# Cours de Mathématiques.

Claude PETIT
Département Réseaux & Télécommunications
IUT de Saint Malo

27 septembre 2011

# Chapitre 21

# Dénombrement

## 21.1 Opérations sur les ensembles

### 21.1.1 réunion, intersection, produit cartésien

On rappelle les définitions suivantes (pour plus de précisions, se reporter à l'annexe « théorie des ensembles ») :

Un ensemble est une collection d'objets (qui sont appelés les éléments de l'ensemble).
$x \in A$ indique que l'élément $x$ appartient à l'ensemble $A$.
$\emptyset$ est le seul ensemble ne contenant aucun élément. On l'appelle ensemble vide.

Etant donné un ensemble $\Omega$ note $\mathfrak{P}(\Omega)$ l'ensemble des parties de $\Omega$.
Un élément de $\mathfrak{P}(\Omega)$ est donc un sous ensemble $A$ de $\Omega$.
Dans toute la suite, $A$ et $B$ seront des sous ensembles de $\Omega$.

Définition 76

- $A \cup B = \{x \in \Omega / \ x \in A \ ou \ x \in B\}$
- $A \cap B = \{x \in \Omega / \ x \in A \ et \ x \in B\}$
- $\bar{A}^{\Omega} = \{x \in \Omega / \ x \notin A\}$
- $A/B = \{x \in \Omega / \ x \in A \ et \ x \notin B\}$
- $A \Delta B = (A \cup B)/(A \cap B) = \{x \in A \cup B / \ x \notin A \cap B$
- $A/B = \{x \in \Omega / \ x \in Aetx \notin B$

$A \cup B$ est la réunion de $A$ et$B$. C'est l'ensemble des éléments qui sont ou dans $A$ ou dans $B$.
$A \cap B$ est l'intersection de $A$ et $B$. C'est l'ensemble des éléments communs à $A$ et$B$.
$\bar{A}^{\Omega}$ est le complémentaire de $A$ dans $\Omega$. C'est l'ensemble des éléments de $\Omega$ qui ne sont pas dans $A$.
$A/B$ est l'ensemble $A$ moins $B$ (éléments de $A$ qui ne sont pas dans $B$. On a aussi $A/B = \bar{A}^{B}$).
$A \Delta B$ s'appelle la différence symétrique de $A$ et $B$. Il s'agit de l'ensemble $A \cup B$ dont on a retiré les éléments communs à $A$ et $B$.

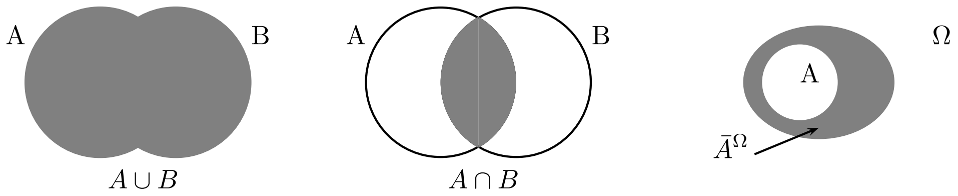


Figure 21.1: Union, intersection, complémentaire

Lorsque $A \cap B = \emptyset$, on dit que $A$ et $B$ sont disjoints (ils n'ont pas d'éléments communs).
$A = B$ si et seulement si $A \subset B$ et $B \subset A$.

Définition 77

> *Une famille non vide $A_1, A_2, ..., A_n$ de sous ensembles de $\Omega$ est une* **partition** *de $\Omega$ si:*
> - *les $A_i$ sont deux à deux disjoints: $A_i \cap A_j = \emptyset \ \forall i \neq j$*
> - *leur réunion forme $\Omega$: $\bigcup_{i=1}^{n} A_i = \Omega$*

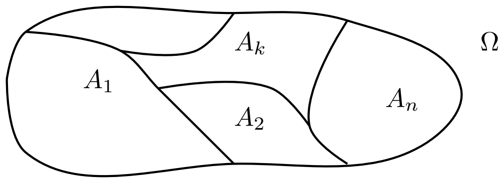


Figure 21.2: Partition d'un ensemble

**Ex** $\Omega = \{\ 1,2,3\ \}$ $A = \{\ 1,3\ \}$ $B = \{\ 2,3\ \}$
$\mathfrak{P}(\Omega) = \{\ \emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\ \}$
$A \cup B = \Omega$, $A \cap B = \{\ 3\ \}$, $A \Delta B = \{1,2\}$, $A/B = \{1\}$ et $\bar{A} = \{\ 2\ \}$

Propriété 73 (Lois de De Morgan)

> - $\overline{A \cup B} = \bar{A} \cap \bar{B}$
> - $\overline{A \cap B} = \bar{A} \cup \bar{B}$

Démo
- Soit $x \in \overline{A \cup B}$, alors $x \notin A \cup B$ donc $x \notin A$ et $x \notin B$ donc $x \in \bar{A} \cap \bar{B}$. La réciproque est la même.
- Soit $x \in \overline{A \cap B}$, alors $x \notin A \cap B$ donc $x \notin A$ ou $x \notin B$ donc $x \in \bar{A} \cup \bar{B}$. Là encore, la réciproque est identique.
□

Propriété 74 (Distributivité de $\cup$ et $\cap$)

> - $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
> - $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Démo
Elle se fait exactement de la même façon que ci-dessus (en exercice).
□

Définition 78

> *On appelle produit cartésien de $A$ par $B$ l'ensemble noté $A \times B$ défini par:*
> $A \times B = \{(a,b)/\ a \in A \text{ et } b \in B\}$

Si $A_1, A_2, ..., A_n \in \mathfrak{P}(\Omega)$ on peut généraliser ce produit à $A_1 \times A_2 \times ... \times A_n = \{(a_1, a_2, ..., a_n)/\ a_i \in A_i\}$
En particulier, $\underbrace{A \times A \times ... \times A}_{n\times}$ se note $A^n$

**Ex:** $A = \{\ 1,2,3\ \}$ $B = \{\ a,b\ \}$ $\Rightarrow$ $A \times B = \{\ (1,a),(2,a),(3,a),(1,b),(2,b),(3,b)\ \}$

**Ex:** $A = \{\ 7,8,9,10,V,D,R,A\ \}$ $B = \{\ \clubsuit, \diamondsuit, \heartsuit, \spadesuit\ \}$ $\Rightarrow$ $A \times B$ est un jeu de 32 cartes.

**Ex:** $A = B = \mathbb{R}$ $\Rightarrow$ $A \times B = \mathbb{R}^2$ est l'ensemble des points du plan

**Ex:** $H = \{$ Marcel,Lucien,Honoré $\}$ $F = \{$ Huguette,Juliette,Cunégonde $\}$
$\Rightarrow$ $H \times F$ est l'ensemble de tous les couples possibles.

### 21.1.2 Cardinal d'un ensemble

Définition 79

> *Le cardinal d'un ensemble $\Omega$ est le nombre de ses éléments. On le note Card $\Omega$ ou $|\Omega|$*

Propriété 75

- *Card* $A \cup B =$ *Card* $A +$ *Card* $B -$ *Card* $A \cap B$
- *Card* $A \times B =$ *Card* $A \times$ *Card* $B$
- *Card* $\Omega = n \ \Rightarrow$ *Card* $\mathfrak{P}(\Omega) = 2^n$

Démo
• Le nombre d'éléments dans $A \cup B$ est égal au nombre d'éléments dans $A$ plus le nombre d'éléments dans $B$ moins le nombre d'éléments communs que l'on a compté deux fois !!!

• Il s'agit du principe multiplicatif.
Pour chaque choix d'un élément de $A$, on a Card $B$ choix d'un élément de $B$, d'où la formule.

• Soit $A$ une partie de $\Omega$. A chaque élément $\omega_1, \omega_2, ..., \omega_n$ de $\Omega$, nous pouvons associer 0 ou 1 selon qu'il appartient ou non à $A$. Ainsi, il y a autant de parties de $\Omega$ que de choix 0 ou 1 possibles.
Autrement dit $\begin{array}{rcl} \phi : \mathfrak{P}(\Omega) & \longrightarrow & \{0,1\}^n \\ A & \longmapsto & \phi(A) = (\epsilon_1, ...\epsilon_n) \end{array}$ avec $\epsilon_k = 1 \ \Leftrightarrow \ w_k \in A$ est une fonction bijective.

□

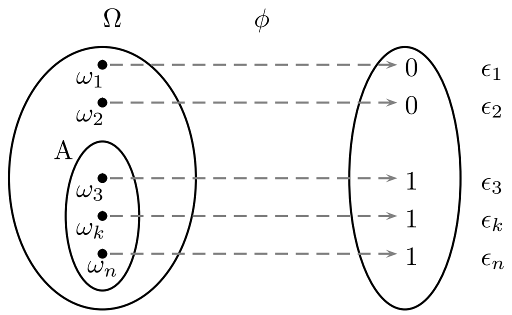


Figure 21.3: Cardinal d'un ensemble

**Ex:** $\Omega = \{ \ a, b, c \ \}$ Card $\Omega = 3$, Card $\mathfrak{P}(\Omega) = 8$.

## 21.2 Listes, arrangements, combinaisons

### 21.2.1 Définition

Définition 80

- *Une p-liste de* $\Omega$ *est un élément de* $\Omega^p$.
- *Un p-arrangement de* $\Omega$ *est une p-liste d'élémens distincts.*
- *Une permutation de* $\Omega$ *est un n-arrangement.*
- *Une combinaison à* $p$ *éléments est un sous ensemble de cardinal* $p$ *de* $\Omega$.

