# PROBABILITÉS

Les trois théorèmes importants de ce chapitre sont :

- la formule des probabilités composées ;
- la formule des probabilités totales ;
- la formule de Bayes.

## I Expériences aléatoires

Dans cette section, on ne fera pas de mathématiques ! On explicite les problèmes non mathématiques que la théorie qui suit souhaite modéliser. On n'a pas souvent l'occasion de le faire en CPGE donc il faut en profiter.

### I.1 Définition

*Définition (informelle) 1 : Expérience aléatoire.*

Une expérience aléatoire est une expérience dont l'ensemble des issues possibles est connue mais dont l'issue n'est pas connue avant de réaliser l'expérience.

**Exemples 1**

1. Une expérience aléatoire qu'on réalise fréquemment est l'expérience « `Choisir(MPSI)` ».
2. Une expérience aléatoire assez peu originale consiste à lancer deux dés verts (non pipés).

### I.2 L'expérience aléatoire « `Choisir(MPSI)` »

Réalisons ensemble l'expérience un grand nombre de fois et observons.

**Observations :**

/!\ Des tests précédents on déduit RIEN. La théorie des probabilités NE PERMET PAS de savoir si l'expérience `Choisir(MPSI)` est équitable ou pas, elle peut au mieux permettre de dire qu'il est probable qu'elle le soit (et on le fera, mais plus tard). Par contre, regarder le code source permet de s'en faire une idée, de même qu'on pourra établir le caractère équillibré d'un dé en examinant sa forme et sa composition, pas en le lançant un grand nombre de fois.

### I.3 Modélisation

**Démarche naturelle :** à une expérience aléatoire, on associe l'ensemble $\Omega$ des issues possibles, ou du moins des issues qu'on peut observer.

**Remarque 1**

Cette année, on considèrera uniquement le cas où $\Omega$ est fini. C'est très restrictif mais philosophiquement plus honnête.

Qu'est-ce qui nous intéresse ?

1. Les éléments de $\Omega$ (par exemple : les deux dés ont doné 1)
2. Mais aussi les parties de $\Omega$ (par exemple : les deux dés ont donné un résultat pair)

**Exemple 2** Pour l'expérience « `Choisir(MPSI)` », la modélisation la plus simple consiste à prendre pour $\Omega$ l'ensemble de nos élèves ($\#\Omega = 46$ )
On s'attend à ce que chaque élément de $\Omega$ soit équiprobable.

**Exemple 3** Pour l'expérience « lancé de deux dés verts (non pipés) » :

1. Une première modélisation, qui est elle-même déjà simpliste, consiste à prendre pour $\Omega$ l'ensemble $[\![1,6]\!]^2$
On s'attend à ce que chaque couple soit équiprobable. Mais ce n'est pas réaliste : on ne sait pas quel dé correspond au premier.
2. Autre modélisation : on prend pour $\Omega$ $\{\{i,j\}, i,j \in [\![1,6]\!]\}$ Mais alors on s'attend à ce que les issues de la forme $\{i,j\}$ soient deux fois moins probable que les autres.

# II Espaces probabilisés finis

## II.1 Espaces probabilisables finis

*Définition 2 .*
On appelle espace probabilisable fini un couple $(\Omega, \mathcal{P}(\Omega))$ où $\Omega$ est un ensemble fini.

**Terminologie** : les probabilistes ont coutume d'appeler :

- « univers » l'ensemble $\Omega$ ;
- « issues » les éléments de $\Omega$ ;
- « événements » les parties de $\Omega$ ;
- « événements élémentaires » les singletons de $\mathcal{P}(\Omega)$.

Si $A$ et $B$ sont deux événements d'un espace probabilisable fini, les probabilistes ont coutume d'appeler :

- « événement **contraire** »de $A$ voire (sacrilège !) « non $A$ » l'événement ${}^cA$ ;
- « $A$ ou $B$ » l'événement $A \cup B$ (horreur !) et « $A$ et $B$ » l'événement $A \cap B$ (ignominie !) ;
- « événements **incompatibles** » des événements disjoints.

Cette terminologie étant importante, elle figure dans la fiche de synthèse.

**Remarque 2**

Comme annoncé dans la remarque 1, on se limite donc au cas fini. C'est très embêtant car on s'interdit des expériences aléatoires très simples [1] dont on a envie de parler, comme par exemple :

- lancer un dé jusqu'à obtenir 6 ;
- choisir un réel uniformément au hasard entre 0 et 1 ;
- ...

