Probabilités

Table des matières

1. Cadre général de la théorie des probabilités

1.1. Espace probabilisé général

Définition 1.1. Soit Ω un ensemble. On appelle *tribu* sur Ω une famille $\mathcal{F}$ de parties de Ω vérifiant :

(1) $\mathcal{F}$ est non-vide : $\emptyset \in \mathcal{F}$,
(2) la stabilité par passage au complémentaire : $\forall A \in \mathcal{F}, A^c \in \mathcal{F}$,
(3) la stabilité par union dénombrable : $\forall (A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}, \bigcup_{n\geq 1} A_n \in \mathcal{F}$.

Définition 1.2. Soit Ω un ensemble et $\mathcal{F}$ une tribu sur Ω. On appelle *mesure de probabilité* une mesure $\mathbb{P} : \mathcal{F} \to \mathbb{R}_+$ vérifiant $\mathbb{P}(\Omega) = 1$.

Définition 1.3. Soit Ω un ensemble, $\mathcal{F}$ une tribu sur Ω et $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{F})$. On appelle *espace probabilisé* le triplet $(\Omega, \mathcal{F}, \mathbb{P})$, on dit que Ω est l'univers et que $\mathcal{F}$ sont les événements.

Remarque 1.4. Dans le cadre discret, on avait souvent $\mathcal{F} := \mathcal{P}(\Omega)$. Dans le cadre général, on aura souvent $\mathcal{F} \subsetneq \mathcal{P}(\Omega)$.

Définition 1.5. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements sur $(\Omega, \mathcal{F}, \mathbb{P})$. On dit que $(A_n)_{n\in\mathbb{N}}$ est un *système complet* si elle vérifie :

(1) les A_n sont disjoints deux à deux,
(2) la probabilité de l'union des A_n est 1.

Proposition 1.6. Soit $(A_n)_{n\in\mathbb{N}}$ un système complet sur $(\Omega, \mathcal{F}, \mathbb{P})$. Alors on a

$$\forall B \in \mathcal{F}, \mathbb{P}(B) = \sum_{n=1}^{+\infty} \mathbb{P}(B \cap A_n).$$

Démonstration. On pose $C := \bigcup_{n\geq 1} A_n$, puisque $\mathbb{P}(C) = 1$, on a $\mathbb{P}(C^c) = 0$ d'où $\mathbb{P}(B \cap C^c) = 0$. Soit $B \in \mathcal{F}$, on en déduit

$$\mathbb{P}(B) = \mathbb{P}(B \cap C) + \underbrace{\mathbb{P}(B \cap C^c)}_{=0} = \mathbb{P}\left(\bigcup_{n\geq 1} B \cap A_n\right) = \sum_{n=1}^{+\infty} \mathbb{P}(B \cap A_n).$$

$\square$

Corollaire 1.7. Soit $(A_n)_{n\in\mathbb{N}}$ un système complet sur $(\Omega, \mathcal{F}, \mathbb{P})$. Alors pour tout $B \in \mathcal{F}$ on a

(1) $\mathbb{P}(B) = \sum_{n=1}^{+\infty} \mathbb{P}(A_n) P(B|A_n)$,
(2) $\forall i \geq 1, \mathbb{P}(A_i|B) = \dfrac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\sum_{n=1}^{+\infty} \mathbb{P}(A_n)\mathbb{P}(B|A_n)}$.

Théorème 1.8. (Continuité de la mesure de probabilité) Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

(1) Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'événements. Alors on a

$$\lim_{n\to+\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n\geq 1} A_n\right).$$

(2) Soit $(A_n)_{n\in\mathbb{N}}$ une suite décroissante d'événements. Alors on a

$$\lim_{n\to+\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n\geq 1} A_n\right).$$

Démonstration.

(1) Pour tout $n \geq 1$, on pose $B_n := A_n \setminus A_{n-1}$ avec $A_0 = \emptyset$, tel que les $(B_n)_{n\in\mathbb{N}}$ forme un système complet sur $\bigcup_{n\geq 1} A_n$, on en déduit alors

$$\mathbb{P}\left(\bigcup_{n\geq 1} A_n\right) = \mathbb{P}\left(\bigcup_{n\geq 1} B_n\right) = \sum_{n=1}^{+\infty} \mathbb{P}(B_n) = \sum_{n=1}^{+\infty} \mathbb{P}(A_n) - \mathbb{P}(A_{n-1})$$

on reconnait une somme téléscopique et on a donc

$$\mathbb{P}\left(\bigcup_{n\geq 1} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n) - \mathbb{P}(A_0) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

(2) On obtient directement le résultat par passage au complémentaire.

