

Intégration et probabilités

César Almecija
étudiant en première année aux MINES ParisTech
cesar.almecija@mines-paristech.fr

28 avril 2021

Table des matières

I	Théorie de la mesure et intégration	3
1	Introduction à la théorie de la mesure	4
1.1	Ensembles mesurables	4
1.2	Définition de la mesure	5
1.3	Exemples importants	6
1.4	Fonctions mesurables	7
1.5	Approximation des fonctions mesurables	9
2	Construction de l'intégrale de Lebesgue	10
2.1	Pourquoi définir une nouvelle intégrale ?	10
2.2	Définition de l'intégrale de Lebesgue pour des fonctions positives	11
2.3	Propriétés de l'intégrale positive	12
2.4	Définition de l'intégrale de Lebesgue pour des fonctions signées .	12
2.5	Démontrer en pratique des résultats sur les intégrales	13
2.6	Généralisation aux fonctions à valeurs dans \mathbb{R}^m	13
3	L'intégrale de Lebesgue munie de la mesure de Lebesgue	15
II	Probabilités	16
4	Généralités sur les probabilités	17
4.1	Premières définitions	17
4.2	Variables aléatoires	17
4.2.1	Loi d'une variable aléatoire	18
4.2.2	Moments d'une variable aléatoire réelle ou discrète	19
4.2.3	Formule de transfert	21
4.2.4	Moments d'un vecteur aléatoire	22

Avant-propos

Ce document vise à résumer les notions essentielles d'intégration et de probabilités. Son objectif est de retracer les cheminements de pensée qui mènent aux résultats essentiels, afin de mieux les retenir. Ainsi, il pourra être avantageusement utilisé pour se refamiliariser avec ces notions, au cas-où elles auraient été oubliées. Ce travail peut également servir de préambule à une étude plus approfondie du sujet, pour découvrir les fondements de ce domaine des mathématiques.

Néanmoins, ce document ne saurait se substituer à un ouvrage de référence. En effet, par soucis de clarté et de concision, il ne détaille pas toutes les démonstrations. Celles-ci sont néanmoins nécessaires pour saisir l'entièreté des résultats énoncés ici. De cette manière, **il faut le lire comme un aide-mémoire ou un résumé, et non comme un cours.**

Ce travail se partage en trois grandes parties :

- Une première partie commençant par une introduction à la théorie de la mesure, suivie de la construction rapide de l'intégrale de Lebesgue, et qui finit par s'attarder sur les résultats essentiels d'intégration
- Une deuxième partie qui applique la partie précédente à la théorie des probabilités
- Une troisième partie rapide qui explore la théorie des séries à l'aide des résultats précédents.

Ce document s'est grandement inspiré du cours Calcul Différentiel, Intégral et Stochastique (CDIS) de l'Ecole des MINES ParisTech. L'auteur recommande aux lecteurs intéressés de s'y référer régulièrement.

Première partie

**Théorie de la mesure et
intégration**

Chapitre 1

Introduction à la théorie de la mesure

Pour comprendre cette notion, partons d'un constat évident. On sait mesurer la taille d'un segment : la manière de mesurer cette grandeur est d'utiliser la *longueur*. Par exemple, la longueur du segment $[0, 1]$ vaut 1. Plus généralement, λ , l'application qui à un segment $[a, b]$ associe sa longueur $b - a$, pourrait s'appeler dans le langage courant une *mesure*, c'est-à-dire une fonction permettant d'obtenir la mesure (la taille) d'un objet.

Se posent alors les questions suivantes :

- dans \mathbb{R} , comment mesurer n'importe-quel ensemble ?
- plus généralement, comment mesurer dans \mathbb{R}^n ?
- mais est-il possible de *tout* mesurer ? Ou faut-il se restreindre à une catégorie d'ensembles *mesurables* ?

On commencera par répondre à cette dernière question.

1.1 Ensembles mesurables

Il n'est pas possible de tout mesurer ainsi. Il faut introduire une classe d'ensembles, appelée **tribu**, qui va représenter l'ensemble des ensembles **mesurables**.

Une tribu doit cependant garantir certaines propriétés de stabilité par opérations ensemblistes. En effet, si deux ensembles sont mesurables, on aimerait pouvoir donner un sens à la mesure de leur intersection ou de leur union. De même, si un ensemble est mesurable, il est légitime de demander que son complémentaire soit mesurable à son tour. Enfin, demander que l'ensemble vide soit mesurable est raisonnable, et on verra par la suite que la mesure du vide doit valoir 0, conformément à l'intuition.

Ces considérations étant faites, on définit alors une tribu comme suit :

Définition 1.1.1. Une **tribu** (ou σ -**algèbre**) \mathcal{A} d'un ensemble E est une col-

lection d'ensembles $\mathcal{A} \subset \mathcal{P}(E)$ vérifiant les hypothèses suivantes :

1. $\emptyset \in \mathcal{A}$
2. \mathcal{A} est stable par passage au complémentaire
3. \mathcal{A} est stable par union dénombrable

Définition 1.1.2. Un **espace mesurable** (E, \mathcal{A}) est un ensemble E muni d'une tribu \mathcal{A} .

Définition 1.1.3. Soit (E, \mathcal{A}) un espace mesurable. Un ensemble $X \in \mathcal{A}$ est dit **\mathcal{A} -mesurable**.

Remarque 1.1.1. Si le contexte est clair (*ie* il n'y a qu'une seule tribu), on pourra omettre \mathcal{A} et parler simplement d'ensemble mesurable

Remarque 1.1.2. On remarquera qu'il faut distinguer les *espaces* mesurables (définition 1.1.2) des *ensembles* mesurables (définition 1.1.3).

