Espaces probabilisés

1) Tribus et événements

a) Tribus

Def : *Une tribu* sur un ensemble Ω est une partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ telle que

(i) $\Omega \in \mathcal{A}$

(ii) $\mathcal{A}$ est stable par passage au complémentaire : si $A \in \mathcal{A}$, alors $\overline{A} = \Omega \setminus A$ appartient à $\mathcal{A}$.

(iii) $\mathcal{A}$ est stable par σ-réunion, c'est-à-dire par union finie ou dénombrable :
si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$, alors $\bigcup_{n=0}^{+\infty} A_n$ appartient à $\mathcal{A}$.

Important : Par (ii), il en est donc de même de $\bigcap_{n=0}^{+\infty} A_n$.

Terminologie : Les éléments de $\mathcal{A}$ sont appelés événements et Ω est l'univers (rarement explicité en pratique).

Def : Deux évenements sont incompatibles ssi $P(A \cap B) = 0$.

Exemples : L'ensemble $\mathcal{P}(\Omega)$ des parties de Ω est une tribu.
Par exemple, si $\Omega = \mathbb{N}$, l'ensemble des parties de $\mathbb{N}$ est une tribu sur Ω.

b) Tribu associée à une partition (finie ou) dénombrable, système complet

Soit $(B_n)_{n\in I}$ une partition finie ou dénombrable de Ω, c'est-à-dire que les B_n sont non vides, deux à deux disjoints et leur union est Ω, et I est une partie finie ou dénombrable.
Alors la plus petite tribu contenant les B_n est l'ensemble $\mathcal{A}$ des parties A de Ω de la forme

$$A = \bigcup_{i\in J} B_i, \text{ où } J \text{ est une partie de } I$$

On dit que la tribu $\mathcal{A}$ est engendré par les B_n, appelés atomes de $\mathcal{A}$.

Exemple : Si Ω est fini ou dénombrable, $\mathcal{P}(\Omega)$ est la tribu engendré par les singletons $\{a\}$, où a décrit Ω.

Terminologie : *Un système complet* $(B_n)_{n\in I}$ sur un espace probabilisable $(\Omega, \mathcal{A})$ est une famille finie ou dénombrable d'événements deux à deux incompatibles dont la réunion est Ω (cf notion de partition).

2) Probabilités

a) Définition

Def : *Une loi de probabilité* (ou *probabilité*) sur $(\Omega, \mathcal{A})$ est une application $P : \mathcal{A} \to [0, 1]$ telle que

(i) $P(\Omega) = 1$

(ii) **Propriété d'additivité** : Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles, on a

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$$

Terminologie : On dit que $(\Omega, \mathcal{A}, P)$ est un espace probabilisé.

Remarques :

- La propriété est naturellement vraie pour toute famille finie $(A_n)_{1\leq n\leq N}$.
- $P(\overline{A}) = 1 - P(A)$, où $\overline{A}$ désigne le complémentaire de A.
- Si $A \subset B$, alors $P(A) \leq P(B)$, et $P(B \smallsetminus A) = P(B) - P(A)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, car $A \cup B$ est union disjointe de A et de $B \setminus (A \cap B)$.

Propriétés :

- **Continuité croissante** :

Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements *croissante* (pour l'inclusion), alors $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n)$.

Preuve : On pose $B_n = A_n \setminus A_{n-1}$, et on a $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = P\left(\bigcup_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} P(B_n) = \lim_{n\to+\infty} P(A_n)$.

- **Continuité décroissante**

Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements *décroissante* (pour l'inclusion), alors $P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n)$.

Remarque : En particulier, pour *toute* suite $(A_n)_{n\in\mathbb{N}}$, on a $P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P\left(\bigcap_{k=0}^{n} A_k\right)$.

- *Sous-additivité* : Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'évenements, $P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n)$.

c) Cas particulier fondamental : Tribu engrendrée par une partition

Soit $(B_n)_{n\in I}$ une partition finie ou dénombrable de Ω, et $\mathcal{A}$ la tribu engendré par les atomes B_n.
Alors une probabilité sur $(\Omega, \mathcal{A})$ est entièrement définie par les $a_n = P(B_n)$, vérifiant $\sum_{n\in J} a_n = 1$.
En particulier, pour tout $J \subset I$, on a $P\left(\bigcup_{n\in J} B_n\right) = \sum_{n\in J} a_n$

d) Partie négligeable, propriété vraie presque partout

Def : Un événement $A \in \mathcal{A}$ est dit négligeable ssi $P(A) = 0$, et dit presque sûr si $P(A) = 1$.
Exemple : L'ensemble des suites presque nulles dans $\{0,1\}^{\mathbb{N}}$ est négligeable.