On verra quand même en TD une astuce permettant d'intégrer malgré tout les expériences du type « lancer un dé jusqu'à obtenir 6 » au programme de MPSI. [2].

**Exemple 4** On a vu dans les exemples 2 et 3 comment associer des espaces probabilisables à des expériences aléatoires.

1. $(\Omega, \mathcal{P}(\Omega))$ avec $\#\Omega = 46$
2. Meth 1 $([\![1,6]\!]^2, \mathcal{P}([\![1,6]\!]^2)$ Meth 2 $(\Omega, \mathcal{P}(\Omega))$ avec $\Omega = \{\{i,j\}, i,j \in [\![1,6]\!]\}$

La définition la plus importante dans les espaces probabilisables est la suivante :

1. Mais non accessibles à l'être humain.
2. C'est comme d'habitude, on prend un gros marteau et on tape fort.

*Définition 3 .*

On appelle « système complet d'événements » (en abrégé sce) une famille $(A_1, \ldots, A_n)$ telle que

$$\begin{cases} \forall i \neq j,\ A_i \text{ et } A_j \text{ sont incompatibles;} \\ \text{et } A_1 \cup \cdots \cup A_n = \Omega\ ; \end{cases}$$

autrement dit telle que $A_1 \amalg A_2 \amalg \cdots \amalg A_n = \Omega$.

C'est quasiment la même chose qu'une partition, sauf qu'on n'impose pas que les $A_i$ soient non vides (mais dans la formule des probabilités totales ce sera impliqué par une hypothèse plus forte).

**Exemple 5** *Le plus simple des sce*

- $(\{\omega\})_{\omega\in\Omega}$
- $(A, {}^cA)$

## II.2 Probabilité

*Définition 4 .*

Soit $\Omega$ un ensemble fini. On appelle probabilité sur $\Omega$ une application $P : \mathcal{P}(\Omega) \to [0,1]$ telle que :

1. $P(\Omega) = 1$ ;
2. pour tout couple $(A, B)$ d'événements incompatibles, $P(A \amalg B) = P(A) + P(B)$.

**Terminologie** : On dit que $\big(\Omega, \mathcal{P}(\Omega), P\big)$ est un <u>espace probabilisé fini</u>.

**Remarque 3**

Lorsqu'on modélise une expérience aléatoire, on ne la modélise pas seulement par $\Omega$, mais par l'espace probabilisé $\big(\Omega, \mathcal{P}(\Omega), P\big)$.

**Exemple 6**

Tout espace probabilisable fini $(\Omega, \mathcal{P}(\Omega))$ peut être muni de sa <u>probabilité uniforme</u> $P_U : \begin{cases} \mathcal{P}(\Omega) & \to & [0,1] \\ A & \mapsto & \dfrac{|A|}{|\Omega|}. \end{cases}$

DÉMONSTRATION. 1. $P_U(\Omega) = \frac{\#\Omega}{\#\Omega} = 1$

2. Soient $A, B \in \mathcal{P}(\Omega)$ <u>incompatibles</u>.

$$\begin{aligned} P_U(A \amalg B) &= \frac{|A \amalg B|}{|\Omega|} \\ &= \frac{|A| + |B|}{|\Omega|} \\ &= \frac{|A|}{|\Omega|} + \frac{|B|}{|\Omega|} \\ &= P_U(A) + P_U(B) \end{aligned}$$

□

On a vu que cette probabilité est naturelle pour modéliser :

1. `Choisir(MPSI)`
2. Loi de 2 dés verts <u>modélisé par des couples</u>

mais aussi qu'elle est totalement à exclure si :

3. Loi de 2 dés verts <u>modélisé par des paires</u>

ou encore si :

4. Le dé est pipé ou si la pièce est truquée.

D'où un premier intérêt des autres probabilités.

*Théorème 1 : théorème de caractérisation.*

Soit $\Omega$ un ensemble fini.

1. Si $P : \mathcal{P}(\Omega) \to [0,1]$ est une probabilité alors $\sum\limits_{\omega\in\Omega} P(\{\omega\}) = 1$.
2. Réciproquement, si $p : \Omega \to [0,1]$ vérifie $\sum\limits_{\omega\in\Omega} p(\omega) = 1$ alors l'application naturellement induite $P : \begin{cases} \mathcal{P}(\Omega) & \to & [0,1] \\ A & \mapsto & \sum\limits_{\omega\in A} p(\omega) \end{cases}$ est bien définie et c'est une probabilité.

