

Probabilités discrètes

Dans ce chapitre, nous franchissons un cap majeur : nous passons du monde confortable du fini au monde vertigineux de l'infini. En première année, on se plaçait souvent en situation d'équiprobabilité ; calculer une probabilité consistait souvent à dénombrer les cas favorables et les cas possibles. Désormais, notre univers d'étude Ω pourra contenir une infinité de résultats (comme le nombre de lancers nécessaires pour obtenir un pile). Pour dompter cet infini, nous allons devoir marier l'analyse et l'algèbre ; l'étude des probabilités discrètes sera intimement liée au chapitre précédent sur les séries numériques.

SOMMAIRE

1	Espaces probabilisables discrets.	2
2	Espaces probabilisés	5
2.1	Probabilité sur un espace probabilisable	5
2.2	Propriétés des probabilités.	6
3	Probabilités conditionnelles	7
3.1	Définition	7
3.2	Propriétés.	8
4	Grands théorèmes de probabilités.	9
4.1	Formule des probabilités totales	9
4.2	Formule des probabilités composées	11
4.3	Formule de Bayes	12
5	Indépendance d'événements.	13
5.1	Indépendance de deux événements	13
5.2	Indépendance d'une famille d'événements	14
6	Méthodes	15
6.1	Calculer la probabilité d'un événement	15
6.2	Dénombrer en situation d'équiprobabilité	16

OBJECTIFS

- Calculer une probabilité.
- Identifier quand utiliser les grands théorèmes de probabilités.
- Identifier la bonne situation de dénombrement.

1 Espaces probabilisables discrets

Définition 1.1

On dit qu'un ensemble est **dénombrable** s'il est en bijection avec l'ensemble $\mathbb{N}$.

En termes plus simples, un ensemble dénombrable est un ensemble dont on peut numéroter tous les éléments sans en oublier un seul (surjection) et sans numéroter deux fois le même élément (injection).

Exemple 1.2

Les ensembles $\mathbb{N}$, $\mathbb{N}^*$ sont dénombrables.

Plus généralement, $\mathbb{Z}$ et tous les sous-ensembles de la forme $\llbracket n_0, +\infty \llbracket$ ($n_0 \in \mathbb{Z}$) sont dénombrables.

Les ensembles $\mathbb{N}^2$ et $\mathbb{Q}$ sont dénombrables, mais $\mathbb{R}$ n'est pas dénombrable.

La non-dénombrabilité de $\mathbb{R}$ a été démontrée par le mathématicien allemand Georg Cantor en 1874 puis en 1891 où il utilisa le célèbre « argument de la diagonale » qui porte son nom. Ce résultat a eu de nombreuses répercussions sur la naissante théorie des ensembles... Elle montre notamment que tous les ensembles infinis n'ont pas le même nombre d'éléments et que les ensembles dénombrables sont les plus « petits » ensembles infinis (encore faut-il définir rigoureusement comment comparer les cardinaux d'ensembles infinis) !

Remarque 1.3

Les indices de sommation d'une série forment un ensemble dénombrable.

Question à se poser à la maison : peut-on donner un sens à $\sum_{x \in \mathbb{R}} f(x)$?

Définition 1.4

On appelle **univers des possibles** l'ensemble Ω des résultats possibles décrivant une expérience aléatoire.

On ne considérera dans ce chapitre que des univers dits **discrets**, i.e. finis ou dénombrables.

Exemple 1.5

- On lance un dé. On peut alors choisir $\Omega = \llbracket 1, 6 \llbracket$.
- On lance une pièce. En notant P l'événement « on obtient pile », on choisit $\Omega = \{P, \overline{P}\}$.
- On lance une pièce n fois. On peut choisir $\Omega = \{P, F\}^n$ en notant $F = \overline{P}$.
- On lance une pièce indéfiniment. On choisit alors $\Omega = \{P, F\}^{\mathbb{N}^*}$.
- Un urne contient une boule blanche, notée B , et quatre boules rouges, notées R .
 - ★ On tire successivement deux boules avec remise : $\Omega = \{(B, B), (B, R), (R, B), (R, R)\}$.
 - ★ On tire successivement deux boules sans remise : $\Omega = \{(B, R), (R, B), (R, R)\}$.
 - ★ On tire simultanément deux boules : $\Omega = \{\{B, R\}, \{R, R\}\}$.

Une fois l'univers fixé, on veut pouvoir définir des sous-ensembles réalisables de l'univers ; les *événements*. On attend de l'univers qu'il constitue un événement. On attend aussi de tous ces événements que leur complémentaire, réunion et intersection soient encore des événements. Cela nous amène à définir la notion de tribu.

 Définition 1.6

On appelle **tribu** (ou **σ -algèbre**) sur un univers Ω tout sous-ensemble $\mathcal{T}$ de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω vérifiant les propriétés suivantes :

1. la tribu $\mathcal{T}$ contient l'univers : $\Omega \in \mathcal{T}$;
2. la tribu $\mathcal{T}$ est stable par complémentaire : $\forall A \in \mathcal{T}, \bar{A} \in \mathcal{T}$;
3. la tribu $\mathcal{T}$ est stable par réunion d'une suite d'événements de la tribu :

$$\forall (B_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}, \bigcup_{n \in \mathbb{N}} B_n \in \mathcal{T}.$$

On définit alors les éléments de la tribu $\mathcal{T}$:

- les éléments de $\mathcal{T}$ sont appelés **événements** de l'expérience aléatoire,
- les singletons de $\mathcal{T}$ sont appelés **événements élémentaires**,
- l'événement Ω est appelé **l'événement certain**, l'événement $\emptyset$ est appelé **l'événement impossible**,

On dit alors que le couple $(\Omega, \mathcal{T})$ est un **espace probabilisable**.

Même si la notion de tribu constitue le socle de la théorie des probabilités, nous ne l'utiliserons pas cette année.

 Exemple 1.7

Soit Ω un univers et soit $A \subset \Omega$.

