

Intégration L3 Actuariat

Chapitre I: Tribu, mesure et applications mesurables

Pierre-Olivier Goffard

Université de Lyon 1
ISFA
pierre-olivier.goffard@univ-lyon1.fr

ISFA
September 9, 2020

I. Tribus

1. Tribus sur un ensemble quelconque

Soit Ω un ensemble.

Exemple 1

On considère l'expérience aléatoire qui consiste à jeter une pièce en l'air. Il est possible de proposer les espaces suivants

- $\Omega_1 = \{\text{Pile, Face}\}$
- $\Omega_2 = \mathbb{R}^3$ correspondant à la localisation du centre de gravité de la pièce à l'issu du lancer
- $\Omega_3 = (\mathbb{R}^3)^{[0, T]}$ correspondant à la suite des positions de la pièces à tout instant entre 0 et T

La définition de l'espace d'état va dépendre de ce qui nous intéresse. Ω_3 est un espace fonctionnelle, espace des fonctions continues sur $[0, T]$ à valeurs dans \mathbb{R}^3 !

Une fois l'ensemble Ω définit, on introduit les évènements A comme des parties de Ω . On a $A \in \mathcal{P}(\Omega)$, il s'agit de l'ensemble des résultats ω de l'expérience qui conduisent à la réalisation de A .

Definition 1 (Terminologie)

- 1 Ω est un évènement certain, \emptyset correspond à un évènement impossible.
- 2 Pour $A, B \subset \Omega$ deux évènements,

$A \cup B$ se réalise si A ou B se réalisent

et

$A \cap B$ se réalise si A et B se réalisent simultanément

- 3 Pour tout $A \subset \Omega$, on définit par

$$A^c = \{x \in \Omega ; x \notin A\}$$

son complément dans Ω , appelé aussi évènement contraire de A .

- 4 Soit $B \subset A$, on définit par

$$A/B = A \cap B^c$$

la différence entre A et B qui se réalise en cas de réalisation de A mais pas de B .

- 5 Deux évènements sont incompatibles si $A \cap B = \emptyset$

Definition 2 (Suite d'évènements)

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'évènements.

- ① Si $A_1 \subset A_2 \subset \dots$ alors $(A_n)_{n \in \mathbb{N}}$ est une suite croissante d'évènements et

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n.$$

- ② Si $A_1 \supset A_2 \supset \dots$ alors $(A_n)_{n \in \mathbb{N}}$ est une suite décroissante d'évènements et

$$\lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$$

Definition 3 (Limite supérieure et inférieure)

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'évènements de Ω , on définit les limites sup et inf par

$$\overline{\lim}_{n \rightarrow +\infty} A_n = \bigcap_{k=1}^{+\infty} \bigcup_{n \geq k} A_n \quad \text{et} \quad \underline{\lim}_{n \rightarrow +\infty} A_n = \bigcup_{k=1}^{+\infty} \bigcap_{n \geq k} A_n.$$

Pour écrire qu'une infinité de A_n se réalisent, on écrit qu'à partir de n'importe quel rang, il existe des évènements qui se réalisent ce qui correspond à la limite sup de (A_n)

$$\overline{\lim}_{n \rightarrow +\infty} A_n = \bigcap_{k=1}^{+\infty} \bigcup_{n \geq k} A_n.$$

Pour écrire que seul un nombre fini de A_n se réalisent, on écrit qu'il existe un rang à partir duquel seul les évènements contraires aux A_n se réalisent. Cela correspond à la limite inf de la suite (A_n^c)

$$\underline{\lim}_{n \rightarrow +\infty} A_n^c = \bigcup_{k=1}^{+\infty} \bigcap_{n \geq k} A_n^c.$$

Le signe intersection s'interprète de la même façon que "pour tout" et le signe union joue le rôle d'"il existe"

Exemple 2

Soit $\Omega = \{1, 2, 3\}$, on définit une suite d'évènements $(A_n)_{n \in \mathbb{N}^*}$ avec

$$A_{2n-1} = \{1, 2\} \text{ et } A_{2n} = \{2, 3\}$$

alors on a

$$\overline{\lim}_{n \rightarrow +\infty} A_n = \{1, 2, 3\} \text{ et } \underline{\lim}_{n \rightarrow +\infty} A_n = \{2\}.$$

Exemple 3

Soit $\Omega = \mathbb{R}$, considérons une suite d'intervalles fermés défini par

$$A_n = [(-1)^n, 1 + 2^{-n}]$$

on a

$$\overline{\lim}_{n \rightarrow +\infty} A_n = [-1, 1] \text{ et } \underline{\lim}_{n \rightarrow +\infty} A_n = \{1\}.$$

Pour la limite supérieur on constate que le fait que k soit pair ou impair ne change rien car

- si k est pair alors

$$\bigcup_{n \geq k} [(-1)^n, 1 + 2^{-n}] = [1, 1 + 2^{-k}] \cup [-1, 1 + 2^{-(k+1)}] \cup \dots = [-1, 1 + 2^{-k}]$$

- si k est impair alors $\bigcup_{n \geq k} [(-1)^n, 1 + 2^{-n}] = [-1, 1 + 2^{-k}]$

Definition 4 (Tribu, espace mesurable)

Un sous-ensemble \mathcal{A} de $\mathcal{P}(\Omega)$ est une tribu sur Ω si

- 1 $\Omega \in \mathcal{A}$
- 2 \mathcal{A} est stable par passage au complémentaire,

$$A \in \mathcal{A} \Rightarrow A^c = \Omega \setminus A \in \mathcal{A}.$$

- 3 \mathcal{A} est stable par réunion dénombrable,

$$(A_n)_{n \in \mathbb{N}^*} \in \mathcal{A} \Rightarrow \bigcup_{i \in \mathbb{N}^*} A_i \in \mathcal{A}.$$

\mathcal{A} est parfois appelée σ -algèbre.

Exemple 4 (Exemples de tribus)

- $\{\Omega, \emptyset\}$ est la tribu triviale
- $\mathcal{P}(\Omega)$ est une tribu
- Soit $\Omega = \{a, b, c, d\}$ alors $\{\Omega, \emptyset, a, \{b, c, d\}\}$ est la plus petite tribu contenant a .

Les propriétés de stabilité de cette classe permettent de combiner des évènements pour en créer des nouveaux qui appartiendront eux aussi à la tribu.

Exemple 5

On reprend l'expérience du pile ou face, soit

$A =$ "Le nombre de lancer nécessaire pour obtenir Pile est pair"

A est la réunion dénombrable des évènements

$A_p =$ "Pile apparaît pour la première fois au $2p$ -ième lancer".

Proposition 1

Soit \mathcal{A} une tribu de Ω et (A_n) une suite d'éléments de \mathcal{A} , on a

- 1 $\emptyset \in \mathcal{A}$
- 2 $\bigcap_{i \in \mathbb{N}^*} A_i \in \mathcal{A}$
- 3 $\bigcap_{i=1}^n A_i \in \mathcal{A}$
- 4 $\bigcup_{i=1}^n A_i \in \mathcal{A}$
- 5 $\overline{\lim_{n \rightarrow +\infty} A_i} \in \mathcal{A}$
- 6 $\underline{\lim_{n \rightarrow +\infty} A_i} \in \mathcal{A}$

Il s'agit de conséquences assez immédiates des axiomes des bases

Definition 5 (Sous-tribus)

Une sous-tribu de \mathcal{A} est une tribu de Ω telle que $\mathcal{B} \subset \mathcal{A}$.

Proposition 2

L'intersection de deux tribus de Ω est une tribu.

Definition 6 (Tribu engendrée)

La tribu engendrée par $\mathcal{E} \subset \mathcal{P}(\Omega)$ (famille de parties de Ω), notée $\sigma(\mathcal{E})$ est l'intersection de toutes les tribus contenant \mathcal{E} .

$\sigma(\mathcal{E})$ est la plus petite tribu (au sens de l'inclusion) contenant \mathcal{E} .

Exemple 6

- 1 Soit $A \in \Omega$ alors $\sigma(A) = \{A, A^c, \emptyset, \Omega\}$
- 2 Soit $\mathcal{S} = \{S_1, \dots, S_n\}$ une partition de Ω , c'est à dire que

$$\bigcup_{k=1}^n S_k = \Omega, \text{ et } S_i \cap S_j = \emptyset \text{ pour } i \neq j$$

Alors $\sigma(\mathcal{S}) = \{\bigcup_{k \in T} S_k ; T \subset \{1, 2, \dots, n\}\}$

2. Tribu Borélienne

Definition 7 (Espace topologique)

Soit E un ensemble. Soit \mathcal{O} une famille de parties de E , appelée ouverts de E , vérifiant

- $\emptyset, E \in \mathcal{O}$,
- Stable par réunion quelconque,
- Stable par intersection finie.