Définition 1.1.4. Soit E un ensemble, et B une collection d'ensembles de E . La **tribu engendrée par B** est la plus petite tribu sur E contenant B . De manière équivalente, c'est l'intersection de toutes les tribus de E contenant B .

Définition 1.1.5. Soit E un ensemble muni d'une topologie. La **tribu de Borel** (ou plus simplement, les **boréliens**) est la tribu engendrée par les ouverts de E (ou de manière équivalente, par les fermés de E). Elle est notée $\mathcal{B}(E)$.

1.2 Définition de la mesure

Abordons désormais la notion de mesure. Un ensemble mesurable doit intuitivement admettre une mesure positive. De plus, si deux ensembles sont disjoints, il est légitime de demander que la mesure de leur union soit la somme des mesures.

Cela nous amène à définir une mesure ainsi :

Définition 1.2.1. Une **mesure** μ sur un espace mesurable (E, \mathcal{A}) est une application $\mu : \mathcal{A} \longrightarrow \mathbb{R}_+ \cup \{+\infty\}$ telle que :

1. $\mu(\emptyset) = 0$
2. Pour une famille au plus dénombrable d'ensembles mesurables $(A_i)_{i \in I}$,

$$\mu \left(\bigcup_{i \in I} A_i \right) = \sum_{i \in I} \mu(A_i)$$

Remarque 1.2.1. Cette deuxième propriété porte le nom de **σ -additivité**.

Remarque 1.2.2. Si $\mu(\emptyset) \neq 0$, le lecteur pourra vérifier que la mesure de n'importe-quel ensemble mesurable (*a fortiori* du vide) vaut $+\infty$. On comprend alors pourquoi on impose que la mesure du vide soit nulle : si ce n'était pas le cas, mesurer n'aurait pour ainsi dire aucun intérêt.

Définition 1.2.2. Un **espace mesuré** (E, \mathcal{A}, μ) est un espace mesurable (E, \mathcal{A}) muni d'une mesure μ .

Définition 1.2.3. Soit (E, \mathcal{A}, μ) un espace mesuré. Un ensemble $X \subset E$ est dit **μ -négligeable** si :

$$\exists Y \in \mathcal{A}, \begin{cases} \mu(Y) = 0 \\ X \subset Y \end{cases}$$

Remarque 1.2.3. On remarque qu'un ensemble négligeable n'est pas forcément mesurable. Il est cependant commode d'imposer qu'un ensemble négligeable soit nécessairement mesurable. Modifier une mesure et la tribu respective pour arriver à ce résultat porte le nom de **complétion d'une mesure**.

Ce processus ne sera pas détaillé ici, mais le lecteur peut néanmoins retenir qu'une mesure complétée vérifie l'équivalence suivante : **un ensemble est négligeable ssi il est mesurable, de mesure nulle**

Remarque 1.2.4. On notera que la notion d'ensemble négligeable dépend de la mesure choisie. Un ensemble négligeable pour une certaine mesure peut très bien ne pas l'être pour une autre mesure.

1.3 Exemples importants

Il existe trois mesures importantes à connaître.

Définition 1.3.1. La **mesure de Lebesgue** λ est l'unique mesure sur \mathbb{R}^n qui prolonge la notion de volume. Habituellement, sa tribu de définition est la **tribu de Lebesgue** ou la **tribu de Borel**.

Remarque 1.3.1. En définissant la mesure de Lebesgue sur la tribu de Borel, elle ne serait en fait pas complète (voir 1.2.3). La tribu de Lebesgue (notée $\mathcal{L}(\mathbb{R}^n)$) est alors définie comme étant la plus petite tribu de \mathbb{R}^n permettant à cette mesure d'être complète.

On ne retiendra pas plus de détails sur la tribu de Lebesgue, mais il est bon de retenir qu'**un élément Borel-mesurable est Lebesgue-mesurable** (la réciproque est fausse).

Remarque 1.3.2. On remarque immédiatement que :

- lorsque $n = 1$, cette mesure prolonge la notion de longueur vue en introduction du chapitre.
- lorsque $n = 2$, cette mesure prolonge la notion de surface.
- lorsque $n = 3$, cette mesure prolonge la notion de volume.

Par ailleurs, cette mesure sera centrale dans la partie ?? concernant les intégrales.

Définition 1.3.2. Soit (E, \mathcal{A}) un espace mesurable. La **mesure de comptage** c est définie sur cette espace comme suit :

$$c : \begin{cases} \mathcal{A} & \longrightarrow & \mathbb{R}_+ \cup \{+\infty\} \\ A & \longmapsto & \begin{cases} 0 & \text{si } A = \emptyset \\ n & \text{si } A \text{ est fini de cardinal } n \\ +\infty & \text{si } A \text{ est infini} \end{cases} \end{cases}$$

Remarque 1.3.3. Comme son nom l'indique, cette mesure compte les éléments présents dans l'ensemble que l'on mesure.

Remarque 1.3.4. Bien que cette mesure soit définie pour un espace mesurable quelconque, elle sera particulièrement utile dans $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ (notamment pour la partie ?? concernant les séries).

Définition 1.3.3. Soit (E, \mathcal{A}) un espace mesurable. La **mesure de Dirac en** $x \in E$ δ_x est définie comme suit :

$$\delta_x : \begin{cases} \mathcal{A} & \longrightarrow & \mathbb{R}_+ \\ A & \longmapsto & \begin{cases} 0 & \text{si } x \in A \\ 1 & \text{si } x \notin A \end{cases} \end{cases}$$

Remarque 1.3.5. La mesure de Dirac et la fonction indicatrice sont reliées de la manière suivante :

$$\forall A \in \mathcal{A}, \forall x \in E, \delta_x(A) = \mathbf{1}_A(x)$$

1.4 Fonctions mesurables

Lorsque deux espaces mesurables sont définis, il peut être utile de définir une classe de fonctions qui font bon ménage avec les tribus des espaces mis en jeu.

