Calcul des probabilités sur un univers fini

28

Nous développons dans ce chapitre des outils pour le calcul des probabilités.
Bien que vous ayez déjà développé une excellente compréhension de ce qu'est[1] une probabilité, la principale difficulté étant de comprendre le formalisme utilisé, qui s'inscrit toujours dans le cadre[2] de la théorie des ensembles.
Il n'y aura donc aucun hasard dans la suite, nous ne ferons que le **modéliser**, et tous les dés que nous évoquerons, toutes les urnes dans lesquelles nous tirerons des boules ne seront que purement imaginaires.

[1] Ou au moins de ce que doit être.

[2] Totalement déterministe.

S'il serait extrêmement réducteur de cantonner la théorie des probabilités aux lancers de dés, tirages de boules dans des urnes et autre jeux de hasard plus ou moins tordus, il me semble que ces situations faciles à comprendre présentent tout de même un vrai intérêt pédagogique de par leur simplicité.
Je ne me priverai donc pas d'utiliser de tels exemples, même s'il faudra garder à l'esprit qu'il ne s'agit là que de *toy models*.
La physique regorge de situations intéressantes où il faut être capable de calculer des probabilités, ce qui nécessite déjà une certaine familiarité avec le concept de probabilité pour manipuler ces exemples.

Enfin, conformément aux programmes en vigueur, nous ne manipulerons cette année que des expériences qui ne possèdent qu'un nombre fini d'issues, et il vous faudra attendre l'année prochaine pour aller plus loin.

28.1 Espaces probabilisés finis

28.1.1 Univers, événement

Définition 28.1 – On appelle **univers fini** un ensemble fini Ω non vide.
Les éléments de Ω sont appelés **événements élémentaires**.

Par exemple, un lancer de pièce sera modélisé par $\{P, F\}$ ou $\{0, 1\}$, un lancer de dé sera modélisé par $\Omega = [\![1, 6]\!]$, le lancer de deux dés sera modélisé par $\Omega = [\![1, 6]\!]^2$ et le tirage d'une main de 5 cartes d'un jeu de 32 cartes sera modélisé par $\Omega = \mathscr{P}_5\left(\left\{\boxed{7\spadesuit}, \boxed{7\diamondsuit}, \dots, \boxed{A\clubsuit}, \boxed{A\heartsuit}\right\}\right)$, l'ensemble des parties à 5 éléments de l'ensemble des 32 cartes.

Notez bien que l'univers qui va modéliser l'expérience est le même que la pièce soit équilibrée ou non, nous n'avons pas encore parlé de probabilité !
Et la plupart du temps, l'univers sur lequel on travaille n'est pas complètement clair, ou tout du moins, il n'est pas nécessaire de l'expliciter pour pouvoir faire des calculs.
L'idée est que l'univers est l'ensemble de toutes les issues possibles à une expérience aléatoire, et plus l'expérience est complexe, plus l'univers risque de l'être.
Par exemple si une expérience consiste à lancer un dé, puis si on a obtenu la face portant le numéro k, à placer $2k$ boules blanches et $3k$ boules rouges dans une urne, puis en tirer avec remise k boules, on n'a pas vraiment envie d'expliciter Ω. Ce qui ne nous empêcherait pas de calculer des probabilités avec cette expérience !

La restriction sur le cardinal de Ω n'est pas une vraie limitation, même si le cas des univers infinis requiert un peu plus de précautions. Conformément au programme en vigueur, c'est le cadre auquel on se limitera en sup, afin de prendre le temps de s'habituer au formalisme. Vous manipulerez des univers infinis en spé, mais pas encore dans le cadre le plus général[3]. En particulier, nous n'étudierons pas cette année de situations où il y a une infinité d'issues possibles, par exemple le fait de lancer une pièce jusqu'à obtention du premier pile.

[3] Et par exemple vous ne rencontrerez pas en prépa les variables à densité (loi exponentielle, loi normale) que vous avez brièvement manipulées en terminale.

Dans toute la suite du chapitre, on considère un univers fini Ω fixé.

Définition 28.2 – On appelle **événement** toute partie de l'univers Ω, c'est-à-dire tout élément de $\mathscr{P}(\Omega)$.

Par abus de langage, on appellera encore événement élémentaire tout singleton $\{\omega\} \in \mathscr{P}(\Omega)$, sans vraiment distinguer ω et $\{\omega\}$.
Ainsi, un événement est un ensemble d'événements élémentaires, ou pour le dire autrement, une collection d'issues possibles de l'expérience considérée.

Exemples 28.3

- Lorsqu'on lance un dé, l'événement «obtenir un nombre pair» est $\{2, 4, 6\}$.
- Lorsqu'on lance deux dés, l'événement «la somme des dés vaut 6» est $\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}$.
- Lorsqu'on tire au main au poker, l'événement «avoir un full» est un ensemble à 1344 éléments[4] : $\left\{\left\{\boxed{A\spadesuit}, \boxed{A\diamondsuit}, \boxed{A\heartsuit}, \boxed{J\spadesuit}, \boxed{J\clubsuit}\right\}, \dots, \left\{\boxed{7\clubsuit}, \boxed{7\diamondsuit}, \boxed{7\heartsuit}, \boxed{Q\spadesuit}, \boxed{Q\heartsuit}\right\}\right\}$

[4] Voir le cours de dénombrement.

Définition 28.4 – Le complémentaire dans Ω d'un événement A est appelé le **contraire** de A, et noté $\overline{A}$.
Si A et B sont deux événements, on dit que $A \cup B$ est la disjonction de A et B, plus souvent appelé «A **ou** B», et $A \cap B$ est la conjonction de A et B, plus souvent appelé «A **et** B».
On dit que A et B sont **incompatibles** si $A \cap B = \emptyset$.

Classique
L'exemple le plus typique d'événements incompatibles est celui d'un événement A et de son contraire $\overline{A}$.

Il faut bien comprendre que $A \cup B$ est l'événement «l'un au moins des deux événements A et B est réalisé».
Et de même, $A \cap B$ est l'événement «A et B sont réalisés simultanément».
Pour rester sur des exemples simples, lançons un dé, et soit $A = \{2, 4, 6\}$ l'événement «obtenir un numéro pair» et $B = \{4, 5, 6\}$ l'événement «obtenir un numéro supérieur ou égal à 4».
Alors $A \cap B = \{4, 6\}$ est bien l'événement «obtenir un numéro pair supérieur ou égal à 4», alors que $A \cup B = \{2, 4, 5, 6\}$ contient bien toutes les issues de l'expérience qui réalisent soit A soit B (soit les deux en même temps).
Des événements incompatibles sont donc des événements qui ne peuvent pas se réaliser en même temps.