Les ensembles $\mathcal{T}_1 = \{\Omega, \emptyset\}$, $\mathcal{T}_2 = \{\Omega, A, \bar{A}, \emptyset\}$ et $\mathcal{P}(\Omega)$ sont des tribus sur Ω .

 Propriété 1.8

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Alors :

1. $\emptyset \in \mathcal{T}$,
2. Pour tout $(A, B) \in \mathcal{T}^2$, $A \cap B$, $A \cup B$ et $A \setminus B$ sont des événements, i.e. $A \cap B \in \mathcal{T}$, $A \cup B \in \mathcal{T}$ et $A \setminus B \in \mathcal{T}$.
3. la tribu $\mathcal{T}$ est stable par intersection dénombrable.

Démonstration admise.

 Définition 1.9 – Vocabulaire des opérations sur les événements

Soient A et B deux événements d'un univers des possibles Ω .

- L'événement $\bar{A}$ est appelé **l'événement contraire** de A .
- L'événement $A \cap B$ est l'événement " **A et B** ".
- L'événement $A \cup B$ est l'événement " **A ou B** ".
- L'événement $A \setminus B$ est l'événement " **A privé de B** ", ou encore " **A mais pas B** ".
- On dit que A **implique** B si $A \subset B$,
- On dit que A et B sont **disjoints** si $A \cap B = \emptyset$.

Propriété 1.10

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements d'un espace probabilisable $(\Omega, \mathcal{T})$.

1. L'événement $\bigcap_{n=0}^{+\infty} A_n$ est réalisé si, et seulement si, tous les A_n sont réalisés.
2. L'événement $\bigcup_{n=0}^{+\infty} A_n$ est réalisé si, et seulement si, au moins un A_n est réalisé.
3. L'événement $\bigcup_{n=0}^{+\infty} \bigcap_{p=n}^{+\infty} A_p$ est réalisé si, et seulement si, tous les A_n sont réalisés à partir d'un certain rang.
4. L'événement $\bigcap_{n=0}^{+\infty} \bigcup_{p=n}^{+\infty} A_p$ est réalisé si, et seulement si, une infinité de A_n est réalisée.

Exemple 1.11 – ♥

On lance une infinité de fois un dé et on note, pour tout $k \in \mathbb{N}^*$, S_k l'événement « on obtient 6 au k -ème lancer ». Traduire les événements suivants à l'aide de la symbolique mathématique de la théorie des ensembles :

1. « On obtient jamais de 6 ».
2. « On obtient au moins une fois 6 lors de la partie ».
3. « On obtient au moins une fois 6 à partir du cinquième lancer ».
4. « On n'obtient que des 6 à partir du cinquième lancer ».
5. « On n'obtient que des 6 à partir d'un certain rang ».
6. « On obtient une infinité de 6 ».

Démonstration.

1. On n'obtient pas 6 au premier lancer **et** on n'obtient pas 6 au deuxième lancer **et** on n'obtient pas 6 au troisième lancer, etc : $\bigcap_{n=1}^{+\infty} \overline{S_n}$.
2. Il y a deux façons de voir les choses :
 - On n'obtient pas 6 au premier lancer **ou** on n'obtient pas 6 au deuxième lancer **ou** on n'obtient pas 6 au troisième lancer, etc : $\bigcup_{n=1}^{+\infty} \overline{S_n}$.
 - On est dans la situation complémentaire de la situation 1 : $\overline{\bigcap_{n=1}^{+\infty} \overline{S_n}} = \bigcup_{n=1}^{+\infty} S_n$.
3. On s'inspire de ce qui précède : $\bigcup_{n=5}^{+\infty} S_n$.
4. On s'inspire de ce qui précède : $\bigcap_{n=5}^{+\infty} S_n$.
5. Il existe un entier n tel que, pour tout $k \geq n$, S_k est réalisé : $\bigcup_{n=1}^{+\infty} \bigcap_{k=n}^{+\infty} S_k$.
6. Pour tout entier n , il existe $k \geq n$ tel que S_k est réalisé : $\bigcap_{n=1}^{+\infty} \bigcup_{k=n}^{+\infty} S_k$.

□

2 Espaces probabilisés

2.1 Probabilité sur un espace probabilisable

Définition 2.1

Soit $(\Omega, \mathcal{T})$ un espace probabilisable fini ou dénombrable.

On appelle **loi de probabilité** sur $(\Omega, \mathcal{T})$ toute application $\mathbb{P} : \mathcal{T} \rightarrow \mathbb{R}_+$ vérifiant :

1. $\mathbb{P}(\Omega) = 1$
2. Pour toute suite finie ou dénombrable $(A_i)_{i \in I}$ d'événements deux à deux disjoints,

$$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mathbb{P}(A_i).$$

Cette propriété est appelée **l'axiome de σ -additivité**.

Pour tout événement A de $\mathcal{T}$, le réel $\mathbb{P}(A)$ est appelé la **probabilité** de l'événement A .

On dit alors que le triplet $(\Omega, \mathcal{T}, \mathbb{P})$ est un **espace probabilisé**.

- Si A et B sont deux événements disjoints, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- Si $(A_i)_{0 \leq i \leq n}$ est une suite *finie* d'événements deux à deux disjoints, on a :

$$\mathbb{P}\left(\bigcup_{i=0}^n A_i\right) = \sum_{i=0}^n \mathbb{P}(A_i).$$

- Si $(A_i)_{i \in \mathbb{N}}$ est une suite *infinie* (dénombrable !) d'événements deux à deux disjoints, on a :

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

Encore faut-il prouver la convergence de cette série, ce qui sera fait à la propriété 2.2 ci-dessous.

Propriété 2.2 – Justification de la convergence (σ -additivité)

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements deux à deux disjoints d'un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.

Alors la série $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ est convergente.