**Exemple:** $\Omega = \{ \ a, b, c \ \}$

Listes:
(a,b),(b,c),(a,a) sont des 2-listes.
(a,b,c), (a,c,b), (a,a,b), (c,c,c) sont des 3-listes.
(a,a,a,b,c,b,a,c,c,b) est une 10-liste.

Arrangements:
(a,b), (b,c) sont des 2-arrangements mais (a,a) n'en est pas un.
(a,b,c), (c,b,a), (a,c,b) sont des 3-arrangements.
Il ne peut pas exister de 4-arrangements ou de p-arrangements pour $p > n$ car les éléments doivent être distincts.

Permutations:
(a,b,c), (a,c,b), (b,a,c), (b,c,a), (c,a,b) et (c,b,a) sont les 6 permutations de $\Omega$

Combinaisons:
{a,b}, {b,c} et {c,a} sont les combinaisons à 2 éléments de $\Omega$ .
{a,b,c} est la seule combinaison à 3 éléments de $\Omega$.

Les combinaisons et permutations sont, par définition, sans répétition des éléments. Nous serons amenés à définir également des combinaisons et des permutations à répétition.

Définition 81

> *Une combinaison avec répétition de $p$ éléments choisis parmi les $n$ éléments de $\Omega$ est un sous-ensemble non ordonné de cardinal $p$ dont les éléments sont choisis dans $\Omega$.*

Par exemple, si $\Omega = \{ \ a, b, c \ \}$, alors $\{a, a\}$, $\{a, b\}$, $\{c, a\}$, $\{c, c\}$ sont des combinaisons (à répétition) de 2 éléments de $\Omega$.

Définition 82

> *Considérons un ensemble $\Omega$ dont les $n$ éléments sont regroupés en $k$ catégories,*
> *contenant respectivement $n_1, n_2, ..., n_k$ éléments (avec $n_1 + n_2 + ... + n_k = n$).*
> *Une permutation avec répétition de $\Omega$ est une liste ordonnée de $n$ éléments,*
> *dans laquelle figure $n_i$ éléments de la catégorie $i$ pour tout $i = 1, .., k$.*

Par exemple, le nombre d'anagrammes (avec ou sans signification) du mot LIVRE est 5!. Par contre, le mot TRALALA contient 2 L et 3 A. En ce cas, les 7 lettres se décomposent en 4 catégories de lettres distinctes et un anagramme du mot est une permutation à répétition.

### 21.2.2 Propriétés

Théorème 73 (Principe Multiplicatif)

> *Si Card $\Omega = n$:*
> - *Le nombre de p-listes de $\Omega$ est $n^p$*
> - *Le nombre de p-arrangements est $A_n^p = n(n-1)(n-2)...(n-p+1)$*
> - *Le nombre de permutations est $n!$*
> - *Le nombre de combinaisons à $p$ éléments est $C_n^p = A_n^p / p\ !$*
> - *Le nombre de combinaisons à répétition de $p$ éléments est $\Gamma_n^p = C_{n+p-1}^p$*

Si $\Omega$ est formé de $k$ catégories $n_1, ..., n_k$, le nombre de permutation à répétition de $\Omega$ est

$$\begin{pmatrix} n \\ n_1; n_2; ...; n_k \end{pmatrix} = \frac{n!}{n_1! n_2! ... n_k!}$$

On rappelle que la factorielle d'un entier $n$ est $n! = 1 \times 2 \times ... \times n$ et que par convention, $0! = 1$.

Démo

- Une p-liste étant un élément du produit cartésien $\Omega^p$, il en existe $(\text{Card } \Omega)^p = n^p$
- On applique encore le principe multiplicatif:
On a $n$ choix pour le premier élément de la liste, puis pour chaque choix effectué il reste $n-1$ choix possibles pour le second (qui doit être différent du premier), puis $n-2$ choix pour le troisième, etc....
- Pour le nombre de permutations, il suffit d'appliquer la formule ci dessus avec $n = p$:
$A_n^n = n(n-1)(n-2)...(n-n+1) = n!$
- Etant donné un p-arrangement de $\Omega$, on peut effectuer $p!$ permutation de ses éléments qui conduisent à un p-arrangement distinct. A ces $p!$ p-arrangements, est associée une unique combinaison formée des éléments des p-arrangements. Il y a donc $p!$ fois moins de p-arrangements que de combinaisons.
- On place les $n$ éléments de $\Omega$ en ligne. Entre chaque élément, on place une boule noire (on a donc $n-1$ boules noires). Une combinaison à répétition est définie par le nombre de fois que chaque élément de $\Omega$ apparaît. A la place de chaque élément, on met autant de boules blanches que le nombre de fois où cet élément apparaît dans la combinaison. On se retrouve alors avec $n+p-1$ boules dont $p$ sont blanches et $n-1$ sont noires et pour chaque combinaison à $p$ éléments, il existe une unique configuration des boules. Définir une combinaison est donc équivalent à définir la position de $n-1$ boules noires parmi $n+p-1$, c'est à dire $C_{n+p-1}^{n-1} = C_{n+p-1}^p$.
- On considère que l'on peut numéroter tous les objets de $\Omega$ pour les distinguer. Le nombre de permutations possibles est $n!$. Lorsque l'on efface les numéros d'une catégorie $i$, toutes les boules de cette catégorie deviennent identiques et l'on peut donc regrouper les permutations en $n_i!$ paquets qui contiennent tous les mêmes mots. En opérant de la même façon pour toutes les catégories, on voit donc qu'il faut diviser $n!$ par $n_1! \times ... \times n_k!$ pour avoir le nombre de permutations à répétition.

$\square$

**Exemple de référence:** Le modèle de l'urne

On considère une urne contenant $n$ boules numérotées de 1 à $n$.

On effectue successivement $p$ tirages d'une boule de l'urne, en remettant la boule après chaque tirage, et l'on observe la suite des numéros sortis en tenant compte de l'ordre. Un tirage correspond alors à une $p$-liste de l'ensemble des boules.

On effectue successivement $p$ tirages d'une boule de l'urne, sans remise, et l'on observe la suite des numéros sortis en tenant compte de l'ordre. Un tirage correspond alors à une $p$-arrangement de l'ensemble des boules.

On effectue successivement $n$ tirages d'une boule de l'urne, sans remise, et l'on observe la suite des numéros sortis en tenant compte de l'ordre. Un tirage correspond alors à une permutation (sans répétition) de l'ensemble des boules.

On effectue le tirage simultané de $p$ boules et l'on observe les numéros sortis, sans tenir compte de l'ordre. Il s'agit alors d'une combinaison (sans répétition) à $p$ éléments.

On tire $p$ successivement dans l'urne, en les remttant dans l'urne après chaque tirage. On observe les numéros sortis, mais sans tenir compte de l'ordre. Il s'agit alors d'une combinaison à répétition.

Si l'urne contient $n_i$ boules de couleur $C_i$ pour $i = 1, ..., k$, que l'on effectue $n$ tirages successifs et sans remise et que l'on observe la couleur de la séquence obtenue, alors cette séquence correspond à une permutation à répétition.

Résumons ces situations par un tableau:

| | avec ordre | sans ordre |
|---|---|---|
| avec répétition | $n^p$ | $\Gamma_n^p$ |
| sans répétition | $A_n^p$ | $C_n^p$ |

Remarque 19
*Le principe multiplicatif indique qu'un* ET *se traduit en terme de dénombrement par une* $\boxed{\times}$ *et un* OU *par une* $\boxed{+}$ *(si les ensembles correspondants sont disjoints).*

**Exemples de dénombrement:**

• On lance $n$ fois de suite une pièce pour jouer à pile ou face. On note à chaque lancer le résultat ($\pi$ ou $\phi$). L'ensemble des configurations est $\{\pi, \phi\}^n$. Un résultat est une n-liste de $\{\pi, \phi\}$ et il y en a en tout $2^n$.

• Combien existe-t-il de tiercés différents dans l'ordre avec 20 chevaux au départ?
$A_{20}^3 = 20 \times 19 \times 18$ car un tiercé est un 3-arrangement de l'ensemble des chevaux.
Dans le désordre, il existe $C_{20}^3$ possibilités.

• Combien y a-t-il de réponses possibles au loto sportif: $3^{16}$. Une réponse est une 16-liste de $\{1, N, 2\}$

• Au poker, l'ensemble des mains de 5 cartes est une combinaison à 5 éléments. Il y en a $C_{32}^5$

• Un message binaire contien $n$ bits égaux à 0 ou 1. 3 erreurs se sont glissées parmi les bits. De combien de façon peuvent se placer ces 3 erreurs ?
On doit choisir la place de 3 cases parmi $n$, ce qui se fait de $C_n^3$ façons.

• Combien d'anagrammes peut-on former avec les lettres du mot « MATHS » ?
Un anagramme (sans sens particulier) correspond à une permutation des lettres de MATHS. Il y a donc $5! = 120$ possibilités.