Le couple (E, \mathcal{O}) est un espace topologique

Exemple 7 (Ouvert dans un espace métrique)

Si E est un espace métrique alors on peut définir une distance entre $x \in E$ et $y \in E$ par $d(x, y)$. Un ouvert O est une partie de E dont la frontière est vide, ou dont tout les point appartient à l'intérieur de O . Concrètement,

$$\forall x \in O, \exists r > 0 \text{ tel que } B(x, r) = \{y \in E ; d(x, y) < r\} \subset O$$

Pour $E = \mathbb{R}$, les ouverts sont les parties qui pour chaque point x contiennent un intervalle du type $]x - \epsilon, x + \epsilon[$. On note

$$\mathcal{I}_{\mathbb{R}} = \{]a, b[, -\infty < a \leq b < +\infty\},$$

l'ensemble des intervalles ouverts bornés. Il contient \emptyset (cas $a=b$).

Definition 8 (Tribu borélienne, borélien)

La tribu borélienne est la tribu $\mathcal{B}(E)$ engendré par les ouverts de E . On appelle borélien un ensemble appartenant à cette tribu.

La tribu borélienne $\mathcal{B}(E)$ contient tout les ouverts de E , ainsi que tout les fermés (par passage au complémentaire), les intersections et réunions de suites d'ouverts et de fermés. La tribu borélienne $\mathcal{B}(\mathbb{R})$ est engendrée par intervalles ouverts de \mathbb{R} , c'est la conséquence du lemme suivant.

Lemme 1

Tout ouvert de \mathbb{R} est la réunion d'une suite d'intervalles ouverts

preuve:

Remarquons que l'ensemble

$$\mathcal{I}^* = \left\{ \left] r - \frac{1}{n}, r + \frac{1}{n} \right[; r \in \mathbb{Q}, n \in \mathbb{N}^* \right\}$$

est dénombrable puisqu'il existe une surjection de $\mathbb{Q} \times \mathbb{N}^*$ sur \mathcal{I}^* . Soit U un ouvert de \mathbb{R} , supposé non vide, et soit $x \in U$. Il existe un $\epsilon > 0$ tel que $]x - \epsilon, x + \epsilon[\subset U$, puis $\exists n \geq 0$ tel que $\frac{1}{n} \leq \frac{\epsilon}{2}$ et enfin un

$$r \in \mathbb{Q} \cap \left] x - \frac{1}{n}, x + \frac{1}{n} \right[.$$

On voit alors que

$$x \in \left] r - \frac{1}{n}, r + \frac{1}{n} \right[.$$

A chaque $x \in U$ est associé un intervalle $I_x \in \mathcal{I}^*$ tel que $x \in I_x \subset U$ si bien que $U = \bigcup_{x \in U} \{x\} \subset \bigcup_{x \in U} I_x \subset U$ et, par suite $\bigcup_{x \in U} I_x = U$. On écrit donc U comme la réunion d'une suite $(I_n) \in \mathcal{I}^*$ qui est aussi une suite de $\mathcal{I}_{\mathbb{R}}$ puisque $\mathcal{I}^* \subset \mathcal{I}_{\mathbb{R}}$.

□

La tribu borélienne $\mathcal{B}(\mathbb{R})$ peut donc être générée par différents types d'intervalles dont

- $[a, b]$
- $[a, +\infty[$
- $]a, +\infty[$
- $]a, b]$

Definition 9

Soient Ω et Ω' deux ensembles. La tribu engendrée par les ensembles $A \times B \in \mathcal{A} \times \mathcal{B}$, où \mathcal{A} et \mathcal{B} sont des tribus de Ω et Ω' respectivement, est la tribu produit $\mathcal{A} \otimes \mathcal{B}$

Proposition 3

① Si $\mathcal{A} = \sigma(A_i, i \in I)$ et $\mathcal{B} = \sigma(B_j, j \in J)$ alors

$$\mathcal{A} \otimes \mathcal{B} = \sigma(A_i \times B_j, (i, j) \in I \times J)$$

② $\mathcal{B}(\mathbb{R}^d) = \mathcal{B}(\mathbb{R})^{\otimes d}$

II. Mesures

1. Définition et propriétés

Le couple (Ω, \mathcal{A}) est un espace mesurable.

Definition 10 (Mesure (positive))

On appelle mesure (positive) une application $\mu: \mathcal{A} \rightarrow \overline{\mathbb{R}}_+$ telle que:

- (i) $\mu(\emptyset) = 0$,
- (ii) pour toute suite $(A_n)_{n \in \mathbb{N}^*}$ de parties disjointes de \mathcal{A} , on a

$$\mu\left(\bigcup_{n \in \mathbb{N}^*} A_n\right) = \sum_{n \in \mathbb{N}^*} \mu(A_n). \quad (\sigma\text{-additivité})$$

Le triplet $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré.

Definition 11 (Terminologie)

- 1 Si $\mu(\Omega) < +\infty$ alors μ est une mesure finie
- 2 Si $(A_n)_{n \geq 1} \in \mathcal{A}$ tels que

$$n \geq 1, \mu(A_n) < \infty \text{ et } \bigcup_{n \geq 1} A_n = \Omega$$

alors μ est σ -finie.

- 3 Si $\mu(\Omega) = 1$ alors μ est une mesure de probabilité. D'ailleurs on désigne parfois (Ω, \mathcal{A}) comme un espace probabilisable.
- 4 Le triplet $(\Omega, \mathcal{A}, \mu)$ est appelé espace mesuré (ou probabilisé si μ est une mesure de probabilité).
- 5 Une mesure signée est une mesure définie comme la différence de deux mesures positives.
- 6 Une propriété \mathcal{P} est vraie μ -presque partout s'il existe $A \in \mathcal{A}$ tel que

$$\forall x \in \Omega/A, \mathcal{P} \text{ est vraie et } \mu(A) = 0$$

- 7 Soit $A \in \mathcal{A}$, on dit que μ est portée par A si $\mu(A^c) = 0$.
- 8 μ est une mesure atomique si elle est portée par les atomes $\{\omega \in \Omega\}$
- 9 μ est une mesure diffuse si $\mu(\{\omega\}) = 0$ (les atomes $\{\omega\}$ sont des parties négligeables)

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

Proposition 4 (Propriété d'un mesure)

Soient $(A_n)_{n \in \mathbb{N}^*}$ une suite d'évènements de \mathcal{A} .

- 1 $A_1 \cap A_2 = \emptyset \implies \mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$
- 2 Si $A_1 \subset A_2$ alors $\mu(A_1) \leq \mu(A_2)$ (monotonie de μ), de plus on a

$$\mu(A_2/A_1) = \mu(A_2) - \mu(A_1)$$

- 3 $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2)$ (formule inclusion-exclusion)
- 4 $\mu(A_1 \cap A_2) \leq \min(\mu(A_1), \mu(A_2))$ et $\mu(A_1 \cup A_2) \geq \max(\mu(A_1), \mu(A_2))$
- 5

$$\mu\left(\bigcup_{k=1}^{+\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k) \text{ (sous } \sigma\text{-additivité)}$$

- 6 Si (A_n) est une suite croissante ($A_i \subset A_{i+1}$, $i \in \mathbb{N}^*$) et que $\bigcup_{n \in \mathbb{N}^*} A_n = A$, alors $(\mu(A_i))_{i \in \mathbb{N}^*}$ est une suite croissante qui converge vers $\mu(A)$.
- 7 Si (A_n) est une suite décroissante ($A_{i+1} \subset A_i$, $i \in \mathbb{N}^*$) telle que $\mu(A_1) < \infty$ et que $\bigcap_{n \in \mathbb{N}^*} A_n = A$, alors $(\mu(A_i))_{i \in \mathbb{N}^*}$ est une suite décroissante qui converge vers $\mu(A)$.

preuve:

- 1 On suppose que $A_i = \emptyset$ pour tout $i > 2$ et on exploite la σ -additivité de μ .
- 2 Soit $A_1 \subset A_2 \subset \mathcal{A}$, on a

$$\mu(A_2) = \mu(\{A_2/A_1\} \cup A_1) = \mu(A_2/A_1) + \mu(A_1) \geq \mu(A_1) \quad (\text{car } \mu \text{ est une mesure positive})$$

On déduit immédiatement de ce qui précède que $\mu(\{A_2/A_1\}) = \mu(A_2) - \mu(A_1)$

- 3 Soit $A_1, B_2 \in \mathcal{A}$, on a

$$\begin{aligned}\mu(A_1 \cup A_2) &= \mu\{A_1 \cup [A_2/(A_1 \cap A_2)]\} \\ &= \mu(A_1) + \mu[A_2/(A_1 \cap A_2)] \\ &= \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2)\end{aligned}$$

