Probabilités

Simone Naldi

Université de Limoges
L2 - Semestre 3

Avant-propos

Ce document constitue les notes du cours de Probabilités donné par Simone Naldi à l'Université de Limoges, années académiques 2018/2019, 2019/2020 (en collaboration avec Francisco Silva) et 2020/2021 (en collaboration avec Hamza Ennaji).

Pour toutes erreurs [1] ou imprécisions dans le texte, merci de contacter Simone Naldi par courriel à l'adresse

`simone.naldi@unilim.fr`

La page web Community du cours de Probabilités, où vous pourrez trouver toutes les informations sur le déroulement de l'UE, est disponible au lien :

`https://community-sciences.unilim.fr/course/view.php?id=2990`

1. Nous remercions les étudiant.e.s listé.e.s ci-dessous de nous avoir signalé des erreurs d'orthographe :
   — Yann Denichou, L2 Mathématiques, année 2019/2020
   — Esteban Lajoux, L2, année 2020/2021
   — Adrien Aumaitre, L2, année 2020/2021

Table des matières

Chapitre I

Éléments de combinatoire

La combinatoire (ou analyse combinatoire) est l'étude et, plus en particulier, le dénombrements des configurations d'une collection finie d'objets. Par exemple, voici des questions typiques qui peuvent être résolues à l'aide de la combinatoire :

- Quel est le nombre de parties d'un ensemble de n éléments contenant exactement k éléments ?
- Quel est le nombre des combinaisons de 6 numéros au Loto français, avec au moins 3 numéros gagnants ?
- Quel est le nombre de triangulations d'un polygone convexe dans le plan ? (voir Figure 1.1)

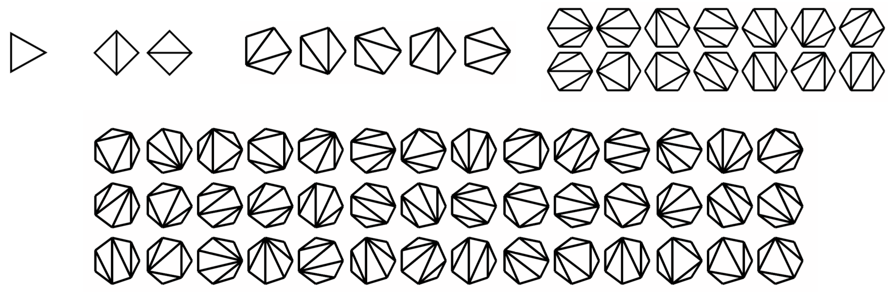


FIGURE 1.1. Triangulations d'un polygone convexe (extrait de P. Flajolet "Analytic Combinatorics")

On utilisera des résultats de combinatoire de base pour la théorie des probabilités sur les ensembles finis. Nous renvoyons à l'Appendice pour des rappels de théorie des ensembles et des fonctions, qui est un prérequis de ce chapitre.

1.1. Dispositions

Soit U un ensemble de cardinal $|U| = n$, on peut penser par exemple à l'ensemble $U = \{1, 2, \ldots, n\}$ des nombres naturels inférieurs à n. Supposons de devoir choisir k éléments parmi les n éléments de U.

On parle de *disposition* si on est interessés à l'*ordre* des éléments choisis, c'est-à-dire en considérant l'arrangement *ordonné* (par exemple, les arrangements 1 2 3 et 2 1 3 sont des dispositions différentes). Nous faisons une distinction en 2 sous-cas distincts : dispositions sans et avec *répétition*.

1.1.1. Dispositions sans répétition

Nous choisissons k éléménts *sans répétition* (c'est-à-dire qu'on ne peut pas choisir le même élément plusieurs fois) : un tel arrangement, si ordonné, est dit une *disposition sans répétition*.

n	0	1	2	3	4	5	...	10	...	20
$n!$	1	1	2	6	24	120	...	3628800	...	2432902008176640000

TABLE 1.1. Valeurs de la factorielle de n, pour n petit.

Théorème 1.1.1 (Dispositions sans répétition)**.** *Soit $k \leq n$, et soit $D_{n,k}$ le nombre de dispositions k-à-k sans répétition de n éléments. Alors*

$$D_{n,k} = \frac{n!}{(n-k)!} = n \cdot (n-1) \cdot (n-2) \cdots (n-k+1).$$

Démonstration. Nous avons n choix pour le premier élément. Pour chacun de ces choix, nous avons $n-1$ choix pour le deuxième éléments, puisque nous ne pouvons pas répéter le premier élément. Donc en totale $n(n-1)$ choix pour les premiers deux éléments. En continuant, nous avons $n-2$ pour le troisième, $n-3$ pour le quatrième etc. pour k fois, ce qui donne la formule. ∎

Dans la formule du Théorème 1.1.1, la valeur $n!$ est appelée la *factorielle* de n (premières valeurs de $n!$ en Table 1.1). La factorielle de n correspond au produit des premiers n nombres naturels, et peut être définie récursivement comme suit :

$$0! = 1 \quad \text{et} \quad n! = n \cdot (n-1)!.$$

Parfois, il peut être utile de calculer l'ordre de grandeur de la valeur $n!$. Le résultat suivant donne une approximation de $n!$ pour n assez grand.

Théorème 1.1.2 (Approximation de Stirling)**.** *La suite $a_n = \frac{n!}{(\frac{n}{e})^n \sqrt{n}}$ est convergente, c'est-à-dire, la limite $\kappa := \lim_n a_n$ est finie. En particulier, on a que*

$$n! \approx \kappa (\tfrac{n}{e})^n \sqrt{n}$$

pour $n \gg 0$. (Pour information, $\kappa = \sqrt{2\pi}$).

Exemple 1.1.1. Nous souhaitons nous habiller différemment pendant une semaine, en évitant de choisir la même couleur de chemise deux jours différents. Nous disposons de 10 chemises de couleurs différentes disponibles. Il y a exactement $D_{10,7} = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 = 604800$ dispositions possibles. Dit autrement, nous avons la possibilité de nous habiller de manière différente pendant 604800 semaines (correspondant à environ 11630 ans). ⋆

Dans le langage des ensembles (voir Appendice), une disposition sans répétition de k éléments d'un ensemble U, $|U| = n$, est une fonction *injective* d'un ensemble A de cardinal k vers U, ce qui correspond à *choisir*, ou *extraire* k éléments de U. La propriété d'injectivité "pour tout $a_1 \neq a_2$ dans A, $f(a_1) \neq f(a_2)$" garantit la non-répétition.

1.1.2. Dispositions avec répétition

Cette fois-ci, nous pouvons choisir k éléménts *avec répétition* mais en considérant encore la disposition *ordonnée* (par exemple, les arrangements $1\,2\,3\,2$ et $2\,2\,1\,3$ sont deux dispositions avec répétition différentes). Remarquons que dans ce cas k peut être superieur à n.

Théorème 1.1.3 (Dispositions avec répétition)**.** *Soit $D'_{n,k}$ le nombre de dispositions k-à-k avec répétition de n éléments. Alors*

$$\boxed{D'_{n,k} = n^k.}$$

Démonstration. Même raisonnement que pour la démonstration du Théorème 1.1.1, sauf que les éléments peuvent être répétés. Donc le nombre de disposition avec répétition est $n \cdot n \cdot n \cdots n$ (k fois) c'est-à-dire n^k. ■

Exemple 1.1.2 (Écriture en base [2] n avec k chiffres). Combien de nombre de au plus k chiffres peut-on écrire en base n ? Supposons pour simplicité que $n = 2$ et $k = 3$: nous nous intéressons aux nombres qui peuvent s'écrire en binaire (c-à-d. avec les symboles 0 et 1) avec 3 chiffres ou moins (les zeros apparaissant à gauche sont normalement éliminés, ce qui donne une écriture avec moins de chiffres). Il y a exactement $8 = 2^3 = n^k$ nombres ayant cette propriété, et précisément :

$$\begin{array}{cccc} 0 = 000 & 1 = 001 & 2 = 010 & 3 = 011 \\ 4 = 100 & 5 = 101 & 6 = 110 & 7 = 111 \end{array}$$

⋆

Une disposition avec répétition correspond à une fonction (n'importe laquelle) d'un ensemble A avec k éléments à valeurs dans U. L'ensemble A contient les k places dans l'arrangement ordonné. Du coup on retrouve bien $|\{f : A \to U\}| = |U^A| = |U|^{|A|} = n^k$.

1.2. Combinaisons

Dans plusieurs cas pratiques, nous sommes intéressés à des suites finies d'objets *sans spécifier un ordre*, et on parle dans ce cas de *combinaisons*. Par exemple les arrangements $1\,5\,4\,2\,3$ et $1\,2\,3\,4\,5$ doivent être considerés comme la même combinaison des cinq premiers nombres naturels (alors que ce sont des dispositions différentes).

1.2.1. Combinaisons sans répétition

Si l'arrangement est *non ordonné* et *sans répétition*, on parle de *combinaison sans répétition*.

Théorème 1.2.1 (Combinaisons sans répétition)**.** *Soit $k \leq n$ et soit $C_{n,k}$ le nombre de combinaisons k-à-k sans répétition de n éléments. Alors*

2. Écrire en base n un certain nombre entier N signifie décomposer N en somme de puissances entières successives de n : $N = a_t n^t + a_{t-1} n^{t-1} + \cdots + a_1 n + a_0 =$ "$a_t\, a_{t-1} \ldots a_0$" (écriture en chiffre). Par exemple en base 2, on a $6 = 4 + 2 = 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 110$. L'écriture est unique.

$$C_{n,k} = \binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

Le symbole $\binom{n}{k}$ est appelé coefficient binomial (de n éléments k-à-k) et est prononcé “k parmi n”.

Démonstration. Soit $D_{n,k}$ le nombre de disposition sans répétition, donné par le Théorème 1.1.1. Remarquons que, si k éléments distincts parmi les n sont fixés, il y a exactement $k!$ arrangement ordonnés (et sans répétition) de ces éléments, c'est-à-dire, il y a exactement $k!$ dispositions sans répétition de ces éléments, alors que il y a *une seule* combinaison (sans répétition) de ces objets. Nous déduisons que le nombre de combinaisons k-à-k sans répétition de n éléments est est $\frac{D_{n,k}}{k!}$ c'est-à-dire $\frac{n!}{k!(n-k)!}$, ce qui donne le résultat. ■

Exemple 1.2.1 (Tirage sans remise). On considère une urne avec 10 boules, numérotées de 1 à 10. On extrait $k = 4$ fois une boule *sans remise*. Si nous ne sommes pas interessés par l'ordre des tirages, l'expérience est équivalente à celle d'extraire un groupe de 4 boules en même temps. Le nombre des possibles résultats du tirage est le nombre de combinaisons de 4 éléménts parmi 10, c'est-à-dire $C_{10,4} = 210$. ⋆

Le nombre $\binom{n}{k}$ est appelé *coefficient binomial*. Voici les propriétés principales des coefficients binomiaux :

— $\binom{n}{0} = \binom{n}{n} = 1$ pour tout n

— $\binom{n}{k} = \binom{n}{n-k}$ (symétrie)

— $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$ (Triangle de Pascal)

Une méthode *récursive* pour calculer avec les coefficients binomiaux est suggérée par la troisième propriété décrite dessus, donnée par ce qu'on appelle le *Triangle de Pascal*[3] (Figure 1.2).

$$\begin{array}{lccccccccccccc}
n=0 & & & & & & & 1 & & & & & & \\
n=1 & & & & & & 1 & & 1 & & & & & \\
n=2 & & & & & 1 & & 2 & & 1 & & & & \\
n=3 & & & & 1 & & 3 & & 3 & & 1 & & & \\
n=4 & & & 1 & & 4 & & \mathbf{6} & & \mathbf{4} & & 1 & & \\
n=5 & & 1 & & 5 & & 10 & & \mathbf{10} & & 5 & & 1 & \\
n=6 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}$$

FIGURE 1.2. Triangle de Pascal

Le triangle de Pascal doit son nom au mathématicien français Blaise Pascal.

Le comportement symétrique de $\binom{n}{k}$ est montré en Figure 1.4 pour $n = 20$. On peut remarquer que le maximum est atteint pour k proche de la moitié de n (précisement pour $k = \frac{n}{2}$ si n est pair, et pour $k = \lfloor \frac{n}{2} \rfloor$ et $k = \lceil \frac{n}{2} \rceil$ si n est impair).

3. Le niveau est donné par n, alors que la profondité dans chaque niveau est donnée par k : $\binom{n+1}{k}$ est donné par la somme des coefficients binomiaux à niveau n et profondité k et $k-1$ (en Figure 1.2, le **10** est la somme de **6** et **4**).

FIGURE 1.3. Le mathématicien Blaise Pascal

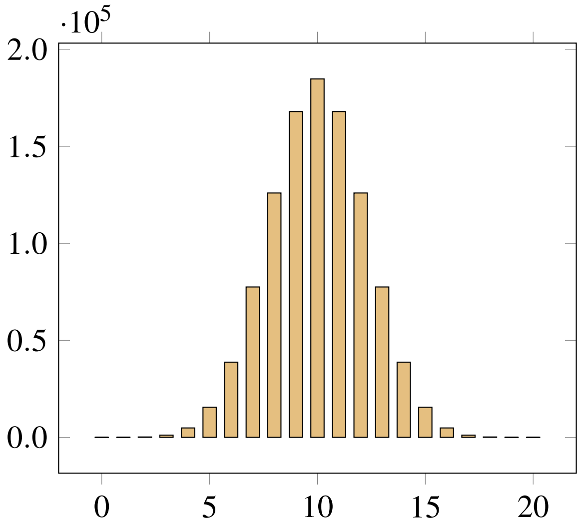


FIGURE 1.4. Valeurs du coefficient binomial $\binom{20}{k}$, pour $k = 0, \ldots, 20$.

1.2.2. Combinaisons avec répétition

On dispose de n éléments appartenants à un ensemble U, et on souhaite un arrangement de taille k *non ordonné* et *avec répétition*, ce qu'on appelle une combinaison avec répétition.

Théorème 1.2.2 (Combinaisons avec répétition)**.** *Soit $C'_{n,k}$ le nombre des combinaisons avec répétition de n objets, k-à-k. Alors*

$$\boxed{C'_{n,k} = \binom{n+k-1}{k}}$$

(Autrement dit, $C'_{n,k} = C_{n+k-1,k}$).

Démonstration. Toute combinaison de k éléments avec répétition à choisir parmi n éléments $a_1, \ldots, a_n$ donnés peut s'écrire (sans perte de généralité) comme suit :

$$\underbrace{a_1 \ldots a_1}_{k_1 \text{ fois}}, \underbrace{a_2 \ldots a_2}_{k_2 \text{ fois}}, \ldots, \underbrace{a_n \ldots a_n}_{k_n \text{ fois}}$$

avec $k_1 + \cdots + k_n = k$. Cela correspond à placer k objets dans n boîtes, et, après, à appeler a_i tous les objets placés dans la $i-$ème boîte. De manière équivalente, à placer $n-1$ "cloisons" pour séparer k objets :

$$\underbrace{a_1 \dots a_1}_{k_1 \text{ fois}} \underbrace{\theta_1}_{\text{cloison}} \underbrace{a_2 \dots a_2}_{k_2 \text{ fois}} \underbrace{\theta_2}_{\text{cloison}} \dots \underbrace{\theta_{n-1}}_{\text{cloison}} \underbrace{a_n \dots a_n}_{k_n \text{ fois}}.$$

Du coup le nombre de combinaisons avec répétition des éléments $a_1, \dots, a_n$ pris k à k est le nombre de combinaisons de $n-1$ objets parmi $k+n-1$, c'est-à-dire $C_{n+k-1,n-1}$ ($= C_{n+k-1,k}$). ■

Exemple 1.2.2 (Nombre de monômes de degré k). Soit $x = (x_1, \dots, x_n)$ un vecteur de n variables, et soit $k \in \mathbb{N}$. Un *monôme* est un produit de la forme $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$, où $i_1, \dots, i_n$ sont des entiers positifs, et son degré est $i_1 + \cdots + i_n$ (par exemple x^2yz^5 et xyz sont deux monômes en trois variables, de degré 8 et 3, respectivement). Le nombre de monômes de degré k en x est donné par le nombre de combinaisons avec répétition pris k à k, des éléments de $U = \{x_1, \dots, x_n\}$ (combinaisons car le produit de variables étant commutatif, on n'est pas interessés par l'ordre ; avec répétition car une variable peut apparaître avec une puissance ≥ 2). Par exemple, il y a exactement dix monômes de degré trois en trois variables (ce sont $x^3, x^2y, x^2z, xy^2, xyz, xz^2, y^3, y^2z, yz^2, z^3$) et on vérifie bien que $\binom{3+3-1}{3} = \binom{5}{3} = 10$. ⋆

La propriété $C'_{n,k} = C'_{n,k-1} + C'_{n-1,k}$ est une consequence directe du Triangle de Pascal appliqué à $C_{n+k-1,k}$, car $C_{n+k-1,k} = C_{n+k-2,k-1} + C_{n+k-2,k}$.