Définition 1.4.1. Soit (E, \mathcal{A}) et (F, \mathcal{B}) deux espaces mesurables. Une **fonction \mathcal{A}/\mathcal{B} -mesurable** est une fonction $f : E \longrightarrow F$ telle que :

$$\forall Y \in \mathcal{B}, f^{-1}(Y) \in \mathcal{A}$$

Remarque 1.4.1. Cette caractérisation des fonctions mesurables porte le nom de **critère de l'image réciproque**.

Remarque 1.4.2. Cette notation est lourde, et c'est pourquoi elle est souvent simplifiée.

- dans le cas où $\mathcal{B} = \mathcal{B}(F)$ (la tribu d'arrivée est celle des boréliens), on pourra simplifier la notation et dire simplement que :

f est \mathcal{A} -mesurable

- si, de plus, la tribu de l'espace de départ n'est pas ambiguë, on pourra alors simplement écrire que :

f est mesurable

Remarque 1.4.3. On remarquera que la notion de fonction mesurable est **indépendante des mesures choisies**. Elle ne dépend que des tribus de l'espace de départ et d'arrivée.

Voici un lien bien utile entre ensembles et fonctions mesurables :

Proposition 1.4.1. *Soit (E, \mathcal{A}) un espace mesurable. Soit $A \subset E$. Alors A est mesurable ssi $\mathbb{1}_A$ est mesurable.*

Remarque 1.4.4. On remarquera que l'on a bien utilisé le formalisme de la remarque précédente :

- la tribu de l'espace d'arrivée est la tribu borélienne de \mathbb{R} .
- la tribu de l'espace de départ est \mathcal{A} , non-ambigüe.

La proposition suivante permet de simplifier considérablement la vérification de la mesurabilité d'une fonction, dans le cas où l'espace d'arrivée est muni de la tribu des boréliens.

Elle découle directement du fait que la tribu des boréliens soit la tribu engendrée par les ouverts : pour vérifier que le résultat est vrai pour tout élément de la tribu, il suffit de le vérifier sur les éléments qui engendrent la tribu (en l'occurrence, les ouverts).

Proposition 1.4.2. *Soit (E, \mathcal{A}) et $(F, \mathcal{B}(F))$ deux espaces mesurables. f est \mathcal{A} -mesurable ssi pour tout ouvert U de F , $f^{-1}(U) \in \mathcal{A}$.*

Remarque 1.4.5. En vertu de la définition 1.1.5, on peut remplacer dans la proposition 1.4.2 « ouvert » par « fermé ».

Le résultat suivant est un résultat important concernant la composition des fonctions mesurables :

Proposition 1.4.3. *Soit (E, \mathcal{A}) , (F, \mathcal{B}) et (G, \mathcal{C}) trois espaces mesurables, $f : E \rightarrow F$ une fonction \mathcal{A}/\mathcal{B} -mesurable et $g : F \rightarrow G$ une fonction \mathcal{B}/\mathcal{C} -mesurable. Alors $g \circ f : E \rightarrow G$ est une fonction \mathcal{A}/\mathcal{C} -mesurable*

Enfin, voici une dernière définition qui sera utile dans le chapitre concernant les probabilités :

Définition 1.4.2. Une fonction $f : E \rightarrow F$ est dite **borélienne** lorsqu'elle est $\mathcal{B}(E)/\mathcal{B}(F)$ -mesurable.

Proposition 1.4.4. *Une fonction continue est borélienne.*

Remarque 1.4.6. Cette proposition est une conséquence directe de la caractérisation de la continuité d'une fonction par l'image réciproque des ouverts. Attention, la réciproque de ce résultat est fausse : il existe des fonctions boréliennes qui ne sont pas continues.

1.5 Approximation des fonctions mesurables

Les fonctions mesurables, en plus de faire bon ménage avec les ensembles mesurables, peuvent être approximées facilement par des fonctions dites « étagées ». Nous allons commencer par définir cette classe de fonctions, puis nous verrons en quoi consiste cette approximation.

Définition 1.5.1. Une **fonction étagée** est une fonction dont l'image est finie.

La proposition suivante fournit une décomposition pratique des fonctions étagées :

Proposition 1.5.1. *Soit f une fonction étagée. Alors :*

$$\exists N \in \mathbb{N}, \exists (A_i)_{i \in [1, N]} \subset \mathbb{R}^N, \exists (y_i)_{i \in [1, N]} \in \mathbb{R}^N, f = \sum_{i=1}^N y_i \mathbb{1}_{A_i}$$

Remarque 1.5.1. Lorsque la fonction étagée que l'on manipule est mesurable, il existe toujours une décomposition de la forme précédente où les $(A_i)_{i \in [1, N]}$ ont le bon goût d'être mesurables.

C'est une propriété essentielle des fonctions étagées mesurables, qui sera utile dans prochain chapitre.

