M61 Intégration-Probabilités

Fascicule de cours

L3 Mathématiques

Université de Lille

Année 2020-2021

mylene.maida@univ-lille.fr

AJOUTER CREDIT

Préambule : suivre un cours à distance

Suivre un cours en distanciel est exigeant. Il faut essayer de se mettre autant que possible dans les mêmes – et même idéalement de meilleures – conditions que pour un cours en présentiel.

- Sauf circonstance grave, **éteignez et rangez tous les appareils électroniques** (téléphone, tablette etc) sauf bien sûr celui qui vous sert à suivre le cours.
- Même sur l'appareil qui vous sert à suivre le cours, **ne lisez ni emails ni sms** pendant le cours
- **Prenez des notes**, de préférence sur papier ; cela aide à fixer votre attention
- Ayez près de vous un peu de brouillon pour faire des calculs, chercher des démonstrations
- **Relisez systématiquement le cours précédent** avant d'assister à la séance ; c'est encore plus important pour un cours à distance
- **Cherchez les exercices** qui vous seront indiqués d'une séance sur l'autre
- N'hésitez pas à **poser des questions** à l'enseignante pendant le cours ou après
- **Travaillez** si possible **en petit groupe** avec quelques camarades, c'est important pour la motivation
- **Demandez de l'aide** sans attendre à l'enseignante si vous vous sentez décrocher

Quelques références bibliographiques

- **De l'intégration aux probabilités,** O. Garet et A. Kurtzmann, Editions Ellipses
- **Mesure, intégration, probabilités,** T. Gallouët et R. Herbin, Editions Ellipses
- **Probbabilités L3-M1,** P. Barbe et M. Ledoux, Editions EDP Sciences.

Introduction

Vous avez déjà rencontré, au lycée puis en L2, des exemples d'expériences aléatoires (lancers de pièces, de dés, tirs de fléchettes sur une cible etc) et de variables aléatoires qui leur sont attachées (nombres de pile au bout de n lancers, somme de deux dés, score obtenu aux fléchettes etc). Certaines de ces expériences aléatoires ont un nombre dénombrable de résultats possibles. Leur étude a fait l'objet du cours **Probabilités discrètes** du S4. Mais ce n'est pas toujours le cas et nous nous attacherons dans ce cours à développer une théorie générale des probabilités, qui s'appuie principalement sur la théorie de la mesure. En outre, la théorie de la mesure sera fondamentale pour la construction d'une théorie robuste de l'intégration. Vous connaissez déjà l'intégrale de Riemann (celle du lycée et de la L1) mais nous aurons besoin d'introduire et d'utiliser une autre intégrale, développée notamment par Lebesgue.

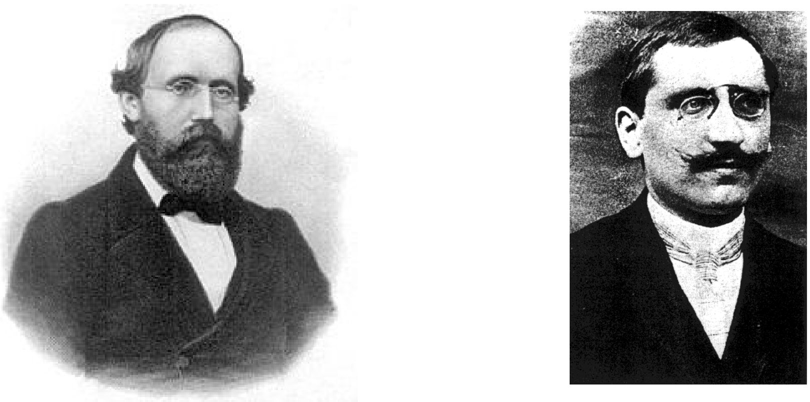
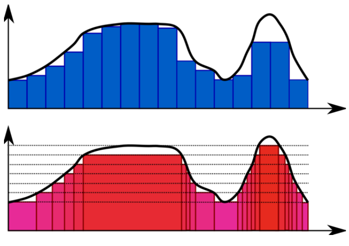


FIGURE 1 – A gauche, Bernhard Riemann, 1826-1866, mathématicien allemand ; à droite, Henri Lebesgue, 1875-1941, mathématicien français

En concevant son intégrale (autour de 1900), Henri Lebesgue l'a lui-même comparée à l'intégrale de Riemann : ≪ Imaginez que je doive payer une certaine somme ; je peux sortir les pièces de mon porte-monnaie comme elles viennent pour arriver à la somme indiquée, ou sortir toutes les pièces et les choisir selon leur valeur. La première méthode est l'intégrale de Riemann, la deuxième correspond à mon intégrale. ≫ Pour comprendre cette phrase, il faut préciser que l'intégration de Riemann ≪ parcourt ≫ le segment et exploite au fur et à mesure la ≪ hauteur ≫ y de la fonction, alors que l'intégration de Lebesgue exploite la ≪ taille ≫ des ensembles de niveau $f = y$ pour toutes les valeurs de y comme l'illustre la figure ci-dessous.



- En bleu : intégrale de Riemann. Avantage : calculs "faciles" à l'aide de primitives. Inconvénients : définition plus compliquée en dimension supérieure. Utilisation théorique compliquée pour certains usages (intégrales généralisées si l'intervalle n'est pas borné ou si la fonction n'est pas bornée)
- En rouge : intégrale de Lebesgue. On a besoin de mesurer $\{x \in \mathbb{R} | f(x) \in [c_i, c_{i+1}]\}$. Inconvénient : moins intuitif. Avantages : la construction est la même sur $\mathbb{R}$, $\mathbb{R}^2$, ... et même sur un espace de départ abstrait.

Utilisation simple grâce à des outils théoriques puissants même sur un ensemble de départ non borné et une fonction non bornée (convenable).

Ces deux intégrales coïncident sur une large classe de ≪ bonnes ≫ fonctions mais la théorie de l'intégration par rapport à une mesure générale, développée par Lebesgue, présente (au moins) les avantages importants suivants :

— elle permet d'intégrer des fonctions plus générales, pour lesquelles l'intégrale de Riemann n'est pas définie, par exemple $\mathbf{1}_{\mathbb{Q}}$,

— elle permet d'avoir des théorèmes de convergence plus robustes,

— dans ce nouveau cadre, l'ensemble des fonctions intégrables muni de la norme L^1 ou L^2 a une structure d'espace de Banach, en particulier il est complet,

— elle permet de donner une définition simple et générale de l'espérance d'une variable aléatoire (dont vous avez déjà vu des exemples dans le cas discret)

La théorie de la mesure est également à la base de la théorie moderne des probabilités, développée notamment dans les années 30 par Kolmogorov.

Figure 2 – Andrei Kolmogorov, 1903-1987, mathématicien russe

Table des matières

Chapitre 1

Tribus, fonctions mesurables

Nous allons dans ce chapitre, introduire les concepts fondamentaux de la théorie de la mesure, que vous avez peut-être déjà rencontrés dans le cours de L2 **Probabilités discrètes**. Comme nous l'avons expliqué dans l'introduction, pour définir l'intégrale de Lebesgue d'une fonction, il nous faudra **mesurer** les sous-ensembles de niveau d'une fonction. Malheureusement, il est impossible de définir une bonne notion de mesure sur tous les sous-ensembles de $\mathbb{R}$: l'ensemble des sous-ensembles de $\mathbb{R}$ est beaucoup trop gros. C'est en fait le cas pour tous les univers Ω qui ne sont pas dénombrables. On va donc introduire la notion de tribu : ce sont des sous-ensembles de l'univers Ω que l'on saura mesurer, c'est-à-dire sur lesquels une **mesure** pourra être définie.

La famille d'ensembles qu'on peut mesurer doit satisfaire des conditions qui nous semblent naturelles : si on peut mesurer deux sous-ensembles d'un ensemble Ω, alors on doit pouvoir mesurer leur réunion, leur intersection et leur différence : la famille d'ensembles qu'on peut mesurer doit donc être stable par rapport à la réunion, à l'intersection et à la différence de deux de ses ensembles. La définition suivante requiert un peu plus, à savoir que l'ensemble Ω lui-même appartienne à cette famille.

Dans toute la suite, si Ω est un ensemble, on notera $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω.

Définition. Un ensemble $\mathcal{A} \subset \mathcal{P}(\Omega)$ de parties de Ω est une **algèbre (d'ensembles)** ou **algèbre booléenne** sur Ω si les trois conditions suivantes sont vérifiées :

1. $\Omega \in \mathcal{A}$;
2. si $A \in \mathcal{A}$ et $B \in \mathcal{A}$ alors $A \cap B \in \mathcal{A}$;
3. si $A \in \mathcal{A}$ alors $\Omega \setminus A \in \mathcal{A}$.

Par récurrence, il s'ensuit qu'une algèbre d'ensembles sur X est stable par rapport à la réunion et à l'intersection d'un nombre fini quelconque de ses éléments.

Il se trouve toutefois que se limiter aux opérations ensemblistes **finies** est une restriction trop forte pour qu'on puisse obtenir une théorie de la mesure suffisamment riche. Notre but est de dépasser la théorie de l'intégration de Riemann et d'obtenir des théorèmes de passage au limite dont cette théorie manque cruellement ; il faudra alors admettre que la famille d'ensembles qu'on peut mesurer soit stable par rapport aux réunions et intersections **dénombrables**. On arrive ainsi définition de **tribu**, la notion naturelle sur laquelle fonder la théorie de la mesure.

Tribus

Définition, tribu engendrée

Définition. (tribu, ensemble mesurable).

> Un ensemble $\mathcal{F}$ de parties de Ω est une **tribu** ou **σ-algèbre** sur Ω si
> 1. $X \in \mathcal{F}$.
> 2. Si $A \in \mathcal{F}$, alors $\Omega \setminus A \in \mathcal{F}$.
> 3. Si $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$ alors $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.
>
> Un **espace mesurable** est un couple $(\Omega, \mathcal{F})$ où Ω est un ensemble et $\mathcal{F}$ est un tribu sur Ω. Un élément $A \in \mathcal{F}$ est appelée **partie (ou ensemble) mesurable** de Ω (relativement à la tribu $\mathcal{F}$).

Remarque. Dans les ouvrages de probabilité on utilise des notations et des termes un peu différents des ceux utilisés dans les ouvrages sur la théorie de la mesure.

L'espace Ω est souvent appelé **univers**, une tribu sur Ω notée $\mathcal{A}$ ou $\mathcal{B}$, tandis que les parties mesurables, c.-à-d. les éléments de la tribu donnée sur Ω, sont appelés les **événements**.

Exemple.

1. La plus petite tribu sur Ω est la **tribu triviale (ou grossière)** $\{\emptyset, \Omega\}$.
2. La plus grande tribu sur Ω est la **tribu discrète** $\mathcal{P}(\Omega)$.

Exercice. Donner toutes les tribus sur l'ensemble $\Omega = \{1, 2, 3\}$. (Il y en a 5)

Proposition 1 *(Intersection de tribus).*

> *Soit I un ensemble **non vide** de tribus sur l'ensemble Ω. Alors $\bigcap_{\mathcal{F}\in I} \mathcal{F}$ est une tribu sur Ω.*

Preuve.