• De combien de façons peut-on répartir 7 personnes sur 7 chaises, si les chaises sont alignées ? Si elles sont en cercle ?
Une configuration est une permutation des 7 personnes: Il y en a 7! et 6! si les chaises sont en cercle.

• De combien de façons peut-on répartir 12 étudiants en 3 groupes de telle sorte que chaque groupe comprenne 4 personnes ?
Une configuration est une permutation à répétition. Il en existe $\dfrac{12!}{4!4!4!}$

**Propriétés des $C_n^p$ et des $A_n^p$:**

$$\boxed{A_n^p = \frac{n!}{(n-p)!}} \quad \boxed{C_n^p = \frac{n!}{p!(n-p)!}}$$

En particulier: $A_n^0 = 1$, $A_n^n = n!$, $A_n^1 = n$, $C_n^0 = 1$, $C_n^n = 1$, $C_n^1 = n$ et $C_n^2 = n(n-1)/2$

DÉMO
Première démonstration:
Choisir $p$ éléments de $\Omega$ pour former une combinaison, c'est aussi rejeter les $n-p$ éléments que l'on ne prend pas. Il y a donc autant de choix de $p$ éléments que de choix de $n-p$ éléments.
Seconde démonstration:
Dans la formule, la substitution de $p$ en $n-p$ donne immédiatement la même formule.
$\square$

THÉORÈME 74 (TRIANGLE DE PASCAL)

$$C_{n+1}^{p+1} = C_n^{p+1} + C_n^p \quad \forall\ 0 \leq p < n$$

DÉMO

$$C_n^{p+1} + C_n^p = \frac{n!}{(p+1)!(n-p-1)!} + \frac{n!}{p!(n-p)!} = \frac{n!}{(n-p-1)!}\{\frac{1}{p+1} + \frac{1}{n-p}\}$$

$$= \frac{n!}{p!(n-p-1)!}\frac{n-p+p+1}{(p+1)(n-p)} = \frac{(n+1)!}{(p+1)!(n-p)!} = C_{n+1}^{p+1}$$

Seconde démonstration:
On fixe un élément $\omega$ d'un ensemble $\Omega$ à $n+1$ éléments. Nous allons décomposer en deux l'ensemble de toutes les parties à $p+1$ éléments de $\Omega$. Considérons les parties à $p+1$ éléments qui contiennent $\omega$. Il en existe $C_n^p$, puisque l'on ne doit choisir que les $p$ éléments qui ne sont pas égaux à $\omega$. Maintenant, les parties à $p+1$ éléments qui ne contiennent pas $\omega$ sont au nombre de $C_n^{p+1}$ car on doit choisir $p+1$ éléments parmi les $n$ qui ne sont pas égaux à $\omega$. Finalement, l'ensemble de toutes les parties à $p+1$ éléments est la réunion disjointe de celui des parties à $p+1$ éléments dont $\omega$ et celui des parties à $p+1$ éléments sans $\omega$. Son cardinal est la somme des cardinaux des deux ensembles et l'on en déduit la formule de Pascal.$\square$

Cette formule permet le calcul de proche en proche des valeurs des $C_n^p$ et permet de les ranger dans un tableau appelé triangle de Pascal:

TRIANGLE DE PASCAL

| $n$ $p$ | 0 | 1 | 2 | 3 | 4 | ... |
|---|---|---|---|---|---|---|
| 0 | 1 | | | | | |
| 1 | 1 | 1 | | | | |
| 2 | 1 | 2 | 1 | | | |
| 3 | 1 | 3 | 3 | 1 | | |
| 4 | 1 | 4 | 6 | 4 | 1 | |
| 5 | 1 | 5 | 10 | 10 | 5 | 1 |

Sauf sur la diagonale ou sur la première colonne, un élément est égal à la somme de l'élément au dessus de lui et de l'élément à gauche de celui du dessus.

$$\begin{tikzcd}
 & \text{colonne } p+1 \arrow[d] & \\
C_n^p & C_n^{p+1} \arrow[r] & \text{ligne } n \\
 & C_{n+1}^{p+1} &
\end{tikzcd}$$

THÉORÈME 75 (FORMULE DU BINÔME DE NEWTON)

$\forall\ a, b \in \mathbb{R}, n\ \in \mathbb{N}$,

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

Démo

- Première démo par récurrence sur $n$:
La formule est vraie si $n=1$. Supposons là exacte au rang $n$.

$$(a+b)^{n+1}=(a+b)(a+b)^n=(a+b)\sum_{k=0}^{n}C_n^k a^k b^{n-k}=\sum_{k=0}^{n}C_n^k a^{k+1}b^{n-k}+\sum_{k=0}^{n}C_n^k a^k b^{n-k+1}=\alpha+\beta$$

Par ailleurs, $\displaystyle\sum_{k=0}^{n+1}C_{n+1}^k a^k b^{n-k+1}=a^{n+1}+b^{n+1}+\sum_{k=1}^{n}C_{n+1}^k a^k b^{n-k+1}=a^{n+1}+b^{n+1}+\sum_{k=0}^{n-1}C_{n+1}^{k+1}a^{k+1}b^{n-k}$

$$=a^{n+1}+b^{n+1}+\sum_{k=0}^{n-1}(C_n^{k+1}+C_n^k)\ a^{k+1}b^{n-k}=a^{n+1}+b^{n+1}+\sum_{k=0}^{n-1}C_n^{k+1}a^{k+1}b^{n-k}+\sum_{k=0}^{n-1}C_n^k a^{k+1}b^{n-k}$$

$$=\alpha+b^{n+1}+\sum_{k=0}^{n-1}C_n^{k+1}a^{k+1}b^{n-k}=\alpha+b^{n+1}+\sum_{k=1}^{n}C_n^k a^k b^{n-k+1}=\alpha+\beta$$

Les deux expressions coïncident, d'où la formule.

- Seconde démo par des arguments de dénombrement:
$(a+b)^n=\underbrace{(a+b)(a+b)....(a+b)}_{n\times}$.

Le développement de $(a+b)^n$ est donc une somme de termes de la forme $a^k b^i$. Puisqu'il y a en tout $n$ facteurs, lorsque $k$ facteurs $a$ sont présents dans l'un des terme, le reste du terme est formé de $n-k$ facteurs $b$. Le développement est donc formé de $a^k b^{n-k}$ pour $k$ variant de 0 à $n$.
Considérons tous les termes ayant $k$ facteurs a et donc $n-k$ facteurs b. Dénombrer ces termes revient à choisir la place des $k$ a parmi les $n$ termes possibles. Il y a donc $C_n^k$ positions possibles.

Ainsi, $\displaystyle(a+b)^n=\sum_{k=0}^{n}C_n^k a^k b^{n-k}$

□

**Ex:** $(a+b)^4=a^4+4a^3b+6a^2b^2+4ab^3+b^4$

### 21.2.3 Exemples d'applications

**Linéarisation d'expression trigonométriques**

Soit $F(x)=\cos^4 x\times\sin^2 x$ que nous souhaiterions écrire sous la forme d'une somme de sinus et de cosinus.

$$F(x)=(\frac{e^{ix}+e^{-ix}}{2})^4\times(\frac{e^{ix}-e^{-ix}}{2i})^2=\frac{1}{2^6}(e^{4ix}+4e^{3ix}e^{-ix}+6+4e^{ix}e^{-3ix}+e^{-4ix})(e^{2ix}-2+e^{-2ix})=\text{etc....}$$

**Une autre démonstration de Card $\mathfrak{P}(\Omega)=2^n$**

Les parties de $\Omega$ se décomposent en une partition formée des parties de cardinal $k$, $k$ variant de 0 à $n$, autrement dit:

$$\mathfrak{P}(\Omega)=\coprod_{k=0}^{n}(\coprod_{A\in\Omega/|A|=k}A)$$

Le nombre de parties de $\Omega$ à $k$ éléments est par définition $C_n^k$. Ainsi:

$$\text{Card }\mathfrak{P}(\Omega)=\sum_{k=0}^{n}C_n^k=\sum_{k=0}^{n}C_n^k 1^k 1^{n-k}=(1+1)^n=2^n$$

□

Dans un groupe de zéro mathématiciens, n'importe lequel sait changer une ampoule.

Un mathématicien, un physicien et un biologiste sont dans un train en Irlande. Par la fenêtre, ils aperçoivent un mouton noir.
- Comme c'est intéressant, dit le biologiste, en Irlande, les moutons sont noirs !
- La seule chose que l'on peut dire, réplique le physicien, c'est qu'il existe au moins un mouton noir en Irlande.
- Non, dit le mathématicien, on peut juste affirmer qu'en Irlande, il existe au moins un mouton dont une face est noir.

Le mathématicien, le physicien et le biologiste sont attablés à la terrasse d'un café lorsqu'ils voient deux personnes entrer dans l'immeuble en face. Quelques instants plus tard, trois personnes en sortent:
- Le biologiste : « Oh! Ils se sont reproduits! »
- Le physicien : « Mais non, il y a une erreur d'expérience ! »
- Le mathématicien : « Si une personne entre dans l'immeuble, il sera à nouveau vide. »

Le Livre de sable (Jorge Luis Borges).