Exemples 28.5

Considérons une urne, qui contient a boules blanches et b boules noires. On tire n boules de cette urne, avec remise entre les tirages.
Pour tout $i \in [\![1, n]\!]$, notons B_i (resp. N_i) l'événement «la $i^{\text{ème}}$ boule tirée est blanche (resp. noire)».
Alors :

- $B_i = \overline{N_i}$
- $\displaystyle\bigcup_{i=1}^{n} N_i$ est l'événement «obtenir au moins une boule noire»
- $\displaystyle\bigcap_{i=1}^{n} B_i$ est l'événement «n'obtenir que des blanches», contraire du précédent

Univers
Un univers possible pour cette expérience est $\{B, N\}^n$, l'ensemble des suites de n éléments à valeurs dans $\{B, N\}$ (pour Noir et Blanc).
B_i est alors l'ensemble de tous les n-uplets dont le $i^{\text{ème}}$ élément est un B.
Vous remarquerez que le nombre de boules n'entre pas en jeu tant qu'on n'a pas parlé de probabilité de chacun des événements.

- $\bigcup_{i=1}^{n-1}(B_i \cap B_{i+1})$ est l'événement «obtenir deux boules blanches consécutives».
- les événements

$$\overline{\left(\bigcup_{i=1}^{n-1}(B_i \cap B_{i+1})\right) \cup \left(\bigcup_{i=1}^{n-1}(N_i \cap N_{i+1})\right)} \text{ et } \left(\bigcap_{\substack{i=1\\ i \text{ pair}}}^{n} N_i \cap \bigcap_{\substack{j=1\\ j \text{ impair}}}^{n} B_j\right) \cup \left(\bigcap_{\substack{i=1\\ i \text{ pair}}}^{n} B_i \cap \bigcap_{\substack{j=1\\ j \text{ impair}}}^{n} N_j\right)$$

sont égaux : ce sont tous les deux l'événement «deux boules consécutives sont de couleurs différentes».

Définition 28.6 – On appelle **système complet d'événements** tout ensemble $\{A_1, \dots, A_n\}$ d'événements, deux à deux incompatibles, et tels que $\bigcup_{i=1}^{n} A_i = \Omega$.
Autrement dit, un système complet d'événements est une partition de Ω, sauf que l'on ne demande pas aux A_i d'être non vides.

Abréviation
Il est courant d'abréger *système complet d'événements* en *s.c.e.*.
Bien entendu cette abréviation n'est pas tolérée dans une copie... du moins c'est le discours officiel.

Exemples 28.7

- Si A est un événement, alors $\{A, \overline{A}\}$ est un système complet d'événements.
- $\{\{\omega\}, \omega \in \Omega\}$ est un système complet d'événements.
- Lors d'un lancer de dé, si on note A l'événement «obtenir un numéro pair» et B l'événement «obtenir un numéro impair», alors $\{A, B\}$ est un système complet d'événements.
- Lors du lancer de deux dés, si on note A_i l'événement «la somme des deux dés vaut i», alors $\{A_i, 2 \leqslant i \leqslant 12\}$ est un système complet d'événements.

28.1.2 Probabilités

Définition 28.8 – Soit Ω un univers fini. On appelle **probabilité** sur Ω toute application $\mathbf{P} : \mathscr{P}(\Omega) \to [0, 1]$ telle que :

- $\mathbf{P}(\Omega) = 1$
- $\forall (A, B) \in \mathscr{P}(\Omega)^2,\ A \cap B = \emptyset \Rightarrow \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$ *(additivité).*

Pour $A \in \mathscr{P}(\Omega)$, le réel $\mathbf{P}(A)$ est appelé probabilité de l'événement A.
On appelle alors **espace probabilisé fini**[5] un couple $(\Omega, \mathbf{P})$ où Ω est un univers fini et $\mathbf{P}$ une probabilité sur Ω.

⚠ Attention !
Ne pas confondre la probabilité $\mathbf{P}$, qui est une application, de la probabilité $\mathbf{P}(A)$ d'un événement A, qui est un réel.

[5] En abrégé «espace probabilisé» puisque cette année nous ne parlerons que du cas fini.

Remarque. Demander que $\mathbf{P}$ soit à valeurs dans $\mathbf{R}_+$ plutôt que $[0, 1]$ suffirait, car alors pour tout $A \in \mathscr{P}(\Omega)$, $1 = \mathbf{P}(\Omega) = \mathbf{P}(A) + \underbrace{\mathbf{P}(\overline{A})}_{\geqslant 0}$ et donc $\mathbf{P}(A) \leqslant 1$.

Quand on lance un dé, c'est au moment de choisir la probabilité que l'on met sur $\llbracket 1, 6 \rrbracket$ que l'on décide si notre dé est ou non équilibré.
Par exemple, dans le cas[6] où le dé tombe absolument toujours sur 6, la probabilité qu'on utilise sur $\mathscr{P}(\llbracket 1, 6 \rrbracket)$ est définie par

$$\mathbf{P}(A) = \begin{cases} 0 & \text{si } 6 \notin A \\ 1 & \text{si } 6 \in A \end{cases}$$

[6] Un peu extrême et complètement inintéressant je vous le concède.

Je ne prouve pas qu'il s'agit bien là d'une probabilité, mais nous y reviendrons.

L'exemple le plus naturel de probabilité, qui est celui qui nous vient en premier à l'esprit est celui de probabilité uniforme, qui correspond à ce qu'on appelle une situation d'équiprobabilité.

Proposition 28.9 : *Soit Ω un univers fini. Alors il existe une unique probabilité $\mathbf{P}$ sur Ω telle que $\omega \mapsto \mathbf{P}(\{\omega\})$ soit constante*[7].
*Cette probabilité, appelée **probabilité uniforme sur** Ω est donnée par*

$$\forall A \in \mathscr{P}(\Omega),\ \mathbf{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}.$$

[7] C'est-à-dire telle que tous les événements élémentaires (= toutes les issues de l'expérience) soient de même probabilité.