□

Définition 1.9. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements de $(\Omega, \mathcal{F}, \mathbb{P})$.

- On appelle *limite supérieure* de la suite $(A_n)_{n\in\mathbb{N}}$ la valeur

$$\limsup_{n\to+\infty} A_n := \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k$$

intuitivement on considère les éléments qui appartiennent à une infinité d'événements.

- On appelle *limite inférieure* de la suite $(A_n)_{n\in\mathbb{N}}$ la valeur

$$\limsup_{n\to+\infty} A_n := \bigcup_{n\geq 1} \bigcap_{k\geq n} A_k.$$

Corollaire 1.10. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements de $(\Omega, \mathcal{F}, \mathbb{P})$. Alors on a

$$\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = \lim_{m\to+\infty} \lim_{n\to+\infty} \mathbb{P}\left(\bigcup_{k=m}^{n} A_k\right)$$

$$\mathbb{P}\left(\liminf_{n\to+\infty} A_n\right) = \lim_{m\to+\infty} \lim_{n\to+\infty} \mathbb{P}\left(\bigcap_{k=m}^{n} A_k\right)$$

Proposition 1.11. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements de $(\Omega, \mathcal{F}, \mathbb{P})$. Alors on a

$$\mathbb{P}\left(\bigcup_{n\geq 1} A_n\right) \leq \sum_{n=1}^{+\infty} A_n.$$

Démonstration. On sait que le résultat est vérifié pour un nombre fini d'événements. Par passage à la limite et par continuité de la mesure $\mathbb{P}$ on a

$$\mathbb{P}\left(\bigcup_{n\geq 1} A_n\right) = \lim_{m\to+\infty} \mathbb{P}\left(\bigcup_{n=1}^{m} A_n\right) \leq \lim_{m\to+\infty} \sum_{n=1}^{m} \mathbb{P}(A_n) = \sum_{n=1}^{+\infty} \mathbb{P}(A_n).$$

□

Définition 1.12. Soit A un événement de $(\Omega, \mathcal{F}, \mathbb{P})$.

- On dit que A est *négligeable* si $\mathbb{P}(A) = 0$.
- On dit que A est *presque-sûr* si $\mathbb{P}(A) = 1$.

Corollaire 1.13. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Alors
- L'union dénombrable d'événements négligeables est négligeable.
- L'intersection dénombrable d'événements presque-sûrs est presque-sûre.

Proposition 1.14. Soit $\mathcal{A}$ une famille d'événements de $(\Omega, \mathcal{F}, \mathbb{P})$. Alors il existe une unique tribu $\sigma(\mathcal{A})$ telle que $\sigma(\mathcal{A})$ soit la plus petite tribu contenant $\mathcal{A}$.

Démonstration. Il existe au moins une tribu contenant $\mathcal{A}$, à savoir $\mathcal{P}(\Omega)$. Alors l'intersection de toutes les tribus contenant $\mathcal{A}$ est une tribu et convient. □

Définition 1.15. Soit $\mathcal{A}$ une famille d'événements de $(\Omega, \mathcal{F}, \mathbb{P})$. On appelle *tribu engendrée* par $\mathcal{A}$, notée $\sigma(\mathcal{A})$, la tribu de la Proposition 1.14.

Exemple 1.16. Soit A un événement de $(\Omega, \mathcal{F}, \mathbb{P})$. Alors $\sigma(\{A\}) = \{\emptyset, A, A^c, \Omega\}$.

1.2. Exemples d'espace probabilisés

Définition 1.17. Soit $(E, \mathcal{O})$ un espace topologique. On appelle *tribu borélienne* sur E, notée $\mathcal{B}(E)$, la tribu engendrée par les intervalles ouverts de E, c'est-à-dire $\mathcal{B}(E) := \sigma(\mathcal{O})$.

Lemme 1.18. Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de probabilité sur $(\Omega, \mathcal{F})$ et $(\lambda_n)_{n\in\mathbb{N}}$ une suite de nombres réels positifs telle que $\sum_{n=1}^{+\infty} \lambda_n = 1$. Alors $\mu = \sum_{n=1}^{+\infty} \lambda_n \mu_n$ est une mesure de probabilité sur $(\Omega, \mathcal{F})$.