- 4 Examen

5 On définit la suite $(B_n)_{n \geq 1} \in \mathcal{A}$ telle que

$$\begin{aligned} B_1 &= A_1, \\ B_2 &= A_2 \cap A_1^c \\ &\vdots \\ B_n &= A_n \cap A_{n-1}^c \cap \dots \cap A_1^c \end{aligned}$$

Les B_n sont disjoints et vérifient $B_n \subset A_n$. On vérifie que

$$\bigcup_{k=1}^{+\infty} B_k = \bigcup_{k=1}^{+\infty} A_k$$

par double inclusion. On a d'une part

$$\bigcup_{k=1}^{+\infty} B_k \subset \bigcup_{k=1}^{+\infty} A_k$$

et de plus pour $\omega \in \bigcup_{k=1}^{+\infty} A_k$, il existe un plus petit n_0 tel que $\omega \in A_{n_0}$ puis $\omega \in B_{n_0}$ et $\omega \in \bigcup_{k=1}^{+\infty} B_k$. On en déduit que

$$\bigcup_{k=1}^{+\infty} B_k \supset \bigcup_{k=1}^{+\infty} A_k$$

puis l'égalité. On peut alors écrire

$$\mu\left(\bigcup_{k=0}^{+\infty} A_k\right) = \mu\left(\bigcup_{k=0}^{+\infty} B_k\right) = \sum_{k=0}^{+\infty} \mu(B_k) \leq \sum_{k=0}^{+\infty} \mu(A_k).$$

- ⑥ Comme $\mu(A_{i+1}) \geq \mu(A_i)$ et $\mu(A_i) < \mu(A)$ alors $(\mu(A_n))_{n \in \mathbb{N}}$ est une suite croissante bornée, donc qui converge. Soit $B_1 = A_1$ et $B_n = A_n \setminus A_{n-1}$ pour $k \geq 2$, les B_k sont disjoints et vérifient $\bigcup_{n=1}^{+\infty} B_n = A$. On a

$$\mu(A) = \mu\left(\bigcup_{k=1}^{+\infty} B_k\right) = \sum_{k=1}^{\infty} \mu(B_k) = \lim_{n \rightarrow +\infty} \sum_{k=1}^n \mu(B_k) = \lim_{n \rightarrow +\infty} \mu\left(\bigcup_{k=1}^n B_k\right) = \lim_{n \rightarrow +\infty} \mu(A_n).$$

- ⑦ Examen



Proposition 5

Soient μ et ν deux mesures définies sur un espace mesurable (Ω, \mathcal{A}) et $\alpha > 0$ alors

- ① $\mu + \nu$ est une mesure
- ② $\alpha \times \mu$ est une mesure

2. Mesure de comptage

Soient (Ω, \mathcal{A}) un espace mesurable,

Definition 12 (Mesure de Dirac)

Soient $x \in \Omega$ et $A \in \mathcal{A}$. La mesure définie par

$$\delta_x(A) = \begin{cases} 1, & x \in A, \\ 0, & \text{sinon.} \end{cases}$$

est appelée mesure de Dirac en x .

Montrons que $A \mapsto \delta_x(A)$ définit bien une mesure.

- 1 $\delta_x(\emptyset) = 0$
- 2 Soit $(A_j)_{j \geq 0}$ une suite d'évènements disjoints de \mathcal{A} .
 - S'il existe i tel que $x \in A_i$ alors $x \in \bigcup_j A_j$ et $\delta_x(\bigcup_j A_j) = 1$. De plus, comme les A_j sont disjoints alors $x \notin A_j$ pour $j \neq i$. On a donc

$$\sum_j \delta_x(A_j) = \delta_x(A_i) = \delta_x\left(\bigcup_i A_i\right) = 1.$$

- Si $x \notin A_j, \forall i$ alors $x \notin \bigcup_i A_i$ et

$$\delta_x\left(\bigcup_i A_i\right) = \sum_i \delta_x(A_i) = 0.$$

Definition 13 (Mesure de comptage)

Si Ω est un ensemble dénombrable alors

$$C(A) = \text{Card}(A), \quad A \in \mathcal{A}$$

définie une mesure appelée mesure de comptage. Il est possible d'écrire

$$C(A) = \sum_{x \in \Omega} \delta_x(A).$$

2. Mesure de probabilité

Une expérience aléatoire est répétée n fois, supposons que A s'est réalisé $k \leq n$ au cours de ces expériences. k est la fréquence absolue d'occurrence de A et k/n est sa fréquence d'occurrence relative. Lorsque n devient grand, la fréquence relative se stabilise autour d'un nombre $\mathbb{P}(A)$ appelé probabilité de A .

A partir des fréquences relatives on constate que

- $0 \leq \mathbb{P}(A) \leq 1$
- $\mathbb{P}(\Omega) = 1$
- Si $A \subset B$ alors $\mathbb{P}(A) \leq \mathbb{P}(B)$
- $A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

propriétés cohérentes avec la définition d'une mesure. Il s'agit de l'interprétation dite fréquentiste des probabilités.

Exemple 8 (Evènements élémentaires équiprobables)

Soit une expérience aléatoire dont les résultats ω sont équiprobables et forment un ensemble Ω fini. L'application $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0, 1]$ définie par

$$\mathbb{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$$

est une mesure de probabilité. On a

$$\mathbb{P}(\{\omega\}) = \frac{1}{\text{Card}(\Omega)}$$

et le calcul des probabilités se résume à des problèmes de dénombrement.

Montrer que $\mathbb{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$ est une mesure de probabilité \Rightarrow Examen.

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.

Lemme 2 (Borel-Cantelli, première partie)

Si $(A_n)_{n \geq 1}$ une suite d'évènements telle que $\sum_{n \geq 1} \mathbb{P}(A_n) < \infty$ alors

$$\mathbb{P}\left(\overline{\lim}_{n \rightarrow +\infty} A_n\right) = 0.$$

preuve:

Notons que

$$\mathbb{P}\left(\bigcup_{n \geq i} A_n\right) \leq \sum_{n \geq i} \mathbb{P}(A_n), \text{ pour tout } i \geq 1,$$

d'après la Proposition 5. De plus, $\overline{\lim}_{n \rightarrow +\infty} A_n = \bigcap_{k \geq 1} \bigcup_{n \geq k} A_n \subset \bigcup_{n \geq i} A_n$ pour tout $i \geq 1$.

Donc

$$0 \leq \mathbb{P}\left(\overline{\lim}_{n \rightarrow +\infty} A_n\right) \leq \sum_{n \geq i} \mathbb{P}(A_n), \text{ pour tout } i \geq 1.$$

Le membre de droite tend vers 0 comme reste d'une série convergente.

□

La probabilité qu'une infinité d'évènements se réalisent est nulle. $\overline{\lim}_{n \rightarrow +\infty} A_n$ est un évènement presque impossible (ou de mesure de probabilité négligeable). On a de manière équivalente

$$\mathbb{P}\left(\underline{\lim}_{n \rightarrow +\infty} A_n^c\right) = 1$$

$\varliminf_{n \rightarrow +\infty} A_n^c$ est un évènement presque certain.

Definition 14 (Probabilité conditionnelle)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soit A et B deux évènements, tel que $\mathbb{P}(B) > 0$, alors on définit la probabilité conditionnelle de A sachant B par

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

L'application $\mathbb{P}(\cdot|B)$ définit bien une probabilité sur (Ω, \mathcal{A}) , en effet,

- 1 $\mathbb{P}(\emptyset|B) = \frac{\mathbb{P}(\emptyset \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(B)} = 0.$
- 2 $\mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1.$
- 3 Soit $(A_n)_{n \geq 1}$ une suite d'évènements disjoints alors

$$\mathbb{P}\left(\bigcup_{n \geq 1} A_n | B\right) = \frac{\mathbb{P}(\bigcup_{n \geq 1} A_n \cap B)}{\mathbb{P}(B)} = \sum_{n \geq 1} \frac{\mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_{n \geq 1} \mathbb{P}(A_n | B).$$

On peut montrer que l'ensemble $\mathcal{A}_B = \{A \cap B, A \in \mathcal{A}\}$ est une tribu, appelée tribu trace, et définir un nouvel espace probabilisé avec $(\Omega, \mathcal{A}_B, \mathbb{P}(\cdot|B))$

Theoreme 1 (Loi des probabilités totales)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, soit $(A_i)_{1 \leq i \leq n}$ une partition de Ω , telle que $\mathbb{P}(A_i) > 0, \forall i \geq 1$, alors

$$\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(B|A_i)\mathbb{P}(A_i), \text{ pour tout } B \in \mathcal{A}.$$

preuve: On a, pour tout $B \in \mathcal{A}$,

$$\mathbb{P}(B) = \mathbb{P}(B \cap \Omega) = \mathbb{P}\left(B \cap \bigcup_{i=1}^n A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^n B \cap A_i\right) = \sum_{i=1}^n \mathbb{P}(B \cap A_i) = \sum_{i=1}^n \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

Theoreme 2 (Bayes)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, soit $(A_i)_{1 \leq i \leq n}$ une partition de Ω , telle que $\mathbb{P}(A_i) > 0 \forall i \geq 1$, et $B \in \mathcal{A}$ un évènement de probabilité non nulle. Alors,

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{j=1}^n \mathbb{P}(B|A_j)\mathbb{P}(A_j)}, \text{ pour tout } i = 1, \dots, n.$$

preuve:

examen.

