Fondements des probabilités

À l'intention des étudiants de STT-1500



Jérôme Soucy

Département de mathématiques et de statistique
Université Laval

Hiver 2025

Résumé

1. Pourquoi revenir aux fondements ?
2. Tribus d'événements
3. Les axiomes de Kolmogorov
4. Conséquences des axiomes
5. Continuité des probabilités

1

Pourquoi revenir aux fondements ?

Ce que nous savons faire

Ce que nous savons faire

Dans le module d'introduction, une probabilité sur un espace fondamental *discret* Ω était construite à partir de poids :

$$p(\omega) \geq 0, \qquad \sum_{\omega \in \Omega} p(\omega) = 1, \qquad \mathbb{P}(A) = \sum_{\omega \in A} p(\omega).$$

Ce que nous savons faire

Dans le module d'introduction, une probabilité sur un espace fondamental *discret* Ω était construite à partir de poids :

$$p(\omega) \geq 0, \qquad \sum_{\omega \in \Omega} p(\omega) = 1, \qquad \mathbb{P}(A) = \sum_{\omega \in A} p(\omega).$$

Nous en avons déduit les règles de calcul suivantes :

- $\mathbb{P}(A) \geq 0$ et $\mathbb{P}(\Omega) = 1$;

Ce que nous savons faire

Dans le module d'introduction, une probabilité sur un espace fondamental *discret* Ω était construite à partir de poids :

$$p(\omega) \geq 0, \qquad \sum_{\omega \in \Omega} p(\omega) = 1, \qquad \mathbb{P}(A) = \sum_{\omega \in A} p(\omega).$$

Nous en avons déduit les règles de calcul suivantes :

- $\mathbb{P}(A) \geq 0$ et $\mathbb{P}(\Omega) = 1$;
- si $A_1, A_2, \ldots$ sont deux à deux incompatibles, $\mathbb{P}\big(\bigcup_k A_k\big) = \sum_k \mathbb{P}(A_k)$.

Ce que nous savons faire

Dans le module d'introduction, une probabilité sur un espace fondamental *discret* Ω était construite à partir de poids :

$$p(\omega) \geq 0, \qquad \sum_{\omega \in \Omega} p(\omega) = 1, \qquad \mathbb{P}(A) = \sum_{\omega \in A} p(\omega).$$

Nous en avons déduit les règles de calcul suivantes :

- $\mathbb{P}(A) \geq 0$ et $\mathbb{P}(\Omega) = 1$;
- si $A_1, A_2, \ldots$ sont deux à deux incompatibles, $\mathbb{P}\big(\bigcup_k A_k\big) = \sum_k \mathbb{P}(A_k)$.

Toutes les autres propriétés (complémentarité, probabilité d'une union, monotonicité, Boole, Bonferroni) ont été obtenues à partir de ces règles seulement, sans jamais revenir aux poids.

Ce que nous ne savons pas encore faire

Ce que nous ne savons pas encore faire

On choisit un nombre au hasard dans $\Omega = [0, 1]$.

Ce que nous ne savons pas encore faire

On choisit un nombre au hasard dans $\Omega = [0, 1]$.

- Chaque nombre x doit avoir une probabilité nulle : $\mathbb{P}(\{x\}) = 0$.

Ce que nous ne savons pas encore faire

On choisit un nombre au hasard dans $\Omega = [0, 1]$.

- Chaque nombre x doit avoir une probabilité nulle : $\mathbb{P}(\{x\}) = 0$.
- Pourtant, on s'attend à ce que $\mathbb{P}([a, b]) = b - a$ pour $0 \leq a \leq b \leq 1$.

Ce que nous ne savons pas encore faire

On choisit un nombre au hasard dans $\Omega = [0, 1]$.

- Chaque nombre x doit avoir une probabilité nulle : $\mathbb{P}(\{x\}) = 0$.
- Pourtant, on s'attend à ce que $\mathbb{P}([a, b]) = b - a$ pour $0 \leq a \leq b \leq 1$.
- Les poids ne permettent pas de retrouver la probabilité d'un intervalle : une somme de zéros vaut zéro.

Ce que nous ne savons pas encore faire

On choisit un nombre au hasard dans $\Omega = [0, 1]$.

- Chaque nombre x doit avoir une probabilité nulle : $\mathbb{P}(\{x\}) = 0$.
- Pourtant, on s'attend à ce que $\mathbb{P}([a, b]) = b - a$ pour $0 \leq a \leq b \leq 1$.
- Les poids ne permettent pas de retrouver la probabilité d'un intervalle : une somme de zéros vaut zéro.

Idée de Kolmogorov (1933) : oublier les poids et prendre comme *définition* les règles de calcul que l'on veut voir respectées. Une probabilité devient une fonction définie directement sur les événements.

Un peu d’histoire

Un peu d'histoire

- En 1900, au Congrès international des mathématiciens, David Hilbert propose 23 problèmes.

Un peu d'histoire

- En 1900, au Congrès international des mathématiciens, David Hilbert propose 23 problèmes. Le sixième demande, entre autres, une axiomatisation du calcul des probabilités.

Un peu d'histoire

- En 1900, au Congrès international des mathématiciens, David Hilbert propose 23 problèmes. Le sixième demande, entre autres, une axiomatisation du calcul des probabilités.
- En 1933, Andreï Kolmogorov publie *Grundbegriffe der Wahrscheinlichkeitsrechnung* (*Fondements de la théorie des probabilités*).