La ligne est composée d'un nombre infini de points ; le plan, d'un nombre infini de lignes ; le volume, d'un nombre infini de plans ; l'hypervolume, d'un nombre infini de volumes... Non, décidément, ce n'est pas là, more geometrico, la meilleure façon de commencer mon récit. C'est devenu une convention aujourd'hui d'affirmer de tout conte fantastique qu'il est véridique ; le mien, pourtant, est véridique.
Je vis seul, au quatrième étage d'un immeuble de la rue Belgrano. II y a de cela quelques mois, en fin d'après-midi, j'entendis frapper à ma porte. J'ouvris et un inconnu entra. C'était un homme grand, aux traits imprécis. Peut-être est-ce ma myopie qui me les fit voir de la sorte. Tout son aspect reflétait une pauvreté décente. II était vêtu de gris et il tenait une valise à la main. Je me rendis tout de suite compte que c'était un étranger. Au premier abord, je le pris pour un homme âgé ; ensuite je constatai que j'avais été trompé par ses cheveux clairsemés, blonds, presque blancs, comme chez les Nordiques. Au cours de notre conversation, qui ne dura pas plus d'une heure, j'appris qu'il était originaire des Orcades. Je lui offris une chaise. L'homme laissa passer un moment avant de parler. II émanait de lui une espèce de mélancolie, comme il doit en être de moi aujourd'hui.
- Je vends des bibles, me dit-il.
Non sans pédanterie, je lui répondis :
- II y a ici plusieurs bibles anglaises, y compris la première, celle de Jean Wiclef. J'ai également celle de Cipriano de Valera, celle de Luther, qui du point de vue littéraire est la plus mauvaise, et un exemplaire en latin de la Vulgate. Comme vous voyez, ce ne sont pas précisément les bibles qui me manquent.
Après un silence, il me rétorqua :
- Je ne vends pas que des bibles. Je puis vous montrer un livre sacré qui peut-être vous intéressera. Je l'ai acheté à la frontière du Bikanir.
- II ouvrit sa valise et pose l'objet sur la table. C'était un volume in-octavo, relié en toile. II avait sans aucun doute passé par bien des mains. Je l'examinai ; son poids insolite me surprit. En haut du dos je lus Holy Writ et en bas Bombay.
- II doit dater du dix-neuvième siècle, observai-je.
- Je ne sais pas. Je ne l'ai jamais su, me fut-il répondu.
Je l'ouvris au hasard. Les caractères m'étaient inconnus. Les pages, qui me parurent assez abîmées et d'une pauvre typographie, étaient imprimées sur deux colonnes à la façon d'une bible. Le texte était serré et disposé en versets. A l'angle supérieur des pages figuraient des chiffres arabes. Mon attention fut attirée sur le fait

qu'une page paire portait, par exemple, le numéro 40514 et 1'impaire, qui suivait, le numéro 999. Je tournai cette page; au verve la pagination comportait huit chiffres. Elle était ornée d'une petite illustration, comme on en trouve dans les dictionnaires : une ancre dessinée à la plume, comme par la main malhabile d'un enfant.
L'inconnu me dit alors:
- Regardez-la bien. Vous ne la verrez jamais plus.
II y avait comme une menace dans cette affirmation, mais pas dans la voix.
Je repérai sa place exacte dans le livre et fermai le volume. Je le rouvris aussitôt. Je cherchai en vain le dessin de l'ancre, page par page. Pour masquer ma surprise, je lui dis :
- II s'agit d'une version de l'Ecriture Sainte dans une des langues hindoues, n'est-ce pas ?
- Non, me répondit-il.
Puis, baissant la voix comme pour me confier un secret :
- J'ai acheté ce volume, dit-il, dans un village de la plaine, en échange de quelques roupies et d'une bible. Son possesseur ne savait pas lire. Je suppose qu'il a pris le Livre des Livres pour une amulette. II appartenait à la caste la plus inférieure; on ne pouvait, sans contamination, marcher sur son ombre. II me dit que son livre s'appelait le livre de sable, parce que ni ce livre ni le sable n'ont de commencement ni de fin.
II me demanda de chercher la première page.
Je posai ma main gauche sur la couverture et ouvris le volume de mon pouce serré contre l'index. Je m'efforçai en vain : il restait toujours des feuilles entre la couverture et mon pouce. Elles semblaient sourdre du livre.
- Maintenant cherchez la dernière.
Mes tentatives échouèrent de même; à peine pus-je balbutier d'une voix qui n'était plus ma voix :
Cela n'est pas possible.
Toujours à voix basse le vendeur de bibles me dit :
- Cela n'est pas possible et pourtant cela est. Le nombre de pages de ce livre est exactement infini. Aucune n'est la première, aucune n'est la dernière. Je ne sais pourquoi elles sont numérotées de cette façon arbitraire. Peut-être pour laisser entendre que les composants d'une série infinie peuvent être numérotés de façon absolument quelconque.
Puis, comme s'il pensait à voix haute, il ajouta :
- Si l'espace est infini, nous sommes dans n'importe quel point de l'espace. Si le temps est infini, nous sommes dans n'importe quel point du temps.
Ses considérations m'irritèrent.
- Vous avez une religion, sans doute? lui demandai-je.
- Oui, je suis presbytérien. Ma conscience est tranquille. Je suis sûr de ne pas avoir roulé l'indigène en lui donnant la Parole du Seigneur contre son livre diabolique.
Je l'assurai qu'il n'avait rien à se reprocher et je lui demandai s'il était de passage seulement sous nos climats. II me répondit qu'il pensait retourner prochainement dans sa patrie. C'est alors que j'appris qu'il était écossais, des îles Orcades. Je lui dis que j'aimais personnellement l'Écosse, ayant une véritable passion pour Stevenson et pour Hume.
- Et pour Robbie Burns, corrigea-t-il.
Tandis que nous parlions je continuais à feuilleter le livre infini.
- Vous avez l'intention d'offrir ce curieux spécimen au British Museum ? lui demandai-je, feignant l'indifférence.
- Non. C'est à vous que je l'offre, me répliqua-t-il, et il énonça un prix élevé.
Je lui répondis, en toute sincérité, que cette somme n'était pas dans mes moyens et je me mis à réfléchir. Au bout de quelques minutes, J'avais ourdi mon plan.
- Je vous propose un échange, lui dis-je. Vous, vous avez obtenu ce volume contre quelques roupies et un exemplaire de l'Écriture Sainte; moi, je vous offre le montant de ma retraite, que je viens de toucher, et la bible de Wiclef en caractères gothiques. Elle me vient de mes parents.
- A black letter Wiclef ! murmura-t-il.
J'allai dans ma chambre et je lui apportai l'argent et le livre. Il le feuilleta et examine la page de titre avec une ferveur de bibliophile.
- Marché conclu, me dit-il.
Je fus surpris qu'il ne marchandât pas. Ce n'est que par la suite que je compris qu'il était venu chez moi décidé à me vendre le livre. Sans même les compter, il mit les billets dans sa poche.
Nous parlâmes de l'Inde, des Orcades et des jarls norvégiens qui gouvernèrent ces îles. Quand l'homme s'en alla, il faisait nuit. Je ne l'ai jamais revu et j'ignore son nom.
Je comptais ranger le livre de sable dans le vide qu'avait laissé la bible de Wiclef, mais je décidai finalement de le dissimuler derrière des volumes dépareillés des Mille et Une Nuits.
Je me couchai mais ne dormis point. Vers trois ou quatre heures du matin, j'allumai. Je repris le livre impossible et me mis à le feuilleter. Sur l'une des pages, je vis le dessin d'un masque. Le haut du feuillet portait un chiffre, que j'ai oublié, élevé à la puissance 9.
Je ne montrai mon trésor à personne. Au bonheur de le posséder s'ajouta la crainte qu'on ne me le volât, puis le soupçon qu'il ne fût pas véritablement infini. Ces deux soucis vinrent accroître ma vieille misanthropie. J'avais encore quelques amis ; je cessai de les voir. Prisonnier du livre, je ne mettais pratiquement plus le pied dehors. J'examinai à la loupe le dos et les plats fatigués et je repoussai l'éventualité d'un quelconque artifice. Je constatai que les petites illustrations se trouvaient à deux mille pages les unes des autres. Je les notai dans un répertoire alphabétique que je ne tardai pas à remplir. Elles ne réapparurent jamais. La nuit, pendant les rares intervalles que m'accordait l'insomnie, je rêvais du livre.

L'été déclinait quand je compris que ce livre était monstrueux. Cela ne me servit à rien de reconnaître que j'étais moi-même également monstrueux, moi qui le voyais avec mes yeux et le palpais avec mes dix doigts et ongles. Je sentis que c'était un objet de cauchemar, une chose obscène qui diffamait et corrompait la réalité. Je pensai au feu, mais je craignis que la combustion d'un livre infini ne soit pareillement infinie et n'asphyxie la planète par sa fumée. Je me souvins d'avoir lu quelque part que le meilleur endroit où cacher une feuille c'est une forêt. Avant d'avoir pris ma retraite, je travaillais à la Bibliothèque nationale, qui abrite neuf cent mille livres; je sais qu'à droite du vestibule, un escalier en colimaçon descend dans les profondeurs d'un sous-sol où sont gardés les périodiques et les cartes. Je profitai d'une inattention des employés pour oublier le livre de sable sur l'un des rayons humides. J'essayai de ne pas regarder à quelle hauteur ni à quelle distance de la porte. Je suis un peu soulagé mais je ne veux pas même passer rue Mexico.

