

# PROBABILITES

Anito Kodama

23 novembre 2024

Ce document PDF n'a pas pour vocation à être vendu ni diffusé à des fins commerciales. Il s'agit de notes personnelles, combinées avec des extraits et des réflexions basés sur des lectures. Toute utilisation ou diffusion doit respecter le caractère privé et non commercial de ces contenus. Les idées présentées ici peuvent ne pas être complètes ni totalement exactes, car elles sont le fruit de notes prises à titre personnel.

## Table des matières

<b>1</b>	<b>Les outils pour la proba</b>	<b>3</b>
1.1	Grand $\mathcal{O}$	3
1.2	Petit $o$	3
1.3	Les suites	3
1.3.1	Théorème de Cesàro	3
1.4	Algèbre et tribu ❤️	3
1.4.1	Définitions	3
1.4.2	Éléments générateurs	3
1.4.3	Tribu Borélienne	4
1.4.4	Espace produit	4
1.5	Mesurabilité	4
1.6	Ensembles de fonctions mesurables	5
1.7	Fonction étagée	5
1.8	Classes monotones ???	6
1.9	Mesure	6
1.9.1	Mesure de dirac	7
1.9.2	Mesure de comptage	7
1.9.3	Mesure image	7
1.9.4	Mesure de Lebesgue 😊	7
1.10	Intégration	7
1.10.1	Trucs simples	7
1.10.2	Intégrale d'une fonctions positives	8
1.10.3	Intégrale d'une fonctions quelconques	8
1.10.4	Convergence monotone de Beppo Levi	8
1.10.5	Convergence dominée de Lebesgue	8
1.10.6	Inégalité de Jensen	8
1.10.7	Théorème de Radon-Nikodym	8
1.10.8	Intégration par rapport à une mesure image	8
1.10.9	Théorème de Fubini-Tonelli	8
1.10.10	Espaces $L^p$	8

2	Espace probabilisé	9
3	Espérance conditionnelle	9

## Preamble

Si on considère un ensemble  $X$  et une tribu  $\mathcal{A}$  sur  $X$ , on nomme alors la paire  $(X, \mathcal{A})$  un **espace mesurable** ou **probabilisable** dans notre cas.

Pour quantifier le « poids » de chaque événement de  $\mathcal{A}$ , on introduit la notion de **probabilité**.

## 1 Les outils pour la proba

### 1.1 Grand $\mathcal{O}$

### 1.2 Petit $o$

### 1.3 Les suites

#### 1.3.1 Théorème de Cesàro

**Proposition 1.** Soit  $(a_n)_n$  une suite de nombres réels ou complexes. Si elle converge vers  $l$ , alors la suite de ses moyennes de Cesàro, de terme general

$$c_n = \frac{1}{n} \sum_{k=1}^n a_k,$$

converge également vers  $l$ .

**Exercice 1.**  Trouver le terme général de la suite  $u_{n+1} = \sin(u_n)$ , avec  $u_0 \in ]0, 1]$ .

### 1.4 Algèbre et tribu

#### 1.4.1 Définitions

**Définition 1** (Algèbre de Boole).

**Définition 2** ( $\sigma$ -Algèbre (Tribu)).

#### 1.4.2 Éléments générateurs

En général, il est difficile d'expliciter tous les éléments d'une tribu. Les algèbres et les tribus se décrivent le plus souvent par leurs éléments **générateurs**.

**Définition 3** (Générateurs).

On peut parler de la tribu engendrée par deux tribus  $\mathcal{A}_1$  et  $\mathcal{A}_2$ , que l'on note

$$\mathcal{A}_1 \vee \mathcal{A}_2 = \sigma(\mathcal{A}_1 \cup \mathcal{A}_2).$$

**Exemple 1.** Soit  $A$  une partie de  $\Omega$ . L'algèbre  $\mathbb{C}(\{A\})$  et la tribu  $\sigma(\{A\})$  sont  $\{\emptyset, \Omega, A, A^c\}$

### 1.4.3 Tribu Borélienne

**Définition 4** (Tribu borélienne). Si  $\Omega$  est un espace topologique, on appelle tribu borélienne, notée  $\mathcal{B}(\Omega)$ , la tribu engendrée par les ouverts de  $\Omega$ . Un borélien est un ensemble à la tribu borélienne.



La tribu borélienne est aussi engendrée par les fermés puisque la tribu est stable par passage au complémentaire.

