{-x/\lambda}\mathbf{1}_{x>0}/\ \lambda_1(dx) = 2\lambda^2$, donc <u>Var($X$) est bien définie</u>. De plus, $\mathbb{E}(X) = \lambda$, donc

$$\text{Var(X)} = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda^2.$$

2. $Y^2 = X^2 + 2X + 1$, avec $X \geq 0$ presque sûrement, donc

$$\mathbb{E}(Y^2) = \mathbb{E}(X^2) + 2\mathbb{E}(X) + 1 = 2\lambda^2 + 2\lambda + 1 < \infty.$$

De plus,

$$\mathbb{E}(Y) = \mathbb{E}(X + 1) = \mathbb{E}(X) + 1 = \lambda + 1,$$
$$\mathbb{E}(XY) = \mathbb{E}(X(X + 1)) = \mathbb{E}(X^2) + \mathbb{E}(X) = 2\lambda^2 + \lambda.$$

En définitive,

$$\boxed{\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 2\lambda^2 + \lambda - \lambda^2 - \lambda = \lambda^2.}$$

Les espaces L^p, $1 \leq p \leq \infty$

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité.

Definition/Propriété

Soit $p \in [1, +\infty[$. On appelle *espace* L^p l'ensemble des variables aléatoires à valeurs réelles X telles que

$$\|X\|_p := \left(\int_\Omega |X(\omega)|^p \, \mathbb{P}(d\omega) \right)^{\frac{1}{p}} = \left(\mathbb{E}(|X|^p) \right)^{\frac{1}{p}} < \infty.$$

L'ensemble L^p est un espace vectoriel et l'application $X \mapsto \|X\|_p$ est une norme sur L^p, en particulier

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p, \, \forall X, Y \in L^p.$$

$\rightarrow$ On pourrait considérer des variables aléatoires dans $\mathbb{R}^d$ en remplaçant $|\cdot|$ par la norme euclidienne.

Definition/Propriété

On appelle *espace* L^∞ l'ensemble des variables aléatoires à valeurs réelles X bornées par une constante, c'est-à-dire telles que

$$\exists c \geq 0 \text{ tel que } \mathbb{P}(|X| \leq c) = 1.$$

L'ensemble L^∞ est un espace vectoriel muni de la norme

$$\begin{aligned} \|X\|_\infty &= \inf\{c \geq 0, \text{ t.q. } \mathbb{P}(|X| \leq c) = 1\} \\ &= \min\{c \geq 0, \text{ t.q. } \mathbb{P}(|X| \leq c) = 1\} \end{aligned}$$

→ La formulation suivante est équivalente

$$\|X\|_\infty = \max\{A \geq 0, \text{ t.q. } \mathbb{P}(|X| > A) > 0\}.$$

→ On utilise parfois la notation $\|X\|_p$, même si $X \notin L^p$. Dans ce cas, on a $\|X\|_p := +\infty$.

Exemple 2. Soit X une variable aléatoire discrète de loi

$$\mathbb{P}_X = \sum_{n=1}^{\infty} \frac{1}{2^n} \delta_n.$$

Déterminer $p \in [1, +\infty]$ tels que $X \in L^p$.

Solution. Soit $p \in [1, +\infty[$. On a, d'après le théorème du transport,

$$\mathbb{E}(|X|^p) = \sum_{n=1}^{\infty} \frac{1}{2^n} |n|^p < \infty$$

En conséquence, $\boxed{X \in L^p, \forall p \in [1, +\infty[.}$

Cas $p = \infty$. Pour tout $c \geq 0$,

$$\mathbb{P}(|X| > c) = \sum_{n>c} \frac{1}{2^n} > 0.$$

En définitive, $\mathbb{P}(|X| \leq c) < 1, \forall c > 0$, donc $\boxed{X \notin L^\infty.}$

Exemple 3. Soit X une v.a. de loi absolument continue de densité

$$f_X(x) = \mathbf{1}_{x\in]0,1]} \frac{2}{\sqrt{x}}, \ \forall x \in \mathbb{R}.$$

Déterminer $p \in [1, +\infty]$ tels que $X \in L^p$.

Solution. Soit $p \in [1, +\infty[$. On a, d'après le théorème du transport,

$$\mathbb{E}(|X|^p) = \int_{\mathbb{R}} |x|^p f_X(x) d\lambda_1(x) = \int_{]0,1]} 2\, x^{p-1/2} d\lambda_1(x).$$

En conséquence, $\boxed{X \in L^p, \forall p \in [1, +\infty[.}$

Cas $p = \infty$. On a $\mathbb{P}(|X| \leq 1) = \int_{\mathbb{R}} \mathbf{1}_{|x|\leq 1} f_X(x) d\lambda_1(x) = 1$,
donc $\boxed{\|X\|_\infty \leq 1 \text{ et } X \in L^\infty.}$
De plus, $\forall c < 1$, $\mathbb{P}(|X| \leq c) = \int_{\mathbb{R}} \mathbf{1}_{|x|\leq c} f_X(x) d\lambda_1(x) < 1$,
donc $\boxed{\|X\|_\infty = 1}$.

Exemple 4. Soit X une v.a. réelle de loi mixte

$$\mathbb{P}_X(dx) = \sum_{n=1}^{\infty} \frac{1}{3^n} \delta_{1/n}(dx) + \frac{1}{2} \mathbf{1}_{x \in [0,1]} \, \lambda_1(dx).$$

Montrer que $X \in L^\infty$ puis calculer $\|X\|_\infty$.

Solution. On a $\mathbb{P}(|X| \leq 1) = 1$,donc $\boxed{\|X\|_\infty \leq 1 \text{ et } X \in L^\infty}$.

Montrons que $\|X\|_\infty = 1$. Soit $A<1$. On a

$$\mathbb{P}(|X| > A) \geq \int_{\mathbb{R}} \mathbf{1}_{|x| \in]A,+\infty[} \frac{1}{2} \mathbf{1}_{x \in [0,1]} \, d\lambda_1(x) > 0.$$

Par conséquent, $\|X\|_\infty \geq A$, $\forall A < 1$, donc $\|X\|_\infty \geq 1$.

En définitive, $\boxed{\|X\|_\infty = 1.}$

Quelques inégalités utiles...

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité et X une v.a.

Inégalité de Jensen

Soit $\Phi : \mathbb{R} \to \mathbb{R}$ convexe. Alors, si les termes ci-dessous sont bien définis,

$$\Phi(\mathbb{E}(X)) \leq \mathbb{E}(\Phi(X)).$$

$\to$ "Φ de la moyenne est plus petit que la moyenne des Φ"

$\to$ On en déduit les inclusions suivantes

$$L^1 \supset L^p \supset L^{p'} \supset L^\infty, \forall 1 \leq p \leq p'$$

Inégalité de Hölder

Soient $p,q \in]1,+\infty[$ tels que $1/p + 1/q = 1$. Nous avons

$$\mathbb{E}(|XY|) \leq \mathbb{E}\left(|X|^p\right)^{1/p} \mathbb{E}\left(|Y|^q\right)^{1/q} = \|X\|_p \|Y\|_q.$$

Soient $p = 1$ et $q = \infty$. Nous avons

$$\mathbb{E}(|XY|) \leq \mathbb{E}(|X|)\|Y\|_\infty.$$

Extension. Soient $f,g : \Omega \to \mathbb{R}$ deux fonctions et μ une mesure, alors, pour tous $p,q \in]1,+\infty[$ tels que $1/p + 1/q = 1$,

$$\int_\Omega |f(\omega)g(\omega)|\,\mu(d\omega) \leq \left(\int_\Omega |f(\omega)|^p\,\mu(d\omega)\right)^{\frac{1}{p}} \left(\int_\Omega |g(\omega)|^q\,\mu(d\omega)\right)^{\frac{1}{q}}$$

Inégalité de Markov

Pour tout $a > 0$,

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(|X|)}{a}.$$

Inégalité de Bienaymée-Tchebychev

Si $X \in \mathcal{L}^2$, alors, pour tout $a > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathsf{Var}(X)}{a^2}.$$

Preuve de l'inégalité de Markov. $\mathbf{1}_{|X|\geq a} \leq \frac{|X|}{a}$ p.s., donc

$$\mathbb{P}(|X| \geq a) = \mathbb{E}\left(\mathbf{1}_{|X|\geq a}\right) \leq \mathbb{E}\left(\frac{|X|}{a}\right) \qquad \square$$

Exemple 5. Soit X de loi normale $\mathcal{N}(0,\sigma^2)$. Alors $\mathbb{E}(|X|) = \sqrt{\frac{2}{\pi}}\,\sigma$, donc

$$\boxed{\mathbb{P}(|X| \geq a) \leq \frac{\sigma}{a}\sqrt{\frac{2}{\pi}}} \quad \text{(inégalité de Markov).}$$

De plus, $\mathbb{E}(X) = 0$ et $\text{Var}(X) = \sigma^2$, donc

$$\boxed{\mathbb{P}(|X| \geq a) \leq \frac{\sigma^2}{a^2}} \quad \text{(inégalité de B-T).}$$

Ici, la seconde inégalité est meilleure pour $a > \sigma\sqrt{\frac{\pi}{2}}$.

L^2 est un espace de Hilbert

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité.