Exemple 2.3 – Univers fini

Si Ω est un univers fini et non vide, l'application définie par :

$$\begin{aligned} \mathbb{P} : \mathcal{P}(\Omega) &\rightarrow \mathbb{R}_+ \\ A &\mapsto \frac{\text{card}(A)}{\text{card}(\Omega)} \end{aligned}$$

est une loi de probabilité, appelée **probabilité uniforme** sur $(\Omega, \mathcal{P}(\Omega))$.

On dit encore qu'on est **en situation d'équiprobabilité** car tous les événements *élémentaires* sont équiprobables.

Remarque 2.4

Le modèle probabiliste (i.e. l'espace probabilisé) est généralement imposé par l'énoncé. Par exemple, la loi de probabilité associée à un lancer de dé non pipé est la loi de probabilité uniforme, ce qui n'est pas le cas lorsque le dé est pipé.

2.2 Propriétés des probabilités

Dans toute la suite, on se place - sauf mention contraire - dans un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.

Propriété 2.5

Soient A et B deux événements.

1. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$ (en particulier $\mathbb{P}(\emptyset) = 0$).
2. $\mathbb{P}(A) \in [0, 1]$.
3. si $A \subset B$, $\mathbb{P}(A) \leq \mathbb{P}(B)$.
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Démonstration admise.

Définition 2.6

Soit A un événement.

1. Si $\mathbb{P}(A) = 1$, on dit alors que A est **certain presque-sûrement**, ou **presque-sûr**.
2. Si $\mathbb{P}(A) = 0$, on dit alors que A est **négligeable**.

Remarque 2.7 – Paradoxe des événements négligeables dans un univers infini

Dès lors que l'univers est infini, un événement de probabilité nulle n'est pas nécessairement impossible (i.e. irréalisable), tout comme un événement de probabilité 1 n'est pas nécessairement l'univers (i.e. réalisé).

Exemple 2.8 – Événements possibles de probabilité nulle - événements non certains de probabilité 1

On répète un tirage de pile ou face d'une pièce équilibrée jusqu'à obtenir *pour la première fois* pile. Bien que l'événement « on obtient face à chaque tirage » est réalisable, montrer de deux façons différentes que la probabilité que le jeu ne s'arrête jamais est égale à 0 (on dit aussi que le jeu le jeu s'arrête presque-sûrement).

Démonstration. Notons, pour tout $k \in \mathbb{N}^*$, F_k l'événement « on obtient face au k -ème lancer ». L'événement $J = \bigcup_{k=1}^{+\infty} F_k$ est l'événement « le jeu ne s'arrête jamais ». Calculons sa probabilité de deux manières : par calcul direct et majoration ou par passage au complémentaire et application de la σ -additivité.

- Soit $n \in \mathbb{N}^*$. Remarquons que :

$$J = \bigcup_{k=1}^{+\infty} F_k \subset \bigcup_{k=1}^n F_k.$$

Par croissance de $\mathbb{P}$ et indépendance des lancers, on trouve que

$$0 \leq \mathbb{P}(J) \leq \mathbb{P}\left(\bigcup_{k=1}^n F_k\right) = \frac{1}{2}^n$$

Puisque cet encadrement est vrai pour tout $n \in \mathbb{N}^*$, on peut passer à la limite lorsque n tend vers 0 : on obtient que $\mathbb{P}(J) = 0$. L'événement J est donc négligeable.

- L'événement $\overline{J}$ « le jeu s'arrête » (à un moment donné) s'écrit comme la réunion des événements A_n : « le jeu s'arrête au n -ème lancer », qui eux-mêmes peuvent être décrits à l'aide des événements F_k :

$$\overline{J} = \bigcup_{n=1}^{+\infty} A_n = \bigcup_{n=1}^{+\infty} F_1 \cap \dots \cap F_{n-1} \cap \overline{F_n}.$$

Puisque la réunion est disjointe, on trouve que :

$$\begin{aligned}
 \mathbb{P}(\bar{J}) &= \sum_{n=1}^{+\infty} \mathbb{P}(A_n) \quad (\text{la série converge par } \sigma\text{-additivité}) \\
 &= \sum_{n=1}^{+\infty} \mathbb{P}(F_1 \cap \dots \cap F_{n-1} \cap \bar{F}_n) \\
 &= \sum_{n=1}^{+\infty} \frac{1}{2^n} \quad (\text{par indépendance des lancers}) \\
 &= 1.
 \end{aligned}$$

L'événement $\bar{J}$ est donc réalisé presque-sûrement, i.e. J est négligeable.

□

Exemple 2.9

Une urne contient une boule rouge. Un joueur lance un dé équilibré : s'il obtient un six, il tire une boule dans l'urne. Sinon, il rajoute une boule blanche dans l'urne et répète la manipulation. Pour tout $k \in \mathbb{N}^*$, on note G_k l'événement « on obtient pour la première fois un six au k -ème lancer. » Calculer la probabilité de l'événement G_k pour tout $k \in \mathbb{N}^*$.

★ Théorème 2.10

Soit $\Omega = \{\omega_n\}_{n \in \mathbb{N}}$ un univers dénombrable et soit $(p_n)_{n \in \mathbb{N}}$ une suite à valeurs dans $\mathbb{R}_+$. Si la série $\sum_{n \in \mathbb{N}} p_n$ converge et a pour somme 1, il existe une seule et unique loi de probabilité $\mathbb{P}$ vérifiant :

$$\forall n \in \mathbb{N}, \mathbb{P}(\{\omega_n\}) = p_n.$$

Démonstration admise.

Au delà de l'importance théorique de ce théorème, ce résultat permet de générer automatiquement des lois : n'importe quelle série convergente à termes positifs va fournir une loi de probabilité (après avoir divisé chaque terme par la somme de la série).

Remarque 2.11 – Impossibilité de loi uniforme sur $\mathbb{N}$

Le résultat précédent implique qu'il est impossible de choisir **de manière équiprobable** un nombre parmi tous les entiers naturels.