1. Comme $\Omega \in \mathcal{F}$ pour tout $\mathcal{F} \in I$, on a $\Omega \in \bigcap_{\mathcal{F}\in I} \mathcal{F}$.
2. Si $A \in \bigcap_{\mathcal{F}\in I} \mathcal{F}$ alors $A \in \mathcal{F}$ pour tout $\mathcal{F} \in I$. Puisque $\mathcal{F}$ une tribu sur l'ensemble Ω, on a $\Omega \setminus A \in \mathcal{F}$, pour tout $\mathcal{F} \in I$. Donc $\Omega \setminus A \in \bigcap_{\mathcal{F}\in I} \mathcal{F}$.
3. Si $(A_n)_{n\in\mathbb{N}}$ est une suite de parties de Ω telle que $A_n \in \bigcap_{\mathcal{F}\in I} \mathcal{F}$, pour tout $n \in \mathbb{N}$. Alors $A_n \in \mathcal{F}$ pour tout $\mathcal{F} \in I$ et tout $n \in \mathbb{N}$. Puisque $\mathcal{F}$ une tribu sur l'ensemble Ω, on a $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$, pour tout $\mathcal{F} \in I$. Donc $\bigcup_{n\in\mathbb{N}} A_n \in \bigcap_{\mathcal{F}\in I} \mathcal{F}$.

Exercice. Soit $\Omega = \{1, 2, 3\}$. Trouver deux tribus $\mathcal{F}_1$ et $\mathcal{F}_2$ telles que $\mathcal{F}_1 \cup \mathcal{F}_2$ n'est pas une tribu.

Corollaire (tribu engendrée).

> Si $\mathcal{A} \subset \mathcal{P}(\Omega)$ est une famille de parties de l'ensemble Ω, il existe une unique plus petite tribu sur Ω contenant $\mathcal{A}$. Elle est notée $\sigma(\mathcal{A})$ et dite la **tribu engendrée par** $\mathcal{A}$.

Preuve. Soit I l'ensemble de toutes les tribus $\mathcal{F} \subset \mathcal{P}(\Omega)$ telles que $\mathcal{A} \subset \mathcal{F}$. L'ensemble I n'est pas vide car $\mathcal{P}(\Omega) \in I$. Par la Proposition 1, $\mathcal{F}_0 := \bigcap_{\mathcal{F}\in I} \mathcal{F}$ est une tribu sur Ω. Évidemment, $\mathcal{A} \subset \mathcal{F}_0$, car $\mathcal{A} \subset \mathcal{F}$ pour tout $\mathcal{F} \in I$. Par ailleurs, si $\mathcal{N} \subset \mathcal{P}(\Omega)$ est une tribu telle que $\mathcal{A} \subset \mathcal{N}$, on a $\mathcal{N} \in I$ et donc $\mathcal{N} \supset \mathcal{F}_0$. Donc $\mathcal{F}_0$ est la plus petite tribu sur Ω contenant $\mathcal{A}$.

Il est rare qu'on puisse expliciter la tribu $\sigma(\mathcal{A})$ engendrée par $\mathcal{A}$.

Exercice. Soit $\Omega = \{1, 2, 3\}$. Décrivez $\sigma(\{1\}), \sigma(\{1, 2\}), \sigma(\{1\}, \{2\})$.

Une définition utile est la suivante.

Définition.

Soient E une partie de Ω et $\mathcal{F}$ une tribu sur Ω. La famille $\{A \cap E \mid A \in \mathcal{F}\}$ est une tribu sur E, appelée **trace de** $\mathcal{F}$ **sur** E, notée $\mathcal{F}_E$.

Exercice. Vérifier que la famille $\{A \cap E \mid A \in \mathcal{F}\}$ est bien une tribu sur E.

On peut facilement construire une tribu sur un espace produit.

Définition. (Tribu produit).

Si $\mathcal{F}_1$ est une tribu sur Ω_1 et $\mathcal{F}_2$ est une tribu sur Ω_2 alors la tribu produit $\mathcal{F}_1 \otimes \mathcal{F}_2$ sur $\Omega_1 \times \Omega_2$ est la tribu engendrée par la famille

$$\{A \times B \in \mathcal{P}(\Omega_1 \times \Omega_2) \mid A \in \mathcal{F}_1, B \in \mathcal{F}_2\}.$$

La définition se généralise aisément à un produit d'un nombre fini de tribus.

Exercice. Montrer que la tribu $\mathcal{F}_1 \otimes \mathcal{F}_2$, produit de deux tribus $\mathcal{F}_1$ et $\mathcal{F}_2$ données sur les ensembles Ω_1 et Ω_2 respectivement, est engendrée par la famille d'ensembles

$$\{A \times \Omega_2 \in \mathcal{P}(\Omega_1 \times \Omega_2) \mid A \in \mathcal{F}_1\} \cup \{\Omega_1 \times B \in \mathcal{P}(\Omega_1 \times \Omega_2) \mid B \in \mathcal{F}_2\}.$$

Tribu de Borel (ou borélienne)

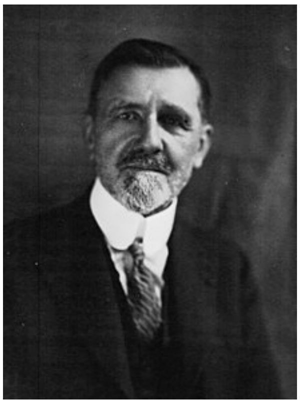


FIGURE 1.1 – Le mot « borélien » vient du mathématicien français Émile Borel (1871-1956) qui a mis en place (avec d'autres) la théorie de la mesure. Il fut maître de conférences à la Faculté des Sciences de Lille.

Définition. (Tribu borélienne).

Si Ω est un espace topologique la **tribu de Borel**, notée $\mathcal{B}_\Omega$ ou $\mathcal{B}(\Omega)$, est la tribu engendrée par les ouverts de Ω. Un ensemble mesurable relativement à la tribu de Borel est appelé **un borélien**.

On s'intéressera en particulier à la tribu borélienne sur $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^n$.

Proposition 2 *(La tribu borélienne de $\mathbb{R}^n$ est engendrée par les pavés).*

La famille (dénombrable) $\mathcal{I}^n(\mathbb{Q})$ formée par les pavés ouverts de centre rationnel et cotés rationnels engendre la tribu de Borel $\mathcal{B}_{\mathbb{R}^n}$ de $\mathbb{R}^n$.

Preuve. On a
$$\mathcal{I}^n(\mathbb{Q}) = \{]a_1, b_1[\times \cdots \times]a_n, b_n[\subset \mathbb{R}^n \mid a_i \in \mathbb{Q}, b_i \in \mathbb{Q}, i = 1, \ldots, n\}.$$
Comme tout $A \in \mathcal{I}^n(\mathbb{Q})$ est ouvert, $\sigma(\mathcal{I}^n(\mathbb{Q})) \subset \mathcal{B}_{\mathbb{R}^n}$.

Pour tout ouvert $U \subset \mathbb{R}^n$ et tout $x \in \mathbb{R}^n$ il existe un pavé ouvert P tel que $x \in P$ et $P \subset U$. Par la densité de $\mathbb{Q}$ dans $\mathbb{R}$, quitte à rapetisser P on peut supposer que $P \in \mathcal{I}^n(\mathbb{Q})$. On a donc

$$U = \bigcup_{P \in \mathcal{I}^n(\mathbb{Q}), P \subset U} P$$

Comme l'ensemble $I^n(\mathbb{Q})$ est dénombrable, la réunion ci-dessus est une réunion dénombrable. Il s'ensuit que tout ouvert U appartient à $\sigma(\mathcal{I}^n(\mathbb{Q}))$. Donc $\sigma(\mathcal{I}^n(\mathbb{Q})) \supset \mathcal{B}_{\mathbb{R}^n}$.

La preuve de la proposition suivante est laissée en exercice.

Proposition 3 *(Descriptions de la tribu borélienne sur $\mathbb{R}$).*

La tribu de Borel de $\mathbb{R}$ est également engendrée par

1. *Les demi-droites $]a, \infty[$ ouvertes à gauche ou les demi-droites $]-\infty, a[$ ouvertes à droite avec $a \in \mathbb{Q}$.*
2. *Les demi-droites $[a, \infty[$ fermées à gauche ou les demi-droites $]-\infty, a]$ fermées à droite, avec $a \in \mathbb{Q}$.*
3. *Les intervalles ouverts $]a, b[$, $a \in \mathbb{Q}$, avec $b \in \mathbb{Q}$.*
4. *Les intervalles fermés $[a, b]$, avec $a \in \mathbb{Q}$ et $b \in \mathbb{Q}$.*
5. *Les intervalles semi-ouverts (à gauche ou bien à droite).*

Proposition 4 *La tribu $\mathcal{B}_{\mathbb{R}^2}$ des boréliens de $\mathbb{R}^2$ est le produit $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$. La proposition se généralise aisément à $\mathcal{B}_{\mathbb{R}^n}$.*

Preuve. Soit $\tau_{\mathbb{R}}$ l'ensemble des parties ouvertes de $\mathbb{R}$. La tribu $\mathcal{B}_{\mathbb{R}}$ est engendré par les ouverts de $\tau_{\mathbb{R}}$. Donc la famille $\mathcal{F}_1 = \{A \times \mathbb{R} \mid A \in \tau_{\mathbb{R}}\}$ engendre la tribu $\mathcal{B}_{\mathbb{R}} \otimes \{\emptyset, \mathbb{R}\}$ et la famille $\mathcal{F}_2 = \{\mathbb{R} \times A \mid A \in \tau_{\mathbb{R}}\}$ engendre la tribu $\{\emptyset, \mathbb{R}\} \otimes \mathcal{B}_{\mathbb{R}}$. Par l'exercice 1.1.1 la famille $\mathcal{F}_1 \bigcup \mathcal{F}_2$ engendre la tribu $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$. Comme tout élément de $\mathcal{F}_1 \bigcup \mathcal{F}_2$ est ouvert dans $\mathbb{R}^2$, on obtient que $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{F}_1 \bigcup \mathcal{F}_2) \subset \mathcal{B}_{\mathbb{R}^2}$.

Soit U ouvert dans $\mathbb{R}^2$. Tout $x \in U$ appartient à un rectangle ouvert $A \times B$ contenu dans U et d'extrémités rationnelles. La famille de tels rectangles est dénombrable. Soit $A \times B$ un tel rectangle. Comme $A \times B = (A \times \mathbb{R}) \cap (\mathbb{R} \times B)$, on a $A \times B \in \sigma(\mathcal{F}_1 \bigcup \mathcal{F}_2)$. Il s'ensuit que tout ouvert $U \subset \mathbb{R}^2$ est une réunion dénombrable d'éléments de $\sigma(\mathcal{F}_1 \bigcup \mathcal{F}_2) = \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$. Donc tout ouvert de $\mathbb{R}^2$ appartient à la tribu $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$. On conclut que $\mathcal{B}_{\mathbb{R}^2} \subset \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$.

La généralisation à $\mathbb{R}^n$ est laissée en exercice au lecteur.

Une autre façon de contruire des tribus est de considérer l'image réciproque d'une tribu par une application.