1.3. Permutations

Nous donnons un nom spécial aux arrangements de n éléments parmi n (c'est-à-dire au cas $k = n$ en Section 1.1) : permutation. La possibilité de répétition joue un rôle aussi pour les permutations, qui seront donc elle aussi sans ou avec répétition.

1.3.1. Permutations sans répétition

Il s'agit d'une disposition sans répétition de n éléments parmi n.

Théorème 1.3.1 (Permutations sans répétition)**.** *Le nombre P_n des permutations sans répétition de n objets est*

$$P_n = n!$$

Démonstration. Voir démonstration du Théorème 1.1.1. ■

Exemple 1.3.1 (Code confidentiel). Le code confidentiel d'une carte bleue est formé par quatre chiffres. Le nombre de codes possibles est évidemment très élévé : il s'agit du nombre de dispositions avec répétitions (voir Section 1.1.2) de 10 chiffres 4 à 4, donc c'est bien $D'_{10,4} = 10^4 = 10000$. Mais supposons pour un moment de savoir que le code est formé par les nombres $1, 2, 6, 9$. Combien de tentatives faut-il faire pour être sûrs de saisir le code correct (code de 4 chiffres) ? Exactement $P_4 = 4! = 24$ tentatives. Pour cette raison, et pour protéger les clients de la banque contre l'usage frauduleux de la carte, normalement le nombre de tentatives est limité. ⋆

1.3.2. Permutations avec répétition

Nous considérons un ensemble U de n éléments, et des nombres naturels $k_1,\ldots,k_n$. Nous souhaitons obtenir un arrangement de $k_1+\cdots+k_n$ éléments ou n sont distincts (les éléments de U) et où les éléments de U se répètent respectivement $k_1,\ldots,k_n$ fois.

Théorème 1.3.2 (Permutations avec répétition)**.** *Le nombre $P'_{n,k_1,\ldots,k_n}$ des permutations avec répétition de $k=k_1+\cdots+k_n$ éléments, où n sont distincts (ou permutations avec répétitions $k_1,\ldots,k_n$) est*

$$\boxed{P'_{n,k_1,\ldots,k_n} = \tfrac{(k_1+\cdots+k_n)!}{k_1!\cdots k_n!}}$$

Démonstration. Nous montrons d'abord que le nombre de permutations de n objets, avec répétitions $k_1,\ldots,k_n$, est égal à

$$C_{k_1+\cdots+k_n,k_1}\cdot C_{k_2+\cdots+k_n,k_2}\cdots C_{k_n,k_n}.$$

En effet, pour construire une suite de $k_1+\cdots+k_n$ éléments avec les répétitions données, nous devons choisir (combinaisons, car l'ordre n'intervient pas)
- les k_1 emplacements du premier élément de U, parmi $k_1+\cdots+k_n$ places
- les k_2 emplacements du deuxième élément de U, parmi $k_2+\cdots+k_n$ places
- ...
- les k_n emplacements du dernier élément de U, parmi k_n places

ce qui donne le produit des coefficients $C_{k_i+\cdots+k_n,k_i}$, pour $i=1,\ldots,n$. Le fait que ça coincide avec la formule du théorème, suit simplement en utilisant la définition de $C_{k_i+\cdots+k_n,k_i}$:

$$\frac{(k_1+\cdots+k_n)!}{k_1!\cancel{(k_2+\cdots+k_n)!}}\frac{\cancel{(k_2+\cdots+k_n)!}}{k_2!\cancel{(k_3+\cdots+k_n)!}}\cdots\frac{\cancel{(k_{n-1}+k_n)!}}{k_{n-1}!\cancel{(k_n)!}}\frac{\cancel{(k_n)!}}{k_n!}=\frac{(k_1+\cdots+k_n)!}{k_1!\cdots k_n!}$$

■

Exemple 1.3.2 (Nombre d'anagrammes). Un anagramme d'un mot donné est un autre mot (pas forcément listé dans le vocabulaire) qui utilise les mêmes lettres. Par exemple, PARISIEN est un anagramme (avec sens) de ASPIRINE : dans ce cas $U=\{A,S,P,I,R,N,E\}$, $n=|U|=7$, $k_1+\cdots+k_7=8=$ longueur du mot. La lettre I apparait 2 fois, du coup il faut diviser par $2!=2$ le nombre total de permutations de Section 1.3.1, donc le nombre d'anagrammes est $=8!/2!=20160$. ⋆

Voici un tableau récapitulatif des types combinatoires qui ont été introduits :

Arrangement	Répétition	Nombre	Exemple
Dispositions	sans	$D_{n,k} = \frac{n!}{(n-k)!}$	*Ex. 1.1.1* : Façons de s'habiller.
Dispositions	avec	$D'_{n,k} = n^k$	*Ex. 1.1.2* : Écriture en base n.
Combinaisons	sans	$C_{n,k} = \binom{n}{k}$	*Ex. 1.2.1* : Tirages sans remise.
Combinaisons	avec	$C'_{n,k} = \binom{n+k-1}{k}$.	*Ex. 1.2.2* : Nombre de monômes.
Permutations	sans	$P_n = n!$	*Ex. 1.3.1* : Code carte bleue.
Permutations	avec	$P'_{n,k_1,\dots,k_n} = \frac{(k_1+\cdots+k_n)!}{k_1!k_2!\cdots k_n!}$	*Ex. 1.3.2* : Anagrammes.

1.4. Théorème du binôme

Les coefficients binomiaux font l'objet d'un résultat centrale en combinatoire.

Théorème 1.4.1 (Théorème du binôme)**.** *Pour tout $a,b \in \mathbb{R}$ et $n \in \mathbb{N}$*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Démonstration. Soit P_n la propriété à prouver pour $n \in \mathbb{N}$. La preuve est par *récurrence*[4] sur n. Pour $n = 0$, la propriété est verifiée puisque

$$(a+b)^0 = 1 \quad \text{et} \quad \binom{0}{0} a^0 b^0 = 1.$$

Supposons P_n vraie. On a que

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n = (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} = \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \overset{p=k+1}{=} \\
&= \left(\textstyle\sum_{p=1}^{n}\binom{n}{p-1}a^p b^{n-p+1} + \binom{n}{n}a^{n+1}b^0\right) + \left(\binom{n}{0}a^0 b^{n+1} + \textstyle\sum_{k=1}^{n}\binom{n}{k}a^k b^{n-k+1}\right) = \\
&= a^{n+1} + b^{n+1} + \sum_{p=1}^{n}\left(\binom{n}{p-1} + \binom{n}{p}\right)a^p b^{n-p+1} \overset{Pascal}{=} \\
&= a^{n+1} + b^{n+1} + \sum_{p=1}^{n}\binom{n+1}{p}a^p b^{n-p+1} = \\
&= \sum_{p=0}^{n+1}\binom{n+1}{p}a^p b^{n-p+1}
\end{aligned}$$

ce qui montre que P_{n+1} est vraie. ■

4. Soit P_n une propriété paramétrée par $n \in \mathbb{N}$. Une preuve de P_n par recurrence sur n est donnée par : (1) une preuve de P_0 et (2) une preuve de l'implication $P_n \to P_{n+1}$.

Exemples de développements du binôme $(a+b)^n$ à comparer avec le Triangle de Pascal de page 8 :

$$
\begin{aligned}
(a+b)^3 &= a^3 + \mathbf{3}a^2b + \mathbf{3}ab^2 + b^3 \\
(a+b)^4 &= a^4 + \mathbf{4}a^3b + \mathbf{6}a^2b^2 + \mathbf{4}ab^3 + b^4 \\
(a+b)^5 &= a^5 + \mathbf{5}a^4b + \mathbf{10}a^3b^2 + \mathbf{10}a^2b^3 + \mathbf{5}ab^4 + b^5 \\
(a+b)^6 &= a^6 + \mathbf{6}a^5b + \mathbf{15}a^4b^2 + \mathbf{20}a^3b^3 + \mathbf{15}a^2b^4 + \mathbf{6}ab^5 + b^6
\end{aligned}
$$

1.5. Exercices du Chapitre I

1. Quel est le nombre de parties de cardinal $n-k$ d'un ensemble de n éléments (pour $k \leq n$) ? Comparer avec les combinaisons sans répétition de n éléments pris k à k. Que peut-on remarquer ?
2. Quel est le nombre des combinaisons de 6 numéros (5 numéros de 1 à 49 + un numéro *joker* entre 1 et 10) au Loto français, avec au moins 3 numéros gagnants (en sachant que le joker *peut* être un des numéros gagnantes) ?
3. Démontrer les *Formules de de Morgan* : si A, B sont des parties d'un ensemble U, alors

$$(A \cap B)^c = A^c \cup B^c$$
$$(A \cup B)^c = A^c \cap B^c$$

   où $A^c = U \setminus A$ est le *complémentaire* de A dans U.
4. Combien le mot KOLMOGOROV a-t-il d'anagrammes ? Et POISSONS ?
5. Avec les chiffres $0, 1, 2, 6$ et 9 utilisés une et une seule fois, combien peut–on écrire de nombres à *exactement* 5 chiffres ? (Évidemment le nombre 01692 n'a pas cinq chiffres.)
6. Montrer pour tout $n \in \mathbb{N}$ l'égalité $\sum_{k=0}^{n} \binom{n}{k} = 2^n$ en utilisant le Théorème du Binôme.
7. Si A est un ensemble de cardinal n, montrer que $|\mathscr{P}(A)| = 2^n$.
8. Montrer les trois propriétés des coefficients binomiaux de page 8 en utilisant la définition.
9. Expliquer la propriété $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$ du Triangle de Pascal à l'aide de la théorie des ensembles (se rappeler que $\binom{n}{k}$ est le nombre de parties de cardinal k de l'ensemble $\{1, \ldots, n\}$).
10. Un voyageur veut visiter n villes françaises, chacune une et une seule fois, en commençant d'une ville donnée. Combien de parcours possibles peut-il emprunter ? Et combien si on ne fixe pas une ville de départ ?
11. Remarquer que pour tout n, k, on a $k!C_{n,k} = D_{n,k}$. Comment peut-on expliquer cette relation ?
12. Une urne contient 10 boules noires et 10 boules rouges. On tire deux boules sans remise. Quel est le nombre
    a. de tirages possibles ?
    b. de tirages où les deux boules sont noires ?
    c. de tirages où au moins une des boules est noire ?
    d. de tirages où les deux boules sont de même couleur ? de couleur differente ?
13. Dans une lotterie, les participants doivent miser sur 6 numéros compris entre 1 et 50. Quel est le nombre total de sextuplets ? Quel est le nombre de sextuplets avec *exactement* trois numéros gagnants ? Quel est le nombre de sextuplets avec *au moins* quatre numéros gagnants ?

14. Une course hippique compte 20 partants (les chevaux sont identifiés par un numéro entre 1 et 20). Un tiercé est la liste des trois premiers chevaux de l'arrivée. Nous parions sur un des chevaux.
    a. Combien a-t-on de tiercé possibles ?
    b. Combien a-t-on de tiercé où notre cheval est gagnant ? Où il est second ?
    c. Combien a-t-on de tiercé où figure notre cheval ?

Chapitre II

Introduction aux probabilités

2.1. Vocabulaire de base

Dans ce chapitre, nous allons donner une définition rigoureuse d'incertitude et d'événements incertains (ou aléatoires). L'incertitude est une propriété intrinsèque des événements naturels, et est mesurée par une fonction appelée *probabilité*.

2.1.1. Expérience aléatoire

Nous donnons d'abord cette définition d'expérience aléatoire : elle est d'un côté "abstraite" et de l'autre elle a un aspect "physique" que nous souhaitons souligner.

Definition 2.1.1. *Une* expérience aléatoire *(ou* épreuve aléatoire*) est une expérience qui peut être répétée dans des conditions identiques et dont le résultat n'est pas prévisible car il dépend du hasard ou de facteurs qui ne peuvent pas être determinés ou mesurés* à priori.

Voici quelques exemples d'expériences aléatoires :
- Lancer d'un ou plusieurs dés, de monnaies. Plus en général, les jeux de hasard, jeux de cartes, loterie ...
- Tirage aléatoire d'une urne ou une boîte, contenant objets discernables.
- La production industrielle ou artisanale d'un objet (contexte où l'on veut mesurer la probabilité d'avoir d'erreurs de production).

Le point de vue *probabiliste* ou *aléatoire* d'une expérience, par exemple en physique, doit être comparé (on pourrait même dire mis en opposition) avec le point de vue *déterministe* qui fait référence plutôt aux théories classiques : d'un point de vue classique, le résultat d'un lancer d'un dé peut être décrit *à priori* par une fonction de l'état initial du dé, de la force appliquée par le lancer et des contraintes physiques que le dé rencontrera du moment où il est lancé au moment où il arrête son mouvement. C'est évident que le schéma classique nécessite d'une description très complèxe de la dynamique du dé et notamment des facteurs extérieurs.

2.1.2. Événements et Ensemble univers Ω

Fixée une épreuve, nous allons déterminer un modèle mathématique pour en représenter les résultats (même si aléatoires).

Definition 2.1.2. *L'ensemble des résultats possibles d'une épreuve aléatoire est appelé* ensemble univers*, et noté classiquement* Ω *(*omega*). Nous utiliserons cette notation. Les parties de* Ω *(c'est-à-dire les éléments de* $\mathscr{P}(\Omega)$*) qui contiennent un seul élément sont appelées* événements élémentaires*, et plus en général les parties de* Ω *simplement* événements.

Les événements correspondent donc à une collection de résultats possibles, et sont typiquement denotés avec les lettres $E,F,G\dots$ Par exemple, l'ensemble vide $\emptyset$ est un évenement (on verra plus tard que $\emptyset$ correspond à un résultat impossible).

En tant que parties de Ω, donc en tant que ensembles, on peut faire des opérations tel que l'union ou l'intersection, en obtenant d'autres évenements.

Definition 2.1.3. *Deux événements E,F tel que $E\cap F=\emptyset$ sont dits* incompatibles.

On verra plus tard que le fait que deux événements sont incompatibles est vérifié si les événements qu'ils répresentent ne peuvent pas se vérifier au même temps. Par exemple, si on jette un dé régulier, les événements « le résultat est pair » et « le résultat est impair » sont évidemment incompatibles.

Voici des exemples d'ensemble univers.