© Jorge Luis Borges in « Le livre de sable. »

# Chapitre 22

# Probabilité sur un ensemble fini

## Historique

La première tentative pour appliquer les mathématiques à un jeu de hasard remonte à 1654 lorsque le chevalier de Méré expose à Pascal (1623-1662) son célèbre problème de lancer de dés. De la correspondance qui va suivre entre Pascal et Pierre de Fermat est né le calcul des probabilités. Hasard provient de l'arabe "az-zahr" (dé à jouer). Depuis le XVIIème siècle, de nombreux mathématiciens ont contribué au développement de cette théorie: De Moivre (1667-1754), Thomas Bayes (1702-1761), Laplace (1749-1827), Condorcet, Denis Poisson, Emile Borel, etc.

La théorie actuelle des probabilités doit beaucoup aux mathématiciens russes:
Pafnouti Tchebychev (1821-1894), Andreï Markov (1856-1922) ou encore Andreï Kolmogorov (1903-1987) qui a axiomatisé la théorie des probabilités dans les années 1920 en faisant le lien avec la théorie des ensembles. Kolmogorov était par contre un proche du pouvoir (son père a été ministre en URSS) qui a co-signé en 1974 un article dans la Pravda où il se félicitait de l'expulsion du prix Nobel de littérature Alexandre Soljenitsyne (celui-ci fût d'ailleurs professeur de mathématiques durant plusieurs années).

Paul Levy (1886-1971) s'intéressa également beaucoup aux probabilités; il publia en 1925 un traité de calcul des probabilités célèbre et étudia en 1948 le mouvement brownien dont Einstein avait découvert l'origine. Ce mouvement brownien permettra à Ito de définir, après la seconde guerre mondiale, l'intégrale stochastique et les équations différentielles stochastiques. Ces outils sont très utilisés aujourd'hui dans les problèmes de physique et d'économie, par exemple pour la modélisation des marchés financiers.

## 22.1 Expérience aléatoire

### 22.1.1 Introduction

Une expérience aléatoire est une expérience dont l'issue dépend du hasard. On choisit un aspect du phénomène observé sous la forme d'un caractère que l'on va étudier pour prévoir les différentes valeurs qu'il peut prendre.

L'ensemble de toutes les issues possibles de l'expérience (c'est à dire l'ensemble de toutes les valeurs possibles du caractère étudié) est appelé ensemble univers ou univers des possibles. On le notera généralement $\Omega$.

Une éventualité de l'expérience (appelée aussi évènement élémentaire ou épreuve) est une des issues possibles. C'est donc un élément $\omega$ de $\Omega$. A chaque éventualité va être associé un nombre compris entre 0 et 1 qui va mesurer sa vraisemblance, sa capacité à se réaliser ou non. Ce sera la probabilité de cette éventualité.

Un évènement de l'expérience aléatoire est un sous ensemble de $\Omega$; c'est donc un ensemble de plusieurs éventualités.

### 22.1.2 Quelques exemples d'expériences aléatoires.

**Ex1:** Epreuve de Bernoulli.

L'expérience consiste en un jeu de pile ou face. Le caractère étudié est le dessin de la face supérieure de la pièce. Nous pourrons les noter $\pi$ et $\phi$ ou encore 1 et 0.
L'ensemble univers est donc $\Omega = \{\text{pile}, \text{face}\} = \{\pi, \phi\} = \{0, 1\}$.
Les éventualités sont "pile" ou "face". De façon plus générale, toute expérience avec deux issues possibles (1 ou 0, un succès ou un échec, etc.) s'appellera une épreuve de Bernoulli.

**Ex2:** Lancer d'un dé à six faces normal.

On étudie le numéro de la face supérieure: $\Omega = \{\ 1, 2, 3, 4, 5, 6\ \}$.
Une éventualité est l'un des six numéros qui compose $\Omega$.
L'évènement "le numéro sorti est pair" est formé par $A = \{\ 2, 4, 6\ \}$.

**Ex3:** Loi équirépartie.

On tire au hasard dans une urne contenant $n$ objets numérotés de 1 à $n$. On a $\Omega = \{\ 1, 2, ..., n\ \}$.
L'évènement "l'objet tiré est inférieur ou égal à 2" est formé par le sous ensemble $A = \{\ 1, 2\ \}$ de $\Omega$.
Cette expérience généralise les deux précédentes qui en sont des cas particuliers (avec $n = 2$ et $n = 6$).

**Ex4:** Lancer de deux dés à six faces.

On lance deux dés (un vert et un rouge, mettons) et l'on observe les numéros des faces supérieures.
Une éventualité de cette expérience est un couple $(i, j)$ où $i \in \{1, 2, .., 6\}$ est le numéro du dé vert et $j$ celui du dé rouge. Ainsi, $\Omega = \{1, 2, ..., 6\} \times \{1, 2, ..., 6\} = \{(i, j)/\ i, j = 1, 2, ..., 6\}$

Dans cette expérience, l'ordre a de l'importance (on peut distinguer les deux dés) et les 36 issues possibles sont également probables.

**Ex5:** Expérience composée.

On dispose d'une pièce et de deux urnes contenant des boules blanches et noires. On tire à pile ou face (avec la pièce). Si pile apparait, on choisit une boule dans la première urne. Si face apparait, on fait de même dans la seconde. On s'intéresse au résultat du pile ou face et à la couleur de la boule obtenue.
Une éventualité est donc un couple de la forme $(\epsilon, C)$ où $\epsilon$ est le résultat du lancer de la pièce et appartient à l'ensemble $\{\pi, \phi\}$ et $C$ est la couleur de la boule et appartient à l'ensemble $\{B, N\}$.
On a alors $\Omega = \{\pi, \phi\} \times \{B, N\}$

L'ensemble univers d'une expérience composée sera représenté par un produit cartésien.

**Ex6:** Succession de tirages à pile ou face.

On effectue des tirages successifs à pile ou face. On s'intéresse au numéro du lancer où apparait pile pour la première fois, de sorte que $\Omega = \{1, 2, 3, ....\} = \mathbb{N}^*$
Il y a dans cette expérience une infinité d'éventualités (il se peut même que pile n'apparaisse jamais, auquel cas le caractère étudié vaut $+\infty$). Cet exemple sort du cadre de la leçon puisque $\Omega$ y est infini.

**Ex7:** Tirage aléatoire entre 0 et1.

La touche random d'une calculatrice peut à priori donner n'importe quel nombre entre 0 et 1 (en réalité cela est inexact car une machine ne peut représenter qu'un nombre fini de nombres et ceux-ci ne peuvent être générés que de façon pseudo-aléatoire). En ce cas, $\Omega = [\ 0, 1\ ]$ est un ensemble infini et même non dénombrable.

**Ex8:** Aiguille de Buffon.

On trace sur une feuille des droites parallèles séparées d'une distance $L$. On lance sur cette feuille une aiguille de longueur $L$ et l'on s'intéresse à l'angle formé par l'aiguille et l'une des droites.
Cet angle est aléatoire et peut prendre n'importe quelle valeur entre 0 et $\frac{\pi}{2}$.
Là encore, $\Omega$ est un ensemble infini non dénombrable.

**Ex9:** Désintégration d'un atome.

La durée entre deux désintégrations dans un atome radioactif est un phénomène parfaitement aléatoire qui prend ses valeurs dans l'intervalle $\Omega = [0, +\infty[$.

### 22.1.3 Lien avec la théorie des ensembles

Le point essentiel du cours est le lien entre la description d'une expérience aléatoire et la théorie des ensembles. C'est Kolmogorov dans les années 1930 qui a utilisé cette modélisation des phénomènes aléatoires. Elle permet de décrire en termes ensemblistes les notions d'ensemble univers, d'évènement, d'éventualité.

| DESCRIPTION DE L'EXPÉRIENCE | VOCABULAIRE ENSEMBLISTE |
|---|---|
| Expérience aléatoire | Ensemble $\Omega$ |
| Eventualité | Elément $\omega \in \Omega$ |
| Evènement | Sous ensemble $A \subset \Omega$ |
| Evènement contraire de $A$ | Complémentaire $\bar{A}^{\Omega}$ |
| Evènement $A$ et $B$ | $A \cap B$ |
| Evènement $A$ ou $B$ | $A \cup B$ |
| Evènement impossible | $\emptyset$ |
| Evènement certain | $\Omega$ |

$\boxed{\text{Evènement de l'expérience } \Leftrightarrow \text{ sous ensemble de } \Omega}$

## 22.2 Probabilité sur un ensemble

### 22.2.1 Mesure de probabilité

DÉFINITION 83 (AXIOMES DE KOLMOGOROV)

*Une probabilité sur un ensemble $\Omega$ est une fonction* $\begin{array}{rccl} \mathbb{P} : \mathfrak{P}(\Omega) & \longrightarrow & [0,1] \\ A & \longmapsto & \mathbb{P}(A) \end{array}$ *vérifiant:*

- $\mathbb{P}(\Omega) = 1$
- $\forall (A_n)_{n\in\mathbb{N}}$ *suite d'évènements de $\Omega$ 2 à 2 disjoints,* $\mathbb{P}(\bigcup_{n=0}^{+\infty} A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$

$\mathfrak{P}(\Omega)$ représente l'ensemble des parties de $\Omega$. Cet ensemble possède une structure de tribu (cf. annexes).