Théorème 1.5.1. *Toute fonction mesurable positive est limite simple d'une suite croissante de fonctions étagées mesurables positives.*

Remarque 1.5.2. Ici, la croissance d'une suite de fonctions $(f_k)_{k \in \mathbb{N}}$ est à comprendre dans le sens suivant :

$$\forall k \in \mathbb{N}, \forall x \in E, f_k(x) \leq f_{k+1}(x)$$

Un corollaire existe pour les fonctions mesurables signées, mais il est assez peu utile. Le voici à titre culturel seulement :

Corollaire 1.5.1. *Toute fonction mesurable est limite simple d'une suite absolument croissante de fonctions étagées mesurables.*

Remarque 1.5.3. Ici, l'absolue croissance d'une suite de fonctions $(f_k)_{k \in \mathbb{N}}$ est à comprendre dans le sens suivant :

$$\forall k \in \mathbb{N}, \forall x \in E, |f_k(x)| \leq |f_{k+1}(x)|$$

Chapitre 2

Construction de l'intégrale de Lebesgue

Maintenant que nous connaissons les notions de base concernant la théorie de la mesure, nous sommes en mesure de définir l'intégrale de Lebesgue. Cette intégrale est en quelque sorte une généralisation de l'intégrale de Riemann.

2.1 Pourquoi définir une nouvelle intégrale ?

Deux raisons principales peuvent nous amener à définir une nouvelle intégrale.

D'une part, l'intégrale de Riemann ne permet de calculer l'intégrale que de fonctions **continues par morceaux**. Or, cela est très restrictif : par exemple, prenons la fonction $\mathbb{1}_{\mathbb{Q}}$. Elle n'est pas continue par morceaux, donc n'est pas intégrable au sens de Riemann. Pourtant, pour la mesure de Lebesgue, \mathbb{Q} est négligeable (voir l'exemple ??). Ainsi, $\mathbb{1}_{\mathbb{Q}}$ est Lebesgue-presque-partout égale à 0. On aurait donc envie de dire que « l'influence de \mathbb{Q} est si petite, qu'elle n'a aucun impact dans l'intégrale ». En disant cela, on veut en fait dire que l'intégrale de $\mathbb{1}_{\mathbb{Q}}$ serait égale à l'intégrale de la fonction nulle, qui vaut zéro. Ce faisant, **on pourrait donner un sens à l'intégrale d'une fonction qui n'est pas continue par morceaux !**

Plus généralement, l'intégrale de Lebesgue permettra de définir l'intégrale d'une fonction mesurable, et non plus d'une fonction continue par morceaux.

D'autre part, l'intégrale de Riemann utilise la notion de longueur : par exemple, si f est la fonction constante égale à 1, nous avons

$$\int_0^1 f(x)dx = 1 > 2 = \int_0^2 f(x)dx$$

Mais il est impossible de « mesurer différemment ». En effet, si une mesure μ vérifie $\mu([1, 2]) = 0$, alors on aimerait pouvoir dire que les deux intégrales ci-dessus, calculées avec μ , seraient égales. Plus généralement, **on aimerait pou-**

voir intégrer dans n'importe quel espace par n'importe quelle mesure. L'intégrale de Lebesgue répondra à ce besoin.

2.2 Définition de l'intégrale de Lebesgue pour des fonctions positives

Les fonctions positives ont une place privilégiée dans la théorie de l'intégration de Lebesgue. En effet, leurs intégrales sont toujours définies (quitte à valoir $+\infty$) :

Définition 2.2.1. Soit (E, \mathcal{A}, μ) un espace mesuré et $f : E \rightarrow \mathbb{R}_+$ mesurable.

La μ intégrale de Lebesgue $\int_E f(x) d\mu(x)$ existe toujours et est l'unique élément de $\mathbb{R}_+ \cup \{+\infty\}$ vérifiant :

- **la relation mesure-intégrale** : pour tout ensemble X \mathcal{A} -mesurable,

$$\int_E \mathbf{1}_X(x) d\mu(x) = \mu(X)$$

- **l'hypothèse de linéarité** : pour toute fonction mesurable $g : E \rightarrow \mathbb{R}_+$ et pour tous $(\alpha, \beta) \in (\mathbb{R}_+^*)^2$,

$$\int_E (\alpha f + \beta g)(x) d\mu(x) = \alpha \int_E f(x) d\mu(x) + \beta \int_E g(x) d\mu(x)$$

- **l'hypothèse de convergence monotone** : pour toute suite croissante de fonctions mesurables positives $f_n : E \rightarrow \mathbb{R}_+$ convergeant simplement vers f ,

$$\lim_{n \rightarrow +\infty} \int_E f_n(x) d\mu(x) = \int_E f(x) d\mu(x)$$

Remarque 2.2.1. L'intégrale d'une fonction mesurable positive vérifie immédiatement, par définition, ces trois hypothèses fondamentales, qui seront centrales dans la suite de ce document. Attention cependant : les deux hypothèses de mesurabilité et de positivité sont absolument nécessaires.

Définition 2.2.2. Soit (E, \mathcal{A}, μ) un espace mesuré et $f : E \rightarrow \mathbb{R}_+$ mesurable.

f est dite **μ -intégrable** lorsque $\int_E f(x) d\mu(x) < +\infty$.

Remarque 2.2.2. On fera donc bien attention au sens précis du mot « intégrable ». Il est plus fort de dire qu'une fonction est intégrable, que de dire que l'intégrale d'une fonction existe. Par exemple, une fonction mesurable positive admet toujours une intégrale, mais elle n'est pas forcément intégrable.

Définition 2.2.3. Soit $f : E \rightarrow \mathbb{R}$, et $F \subset E$. On dit que f est **intégrable sur le sous ensemble F** lorsque $f\mathbf{1}_F$ est intégrable.

Remarque 2.2.3. Pour avoir une chance que f soit intégrable sur F , il est nécessaire que $f\mathbf{1}_F$ soit mesurable.