Exemple 2.1.1. $\Omega=\{1,2,3,4,5,6\}$ est l'ensemble univers pour un lancer d'un dé regulier à 6 faces. L'événement élémentaire $\{4\}$ est le résultat « 4 » alors que l'événement $E=\{1,3,5\}\in\mathscr{P}(\Omega)$ est le résultat « nombre impair ». Il y a en total $2^6=64$ événements qui peuvent être définis pour cette expérience (car tel est le cardinal de $\mathscr{P}(\Omega)$). ⋆

Exemple 2.1.2. $\Omega=\{P,F\}$ est l'ensemble univers pour un lancer d'une monnaie (pile, face). Il y a deux événements élémentaires possibles : $\{P\}$ (pile) ou $\{F\}$ (face). Dans $\mathscr{P}(\Omega)=\{\emptyset,\{P\},\{F\},\Omega\}$ il y a exactement quatre événements possibles :
$E=\emptyset$ (est-il possible ?)
$E=\{P\}$: le résultat est pile
$E=\{F\}$: le résultat est face
$E=\Omega$: le résultat est pile ou face (est-il possible ?)

⋆

Nous avons vu dans les exemples précédents que un événement correspond à un sous-ensemble de Ω (donc à une collection d'événements élémentaires). Donc on peut se demander quel est le résultat des opérations basiques sur les ensembles correspondants, par exemple, à quel événement corréspond l'union de deux événements.

Definition 2.1.4. *Soient $A\subset\Omega$ et $\omega\in\Omega$. On dit que ω* vérifie *A (ou que A est vérifié par ω) si $\omega\in A$.*

Union. Soient $E,F\subset\Omega$. Alors $E\cup F$ est vérifié par $\omega\in\Omega$ si et seulement si (par définition) ω est un élément de $E\cup F$. Vu que $\omega\in E\cup F$ si et seulement si $\omega\in E$ **ou** $\omega\in F$, on conclut que *$E\cup F$ est vérifié (par ω) si et seulement si E est vérifié (par ω) ou F est vérifié (par ω).* Le mot **ou** ici est non-exclusif, comme le latin **vel** (par exemple on peut dire que "2 est un nombre pair **ou** prémier", voir l'Appendice).

Intersection. De manière similaire, on conclut que *$E\cap F$ est vérifié si et seulement si E est vérifié* **et** *F l'est.*

Différence et différence symétrique. La différence $A\setminus B$ est par définition l'ensemble des éléments qui appartiennent à A mais pas à B, donc : *$A\setminus B$ est vérifié si et seulement si A est* *vérifié et B ne l'est pas*. L'événement différence symétrique de A et B, noté $A \Delta B$, corréspond à l'union $(A \setminus B) \cup (B \setminus A)$, et donc *est vérifiée si et seulement si A est vérifié et B ne l'est pas ou B est vérifié et A ne l'est pas*.

Un événement E qui n'est pas élémentaire peut être decomposé comme union de deux événements, chacun d'eux étant différent de E. Par exemple chacque événement est union d'événements élémentaires car $E = \cup_{\omega \in E} \{\omega\}$.

Exemple 2.1.3. L'événement $E =$ « obtenir 10 » en lançant deux dés est un événement composé, car il peut se vérifier en obtenant ⚃⚅ ou ⚅⚃ ou ⚄⚄, donc $E = \{$⚃⚅, ⚅⚃, ⚄⚄$\}$ alors que l'événement « obtenir 12 » est élémentaire, car il correspond au singleton $F = \{$⚅⚅$\}$. ⋆

Pour l'instant nous avons vu des exemples d'expériences aléatoires simples, pourtant nous allons souvent utiliser des expériences répétées. Voici deux exemples.

Exemple 2.1.4. Pour deux lancers consecutifs d'une pièce de monnaie, l'ensembles des événements élémentaires est $\Omega = \{(P,P),(F,P),(P,F),(F,F)\}$. Les éléments sont les couples ordonnées où le premier élément est le résultat du premier lancer, et le deuxième est le résultat du deuxième lancer. ⋆

Exemple 2.1.5. L'ensemble $\Omega = \{\{r,r\},\{r,b\},\{b,b\}\}$ est l'ensemble univers pour l'épreuve suivante : deux tirages (avec ou sans remise, peu importe pour l'instant), d'une urne contenant boules rouges (r) ou bleues (b), sans tenir compte de l'ordre. En revanche, $\Omega = \{(r,r),(r,b),(b,r),(b,b)\}$ est l'ensemble univers pour l'épreuve suivante : deux tirages (avec ou sans remise), d'une urne contenant boules rouges (r) ou bleues (b), en tenant compte de l'ordre. Du coup dans le deuxième cas, chaque événement élémentaire est une *couple ordonnée* formée par les symboles r et b, alors que dans le premier cas, nous ne nous interessions pas à l'ordre des résultats. ⋆

2.1.3. Probabilité sur un ensemble fini

Soit Ω un ensemble fini, c'est à dire $\Omega = \{\omega_1, \dots, \omega_n\}$ où (par définition d'ensemble) $\omega_i \neq \omega_j$ pour $i \neq j$. Nous supposons d'attribuer des *poids* $p_1, \dots, p_n$ aux éléments ω_i respectivement, vérifiant $p_1 + \cdots + p_n = 1$ et $p_i \geq 0$, $i = 1, \dots, n$.

Dans ce contexte, et de manière naïve, la probabilité d'un événement $E \in \mathscr{P}(\Omega)$ peut se définir comme suit :

$$\mathbf{P}(E) = \sum_{\omega_i \in E} p_i$$

c'est-à-dire la somme des poids des éléments de E.

Dans le cas d'*equiprobabilité*, c'est-à-dire quand $p_i = \frac{1}{|\Omega|} = \frac{1}{n}$ pour tout i, on trouve bien une définition classique de probabilité :

$$\mathbf{P}(E) = \frac{|E|}{|\Omega|} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

Exemple 2.1.6. Dans le cas du tirage d'un dé regulier à six faces, si E est l'événement $\{2,4,6\}$ (« le résultat est un nombre pair »), alors la probabilité de E sera $\frac{|E|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$. ⋆

Exemple 2.1.7. D'une caisse qui contient n pièces numérotées de 1 à n, nous tirons au hasard l'une après l'autre toutes les pièces (sans remise). Quelle est la probabilité pour que les numéros des pièces extraites observent l'ordre $1,2,\dots,n$? Ici Ω est l'ensemble des dispositions (on tient compte de l'ordre) de n objets pris n à n, c'est-à-dire les permutations de n objets. Du coup les cas possibles sont $n!$ et la probabilité $\frac{1}{n!}$ car toutes les permutations ont le même poids, c'est-à-dire, la même probabilité. ⋆

Exemple 2.1.8. On considère la même caisse que dans l'exemple précédent, mais après le tirage le numéro de chaque pièce est noté et elle est remise dans la caisse et mélangée avec les autres. La probabilité pour que la succession enregistrée des numéros soit $1,2,\dots,n$ est maintenant $\frac{1}{n^n}$, puisque il s'agit de disposition avec répétition de n objets parmi n, et $D'_{n,n} = n^n$ (et nous avons encore equiprobabilité des arrangements). ⋆

Exemple 2.1.9. Au championnat de basket-ball participent 18 équipes dont on forme au hasard deux groupes de 9 équipes chacun. Parmi les concurrents, il y a 5 équipes d'extra-classe. Quelle est la probabilité des événements suivants ?

$$A = \{\text{toutes les équipes d'extra-classe entrent dans le même groupe}\};$$

$$B = \{\text{deux équipes d'extra-classe entrent dans un groupe, et trois dans l'autre}\}.$$

Pour l'événement A, on a $\mathbf{P}(A) = \frac{C_{13,4}+C_{13,9}}{C_{18,9}} = \frac{1}{34}$. En fait, les cas possibles sont donnés par le choix de 9 équipes parmi 18 pour le premier groupe (le deuxième est consequent), donc $C_{18,9}$ (sans ordre ni répétition). Les cas favorables sont : les 5 équipes entrent dans le groupe choisi (et il faut en choisir 4 en plus parmi ce qui restent) ou bien on a choisi 9 équipes parmi les 13 équipes pas bonnes (montrer que on ne compte pas deux fois le même arrangement, et donc que le numerateur est bien $C_{13,4}+C_{13,9}$).
De manière similaire $\mathbf{P}(B) = \frac{C_{5,2}C_{13,7}+C_{5,3}C_{13,6}}{C_{18,9}} = \frac{12}{17}$: ici il faut choisir 2 équipes extra-classe parmi 5 et pour chaque choix il faut encore choisir 7 équipes dans les 13 non-extra-classe ou sinon 3 équipes extra-classe parmi 5 et 6 équipes dans les 13 non-extra-classe (montrer que on ne compte pas deux fois le même arrangement). ⋆

2.1.4. Loi de probabilité et espace probabilisé

Évidemment, la définition naïve dans le cas d'ensembles finis, ne peut pas s'étendre naturellement au cas d'ensembles infinis. Plus rigoureusement, pour un ensemble Ω quelconque, nous allons considérer une collection d'événements, c'est-à-dire une partie $\mathscr{C} \subset \mathscr{P}(\Omega)$, satisfaisante les axioms de $\sigma-$*algèbre*.

Definition 2.1.5. *Soit Ω un ensemble. Alors $\mathscr{C} \subset \mathscr{P}(\Omega)$ est une $\sigma-$*algèbre d'événements *pour Ω si :*

- $\emptyset \in \mathscr{C}$
- *Pour tout $A \in \mathscr{C}$ on a que $A^c \in \mathscr{C}$*
- *Si $A_n \in \mathscr{C}$, $n \in \mathbb{N}$, alors $\bigcup_n A_n \in \mathscr{C}$.*

Autrement dit, une $\sigma-$algèbre est une collection de parties de Ω qui contient l'ensemble vide et qui est stable par passage au complément et par union dénombrable.

Par exemple, $\mathscr{P}(\Omega)$ est une $\sigma-$algèbre sur Ω. Remarquez que Ω est contenu dans toute $\sigma-$algèbres sur Ω (en utilisant les premières deux propriétés), et que les $\sigma-$algèbres sont stables par intersection (en applicant les deux dernières propriétés).

Avec une structure de $\sigma-$algèbre d'événements, nous pouvons donner une définition rigoureuse de probabilité.

Definition 2.1.6. *Soit Ω un ensemble et soit $\mathscr{C} \subset \mathscr{P}(\Omega)$ une $\sigma-$algèbre sur Ω. Une loi de probabilité sur $(\Omega, \mathscr{C})$ (ou simplement une probabilité) est une application $\mathbf{P} : \mathscr{C} \to [0,1]$ telle que*

— $\mathbf{P}(\Omega) = 1$

— *Pour une collection dénombrable $T = \{A_1, A_2, A_3, \ldots\}$ d'éléments de $\mathscr{C}$, incompatibles deux à deux (c'est-à-dire vérifiant $A_i \cap A_j = \emptyset$ pour tout $i \neq j$), on a*

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

La deuxième propriété de la définition de probabilité est appelée $\sigma-$*additivité* de $\mathbf{P}$. Remarquez que, par hypothèse, la famille T est une partition de $\bigcup_{i=1}^{+\infty} A_i$, donc la propriété dit que si une union est disjointe (donc une partition) la probabilité de l'union est la somme des probabilités. Évidemment l'additivité n'est en général pas valable pour des familles arbitraires d'événements.

Il faut faire attention à la somme $\sum_{i=1}^{+\infty} \mathbf{P}(A_i)$ qui peut faire intervenir un nombre infini de termes, comme dans l'exemple qui suit (il s'agit d'une *série numérique*, voir l'Appendice).

Exemple 2.1.10. Soit $\Omega = \mathbb{N} = \{0, 1, 2 \ldots\}$, et soit $\mathbf{P}$ la fonction qui associe $\mathbf{P}(n) = \frac{1}{2^{n+1}}$. Alors $\mathbf{P}$ est une probabilité sur $(\Omega, \mathscr{P}(\Omega))$. On peut verifier aussi l'égalité

$$\sum_{n=0}^{+\infty} \frac{1}{2^{n+1}} = 1$$

en choisissant $T = \mathscr{P}_1(\Omega) = \{\{0\}, \{1\}, \ldots\}$. Plus en général, on peut associer $\mathbf{P}(n) = p_n$ où $p_n \geq 0$ pour tout n, et $\sum_{n=0}^{+\infty} p_n = 1$. La fonction $\mathbf{P}(n) = \frac{1}{n^2}$ n'est pas une probabilité sur $\mathbb{N} \setminus \{0\}$, par contre $\frac{6}{\pi^2} \frac{1}{n^2}$ l'est (pourquoi ?). ⋆

Exemple 2.1.11. Soit $a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$ et soit f une fonction continue et positive sur (a, b) tel que $\int_a^b f(t)dt = 1$. Alors la fonction $\mathbf{P}(S) = \int_S f(t)dt$ est une loi de probabilité définie sur la $\sigma-$algèbre $\mathscr{C} \subset \mathscr{P}((a,b))$ contenante les sous-ensembles mésurables de (a, b). L'idée ici est que nous *mésurons* S à partir d'une fonction définie sur (a, b), en calculant l'intégrale dessus. ⋆

Exemple 2.1.12. On doit choisir au hasard un point du carré $Q \subset \mathbb{R}^2$ de sommets $(-1, -1)$, $(-1, 1)$,$(1, -1)$ et $(1, 1)$. On suppose pour l'instant que tous les points sont equiprobables (on dira plus tard que la distribution sur Q est uniforme). L'ensemble univers est donc l'ensemble Q

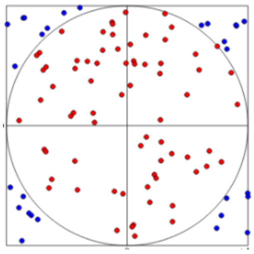


FIGURE 1.5.

même : $\Omega = Q$. Le disque unitaire $D = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 \le 1\}$ est un sous-ensemble de Ω, donc $D \in \mathscr{P}(\Omega)$, et donc on peut s'intéresser à la probabilité que le point choisi appartienne à D, c'est-à-dire $\mathbf{P}(D)$ (voir Figure 1.5). ⋆

Definition 2.1.7. *Un espace probabilisé est une triple $(\Omega, \mathscr{C}, \mathbf{P})$ où Ω est un ensemble, $\mathscr{C}$ est une $\sigma-$algèbre sur Ω et $\mathbf{P}$ est une loi de probabilité sur $(\Omega, \mathscr{C})$.*

Propriétés des lois de probabilités :

$\mathbf{P}(\emptyset) = 0$. *Démonstration :* Pour tout événements E, on a $E = E \cup \emptyset$. Du coup $\mathbf{P}(E) = \mathbf{P}(E \cup \emptyset) = \mathbf{P}(E) + \mathbf{P}(\emptyset)$ (car $E \cap \emptyset = \emptyset$). Donc $\mathbf{P}(\emptyset) = 0$.

$\mathbf{P}(E^c) = 1 - \mathbf{P}(E)$. *Démonstration :* On applique $E \cup E^c = \Omega$ et $E \cap E^c = \emptyset$. Donc $1 = \mathbf{P}(\Omega) = \mathbf{P}(E \cup E^c) = \mathbf{P}(E) + \mathbf{P}(E^c)$.

$\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F)$. En particulier, remarquez que si E et F sont incompatibles, nous retrouvons $\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F)$. *Démonstration :* On applique les propriétés de $\mathbf{P}$ à la partition $E \cup F = (E \cap F^c) \cup (E^c \cap F) \cup (E \cap F)$.

$\mathbf{P}(E) \le \mathbf{P}(F)$ si $E \subset F$. *Idée de la démonstration :* Utilisez la partition $F = E \cup (F \setminus E)$.