Démonstration. Supposons qu'une telle probabilité $\mathbf{P}$ existe, et soit $\lambda \in [0,1]$ la valeur commune à tous les $\mathbf{P}(\{\omega\})$, $\omega \in \Omega$.
Alors

$$P(\Omega) = P\left(\bigcup_{\omega\in\Omega}\{\omega\}\right) = \sum_{\omega\in\Omega}\mathbf{P}(\{\omega\}) = \sum_{\omega\in\Omega}\lambda = \lambda\text{Card}(\Omega).$$

Puisque $\mathbf{P}(\Omega) = 1$, on a donc $\lambda = \dfrac{1}{\text{Card}(\Omega)}$.
Et en particulier, pour tout événement $A \in \mathscr{P}(\Omega)$, il vient

$$\mathbf{P}(A) = \mathbf{P}\left(\bigcup_{\omega\in A}\{\omega\}\right) = \sum_{\omega\in A}\mathbf{P}(\{\omega\}) = \text{Card}(A)\frac{1}{\text{Card}(\Omega)}.$$

À présent, définissons une application $\mathbf{P}$ sur $\mathscr{P}(\Omega)$ par $\mathbf{P}(A) = \dfrac{\text{Card}(A)}{\text{Card}(\Omega)}$.
Alors il est évident que $\mathbf{P}$ est à valeurs dans $[0,1]$ (puisqu'un cardinal est positif, et que le cardinal d'une partie est inférieur ou égal à celui de l'ensemble tout entier), et on a bien $P(\Omega) = \dfrac{\text{Card}(\Omega)}{\text{Card}(\Omega)} = 1$.
Enfin, si A et B sont deux événements incompatibles,

$$\mathbf{P}(A \cup B) = \frac{\text{Card}(A \cup B)}{\text{Card}(\Omega)} = \frac{\text{Card}(A) + \text{Card}(B)}{\text{Card}(\Omega)}.$$

Donc $\mathbf{P}$ est bien une probabilité sur Ω. □

Évident ?
Le fait de passer de la proba de l'union à la somme des probas doit vous sembler logique, la vraie justification se trouve à la proposition suivante.

Exemples 28.10

► Au poker, si l'on considère que toutes les mains sont équiprobables[8], la probabilité d'avoir un carré est $\dfrac{224}{201\,316} \simeq 0.00111$.
► Si je choisis au hasard un élève de MP2I, la probabilité qu'il soit né un 1er janvier est $\frac{1}{46}$.

[8] Ce qui semble légitime si vous ne vivez pas dans un album de Lucky Luke.

Théorème 28.11 (Calculs avec des probabilités) : *Soit $\mathbf{P}$ une probabilité sur Ω. Alors :*

1. $\mathbf{P}(\emptyset) = 0$
2. $\forall A \in \mathscr{P}(\Omega)$, $\mathbf{P}\left(\overline{A}\right) = 1 - \mathbf{P}(A)$
3. $\forall (A,B) \in \mathscr{P}(\Omega)^2$, $A \subset B \Rightarrow \mathbf{P}(A) \leqslant \mathbf{P}(B)$ *(croissance de la probabilité)*
4. $\forall (A,B) \in \mathscr{P}(\Omega)^2$, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.
5. $\forall (A_1,\dots,A_n) \in \mathscr{P}(\Omega)^n$, $\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) \leqslant \sum_{i=1}^n \mathbf{P}(A_i)$ *(sous-additivité).*
   Si de plus les A_i sont deux à deux disjoints, alors on a égalité :

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

Autrement dit
$\mathbf{P}$ est une application croissante de l'ensemble ordonné $(\mathscr{P}(\Omega), \subset)$ dans $([0,1], \leqslant)$.

Démonstration. 1. On a $\Omega \cap \emptyset = \emptyset$, et donc par additivité,

$$\mathbf{P}(\Omega) = \mathbf{P}(\Omega \cup \emptyset) = \mathbf{P}(\Omega) + \mathbf{P}(\emptyset) \Leftrightarrow \mathbf{P}(\emptyset) = 0.$$

2. On a $1 = \mathbf{P}(\Omega) = \mathbf{P}(A \cup \overline{A}) = \mathbf{P}(A) + \mathbf{P}(\overline{A})$ car A et $\overline{A}$ sont incompatibles.
Et donc $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$.
3. Si $A \subset B$, alors $B = A \cup (B \setminus A)$, et l'union est disjointe.
Donc $\mathbf{P}(B) = \mathbf{P}(A) + \underbrace{\mathbf{P}(B \setminus A)}_{\geqslant 0} \geqslant \mathbf{P}(A)$.
4. Puisque $A \cup B = B \cup (A \setminus B)$, il vient $\mathbf{P}(A \cup B) = \mathbf{P}(B) + \mathbf{P}(A \setminus B)$.
Mais $A = (A \setminus B) \cup (A \cap B)$, et encore une fois il s'agit d'une union disjointe, donc

$$\mathbf{P}(A) = \mathbf{P}(A \setminus B) + \mathbf{P}(A \cap B) \Leftrightarrow \mathbf{P}(A \setminus B) = \mathbf{P}(A) - \mathbf{P}(A \cap B)$$

de sorte qu'on a bien $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.
5. Par récurrence sur n, le nombre d'événements prouvons $\mathscr{P}(n)$:
«$\forall (A_1, \dots, A_n) \in \mathscr{P}(\Omega)^n$, $\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} \mathbf{P}(A_i)$ et si les A_i sont deux à deux incompatibles, alors $\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbf{P}(A_i)$.»
Pour $n = 2$, on a bien $\mathbf{P}(A_1 \cup A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2) - \mathbf{P}(A_1 \cap A_2) \leqslant \mathbf{P}(A_1) + \mathbf{P}(A_2)$.
Et le cas d'incompatibilité découle directement de la définition de probabilité.
Supposons donc $\mathscr{P}(n)$ vraie, et soient $A_1, \dots, A_{n+1} \in \mathscr{P}(\Omega)$. Alors

$$\mathbf{P}\left(\bigcup_{i=1}^{n+1} A_i\right) = \mathbf{P}\left(A_1 \cup \bigcup_{i=2}^{n+1} A_i\right) \leqslant \mathbf{P}(A_1) + \mathbf{P}\left(\bigcup_{i=2}^{n+1} A_i\right) \leqslant \mathbf{P}(A_1) + \sum_{i=2}^{n+1} \mathbf{P}(A_i) \leqslant \sum_{i=1}^{n+1} \mathbf{P}(A_i).$$

Dans le cas où les A_i sont deux à deux incompatibles, alors A_1 est incompatible avec $\bigcup_{i=2}^{n+1} A_i$, et donc les deux inégalités ci-dessus sont des égalités[9].
Ainsi, $\mathscr{P}(n+1)$ est vraie, et par le principe de récurrence, la propriété est vraie pour tout n.

