Chapitre 1

Savez-vous prévoir ?

« *We cannot predict whether a given photon will arrive at A or B. All we can predict is that out of 100 photons that come down, an average of 4 will be reflected by the front surface. Does this mean that physics, a science of great exactitude, has been reduced to calculating only the probability of an event, and not predicting exactly what will happen ? Yes. That's a retreat, but that's the way it is : Nature permits us to calculate only probabilities. Yet science has not collapsed.* »

Richard P. FEYNMAN - The strange theory of light and matter (1985)

Les probabilités sont omniprésentes dans le monde scientifique, c'est pourquoi elles ont fait leur apparition ...tardive en CPGE à la rentrée 2013. Elles tiennent un rôle très important au lycée également mais malheureusement, des notions compliquées sont souvent introduites sans preuve ni rigueur. Vous allez donc reprendre les choses au début en prépa tout en vous restreignant aux probabilités discrètes. Nous allons en faire de même dans ce chapitre : vous pouvez vous contenter d'un survol en première lecture mais il sera utile de revenir sur ce cours avec plus d'attention par la suite pour passer d'une vision lycéenne à un travail de préparationnaire... Le travail sur le cours est une nouvelle fois fondamental : chaque définition doit être comprise, chaque démonstration travaillée.

1.1 Rappels de théorie des ensembles

1.1.1 Ensembles finis

Il est aisé de comprendre intuitivement ce qu'est un ensemble fini. En voici une définition :

Définition (Ensemble fini)
On dit qu'un ensemble E est fini si, et seulement si, il existe un entier naturel n et une bijection φ *de* $[\![1, n]\!]$ *sur E.*

Cela signifie en fait que chaque élément d'un ensemble fini E porte un numéro de dossard et qu'on connaît le nombre total de dossards : n. Ce nombre jouant un rôle important, on lui donne un nom :

Définition (Cardinal d'un ensemble fini)
L'entier n défini précédemment est appelé cardinal de E *et noté Card(E) ou* $|\mathrm{E}|$.

Si on considère l'ensemble formé des pages de cet ouvrage, il constitue un ensemble fini : quel est son cardinal ?
Par convention, l'ensemble vide est de cardinal nul : $|\varnothing| = 0$.

1.1.2 Parties d'un ensemble

Définition (Parties d'un ensemble)
Soit E un ensemble. L'ensemble formé de tous les sous-ensembles de E est appelé ensemble des parties de E *et noté* $\mathfrak{P}(E)$.

Il faudra donc faire bien attention au moment d'utiliser les symboles $\in$ et $\subseteq$. Ainsi :

$$A \in \mathfrak{P}(E) \iff A \subseteq E$$

Si $E_2 = \{a, b\}$ alors $\mathfrak{P}(E_2) = \Big\{\varnothing, \{a\}, \{b\}, \{a, b\}\Big\}$.
Nous aurons également besoin de parler des éléments de E qui n'appartiennent pas à une de ses parties.

Définition (Complémentaire d'une partie)
Soit A une partie de E. On appelle ***complémentaire*** *de A dans E l'ensemble des éléments de E qui n'appartiennent pas à A. On note cet ensemble* $\overline{A}_E$, $\complement_E A$ *ou plus simplement* $\overline{A}$ *quand il n'y a pas d'ambiguïté.*

Par exemple, le complémentaire de l'ensemble des filles dans une classe est l'ensemble des garçons de cette classe. La notation permet de faire le lien avec la négation d'une proposition : la négatiion de « $x \in A$ » est « $x \in \overline{A}$ »...

1.2 Une dose d'algèbre générale

1.2.1 Relation d'ordre

Soit A, B et C trois éléments de $\mathfrak{P}(E)$.
Considérons la relation d'inclusion ($\subseteq$) qui relie deux éléments quelconques de $\mathfrak{P}(E)$.
On obtient facilement que :
- $A \subseteq A$: on dit que la relation $\subseteq$ est *réflexive* ;
- $A \subseteq B \wedge B \subseteq A \longrightarrow A = B$: on dit que la relation $\subseteq$ est *antisymétrique* ;
- $A \subseteq B \wedge B \subseteq C \longrightarrow A \subseteq C$: on dit que la relation $\subseteq$ est *transitive.*

Théorème (la relation d'ordre $\subseteq$)
Ces trois propriétés étant vérifiées, on dit que $\subseteq$ *est une* ***relation d'ordre*** *sur* $\mathfrak{P}(E)$.

Vous vérifierez que la relation $\leqslant$ sur $\mathbb{R}$ a exactement les mêmes propriétés : vous commencez à entrevoir l'essence de l'algèbre qui est de dégager et d'étudier des structures les plus générales possibles afin de simplifier l'étude de phénomènes qui paraissaient différents.

1.2.2 Loi de composition interne

Définition (réunion d'ensembles)
La réunion de deux ensembles A et B notée $A \cup B$ *est l'ensemble des éléments appartenant à A* ***OU*** *B.*

Le « ou » correspond au ou inclusif.

— Comme $A \cup B$ est encore un élément de $\mathfrak{P}(E)$, on dit que $\cup$ est une **loi de composition interne** ;
— comme $A \cup (B \cup C) = (A \cup B) \cup C$, on dit que la loi $\cup$ est **associative** ;
— comme $A \cup B = B \cup A$, on dit que la loi $\cup$ est **commutative** ;
— comme $A \cup \emptyset = A$, on dit que $\emptyset$ est une **élément neutre** pour la loi $\cup$;
— comme $A \cup A = A$, on dit que la loi $\cup$ est **idempotente**.

On peut remarquer que ces propriétés sont vérifiées par l'addition sur $\mathbb{N}$, l'élément neutre étant 0, sauf une : laquelle ?

Définition (intersection d'ensembles)
L'intersection de deux ensembles A et B notée $A \cap B$ *est l'ensemble des éléments appartenant à A* ***ET*** *B.*

Vous vérifierez que la loi $\cap$ a les mêmes propriétés que $\cup$ avec cette fois comme élément neutre E.

1.2.3 Lien avec les opérateurs logiques

Comme souvent en mathématiques, on retrouve des structures semblables dans des domaines différents.
Nous avons déjà découvert les liens entre $\cup$ et OU puis entre $\cap$ et ET.
On peut également faire le rapprochement entre l'implication des propositions et l'inclusion des ensembles, « vrai » correspondant « $\in$ ».
Ainsi, si deux ensembles A et B vérifient $A \subseteq B$, alors il est impossible de trouver un élément x tel que $x \in A$ et $x \notin B$. Tous les autres cas sont possibles.
De même, si $A \longrightarrow B$, il est impossible que A soit vraie et B soit fausse. Toutes les autres propositions sont vraies.
On fait le même rapprochement entre l'équivalence des propositions et l'égalité des ensembles. Merveilleuses mathématiques...

1.2.4 Partition

On a souvent besoin de découper un ensemble en tranches (souvenez-vous de votre cours de terminale sur les probabilités totales...). Voici une définition utile :

Définition (Partition d'un ensemble)
La famille $(A_i)_{i \in [\![1,n]\!]}$ *de parties non vides d'un ensemble E réalise une* ***partition*** *de cet ensemble si, et seulement si :*

$$\forall i \neq j \; A_i \cap A_j = \emptyset \quad et \quad \bigcup_{i=1}^{n} A_i = E$$

Par exemple, une partie A de E et son complémentaire dans E réalise une partition de E car $A \cap \overline{A} = \emptyset$ et $A \cup \overline{A} = E$.

1.2.5 Produit cartésien

Définition (Produit cartésien)
Soit $(A_i)_{i \in [\![1,n]\!]}$ *des parties d'un ensemble E. Le produit* $\prod_{i=1}^{n} A_i = A_1 \times A_2 \times \cdot \times A_n$ *est l'ensemble des n-listes* $(x_1, x_2, \ldots, x_n)$ *telles que* $x_1 \in A_1, \ldots, x_n \in A_n$.

Soit $A = \{1,3,5,7,9\}$ et $B = \{0,2,4,6,8\}$. Alors $(1,0)$ est un élément de $A \times B$ et $(0,1)$ est un élément de $B \times A$. L'ordre des ensembles dans le produit et donc des éléments dans les n-listes est important !

1.3 Quelques résultats sur les cardinaux

1.3.1 Cardinal de $\mathfrak{P}(\mathrm{E})$

Soit $\mathrm{E}_2 = \{a, b\}$ et $\mathrm{E}_3 = \{a, b, c\}$. En fait, $\mathrm{E}_3 = \mathrm{E}_2 \cup \{c\}$. Les parties de E_3 sont donc les parties de E_2 et ces mêmes parties auxquelles on adjoint c.
En effet, $\mathscr{P}(\mathrm{E}_3) = \Big\{\varnothing, \{a\}, \{b\}, \{a,b\}, \varnothing \cup \{c\}, \{a\} \cup \{c\}, \{b\} \cup \{c\}, \{a,b\} \cup \{c\}\Big\}$.
Ainsi $|\mathscr{P}(\mathrm{E}_3)| = 2 \times |\mathscr{P}(\mathrm{E}_2)|$.
Notons de manière générale E_n un ensemble à n éléments avec $n \in \mathbb{N}$ et $\mathrm{E}_{n+1} = \mathrm{E}_n \cup \{\alpha\}$ avec $\alpha \notin \mathrm{E}_n$. On montre facilement que $|\mathscr{P}(\mathrm{E}_{n+1})| = 2 \times |\mathscr{P}(\mathrm{E}_n)|$. Or $\mathscr{P}(\mathrm{E}_0) = \{\varnothing\}$ donc $|\mathfrak{P}(\mathrm{E}_0)| = 1$.
une rapide récurrence permet donc de démontrer que :

Théorème (Cardinal de $\mathfrak{P}$(E))
Si $|\mathrm{E}| = n$ *alors* $|\mathfrak{P}(\mathrm{E})| = 2^n$

1.3.2 Cardinal d'une partition

Il est assez immédiat d'obtenir le résultat suivant :

Théorème (Cardinal d'une partition)
Si la famille $(\mathrm{A}_i)_{i \in [\![1,n]\!]}$ *de parties non vides d'un ensemble E réalise une* ***partition*** *de cet ensemble alors :*

$$|\mathrm{E}| = \sum_{i=1}^{n} |\mathrm{A}_i|$$

Pour compter les élèves du lycée, il suffit de compter les élèves de chaque classe et d'additionner les résultats...

1.3.3 Cardinaux et inclusion

Théorème ()
$\forall (\mathrm{A}, \mathrm{B}) \in \mathfrak{P}(\mathrm{E}) \times \mathfrak{P}(\mathrm{E}),\ \mathrm{A} \subseteq \mathrm{B} \longrightarrow |\mathrm{A}| \leqslant |\mathrm{B}|$

Pour le démontrer, il suffit d'utiliser une petite ruse (bientôt) habituelle : $\mathrm{A} \cap \mathrm{B}$ et $\overline{\mathrm{A}} \cap \mathrm{B}$ donc A et $\overline{\mathrm{A}} \cap \mathrm{B}$ forment une partition de B (car $\mathrm{A} \subseteq \mathrm{B}$) donc, d'après le / théorème 3, $|\mathrm{A}| + \left|\overline{\mathrm{A}} \cap \mathrm{B}\right| = |\mathrm{B}|$, donc $|\mathrm{A}| \leqslant |\mathrm{B}|$
Cette application n'est pas une équivalence ! Prenez $\mathrm{A} = \{1\}$ et $\mathrm{B} = \{3; 4\}$

1.3.4 Cardinal d'un produit cartésien

Voici un autre résultat qui sera bien utile en probabilités :

Théorème (Cardinal d'un produit cartésien)
Soit $(\mathrm{A}_i)_{i \in [\![1,n]\!]}$ *des parties d'un ensemble E.*

$$\left|\prod_{i=1}^{n} \mathrm{A}_i\right| = \prod_{i=1}^{n} |\mathrm{A}_i|$$

En effet, pour chaque n-liste $(x_1, \ldots, x_n)$, il y a $|\mathrm{A}_1|$ choix pour la première composante, $|\mathrm{A}_2|$ choix pour la deuxième, etc.