$\mathbf{P}\left(\bigcup_{i=1} E_i\right) \le \sum_{i=1} \mathbf{P}(E_i)$ (avec égalité si et seulement si les évenements sont incompatibles 2 à 2). *Idée de la démonstration :* Utilisez la troisième propriété "$\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F)$" de manière récursive.

Si $E_1 \supset E_2 \supset \cdots \supset E_n \supset \cdots$, et $\lim_i E_i = \emptyset$, alors $\lim_i \mathbf{P}(E_i) = 0$. Sans démonstration.

Exemple 2.1.13. On dispose d'un jeu de 52 cartes (13 cartes pour chacune des 4 enseignes $\clubsuit, \diamondsuit, \heartsuit, \spadesuit$). On pioche une carte au hasard. Quelle est la probabilité qu'elle soit le roi de $\clubsuit$ ou une dame (d'une couleur quelconque) ? On a les événements

$$E = \text{« la carte est le roi de } \clubsuit \text{ »} = \{\mathscr{K}_\clubsuit\};$$

$$F = \text{« la carte est une dame »} = \{\mathscr{Q}_\clubsuit, \mathscr{Q}_\diamondsuit, \mathscr{Q}_\heartsuit, \mathscr{Q}_\spadesuit\}.$$

Vu que $E \cap F = \emptyset$, on a que $\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) = \frac{1}{52} + \frac{4}{52} = \frac{5}{52}$. ⋆

2.2. Indépendance

Deux événements E et F sont *indépendants* si la réalisation d'un des deux n'affecte pas la réalisation de l'autre. Ceci est une définition qualitative et non mathématique. Nous donnons maintenant une définition rigoureuse :

Definition 2.2.1. *Deux événements E et F sont indépendants si* $\mathbf{P}(E \cap F) = \mathbf{P}(E)\mathbf{P}(F)$.

Remarquez la différence avec les événements incompatibles :

Remarque 2.2.1. *Deux événements E et F incompatibles ($E \cap F = \emptyset$) sont indépendants si et seulement si* $\mathbf{P}(E)\mathbf{P}(F) = 0$ *donc si et seulement si* $\mathbf{P}(E) = 0$ *ou* $\mathbf{P}(F) = 0$.

Du coup en général deux événements incompatibles peuvent avoir une dépendance. En revanche, tout événement E est incompatible et indépendant avec $\emptyset$, car $E \cap \emptyset = \emptyset$, donc $\mathbf{P}(E \cap \emptyset) = 0 = \mathbf{P}(E)\mathbf{P}(\emptyset)$. En plus, un événement E est indépendant avec E (lui même) si et seulement si $\mathbf{P}(E) = 0$ ou $\mathbf{P}(E) = 1$, puisque pour tout événement E on a $\mathbf{P}(E) = \mathbf{P}(E \cap E)$ et donc l'indépendance impliquerait $\mathbf{P}(E) = \mathbf{P}(E)^2$.

Exemple 2.2.1 (Événements indépendants). On jette deux dés réguliers à 6 faces (sans intéraction entre les deux lancers). Toute couples E, F d'événements, où E concerne le premier dé (ex. $E =$ « le résultat du premier dé est 6 ») et F le deuxième (ex. $F =$ « le résultat du deuxième dé est impair »), sont évidemment indépendants, car on lance des dés différents et pour les informations qu'on possède, le résultat du premier n'a aucun impact sur le résultat du deuxième. Dit autrement, l'expérience aléatoire est composée de deux expériences qui n'ont rien à voir entre elles. On peut quand même vérifier l'indépendance de E et F en calculant les probabilités $\mathbf{P}(E \cap F)$, $\mathbf{P}(E)$ et $\mathbf{P}(F)$, et en vérifiant que $\mathbf{P}(E \cap F) = \mathbf{P}(E)\mathbf{P}(F)$. En fait, pour les événements décrits avant en parenthèse, nous avons

$$E = \{(6,y) : y = 1,\dots,6\} \quad \text{et} \quad F = \{(x,1),(x,3),(x,5) : x = 1,\dots,6\}$$

et donc $E \cap F = \{(6,1),(6,3),(6,5)\}$. Donc $\mathbf{P}(E \cap F) = \frac{3}{36} = \frac{1}{12}$, $\mathbf{P}(E) = \frac{6}{36} = \frac{1}{6}$ et que $\mathbf{P}(F) = \frac{18}{36} = \frac{1}{2}$. On vérifie bien que $\frac{1}{6}\frac{1}{2} = \frac{1}{12}$, et donc que E et F sont indépendants. ⋆

Un exemple d'événements non-indépendants (on pourra dire, dépendants) est discuté dans l'exemple qui suit.

Exemple 2.2.2 (Événements non indépendants). Soit $E =$ « le résultat est ≤ 3 » et $F =$ « le résultat est pair », pour le lancer d'un dé régulier à 6 faces. L'intersection est $E \cap F = \{2\}$, donc $\mathbf{P}(E \cap F) = \frac{1}{6} \neq \frac{1}{2}\frac{1}{2} = \mathbf{P}(E)\mathbf{P}(F)$, ce qui montre que E et F ne sont pas indépendants. ⋆

2.3. Théorème des probabilités totales

Nous considerons un ensemble univers Ω et une partition $A_1, \dots, A_n, \dots$ de Ω. Une telle partition est aussi appelée un *système complet d'événements*. Dans ce contexte, on a le résultat suivant.

Théorème 2.3.1 (Théorème des probabilités totales)**.** *Soit Ω un ensemble et $A_1, \dots, A_n, \dots$ une partition de Ω. Alors pour tout $E \subset \Omega$,* $\mathbf{P}(E) = \sum_{n \in \mathbb{N}} \mathbf{P}(E \cap A_n)$.

Exemple 2.3.1. Soit $E = \{3,4\}$ pour le lancer d'un dé, et soient $P =$ « résultat pair », $I =$ « résultat impair ». On a que $P \cup I = \Omega = \{1,2,3,4,5,6\}$, donc P, I est un système complet d'événements. On obtient par le théorème des probabilités totales que

$$\mathbf{P}(E) = \mathbf{P}(E \cap P) + \mathbf{P}(E \cap I) = \mathbf{P}(\{4\}) + \mathbf{P}(\{3\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

⋆

2.4. Probabilités conditionnelles

Le contexte de cette section est le suivant : nous considerons un événement F appartenant à une $\sigma-$algèbre $\mathscr{C}$ sur un univers Ω. Nous souhaitons déduire la probabilité d'un autre événement E en sachant que l'événement F est réalisé. On appelle donc la *probabilité conditionnelle* (ou *de Bayes*) *de E par rapport à F* la valeur

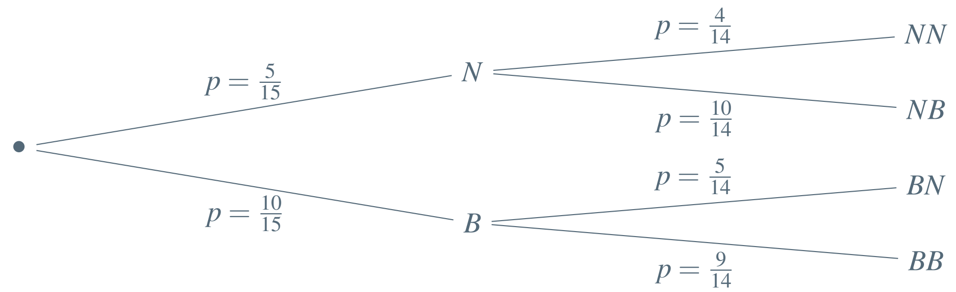
$$\mathbf{P}(E|F) = \frac{\mathbf{P}(E\cap F)}{\mathbf{P}(F)}.$$

Le Théorème des probabilités totales (Théorème 2.3.1) peut donc se reécrire de la manière suivante.

Théorème 2.4.1 (Théorème des probabilités totales, version deux)**.** *Soit Ω un ensemble et $A_1,\ldots,A_n,\ldots$ une partition de Ω. Alors pour tout $E\subset\Omega$, $\mathbf{P}(E) = \sum_{n\in\mathbb{N}} \mathbf{P}(E|A_n)\mathbf{P}(A_n)$.*

Nous donnons une application du Théorème 2.4.1 dans l'exemple qui suit.

Exemple 2.4.1. Une urne contient 5 boules noires et 10 boules blanches. On tire deux boules sans remise. Quel est la probabilité d'obtenir une boule blanche au deuxième tirage ? Il faut considerer l'*arbre des événements* suivant, qui correspond au tirage de deux boules sans remise :



On est interessés à l'événement B_2 = « blanche au deuxième tirage », donc aux événements BB ou NB. Les événements N_1 et B_1 (resp. « noire au premier tirage » et « blanche au premier tirage ») constituent un système complet d'événements. On applique donc la formule des probabilités totales (version conditionnelle) :

$$\mathbf{P}(B_2) = \mathbf{P}(B_2|B_1)\mathbf{P}(B_1) + \mathbf{P}(B_2|N_1)\mathbf{P}(N_1) = \frac{9}{14}\frac{10}{15} + \frac{10}{14}\frac{5}{15} = \frac{140}{210} = \frac{2}{3}.$$

⋆

On peut mettre en rélation la probabilité conditionnelle de E par rapport F avec la probabilité conditionnelle de F par rapport E avec la Formule de Bayes.

Théorème 2.4.2 (Formule de Bayes)**.** *Pour des événements $E,F\in\mathscr{C}$:*

$$\frac{\mathbf{P}(E|F)}{\mathbf{P}(F|E)} = \frac{\mathbf{P}(E)}{\mathbf{P}(F)}$$

Théorème 2.4.3 (Formules de Bayes généralisée)**.** *Soit $\{B_1,B_2,\ldots\}$ une partition dénombrable[5] d'un ensemble univers Ω. Soit $A\subset\Omega$ avec $\mathbf{P}(A)>0$. Alors pour tout i*

$$\mathbf{P}(B_i|A) = \frac{\mathbf{P}(A|B_i)\mathbf{P}(B_i)}{\sum_j \mathbf{P}(A|B_j)\mathbf{P}(B_j)}$$

5. Nous rappélons qu'un ensemble S est dénombrable s'il existe une partie $K\subset\mathbb{N}$ et une fonction bijective $\varphi : S\to K$. Par exemple $\mathbb{N}$ est dénombrable, une partie de $\mathbb{N}$ est dénombrable, $\mathbb{Q}$ est dénombrable, $\mathbb{R}$ n'est pas dénombrable.

2.5. Exercices du Chapitre II

1. (Foata, Fuchs) Soient A, B, C trois événements. Exprimer en fonction de A, B, C et des opérations ensemblistes les événements ci-après :
   — A seul se produit ;
   — A et C se produisent, mais non B ;
   — les trois événements se produisent ;
   — l'un au moins des événements se produit ;
   — deux événements au moins se produisent ;
   — un événement au plus se produit ;
   — aucun des trois événements ne se produit ;
   — deux événements exactement se produisent ;
   — pas plus de deux événements ne se produisent.

2. Soit $F \subset \Omega$ un événement fixé. Montrer que la fonction $E \mapsto \mathbf{P}(E|F)$ définit bien une loi de probabilité.

3. D'un jeu de 52 cartes (13 cartes de chacune des couleurs $\diamondsuit$ $\heartsuit$ $\clubsuit$ and $\spadesuit$), on tire 5 cartes sans remise. Quelle est la probabilité de tirer
   (a) 5 coeurs ? 2 piques et 3 coeurs ? 5 trèfles ou 5 coeurs ?
   (b) 5 cartes de la même couleur ?
   (c) 3 cartes d'une couleur et 2 d'une autre ?
   (d) les 4 as et une autre carte ?

4. Supposons de disposer d'un dé régulier (faces équiprobables) avec 9 faces : une face pour les nombres $1, 2, 3$ et deux faces pour les nombres $4, 5, 6$. On lance le dé une fois.
   a. Décrire un espace probabilisé correspondant à cette experience aléatoire.
   b. Quelle est la probabilité d'obtenir un nombre impair ? un nombre pair ?

5. *Paradoxe des anniversaires.* On cherche à estimer le nombre de personnes que l'on doit réunir pour que la probabilité que deux personnes de ce groupe aient leur anniversaire le même jour soit un certain nombre p fixé à priori.
   — Dans un groupe de TD de Probabilité à l'Université de Limoges, il y a 30 étudiants, tous nés en 1999 (année non bissextile, donc de 365 jours). Quelle est la probabilité de l'événement $E =$ « deux étudiants fêtent l'anniversailre le même jour » ?
   — Donner une formule pour la probabilité de l'événement E pour un groupe de k étudiants.
   — Quelle est la taille d'un groupe de TD pour que $\mathbf{P}(E) \geq 0.99$.

6. Un comité de 5 personnes doit être choisi parmi 20 hommes et 5 femmes, quelle est la probabilité
   — qu'il se compose de 5 femmes ?
   — qu'il se compose de 4 hommes et 1 femme ?

7. On jette une pièce non pipée. Si on obtient pile, on tire une boule dans une urne U_1 contenant 1 boule blanche et 2 boules noires. Si on obtient face, on tire une boule dans l'urne U_2 contenant 3 boules blanches et 2 boules noires.
   — Decrire l'arbre des événements et le probabiliser.
   — Quelle est la probabilité de tirer une boule noire ?
8. Soit E l'événement « il pleut » et F l'événement « je prends mon parapluie ».
   — Decrire les événements $F|E$ et $E|(F^c)$.
   — Je suis les prévisions météo, qui sont fiables au 75% (c'est-à-dire, que elle donnent la réponse correcte 3 fois sur 4). Si la probabilité de pluie est de deux tiers
     — quelle est la probabilité que je prenne mon parapluie ?
     — quelle est la probabilité que je prenne la pluie ?
9. On jette une pièce non pipée. Si on obtient pile, on tire une boule dans une urne U_1 contenant 1 boule blanche et 2 boules noires. Si on obtient face, on tire une boule dans l'urne U_2 contenant 3 boules blanches et 2 boules noires.
   — Decrire l'arbre des événements et le probabiliser.
   — Quelle est la probabilité de tirer une boule noire ?

Chapitre III

Variables aléatoires

Nous allons définir un objet d'intérêt fondamental en probabilités (et en statistiques), qui nous permettra de modéliser certaines expériences aléatoires et d'en déduir le comportement moyen et asymptotique. Le contexte général est encore donné par une épreuve aléatoire éventuellement répétée plusieures fois.

3.1. Variables aléatoires

Soit Ω l'ensemble univers d'une expérience aléatoire donnée. On veut associer à chacun des événements élémentaires une valeur réelle, et ensuite mésurer la probabilité d'obtenir une telle valeur. Nous rappelons que Ω est dénombrable si ses éléments peuvent être indexés par une suite de nombres entiers (ou plus formellement, s'il existe une bijecton de Ω vers un sous-ensemble de $\mathbb{N}$).

Definition 3.1.1. *Soit Ω l'ensemble univers associé à une épreuve aléatoire donnée. Une* variable aléatoire (réelle) *est une fonction $X\colon \Omega \to \mathbb{R}$.*

Donc une variable aléatoire associe un nombre réel $x \in \mathbb{R}$ à un évenement élémentaire $\omega \in \Omega$: en symboles, $X(\omega) = x$. Si $x \in \mathbb{R}$, on note compactement $X = x$ l'événement « X prends valeur x » (formellement, l'événement « il existe $\omega \in \Omega$ satisfaisant $X(\omega) = x$ »). L'ensemble $X(\Omega) \subset \mathbb{R}$, c'est-à-dire, l'image de X, contient toutes les valeurs prises par X.