Def : Une propriété V définie sur Ω est vraie presque sûrement ssi $P(\{\omega \in \Omega \mid V(\omega)\}) = 1$.

3) Loi conditionnelle

a) Probabilité conditionnelle d'un événement sachant un autre événement

Def : $\boxed{\text{Si } P(B) > 0, \text{ alors la probabilité (conditionnelle) de } A \text{ sachant } B \text{ est } P(A \mid B) = \dfrac{P(A\cap B)}{P(B)}}$.
Remarque : La fonction $P_B : A \longmapsto P(A \mid B)$ est une probabilité sur $(\Omega, \mathcal{A})$.
En particulier, $P_B(A) = 1$ pour tout évenement A contenant B.

b) Formule des probabilités composées et formule de Bayes

Formule des probabilités composées : $\boxed{P(A\cap B) = P(A)P(B \mid A)}$

Formule de l'intersection $P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)P(A_2 \mid A_1)...P(A_n \mid A_1 \cap A_2 \cap ... \cap A_{n-1})$.

Remarque : Et par passage à la limite, $P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \prod_{n=0}^{+\infty} P(A_n \mid A_1 \cap A_2 \cap ... \cap A_{n-1})$.

Formule de Bayes : $P(A \mid B)P(B) = P(B \mid A)P(A)$.

c) Formule des probabilités totales

$\boxed{\text{Si } (B_n)_{n\in I} \text{ est un système complet, alors pour tout événement } A, \text{ on a } P(A) = \sum_{n\in I} P(A \mid B_n)P(B_n)}$

Preuve : A est réunion disjointe des $A \cap B_n$.

Remarque : Pour que la relation soit vraie, il suffit que la réunion (disjointe) des B_n contienne A.

Remarque : La formule de Bayes s'écrit donc $\boxed{P(A \mid B) = \dfrac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid \overline{A})P(\overline{A})}}$.

Remarque : Appliquée à une loi conditionnelle, la formule des probabilités totales s'écrit :
$P(A \mid C) = \sum_{n=0}^{+\infty} P(A \mid B_n \cap C)P(B_n \mid C)$, car $P_C(A \mid B_n) = \dfrac{P_C(A\cap B_n)}{P_C(B_n)} = \dfrac{P(A\cap B_n \cap C)}{P(B_n \cap C)}$.

4) Evénements indépendants

a) Evénements indépendants

Def : On dit que $\boxed{\text{deux événement sont indépendants ssi } P(A\cap B)=P(A)P(B)}$.

Def équivalente lorsque $P(B)>0$: $P(A\mid B)=P(A)$.

Remarque : Si A et B sont indépendants, alors leurs complémentaires $\overline{A}$ et $\overline{B}$ sont indépendants.
En effet, si $x=P(A)$ et $y=P(B)$, alors $P(\overline{A})=1-x$ et $P(\overline{B})=1-y$.
Donc $P(\overline{A}\cap\overline{B})=1-P(A)-P(B)+P(A\cap B)=1-x-y+xy=(1-x)(1-y)$.

b) Evénements mutuellement indépendants

Def : Des événements A_i, avec $i\in I$, sont *mutuellement* indépendants ssi pour toute partie finie $J\subset I$, on a

$$P\left(\bigcap_{i\in J}A_i\right)=\prod_{i\in J}P(A_i)$$

Corollaire : Si $(A_n)_{n\in\mathbb{N}}$ est une famille d'événements mutuellement indépendants, alors $P\left(\bigcap_{n=0}^{+\infty}A_n\right)=\prod_{n=0}^{+\infty}P\left(A_n\right)$.

Remarque : Trois événements deux à deux indépendants ne sont pas nécessairement mutuellement indépendants.