Lemme. Soit $(Y, \mathcal{N})$ un espace mesurable et soit $f\colon \Omega \to Y$ une application. La famille

$$f^{-1}(\mathcal{N}) := \{f^{-1}(B) \mid B \in \mathcal{N}\} \subset \mathcal{P}(\Omega)$$

est une tribu sur Ω.

Exercice. Donner la preuve de ce lemme.
Si $\mathcal{F}$ est une tribu sur Ω, est-ce que $f(\mathcal{F}) := \{f(A) \mid A \in \mathcal{F}\}$ est une tribu sur Y ?

Classes monotones

Dans la suite, nous souhaiterons pouvoir mesurer tous les éléments d'une tribu, mais ceux-ci peuvent être très nombreux. On souhaite donc souvent pouvoir d'abord définir une mesure sur une classe plus restreinte d'ensemble et appliquer ensuite un procédé d'extension permettant de la définir sur la tribu toute entière. Dans ce paragraphe, on construit un tel outil.
Définition.

Une famille $\mathcal{M}$ de parties de Ω est appelée une **classe monotone** si

1. $\Omega \in \mathcal{M}$,
2. si $A, B \in \mathcal{M}$ avec $B \subset A$, alors $A \setminus B \in \mathcal{M}$,
3. si $\forall i \in \mathbb{N}$, $A_i \in \mathcal{M}$ et $A_i \subset A_{i+1}$, alors $\cup_{i \in \mathbb{N}} A_i \in \mathcal{M}$.

Si $\mathcal{E}$ est une famille de parties de Ω, l'intersection de toutes les classes monotones contenant $\mathcal{E}$ est encore une classe monotone, notée $\mathcal{M}(\mathcal{E})$ et appelée **classe monotone engendrée** par $\mathcal{E}$.

Exercice. Montrer les propriétés suivantes :

1. une intersection de classes monotones est encore une classe monotone,
2. une tribu est une classe monotone,
3. une classe monotone stable par intersection finie est une tribu.

On peut maintenant énoncer le théorème.

Théorème 5 *(Théorème de la classe monotone)*

Soit $\mathcal{E}$ est une famille de parties de Ω, stable par intersection finie. Alors $\mathcal{M}(\mathcal{E}) = \sigma(\mathcal{E})$.

Preuve. D'après l'exercice, $\sigma(\mathcal{E})$ est une classe monotone qui contient $\mathcal{E}$ donc $\mathcal{M}(\mathcal{E}) \subset \sigma(\mathcal{E})$. Montrons l'inclusion inverse. D'après l'exercice, il suffit de vérifier que $\mathcal{M}(\mathcal{E})$ est stable par intersection finie (car alors, $\mathcal{M}(\mathcal{E})$ sera une tribu contenant $\mathcal{E}$ et on aura $\sigma(\mathcal{E}) \subset \mathcal{M}(\mathcal{E})$.

Il suffit de prouver que si $A, B \in \mathcal{M}(\mathcal{E})$, $A \cap B \in \mathcal{M}(\mathcal{E})$. On pose

$$\mathcal{M}_1 := \{A \in \mathcal{M}(\mathcal{E}); \forall B \in \mathcal{E}, A \cap B \in \mathcal{M}(\mathcal{E})\}.$$

L'ensemble $\mathcal{M}_1$ est une classe monotone qui contient $\mathcal{E}$, et par conséquent $\mathcal{M}(\mathcal{E})$. On pose maintenant :

$$\mathcal{M}_2 := \{B \in \mathcal{M}(\mathcal{E}); \forall C \in \mathcal{E}, B \cap C \in \mathcal{M}(\mathcal{E})\}.$$

C'est une classe monotone. Montrons qu'elle contient $\mathcal{E}$. Soit $B \in \mathcal{E}$ et $C \in \mathcal{M}(\mathcal{E})$, alors $C \in \mathcal{M}_1$ et donc $B \cap C \in \mathcal{M}(\mathcal{E})$, ce qui siginifie que $B \in \mathcal{M}_2$. On a donc $\mathcal{E} \subset \mathcal{M}_2$, ce qui donne que $\mathcal{M}(\mathcal{E})$ est stable par intersection finie.

Fonctions mesurables, variables aléatoires

On aborde ici la notion de fonction mesurable.

Définition. (Fonction mesurable)

Une fonction f d'un espace mesurable $(\Omega, \mathcal{F})$ dans un espace mesurable $(Y, \mathcal{N})$ est dite **mesurable (par rapport aux tribus $\mathcal{F}$ et $\mathcal{N}$)** si la pré-image par f de tout sous-ensemble mesurable de Y est un sous-ensemble mesurable de Ω, autrement dit si $f^{-1}(\mathcal{N}) \subset \mathcal{F}$.

Remarque. Dans le cadre des probabilités, une fonction mesurable sera appelée **variable aléatoire**.

De la définition s'ensuit immédiatement

Proposition 6 1. *La composition de fonctions mesurables est une fonction mesurable. Plus précisément, si $f : (\Omega_1, \mathcal{F}) \to (\Omega_2, \mathcal{N})$ et $g : (\Omega_2, \mathcal{N}) \to (\Omega_3, \mathcal{R})$ alors $g \circ f$ est mesurable (par rapport aux tribus $\mathcal{F}$ et $\mathcal{R}$).*

2. *Soient Ω_1 et Ω_2 deux espaces topologiques, munis de leurs tribus boréliennes respectivse. Toute fonction **continue** de Ω_1 dans Ω_2 est mesurable.*
3. *Si f et g sont deux fonctions mesurables de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, alors la fonction $\omega \in \Omega \mapsto (f(\omega), g(\omega))$ est mesurable de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.*
4. *L'ensemble des fonctions mesurables de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est stable par addition, multiplication, maximum, multiplication par une constante.*

Preuve. Exercice.

La définition de mesurabilité d'une fonction demande de vérifier que pour tout $E \in \mathcal{N}$ on a $f^{-1}(E) \in \mathcal{F}$. En réalité il suffit de vérifier cette condition pour des ensembles E appartenant à une classe beaucoup plus petite que $\mathcal{N}$.

Proposition 7 *(Critère de mesurabilité).*

Soient $(\Omega, \mathcal{F})$ et $(Y, \mathcal{N})$ deux espaces mesurables et soit $\mathcal{S}$ une famille de parties de Y qui engendre la tribu $\mathcal{N}$. Une application $f : \Omega \to Y$ est mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{N}$ si et seulement si pour tout $S \in \mathcal{S}$ on a $f^{-1}(S) \in \mathcal{F}$.

Preuve. La nécessité est évidente. Montrons la suffisance. La famille d'ensembles $\mathcal{N}_1 = \{E \in \mathcal{P}(Y) \mid f^{-1}(E) \in \mathcal{F}\}$ est une tribu sur Y. Par hypothèse, la tribu $\mathcal{N}_1$ contient la famille $\mathcal{S}$. Il s'ensuit que la tribu engendrée par $\mathcal{S}$ est aussi contenue dans $\mathcal{N}_1$. Donc $\mathcal{N} \subset \{E \in \mathcal{P}(Y) \mid f^{-1}(E) \in \mathcal{F}\}$. Mais cela signifie que pour tout $E \in \mathcal{N}$ on a $f^{-1}(E) \in \mathcal{F}$.

Deux questions se posent naturellement :

Si $(\Omega, \mathcal{F})$ est un espace mesurable et f une fonction de Ω dans Y, existe-t-il une tribu $\mathcal{N}$ sur Y telle que f soit mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{N}$?

Ainsi formulée la question est banale car si on munit Y de la tribu triviale $\{\emptyset, Y\}$, toute fonction $\Omega \to Y$ est mesurable. Donc la question est plutôt, quelle est la plus grande tribu sur Y par rapport à laquelle la fonction $f : \Omega \to Y$ est mesurable ? Le lecteur vérifiera sans peine que la famille

$$\mathcal{N}_f = \{A \subset Y \mid f^{-1}(A) \in \mathcal{F}\}$$

est une tribu sur Y. Par sa définition $\mathcal{N}_f$ est l'unique plus grande tribu sur Y par rapport à laquelle f est mesurable (à vérifier).

La question duale est la suivante : si $(f_\alpha)_{\alpha \in I}$ est une famille de fonctions d'un ensemble Ω dans des espaces mesurables $(Y_\alpha, \mathcal{N}_\alpha)$, existe-t-il une tribu $\mathcal{F}$ sur Ω qui rend toutes ces fonctions mesurables ?

Encore une fois, la formulation de cette question est banale car toute fonction de Ω dans un espace mesurable est mesurable si on munit Ω de la tribu totale $\mathcal{P}(\Omega)$. La véritable question est donc : quelle est la plus petite tribu sur Ω par rapport à laquelle toutes les fonctions $f_\alpha : \Omega \to Y_\alpha$ sont mesurables. Cette tribu devra contenir toutes les tribus $f_\alpha^{-1}(\mathcal{N}_\alpha)$. Par conséquent elle est donné par la plus petite tribu contenant toutes les tribus $f_\alpha^{-1}(\mathcal{N}_\alpha)$. En rappelant qu'on note $\bigvee_{\alpha \in I} \mathcal{F}_\alpha$ la tribu engendrée par la réunion des tribus $\mathcal{F}_\alpha$, ($\alpha \in I$), on est arrivé à la conclusion que la plus petite tribu sur Ω par rapport à laquelle toutes les fonctions $f_\alpha : \Omega \to Y_\alpha$ sont mesurables est donnée par

$$\mathcal{F} = \bigvee_{\alpha \in I} f_\alpha^{-1}(\mathcal{N}_\alpha).$$

Ces observations ont des conséquences sur la mesurabilité des fonctions définies ou à valeurs dans un produit d'espaces mesurables.

Proposition 8 *(Retour sur la tribu produit).*

> *Soient $(\Omega_1, \mathcal{F}_1)$ et $(\Omega_2, \mathcal{F}_2)$ deux espaces mesurables. La tribu produit $\mathcal{F}_1 \otimes \mathcal{F}_2$ est la plus petite tribu sur $\Omega_1 \times \Omega_2$ telle que les projections $\pi_i \colon \Omega_1 \times \Omega_2 \to \Omega_i$ soient mesurables.*

Preuve. En effet pour tout $A \subset \Omega_1$ on a $A \times \Omega_2 = \pi_1^{-1}(A)$ et pour tout $B \subset \Omega_2$ on a $\Omega_1 \times B = \pi_2^{-1}(B)$. Donc la tribu engendrée par la famille $\{A \times \Omega_2 \in \mathcal{P}(\Omega_1 \times \Omega_2) \mid A \in \mathcal{F}_1\} \bigcup \{\Omega_1 \times B \in \mathcal{P}(\Omega_1 \times \Omega_2) \mid B \in \mathcal{F}_2\}$ est exactement la tribu engendrée par la réunion $\pi_1^{-1}(\mathcal{F}_1) \cup \pi_2^{-1}(\mathcal{F}_2)$, tribu qu'on a noté $\pi_1^{-1}(\mathcal{F}_1) \vee \pi_2^{-1}(\mathcal{F}_2)$. Par la discussion précédente, cette tribu est la plus petite tribu sur $\Omega_1 \times \Omega_2$ telle que les projections $\pi_i \colon \Omega_1 \times \Omega_2 \to \Omega_i$ soient mesurables.