2.3 Propriétés de l'intégrale positive

2.4 Définition de l'intégrale de Lebesgue pour des fonctions signées

Commençons par des résultats préliminaires qui seront utiles par la suite :

Définition 2.4.1. Soit E un ensemble. Soit $f : E \rightarrow \mathbb{R}$. Les **partie positive** et **partie négative** de f , respectivement f_+ et f_- , sont définies comme suit :

$$\begin{cases} f_+ &:= \max(0, f) \\ f_- &:= \min(0, f) \end{cases}$$

Proposition 2.4.1. Soit E un ensemble et $f : E \rightarrow \mathbb{R}$. Alors :

$$f = f_+ - f_-$$

Le résultat suivant est une conséquence directe de la proposition ?? :

Proposition 2.4.2. Soit (E, \mathcal{A}) un espace mesurable et $f : E \rightarrow \mathbb{R}$ une fonction mesurable. Alors f_+ et f_- sont mesurables

Désormais, passons à la définition de l'intégrale d'une fonction signée. Celle-ci arrive volontairement dans une nouvelle section, pour souligner la différence avec l'intégrale d'une fonction positive. En effet, comme cela a été souligné au début de la section 2.2, les fonctions positives occupent une place privilégiée dans la théorie de l'intégration de Lebesgue : elles vérifient des résultats qui ne sont pas systématiquement vrais pour les fonctions signées.

La définition de l'intégrale pour une fonction signée est un premier exemple soulignant cette différence :

Définition 2.4.2. Soit (E, \mathcal{A}, μ) un espace mesuré et $f : E \rightarrow \mathbb{R}_+$ mesurable.

La μ intégrale de Lebesgue $\int_E f(x) d\mu(x)$ existe dès que f_+ ou f_- est μ -intégrable. Dans ce cas, on a :

$$\int_E f(x) d\mu(x) = \int_E f_+ d\mu(x) - \int_E f_- d\mu(x)$$

Remarque 2.4.1. Contrairement au cas positif, **l'intégrale d'une fonction signée n'existe pas systématiquement !**

Définition 2.4.3. Soit (E, \mathcal{A}, μ) un espace mesuré et $f : E \rightarrow \mathbb{R}$ mesurable. f est dite **μ -intégrable** lorsque f_+ et f_- sont μ -intégrables.

Remarque 2.4.2. On notera que, de manière équivalente, **une fonction signée est intégrable ssi l'intégrale de f existe, et est finie.**

Remarque 2.4.3. Comme dans le cas positif (remarque 2.2.2), le mot « intégrable » a un sens bien précis !

Théorème 2.4.1. Soit (E, \mathcal{A}, μ) un espace mesuré et $f : E \rightarrow \mathbb{R}$ mesurable. f est μ -intégrable ssi $|f|$ est μ -intégrable. On dit que **l'intégrale de Lebesgue est absolue**.

Remarque 2.4.4. Pour vérifier qu'une fonction signée est intégrable, on vérifiera en pratique que sa valeur absolue est intégrable.

Le théorème suivant assure que cette nouvelle intégrale vérifie bien les besoins que l'on avait exprimé en début de chapitre.

Théorème 2.4.2. L'intégrale de Lebesgue prolonge l'intégrale de Riemann.

2.5 Démontrer en pratique des résultats sur les intégrales

Pour montrer des résultats sur des fonctions mesurables positives, on utilise la décomposition vue au théorème 1.5.1.

- On montre le résultat sur des fonctions indicatrices mesurables.
- Puis on le montre sur une somme de fonctions indicatrices mesurables. On a donc montré le résultat pour les fonctions étagées mesurables.
- En vertu de l'hypothèse de convergence monotone de l'intégrale et de la décomposition fournie par le théorème 1.5.1, on passe à la limite. Le résultat est alors démontré pour n'importe-quelle fonction mesurable positive.

En général, les résultats pour les fonctions signées se déduiront des résultats pour les fonctions positives de la manière suivante :

- On décompose une fonction signée mesurable $f : E \rightarrow \mathbb{R}$ en sa partie positive f^+ et sa partie négative f_- (voir la définition 2.4.1).
- Ces deux fonctions étant positives (et encore mesurables par la proposition 2.4.2), on applique leur applique un résultat vrai pour les fonctions positives.
- On utilise l'identité $f = f^+ + f_-$ pour obtenir le résultat final.

On rappelle toutefois qu'un certain nombre de résultats vrais pour les fonctions positives ne le sont pas pour les fonctions signées.

2.6 Généralisation aux fonctions à valeurs dans \mathbb{R}^m

Nos fonctions étaient jusqu'ici à valeurs dans \mathbb{R} . Qu'en est-il si elle sont à valeurs dans \mathbb{R}^m ?

Soit $f : E \rightarrow \mathbb{R}^m$. Décomposons f :

$$f = \sum_{i=1}^m f_i e_i$$

où :

- $(f_i)_{i \in \llbracket 1, m \rrbracket}$ sont les **composantes** de f .
- $(e_i)_{i \in \llbracket 1, m \rrbracket}$ est la base canonique de \mathbb{R}^m .

Posons alors les définitions suivantes :

Définition 2.6.1. L'intégrale de f existe lorsque les m intégrales des f_i existent. Lorsqu'elle existe, l'intégrale de f vaut :

$$\int_E f = \sum_{i=1}^m \int_E f_i$$

(on peut permuter vecteurs et intégrale)

Définition 2.6.2. f est dite **intégrable** lorsque les m fonctions f_i sont intégrables.

Avec ces deux définitions, nous avons très facilement généralisé la notion d'intégrale lorsque les fonctions sont à valeurs dans \mathbb{R}^m . Tous les résultats d'intégrations valables pour des fonctions à valeurs dans \mathbb{R} pourront ainsi se transposer très facilement dans le cas général.