Exemple : Si $\sum_{n=0}^{+\infty}P(A_n)=+\infty$ et si les A_n sont indépendants, alors $P\left(\bigcap_{n=0}^{+\infty}\overline{A_n}\right)=0$.
En effet, en posant $a_n=P(A_n)$, on a $\sum_{n=0}^{+\infty}a_n=+\infty$, alors $\sum_{n=0}^{+\infty}\ln(1-a_n)=-\infty$.
On obtient ainsi $\lim_{n\to+\infty}\prod_{n=0}^{+\infty}(1-a_n)=0$, c'est-à-dire $P\left(\bigcap_{n=0}^{+\infty}\overline{A_n}\right)=0$.

5) Exemples

a) Considérons un tirage de N entiers à valeurs dans $\{0,1\}$.

Ici, l'univers Ω est l'ensemble $\{0,1\}^N$ des tirages possibles.
Une probabilité sur Ω est en fait définie par la probabilité donné à chaque tirage (cf notion d'atome)
Un événement est un ensemble de tirages : leur probabilité est la somme des probabilités des dits tirages.

Si les N tirages sont indépendants les uns des autres, alors tous les événements sont *mutuellement* indépendants.

Supposons que chacun des N entiers a une probabilité $p\in[0,1]$ d'être égal à 1, et $1-p$ d'être égal à 0.

Alors la probabilité qu'un tirage $(\varepsilon_1,...,\varepsilon_N)\in\{0,1\}^N$ vérifie $\varepsilon_1+...+.\varepsilon_N=k$ est $\binom{n}{k}p^k(1-p)^{N-k}$.

En effet, il y a $\binom{n}{k}$ tirages contenant k termes valant 1, et chacun de ces tirages a une probabilité $p^k(1-p)^{N-k}$, puisqu'il y a k termes égaux à 1 et $N-k$ termes égaux à 0.

b) Considérons un tirage d'une suite d'entiers $(\varepsilon_n)_{n\in\mathbb{N}}$ à valeurs dans $\{0,1\}$.

Ici, l'univers Ω est l'ensemble $\{0,1\}^{\mathbb{N}}$ des tirages possibles, c'est-à-dire des suites à valeurs dans $\{0,1\}$.
Un événement est un ensemble de tirages : leur probabilité est la somme des probabilités des dits tirages.

Supposons que chacun des N entiers a une probabilité $p\in[0,1]$ d'être égal à 1, et $1-p$ d'être égal à 0.

La tribu naturellement associée est la plus petite tribu contenant les $A_{m,0}=\{(\varepsilon_n)_{n\in\mathbb{N}}\mid\varepsilon_m=0\}$ (et les $A_{m,1}=\{(\varepsilon_n)_{n\in\mathbb{N}}\mid\varepsilon_m=1\}$). On a $P(A_{m,0})=1-p$ et $P(A_{m,1})=p$.
Alors, si A est l'ensemble des suites vérifiant $\varepsilon_k=0$ pour tout $0\leq k<N$, on a $P(A)=(1-p)^N$.
Considérons A l'ensemble des suites vérifiant $\varepsilon_k=0$ pour tous les k pairs.
Alors $P(A)=\lim_{k\to+\infty}(1-p)^k$, qui vaut donc 0 si $p<1$.

6) Lemme de Borel-Cantelli (*complément culturel*)

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements.

a) Posons $B = \bigcap_{n\in\mathbb{N}} B_n$, où $B_n = (\bigcup_{k\geq n} A_k)$.

Ainsi, B_n est l'ensemble des éléments $\omega \in \Omega$ qui appartiennent à (au moins) un A_k, avec $k \geq n$.

Autrement dit, B est l'ensemble des éléments $\omega \in \Omega$ tels que $\{n \in \mathbb{N} \mid \omega \in A_n\}$ est infini (= non majoré).

Montrer que si $\sum_{n=0}^{+\infty} P(A_n) < +\infty$, alors $P(B) = 0$.

Autrement dit, $\boxed{\text{si } \sum_{n=0}^{+\infty} P(A_n) < +\infty, \text{ alors presque tout élément appartient qu'à un nombre fini de } A_n}$.

b) Posons $C = \bigcup_{n\in\mathbb{N}} C_n$, où $C_n = \bigcap_{k\geq n} \overline{A_k}$.