Démonstration. 1. Soit $P : \mathcal{P}(\Omega) \to [0,1]$ une probabilité

$$\begin{aligned} \sum_{\omega\in\Omega} P(\{\Omega\}) &= P\left(\coprod_{\omega\in\Omega}\{\omega\}\right) && \text{en itérant le point 2 de la déf. d'une proba} \\ &= P(\Omega) && \text{point 1} \\ &= 1 \end{aligned}$$

2. Soit $p : \Omega \to [0,1]$ tel que $\sum_{\omega\in\Omega} p(\omega) = 1$

$$P : \begin{cases} \mathcal{P}(\Omega) & \to [0,1] \\ A & \mapsto \sum_{\omega\in A} p(\omega) \end{cases}$$

est une proba

(a) $P(\Omega) = \sum_{\omega\in\Omega} p(\omega) = 1$

(b) Soient $A, B \in \mathcal{P}(\Omega)$ incompatibles

$$\begin{aligned} P(A \amalg B) &= \sum_{\omega\in A\amalg B} p(\omega) \\ &= \sum_{\omega\in A} p(\omega) + \sum_{\omega\in B} p(\omega) && \text{d'après Chasles} \\ &= P(A) + P(B) \end{aligned}$$

□

Ce théorème peut donc servir de façon très pratique à décrire plus simplement une probabilité.

**Exemple 7**

### II.3 Propriétés immédiates

*Proposition 1 .*

Soit $\big(\Omega, \mathcal{P}(\Omega), P\big)$ un espace probabilisé. On a :

0. $P(\emptyset) = 0$ ;
1. pour tout $A \in \mathcal{P}(\Omega)$, $P({}^cA) = 1 - P(A)$ ;
2. $P$ est croissante pour l'inclusion ;
3. pour tous $A,\ B \in \mathcal{P}(\Omega)$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ;

**Application 1** *: l'exemple historique.*

Est-il plus facile d'obtenir au moins un six en lançant 4 fois un seul dé ou d'obtenir au moins un double six en lançant 24 fois deux dés ? <u>4 lancés d'un dé</u>

$$\begin{cases} \Omega & = [\![1,6]\!]^4 \\ P & = P_U \end{cases}$$

On cherche $P(A)$ où

$$A = "Onaobtenuaumoinsunefoisunsix"$$

On calculera ${}^cA = "Onapasobtenudesix" = [\![1,5]\!]^4$

$$\begin{aligned} P(A) &= 1 - P({}^cA) \\ &= 1 - \frac{|{}^cA|}{|A|} \\ &= 1 - \left(\frac{5}{6}\right)^4 \\ &\approx 0.52 \end{aligned}$$

Double 6 en lançant 24 fois 2 dés

$$\begin{aligned} \Omega &= \left([\![1,6]\!]^2\right)^{24} \\ P &= P_U \\ A &= "Obteniraumoinsundouble6" \\ {}^cA &= "Nejamaisobtenirdedouble6" \\ &= ([\![1,6]\!]^2 \setminus \{6,6\})^{24} \\ P(A) &= 1 - P({}^cA) \\ &= 1 - \frac{|{}^cA|}{|\Omega|} \\ &= 1 - \frac{35^{24}}{36^{24}} \\ &\approx 0.49 \end{aligned}$$

**Conclusion** Le double 6 en lançant 24 fois 2 dés est moins probable

# III Indépendance

**Exemple 8**

Considérons l'expérience aléatoire suivante : « tirer uniformément au hasard une application $f : \{1,2\} \to \{1,2\}$ ».