Exemple 3.1.1. Soit X le nombre de faces obtenues lors d'un lancer de 3 pièces de monnaie (où lors de trois lancers consécutifs d'une même pièce). Donc X peut prendre les valeurs entre 0 et 3. Ici $\Omega = \{3F, 2F+P, F+2P, 3P\}$, $X(3F) = 3, X(2F+P) = 2, X(F+2P) = 1, X(3P) = 0$ et $X(\Omega) = \{0,1,2,3\}$. ⋆

Exemple 3.1.2. Soit X le nombre de lancers d'un dé (régulier, six faces) nécessaires pour obtenir ⚅. Donc X peut prendre une valeur (entière) arbitraire supérieur ou égal à 1. Ici l'ensemble univers est donné par

$$\Omega = \left\{\text{⚅}, \text{⚀⚅}, \text{⚁⚅}, \dots, \text{⚄⚅}, \text{⚀⚀⚅}, \text{⚀⚁⚅}, \dots, \text{⚄⚄⚅}, \dots\right\}$$

et par exemple

$$X\left(\text{⚅}\right) = 1,\ X\left(\text{⚀⚅}\right) = X\left(\text{⚁⚅}\right) = 2,\ X\left(\text{⚁⚁⚂⚀⚅}\right) = 5, \dots.$$

On remarque que ici $X(\Omega) = \mathbb{N}$. ⋆

Exemple 3.1.3. Une mesure physique quelconque (ex. longueur, capacité, poids, taille...) peut être considérée une variable aléatoire réelle. En effet, même si classiquement la valeur d'une mesure est une valeur *déterministe* (la longueur d'un table est une et une seule) le point de vue de la théorie des probabilités implique que une mesure est une valeur *aléatoire*, donc tout à fait une variable aléatoire : pour obtenir la longueur d'un table, on peut penser de demander à 1000 personnes de le mesurer, et après de calculer – par exemple – la moyenne des résultats obtenus (ou une autre fonction raisonnable de ces résultats). ⋆

3.2. Loi de probabilité d'une variable aléatoire

Soit $(\Omega, \mathscr{C}, \mathbf{P})$ un ensemble probabilisé, qu'on notera dorenavant simplement $(\Omega, \mathbf{P})$, en fixant la $\sigma-$algèbre des événements à $\mathscr{C} = \mathscr{P}(\Omega)$.

Definition 3.2.1. *La* loi de probabilité *d'une variable aléatoire X sur un espace probabilisé $(\Omega, \mathbf{P})$ est donnée par*

- *L'ensemble des valeurs prises par X, c'est-à-dire $X(\Omega)$*
- *Les probabilités attribuées à chacune des valeurs prises par X, c'est-à-dire la fonction $x \mapsto P(X = x) := \mathbf{P}(X^{-1}(x))$.*

Pour une variable aléatoire définie sur un ensemble fini Ω, nous répresenterons sa loi de probabilité sous la forme d'un tableau à deux lignes :

$x \in X(\Omega)$	x_1	x_2	x_3	x_4	$\cdots$	x_n
$P(X = x)$	p_1	p_2	p_3	p_4	$\cdots$	p_n

Exemple 3.2.1. Considerons la série de lancers (indépendants) d'un dé régulier à six faces, de l'Exemple 3.1.2, donc X représent le nombre de lancers pour obtenir ⚅. On obtient

$$P(X = 1) = \mathbf{P}(X^{-1}(1)) = \mathbf{P}(\{\text{⚅}\}) = \frac{1}{6}$$

$$P(X = 2) = \mathbf{P}(X^{-1}(2)) = \mathbf{P}(\{\text{⚀⚅}, \ldots, \text{⚄⚅}\}) = \frac{5}{6} \times \frac{1}{6} = \frac{5}{36},$$

et plus généralement $P(X = k) = (\frac{5}{6})^{k-1}\frac{1}{6}$. Nous obtenons pour X la loi de probabilité suivante :

$x \in X(\Omega) = \mathbb{N}$	1	2	3	$\cdots$	k	$\cdots$
$P(X = x)$	$\frac{1}{6}$	$\frac{5}{6} \times \frac{1}{6}$	$(\frac{5}{6})^2 \times \frac{1}{6}$	$\cdots$	$(\frac{5}{6})^{k-1} \times \frac{1}{6}$	$\cdots$

⋆

3.3. Fonction de répartition

Nous rappelons que si X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathbf{P})$, on peut définir une loi de probabilité $P(X = x) := \mathbf{P}(X^{-1}(x))$. Nous allons definir la fonction *cumulée* de P, qu'on appèlle classiquement fonction de répartition de la variable X.

Nous utilisons typiquement ces notations

$$P(X < x) := \mathbf{P}(X^{-1}((-\infty, x))) \text{ et}$$

$$P(X \leq x) := \mathbf{P}(X^{-1}((-\infty, x]))$$

Definition 3.3.1. *La* fonction de répartition *d'une variable aléatoire X sur un espace probabilisé $(\Omega, \mathbf{P})$ est l'application $F_X : \mathbb{R} \to [0, 1]$ définie*[6] *par $F_X(x) = P(X \leq x)$.*

6. Parfois dans la littérature la fonction de répartition est définie avec l'inferieur strict, voir en bas la différence pour des variables discrètes. Dans ce cours, nous utilisons bien la version plus traditionnelle avec l'inferieur non strict.

Autrement dit $F_X(x) = \mathbf{P}(X^{-1}((-\infty, x]))$, donc elle donne la probabilité que la variable X prenne une valeur inférieure ou égale à x.

Remarquez que dans la définition précedente, x est un nombre réel quelconque. Voici la répresentation compacte de la fonction de répartition d'une variable aléatoire définie sur un ensemble fini (*cf.* Figure 1.6), qui souligne bien le fait que F_X est la fonction cumulée de P :

i	1	2	3	$\cdots$	n
$P(X=i)$	p_1	p_2	p_3	$\cdots$	p_n
$F_X(i)$	p_1	p_1+p_2	$p_1+p_2+p_3$	$\cdots$	$\sum_{j=1}^{n} p_j$

Mais attention : F_X est bien définie sur tout $\mathbb{R}$. Par exemple, pour cet exemple, nous avons

$$F_X(3.5) = P(X \leq 3.5) = P(X \leq 3) = F_X(3) = p_1 + p_2 + p_3$$

$$F_X(n+5) = P(X \leq n+5) = p_1 + p_2 + \cdots + p_n = 1$$
$$F_X(-8) = P(X \leq -8) = 0$$

Voici en Figure 1.6 deux exemples de loi de probabilité avec sa fonction de répartition : en haut, une loi *uniforme* sur $\{1, \ldots, 6\}$ (à gauche) pour le lancer d'un dé régulier avec sa fonction de répartition linéaire par morceaux (à droite), en bas un deuxième exemple.

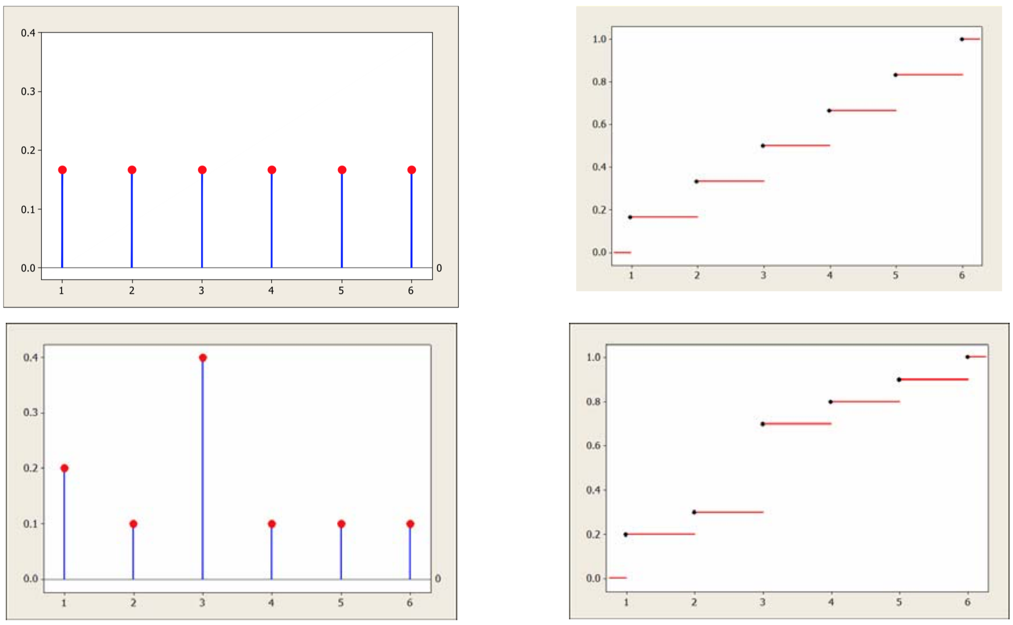


FIGURE 1.6. Densités et fonctions de répartition dans le cas fini

Voici quelques propriétés de la fonction de répartition :

— $P(a < X \leq b) = F_X(b) - F_X(a)$.
— F_X est continue à gauche, c'est-à-dire $F_X(x) \to F_X(x_0)$ si $x \to x_0^-$.
— F_X est non-décroissante.
— Si $x \to +\infty$, alors $F_x(x) \to 1^-$.
— Si $x \to -\infty$, alors $F_x(x) \to 0^+$.

3.4. Variables aléatoires discrètes

Nous rappelons que un ensemble S est dit dénombrable si S est en bijection avec une partie de $\mathbb{N}$. En particulier, tout enseble fini est dénombrable, $\mathbb{N}$ et $\mathbb{Q}$ sont dénombrables, alors que $\mathbb{R}$ ou un intervalle réel (a,b) ne sont pas dénombrables.

Definition 3.4.1. *Une variable aléatoire $X\colon \Omega \to \mathbb{R}$ est dite* discrète *si $X(\Omega)$ est dénombrable.*

Les Exemples 3.1.1 et 3.1.2 correspondent à des variables discrètes. Une variable définie sur un ensemble fini est discrète. Remarquez l'identité suivante concernant la loi de probabilité de la variable alétoire de l'Exemple 3.1.2 :

$$P(X \in \mathbb{N}) = P(X = 1 \text{ ou } X = 2 \text{ ou } ...) = \sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} \left(\frac{5}{6}\right)^{k-1} \frac{1}{6} = \frac{1}{6} \frac{1}{1-\frac{5}{6}} = 1.$$

Nous avons utilisée la convergence de la série géométrique rappelée en Appendice. Plus généralement, si P est une loi de probabilité définie sur une variable discrète X, la série de terme générale $P(X = x_i)$, $x_i \in X(\Omega)$, est convergente, et la limite vaut 1, c'est-à-dire $\sum_i P(X = x_i) = 1$.

Pour une variable aléatoire discrète X, la fonction $x \mapsto P(X = x)$ est souvent appelée *densité discrète* de X, à ne pas confondre avec la densité d'une probabilité continue (voir en bas). De même, $F_X(x) = \sum_{y \leq x} P(X = y)$ est une fonction de répartition pour X. Il faut noter que, dans le cas discret, il y a une différence entre $P(X < x)$ et $F_x(x) = P(X \leq x)$ (voir exemple ci-dessous).

Exemple 3.4.1. Soit X une variable aléatoire discrète, avec $X(\Omega) = \mathbb{N}$. Alors pour tout $n \in \mathbb{N}$ la valeur de la fonction de répartition de X vérifie

$$F_X(n) = P(X \leq n) = \sum_{k=0}^{n} P(X = k) = P(X < n+1)$$

et pour tout $x \in \mathbb{R}$:

$$F_X(x) = P(X \leq x) = \sum_{k=0}^{\lfloor x \rfloor} P(X = k)$$

et

$$P(X < x) = \sum_{k=0}^{\lceil x-1 \rceil} P(X = k).$$

⋆

3.5. Variables aléatoires continues

Une variable aléatoire qui n'est pas discrète est dite *continue*. Pour une variable continue X l'image $X(\Omega)$ n'est donc pas un ensemble dénombrable (il est de cardinal strictement supérieur [7] au cardinal de l'ensemble des nombres entiers).

Dans l'exemple ci-dessous, nous décrivons des exemples simples de variables continues.

7. Quand on dit que il n'y a pas d'autres types de variables aléatoires, si on exclut celles à valeurs dans un ensemble discret et celles à valeurs dans un ensemble continu, nous sommes en train d'utiliser implicitement l'*hypothèse du continu*.

Exemple 3.5.1. Exemple de variables aléatoires continues :

— Une mesure physique (voir Example 3.1.3).

— $\Omega = \{[a,b] : a,b \in \mathbb{R}\}$, avec $X([a,b]) = b-a$, c-à-d., X mesure la longueur de l'intervalle $[a,b]$. Ici nous avons $X(\Omega) = \mathbb{R}_{\geq 0}$.

— le temps d'attente d'un bus est une variable aléatoire continue X à valeurs dans un intervalle, par exemple $X(\Omega) = [0,10]$ (entre 0 et 10 minutes, si on est optimistes).

⋆

Si X est continue il n'y aura aucune différence entre $F_X(x) = P(X < x)$ et $P(X \leq x)$. On peut associer à chaque variable aléatoire continue une fonctions appelés *densité* comme suit.

Théorème 3.5.1. *Soit $f\colon \mathbb{R} \to \mathbb{R}$ une fonction intégrable, telle que*

— $f(t) \geq 0$ *pour tout* $t \in \mathbb{R}$

— $\displaystyle\int_{-\infty}^{+\infty} f(t)\,dt = 1.$

Alors $F(x) = \int_{-\infty}^{x} f(t)\,dt$ est la fonction de répartition d'une variable continue X.

La fonction f du théorème précedent est dite *densité* de X, et noté avec le symbole f_X. La fonction de répartition et la densité d'une variable continue X satisfont donc la propriété suivante :

$$\boxed{F_X(x) = \int_{-\infty}^{x} f_X(t)dt}$$

En particulier, nous déduisons la propriété suivante pour une variable continue X :

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) =$$

$$= F_X(b) - F_X(a) = \int_a^b f_X(t)dt.$$

3.6. Characteristiques d'une variable aléatoire

Nous introduisons des valeurs qu'on peut associer à chaque variable aléatoire (discrète ou continue) et qui ont un interêt notamment en statistiques descriptive. Soit X une variable aléatoire réelle.

Fractiles. Soit $\alpha \in (0,1)$. Nous nous intéressons à la valeur de X tels que la probabilité d'observer une telle valeur ou une valeur plus petite, est egale à α. Plus précisement, le *fractile (ou quantile) d'ordre* α de X est la valeur $x_\alpha \in \mathbb{R}$ telle que

$$F_X(x_\alpha) = \alpha.$$

Autrement dit, x_α est l'image inverse [8] de α par la fonction de répartition de X, c'est-à-dire, $x_\alpha := F_X^{-1}(\alpha)$. Des exemples classiques de fractiles sont :

— la *médiane* $x_{\frac{1}{2}}$, c'est à dire le fractile d'ordre $\alpha = \frac{1}{2}$. Le nom médiane est justifié par le fait que $x_{\frac{1}{2}}$ "divise en deux" l'ensemble des valeurs prises par X.

8. Le calcul du fractile passe donc, essentiellement, par le calcul de la fonction inverse de la fonction de répartition. La fonction inverse des distributions de probabilités le plus courants est implementé dans les calculatrices scientifiques ainsi que dans des logiciels comme Excel ou OpenOffice.

— le *premier quartile* $x_{\frac{1}{4}}$, et le *troisième* quartile $x_{\frac{3}{4}}$; évidemment le deuxième quartile coincide avec la médiane : $x_{\frac{2}{4}} = x_{\frac{1}{2}}$. Les quartiles divisent en quatre l'image de X.

— les *déciles* $x_{\frac{1}{10}}, \ldots, x_{\frac{9}{10}}$, définis pareil, avec $\alpha = k\frac{1}{10}, k = 1, \ldots, 9$.