Théorème

L^2 muni de la norme $\|\cdot\|_2$ est un espace de Hilbert, ce qui signifie

1. $\|\cdot\|_2$ est issue du produit scalaire $\langle X,Y\rangle := \mathbb{E}(XY)$,
2. toute suite de Cauchy converge.

Rappel. $(u_n)_{n\geq 0}$ est appelée une **suite de Cauchy** si

$$\forall \epsilon > 0,\ \exists N_0 \text{ tel que } \forall m,n \geq N_0,\ \|u_m - u_n\| \leq \epsilon.$$

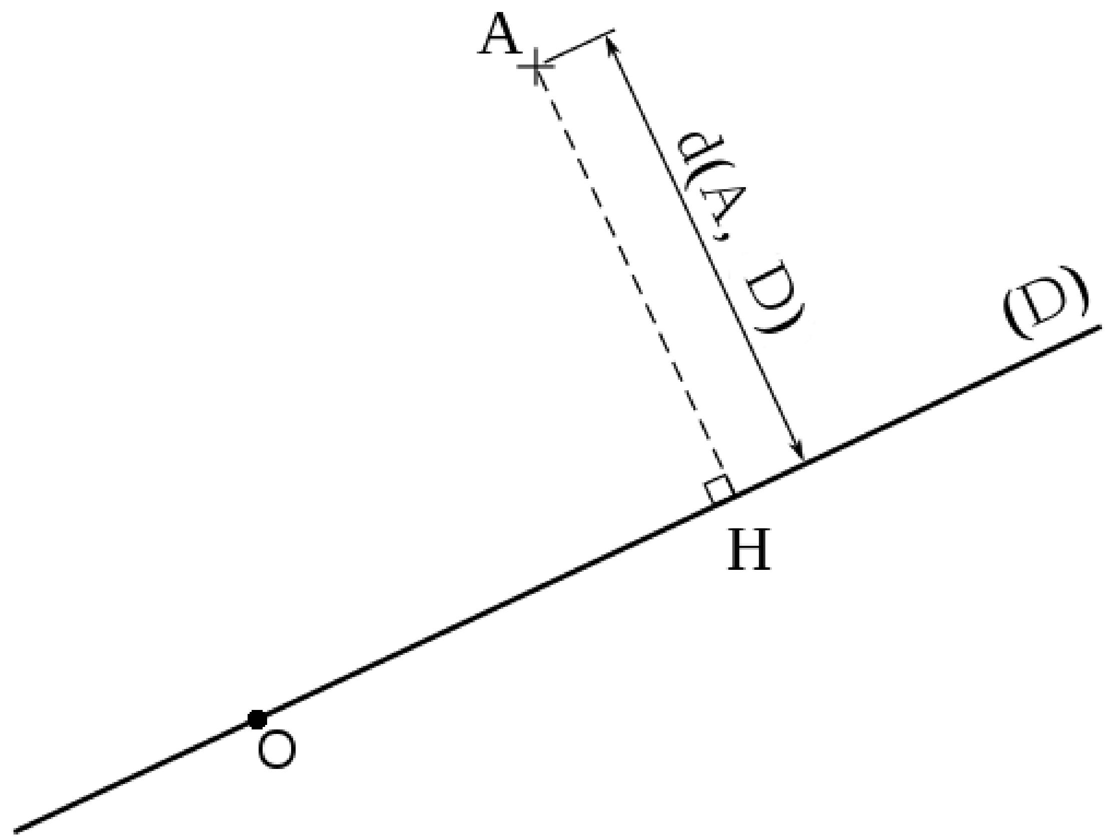
Remarque. Si $X,Y \in L^2$, alors

- X, Y et XY sont intégrables,
- On a l'inégalité de Cauchy-Schwartz

$$\langle X,Y\rangle = \mathbb{E}(XY) \leq \mathbb{E}\left(X^2\right)^{1/2}\mathbb{E}\left(Y^2\right)^{1/2} = \|X\|_2\|Y\|_2.$$

Notion fondamentale : la projection orthogonale.

$\rightarrow$ Dans $\mathbb{R}^2$.



$\rightarrow$ Le point H est l'unique point de D tel que $\vec{AH} \perp D$.

$\rightarrow$ Le point H est caractérisé par $H \in D$ et

$$|AH| = d(A,D) := \min_{M \in D} |AM|.$$

$\rightarrow$ Dans l'espace de Hilbert L^2.

Soit $E \subset L^2$ un sous ev fermé (par ex. de dimension finie).

1. Orthogonalité de $X, Y \in L^2$ signifie
$$X \perp Y \;\Leftrightarrow\; \langle X, Y \rangle = 0 \;\Leftrightarrow\; \mathbb{E}(XY) = 0,$$

2. Distance entre X et l'espace fermé E est
$$d(X,E) = \inf_{Y \in E} \|X - Y\|_2 = \inf_{Y \in E} \mathbb{E}\left((X - Y)^2\right)^{1/2}$$

Théorème

Soit $X \in L^2$. Il existe un unique élément de E, noté $P_E(X)$ et appelé **projection orthogonale de X sur E**, tel que
$$X - P_E(X) \perp Y, \forall Y \in E.$$

De plus $P_E(X)$ est l'unique élément de E tel que
$$\|X - P_E(X)\|_2 = \mathbb{E}\left((X - P_E(X))^2\right)^{1/2} = d(X,E).$$

Remarques.

- Les éléments de E, et *a fortiori* $P_E(X)$, sont des variables aléatoires.
- LA variable aléatoire $P_E(X)$ est la plus proche possible de X, c'est-à-dire que

$$\mathbb{E}\left(X - P_E(X))^2\right)^{1/2} = \min_{Y \in E} \mathbb{E}\left(X - Y)^2\right)^{1/2}.$$

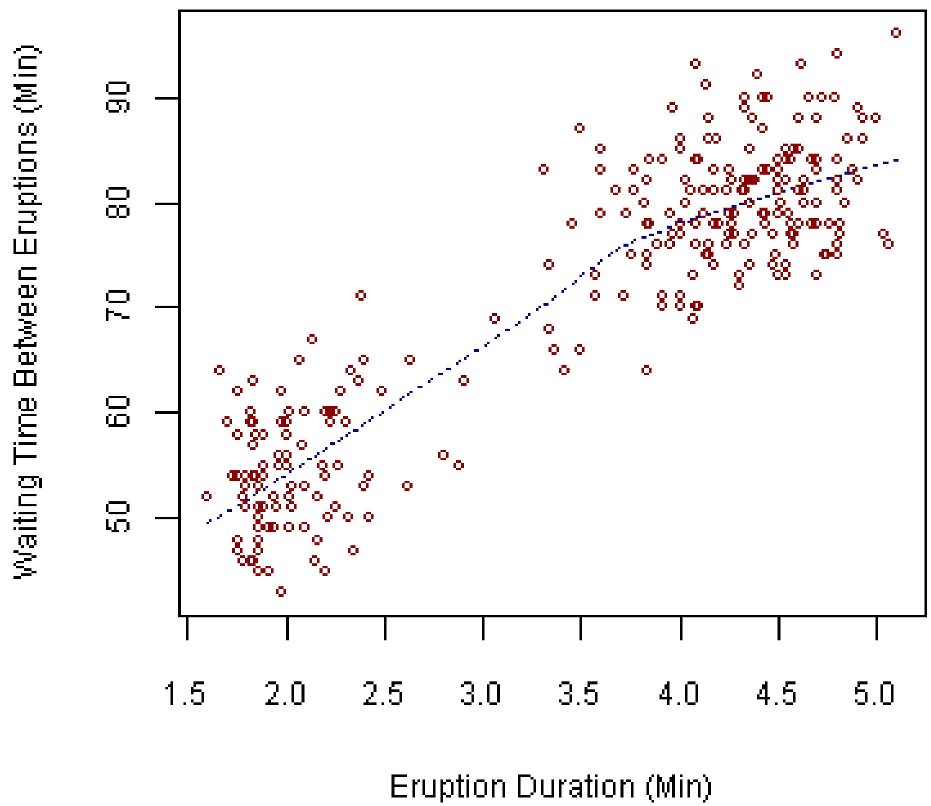
- Si $(Y_k)_{k \geq 0}$ est une base orthonormée de E, alors

$$\boxed{P_E(X) = \sum_{k \geq 0} \langle X, Y_k \rangle\, Y_k := \sum_{k \geq 0} \mathbb{E}\left(XY_k\right) Y_k.}$$

- Pour trouver une base orthonormée d'un espace E, on peut utiliser le procéder d'orthonormalisation de Gram-Schmidt.

Applications de la projection orthogonale

Application 1. Soient (X_1, Y_1),..., (X_n, Y_n) des relevés de données.



Meilleure approximation de X_i comme une fonction de Y_i ?

$\rightarrow$ Modélisation : on suppose que (X_1, Y_1),..., (X_n, Y_n) sont des tirages indépendants d'une même v.a. (X, Y).