Exemple 9

Rey tente de s'échapper des griffes de Kylo Ren et de l'empire. Elle choisit au hasard un véhicule parmi

- le Millenium Falcon (MF)
- Le TIE fighter (TIE)
- Le X-Wing starfighter (XW)

La probabilité qu'elle s'échappe (E) est de

- 0.4 si elle opte pour le Millenium Falcon
- 0.6 si elle opte pour Le TIE fighter
- 0.7 si elle opte pour Le X-Wing starfighter

Quelle est la probabilité qu'elle s'échappe si elle a choisit le Millenium Falcon?

D'après l'énoncé, $\mathbb{P}(MF) = \mathbb{P}(TIE) = \mathbb{P}(XW) = 1/3$ et

$$\mathbb{P}(E|MF) = 0.4, \mathbb{P}(E|TIE) = 0.6, \text{ et } \mathbb{P}(E|XW) = 0.7$$

On applique la formule de Bayes pour obtenir

$$\mathbb{P}(MF|E) = \frac{\mathbb{P}(E|MF)\mathbb{P}(MF)}{\mathbb{P}(E|MF)\mathbb{P}(MF) + \mathbb{P}(E|TIE)\mathbb{P}(TIE) + \mathbb{P}(E|XW)\mathbb{P}(XW)}$$

et on remplace.

Definition 15 (Evènements indépendants)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, Les évènements $A, B \in \mathcal{A}$ sont indépendants sous la probabilité \mathbb{P} si et seulement si

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

On observe directement que

$$\mathbb{P}(A|B) = \mathbb{P}(A).$$

Proposition 6

Si A et B sont indépendants sous \mathbb{P} alors

- 1 A^c et B sont indépendants
- 2 A et B^c sont indépendants
- 3 A^c et B^c sont indépendants

preuve:

- ① On note que

$$A = (A \cap B^c) \cup (A \cap B)$$

puis

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

et finalement

$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(B^c),$$

d'où l'indépendance de A et de B^c .

- ② Même raisonnement
③ idem

Definition 16 (Evènements mutuellement indépendants)

La suite $(A_n)_{n \geq 1} \in \mathcal{A}$ est une suite d'évènement mutuellement indépendants si pour tout sous ensemble $(A_{i_1}, \dots, A_{i_k})$ d'évènement, avec $(i_1, \dots, i_k) \in \mathbb{N}^k$ un ensemble d'indices, on a

$$\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \times \dots \times \mathbb{P}(A_{i_k}).$$

Exemple 10

Il ne faut pas confondre mutuellement indépendant et indépendant deux à deux !
En effet, soit l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ avec

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}, \mathcal{A} = \mathcal{P}(\Omega) \text{ et } \mathbb{P}(\omega_i) = 1/4, \forall i = 1, 2, 3, 4.$$

On définit les évènements $A_1 = \{\omega_1, \omega_4\}$, $A_2 = \{\omega_2, \omega_4\}$ et $A_3 = \{\omega_3, \omega_4\}$ alors on observe que A_1 et A_2 sont indépendants, A_1 et A_3 sont indépendants et A_2 et A_3 sont indépendants. Cependant,

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) \neq \mathbb{P}(A_1) \times \mathbb{P}(A_2) \times \mathbb{P}(A_3).$$

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, et une suite $(A_n)_{n \geq 1}$ d'évènements mutuellement indépendants.

Proposition 7

$$\mathbb{P}\left(\bigcap_{n \geq 1} A_n\right) = \lim_{n \rightarrow +\infty} \prod_{k=1}^n \mathbb{P}(A_k).$$

preuve:

La suite d'évènement $\left(\bigcap_{k=1}^n A_k\right)_{n \geq 1}$ est décroissante, on a donc

$$\mathbb{P}\left[\left(\bigcap_{k=1}^{\infty} A_k\right)\right] = \lim_{n \rightarrow \infty} \mathbb{P}\left(\bigcap_{k=1}^n A_k\right) = \lim_{n \rightarrow \infty} \prod_{k=1}^n \mathbb{P}(A_k).$$

Lemme 3 (Borel-Cantelli deuxième partie)

Si $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, alors

$$\mathbb{P}\left(\overline{\lim_{n \rightarrow +\infty} A_n}\right) = 1.$$

preuve:

Notons que (pour exploiter l'indépendance des A_n)

$$\overline{\lim_{n \rightarrow +\infty} A_n} = \bigcap_{k \geq 1} \bigcup_{n \geq k} A_n = \left(\bigcup_{k \geq 1} \bigcap_{n \geq k} A_n^c\right)^c = \left(\lim_{n \rightarrow +\infty} A_n^c\right)^c.$$

Comme les évènements A_n^c sont indépendants alors

$$\mathbb{P}\left(\bigcap_{n \geq k} A_n^c\right) = \prod_{n \geq k} \mathbb{P}(A_n^c) = \prod_{n \geq k} (1 - \mathbb{P}(A_n)).$$

Comme $1 - x \leq e^{-x}$ pour $0 \leq x \leq 1$, alors

$$\mathbb{P}\left(\bigcap_{n \geq k} A_n^c\right) \leq \exp\left(-\sum_{n \geq k} \mathbb{P}(A_n)\right) = 0.$$

On en déduit que

$$\mathbb{P}\left(\bigcup_{n \geq k} A_n\right) = 1 - \mathbb{P}\left(\bigcap_{n \geq k} A_n^c\right) = 1$$

pour tout $k \geq 1$ et donc valable lorsque $k \rightarrow \infty$. Or $(\bigcup_{n \geq k} A_n)_{k \geq 1}$ est une suite décroissante et donc $\mathbb{P}\left(\overline{\lim}_{n \rightarrow +\infty} A_n\right) = \mathbb{P}\left(\bigcap_{k \geq 1} \bigcup_{n \geq k} A_n\right) = \lim_{k \rightarrow \infty} \mathbb{P}\left(\bigcup_{n \geq k} A_n\right) = 1$.

□

En combinant les deux parties du lemme de Borel-Cantelli, on parvient au résultat suivant

Théorème 3 (Loi du 0-1)

Pour $(A_n)_{n \geq 1}$ une suite d'événements mutuellement indépendants, on a

$$\mathbb{P}\left(\overline{\lim}_{n \rightarrow +\infty} A_n\right) = \begin{cases} 1, & \text{si } \sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty \\ 0, & \text{si } \sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty \end{cases}$$

Exemple 11

Admettons qu'on lance une pièce un nombre infini de fois, si on note

$$A_n = \text{"le } n\text{-ième lancer est pile"}$$

alors $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \frac{1}{2} = \infty$ et donc $\mathbb{P}(\overline{\lim}_{n \rightarrow +\infty} A_n) = 1$, ce qui revient à obtenir de façon certain un nombre infini de pile.

3. Mesure de Lebesgue

Definition 17 (Mesure de Lebesgue)

On appelle mesure de Lebesgue sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, la mesure λ telle que, pour tout intervalle $]a, b]$,

$$\lambda(]a, b]) = b - a.$$

La mesure de Lebesgue est la seule mesure sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ qui mesure un intervalle par sa longueur. Une idée de la preuve est donné en appendice.