Ouhlà. Prenons une minute pour comprendre de quoi ça cause.

$$
\begin{aligned}
\Omega &= \{x \mapsto 1, x \mapsto 2, \mathrm{id}, \mathrm{id}^{-1}\} \\
P &= P_U \\
P(\text{"}f(1) = 1\text{"}) &= \frac{|\{1, \mathrm{id}\}|}{|\Omega|} = \frac{2}{4} = \frac{1}{2} \\
P(\text{"}f(2) = 2\text{"}) &= \frac{|\{2, \mathrm{id}\}|}{|\Omega|} = \frac{2}{4} = \frac{1}{2} P(\text{"}f(2) = 2\text{"} \cap \text{"}f(1) = 1\text{"}) && = P(\text{"}f = \mathrm{id}\text{"}) \\
&= \frac{|\{\mathrm{id}\}|}{|\Omega|} = \frac{1}{4} \\
&= \underbrace{\frac{1}{2}}_{P(\text{"}f(1)=2\text{"}} \times \underbrace{\frac{1}{2}}_{P(\text{"}f(2)=2\text{"})}
\end{aligned}
$$

Ici, le choix de $f(1)$ ne conditionne pas celui de $f(2)$. On dira que les événements sont indépendants.

**Exemple 9**
Considérons l'expérience aléatoire suivante : « tirer uniformément au hasard une **bijection** $f : \{1, 2\} \to \{1, 2\}$ ».

$$
\begin{aligned}
P(\text{"}f(1) = 1\text{"}) &= \frac{|\{\mathrm{id}\}|}{|\Omega|} = \frac{1}{2} \\
P(\text{"}f(2) = 2\text{"}) &= \frac{|\{\mathrm{id}\}|}{|\Omega|} = \frac{1}{2}
\end{aligned}
$$

Mais :

$$
\begin{aligned}
P(\text{"}f(1) = 1\text{"} \cap \text{"}f(2) = 2\text{"}) &= \frac{|\{\mathrm{id}\}|}{|\Omega|} = \frac{1}{2} \\
&\neq \frac{1}{4} = P(\text{"}f(1) = 1\text{"}) \cdot P(\text{"}f(2) = 2\text{"})
\end{aligned}
$$

$$\textbf{On a} \quad P(A \cap B) \neq P(A) \cdot P(B)$$

On dira que les événements $A$ et $B$ ne sont pas indépendants.

*Définition 5 .*

Soit $\big(\Omega, \mathcal{P}(\Omega), P\big)$ un espace probabilisé.
Deux événements $A$ et $B$ sont dit <u>indépendants</u> lorsqu'on a $P(A \cap B) = P(A)P(B)$.

**Remarque 4**

Considérons l'expérience aléatoire « on lance un dé vert et un dé bleu ».
Faute d'information supplémentaire, on la modélise par $\big(\Omega, \mathcal{P}(\Omega), P\big)$ où $\begin{cases} \Omega &= [\![1,6]\!]^2 \\ P &= P_U \end{cases}$ Montrer que cette modélisation implique que les deux lancés sont indépendants (commencer par donner un sens à cette phrase).
On note "couleur $\to i$" := "le dé couleur donne $i$"
Soient $\begin{cases} V_i : &= \text{"le dé vert donne } i\text{"} \\ B_j : &= \text{"le dé bleu donne } j\text{"} \end{cases}$.

$$\begin{aligned}
P(V_i) &= P(\{i\} \times \{1, \ldots, 6\}) \\
&= \frac{|\{i\} \times \{1, \ldots, 6\}|}{|\Omega|} \\
&= \frac{6}{36} = \frac{1}{6} \\
P(B_j) &= P(\{j\} \times \{1, \ldots, 6\}) \\
&= \frac{|\{j\} \times \{1, \ldots, 6\}|}{|\Omega|} \\
&= \frac{6}{36} = \frac{1}{6} \\
P(V_i \cap B_j) &= P(\{i, j\}) \\
&= \frac{1}{36} \\
&= \frac{1}{6} \cdot \frac{1}{6} \\
&= P(V_i)P(B_j)
\end{aligned}$$

## III.1 Indépendance mutuelle

*Définition 6 .*

Soit $\big(\Omega, \mathcal{P}(\Omega), P\big)$ un espace probabilisé et $(A_1, \ldots, A_n)$ une famille d'événements.
On dit que <u>$A_1, \ldots, A_n$ sont mutuellement indépendants</u> lorsqu'on a :
$\forall k \in \{1, \ldots, n\},\ \forall i_1 < i_2 < \cdots < i_k \in \{1, \ldots, n\},\ P(A_{i_1} \cap A_{i_2} \cap \cdots A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$.

