1. *$\forall A \in \mathcal{A}, P(A) \geq 0$ (positivité)*
2. *$P(\Omega) = 1$ (totalité)*
3. *Pour toute suite (A_n) d'éléments deux à deux disjoints,*
   $P\left(\bigcup_n A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$ *(additivité)*

Remarque : avec le vocabulaire de la théorie de la mesure, P est donc une mesure positive finie, de masse totale égale à 1.

Définition 1.7 *Soit $A \in \mathcal{A}$.*
Si $P(A) = 0$, on dit que A est un événement presque impossible.
Si $P(A) = 1$, on dit que A est un événement presque certain.

Remarque : on se place dans le cas d'un univers fini, $\Omega = \{\omega_1; \cdots; \omega_n\}$, et on s'intéresse aux événements élémentaires en posant comme tribu $\mathcal{A} = \mathcal{P}(\Omega)$. La donnée de n nombres réels positifs $p_1, \cdots, p_n$ de somme égale à 1 permet de définir une probabilité P sur $\mathcal{A}$ en posant :

i $\forall i \in [\![1; n]\!], P(\{\omega_i\}) = p_i$

ii $\forall A \in \mathcal{P}(\Omega), P(A) = \sum_{i/\omega_i \in A} p_i$ (puisque P doit être additive)

Prenons le cas du jet d'un dé, où l'on pose $\Omega = [\![1; 6]\!]$.
La donnée de

$$p_1 = p_2 = p_3 = p_4 = p_5 = \frac{1}{10} \text{ et } p_6 = \frac{5}{10}$$

permet de définir une probabilité P comme décrit ci-dessus.
Si on note A l'événement « le résultat est pair », alors :

$$P(A) = \sum_{i/\omega_i \in A} p_i = p_2 + p_4 + p_6 = \frac{7}{10}$$

On procède de la même façon lorsque Ω est dénombrable, en se donnant une suite de réels positifs $(p_i)_{i\in\mathbb{N}}$ tels que la série $\sum_{i\geq 0} p_i$ converge vers 1.

Cas particulier important **: le cas d'équiprobabilité (Ω fini, de cardinal n).**

Des considérations relatives à l'expérience (dé équilibré, boules indiscernables, ...) peuvent conduire à définir une probabilité P sur $\mathcal{P}(\Omega)$ telle que :

$$P(\{\omega_1\}) = P(\{\omega_2\}) = \cdots = P(\{\omega_n\})$$

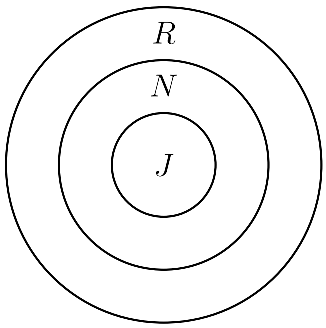
On a alors : $\forall i \in [\![1;n]\!], P(\{\omega_i\}) = \dfrac{1}{n}$, et P est entièrement déterminée, puisqu'elle est connue sur les n événements élémentaires $\{\omega_i\}$. En effet :

$$\begin{aligned}
\forall A \in \mathcal{P}(\Omega), P(A) &= P\left(\bigcup_{\omega\in A}\{\omega\}\right) \\
&= \sum_{\omega\in A} P(\{\omega\}) \text{ (additivité de } P) \\
&= \sum_{\omega\in A} \frac{1}{n} \\
&= \frac{\mathrm{Card}(A)}{n}
\end{aligned}$$

Définition 1.8 *L'application P définie ci-dessus est appelée la probabilité uniforme.*

Dans ce cas, les calculs de probabilité se ramènent à des problèmes de dénombrement, puisque $P(A) = \dfrac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}$.

Exemple de probabilité non uniforme : on lance une fléchette contre la cible suivante :



On note la couleur obtenue : on pose donc $\Omega = \{J; N; R\}$.
Si la probabilité de chacun des trois événements élémentaires est proportionnelle à l'aire de chacune des trois parties colorées, on a :

$$P(\{J\}) = \frac{1}{9}; P(\{N\}) = \frac{3}{9}; P(\{R\}) = \frac{5}{9}$$

L'exemple 1.6 constitue aussi un cas où P n'est pas uniforme.