1.2.1. Univers $\Omega = \mathbb{N}$

Se référer au cours de *Probabilités* de deuxième année.

1.2.2. Univers $\Omega = \mathbb{R}$

Exemple 1.19. (Mesure de Dirac) Soit $x \in \mathbb{R}$, l'application $\delta_x : \mathcal{B}(\mathbb{R}) \to \mathbb{R}_+$ définie par

$$\forall A \in \mathcal{B}(\mathbb{R}), \delta_x(A) = \begin{cases} 0 \text{ si } x \notin A \\ 1 \text{ si } x \in A \end{cases}$$

est une mesure de probabilité sur $\mathbb{R}$.

Exemple 1.20. (Mesure uniforme sur $\{1, ..., n\}$) L'application $\mu = \frac{1}{n} \sum_{k=1}^{n} \delta_k$ est une mesure uniforme sur $\mathbb{R}$.

Exemple 1.21. (Mesure de Poisson) Soit $\lambda > 0$, l'application $\mu = \sum_{n=1}^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} \delta_n$ est une mesure de Poisson sur $\mathbb{R}$.

Définition 1.22. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction borélienne. On dit que f est une *densité de probabilité* sur $\mathbb{R}$ si elle vérifie :
(1) pour λ-presque tout $x \in \mathbb{R}$, $f(x) \geq 0$,
(2) $\int_{\mathbb{R}} f(x) \,\mathrm{d}\lambda(x) = 1$.

Lemme 1.23. Soit f une densité de probabilité sur $\mathbb{R}$. Alors l'application $\mu_f : \mathcal{B}(\mathbb{R}) \to \mathbb{R}_+$ définie par $\forall A \in \mathcal{B}(\mathbb{R}), \mu_f(A) = \int_A f(x) \,\mathrm{d}\lambda(x)$ est une mesure de probabilité sur $\mathbb{R}$.

Démonstration. On a bien $\forall A \in \mathcal{B}(\mathbb{R}), \mu_f(A) \geq 0$. De plus $\mu_f(\mathbb{R}) = 1$. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{B}(\mathbb{R})$ deux à deux disjoints. On pose $A := \bigcup_{n\geq 1} A_n$, alors $\mathbb{1}_A = \sum_{n=1}^{+\infty} \mathbb{1}_{A_n}$ et

$$\mu_f(A) = \int_A f(x) \,\mathrm{d}\lambda(x) = \int_{\mathbb{R}} \mathbb{1}_A(x) f(x) \,\mathrm{d}\lambda(x) = \int_{\mathbb{R}} \sum_{n=1}^{+\infty} \mathbb{1}_{A_n}(x) f(x) \,\mathrm{d}\lambda(x)$$

d'après le théorème de convergence monotone on a

$$\mu_f(A) = \lim_{m\to+\infty} \int_{\mathbb{R}} \sum_{n=1}^{m} \mathbb{1}_{A_n} f(x)\,\mathrm{d}\lambda(x) = \lim_{m\to+\infty} \sum_{n=1}^{m} \mu_f(A_n) = \sum_{n=1}^{+\infty} \mu_f(A_n).$$

Donc μ_f est bien une mesure de probabilité sur $\mathbb{R}$. □

Remarque 1.24. On dit que μ_f est la *mesure de densité f*.

Proposition 1.25. Soit f et g deux densités de probabilités sur $\mathbb{R}$. Alors les mesures de densité μ_f et μ_g sont égales si et seulement si f et g sont égales presque partout.

Démonstration.
$\Rightarrow$: Supposons que $\mu_f = \mu_g$. On pose

$$\begin{aligned} A_+ &:= \{x \in \mathbb{R} \mid f(x) > g(x)\} \\ A_- &:= \{x \in \mathbb{R} \mid f(x) < g(x)\} \end{aligned}$$

ces deux ensembles sont boréliens car f et g sont boréliennes. Par construction

$$\int_{A_+} f - g\,\mathrm{d}\lambda = \mu_f(A_+) - \mu_g(A_+) = 0 = \int_{A_-} f - g\,\mathrm{d}\lambda$$

de plus $A := \{x \in \mathbb{R} \mid |f(x) - g(x)| > 0\} = A_+ \cup A_-$, on en déduit

$$\int_A |f-g|\,\mathrm{d}\lambda = \int_A (f-g)\mathbb{1}_{A_+} + (g-f)\mathbb{1}_{A_-}\,\mathrm{d}\lambda = 0$$

donc $f - g = 0$ presque partout et $f = g$ presque partout.