□

[9] La première par définition d'une probabilité, la seconde par hypothèse de récurrence.

Remarques. ▶ La signification de l'inclusion n'est pas complètement évidente au premier abord. Si $A \subset B$, cela signifie que tous les événements élémentaires qui composent A sont dans B.
Autrement dit, que **dès que A est réalisé, alors B l'est automatiquement.** C'est vraiment ainsi qu'il faut comprendre l'inclusion.
Dès lors, la croissance de la probabilité se comprend bien : B doit donc être plus souvent réalisé que A et donc avoir une probabilité supérieure à celle de A.
▶ Le point 4. concernant la probabilité d'une union ne doit pas du tout vous surprendre dans le cas d'équiprobabilité, puisque c'est alors une conséquence directe de

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B).$$

▶ Sur le même principe que pour les cardinaux, on peut généraliser le point 3 :

$$\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(A \cap B) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) + \mathbf{P}(A \cap B \cap C)$$

et il existe une formule du crible[10] qui est exactement la même que celle vue en TD pour les cardinaux, mais en changeant les cardinaux en probabilités.

[10] Hors-programme.

28.1.3 Construction de probabilités

Une conséquence facile du dernier point de la proposition qui précède est que si $\mathbf{P}$ est une probabilité, alors $\sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) = 1$.
Nous nous intéressons ici à une réciproque à ce résultat :

Proposition 28.12 : *Soit $\Omega = \{\omega_1, \dots, \omega_n\}$ un ensemble fini, et soient $p_1, \dots, p_n$ des réels positifs tels que $\sum_{i=1}^n p_i = 1$.*
Alors il existe une unique probabilité $\mathbf{P}$ sur Ω telle que $\forall i \in [\![1,n]\!]$, $\mathbf{P}(\{\omega_i\}) = p_i$.

Remarque
Si les p_i sont des nombres **positifs** de somme 1, alors tous sont compris entre 0 et 1.

Démonstration. Pour l'unicité, commençons par noter que pour tout $A \in \mathscr{P}(\Omega)$, $\mathbf{P}(A) = \sum_{x \in A} \mathbf{P}(\{x\})$, et donc si deux probabilités coïncident sur tous les événements élémentaires, alors elles sont égales.

Considérons à présent l'application $\mathbf{P}$ définie sur $\mathscr{P}(\Omega)$ par $\mathbf{P}(A) = \sum_{i=1}^n p_i \mathbb{1}_A(\omega_i)$.

Autrement dit
Les p_i qui apparaissent dans la somme sont ceux pour lequels $\omega_i \in A$.

Alors $\mathbf{P}(\Omega) = \sum_{i=1}^n p_i = 1$. Si A et B sont deux événements incompatibles, nous savons que $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \underbrace{\mathbb{1}_{A\cap B}}_{=0}$, et donc

$$\mathbf{P}(A \cup B) = \sum_{i=1}^n p_i \mathbb{1}_{A\cup B}(\omega_i) = \sum_{i=1}^n p_i \mathbb{1}_A(\omega_i) + \sum_{i=1}^n p_i \mathbb{1}_B(\omega_i) = \mathbf{P}(A) + \mathbf{P}(B).$$

Donc $\mathbf{P}$ est bien une probabilité, qui a clairement les propriétés requises. □

Ce que légitime cette proposition, c'est que se donner une probabilité sur Ω revient à se donner les probabilités de tous les événements élémentaires, sous l'hypothèse très raisonnable que ces probabilités soient bien positives et de somme 1.
Par exemple, il n'existe qu'une mesure de probabilité sur $[\![1,6]\!]$ qui permet de simuler un lancer d'un dé qui tomberait sur $2, 4$ ou 6 avec probabilité $\frac{1}{6}$, sur 1 ou 3 avec probabilité $\frac{1}{10}$ et sur 5 avec probabilité $\dfrac{3}{10}$.

28.2 PROBABILITÉS CONDITIONNELLES

28.2.1 Définition

Définition 28.13 – Soient A et B deux événements d'un espace probabilisé fini $(\Omega, \mathbf{P})$, tels que $\mathbf{P}(A) \neq 0$.
On appelle **probabilité conditionnelle de** B **sachant** A (en abrégé probabilité de B sachant A) et on note $\mathbf{P}_A(B)$ (ou parfois $\mathbf{P}(B|A)$) le réel

$$\mathbf{P}_A(B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}.$$

Le plus dur ici est de bien comprendre que cette probabilité correspond à l'intuition qu'on s'en fait lorsque parle de probabilité de B *sachant* A.
Le cas qu'on comprend le mieux est celui d'équiprobabilité.

Exemple 28.14

Je suis convoqué pour faire passer un oral dans un lycée bien particulier, dont le site Web m'indique que tous les élèves sont en sport étude, soit en section escalade (26 élèves), soit en section bridge (47 élèves), et qu'il compte dans ses rangs 13 médaillés aux mondiaux junior (8 en escalade et 5 en bridge).
Avant le premier oral, je fais donc un petit calcul qui m'apprend que la probabilité que le prochain candidat soit un médaillé mondial est $\dfrac{13}{73}$.
J'ouvre la porte, le candidat entre. Ses biceps et son bronzage ne m'autorisent décemment pas à le penser en section bridge. Quelle est donc la probabilité qu'il s'agisse d'un médaillé mondial ? C'est-à-dire, la probabilité, *sachant que c'est un*

grimpeur, qu'il s'agisse d'un médaillé mondial ?
La réponse est évidente, c'est $\frac{8}{26}$, le nombre de médaillés en escalade sur le nombre total de grimpeurs, mais si on réfléchit en termes de probabilités, $\frac{8}{26} = \frac{\frac{8}{77}}{\frac{26}{77}}$, où $\frac{8}{77}$ est la probabilité qu'un élève choisi au hasard soit **à la fois** grimpeur et médaillé, et $\frac{26}{77}$ est la probabilité qu'un élève choisi au hasard soit grimpeur. On trouve donc bien la formule $\frac{\mathbf{P}(A\cap B)}{\mathbf{P}(A)}$.