Propriété 1.1 *Toute probabilité P possède les propriétés suivantes :*

1. $P(\emptyset) = 0$

2. $\forall (A; B) \in \mathcal{A}^2, A \subset B \Rightarrow P(A) \leq P(B)$
3. $\forall A \in \mathcal{A}, P(A) \in [0; 1]$ *et* $P(A) = 1 - P(\overline{A})$
4. *Formule du crible de Poincaré :* $\forall (A_1; \cdots; A_n) \in \mathcal{A}^n$,

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{1\leq i\leq n} P(A_i) - \sum_{1\leq i_1<i_2\leq n} P(A_{i_1} \cap A_{i_2}) \\ &+ \cdots \\ &+(-1)^{k-1} \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \\ &+ \cdots \\ &+(-1)^{n-1} P(A_1 \cap \cdots \cap A_n) \end{aligned}$$

5. *Si* $(A_i)_{i\in\mathbb{N}}$ *est une suite croissante d'événements (au sens de l'inclusion), alors la suite de réels* $(P(A_i))_{i\in\mathbb{N}}$ *est croissante, et :*

$$\lim_{n\to+\infty} P(A_n) = P\left(\bigcup_{i=0}^{+\infty} A_i\right)$$

6. *Si* $(A_i)_{i\in\mathbb{N}}$ *est une suite décroissante d'événements (au sens de l'inclusion), alors la suite de réels* $(P(A_i))_{i\in\mathbb{N}}$ *est décroissante, et :*

$$\lim_{n\to+\infty} P(A_n) = P\left(\bigcap_{i=0}^{+\infty} A_i\right)$$

Remarque : pour $n = 2$, la formule de Poincaré s'écrit :

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Pour $n = 3$, elle s'écrit :

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) \\ &-P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ &+P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

Démonstration des propriétés de P :

1. On utilise l'additivité de P :

$$\begin{aligned} P(\Omega \cup \emptyset) &= P(\Omega) + P(\emptyset) \\ \Rightarrow P(\Omega) &= P(\Omega) + P(\emptyset) \\ \Rightarrow P(\emptyset) &= 0 \end{aligned}$$

2. On décompose B : $A \subset B \Rightarrow B = (B \cap \overline{A}) \cup A$ (union disjointe), d'où :

$$\begin{array}{rcl} P(B) & = & P(B \cap \overline{A}) + P(A) \\ \Rightarrow P(B) - P(A) & = & P(B \cap \overline{A}) \\ \Rightarrow P(B) - P(A) & \geq & 0 \text{ (par positivité de } P) \end{array}$$

3. D'après la propriété précédente :

$$\begin{array}{rcl} A \subset \Omega & \Rightarrow & P(A) \leq P(\Omega) \\ & \Rightarrow & P(A) \leq 1 \end{array}$$

et d'autre part :

$$A \cup \overline{A} = \Omega \Rightarrow P(A) + P(\overline{A}) = 1$$

4. On démontre la formule par récurrence sur n (voir l'exercice 3.1 pour une autre démonstration, utilisant les propriétés de l'espérance).
Soit P_n la propriété : $\forall (A_1; \cdots; A_n) \in \mathcal{A}^n$,

$$\begin{array}{rcl} P\left(\bigcup_{i=1}^{n} A_i\right) & = & \sum_{1 \leq i \leq n} P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2}) \\ & & + \cdots \\ & & +(-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \\ & & + \cdots \\ & & +(-1)^{n-1} P(A_1 \cap \cdots \cap A_n) \end{array}$$

Pour $n = 1$: il est clair que P_1 est vraie.
Avant de montrer que : $\forall n \in \mathbb{N}, P_n \Rightarrow P_{n+1}$, on va démontrer que P_2 est vraie car la méthode sera identique pour montrer l'hérédité :

$$\begin{array}{rcl} A_1 \cup A_2 & = & A_1 \cup (A_2 \cap \overline{A}_1) \text{ (union disjointe)} \\ \Rightarrow P(A_1 \cup A_2) & = & P(A_1) + P(A_2 \cap \overline{A}_1) \\ \text{Or : } A_2 & = & (A_2 \cap A_1) \cup (A_2 \cap \overline{A}_1) \text{ (union disjointe)} \\ \Rightarrow P(A_2) & = & P(A_2 \cap A_1) + P(A_2 \cap \overline{A}_1) \\ \text{d'où : } P(A_1 \cup A_2) & = & P(A_1) + P(A_2) - P(A_1 \cap A_2) \end{array}$$

On démontre à présent que : $\forall n \in \mathbb{N}, P_n \Rightarrow P_{n+1}$.

$$\begin{array}{rcl} \bigcup_{i=1}^{n+1} A_i & = & \left(\bigcup_{i=1}^{n} A_i\right) \cup \left(A_{n+1} \cap \overline{\bigcup_{i=1}^{n} A_i}\right) \text{ (union disjointe)} \\ \Rightarrow P\left(\bigcup_{i=1}^{n+1} A_i\right) & = & P\left(\bigcup_{i=1}^{n} A_i\right) + P\left(A_{n+1} \cap \overline{\bigcup_{i=1}^{n} A_i}\right) \end{array}$$