Sur la même idée, on peut définir la tribu produit pour un produit quelconque d'ensembles mesurables.

Définition. (Tribu produit, cas général)

> Soit $(\Omega_\alpha, \mathcal{F}_\alpha)_{\alpha \in I}$ une famille d'espaces mesurables. La **tribu produit** $\bigotimes_{\alpha \in I} \mathcal{F}_\alpha$ est la plus petite tribu sur $\Omega = \prod_{\alpha \in I} \Omega_\alpha$ telle que toutes les projections $\pi_\alpha \colon \Omega \to \Omega_\alpha$ soient mesurables. C'est donc la tribu
> $$\bigvee_{\alpha \in I} \pi_\alpha^{-1}(\mathcal{F}_\alpha)$$
> engendrée par la famille de parties $\bigcup_{\alpha \in I} \ \pi_\alpha^{-1}(\mathcal{F}_\alpha)$.

Notation. Lorsque tous les espaces $(\Omega_\alpha, \mathcal{F}_\alpha)_{\alpha \in I}$ coïncident avec un même espace $(\Omega, \mathcal{F})$ la tribu $\bigotimes_{\alpha \in I} \mathcal{F}_\alpha$ est noté $\mathcal{F}^{\otimes I}$.

Exemple. L'espace $\{0,1\}^{\mathbb{N}}$, c.-à-d. l'espace des suites infinies de 0 ou 1, est important en probabilité car on peut considérer un point de cet espace comme une suite infinie de tirages de pile ou face ("0"= pile, "1"= face). En effet, si $\omega = (\omega_n)_{n \in \mathbb{N}} \in \{0,1\}^{\mathbb{N}}$, alors au temps n on a tiré pile si $\omega_n = 0$, face si $\omega_n = 1$.

Il est clair qu'on veut pouvoir assigner une probabilité à l'événement ≪ au temps j on a tiré pile (ou face) ≫, c.-à-d. aux ensembles

$$E_j^i = \{(\omega_n)_{n \in \mathbb{N}} \in \{0,1\}^{\mathbb{N}} \mid \omega_j = i\}, \quad i \in \{0,1\}.$$

Or, si on note π_j la projection $p_j \colon (\omega_n)_{n \in \mathbb{N}} \in \{0,1\}^{\mathbb{N}} \mapsto \omega_j \in \{0,1\}$, on a $E_j^i = \pi_j^{-1}(\{i\})$. On veut donc que tous les ensembles de la tribu $\pi_j^{-1}(\mathcal{P}(\{0,1\})$ soient mesurables. La conclusion de ce raisonnement est que la plus petite tribu sur $\{0,1\}^{\mathbb{N}}$ telle que tous les événements E_j^i soient mesurables est la tribu $\bigvee_{\alpha \in I} \ \pi_j^{-1}(\mathcal{P}(\{0,1\})$, c.-à-d. la tribu produit $\big(\mathcal{P}(\{0,1\})\big)^{\otimes \mathbb{N}}$.

Corollaire

> Soit $(Y_\alpha, \mathcal{N}_\alpha)_{\alpha \in I}$ une famille d'espaces mesurables. On pose $Y = \prod_{\alpha \in I} Y_\alpha$ et on note $\pi_\alpha \colon Y \to Y_\alpha$ les projections canoniques. Une fonction $f \colon \Omega \mapsto Y$ d'un espace mesurable $(\Omega, \mathcal{F})$ dans $(Y, \bigotimes_{\alpha \in I} \mathcal{F}_\alpha)$ est mesurable si et seulement si toutes ses composantes $f_\alpha = \pi_\alpha \circ f \colon \Omega \to Y_\alpha$ sont mesurables.

Preuve. Exercice.

Chapitre 2

Mesures

Nous allons maintenant définir la notion de mesure. Ce sont des fonctions positives définies sur une algèbre d'ensembles ou sur une tribu. Notons qu'il est parfois utile d'attribuer à certains ensembles une mesure infinie.

Notion générale de mesure

Définition. Une **mesure finiment additive** sur une algèbre $\mathcal{A}$ est une application $\nu\colon \mathcal{A} \to [0,\infty]$ telle que si A et B sont des éléments disjoints de $\mathcal{A}$, on a

$$\nu\big(A \cup B\big) = \nu(A) + \nu(B)$$

De la définition découle immédiatement que, si ν est une mesure finiment additive sur une algèbre $\mathcal{A}$ et A_1, A_2, ..., A_k sont des éléments deux à deux disjoints de $\mathcal{A}$, alors on a l'égalité suivante

$$\nu\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \nu(A_i)$$

qui explique le terme **finiment** additive attribué à la mesure ν.

Les mesures sur les algèbres booléenne sont faciles à définir. Par exemple on verra plus loin que la famille des réunions finies d'intervalles disjoints de $\mathbb{R}$ forme une algèbre, et que l'application qui assigne à une réunion finie d'intervalles disjoints la somme de leurs longueurs est une mesure finiment additive sur cette algèbre.

Toutefois, pour les passages aux limites nécessaires aux développements de l'analyse et pour l'interprétation courante des probabilités, on ne peut pas se limiter à considérer des réunions finies. Pour cette raison nous nous limiterons à considérer des mesures définies sur des tribus. La définition de mesure sur une tribu est donnée au paragraphe suivant ; dans le reste de ce chapitre nous développerons leurs propriétés.

Pour l'instant nous nous limitons à remarquer que, comme les éléments d'une tribu sont, en général, difficiles à décrire de manière constructive, le problème de définir une mesure sur une tribu, en partant d'une mesure finiment additive facilement définie sur une algèbre, est un problème qui se pose de façon naturelle. Ce problème sera au centre de notre attention dans la suite.

Mesures sur les tribus, espaces mesurés

Définition. (Mesure)

> Une **mesure (positive)** sur un espace mesurable $(\Omega, \mathcal{F})$ est une application $\mu\colon \mathcal{F} \to [0,+\infty]$ satisfaisant :
>
> 1. $\mu(\emptyset) = 0$.

2. μ est **dénombrablement additive (ou σ-additive)** : cela signifie que si $(E_k)_{k\in\mathbb{N}}$ est une famille dénombrable de parties mesurables **deux à deux disjointes** alors

$$\mu\left(\bigcup_{k\in\mathbb{N}} E_k\right) = \sum_{k\in\mathbb{N}} \mu(E_k)$$

Une mesure μ est dite :

— une **(mesure de) probabilité** si $\mu(\Omega) = 1$;

— **finie** si $\mu(\Omega) < \infty$;

— **σ-finie** si l'espace Ω est une réunion dénombrable de parties mesurables de mesure finie ;

Un **espace mesuré** est un triplet $(\Omega, \mathcal{F}, \mu)$ composé d'un espace mesurable $(\Omega, \mathcal{F})$ et d'une mesure μ sur $(\Omega, \mathcal{F})$. Si μ est une mesure de probabilité l'espace mesuré $(\Omega, \mathcal{F}, \mu)$ est appelé **espace de probabilité** ou **triplet probabiliste**.

Remarque. Un espace de probabilité est le plus souvent noté $(\Omega, \mathcal{F}, \mathbb{P})$, où $\mathcal{F}$ est une tribu et, comme on l'a déjà dit, les parties $A \in \mathcal{F}$ sont **événements**. Le nombre $\mathbb{P}(A)$ est la **probabilité de l'événement** A (on considère que ce nombre représente la probabilité que l'événement A se réalise dans un tirage au sort selon la loi $\mathbb{P}$)

Exemple. Soit Ω un ensemble et soit $x \in \Omega$. La mesure définie sur $\mathcal{P}(\Omega)$ par

$$\delta_x(E) = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{sinon} \end{cases}$$

est dite la **masse de Dirac en** x. Il s'agit, bien évidemment, d'une probabilité.

Exemple. Soit Ω un ensemble. La mesure μ définie sur $\mathcal{P}(\Omega)$ par

$$\mu(E) = \begin{cases} n & \text{si } E \text{ est un sous-ensemble fini avec } n \text{ éléments} \\ \infty & \text{sinon} \end{cases} \tag{2.1}$$

est dite la **mesure de comptage** sur Ω. Elle est définie également par la condition $\mu(\{x\}) = 1$, pour tout $x \in \Omega$. En formules on peut écrire aussi

$$\mu(E) = \sum_{x\in E} 1 = \sum_{x\in\Omega} \mathbf{1}_E(\{x\})$$

La mesure de comptage est

— finie si (et seulement si) Ω est fini. Dans ces cas, $\nu = \frac{1}{card(\Omega)}\mu$ est une mesure de probabilité sur Ω.

— σ-finie si (et seulement si) Ω est dénombrable, car dans ce cas Ω est la réunion dénombrable de singletons.

Les probabilités élémentaires nous fournissent de nombreux exemples de mesures. Dans les exemples suivants la tribu sur laquelle la mesure est définie est la tribu discrète $\mathcal{P}(\Omega)$.

Exemple.

1. Si Ω est un ensemble fini la mesure définie par $\mu = \frac{1}{card(\Omega)} \sum_{x\in\Omega} \delta_x$, c'est-à-dire définie par $\mu(A) = \frac{card(A)}{card(\Omega)}$ pour tout $A \subset \Omega$, est appelée la **distribution uniforme sur** Ω.
2. Soit $\Omega = \{0,1\}^n$ et pour $\omega = (\omega_i)_{i=1}^n \in \Omega$ posons $S(\omega) = \sum_{i=1}^n \omega_i$. Clairement $S(\omega)$ est le nombre d'indices $i \in \{1, \ldots, n\}$ tels que $\omega_i = 1$. On peut considérer $S(\omega)$ comme le nombre de **succès** pour la sortie $\omega \in \Omega$. On pose alors $\mu(\{\omega\}) = p^{S(\omega)}(1-p)^{k-S(\omega)}$. Pour tout sous-ensemble $A \subset \Omega$ on pose $\mu(A) = \sum_{\omega\in A} \mu(\{\omega\})$.

3. Soit $n \in \mathbb{N}^*$ et $p \in [0,1]$. La mesure sur $\{0,1,\dots,n\}$ définie par $\mu = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k$ c'est-à-dire $\mu(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}$ est dite la **distribution binomiale de paramètres** n **et** p.
4. Soit $\lambda > 0$. Soit μ la mesure sur $\mathbb{N}$ définie par $\mu = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \delta_k$, c'est-à-dire $\mu(\{k\}) = e^{\lambda} \frac{\lambda^k}{k!}$. Cette mesure est dite la **distribution de Poisson de paramètre** λ.