1.3.5 Formule du crible de Poincaré

Titre

Nous n'en verrons qu'un cas particulier que vous devez déjà connaître :

Théorème (Formule du crible pour deux ensembles)

$\forall (A,B) \in \mathscr{P}(E) \times \mathfrak{P}(E),\ |A \cup B| = |A| + |B| - |A \cap B|$

Pour le démontrer, nous allons former une partition de $A \cup B$ avec $A \cap \overline{B}$, $\overline{A} \cap B$ et $A \cap B$.
Mais on peut également « partitionner » A avec $A \cap \overline{B}$ et $A \cap B$ puis « partitionner » B avec $B \cap \overline{A}$ et $B \cap A$.
On obtient donc que $|A| = \left|A \cap \overline{B}\right| + |A \cap B|$ puis que $|B| = \left|B \cap \overline{A}\right| + |B \cap A|$.
On en déduit que $|A| + |B| = \left|A \cap \overline{B}\right| + \left|B \cap \overline{A}\right| + |A \cap B| + |B \cap A|$ et enfin que :

$$|A| + |B| - |A \cap B| = \left|A \cap \overline{B}\right| + \left|B \cap \overline{A}\right| + |A \cap B| = |A \cup B|$$

1.4 Dénombrement

1.4.1 Nombre de permutations

Définition (Permutation)

Soit E un ensemble. Une permutation de E est une application bijective de E sur E. On note $\mathfrak{S}(E)$ l'ensemble des permutations de E.

Généralement, on note les permutations sous forme d'une matrice où la première ligne correspond aux éléments de E et où la deuxième ligne correspond aux images des éléments de E.
Par exemple, les permutations de $\{a,b,c\}$ sont :

$$\begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix}\begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix}\begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix}\begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}\begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix}\begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}$$

On peut même se contenter d'écrire (a,b,c), (a,c,b), (b,a,c), (b,c,a), (c,a,b), (c,b,a) si l'on est sûr que la première ligne est toujours (a,b,c).

Théorème (Nombre de permutations)

Il y a $n!$ permutations d'un ensemble de n éléments.

La démonstration s'effectue par récurrence. Il y a bien sûr 1 ! façon de permuter 1 élément.
On peut donc supposer qu'il existe au moins un entier k tel qu'il y ait $k!$ permutations d'un ensemble à k éléments.
Considérons un ensemble à $k+1$ éléments. On en choisit 1 : il y a $k+1$ choix possibles.
Ensuite, il s'agit d'ordonner les k éléments restant ce qui laisse $k!$ possibilités.
Il y a donc finalement $(k+1) \times k! = (k+1)!$ permutations au total.
On retrouve dans l'exemple précédent qu'il y a $3! = 6$ permutations de $\{a,b,c\}$.
On conviendra que $0! = 1$.
On a déjà rencontré les permutations lors de notre découverte de la cryptographie. Elles jouent également un rôle important dans les algorithmes de tri par exemple et aussi dans certains algorithmes concernant les graphes que vous étudierez en deuxième année, dans les mélanges de cartes, dans l'étude du Rubik's cube, et bien d'autres domaines encore...

Remarque (Formule de STIRLING)
Pour n « grand », on peut estimer que :

$$n! \underset{+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$$

1.4.2 Nombre d'applications injectives - arrangements sans répétitions

Une source veut envoyer à chacun de ses cinq clients un fichier différent parmi les vingt-six situés dans un certain répertoire.
Combien de possibilités a la source pour envoyer un fichier à chacun des 26 clients ?
On peut modéliser cette situation : il s'agit de compter les applications injectives (i.e. tel que chaque élément de l'ensemble de départ admet au plus une image dans l'ensemble d'arrivée) de l'ensemble E des clients vers l'ensemble F des fichiers. Il ne peut y avoir répétition donc il y a 26 choix de fichiers pour le premier client, 25 pour le deuxième, etc. ce qui donne $26 \times 25 \times 24 \times 22 \times 21$ possibilités.
On peut généraliser.

Définition (Arrangement sans répétition)
Une application injective d'un ensemble E de cardinal p dans un ensemble F de cardinal n (avec $p \leqslant n$) est appelée un arrangement sans répétition de p parmi n.

Un exemple fameux est la recherche de tiercés dans l'ordre (sans ex-æquo...) : pourquoi ?
Nous pouvons formuler cette propriété plus synthétiquement. Notons A_n^p le nombre d'arrangements de p parmi n. Alors :

$$\mathrm{A}_n^p = n(n-1)(n-2)\cdots(n-p+1) \times \frac{(n-p)(n-p-1)(n-p-2)\times\cdots\times 2\times 1}{(n-p)(n-p-1)(n-p-2)\times\cdots\times 2\times 1} = \frac{n!}{(n-p)!}$$

d'où

Théorème (Nombre d'arrangements)
Le nombre d'arrangements sans répétition de p parmi n est

$$\mathrm{A}_n^p = n(n-1)(n-2)\cdots(n-p+1) = \frac{n!}{(n-p)!}$$

1.4.3 Combinaisons sans répétition

Il n'y a pas assez de gagnants au loto. Les règles en sont donc modifiées. Il s'agit maintenant de trouver 3 numéros parmi 5. Combien y a-t-il de grilles (combinaisons) possibles ?
Par exemple, on pourrait dire que j'ai cinq manières de choisir le premier numéro, quatre choix pour le deuxième et trois choix pour le troisième, donc il y a $5\times4\times3$ grilles différentes mais dans ce cas, je compte des 3-listes ordonnées alors que les 3-listes $(1,2,3)$ et $(3,2,1)$ correspondent à la même grille ou combinaison $\{1,2,3\}$.
Posons une petite définition pour clarifier les débats. Donnons en fait un nom à une grille du loto, c'est à dire à un sous-ensemble (une partie) contenant p éléments d'un plus grand ensemble contenant n éléments.

Définition (Combinaison sans répétition)
Soit n et p deux entiers naturels et E *un ensemble contenant n éléments. Un sous-ensemble de* E *contenant p éléments est appelé une* **combinaison** *de p éléments de* E *ou encore une* p-**combinaison** *d'éléments de* E.

Or ce qui nous intéresse, c'est le nombre de ces combinaisons, donc introduisons une notation :

Définition (Nombre de combinaisons sans répétitions)
Le nombre de p-combinaisons d'un ensemble E contenant n éléments est noté $\binom{|E|}{p}$ *ou encore* $C_{|E|}^{p} = C_{n}^{p}$

Revenons à notre mini-loto. Considérons une grille quelconque (i.e. une 3-combinaison de l'ensemble des 5 numéros) : par exemple $\{2,4,5\}$. Nous avons vu dans un paragraphe précédent qu'il y a 3! facons d'ordonner ces nombres. Finalement, il y a $C_5^3 \times 3!$ suites de 3 nombres ordonnées : c'est le nombre d'arrangements de p parmi n. Or nous en avons comptées $5 \times 4 \times 3$ tout à l'heure. Nous en déduisons finalement que

$$C_5^3 = \frac{A_5^3}{3!} = \frac{5 \times 4 \times 3}{3!}$$

Il est alors possible de généraliser la formule suivante :

Théorème (Nombre de combinaisons)

$$C_n^p = \frac{A_n^p}{p!} = \frac{n(n-1)(n-2)\cdots(n-(p-1))}{p!} = \frac{n!}{p!(n-p)!}$$

1.5 Triangle de Pascal - Binôme de Newton

À l'aide des formules précédentes, on prouve facilement (faites-le !) le résultat suivant :

Théorème ()

$$C_n^p = C_n^{n-p}$$

On établit ensuite, toujours par le calcul, la relation suivante, dite **Relation de Pascal** même si les mathématiciens chinois l'avaient mise en évidence avant lui :

Théorème (Relation de PASCAL)

$$C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$$

Il est possible de démontrer cette formule à l'aide d'un raisonnement ensembliste : pour former des groupes de p éléments dans un ensemble en contenant n, on distingue un élément n_0 quelconque :
— soit le groupe le contient et alors il y a C_{n-1}^{p-1} choix des $p-1$ éléments distincts de n_0 parmi les $n-1$ restant ;
— soit le groupe ne le contient pas et il faut donc choisir p éléments parmi les $n-1$ éléments distincts de n_0 : il y a donc cette fois C_{n-1}^p.

Cette relation permet de démontrer la formule du binôme de Newton (que nous admettrons) :

Théorème (Formule du binôme)
Soit $\mathbb{A}$ *un anneau commutatif. Dans* $\mathbb{A}[X]$, *pour tout entier naturel non nul n :*

$$(1+X)^n = \sum_{k=0}^{n} C_n^k X^k$$

et plus généralement :

$$\forall (a,b) \in \mathbb{A}^2, \quad (a+b)^n = \sum_{k=0}^{n} \mathrm{C}_n^k a^k b^{n-k}$$

1.6 Probabilités ?

Le calcul des probabilités est très important dans toute science. Il nous permet de résoudre des problèmes tout en étant confrontés à l'incertitude ou à des systèmes très complexes. Contrairement aux clichés, la théorie des probabilités, en tant que discipline mathématique, est dotée de toute la rigueur et de la précision possibles. Elle puise ses sources dans tous les champs des mathématiques. Il n'est dès lors pas étonnant de constater que les prix mathématiques les plus prestigieux sont attribués depuis quelque temps à des travaux liés aux probabilités.
Il ne faut pas confondre les probabilités avec la statistique qui complète un modèle inconnu à l'aide d'observations (raisonnement inductif) et fait plutôt appel à l'intuition.
Enfin, s'il est tout à fait possible de calculer des probabilités sans faire de statistique.

1.7 Avant la formalisation

Essayons, grâce à notra petite expérience probabiliste acquise au lycée, de dégager les grandes lignes de ce que pourrait être une théorie formalisée, nécessaire au « dialogue » avec une machine.
Lors de l'étude des probabilités, on est amené à effectuer une **expérience aléatoire**, c'est-à-dire un processus dont on ne peut prévoir à l'avance le résultat qui peut donner des résultats différents même si l'on répète l'expérience dans des conditions qui semblent identiques.
Nous allons considérer par la suite le lancement d'un dé cubique équilibré et tout et tout.
On doit d'abord connaître l'ensemble des résultats possibles : c'est l'**univers** que nous noterons par la suite Ω.
Par exemple, si je lance une fois le dé, $\Omega_1 = \{1,2,3,4,5,6\}$.
Si je lance deux fois de suite le dé et que je note dans l'ordre le numéro des faces, j''obtiens dans ce cas $\Omega_2 = \{11,12,13,...,66\}$
Si je lance les dés jusqu'à obtenir deux 6 à la suite, le nombre de lancers n'est maintenant pas fixé à l'avance et l'univers est a priori infini mais en demeurant discret : chaque évènement peut en effet être « numéroté ». Plus rigoureusement, Ω_3 peut être mis en bijection avec $\mathbb{N}$: $\Omega_3 = \{1366, 3546213466, 55236221466, ...\}$.
Il faut maintenant étudier la notion d'« évènement [1] ». Disons dans un premier temps qu'il s'agit d'une assertion vérifiable relativement au résultat d'une expérience aléatoire. Il est donc « associé » à un sous-ensemble (une partie) de Ω. Un élément de Ω sera associé à un **évènement élémentaire**.
Reprenons le deuxième exemple. Voici quelques évènements :

- « Obtenir deux fois 3 » : $\mathrm{E}_a = \{33\}$. C'est un évènement élémentaire ;
- « Ne pas obtenir deux fois 3 » : $\mathrm{E}_b = \Omega_2 \setminus \mathrm{E}_a \subseteq \Omega_2$;
- « Obtenir deux faces dont la somme des numéros fait 11 » : $\mathrm{E}_c = \{56,65\} \subseteq \Omega_2$;

1. Depuis 1990, l'Académie française préconise de mettre un accent grave au deuxième e d'événement qui doit donc s'écrire évènement...

— « Obtenir deux faces dont le numéro est inférieur à 7 » : $E_d = \Omega_2$. Il s'agit donc de l'**évènement certain** ;
— « Obtenir deux faces dont la somme des numéros fait 37 » : $E_e = \emptyset$. Il s'agit cette fois de l'**évènement impossible**.