Autrement dit, C_n est l'ensemble des éléments $\omega \in \Omega$ tels que $\omega \notin A_n$ pour n assez grand.

Comparer $\overline{B}$ et C.

Montrer que si les A_n sont indépendants et si $\sum_{n=0}^{+\infty} P(A_n) = +\infty$, alors $P(B) = 1$.

Autrement dit,

$\boxed{\text{si les } A_n \text{ sont indépendants et si } \sum_{n=0}^{+\infty} P(A_n) = +\infty, \text{ presque tout élément appartient à un nombre infini de } A_n}$.

Remarque : La propriété peut être fausse si les A_n ne sont pas indépendants, par exemple s'ils sont tous égaux !

2) *Exemples d'utilisations*

a) Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de même loi de Bernoulli $\mathcal{B}(p)$, avec $0 < p < 1$.

Montrer que presque toute suite $\omega \in \{0,1\}^{\mathbb{N}}$ contient une infinité de 1.

b) On considère une marche au hasard, c'est-à-dire $S_n = \sum_{k=0}^{n} X_k$, avec $P(X_k = 1) = p$ et $P(X_k = -1) = q = 1 - p$.

On considère l'événement $A_n : S_n = 0$. Montrer que $P(A_{2m}) = \binom{2m}{m} p^m q^m \sim \dfrac{(4pq)^m}{\sqrt{\pi m}}$.

En déduire que si la marche est biaisée, c'est-à-dire $p \neq \frac{1}{2}$, alors presque sûrement, le marcheur passe qu'un nombre fini de fois par l'origine (c'est-à-dire que presque sûrement, $\{n \in \mathbb{N} \mid S_n = 0\}$ est fini).

Corrigé

1) a) On a $P(B_n) \leq \sum_{k=n}^{+\infty} P(A_k)$. Comme $\sum P(A_k)$ converge, alors $\lim_{n\to+\infty} P(B_n) = 0$.

Par le théorème de la limite décroissante, on a donc $P(B) = \lim_{n\to+\infty} P(B_n) = 0$.

b) C est l'ensemble des $\omega \in \Omega$ tels que $\omega \notin A_n$ pour n assez grand, c'est-à-dire l'ensemble des $\omega \in \Omega$ qui n'appartiennent qu'à un nombre fini de A_n. Donc C et B sont complémentaires dans Ω.

c) On a $P\left(\bigcap_{k\geq n} \overline{A_k}\right) = \prod_{k=n}^{+\infty}(1 - a_k)$ car les $\overline{A_k}$ sont indépendants (mutuellement) et par limite décroissante.
Or, comme $\sum a_k$ diverge vers $+\infty$, alors $\prod_{k=n}^{+\infty}(1 - a_k) = 0$ pour tout $n \in \mathbb{N}$.
En effet, la série $\sum \ln(1 - a_k)$ diverge vers $-\infty$, donc il en est de même des séries $\sum_{k\geq n} \ln(1 - a_k)$.
On en déduit par limite croissante que $P(\bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} \overline{A_k}) = 0$, c'est-à-dire $P(C) = 0$. D'où $P(B) = 1$.

2) *Exemples d'utilisations*

a) On prend ici $A_n : X_n = 1$. On a $\sum_{n=0}^{+\infty} P(A_n) = +\infty$, donc par 1) c), $P(B) = 1$.
Donc presque toute suite $\omega \in \{0,1\}^{\mathbb{N}}$ contient une infinité de 1.

b) $P(A_{2k})$ correspond aux sommes $\sum_{k=0}^{2m} X_k$ contenant m termes 1 et m termes -1.
D'où (loi binomiale), $P(A_{2m}) = \binom{2m}{m} p^m q^m$. La formule de Stirling permet d'obtenir $\binom{2m}{m} p^m q^m \sim \dfrac{(4pq)^m}{\sqrt{\pi m}}$.

Lorsque $p \neq \frac{1}{2}$, alors $pq < \frac{1}{4}$, et la série $\sum P(A_{2m})$ converge.
Comme $P(A_n) = 0$ lorsque n est pair, alors la série $\sum P(A_n)$ converge. Par 1) a), on a donc $P(B) = 0$, c'est-à-dire presque sûrement, le marcheur passe un nombre fini de fois par l'origine.