3 Probabilités conditionnelles

3.1 Définition

Définition 3.1

Soit A un événement de probabilité non nulle et soit B un événement quelconque. On appelle **probabilité conditionnelle de B sachant A** le réel $\mathbb{P}_A(B)$, encore noté $\mathbb{P}(B|A)$, défini par :

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

i Remarque 3.2

On pourra remarquer que $\mathbb{P} = \mathbb{P}_\Omega$. On peut ainsi voir le passage de $\mathbb{P}$ à $\mathbb{P}_A$ comme un changement d'univers, de Ω à A . La quantité $\mathbb{P}_A(B)$ représente alors la *mesure* de l'événement $A \cap B$ dans ce nouvel univers A .

i Remarque 3.3 – Attention !

La notation « $B|A$ » n'existe pas ; elle ne désigne pas un événement ! C'est la loi de probabilité qui est *conditionnée* par A , pas l'événement B .

i Remarque 3.4 – Convention d'écriture

Remarquons que $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}_A(B)$ si $\mathbb{P}(A) \neq 0$.
Supposons que $\mathbb{P}(A) = 0$. Puisque $A \cap B \subset A$, on en déduit que $\mathbb{P}(A \cap B) = 0$.
On convient alors de toujours pouvoir écrire $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}_A(B)$, même si $\mathbb{P}_A(B)$ n'existe pas.

3.2 Propriétés**★ Théorème 3.5**

Soit A un événement de probabilité non nulle d'un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. Alors l'application :

$$\begin{aligned} \mathbb{P}_A : \mathcal{T} &\rightarrow \mathbb{R}_+ \\ B &\mapsto \mathbb{P}_A(B) \end{aligned}$$

est une loi de probabilité sur $(\Omega, \mathcal{T})$, appelée **loi de probabilité sachant A** .

Démonstration admise (bon exercice à faire à la maison).

⚙ Propriété 3.6

Pour tous événements A, B et C , avec $\mathbb{P}(A) \neq 0$, on a :

1. $\mathbb{P}_A(\overline{B}) = 1 - \mathbb{P}_A(B)$
2. $\mathbb{P}_A(B \setminus C) = \mathbb{P}_A(B) - \mathbb{P}_A(B \cap C)$
3. $\mathbb{P}_A(B \cup C) = \mathbb{P}_A(B) + \mathbb{P}_A(C) - \mathbb{P}_A(B \cap C)$.

Démonstration. $\mathbb{P}_A$ est une loi de probabilité.

□

4 Grands théorèmes de probabilités

4.1 Formule des probabilités totales

Définition 4.1

On appelle **système complet d'événements** toute suite $(A_i)_{i \in I}$ - finie ou dénombrable - d'événements vérifiant :

1. Les événements sont deux à deux disjoints :

$$\forall (i, j) \in I^2, i \neq j \Rightarrow A_i \cap A_j = \emptyset.$$

2. Ils décrivent l'univers :

$$\bigcup_{i \in I} A_i = \Omega.$$

Définition 4.2

On appelle **système quasi-complet d'événements** toute suite $(A_i)_{i \in I}$ - finie ou dénombrable - d'événements deux à deux disjoints vérifiant :

$$\sum_{i \in I} \mathbb{P}(A_i) = 1.$$

Exemple 4.3 – Suite de l'exercice 2.9

Une urne contient une boule rouge. Un joueur lance un dé équilibré : s'il obtient un six, il tire une boule dans l'urne. Sinon, il rajoute une boule blanche dans l'urne et répète la manipulation. Pour tout $k \in \mathbb{N}^*$, on note S_k l'événement « on obtient pour la première fois un six au k -ème lancer. » Montrer que $(S_k)_{k \in \mathbb{N}^*}$ forme un système quasi-complet d'événements.

Remarque 4.4

Un système complet d'événements est quasi-complet. En effet, la proposition 2.2 assure que si $(A_i)_{i \in I}$ est un système complet d'événements dénombrables, alors la série $\sum_{i \in I} \mathbb{P}(A_i)$ converge.

De plus, que le système complet soit fini ou dénombrable, $\sum_{i \in I} \mathbb{P}(A_i) = 1$.

Théorème 4.5 – Formule des probabilités totales

Soit $(A_i)_{i \in I}$ un système complet ou quasi-complet d'événements et soit B un événement. Alors :

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(A_i \cap B) = \sum_{i \in I} \mathbb{P}(A_i) \mathbb{P}_{A_i}(B).$$

Méthode 4.6 – Utilisation de la formule des probabilités totales

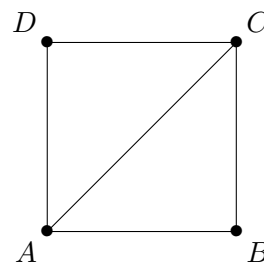
On applique la formule des probabilités totales lorsque le calcul direct de $\mathbb{P}(B)$ est trop complexe, mais que la situation s'éclaire dès que l'on raisonne *par disjonction de cas*. Si l'on sait évaluer la probabilité de B sous différents scénarios (A_i) qui décrivent entièrement l'univers, alors l'analyse de B « au cas par cas » permet de reconstituer sa probabilité globale.

L'exercice qui suit illustre très bien une famille de problèmes qui reviennent souvent en probabilités discrètes : les **chaines de Markov** (dont la théorie est hors-programme). On les rencontrera dès qu'une situation décrit un système qui évolue par étapes successives (instants n , générations n , lancers n , etc) et dont la dynamique est simple : *toute situation à une étape donnée ne dépend uniquement et directement que de ce qui s'est passé à l'étape précédente*. Dans ces cas-là, calculer directement la probabilité d'un événement à l'étape $n+1$ est impossible au premier abord parce qu'il y a trop de *trajectoires*. Le réflexe absolu consiste alors à remonter d'un cran dans le temps en faisant une disjonction de cas (via la formule des probabilités totales) sur toutes les positions possibles à l'étape n .