On comprend que l'univers lui-même et l'ensemble vide sont des évènements. On voudrait que la réunion et l'intersection d'évènements soit encore un évènement et que le complémentaire de la partie associée à un évènement définisse encore un évènement : les mathématiciens (français...) disent alors que l'ensemble de tous ces évènements forment une **tribu**. Dans le reste du monde, on parle plutôt de σ-algèbre de Boole : les ensembles concernés vérifient en effet certaines propriétés *algébriques*.
Il reste enfin à définir un moyen de **mesurer** les probabilités : c'est une sorte de degré de certitude correspondant à une échelle graduée de 0 à 1, 0 correspondant au degré de l'évènement impossible et 1 au degré de l'évènement certain. Il est important de comprendre que ce degré de certitude est donné **a priori** et permet donc de prévoir l'issue de l'expérience avant de la réaliser.
Il faudrait que cette mesure assure une cohérence entre les évènements et les degré.
Par exemple, si $A \subseteq B$, on aimerait que $\mathbb{P}(A) \leqslant \mathbb{P}(B)$. On aimerait également que l'union de deux parties de l'univers se « transforme » en somme de leur degré de certitude, mais on « sent » bien que cela ne pourra être vrai que si les évènements sont totalement **incompatibles** : $A \cap B = \emptyset \Longrightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. On retrouve ici cette notion algébrique de « transport » des opérations : la notion d'*isomorphisme*.

Considérons pour illustrer ces dernières notions le premier exemple du lancer unique d'un dé.
Notre univers Ω_1 contient 6 éléments et prenons comme tribu naturellement associée les $2^6 = 64$ parties de Ω_1 qui forment l'ensemble $\mathbb{P}(\Omega_1)$.
Imposons pour notre modèle que la mesure du degré de certitude ne dépende pas des numéros des faces mais uniquement du cardinal de la partie considérée : chacun des six évènement élémentaires aura donc la même probabilité : $\mathbb{P}(1) = \mathbb{P}(2) = \cdots = \mathbb{P}(6) = \frac{1}{6}$.
La probabilité d'obtenir un numéro pair sera donc, sachant que l'évènement est associé à $A = \{2, 4, 6\}$: $\mathbb{P}(A) = \mathbb{P}(2) + \mathbb{P}(4) + \mathbb{P}(6) = \frac{1}{2}$.
Mais si le dé est truqué ? Si l'univers est infini ? S'il n'est pas dénombrable ?
Il est temps de définir clairement et rigoureusement les notions afin de ne pas dire de bêtises...

1.8 Espace probabilisable - Espace probabilisé

Dans ce chapitre, conformément à votre futur programme de prépa, Ω est un ensemble fini. Nous le considérerons parfois dénombrable infini.

Définition (Tribu)
Une tribu $\mathcal{T}_\Omega$ définie sur un univers Ω est une partie de $\mathbb{P}(\Omega)$ qui vérifie :

1. $\Omega \in \mathcal{T}_\Omega$;
2. $A \in \mathcal{T}_\Omega \Longrightarrow \complement_\Omega A \in \mathcal{T}_\Omega$;
3. $\forall I \subseteq \mathbb{N}, (A_i)_{i \in I} \in \mathcal{T}_\Omega^{|I|} \Longrightarrow \bigcup_{i \in I} A_i \in \mathcal{T}_\Omega$.

Attention ! Une tribu est une partie de $\mathbb{P}(\Omega)$: c'est donc un ensemble de sous-ensembles de Ω !

Définition (Espace probabilisable)
Soit $\mathscr{T}_\Omega$ une tribu définie sur l'univers Ω. On dit que $(\Omega, \mathscr{T}_\Omega)$ est un espace probabilisable (ou espace d'évènements). Les éléments de la tribu sont appelés ***évènements****.*

Attention ! Pour une certaine expérience aléatoire, il peut y avoir de nombreux espaces probabilisables différents.
Reprenons un des exemples introduits dans la section précédente :
— $(\Omega_1, \{\emptyset, \Omega_1\})$ est un espace probabilisable : on ne pourra s'intéresser qu'aux évènements « obtenir un numéro entre 1 et 6 » et « ne pas obtenir de numéro entre 1 et 6 » ce qui est un peu sommaire ;
— $(\Omega_1, \{\emptyset, A, \complement_\Omega A, \Omega_1\})$ avec A associé à « obtenir un 6 » ;
— $(\Omega_1, \mathbb{P}(\Omega_1))$ est ce qu'on peut faire de plus précis.
Vérifiez qu'il s'agit bien d'espaces probabilisables !

Définition (Axiomes de KOLMOGOROV)
Soit $(\Omega, \mathscr{T})$ un espace probabilisable. On appelle ***mesure de probabilité*** *(ou probabilité tout court) une application Pr définie sur $\mathscr{T}$ vérifiant :*

1. $\forall A \in \mathscr{T}$, $\mathbb{P}(A) \in [0,1]$;
2. $\mathbb{P}(\Omega) = 1$;
3. $\forall (A,B) \in \mathscr{T} \times \mathscr{T}$, $A \cap B = \emptyset \Longrightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

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

Vous pouvez vérifier que l'espace probabilisable $(\Omega, \mathbb{P}(\Omega))$ avec $\Omega = \{1,2,3,4,5,6\}$ muni de l'application :

$$\mathbb{P} : \begin{array}{ccc} \mathbb{P}(\Omega) & \to & [0,1] \\ A & \mapsto & \frac{|A|}{6} \end{array}$$

définit un espace probabilisé.
Pour l'axiome 3, on se souviendra de la formule du crible (cf théorème 6 page 9).

Comme Ω est fini, il existe une certaine famille I d'entiers telle que $\Omega = \bigcup_{i\in I}\{\omega_i\}$ et pour tout évènement A de $\mathscr{T}$, il existe une famille $J \subseteq I$ telle que $A = \bigcup_{j\in J}\{\omega_j\}$.
Alors, par application itérée du troisième axiome :

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{j\in J}\{\omega_j\}\right) = \sum_{j\in J}\mathbb{P}\left(\{\omega_j\}\right)$$

Ainsi la donnée des valeurs prises par $\mathbb{P}$ sur les évènements élémentaires suffit à caractériser $\mathbb{P}$ sur $\mathscr{T}$.
Dans le cas où Ω est fini et si tous les évènements élémentaires ont la même probabilité (on dit qu'il y a **équiprobabilité**), alors :

$$\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = \cdots = \mathbb{P}(\omega_{|\Omega|}) = \frac{1}{|\Omega|}$$

Dans ce cas, $\mathbb{P}$ est appelée **probabilité uniforme** sur Ω.
Attention ! Toute probabilité n'est pas uniforme...Votre portable sonne : quelle est la probabilité que cela soit Monica B. qui vous fasse une déclaration d'amour ? La tribu contient deux évènements en plus de l'évènement certain et de l'évènement impossible : « c'est

Monica » et « ce n'est pas Monica ». Le rapport entre cas favorables et cas possibles est bien $\frac{1}{2}$ et pourtant...

Vous noterez enfin que lorsque l'univers est infini dénombrable, il ne peut pas y avoir équiprobabilité : on aurait $\sum_{n\geqslant .} p = p\sum_{n\geqslant .} 1 \neq 1$...

Théorème (Principales propriétés)
Soit $(\Omega, \mathcal{T}, \mathbb{P})$ *un espace probabilisé. La probabilité* $\mathbb{P}$ *vérifie :*

1. $\mathbb{P}\left(\complement_\Omega A\right) = 1 - \mathbb{P}(A)$;
2. $\mathbb{P}(\emptyset) = 0$;
3. $\forall (A, B) \in \mathcal{T} \times \mathcal{T},\ A \subset B \longrightarrow \mathbb{P}(A) \leqslant \mathbb{P}(B)$;
4. $\forall (A, B) \in \mathcal{T} \times \mathcal{T},\ \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ *(cas particulier de la formule du crible de* POINCARÉ*).*

Dans la suite, quand il n'y aura pas ambiguïté, on notera $\complement_\Omega A$ par $\overline{A}$.
Les démonstrations utilisent les axiomes de KOLMOGOROV.

Définition (Système complet d'évènements)
Une famille $A_1, A_2, \dots, A_n$ *d'évènements non vides de* $\mathcal{T}$ *forme un système complet d'évènements si, et seulement si :*

— $\forall i \neq j,\ A_i \cap A_j = \emptyset$ *(évènements incompatibles deux à deux);*

— $\displaystyle\bigcup_{i=1}^{n} A_i = \Omega$.
On dit aussi que la famille (A_i) *forme une* ***partition*** *de* Ω.

Théorème (Théorème des probabilités totales (version 1))
Soit $(A_i)_{1\leqslant i\leqslant n}$ *un système complet d'évènements, alors :*

$$\forall E \in \mathcal{T},\ \mathbb{P}(E) = \sum_{i=1}^{n} \mathbb{P}(E \cap A_i)$$

La preuve s'obtient en utilisant les axiomes de KOLMOGOROV.

1.9 Probabilités conditionnelles

Dans toute cette section on travail dans l'espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.

1.9.1 Un exemple pour comprendre

Considérons l'expérience simplissime consistant à lancer deux fois un dé à six faces. L'univers Ω est donc constitué de l'ensemble des couples (i, j), avec i et j appartenant à l'ensemble $[\![1, 6]\!]$: il y a donc 36 éléments dans Ω. Intéressons nous à la somme des deux chiffres et soit A l'événement « le total fait neuf ». On prendra $\mathcal{T} = \mathbb{P}(\Omega)$.

$$A = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$$

donc

$$\mathbb{P}(A) = \frac{4}{36} = \frac{1}{9}$$

Soit B l'événement : « on obtient 3 au premier lancer », alors

$$\mathrm{B} = \{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\}$$

donc

$$\mathbb{P}(\mathrm{B}) = \frac{6}{36} = \frac{1}{6}$$

Jusqu'ici, tout roule. Posons-nous maintenant le problème suivant : sachant que B est réalisé, i.e. que l'on obtient 3 au premier lancer, quelle est la probabilité que le total fasse 9 ? On ne considère plus tous les éléments de Ω. Il semble alors nécessaire de définir un nouvel univers (notre nouvel ensemble des possibles) :

$$\Omega' = \{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\} = \Omega \cap \mathrm{B}$$

Puisque l'univers change, la probabilité aussi. L'événement A' « le total fait neuf » dans ce nouveau modèle s'écrit

$$\mathrm{A}' = \{(3,6)\} = \mathrm{A} \cap \mathrm{B}$$

et donc

$$\mathbb{P}'(\mathrm{A}') = \frac{1}{6} = \frac{\text{nombre d'éléments de } \mathrm{A}\cap\mathrm{B}}{\text{nombre d'éléments de } \mathrm{B}} = \frac{\mathbb{P}(\mathrm{A}\cap\mathrm{B})}{\mathbb{P}(\mathrm{B})}$$

1.9.2 Définition

Cela nous incite à poser la définition suivante :

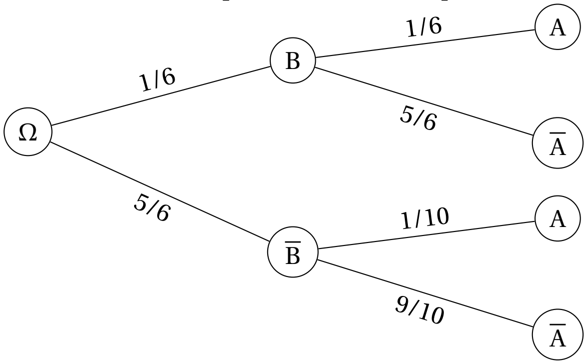
Définition (Probabilité conditionnelle)
Soient A *et* B *deux événements d'un espace probabilisé muni d'une probabilité* $\mathbb{P}$*, avec* $\mathbb{P}(\mathrm{B}) \neq 0$. ***La probabilité de A sachant B*** *est définie par*

$$\mathbb{P}(\mathrm{A} \mid \mathrm{B}) = \frac{\mathbb{P}(\mathrm{A}\cap\mathrm{B})}{\mathbb{P}(\mathrm{B})}$$

On parle de *probabilité* conditionnelle. Mais est-ce bien une probabilité ? C'est-à-dire est-ce que $(\Omega, \mathscr{T}, \mathbb{P}(\cdot \mid \mathrm{B}))$ est un espace probabilisé ? D'après la définition 15 page 14, il y a trois conditions à vérifier (faites-le !...).

1.9.3 Arbre

Reprenons l'exemple de la section 1.9.1 page précédente en répondant aux mêmes questions à l'aide d'un arbre *pondéré*, c'est à dire un arbre dont chaque branche est marquée de la probabilité (du *poids*) correspondant.
Alors la somme des probabilités de chaque « ramification » est égale à 1.