Chapitre 3

L'intégrale de Lebesgue munie de la mesure de Lebesgue

Dans ce chapitre, nous allons voir les principaux résultats d'intégration, lorsque l'on muni l'intégrale de Lebesgue de la mesure de Lebesgue. Dans la pratique, c'est souvent cette catégorie d'intégrale qui sera utilisée, puisqu'elle a un comportement très proche de l'intégrale de Riemann.

On rappelle que, par définition, la mesure de Lebesgue est définie sur l'espace mesurable $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ ou $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n))$. Nos fonctions prendront donc leurs valeurs dans \mathbb{R}^n .

Pour faciliter les notations, et souligner cette ressemblance, on notera désormais :

$$\int_A f(x) d\lambda(x) = \int_A f(x) dx$$

Deuxième partie

Probabilités

Chapitre 4

Généralités sur les probabilités

4.1 Premières définitions

Commençons par définir les notions fondamentales de probabilités, à l'aide du formalisme introduit dans la partie précédente.

Définition 4.1.1. On appelle **espace probabilisable** un espace mesurable, dans le contexte des probabilités.

Définition 4.1.2. Soit (Ω, \mathcal{A}) un espace probabilisable. Une **mesure de probabilité** (ou **loi de probabilité**) \mathbb{P} sur (Ω, \mathcal{A}) est une mesure sur (Ω, \mathcal{A}) , à valeurs dans $[0, 1]$, telle que $\mathbb{P}(\Omega) = 1$.

Définition 4.1.3. On appelle **espace probabilisé** un espace probabilisable muni d'une mesure de probabilité.

4.2 Variables aléatoires

Voyons maintenant ce qu'est une variable aléatoire. C'est une appellation étrange, puisqu'une variable aléatoire n'est en fait ni une variable, ni quelque chose d'aléatoire. En réalité, il s'agit simplement d'une fonction qui possède certaines propriétés (notamment la mesurabilité). Nous nous limitons ici à l'étude des variables aléatoires réelles, des vecteurs aléatoires et des variables aléatoires discrètes.

Dans toute cette section, on se donne un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On rappelle que $\mathcal{B}(E)$ désigne la tribu des boréliens de l'espace topologique E .

Définition 4.2.1. Une **variable aléatoire réelle** est une fonction mesurable

$$X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

Définition 4.2.2. Un **vecteur aléatoire** est une fonction mesurable

$$X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$$

Définition 4.2.3. Une **variable aléatoire discrète** est une fonction mesurable

$$X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow (\mathbb{N}, \mathcal{P}(\mathbb{N}))$$

On peut également introduire la définition suivante, qui ne sera pas utile en pratique, mais qui a le mérite d'unifier les trois définitions précédentes :

Définition 4.2.4. Plus généralement, si E est un espace topologique quelconque, une **variable aléatoire** est une fonction mesurable

$$X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow (E, \mathcal{B}(E))$$

Remarque 4.2.1. On remarquera que, exceptés l'espace d'arrivée et la tribu de l'espace d'arrivée, ces définitions sont identiques.

Remarque 4.2.2. Un vecteur aléatoire ne contient pas « variable » dans son nom, mais c'est bien une variable aléatoire.

Par ailleurs, on vérifie aisément qu'un vecteur aléatoire n'est qu'un vecteur de n variables aléatoires réelles. Ainsi, dès que l'on énoncera des résultats d'intégration pour des variables aléatoires réelles, on fera l'économie de ne pas les énoncer pour les vecteurs aléatoires (voir la section ??)

Remarque 4.2.3. On fera attention au fait que :

- une variable aléatoire est une fonction qui prend ses valeurs dans Ω ;
- une mesure de probabilité est une fonction qui prend ses valeurs dans \mathcal{A} .

Ainsi, si $\omega \in \Omega$:

- écrire $X(\omega)$ a du sens.
- écrire $\mathbb{P}(\omega)$ n'a pas de sens (ω n'a aucune chance d'être dans \mathcal{A} , puisque ce n'est même pas une partie de Ω).
- écrire $\mathbb{P}(\{\omega\})$ a du sens dès que $\{\omega\} \in \mathcal{A}$ (ie dès que $\{\omega\}$ est mesurable).
- écrire $X(\{\omega\})$ n'a pas de sens.

Proposition 4.2.1. Soit E et F deux espaces topologiques. Soit $X : \Omega \longrightarrow E$ une variable aléatoire à valeurs dans E . Soit $g : E \longrightarrow F$ une fonction borélienne. Alors $g \circ X$ est une variable aléatoire à valeurs dans F .

Remarque 4.2.4. Cette proposition est une conséquence immédiate du résultat de composition de fonctions mesurables.

4.2.1 Loi d'une variable aléatoire

Dans toute cette section, on se donne

- un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$;
- une variable aléatoire quelconque X à valeurs dans un espace topologique (E, \mathcal{B}) .

Ces définitions peuvent être données pour des variables aléatoires quelconques, mais seront utiles en pratique pour les trois types de variables aléatoires que nous avons vues.

Définition 4.2.5. La loi de probabilité de X \mathbb{P}_X est définie comme :

$$\mathbb{P}_X = \mathbb{P} \circ X^{-1}$$

C'est la mesure image de \mathbb{P} par X . Il s'agit d'une loi de probabilité sur (E, \mathcal{B})

Définition 4.2.6. Soit \mathbb{Q} une loi de probabilité sur (E, \mathcal{B}) . On dit que X **suit la loi** \mathbb{Q} lorsque $\mathbb{P}_X = \mathbb{Q}$.

Remarque 4.2.5. Attention, ce n'est pas parce que deux variables aléatoires ont la même loi qu'elles sont égales ! Voici un contre-exemple. Il utilise cependant des notions qui n'ont pas encore été abordées à ce stade.