La notation $\mathbf{P}(B|A)$ (qu'on essaiera autant que possible d'éviter, bien qu'elle figure au programme officiel) est trompeuse, et laisse penser qu'il existe un événement $B|A$, qui serait «B sachant A».
Un tel «événement conditionnel» **n'existe pas**, et lorsque vous aurez envie de manipuler de tels événements[11], c'est en fait bien souvent $A\cap B$ que vous aurez derrière la tête.

[11] Ce que vous vous abstiendrez bien entendu de faire !

Proposition 28.15 : *Si A est un événement tel que $\mathbf{P}(A) \neq 0$, alors l'application* $\mathbf{P}_A : \left|\begin{array}{ccc} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ B & \longmapsto & \mathbf{P}_A(B)\end{array}\right.$ *est une probabilité sur Ω.*

Une conséquence importante de ce fait est que toutes les règles énoncées pour le calcul des probabilités restent valables pour le calcul des probabilités conditionnelles !
Par exemple, $\mathbf{P}_A(\emptyset) = 0$, $\mathbf{P}_A\left(\overline{B}\right) = 1 - \mathbf{P}_A(B)$ et $\mathbf{P}_A(B\cup C) = \mathbf{P}_A(B) + \mathbf{P}_A(C) - \mathbf{P}_A(B\cap C)$.

Démonstration. Puisque la probabilité $\mathbf{P}$ est positive, croissante[12] et que $A\cap B\subset A$, on a bien pour tout $B\in\mathscr{P}(\Omega)$, $\mathbf{P}_A(B) = \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(A)} \in [0,1]$.
Par ailleurs, $\mathbf{P}_A(\Omega) = \frac{\mathbf{P}(A\cap\Omega)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1$.
Enfin, si B et C sont deux événements incompatibles, alors $A\cap B$ et $A\cap C$ le sont encore. Donc

$$\mathbf{P}_A(B\cup C) = \frac{\mathbf{P}(A\cap(B\cup C))}{\mathbf{P}(A)} = \frac{\mathbf{P}\left((A\cap B)\cup(A\cap C)\right)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A\cap B)+\mathbf{P}(A\cap C)}{\mathbf{P}(A)} = \mathbf{P}_A(B)+\mathbf{P}_A(C).$$

□

[12] Pour l'inclusion.

28.2.2 La formule des probabilités composées

Sur la définition de probabilité conditionnelle, il est assez clair que $\mathbf{P}(A\cap B) = \mathbf{P}(A)\mathbf{P}_A(B)$.
Ce résultat se généralise de la manière suivante :

Théorème 28.16 (Formule des probabilités composées) : *Soient $A_1,\dots,A_n$ des événements d'un espace probabilisé fini $(\Omega,\mathbf{P})$, tels que $\mathbf{P}(A_1\cap\cdots\cap A_{n-1})\neq 0$. Alors*

$$\begin{aligned}\mathbf{P}(A_1\cap\cdots\cap A_n) &= \mathbf{P}(A_1)\mathbf{P}_{A_1}(A_2)\mathbf{P}_{A_1\cap A_2}(A_3)\cdots\mathbf{P}_{A_1\cap\cdots\cap A_{n-1}}(A_n)\\ &= \prod_{i=0}^{n-1}\mathbf{P}_{A_1\cap\cdots\cap A_i}(A_{i+1}).\end{aligned}$$

Remarque
La formule s'apprend et se comprend bien mieux avec les pointillés.

Démonstration. Notons que puisque pour tout $i\in[\![1,n-1]\!]$, $A_1\cap\cdots\cap A_{n-1}\subset A_1\cap\cdots\cap A_i$, et donc par croissance de la probabilité,

$$\mathbf{P}(A_1\cap\cdots\cap A_i) \geqslant \mathbf{P}(A_1\cap\cdots\cap A_{n-1}) > 0$$

de sorte que la probabilité conditionnelle $\mathbf{P}_{A_1\cap\cdots\cap A_i}(A_{i+1})$ est bien définie[13].
Sinon, il s'agit de remarquer que le produit donné dans l'énoncé est télescopique :

$$\prod_{i=0}^{n-1}\mathbf{P}_{A_1\cap\cdots\cap A_i}(A_{i+1})$$

[13] C'est d'ailleurs la seule raison d'être de cette hypothèse, qui ne pose en pratique aucun soucis.

$$\begin{aligned}&= \cancel{\mathbf{P}(A_1)}\frac{\cancel{\mathbf{P}(A_1\cap A_2)}}{\cancel{\mathbf{P}(A_1)}}\frac{\cancel{\mathbf{P}(A_1\cap A_2\cap A_3)}}{\cancel{\mathbf{P}(A_1\cap A_2)}}\cdots\frac{\cancel{\mathbf{P}(A_1\cap\cdots\cap A_{n-1})}}{\cancel{\mathbf{P}(A_1\cap\cdots\cap A_{n-2})}}\frac{\mathbf{P}(A_1\cap\cdots\cap A_n)}{\cancel{\mathbf{P}(A_1\cap\cdots\cap A_{n-1})}}\\&= \mathbf{P}(A_1\cap\cdots\cap A_n).\end{aligned}$$

□

Cette formule est absolument fondamentale lorsqu'on manipule des expériences successives telles que le résultat d'une expérience influe sur la suivante. L'exemple typique étant les tirages successifs sans remise.

Exemple 28.17

On effectue des tirages sans remise dans une urne qui contient 2 boules rouges et $n-2$ boules vertes.
Pour $k \in [\![1, n-1]\!]$, on note A_k l'événement «la première boule rouge sort au $k^{\text{ème}}$ tirage».
Notons également R_i l'événement «la boule obtenue au $i^{\text{ème}}$ tirage est rouge».
Nous ne disposons pas directement des probabilités des R_i, mais plutôt de leurs probabilités conditionnelles connaissant les résultats des tirages précédents !

On a $A_k = \left(\displaystyle\bigcap_{i=1}^{k-1}\overline{R_i}\right)\cap R_k$.