À l'aide de la formule des probabilités totales, il est aisé d'obtenir les résultats suivant :

1. **Calcul de** $\mathbb{P}(A)$
   B et $\overline{B}$ forment une partition. D'après la formule des probabilités totales,

   $$\mathbb{P}(A) = \mathbb{P}(B \cap A) + \mathbb{P}(\overline{B} \cap A) = \mathbb{P}(B) \times \mathbb{P}(A \mid B) + \mathbb{P}(\overline{B}) \times \mathbb{P}(A \mid \overline{B})$$

   $$\mathbb{P}(A) = \tfrac{1}{6} \times \tfrac{1}{6} + \tfrac{5}{6} \times \tfrac{1}{10} = \tfrac{1}{9}$$

2. **Calcul de** $\mathbb{P}(\overline{A})$
   B et $\overline{B}$ forment une partition. D'après la formule des probabilités totales,

   $$\mathbb{P}(\overline{A}) = \mathbb{P}(B \cap \overline{A}) + \mathbb{P}(\overline{B} \cap \overline{A}) = \mathbb{P}(B) \times \mathbb{P}(\overline{A} \mid B) + \mathbb{P}(\overline{B}) \times \mathbb{P}(\overline{A} \mid \overline{B})$$

   $$\mathbb{P}(\overline{A}) = \tfrac{1}{6} \times \tfrac{5}{6} + \tfrac{5}{6} \times \tfrac{9}{10} = \tfrac{8}{9}$$

3. **Probabilités conditionnelles**

   $$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\frac{1}{36}}{\frac{1}{9}} = \frac{1}{4} \qquad \mathbb{P}(B \mid \overline{A}) = \frac{\mathbb{P}(B \cap \overline{A})}{\mathbb{P}(\overline{A})} = \frac{\frac{5}{36}}{\frac{8}{9}} = \frac{5}{32}$$

   $$\mathbb{P}(\overline{B} \mid A) = \frac{\mathbb{P}(\overline{B} \cap A)}{\mathbb{P}(A)} = \frac{\frac{1}{12}}{\frac{1}{9}} = \frac{3}{4} \qquad \mathbb{P}(\overline{B} \mid \overline{A}) = \frac{\mathbb{P}(\overline{B} \cap \overline{A})}{\mathbb{P}(\overline{A})} = \frac{\frac{3}{4}}{\frac{8}{9}} = \frac{27}{32}$$

1.9.4 Formule de BAYES

Cette formule correspond au « retournement » d'un arbre.

Théorème (Deuxième formule de BAYES (probabilité des causes))
Soit $(\Omega, \mathcal{T}, \mathbb{P})$ *un espace probabilisé et* $(A_i)_{i \in I \subseteq \mathbb{N}}$ *un système complet d'évènements.*

$$\forall B \in \mathcal{T}, \quad \mathbb{P}(A_k \mid B) = \frac{\mathbb{P}(B \mid A_k) \times \mathbb{P}(A_k)}{\sum\limits_{i \in I} \mathbb{P}(B \mid A_i) \times \mathbb{P}(A_i)}$$

Preuve ?
On utilisera souvent ce théorème dans le cas particulier de deux évènements :

Théorème (Première formule de BAYES)

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A) \times \mathbb{P}(A)}{\mathbb{P}(B)}$$

On aurait pu écrire :

$$\mathbb{P}(\text{cause} \mid \text{effet}) = \frac{\mathbb{P}(\text{effet} \mid \text{cause}) \times \mathbb{P}(\text{cause})}{\mathbb{P}(\text{effet})}$$

Statistiquement, on a accès aux effets connaissant les causes (c'est le déterminisme...) mais dans la réalité, on subit les effets et on voudrait en connaître la cause.
Le cas le plus célèbre d'utilisation de la *statistique bayesienne* est le test de grossesse. Les médecins « disposent » d'un échantillon de femmes enceintes ou non (et donc de $\mathbb{P}(E)$ et $\mathbb{P}(\overline{E})$) et d'un test de grossesse. Ils font subir le test à tout l'échantillon et disposent donc de statistiques sur $\mathbb{P}(T \mid E)$, $\mathbb{P}(\overline{T} \mid \overline{E})$, $\mathbb{P}(T \mid \overline{E})$ et $\mathbb{P}(\overline{T} \mid E)$ mais ce qui intéresse le médecin et surtout la femme qui utilise le test c'est $\mathbb{P}(E \mid T)$ et $\mathbb{P}(\overline{E} \mid \overline{T})$...

1.9.5 Indépendance

La notions d'évènements indépendants est l'une des difficultés du calcul des probabilités. Comme souvent, cette notion purement mathématique renvoie, par son appellation, à une notion intuitive utilisée dans le langage courant. Il faut bien garder en mémoire que le mot du vocabulaire courant est souvent un « faux ami ». Donnons tout d'abord sa définition.

Définition (Évènements indépendants)
Les évènements A *et* B *sont dits* ***indépendants*** *si* $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$

Ainsi, dans notre exemple de dés, on a $\mathbb{P}(A) \times \mathbb{P}(B) = \frac{1}{6} \times \frac{6}{36} = \frac{1}{6} = \mathbb{P}(A \cap B)$, donc le lancer du $2^{\text{ème}}$ dé est indépendant du premier, ce qui est rassurant.
Cette notion est purement abstraite et ne renvoie qu'à des propriétés mathématiques dont la principale est :

Théorème (Indépendance et probabilités conditionnelles)
Soient A *et* B *deux évènements tels que* $\mathbb{P}(B) \neq 0$. *Les évènements* A *et* B *sont indépendants si et seulement si* $\mathbb{P}(A \mid B) = \mathbb{P}(A)$

On peut faire dire ce que l'on veut à des probabilités selon le modèle choisi.
Supposons que sur un groupe de 100 personnes, 20 portent des sous-vêtements bleus, 50 jouent au football et 10 font les deux à la fois. On met ces 100 personnes dans une boîte et on en tire une au hasard. Vérifiez que les évènements « la personne tirée porte des sous-vêtements bleus » et « la personne tirée joue au football » sont indépendants.
Étudiez le même problème en considérant cette fois-ci que 15 personnes jouent au football (pourquoi pas). [2]
Veillez à ne pas confondre évènements *indépendants* et évènements *incompatibles*.
On peut montrer d'ailleurs que deux évènements incompatibles de probabilité non nulle ne sont jamais indépendants.
En effet, A et B sont incompatibles si, et seulement si, $A \cap B = \emptyset$, donc $\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0$. Si $\mathbb{P}(A) \neq 0$ et $\mathbb{P}(B) \neq 0$, on a forcément $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \times \mathbb{P}(B)$.
La seule idée à retenir est que , si A et B sont indépendants, **avoir observé la réalisation de *A* ne modifie pas la probabilité d'une éventuelle réalisation de *B*.**
Ainsi, en supposant que la Française des Jeux n'utilise pas de boules truquées, on peut considérer que deux tirages successifs du loto sont indépendants.

Théorème (Indépendance mutuelle)
Soit $(A_i)_{i \in I \subseteq \mathbb{N}}$ *une suite d'événements de* $\mathscr{T}$.

1. *On dit que ces événements sont* ***indépendants deux à deux*** *si, et seulement si :*

$$\forall (i, j) \in I^2, i \neq j, \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \times \mathbb{P}(A_j)$$

2. *On dit que ces événements sont* ***mutuellement indépendants*** *si, et seulement si, pour toute partie J de I on a :*

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

2. De façon générale, la définition probabiliste de l'indépendance est plus large que la notion intuitive.

1.10 Variables aléatoires réelles finies

On travaillera *en général* dans l'espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$ avec $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$ un univers fini de cardinal n.

1.10.1 Définition

L'ensemble $\mathbb{P}(\mathbb{R})$ est bien une tribu mais il pose quelques problèmes lors d'une étude plus approfondie des probabilités. Dans $\mathbb{R}$, on utilisera donc une tribu particulière qui correspond en fait à la tribu à laquelle on pense « naturellement », celle crée à partir des intervalles :

Définition (Tribu borélienne)
On note $\mathbf{B}(\mathbb{R})$ *l'ensemble des parties engendrées (par intersections et réunions dénombrables) par les intervalles de* $\mathbb{R}$.

Vous vérifierez qu'il s'agit d'une tribu.

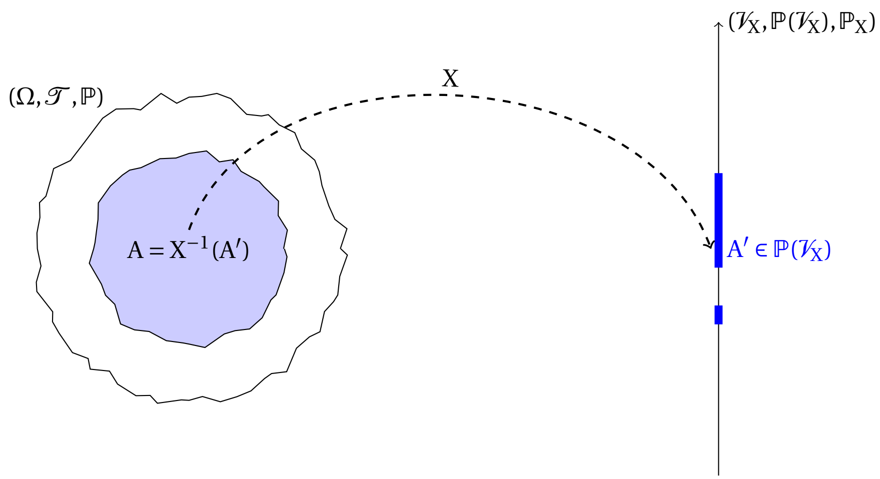
Définition (Variable aléatoire réelle)
On appelle variable aléatoire réelle finie toute application de Ω *dans* $\mathbb{R}$ *vérifiant :*

$$\forall I \subseteq \mathbb{R},\ X^{-1}(I) \in \mathscr{T}$$

Cette définition nous assure que $X^{-1}(I)$ est un événement dont on peut calculer la probabilité.
Une v.a. est *réelle* car X est à valeurs dans $\mathbb{R}$ et *finie* car l'univers est fini donc l'ensemble image de X aussi.

Attention!
Vous noterez qu'une variable *aléatoire réelle (v.a.r. pour les intimes) est en fait une* fonction *!...*

Nous noterons $\mathscr{V}_X = X(\Omega)$ l'ensemble des valeurs prises par la v.a. X.



Soit $x_1, \cdots, x_k$ les différentes valeurs prises par la fonction X.
On note $\{X = x_i\}$ l'événement « la variable aléatoire prend la valeur x_i ». Il se note rigoureusement $\{\omega \in \Omega \mid X(\omega) = x_i\}$, ce qui se lit « l'ensemble des ω tels que $X(\omega) = x_i$ » ou encore $X^{-1}(\{x_i\})$.

La v.a.r. X induit une nouvelle probabilité (ou plutôt loi de probabilité comme nous allons le voir) $\mathbb{P}_X$ sur $(\mathbb{R}, \mathbf{B}(\mathbb{R}))$. Par exemple, $\mathbb{P}_X(\{2\}) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = 2\})$.

La définition précédente n'est pas très parlante, il faut essentiellement retenir que l'on va s'intéresser à des événements décrits par des nombres, les variables aléatoires nous permettent de décrire, et donc de noter, facilement des événements à l'aide de nombres.
Si $\mathcal{V}_X$ est réduit à un singleton, $\mathcal{V}_X = \{\alpha\}$, on dit que X est une variable certaine, on note alors $X = \alpha$.

