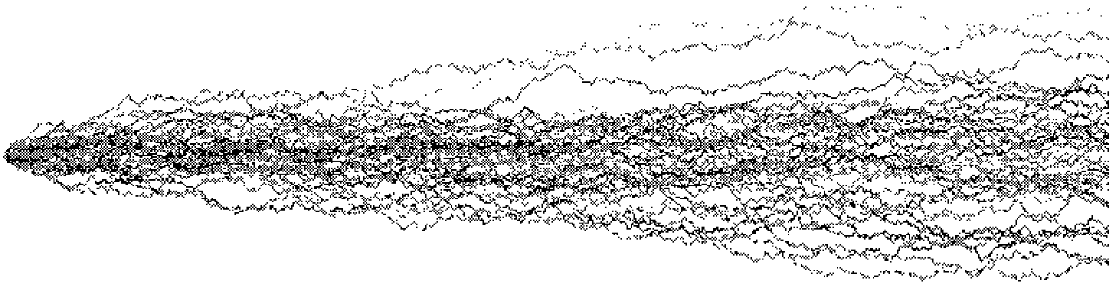
Chapitre 28

Probabilités



Trajectoires de marches aléatoires en dimension 1

Le calcul des probabilités remonte à Blaise Pascal (1623 – 1662), qui était à la fois mathématicien, physicien, philosophe et inventeur (il a créé la première machine à calculer).

Cependant, il faut attendre 1931 (!) et les travaux du mathématicien russe ***Andreï Kolmogorov*** *pour donner une axiomatique précise à la théorie des probabilités. Cette axiomatique a permis de poursuivre l'étude mathématique des phénomènes aléatoires, et notamment des marches aléatoires. Cette théorie est entre autres appliquée à l'étude de l'évolution des marchés boursiers.*

1 Espaces probabilisés

1.1 Définition

Définition
Un espace probabilisé fini *est un couple* (Ω, P) *où*

- Ω *est un ensemble fini non vide, appelé* univers ;
- P *est une application* $P : \mathscr{P}(\Omega) \longrightarrow [0,1]$*, appelée* probabilité *ou* mesure de probabilité *et vérifiant les axiomes suivants :*
  - *(i)* Normalisation *:*
  $$\boxed{P(\Omega) = 1.}$$
  - *(ii)* Additivité *: Si* A *et* B *sont deux parties de* Ω *disjointes, on a*
  $$\boxed{P(A \sqcup B) = P(A) + P(B).}$$

1.2 Notation

Dans la suite de ce document, quand ce n'est pas précisé, Ω désigne un ensemble non vide et $P(\cdot)$ une probabilité sur Ω.

1.3 Exemples importants

a) Soit Ω un ensemble fini non vide.
On appelle *probabilité uniforme sur* Ω l'application $P : \mathscr{P}(\Omega) \longrightarrow [0,1]$ définie par
$$\forall X \in \mathscr{P}(\Omega),\ P(X) = \frac{\mathrm{Card}(X)}{\mathrm{Card}(\Omega)}.$$

b) Soient $\Omega = \big\{0,1\big\}$ et $p \in [0,1]$. L'application
$$P : \begin{cases} \mathscr{P}(\Omega) \longrightarrow [0,1] \\ X \longmapsto \begin{cases} 0 & \text{si } X = \varnothing \\ 1 & \text{si } X = \Omega \\ p & \text{si } X = \{1\} \\ 1-p & \text{si } X = \{0\} \end{cases} \end{cases}$$
est une probabilité, appelée *probabilité de Bernoulli de paramètre p*. Il s'agit de la probabilité uniforme si et seulement si $p = 1/2$.

1.4 Propriétés fondamentales

Proposition
Soit (Ω, P) un espace probabilisé fini. On a les propriétés suivantes :

(i) $P(\varnothing) = 0$.

(ii) Si A est une partie de Ω, alors $P(\Omega \setminus A) = 1 - P(A)$.