Conséquence:

- λ est une mesure σ -finie
- $\lambda(\{a\}) = 0$ pour tout $a \in \mathbb{R}$, λ est une mesure diffuse.
- La mesure de Lebesgue d'un ensemble dénombrable est nulle
- Pour tout $a, b \in \mathbb{R}$, on a

$$\lambda(]a, b]) = \lambda(]a, b[) = \lambda([a, b]) = \lambda([a, b[)$$

- La mesure de Lebesgue est invariante par symétrie et translation, précisément si on pose

$$A^{-1} = \{y \in \mathbb{R} : -y \in A\} \text{ et } A+x = \{y \in \mathbb{R} : y = x+z, z \in A\}$$

alors $\lambda(A) = \lambda(A^{-1}) = \lambda(A+x)$

Definition 18 (Mesure de Lebesgue sur les pavés)

On appelle mesure de Lebesgue sur $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ la mesure λ_k telle que, pour $A = \prod_{i=1}^k]a_i, b_i[$,

$$\lambda_k(A) = \prod_{i=1}^k (b_i - a_i).$$

III. Applications mesurables

1. Rappels et définition

Soit (Ω, \mathcal{A}) et (E, \mathcal{B}) deux espaces mesurables, et $f : \Omega \rightarrow E$ une application. On définit l'application inverse de f par $f^{-1} : \mathcal{P}(E) \rightarrow \mathcal{P}(\Omega)$ par

$$f^{-1}(B) = \{\omega \in \Omega : f(\omega) \in B\}, \text{ pour } B \in \mathcal{P}(E).$$

On écrit aussi $f^{-1}(B) = \{f \in B\}$. Elle vérifie, ,

- $f^{-1}(\cup_{i \in I} B_i) = \cup_{i \in I} f^{-1}(B_i)$,
- $f^{-1}(\cap_{i \in I} B_i) = \cap_{i \in I} f^{-1}(B_i)$
- $f^{-1}(B^c) = f^{-1}(B)^c$

où $B, (B_i)_{i \in I} \in \mathcal{P}(E)$ et I est un ensemble d'indice.

Proposition 8 $((g \circ f)^{-1})$

Considérons trois ensembles non vides E_1, E_2 et E_3 , et deux fonctions $f : E_1 \rightarrow E_2$ et $g : E_2 \rightarrow E_3$. Alors pour tout $A_3 \subset E_3$, on a

$$(g \circ f)^{-1}(A_3) = f^{-1} \left[g^{-1}(A_3) \right].$$

Definition 19 (Application mesurable)

Soient (Ω, \mathcal{A}) et (E, \mathcal{B}) deux espaces mesurables. $f : \Omega \rightarrow E$ une application mesurable si

$$f^{-1}(B) \in \mathcal{A} \text{ pour tout } B \in \mathcal{B}.$$

Definition 20 (Variable aléatoire réelle)

Une variable aléatoire réelle est une application mesurable $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Les variables aléatoires réelles permettent de quantifier les évènements d'une expérience aléatoire. On peut définir des vecteurs aléatoires (X_1, \dots, X_p) comme applications mesurables de (Ω, \mathcal{A}) vers $(\mathbb{R}^p, \mathcal{B}_{\mathbb{R}^p})$.

Exemple 12 (Fonction indicatrice)

Soient $A \subset \Omega$ l'application $\mathbb{1}_A : \Omega \rightarrow \{0, 1\}$, définie, par

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \text{si } \omega \in A \\ 0, & \text{sinon.} \end{cases}$$

Ici $\mathcal{B} = \mathcal{P}(\{0, 1\})$. Soit $B \in \mathcal{B}$, on a

- $\mathbb{1}_A^{-1}(B) = \emptyset$ si B ne contient ni 0, ni 1. En fait $B = \emptyset$
- $\mathbb{1}_A^{-1}(B) = A$ si B contient 1 et pas 0
- $\mathbb{1}_A^{-1}(B) = A^c$ si B contient 0 mais pas 1
- $\mathbb{1}_A^{-1}(B) = \Omega$ si B contient 0 et 1

Si $A \in \mathcal{A}$ alors l'application $\mathbb{1}_A$, aussi appelé fonction indicatrice sur A est mesurable.

Si $A_1, A_2, \dots, A_n \in \mathcal{A}$ forment une partition de Ω alors l'application

$$f = \sum_{i=1}^n x_i \mathbb{1}_{A_i}$$

où $x_1, \dots, x_n \in \mathbb{R}$ est une application mesurable de (Ω, \mathcal{A}) vers $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Theoreme 4

Soit Ω un ensemble. Soit (E, \mathcal{B}) un espace mesurable et soit $f : \Omega \rightarrow E$ une application. On a

- 1 $f^{-1}(\mathcal{B}) = \{f^{-1}(B) \mid B \in \mathcal{B}\}$ est une tribu sur Ω
- 2 Pour tout $\mathcal{C} \subset \mathcal{P}(E)$: $f^{-1}(\sigma(\mathcal{C})) = \sigma(f^{-1}(\mathcal{C}))$

preuve:

- 1 On exploite les propriétés ensemblistes de f^{-1} ,
 - (i) $f^{-1}(E) = \Omega$ donc $\Omega \in f^{-1}(\mathcal{B})$
 - (ii) Soit $A \in f^{-1}(\mathcal{B}) \subset \mathcal{A}$. Il existe $B \in \mathcal{B}$ tel que $A = f^{-1}(B)$. On a $f^{-1}(B^c) = f^{-1}(B)^c \in \mathcal{A}$ donc $f^{-1}(B)^c \in f^{-1}(\mathcal{B})$
 - (iii) Soit $(A_n)_{n \in \mathbb{N}} \in f^{-1}(\mathcal{B}) \subset \mathcal{A}$. Il existe $B_n \in \mathcal{B}$ tel que $A_n = f^{-1}(B_n)$, $\forall n \in \mathbb{N}$. On a $\bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n \in \mathbb{N}} f^{-1}(B_n) = f^{-1}(\bigcup_{n \in \mathbb{N}} B_n) \in f^{-1}(\mathcal{B})$.
- 2 On remarque que $f^{-1}(\sigma(\mathcal{C}))$ est une tribu qui contient $f^{-1}(\mathcal{C})$ donc

$$\sigma(f^{-1}(\mathcal{C})) \subset f^{-1}(\sigma(\mathcal{C})).$$

On définit

$$\mathcal{F} = \left\{ B \subset E \mid f^{-1}(B) \in \sigma(f^{-1}(\mathcal{C})) \right\}$$

et on montre qu'il s'agit d'une tribu sur E qui contient \mathcal{C} .

- (i) $E \in \mathcal{F}$ puisque $f^{-1}(E) = \Omega \in \sigma(f^{-1}(\mathcal{C}))$
- (ii) Soit $B \in \mathcal{F}$, on a $f^{-1}(B^c) = f^{-1}(B)^c \in \sigma(f^{-1}(\mathcal{C}))$ donc $B^c \in \mathcal{F}$
- (iii) Soit $(B_n)_{n \in \mathbb{N}}$, on a $f^{-1}(\cup_{n \in \mathbb{N}} B_n) = \cup_{n \in \mathbb{N}} f^{-1}(B_n) \in \sigma(f^{-1}(\mathcal{C}))$ donc $\cup_{n \in \mathbb{N}} B_n \in \mathcal{F}$

On observe ainsi que \mathcal{F} est une tribu sur E contenant \mathcal{C} et par conséquent $\sigma(\mathcal{C}) \subset \mathcal{F}$. On observe alors que

$$f^{-1}(\sigma(\mathcal{C})) \subset f^{-1}(\mathcal{F}) \subset \sigma(f^{-1}(\mathcal{C})).$$

□

Definition 21 (Tribu engendrée par f)

$f^{-1}(\mathcal{B})$ est la tribu engendrée par f . Il s'agit de la plus petite tribu \mathcal{T} de Ω pour laquelle f est une application mesurable de (Ω, \mathcal{T}) vers (E, \mathcal{B}) .

Corollaire 1 (Caractérisation de la mesurabilité)

Soit $\mathcal{C} \subset \mathcal{P}(E)$ vérifiant $\sigma(\mathcal{C}) = \mathcal{B}$. Soit $f : (\Omega, \mathcal{A}) \rightarrow (E, \mathcal{B})$ une application.

$$f \text{ est mesurable} \Leftrightarrow f^{-1}(\mathcal{C}) \subset \mathcal{A}.$$

preuve:

\Rightarrow Supposons que f soit mesurable, alors $f^{-1}(\mathcal{C}) \subset \mathcal{A}$ découle de la définition de la mesurabilité.

\Leftarrow Supposons que $f^{-1}(\mathcal{C}) \subset \mathcal{A}$ alors on a

$$f^{-1}(\mathcal{B}) = f^{-1}(\sigma(\mathcal{C})) = \sigma(f^{-1}(\mathcal{C})) \subset \mathcal{A}$$

car $f^{-1}(\mathcal{C}) \subset \mathcal{A}$ et $\sigma(f^{-1}(\mathcal{C}))$ est la plus petite tribu de Ω contenant $f^{-1}(\mathcal{C})$. Cela implique que f est mesurable.