- La meilleure approximation constante est donnée par

$$P_{E_1}(X), \text{ où } E_1 = \{a,\ a \in \mathbb{R}\}.$$

- La meilleure approximation linéaire est donnée par

$$P_{E_2}(X), \text{ où } E_2 = \{a + bY,\ a,b \in \mathbb{R}\}.$$

- La meilleure approximation quadratique est donnée par

$$P_{E_3}(X), \text{ où } E_3 = \{a + bY + cY^2,\ a,b,c \in \mathbb{R}\}.$$

Problème 1. Soit $X \in L^2$ et E un sous-espace vectoriel fermé qui contient les fonctions constantes. Montrer que

$$\mathbb{E}(P_E(X)) = \mathbb{E}(X).$$

Solution. Nous avons $X - P_E(X) \perp E$, donc $X - P_E(X) \perp \mathbf{1}$, car la fonction constante $\mathbf{1} \in E$. C'est-à-dire

$$\mathbb{E}(X - P_E(X)) = \langle X - P_E(X), 1 \rangle = 0.$$

De plus, $X \in L^2$ et $P_E(X) \in L^2$, donc X et $P_E(X)$ sont intégrables. Par linéarité de l'intégrale, on en déduit que

$$\mathbb{E}(P_E(X)) = \mathbb{E}(X).$$

Problème 2. Soit $X \in L^2$ et E un sous-espace vectoriel fermé qui contient les constantes. Montrer que

$$\mathsf{Var}(P_E(X)) \leq \mathsf{Var}(X).$$

Solution. Par Pythagore, on a

$$\mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}((X - P_E(X))^2) + \mathbb{E}((P_E(X) - \mathbb{E}(X))^2).$$

Or

$$\mathbb{E}((P_E(X) - \mathbb{E}(X))^2) = \mathbb{E}((P_E(X) - \mathbb{E}(P_E(X)))^2) = \mathsf{Var}(P_E(X)).$$

En définitive, $\mathsf{Var}(X) = \mathbb{E}((X - P_E(X))^2) + \mathsf{Var}(P_E(X))$, donc

$$\boxed{\mathsf{Var}(P_E(X)) \leq \mathsf{Var}(X).}$$

Application 2. Soit X une variable aléatoire. Estimation de $\mathbb{E}(X)$?
On simule $X_1, X_2, \ldots, X_n$ des copies indépendantes de X.

$\rightarrow$ Loi des grands nombres

$$\frac{1}{n}\sum_{i=1}^{n} X_i \simeq \mathbb{E}(X).$$

$\rightarrow$ Erreur sur l'estimation

$$\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}(X)\right| \sim \mathrm{Var}(X).$$

$\rightarrow$ *Meilleure* méthode : on simule $Y_1, Y_2, \ldots, Y_n$ des copies indépendantes de $P_E(X) \neq 0$.

$\rightarrow$ Loi des grands nombres

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \;\simeq\; \mathbb{E}(P_E(X)) = \mathbb{E}(X).$$

$\rightarrow$ Erreur sur l'estimation

$$\left|\frac{1}{n}\sum_{i=1}^{n} Y_i - \mathbb{E}(X)\right| \;\sim\; \mathsf{Var}(P_E(X)) \leq \mathsf{Var}(X).$$

Plus faible que dans le cas précédent !

Procédé appelé *Méthode de réduction de variance*.

Problème 3. Soit X une variable aléatoire de carré intégrable. Montrer que $P_{\mathbb{R}}(X) = \mathbb{E}(X)$, c'est-à-dire que

$$\mathbb{E}((X - \mathbb{E}(X))^2) = \min_{a \in \mathbb{R}} \mathbb{E}((X - a)^2).$$

Solution. La v.a. constante 1 est une base orthonormale de $\mathbb{R}$. On en déduit que

$$P_{\mathbb{R}}(X) = \langle X, 1 \rangle \, 1 = \mathbb{E}(X \cdot 1) \, 1 = \mathbb{E}(X).$$

$\Longrightarrow$ $\mathbb{E}(X)$ est la constante qui minimise $\mathbb{E}((X - a)^2)$.

Problème 4. Soient X, Y de carrés intégrables et $E = \{a + bY,\ a, b \in \mathbb{R}\}$. Calculer $P_E(X)$, la meilleure approximation de X par une fonction linéaire de Y.

Solution. Les v.a. $e_1 = 1$ et $e_2 = (Y - \mathbb{E}(Y))/\sqrt{\mathsf{Var}(Y)}$ forment une base orthonormale de E.
Donc

$$p_E(X) = \langle X, e_1 \rangle e_1 + \langle X, e_2 \rangle e_2$$

En reportant les expressions, on obtient

$$\boxed{p_E(X) = \mathbb{E}(X) + \frac{\mathsf{Cov}(X, Y)}{\mathsf{Var}(Y)}(Y - \mathbb{E}(Y))}.$$

Problème 5. Soit (X, Y) de densité $f_{(X,Y)}(x,y)$.
On définit l'espace vectoriel

$$E = \{f(Y), f \text{ mesurable bornée}\}.$$

1. Montrer que, pour tout $y \in \mathbb{R}$,

$$g_y : x \mapsto \frac{f_{(X,Y)}(x,y)}{\int_{\mathbb{R}} f_{(X,Y)}(x,y) d\lambda_1(x)}$$

est une densité sur $\mathbb{R}$.

2. Soit $h : \mathbb{R} \to \mathbb{R}$ définie par

$$h(y) = \int_{\mathbb{R}} x \, g_y(x) d\lambda_1(x).$$

Montrer que $P_E(X) = h(Y)$.

$\Longrightarrow$ On sait donc calculer la meilleure approximation de X comme une fonction de Y !

Solution.

1. Pour tout y, g_y est mesurable positive, d'intégrale 1 (Fubini).

2. Pour tout y, soit X_y de densité g_y.
Ainsi,

$$\mathbb{E}(X_y) = \int_{\mathbb{R}} x\, g_y(x)\, d\lambda_1(x) \overset{\text{déf}}{=} h(y).$$

Donc, pour tout $f(y)$,

$$\int_{\mathbb{R}} (x - h(y))^2\, g_y(x)\, d\lambda_1(x) \leq \int_{\mathbb{R}} (x - f(y))^2\, g_y(x)\, d\lambda_1(x).$$

donc, en multipliant par $\int_{\mathbb{R}} f_{(X,Y)}(x,y)\, d\lambda_1(x)$,

$$\int_{\mathbb{R}} (x - h(y))^2 f_{(X,Y)}(x,y)\, d\lambda_1(x) \leq \int_{\mathbb{R}} (x - f(y))^2 f_{(X,Y)}(x,y)\, d\lambda_1(x).$$

En intégrant par rapport à $d\lambda_1(y)$, on obtient

$$\boxed{\mathbb{E}((X - h(Y))^2) \leq \mathbb{E}((X - f(Y))^2),\ \forall f.}$$

Probabilités
8. Suites de variables aléatoires

´

1. Convergence presque-sûre

Définition

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire définies sur le même espace $(\Omega,\mathcal{F},\mathbb{P})$, à valeurs dans $\mathbb{R}^d$.

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X si et seulement si

$$\lim_{n\to\infty} X_n(\omega) = X(\omega), \qquad \mathbb{P}(\omega)\text{–ps}$$

c'est-à-dire si et seulement si $\mathbb{P}(X_n \xrightarrow[n\to\infty]{} X) = 1$.

On note souvent

$$\boxed{X_n \xrightarrow[n\to\infty]{ps} X.}$$

$\rightarrow$ Mêmes techniques que pour établir la convergence d'une suite déterministe $(u_n)_{n\in\mathbb{N}}$, car on raisonne à ω fixé.

$\rightarrow$ Généralisation des propriétés classiques connues pour la convergence d'une suite déterministe :

- Somme, multiplication par un scalaire,
- Multiplication de deux suites réelles convergeant p.s.,
- Suite image par une fonction continue.

$\rightarrow$ Exemple. $X_n \to X$ et $Y_n \to Y$ ps, alors

- $(X_n, Y_n) \to (X, Y)$ presque sûrement
- Si de plus $\mathbb{P}(Y_n = 0) = \mathbb{P}(Y = 0) = 0$ pour tout n, alors

$$\frac{X_n}{Y_n} \xrightarrow[n\to\infty]{ps} \frac{X}{Y}$$

$\rightarrow$ Exemple. $X_n \to X$ et $f : \mathbb{R} \to \mathbb{R}$ telle que $f(X_n)$ est bien définie ps et f continue en X presque sûrement, alors

$$f(X_n) \xrightarrow[n\to\infty]{ps} f(X).$$

Proposition (une caractérisation de la convergence presque sûre)

Une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X si et seulement si

$$\forall \varepsilon > 0, \mathbb{P}\left(\bigcap_{n=0}^{+\infty} \bigcup_{k=n}^{+\infty} \{|X_k - X| \geq \varepsilon\}\right) = 0,$$

c'est-à-dire que, pour tout $\epsilon > 0$,

$$\mathbb{P}\left(\text{il existe une infinité d'indices tels que } |X_k - X| \geq \varepsilon\right) = 0.$$

- On peut remplacer
  - $\geq \varepsilon$ par $> \varepsilon$,
  - $\forall \varepsilon > 0$ par $\forall \varepsilon \in]0, r[$,
  - $\forall \varepsilon > 0$ par une suite $(\varepsilon_p)_{p\in\mathbb{N}}$ décroissante convergeant vers 0, par exemple $\varepsilon_p = 1/p$.

Proposition (Lemme de Borel-Cantelli et Convergence p.s.)

Soient $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires.

- Si, pour tout $\varepsilon > 0$,
$$\sum_{n\in\mathbb{N}} \mathbb{P}\left(|X_n - X| > \varepsilon\right) < +\infty,$$
alors, (X_n) converge presque sûrement vers X.