Considérons l'espace probabilisé $\Omega := \{\text{beau temps}, \text{mauvais temps}\}$ muni de l'ensemble de ses parties, et de la mesure de probabilité \mathbb{P} uniforme. On a donc :

$$\mathbb{P}(\{\text{beau temps}\}) = \mathbb{P}(\{\text{mauvais temps}\}) = \frac{1}{2}$$

Puis posons la variable aléatoire discrète X à valeurs dans $\{0, 1\}$, telle que $X(\{\text{beau temps}\}) = 1$ et $X(\{\text{mauvais temps}\}) = 0$.

X suit alors la loi uniforme sur l'espace $\{0, 1\}$.

Considérons maintenant l'espace probabilisé $\Omega' := \{\text{temps chaud}, \text{temps froid}\}$ muni de l'ensemble de ses parties, et de la mesure de probabilité \mathbb{P}' uniforme. On a donc :

$$\mathbb{P}'(\{\text{temps chaud}\}) = \mathbb{P}'(\{\text{temps froid}\}) = \frac{1}{2}$$

Puis posons la variable aléatoire discrète Y à valeurs dans $\{0, 1\}$, telle que $Y(\{\text{temps chaud}\}) = 1$ et $Y(\{\text{temps froid}\}) = 0$.

Y suit également la loi uniforme sur l'espace $\{0, 1\}$.

Nous avons donc explicité deux variables aléatoires qui ont même loi, mais qui ne sont pas égales (elles n'ont pas le même espace de départ).

4.2.2 Moments d'une variable aléatoire réelle ou discrète

Dans toute cette section, on se donne

- un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$;
- une variable aléatoire réelle ou discrète.

On ne pourra pas donner de sens à ces notions dans le cas général d'une variable aléatoire quelconque. Nous verrons néanmoins par la suite comment généraliser ces notions aux vecteurs aléatoires.

Définition 4.2.7. Soit $r \in \mathbb{N}$. On dit que X **admet un moment d'ordre** r lorsque $\omega \mapsto X^r(\omega)$ est \mathbb{P} -intégrable.

Remarque 4.2.6. On rappelle que l'intégrale de Lebesgue est absolue. Ainsi, cette définition et les assertions suivantes sont équivalentes :

- $\omega \mapsto |X(\omega)|^r$ est intégrable.
- $\int_{\Omega} |X(\omega)|^r d\mathbb{P}(\omega) < +\infty$

Définition 4.2.8. Soit $r \in \mathbb{N}$. Lorsque X admet un moment d'ordre r , on définit son **moment d'ordre r** comme

$$m_r := \int_{\Omega} X^r(\omega) d\mathbb{P}(\omega)$$

Remarque 4.2.7. On vérifie aisément, par la relation mesure-intégrale, que toute variable aléatoire admet un moment d'ordre 0, et que celui-ci vaut toujours 1.

Définition 4.2.9. On dit que X **admet une espérance** lorsqu'elle admet un moment d'ordre 1 (*ie* lorsque $\omega \mapsto X(\omega)$ est \mathbb{P} -intégrable). Son **espérance** est alors définie comme son moment d'ordre 1. Elle est notée $\mathbb{E}(X)$.

On note $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ l'ensemble des variables aléatoires qui admettent une espérance (*ie* l'ensemble des variables aléatoires \mathbb{P} -intégrables).

Remarque 4.2.8. Par abus de notation, on pourra écrire $\mathbb{E}(X) = +\infty$ lorsque X n'admet pas d'espérance et qu'elle est positive. Ainsi, on pourra écrire $\mathbb{E}(X)$ dès que X est positive, qu'elle admette une espérance ou non. Vue la place privilégiée qu'occupent les fonctions positives dans la théorie d'intégration de Lebesgue, cette notation ne surprend pas.

Attention cependant, si X n'admet pas d'espérance et n'est pas positive, on s'interdira formellement d'écrire $\mathbb{E}(X)$ (car l'intégrale sous-jacente n'existe peut-être même pas!).

Remarque 4.2.9. Ces considérations étant faites, on remarque que :

- X admet une espérance ssi $\mathbb{E}(|X|) < +\infty$;
- X admet un moment d'ordre r ssi $\mathbb{E}(|X|^r) < +\infty$;
- si X admet un moment d'ordre r , alors $m_r = \mathbb{E}(X^r)$.

Proposition 4.2.2. Soit $r \in \mathbb{N}$. Si X admet un moment d'ordre r , alors elle admet un moment d'ordre k pour tout $k \in \llbracket 0, r \rrbracket$. Autrement dit, si X^r est \mathbb{P} -intégrable, alors X^k est \mathbb{P} -intégrable pour tout $k \in \llbracket 0, r \rrbracket$.

Démonstration. Soit $k \in \llbracket 0, r \rrbracket$. Nous avons $|X|^k \leq |X|^r + 1$, donc par croissance et linéarité de l'intégrale :

$$\int_{\Omega} |X(\omega)|^k d\mathbb{P}(\omega) \leq \int_{\Omega} |X(\omega)|^r d\mathbb{P}(\omega) + \int_{\Omega} d\mathbb{P}(\omega)$$

Comme X admet un moment d'ordre r , le premier terme de la somme est fini. Puis, d'après la relation mesure-intégrale, le deuxième terme de la somme vaut 1 (\mathbb{P} est une mesure de probabilité).

Le terme de gauche est donc fini, donc X admet un moment d'ordre k . \square

Remarque 4.2.10. Attention, on remarquera que ceci n'est vrai que parce que \mathbb{P} est une mesure de probabilité! Dans le cas d'une mesure quelconque (par

exemple, la mesure de Lebesgue), cette démonstration ne permet pas de conclure (car $\lambda(\mathbb{R}) = +\infty$)... et heureusement, car ce résultat est faux !