Un peu d'histoire

- En 1900, au Congrès international des mathématiciens, David Hilbert propose 23 problèmes. Le sixième demande, entre autres, une axiomatisation du calcul des probabilités.
- En 1933, Andreï Kolmogorov publie *Grundbegriffe der Wahrscheinlichkeitsrechnung* (*Fondements de la théorie des probabilités*).
- Il s'appuie sur la théorie de la mesure (Borel, Lebesgue), développée au début du xxe siècle pour mesurer longueurs, aires et volumes d'ensembles très généraux.

Un peu d'histoire

- En 1900, au Congrès international des mathématiciens, David Hilbert propose 23 problèmes. Le sixième demande, entre autres, une axiomatisation du calcul des probabilités.
- En 1933, Andreï Kolmogorov publie *Grundbegriffe der Wahrscheinlichkeitsrechnung* (*Fondements de la théorie des probabilités*).
- Il s'appuie sur la théorie de la mesure (Borel, Lebesgue), développée au début du xxe siècle pour mesurer longueurs, aires et volumes d'ensembles très généraux.



Une probabilité est une manière de « mesurer » les événements, comme la longueur mesure les intervalles, avec la contrainte que la mesure totale vaut 1.

2

Tribus d’événements

Quels ensembles sont des événements ?

Quels ensembles sont des événements ?

Lorsque Ω est discret, tout sous-ensemble de Ω est un événement : on peut lui attribuer une probabilité.

Quels ensembles sont des événements ?

Lorsque Ω est discret, tout sous-ensemble de Ω est un événement : on peut lui attribuer une probabilité.

Lorsque Ω n'est pas dénombrable, on peut démontrer qu'il est *impossible* d'attribuer une « longueur » à *tous* les sous-ensembles de $[0, 1]$ en respectant les règles de calcul habituelles (résultat de Vitali, 1905).

Quels ensembles sont des événements ?

Lorsque Ω est discret, tout sous-ensemble de Ω est un événement : on peut lui attribuer une probabilité.

Lorsque Ω n'est pas dénombrable, on peut démontrer qu'il est *impossible* d'attribuer une « longueur » à *tous* les sous-ensembles de $[0, 1]$ en respectant les règles de calcul habituelles (résultat de Vitali, 1905).

> On se restreint donc à une famille $\mathcal{F}$ de sous-ensembles de Ω, appelés **événements**.

Quels ensembles sont des événements ?

Lorsque Ω est discret, tout sous-ensemble de Ω est un événement : on peut lui attribuer une probabilité.

Lorsque Ω n'est pas dénombrable, on peut démontrer qu'il est *impossible* d'attribuer une « longueur » à *tous* les sous-ensembles de $[0, 1]$ en respectant les règles de calcul habituelles (résultat de Vitali, 1905).

> On se restreint donc à une famille $\mathcal{F}$ de sous-ensembles de Ω, appelés **événements**. Cette famille doit être assez riche pour que les opérations ensemblistes (complémentaire, union, intersection) appliquées à des événements donnent encore des événements.

Tribu

Tribu

Définition. Soit Ω un ensemble. Une **tribu** (ou **σ-algèbre**) sur Ω est une famille $\mathcal{F}$ de sous-ensembles de Ω telle que

Le couple $(\Omega, \mathcal{F})$ est appelé un **espace probabilisable**. Les éléments de $\mathcal{F}$ sont les **événements**.

Tribu

Définition. Soit Ω un ensemble. Une **tribu** (ou **σ-algèbre**) sur Ω est une famille $\mathcal{F}$ de sous-ensembles de Ω telle que

1. $\Omega \in \mathcal{F}$;

Le couple $(\Omega, \mathcal{F})$ est appelé un **espace probabilisable**. Les éléments de $\mathcal{F}$ sont les **événements**.

Tribu

Définition. Soit Ω un ensemble. Une **tribu** (ou **σ-algèbre**) sur Ω est une famille $\mathcal{F}$ de sous-ensembles de Ω telle que

1. $\Omega \in \mathcal{F}$;
2. si $A \in \mathcal{F}$, alors $A^{\mathsf{c}} \in \mathcal{F}$;

Le couple $(\Omega, \mathcal{F})$ est appelé un **espace probabilisable**. Les éléments de $\mathcal{F}$ sont les **événements**.

Tribu

Définition. Soit Ω un ensemble. Une **tribu** (ou **σ-algèbre**) sur Ω est une famille $\mathcal{F}$ de sous-ensembles de Ω telle que

1. $\Omega \in \mathcal{F}$;
2. si $A \in \mathcal{F}$, alors $A^{\mathsf{c}} \in \mathcal{F}$;
3. si $A_1, A_2, \ldots \in \mathcal{F}$, alors $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.

Tribu

Définition. Soit Ω un ensemble. Une **tribu** (ou **σ-algèbre**) sur Ω est une famille $\mathcal{F}$ de sous-ensembles de Ω telle que

1. $\Omega \in \mathcal{F}$;
2. si $A \in \mathcal{F}$, alors $A^{\mathsf{c}} \in \mathcal{F}$;
3. si $A_1, A_2, \ldots \in \mathcal{F}$, alors $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.