- Si les variables aléatoires $(X_n - X)$ sont mutuellement indépendantes, alors la réciproque est vraie.

2. Convergence en probabilité et dans L^p

Définition (Convergence en probabilité)

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire définies sur le même espace $(\Omega,\mathcal{F},\mathbb{P})$, à valeurs dans $\mathbb{R}^d$.

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X si et seulement si, pour tout $\epsilon > 0$,

$$\mathbb{P}(|X_n - X| \geq \epsilon) \xrightarrow[n\to\infty]{} 0.$$

On note souvent

$$\boxed{X_n \xrightarrow[n\to\infty]{Proba.} X.}$$

$\rightarrow$ Remarque. La limite est unique presque sûrement, c'est-à-dire

$$\begin{cases} X_n \xrightarrow[n\to\infty]{Proba.} X \\ X_n \xrightarrow[n\to\infty]{Proba.} Y \end{cases} \Longrightarrow X = Y \text{ p.s.}$$

Définition (Convergence L^p)

Soit $p \in]0,\infty]$. Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire de $L^p(\Omega,\mathcal{F},\mathbb{P})$.

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge dans L^p vers X si et seulement si

$$\|X_n - X\|_p \xrightarrow[n\to\infty]{} 0.$$

On note souvent

$$\boxed{X_n \xrightarrow[n\to\infty]{L^p} X.}$$

$\rightarrow$ Remarque. La limite est unique presque sûrement.

Liens entre les convergences ps, proba et L^p

Proposition

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire définies sur le même espace $(\Omega,\mathcal{F},\mathbb{P})$, à valeurs dans $\mathbb{R}^d$.

- Si $X_n \xrightarrow[n\to\infty]{p.s} X$, **alors** $X_n \xrightarrow[n\to\infty]{Proba} X$.
- Si $\exists p \in]0,\infty]$ tq $X_n \xrightarrow[n\to\infty]{L^p} X$, **alors** $X_n \xrightarrow[n\to\infty]{Proba} X$.

De plus,

- Si $X_n \xrightarrow[n\to\infty]{p.s} X$ et si $\exists Y \in L^p$ tel que $|X_n| \leq |Y|$ pour tout $n \geq 1$, **alors** $X_n \xrightarrow[n\to\infty]{L^p} X$.
- Si $\exists p \in]0,\infty]$ tq $X_n \xrightarrow[n\to\infty]{L^p} X$, **alors** $X_n \xrightarrow[n\to\infty]{L^q} X$ $\forall q \in]0,p]$.

3. Convergence en loi

Définition

Soient (X_n) une suite de variables aléatoires et X une variable aléatoire à valeurs dans $\mathbb{R}^d$.

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X si et seulement si, pour toute fonction continue bornée $f : \mathbb{R}^d \to \mathbb{R}$,

$$\mathbb{E}(f(X_n)) \xrightarrow[n\to\infty]{} \mathbb{E}(f(X)).$$

On note souvent

$$\boxed{X_n \xrightarrow[n\to\infty]{\textit{Loi}} X.}$$

$\rightarrow$ Remarque. La limite n'est pas unique presque sûrement ! Seule la loi de la limite est uniquement déterminée.

Proposition (Liens avec les autres types de convergence)

Soient (X_n) une suite de variables aléatoires et X une variable aléatoire définies sur le même espace $(\Omega,\mathcal{F},\mathbb{P})$. Si

$$X_n \xrightarrow[n\to+\infty]{\text{p.s. ou } L^p \text{ ou Proba.}} X,$$

alors

$$\boxed{X_n \xrightarrow[n\to+\infty]{Loi} X.}$$

$\rightarrow$ Remarques.

- La réciproque est fausse en général.
- Toutefois, si les X_n sont toutes définies sur un même espace de proba $(\Omega,\mathcal{F},\mathbb{P})$ et si $X = a \in \mathbb{R}^d$ est déterministe, alors

$$X_n \xrightarrow[n\to+\infty]{Loi} a \iff X_n \xrightarrow[n\to+\infty]{Proba} a.$$

Caractérisation...

... par la fonction caractéristique (Théorème de Paul Lévy)

Soient (X_n) une suite de variables aléatoires et X une variable aléatoire à valeurs dans $\mathbb{R}^d$. On a $X_n \xrightarrow[n\to+\infty]{Loi} X$ si et seulement si

$$\forall u \in \mathbb{R}^d, \varphi_{X_n}(u) := \mathbb{E}\left(e^{i\langle u, X_n\rangle}\right) \xrightarrow[n\to+\infty]{} \varphi_X(u) = \mathbb{E}\left(e^{i\langle u, X\rangle}\right).$$

... par la fonction de répartition

Soient (X_n) une suite de variables aléatoires et X une variable aléatoire à valeurs dans $\mathbb{R}$. On a $X_n \xrightarrow[n\to+\infty]{Loi} X$ si et seulement si

$$F_{X_n}(t) := \mathbb{P}(X_n \leq t) \xrightarrow[n\to+\infty]{} F_X(t) := \mathbb{P}(X \leq t)$$

pour tout $t \in \mathbb{R}$ en lequel F_X est continue.

4. Les théorèmes de convergence

4a. La loi forte des grands nombres

Loi forte des grands nombres

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d. (indépendantes identiquement distribuées) à valeurs dans $\mathbb{R}^k$. Pour tout $n \in \mathbb{N}^*$, posons

$$\overline{X_n} = \frac{X_1 + \ldots + X_n}{n}.$$

1. Si $\mathbb{E}(|X_1|) < +\infty$, alors

$$\overline{X_n} \xrightarrow[n\to\infty]{p.s.,\ L^1} \mathbb{E}(X_1).$$

2. Si $\mathbb{E}(|X_1|) = +\infty$, alors, p.s. la suite $(\overline{X_n})_{n\in\mathbb{N}^*}$ diverge.

Ex 3. Estimation des moments. Soit $(X_n)_{n\geq 0}$ est une suite de variables aléatoires i.i.d. et $p>0$, montrer que

$$\frac{1}{n}\sum_{i=1}^{n}|X_i|^p \xrightarrow[n\to\infty]{p.s.} \mathbb{E}(|X_0|^p).$$

Solution. Deux situations

1. Si $X_0 \in L^p$, alors il suffit d'appliquer la loi des grands nombres à $Y_n = |X_n|^p$
2. Si $X_0 \notin L^p$, alors on pose, pour tout $K \geq 0$, $Y_n^K = |X_n|^p \wedge K$. Ainsi, $Y_n^K \in L^1$ et, d'après la loi forte des grands nombres,

$$\frac{1}{n}\sum_{i=1}^{n} Y_i^K \xrightarrow[n\to\infty]{p.s.} \mathbb{E}(Y_i^K) \xrightarrow[K\to\infty]{} \mathbb{E}(|X_0|^p) = \infty.$$

Or $\frac{1}{n}\sum_{i=1}^{n}|X_i|^p \geq \frac{1}{n}\sum_{i=1}^{n} Y_i^K$ donc

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}|X_i|^p = \infty = \mathbb{E}(|X_0|^p) \text{ p.s. } \to \square$$

4. Les théorèmes de convergence

4b. Le théorème central limite

Théorème (Théorème central limite)

Supposons $E = \mathbb{R}$. Si les v.a. $(X_n)_{n\in\mathbb{N}}$ sont i.i.d. de carré intégrable, alors

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}(X_1)\right) \xrightarrow[n\to+\infty]{Loi} Z \sim \mathcal{N}(0,\text{Var } X_1).$$

- Pour n grand, l'erreur entre $\overline{X_n}$ et $\mathbb{E}(X_1)$ est proche d'une variable aléatoire gaussienne divisée par $\sqrt{n}$.
- Généralisation à $E = \mathbb{R}^d$ en supposant $\mathbb{E}(\|X_n\|^2) < +\infty$
  $\rightarrow$ Dans ce cas, Z est un vecteur gaussien.

Ex 4. Intervalles de confiance. Soient $X \sim \mathcal{N}(m,\sigma^2)$ et $\alpha \in [0,1]$. Trouver $\beta > 0$ tel que

$$(*) \quad \mathbb{P}(X \in [m-\beta; m+\beta]) \simeq \alpha.$$

Solution. On sait que que $(X-m)/\sigma \sim \mathcal{N}(0,1)$ donc

$$(*) = \mathbb{P}((X-m)/\sigma \in [-\beta/\sigma, \beta/\sigma]) = \frac{1}{\sqrt{2\pi}} \int_{-\beta/\sigma}^{\beta/\sigma} e^{-x^2/2}\,dx.$$

Il reste simplement à trouver β tel que $(*) \simeq \alpha$.

Exemple. $\alpha = 95\%$. On sait (grâce aux tables) que

$$\frac{1}{\sqrt{2\pi}} \int_{-1{,}96}^{1{,}96} e^{-x^2/2}\,dx \simeq 95\%.$$

donc, avec $\beta = 1{,}96\sigma$, $\boxed{\mathbb{P}(X \in [m-1{,}96\sigma; m+1{,}96\sigma]) \simeq 95\%.}$

$[m-1{,}96\sigma; m+1{,}96\sigma]$ est l'<u>intervalle de confiance à 95% de X</u>.