□

Proposition 9 (Mesurabilité de $g \circ f$)

La composée de deux fonctions mesurables est mesurable.

preuve:

Soit $(\Omega_i, \mathcal{A}_i), i = 1, 2, 3$ des espaces mesurables et $f : \Omega_1 \mapsto \Omega_2$ et $g : \Omega_2 \mapsto \Omega_3$. Pour tout $A_3 \in \mathcal{A}_3$, on a

$$(g \circ f)^{-1}(A_3) = f^{-1}(g^{-1}(A_3))$$

avec $g^{-1}(A_3) \in \mathcal{A}_2$ puis $f^{-1}(g^{-1}(A_3)) \in \mathcal{A}_1$, ce qui permet de conclure que $(g \circ f)$ est mesurable.

□

Theoreme 5 (μ_f)

Soit (Ω, \mathcal{A}) et (E, \mathcal{B}) deux espaces mesurables, et $f : (\Omega, \mathcal{A}) \mapsto (E, \mathcal{B})$ une application mesurable. A toute mesure μ sur (Ω, \mathcal{A}) on peut associer une mesure μ_f sur (E, \mathcal{B}) définie par

$$\mu_f(B) = \mu \left[f^{-1}(B) \right], \quad B \in \mathcal{B}.$$

preuve:

μ_f est à valeurs positives, comme μ . De plus,

$$\mu_f(\emptyset) = \mu \left[f^{-1}(\emptyset) \right] = \mu(\emptyset) = 0.$$

Soit $(B_n)_{n \geq 0}$ une suite d'évènements de \mathcal{B} deux à deux disjoints. La suite d'évènements $\{f^{-1}(B_n)\}_{n \geq 0}$ est une suite d'évènements disjoints de \mathcal{A} . En effet, supposons l'existence de $\omega \in f^{-1}(B_1) \cap f^{-1}(B_2)$, alors $\omega \in f(B_1)$ et $\omega \in f(B_2)$ ce qui contredit l'hypothèse $B_1 \cap B_2 = \emptyset$. Par suite,

$$\mu_f \left(\bigcup_{n \geq 0} B_n \right) = \mu \left[f^{-1} \left(\bigcup_{n \geq 0} B_n \right) \right] = \mu \left[\bigcup_{n \geq 0} f^{-1}(B_n) \right] = \sum_{n \geq 0} \mu \left[f^{-1}(B_n) \right] = \sum_{n \geq 0} \mu_f(B_n).$$

□

Definition 22 (Mesure image)

μ_f est appelée mesure image de μ par f .

Une application permet de passer d'un espace mesuré $(\Omega, \mathcal{A}, \mu)$ à un autre espace mesuré (E, \mathcal{B}, μ_f)

Definition 23 (Loi de probabilité d'une variable aléatoire réelle)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. La loi de probabilité \mathbb{P}_X de la variable aléatoire $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est une mesure de probabilité définie par

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(X^{-1}(B)), \quad \forall B \in \mathcal{B}_{\mathbb{R}}.$$

Il s'agit de la mesure image de \mathbb{P} par X .

Corollaire 2 (Continuité et mesurabilité)

Soient que (E_1, \mathcal{O}_1) et (E_2, \mathcal{O}_2) deux espaces topologiques et \mathcal{B}_1 et \mathcal{B}_2 leur tribu borélienne associée, et $f : E_1 \rightarrow E_2$ une application. On a

f est continue $\Rightarrow f$ est mesurable .

preuve:

On note simplement que $\mathcal{O}_2 \subset \mathcal{B}_2$ et $\sigma(\mathcal{O}_2) \subset \mathcal{B}_2$ puis

$$f^{-1}(\mathcal{O}_2) \subset \mathcal{O}_1 \subset \mathcal{B}_1$$

f est mesurable d'après le corollaire 1.

□

2. Produits d'espace mesurable

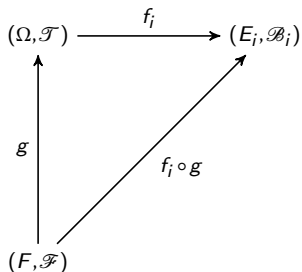
Soit Ω un ensemble, et $(E_i, \mathcal{B}_i)_{i \in I}$ une famille d'espace mesurable et $(f_i)_{i \in I}$ une famille d'applications

$$f_i : \Omega \rightarrow (E_i, \mathcal{B}_i).$$

Definition 24

La tribu engendrée par la famille $(f_i)_{i \in I}$ est la plus petite tribu sur Ω pour laquelle les f_i sont mesurables.

Il s'agit de la plus petite tribu contenant $\{f_i^{-1}(B) ; i \in I, B \in \mathcal{B}_i\}$. On la notera \mathcal{T} .
Soit $g : F \rightarrow \Omega$.



Theoreme 6

Une condition nécessaire et suffisante pour que g soit mesurable est que , pour tout $i \in I$, $f_i \circ g$ soit mesurable.

preuve:

Si g est mesurable alors les composées $f_i \circ g$ sont mesurables puisque \mathcal{T} rend les f_i mesurable.

Réciproquement, considérons

$$\mathcal{G} = \{f_i^{-1}(B) ; i \in I , B \in \mathcal{B}_i\}$$

et supposons que pour tout $i \in I$, $f_i \circ g$ soit mesurable. Alors, pour tout $B \in \mathcal{B}_i$

$$(f_i \circ g)^{-1}(B) \in \mathcal{F}$$

puis

$$g^{-1}(f_i^{-1}(B)) \in \mathcal{F}$$

On en déduit que $g^{-1}(\mathcal{G}) \subset \mathcal{F}$ et finalement

$$\sigma[g^{-1}(\mathcal{G})] = g^{-1}[\sigma(\mathcal{G})] = g^{-1}(\mathcal{T}).$$

g est bien mesurable.

□

Soient $(\Omega_1, \mathcal{A}_\infty)$ et $(\Omega_2, \mathcal{A}_\epsilon)$ deux espaces mesurables. On désigne par

$$p_1 : (\omega_1, \omega_2) \mapsto \omega_1 \text{ et } p_2 : (\omega_1, \omega_2) \mapsto \omega_2, (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2.$$

les applications projections canoniques.

Definition 25

La tribu produit $\mathcal{A}_1 \otimes \mathcal{A}_2$ sur $\Omega_1 \times \Omega_2$ est la tribu engendré par les applications p_1 et p_2 , c'est à dire la plus petite tribu rendant mesurables les applications projections.

Proposition 10

Soit g une application définie sur un espace mesurable (F, \mathcal{F}) , à valeur dans un espace produit $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$. Une condition nécessaire et suffisante pour que g soit mesurable est que

$$p_1 \circ g : (F, \mathcal{F}) \mapsto (\Omega_1, \mathcal{A}_1) \text{ et } p_2 \circ g : (F, \mathcal{F}) \mapsto (\Omega_2, \mathcal{A}_2)$$

soient mesurables.

Deux conséquences immédiates:

- Une application mesurable de (Ω, \mathcal{A}) à valeurs dans $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$ n'est rien d'autre qu'un couple d'applications mesurables de (Ω, \mathcal{A}) à valeur dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.
- Une application de (Ω, \mathcal{A}) à valeur dans $(\mathbb{C}, \mathcal{B}_{\mathbb{C}})$ est mesurable si et seulement si $\Re f$ et $\Im f$ sont mesurables de (Ω, \mathcal{A}) dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$

3. Propriétés des applications mesurables (numériques)

Soient f et g sont deux applications mesurables de (Ω, \mathcal{A}) dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Proposition 11

①

f est mesurable $\Leftrightarrow \forall a \in \mathbb{R}, \{f < a\} \in \mathcal{A}$

Valide aussi avec $\{f \leq a\}$, $\{f > a\}$, et $\{f \geq a\}$.

②

f, g mesurables $\Rightarrow \{f < g\}, \{f \leq g\}, \{f = g\}, \{f \neq g\} \in \mathcal{A}$

preuve:

① On remarque simplement que $\{f < a\} = f^{-1}(]-\infty, a[)$

② Soit $\omega \in \{f < g\}$ alors

$$\begin{aligned} f(\omega) < g(\omega) &\Leftrightarrow \exists r \in \mathbb{Q}, f(\omega) < r < g(\omega) \\ &\Leftrightarrow \exists r \in \mathbb{Q} \omega \in \{f < r\} \cap \{g > r\} \\ &\Leftrightarrow \omega \in \bigcup_{r \in \mathbb{Q}} \{f < r\} \cap \{g > r\} \end{aligned}$$

On en déduit que $\{f < g\} = \bigcup_{r \in \mathbb{Q}} \{f < r\} \cap \{g > r\} \in \mathcal{A}$. Les autres propriétés se déduisent des observations suivantes

$$\{f \leq g\} = \Omega / \{f > g\}, \{f = g\} = \{f \geq g\} \cap \{f \leq g\} \text{ et } \{f \neq g\} = \Omega / \{f = g\}$$

□

Proposition 12 (Vecteur de fonctions mesurables)

$h: \omega \in \Omega \mapsto (f(\omega), g(\omega))$ est une fonction mesurable de (Ω, \mathcal{A}) dans $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$

preuve:

Soit $A \times B$ un pavé dans $\mathcal{B}(\mathbb{R}^2)$, on a

$$h^{-1}(A \times B) = f^{-1}(A) \cap g^{-1}(B) \in \mathcal{A}$$

Comme $\sigma(A \times B) = \mathcal{B}(\mathbb{R}^2)$ alors h est mesurable par application du corollaire 1. □

Proposition 13 (Opérations sur les fonctions mesurables)

① Les applications

$$f + g; \alpha \times f, \text{ avec } \alpha \in \mathbb{R}; f \times g;$$

sont mesurables.