Et donc

$$\begin{aligned}\mathbf{P}(A_k) &= \mathbf{P}\left(\left(\bigcap_{i=1}^{k-1}\overline{R_i}\right)\cap R_k\right)\\&= \mathbf{P}(\overline{R_1})\mathbf{P}_{\overline{R_1}}(\overline{R_2})\mathbf{P}_{\overline{R_1}\cap\overline{R_2}}(\overline{R_3})\cdots\mathbf{P}_{\overline{R_1}\cap\cdots\cap\overline{R_{k-2}}}(\overline{R_{k-1}})\mathbf{P}_{\overline{R_1}\cap\cdots\cap\overline{R_{k-1}}}(R_k)\\&= \frac{n-2}{n}\frac{n-3}{n-1}\frac{n-4}{n-2}\cdots\frac{n-2-(k-2)}{n-(k-2)}\frac{2}{n-(k-1)}\\&= \frac{2(n-k)}{n(n-1)}.\end{aligned}$$

28.2.3 La formule des probabilités totales

La formule des probabilités totales sert tout le temps, et vous l'avez déjà beaucoup utilisée, en vous gardant bien de le dire : elle sert précisément à faire des distinctions de cas afin de se ramener à des cas où les calculs de probabilités sont plus simples, puis de «remettre ensemble» toutes les probabilités obtenues.

Théorème 28.18 (Formule des probabilités totales) : *Soit $\{A_1, \dots, A_n\}$ un système complet d'événements. Alors pour tout $B \in \mathcal{P}(\Omega)$,*

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

Si de plus, tous les A_i sont de probabilités non nulles, alors

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

Méthode
Il faut connaître les deux formes données ici, et être capable de décider laquelle on souhaite utiliser dans une situation concrète. Ce qui dépend essentiellement de ce que vous pouvez calculer le plus facilement : des probabilités d'intersection ou des probabilités conditionnelles ?

Démonstration. Puisque $\Omega = \displaystyle\bigcup_{i=1}^{n}A_i$, et que ces événements sont deux à deux incompatibles, alors $B = B\cap\Omega = \displaystyle\bigcup_{i=1}^{n}(B\cap A_i)$, et les $B\cap A_i$ sont deux à deux incompatibles.
Donc il vient $\mathbf{P}(B) = \displaystyle\sum_{i=1}^{n}\mathbf{P}(B\cap A_i)$.

La seconde forme, dans le cas où les $\mathbf{P}(A_i)$ sont tous non nuls s'en déduit immédiatement en notant que $\mathbf{P}(B \cap A_i) = \mathbf{P}(A_i)\mathbf{P}_{A_i}(B)$. □

Si certains des $\mathbf{P}(A_i)$ sont nuls, ce qui normalement ne devrait pas trop arriver sur un univers fini, on a

$$\mathbf{P}(B) = \sum_{i \mid \mathbf{P}(A_i) \neq 0} \mathbf{P}(A_i)\mathbf{P}_{A_i}(B).$$

Exemples 28.19

► On lance 3 fois une pièce qui tombe sur *pile* avec probabilité p et sur *face* avec probabilité $q = 1 - p$.
On note P_i (resp. F_i) l'événement «le $i^{\text{ème}}$ lancer donne pile (resp. face)». On note A l'événement «deux lancers consécutifs n'ont jamais donné le même résultat».
Utilisons alors la formule des probabilités totales à l'aide du système complet d'événements $\{P_2, F_2\}$.

$$\mathbf{P}(A) = \mathbf{P}(A \cap P_2) + \mathbf{P}(A \cap F_2) = \mathbf{P}(F_1 \cap P_2 \cap F_3) + \mathbf{P}(P_1 \cap F_2 \cap P_3) = q^2 p + p^2 q.$$

Remarque

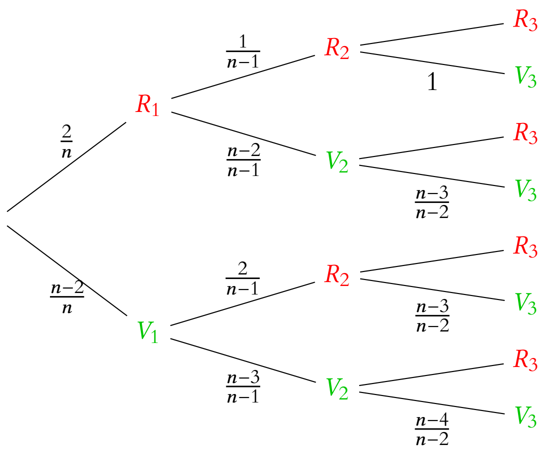
La véritable justification des produits, qui vient de l'indépendance des lancers, viendra un peu plus tard.

► Une urne contient $n \geqslant 3$ boules : 2 boules rouges et $n-2$ boules vertes. On tire successivement et sans remise 3 boules. Quelle est la probabilité que la troisième boule soit verte ?
Avec des notations évidentes, $\{R_1 \cap R_2, R_1 \cap V_2, V_1 \cap R_2, V_1 \cap V_2\}$ est un système complet d'événements.
On a alors $\mathbf{P}(R_1 \cap R_2) = \mathbf{P}(R_1)\mathbf{P}_{R_1}(R_2) = \dfrac{2}{n}\dfrac{1}{n-1}$.
Et de même,

$$\mathbf{P}(R_1 \cap V_2) = \frac{2}{n}\frac{n-2}{n-1},\ \mathbf{P}(V_1 \cap R_2) = \frac{n-2}{n}\frac{2}{n-1},\ \mathbf{P}(V_1 \cap V_2) = \frac{n-2}{n}\frac{n-3}{n-1}.$$

Donc par la formule des probabilités totales, avec le système complet d'événements susmentionné

$$\begin{aligned}
\mathbf{P}(V_3) &= \mathbf{P}(R_1 \cap R_2)\mathbf{P}_{R_1 \cap R_2}(V_3) + \mathbf{P}(R_1 \cap V_2)\mathbf{P}_{R_1 \cap V_2}(V_3) + \mathbf{P}(V_1 \cap R_2)\mathbf{P}_{V_1 \cap R_2}(V_3) + \mathbf{P}(V_1 \cap V_2)\mathbf{P}_{V_1 \cap V_2}(V_3) \\
&= \frac{2}{n(n-1)} \times 1 + \frac{2(n-2)}{n(n-1)}\frac{n-3}{n-2} + \frac{2(n-2)}{n(n-1)}\frac{n-3}{n-2} + \frac{(n-2)(n-3)}{n(n-1)}\frac{n-4}{n-2} \\
&= \frac{1}{n(n-1)(n-2)} \times (2 + 4(n-2)(n-3) + (n-2)(n-3)(n-4)) \\
&= \frac{n^3 - 5n^2 + 6n + 2}{n(n-1)(n-2)}.
\end{aligned}$$



La formule des probabilités totales et la formule des probabilités composées sont les deux fondements des arbres de probabilités que vous avez pu utiliser au lycée : les probabilités indiquées sur les branches sont des probabilités conditionnelles, la formule des probabilités composées justifier qu'on fasse le produit des probabilités situées sur des branches qui se suivent, et la formule des probabilités totales justifie qu'on finisse par sommer toutes les probabilités aboutissant à la réalisation d'un événement donné (ici V_3).