Ex 5. Intervalles de confiance empiriques. Soit $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, avec $X_i \sim \mathcal{B}(p)$. On cherche $\beta > 0$ tel que

$$\mathbb{P}(p \in [\bar{X} - \beta, \bar{X} + \beta]) \simeq 95\%.$$

Solution. D'après le théorème central limite,

$$\sqrt{n/\text{Var}(X_1)}(\bar{X} - \mathbb{E}(X_1)) = \sqrt{\frac{n}{p(1-p)}}(\bar{X} - p) \simeq Z \sim \mathcal{N}(0,1)$$

De plus, $p(1-p) = \mathbb{E}(X_1)(1 - \mathbb{E}(X_1)) \simeq \bar{X}(1 - \bar{X})$ (LGN).

En conséquence, $\boxed{\sqrt{\frac{n}{\bar{X}(1-\bar{X})}}(\bar{X} - p) \simeq Z \sim \mathcal{N}(0,1)}$.

Donc

$$\mathbb{P}\left(-1{,}96 \leq \frac{\sqrt{n}(X - p)}{\sqrt{X(1-X)}} \leq 1{,}96\right) \simeq 95\%.$$

Soit, de manière équivalente,

$$\boxed{\mathbb{P}\left(X - \frac{1{,}96\sqrt{X(1-X)}}{\sqrt{n}} \leq p \leq X + \frac{1{,}96\sqrt{X(1-X)}}{\sqrt{n}}\right) \simeq 95\%.}$$

Application : Intervalles de confiance pour un sondage

Ex 6. Un sondage pré-élection avec 1000000 de votants potentiels effectué sur un échantillon de 1000 personnes donne à un candidat 23% d'intention de vote. Quelle est l'intervalle de confiance à 95% de ce résultat ?

Solution. On suppose que chaque votant est une réalisation indépendante d'une loi de Bernoulli de paramètre p, où p est le score réel du candidat.

On applique les résultats précédents à

$$X = 23\% \text{ et } n = 1000.$$

On en déduit que l'intervalle de confiance à 95% de p est

$$\boxed{\left[X - \frac{1{,}96\sqrt{X(1-X)}}{\sqrt{n}}, X + \frac{1{,}96\sqrt{X(1-X)}}{\sqrt{n}}\right] = [20.4\%, 25.6\%]}$$

Intégration et Probabilités
Chapitre 9. Vecteurs gaussiens

´

1. Rappels d'algèbre linéaire

Définition d'une matrice symétrique positive

Une matrice réelle A de taille $d \times d$ est dite symétrique positive si et seulement si

$$A = {}^t A \qquad \text{(symétrie)}$$

et, pour tout vecteur colonne $x \in \mathbb{R}^d$,

$${}^t x A x \geq 0. \qquad \text{(positive)}$$

Définition d'une matrice symétrique positive

Une matrice réelle A de taille $d \times d$ est dite <u>symétrique positive</u> si et seulement si

$$A = {}^tA \qquad \text{(symétrie)}$$

et, pour tout vecteur colonne $x \in \mathbb{R}^d$,

$${}^txAx \geq 0. \qquad \text{(positive)}$$

$\rightarrow$ Calcul

- Si $x = {}^t(x_1, \ldots, x_d) \in \mathbb{R}^d$, alors

$${}^txAx = \sum_{i=1}^{d} a_{ii}x_i^2 + \sum_{i \neq j}^{d} a_{ij}\, x_i x_j\,.$$

Ex. 1. Matrice diagonale à coefficients positifs. Montrer que

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ est symétrique positive.}$$

Ex. 1. Matrice diagonale à coefficients positifs. Montrer que

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ est symétrique positive.}$$

Solution.

- A_1 est bien symétrique,
- Pour tout $x = {}^t(x_1, x_2, x_3)$, on a

$$\begin{aligned} {}^t x\, A_1\, x &= (x_1 \;\; x_2 \;\; x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= x_1^2 + 3x_2^2 + 2x_3^2 \;\geq 0. \end{aligned}$$

Par conséquent,

$$\boxed{A_1 \text{ est une matrice symétrique positive.}}$$

Caractérisation des matrices symétriques positives I

Soit A une matrice symétrique. La matrice A est symétrique positive si et seulement si toutes ses valeurs propres sont positives.

Caractérisation des matrices symétriques positives I

Soit A une matrice symétrique. La matrice A est symétrique positive si et seulement si toutes ses valeurs propres sont positives.

Caractérisation des matrices symétriques positives II

Soit A une matrice symétrique. La matrice A est symétrique positive si et seulement si il existe une matrice B symétrique telle que $A = B^2$.

Caractérisation des matrices symétriques positives I

Soit A une matrice symétrique. La matrice A est symétrique positive si et seulement si toutes ses valeurs propres sont positives.

Caractérisation des matrices symétriques positives II

Soit A une matrice symétrique. La matrice A est symétrique positive si et seulement si il existe une matrice B symétrique telle que $A = B^2$.

$\longrightarrow$ Remarques

- B est appelée une racine carrée de A
- Étant donné A symétrique positive, on peut toujours une racine carrée qui est symétrique positive.
- Cependant il existe des racines carrées de A qui ne sont pas positives.

Méthode : Réduction d'une forme quadratique. En pratique, il est souvent plus facile de démontrer à la main que ${}^t xAx$ est positif pour tout $x \in \mathbb{R}^d$. Pour cela,

1. On calcule explicitement ${}^t xAx$ en fonction de $x_1, \ldots, x_d$.
2. On cherche à écrire cette expression sous la forme d'une combinaison linéaire de termes au carré
3. Deux possibilités :
   - Soit chaque terme est à coefficient positif, alors ${}^t xAx$ est positif pour tout x, donc A est symétrique positive,
   - Soit il existe au moins un terme à coefficient négatif, alors on peut chercher x tel que ${}^t xAx < 0$, pour montrer que A n'est pas symétrique positive.

Ex 2. La matrice

$$A_2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \text{ est-elle symétrique positive ?}$$

Ex 2. La matrice

$$A_2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \text{ est-elle symétrique positive ?}$$

Solution. La matrice A_2 est bien symétrique.

1. Pour tout $x = {}^t(x_1,x_2,x_3) \in \mathbb{R}^3$, nous avons

$$\begin{aligned} {}^t xAx &= (x_1 \ x_2 \ x_3) \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= 2x_1^2 + 2x_1x_2 + 2x_1x_3 + 2x_2^2 + 2x_2x_3 + 2x_3^2 \\ &= 2(x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2). \end{aligned}$$

   Il nous reste à *réduire* $x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2$.

Il nous reste à *réduire* $(*)$ $x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2$.

2. Nous avons, pour tout $x = {}^t(x_1, x_2, x_3) \in \mathbb{R}^3$,

$$x_1^2 + x_1x_2 + x_1x_3 = (x_1 + x_2/2 + x_3/2)^2 - x_2^2/4 - x_2x_3/2 - x_3^2/4.$$

Il nous reste à *réduire* $(*)\ x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2$.

2. Nous avons, pour tout $x = {}^t(x_1, x_2, x_3) \in \mathbb{R}^3$,

$$x_1^2 + x_1x_2 + x_1x_3 = (x_1 + x_2/2 + x_3/2)^2 - x_2^2/4 - x_2x_3/2 - x_3^2/4.$$

Donc

$$(*) = (x_1 + x_2/2 + x_3/2)^2 + \frac{3}{4}x_2^2 + \frac{1}{2}x_2x_3 + \frac{3}{4}x_3^2.$$

Il nous reste à *réduire* $(*)\ x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2$.

2. Nous avons, pour tout $x = {}^t(x_1, x_2, x_3) \in \mathbb{R}^3$,

$$x_1^2 + x_1x_2 + x_1x_3 = (x_1 + x_2/2 + x_3/2)^2 - x_2^2/4 - x_2x_3/2 - x_3^2/4.$$

Donc

$$(*) = (x_1 + x_2/2 + x_3/2)^2 + \frac{3}{4}x_2^2 + \frac{1}{2}x_2x_3 + \frac{3}{4}x_3^2.$$

Or

$$\frac{3}{4}x_2^2 + \frac{1}{2}x_2x_3 = \frac{3}{4}(x_2 + x_3/3)^2 - x_3^2/9.$$

Il nous reste à *réduire* $(*)\ x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2$.

2. Nous avons, pour tout $x = {}^t(x_1,x_2,x_3) \in \mathbb{R}^3$,

$$x_1^2 + x_1x_2 + x_1x_3 = (x_1 + x_2/2 + x_3/2)^2 - x_2^2/4 - x_2x_3/2 - x_3^2/4.$$

Donc

$$(*) = (x_1 + x_2/2 + x_3/2)^2 + \frac{3}{4}x_2^2 + \frac{1}{2}x_2x_3 + \frac{3}{4}x_3^2.$$

Or

$$\frac{3}{4}x_2^2 + \frac{1}{2}x_2x_3 = \frac{3}{4}(x_2 + x_3/3)^2 - x_3^2/9.$$

Donc

$$(*) = (x_1 + x_2/2 + x_3/2)^2 + \frac{3}{4}(x_2 + x_3/3)^2 + \frac{23}{36}x_3^2 \quad \geq 0.$$

Il nous reste à *réduire* $(*)$ $x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2$.