Il est fondamental de reconnaître un tel schéma, même si l'expérience change (tirages successifs dans des urnes, lancers de pièces, déplacement de particules, étude de population).

Exemple 4.7 – Chaîne de Markov

On considère une particule se déplaçant aléatoirement sur les arêtes de la figure ci-contre délimitée par les points $A(0,0)$, $B(1,0)$, $C(1,1)$ et $D(0,1)$. À l'instant $n=0$, elle se situe en A et à chaque étape, elle se déplace aléatoirement et de manière équiprobable vers l'un des sommets qu'elle peut atteindre en une étape. Par exemple, du point A , la particule peut aller en B , C ou D mais du point D , elle ne peut aller qu'en A ou C (et pas en B). Pour tout $n \in \mathbb{N}$, on note A_n (resp. B_n , C_n et D_n) l'événement « la particule se trouve en A (resp. en B , C ou D) à l'instant n » et a_n (resp. b_n , c_n et d_n) sa



probabilité ; on note enfin $X_n = \begin{pmatrix} a_n \\ b_n \\ c_n \\ d_n \end{pmatrix}$.

1. Déterminer une matrice $A \in \mathcal{M}_4(\mathbb{R})$ telle que, pour tout $n \in \mathbb{N}$, $X_{n+1} = AX_n$.
2. En déduire que, pour tout $n \in \mathbb{N}$, $X_n = A^n X_0$.

Démonstration.

1. Soit $n \in \mathbb{N}$. À l'instant n , la particule est nécessairement en A , B , C ou D . Ainsi les événements (A_n, B_n, C_n, D_n) forment un système complet d'événements. En appliquant la formule des probabilités totales avec ce système, on trouve que :

$$\begin{aligned} a_{n+1} &= \mathbb{P}(A_{n+1}) \\ &= \mathbb{P}(A_n)\mathbb{P}_{A_n}(A_{n+1}) + \mathbb{P}(B_n)\mathbb{P}_{B_n}(A_{n+1}) + \mathbb{P}(C_n)\mathbb{P}_{C_n}(A_{n+1}) + \mathbb{P}(D_n)\mathbb{P}_{D_n}(A_{n+1}) \\ &= \frac{1}{2}b_n + \frac{1}{3}c_n + \frac{1}{2}d_n. \end{aligned}$$

De la même manière, on trouve que :

$$\begin{aligned} b_{n+1} &= \frac{1}{3}a_n + \frac{1}{3}c_n \\ c_{n+1} &= \frac{1}{3}a_n + \frac{1}{2}b_n + \frac{1}{2}d_n \\ d_{n+1} &= \frac{1}{3}a_n + \frac{1}{3}c_n. \end{aligned}$$

Ainsi :

$$X_{n+1} = \begin{pmatrix} a_{n+1} \\ b_{n+1} \\ c_{n+1} \\ d_{n+1} \end{pmatrix} = \begin{pmatrix} 0a_n + \frac{1}{2}b_n + \frac{1}{3}c_n + \frac{1}{2}d_n \\ \frac{1}{3}a_n + 0b_n + \frac{1}{3}c_n + 0d_n \\ \frac{1}{3}a_n + \frac{1}{2}b_n + 0c_n + \frac{1}{2}d_n \\ \frac{1}{3}a_n + 0b_n + \frac{1}{3}c_n + 0d_n \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \\ c_n \\ d_n \end{pmatrix} = AX_n.$$

$$\text{en posant } A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

2. Montrons le résultat par récurrence sur $n \in \mathbb{N}$.

La propriété est trivialement initialisée : $A^0 X_0 = X_0$.

Soit $n \in \mathbb{N}$. Supposons que $X_n = A^n X_0$. Alors $X_{n+1} = AX_n = A \times A^n X_0 = A^{n+1} X_0$, ce qui conclut la récurrence. □

Exemple 4.8 – Suite des exercices 2.9 et 4.3

Une urne contient une boule rouge. Un joueur lance un dé équilibré : s'il obtient un six, il tire une boule dans l'urne. Sinon, il rajoute une boule blanche dans l'urne et répète la manipulation. Calculer la probabilité que la boule tirée soit rouge.

Exercice à la maison 4.9

On dispose de six urnes numérotées de 1 à 6. Pour tout $k \in \llbracket 1, 6 \rrbracket$, l'urne numéro k comporte k boules blanches et une boule rouge. Un joueur lance un dé équilibré puis choisit une boule dans l'urne correspondant au résultat du dé.

Montrer que la probabilité que la boule tirée soit blanche est égale à $\frac{1}{6} \sum_{k=1}^6 \frac{k}{k+1}$.

4.2 Formule des probabilités composées

Propriété 4.10 – Formule des probabilités composées

Soient $(A_1, \dots, A_n)$ une famille d'événements telle que $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$. Alors :

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \mathbb{P}_{A_1 \cap A_2}(A_3) \times \dots \times \mathbb{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n).$$

Démonstration admise.

Remarque 4.11

La formule des probabilités composées s'utilise par exemple lorsqu'on tire **successivement** et **sans remise** des boules dans une urne.

Exemple 4.12 –

Une urne contient n boules blanches et n boules rouges. On tire successivement et sans remise n boules dans cette urne. Montrer que la probabilité qu'une boule rouge figure dans ce tirage est égale à $1 - \frac{(n!)^2}{(2n)!}$.