### 1.4.4 Espace produit

## 1.5 Mesurabilité

En mathématique lorsqu'une structure est définie sur un espace, on souhaite pouvoir la **transporter** sur d'autres espaces par des fonctions. En général, on utilise d'ailleurs les **images réciproques** par les fonctions.

**Exemple 2.** Par exemple :

1. Sur  $\mathbb{R}$ , la structure d'ordre est préservée par la réciproque d'une application **croissante** : Si  $x < y$  sont dans l'image de  $\mathbb{R}$  par une fonction  $f$  croissante, alors  $f^{-1}(x) < f^{-1}(y)$ .
2. La structure topologique est préservée par une application de la réciproque d'une application **continue** :  $f$  est continue si  $f^{-1}(U)$  est ouvert pour tout ouvert  $U$ .

La notion analogue dans le contexte de la théorie de la mesure est celle de **mesurabilité**.

**Définition 5** (Fonctions mesurables). Pour rappel le couple  $(\Omega, \mathcal{A})$  formé d'un ensemble  $\Omega$  et d'une tribu  $\mathcal{A}$  est un **espace mesurable**. Les éléments de  $\mathcal{A}$  sont appelés **ensembles mesurables**.

- (i) Soit  $(\Omega, \mathcal{A})$  et  $(E, \mathcal{B})$ , deux espaces mesurables. Soit  $f$  une fonction de  $\Omega$  dans  $E$ . On dit que  $f$  est mesurable (pour  $\mathcal{A}$  et  $\mathcal{B}$ ) si

$$f^{-1}(\mathcal{B}) \subset \mathcal{A}.$$

C'est à dire,  $f^{-1}(B) \in \mathcal{A}$  pour tout  $B \in \mathcal{B}$ .

**Définition 6** (Fonctions borélienne). Une fonction mesurable de  $(\Omega, \mathcal{A})$  dans un espace topologique muni de sa tribu borélienne  $(E, \mathcal{B}(E))$  est dite borélienne.

La proposition suivante montre que pour qu'une fonction soit **mesurable**, il suffit de vérifier sa propriété caractéristique sur une famille génératrice de la tribu d'arrivée.

**Proposition 2** (Caractéristique d'une fonction mesurable). Soit  $\Omega$  et  $E$  deux ensembles. Soit  $\mathcal{E} \subset \mathcal{P}(E)$  et soit  $\mathcal{B} = \sigma(\mathcal{E})$ .

La tribu engendrée par une fonction  $f$  de  $\Omega$  dans  $(E, \mathcal{B})$  est

$$\sigma(f) = \sigma(f^{-1}(\mathcal{E})) = \sigma(\{f^{-1}(C) : C \in \mathcal{E}\}).$$

Plus généralement, si  $\mathcal{F}$  est une famille de fonctions de  $\Omega$  dans  $(E, \mathcal{B})$ , alors

$$\sigma(\mathcal{F}) = \sigma(\{f^{-1}(C) : C \in \mathcal{E}; f \in \mathcal{F}\}).$$

En particulier, pour qu'une fonction  $f$  de  $(\Omega, \mathcal{A})$  dans  $(E, \sigma(\mathbb{E}))$  soit mesurable, il suffit que  $f^{-1}(\mathbb{E})$  soit inclus dans  $\mathcal{A}$ .

**Démonstration :** Soit

$$\mathcal{T} = \{B \subset E : f^{-1}(B) \in \sigma(f^{-1}(\mathbb{E}))\}$$

**Proposition 3** (Théorème fondamental de la mesurabilité). Soit  $f$  une application quelconque d'un ensemble  $\Omega$  dans un ensemble  $\Omega'$ . Alors

1. Pour toute tribu  $\mathcal{A}$  sur  $\Omega'$ ,  $f^{-1}(\mathcal{A})$  est une tribu sur  $\Omega$ , où  $f^{-1}(\mathcal{A}) = \{f^{-1}(A) : A \in \mathcal{A}\}$
2. Pour tout  $\mathcal{A} \in \mathcal{P}(\mathcal{P}(\Omega'))$ , on a  $\sigma(f^{-1}(\mathcal{A})) = f^{-1}(\sigma(\mathcal{A}))$ .

On appelle  $\sigma(f^{-1}(\mathcal{A})) = f^{-1}(\sigma(\mathcal{A}))$  la tribu image de  $\mathcal{A}$  par  $f$ .