2. Nous avons, pour tout $x = {}^t(x_1,x_2,x_3) \in \mathbb{R}^3$,

$$x_1^2 + x_1x_2 + x_1x_3 = (x_1 + x_2/2 + x_3/2)^2 - x_2^2/4 - x_2x_3/2 - x_3^2/4.$$

Donc

$$(*) = (x_1 + x_2/2 + x_3/2)^2 + \frac{3}{4}x_2^2 + \frac{1}{2}x_2x_3 + \frac{3}{4}x_3^2.$$

Or

$$\frac{3}{4}x_2^2 + \frac{1}{2}x_2x_3 = \frac{3}{4}(x_2 + x_3/3)^2 - x_3^2/9.$$

Donc

$$(*) = (x_1 + x_2/2 + x_3/2)^2 + \frac{3}{4}(x_2 + x_3/3)^2 + \frac{23}{36}x_3^2 \quad \geq 0.$$

En définitive, $\boxed{A_2 \text{ est symétrique positive.}}$

Ex 3. Recherche d'un contre-exemple. La matrice

$$A_3 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix} \text{ est-elle symétrique positive ?}$$

Solution. La matrice A_3 est bien symétrique.

1. Nous avons ${}^t x A_3 x = x_1^2 + 2x_1x_2 + x_2^2 + 2x_2x_3 + x_3^2$.

Ex 3. Recherche d'un contre-exemple. La matrice

$$A_3 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix} \text{ est-elle symétrique positive ?}$$

Solution. La matrice A_3 est bien symétrique.

1. Nous avons ${}^t x A_3 x = x_1^2 + 2x_1x_2 + x_2^2 + 2x_2x_3 + x_3^2$.
2.

$$\begin{aligned} {}^t x A_3 x &= (x_1 + x_2)^2 + 2x_2x_3 + x_3^2 \\ &= (x_1 + x_2)^2 + (x_2 + x_3)^2 - x_2^2. \end{aligned}$$

Ex 3. Recherche d'un contre-exemple. La matrice

$$A_3 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix} \text{ est-elle symétrique positive ?}$$

Solution. La matrice A_3 est bien symétrique.

1. Nous avons ${}^t x A_3 x = x_1^2 + 2x_1x_2 + x_2^2 + 2x_2x_3 + x_3^2$.
2. $$\begin{aligned} {}^t x A_3 x &= (x_1 + x_2)^2 + 2x_2x_3 + x_3^2 \\ &= (x_1 + x_2)^2 + (x_2 + x_3)^2 - x_2^2. \end{aligned}$$
3. Attention, ça ne suffit pas, *a priori*, pour conclure !

   $\rightarrow$ Pour $x = (-1,1,-1)$, nous avons ${}^t x A x = -1 < 0$.

Ex 3. Recherche d'un contre-exemple. La matrice

$$A_3 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix} \text{ est-elle symétrique positive ?}$$

Solution. La matrice A_3 est bien symétrique.

1. Nous avons ${}^t x A_3 x = x_1^2 + 2x_1x_2 + x_2^2 + 2x_2x_3 + x_3^2$.
2. $$\begin{aligned} {}^t x A_3 x &= (x_1 + x_2)^2 + 2x_2x_3 + x_3^2 \\ &= (x_1 + x_2)^2 + (x_2 + x_3)^2 - x_2^2. \end{aligned}$$
3. Attention, ça ne suffit pas, *a priori*, pour conclure !

   $\rightarrow$ Pour $x = (-1,1,-1)$, nous avons ${}^t x A x = -1 < 0$.

   Par conséquent, $\boxed{A_3 \text{ n'est pas symétrique positive.}}$

Ex 4. Recherche d'un contre-exemple. La matrice

$$A_4 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \text{ est-elle symétrique positive ?}$$

Solution. La matrice A_4 est bien symétrique.

1. Nous avons ${}^t x A_4 x = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$.

Ex 4. Recherche d'un contre-exemple. La matrice

$$A_4 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$ est-elle symétrique positive ?

Solution. La matrice A_4 est bien symétrique.

1. Nous avons ${}^t x A_4 x = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$.

   Nous remarquons que, pour $x = (0,1,-1)$, nous avons

$$ {}^t x A_4 x = -2 < 0. $$

Ex 4. Recherche d'un contre-exemple. La matrice

$$A_4 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$ est-elle symétrique positive ?

Solution. La matrice A_4 est bien symétrique.

1. Nous avons ${}^t x A_4 x = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$.

   Nous remarquons que, pour $x = (0{,}1, -1)$, nous avons

   $$ {}^t x A_4 x = -2 < 0. $$

   Par conséquent,

   $$\boxed{A_4 \text{ n'est pas symétrique positive.}}$$

Ex 4. Recherche d'un contre-exemple. La matrice

$$A_4 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$ est-elle symétrique positive ?

Solution. La matrice A_4 est bien symétrique.

1. Nous avons ${}^t x A_4 x = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$.

   Nous remarquons que, pour $x = (0,1,-1)$, nous avons

$$ {}^t x A_4 x = -2 < 0. $$

   Par conséquent,

$$\boxed{A_4 \text{ n'est pas symétrique positive.}}$$

2 & 3. Étapes de calcul inutiles ici.

2. Définitions

Définition

Un vecteur aléatoire $X = {}^t(X_1, \ldots, X_d)$ à valeurs dans $\mathbb{R}^d$ est appelé un **vecteur gaussien** si et seulement si, pour tout $(a_1, \ldots, a_d) \in \mathbb{R}^d$,

$$\sum_{i=1}^{d} a_i X_i \text{ est une variable aléatoire réelle gaussienne.}$$

Définition

Un vecteur aléatoire $X = {}^t(X_1, \dots, X_d)$ à valeurs dans $\mathbb{R}^d$ est appelé un **vecteur gaussien** si et seulement si, pour tout $(a_1, \dots, a_d) \in \mathbb{R}^d$,

$$\sum_{i=1}^{d} a_i X_i \text{ est une variable aléatoire réelle gaussienne.}$$

→ Remarques

- Une variable aléatoire réelle constante égale à $m \in \mathbb{R}$ presque sûrement est considérée de loi gaussienne $\mathcal{N}(m,0)$.

Définition

Un vecteur aléatoire $X = {}^t(X_1, \ldots, X_d)$ à valeurs dans $\mathbb{R}^d$ est appelé un **vecteur gaussien** si et seulement si, pour tout $(a_1, \ldots, a_d) \in \mathbb{R}^d$,

$$\sum_{i=1}^{d} a_i X_i \text{ est une variable aléatoire réelle gaussienne.}$$

→ Remarques

- Une variable aléatoire réelle constante égale à $m \in \mathbb{R}$ presque sûrement est considérée de loi gaussienne $\mathcal{N}(m,0)$.
- Pour tout i, X_i est une variable aléatoire gaussienne.

Définition

Un vecteur aléatoire $X = {}^t(X_1, \ldots, X_d)$ à valeurs dans $\mathbb{R}^d$ est appelé un **vecteur gaussien** si et seulement si, pour tout $(a_1, \ldots, a_d) \in \mathbb{R}^d$,

$$\sum_{i=1}^{d} a_i X_i \text{ est une variable aléatoire réelle gaussienne.}$$

$\rightarrow$ Remarques

- Une variable aléatoire réelle constante égale à $m \in \mathbb{R}$ presque sûrement est considérée de loi gaussienne $\mathcal{N}(m,0)$.
- Pour tout i, X_i est une variable aléatoire gaussienne.
- Pour tout $i_1, \ldots, i_k \subset \{1, \ldots, d\}$, $Y = (X_{i_1}, \ldots, X_{i_k})$ est un vecteur gaussien.

Définition

Un vecteur aléatoire $X = {}^t(X_1,\ldots,X_d)$ à valeurs dans $\mathbb{R}^d$ est appelé un **vecteur gaussien** si et seulement si, pour tout $(a_1,\ldots,a_d) \in \mathbb{R}^d$,

$$\sum_{i=1}^{d} a_i X_i \text{ est une variable aléatoire réelle gaussienne.}$$

→ Remarques

- Une variable aléatoire réelle constante égale à $m \in \mathbb{R}$ presque sûrement est considérée de loi gaussienne $\mathcal{N}(m,0)$.
- Pour tout i, X_i est une variable aléatoire gaussienne.
- Pour tout $i_1,\ldots,i_k \subset \{1,\ldots,d\}$, $Y = (X_{i_1},\ldots,X_{i_k})$ est un vecteur gaussien.
- Un vecteur composé de d variables aléatoires gaussiennes mutuellement indépendantes est un cas particulier de vecteur gaussien (voir TD chap. 6).

Définition

Soit $X = {}^t(X_1, \ldots, X_d)$ un vecteur (gaussien). On appelle **vecteur moyenne** de X le vecteur

$$m_X = {}^t\left(\mathbb{E}(X_1), \ldots, \mathbb{E}(X_d)\right).$$

On appelle **matrice de covariance** de X la matrice $d \times d$

$$\Gamma_X = \left[\mathrm{Cov}(X_i, X_j)\right]_{i,j \in \{1,\ldots,d\}}.$$

$\rightarrow$ Remarques

- La matrice de covariance est symétrique positive (montrez le !).
- Si les $X_1, \ldots, X_d$ sont indépendants, alors Γ_X est diagonale.

Ex 5. Un cas particulier. Soient $X_1, \ldots, X_d$ des variables aléatoires réelles indépendantes de loi $\mathcal{N}(0,1)$. Calculer le vecteur moyenne et la matrice de covariance de $X = (X_1, \ldots, X_d)$.