**Remarque 5**

1. Si $A_1, \ldots, A_n$ sont mutuellement indépendants, alors
   - ils sont indépendants 2 à 2 (*i. e.* $\forall i \neq j,\ P(A_i \cap A_j) = P(A_i)P(A_j)$) ;
   - et ils sont indépendants $n$ à $n$ (*i. e.* $P(A_1 \cap \cdots \cap A_n) = P(A_1) \times \cdots \times P(A_n)$).
2. L'indépendance 2 à 2 n'implique pas l'indépendance mutuelle.
3. L'indépendance $n$ à $n$ n'implique pas l'indépendance mutuelle.

Démonstration. 1. On suppose $A_1, A_2, \ldots, A_n$ <u>mutuellement</u> indépendants. Pour $i, j \in [\![, n]\!]$ [3] On prend $k = 2$.
$\begin{cases} i_1 &= i \\ i_2 &= j \end{cases}$ (ÀRP)

3. Notation perso (ewen) pour $[\![1, n]\!]$

$$P(A_{i_1} \cap A_{i_2}) = P(A_{i_1})P(A_{i_2})$$
$$\text{ie } P(A_i \cap A_j) = P(A_i)P(A_j)$$

2. On suppose $A_1, A_2, \ldots, A_n$ mutuellement indépendants. Pour $i, j \in [\![, n]\!]$ [4] On prend $k = 2$. $\begin{cases} i_1 & = 1 \\ & \vdots \\ i_2 & = n \end{cases}$

$$P(\bigcap_{k \in [\![, n]\!]} A_{i_k}) = \prod_{k \in [\![, n]\!]} A_{i_k} \text{ie } P(\bigcap_{k \in [\![, n]\!]} A_k) \qquad = \prod_{k \in [\![, n]\!]} A_k$$

3. On considère l'expérience aléatoire *On lance deux pièces équilibrées* $\begin{cases} \Omega \\ P \\ P(\text{"la 1ère donne T"}) = P(\text{"la 2ème donne T"}) = P(\text{"} \\ P(\text{"1ère T"} \cap \text{"2ème H"}) = P(\text{"1ère T"} \cap \text{"deux différe} \\ \implies \text{indépendance 2 à 2} \\ P(\text{"1ère T"} \cap \text{"2ème H"} \cap \text{"deux différents"}) = \frac{1}{4} \neq \frac{1}{8} \end{cases}$

4. $A_1 \neq \emptyset$, $A_2, \ldots, A_n \neq \emptyset$

$\square$

# IV Probabilités conditionnelles

L'idée est de quantifier le conditionnement d'un événement par un autre lorsqu'ils sont potentiellement non indépendants.

## IV.1 Définition

**Exemple 10** On réalise l'expérience « Choisir MPSI » mais on oublie le résultat obtenu. Cependant, **on se rappelle que c'est un garçon qui a été tiré au sort**. Est-il probable que l'élève tiré au sort soit un élève du groupe option info ?

On a envie de définir

4. Notation perso (ewen) pour $[\![1, n]\!]$

$$\begin{aligned}
P(\text{"option info"}|\text{"garçon"}) &= \frac{|\text{"option info"} \cap \text{"garçon"}|}{|\text{"garçon"}|}\\
&= \frac{\frac{|\text{"option info"}\cap\text{"garçon"}|}{46}}{\frac{|\text{"garçons"}|}{46}}\\
&\approx 0.529
\end{aligned}$$

*Définition 7 .*
Soit $\big(\Omega, \mathcal{P}(\Omega), P\big)$ un espace probabilisé, $A$ un événement et $B$ un événement de probabilité non nulle. On appelle <u>probabilité de $A$ sachant $B$</u> et on note $P_B(A)$ ou $P(A \,|\, B)$ le quotient $P_B(A) = \dfrac{P(A\cap B)}{P(B)}$.

/!\ « $A \,|\, B$ » tout seul n'a aucun sens, et en particulier ce n'est pas un événement !

Le point essentiel (et qui fait que je préfère la notation $P_B(A)$ à l'autre) est le suivant :

*Théorème 2 .*
Soit $B$ un événement de probabilité non nulle.
L'application $P_B : \left\{ \begin{array}{ccc} \mathcal{P}(\Omega) & \to & [0,1] \\ A & \mapsto & P_B(A) \end{array} \right.$ est une probabilité.