(iii) Si $A_1, \dots, A_n$ sont des parties de Ω deux à deux disjointes, alors on a

$$P\left(\bigsqcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

(iv) Si $A \subset B$ sont deux parties de Ω, alors $P(A) \leqslant P(B)$.

(v) Si A et B sont deux parties de Ω, on a

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

(vi) ***Formule des probabilités totales***
Si $A_1, \dots, A_n, B$ sont des parties de Ω telles que les A_i sont deux à deux disjointes telles

$$\bigsqcup_{i=1}^{n} A_i = \Omega,$$

alors on a

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B).$$

(On dit que les A_i forment un système complet *d'événements.)*

Cette proposition est à connaître par cœur.
Sa démonstration est à maîtriser totalement.
Tous les points de la preuve doivent devenir des réflexes absolus.

Démonstration. —

(i) On a $\varnothing = \varnothing \sqcup \varnothing$, donc la propriété d'additivité entraîne $P(\varnothing) = P(\varnothing) + P(\varnothing)$, donc $P(\varnothing) = 0$.

(ii) Il suffit de remarquer que $A \sqcup (\Omega \setminus A) = \Omega$.

(iii) Si A, B et C sont deux à deux disjointes, alors C et $A \sqcup B$ sont disjoints. On a alors

$$P(A \sqcup B \sqcup C) = P\Big((A \sqcup B) \sqcup C\Big) = P(A \sqcup B) + P(C) = P(A) + P(B) + P(C).$$

Le cas général se démontre par récurrence, en suivant la même idée *(exercice)*.

(iv) On a $P(B) = P(A \sqcup (B \setminus A)) = P(A) + P(B \setminus A) \geqslant P(A)$ car $P(\cdot)$ est à valeurs dans $[0, 1] \subset \mathbb{R}_+$.

(v) On a $(A \cap B) \sqcup (A \setminus B) \sqcup (B \setminus A) = A \cup B$, donc

$$\begin{aligned} P(A \cup B) &= P(A \cap B) + P(A \setminus B) + P(B \setminus A) \\ &= P(A \cap B) + P(A \setminus B) + P(A \cap B) + P(B \setminus A) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

(vi) Il suffit de remarquer que $\bigsqcup_{i=1}^{n} (A_i \cap B) = B$.

■

1.5 Caractérisation d'une probabilité

Proposition

a) Soit (Ω, P) un espace probabilisé fini. Alors, on a

$$\sum_{\omega\in\Omega} P\Big(\{\omega\}\Big) = 1.$$

b) Réciproquement, si Ω est un ensemble fini non vide, et si $(p_\omega)_{\omega\in\Omega}$ est une famille de nombres appartenant à $[0,1]$ telle que $\sum\limits_{\omega\in\Omega} p_\omega = 1$, alors il existe une unique probabilité $P(\cdot)$ telle que

$$\forall\omega\in\Omega,\ P\Big(\{\omega\}\Big) = p_\omega.$$

Autrement dit : une probabilité est caractérisée par la probabilité qu'elle donne aux singletons.

Démonstration. · · ·

a) Les singletons $\Big(\{\omega\}\Big)_{\omega\in\Omega}$ sont deux à deux disjoints, donc

$$\sum_{\omega\in\Omega} \Big(\{\omega\}\Big) = P\Bigg(\bigsqcup_{\omega\in\Omega}\{\omega\}\Bigg) = P(\Omega) = 1.$$

b) Commençons par montrer l'existence. Soit

$$P : \left\{\begin{array}{rcl} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ A & \longmapsto & \sum\limits_{\omega\in A} p_\omega \end{array}\right. .$$

Montrons que P est une probabilité.

(i) Déjà, on a $\forall\omega\in\Omega,\ p_\omega \geqslant 0$. Donc, si $A\in\mathscr{P}(\Omega)$, on a

$$0 \leqslant P(A) = \sum_{\omega\in A} p_\omega \leqslant \sum_{\omega\in\Omega} p_\omega = 1.$$

(ii) On a $P(\Omega) = \sum\limits_{\omega\in\Omega} p_\omega = 1$.