1.10.2 Loi de probabilité

Définition (Loi de probabilité d'une v.a.r.)
Soit X *une variable aléatoire réelle sur* Ω*. On appelle* ***loi de probabilité*** *de* X *la fonction* $\mathbb{P}_X$ *de* $\mathbf{B}(\mathbb{R})$ *dans* $[0,1]$ *définie par*

$$\mathbb{P}_X : \begin{array}{rcl} \mathbf{B}(\mathbb{R}) & \to & [0,1] \\ B & \mapsto & \mathbb{P}(\{X^{-1}(B)\}) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\}) = \mathbb{P}(X \in B) = \mathbb{P}_X(B) \end{array}$$

Vérifiez que $\mathbb{P}_X$ vérifie les axiomes de Kolmogorov donc que $\mathbb{P}_X$ est bien une mesure de probabilité sur $(\mathbb{R}, \mathbf{B}(\mathbb{R}))$.
Dans la plupart des cas, on étudiera $\mathbb{P}(X = x)$, c'est-à-dire $\mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x\})$.
Si $x \notin X(\Omega)$, alors $\{X = x\} = \emptyset$ et donc $\mathbb{P}(\{X = x\}) = 0$.
Définir *la loi de probabilité d'une expérience aléatoire* reviendra donc à :

- déterminer toutes les valeurs possibles $x_1, \cdots, x_n$ prises par X;
- calculer les probabilités $p_1, \cdots, p_n$ des événements correspondants;
- regrouper les résultats dans un tableau du type :

Valeurs prises par X	x_1	x_2	$\cdots$	x_n
Probabilité correspondante $\mathbb{P}(\{X = x_i\})$	p_1	p_2	$\cdots$	p_n

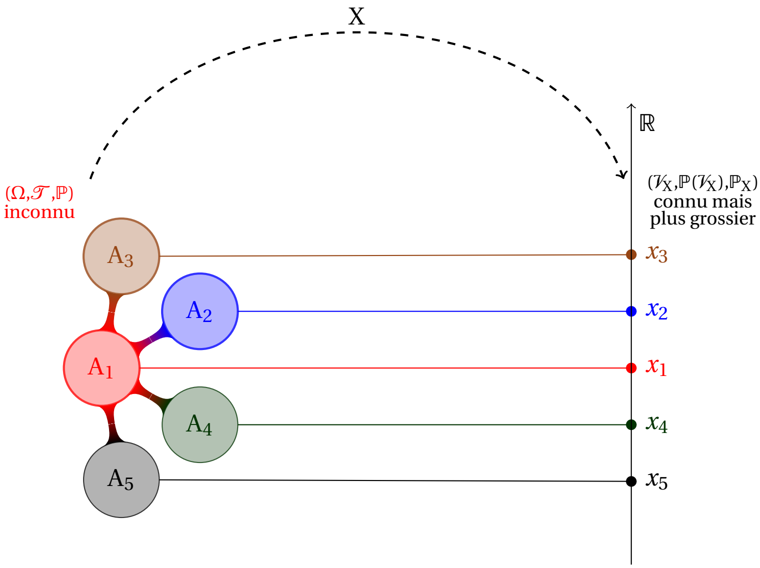
Vous n'oublierez pas de vérifier que $p_1 + \cdots + p_n = 1$ d'après la loi des probabilités totales.
En général, on se trouve dans la situation suivante :

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[d, "\mathbb{P} \text{ inconnue}"'] & \mathbb{R} \arrow[d, "\mathbb{P}_X \text{ connue}"] \\ \mathbb{R} & \mathbb{R} \end{tikzcd}$$

Remarque
Dans la vraie vie, on connaît les valeurs prises par la v.a.r. mais on ne connaît pas bien Ω*, voire pas du tout. On réfléchit donc sur les lois et non pas sur des ensembles comme au lycée. Par exemple, on peut vouloir réfléchir au nombre de plantages d'un réseau. On estime statistiquement les probabilités de tomber 0, 1, 2, etc. fois en panne dans une journée. On ne sait pas trop ce qu'est* Ω *mais on peut parler de la loi de* X*.*
En fait, une variable aléatoire « transporte » les calculs de probabilités d'un univers inconnu vers des valeurs réelles connues.
La v.a.r. a pour effet de remplacer la tribu $\mathcal{T}$ *par une tribu* $X^{-1}(\mathbb{P}(\mathcal{V}_X))$ *en général plus petite : elle condense l'information sur l'univers en une information plus grossière (numérique) mais c'est déjà bien car en général on ne connaît de toute manière pas ou peu l'univers de départ!...*

On retiendra donc que le « praticien » travaille la plupart du temps avec des lois de probabilités et non pas des ensembles : cela simplifie son travail car il peut travailler sur des espaces complexes (l'univers de départ est souvent trop complexe voire inconnu) sans les connaître.

L'information donnée par la loi est souvent plus grossière : par exemple, on lance deux dés et on regarde la somme des deux faces. La tribu de départ contient 2^{36} évènements alors que la tribu de la loi n'en contient que 2^{11} (il y a 11 sommes possibles).



À partir de la loi, on induit un système complet d'évènements sur Ω : la connaissance de la loi nous donne des renseignements sur l'univers mais souvent de manière grossière (par « paquets »).

Attention!
Deux v.a.r. peuvent avoir la même loi sans être égales! On lance n fois une pièce non truquée. Considérez X le nombre de pile tombés et $n - X$ le nombre de face. Ces deux v.a.r. suivent la même loi binomiale de mêmes paramètres n et 1/2 et pourtant elles ne sont pas égales (ni même égales presque partout ce qui signifie sauf sur un évènement de probabilité nulle mais ceci est une autre histoire).

1.10.3 Variables aléatoires indépendantes

La notation $\mathbb{P}(X = x_i, Y = y_j)$ signifie usuellement $\mathbb{P}(\{X = x_i\} \cap \{Y = y_j\}) = \mathbb{P}(X = x_i \text{ et } Y = y_j)$.
On peut donc définir un **vecteur aléatoire** $Z = (X, Y)$ tel que

$$\mathbb{P}(Z = (x_i, y_j)) = \mathbb{P}(X = x_i, Y = y_j)$$

On peut se représenter l'espace produit à l'aide d'une représentation cartésienne : on parle alors d'évènements « parallèles » à des « axes » (les Ω_i), ces « axes » étant « orthogonaux ». Un système complet d'évènements est alors une « base ». Tous ces ponts de vocabulaire reliant plusieurs théories est un plaisir évident pour le mathématicien mais vous pouvez laisser cela de côté dans un premier temps...

Définition (Couple de variables indépendantes)
Les v.a.r. X et Y sont dites indépendantes si, et seulement si, pour toute partie $(A,B) \in \mathcal{V}_X \times \mathcal{V}_Y$,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B)$$

Définition (Variables aléatoires mutuellement indépendantes)
Si $X_1, X_2, X_3, \cdots, X_k$ *sont* k *v.a. définies sur* Ω*, on dit qu'elles sont mutuellement indépendantes si, et seulement si, quels que soit* $(I_1, I_2, \cdots, I_k) \in \mathcal{V}_{X_1} \times \mathcal{V}_{X_2} \times \cdots \times \mathcal{V}_{X_k}$ *on a :*

$$\mathbb{P}(X_1 \in I_1, X_2 \in I_2, \ldots, X_k \in I_k) = \mathbb{P}(X_1 \in I_1) \times \mathbb{P}(X_2 \in I_2) \times \cdots \times \mathbb{P}(X_k \in I_k)$$

Attention! L'indépendance deux à deux de variables aléatoires n'entraine pas leur indépendance mutuelle.
ON ENLÈVE LE PARAGRAPHE SUR LES FONCTIONS DE RÉPARTITION

1.10.4 Espérance mathématique

Définition (Espérance mathématique)
On appelle espérance mathématique de la variable aléatoire discrète X *le nombre noté* $\mathbb{E}(X)$ *défini par*

$$\mathbb{E}(X) = x_1 \cdot \mathbb{P}_X(x_1) + x_2 \cdot \mathbb{P}_X(x_2) + \cdots + x_n \cdot \mathbb{P}_X(x_n) = \sum_{i=1}^{n} x_i \cdot \mathbb{P}_X(x_i)$$

Exemple : on lance un dé honnête. On définit la variable aléatoire X qui prend la valeur 2 si le numéro du dé est pair et 1 sinon.
Notons ω_i l'événement « le numéro de la face est i », alors $\Omega = \{\omega_1, \omega_2, \omega_3, \emptyset_4, \emptyset_5, \omega_6\}$ et

i	1	2	3	4	5	6
$\mathbb{P}(\omega_i)$	1/6	1/6	1/6	1/6	1/6	1/6
$X(\omega_i)$	1	2	1	2	1	2

D'après le théorème précédent, on a $\mathbb{E}(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 1 \times \frac{1}{6} + 2 \times \frac{1}{6} = \frac{3}{2}$
Nous aurions pu procéder autrement pour définir la loi de probabilité :

Valeurs x_i prises par X	1	2
$\mathbb{P}_X(x_i)$	1/2	1/2

alors, d'après la définition, $\mathbb{E}(X) = 1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}$

Théorème (Espérance de $f \circ X$)
Soit X une v.a.r. Soit $f : \mathbb{R} \to \mathbb{R}$ *telle que* $f \circ X$ *soit une v.a.r. Alors*

$$\mathbb{E}(f \circ X) = \sum_{n \in \mathbb{N}} f(x_n) \cdot \mathbb{P}(X = x_n) = \sum_{n \in \mathbb{N}} f(x_n) \cdot \mathbb{P}_X(x_n)$$

1.10.5 Variance

La variance mesure l'« écart » par rapport à l'espérance en faisant la somme des carrés des « distances »entre chaque valeur prise par la variable aléatoire et l'espérance pondérés par la probabilité correspondante, ce qui donne :

Définition (Variance)
$\mathbb{V}(\mathrm{X}) = \sum_{i=1}^{n} \left(x_i - \mathbb{E}(\mathrm{X})\right)^2 \cdot \mathbb{P}(\mathrm{X} = x_i) = \sum_{i=1}^{n} \left(x_i - \mathbb{E}(\mathrm{X})\right)^2 \cdot \mathbb{P}_{\mathrm{X}}(x_i)$

On a choisi d'utiliser les carrés de manière arbitraire pour ne pas avoir de problèmes de signes; on aurait pu choisir une autre méthode mais celle-ci a l'avantage de rappeler la distance euclidienne bien connue. La variance est en ce sens homogène au carré d'une distance. On va donc définir une distance proprement dite en en prenant la racine carrée : c'est ce qu'on appelle l'**écart-type**.

Définition (Écart-type)
$\sigma(\mathrm{X}) = \sqrt{\mathbb{V}(\mathrm{X})}$

Vous pouvez obtenir espérance, variance et écart-type très simplement à l'aide des modules statistiques de vos calculatrices. Il suffit de rentrer les valeurs prises par la variable aléatoire en liste 1, les probabilités correspondantes en liste 2.

1.10.6 Linéarité de l'espérance

À partir de variables aléatoires existantes, on peut en créer de nouvelles.
Avec des notations usuelles on obtient

- $a\mathrm{X} + b : \omega_i \mapsto a\mathrm{X}(\omega_i) + b$ avec a et b des réels.
- $\mathrm{X} + \mathrm{Y} : \omega_i \mapsto \mathrm{X}(\omega_i) + \mathrm{Y}(\omega_i)$

On peut alors démontrer les propriétés suivantes :

Théorème (Linéarité de l'espérance)

- $\mathbb{E}(a\mathrm{X} + b) = \sum_{i=1}^{n} (ax_i + b) \cdot \mathbb{P}_{\mathrm{X}}(x_i) = a\mathbb{E}(\mathrm{X}) + b$
- $\mathbb{E}(\mathrm{X} + \mathrm{Y}) = \mathbb{E}(\mathrm{X}) + \mathbb{E}(\mathrm{Y})$

Les démonstrations sont similaires. Notons $\mathrm{Y} = a\mathrm{X} + b$.
Pour tout $\omega_i \in \Omega$ on a $\mathrm{Y}(\omega_i) = a\mathrm{X}(\omega_i) + b$ et donc $y_i = ax_i + b$ avec les notations usuelles donc $\left\{\omega \in \Omega | \mathrm{Y}(\omega) = y_i\right\} = \left\{\omega \in \Omega | a\mathrm{X}(\omega) + b = ax_i + b\right\} = \left\{\omega \in \Omega | \mathrm{X}(\omega) = x_i\right\}$
Alors

$$\mathbb{E}(\mathrm{Y}) = \sum_{i=1}^{n} y_i \mathbb{P}(\mathrm{Y} = y_i) = \sum_{i=1}^{n} (ax_i + b)\mathbb{P}(\mathrm{X} = x_i) = a\sum_{i=1}^{n} x_i \mathbb{P}(\mathrm{X} = x_i) + b\sum_{i=1}^{n} \mathbb{P}(\mathrm{X} = x_i)$$

On en déduit que $\mathbb{E}(\mathrm{Y}) = a\mathbb{E}(\mathrm{X}) + b$.