Or A_{n+1} se décompose sous la forme suivante :

$$A_{n+1} = \left(A_{n+1} \cap \bigcup_{i=1}^{n} A_i\right) \cup \left(A_{n+1} \cap \overline{\bigcup_{i=1}^{n} A_i}\right) \text{ (union disjointe)}$$

On a donc :

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P\left(\bigcup_{i=1}^{n} A_i\right) + P(A_{n+1}) - P\left(A_{n+1} \cap (\bigcup_{i=1}^{n} A_i)\right)$$

On développe ensuite en utilisant P_n :

$$\begin{aligned} P\left(A_{n+1} \cap (\bigcup_{i=1}^{n} A_i)\right) &= P\left(\bigcup_{i=1}^{n}(A_i \cap A_{n+1})\right) \\ &= \sum_{i=1}^{n} P(A_i \cap A_{n+1}) \\ &\quad - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2} \cap A_{n+1}) + \cdots \\ &\quad + (-1)^{n-1} P(A_1 \cap \cdots \cap A_{n+1}) \end{aligned}$$

Donc finalement :

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= \sum_{1 \leq i \leq n+1} P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n+1} P(A_{i_1} \cap A_{i_2}) \\ &\quad + \cdots \\ &\quad + (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n+1} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \\ &\quad + \cdots \\ &\quad + (-1)^n P(A_1 \cap \cdots \cap A_{n+1}) \end{aligned}$$

5. La suite de réels $(P(A_n))$ est croissante (d'après la propriété 2), et majorée par 1 (d'après la propriété 3). Elle est donc convergente, et sa limite est inférieure ou égale à 1 (c'est un résultat important d'analyse : une suite croissante majorée de réels converge, à savoir démontrer).
L'idée de la démonstration repose sur l'introduction d'une suite d'événements deux à deux disjoints, ce qui va permettre d'utiliser l'additivité de P.
La suite (A_n) étant croissante (pour l'inclusion), on a, pour tout entier n, $\bigcup_{i=0}^{n} A_i = A_n$; on définit une suite d'événements $(B_i)_{i \in \mathbb{N}}$ par :

$$B_0 = A_0 \text{ et } \forall i \geq 1, B_i = A_i \backslash A_{i-1}$$

(la notation \ signifie « privé de », autrement dit $A \backslash B = A \cap \overline{B}$).
On a :

$$\forall n \geq 0, \bigcup_{i=0}^{n} A_i = \bigcup_{i=0}^{n} B_i$$

En effet : soit $x \in \bigcup_{i=0}^{n} A_i$ et soit $i_0 = \min\{i \in [\![0; n]\!] / x \in A_i\}$.

Si $i_0 = 0$, $x \in A_0 = B_0$, donc $x \in \bigcup_{i=0}^{n} B_i$.

Si $i_0 > 0$, $x \in A_{i_0}$ et $x \notin A_{i_0 - 1}$, donc $x \in B_{i_0} \subset \bigcup_{i=0}^{n} B_i$.

D'où :

$$\bigcup_{i=0}^{n} A_i \subset \bigcup_{i=0}^{n} B_i$$

L'autre inclusion est évidente puisque pour tout entier i, $B_i \subset A_i$. On en déduit, grâce à l'additivité de P :

$$P\left(\bigcup_{i=0}^{+\infty} A_i\right) = P\left(\bigcup_{i=0}^{+\infty} B_i\right) = \sum_{i=0}^{+\infty} P(B_i)$$

On revient ensuite aux sommes partielles :

$$\begin{aligned}
\sum_{i=0}^{+\infty} P(B_i) &= \lim_{n \to +\infty} \sum_{i=0}^{n} P(B_i) \\
&= \lim_{n \to +\infty} P\left(\bigcup_{i=0}^{n} B_i\right) \\
&= \lim_{n \to +\infty} P\left(\bigcup_{i=0}^{n} A_i\right) \\
&= \lim_{n \to +\infty} P(A_n)
\end{aligned}$$

6. Passer au complémentaire.

Remarque : les propriétés qui viennent d'être démontrées sont communes à toutes les mesures positives (sauf la troisième pour laquelle il faut supposer en plus que la mesure de Ω vaut 1).