Le couple $(\Omega, \mathcal{F})$ est appelé un **espace probabilisable**. Les éléments de $\mathcal{F}$ sont les **événements**.

Exemples de tribus

Exemples de tribus

Exemples. Soit Ω un ensemble et $A \subseteq \Omega$.

- $\{\emptyset, \Omega\}$ est une tribu : la plus petite possible.

Dans le cas discret, on prend toujours $\mathcal{F} = \mathcal{P}(\Omega)$. C'est ce que nous avons fait implicitement jusqu'ici.

Exemples de tribus

Exemples. Soit Ω un ensemble et $A \subseteq \Omega$.

- $\{\emptyset, \Omega\}$ est une tribu : la plus petite possible.
- $\{\emptyset, A, A^c, \Omega\}$ est une tribu.

Dans le cas discret, on prend toujours $\mathcal{F} = \mathcal{P}(\Omega)$. C'est ce que nous avons fait implicitement jusqu'ici.

Exemples de tribus

Exemples. Soit Ω un ensemble et $A \subseteq \Omega$.

- $\{\emptyset, \Omega\}$ est une tribu : la plus petite possible.
- $\{\emptyset, A, A^c, \Omega\}$ est une tribu.
- $\mathcal{P}(\Omega)$, l'ensemble de *tous* les sous-ensembles de Ω, est une tribu : la plus grande possible.

Exemples de tribus

Exemples. Soit Ω un ensemble et $A \subseteq \Omega$.

- $\{\emptyset, \Omega\}$ est une tribu : la plus petite possible.
- $\{\emptyset, A, A^c, \Omega\}$ est une tribu.
- $\mathcal{P}(\Omega)$, l'ensemble de *tous* les sous-ensembles de Ω, est une tribu : la plus grande possible.

Dans le cas discret, on prend toujours $\mathcal{F} = \mathcal{P}(\Omega)$. C'est ce que nous avons fait implicitement jusqu'ici.

Propriétés des tribus

Propriétés des tribus

Proposition. Soit $\mathcal{F}$ une tribu sur Ω. Alors

1. $\emptyset \in \mathcal{F}$;

Démontrez la propriété 2 à l'aide des lois de De Morgan : $\bigcap_k A_k = \left(\bigcup_k A_k^{\mathsf{c}}\right)^{\mathsf{c}}$.

Propriétés des tribus

Proposition. Soit $\mathcal{F}$ une tribu sur Ω. Alors

1. $\emptyset \in \mathcal{F}$;
2. si $A_1, A_2, \ldots \in \mathcal{F}$, alors $\bigcap_{k=1}^{\infty} A_k \in \mathcal{F}$;

Démontrez la propriété 2 à l'aide des lois de De Morgan : $\bigcap_k A_k = \left(\bigcup_k A_k^{\mathsf{c}}\right)^{\mathsf{c}}$.

Propriétés des tribus

Proposition. Soit $\mathcal{F}$ une tribu sur Ω. Alors

1. $\emptyset \in \mathcal{F}$;
2. si $A_1, A_2, \ldots \in \mathcal{F}$, alors $\bigcap_{k=1}^{\infty} A_k \in \mathcal{F}$;
3. si $A, B \in \mathcal{F}$, alors $A \cup B$, $A \cap B$ et $A \setminus B$ sont dans $\mathcal{F}$.

Propriétés des tribus

Proposition. Soit $\mathcal{F}$ une tribu sur Ω. Alors

1. $\emptyset \in \mathcal{F}$;
2. si $A_1, A_2, \ldots \in \mathcal{F}$, alors $\bigcap_{k=1}^{\infty} A_k \in \mathcal{F}$;
3. si $A, B \in \mathcal{F}$, alors $A \cup B$, $A \cap B$ et $A \setminus B$ sont dans $\mathcal{F}$.

Démontrez la propriété 2 à l'aide des lois de De Morgan : $\bigcap_k A_k = \left(\bigcup_k A_k^{\mathsf{c}}\right)^{\mathsf{c}}$.

Tribu engendrée et tribu borélienne

Tribu engendrée et tribu borélienne

Proposition. Soit $\mathcal{C}$ une famille de sous-ensembles de Ω. Il existe une plus petite tribu contenant $\mathcal{C}$. On l'appelle la **tribu engendrée** par $\mathcal{C}$ et on la note $\sigma(\mathcal{C})$.

Tribu engendrée et tribu borélienne

Proposition. Soit $\mathcal{C}$ une famille de sous-ensembles de Ω. Il existe une plus petite tribu contenant $\mathcal{C}$. On l'appelle la **tribu engendrée** par $\mathcal{C}$ et on la note $\sigma(\mathcal{C})$.

Définition. La **tribu borélienne** $\mathcal{B}(\mathbb{R})$ est la tribu sur $\mathbb{R}$ engendrée par les intervalles. Ses éléments sont appelés les **boréliens**.

Tribu engendrée et tribu borélienne

Proposition. Soit $\mathcal{C}$ une famille de sous-ensembles de Ω. Il existe une plus petite tribu contenant $\mathcal{C}$. On l'appelle la **tribu engendrée** par $\mathcal{C}$ et on la note $\sigma(\mathcal{C})$.