— les *centiles* ou *percentiles*, utilisés notamment dans les statistiques descriptives du poids ou de la taille des enfants (voir Figure 1.7), ce sont les fractiles d'ordre multiple de $\frac{1}{100}$.

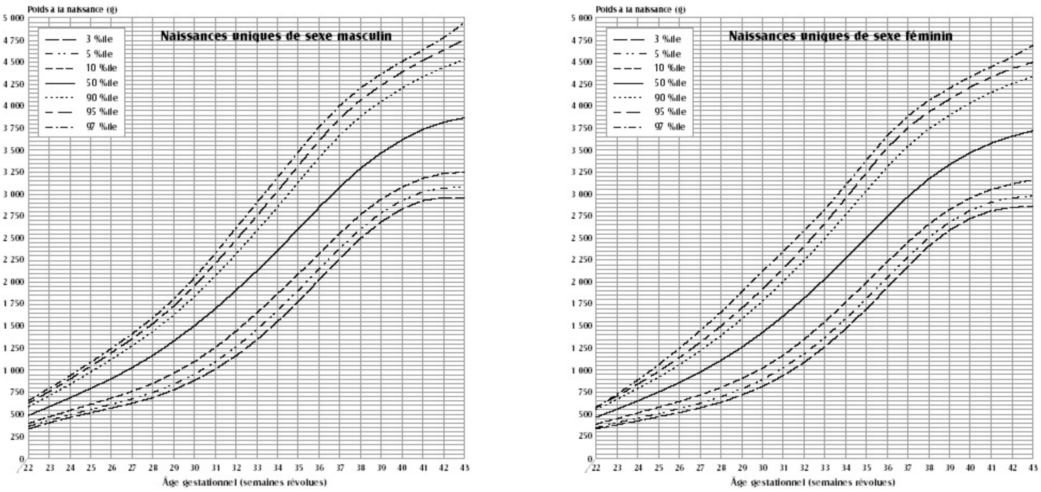


FIGURE 1.7. Percentiles du poids à la naissance pour un nouveau né (masculin et feminin)

Mode. La *mode* de X est la valeur de plus grande probabilité. Si X est discrète, cela correspond à la valeur x telle que $P(X = x)$ est maximum. Si X est continue, la mode est la valeur x qui maximise $f_X(x)$.

Espérance mathématique. Le mot *espérance* est motivé par son usage dans l'estimation de l'équitabilité des jeux de hasard. Partons d'un exemple.

Exemple 3.6.1. Nous nous démandons si le jeu suivant est équitable. On vous démande d'y participer en payant une somme de t€, pour un $t \geq 0$ donné (par exemple, $t = 0$ signifie "gratuitement", alors que pour $t = 1$ le coût est de 1€) qui ne sera pas récupérée.
Nous lançons un dé régulier, six faces. Si le résultat est pair, vous gagnez 1€; si le résultat est 1 ou 3, vous gagnez 2€; sinon, vous perdez 3€. Êtes-vous disponibles à jouer ? C'est clair que ça va dependre du coût t. L'idée est de comparer t avec ce que nous *espérons* de gagner, c'est à dire (de manière naïve) :

$$(1€) \times \frac{1}{2} + (2€) \times \frac{1}{3} + (-3€) \times \frac{1}{6} = \frac{2}{3}€ \approx 67€ cents.$$

Donc on peut deduire que pour une mise t inférieure à $\frac{2}{3}$€ le jeu est *équitable*, sinon le gain moyen est inférieur à la mise, donc le jeu n'est pas équitable. ⋆

Definition 3.6.1. *Soit X une variable aléatoire discrète. L'espérance (mathématique) de X est la valeur*

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \times P(X = x)$$

lorsque la somme est convergente. En particulier, si $X(\Omega) = \{x_1, \ldots, x_n\}$ est fini, la somme converge et l'espérance vaut :

$$\mathbb{E}(X) = x_1 \times P(X = x_1) + \cdots + x_n \times P(X = x_n).$$

Si X est une variable continue, l'espérance de X est définie comme suit :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} (t \times f(t))\,dt.$$

Quelques propriétés de l'espérance :

— Cas constant. Si $c \in \mathbb{R}$, alors $\mathbb{E}(c) = c$. En effet, une "constante" est une variable aléatoire très speciale (sa valeur n'est pas vraiment aléatoire).

— Linéarité. Soient $c_1, \ldots, c_N \in \mathbb{R}$ et soient $X_1, \ldots, X_N$ des variables aléatoires d'espérance bien définie, alors

$$\mathbb{E}\left(\sum_{k=1}^{N} c_k X_k\right) = \sum_{k=1}^{N} c_k \mathbb{E}(X_k).$$

Théorème 3.6.1 (Somme aléatoire de variables aléatoires)**.** *Soit $\{X_i\}_{i\in\mathbb{N}}$ une famille dénombrable de variables aléatoires, de même espérance μ, et soit N une variable aléatoire discrète ($N(\Omega) = \mathbb{N}$), indépendante de chacune des X_k. Alors*

$$\mathbb{E}\left(\sum_{k=1}^{N} X_k\right) = \mu \mathbb{E}(N)$$

Théorème 3.6.2 (Inégalité de Schwartz probabiliste)**.** *Soient X, Y v.a. discrètes, telles que $\mathbb{E}(X^2), \mathbb{E}(Y^2)$ existent. Alors*

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$$

Théorème 3.6.3. *Soit X une variable aléatoire positive ($X(\Omega) \subset \mathbb{R}_+$). Alors*[9] *pour tout nombre réel $t > 0$*

$$\mathbf{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}$$

Variance et écart-type. Si l'espérance donne une idée de "moyenne" de la variable X, il existe des mesures de "variabilité autour de la moyenne".

Definition 3.6.2. *Soit X une variable aléatoire, d'espérance $\mu = \mathbb{E}(X)$. La variance de X est la valeur*

$$Var(X) = \mathbb{E}\left((X-\mu)^2\right) = \sum_{k \in X(\Omega)} (k-\mu)^2 \times P(X = k)$$

lorsque la somme est convergente. En particulier, si $X(\Omega) = \{k_1, \ldots, k_n\}$ est fini, la somme converge et la variance vaut :

$$Var(X) = (k_1-\mu)^2 \times P(X = k_1) + \cdots + (k_n-\mu)^2 \times P(X = k_n).$$

9. Un corollaire direct du Théorème 3.6.3 est :

$$\mathbf{P}(X < t) = 1 - \mathbf{P}(X \geq t) \geq 1 - \frac{\mathbb{E}(X)}{t}.$$

Si X est une variable continue, la variance de X est la valeur définie comme suit :

$$Var(X) = \int_{-\infty}^{+\infty} ((t-\mu)^2 \times f(t))\, dt$$

si l'integral est convergent. L'écart-type est la racine carrée de la variance, c'est-à-dire :

$$\sigma(X) = \sqrt{Var(X)}.$$

Voici des propriétés de la variance (voir aussi les exercices à page 34). Soit X une variable aléatoire réelle admettant une variance.

— $Var(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2$.
— $Var(X) = \mathbb{E}(X(X-1)) - \mathbb{E}(X)(\mathbb{E}(X)-1)$.
— $Var(aX+b) = Var(aX) = a^2 Var(X)$ pour tout $a, b \in \mathbb{R}$.
— Si $\mathbb{E}(X) = \mu \in \mathbb{R}$ et $Var(X) = \sigma^2 > 0$, la variable aléatoire $Y = \frac{X-\mu}{\sigma}$ a espérance nulle et variance egale à 1.
— $Var(X) = \min_{a \in \mathbb{R}} \mathbb{E}\left((X-a)^2\right)$

L'inegalité suivante donne une borne supérieure pour la probabilité que une variable aléatoire X s'écarte de sa moyenne d'une valeure fixée :

Théorème 3.6.4 (Inegalité de Bienaymé-Tchebychev)**.** *Soit X une variable aléatoire. Alors pour tout nombre réel $t \geq 0$*

$$\mathbf{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{Var(X)}{t^2}$$

3.7. Fonctions d'une variable aléatoire

Une v.a. est en particulier une fonction, donc on peut *a priori* la composer avec une autre fonction. Nous allons pour l'instant considérer la composition d'une v.a. avec une fonction déterministe.

Théorème 3.7.1. *Soit X une variable aléatoire et soit g une fonction continue* [10] *définie sur $X(\Omega)$. Alors $g(X)$ est une variable aléatoire.*

10. Le même résultat est valable pour une fonction mesurable au sens de Borel (*cf.* cours de topologie).

3.8. Exercices du Chapitre III

1. On lance trois fois de suite une pièce de monnaie équilibrée. On gagne 2 euros pour chaque résultat « pile » et on perd 1 euro pour chaque résultat « face ».
   (a) Quel est l'ensemble Ω des résultats possibles ?
   (b) Soit $X: \Omega \to \mathbb{R}$ l'application qui associe le gain à chaque événement. Quelles sont les valeurs prises par X ?
   (c) Quelle est la probabilité de gagner 3 euros ? De ne rien gagner ?
2. On lance 2 dés equilibrés à 6 faces et on s'intéresse à la somme des résultats. Decrire une variable aléatoire X (donc un ensemble univers Ω, une fonction $X: \Omega \to \mathbb{R}$ et l'ensemble $X(\Omega)$) qui représente cette expérience aléatoire. Est-elle discrète ou continue ? Calculer l'espérance et la mode de X. Répeter l'exercice avec deux dés equilibrés à n faces.)
3. Une urne contient 6 boules blanches et n boules rouges, et on tire deux boules au même temps, en gagnant 2 euros à chaque boule blanche tirée, en perdant 3 euros à chaque boule rouge tirée.
   — Decrire une variable aléatoire associée à cette expérience aléatoire.
   — Pour quelles valeurs de n seriez-vous disponibles à jouer ? Pour quelle valeur le joueur ne gagne et ne perd rien en moyenne ?
4. **Roulette française.** Dans le jeu de la roulette une petite bille est lancée sur un plateau concave, et elle tombe sur un cylindre numeroté de 0 et 36, qui tourne en sens inverse. Il y a 18 nombres rouges et 18 noirs, le 0 est vert. Le croupier vous paye le montant exact de votre mise si vous devinez la couleur (c'est-à-dire, si vous misez 1€, il vous donne 1€ en plus, donc vous sortez du casino avec 2€) et il vous paye 35 fois la mise si vous devinez le nombre exact (donc vous sortez avec 36€). On se demande si la roulette est un jeu équitable.
   — Parieriez-vous sur le rouge ? Parieriez-vous sur le jour de votre naissance ?
   — En supposant que vous disposiez d'un portefeuille illimité, et que le casinó ne prevoit aucun limite supérieur pour les mises, quelle stratégie proposeriez-vous pour avoir des chances de gagner ?

   Répeter l'exercice pour la roulette americaine, qui contient aussi le 00 vert.)
5. **Saint-Pétersbourg.** Supposons de participer au jeu suivant : On lance une pièce, et on gagne la première fois que on obtient pile. Le gain est donné par ce tableau [11] :

apparition première P	1	2	3	4	5	$\cdots$	k	$\cdots$
gain (€)	2	1	$\frac{8}{9}$	1	$\frac{32}{25}$	$\cdots$	$\frac{2^k}{k^2}$	$\cdots$

   Pour participer au jeu, on nous demande de payer un montant initial fixe (qui ne sera pas rendu en cas de victoire). Seriez-vous disponibles à payer 2€ pour participer au jeu ?
6. Soit X une variable aléatoire réelle admettant une variance.

11. Par exemple : on gagne 2€ si on obtient P au premier lancer, on gagne $\frac{32}{25}$€ si la première pile apparaît au cinquième lancer (c-à-d. on lance 4 fois F avant d'obtenir P).

— Montrer que $Var(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2$.
— Montrer que $Var(X) = \mathbb{E}(X(X-1)) - \mathbb{E}(X)(\mathbb{E}(X)-1)$.
— Montrer que $Var(aX+b) = Var(aX) = a^2 Var(X)$ pour tout $a, b \in \mathbb{R}$.
— Si $\mathbb{E}(X) = \mu \in \mathbb{R}$ et $Var(X) = \sigma^2 > 0$, calculer l'espérance et la variance de $Y = \frac{X-\mu}{\sigma}$.
— Montrer que $Var(X) = \min_{a \in \mathbb{R}} \mathbb{E}\left((X-a)^2\right)$

7. Soient X et Y deux variables aléatoires réelles, admettant une variance. On suppose que X et Y sont indépendantes. Montrer que $Var(X+Y) = Var(X) + Var(Y)$. Donner ensuite un exemple de deux variables aléatoires X et Y telles que $Var(X+Y) \neq Var(X) + Var(Y)$.

Chapitre IV

Lois discrètes usuelles

Dans le chapitre précedent, nous avons défini les variables aléatoires. Nous allons maintenant définir des classes de variables aléatoires, appélées *lois* ou *distributions* de probabilité. Typiquement ces familles de v.a. dépendent de certains *paramètres* réels.

L'idée principale est que deux variables aléatoires différentes peuvent *suivre* la même distribution, et même avec les mêmes paramètres, tout en gardant leur différence (et peut être leur indépendance). Nous allons discuter des familles classiques de variables aléatoires qu'on retrouve dans plusieurs contextes applicatifs.

4.1. Loi uniforme discrète $\mathscr{U}(n)$

Le premier exemple est le cas simple d'une variable aléatoire où les valeurs observés ont la même probabilité. Soit donc $\Omega = \{\omega_1, \ldots, \omega_n\}$ un ensemble fini, de cardinal n, et soit $E = \{x_1, \ldots, x_n\} \subset \mathbb{R}$. Soit $X : \Omega \to \mathbb{R}$ la variable aléatoire telle que $X(\omega_i) = x_i, \quad i = 1, \ldots, n$ avec loi de probabilité

x	x_1	x_2	x_3	$\cdots$	x_n
$P(X = x)$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\cdots$	$\frac{1}{n}$

On dit que X suit une *loi uniforme (discrète)* sur $E = \{x_1, \ldots, x_n\}$, en symbole $X \sim \mathscr{U}(E)$, ou plus souvent simplement $X \sim \mathscr{U}(n)$, car E sera identifié avec $\{1, \ldots, n\}$

On remarque que la fonction de répartition de X est donnée par

x	x_1	x_2	x_3	$\cdots$	x_n
$F_X(x) = P(X \leq x)$	$\frac{1}{n}$	$\frac{2}{n}$	$\frac{3}{n}$	$\cdots$	1

L'espérance est bien la moyenne arithmétique des valeurs prises par X, c'est-à-dire

$$\mathbb{E}(X) = x_1 \frac{1}{n} + \cdots + x_n \frac{1}{n} = \frac{x_1 + \cdots + x_n}{n}.$$

Le calcul de la variance est aussi très direct :

$$Var(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2 = \sum_i \frac{1}{n} x_i^2 - \left(\frac{x_1 + \cdots + x_n}{n}\right)^2$$

Par exemple, en supposant que $x_i = i$, pour $i = 1, 2, \ldots, n$, on obtient $Var(X) = \frac{n^2-1}{12}$.

Exemple 4.1.1. Nous lançons un dé à 6 faces équiprobables, soit X le résultat. Donc $X \sim \mathscr{U}(6)$. La loi de probabilité ainsi que la fonction de répartition de X sont données par

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$F_X(x)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{5}{6}$	1

Remarquez que la probabilité d'obtenir un résultat inferieur ou égal à 4 est $F_X(4) = \frac{2}{3}$ et supérieur ou égal à 2 est $1 - F_X(1) = \frac{5}{6}$. ⋆

Nous pouvons facilement remarquer que une variable discrète *infinie* ne peut pas suivre une loi uniforme. Supposons en fait que ça soit possible : nous aurions une variable X telle que $X(\Omega) = \{x_i : i \in \mathbb{N}\}$, et telle que $P(X = x_i) = p$ pour tout i, pour un valeur donné de probabilité $p > 0$. On aurait donc une contradiction par l'égalité $1 = \sum_{n=0}^{+\infty} p = +\infty$.