Proposition 9 *(Propriétés générales d'une mesure).*

Soit $(\Omega, \mathcal{F}, \mu)$ un espace mesuré. Les propriétés suivantes sont alors satisfaites :

1. *(Monotonie) Si A et B sont mesurables et $A \subset B$ alors $\mu(A) \leq \mu(B)$.*
2. *(Sous-additivité) si $(E_k)_{k \in I}$ est une famille dénombrable de parties mesurables alors*
$$\mu\Big(\bigcup_{k\in I} E_k\Big) \leq \sum_{k \in I} \mu(E_k).$$
3. *(Continuité par rapport aux réunions croissantes). Si $(E_k)_{k\in\mathbb{N}}$ est une suite dénombrable croissante de parties mesurables, c-à-d. si $E_0 \subseteq E_1 \subseteq \cdots \subseteq E_k \subseteq \cdots$, alors*
$$\mu\Big(\bigcup_{k\in\mathbb{N}} E_k\Big) = \lim_{k\to\infty} \mu(E_k).$$
4. *(Continuité par rapport aux intersections décroissantes). Si $(E_k)_{k\in\mathbb{N}}$ est une suite dénombrable décroissante de parties mesurables, c-à-d. si $E_0 \supseteq E_1 \supseteq \cdots \supseteq E_k \supseteq \cdots$, et si en plus **un des** E_i **a mesure finie** alors*
$$\mu\Big(\bigcap_{k\in\mathbb{N}} E_k\Big) = \lim_{k\to\infty} \mu(E_k).$$

Preuve. Les preuves des affirmations (1) et (2) sont immédiates et sont laissées en exercice au lecteur.

Preuve de (3). Comme $E_0 \subseteq E_1 \subseteq \cdots \subseteq E_k \subseteq \cdots$, les ensembles $E_j \setminus E_{j-1}$ sont disjoints deux à deux et $E_k = \bigcup_{j=0}^{k} (E_j \setminus E_{j-1})$ (on a posé $E_{-1} = \emptyset$). Il s'ensuit que $\bigcup_{k=0}^{\infty} E_k = \bigcup_{k=0}^{\infty} (E_k \setminus E_{k-1})$. Puisque les ensembles $E_k \setminus E_{k-1}$ sont disjoints deux à deux, par additivité dénombrable, on obtient

$$\begin{aligned}\mu\Big(\bigcup_{k=0}^{\infty} E_k\Big) &= \mu\Big(\bigcup_{k=0}^{\infty} E_k \setminus E_{k-1}\Big) = \sum_{k=0}^{\infty} \mu(E_k \setminus E_{k-1}) \\ &= \lim_{N\to\infty} \sum_{k=0}^{N} \mu(E_k \setminus E_{k-1}) = \lim_{N\to\infty} \mu\Big(\bigcup_{k=0}^{N} E_k \setminus E_{k-1}\Big) = \lim_{N\to\infty} \mu(E_N).\end{aligned}$$

Preuve de (4). Soit $\mu(E_i) < \infty$. Comme $\bigcap_{k\in\mathbb{N}} E_k = \bigcap_{k\geq i} E_k$ on peut se réduire au cas $\mu(E_0) < \infty$, en remplaçant la suite (E_k) par la suite (E_{k+i}). Soit donc $\mu(E_0) < \infty$.

Les inclusions $E_0 \supseteq E_1 \supseteq \cdots \supseteq E_k \supseteq \cdots$ impliquent les inclusions inverses $(E_0 \setminus E_1) \subseteq (E_0 \setminus E_2) \subseteq \cdots \subseteq (E_0 \setminus E_k)$. En appliquant le résultat (3) on obtient

$$\mu\Big(E_0 \setminus \bigcap_{k=0}^{\infty} E_k\Big) = \mu\Big(\bigcup_{k=0}^{\infty} (E_0 \setminus E_k)\Big) = \lim_{N\to\infty} \mu(E_0 \setminus E_N)$$

Comme $\mu(E_0) < \infty$ on a $\mu(E_0 \setminus \bigcap_{k=0}^{\infty} E_k) = \mu(E_0) - \mu(\bigcap_{k=0}^{\infty} E_k)$ et $\mu(E_0 \setminus E_N) = \mu(E_0) - \mu(E_N)$ car $\bigcap_{k=0}^{\infty} E_k \subseteq E_0$ et $E_N \subseteq E_0$. On obtient

$$\mu(E_0) - \mu\Big(\bigcap_{k=0}^{\infty} E_k\Big) = \lim_{N\to\infty} \big(\mu(E_0) - \mu(E_N)\big) = \mu(E_0) - \lim_{N\to\infty} \mu(E_N),$$

ce qui conlut la preuve.

Ensembles de mesure nulle

Les ensembles de mesure nulle dans un espace mesuré $(\Omega, \mathcal{F}, \mu)$ jouent un rôle tout à fait particulier dans la théorie de la mesure. D'abord on a

Proposition 10 *Soit $(\Omega, \mathcal{F}, \mu)$ un espace mesuré. Si $(N_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$ est une suite d'ensembles de mesure nulle, alors la réunion $\bigcup_n N_n$ est encore un ensemble (mesurable) de mesure nulle.*

Si $N \in \mathcal{F}$ est un ensemble de mesure nulle, alors, pour tout $E \in \mathcal{F}$, l'ensemble $E \cap N$ est encore un ensemble (mesurable) de mesure nulle.

Les propriétés ci-dessus ont une ressemblance avec la notion d'idéal en algèbre (la somme étant représentée par la réunion dénombrable et le produit par l'intersection) Pour cette raison on dit que la famille $\{N \in \mathcal{F} \mid \mu(N) = 0\}$ est un **σ-idéal** de la tribu $\mathcal{F}$.

Définition. (Ensemble négligeable)

On dit qu'un ensemble A dans un espace mesuré est **négligeable (par rapport à μ) ou μ-négligeable** si A est contenu dans un ensemble **mesurable** de mesure nulle.

Si une propriété est vérifiée pout tout $\omega \in A^c$, avec A un ensemble μ-négligeable, on dit que cette propriété est vérifiée **μ-presque-partout** (p.p.).

De la proposition précédente on a immédiatement qu'une réunion dénombrable d'ensembles négligeables est négligeable.

Définition. Un espace mesuré $(\Omega, \mathcal{F}, \mu)$ est dit **complet** si toute partie négligeable par rapport à μ est mesurable : si $E \in \mathcal{F}$ et $\mu(E) = 0$ alors pour tout $F \subseteq E$ on a $F \in \mathcal{F}$.

Théorème 11 *Soit $(\Omega, \mathcal{F}, \mu)$ un espace mesuré. On pose*

$$\mathcal{F}' = \{A \cup B \mid A \in \mathcal{F},\ B \ \mu\text{-négligeable}\}.$$

Si $C \in \mathcal{F}'$ et $C = A \cup B \in \mathcal{F}'$ avec $A \in \mathcal{F}$, et B μ-négligeable on pose aussi

$$\mu'(C) = \mu(A).$$

Le triplet $(\Omega, \mathcal{F}', \mu')$ ainsi obtenu est un espace mesuré complet satisfaisant

$$\mathcal{F} \subset \mathcal{F}' \quad \text{et} \quad \mu'_{|\mathcal{F}} = \mu.$$

*L'espace $(\Omega, \mathcal{F}', \mu')$ est dit l'**extension complète de** $(\Omega, \mathcal{F}, \mu)$.*

Preuve. Il est clair de la définition que $\mathcal{F} \subset \mathcal{F}'$ car l'ensemble vide est négligeable. Donc $\Omega \in \mathcal{F}'$. Si $(C_i)_{i\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{F}'$, pour tout $i \in \mathbb{N}$, on a $C_i = A_i \cup B_i$ avec avec $A_i \in \mathcal{F}$, et B_i μ-négligeable. Comme la réunion des A_i est mesurable et la réunion des b_i est négligeable, la réunion $\bigcup_{i\in\mathbb{N}} C_i = \bigcup_{i\in\mathbb{N}} A_i \cup \bigcup_{i\in\mathbb{N}} B_i$ appartient à $\mathcal{F}'$. Finalement si $C = A \cup B$ avec $A \in \mathcal{F}$, et B μ-négligeable,il existe $N \in \mathcal{F}$ tel que $\mu(N) = 0$ et $B \subset N$. Puisque

$$(A \cup B)^c = (A \cup N)^c \cup (N \setminus (A \cup B))$$

on a $(A \cup B)^c \in \mathcal{F}'$ car $(A \cup N)^c$ est mesurable et $(N \setminus (A \cup B) \subset N$ est négligeable.

En conclusion, $\mathcal{F}'$ est une tribu.

La fonction μ est bien définie sur $\mathcal{F}'$: supposons que $C = A \cup B = A' \cup B'$ avec $A, A' \in \mathcal{F}$, et B, B' négligeables. On a alors $A \subset C = A' \cup B' \subset A' \cup N'$, où N' est un ensemble mesurable de mesure nulle. Donc $\mu(A) \leq \mu(A' \cup N') \leq \mu(A') + \mu(N') = \mu(A')$. Par le même raisonnement, $\mu(A') \leq \mu(A)$ et donc $\mu(A') = \mu(A)$ On conclut que la définition de $\mu'(C)$ ne dépend pas de la décomposition particulière comme réunion d'un ensemble mesurable et un ensemble négligeable. Il est clair que si $A \in \mathcal{F}$, $\mu'(A) = \mu(A)$.

Finalement si $C \in \mathcal{F}'$ satisfait $\mu'(C) = 0$ et $D \subset C$ on a $C = A \cup B \subset\subset A \cup N$ avec A, N mesurable et satisfaisant $\mu(A) = \mu(N) = 0$. Donc D et C sont μ-négligeable, en particulier $D \in \mathcal{F}'$. Cela montre que $(\Omega, \mathcal{F}', \mu')$ est un espace mesuré complet. On a aussi montré, en passant, que tout ensemble de mesure μ' nulle est μ-négligeable. La réciproque est immédiate de la définition de μ'.

Remarque. Un espace mesuré $(\Omega, \mathcal{F}_2, \mu_2)$ est une extension (ou un prolongement) d'un espace mesuré $(\Omega, \mathcal{F}_1, \mu_1)$ si $\mathcal{F}_1 \subset \mathcal{F}_2$ et $\mu_1 = \mu_2$ sur $\mathcal{M}_1$.

Si $(\Omega, \mathcal{F}'', \mu'')$ est un autre extension complète de $(\Omega, \mathcal{F}, \mu)$, alors tout ensemble μ-négligeable appartient à $\mathcal{F}''$ (en vertu de la complétude de $(\Omega, \mathcal{F}'', \mu'')$) et sa mesure μ'' est nulle. Donc $\mathcal{F}' \subset \mathcal{F}''$. De plus si $C = A \cup B \in \mathcal{F}'$, avec $A \in \mathcal{F}$ et B μ-négligeable on a $\mu'(C) = \mu(A)$ et $\mu(A) = mu''(A) \leq \mu''(C) \leq \mu''(A) + \mu''(B) = \mu(A)$, d'où $\mu'(C) = \mu''(C)$. En conclusion le complété $(\Omega, \mathcal{F}', \mu')$ est la **plus petite** extension complète de $(\Omega, \mathcal{F}, \mu)$, dans le sens que toute autre extension complète $(\Omega, \mathcal{F}'', \mu'')$ de $(\Omega, \mathcal{F}, \mu)$ est une extension de $(\Omega, \mathcal{F}', \mu')$.