Démonstration. Notons R l'événement dont on cherche la probabilité et, pour tout $k \in \llbracket 1, n \rrbracket$, B_k l'événement « on obtient une boule blanche au k -ème tirage ». Remarquons que $\bar{R} = B_1 \cap B_2 \cap \dots \cap B_{n-1} \cap B_n$. En appliquant

la formule des probabilités composées, on trouve que :

$$\begin{aligned}
 \mathbb{P}(\overline{R}) &= \mathbb{P}(B_1 \cap B_2 \cap \dots \cap B_{n-1} \cap B_n) \\
 &= \mathbb{P}(B_1) \mathbb{P}_{B_1}(B_2) \dots \mathbb{P}_{B_1 \cap \dots \cap B_{n-2}}(B_{n-1}) \mathbb{P}_{B_1 \cap \dots \cap B_{n-1}}(B_n) \\
 &= \frac{n}{2n} \times \frac{n-1}{2n-1} \times \dots \times \frac{n-(n-2)}{2n-(n-2)} \times \frac{n-(n-1)}{2n-(n-1)} \\
 &= \frac{n}{2n} \times \frac{n-1}{2n-1} \times \dots \times \frac{2}{n+2} \times \frac{1}{n+1} \\
 &= \frac{n!}{(2n)!} = \frac{(n!)^2}{(2n)!}.
 \end{aligned}$$

On en déduit le résultat attendu : $\mathbb{P}(R) = 1 - \frac{(n!)^2}{(2n)!}$. □

4.3 Formule de Bayes

★ Théorème 4.13 – Formule de Bayes

Soit $(A_i)_{i \in I}$ un système complet ou quasi-complet d'événements et soit B un événement de probabilité non nulle. Alors :

$$\forall i \in I, \mathbb{P}_B(A_i) = \frac{\mathbb{P}(A_i) \mathbb{P}_{A_i}(B)}{\sum_{j \in I} \mathbb{P}(A_j) \mathbb{P}_{A_j}(B)}.$$

i Remarque 4.14

La formule de Bayes ne doit absolument pas être apprise, il faut savoir la retrouver via sa démonstration. Cette formule est utilisée pour *inverser* le conditionnement entre *causes* et *conséquences*.

✎ Exemple 4.15 – ♥

Un élève vient au lycée en RER dans 10 % des cas, en voiture dans 40 % des cas, et à vélo dans 50 % des cas. On sait que la probabilité d'arriver en retard est de 0,6 en RER, 0,2 en voiture, et 0,1 à vélo. L'élève est en retard aujourd'hui. Quelle est la probabilité qu'il soit venu en voiture ?

Démonstration. On définit les événements suivants :

T : « l'élève vient en RER (train) » ;

V : « l'élève vient en voiture » ;

B : « l'élève vient en vélo (bicyclette) » ;

R : « l'élève est en retard ».

On cherche $\mathbb{P}_V(R)$ et on sait que $\mathbb{P}(T) = 0,1$; $\mathbb{P}(V) = 0,4$; $\mathbb{P}(B) = 0,5$; $\mathbb{P}_T(R) = 0,6$; $\mathbb{P}_T(V) = 0,2$; $\mathbb{P}_T(B) = 0,1$. Appliquons la formule de Bayes (et la formule des probabilités totales avec (T, V, B) comme système complet) :

$$\begin{aligned}
 \mathbb{P}_R(V) &= \frac{\mathbb{P}(V \cap R)}{\mathbb{P}(R)} \\
 &= \frac{\mathbb{P}(V) \mathbb{P}_V(R)}{\mathbb{P}(T) \mathbb{P}_T(R) + \mathbb{P}(V) \mathbb{P}_V(R) + \mathbb{P}(B) \mathbb{P}_B(R)} \\
 &= \frac{0,08}{0,06 + 0,08 + 0,05} = \frac{8}{19}.
 \end{aligned}$$

□

5 Indépendance d'événements

5.1 Indépendance de deux événements

Définition 5.1

On dit que deux événements A et B sont **indépendants** (pour la probabilité $\mathbb{P}$) si :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Remarque 5.2 – D'où vient cette définition ?

Intuitivement, c'est exactement par la relation $\mathbb{P}_A(B) = \mathbb{P}(B)$ que l'on a envie de définir l'indépendance. Elle traduit parfaitement l'idée concrète : « *Savoir que l'événement A s'est réalisé ne modifie en rien la probabilité que B se réalise* ».

Cependant, cette égalité pose deux problèmes majeurs pour être choisie comme définition :

- Elle n'est pas symétrique : elle met en avant le rôle de A par rapport à B , alors que l'indépendance est une notion symétrique (si A n'influence pas B , alors B n'influence pas A).
- Elle est restrictive : elle impose d'avoir $\mathbb{P}(A) \neq 0$ pour pouvoir simplement considérer la probabilité conditionnelle $\mathbb{P}_A(B)$.

Pour s'affranchir de ces problèmes, on utilise la définition d'une probabilité conditionnelle :

$\mathbb{P}_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$. En remplaçant $\mathbb{P}_A(B)$ par $\mathbb{P}(B)$ dans cette égalité et en « passant le dénominateur de l'autre côté », on obtient la relation $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. C'est cette formule qui est retenue comme définition générale. Elle a l'immense avantage d'être symétrique et toujours définie.

Remarque 5.3

Attention à ne pas confondre l'incompatibilité (on préférera parler de « disjonction ») et l'indépendance de deux événements ! On pourra par exemple vérifier que deux événements disjoints et de probabilités non nulles ne sont jamais indépendants...

Sans contexte, l'indépendance de deux événements n'est en général pas démontrable : elle résulte souvent d'un choix de modélisation.

Propriété 5.4

Soient A et B deux événements indépendants (pour une probabilité $\mathbb{P}$). Alors les événements A et $\overline{B}$ sont indépendants (pour la probabilité $\mathbb{P}$), de même que les événements $\overline{A}$ et B , et les événements $\overline{A}$ et $\overline{B}$.

Démonstration admise.

5.2 Indépendance d'une famille d'événements

Définition 5.5

On dit que les événements d'une famille $(A_i)_{i \in I}$ sont :

- **deux à deux indépendants** si, pour tout i et j dans I tels que $i \neq j$, A_i et A_j sont indépendants, i.e. :

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j).$$

- **mutuellement indépendants** si, pour tout sous-ensemble fini $J \subset I$,

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

Remarque 5.6

L'indépendance mutuelle implique l'indépendance deux à deux (il suffit de choisir $J = \{i, j\}$) mais la réciproque est fautive à partir d'une famille à trois événements.