Définition. La **tribu borélienne** $\mathcal{B}(\mathbb{R})$ est la tribu sur $\mathbb{R}$ engendrée par les intervalles. Ses éléments sont appelés les **boréliens**.

Les intervalles, les points, les ensembles finis ou dénombrables (comme $\mathbb{Q}$), les unions dénombrables d'intervalles, et leurs complémentaires sont des boréliens. En pratique, tous les ensembles que vous rencontrerez dans ce cours sont des boréliens.

Exemple : Pourquoi $\{x\}$ et $\mathbb{Q}$ sont des boréliens

Exemple : Pourquoi $\{x\}$ et $\mathbb{Q}$ sont des boréliens

- Pour $x \in \mathbb{R}$, $\{x\} = [x, x]$ est un intervalle fermé.

Exemple : Pourquoi $\{x\}$ et $\mathbb{Q}$ sont des boréliens

- Pour $x \in \mathbb{R}$, $\{x\} = [x, x]$ est un intervalle fermé. On peut aussi l'écrire

$$\{x\} = \bigcap_{n=1}^{\infty}]x - \tfrac{1}{n}, x + \tfrac{1}{n}[.$$

Exemple : Pourquoi $\{x\}$ et $\mathbb{Q}$ sont des boréliens

- Pour $x \in \mathbb{R}$, $\{x\} = [x, x]$ est un intervalle fermé. On peut aussi l'écrire

$$\{x\} = \bigcap_{n=1}^{\infty}]x - \tfrac{1}{n}, x + \tfrac{1}{n}[.$$

- $\mathbb{Q}$ est dénombrable : $\mathbb{Q} = \{q_1, q_2, q_3, ...\}$.

Exemple : Pourquoi $\{x\}$ et $\mathbb{Q}$ sont des boréliens

- Pour $x \in \mathbb{R}$, $\{x\} = [x, x]$ est un intervalle fermé. On peut aussi l'écrire

$$\{x\} = \bigcap_{n=1}^{\infty}]x - \tfrac{1}{n}, x + \tfrac{1}{n}[.$$

- $\mathbb{Q}$ est dénombrable : $\mathbb{Q} = \{q_1, q_2, q_3, ...\}$. Ainsi

$$\mathbb{Q} = \bigcup_{k=1}^{\infty} \{q_k\}$$

est une union dénombrable de boréliens, donc un borélien.

Exemple : Pourquoi $\{x\}$ et $\mathbb{Q}$ sont des boréliens

- Pour $x \in \mathbb{R}$, $\{x\} = [x, x]$ est un intervalle fermé. On peut aussi l'écrire

$$\{x\} = \bigcap_{n=1}^{\infty}]x - \tfrac{1}{n}, x + \tfrac{1}{n}[.$$

- $\mathbb{Q}$ est dénombrable : $\mathbb{Q} = \{q_1, q_2, q_3, ...\}$. Ainsi

$$\mathbb{Q} = \bigcup_{k=1}^{\infty} \{q_k\}$$

  est une union dénombrable de boréliens, donc un borélien.
- Par conséquent, $\mathbb{R} \setminus \mathbb{Q}$, l'ensemble des irrationnels, est aussi un borélien.

3

Les axiomes de Kolmogorov

Une mesure de probabilité

Une mesure de probabilité

Définition. Soit $(\Omega, \mathcal{F})$ un espace probabilisable. Une **mesure de probabilité** sur $(\Omega, \mathcal{F})$ est une fonction $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ satisfaisant les conditions suivantes :

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

Une mesure de probabilité

Définition. Soit $(\Omega, \mathcal{F})$ un espace probabilisable. Une **mesure de probabilité** sur $(\Omega, \mathcal{F})$ est une fonction $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ satisfaisant les conditions suivantes :

1. $\mathbb{P}(A) \geq 0$ pour tout $A \in \mathcal{F}$;

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

Une mesure de probabilité

Définition. Soit $(\Omega, \mathcal{F})$ un espace probabilisable. Une **mesure de probabilité** sur $(\Omega, \mathcal{F})$ est une fonction $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ satisfaisant les conditions suivantes :

1. $\mathbb{P}(A) \geq 0$ pour tout $A \in \mathcal{F}$;
2. $\mathbb{P}(\Omega) = 1$;

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

Une mesure de probabilité

Définition. Soit $(\Omega, \mathcal{F})$ un espace probabilisable. Une **mesure de probabilité** sur $(\Omega, \mathcal{F})$ est une fonction $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ satisfaisant les conditions suivantes :

1. $\mathbb{P}(A) \geq 0$ pour tout $A \in \mathcal{F}$;
2. $\mathbb{P}(\Omega) = 1$;
3. **(σ-additivité)** si $A_1, A_2, \ldots \in \mathcal{F}$ sont deux à deux incompatibles, alors

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k).$$

Une mesure de probabilité

Définition. Soit $(\Omega, \mathcal{F})$ un espace probabilisable. Une **mesure de probabilité** sur $(\Omega, \mathcal{F})$ est une fonction $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ satisfaisant les conditions suivantes :

1. $\mathbb{P}(A) \geq 0$ pour tout $A \in \mathcal{F}$;
2. $\mathbb{P}(\Omega) = 1$;
3. **(σ-additivité)** si $A_1, A_2, \ldots \in \mathcal{F}$ sont deux à deux incompatibles, alors

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k).$$

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

Le cas discret est un cas particulier

Le cas discret est un cas particulier

Proposition. Soit Ω un ensemble discret et $\mathcal{F} = \mathcal{P}(\Omega)$.