Une probabilité est donc une collection de nombres positifs dont la somme est 1.
$\mathbb{P}$ est une fonction qui à un ensemble associe un nombre mesurant sa **vraisemblance** ou son **poids**.

Le second axiome, dans le cas de deux évènements **disjoints** $A$ et $B$, indique que $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ et porte le nom de propriété d'additivité.

La donnée d'une expérience aléatoire modélisée par $\Omega$ et d'une probabilité $\mathbb{P}$ sur cet ensemble s'appelle un espace probabilisé et se note $(\Omega, \mathbb{P})$.

PROPRIÉTÉ 76

- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- $A \subset B \;\Rightarrow\; \mathbb{P}(A) \leq \mathbb{P}(B)$
- $\mathbb{P}(\emptyset) = 0$
- $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$

DÉMO
• Nous pouvons écrire sous la forme d'une réunion disjointe $(A \cup B) = (A/B) \cup (A\Delta B) \cup (B/A)$. On peut donc appliquer la propriété d'additivité et conclure que $\mathbb{P}(A \cup B) = \mathbb{P}(A/B) + \mathbb{P}(A\Delta B) + \mathbb{P}(B/A)$. Mais on a également $A = (A/B) \cup (A\Delta B)$ et $B = (B/A) \cup (A\Delta B)$, donc:
$\mathbb{P}(A) = \mathbb{P}(A/B) + \mathbb{P}(A\Delta B)$ et $\mathbb{P}(B) = \mathbb{P}(B/A) + \mathbb{P}(A\Delta B)$. Ainsi, $\mathbb{P}(B/A) = \mathbb{P}(B) - \mathbb{P}(A\Delta B)$ et $\mathbb{P}(A/B) = \mathbb{P}(A) - \mathbb{P}(A\Delta B)$. En recoupant les trois équations, on obtient l'égalité recherchée.

• Supposons $A \subset B$. Alors $B = A \cup \bar{A}^B$ et cette réunion est disjointe. Donc $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(\bar{A}^B)$. Puisqu'une probabilité est positive, on en déduit que $\mathbb{P}(B) \geq \mathbb{P}(A)$.

• On a $\Omega = \Omega \cup \emptyset \Rightarrow \mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset)$ et le résultat en découle.

• Là encore, $\Omega = A \cup \bar{A} \Rightarrow 1 = \mathbb{P}(A) + \mathbb{P}(\bar{A}) \Rightarrow \mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$

□

**Ex1:** Epreuve de Bernoulli

Considérons deux nombres $p$ et $q$ positifs tels que $p + q = 1$.

Nous définissons une mesure de probabilité en posant $\mathbb{P}(\pi) = p$ et $\mathbb{P}(\phi) = q$.
Cette loi est appelée loi de Bernoulli de paramètre $p$. Elle représente un tirage à pile ou face dans lequel la probabilité de faire pile est $p$ et celle de faire face est $q = 1 - p$.

**Ex3:** Loi équiprobable

Soit $\Omega = \{ \omega_1, \omega_2, ..., \omega_n \}$ un ensemble de $n$ objets. En posant $\mathbb{P}(\omega_i) = \frac{1}{n}$ on définit une loi appelée loi équiprobable qui représente une expérience équitable dans laquelle toutes les éventualités ont la même probabilité d'apparaître. Cette loi est très importante car elle modélise les expériences équitables. Le lancer d'un pile ou face honnête (avec $p = q = 1/2$) est un cas particulier de loi équiprobable. Le lancer d'un dé à six face normal aussi: on a alors $n = 6$ et toute issue a pour probabilité $1/6$.

Lorsque l'on veut calculer la probabilité d'un évènement quelconque avec la loi équiprobable, il suffit de calculer son cardinal. La probabilité de cet évènement est alors

$$\mathbb{P}(A) = \frac{Card(A)}{Card(\Omega)}$$

que l'on peut également traduire par la formule bien connue

$$\mathbb{P}(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

**Ex4:** Lancer de deux dés

On s'intéresse à la somme des deux lancers. Les valeurs possibles vont de 2 à 12 et l'on pourrait choisir comme probabilité l'équiprobabilité sur $\{2, ..., 12\}$. Ce serait une erreur, car on s'aperçoit assez vite qu'un 7 se produit beaucoup plus souvent qu'un 12: Il n'y a pas équiprobabilité dans cet ensemble (c'est d'ailleurs sur cet exemple que le chevalier de Méré a questionné Pascal). En fait, pour pouvoir poser l'équiprobabilité (qui simplifie grandement les calculs), il faut modéliser l'expérience aléatoire sous la forme suivante $\Omega = \{(i, j); i, j = 1..6\}$; Chacun des 36 couples possibles se produit manifestement avec la même probabilité et l'on peut donc choisir l'équiprobabilité sur cet ensemble. L'évènement $A$="la somme des lancers vaut 7" a alors pour probabilité $\mathbb{P}(A) = 6/36 = 1/6$ tandis que l'évènement $B$="la somme des lancers vaut 12" a pour probabilité $\mathbb{P}(B) = 1/36$

### 22.2.2 Probabilité conditionnelle

Considérons une expérience aléatoire sur laquelle a été définie un ensemble univers $\Omega$ et une mesure de probabilité $\mathbb{P}$. Soit $B$ un évènement de $\Omega$ tel que $\mathbb{P}(B) > 0$.
L'évènement $A/B$ s'appelle "A sachant B" et sa probabilité est la probabilité conditionnelle de $A$ sachant que $B$ est réalisé. On pose:

$$\boxed{\mathbb{P}(A/B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}}$$

$A/B$ représente le résidu d'information sur $A$ contenu dans l'évènement $B$.
La formule ci-dessus s'appelle formule des probabilités conditionnelles. On peut l'exprimer sous la forme suivante qui ne nécessite plus de supposer $\mathbb{P}(B) > 0$:

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

La fonction $\mathbb{P}(./B)$ qui à un évènement $A$ associe $\mathbb{P}(A/B)$ est une mesure de probabilité appelée probabilité conditionnelle sachant $B$. L'ensemble $\Omega$ sur lequel elle est définie est $B$ et non plus l'ensemble univers initial de l'expérience.

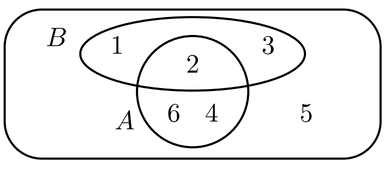


Figure 22.1: Lancer de dés et probabilité conditionnelle

**Ex8:**

Considérons le lancer d'un dé normal à 6 faces. $\Omega = \{1, 2, ..., 6\}$
Soit $A$ l'évènement "le résultat est pair" et $B$ l'évènement "le résultat est inférieur ou égal à 3"
La probabilité d'un évènement sachant $B$ défini un nouvel espace de probabilité qui est $\Omega' = \{1, 2, 3\}$
dans lequel la probabilité de $A$ est modifiée en $\mathbb{P}(A/B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \dfrac{1}{3}$

**Ex9:**

Une famille a deux enfants dont une fille. Quelle est la probabilité que le 2nd enfant soit un garçon ?

$\Omega = \{\ (G, G), (F, F), (F, G), (G, F)\ \}$ est l'ensemble univers sur lequel nous choisissons l'équiprobabilité.

Soit $A =$ " l'un des enfant est une fille " et $B =$ " l'un des enfant est un garçon ".

$A = \{\ (F, G), (G, F), (F, F)\ \}$ et $B = \{\ (G, G), (F, G), (G, F)\ \}$. Ainsi, $A \cap B = \{(F, G), (G, F)\}$

$\mathbb{P}(B/A) = \dfrac{2}{4} \times \dfrac{4}{3} = \dfrac{2}{3}$

On remarquera que $\boxed{\overline{A/B} = \bar{A}/B}$ et qu'il ne faut pas confondre $A \cap B$ et $A/B$.

Propriété 77 (Formule des probabilités totales)

*Soit $A_1$, $A_2$, ..., $A_n$ une partition de $\Omega$ et $B \subset \Omega$.*

$\mathbb{P}(B) = \displaystyle\sum_{i=1}^{n} \mathbb{P}(A_i)\mathbb{P}(B/A_i)$

Démo
Ce théorème montre que l'on peut calculer la probabilité d'un évènement en le "découpant" dans les différents ensembles d'une partition et en considérant l'influence de chaque ensemble sur cet évènement.