1.10.7 Théorème de König-Huygens

Johann KÖNIG (1712-1757) fut un mathématicien allemand, élève de Jean BERNOULLI et Christian HUYGENS (1629-1695) était lui néerlandais.

Théorème (Théorème de König-Huygens)

$$\mathbb{V}(\mathrm{X}) = \mathbb{E}\left((\mathrm{X} - \mathbb{E}(\mathrm{X}))^2\right) = \mathbb{E}\left(\mathrm{X}^2\right) - \left(\mathbb{E}(\mathrm{X})\right)^2$$

À démontrer en exercice...

1.11 Quelques lois discrètes classiques

1.11.1 Loi uniforme (discrète)

C'est la loi qui correspond au tirage équiprobable d'un jeton blanc dont on note le numéro dans une urne qui contient n jetons blancs numérotés.

Définition (Loi uniforme)
On dit que $X : \Omega \longrightarrow [\![1, n]\!]$ *suit une* ***loi uniforme*** *sur* $[\![1, n]\!]$ *et on note* $X \rightsquigarrow \mathcal{U}([\![1, n]\!])$ *si :*

$$X(\Omega) = [\![1, n]\!] \quad et \quad \forall k \in [\![1, n]\!],\ \mathbb{P}(X = k) = \frac{1}{n}$$

Calculons l'espérance et la variance de $X \rightsquigarrow \mathcal{U}([\![1, n]\!])$.
$\mathbb{E}(X) = \sum_{k=1}^{n} k\mathbb{P}(X = k) = \sum_{k=1}^{n} k\frac{1}{n} = \frac{1}{n}\sum_{k=1}^{n} k = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}$.
De manière analogue, on obtient que $\mathbb{V}(X) = \frac{n^2-1}{12}$ sachant que $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

1.11.2 Loi de Bernoulli : to be or not to be

On place des jetons blancs et des noirs dans une urne. On suppose que toutes les conditions sont réunies pour effectuer le tirage d'un jeton dans le cadre d'équiprobabilité et que la proportion de jetons blancs est p. Que dire de la variable aléatoire qui vaut 1 si le jeton tiré est blanc et 0 s'il est noir ?

Définition (Loi de Bernoulli)
On dit que la v.a.r. finie X suit la ***loi de Bernoulli de paramètre p*** *si :*
- $p \in]0;1[$, $X(\Omega) = \{0, 1\}$;
- $\mathbb{P}(X = 1) = p$.

On note alors $X \rightsquigarrow \mathcal{B}(1, p)$

On dit aussi que X est une **variable aléatoire de Bernoulli**. Sa loi de probabilité est simple à décrire :

x_i	0	1
$\mathbb{P}(X = x_i)$	1-p	p

Il est simple de déterminer son espérance et sa variance :
- $\mathbb{E}(X) = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) = p$;
- $\mathbb{V}(X) = (0 - p)^2 \times \mathbb{P}(X = 0) + (1 - p)^2 \times \mathbb{P}(X = 1) = p^2(1 - p) + (1 - p)^2 p = p(1 - p)$.

1.11.3 Loi binomiale

Dans les mêmes conditions que précédemment, on effectue n tirages d'un jeton avec remise. À chaque tirage on associe une v.a.r. indépendante des autres qui suit la loi de Bernoulli de paramètre p. Les issues élémentaires ω de ces n tirages sont des « mots » de n lettres, chaque lettre étant un B (pour blanc) ou un $\overline{B}$. On définit alors la variable aléatoire X à valeurs dans $\{0, 1, 2, ..., n\}$, donnant le nombre de jetons blancs de ces issues élémentaires, c'est à dire que X est la somme des v.a.r. finies élémentaires suivant la loi de Bernoulli de manière indépendante.

Définition (Loi binomiale)
On dit que la v.a.r. finie X suit la ***loi binomiale de paramètres n et p*** *si* $n \in \mathbb{N} \setminus \{0\}$, $p \in]0\,;1[$, $\mathrm{X}(\Omega) = [\![0\,, n]\!]$ *et si l'expérience est une suite de n expériences de Bernoulli de paramètre p menées indépendamment.*
On note alors $\mathrm{X} \rightsquigarrow \mathscr{B}(n, p)$

On s'intéresse maintenant à l'événement $(\mathrm{X} = k)$, c'est à dire l'ensemble des mots de n lettres écrits avec k B et $n-k$ $\overline{\mathrm{B}}$. Notons B_j l'évènement « obtenir un jeton blanc au $\mathrm{j}^{\text{ème}}$ tirage ». Chaque tirage d'une boule étant indépendant des autres, on a

$$\mathbb{P}(\omega) = \mathbb{P}(\mathrm{B}_1) \times \cdots \times \mathbb{P}(\mathrm{B}_k) \times \mathbb{P}(\overline{\mathrm{B}_{k+1}}) \times \cdots \times \mathbb{P}(\overline{\mathrm{B}_n}) = p^k (1-p)^{n-k}$$

Il reste à déterminer combien il y a de tels mots. Ce sont les anagrammes (permutations distinctes) du mot de n lettres BB...B$\overline{\mathrm{B}}\overline{\mathrm{B}}$...$\overline{\mathrm{B}}$. Il y en a $n!$ qu'il faut diviser par les $k!$ permutations des B et les $(n-k)!$ permutations des $\overline{\mathrm{B}}$, c'est-à-dire C_n^k. Finalement, on obtient :

Théorème ()
Pour toute v.a.r. finie $\mathrm{X} \rightsquigarrow \mathscr{B}(n, p)$, *on a, pour tout* $k \in \mathrm{X}(\Omega)$,

$$\mathbb{P}(\mathrm{X} = k) = \mathrm{C}_n^k p^k (1-p)^{n-k}$$

Vous démontrerez que $\mathbb{E}(\mathrm{X}) = np$ et $\mathbb{V}(\mathrm{X}) = np(1-p)$.

1.11.4 Loi hypergéométrique

Modèle : on considére une urne contenant N boules ($\mathrm{N} \in \mathbb{N}^\star$), la proportion de boules blanches dans l'urne étant p. On tire simultanément n boules (les boules sont choisies au hasard) et X est la v.a. qui indique le nombre de boules blanches tirées. On dit que X suit la loi hypergéométrique de paramètres N, n et p notée $\mathscr{H}(\mathrm{N}, n, p)$.
On retiendra :
- X peut prendre toute valeur entière de l'intervalle $\{\max\{0, n - \mathrm{N}(1-p)\}, \min\{n, \mathrm{N}p\}\}$
- $k \in \mathscr{V}_{\mathrm{X}}$, $\mathbb{P}(\{\mathrm{X} = k\}) = \dfrac{\mathrm{C}_{\mathrm{N}p}^k \mathrm{C}_{\mathrm{N}(1-p)}^{n-k}}{\mathrm{C}_{\mathrm{N}}^n}$
- $\mathbb{E}(\mathrm{X}) = np$, $\mathbb{V}(\mathrm{X}) = np(1-p)\dfrac{\mathrm{N}-n}{\mathrm{N}-1}$

1.11.5 Loi géométrique

Modèle : on répète indéfiniment une épreuve de Bernoulli dont la probabilité de succès est p.
Les répétitions étant mutuellement indépendantes, X désigne le rang de l'essai donnant pour la première fois un succès. On dit que X suit la loi géométrique de paramètre p et on écrit $\mathrm{X} \rightsquigarrow \mathscr{G}(p)$. On retiendra :

$$\begin{aligned}
&\mathscr{V}_{\mathrm{X}} = \mathbb{N}^\star \\
&k \in \mathbb{N}^\star, \mathbb{P}(\{\mathrm{X} = k\}) = p(1-p)^{k-1} \\
&\mathbb{E}(\mathrm{X}) = \tfrac{1}{p}, \mathbb{V}(\mathrm{X}) = \frac{1-p}{p^2}
\end{aligned}$$

La loi géométrique modélise donc le **temps d'attente jusqu'au premier succès**.

Problème du collectionneur

Voici une application classique de la loi géométrique : *en Syldavie,le petit Ivan collectionne les images du chocolat Wonka. Il y a n images différentes. Lorsqu'il aura obtenu les n images, il aura droit de visiter les usines Wonka. À chaque fois qu'il achète une tablette, la probabilité d'obtenir une image quelconque est* $1/n$. *Soit* T_n *le temps d'attente (en nombre de tablettes achetées) pour obtenir les n images. Calculez* $\mathbb{E}(\mathrm{T}_n)$.

Soit X_i la variable aléatoire correspondant au temps d'attente pour passer de $i-1$ images différentes à i images différentes. Quand on est dans l'état $i-1$, on dispose de $i-1$ images différentes. Si on tombe sur l'une des $i-1$ images dont on dispose (avec la probabilité $\frac{i-1}{n}$), on reste dans l'état $i-1$, sinon, on passe à l'état i (avec la probabilité $\frac{n-(i-1)}{n}$).
Ainsi, $\mathrm{X}_i \rightsquigarrow \mathcal{G}(p_i)$ avec $p_i = \frac{n-(i-1)}{n}$. On en déduit que $\mathbb{E}(\mathrm{X}_i) = \frac{1}{p_i} = \frac{n}{n+1-i}$.
Or, $\mathrm{T}_n = \mathrm{X}_1 + \mathrm{X}_2 + \cdots + \mathrm{X}_n$, donc $\mathbb{E}(\mathrm{T}_n) = \mathbb{E}(\mathrm{X}_1) + \mathbb{E}(\mathrm{X}_2) + \cdots + \mathbb{E}(\mathrm{X}_n)$:

$$\mathbb{E}(\mathrm{X}_n) = \sum_{i=1}^{n} \frac{1}{p_i} = n \sum_{i=1}^{n} \frac{1}{n+1-i} = n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{2} + 1\right)$$

Par exemple, avec $n = 10$:
Il faut en moyenne acheter 30 tablettes pour obtenir les dix images.

1.12 EXERCICES

Dénombrements

Test 1.1 (Octets - ♡)
Combien y a-t-il d'octets commençant par 1 ou se terminant par 00 ?

Test 1.2 (Téléphone - ♡)

1. On suppose que les numéros de téléphone sont de la forme 0Z-YX-XX-XX-XX avec X un nombre entier entre 0 et 9, Y un nombre entier entre 2 et 9 et Z un entier entre 1 et 9.

   En France (données du quatrième trimestre 2011), il y a environ 40 millions d'abonnements à un « fixe » (Z étant compris entre 1 et 5) et 69 millions à un « mobile » (Z valant 6 ou 7).

   Est-on tranquille ? Pourquoi existe-t-il des numéros de mobile commençant par 07 ?

2. Supposons qu'en Syldavie, il y ait 109 millions d'abonnés à un téléphone fixe et que le Y soit fixé par opérateur et qu'il n'y ait que deux opérateurs. De combien de zones a-t-on besoin ?

Test 1.3 (Nombre d'applications - ♡♡)
Combien y a-t-il d'applications d'un ensemble E de cardinal n vers un ensemble F de cardinal p ?

Test 1.4 (Barbie - ♡♡)
Un étudiant de Math Sup du lycée Ottokar de Klow en Syldavie a reçu pour Noël sept Barbie et sept séries de dix robes de couleurs différentes. Il habille chaque Barbie d'une de ces robes.

1. Combien y a-t-il d'habillages possibles selon la couleur ?
2. Combien y a-t-il d'habillages de sorte que chaque Barbie ait une robe de couleur différente des autres ?
3. Combien y a-t-il d'habillages de sorte qu'au moins deux Barbie aient une robe de la même couleur ?
4. Combien y a-t-il d'habillages de sorte qu'exactement deux Barbie aient une robe de la même couleur ?

Test 1.5 (Jus - ♡♡)
Lors de la soirée d'intégration des Math Sup du lycée Ottokar de Klow, 9 étudiants ont bu un verre d'eau, 25 un verre de jus de tomate, 14 un verre de jus de betterave, 7 un verre de jus de tomate et un verre d'eau, 4 un verre d'eau et un verre de jus de betterave, 10 un verre de jus de betterave et un verre de jus de tomate. Les étudiants de Math Sup s'occupaient des boissons et ceux de BCPST des entrées or ces derniers ont perdu leur liste et ne savent plus combien d'étudiants ont participé à la soirée. Les étudiants de Math Sup, connaissant les statistiques sur les consommations de boissons, pourront-ils corriger la gaffe des BCPST sachant que tout le monde a bu et que lors de la valse, tous les participants dansaient en couple ?