Chapitre 3

Mesures extérieures, mesures de Lebesgue-Stieltjes

Dans le chapitre précédent nous avons étudié les propriétés générales des mesures définies sur des tribus. Toutefois les exemples d'espaces mesurés que nous avons jusqu'ici sont assez banales et le doute pourrait surgir que, contrairement aux mesures finiment additives sur les algèbres d'ensembles, les espaces mesurés sont plutôt rares ou peu intéressants. Pour dissiper ce doute dans ce cours nous penchons sur une méthode générale qui permet de définir des véritables mesures, c.-à-d. des mesures (dénombrablement additives) sur des tribus.

L'exemple principal sera la construction de la mesure de Lebesgue sur $\Omega = [0,1]$. Voilà les grandes lignes de cette construction.

On va partir d'une « pré-mesure » qui, pour un intervalle $I \subseteq [0,1]$ n'est rien d'autre que sa longueur. Pour estimer la mesure d'un ensemble quelconque $A \subseteq [0,1]$ on pourrait le recouvrir d'intervalles et sommer les pré-mesures des intervalles de ce recouvrement. On obtient ainsi un nombre qui dépend du recouvrement choisi. Il y a maintes façons de recouvrir A par des intervalles et il est donc raisonnable de choisir des recouvrements de plus en plus efficaces pour obtenir la plus petite estimation[1] de la "veritable" mesure de A. De cette façon on obtient un nombre qui représente la meilleure estimation de la mesure de A, estimation **faite de l'extérieur** car on a considéré des **recouvrements** de l'ensemble A. Ce nombre s'appelle la **mesure extérieure** de A et on le notera $\lambda^*(A)$; il est défini pour tout $A \subseteq [0,1]$ et il n'est pas difficile de voir qu'il satisfait une propriété de sous-additivité : $\lambda^*(A \cup B) \leq \lambda^*(A) + \lambda^*(B)$ et en particulier

$$\lambda^*(A) + \lambda^*(A^c) \geq \lambda^*([0,1]) = 1.$$

Toutefois si A est mesurable et $\lambda^*(A)$ est sa veritable mesure, on s'attend que A^c soit aussi mesurable et que $\lambda^*(A^c)$ soit la mesure de A^c ; de cela découle l'égalité $\lambda^*(A) + \lambda^*(A^c) = 1$ car une mesure est additive.

On est donc amené à proposer la définition d'ensemble mesurable suivante : A est mesurable si

$$\lambda^*(A) + \lambda^*(A^c) \geq \lambda^*([0,1]) = 1.$$

Dans ce cours, après avoir introduit la notion abstraite de mesure extérieure, définie par des propriétés inspirées de la discussion prédédente, et après avoir donné des méthodes pour construire des mesures extérieures, nous nous penchons sur le Théorème de Carathéodory, qui nous permet d'obtenir un espace mesuré à partir d'une mesure extérieure.

Comme application de ce théorème nous donnerons une description complète de toutes les mesures sur l'espace mesurable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ qui sont finies sur les parties compactes. Un exemple fondamental sera la mesure de Lebesgue sur la droite réelle.

1. Formellement cela signifie qu'on considère la borne inférieure des toutes les estimations obtenues par des recouvrements de A

Mesures extérieures

Au vu de la discussion ci-dessus, on introduit la définition suivante :
Définition. (Mesure extérieure)

> Une **mesure extérieure** sur un ensemble Ω est une application
>
> $$\mu^* : \mathcal{P}(\Omega) \to [0, +\infty]$$
>
> satisfaisant le propriétés suivantes :
>
> 1. $\mu^*(\emptyset) = 0$;
> 2. μ^* est **monotone** : si $A \subset B \subset \Omega$ alors $\mu^*(A) \leq \mu^*(B)$.
> 3. μ^* est **dénombrablement sous-additive** : pour toute collection dénombrable $\{A_j\}_{j\in\mathbb{N}}$ de parties de Ω on a $\mu^*(\bigcup_{j\in\mathbb{N}} A_j) \leq \sum_{j\in\mathbb{N}} \mu^*(A_j)$.

Remarquons qu'une mesure extérieure **n'est pas une mesure** et qu'elle est définie pour tout sous-ensemble de Ω.

Exercice. Montrer que les propriétés (2) et (3) de la définition de mesure extérieure sont équivalentes à la propriété suivante :

2′. Si $A \subset \bigcup_{j\in\mathbb{N}} A_j$ alors $\mu^*(A) \leq \sum_{j\in\mathbb{N}} \mu^*(A_j)$.

Mesures extérieures définies par une pré-mesure

Il est très facile de définir des mesures extérieures en utilisant des recouvrements et une fonction positive définie sur une famille arbitraire de parties de Ω.

Définition. Une **pré-mesure sur** Ω est un triplet $(\Omega, \mathcal{C}, \rho)$ avec $\mathcal{C}$ une collection de parties de Ω telle que $\emptyset \in \mathcal{C}$ et $\mu\colon \mathcal{C} \to [0, \infty]$ une fonction telle que $\mu(\emptyset) = 0$.

Exemple.

1. Soit $\mathcal{C}$ l'ensemble ses intervalles ouverts de $\mathbb{R}$ et $\rho(]a, b[) = b - a$.
2. Soit $\mathcal{C}$ l'ensemble des parties compactes de $\mathbb{R}^n$ et $\rho(K) = \mathrm{diam}(K)^\alpha$ ($\alpha \in]0, n]$).
3. Soit $\mathcal{C}$ l'ensemble des boules ouvertes de $\mathbb{R}^n$ et $\rho(B) = \mathrm{diam}(B)^n$.
4. Soit $\mathcal{C}$ l'ensemble formé par le vide et les singletons d'un ensemble Ω. On pose $\rho(\{x\}) = 1$, pour tout $x \in \Omega$.

Les exemples précédents montrent qu'il y a une très grande variété de pré-mesures, car la définition de pré-mesure pose très peu de limitations.

> **Proposition 12** *Soit $(\Omega, \mathcal{C}, \rho)$ une pré-mesure. On pose pour tout $E \subset \Omega$*
>
> $$\mu^*(E) = \inf \left\{ \sum_{j\in\mathbb{N}} \rho(A_j) \;\middle|\; (A_j)_{j\in\mathbb{N}} \in \mathcal{C}^{\mathbb{N}},\ E \subset \bigcup_{j\in\mathbb{N}} A_j \right\}.$$
>
> *Alors μ^* est une mesure extérieure sur Ω dite la* ***mesure extérieure induite par la pré-mesure*** $(\Omega, \mathcal{C}, \rho)$.

Preuve. Vérifions que les trois propriétés d'une mesure extérieure sont satisfaites :

1. Comme $\emptyset \in \mathcal{C}$ et $\rho(\emptyset) = 0$ on a $\mu^*(\emptyset) = 0$.

2. Si $A \subset B$ tout recouvrement de B par une famille dénombrable d'éléments de $\mathcal{C}$ est aussi un recouvrement de A. Donc $\mu^*(A) \leq \mu^*(B)$.
3. Soit $(A_j)_{j\in\mathbb{N}}$ est une famille dénombrable de parties de Ω.
   On pourra supposer que $\mu^*(A_j) < \infty$ pour tout $j \in \mathbb{N}$ car, sinon, l'inégalité $\mu^*(\bigcup_{j\in\mathbb{N}} A_j) \leq \sum_{j\in\mathbb{N}} \mu^*(A_j)$ est trivialement vraie.
   Il s'ensuit que pour tout $\epsilon > 0$ et pour tout $j \in \mathbb{N}$ il existe une famille dénombrable $(C_{jk})_{k\in\mathbb{N}} \in \mathcal{C}^{\mathbb{N}}$ telle que
   $$A_j \subset \bigcup_{k\in\mathbb{N}} C_{jk} \qquad \text{et} \qquad \mu^*(A_j) + \epsilon 2^{-j} > \sum_{k\in\mathbb{N}} \mu^*(C_{jk}).$$
   On a alors $\bigcup_{j\in\mathbb{N}} A_j \subset \bigcup_{j\in\mathbb{N}} \bigcup_{k\in\mathbb{N}} C_{jk}$ et par conséquent
   $$\mu^*(\bigcup_{j\in\mathbb{N}} A_j) \leq \sum_{j\in\mathbb{N}} \sum_{k\in\mathbb{N}} \rho(C_{jk}) < \sum_{j\in\mathbb{N}} (\mu^*(A_j) + \epsilon 2^{-j}) = \sum_{j\in\mathbb{N}} \mu^*(A_j) + 2\epsilon.$$
   Comme $\epsilon > 0$ est arbitraire on conclut que $\mu^*(\bigcup_{j\in\mathbb{N}} A_j) \leq \sum_{j\in\mathbb{N}} \mu^*(A_j)$.

Le théorème de Carathéodory

Définition.

> Soit μ^* est une mesure extérieure sur Ω. On dit que $A \subset \Omega$ est μ^***-mesurable** si pour tout $E \subset \Omega$ on a
> $$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \tag{3.1}$$

Remarque. Si μ^* est une mesure extérieure sur Ω, lorsque on cherche à démontrer l'égalité (3.1), on peut se limiter à démontrer l'inégalité en apparence plus faible
$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c) \tag{3.2}$$
En effet l'inégalité opposée $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c)$ découle de la définition de mesure extérieure et elle est donc toujours vraie.

L'inégalité (3.2) est trivialement vérifié si $\mu^*(E) = +\infty$. Pour cette raison dans la définition (3.1) de μ^*-mesurabilité on peux se limiter à considérer les ensemble E avec $\mu^*(E) < \infty$

Théorème 13 *(Carathéodory)*

> *Si μ^* est une mesure extérieure sur Ω, la famille d'ensembles μ^*-mesurable est une tribu Σ_{μ^*}. La restriction $\bar{\mu}$ de μ^* à la tribu Σ_{μ^*} est une mesure complète sur (Ω, Σ_{μ^*}), dite **mesure de Carathéodory associée à** μ^*.*

Preuve. Notons Σ_{μ^*} la famille des ensembles μ^*-mesurables et vérifions les propriétés d'une tribu. Évidemment $\Omega \in \Sigma_{\mu^*}$ car
$$\mu^*(E \cap \Omega) + \mu^*(E \cap \Omega^c) = \mu^*(E) + \mu(\emptyset) = \mu^*(E).$$
Puisque la définition de μ^*-mesurabilité de $A \subset \Omega$ est symétrique en A et A^c, si $A \in \Sigma_{\mu^*}$ alors $A^c \in \Sigma_{\mu^*}$.