4.2. Loi de Bernoulli $\mathscr{B}(p)$

La plus simple parmi les lois non-uniformes est sans doutes la loi de Bernoulli. Le contexte est une expérience aléatoire simple avec deux résultats possibles, qu'on peut appéler *succès* et *échec*. On identifie classiquement le succès avec 1 et l'échec avec 0, car il est sous-entendu qu'on privilégie la survenue d'un événement.

Plus formellement, soit Ω un ensemble (univers pour une épreuve fixée) et soit $A \subset \Omega$. On denote par $p = \mathbf{P}(A)$. On dit que une v.a. $X : \Omega \to \mathbb{R}$ *suit une loi de Bernoulli de paramètre* p (en symbole, $X \sim \mathscr{B}(p)$) si

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon.} \end{cases}$$

Donc $X(\Omega) = \{0, 1\}$. On vérifie bien que

$$P(X = 1) = \mathbf{P}(X^{-1}(1)) = \mathbf{P}(\{\omega \in \Omega : \omega \in A\}) = \mathbf{P}(A) = p$$

et donc que $P(X = 0) = 1 - P(X = 1) = 1 - p$.

On déduit que la loi de probabilité et la fonction de répartition de X sont données par

x	0	1
$P(X = x)$	$1 - p$	p
$F_X(x)$	$1 - p$	1

On a aussi que $\mathbb{E}(X) = p$ et $Var(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2 = \mathbb{E}(X) - \mathbb{E}(X)^2$ car $X = X^2$ pour une variable de Bernoulli (le vérifier) et donc $Var(X) = p(1 - p)$.

Exemple 4.2.1. Quelques exemples classiques d'épreuve de Bernoulli :

— lancer d'une pièce de monnaie pile/face. On fait correspondre le succès à $A =$ « pile », donc $p = 1/2$ si le dé est régulier.

— lancer d'un dé régulier à 6 faces, et $A = \{1,4\}$; ici $p = 1/3$ et le succès corréspond à l'événement « le résultat est 1 ou 4 ».

— production industrielle d'un objet, où l'on connait (ou l'on peut estimer) la probabilité p que l'objet soit endommagé ; dans ce cas le succès est, par exemple, $A =$ « l'objet est endommagé ».

⋆

4.3. Loi binomiale $\mathscr{B}(n,p)$

Dans les chapitres précedents nous avons souvent considéré des expériences aléatoires *répétée*. La distribution classiquement associée à la répétition pour n fois d'une même épreuve de Bernoulli de paramètre p, est dite *binomiale*.

Plus précisement, soient $X_1, X_2, \ldots, X_n$ des variables aléatoires suivant une loi de Bernoulli *de même paramètre* p, et *deux à deux indépendantes* (X_i et X_j sont indépendantes, pour tout $i \neq j$). Alors la variable

$$Y = X_1 + X_2 + \ldots + X_n$$

suit une loi binomiale de paramètre n *et* p, en symbole $Y \sim \mathscr{B}(n,p)$. La v.a. Y *compte le nombre de succès* parmi les n répétitions d'une expérience aléatoire de type Bernoulli $\mathscr{B}(p)$. En particulier, remarquez que $\mathscr{B}(1,p) = \mathscr{B}(p)$.

Remarquons que si Ω est l'ensemble univers des X_i, alors l'ensemble univers de Y est bien

$$\Omega^n = \Omega \times \Omega \times \cdots \times \Omega$$

et l'image $Y(\Omega^n) = \{0,1,2,\ldots,n\}$. Par exemple, pour $n = 5$, la suite $(1,1,0,0,1)$ correspond à une expérience répétée 5 fois, avec 3 succès au premier, deuxième et cinquième tentatif, donc la valeur de Y est bien 3. Évidemment on peut avoir 3 succès distribués différemment (par exemple $(1,1,1,0,0)$ ou $(0,1,0,1,1)$ ou ...).

Théorème 4.3.1. *Soit $Y \sim \mathscr{B}(n,p)$. Alors $P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$.*

Démonstration. Soit A l'événement correspondant au succès pour X_i, et pour $k \in Y(\Omega^n)$ soit $B_k \subset \Omega^n$ l'événement « A s'est vérifié exactement k fois ». Alors les événements $B_0, B_1, B_2, \ldots, B_n$ forment un système complet d'événements, donc une partition, de Ω^n. L'image de B_k est representée par l'ensemble

$$Y(B_k) = \{(a_1, \ldots, a_n) \in \{0,1\}^n : \sum_{i=1}^{n} a_i = k\}$$

contenant les vecteurs de longueur n contenant k uns et $n-k$ zeros. C'est clair que pour obtenir un de ces vecteurs, il faut "choisir" k entrées du vecteur parmi n, sans répétitions et sans tenir compte de l'ordre, donc c'est des combinaisons sans répétition. On déduit que $|B_k| = C_{n,k} = \binom{n}{k}$. La probabilité de chacun des événements $\omega \in B_k$ est la même et egale à $p^k(1-p)^{n-k}$ car on veut exactement k succès et $n-k$ échecs. On conclut que

$$P(Y = k) = \mathbf{P}(B_k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

■

Quelques exemples de loi binomiale :

Exemple 4.3.1. En répétant l'experience pile/face de l'Exemple 4.2.1 pour 5 fois, on obtient une expérience de type binomial : soit Y le nombre de piles obtenus lors de 5 lancers, donc $Y \sim \mathscr{B}(5, \frac{1}{2})$. ⋆

Exemple 4.3.2. Dans l'Exemple 4.2.1 concernant les objets endommagés, en supposant que n objets sont produits de façon indépendante, la probabilité d'obtenir k objets endommagés parmi les n est donc de $\binom{n}{k} p^k (1-p)^{n-k}$. Typiquement pour ce type d'exemples $p \ll 1$ (c'est-à-dire p est une probabilité faible) et nous verrons plus tard que l'on peut aussi utiliser la variable de Poisson (voir Section 4.8). ⋆

La loi de probabilité de $Y \sim \mathscr{B}(n, p)$ donnée par le Théorème 4.3.1 est donc

y	0	1	2	$\cdots$	$n-1$	n
$P(Y=y)$	$(1-p)^n$	$np(1-p)^{n-1}$	$\binom{n}{2} p^2 (1-p)^{n-2}$	$\cdots$	$\binom{n}{n-1} p^{n-1}(1-p)$	p^n

L'esperance d'une v.a. $Y \sim \mathscr{B}(n, p)$ vaut

$$\mathbb{E}(Y) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = np$$

puisque pour le Théorème du binôme (Théorème 1.4.1) on a $\sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = (p+1-p)^{n-1} = 1$, mais aussi parce que $Y = \sum_i X_i$ et donc $\mathbb{E}(X) = \sum_i \mathbb{E}(X_i) = p + p + \cdots + p$. La variance vaut

$$Var(Y) = np(1-p)$$

puisque les X_i sont deux-à-deux indépendantes, et $Var(X_i) = p(1-p)$

Exemple 4.3.3 (Rényi). Dans une désintégration radioactive, le nombre X des atomes désintégrés pendant un intervalle de temps de longueur t a une distribution binomiale : la probabilité pour que exactement k atomes donnés se désintègrent pendant ce temps t est en effet $\binom{N}{k} p^k (1-p)^{N-k}$ où N est le nombre d'atomes existant au début de l'intervalle de temps, et $p = 1 - e^{-\lambda t}$ où λ est la constante de désintégration. Donc on voit bien que l'espérance $\mathbb{E}(X) = Np = N(1 - e^{-\lambda t})$ tends vers N. ⋆

Théorème 4.3.2 (Stabilité de la loi binomiale)**.** *Soient $X \sim \mathscr{B}(m, p)$ et $Y \sim \mathscr{B}(n, p)$, et supposons que X et Y sont indépendantes. Alors $X + Y \sim \mathscr{B}(m+n, p)$.*

Évidemment Théorème 4.3.2 ne peut pas s'étendre au cas de non-indépendance : si $X \sim \mathscr{B}(n, p)$ alors $X + X = 2X$ ne suit pas une loi $\mathscr{B}(2n, p)$ – car par exemple l'image de $2X$ est $\{2, 4, 6, \ldots, 2n\}$ et non pas $\{1, 2, \ldots, 2n\}$. Pour se convaincre, considerez le cas $n = 1$ donc d'une variable Bernoulli :

Exemple 4.3.4. Soit $X \sim \mathscr{B}(p) = \mathscr{B}(1, p)$. Alors $2X$ est donnée par

$X =$	0	1
$2X =$	0	2
$P(X = x) = P(2X = 2x) =$	$1-p$	p

et donc $2X \not\sim \mathscr{B}(2, p)$. La raison est que X n'est évidemment pas indépendant avec X. ⋆

4.4. Loi hypergéométrique $\mathscr{H}(N,n,p)$

Supposons de devoir resoudre le problème decrit dans l'exemple suivant.

Exemple 4.4.1. Nous disposons d'une population de taille N (par exemple N cellules, N personnes...) dont on sait que un percentage p d'individus possèdent un caractère fixé (par exemple $N = 100$ familles de Limoges, dont on sait que $30 = Np$ parmi elles ont trois enfants – ici $p = 3/10$).
On extrait un échantillon de n individus distincts (donc ici $n \leq N$). Quelle est la probabilité que exactement k parmi les n individus possèdent le caractère ? ⋆

Soit X la variable aléatoire qui compte le nombre d'individus possedant le caractère, parmi les n individus qui ont été chosis. La variable X suit une loi qu'on appèlle *hypergéométrique* de paramètres N,n,p et qu'on dénote par $X \sim \mathscr{H}(N,n,p)$.

En utilisant la combinatoire, on obtient que

$$\mathbb{P}(X = k) = \frac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}}.$$

En effet :
- Le nombre total de choix possibles est $\binom{N}{n}$, car c'est un "tirage sans remise" de n individus parmi N.
- Par rapport au nombre de cas favorables : il y a Np individus possedant le caractère, et il faut en choisir k ; pour chacun de ces choix, on doit chosir les autres $n-k$ parmi les $N(1-p)$ individus qui ne possèdent pas le caractère. Donc le nombre de cas favorables est bien le produit $\binom{Np}{k}\binom{N(1-p)}{n-k}$.

On voit enfin que l'esperance de X et sa variance valent

$$\mathbb{E}(X) = np \quad \text{et} \quad Var(X) = \frac{N-n}{N-1}np(1-p).$$

4.5. Loi géométrique $\mathscr{G}(p)$

La *loi géométrique* est la loi d'une variable aléatoire qui compte le nombre d'essais nécessaires pour que un événement de probabilité p se vérifie. Le contexte est donc le même que pour la loi binomiale, il s'agit d'une répétition d'épreuves de Bernoulli indépendantes, mais ici on veut compter le nombre de tirages necessaires pour avoir le premier succès.

Si X suit une loi géométrique, on écrit $X \sim \mathscr{G}(p)$. On a que $X(\Omega) = \{1,2,3,4,\dots\}$ et que par indépendance des épreuves :

$$P(X = k) = (1-p)^{k-1}p$$

car on aura comme événement favorable seulement le $k-$uplet $(0,0,0,0,\dots,0,1)$, où encore une fois $0 =$ échec et $1 =$ succès.

Un exercice simple qui utilise les proprietés des séries géométriques (d'où le nom de la distribution) est de montrer que $\mathbb{E}(X) = \frac{1}{p}$ et que $Var(X) = \frac{1-p}{p^2}$.

4.6. Loi de Pascal $\mathscr{P}(r,p)$

On se pose dans le même contexte de la section precedente, celle de la loi géométrique. Supposons que maintenant nous souhaitons compter le nombre d'essais nécessaires pour obtenir exactement r succès (dit autrement, pour obtenir le r-ème succès). La variables correspondante est la somme de r variables géométriques de même paramètre :

$$Y = X_1 + X_2 + \cdots + X_r, \quad X_i \sim \mathscr{G}(p).$$

En effet, la variable X_1 compte le nombre nécessaires k_1 d'essais pour obtenir le premier succès ; la X_2 compte le nombre nécessaires k_2 pour le deuxième succès, après les premiers k_1, donc après $k_1 + k_2$ essais on a deux succès exactement, et on continue ... jusqu'à obtenir exactement r succès, après $k = k_1 + k_2 + \cdots + k_r$ essais.

On voit bien que

$$P(Y = k) = \binom{r-1}{k-1} p^r (1-p)^{k-r}$$

et que $\mathbb{E}(Y) = \frac{r}{p}$ et $Var(Y) = r\frac{1-p}{p^2}$.

4.7. Loi de Poisson $\mathscr{P}(\lambda)$

La distribution de Poisson est parmi le plus communes et importantes lois du calcul des probabilités. Traitons d'abord un exemple.

Exemple 4.7.1 (Renyi). Une usine produit des bouteilles de verre, et le problème suivant se présente. Dans le procés de fusion du verre, il est possible que des corpuscules solides (des pierres) peuvent être présents et si une de telles pierres est incorporée au verre d'une bouteille, la bouteille est inutilisable. On a raison de supposer que dans des conditions normales de fabrication, une même masse de verre liquide contient en moyenne le même nombre de pierres, disons x pierre pour un quintal de verre (donc pour 100 pierres, en sachant que le poids d'une bouteille est d'environ 1 kg). Quel est le nombre attendu de bouteilles non utilisables ? ⋆

En se basant sur l'Exemple 4.7.1, nous allons construire un modèle mathématique qui sera, pour sa nature, une approximation de la realité : Supposons que chaque pierre a la même probabilité de se trouver dans le matériau de chacune des bouteilles. Nous avons donc à distribuer n pierres en N bouteilles, et on s'intéresse à la probabilité qu'une bouteille choisie au hasard contienne exactement k pierres.