② Les applications

$$\inf(f, g); \sup(f, g); f^+ = \sup(f, 0); f^- = \inf(f, 0); |f|$$

sont mesurables.

③ Soit $(f_n)_{n \geq 0}$ une suite d'applications mesurables de (Ω, \mathcal{A}) à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Si $\sup_{n \geq 0} f_n$ et $\inf_{n \geq 0} f_n$ sont finies alors

$$\sup_{n \geq 0} f_n; \inf_{n \geq 0} f_n; \overline{\lim}_{n \rightarrow +\infty} f_n; \underline{\lim}_{n \rightarrow +\infty} f_n.$$

sont mesurables. En particulier, si $\lim_{n \rightarrow +\infty} f_n = f$ alors f est mesurable.

preuve:

- 1 L'application $\Psi: \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $\Psi(x, y) = x + y$ est continue donc mesurable. On remarque que l'application $f + g$ est la composée de $\Psi \circ h$, où $h: (x, y) \mapsto (f(x), g(x))$, ce qui la rend mesurable. Le raisonnement est similaire pour af et fg .
- 2 On remarque simplement que $\{\sup(f, g) > a\} = f^{-1}([a, +\infty[) \cap g^{-1}([a, +\infty[) \in \mathcal{A}$. Le raisonnement est similaire pour $\inf(f, g)$, $\sup(f, 0)$ et $\inf(f, 0)$. On garde $|f|$ pour l'examen :).
- 3 Par définition $\overline{\lim}_{n \rightarrow +\infty} f_n = \inf_{k \geq 0} \sup_{n \geq k} f_n$ qui est mesurable en vertu du point précédent. De même $\underline{\lim}_{n \rightarrow +\infty} f_n$. Enfin si f_n tend vers f alors

$$f = \lim_{n \rightarrow +\infty} f_n = \overline{\lim}_{n \rightarrow +\infty} f_n = \underline{\lim}_{n \rightarrow +\infty} f_n.$$

est mesurable.

A.1 Existence et unicité de la mesure de Lebesgue

Il est naturel de mesurer un intervalle de \mathbb{R} par sa longueur ou une union d'intervalles disjoints par la somme de leur longueur respective.

Definition 26 ($\mathcal{I}_{\mathbb{R}}$, application longueur)

L'application longueur $l: \mathcal{I}_{\mathbb{R}} \rightarrow \mathbb{R}_+$ définie par

$$l(]a, b]) = b - a, \text{ et } l(\emptyset) = 0.$$

L'objectif est de définir une application permettant de mesurer une partie quelconque de \mathbb{R} ou pour être précis les ouverts de \mathbb{R} . Comme $\mathbb{R} = \bigcup_{k=1}^{+\infty}]-k, k[$ alors toute partie de \mathbb{R} peut être recouverte. Cette application sera une mesure sur $\mathcal{B}(\mathbb{R})$ coïncidant avec l'application longueur sur les intervalles ouverts.

Theoreme 7 (Carathéodory)

Il existe une et une seule mesure sur $\mathcal{B}(\mathbb{R})$, notée λ , appelée mesure de Lebesgue, telle que

$$\lambda(]a, b]) = b - a, \text{ pour tout } -\infty < a < b < +\infty.$$

preuve (synthétique):

Existence:

Pour une partie $A \subset \mathcal{P}(\mathbb{R})$ on introduit l'instrument de mesure suivant.

Definition 27 (Mesure extérieure de Lebesgue)

On appelle mesure extérieure de Lebesgue dans \mathbb{R} l'application $\lambda^* : \mathbb{R} \rightarrow \overline{\mathbb{R}}^+$ définie, pour tout $A \in \mathcal{P}(\mathbb{R})$, par

$$\lambda^* = \inf \left\{ \sum_{n=0}^{+\infty} l(I_n) ; (I_n)_{n \in \mathbb{N}} \in \mathcal{I}_{\mathbb{R}} \text{ et } A \subset \bigcup_{n=1}^{\infty} I_n \right\}$$

Proposition 14 (Propriétés de λ^*)

L'application λ^* vérifie les propriétés suivantes

- 1 $\lambda^*(\emptyset) = 0$
- 2 $\lambda^*(A) \leq \lambda^*(B)$ pour $A, B \subset \mathbb{R}$ telles que $A \subset B$ (λ^* est monotone).
- 3 Soit $(A_n)_{n \in \mathbb{N}} \in \mathcal{P}(\mathbb{R})$ et $A = \bigcup_{n \in \mathbb{N}} A_n$ alors

$$\lambda^*(A) \leq \sum_{n \in \mathbb{N}} \lambda^*(A_n).$$

(λ^* est sous σ -additive)

- 4 $\lambda^*([a, b]) = b - a$ pour tout $a, b \in \mathbb{R}$ tels que $-\infty < a < b < +\infty$.

λ^* n'est pas σ -additive et n'est donc pas une mesure sur $\mathcal{P}(\mathbb{R})$. On va montrer que λ^* est une mesure si on restreint l'application à $\mathcal{B}(\mathbb{R})$. Concrètement, on montre que λ^* est une mesure sur une tribu \mathcal{L} qui englobe $\mathcal{B}(\mathbb{R})$.

Definition 28 (La tribu de Lebesgue \mathcal{L})

Soit

$$\mathcal{L} = \{E \in \mathcal{P}(\mathbb{R}) ; \lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \cap E^c)\}, \text{ pour tout } A \subset \mathcal{P}(\mathbb{R}),$$

un sous-ensemble de $\mathcal{P}(\mathbb{R})$, appelé tribu de Lebesgue.

Proposition 15 (Propriétés de \mathcal{L})

- 1 \mathcal{L} est une tribu sur \mathbb{R} ,
- 2 $\lambda^*|_{\mathcal{L}} : \mathcal{L} \rightarrow \overline{\mathbb{R}}_+$ est une mesure.

Les membres de la tribu \mathcal{L} réalisent un bon partage des parties de \mathbb{R} .

preuve:

Il est immédiat que $\mathbb{R} \in \mathcal{L}$ et que \mathcal{L} est stable par passage au complémentaire. De même, on remarque que $\lambda^*(\emptyset) = 0$.

Etape 1. On va montrer que \mathcal{L} est stable par réunion finie et que λ^* vérifie, pour $(E_i)_{i=1,\dots,n}$ telles que $E_i \cap E_j = \emptyset$ pour $i \neq j$,

$$\lambda^* \left(A \cap \bigcup_{i=1}^n E_i \right) = \sum_{i=1}^n \lambda^*(A \cap E_i).$$

Soit $E_1, E_2 \subset \mathcal{L}$ et $E = E_1 \cup E_2$. On rappelle que $E \subset \mathcal{L}$ si

$$\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \cap E^c)$$

Nous savons que $\lambda^*(A) \leq \lambda^*(A \cap E) + \lambda^*(A \cap E^c)$ du fait de la σ sous-additivité de λ^* .
Notons que

$$\begin{aligned} \lambda^*(A \cap E) &= \lambda^*[A \cap (E_1 \cup E_2)] \\ &= \lambda^*[(A \cap E_1) \cup (A \cap E_2)] \\ &= \lambda^*[(A \cap E_1) \cup (A \cap E_2 \cap E_1^c)] \\ &\leq \lambda^*(A \cap E_1) + \lambda^*(A \cap E_2 \cap E_1^c). \end{aligned} \quad (1)$$

Comme $E_2 \subset \mathcal{L}$ et $A \cap E_1^c \in \mathcal{P}(\mathbb{R})$ alors

$$\lambda^*(A \cap E_1^c) = \lambda^*(A \cap E_1^c \cap E_2) + \lambda^*(A \cap E_1^c \cap E_2^c) = \lambda^*(A \cap E_1^c \cap E_2) + \lambda^*(A \cap E^c). \quad (2)$$

On a également

$$\lambda^*(A) = \lambda^*(A \cap E_1) + \lambda^*(A \cap E_1^c). \quad (3)$$

puisque $E_1 \subset \mathcal{L}$. En ré-injectant (2) et (3) dans l'inégalité (1), on obtient

$$\lambda^*(A) \geq \lambda^*(A \cap E) + \lambda^*(A \cap E^c).$$

Supposons que $E_1 \cap E_2 = \emptyset$ alors

$$\begin{aligned} \lambda^*(A \cap E) &= \lambda^*[A \cap (E_1 \cup E_2)] \\ &= \lambda^*[(A \cap E_1) \cup (A \cap E_2)] \\ &= \lambda^*\{[(A \cap E_1) \cup (A \cap E_2)] \cap E_1\} + \lambda^*\{[(A \cap E_1) \cup (A \cap E_2)] \cap E_1^c\} \\ &= \lambda^*(A \cap E_1) + \lambda^*(A \cap E_2) \end{aligned}$$

Les deux propriétés se généralisent pour une suite $(E_n)_{n=1, \dots, n}$ par récurrence.