(iii) Soit A et B deux parties disjointes. On a

$$P(A\sqcup B) = \sum_{\omega\in A\sqcup B} p_\omega = \sum_{\omega\in A} p_\omega + \sum_{\omega\in B} p_\omega = P(A) + P(B).$$

Soit maintenant P_1 et P_2 deux probabilités telles que

$$\forall\omega\in\Omega,\ P_1\Big(\{\omega\}\Big) = P_2\Big(\{\omega\}\Big).$$

On a alors, quel que soit $A\in\mathscr{P}(\Omega)$,

$$P_1(A) = \sum_{\omega\in A} P_1\Big(\{\omega\}\Big) = \sum_{\omega\in A} P_2\Big(\{\omega\}\Big) = P_2(A).$$

Donc, on a bien $P_1 = P_2$.

■

Exemples

- La probabilité uniforme sur Ω est caractérisée par :
$$\forall \omega \in \Omega, \quad P\Big(\{\omega\}\Big) = \frac{1}{|\Omega|}.$$
- De même, la probabilité uniforme sur Ω est caractérisée par le fait qu'elle est constante sur les singletons (d'où son nom) :
$$\forall \omega, \omega' \in \Omega, \quad P\Big(\{\omega\}\Big) = P\Big(\{\omega'\}\Big).$$
- La probabilité de Bernoulli de paramètre p est caractérisée par le fait que $P(\{1\}) = p$ (qui implique à son tour que $P(\{0\}) = 1 - p$).

1.6 Vocabulaire !!

Ce qui suit est à maîtriser impérativement et absolument.

Soit (Ω, P) est un espace probabilisé.

1) Un élément $\omega \in \Omega$ est appelé une *issue* (sous-entendu de (Ω, P)).
Une partie de Ω est appelée un *événement* (sous-entendu de (Ω, P)). Si A est un événement, on dit que $P(A)$ est la *probabilité de* A. Ainsi, P est une application qui associe un élément de $[0, 1]$ à chaque événement.

2) La partie $\varnothing \in \mathscr{P}(\Omega)$ est appelée l'*événement impossible.*
La partie $\Omega \in \mathscr{P}(\Omega)$ est appelée l'*événement certain* (sous-entendu de (Ω, P)).

3) Soit A un événement.
Le complémentaire de A dans Ω (*ie* $\Omega \setminus A$) est appelé *événement contraire de* A. On le note $\overline{A}$.

4) Si A et B sont deux événements, on dit que *A et B sont incompatibles* $\overset{\Delta}{ssi}$ $A \cap B = \varnothing$.

5) Un *système complet d'événements* (sous-entendu de (Ω, P)) est une famille d'événements $(A_1, \dots, A_n)$ telle que
   a) les parties A_i sont disjointes *ie*
$$\boxed{\forall i, j \in [\![1, n]\!], \quad i \neq j \implies A_i \cap A_j = \varnothing}$$
   b) les A_i recouvrent tout Ω, *ie*
$$\boxed{\bigsqcup_{i=1}^{n} A_i = A_1 \sqcup \cdots \sqcup A_n = \Omega.}$$

6) Soit A un événement.
On dit que *A est négligeable* $\overset{\Delta}{ssi}$ $P(A) = 0$.
On dit que *A est presque sûr* ou que *A est quasi certain* $\overset{\Delta}{ssi}$ $P(A) = 1$.

7) On décrit la probabilité uniforme comme celle dont les issues sont *équiprobables*, ce qui signifie que les événements $\{\omega\}$ ont tous la même probabilité.

Fait
Soit A un événement. Alors, on a
a) La famille $(A, \overline{A})$ est un sytème complet d'événements.
b) A quasi certain $\iff$ $\overline{A}$ négligeable.