Démonstration. 1.

$$\begin{aligned}
P(\Omega|B) &= \frac{P(\Omega \cap B}{P(B)}\\
&= \frac{P(B)}{P(B)}\\
&= 1
\end{aligned}$$

2. Soient $A_1 \amalg A_2$.

$$\begin{aligned}
P(A_1 \amalg A_2|B) &= \frac{P((A_1 \amalg A_2)\cap B)}{P(B)}\\
&= \frac{P((A_1\cap B) \amalg (A_2 \cap B))}{P(B)}\\
&= \frac{P(A_1\cap B) + P(A_2\cap B)}{P(B)} && \text{car } P \text{ est une probabilité}\\
&= \frac{P(A_1\cap B)}{P(B)} + \frac{P(A_2\cap B)}{P(B)}\\
&= P(A_1|B) + P(A_2|B)
\end{aligned}$$

□

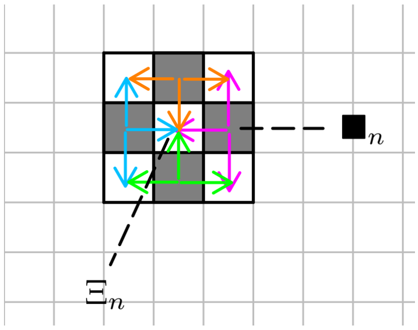
**Remarque 6**

Si $A$ et $B$ sont deux événement de probabilités non nulles alors :

1. $P(A\cap B) = P(A) \times P(B|A)$ ;
2. $P(A\cap B) = P(B) \times P(A|B)$.

**Application 2** Un bidule se déplace sur une grille de morpion (*i. e.* un échiquier $3\times 3$). À l'instant initial $n = 0$, le bidule se trouve sur la case située en haut à droite. À chaque instant $n$, le bidule se déplace d'exactement une case,

horizontalement ou verticalement (mais pas en diagonale et sans sortir de la grille), avec équiprobabilité pour chacune des cases accessibles. Quelle est la probabilité que le bidule se trouve au centre de la grille à un instant $n \in \mathbb{N} \setminus \{0\}$ ?
On admet [5] qu'on peut modéliser cette expérience aléatoire par un espace probabilisé fini $\big(\Omega, \mathcal{P}(\Omega), P\big)$.
Notons $\Xi_n :=$ "le bidule est au centre à l'issue de $n$ déplacements"
On cherche $P(\Xi_n)$
On colorie le damier de sorte que la case centrale soit blanche.
On note $\blacksquare_n :=$ "le bidule est sur une case noire après $n$ déplacements"




Par récurrence immédiate, $\forall n \in 2\mathbb{N}+1, \blacksquare_n = \Omega$
Les évènements $\Xi_n$ et $\blacksquare_n$ sont incompatibles.
Donc $n \in 2\mathbb{N}+1 \implies P(\Xi_n) = 0$
De plus $P(\Xi_0) = 0$ d'après l'énoncé
Soit $n \in 2\mathbb{N}^*$ (donc $n-1 \in 2\mathbb{N}+1$)
D'après l'énoncé, $P(\Xi_n | \blacksquare_{n-1}) = \frac{1}{3}$
Et comme $n-1 \in 2\mathbb{N}+1$, $\blacksquare_{n-1} = \Omega$
Donc

$$\begin{aligned}\frac{1}{3} &= P(\Xi_n | \blacksquare_{n-1}) \\ &= P(\Xi_n | \Omega) \\ &= \frac{P(\Xi_n \cap \Omega)}{P(\Omega)} \\ &= P(\Xi_n)\end{aligned}$$

Conclusion :

$$P(\Xi_n) = \begin{cases} 0 & \text{si} n \in (2\mathbb{N}+1) \cup \{0\} \\ \frac{1}{3} & \text{si } n \in 2\mathbb{N}^* \end{cases}$$

## IV.2 Formule des probabilités composées

*Théorème 3 : Formule des probabilités composées.*

Soit $\big(\Omega, \mathcal{P}(\Omega), P\big)$ un espace probabilisé et $(A_1, \dots, A_n)$ une famille d'événements tels que $P(A_1 \cap \dots \cap A_{n-1}) \neq 0$.
Alors :

1. $\forall k \leqslant n-1,\ P(A_1 \cap \dots \cap A_k) \neq 0$ ;
2. $P(A_1 \cap \dots \cap A_{n-1} \cap A_n) = P(A_1) P_{A_1}(A_2) P_{A_1 \cap A_2}(A_3) \cdots P_{A_1 \cap \dots \cap A_{n-1}}(A_n)$.