► On dispose de $n+1$ urnes $U_0, U_1, \dots, U_n$, qui contiennent chacune n boules : l'urne U_k contient k boules bleues et $n-k$ boules rouges.

On choisit une urne au hasard (et de manière équiprobable), et on tire simultanément deux boules dans cette urne. Quelle est la probabilité que les deux boules soient bleues ?

Notons[14] $U_0, \dots, U_n$ les événements «le tirage a lieu dans l'urne U_i», de sorte que $\{U_0, \dots, U_n\}$ est un système complet d'événements.

Notons également A l'événement «les deux boules sont bleues».

Il est évident que la probabilité que A se réalise dépend de l'urne dans laquelle le tirage a lieu.

Par la formule des probabilités totales appliquée avec le système complet d'événements $\{U_0, \dots, U_n\}$, il vient

$$\mathbf{P}(A) = \sum_{k=0}^{n} \mathbf{P}(U_k)\mathbf{P}_{U_k}(A) = \sum_{k=0}^{n} \frac{1}{n+1}\mathbf{P}_{U_k}(A).$$

Or, il est facile de constater que $\mathbf{P}_{U_0}(A) = \mathbf{P}_{U_1}(A) = 0$.

Et pour $k \geqslant 2$, alors l'urne U_k contient n boules, donc il y a $\binom{n}{2}$ manières d'y tirer 2 boules, et puisqu'il y a k boules bleues, il y a $\binom{k}{2}$ manières d'y tirer 2 boules bleues.

Autrement dit $\mathbf{P}_{U_k}(A) = \dfrac{\binom{k}{2}}{\binom{n}{2}} = \dfrac{k(k-1)}{n(n-1)}$.

Et donc

$$\begin{aligned}\mathbf{P}(A) &= \frac{1}{n+1}\sum_{k=2}^{n}\frac{k(k-1)}{n(n-1)} = \frac{1}{(n+1)n(n-1)}\sum_{k=1}^{n} k(k-1)\\ &= \frac{1}{(n+1)n(n-1)}\left(\sum_{k=1}^{n} k^2 - \sum_{k=1}^{n} k\right)\\ &= \frac{1}{(n+1)n(n-1)}\left(\frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}\right)\\ &= \frac{1}{n-1}\frac{2n-2}{6} = \frac{1}{3}.\end{aligned}$$

Grow up !
Si un arbre est un excellent support au brouillon, il ne peut plus en aucun cas être considéré comme une preuve, et vous devez justifier tous vos calculs !

[14] Abusivement.

⚠ Attention !
Pour autant $\mathbf{P}(A)$ n'a qu'une et qu'une seule valeur.
Je ne veux surtout pas voir de «si le tirage a lieu dans l'urne i, alors $\mathbf{P}(A) = \dots$.
Ce qu'on donnerait ainsi, ce sont des probabilités conditionnelles $\mathbf{P}_{U_i}(A)$.
La formule des probabilités totales est justement le moyen de rassembler toutes ces probabilités pour déterminer **la** valeur de $\mathbf{P}(A)$.

Le terme $k=1$ est nul.

28.2.4 Les formules de Bayes

Proposition 28.20 (Formule de Bayes) : *Soient A et B deux événements de probabilités non nulles. Alors*

$$\mathbf{P}_B(A) = \frac{\mathbf{P}_A(B)\mathbf{P}(A)}{\mathbf{P}(B)}.$$

Démonstration.

$$\mathbf{P}_B(A) = \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}_A(B)\mathbf{P}(A)}{\mathbf{P}(B)}.$$

□

Exemple 28.21

Les situations où il y a besoin de la formule de Bayes sont assez facilement reconnaissables : ce sont celles où on cherche une probabilité conditionnelle $\mathbf{P}_B(A)$, mais où on connaît la probabilité conditionnelle «dans l'autre sens» $\mathbf{P}_A(B)$.

Reprenons par exemple les urnes $U_0, \dots, U_n$ de l'exemple 28.19.

On tire deux boules, et les deux sont bleues. Quelle est la probabilité que le tirage ait eu lieu dans l'urne U_k ?
Avec les mêmes notations que précédemment, il s'agit donc de calculer $\mathbf{P}_A(U_k)$. Par la formule de Bayes,

$$\mathbf{P}_A(U_k) = \frac{\mathbf{P}(U_k)\mathbf{P}_{U_k}(A)}{\mathbf{P}(A)} = \frac{\frac{1}{n+1}\frac{k(k-1)}{n(n-1)}}{\frac{1}{3}} = \frac{3k(k-1)}{(n+1)n(n-1)}.$$

Un bon moyen de vérifier le résultat est souvent de regarder ce qui se passe pour les valeurs extrêmes. Ici, on sait que si le tirage a donné deux boules bleues, il ne peut avoir eu lieu ni dans l'urne U_0, ni dans l'urne U_1.
Or notre formule donne une probabilité nulle pour $k \in \{0, 1\}$, ce qui est cohérent. De plus, le fait que la probabilité soit une fonction croissante de k est aussi conforme à l'intuition.

Une autre forme parfois utile de la formule de Bayes est la suivante :

Remarque
Je la mentionne parce qu'elle est au programme, mais comme il s'agit simplement de coupler la formule de Bayes ci-dessus à la formule des probabilités totales, normalement vous saurez toujours la retrouver dans les cas où vous en aurez besoin, sans qu'il soit nécessaire de connaître un énoncé général.