Démonstration du fait. — Laissée en exercice. ■

1.7 Remarques

On peut utiliser ce vocabulaire pour reformuler les propriétés de la section précédente.
Par exemple,

- Deux événements contraires ont des probabilités dont la somme fait 1.
- La propriété d'additivité dit que la probabilité de l'union de deux événements incompatibles est la somme de leur probabilité.
- La formule des probabilités totales parle d'un système complet d'événements.

Par ailleurs,

- L'événement $\varnothing$ est négligeable, l'événement Ω est quasi certain.
- Les réciproques sont, en général, fausses. Si l'on prend la mesure de probabilités de Bernoulli de paramètre $p = 0$, l'événement $\{1\}$ est négligeable sans être vide, et $\{0\}$ est presque sûr sans être $\Omega = \{0, 1\}$ tout entier.

Les questions de probabilité sont souvent donnés dans un vocabulaire intuitif, en fonction *d'expériences aléatoires*. Pour transformer le problème en mathématiques, il est alors important de *modéliser* le problème, c'est-à-dire de trouver un espace probabilisé fini qui retranscrit plus ou moins fidèlement le problème.

1.8 Un exemple avec des cartes

On considère un jeu de cartes posé devant nous, faces contre la table. On fait l'expérience aléatoire où l'on tire la première carte du paquet mélangé.
On peut modéliser mathématiquement cette expérience aléatoire de plusieurs façons. À chaque fois, il s'agit de trouver un univers Ω et une mesure de probabilité $P(\cdot)$ qui « corresponde » à notre expérience. Dans cet exemple, il y a au moins deux modélisations raisonnables, qu'on présente ici.

1.8.1 Une première modélisation

Dans le premier cas, on note

$$\begin{aligned}\Omega_1 := \Big\{&\mathrm{As}^{\spadesuit}, \mathrm{Roi}^{\spadesuit}, \dots, 3^{\spadesuit}, 2^{\spadesuit},\\ &\mathrm{As}^{\clubsuit}, \mathrm{Roi}^{\clubsuit}, \dots, 3^{\clubsuit}, 2^{\clubsuit},\\ &\mathrm{As}^{\diamondsuit}, \mathrm{Roi}^{\diamondsuit}, \dots, 3^{\diamondsuit}, 2^{\diamondsuit},\\ &\mathrm{As}^{\heartsuit}, \mathrm{Roi}^{\heartsuit}, \dots, 3^{\heartsuit}, 2^{\heartsuit}\Big\}.\end{aligned}$$

C'est l'univers qu'on considère. Il possède 52 issues qui représentent les possibilités. On le munit de la probabilité uniforme P_1. L'événement (au sens intuitif) « tirer un cœur » est alors représenté par l'événement (au sens mathématique)

$$A = \Big\{\mathrm{As}^{\heartsuit}, \mathrm{Roi}^{\heartsuit}, \dots, 3^{\heartsuit}, 2^{\heartsuit}\Big\},$$

qui a 13 éléments. On a donc

$$P_1(\text{« tirer un cœur »}) = P_1(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega_1)} = \frac{13}{52} = \boxed{\frac{1}{4}}.$$

Proposition (Indépendance et produit cartésien)
Soient $E_1, \dots, E_n$ des ensembles finis non vide.
On munit $\Omega := E_1 \times E_2 \times \cdots \times E_n$ de la probabilité uniforme.
Soient $F_1 \subset E_1, \dots, F_n \subset E_n$ des parties. Les événements

$$A_i := \left\{(x_1, \dots, x_n) \in \Omega \;\middle|\; x_i \in F_i\right\} = E_1 \times \cdots \times E_{i-1} \times F_i \times E_{i+1} \times E_n$$

sont indépendants.