Voici un contre-exemple :

$$x \mapsto \mathbb{1}_{\mathbb{R}_+^*}(x) \frac{1}{x}$$

n'est pas intégrable sur \mathbb{R} , alors que son carré l'est.

Définition 4.2.10. On dit que X **admet une variance** lorsqu'elle admet un moment d'ordre 2 (ie lorsque $\omega \mapsto X^2(\omega)$ est \mathbb{P} -intégrable). Sa **variance** est alors définie comme :

$$\mathbb{V}(X) := \mathbb{E} \left((X - \mathbb{E}(X))^2 \right) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

On note $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ l'ensemble des variables aléatoires qui admettent une variance (ie l'ensemble des variables aléatoires qui sont de carré \mathbb{P} -intégrables).

Remarque 4.2.11. Cette définition impose de faire trois remarques :

- le théorème qui fournit l'égalité de ces deux expressions s'appelle **le théorème de König-Huygens**. Sa démonstration est immédiate par linéarité de l'intégrale.
- d'après la deuxième expression, l'existence de $\mathbb{V}(X)$ est assurée dès que X admet un moment d'ordre 1 et 2. Cependant, bien que la définition demande que X admette un moment d'ordre 2, elle ne demande pas qu'elle admette un moment d'ordre 1. Nous sommes en fait sauvés par la proposition 4.2.2 !
- d'après la proposition 4.2.2, $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P}) \subset \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$. A nouveau, on rappelle que ceci n'est vrai que parce que \mathbb{P} est une mesure de probabilité (voir la remarque 4.2.10)

4.2.3 Formule de transfert

Dans toute cette section, on se donne

- un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$;
- un espace topologique quelconque E ;
- une variable aléatoire quelconque $X : \Omega \longrightarrow E$.

La formule de transfert peut s'exprimer de deux manières (selon si l'on fait face à des variables aléatoires réelles ou discrètes). On pourrait les exprimer en un seul théorème, mais les séparer en deux permet de bien repérer dans quels cas le théorème peut s'appliquer.

Théorème 4.2.1. Soit $g : E \longrightarrow \mathbb{R}$ une fonction borélienne. Par la proposition 4.2.1, $g \circ X$ est encore une variable aléatoire ; c'est d'ailleurs une variable aléatoire réelle. On sait donc donner un sens à son espérance.

La variable aléatoire réelle $g \circ X$ sur $(\Omega, \mathcal{A}, \mathbb{P})$ admet une espérance (ie $\omega \mapsto (g \circ X)(\omega)$ est \mathbb{P} -intégrable) ssi la variable aléatoire réelle g sur $(E, \mathcal{B}(E), \mathbb{P}_X)$ admet une espérance (ie $x \mapsto g(x)$ est \mathbb{P}_X -intégrable).

Si c'est le cas, alors on a la **formule de transfert** :

$$\mathbb{E}(g \circ X) = \int_E g(x) d\mathbb{P}_X(x)$$

Théorème 4.2.2. Soit $g : E \longrightarrow \mathbb{N}$ une fonction borélienne. Par la proposition 4.2.1, $g \circ X$ est encore une variable aléatoire ; c'est d'ailleurs une variable aléatoire discrète. On sait donc donner un sens à son espérance.

La variable aléatoire discrète $g \circ X$ sur $(\Omega, \mathcal{A}, \mathbb{P})$ admet une espérance (ie $\omega \mapsto (g \circ X)(\omega)$ est \mathbb{P} -intégrable) ssi la variable aléatoire discrète g sur $(E, \mathcal{B}(E), \mathbb{P}_X)$ admet une espérance (ie $x \mapsto g(x)$ est \mathbb{P}_X -intégrable).

Si c'est le cas, alors on a la **formule de transfert** :

$$\mathbb{E}(g \circ X) = \int_E g(x) d\mathbb{P}_X(x)$$

Remarque 4.2.12. On comprend pourquoi cette formule porte ce nom, lorsque l'on remplace l'espérance par sa définition, dans les expressions précédentes :

$$\int_{\Omega} (g \circ X)(\omega) d\mathbb{P}(\omega) = \int_E g(x) d\mathbb{P}_X(x)$$

Cette formule transfère le calcul sur un autre espace, muni d'une autre mesure. Pour peu que la mesure \mathbb{P}_X soit facile à utiliser, on a intérêt à utiliser cette nouvelle formulation.

Voici quelques applications de ces théorèmes :

Exemple 4.2.1. Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire réelle. La fonction identité $\text{Id} : \mathbb{R} \longrightarrow \mathbb{R}$ est continue donc borélienne.

En appliquant le théorème précédent, X admet une espérance ssi Id est \mathbb{P}_X -intégrable. Dans ce cas, on aura :

$$\mathbb{E}(X) = \mathbb{E}(\text{Id} \circ X) = \int_{\mathbb{R}} x d\mathbb{P}_X(x)$$

Pour peu que l'on sache expliciter \mathbb{P}_X , cette nouvelle l'expression de l'espérance nous simplifie grandement le calcul.

Exemple 4.2.2. On obtient le même résultat avec une variable aléatoire discrète.

Exemple 4.2.3. De même, en choisissant la fonction carrée pour g on obtient (si l'espérance de X^2 existe) :

$$\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 d\mathbb{P}_X(x)$$

Notons qu'en choisissant l'identité pour g , mais la variable aléatoire X^2 , on obtient (à nouveau sous réserve d'existence) :

$$\mathbb{E}(X^2) = \int_{\mathbb{R}} x d\mathbb{P}_{X^2}(x)$$

4.2.4 Moments d'un vecteur aléatoire