D'après la loi binomiale et la combinatoire du Chapitre I, nous avons que si X_n est le nombre de pierres (parmi les n disponibles) qui finissent dans une bouteille fixée, on a

$$P(X_n = k) = \binom{n}{k} \frac{1}{N^k} \left(1 - \frac{1}{N}\right)^{n-k}.$$

Supposons qu'on veut maintenant connaître le percentage de rebut dans la fabrication de N bouteilles avec M quintaux de verre liquide, donc ici $N = 100M$ et $n = xM$ (chacune des bouteilles d'1 kg, et x est défini en Exemple 4.7.1). En posant $\lambda = \frac{x}{100}$ (le nombre de pierre qu'on attenderait en moyenne par bouteille, si les pierres étaient distribuées uniformement), on a que

$$P(X_n = k) = \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^{n-k}\prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right)$$

On s'intéresse à la limite pour $n \to +\infty$ (nombre de pierre grand), qui donne

$$\lim_{n\to+\infty} P(X_n = k) = \frac{\lambda^k}{k!}e^{-\lambda}$$

Une variable aléatoire X telle que $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$ pour un certain $\lambda > 0$ est dite suivre une *loi de Poisson de paramètre* λ. En symboles, nous allons dire que $X \sim \mathscr{P}(\lambda)$. Elle est bien une loi de probabilité car $\frac{\lambda^k}{k!}e^{-\lambda} \geq 0$ et

$$\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1.$$

En utilisant les propriétés de la série $e^x = \sum_k \frac{x^k}{k!}$, on peut facilement montrer (voir feuilles de TD) pour une variable $X \sim \mathscr{P}(\lambda)$:

$$\mathbb{E}(X) = \lambda,\quad \mathbb{E}\left(X^2\right) = \lambda^2+\lambda \quad \text{et que} \quad Var(X) = \lambda.$$

Théorème 4.7.1 (Stabilité de la loi de Poisson)**.** *Soient $X \sim \mathscr{P}(\lambda)$ et $Y \sim \mathscr{P}(\mu)$, et supposons que X et Y sont indépendantes. Alors $X+Y \sim \mathscr{P}(\lambda+\mu)$.*

Démonstration. On sait que $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$ et que $P(Y = \ell) = \frac{\mu^\ell}{\ell!}e^{-\mu}$. Maintenant, on sait que $X(\Omega) = Y(\Omega) = (X+Y)(\Omega) = \mathbb{N}$. Soit $n \in (X+Y)(\Omega)$, donc on déduit la décomposition d'événements

$$«X+Y = n» = \bigcup_{k+\ell=n} «X = k» \cap «Y = \ell»$$

et cette décomposition est disjointe, donc les probabilités s'ajoutent. On déduit par indépendance que

$$\begin{aligned}
P(X+Y = n) &= \sum_{k+\ell=n} P(«X = k» \cap «Y = \ell») = \sum_{k+\ell=n} P(X = k)P(Y = \ell) = \\
&= \sum_{k+\ell=n}\frac{\lambda^k}{k!}e^{-\lambda}\frac{\mu^\ell}{\ell!}e^{-\mu} = e^{-(\lambda+\mu)}\sum_{k+\ell=n}\frac{1}{(k+\ell)!}\frac{(k+\ell)!}{k!\ell!}\lambda^k\mu^\ell \\
&= \frac{e^{-(\lambda+\mu)}}{n!}\sum_{k=0}^{n}\binom{n}{k}\lambda^k\mu^{n-k} = \frac{(\lambda+\mu)^n}{n!}e^{-(\lambda+\mu)}
\end{aligned}$$

Donc on déduit que $X+Y$ suit bien une loi de Poisson de paramètre $\lambda+\mu$. ∎

4.8. Approximation de la loi binomiale par une loi de Poisson

Nous avons défini les lois binomiales et de Poisson. Dans cette section nous allons voir que dans certains cas la première peut être approximée par la deuxième. Discutons d'abord un exemple numérique.

Exemple 4.8.1. Une personne est née le 25 avril. Il/elle décide de demander le jour de naissance à 500 personnes au hasard, et de noter le nombre de personnes nées le 25 avril. La situation peut être assimilée à une suite de $n = 500$ épreuves indépendantes répétées, avec une probabilité de succès $p = \frac{1}{365}$ (on néglige les effets des années bissextiles).
Soit X la variable aléatoire qui compte le nombre de personnes nées le 25 avril, parmi les 500 interviewées. Nous avons vu en Section 4.3 que X suit une loi binomiale $\mathscr{B}(500, \frac{1}{365})$.
La loi exacte de X est donc $P(X = k) = C_{500,k}p^k(1-p)^{500-k}$, c'est-à-dire elle est donnée par les valeurs de probabilités suivantes (pour k petit) :

k	0	1	2	3	4	5	...
$P(X = k)$	$0,2537$	$0,3484$	$0,2388$	$0,1089$	$0,0372$	$0,0101$	...

On considère aussi une variable aléatoire Y suivant une distribution de Poisson $P(\lambda)$, de paramètre $\lambda = np = \frac{500}{365}$, dont la distribution est donc $P(Y = k) = e^{-np}\frac{(np)^k}{k!}$, c'est-à-dire :

k	0	1	2	3	4	5	...
$P(Y = k)$	$0,2541$	$0,3481$	$0,2385$	$0,1089$	$0,0373$	$0,0102$	...

On constate que les valeurs approchées ($P(Y = k)$) sont très proches des valeurs exactes ($P(X = k)$). Les deux distributions sont comparées en Figure 1.8. ⋆

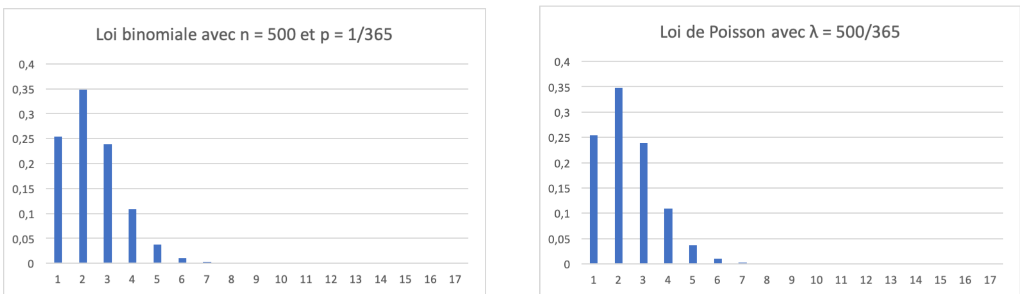


FIGURE 1.8. Comparaison de $\mathscr{B}(500, \frac{1}{365})$ et $\mathscr{P}(\frac{500}{365})$

Le phenomène identifié par l'Exemple 4.8.1 est décrit par le théorème suivant.

Théorème 4.8.1. *Soit X_n une suite de variables aléatoires suivant une loi binomiale $B(n, p_n)$. On suppose que $np_n \to \lambda > 0$. Alors (X_n) converge en loi vers une variable aléatoire de Poisson de paramètre λ, c'est-à-dire, pour tout $k \in \mathbb{N}$, on a*

$$P(X_n = k) \to \frac{e^{-\lambda}\lambda^k}{k!}$$

4.9. Exercices du Chapitre IV

1. Calculer l'espérance et la variance d'une variable aléatoire réelle X
   — de loi de Bernouilli de paramètre p.
   — de loi binomiale $\mathscr{B}(n,p)$.
   — de loi géométrique $\mathscr{G}(p)$.
   — de loi binomiale négative $\mathscr{B}\mathscr{N}(n,p)$.
   — de loi de Poisson $\mathscr{P}(\lambda)$.

2. Soit $n \in \mathbb{N}^*$. On effectue une série de n tirages au hasard, sans remise, dans une urne qui contient n_1 boules blanches et n_2 boules noires, avec $n \leq \min(n_1, n_2)$. On note $p = \frac{n_1}{n_1+n_2} \in]0,1[$ la proportion de boules blanches dans l'urne. On note X le nombre de boules blanches obtenues au cours de ces n tirages. On a vu en cours que la loi de X est hypergéométrique : $X \sim H(n_1+n_2, n, p)$.

3. Soit $n \in \mathbb{N}^*$. On effectue une série de n tirages au hasard, sans remise, dans une urne qui contient n_1 boules blanches et n_2 boules noires, avec $n \leq \min(n_1, n_2)$. On note $p = \frac{n_1}{n_1+n_2} \in]0,1[$ la proportion de boules blanches dans l'urne. On note X le nombre de boules blanches obtenues au cours de ces n tirages. Montrer que X suit une loi hypergéométrique : $X \sim H(n_1+n_2, n, p)$. Après, pour tout $i \in \{1,\dots,n\}$, on note Y_i la variable Bernoulli qui vaut 1 si le $i-$ième tirage est une boule blanche, sinon 0.
   — Déterminer la loi de Y_1.
   — Calculer $P[Y_2 = 1]$ (à l'aide de la formule des probabilités totales) et donner la loi de Y_2.
   — En déduire la loi de Y_i pour tout $i \in \{1,\dots,n\}$.
   — Donner une relation entre X et les Y_i.
   — En déduire l'espérance de X.

4. Soient deux avions : un biréacteur B et un quadriréacteur Q. On suppose que tous les réacteurs ont la même probabilité $p \in]0,1[$ de tomber en panne lors d'un voyage fixé. On estime qu'un avion peut achever son vol si et seulement si la moitié au moins de ses réacteurs fonctionnent. En supposant que les réacteurs sont indépendants, indiquer selon les valeurs de p celui des deux avions B et Q qui offre la meilleure sécurité.

5. Soit X une variable aléatoire réelle de loi de Poisson de paramètre $\lambda > 0$.
   (a) Déterminer l'espérance de $\frac{1}{1+X}$ et de $\frac{1}{2+X}$.
   (b) Déterminer la loi et l'espérance de $Y = (-1)^X$.

6. Montrer que si $X \sim \mathscr{B}(p)$, alors $X^2 = X$.

7.

Appendice

7.1. Rappels de théorie des ensembles et des fonctions

Nous considérons une approche naïve à la théorie des ensembles : on dira qu'un ensemble est une collection d'objets que l'on appelle *éléments* de l'ensemble. Les ensembles seront notés avec les lettres majuscules $A, B, C\dots$, exception faite des ensembles de nombres naturels $\mathbb{N}$, entiers $\mathbb{Z}$ ou réels $\mathbb{R}$. Si $a, b, c\dots$ appartiennent à l'ensemble A, on notera $A = \{a, b, c\dots\}$, ou simplement $a, b, c \in A$. On n'admet pas de répétition dans les éléments d'un ensemble, par exemple les ensembles $\{1, 1, 2\}$ et $\{1, 2\}$ coïncident.

On utilisera les notations classiques pour la *réunion* $A \cup B$, l'*intersection* $A \cap B$, la *différence* $A \setminus B$ et la *différence symétrique* $A \Delta B := (A \cup B) \setminus (A \cap B)$ (voir Figure 1.9) de deux ensembles A, B. Par exemple, si $A = \{1, 2, 3, 4, 5, 6\}$ et si B est l'ensemble des nombres impairs ($B \subset \mathbb{N}$), on aura $A \cap B = \{1, 3, 5\}$ et $A \setminus B = \{2, 4, 6\}$. Réunion et intersection s'étendent à des collections infinies d'ensembles.

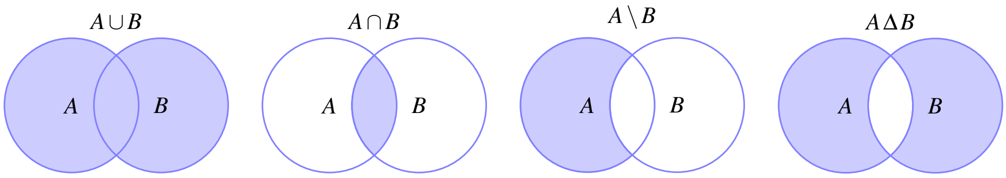


FIGURE 1.9. Réunion, intersection, différence et différence symétrique d'ensembles

Une *partie* ou *sous-ensemble* d'un ensemble A est un ensemble B dont les éléments appartiennent à A (noté $B \subset A$). Deux ensembles A, B sont égaux si et seulement si $A \subset B$ et $B \subset A$ (A est une partie de B et B est une partie de A). Un ensemble A contenant un nombre fini d'éléments est dit *fini*, et le nombre de ses éléments (son *cardinal*) est noté $|A|$. Si A est un ensemble, on appelle *ensemble des parties* de A l'ensemble $\mathscr{P}(A)$ dont les éléments sont les parties de A. Par exemple

$$\mathscr{P}(\{1,2,3\}) = \Big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\Big\}.$$

Si A est fini, $|\mathscr{P}(A)| = 2^{|A|}$ (exercice).

Une fonction $f : A \to B$ est dite *injective* si $f(a_1) = f(a_2)$ implique $a_1 = a_2$, pour tout $a_1, a_2 \in A$ et *surjective* si pour tout $b \in B$ il existe $a \in A$ tel que $f(a) = b$. Une fonction injective et surjective est dite *bijective*. Si A est B sont des ensembles finis, et $f : A \to B$ est bijective, alors $|A| = |B|$. L'ensemble des fonctions de A à B est noté B^A ; si A, B sont des ensembles finis, alors $|B^A| = |B|^{|A|}$.

Un ensemble A est dit *dénombrable* s'il existe $B \subset \mathbb{N}$ et une fonction bijective $f : A \to B$. Par exemple, tout ensemble fini est dénombrable. Soit A un ensemble, alors une collection dénombrable $A_1, \dots, A_n \dots \in \mathscr{P}(A)$ de parties de A est appelée une *partition* de A si $A_i \cap A_j = \emptyset$, pour tout $i \neq j$, et $A = \cup_i A_i := A_1 \cup A_2 \cup \dots \cup A_n \cup \cdots$. Par exemple $\{1,2,3\}, \{4\}, \{5,6,7\}$ est une partition (finie) de $\{1,2,3,4,5,6,7\}$.

Le résultat suivant est très important en combinatoire et sera utilisé dans la probabilité sur les ensembles finis. Il permet d'exprimer le cardinal de la réunion d'une collection finie d'ensembles, en fonction du cardinal de ces ensembles et de leurs intersections.

Théorème 7.1.1 (Principe d'inclusion-exclusion)**.** *Soient $A_1, \dots, A_n$ ensembles finis. Alors*

$$|A_1 \cup \dots \cup A_n| = \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{1 \le i_1 < \dots < i_k \le n} |A_{i_1} \cap \dots \cap A_{i_k}| \right)$$

Le principe d'inclusion-exclusion pour n petit :

$n = 2$: $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$

$n = 3$: $|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|$

Explication pour le cas $n = 3$ (et en général pour tout n). On veut calculer $|A_1 \cup A_2 \cup A_3|$, et pour faire cela on commence à additionner les cardinaux $|A_1| + |A_2| + |A_3|$. En faisant ça, nous avons compté 2 fois chacune des intersections 2 à 2, et trois fois l'intersection 3 à 3. En soustrayant les intersections 2 à 2, on les a donc comptées une fois chacune, mais nous avons soustrait 3 fois l'intersection 3 à 3, qu'il faut donc réadditionner pour obtenir le bon résultat.

Nous rappélons que une réunion $A \cup B$ est dite *disjointe* si $A \cap B = \emptyset$ (même chose pour une réunion de plus de deux ensembles ou infinie, dans ce cas on demandera que les intersections deux à deux sont vides).

Théorème 7.1.2. *Soit $\{A_i\}_{i=1,\dots,+\infty}$ une suite de parties d'un ensemble Ω. Alors*

$$\bigcup_{i=1}^{+\infty} A_i = \bigcup_{n=1}^{+\infty} \left(A_n \setminus \bigcup_{j=0}^{n-1} A_j \right).$$

Autrement dit, toute réunion peut s'exprimer sous forme de réunion disjointe.

7.2. Suites et séries numériques

Une *suite numérique à valeurs réelles* est une application

$$u \colon \mathbb{N} \to \mathbb{R},$$

éventuellement non-définie sur une partie finie de $\mathbb{N}$.

On note de manière compacte $u_n := u(n) \in \mathbb{R}$, et avec $\{u_n\}$ la suite u. Des exemples classiques de suites numériques sont par exemple

— $u_n = 1/n$, pour $n \in \mathbb{N}$, dite *suite harmonique*. Dans ce cas l'image $u(\mathbb{N})$ est une partie de $[0, 1]$ (suite bornée).

— $u_n = a^n$, pour un réel $a \in \mathbb{R}$ fixé, et pour $n \in \mathbb{N}$. Cette suite est dite *suite géométrique de raison a*. Si a est positif, u_n est une suite à termes positives ; autrement c'est une suite de signes alternés. Par exemple pour $a = -1$ nous obtenons la suite alternante $u_n = (-1)^n$.

La suite $\{u_n\}$ est dite *convergente* s'il existe $\ell \in \mathbb{R}$ tel que :

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \ \text{ tel que } \ |u_n - \ell| < \varepsilon \ \text{ pour tout } n > N,$$

Si $\{u_n\}$ est convergente, tout $\ell \in \mathbb{R}$ vérifiant la propriété décrite dessus est dit une *limite* de $\{u_n\}$. La limite d'une série convergente est unique. Plusieures *points limites* peuvent exister pour une suite non convergente (limites de sous-suites).

Propriété de la série géométrique :

$$\boxed{\sum_{n=0}^{\infty} a^n = \frac{1}{1-a} \quad \text{pour tout } \ -1 < a < 1}$$

Index