Etape 2.

Considérons $(E_n)_{n \in \mathbb{N}} \subset \mathcal{L}$ et $E = \bigcup_{n \in \mathbb{N}} E_n$. Soit

$$F_0 = E_0 \text{ et } F_n = E_n \setminus \left(E_n \cap \bigcup_{p=0}^{n-1} F_p \right)$$

de sorte que F_0, F_1, \dots appartiennent à \mathcal{L} , soient disjoints, et vérifient $E = \bigcup_{n \in \mathbb{N}} F_n$. On a

$$\begin{aligned} \lambda^*(A) &= \lambda^* \left(A \cap \bigcup_{p=0}^n F_p \right) + \lambda^* \left[A \cap \left(\bigcup_{p=0}^n F_p \right)^c \right] \\ &\geq \lambda^* \left(A \cap \bigcup_{p=0}^n F_p \right) + \lambda^* [A \cap E] \\ &\geq \sum_{p=0}^n \lambda^*(A \cap F_p) + \lambda^* [A \cap E] \end{aligned}$$

En passant à la limite lorsque $n \rightarrow +\infty$, on obtient

$$\begin{aligned} \lambda^*(A) &\geq \sum_{p=0}^{+\infty} \lambda^*(A \cap F_p) + \lambda(A \cap E^c) \\ &\geq \lambda^*(A \cap E) + \lambda(A \cap E^c) \text{ sous } \sigma\text{-additivité.} \end{aligned}$$

ce qui prouve que $E \subset \mathcal{L}$.

On montre maintenant que λ^* est bien une mesure sur \mathcal{L} . Soit $(E_n)_{n \in \mathbb{N}}$ telle que $E_n \cap E_m = \emptyset$ si $n \neq m$. Comme $\bigcup_{p=0}^n E_p \subset E$ alors

$$\begin{aligned}\lambda^*(A \cap E) &\geq \lambda\left(\bigcup_{p=0}^n A \cap E_p\right) \\ &= \sum_{p=0}^n \lambda(A \cap E_p).\end{aligned}$$

On obtient $\lambda^*(E) \geq \sum_{p=0}^{+\infty} \lambda^*(E_p)$ en choisissant $A = E$ puis en passant à la limite lorsque $n \rightarrow +\infty$. De plus $\lambda^*(E) \leq \sum_{p=0}^{+\infty} \lambda^*(E_p)$ en vertu de la sous σ -additivité. On a donc

$$\lambda^*(E) = \sum_{n=0}^{+\infty} \lambda^*(E_n).$$

Pour montrer l'existence du théorème (7), il suffit de montrer que $]a, +\infty[\subset \mathcal{L}$ pour tout $a \in \mathbb{R}$ car dans ce cas $\mathcal{B}(\mathbb{R}) \subset \mathcal{L}$ puisque $]a, +\infty[$ engendre $\mathcal{B}(\mathbb{R})$. Soit $E =]a, +\infty[$, pour $a \in \mathbb{R}$ et $A \in \mathcal{P}(\mathbb{R})$, on veut montrer que

$$\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \cap E^c). \quad (4)$$

D'après la définition de λ^* , il existe $(I_n)_{n \in \mathbb{N}} \in \mathcal{A}_{\mathbb{R}}$, telle que $A \subset \bigcup_{n \in \mathbb{N}} I_n$ et $\lambda^*(A) = \sum_{n \in \mathbb{N}} l(I_n) - \epsilon$. Comme

$$\begin{cases} A \cap E \subset \bigcup_{n \in \mathbb{N}} I_n \cap E, \\ A \cap E^c \subset \bigcup_{n \in \mathbb{N}} I_n \cap E^c, \end{cases}$$

alors la σ sous-additivité implique que

$$\begin{cases} \lambda^*(A \cap E) \leq \sum_{n \in \mathbb{N}} \lambda^*(I_n \cap E), \\ \lambda^*(A \cap E^c) \leq \sum_{n \in \mathbb{N}} \lambda^*(I_n \cap E^c). \end{cases}$$

On a

$$\begin{aligned} \lambda^*(A \cap E) + \lambda^*(A \cap E^c) &\leq \sum_{n \in \mathbb{N}} \lambda^*(I_n \cap E) + \lambda^*(I_n \cap E^c) \\ &= \sum_{n \in \mathbb{N}} l(I_n), \end{aligned}$$

puis $\lambda^*(A \cap E) + \lambda^*(A \cap E^c) \leq \lambda^*(A) + \epsilon$, où ϵ peut être choisi arbitrairement petit. Finalement, $\lambda^*(A) \leq \lambda^*(A \cap E) + \lambda^*(A \cap E^c)$ est une conséquence de la σ sous-additivité, ce qui permet de conclure à l'égalité (4).

Pour l'unicité, on montre que s'il existe une autre mesure m sur $\mathcal{B}(\mathbb{R})$ telle que $m(]a, b[) = b - a$ alors elle coïncide avec λ^* . La proposition suivante est dès lors très utile.

Proposition 16 (Condition suffisante pour l'égalité de deux mesures)

Soit (Ω, \mathcal{A}) un espace mesurable et m, μ deux mesures sur \mathcal{A} . Supposons qu'il existe $\mathcal{C} \subset \mathcal{A}$ tel

- 1 \mathcal{C} engendre \mathcal{A}
- 2 \mathcal{C} est stable par intersection fini
- 3 Il existe $(C_n)_{n \in \mathbb{N}} \in \mathcal{C}$ telle que $C_n \cap C_m = \emptyset$ si $n \neq m$, et $\Omega = \bigcup_{n \in \mathbb{N}} C_n$.
- 4 $m(C) = \mu(C)$ pour tout $C \subset \mathcal{C}$

On a alors $m = \mu$

La preuve se termine en appliquant la proposition (16), avec $\mathcal{C} = \{]a, b[, -\infty < a < b < +\infty\}$. On vérifie que

- $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R})$
- \mathcal{C} est stable par intersection
- Considérons la suite

$$F_n =]n, n+1], n \in \mathbb{Z}$$

est dénombrable, disjointe et telle que $\bigcup_{n \in \mathbb{Z}} F_n = \mathbb{R}$

- On a par continuité décroissante

$$\begin{aligned}
 m(]a, b]) &= \lim_{n \rightarrow +\infty} m(]a, b + \frac{1}{n}) \\
 &= \lim_{n \rightarrow +\infty} b - a + \frac{1}{n} \\
 &= b - a \\
 &= \lambda^*(]a, b])
 \end{aligned}$$

On définit alors $\lambda := \lambda^*_{|\mathcal{B}(\mathbb{R})}$

En résumé,

- $\lambda^* : \mathcal{P}(\mathbb{R}) \rightarrow \overline{\mathbb{R}}^+ \Rightarrow \lambda^*_{|\mathcal{L}} : \mathcal{L} \rightarrow \overline{\mathbb{R}}_+$ est une mesure
- $\Rightarrow \lambda := \lambda^*_{|\mathcal{B}(\mathbb{R})} : \mathcal{B}(\mathbb{R}) \rightarrow \overline{\mathbb{R}}_+$ est la seule mesure telle que $\lambda(]a, b]) = b - a$
- \Rightarrow La mesure de Lebesgue

Références bibliographiques I

Mes notes se basent sur les documents suivants [1, 3, 2]



Michel Carbon.

Probabilités 1 et 2.

Note de cours ENSAI, 2009.



Olivier Garet and Aline Kurtzmann.

De l'intégration aux probabilités, volume 470.

Ellipses, 2011.



Jean-François Le Gall.

Intégration, probabilités et processus aléatoires.

Ecole Normale Supérieure de Paris, 2006.