## 1.6 Ensembles de fonctions mesurables

**Proposition 4.** La composée de deux fonctions mesurables est mesurable.

**Démonstration :**

**Lemma 1.** Si  $f, g$  sont des fonctions mesurables de  $(\Omega, \mathcal{A})$  dans  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , alors  $\omega \in \Omega \rightarrow (f(\omega), g(\omega)) \in \mathbb{R}^2$  est mesurable de  $(\Omega, \mathcal{A})$  dans  $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ .

**Démonstration :**

**Proposition 5.** Soit  $\Omega_1, \Omega_2$  deux espaces topologiques munis de leurs tribu borélienne. Toute fonction **continue** de  $\Omega_1$  dans  $\Omega_2$  est **mesurable** X (ou borélienne) 🙌.

**Démonstration :**



Pour  $x, y \in \mathbb{R}$  on note leur **maximum** :  $x \vee y$

**Proposition 6.** L'espace des fonctions mesurables (boréliennes) de  $(\Omega, \mathcal{A})$  dans  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  est stable pour les opérations de :

1. **multiplication par une constante** :  $(\lambda f)(\omega) = \lambda f(\omega)$  pour  $\lambda \in \mathbb{R}$  ;
2. **addition** :  $(f + g)(\omega) = f(\omega) + g(\omega)$  ;
3. **multiplication** :  $(fg)(\omega) = f(\omega)g(\omega)$  ;
4. **maximum** :  $(f \vee g)(\omega) = f(\omega) \vee g(\omega)$ .

**Démonstration :**

## 1.7 Fonction étagée

On peut approcher **toute** fonction mesurable par des fonctions mesurables plus simples.

**Définition 7** (Fonction étagée). Soit  $(\Omega, \mathcal{A})$  un espace mesurable. On appelle fonction étagée (à valeurs dans  $\mathbb{R}^d$ ) une fonction de la forme :

$$f(\omega) = \sum_{i=1}^k a_i \mathbb{1}_{A_i}(\omega)$$

où les  $A_i$  sont des éléments **disjoints** de  $\mathcal{A}$ , et où les coefficients  $a_i$  appartiennent à  $\mathbb{R}^d$ .

**Proposition 7.** Toute fonction  $f$  mesurable de  $(\Omega, \mathcal{A})$  dans  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  est **limite simple** de fonctions étagées.

Si  $f$  est positive, la limite peut être choisie croissante.

**Démonstration :** Prenons d'abord  $f$  positive. Définissons pour  $n, k \geq 1$ ,

$$A_{n,k} = \left\{ \omega : \frac{k-1}{2^n} \leq f(\omega) \leq \frac{k}{2^n} \right\}$$

Les  $A_{n,k}$  sont éléments de  $\mathcal{A}$  en tant qu'images réciproques par la fonction mesurable  $f$  d'intervalles.

La suite

$$f_n(\omega) = \sum_{k=1}^{2^n} \frac{k-1}{2^n} \mathbb{1}_{A_{n,k}}(\omega)$$

converge en croissant vers  $f$ .

Si  $f$  est quelconque, écrivons  $f = f^+ - f^-$  avec  $f^+ = f \vee 0$  et  $f^- = (-f) \vee 0$ , et approximations les fonctions positives  $f^+$  et  $f^-$  par la méthode précédente. ■

Autre formulation :

**Lemme 2 (Lemme d'approximation).** Soit  $f$  une fonction mesurable de  $(\Omega, \mathcal{A})$  dans  $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$  avec  $\mathbb{K} = \mathbb{R}, \bar{\mathbb{R}}, \mathbb{C}$ . Il existe une suite de fonctions étagées  $(f_n)_n$  telle que

$$f(\omega) = \lim_{n \rightarrow \infty} f_n(\omega), \text{ pour tout } \omega \in \Omega.$$

Si  $f \geq 0$ ,  $(f_n)_n$  peut être choisie croissante et positive, i.e  $0 \leq f_n \leq f_{n+1}$  pour tout  $n \in \mathbb{N}^*$ .



Pour prouver une propriété concernant l'intégrale d'une variable aléatoire positive, la recette est toujours la même :

1. Commencer par la prouver pour une indicatrice.
2. L'obtenir pour les fonctions étagées positives, par linéarité de l'intégrale sur l'ensemble de ces fonctions.
3. Utiliser alors le lemme d'approximation pour écrire toute fonction mesurable dans  $\bar{\mathbb{R}}_+$  comme limite croissante de fonction étagées.
4. Obtenir la propriété souhaitée pour les fonctions mesurables positives grâce au théorème de la convergence monotone.