Ex 5. Un cas particulier. Soient $X_1, \dots, X_d$ des variables aléatoires réelles indépendantes de loi $\mathcal{N}(0,1)$. Calculer le vecteur moyenne et la matrice de covariance de $X = (X_1, \dots, X_d)$.

Solution. Pour tout $i \in \{1, \dots, d\}$, $\mathbb{E}(X_i) = 0$, donc

$$\boxed{m_X = {}^t(0, \dots, 0).}$$

Ex 5. Un cas particulier. Soient $X_1,\dots,X_d$ des variables aléatoires réelles indépendantes de loi $\mathcal{N}(0,1)$. Calculer le vecteur moyenne et la matrice de covariance de $X = (X_1,\dots,X_d)$.

Solution. Pour tout $i \in \{1,\dots,d\}$, $\mathbb{E}(X_i) = 0$, donc

$$\boxed{m_X = {}^t(0,\dots,0).}$$

De plus, pour tous $i,j \in \{1,\dots,d\}$,

$$\mathrm{Cov}(X_i,X_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Ex 5. Un cas particulier. Soient $X_1, \ldots, X_d$ des variables aléatoires réelles indépendantes de loi $\mathcal{N}(0,1)$. Calculer le vecteur moyenne et la matrice de covariance de $X = (X_1, \ldots, X_d)$.

Solution. Pour tout $i \in \{1, \ldots, d\}$, $\mathbb{E}(X_i) = 0$, donc

$$\boxed{m_X = {}^t(0, \ldots, 0).}$$

De plus, pour tous $i,j \in \{1, \ldots, d\}$,

$$\text{Cov}(X_i, X_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Par conséquent,

$$\boxed{\Gamma_X = I_d \text{ (la matrice identité de } \mathcal{M}_{d \times d}(\mathbb{R})).}$$

Proposition

Soit X un vecteur gaussien de moyenne $m_X \in \mathbb{R}^d$ et de matrice de covariance $\Gamma_X \in \mathcal{M}_{d\times d}$. La fonction caractéristique de X est donnée, pour tout $u \in \mathbb{R}^d$, par

$$\varphi_X(u) = \mathbb{E}\left(e^{i\langle u,X\rangle}\right) = \exp\left(i\langle u,m_X\rangle - \frac{{}^t u\,\Gamma_X\,u}{2}\right).$$

Proposition

Soit X un vecteur gaussien de moyenne $m_X \in \mathbb{R}^d$ et de matrice de covariance $\Gamma_X \in \mathcal{M}_{d\times d}$. La fonction caractéristique de X est donnée, pour tout $u \in \mathbb{R}^d$, par

$$\varphi_X(u) = \mathbb{E}\left(e^{i\langle u,X\rangle}\right) = \exp\left(i\langle u,m_X\rangle - \frac{{}^t u\,\Gamma_X\,u}{2}\right).$$

Proposition

Soient X et Y deux vecteurs aléatoires gaussiens. Ils ont même loi si et seulement si

$$m_X = m_Y \quad \text{et} \quad \Gamma_X = \Gamma_Y.$$

En d'autres termes le **vecteur moyenne et la matrice de covariance** d'un vecteur gaussien caractérisent sa loi.

Définition

La loi d'un vecteur gaussien de vecteur moyenne $m = {}^t(m_1, \ldots, m_d) \in \mathbb{R}^d$ et de matrice de covariance $\Gamma \in \mathcal{M}_{d \times d}$ est notée

$$\mathcal{N}(m, \Gamma).$$

Définition

La loi d'un vecteur gaussien de vecteur moyenne $m = {}^t(m_1, \ldots, m_d) \in \mathbb{R}^d$ et de matrice de covariance $\Gamma \in \mathcal{M}_{d\times d}$ est notée

$$\mathcal{N}(m,\Gamma).$$

Proposition

Soit X un vecteur gaussien de moyenne $m_X \in \mathbb{R}^d$ et de matrice de covariance $\Gamma_X \in \mathcal{M}_{d\times d}$. La loi de X admet **une densité par rapport à la mesure de Lebesgue** λ_d si et seulement si Γ_X **est inversible**, c'est-à-dire si

$$\det \Gamma_X \neq 0.$$

Dans le cas contraire, X est dit dégénéré.

3. Propriétés d'un vecteur gaussien

Propriété

Soient $X = {}^t(X_1, \ldots, X_d)$ un vecteur gaussien, $A \in \mathcal{M}_{n\times d}$ et $b \in \mathbb{R}^n$. Alors le vecteur aléatoire de dimension n défini par

$$Y = AX + b$$

est un vecteur gaussien de moyenne et de matrice de covariance

$$m_Y = Am_X + b \quad \text{et} \quad \Gamma_Y = A\Gamma_X {}^t A.$$

Remarque

- Pour se souvenir de l'emplacement de la transposée dans $\Gamma_Y = A\Gamma_X {}^t A$, il suffit de regarder les dimensions. En effet, le nombre de lignes de A donne la dimension de Y et par conséquent la dimension de la matrice carrée Γ_Y. Ainsi, A et Γ_Y ont le même nombre de lignes.

Ex 6. Soit $X = {}^t(X_1, \dots, X_4)$ un vecteur gaussien tel que

$$m_X = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix} \text{ et } \Gamma_X = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{pmatrix}.$$

Montrer que $Y = {}^t(X_1 + 2X_2 + 1, X_2 - X_3)$ est un vecteur gaussien et donner sa loi en fonction de m_X et Γ_X.

Ex 6. Soit $X = {}^t(X_1, \ldots, X_4)$ un vecteur gaussien tel que

$$m_X = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix} \text{ et } \Gamma_X = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{pmatrix}.$$

Montrer que $Y = {}^t(X_1 + 2X_2 + 1, X_2 - X_3)$ est un vecteur gaussien et donner sa loi en fonction de m_X et Γ_X.

Solution. Nous avons

$$Y = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} X + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = AX + b.$$

Ex 6. Soit $X = {}^t(X_1, \dots, X_4)$ un vecteur gaussien tel que

$$m_X = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix} \text{ et } \Gamma_X = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{pmatrix}.$$

Montrer que $Y = {}^t(X_1 + 2X_2 + 1, X_2 - X_3)$ est un vecteur gaussien et donner sa loi en fonction de m_X et Γ_X.

Solution. Nous avons

$$Y = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} X + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = AX + b.$$

Donc Y est un vecteur gaussien de loi $\mathcal{N}(m_Y, \Gamma_Y)$, où

$$m_Y = Am_X + b = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ et } \Gamma_Y = A\Gamma_X {}^tA = \begin{pmatrix} 14 & 3 \\ 3 & 2 \end{pmatrix}.$$

Proposition

Soient $m \in \mathbb{R}^d$ et $\Gamma \in \mathcal{M}_{d\times d}$ une matrice symétrique positive. Alors il existe un vecteur gaussien X de vecteur moyenne m et de matrice de covariance Γ.

Proposition

Soient $m \in \mathbb{R}^d$ et $\Gamma \in \mathcal{M}_{d \times d}$ une matrice symétrique positive. Alors il existe un vecteur gaussien X de vecteur moyenne m et de matrice de covariance Γ.

Preuve. Soit $Z = (Z_1, \ldots, Z_d)$ un vecteur de variables aléatoires réelles i.i.d. de loi $\mathcal{N}(0,1)$.

- Z est un vecteur gaussien de vecteur moyenne $m_Z = 0_d$ et de matrice de covariance $\Gamma_Z = I_d$.
- Γ étant symétrique positive, il existe $B \in \mathcal{M}_{d \times d}$ symétrique telle que $\Gamma = B^2$.

Posons $X = BZ + m$.

- X est un vecteur gaussien de moyenne et de matrice de covariance

$$\boxed{m_X = Bm_Z + m = m \quad \text{et} \quad \Gamma_X = B\Gamma_Z{}^tB = B^2 = \Gamma.}$$

Ex 7. Montrer qu'il existe un vecteur gaussien de matrice de covariance et de vecteur moyenne

$$\Gamma = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \text{ et } m = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

Ex 7. Montrer qu'il existe un vecteur gaussien de matrice de covariance et de vecteur moyenne

$$\Gamma = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \text{ et } m = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

Solution. La matrice Γ est symétrique positive, comme nous l'avons montré à l'exemple 1.

D'après la proposition précédente, nous en déduisons qu'il existe un vecteur gaussien de vecteur moyenne m et de matrice de covariance Γ.

Proposition : indépendance dans un vecteur gaussien

Soit $X = {}^t(X_1, \ldots, X_d)$ un vecteur gaussien.

1. Pour tous $i \neq j \in \{1, \ldots, d\}$, X_i et X_j sont indépendants si et seulement si $\mathrm{Cov}(X_i, X_j) = 0$.

Proposition : indépendance dans un vecteur gaussien

Soit $X = {}^t(X_1, \ldots, X_d)$ un vecteur gaussien.