1. Si p est une loi de probabilité sur Ω (au sens du module d'introduction), alors $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)$ définit une mesure de probabilité sur $(\Omega, \mathcal{F})$.

Sur un espace discret, se donner une mesure de probabilité ou se donner des poids, c'est la même chose. Tout ce qui a été fait dans les modules précédents s'inscrit donc dans le cadre de Kolmogorov.

Le cas discret est un cas particulier

Proposition. Soit Ω un ensemble discret et $\mathcal{F} = \mathcal{P}(\Omega)$.

1. Si p est une loi de probabilité sur Ω (au sens du module d'introduction), alors $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)$ définit une mesure de probabilité sur $(\Omega, \mathcal{F})$.
2. Réciproquement, si $\mathbb{P}$ est une mesure de probabilité sur $(\Omega, \mathcal{F})$, alors $p(\omega) = \mathbb{P}(\{\omega\})$ est une loi de probabilité et $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)$.

Le cas discret est un cas particulier

Proposition. Soit Ω un ensemble discret et $\mathcal{F} = \mathcal{P}(\Omega)$.

1. Si p est une loi de probabilité sur Ω (au sens du module d'introduction), alors $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)$ définit une mesure de probabilité sur $(\Omega, \mathcal{F})$.
2. Réciproquement, si $\mathbb{P}$ est une mesure de probabilité sur $(\Omega, \mathcal{F})$, alors $p(\omega) = \mathbb{P}(\{\omega\})$ est une loi de probabilité et $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)$.

Sur un espace discret, se donner une mesure de probabilité ou se donner des poids, c'est la même chose. Tout ce qui a été fait dans les modules précédents s'inscrit donc dans le cadre de Kolmogorov.

Démonstration de la réciproque

Démonstration de la réciproque

Soit $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{P}(\Omega))$, avec Ω discret, et posons $p(\omega) = \mathbb{P}(\{\omega\})$.

Démonstration de la réciproque

Soit $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{P}(\Omega))$, avec Ω discret, et posons $p(\omega) = \mathbb{P}(\{\omega\})$.

- Par l'axiome 1, $p(\omega) \geq 0$.

Démonstration de la réciproque

Soit $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{P}(\Omega))$, avec Ω discret, et posons $p(\omega) = \mathbb{P}(\{\omega\})$.

- Par l'axiome 1, $p(\omega) \geq 0$.
- Soit $A = \{a_1, a_2, ...\} \subseteq \Omega$ (fini ou dénombrable). Les singletons $\{a_k\}$ sont deux à deux incompatibles et leur union est A.

Démonstration de la réciproque

Soit $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{P}(\Omega))$, avec Ω discret, et posons $p(\omega) = \mathbb{P}(\{\omega\})$.

- Par l'axiome 1, $p(\omega) \geq 0$.
- Soit $A = \{a_1, a_2, ...\} \subseteq \Omega$ (fini ou dénombrable). Les singletons $\{a_k\}$ sont deux à deux incompatibles et leur union est A. Par σ-additivité,

$$\mathbb{P}(A) = \sum_k \mathbb{P}(\{a_k\}) = \sum_{\omega \in A} p(\omega).$$

Démonstration de la réciproque

Soit $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{P}(\Omega))$, avec Ω discret, et posons $p(\omega) = \mathbb{P}(\{\omega\})$.

- Par l'axiome 1, $p(\omega) \geq 0$.
- Soit $A = \{a_1, a_2, ...\} \subseteq \Omega$ (fini ou dénombrable). Les singletons $\{a_k\}$ sont deux à deux incompatibles et leur union est A. Par σ-additivité,

$$\mathbb{P}(A) = \sum_k \mathbb{P}(\{a_k\}) = \sum_{\omega \in A} p(\omega).$$

- Avec $A = \Omega$ et l'axiome 2, on obtient $\sum_{\omega \in \Omega} p(\omega) = 1$.

Démonstration de la réciproque

Soit $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{P}(\Omega))$, avec Ω discret, et posons $p(\omega) = \mathbb{P}(\{\omega\})$.

- Par l'axiome 1, $p(\omega) \geq 0$.
- Soit $A = \{a_1, a_2, ...\} \subseteq \Omega$ (fini ou dénombrable). Les singletons $\{a_k\}$ sont deux à deux incompatibles et leur union est A. Par σ-additivité,

$$\mathbb{P}(A) = \sum_k \mathbb{P}(\{a_k\}) = \sum_{\omega \in A} p(\omega).$$

- Avec $A = \Omega$ et l'axiome 2, on obtient $\sum_{\omega \in \Omega} p(\omega) = 1$.

C'est l'étape « décomposer A en singletons » qui échoue si Ω n'est pas dénombrable.