Définition 1.9 *Le triplet* $(\Omega; \mathcal{A}; P)$ *est appelé espace probabilisé.*

I.4 - Probabilité conditionnelle

On se donne un événement B qui est réalisé et, disposant de cette information, on souhaite modifier la probabilité affectée aux événements A de la tribu $\mathcal{A}$ (ceux qui ont une intersection vide avec B auront alors naturellement une probabilité nulle, et B aura bien sûr une probabilité égale à 1).
On ne change donc pas d'espace probabilisable (c'est toujours le couple $(\Omega;\mathcal{A})$), on modifie seulement l'application P.

Définition 1.10 *Soit B un événement de probabilité non nulle. On appelle probabilité conditionnelle à B, ou probabilité sachant B, associée à P l'application :*

$$P_B : \begin{cases} \mathcal{A} \to \mathbb{R} \\ A \mapsto \dfrac{P(A\cap B)}{P(B)} \end{cases}$$

Notation : on note aussi $P(A/B)$ pour $P_B(A)$. On sera alors attentif au fait que c'est une simple notation, et que A/B n'est pas un événement. On utilise dans la suite indifféremment l'une de ces deux notations.

Propriété 1.2 *L'application P_B est une probabilité sur $(\Omega;\mathcal{A})$.*

Démonstration : laissée au lecteur (vérifier les trois conditions).

Cas particulier de l'équiprobabilité :
on se place dans le cas où l'univers Ω est fini, la tribu $\mathcal{A}$ est égale à $\mathcal{P}(\Omega)$ et P est la probabilité uniforme. Soit B un événement de probabilité non nulle, distinct de Ω. On a alors :

$$P_B(A) = \frac{P(A\cap B)}{P(B)} = \frac{\dfrac{\mathrm{Card}(A\cap B)}{\mathrm{Card}(\Omega)}}{\dfrac{\mathrm{Card}(B)}{\mathrm{Card}(\Omega)}} = \frac{\mathrm{Card}(A\cap B)}{\mathrm{Card}(B)}$$

On en déduit la valeur de P_B sur les événements élémentaires :

$$\begin{aligned} \forall \omega \in \overline{B}, P_B(\{\omega\}) &= 0 \\ \forall \omega \in B, P_B(\{\omega\}) &= \frac{1}{\mathrm{Card}(B)} \end{aligned}$$

La probabilité P_B n'est donc pas uniforme, mais sa restriction à $\mathcal{P}(B)$ est uniforme, et elle est bien sûr toujours nulle sur les événements élémentaires qui ne sont pas inclus dans B (voir l'exemple 1.6).

Propriété 1.3 *Soit B un événement de probabilité non nulle. Alors :*

$$\forall A \in \mathcal{A}, P(A \cap B) = P(A/B)P(B)$$

Par récurrence, on obtient, sous réserve que toutes les probabilités conditionnelles qui suivent soient bien définies, la formule des probabilités enchaînées :

$$\begin{aligned} P(A_1 \cap \cdots \cap A_n) &= P(A_1/(A_2 \cap \cdots \cap A_n))P(A_2/(A_3 \cap \cdots \cap A_n)) \\ &\quad \ldots P(A_{n-1}/A_n)P(A_n) \end{aligned}$$

I.5 - Formule des probabilités totales

Définition 1.11 *On appelle système complet d'événements toute famille d'événements $(A_i)_{i \in I}$ (I étant une partie non vide de $\mathbb{N}$) telle que :*

1. $\forall i \in I, A_i \neq \emptyset$
2. $\forall (i;j) \in I^2, i \neq j \Rightarrow A_i \cap A_j = \emptyset$
3. $\bigcup\limits_{i \in I} A_i = \Omega$

Propriété 1.4 *Soit $(A_i)_{i \in I}$ un système complet d'événements. Alors on a la formule des probabilités totales :*

$$\forall A \in \mathcal{A}, P(A) = \sum_{i \in I} P(A/A_i)P(A_i)$$

(sous réserve que pour tout i, $P(A_i)$ soit non nulle, afin que les probabilités conditionnelles soient bien définies).

Démonstration : comme $(A_i)_{i \in I}$ est un système complet, on peut écrire $A = A \cap \left(\bigcup\limits_{i \in I} A_i\right)$ et on a donc la décomposition de A sous forme d'union disjointe :

$$A = \bigcup_{i \in I} \Big(A \cap A_i \Big)$$

Par additivité de P, $P(A) = \sum\limits_{i \in I} P(A \cap A_i) = \sum\limits_{i \in I} P(A/A_i)P(A_i)$.

Remarque : on écrit souvent la formule des probabilités totales relativement au système complet $\{A_1; \overline{A}_1\}$:

$$P(A) = P(A/A_1)P(A_1) + P(A/\overline{A}_1)P(\overline{A}_1)$$