$B = B \cap \Omega = B \cap (\displaystyle\bigcup_{i=1}^{n} A_i) = \displaystyle\bigcup_{i=1}^{n} B \cap A_i$ d'après les lois de De Morgan.
Pour $i$ variant de 1 à $n$, les ensembles $B \cap A_i$ sont tous disjoints et l'on a donc:
$\mathbb{P}(B) = \displaystyle\sum_{i=1}^{n} \mathbb{P}(B \cap A_i) = \displaystyle\sum_{i=1}^{n} \mathbb{P}(B/A_i)\mathbb{P}(A_i)$

□

Propriété 78 (Formule de Thomas Bayes (1702-1761))

*Si $A_1$, $A_2$, ..., $A_n$ forment une partition de $\Omega$, alors $\forall B \subset \Omega$:*

$\mathbb{P}(A_i/B) = \dfrac{\mathbb{P}(A_i)\mathbb{P}(B/A_i)}{\mathbb{P}(B)}$

En utilisant la formule des probabilités totales, cette formule peut aussi s'écrire:

$$\mathbb{P}(A_i/B) = \frac{\mathbb{P}(A_i)\mathbb{P}(B/A_i)}{\displaystyle\sum_{i=1}^{n} \mathbb{P}(A_i)\mathbb{P}(B/A_i)}$$

Démo
$\mathbb{P}(B \cap A_i) = \mathbb{P}(A_i \cap B) = \mathbb{P}(A_i/B)\mathbb{P}(B) = \mathbb{P}(B/A_i)\mathbb{P}(A_i)$
□

La formule de Bayes s'appelle aussi $\boxed{\text{formule des causes}}$ et peut s'exprimer comme suit:
Supposons connus $\mathbb{P}(B)$ et $B$ dont les causes possibles de réalisation sont $A_1$, $A_2$, ..., $A_n$.
Supposons connue l'influence de $A_1$, $A_2$, ..., $A_n$ sur la réalisation de $B$: $\mathbb{P}(B/A_1)$, $\mathbb{P}(B/A_2)$, ..., $\mathbb{P}(B/A_n)$
Supposons enfin que $B$ se réalise. Alors la formule donne la probabilité que $A_i$ en soit la cause.

On peut voir cette formule comme une sorte de formule d'inversion des probabilités.

**Ex10:**
Dans un élevage, on effectue un test de dépistage sur une maladie $M$. Le test n'est pas fiable à 100%.

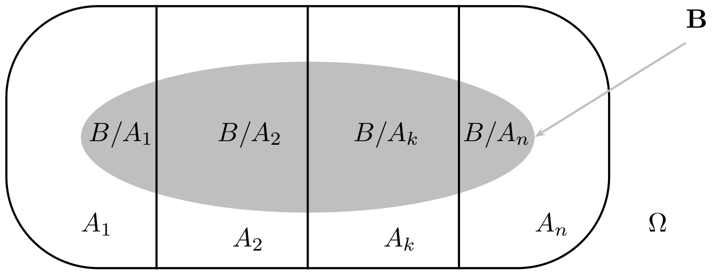


Figure 22.2: Formule de Bayes

Il peut ne pas détecter la maladie ou la détecter à tort (on appelle ces probabilités la spécificité et la sensibilité du test). On cherche la probabilité qu'un animal soit malade sachant que le test est positif (ce qui caractérisera l'efficacité du test).

On donne $\mathbb{P}(M) = 0.3$, $\mathbb{P}(-/\bar{M}) = 0.9$, $\mathbb{P}(+/M) = 0.8$ et l'on cherche $\mathbb{P}(M/+)$.

$$\mathbb{P}(M/+) = \frac{\mathbb{P}(M)\mathbb{P}(+/M)}{\mathbb{P}(M)\mathbb{P}(+/M) + \mathbb{P}(\bar{M})\mathbb{P}(+/\bar{M})} = \frac{0.3 \times 0.8}{0.24 + 0.7 \times 0.1} = 0.77$$

**Ex:10**
On considère un appareil électronique possédant trois organes $A$, $B$, $C$ indispensables à son fonctionnement. L'appareil tombe en panne. On cherche l'organe responsable de la panne. On note $P$ l'évènement "l'appareil tombe panne" et l'on donne $\mathbb{P}(P/A) = 0.1$, $\mathbb{P}(P/B) = 0.1$ et $\mathbb{P}(P/C) = 0.2$. En notant $I$ l'évènement "l'organe I tombe en panne", on sait par ailleurs que $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 0.1$.

D'après la formule de Bayes, $\mathbb{P}(A/P) = \dfrac{\mathbb{P}(P/A)\mathbb{P}(A)}{\mathbb{P}(P/A)\mathbb{P}(A) + \mathbb{P}(P/B)\mathbb{P}(B) + \mathbb{P}(P/C)\mathbb{P}(C)} = \dfrac{1}{4}$
De même, $\mathbb{P}(B/P) = \dfrac{1}{4}$ et $\mathbb{P}(C/P) = \dfrac{1}{2}$

Le plus probable est donc que la panne provienne de $C$ (on pouvait s'en douter).

Propriété 79 (Formule de la multiplication)

> *Pour une famille d'évènements $A_1$, $A_2$, ..., $A_n$:*
> $\mathbb{P}(A_A \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2/A_1)\mathbb{P}(A_3/A_1 \cap A_2)...\mathbb{P}(A_n/A_1 \cap ... \cap A_{n-1})$

Démo
Il s'agit d'une application directe de la formule des probabilités conditionnelles.
□

**Ex: 11**
Une urne contient 3 boules rouges et 1 boule verte. On tire successivment et sans remise une boule dans l'urne. On cherche la probabilité de tirer trois boules rouges lors des trois premiers tirages. Nous noterons évidemment $R_i$ l'évènement "on a tiré une boule rouge" au tirage n$^\circ$ $i$ et $V_i$ l'évènement "on a tiré une boule verte" au tirage numéro $i$. On a:

$$\mathbb{P}(R_1 \cap R_2 \cap R_3) = \mathbb{P}(R_1)\mathbb{P}(R_2/R_1)\mathbb{P}(R_3/R_1 \cap R_2) = \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{4}$$

### 22.2.3 Indépendance stochastique

Définition 84

> *On dit que deux évènements $A$ et$B$ sont indépendants si* $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$

En ce cas, $\mathbb{P}(B/A) = \mathbb{P}(B)$. Autrement dit, la connaissance de l'un n'influe pas sur l'autre.

**Ex12:**
On lance deux dés (un vert et un rouge, comme d'habitude). On note $A$ l'évènement "le dé vert fait six" et $B$ l'évènement "le dé rouge donne un résultat pair". On modélise l'expérience par l'ensemble des 36 couples de résulats possibles, qui sont équiprobables. On a alors:

$\mathbb{P}(A) = \frac{6}{36} = \frac{1}{6}$, $\mathbb{P}(B) = \frac{18}{36} = \frac{1}{2}$ et $\mathbb{P}(A \cap B) = \frac{3}{36} = \frac{1}{12}$
Par suite $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$ et les évènements sont bien indépendants.

On peut généraliser à plusieurs évènements:

Définition 85

*Soit $I = \{1, 2, ..., n\}$*
*Des évènements $A_1, A_2, ..., A_n$ sont indépendants dans leur ensemble*
*(ou mutuellement indépendants) si $\forall J \subset I$:*
$\mathbb{P}(\bigcap_{k\in J} A_k) = \prod_{k\in J} \mathbb{P}(A_k)$

Par exemple, pour 3 évènements, $A, B, C$ indépendants $\Leftrightarrow$

$$\begin{cases} \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \times \mathbb{P}(B) \times \mathbb{P}(C) & (1) \\ \mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B) & (2) \\ \mathbb{P}(A \cap C) = \mathbb{P}(A) \times \mathbb{P}(C) & (3) \\ \mathbb{P}(B \cap C) = \mathbb{P}(B) \times \mathbb{P}(C) & (4) \end{cases}$$

En particulier, la condition (1) ne suffit pas à prouver l'indépendance. L'indépendance deux à deux n'est pas non plus suffisante pour démontrer l'indépendance mutuelle.

**Ex2:** Si $\Omega = \{\ 1, 2, 3, 4, 5, 6\ \}$, $A = \{\ 1, 2, 3\ \}$, $B = \{\ 2, 4, 6\ \}$ et $C = \{\ 1, 2, 4, 5\ \}$

Les évènements $A, B, C$ ne sont pas indépendants car $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \times \mathbb{P}(B)$
et pourtant la condition (1) est vérifiée. Qu'en est-il de l'indépendance de $A$ et $B$, de $A$ et $C$, de $B$ et $C$ et de de $A$, $B$, $C$ ?

⚠ ATTENTION

Disjoints et indépendants sont deux notions différentes:
Dans l'exemple 9, $M$ et $\bar{M}$ sont disjoints sans être indépendants.
De même, des évènements peuvent être indépendants sans être disjoints: $\mathbb{P}(\Omega \cap A) = \mathbb{P}(A) \times \mathbb{P}(\Omega)$

Vous pouvez également comparer ces deux notions sur l'exemple précédent.

Propriété 80

*Soient $A$ et $B$ deux évènements indépendants. Alors les évènements suivants sont indépendants:*
- $\bar{A}$ *et* $\bar{B}$.
- $A$ *et* $\bar{B}$.
- $\bar{A}$ *et* $B$.

Démo
Montrons par exemple la première assertion:
$\mathbb{P}(\bar{A} \cap \bar{B}) = \mathbb{P}(\overline{A \cup B}) = 1 - \mathbb{P}(A \cup B) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B)$ par indépendance de $A$ et $B$.
Par ailleurs, $\mathbb{P}(\bar{A})\mathbb{P}(\bar{B}) = (1 - \mathbb{P}(A))(1 - \mathbb{P}(B))$.
En développant cette dernière expression, on retrouve l'égalité du dessus.