5. Cf TD.

DÉMONSTRATION. 1. Soit $k \leq n-1$. On a

$$\begin{aligned}
\bigcap_{1\leq i\leq k} A_i &\supset \bigcap_{1\leq i\leq n-1} A_i \\
\text{donc } P\left(\bigcap_{1\leq i\leq k} A_i\right) &\geq P\left(\bigcap_{1\leq i\leq n-1} A_i\right) \qquad \text{par croissance de } \subset \\
\text{or } &\begin{cases} P\left(\bigcap_{1\leq i\leq k} A_i\right) & \neq 0 \\ P\left(\bigcap_{1\leq i\leq n-1} A_i\right) & \in [0,1] \end{cases} \\
\text{donc } P\left(\bigcap_{1\leq i\leq k} A_i\right) &\geq P\left(\bigcap_{1\leq i\leq n-1} A_i\right) > 0
\end{aligned}$$

2.

$$\begin{aligned}
P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)\cdots P(A_n|A_1\cap\ldots\cap A_{n-1}) &= \cancel{P(A_1)}\frac{P(A_1\cap A_2)}{P(A_1)} \times\cdots\times \frac{P(A_1\cap\ldots\cap A_n)}{P(A_1\cap\ldots\cap A_{n-1})} \\
&= P(A_1\cap\ldots\cap A_n)
\end{aligned}$$

□

**Application 3** On effectue $n$ tirages successifs dans une urne qui contient initialement une boule rouge et une boule blanche. Si la boule tirée est blanche, on la remet et on passe au tirage suivant. Si la boule tirée est rouge, on la remet ainsi qu'une autre boule rouge.

1. Calculer la probabilité de ne piocher que des boules blanches.
2. Calculer la probabilité de ne piocher que des boules rouges.

On note $\begin{cases} W_k & = \text{On tire une boule blanche au } k\text{-ième tirage} \\ R_k & = \text{La } k\text{-ième boule tirée est rouge} \end{cases}$

On a $P(W_1\cap\ldots\cap W_{n-1}) \neq 0$

D'où

$$P(W_1\cap\ldots\cap W_n) = P(W_1)P(W_2|W_1)\ldots P(W_n|W_1\cap\ldots\cap W_{n-1})$$

En général,

$$P(W_k|W_1\cap\ldots\cap W_{k-1}) = \frac{1}{2}$$

D'où

$$P(W_1\cap\ldots\cap W_{n-1}) = \frac{1}{2^n}$$

On a $P(R_1\cap\ldots\cap R_{n-1}) \neq 0$

D'où d'après la FPC

$$P(R_1\cap\ldots\cap R_n) = P(R_1)P(R_2|R_1)\ldots P(R_n|R_1\cap\ldots\cap R_{n-1})$$

En général,

$$P(R_k|R_1 \cap \ldots \cap R_{k-1}) = \frac{k}{k+1}$$

D'où

$$\begin{aligned} P(R_1 \cap \ldots \cap R_n) &= \frac{1}{2} \times \frac{2}{3} \times \cdots \frac{n}{n+1} \\ &= \frac{1}{n+1} \end{aligned}$$

## IV.3 Formule des probabilités totales

*Théorème 4 : Formule des probabilités totales.*

Soit $(A_1, A_2, \ldots, A_n)$ un système complet d'événements de probabiltés non nulles et $B$ un événement. On a : $P(B) = \displaystyle\sum_{k=1}^{n} P(A_k) P_{A_k}(B)$.

**Remarque 7**

La formule reste vraie si certains $A_i$ sont de probabilité nulle avec la convention $0 \cdot P(B|A) = 0$

Démonstration.

$$\begin{aligned} P(B) &= P(B \cap \Omega) \\ &= P(B \cap (A_1 \amalg \cdots \amalg A_n)) \\ &= P(B \cap A_1 \amalg B \cap A_2 \amalg \cdots \amalg B \cap A_n) \\ &= P(B \cap A_1) + \cdots + P(B \cap A_n) \\ &= P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n) \end{aligned}$$

□

**Application 4** Notons pour $k \leq n$ $\begin{cases} E_k & = \text{la somme des } k \text{ premiers résultats est paire} \\ {}^cE_k & = \text{la somme des } k \text{ premiers résultats est impaire} \end{cases}$ $(E_{n-1}, {}^c E_{n-1})$ forme un sce d'évènements de probabilités non-nulle donc d'après la FPT