## 1.8 Classes monotones ???

à voir

## 1.9 Mesure

**Définition 8 (Axiomes).**

**Proposition 8.**

### 1.9.1 Mesure de dirac

Soit  $(\Omega, \mathcal{F})$  un espace muni d'une tribu. Soit  $x \in \Omega$ . On appelle mesure de Dirac (ou masse de Dirac) en  $x$  et on note  $\delta_x$  la mesure définie par :

$$\forall A \in \mathcal{F}, \delta_x(A) = \mathbb{1}_A(x).$$



La vérification du fait que  $\delta_x$  est une mesure est évidente mais utile à faire une fois dans sa vie 😊.

### 1.9.2 Mesure de comptage

Soit  $(\Omega, \mathcal{F})$  un espace mesurable. On appelle mesure de comptage sur  $\Omega$  la mesure  $C$  définie par :

$$\forall A \in \mathcal{F}, C(A) = |A|,$$

Où  $|A|$  est le cardinal de  $A$ .



La vérification du fait que  $C$  est une mesure est évidente mais utile à faire une fois dans sa vie 😊.

### 1.9.3 Mesure image

**Définition 9** (Mesure image).

### 1.9.4 Mesure de Lebesgue 🧐

**Définition 10.** La mesure de Lebesgue est l'**unique** mesure définie sur  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , stable par **translation**, et telle que  $\lambda([0, 1]) = 1$ . Elle existe, et la mesure de Lebesgue d'un intervalle correspond à sa **longueur**.

## 1.10 Intégration

### 1.10.1 Trucs simples



#### Indicatrice

Soit  $A \in \mathcal{A}$ , nous définissons

$$\int_{\Omega} \mathbb{1}_A d\mu =: \mu(A).$$

Pour  $\mu$  une mesure positive.

## Fonctions étagées

Pour rappel une fonction de  $(\Omega, \mathcal{A})$  dans  $(\mathbb{K}, \mathcal{B}(\mathbb{K}))$  est dite étagée si elle prend un nombre fini de valeurs notées  $(a_i)_i$ , les  $a_i$  étant supposés tous distincts. Elle s'écrit donc :

$$f = \sum_{k=1}^n a_k \mathbb{1}_{A_k},$$

où  $A_k = f^{-1}(\{a_k\})$ .

Dans le cas où  $\Omega$  est fini ou dénombrable, nous définissons ainsi :

$$\int_{\Omega} f d\mu := \sum_{k=1}^n a_k \int_{\Omega} \mathbb{1}_{A_k} d\mu = \sum_{k=1}^n a_k \mu(A_k)$$

Dans le cas où  $\mu$  est une mesure de probabilité notée  $\mathbb{P}$ , l'équation précédente se réécrit pour  $X = \sum_{k=1}^n a_k \mathbb{1}_{X=a_k}$

$$\mathbb{E}[X] = \sum_{k=1}^n a_k \mathbb{P}(\{X = a_k\}) = \sum_{k=1}^n a_k \mathbb{P}_X(\{a_k\}).$$



L'espérance ne dépend de  $X$  qu'au travers de sa loi.  
Ainsi, deux variables aléatoires ayant la même loi,  
même potentiellement "très différente" ont même  
espérance !!

### 1.10.2 Intégrale d'une fonctions positives

### 1.10.3 Intégrale d'une fonctions quelconques

### 1.10.4 Convergence monotone de Beppo Levi

### 1.10.5 Convergence dominée de Lebesgue

### 1.10.6 Inégalité de Jensen

### 1.10.7 Théorème de Radon-Nikodym

### 1.10.8 Intégration par rapport à une mesure image

**Proposition 9** (de transport).

### 1.10.9 Théorème de Fubini-Tonelli

### 1.10.10 Espaces $L^p$

**Proposition 10** (Inégalité de Hölder).

**Proposition 11** (Inégalité de Minkowski).

**Proposition 12.** Pour tout  $p \geq 1$ , l'espace  $L^p$  est complet.



## 2 Espace probabilisé

Les choses nouvelles commencent 🤔

On appelle probabilité, ou mesure de probabilité, ou loi sur  $(\Omega, \mathcal{F})$  toute application

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$$

vérifiant les propriétés suivantes :

1.  $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$  ;
2. Pour toute suite  $(A_i)_{i \geq 1}$  d'éléments de  $\mathcal{F}$  deux à deux disjoints, on a

$$\mathbb{P} \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Le triplet  $(\Omega, \mathcal{F}, \mathbb{P})$  est alors appelé **espace probabilisé**.