Corollaire 28.22 – *Soit $\{A_1, \dots, A_n\}$ un système complet d'événements de probabilités non nulles, et soit B un événement de probabilité non nulle. Alors*

$$\forall j \in [\![1, n]\!],\ \mathbf{P}_B(A_j) = \frac{\mathbf{P}_{A_j}(B)\mathbf{P}(A_j)}{\sum_{i=1}^{n} \mathbf{P}(A_i)\mathbf{P}_{A_i}(B)}.$$

Démonstration. Il suffit juste de partir de la formule de Bayes, avec $A = A_j$, et d'appliquer la formule des probabilités totales au dénominateur afin de calculer $\mathbf{P}(B)$. □

28.3 Indépendance

Définition 28.23 – Deux événements A et B d'un même espace probabilisé fini $(\Omega, \mathbf{P})$ sont dits **indépendants** si $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

Notons bien que ceci fait appel à la probabilité $\mathbf{P}$, et ne dépend donc pas que de l'univers Ω, mais bien de la probabilité $\mathbf{P}$ que l'on met dessus.
La plupart du temps, l'indépendance est supposée par l'énoncé, ou alors est implicite, notamment lors de la répétition d'expériences telles que l'issue d'une expérience n'influe pas sur les suivantes. Typiquement lors de plusieurs lancers d'une pièce/d'un dé/whatever ou de tirages successifs **avec remise**.

Rédaction ✎
Dans ces exemples, un énoncé ne mentionnera jamais l'indépendance des répétitions, ce sera à vous de la comprendre et de la mentionner si vous l'utilisez.

⚠ En dehors des exemples ci-dessus, ne vous fiez pas trop à l'intuition pour l'indépendance, c'est une notion parfois imprévisible qui peut contredire l'intuition.

Exemple 28.24

On lance deux fois une pièce de monnaie qui tombe sur «face» avec probabilité $p \in [0, 1]$.
Notons P_i (resp. F_i) l'événement «le $i^{\text{ème}}$ lancer donne pile», et A l'événement «les deux lancers donnent le même résultat».
Alors $\mathbf{P}(F_1) = p$ et

$$\mathbf{P}(A) = \mathbf{P}\left((F_1 \cap F_2) \cup (P_1 \cap P_2)\right) = \mathbf{P}(F_1 \cap F_2) + \mathbf{P}(P_1 \cap P_2) = \mathbf{P}(F_1)\mathbf{P}(F_2) + \mathbf{P}(P_1)\mathbf{P}(P_2) = p^2 + (1-p)^2.$$

Enfin, on a $\mathbf{P}(A \cap F_1) = \mathbf{P}(F_1 \cap F_2) = p^2$.
Donc pour $p \notin \{\frac{1}{2}, 1\}$, les événements A et F_1 ne sont pas indépendants, alors qu'ils le sont pour $p = \frac{1}{2}$.
Pourtant, dans les deux cas, l'univers considéré est le même : c'est $\{P, F\}^2$.

Notons que deux événements A et B incompatibles et de probabilités non nulles ne peuvent pas être indépendants, puisqu'alors $\mathbf{P}(A \cap B) = 0 \neq \mathbf{P}(A)\mathbf{P}(B)$.
C'est souvent là un bon moyen de prouver la non-indépendance de deux événements.

Proposition 28.25 : *Deux événements A et B avec $\mathbf{P}(A) \neq 0$ sont indépendants si et seulement si $\mathbf{P}_A(B) = \mathbf{P}(B)$.*

Démonstration. C'est évident puisque $\mathbf{P}_A(B) = \dfrac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}$ et donc

$$\mathbf{P}_A(B) = \mathbf{P}(B) \Leftrightarrow \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}(A \cap B).$$

□

L'intuition là-derrière est que si A et B sont indépendants, alors savoir que A réalisé ne change rien à la probabilité que B le soit.

Le cas où $\mathbf{P}(A) = 0$ est peu intéressant, car un tel événement[15] est indépendant de tout autre événement B, puisque $A \cap B \subset A$ et donc $\mathbf{P}(A \cap B) \leqslant \mathbf{P}(A) = 0$, de sorte que $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

[15] De probabilité nulle signifiant qu'il ne se produit jamais ou presque...

Proposition 28.26 : *Si A et B sont indépendants, alors A et $\overline{B}$ sont indépendants.*

Corollaire
Et donc $\overline{A}$ et $\overline{B}$ sont également indépendants.

Démonstration. Puisque $\{B, \overline{B}\}$ est un système complet d'événements, par la formule des probabilités totales, $\mathbf{P}(A) = \mathbf{P}(A \cap B) + \mathbf{P}(A \cap \overline{B})$.
Et donc

$$\mathbf{P}(A \cap \overline{B}) = \mathbf{P}(A) - \mathbf{P}(A \cap B) = \mathbf{P}(A) - \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}(A)(1 - \mathbf{P}(B)) = \mathbf{P}(A)\mathbf{P}(\overline{B})$$

de sorte que A et $\overline{B}$ sont bien indépendants. □

Définition 28.27 – Des événements $A_1, \dots, A_n$ d'un espace probabilisé $(\Omega, \mathbf{P})$ sont dits **mutuellement indépendants** si $\forall I \subset [\![1, n]\!]$, $\mathbf{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbf{P}(A_i)$.

⚠ Pour $n \geqslant 3$, l'indépendance mutuelle de n événements implique bien qu'ils sont deux à deux indépendants (il suffit de prendre I de cardinal 2 dans la définition), mais la réciproque est fausse.
Par exemple, lançons deux dés[16], et notons A «le numéro du premier dé est pair», B «le numéro du second dé est pair» et C «la somme des deux dés est paire».
Alors on a $\mathbf{P}(A) = \mathbf{P}(B) = \mathbf{P}(C) = \dfrac{1}{2}$.
Par ailleurs, il est clair que A et B sont indépendants, et

$$\mathbf{P}(A \cap C) = \mathbf{P}(A \cap B) = \frac{1}{4}.$$

[16] Équilibrés, à 6 faces.

Donc A, B, C sont deux à deux indépendants.
En revanche, $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A \cap B) = \dfrac{1}{4}$, alors que $\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = \dfrac{1}{8}$.
Donc A, B et C ne sont pas mutuellement indépendants.

Détails
Si A et B sont réalisés, alors C l'est.
Autrement dit, $A \cap B \subset C$.