Démonstration. – On a déjà

$$P(A_i) = \frac{|E_1 \times \cdots \times E_{i-1} \times F_i \times E_{i+1} \times \cdots \times E_n|}{|E_1 \times \cdots \times E_{i-1} \times E_i \times E_{i-1} \times \cdots \times E_n|} = \frac{|F_i|}{|E_i|}.$$

Si $I \subset [\![1, n]\!]$ l'intersection $\bigcap_{i \in I} A_i$ est l'ensemble des n-uplets $(x_1, \dots, x_n)$ tels que $x_i \in F_i$ si $i \in I$. Si on écrit $I = \{i_1, \dots, i_r\}$ avec $i_1 < i_2 < \cdots < i_r$, alors on a

$$\begin{aligned}
P\left(\bigcap_{k=1}^{r} A_{i_k}\right) &= \frac{|E_1 \times \cdots \times E_{i_1-1} \times F_{i_1} \times \cdots \times F_{i_r} \times E_{i_r+1} \times \cdots \times E_n|}{|E_1 \times \cdots \times E_n|} \\
&= \frac{|E_1|}{|E_1|} \times \cdots \times \frac{|E_{i_1-1}|}{|E_{i_1-1}|} \times \frac{|F_{i_1}|}{|E_{i_1}|} \times \cdots \times \frac{|F_{i_r}|}{|E_{i_r}|} \times \frac{|E_{i_r+1}|}{|E_{i_r+1}|} \times \cdots \times \frac{|E_n|}{|E_n|} \\
&= \prod_{k=1}^{r} \frac{|F_{i_k}|}{|E_{i_k}|} \\
&= \prod_{k=1}^{r} P(A_{i_k}),
\end{aligned}$$

ce qui montre que les événements sont indépendants. ∎

2.6 Probabilités conditionnelles

Définition
Soit (Ω, P) *un espace probabilisé fini.*
Soient A *et* $B \in \mathscr{P}(\Omega)$ *deux événements tels que* $P(B) > 0$.
La probabilité conditionnelle de A *sachant* B*, notée* $P(A|B)$ *ou* $P_B(A)$ *est définie par*

$$P(A|B) = P_B(A) = \frac{P(A \cap B)}{P(B)}.$$

Proposition
Soit (Ω, P) *un espace probabilisé fini et* $B \in \mathscr{P}(\Omega)$ *un événement tel que* $P(B) > 0$.
Alors, l'application

$$P(\cdot|B) = P_B : \left\{ \begin{array}{ccc} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ X & \longmapsto & P_B(X) \end{array} \right.$$

est une probabilité sur Ω.
Pour cette probabilité, B *est presque sûr : on a* $P(B|B) = 1$.

Démonstration. —

- Soit $A \in \mathscr{P}(\Omega)$. On a déjà clairement $P(A|B) \geqslant 0$. La croissance de P montre que $P(A \cap B) \leqslant P(B)$, et donc que $P(A|B) \leqslant 1$. La fonction $P(\cdot|B)$ est donc bien à valeurs dans $\mathbb{R}_+$.
- On a $P(\Omega|B) = \dfrac{P(\Omega \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$. De même, $P(B|B) = \dfrac{P(B \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$.
- Soit A et $A' \in \mathscr{P}(\Omega)$ deux événements incompatibles. On a

$$\begin{aligned} P(A \sqcup A'|B) &= \frac{P((A \sqcup A') \cap B)}{P(B)} \\ &= \frac{P((A \cap B) \sqcup (A' \cap B))}{P(B)} \\ &= \frac{P(A \cap B) + P(A' \cap B)}{P(B)} \\ &= P(A|B) + P(A'|B). \end{aligned}$$

■

2.7 Formule des probabilités composées

Proposition

Soient $A_1, \ldots, A_n \in \mathscr{P}(\Omega)$ *des événements tels que* $\forall i \in [\![1, n]\!]$, $P\left(\bigcap_{i=1}^{n} A_i\right) > 0$.

Alors, on a

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times \cdots \times P\left(A_n \middle| \bigcap_{i=1}^{n-1} A_i\right).$$

2.12 Application

Question

Un test médical est appliqué à l'ensemble de la population, pour détecter une maladie dont la *prévalence* est 1% (autrement dit, 1% de la population a la maladie).
Quand un patient est malade, le test est positif avec probabilité 80%. Quand il ne l'est pas, le test est positif avec une probabilité de 10%.
Quelle est la probabilité d'être malade sachant que votre test était positif ?