Ce qui fait qu'un espace probabilisé est très exactement un espace mesuré associé à une mesure positive de masse totale 1 !!



Dans le contexte des probabilités, on appelle événement tout élément de la tribu  $\mathcal{F}$ .

## 3 Espérance conditionnelle

**Définition 11** (Espérance conditionnelle). L'espérance conditionnelle de  $X$  sachant  $\mathcal{F}$  notée  $\mathbb{E}[X | \mathcal{F}]$  est définie par :

1.  **$\mathcal{F}$ -mesurabilité** :  $\mathbb{E}[X | \mathcal{F}]$  est  $\mathcal{F}$ -mesurable ;
2. **Propriété orthogonale** :  $\forall Z$   $\mathcal{F}$ -mesurable,  $\mathbb{E}[(X - \mathbb{E}[X | \mathcal{F}]) \times Z] = 0$  si et seulement si  $\mathbb{E}[X \times Z] = \mathbb{E}[\mathbb{E}[X | \mathcal{F}] \times Z]$ .



Cette deuxième propriété signifie que l'écart entre  $X$  et son espérance conditionnelle  $\mathbb{E}[X | \mathcal{F}]$  est indépendant (par analogie avec l'algèbre linéaire on pourrait dire "orthogonal") à toute information contenue dans la tribu  $\mathcal{F}$ , i.e que  $\mathbb{E}[X | \mathcal{F}]$  est la meilleure approximation de  $X$  à partir des informations contenues dans  $\mathcal{F}$  et ce qui reste de  $X$  après avoir enlevé  $\mathbb{E}[X | \mathcal{F}]$  n'est pas corrélé avec aucune information dans  $\mathcal{F}$ .


👉 En fait pour revenir de nouveau sur la deuxième propriété de la définition on peut dire qu'elle découle de l'idée que  $X$  peut être décomposé en deux parties :

$$X = \mathbb{E}[X | \mathcal{F}] + \text{residu}$$


où le résidu  $(X - \mathbb{E}[X | \mathcal{F}])$  est "orthogonal" à tout événement mesurable dans  $\mathcal{F}$ . Cela garantit que  $\mathbb{E}[X | \mathcal{F}]$  capture toute l'information "reliée" à  $\mathcal{F}$ , et le résidu n'est plus corrélé à  $\mathcal{F}$ .

**Proposition 13.** On a alors les trois propriétés suivantes :

1. Si  $X$  est  $\mathcal{F}$ -mesurable,  $\mathbb{E}[X \mid \mathcal{F}] = X$  ;
2.  $\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}]] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{F}] \times 1] = \mathbb{E}[X \times 1] = \mathbb{E}[X]$  ;
3. Si  $\mathcal{F}_1 \subset \mathcal{F}_2$ , on a  $\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}_1] \mid \mathcal{F}_2] = \mathbb{E}[X \mid \mathcal{F}_1]$ .

**Démonstration esquisse :** 1. Nous n'avons pour l'heure rien d'autre que la définition précédente donc on la regarde  ; on a dans un premier temps, par hypothèse,  $X$   $\mathcal{F}$ -mesurable, donc déjà  $X$  satisfait la première condition de l'espérance conditionnelle, ensuite on doit vérifier "l'orthogonalité", *i.e* vérifier si pour tout  $Z$   $\mathcal{F}$ -mesurable, on a  $\mathbb{E}[(X - \mathbb{E}[X \mid \mathcal{F}]) \times Z] = 0$ , cela nous donne  $\mathbb{E}[0 \times Z] = \mathbb{E}[0] = 0$ , donc  $X$  satisfait également la deuxième condition de la définition.



En fait en y pensant bien, l'espérance conditionnelle  $\mathbb{E}[X \mid \mathcal{F}] = X$  peut être vue comme une "approximation" de  $X$  en utilisant uniquement les informations contenues dans la tribu  $\mathcal{F}$ . Donc si  $X$  est déjà entièrement déterminé par  $\mathcal{F}$  (*i.e* si  $X$  est  $\mathcal{F}$ -mesurable), alors cette approximation est parfaite, il n'y a aucune perte d'information et donc on a l'égalité! 

2. La deuxième propriété exprime que la moyenne globale d'une espé