1. On lance $n \geqslant 1$ dé(s) à 20 faces. Quelle est la probabilité que la somme des numéros obtenus soit paire ?

$$\begin{aligned} P(E_k) &= P(E_{n-1})P(E_n|E_{n-1}) + P({}^cE_{n-1})P(E_n|{}^cE_{n-1}) \\ &= P(E_{n-1}) \times \frac{1}{2} + P({}^cE_{n-1}) \times \frac{1}{2} \\ &= \frac{1}{2}\left(\underbrace{P(E_{n-1}) + P({}^cE_{n-1})}_{1}\right) \\ &= \frac{1}{2} \end{aligned}$$

2. On lance $n \geqslant 1$ dé(s) à 3 faces. Quelle est la probabilité que la somme des numéros obtenus soit impaire ?

$$\begin{aligned} P({}^cE_k) &= P(E_{n-1})P({}^cE_n|E_{n-1}) + P({}^cE_{n-1})P({}^cE_n|{}^cE_{n-1}) \\ &= P(E_{n-1}) \times \frac{2}{3} + P({}^cE_{n-1}) \times \frac{1}{3} \\ &= (1 - P({}^cE_{n-1}))\frac{2}{3} + P({}^cE_{n-1})\frac{1}{3} \\ &= \frac{1}{2} \end{aligned}$$

On a une relation de récurrence sur $P({}^cE_n)$ :

$$\forall n \geq 1,\ P({}^cE_n) = -\frac{1}{3}P({}^cE_{n-1}) + \frac{2}{3}$$

C'est une suite arithmético-géométrique : $l = -\frac{1}{3}l + \frac{2}{3}$

$$\begin{aligned} P({}^cE_n) - l &= -\frac{1}{3}(P({}^cE_n) - l) \\ &= \left(-\frac{1}{3}\right)^n (P({}^cE_0) - l) \end{aligned}$$

On a $\begin{cases} l &= \frac{1}{2} \\ P({}^cE_0) &= 0 \end{cases}$ donc $P({}^cE_n) = \left(-\frac{1}{3}\right)^n \left(-\frac{1}{2}\right) + \frac{1}{2}$

## IV.4 Formule de Bayes

*Théorème 5 : Petite formule de Bayes.*

Soient $A$ et $B$ deux événements de probabiltés non nulles. On a : $P_B(A) = \dfrac{P(A)P_A(B)}{P(B)}$.

Démonstration.

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

□

**Application 5** Un questionnaire à choix multiples propose $m$ réponses pour chaque question. Soit $p$ la probabilité qu'un étudiant connaisse la bonne réponse à une question donnée. S'il ignore la réponse, il choisit au hasard l'une des réponses proposées. Quelle est la probabilité qu'un étudiant connaisse vraiment la bonne réponse lorsqu'il l'a donnée ?

Notons $\begin{cases} K & := \text{iel connaissait la bonne répons.e.s} \\ R & := \text{iel a donné.e.s la bonne répons.e.s} \end{cases}$.

$$P(K|R) = \frac{P(K)P(R|K)}{P(R)}$$

$$\begin{cases} P(R|K) & = 1 \\ P(R|^cK) & = \frac{1}{n} \end{cases}$$

$$\begin{aligned} P(R) &= P(R)P(R|K) + P(^cK)P(R|^cK) \\ &= p \cdot 1 + (1-p) \cdot \frac{1}{n} \end{aligned}$$

$$P(K|R) = \frac{p \cdot n}{p \cdot n + (A-p) \cdot 1}$$

On remarquera que, dans cet exercice, on a obtenu la probabilité que l'étudiant donne la bonne réponse en utilisant la formule des probabilités totales. Il en va souvent ainsi, d'où l'idée d'énoncer la reformulation suivante :

*Théorème 6 : Grosse formule de Bayes.*

Soit $(A_1, A_2, \ldots, A_n)$ un système complet d'événements de probabiltés non nulles et $B$ un événement de probabilité non nulle. On a, pour tout $k \in \{1, \ldots, n\}$, $P_B(A_k) = \dfrac{P(A_k)P_{A_k}(B)}{\sum\limits_{i=1}^{n} P(A_i)P_{A_i}(B)}$.

Sur l'exemple précédent, on l'a utilisée sans le dire avec $(A_1, \ldots, A_k) = (A_1, A_2)$ et $B = B$