Test 1.6 (*k*-listes - ♡♡)
Soit E un ensemble fini non vide de cardinal n.

a) Combien y-a-t-il de parties de E formées de k éléments ?
b) Combien y-a-t-il de k-uplets d'éléments de E ?
c) Combien y-a-t-il de k-uplets d'éléments deux à deux distincts de E ?

d) Combien y-a-t-il de k-uplets d'éléments deux à deux distincts de E, tel que le premier élément est le plus petit et le dernier élément est le plus grand ?
e) Combien y-a-t-il de k-uplets d'éléments de E ordonnés dans l'ordre strictement croissant ?

Test 1.7 (Poker - ♡♡)
Une main au poker est constituée de 5 cartes tirées d'un jeu de 32 cartes. Combien y a-t-il de mains contenant des carrés (XXXXY) ? des fulls (XXXYY) ? des brelans (XXXYZ) ? des doubles paires (XXYYZ) ? des paires (XXYZA) ?
Deux lettres identiques (par exemple XX) correspondent à deux cartes de même hauteur (par exemple deux dames).

Test 1.8 (Mélange de cartes - ♡♡♡)
On dispose d'un jeu de 52 cartes et on le mélange en *queue d'aronde* : on divise le jeu en deux parties de 26 cartes et on intercale les cartes de la moitié gauche dans la moitié droite sans modifier l'ordre de chaque partie bien sûr. Est-ce que quatre mélanges en queue d'aronde successifs permettent de fournir un ordre aléatoire des cartes ?
Il faudrait préciser d'abord cette notion : combien y a-t-il d'ordres possibles des cartes ? Existe-t-il alors des ordres qui ne pourront être obtenus par notre succession de mélanges ? Pour répondre à cette question, il faudrait calculer le nombre d'ordres qu'on peut obtenir par la succession de quatre mélanges...Imaginez que vous disposiez de 26 cartes bleues et 26 rouges. Vous les mélangez avec la méthode de la queue d'aronde : pouvez-vous reconstituer le jeu initial ? Combien y a-t-il alors d'ordres possibles après le premier mélange ?

Test 1.9 (Suissesses multinomiales - ♡♡♡)
Combien existe-il d'anagrammes au mot « suissesses » ?
Pour répondre à cette question on définit un nouveau type de coefficient :

Définition (Coefficient multinomial)
Soit k_1, k_2,...,k_p des entiers naturels tels que $\sum_{i=1}^{p} k_i = n$. On note alors :

$$\mathrm{C}_n^{k_1,k_2,\ldots,k_p} = \frac{n!}{k_1!k_2!\cdots k_p!}$$

1. Quel est le coefficient multinomial égal à $\binom{n}{p}$?
2. Comment interpréter ces coefficients en termes combinatoires ? Quel est le rapport avec les suissesses ?
3. Il existe un pendant multinomial à la formule du binôme :

   Théorème (Formule du multinôme)

   $$(x_1 + x_2 + \cdots + x_p)^n = \sum_{k_1+k_2+\cdots+k_p=n} \mathrm{C}_n^{k_1,k_2,\ldots,k_p} x_1^{k_1} x_2^{k_2} \cdots x_p^{k_p}$$

   Par exemple, quel est le coefficient de x^2y^5z dans $(x+y+z)^8$?

Test 1.10 (L'âge du capitaine - ♡♡)
Le capitaine des pompiers de New-York est marié et a quatre enfants dont un mange un yaourt aux fruits tous les matins. Il ne fume pas, aime regarder des films de gladiateurs et réside à l'angle de la 1ère avenue et de la 33ème rue. La caserne se trouve à l'angle de la 9ème avenue et de la 40ème rue. Il s'y rend tous les jours à pied en sifflant « O du schöner Westerwald » et sans perdre de temps (i.e. dans le sens des numéros croissants aussi bien pour les rues

que pour les avenues). Sachant qu'il a commencé à travailler le jour de ses 18 ans, et sachant qu'il n'est jamais passé deux fois par le même chemin à l'aller, qu'il boîte légèrement de la jambe droite après avoir participé au championnat d'Écosse de lancement d'enclume, qu'il est sourd de l'oreille gauche depuis qu'il a plongé dans un lac gelé pour sauver son petit frère John qui était en train de se noyer après que la glace sur laquelle il patinait craqua, quel est l'âge (maximum) du capitaine ?

Probabilités

Test 1.11 (Formule du crible de Poincaré : cas de trois évènements - ♡)
Déterminez $\mathbb{P}(A \cup B \cup C)$ en fonction des probabilités de A, B, C, $A \cap B$, $A \cap C$, $B \cap C$ et $A \cap B \cap C$.

Test 1.12 (Définition d'une probabilité - ♡♡♡)
Soit $\Omega = \{\omega_1, \omega_2, \ldots\}$ un ensemble infini dénombrable, les ω_i étant distincts deux à deux. On suppose que, pour tout entier naturel non nul k, $\mathbb{P}\left(\{\omega_k\}\right) = \frac{\lambda}{37^k}$ avec λ un réel.
Existe-il une valeur de λ telle que $(\Omega, \mathbb{P}(\Omega), \mathbb{P})$ soit un espace probabilisé ?

Test 1.13 (Problèmes historiques - ♡♡)
Tout a commencé au XVII^{e} siècle, lorsque les Princes demandèrent aux mathématiciens de les aider à gagner au jeu...

1. Le Duc de Toscane demande à GALILÉE pourquoi, en lançant trois dés, on obtient plus souvent un total de 10 qu'un total de 9, alors qu'il y a dans les deux cas exactement 6 façons d'obtenir ces résultats.
2. Le Chevalier de MÉRÉ soutient à PASCAL que les deux jeux suivants sont favorables au joueur : obtenir au moins un 6 en lançant 4 fois de suite un dé, et obtenir au moins un double 6 en lançant 24 fois de suite 2 dés.

Vous simulerez ces expériences à l'aide de Python et vous démontrerez les résultats observés.

Test 1.14 (Formule de B. et Monica B. - ♡)
En Syldavie, 70% des habitants sont de sexe masculin et 60% des hommes et 25% des femmes ont une photo de Monica B. dans leur portefeuille. Un extra-terrestre enlève au hasard une personne syldave et découvre une photo de Monica B. dans son portefeuille : quelle est la probabilité que la personne enlevée soit une femme ?

Test 1.15 (Marylin et les chèvres - ♡♡♡)
Dans le numéro du 9 septembre 1990 de *Parade*, Marilyn VOS SAVANT connue pour avoir un QI de 186, soit l'un des plus élevés au monde, répondit au courrier suivant : *Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, « Do you want to pick door number 2 ? » Is it to your advantage to switch your choice of doors ?*

Craig. F. Whitaker Columbia, MD

Cette lettre décrit en fait le déroulement d'un jeu télévisé des années 1970, *Let's make a deal*, animé par Carol MERRIL et **Monty HALL** : c'est en référence à ce dernier qu'on donna son nom à ce problème célèbre (plus tard, un haut représentant de la culture française reprit ce jeu : ce fut Lagaf et son *Bigdil*...).
Marilyn répondit qu'il valait mieux pour le candidat changer de porte quand le présentateur l'y invitait. Qu'en pensez-vous ?
Vous pourrez également proposer une simulation avec **Python** .

Test 1.16 (Duel - ♡♡♡)
Alexandre P. et Évariste G. s'affrontent en duel au pistolet suivant les règles suivantes :

— Ils tirent chacun leur tour. Le premier qui atteint l'autre a gagné...
— Lorsqu'il tire, Alexandre atteint Évariste avec la probabilité a $(0 < a < 1)$ et il le rate avec la probabilité $\overline{a} = 1 - a$.
— Lorsqu'il tire, Évariste atteint Alexandre avec la probabilité b $(0 < b < 1)$ et il le rate avec la probabilité $\overline{b} = 1 - b$.
— Alexandre tire le premier.*Ainsi, Alexandre (resp. Évariste) n'effectue que des tirs de rang impair (resp. pair).*

On considère, pour tout entier $n \geqslant 1$, les évènements A_{2n-1} : « Alexandre gagne à l'issue du tir numéro $2n-1$ », B_{2n} : « Évariste gagne à l'issue du tir numéro $2n$ ».

1. Calculez, en fonction de a et b, les probabilités des évènements A_1, B_2 et A_3. Plus généralement, calculez $\mathbb{P}(A_{2n-1})$ et $\mathbb{P}(B_{2n})$.
2. Pour $n \geqslant 1$, on note C_n (resp. D_n) l'évènement : « Alexandre (resp. Évariste) gagne à un tir dont le numéro est entre 1 et $2n-1$ (resp. entre 2 et $2n$) ». Calculer $\mathbb{P}(C_n)$ et $\mathbb{P}(D_n)$.
3. Calculer $\alpha = \lim\limits_{n\to+\infty} \mathbb{P}(C_n)$ et $\beta = \lim\limits_{n\to+\infty} \mathbb{P}(D_n)$. Vérifiez que $\alpha + \beta = 1$. Comment interpréter ce résultat ?

Test 1.17 (To Binomiale or not to binomiale ? - ♡)
Dites si les variables aléatoires X définies ci-dessous suivent une loi binomiale ? Donnez quand c'est possible $X(\Omega)$, les paramètres de la loi ainsi que l'espérance.

1. Dans une classe on tire au sort et sans remise 5 élèves. X est le nombre d'élèves travaillant pour les services secrets bordures dans le lot tiré au sort.
2. Dans un sac de 20 billes contenant 7 noires et 13 blanches, on tire avec remise 3 d'entre elles, X étant le nombre de billes noires obtenues.
3. On lance 10 dés, X est le nombre de « 5 » obtenus.
4. Un circuit comprend 32 lampes en série. Pour chacune d'elle, la probabilité qu'elle fonctionne est de $3/100$. X est le nombre de lampes qui s'allument lorsqu'on appuie sur l'interrupteur.
5. Même question avec cette fois des lampes en parallèle.

Test 1.18 (Lois usuelles - ♡)
Dans chacune des expériences aléatoires ci-dessous, reconnaître une loi usuelle pour la variable aléatoire X. Donner alors son espérance.

1. On choisit une carte au hasard dans un jeu de 32 cartes numérotées. X = « hauteur de la carte choisie »
2. Un dieu de l'Olympe tue chaque jour un Achéen (avec probabilité $2/3$) ou un Troyen (avec probabilité $1/3$). On suppose que le choix la victime est indépendant de celle des autres. X = « nombre de Troyens tués sur 10 jours consécutifs ».
3. Sur la plaine d'Ilium se trouvent 15 Achéens et 30 Troyens. Un dieu de l'Olympe foudroie chaque matin un Achéen ou un Troyen, pris parmi les soldats de la plaine. X = « nombre de Troyens tués sur 10 jours consécutifs ».
4. On tire une à une les cartes d'un jeu de 32 cartes, jusqu'à obtenir le valet de pique. X = « nombre de tirage effectuées ».
5. On lance un dé à 6 faces. X = « chiffre obtenu ».

6. Monica a dans sa poche 7 allumettes et 12 pièces de monnaie. Elle met la main dans sa poche et en sort 5 objets. X = « nombre de pièces obtenues ».
7. Monica fait passer une audition pour trouver des partenanires pour son prochain film. 3000 candidats passent l'audition. Chaque candidat réussit son entretien avec probabilité 2/3. X = « nombre de candidats sélectionnés par Monica ».
8. James dort chez son amie Monica qui lui a prêté son appartement. Il dispose d'un trousseau de 7 clés, et ne sait pas laquelle ouvre la porte de l'appartement. Il les essaye une par une, en mettant de côté chaque clé essayée. X = « nombre d'essais nécessaires pour trouver la bonne clé ».