1. Pour tous $i \neq j \in \{1, \ldots, d\}$, X_i et X_j sont indépendants si et seulement si $\mathrm{Cov}(X_i, X_j) = 0$.
2. Soit $I = \{i_1 < \cdots < i_k\} \subset \{1, \ldots, d\}$. Les variables aléatoires $X_{i_1}, \ldots, X_{i_k}$ sont mutuellement indépendantes si et seulement si

$$\mathrm{Cov}(X_i, X_j) = 0, \ \forall i \neq j \in I.$$

Proposition : indépendance dans un vecteur gaussien

Soit $X = {}^t(X_1, \ldots, X_d)$ un vecteur gaussien.

1. Pour tous $i \neq j \in \{1, \ldots, d\}$, X_i et X_j sont indépendants si et seulement si $\mathrm{Cov}(X_i, X_j) = 0$.
2. Soit $I = \{i_1 < \cdots < i_k\} \subset \{1, \ldots, d\}$. Les variables aléatoires $X_{i_1}, \ldots, X_{i_k}$ sont mutuellement indépendantes si et seulement si

$$\mathrm{Cov}(X_i, X_j) = 0, \ \forall i \neq j \in I.$$

→ Remarques

- Pour les composantes d'un vecteur gaussien, l'indépendance mutuelle est équivalente à l'indépendance 2 à 2.
- Pour les composantes d'un vecteur gaussien, l'indépendance est équivalente à la non-corrélation.

Ex 8. Soit $X = {}^t(X_1, X_2, X_3)$ un vecteur gaussien de matrice de covariance

$$\Gamma_X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix}.$$

Montrer que X_1 et $X_2 - X_3$ sont indépendants.

Ex 8. Soit $X = {}^t(X_1, X_2, X_3)$ un vecteur gaussien de matrice de covariance

$$\Gamma_X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix}.$$

Montrer que X_1 et $X_2 - X_3$ sont indépendants.

Solution. Le vecteur $Y = {}^t(X_1, X_2 - X_3)$ est un vecteur gaussien, car

$$Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} X.$$

Ex 8. Soit $X = {}^t(X_1,X_2,X_3)$ un vecteur gaussien de matrice de covariance

$$\Gamma_X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix}.$$

Montrer que X_1 et $X_2 - X_3$ sont indépendants.

Solution. Le vecteur $Y = {}^t(X_1,X_2 - X_3)$ est un vecteur gaussien, car

$$Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} X.$$

De plus,

$$\begin{aligned} \mathrm{Cov}(Y_1, Y_2) &= \mathrm{Cov}(X_1,X_2 - X_3) = \mathrm{Cov}(X_1,X_2) - \mathrm{Cov}(X_1,X_3) \\ &= 1 - 1 = 0. \end{aligned}$$

Ex 8. Soit $X = {}^t(X_1,X_2,X_3)$ un vecteur gaussien de matrice de covariance

$$\Gamma_X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix}.$$

Montrer que X_1 et $X_2 - X_3$ sont indépendants.

Solution. Le vecteur $Y = {}^t(X_1,X_2 - X_3)$ est un vecteur gaussien, car

$$Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} X.$$

De plus,

$$\begin{aligned}\mathrm{Cov}(Y_1, Y_2) &= \mathrm{Cov}(X_1,X_2 - X_3) = \mathrm{Cov}(X_1,X_2) - \mathrm{Cov}(X_1,X_3) \\ &= 1 - 1 = 0.\end{aligned}$$

Par conséquent, $Y_1 = X_1$ et $Y_2 = X_2 - X_3$ sont indépendants.

4. Convergence en loi et vecteurs gaussiens

Stabilité des vecteurs gaussiens

Soit $(X^{(n)})_{n\geq 1}$ une suite de **vecteurs gaussiens** dans $\mathbb{R}^d$. Notons m_n et Γ_n le vecteur moyenne et la matrice de covariance de $X^{(n)}$.

Proposition

La suite de vecteurs aléatoires $(X^{(n)})_{n\geq 1}$ converge en loi vers un vecteur aléatoire X *si et seulement si* les deux suites $(m_n)_{n\in\mathbb{N}}$ et $(\Gamma_n)_{n\in\mathbb{N}}$ convergent vers un vecteur m et une matrice Γ respectivement.

Dans ce cas, X est un vecteur gaussien à valeur dans $\mathbb{R}^d$ de loi $\mathcal{N}(m,\Gamma)$, c'est-à-dire tel que

$$m_X = m = \lim_{n\to\infty} m_n \quad \text{et} \quad \Gamma_X = \Gamma = \lim_{n\to\infty} \Gamma_n.$$

Théorème central limite multi-dimensionnel

Soit $(X^{(n)})_{n\geq 1}$ une suite de **vecteurs aléatoires** dans $\mathbb{R}^d$ **indépendants et identiquement** distribués. Nous supposons que

$$\mathbb{E}\left(\left\|X^{(1)}\right\|^2\right) < \infty.$$

Nous notons m et Γ le vecteur moyenne et la matrice de covariance de $X^{(1)}$.

Théorème central limite multi-dimensionnel

Soit $(X^{(n)})_{n\geq 1}$ une suite de **vecteurs aléatoires** dans $\mathbb{R}^d$ **indépendants et identiquement** distribués. Nous supposons que

$$\mathbb{E}\left(\left\|X^{(1)}\right\|^2\right) < \infty.$$

Nous notons m et Γ le vecteur moyenne et la matrice de covariance de $X^{(1)}$.

Théorème central limite

Nous avons

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X^{(i)} - m\right) \xrightarrow[n\to\infty]{loi} Z \sim \mathcal{N}(0,\Gamma).$$

Ex 9. Considérons une assurance devant couvrir deux types de risques $\{R_1,R_2\}$. Les coûts de dédommagement en cas d'accident sont

$$R_1 : 1000 \text{ €}, \quad R_2 : 10000 \text{ €},$$

et les probabilités de réalisation de ces risques

$$\{R_1,\text{non-}R_2\} : 30\%, \quad \{\text{non-}R_1,R_2\} : 10\%, \quad \{R_1,R_2\} = 2\%.$$

1. Ces risques sont-ils indépendants ?
2. La compagnie assure 10000 clients. Notons $X^{(n)} = (X_1^{(n)},X_2^{(n)})$ la variable aléatoire représentant la somme à verser au n-ième client. Montrer que $\mathbb{E}(X_1^{(n)} + X_2^{(n)}) = 1520$.
3. La compagnie assure 10000 clients, dont la cotisation moyenne s'élève à $1520 + 20$ €. Quelle-est la probabilité d'un déficit ?

Solution.

1. Les événements ne sont pas indépendants car

$$\mathbb{P}(R_1) = 32\%,\ \mathbb{P}(R_2) = 12\%,\ \mathbb{P}(R_1)\mathbb{P}(R_2) = 3{,}84\%,$$
$$\mathbb{P}(R_1, R_2) = 2\% < \mathbb{P}(R_1)\mathbb{P}(R_2).$$

Solution.

1. Les événements ne sont pas indépendants car

$$\mathbb{P}(R_1) = 32\%,\ \mathbb{P}(R_2) = 12\%,\ \mathbb{P}(R_1)\mathbb{P}(R_2) = 3{,}84\%,$$
$$\mathbb{P}(R_1,R_2) = 2\% < \mathbb{P}(R_1)\mathbb{P}(R_2).$$

2. Nous avons $\mathbb{E}(X_1^{(n)} + X_2^{(n)}) = \mathbb{E}(X_1^{(n)}) + \mathbb{E}(X_2^{(n)})$, or

$$\mathbb{E}(X_1^{(n)}) = 1000\,\mathbb{P}(R_1) = 320, \quad \mathbb{E}(X_2^{(n)}) = 10000\,\mathbb{P}(R_2) = 1200.$$

Solution.

1. Les événements ne sont pas indépendants car

$$\mathbb{P}(R_1) = 32\%,\ \mathbb{P}(R_2) = 12\%,\ \mathbb{P}(R_1)\mathbb{P}(R_2) = 3{,}84\%,$$
$$\mathbb{P}(R_1{,}R_2) = 2\% < \mathbb{P}(R_1)\mathbb{P}(R_2).$$

2. Nous avons $\mathbb{E}(X_1^{(n)} + X_2^{(n)}) = \mathbb{E}(X_1^{(n)}) + \mathbb{E}(X_2^{(n)})$, or

$$\mathbb{E}(X_1^{(n)}) = 1000\,\mathbb{P}(R_1) = 320, \quad \mathbb{E}(X_2^{(n)}) = 10000\,\mathbb{P}(R_2) = 1200.$$

3. D'après le théorème central limite, nous avons

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_1^{(i)} - 320 + X_2^{(i)} - 1200) \geq 20\right) \simeq \mathbb{P}\left(\frac{1}{\sqrt{n}}(Z_1 + Z_2) \geq 20\right)$$

où Z est un vecteur gaussien centré de même matrice de covariance que $X^{(n)}$ (les v.a. $X^{(n)}$ sont supposées iid).

La matrice de covariance de $X^{(n)}$ est donnée par

$$\Gamma_n = \begin{pmatrix} 217600 & -164000 \\ -164000 & 10560000 \end{pmatrix}.$$

En particulier, $Z_1 + Z_2$ est une v.a. gaussienne centrée, de variance 10449600

$$\mathbb{P}(Z_1 + Z_2 \geq 2000) \simeq 19{,}1\%.$$

Exercice. Montrer que doubler la prime aura le même effet que de quadrupler le nombre de clients.
Exercice. Quel nombre de clients faut-il réunir pour assurer un risque de inférieur à $0{,}01\%$?