Exemple 5.7

On lance deux dés discernables et on considère les événements A : « le premier dé lancé donne un résultat pair », B : « le second dé lancé donne un résultat pair » et C : « la somme des deux dés est un résultat pair ». Montrer que les événements A , B et C sont deux à deux indépendants, mais pas mutuellement indépendants.

Démonstration. Remarquons que A et B sont indépendants par indépendance des lancers des deux dés puis que

$$A \cap B \cap C = A \cap B \text{ et } \mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}.$$

Ainsi : $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \cap B) = \frac{1}{4}$ mais $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{8}$. Ainsi, les événements A , B et C sont deux à deux indépendants, mais pas mutuellement indépendants. $\square$

Propriété 5.8

Soient $(A_1, \dots, A_n)$ une famille d'événements mutuellement indépendants.

1. On obtient encore une famille d'événements mutuellement indépendants en remplaçant un nombre quelconque des A_k par leurs complémentaires.
2. Chacun des événements A_k est indépendant de tout événement B pouvant s'écrire sous la forme de réunions et d'intersections des autres événements de la famille.

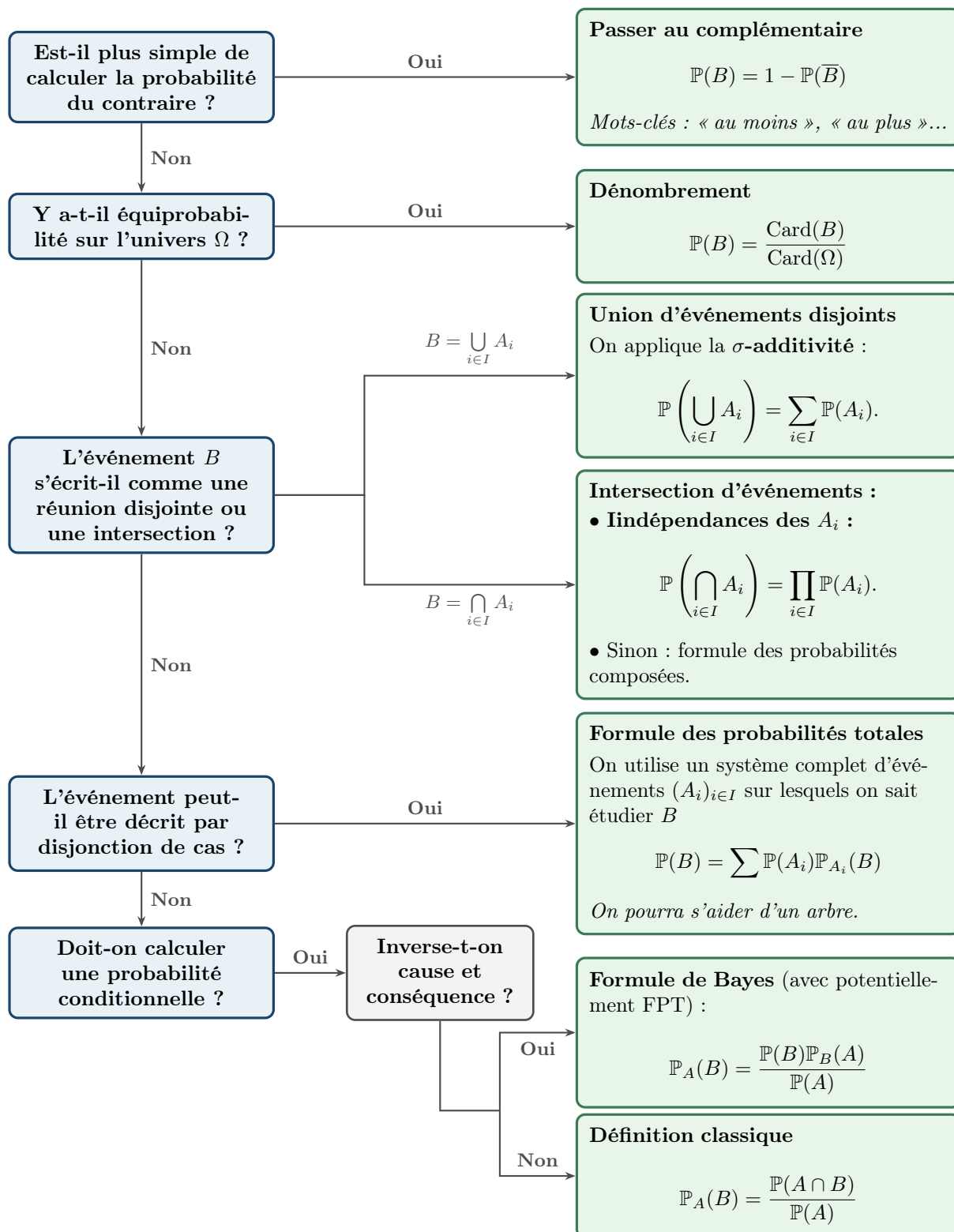
Démonstration admise.