4

Conséquences des axiomes

Premières conséquences

Premières conséquences

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

1. $\mathbb{P}(\emptyset) = 0$;

Premières conséquences

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

1. $\mathbb{P}(\emptyset) = 0$;
2. **(additivité finie)** si $A_1, \dots, A_n \in \mathcal{F}$ sont deux à deux incompatibles,

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k).$$

Premières conséquences

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

1. $\mathbb{P}(\emptyset) = 0$;
2. **(additivité finie)** si $A_1, \dots, A_n \in \mathcal{F}$ sont deux à deux incompatibles,

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k).$$

Premières conséquences — démonstration

1. Appliquons la σ-additivité à $\Omega, \emptyset, \emptyset, \ldots$, qui sont deux à deux incompatibles :

$$\mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \sum_{k=2}^{\infty} \mathbb{P}(\emptyset).$$

Premières conséquences — démonstration

1. Appliquons la σ-additivité à $\Omega, \emptyset, \emptyset, \ldots$, qui sont deux à deux incompatibles :

$$\mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \sum_{k=2}^{\infty} \mathbb{P}(\emptyset).$$

Comme $\mathbb{P}(\emptyset) \geq 0$, cette série ne peut valoir 0 que si $\mathbb{P}(\emptyset) = 0$.

Premières conséquences — démonstration

1. Appliquons la σ-additivité à $\Omega, \emptyset, \emptyset, \ldots$, qui sont deux à deux incompatibles :

$$\mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \sum_{k=2}^{\infty} \mathbb{P}(\emptyset).$$

Comme $\mathbb{P}(\emptyset) \geq 0$, cette série ne peut valoir 0 que si $\mathbb{P}(\emptyset) = 0$.

2. On complète la famille par $A_{n+1} = A_{n+2} = \cdots = \emptyset$ et on applique la σ-additivité, puis le point 1 :

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k) + \sum_{k=n+1}^{\infty} \underbrace{\mathbb{P}(\emptyset)}_{=0}.$$

Les propriétés du module d’introduction restent vraies

Les propriétés du module d'introduction restent vraies

Les démonstrations du module d'introduction n'utilisaient que la positivité, $\mathbb{P}(\Omega) = 1$ et l'additivité.

Les propriétés du module d'introduction restent vraies

Les démonstrations du module d'introduction n'utilisaient que la positivité, $\mathbb{P}(\Omega) = 1$ et l'additivité. Elles valent donc pour *toute* mesure de probabilité.

Les propriétés du module d'introduction restent vraies

Les démonstrations du module d'introduction n'utilisaient que la positivité, $\mathbb{P}(\Omega) = 1$ et l'additivité. Elles valent donc pour *toute* mesure de probabilité.

Propositions. Pour tous $A, B, A_1, A_2, \ldots \in \mathcal{F}$:

- $\mathbb{P}(A^{\mathsf{c}}) = 1 - \mathbb{P}(A)$ et $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;

Les propriétés du module d'introduction restent vraies

Les démonstrations du module d'introduction n'utilisaient que la positivité, $\mathbb{P}(\Omega) = 1$ et l'additivité. Elles valent donc pour *toute* mesure de probabilité.

Propositions. Pour tous $A, B, A_1, A_2, \ldots \in \mathcal{F}$:

- $\mathbb{P}(A^{\mathsf{c}}) = 1 - \mathbb{P}(A)$ et $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
- si $A \subseteq B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$ et $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$;

Les propriétés du module d'introduction restent vraies

Les démonstrations du module d'introduction n'utilisaient que la positivité, $\mathbb{P}(\Omega) = 1$ et l'additivité. Elles valent donc pour *toute* mesure de probabilité.

Propositions. Pour tous $A, B, A_1, A_2, \ldots \in \mathcal{F}$:

- $\mathbb{P}(A^{\mathsf{c}}) = 1 - \mathbb{P}(A)$ et $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
- si $A \subseteq B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$ et $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$;
- $\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mathbb{P}(A_k)$ (Boole) ;

Les propriétés du module d'introduction restent vraies

Les démonstrations du module d'introduction n'utilisaient que la positivité, $\mathbb{P}(\Omega) = 1$ et l'additivité. Elles valent donc pour *toute* mesure de probabilité.

Propositions. Pour tous $A, B, A_1, A_2, \ldots \in \mathcal{F}$:

- $\mathbb{P}(A^{\mathsf{c}}) = 1 - \mathbb{P}(A)$ et $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
- si $A \subseteq B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$ et $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$;
- $\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mathbb{P}(A_k)$ (Boole) ;
- $\mathbb{P}\left(\bigcap_{k=1}^{n} A_k\right) \geq 1 - \sum_{k=1}^{n} \mathbb{P}(A_k^{\mathsf{c}})$ (Bonferroni).

Exemple : Une démonstration à partir des axiomes

À partir des axiomes de Kolmogorov et de l'additivité finie, démontrez que $\mathbb{P}(A^{\mathsf{c}}) = 1 - \mathbb{P}(A)$ pour tout $A \in \mathcal{F}$.

Exemple : Une démonstration à partir des axiomes

À partir des axiomes de Kolmogorov et de l'additivité finie, démontrez que $\mathbb{P}(A^{\mathsf{c}}) = 1 - \mathbb{P}(A)$ pour tout $A \in \mathcal{F}$.

Les événements A et A^{c} sont incompatibles et $A \cup A^{\mathsf{c}} = \Omega$.

Exemple : Une démonstration à partir des axiomes

À partir des axiomes de Kolmogorov et de l'additivité finie, démontrez que $\mathbb{P}(A^{\mathsf{c}}) = 1 - \mathbb{P}(A)$ pour tout $A \in \mathcal{F}$.