Test 1.19 (Loi hypergéométrique - ♡)

NOUVEL EXERCICE

Test 1.20 (HEC 2016 - ♡)

On tire avec remise une boule d'une urne contenant des boules numérotées de 1 à n. Quelle est la variable aléatoire V_n dont la fonction Scilab suivante calcule la simulation ?

```
function compt = V(n)
    A = []; compt = 0;
    while length(A) < n
        u = floor(n*rand() + 1);
        i = find(A == u); // renvoie les positions de u dans le vecteur A
        if length(i) == 0 then
            A = [A, u];
        end;
        compt = compt + 1;
    end;
endfunction
```

NOUVEL EXERCICE

Test 1.21 (CENTRALE 2017 - ♡)

Un pion se trouve à l'instant 0 sur la case 0 d'un parcours linéaire dont les cases sont numérotées par les entiers consécutifs. À chaque étape il avance d'un nombre strictement positif aléatoire de cases. Pour $n \in \mathbb{N}^*$ on note Y_n le nombre de cases dont il avance à la n-ième étape. Ainsi $S_n = Y_1 + Y_2 + \cdots + Y_n$ est sa position à l'instant n avec $S_0 = 0$ par convention. On suppose que les variables $(Y_n)_{n\in\mathbb{N}^*}$ suivent toutes la même loi de probabilité et sont indépendantes.

On suppose que $Y_1 - 1$ suit une loi de Bernoulli de paramètre p et on choisit un entier $k \in \mathbb{N}^*$.

1. Écrire une fonction en **Python** prenant en argument les paramètres p et k et simulant l'expérience jusqu'à ce que le pion dépasse (au sens large) la position k. La fonction renverra 1 si le point a atterri sur la case k et 0 s'il l'a dépassée sans s'y arrêter.
   Pour simuler une loi de Bernoulli, on pourra utiliser l'aide documentaire fournie dont voici un extrait :
   Les fonctions d'échantillonnage et de génération de valeurs pseudo-aléatoires sont regroupées dans la bibliothèque `numpy.random+`.

```python
import numpy.random as rd
```

La fonction binomial+ *permet de simuler une variable aléatoire suivant une loi binomiale de paramètres n et p. Elle permet donc également de simuler une variable aléatoire suivant une loi de Bernoulli de paramètres p en prenant simplement n = 1. Cette fonction prend un troisième paramètre optionnel qui correspond, comme pour les fonctions précédentes, au nombre de valeurs à obtenir.*

```python
In [1]: rd.binomial(10, 0.3, 7)
Out[1]: array([4, 3, 4, 1, 4, 3, 3])

In [2]: rd.binomial(1, 0.6, 20)
Out[2]: array([0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1,
    ↪  1])
```

2. Pour un entier k assez grand et des valeurs de p de votre choix, calculez sur une centaine d'essais la proportion de tentatives pour lesquelles le pion atteint la position k exactement. Comparez avec $1/\mathbb{E}(Y_1)$.

Test 1.22 (Roulette - ♡♡)

Une roulette contient 36 cases numérotées de 1 à 36 dont 18 sont rouges et 18 sont noires, plus une case numérotée 0 de couleur verte.
Un joueur qui mise sur la couleur rouge ou noire gagne deux fois sa mise si la couleur choisie sort.
Un joueur qui mise sur un numéro de 1 à 36 gagne 36 fois sa mise si le numéro sort.
Il est interdit de miser sur le zéro.

1. Un joueur mise a € sur une couleur. Soit C la variable aléatoire correspondant au gain associé. Trouvez la loi de C puis calculez $\mathbb{E}(C)$ et $\sigma(C)$.
2. Un joueur mise a € sur un numéro. Soit N la variable aléatoire correspondant au gain associé. Trouvez la loi de N puis calculez $\mathbb{E}(N)$ et $\sigma(N)$.
3. Vaut-il mieux miser sur une couleur ou un numéro ?

Test 1.23 (« C'est sans danger... » - ♡♡)

Afin d'effectuer des économies dans le budget de l'État, le premier ministre syldave a supprimé la faculté de médecine. Dans le but de promouvoir les liens sociaux, les dentistes sont maintenant choisis parmi les anciens boxeurs aveugles et parkinsoniens. Ils arrachent les dents de leurs patients au hasard.
Suite à quelques plaintes, une étude statistique a été menée dans le cabinet d'un de ces nouveaux dentistes. Les syldaves venant le consulter ont toujours une seule dent de malade parmi les trente-deux qu'ils possèdent encore avant l'intervention des tenailles ou des poings, c'est selon (une étude précédente a en effet montré que les syldaves ne retournent jamais chez leur dentiste après une première consultation).
L'étude a été effectuée sur n patients. On note X la variable aléatoire égale au nombre de dents malades extraites à bon escient.

1. Déterminez la loi de probabilité de la variable aléatoire X. Calculez la probabilité pour qu'aucune dent malade n'ait été extraite.

2. Combien de patients doit-il traiter pour extraire au moins une dent malade avec une probabilité supérieure à 0,9?
3. Le dernier patient est prêt à se sacrifier pour la gloire des statistiques syldaves et se laisse arracher les dents une à une tant que la dent malade n'a pas été extraite. On note Y le nombre de dents saines que ce vaillant patriote voit tomber de ses mâchoires endolories.

   Calculez la probabilité pour qu'il reparte complètement édenté, puis $\mathbb{E}(Y)$ et $\sigma(Y)$.

Test 1.24 (Dé truqué - ♡♡)
On considère 6 dés, cinq étant équilibrés. Le dernier est pipé de manière à ce que lorsque l'on lance ce dé, chacun des chiffres apparaît avec une probabilité proportionnelle à ce chiffre.

1. Donner la loi, l'espérance et la variance de la variable aléatoire égale au chiffre donné par le dé truqué lorsqu'on le lance.
2. On effectue n tirages successifs et indépendants d'un dé parmi les six.

   Quelle est la loi suivie par la variable aléatoire égale au nombre de fois où on a tiré le dé truqué? Combien de tirages doit-on effectuer pour que la probabilité d'avoir obtenu le dé truqué parmi ceux tirés soit supérieure ou égale à $1/2$?
3. On effectue n tirages ($n \leqslant 6$) successifs sans remise d'un dé parmi les six. Quelle est la loi suivie par la variable aléatoire égale au nombre de fois où est tiré le dé truqué? Combien de tirages doit-on effectuer pour que la probabilité d'avoir obtenu le dé truqué parmi ceux tirés soit supérieure ou égale à $1/2$?

Test 1.25 (File d'attente - ♡♡)
On constitue une file d'attente en attribuant au hasard des numéros d'ordre à n personnes dont vous et Monica B.. Quelle est la probabilité que vous soyez distant de p places de Monica (c'est-à-dire séparés par $p-1$ personnes)? Quelle est la distance la plus probable entre vous et Monica?

Test 1.26 (Théorème de KÖNIG-HUYGENS - ♡♡)
Démontrez ce théorème...

Test 1.27 (Espérance d'une loi géométrique - ♡♡)
Notons S la somme définie par :

$$\forall q \in \mathbb{R},\ 0 \leqslant q < 1 \Longrightarrow S = \sum_{n \geqslant 0} q^n$$

Pour $0 \leqslant q < 1$, notons $T = \sum_{n \geqslant 1} nq^{n-1}$ en supposant que cette écriture a un sens.
Calculer $(1-q)T$ en fonction de S et déduisez-en T puis l'espérance d'une loi géométrique.

1.12.1 Coups de pouce

Coup de pouce pour le test 1.1
Un octet est une 8-liste d'éléments de $\{0,1\}$.

Coup de pouce pour le test 1.2
Utiliser le principe multiplicatif : un numéro de téléphone est une 10-liste.

Coup de pouce pour le test 1.3
Une application est une fonction telle que tout élément de l'ensemble de départ E admet une image et une seule dans F. Les informaticiens parlent eux de fonction totale.

Coup de pouce pour le test 1.4 Modéliser un habillage par une 7-liste.

Coup de pouce pour le test 1.5 On note E, T et B l'ensemble des buveurs d'eau, de jus de tomate et de jus de betterave respectivement.
Utilisez la formule du crible à répétitions avec $E \cup (T \cup B)$ ainsi que la distributivité de l'intersection sur la réunion..

Coup de pouce pour le test 1.6
Pour les dénombrements avec ordre final imposé, commencer par chosir les éléments de façon non ordonnée.

Coup de pouce pour le test 1.7 Choisir une carte c'est choisir sa hauteur et sa couleur : pensez à des combinaisons.

Coup de pouce pour le test 1.8
Le jeu est mélangé une fois. Numérotons les carte de 1 à 52. Pour recréer la situation initiale (rouges d'un côté, bleues de l'autre), il faut donc créer un groupe de 26 cartes parmi les 52.

Coup de pouce pour le test 1.9
Souvenez-vous que $n!$ est le nombre de permutations d'une n-liste.

Coup de pouce pour le test 1.10
On peut modéliser un trajet comme un mot constitué de r (pour rue) et de a (pour avenue) : il s'agit de chercher les anagrammes d'un tel mot...

Coup de pouce pour le test 1.11 La démonstration de la formule associée pour les cardinaux a déjà été faite à l'exercice 1.5 page 27...

Coup de pouce pour le test 1.12
Relisez la partie du cours concernant les axiomes de KOLMOGOROV
Étudier la limite quand n tend vers l'infini de $\sum_{i=1}^{n} q^i$ lorsque $|q| < 1$.

Coup de pouce pour le test 1.13
Il **semble** y avoir 6 façons d'obtenir 9 ou 10. Pour modéliser la situation, numérotez les dés. Un lancer correspond donc à un triplet (ordonné !) de chiffres entre 1 et 6.

Coup de pouce pour le test 1.14
La formule de BAYES :

$$\mathbb{P}(\text{cause} \mid \text{effet}) = \frac{\mathbb{P}(\text{effet} \mid \text{cause}) \times \mathbb{P}(\text{cause})}{\mathbb{P}(\text{effet})}$$

Coup de pouce pour le test 1.15
On peut avoir deux idées contradictoires :
— le présentateur ayant ouvert une porte, le candidat a une chance sur deux de tomber sur la voiture ;
— Au départ, le candidat a une probabilité de $\frac{1}{3}$ de tomber sur la voiture. Ensuite, s'il garde son choix initial, il a toujours la même probabilité $\frac{1}{3}$ de gagner donc s'il change d'avis, la probabilité qu'il gagne est donc $1 - \frac{1}{3} = \frac{2}{3}$.

Qui a raison ? Commencez par dresser un arbre et à considérer que les choix des portes sont équiprobables.

Coup de pouce pour le test 1.16
Un arbre permet de visualiser la situation.
Il faut savoir exprimer la somme des termes d'une suite géométrique.

Coup de pouce pour le test 1.17 Rappel : on dit que la v.a.r. finie X suit la **loi binomiale de paramètres n et p** si $n \in \mathbb{N} \setminus \{0\}$, $p \in]0\,;1[$, $X(\Omega) = [\![0, n]\!]$ et si l'expérience est une suite de n expériences de Bernoulli de paramètre p menées indépendamment.

Coup de pouce pour le test 1.18 Étudier l'indépendance des répitions pour distinguer loi binomiale et loi hypergéométrique.

Coup de pouce pour le test 1.19
Pensez au poker...

Coup de pouce pour le test 1.20 On tire une boule tant que les n numéros n'ont pas été tirés.

Coup de pouce pour le test 1.21

1. On obtient Y_1 avec `1+rd.binomial(1, p, 1)[0]`. Comme `binomial` renvoie un tableau numpy+ le crochet permet d'en extraire le premier et seul élément.
2. 1 pour succès et 0 pour échec : la somme des valeurs obtenues donne donc le nombre de succès.

Coup de pouce pour le test 1.22 La variable aléatoire C ne prend que deux valeurs : a et $-a$. Étudiez sa loi de probabilité.

Coup de pouce pour le test 1.23 Pour chaque patient, l'extraction d'une dent est une épreuve de Bernoulli de paramètre $\frac{1}{32}$.

Coup de pouce pour le test 1.24

1. Noter p la probabilité d'obtenir « 1 » avec le dé pipé.
2. Remarquer que les tirages sont indépendants.
3. Remarquer que les tirages ne sont pas indépendants.

Coup de pouce pour le test 1.25 Commencez par compter le nombre de files telles que vous soyez à la distance p de Monica en ayant un numéro plus petit qu'elle et souvenez-vous des sommes usuelles : somme des n premiers entiers, des n premiers carrés d'entiers,...

Coup de pouce pour le test 1.26
Souvenz-vous de $(a+b)^2 = \ldots$ et de la linéarité de l'espérance.

Coup de pouce pour le test 1.27 Nous avons déjà démontré que :

$$\forall q \in \mathbb{R},\ |q| < 1 \Longrightarrow S = \sum_{n \geqslant 0} q^n = \frac{1}{1-q}$$