$\Leftarrow$: Si $f = g$ presque partout, alors il est évident que $\mu_f = \mu_g$. □

Exemple 1.26. (Loi uniforme) Soit $c, d \in \mathbb{R}$ avec $c < d$. Alors la fonction $f : \mathbb{R} \to \mathbb{R}, x \mapsto \frac{\mathbb{1}_{[c,d]}(x)}{d-c}$ est une densité de probabilité. En particulier, pour tout $[a,b] \subset [c,d]$

$$\mu_f([a,b]) = \int_{[a,b]} f(x)\,\mathrm{d}\lambda x = \frac{b-a}{d-c}.$$

On note la probabilité associée $\mathcal{U}([c,d])$.

Exemple 1.27. (Loi exponentielle) Soit $\lambda > 0$. Alors la fonction $f : \mathbb{R} \to \mathbb{R}, x \mapsto \lambda e^{-\lambda x}\mathbb{1}_{\mathbb{R}_+}(x)$ est une densité de probabilité. On note la probabilité associée $\mathcal{E}(\lambda)$.

Exemple 1.28. (Loi normale) La fonction $f : \mathbb{R} \to \mathbb{R}, x \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ est une densité de probabilité. On note la probabilité associée $\mathcal{N}(0,1)$.

1.2.3. Univers $\Omega = \mathbb{R}^d$

On peut étendre les exemples de $\mathbb{R}$, ainsi que les définitions de densité et de mesures de probabilité associée.

1.3. Classe monotone

Définition 1.29. Soit $\mathcal{C}$ une famille de parties d'un ensemble Ω. On dit que $\mathcal{C}$ est une *classe monotone* si elle vérifie :
(1) $\Omega \in \mathcal{C}$,
(2) $\forall A, B \in \mathcal{C}, A \subset B \Rightarrow B \setminus A \in \mathcal{C}$,
(3) $\forall (A_n)_{n\in\mathbb{N}} \in \mathcal{C}^{\mathbb{N}}$ croissante, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{C}$.

Remarque 1.30. Une tribu est une classe monotone, la réciproque est fausse.

Lemme 1.31. Soit $\mathcal{C}$ une classe monotone. Alors $\mathcal{C}$ est une tribu si et seulement si elle est stable par intersection finie, c'est-à-dire :

$$\forall A_1, ..., A_n \in \mathcal{C}, \bigcap_{k=1}^{n} A_n \in \mathcal{C}.$$

Démonstration.
$\Rightarrow$: Si $\mathcal{C}$ est une tribu elle est stable par intersection finie.
$\Leftarrow$: Supposons que $\mathcal{C}$ est stable par intersection finie. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{C}$. Puisque $\mathcal{C}$ est stable par passage au complémentaire, $\mathcal{C}$ est aussi stable par union finie, en effet

$$A, B, \in \mathcal{C} \Rightarrow A^c, B^c \in \mathcal{C} \Rightarrow A^c \cap B^c \in \mathcal{C} \Rightarrow A \cup B = (A^c \cap B^c)^c \in \mathcal{C}$$

on a donc pour tout $N \in \mathbb{N}$, $\bigcup_{n=0}^{N} A_n \in \mathcal{C}$, et par union croissante

$$\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{N\in\mathbb{N}} \underbrace{\bigcup_{n=0}^{N} A_n}_{\text{croissante}} \in \mathcal{C}$$

donc $\mathcal{C}$ est bien une tribu. □

Définition 1.32. Soit $\mathcal{A}$ une famille de parties d'un ensemble Ω. On appelle *classe monotone engendrée* par $\mathcal{A}$, notée $\mathcal{C}(\mathcal{A})$, l'intersection de toutes les classes monotones contenant $\mathcal{A}$.

Théorème 1.33. (Théorème de la classe monotone) Soit $\mathcal{A}$ une famille de partie d'un ensemble Ω. Si $\mathcal{A}$ est stable par intersection finie, alors $\mathcal{C}(\mathcal{A}) = \sigma(\mathcal{A})$.