Il reste à démontrer qu'une réunion dénombrable de parties μ^*-mesurables est μ^*-mesurable. Montrons, d'abord qu'une réunion **finie** de parties μ^*-mesurables est μ^*-mesurable. Or, si A et B sont μ^*-mesurables et $E \subset \Omega$ on a
$$\begin{aligned}
\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\
&= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c) \\
&\geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c) \geq \mu^*(E).
\end{aligned} \tag{3.3}$$
où on a utilisé dans la dernière ligne l'identité $(A \cup B) = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$ et la sous-addivité de μ^*. Nous avons démontré l'égalité $\mu^*(E) = \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$ et donc la

μ^*-mesurabilité de $A \cup B$. Remarquons que puisque toutes les inégalités (3.3) sont des égalités, lorsque $A \cap B = \emptyset$ on a

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap B) + \mu^*(E \cap (A \cup B)^c) \\ &= \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)\end{aligned}$$

et donc pour tout $E \subset \Omega$, avec $\mu^*(E) < \infty$ et pour tout paire A et B d'ensemble μ^*-mesurables disjoints on a

$$\mu^*(E \cap (A \cup B)) = \mu^*(E \cap A) + \mu^*(E \cap B).$$

Par récurrence, on obtient que pour tout $E \subset \Omega$, avec $\mu^*(E) < \infty$ et pour tout famille finie $A_1, \ldots, A_n$ d'ensembles μ^*-mesurables disjoints deux à deux on a

$$\mu^*(E \cap \bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu^*(E \cap A_i). \tag{3.4}$$

Soit maintenant $(A_i)_{i=1}^{\infty}$ une suite d'ensembles μ^*-mesurables. On veut montrer que $\bigcup_{i=1}^{\infty} A_i$ est μ^*-mesurable. On peut supposer alors que les ensembles A_i sont disjoints deux à deux, car nous avons montré que la famille Σ_{μ^*} d'ensembles μ^*-mesurables est fermée par rapport aux réunions finies et au passage aux complémentaires ; la famille $(B_i = A_i \setminus \bigcup_{j<i} A_j)$ est alors une famille d'ensembles μ^*-mesurables disjoints deux à deux dont la réunion coïncide avec la réunion des A_i. Donc supposons les A_i disjoints deux à deux. Pour tout $N \in \mathbb{N}$ et tout $E \subset \Omega$, avec $\mu^*(E) < \infty$, la μ^*-mesurabilité de $\bigcup_{i=1}^{N} A_i$ se traduit dans l'égalité

$$\mu^*(E) = \mu^*(E \cap \bigcup_{i=1}^{N} A_i) + \mu^*(E \cap (\bigcup_{i=1}^{N} A_i)^c)$$

En utilisant l'égalité (3.4), et minorant le terme $\mu^*(E \cap (\bigcup_{i=1}^{N} A_i)^c)$ avec $\mu^*(E \cap (\bigcup_{i=1}^{\infty} A_i)^c)$ on obtient

$$\mu^*(E) \geq \sum_{i=1}^{N} \mu^*(E \cap A_i) + \mu^*(E \cap (\bigcup_{i=1}^{\infty} A_i)^c)$$

Cette inégalité étant vraie pour tout $N \in \mathbb{N}$ on peut passer à la limite et l'on obtient

$$\mu^*(E) \geq \sum_{i=1}^{\infty} \mu^*(E \cap A_i) + \mu^*(E \cap (\bigcup_{i=1}^{\infty} A_i)^c) \geq \mu^*(E \cap \bigcup_{i=1}^{\infty} A_i) + \mu^*(E \cap (\bigcup_{i=1}^{\infty} A_i)^c) \geq \mu^*(E).$$

Les inégalité ci-dessous sont donc des égalités, ce qui démontre à la fois (a) que $\bigcup_{i=1}^{\infty} A_i$ est μ^*-mesurable et que (b) pour tout $E \subset \Omega$ de mesure finie et toute famille $(A_i)_{i=1}^{\infty}$ d'ensembles μ^*-mesurables disjoints deux à deux on a

$$\mu^*(E \cap \bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu^*(E \cap A_i)$$

En particulier si $\mu^*(\bigcup_{i=1}^{\infty} A_i) < \infty$, en posant $E = \bigcup_{i=1}^{\infty} A_i$ dans cette égalité on obtient

$$\mu^*(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu^*(A_i)$$

Si $\mu^*(\bigcup_{i=1}^{\infty} A_i) = \infty$ cette égalité est aussi vraie par la sous-addivité de la mesure extérieure μ^*. En conclusion on a montré que μ^* restreinte à la tribu Σ_{μ^*} des ensembles μ^*-mesurables est une mesure.

Il reste, pour conclure, à montrer que l'espace mesuré $(\Omega, \Sigma_{\mu^*}, \mu^*)$ est complet. Pour cela il suffit de montrer que tout $A \subset \Omega$ avec $\mu^*(A) = 0$ est μ^*-mesurable. Soit A un tel ensemble et $E \subset \Omega$. Puisque $E \cap A^c \subset E$ on a $\mu^*(E \cap A^c) \leq \mu^*(E)$; de l'autre coté $\mu^*(E \cap A) \leq \mu^*(A) = 0$. Donc

$$\mu^*(E \cap A) + \mu^*(E \cap A^c) \leq \mu^*(E),$$

ce qui montre la mesurabilité de A.

Les mesures de Lebesgue-Stieltjes

Si μ est une mesure de probabilité définie sur la tribu de Borel $\mathcal{B}(\mathbb{R})$ de $\mathbb{R}$ on pose

$$F(x) = \mu(\,]-\infty, x]) \tag{3.5}$$

Cette fonction est appelée **fonction de répartition** de la mesure de probabilité μ.

On verra dans cette section que la fonction de répartition détermine de façon unique la mesure μ.

Lorsque μ n'est pas finie, on est obligé de modifier la définition de fonction de répartition : on supposera que μ est finie sur les parties compactes de $\mathbb{R}$ et l'on posera [2]

$$F(x) = \begin{cases} \mu(\,]0, x]) & \text{si } x \geq 0; \\ -\mu(\,]x, 0]) & \text{si } x < 0. \end{cases} \tag{3.6}$$

Proposition 14 *Soit μ une mesure définie sur la tribu de Borel $\mathcal{B}(\mathbb{R})$ de $\mathbb{R}$, finie sur les parties compactes de $\mathbb{R}$. Soit F définie par* (3.5)*, si μ est une probabilité, ou définie par* (3.6) *dans le cas contraire.*

La fonction F est continue à droite et croissante et satisfait

$$\mu(\,]a, b]\,) = F(b) - F(a)$$

pour tout intervalle semi-ouvert à gauche $]a, b]$.

De plus, F est continue à gauche en $x \in \mathbb{R}$ si et seulement si le singleton $\{x\}$ a une mesure nulle.

Si μ est une probabilité et F est définie par (3.5)*, on a*

$$\lim_{t \to -\infty} F(t) = 0, \qquad \lim_{t \to +\infty} F(t) = 1.$$

Preuve. Exercice.

Dans cette section, nous étudions le problème inverse qui consiste à retrouver la mesure μ à partir de la fonction F.

Soit donc $F\colon \mathbb{R} \to \mathbb{R}$ une fonction **croissante et continue à droite**. Pour tout intervalle semi-ouvert à gauche $]a, b]$ on pose

$$\rho_F(\,]a, b]\,) = F(b) - F(a).$$

(En particulier si $F(x) = x$, pour tout $x \in \mathbb{R}$, $\rho_F(\,]a, b]\,)$ est la longueur de l'intervalle $]a, b]$.)

Soit $\mathcal{C}$ la famille des intervalles semi-ouverts à gauche de $\mathbb{R}$ et considérons la pré-mesure $(\mathbb{R}, \mathcal{C}, \rho_F)$. Par la proposition 12, l'application $\mu_F^*\colon \mathbb{R} \to [0, \infty]$ définie pour $E \subset \mathbb{R}$ par

$$\mu_F^*(E) = \inf \left\{ \sum_{j \in \mathbb{N}} \rho_F(\,]a_j, b_j]\,) \;\middle|\; E \subset \bigcup_{j \in \mathbb{N}} \,]a_j, b_j] \right\} \tag{3.7}$$

est une mesure extérieure. Le théorème de Carathéodory nous dit alors que la famille des sous-ensembles de $\mathbb{R}$ qui sont μ_F^*-mesurables est une tribu ; de plus, μ_F^* restreinte à cette tribu est une mesure.

Définition.

Soit $F\colon \mathbb{R} \to \mathbb{R}$ une fonction croissante et continue à droite. Soit μ_F^* la mesure extérieure définie par formule (3.7) et soit $\mathcal{L}_F$ la tribu des ensembles μ_F^*-mesurables. La **mesure de Lebesgue-Stieltjes associée à** F est, par définition, la restriction λ_F de μ_F^* à la tribu $\mathcal{L}_F$.

2. Dans le cas où μ est une mesure de probabilité, la nouvelle fonction F, définie par la formule (3.6), diffère de l'ancienne, définie par la formule (3.5), par la constante $\mu(\,]-\infty, 0])$.

Nous nous proposons maintenant de démontrer que tout borélien est μ_F^*-mesurable et que la mesure μ_F^* d'un intervalle semi-ouvert à gauche coïncide avec ρ_F.

Lemme. Les intervalles ouverts à gauche sont μ_F^*-mesurables.

Preuve. Soit $]a,b] \subset \mathbb{R}$. Il faut montrer que pour tout $E \subset \mathbb{R}$

$$\mu_F^*(E) \geq \mu_F^*(E \cap \,]a,b]\,) + \mu_F^*(E \cap \,]a,b]^c\,). \tag{3.8}$$

On peut donc supposer que $\mu_F^*(E) < \infty$. Fixons donc $\epsilon > 0$ et soit $\big(\,]a_i,b_i]\,\big)_{i\in\mathbb{N}}$ une suite d'intervalles semi-ouverts à gauche telle que

$$E \subset \bigcup_{i\in\mathbb{N}} \,]a_i,b_i]\,, \quad \text{et} \quad \mu_F^*(E) \geq \sum_{i\in\mathbb{N}} \rho_F(\,]a_i,b_i]\,) + \epsilon. \tag{3.9}$$

Or, pour tout $i \in \mathbb{N}$, l'intervalle $]a_i,b_i]$ est la réunion

$$]a_i,b_i] = \,]c_i,d_i] \;\cup\;]d_i,e_i] \;\cup\;]e_i,f_i] \tag{3.10}$$

de trois intervalles disjoints et consécutifs (dont certaines sont peut-être vides) définis par

$$]c_i,d_i] = \,]a_i,b_i] \;\cap\;]-\infty,a]\,, \qquad]d_i,e_i] = \,]a_i,b_i] \;\cap\;]a,b] \quad \text{et}$$
$$]e_i,f_i] = \,]a_i,b_i] \;\cap\;]b,+\infty]\,.$$

On obtient,

$$E \cap \,]a,b] \subset \bigcup_{i\in\mathbb{N}} \,]d_i,e_i]\,, \quad \text{et} \quad E \cap \,]a,b]^c \subset \bigcup_{i\in\mathbb{N}} \,]c_i,d_i] \cup \bigcup_{i\in\mathbb{N}} \,]e_i,f_i]$$

et donc que

$$\mu_F^*(E \cap \,]a,b]\,) \leq \sum_{i\in\mathbb{N}} \rho_F(\,]d_i,e_i]\,) \quad \text{et}$$
$$\mu_F^*(E \cap \,]a,b]^c\,) \leq \sum_{i\in\mathbb{N}} \rho_F(\,]c_i,d_i]\,) + \sum_{i\in\mathbb{N}} \rho_F(\,]e_i,f_i]\,).$$

Comme, par la définition de ρ_F et par la (3.10), on a

$$\rho_F(\,]c_i,d_i]\,) + \rho_F(\,]d_i,e_i]\,) + \rho_F(\,]e_i,f_i]\,) = \rho_F(\,]a_i,b_i]\,) \tag{3.11}$$

on obtient

$$\sum_{i\in\mathbb{N}} \rho_F(\,]a_i,b_i]\,) \geq \mu_F^*(E \cap \,]a,b]\,) + \mu_F^*(E \cap \,]a,b]^c\,)$$

et par l'inégalité (3.9) que

$$\mu_F^*(E) \geq \mu_F^*(E \cap \,]a,b]\,) + \mu_F^*(E \cap \,]a,b]^c\,) + \epsilon.$$

Le nombre ϵ étant arbitraire, on a montré l'inégalité (3.8) et la mesurabilité de l'intervalle $]a,b]$.