6 Méthodes

6.1 Calculer la probabilité d'un événement

Méthode 6.1 – Calcul de la probabilité d'un événement B

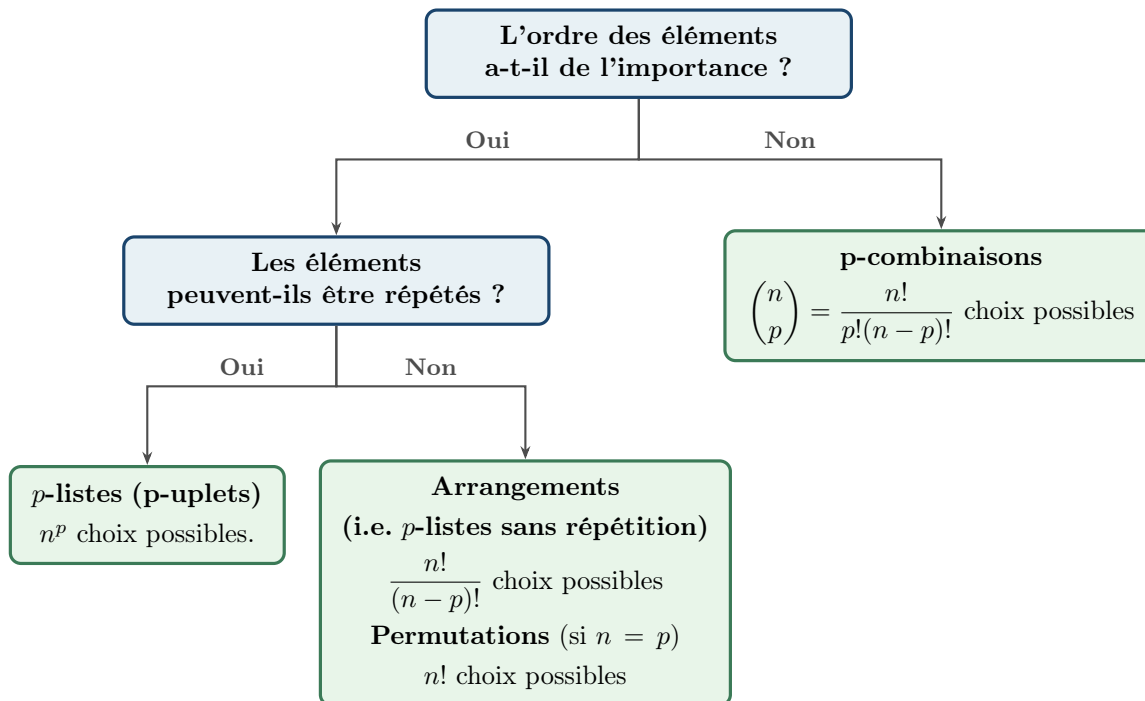
Pour pouvoir évaluer la probabilité d'un événement B , il faut généralement raisonner sur cet événement (et non pas sa probabilité).



6.2 Dénombrer en situation d'équiprobabilité

Méthode 6.2

L'arbre de décision ci-après permet de dénombrer les façons de choisir p éléments dans un ensemble à n éléments où $p \in \llbracket 0, n \rrbracket$.



Le choix d'une p -liste peut être modélisé par p **tirages successifs avec remise** dans un ensemble de cardinal n .

Exemple 6.3 – Dénombrer des p -listes avec répétition

Il y a exactement 2^8 nombres binaires représentables sur 8 bits (soit un octet).

Démonstration. Un nombre binaire sur 8 bits est une liste **ordonnée** de longueur 8 formée de 0 et de 1. Les éléments 0 et 1 pouvant être **répétés**, c'est une 8 liste de l'ensemble $\{0; 1\}$. Il y a donc 2^8 nombres binaires sur 8 bits. On peut même les énumérer : $\overline{00000000}^2$, $\overline{00000001}^2$, $\overline{00000010}^2$, ..., $\overline{11111101}^2$, $\overline{11111110}^2$ et $\overline{11111111}^2$. On peut utiliser une autre méthode : le **principe multiplicatif**. Il y a 2 choix pour le premier bit, 2 pour le second, etc, soit 2^8 choix pour construire un nombre binaire sur 8 bits. $\square$

Le choix d'une p -liste sans répétition peut être modélisé par p **tirages successifs sans remise** dans un ensemble de cardinal n .

Exemple 6.4 – Dénombrer des p -listes sans répétition

Combien y a-t-il de podiums possibles lors d'une course où 8 athlètes s'affrontent (on n'acceptera pas d'ex-æquo) ?

Démonstration. Un podium est une liste ordonnée de trois individus distincts (i.e. sans répétition). Il y a donc $\frac{8!}{5!} = 336$ podiums possibles.

Autre méthode : il y a 8 choix possibles pour la médaille d'or, 7 choix pour celle d'argent puis 6 choix pour celle de bronze. Par le **principe multiplicatif**, il y a donc $8 \times 7 \times 6 = 336$ podiums possibles. $\square$

 Exemple 6.5 – Dénumbrer des permutations ♥

Combien y a-t-il d'anagrammes de la chaîne de caractères "chien" (sans éliminer les mots qui n'existent pas) ?

Démonstration. *Puisqu'aucun caractère n'est répété dans "chien", un anagramme de cette chaîne est une permutation de ses caractères. Il y a donc $5!$ anagrammes de "chien". On pouvait aussi utiliser encore une fois le principe multiplicatif ici.* □

Une p -combinaison d'un ensemble à n éléments est un sous-ensemble à p éléments.

Le choix d'une combinaison à p éléments peut être modélisé par un **tirage simultané de p éléments** (ce qui équivaut à p tirages **sans remise et sans ordre**) dans un ensemble de cardinal n .

 Exemple 6.6 – Dénumbrer des combinaisons ♥

Combien y a-t-il de mains de 5 cartes dans un jeu de 52 cartes ?

Démonstration. *Une main constitue un sous-ensemble de 5 éléments dans un ensemble à 52 éléments. Il y a donc $\binom{52}{5}$ mains de 5 cartes.* □

Démonstrations à connaître

- Exemple 1.11 (décrire des événements complexes).
- Exemple 2.8 (montrer qu'un événement est négligeable ou presque-sûr de deux manières).
- Exemple 4.7 (appliquer la formule des probabilités totales).
- Exemple 4.12 (appliquer la formule des probabilités composées).
- Exemple 4.15 (appliquer la formule de Bayes).
- Exemple 6.3 (dénumbrer des p -listes avec répétition).
- Exemple 6.4 (dénumbrer des p -listes sans répétition).
- Exemple 6.5 (dénumbrer des permutations).
- Exemple 6.6 (dénumbrer des combinaisons).