Solution

On note M l'événement « être malade » et T^+ l'événement « le test est positif ».
L'énoncé donne les probabilités $P(T^+|M) = 8/10$ et $P(T^+|\overline{M}) = 1/10$. On a alors

$$P(T^+) = P(T^+|M)P(M) + P(T^+|\overline{M})P(\overline{M}) = \frac{4}{5}\frac{1}{100} + \frac{1}{10}\frac{99}{100} = \frac{107}{1000}$$

et

$$P(M|T^+) = \frac{P(T^+|M)P(M)}{P(T^+)} = \frac{4/500}{107/1000} = \frac{8}{107} \approx 7{,}48\%.$$

Chapitre 28 : Probabilités

I. 11 : Indépendance

Rq : Soient A et B évène^t
Csq $P(B) \neq 0$
Alors, A et B indépendants $\Leftrightarrow$ $P(A) = P(A|B)$

démo : $P(A|B) = \dfrac{P(A\cap B)}{P(B)} = \dfrac{P(A)\,P(B)}{P(B)} = P(A)$

Fait : $(A_1, \dots, A_n)$ mutuelle^t indépendants
$$\Downarrow$$
les A_i sont 2 à 2 indépendants

démo : Soient $i, j \in [\![1, n]\!]$ tq $i \neq j$
alors $$\underbrace{P\Big(\bigcap_{l\in\{i;j\}} A_l\Big)}_{= P(A_i \cap A_j)} = \underbrace{\prod_{l\in\{i;j\}} P(A_l)}_{= P(A_i)\,P(A_j)}$$

Quasi - partitions

Déf : E ens $\quad (A_i)_{i\in I} \in \mathcal{P}(E)^I$
I ens
On dit que $(A_i)_i$ est une quasi - partition de E
$\overset{\Delta}{\text{ssi}}$. $\forall i, j \,.\, i \neq j \Rightarrow A_i \cap A_j = \emptyset$
. $\bigcup_{i\in I} A_i = E$

Fait : Soit $E \xrightarrow{f} F$ et soit $(B_j)_{j\in J}$ une quasi - part de E
Alors : $(f^{\langle -1\rangle}[B_j])_{j\in J}$ quasi - part de E

démo : Soient $i \neq j$ dans J

Osq $f^{\langle -1 \rangle}[B_i] \cap f^{\langle -1 \rangle}[B_j] \neq \emptyset$

donc $f(x) \in B_i$ et $f(x) \in B_j$

Or $B_i \cap B_j = \emptyset$: c'est absurde

Mq $\bigcup_{j \in J} f^{\langle -1 \rangle}[B_j] = E$

déf de $\bigcup_{i \in I} A_i$ où $\forall i, A_i \subset E$

c'est $\left\{ x \in E \;\middle|\; \exists i \in I : x \in A_i \right\}$

⚠ $\bigcup_j f^{\langle -1 \rangle}[B_j] = f^{\langle -1 \rangle} \bigcup_j [B_j]$

vrai mais pas car f est linéaire

déjà : on a $\bigcup_{j \in J} f^{\langle -1 \rangle}[B_j] \subset E$

Rcpt : soit $x \in E$. On cherche $j \in J$ tq $x \in f^{\langle -1 \rangle}[B_j]$

On a $f(x) \in F$. Or $F = \bigcup_{j \in J} B_j$

Soit donc $j \in J$ tq $f(x) \in B_j$

ie $x \in f^{\langle -1 \rangle}[B_j]$

Contre-exemple : le tiré en arrière d'1 partition n'est pas une partition

$\mathbb{R} \xrightarrow{f} \mathbb{R}$ et $(\mathbb{R}_+, \mathbb{R}_-^*)$

$x \longmapsto x^2$

On a $f^{\langle -1 \rangle}[\mathbb{R}_+] = \mathbb{R}$

$f^{\langle -1 \rangle}[\mathbb{R}_-^*] = \emptyset$