Les événements A et A^{c} sont incompatibles et $A \cup A^{\mathsf{c}} = \Omega$. Par additivité finie et l'axiome 2,

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^{\mathsf{c}}) = \mathbb{P}(A) + \mathbb{P}(A^{\mathsf{c}}).$$

Exemple : Une démonstration à partir des axiomes

À partir des axiomes de Kolmogorov et de l'additivité finie, démontrez que $\mathbb{P}(A^{\mathsf{c}}) = 1 - \mathbb{P}(A)$ pour tout $A \in \mathcal{F}$.

Les événements A et A^{c} sont incompatibles et $A \cup A^{\mathsf{c}} = \Omega$. Par additivité finie et l'axiome 2,

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^{\mathsf{c}}) = \mathbb{P}(A) + \mathbb{P}(A^{\mathsf{c}}).$$

Remarquez que $A^{\mathsf{c}} \in \mathcal{F}$ grâce à la deuxième propriété des tribus : sans elle, $\mathbb{P}(A^{\mathsf{c}})$ ne serait même pas défini.

5

Continuité des probabilités

Suites monotones d’événements

Suites monotones d'événements

Définitions. Soit $A_1, A_2, \ldots$ une suite d'événements.

- La suite est **croissante** si $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$. On note alors $A_n \uparrow A$, où $A = \bigcup_{n=1}^{\infty} A_n$.

Exemples. $[0, 1 - \frac{1}{n}] \uparrow [0, 1[$ et $]x - \frac{1}{n}, x + \frac{1}{n}[\downarrow \{x\}$.

Suites monotones d'événements

Définitions. Soit $A_1, A_2, \ldots$ une suite d'événements.

- La suite est **croissante** si $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$. On note alors $A_n \uparrow A$, où $A = \bigcup_{n=1}^{\infty} A_n$.
- La suite est **décroissante** si $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$. On note alors $A_n \downarrow A$, où $A = \bigcap_{n=1}^{\infty} A_n$.

Suites monotones d'événements

Définitions. Soit $A_1, A_2, \ldots$ une suite d'événements.

- La suite est **croissante** si $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$. On note alors $A_n \uparrow A$, où $A = \bigcup_{n=1}^{\infty} A_n$.
- La suite est **décroissante** si $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$. On note alors $A_n \downarrow A$, où $A = \bigcap_{n=1}^{\infty} A_n$.

Exemples. $[0, 1 - \frac{1}{n}] \uparrow [0, 1[$ et $]x - \frac{1}{n}, x + \frac{1}{n}[\downarrow \{x\}$.

Théorème de continuité

Théorème de continuité

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

1. Si $A_n \uparrow A$, alors $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.

On peut « passer à la limite » à l'intérieur de $\mathbb{P}$ le long d'une suite monotone d'événements. Ce résultat est une conséquence directe de la σ-additivité : l'additivité finie ne suffirait pas.

Théorème de continuité

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

1. Si $A_n \uparrow A$, alors $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.
2. Si $A_n \downarrow A$, alors $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.

Théorème de continuité

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

1. Si $A_n \uparrow A$, alors $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.
2. Si $A_n \downarrow A$, alors $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.

On peut « passer à la limite » à l'intérieur de $\mathbb{P}$ le long d'une suite monotone d'événements. Ce résultat est une conséquence directe de la σ-additivité : l'additivité finie ne suffirait pas.

Démonstration du cas croissant

Démonstration du cas croissant

Posons $B_1 = A_1$ et $B_n = A_n \setminus A_{n-1}$ pour $n \geq 2$.

Démonstration du cas croissant

Posons $B_1 = A_1$ et $B_n = A_n \setminus A_{n-1}$ pour $n \geq 2$. Les B_n sont deux à deux incompatibles et

$$A_n = \bigcup_{k=1}^{n} B_k, \qquad A = \bigcup_{k=1}^{\infty} B_k.$$

Démonstration du cas croissant

Posons $B_1 = A_1$ et $B_n = A_n \setminus A_{n-1}$ pour $n \geq 2$. Les B_n sont deux à deux incompatibles et

$$A_n = \bigcup_{k=1}^{n} B_k, \qquad A = \bigcup_{k=1}^{\infty} B_k.$$

Par σ-additivité, puis additivité finie,

$$\mathbb{P}(A) = \sum_{k=1}^{\infty} \mathbb{P}(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(B_k) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

Démonstration du cas croissant

Posons $B_1 = A_1$ et $B_n = A_n \setminus A_{n-1}$ pour $n \geq 2$. Les B_n sont deux à deux incompatibles et

$$A_n = \bigcup_{k=1}^{n} B_k, \qquad A = \bigcup_{k=1}^{\infty} B_k.$$

Par σ-additivité, puis additivité finie,

$$\mathbb{P}(A) = \sum_{k=1}^{\infty} \mathbb{P}(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(B_k) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

Déduisez le cas décroissant en passant aux complémentaires : si $A_n \downarrow A$, alors $A_n^{\mathsf{c}} \uparrow A^{\mathsf{c}}$.

Exemple : Ne jamais obtenir pile

On lance indéfiniment une pièce de monnaie équilibrée. Quelle est la probabilité de ne jamais obtenir pile ?