Corollaire

> Tout borélien est μ_F^*-mesurable.

Preuve. La famille des intervalles semi-ouverts à gauche engendre la tribu de Borel. Or, par le théorème de Carathéodory, les ensembles μ_F^*-mesurables forment une tribu qui, par le lemme précédent, contient la famille des intervalles semi-ouverts.

Lemme. Pour tout intervalle semi-ouvert $]a,b]$, on a

$$\mu_F^*(\,]a,b]\,) = \rho_F(\,]a,b]\,) = F(b) - F(a).$$

Preuve. Par la définition de la mesure extérieure μ_F^* et la définition de ρ_F on a

$$\mu_F^*(\,]a,b]\,) \le \rho_F(\,]a,b]\,) = F(b) - F(a). \tag{3.12}$$

Montrons d'abord que si $]a,b]$ est recouvert par une famille finie $=\ ]a_1,b_1],\ \ldots,\ ]a_n,b_n]$ d'intervalles semi-ouverts à gauche on a

$$\rho_F(\,]a,b]\,) \le \sum_{i=1}^{n} \rho_F(\,]a_i,b_i]\,)$$

L'affirmation est vraie si $n=1$ car, dans ce cas, $a_1 \le a \le b \le b_1$ et, comme F est croissante, pm a

$$\rho_F(\,]a_1,b_1]\,) = F(b_1) - F(a_1) \ge F(b) - F(a) = \rho_F(\,]a,b]\,).$$

Supposons par récurrence que l'affirmation soit vraie pour $n \le k-1$, et supposons que les k intervalles $]a_1,b_1],\ \ldots,\ ]a_k,b_k]$ recouvrent $]a,b]$. Il existe donc un intervalle $]a_i,b_i]$ parmi les k donnés tel que $b \in\ ]a_i,b_i]$. Les intervalles restantes doivent alors forcement recouvrir l'intervalle $]a,a_i]$. En ce cas, par hypothèse de récurrence on a

$$F(a_i) - F(a) = \rho_F(\,]a,a_i]\,) \le \sum_{j \le k, j \ne i} \rho_F(\,]a_j,b_j]\,)$$

Comme $b \in\ ]a_i,b_i]$, on a $F(b) \le F(b_i)$ et on conclut que

$$F(b) - F(a) \le F(b_i) - F(a) = \rho_F(\,]a_i,b_i]\,) + F(a_i) - F(a) \ \le \sum_{j=1}^{k} \rho_F(\,]a_i,b_i]\,)$$

L'affirmation est donc vraie pour tout $n \in \mathbb{N}$.

Montrons maintenant que si $]a,b] \subset \bigcup_{i=0}^{\infty}\ ]a_i,b_i]$ alors

$$\rho_F(\,]a,b]\,) \le \sum_{i=1}^{\infty} \rho_F(\,]a_i,b_i]\,) \tag{3.13}$$

Il s'ensuivra alors que $\rho_F(\,]a,b]\,) \le \mu_F^*(\,]a,b]\,)$ et la preuve, en vue de l'inégalité (3.12), sera achevée.

Soit $\epsilon > 0$. Par la continuité à droite de la fonction F il existe $a' \in\ ]a,b]$ tel que $F(a') - F(a) < \epsilon$, et donc tel que

$$\rho_F(\,]a,b]\,) < \rho_F(\,]a',b]\,) + \epsilon. \tag{3.14}$$

Pour la même raison, pour tout $i \in \mathbb{N}$, il existe $b_i' > b_i$ tel que $F(b_i') - F(b_i) < 2^{-i}\epsilon$, ce qui donne

$$\rho_F(\,]a_i,b_i']\,) < \rho_F(\,]a_i,b_i]\,) + 2^{-i}\epsilon. \tag{3.15}$$

L'intervalle compact $]a',b\,[$ est recouvert par les intervalles $\big(\,]a_i,b_i]\,\big)_{i\in\mathbb{N}}$ et donc **a fortiori** par les intervalles ouverts $\big(\,]a_i,b_i'[\,\big)_{i\in\mathbb{N}}$. Il existe alors un nombre fini d'intervalles de la collection $\big(\,]a_i,b_i']\,\big)_{i\in J}$ qui recouvre $]a',b\,]$. Donc, par les (3.14) et (3.15) on obtient

$$\rho_F(\,]a,b]\,) - \epsilon \le \rho_F(\,]a',b]\, \le \sum_{i\in J} \rho_F\Big(\,\Big]a_i,b_j'\Big] \le \sum_{i\in\mathbb{N}} \rho_F(\,]a_i,b_j] + \sum_{i\in\mathbb{N}} 2^{-i}\epsilon$$

ou bien $\rho_F(\,]a,b]\,) \le \sum_{i\in\mathbb{N}} \rho_F(\,]a_i,b_j] + 3\epsilon$. Ceci étant valable pour tout $\epsilon > 0$ on a bien montré l'inégalité (3.13) et terminé la preuve.

On peut donc résumer et préciser les conséquences des résultats jusqu'ici obtenus dans le théorème suivant.

Théorème 15 *(a) Soit $F\colon \mathbb{R} \to \mathbb{R}$ une fonction croissante et continue à droite et soit $(\mathbb{R}, \mathcal{L}_F, \lambda_F)$ la mesure de Lebesgue-Stieltjes associée à F. Alors la tribu $\mathcal{L}_F$ contient la tribu de Borel $\mathcal{B}(\mathbb{R})$. La mesure λ_F est* ***finie sur le parties compactes de*** $\mathbb{R}$ *et pour tout intervalle semi-ouvert à gauche $]a, b]$ on a*

$$\lambda_F(\,]a, b]) = F(b) - F(a).$$

pour tout intervalle semi-ouvert à gauche $]a, b] \subset \mathbb{R}$.

(b) Si G est une autre fonction croissante et continue à droite, l'égalité $\lambda_F = \lambda_G$ entraîne que $F - G$ est une constante.

(c) Si ν est une mesure définie sur la tribu de Borel $\mathcal{B}(\mathbb{R})$ telle que $\nu\big(\,]a, b]\,\big) = F(b) - F(a)$ pour tout intervalle semi-ouvert à gauche $]a, b]$ alors $\nu(A) = \lambda_F(A)$ pour tout borélien $A \subset \mathbb{R}$.

Preuve. (a) Que la la tribu de Borel $\mathcal{B}(\mathbb{R})$ soit contenue dans $\mathcal{L}_F$ et que pour tout intervalle semi-ouvert à gauche $]a, b]$ on ait $\lambda_F(\,]a, b]) = F(b) - F(a)$ sont les affirmations du Corollaire 3.2 et du Lemme 3.2.

La mesure λ_F est finie sur le parties compactes de $\mathbb{R}$, car tout compact $K \subset \mathbb{R}$ est contenu dans un intervalle $]-n, n]$ et on a donc $\lambda_F(K) \leq \lambda_F(\,]-n, n]) = F(n) - F(-n) < \infty$.

(b) Si G est une autre fonction croissante et continue à droite, l'égalité $\lambda_F = \lambda_G$ entraîne que

$$F(b) - F(a) = \lambda_F(\,]a, b]) = \lambda_G(\,]a, b]) = G(b) - G(a).$$

Donc $F(x) - G(x) = F(0) - G(0)$, pour tout $x \in \mathbb{R}$.

(c) Soit $A \subset \,]-n, n]$ un borélien. Pour toute famille $\big(\,]a_i, b_i]\,\big)_{i\in\mathbb{N}}$ telle que $A \subset \bigcup_{i\in\mathbb{N}} \,]a_i, b_i]$ on a $\nu(A) \leq \sum_{i\in\mathbb{N}} \nu\big(\,]a_i, b_i]\,\big) = \sum_{i\in\mathbb{N}} \rho_F\big(\,]a_i, b_i]\,\big)$. En prenant le infimum sur toute telles familles $\big(\,]a_i, b_i]\,\big)_{i\in\mathbb{N}}$ on obtient l'inégalité $\nu(A) \leq \mu_F^*(A) = \lambda_F(A)$. Par le même argument on a aussi l'inégalité $\nu\big(\,]-n, n] \setminus A\big) \leq \lambda_F\big(\,]-n, n] \setminus A\big)$. Comme $\nu\big(\,]-n, n]\,\big) = \lambda_F\big(\,]-n, n]\,\big) = F(n) - F(-n)$, et comme $]-n, n]$ est l'union disjointe de A et $]-n, n] \setminus A$ aucune de des deux inégalité peut être stricte. Donc $\nu(A) = \lambda_F(A)$ pour tout borélien $A \subset \,]-n, n]$.

Si $A \subset \mathbb{R}$ est un borélien quelconque, par la Proposition 9, on a $\nu(A) = \lim_n \nu\big(\,]-n, n] \cap A\big) = \lim_n \lambda_F\big(\,]-n, n] \cap A\big) = \lambda_F(A)$.

Un cas particulier est le cas où $F(x) = x$ pour tout $x \in \mathbb{R}$.

Définition.

L'espace mesuré de Lebesgue-Steltjes $(\mathbb{R}, \mathcal{L}_\mathbb{R}, \lambda)$ associé à la fonction identique de $\mathbb{R}$ et donc tel que que

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

pour tout intervalle semi-ouvert à gauche $]a, b]$, est dit la **mesure de Lebesgue**. La mesure λ et la tribu $\mathcal{L}_\mathbb{R}$ sont dites la **mesure de Lebesgue de** $\mathbb{R}$ et la **tribu de Lebesgue de** $\mathbb{R}$.

Remarquons que, par la Proposition 9, pour tout $(a, b) \in \mathbb{R}^2$ avec $-\infty < a < b < +\infty$, le nombre $b - a$ est également la mesure de Lebesgue de tout intervalle d'extrémités a, b car

$$]a, b[\; = \bigcup_{n\in\mathbb{N}^+} \Big]a, b - n^{-1}\Big], \qquad]a, b[\; = \bigcap_{n\in\mathbb{N}^+} \Big]a - n^{-1}, b\Big], \quad [a, b] = \bigcup_{n\in\mathbb{N}^+} \Big]a, b - n^{-1}\Big[.$$

Exercice. On montrera les propriétés suivantes de la mesure de Lebesgue :

1. La mesure de Lebesgue est la seule mesure sur tribu de Borel $\mathcal{B}(\mathbb{R})$ qui est finie sur les compacts et invariante par translation.
2. Pour tout $\alpha > 0$ et tout Borelien $A \subset \mathbb{R}$ la mesure de Lebesgue de l'ensemble $\alpha A = \{\alpha x \mid x \in A\}$ est égale à $\alpha\,\lambda(A)$.