Exemple : Ne jamais obtenir pile

On lance indéfiniment une pièce de monnaie équilibrée. Quelle est la probabilité de ne jamais obtenir pile ?

Soit A_n l'événement *aucun pile lors des n premiers lancers*.

Exemple : Ne jamais obtenir pile

On lance indéfiniment une pièce de monnaie équilibrée. Quelle est la probabilité de ne jamais obtenir pile ?

Soit A_n l'événement *aucun pile lors des n premiers lancers*. On a $\mathbb{P}(A_n) = \frac{1}{2^n}$ et $A_n \downarrow A$, où A est l'événement *ne jamais obtenir pile*.

Exemple : Ne jamais obtenir pile

On lance indéfiniment une pièce de monnaie équilibrée. Quelle est la probabilité de ne jamais obtenir pile ?

Soit A_n l'événement *aucun pile lors des n premiers lancers*. On a $\mathbb{P}(A_n) = \frac{1}{2^n}$ et $A_n \downarrow A$, où A est l'événement *ne jamais obtenir pile*. Par continuité,

$$\mathbb{P}(A) = \lim_{n\to\infty} \frac{1}{2^n} = 0.$$

Exemple : Ne jamais obtenir pile

On lance indéfiniment une pièce de monnaie équilibrée. Quelle est la probabilité de ne jamais obtenir pile ?

Soit A_n l'événement *aucun pile lors des n premiers lancers*. On a $\mathbb{P}(A_n) = \frac{1}{2^n}$ et $A_n \downarrow A$, où A est l'événement *ne jamais obtenir pile*. Par continuité,

$$\mathbb{P}(A) = \lim_{n\to\infty} \frac{1}{2^n} = 0.$$

A n'est pas vide (il contient $F, F, F, \ldots$) : un événement de probabilité nulle n'est pas forcément impossible.

Exemple : Probabilité d’un point

Exemple : Probabilité d'un point

Supposons que $\mathbb{P}$ soit une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $\mathbb{P}([a,b]) = b - a$ pour $0 \leq a \leq b \leq 1$ (choix d'un nombre au hasard dans $[0,1]$).

Exemple : Probabilité d'un point

Supposons que $\mathbb{P}$ soit une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $\mathbb{P}([a,b]) = b - a$ pour $0 \leq a \leq b \leq 1$ (choix d'un nombre au hasard dans $[0,1]$).

Pour $x \in]0,1[$, on a $]x - \frac{1}{n}, x + \frac{1}{n}[\downarrow \{x\}$.

Exemple : Probabilité d'un point

Supposons que $\mathbb{P}$ soit une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $\mathbb{P}([a,b]) = b - a$ pour $0 \leq a \leq b \leq 1$ (choix d'un nombre au hasard dans $[0,1]$).

Pour $x \in]0,1[$, on a $]x - \frac{1}{n}, x + \frac{1}{n}[\downarrow \{x\}$. Par continuité,

$$\mathbb{P}(\{x\}) = \lim_{n\to\infty} \mathbb{P}\big(]x - \tfrac{1}{n}, x + \tfrac{1}{n}[\big) = \lim_{n\to\infty} \frac{2}{n} = 0.$$

Exemple : Probabilité d'un point

Supposons que $\mathbb{P}$ soit une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $\mathbb{P}([a,b]) = b - a$ pour $0 \leq a \leq b \leq 1$ (choix d'un nombre au hasard dans $[0,1]$).

Pour $x \in]0,1[$, on a $]x - \frac{1}{n}, x + \frac{1}{n}[\downarrow \{x\}$. Par continuité,

$$\mathbb{P}(\{x\}) = \lim_{n\to\infty} \mathbb{P}\big(]x - \tfrac{1}{n}, x + \tfrac{1}{n}[\big) = \lim_{n\to\infty} \frac{2}{n} = 0.$$

Chaque point a une probabilité nulle, alors que les intervalles ont une probabilité positive : c'est le point de départ du module sur les probabilités géométriques.

En résumé

En résumé

	Cas discret	**Cas général**
Événements	$\mathcal{P}(\Omega)$	une tribu $\mathcal{F}$
Probabilité	donnée par les poids $p(\omega)$	donnée sur les événements
$\mathbb{P}(A)$	$\sum_{\omega \in A} p(\omega)$	axiomes de Kolmogorov
$\mathbb{P}(\{\omega\})$	$p(\omega)$, souvent > 0	souvent 0

En résumé

	Cas discret	**Cas général**
Événements	$\mathcal{P}(\Omega)$	une tribu $\mathcal{F}$
Probabilité	donnée par les poids $p(\omega)$	donnée sur les événements
$\mathbb{P}(A)$	$\sum_{\omega \in A} p(\omega)$	axiomes de Kolmogorov
$\mathbb{P}(\{\omega\})$	$p(\omega)$, souvent > 0	souvent 0

Les axiomes de Kolmogorov offrent un seul et même cadre pour les expériences aléatoires discrètes et continues. Toutes les règles de calcul (complémentarité, union, Boole, probabilités conditionnelles, indépendance, etc.) s'y appliquent sans modification.

Remarques

Remarques.

- Ce document est rédigé avec Beamer, une classe de LaTeX.

Pour signaler une erreur dans ce document, veuillez écrire à jerome.soucy@mat.ulaval.ca.