□

⚠ Il faut faire attention à la notion d'indépendance. Même si celle-ci correspond à une absence d'influence entre la réalisation de deux évènements, il ne faut pas se contenter de cette notion vague. Seul le calcul permet de démontrer l'indépendance. Par exemple, on peut construire des espaces probabilisés dans lesquels un évènement est indépendant de lui-même !

Crois-moi, seuls ceux qui n'attendent rien du hasard maîtrisent leur destin. (M. Arnold)

# Chapitre 23

# Variables aléatoires discrètes

## 23.1 Loi d'une variable discrète

### 23.1.1 Introduction et définition

**Ex0:** Le jeu de pile ou face.

Considérons un lancer à pile ou face dans lequel le joueur gagne 1 si pile apparaît et 0 sinon.
Soit $X$ le gain du joueur après une partie. $X$ peut prendre deux valeurs 0 ou 1, mais on ne peut prédire laquelle avant le lancer. La valeur de $X$ dépend du hasard, on dit qu'il s'agit d'une variable aléatoire.

On note $[X = 0]$ l'évènement "la pièce est tombée sur face."
On note $[X = 1]$ l'évènement "la pièce est tombée sur pile."
La loi de $X$ est la donnée des probabilités des deux évènements ci-dessus ($\frac{1}{2}$ si la pièce est normale).

**Ex1:** L'épreuve de Bernoulli.

Il s'agit d'une généralisation du lancer à pile ou face:
On considére que la variable aléatoire $X$ peut prendre les valeurs 0 ou 1 avec probabilité $p \in [0, 1]$ et $q = 1 - p$.

Pour indiquer que $X$ suit une loi de Bernoulli de paramètre $p$, nous noterons $X \sim b(p)$.

**Ex2:** La loi équiprobable.

Considérons le lancer d'un dé ordinaire et notons $X$ le numéro de la face supérieure.
$X$ peut prendre comme valeur 1,2,3,4,5 ou 6 de façon équiprobable (si le dé est normal).
La loi de $X$ est donnée par le tableau ci-dessous:

| $k$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P[X = k]$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |

**Ex3:** Loi géométrique.

On effectue une série de lancers à pile ou face indépendants modélisés par des épreuves de Bernoulli de paramètre $p$ et l'on s'arrète lorsque le premier pile apparaît. Soit $X$ le nombre de lancers nécessaires avant d'obtenir ce premier pile. $X$ est une variable aléatoire discrète mais pouvant prendre une infinité de valeurs (1, 2, ... et éventuellement $+\infty$).
Sa loi est la donnée de $P[X = k]$ pour tout $k \in \mathbb{N}^\star \cup \{+\infty\}$.

L'évènement $[X = k]$ se lit "le premier pile apparaît au kième lancer", ce que nous pourrions aussi écrire $(\underbrace{\phi, \phi, ...\phi}_{k-1\times}, \pi, ...)$

Par indépendance des lancers, on a donc $\boxed{\mathbb{P}[X = k] = q^{k-1}p}$ avec $k \geq 1$ et $q = 1 - p$

Nous devons encore vérifier que cette formule définit bien une probabilité. Nous avons vu dans la leçon précédente qu'une loi de probabilité est une collection de nombres positifs dont la somme est 1. Nous avons clairement $0 \leq q^{k-1}p \leq 1$ et par ailleurs

$$\sum_{k=1}^{+\infty} \mathbb{P}[X = k] = \sum_{k=1}^{+\infty} q^{k-1}p = p\sum_{k=1}^{+\infty} q^{k-1} = p \times \frac{1}{1-q} = 1$$

Pour indiquer que $X$ une loi géométrique de paramètre $p$, nous noterons $X \sim \mathcal{G}(p)$.

DÉFINITION 86

*On considère une expérience aléatoire sur un espace $\Omega$.*
*Une variable aléatoire $X$ sur $\Omega$ est une fonction mesurable* $\begin{array}{rccl} X: & \Omega & \longrightarrow & \mathbb{R} \\ & \omega & \longmapsto & X(\omega) \end{array}$
*qui à une éventualité $\omega$ associe le réel $X(\omega)$.*
*On note $X(\Omega)$ l'ensemble des valeurs possibles de $X$.*

A chaque résultat de l'expérience aléatoire la fonction $X$ associe un nombre.

Pour un réel $k \in \mathbb{R}$, $X^{-1}(k)$ est l'image réciproque (l'antécédent) de $k$ par $X$. Nous la noterons également

$$\boxed{[X = k] = X^{-1}(k) = \{\omega \in \Omega / \ X(\omega) = k\} = \text{“}X \text{ a pris la valeur k”}}$$

De même, si $[a, b]$ est un intervalle de $\mathbb{R}$, nous noterons son image réciproque $X^{-1}([a, b])$ sous la forme:

$$\boxed{\begin{array}{rcl} [a \leq X \leq b] & = & X^{-1}([a,b]) = \{\omega \in \Omega / \ a \leq X(\omega) \leq b\} \\ & = & \text{“}X \text{ a pris une valeur comprise entre } a \text{ et } b\text{”.} \end{array}}$$

De façon générale, si $A \subset \mathbb{R}$, l'évènement $[X \in A] \subset \Omega$ signifie

$$\boxed{\begin{array}{rcl} [X \in A] & = & X^{-1}(A) = \{\omega \in \Omega / \ X(\omega) \in A\} \\ & = & \text{“}X \text{ a pris une valeur dans } A\text{.”} \end{array}}$$

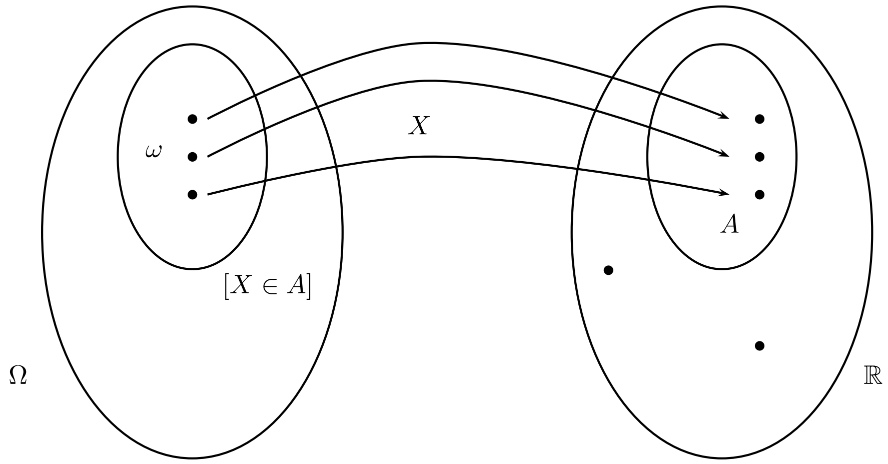


Figure 23.1: Variable aléatoire.

Nous dirons qu'une fonction $X$ est mesurable si $X^{-1}(k) \in \mathcal{P}(\Omega)$ $\forall k$, c'est à dire si l'ensemble $[X = k]$ est un évènement de $\Omega$. Dans le cas d'une probabilité finie ou dénombrable, cette condition sera toujours vérifiée et nous n'aurons donc pas à nous préoccuper de cette notion de mesurabilité et, pour nous, une VA sera simplement une fonction de $\Omega$ dans $\mathbb{R}$ (pour plus de précisions, on pourra consulter le cours de licence / master sur l'intégrale de Lebesgue).

**Ex1:** Loi de Bernoulli.

Le jeu a deux issues possibles: pile ou face que nous noterons $\pi$ et $\phi$.
On a donc $\Omega = \{\pi, \phi\}$ et $X(\Omega) = \{0, 1\}$
$X$ est définie par $X(\pi) = 1$ et $X(\phi) = 0$

**Ex4:** Lancer de deux dés.

On considère l'expérience aléatoire consistant à lancer simultanément deux dés discernables (un bleu et un vert) et à noter le n° du bleu puis celui du vert. Une issue de l'expérience est donc un couple d'entiers compris entre 1 et 6. Ainsi $\Omega = \{(i, j) / \ i, j = 1, 2, ..., 6\}$ est l'ensemble des 36 couples possibles et l'on peut poser l'équiprobabilité sur cet ensemble.

Soit $X$ la variable aléatoire égale à la somme des numéros des 2 dés. On a $X(i, j) = i + j$

$X$ est donc définie par $\begin{array}{rcl} X:\Omega & \longrightarrow & \mathbb{R} \\ \omega=(i,j) & \longmapsto & X(\omega)=i+j \end{array}$

$X(\Omega)=\{2,3,4,...,12\}$ et l'on a, par exemple,

$[X=4]$ = "la somme des numéros vaut 4" = $\{(1,3),(2,2),(3,1)\}$ ou
$[10 \leq X \leq 11]$ = "la somme des numéros est comprise entre 10 et 11" = $\{(4,6),(5,5),(6,4),(5,6),(6,5)\}$

La loi de $X$ est la donnée du tableau ci-dessous:

| $k$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $P[X=k]$ | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |

Nous pourrions également définir la variable aléatoire $S$ égale au maximum des deux lancers. De la même façon que ci-dessus, la loi de cette variable serait donnée par le tableau ci-dessous:

| $k$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P[X=k]$ | 1/36 | 3/36 | 5/36 | 7/36 | 9/36 | 11/36 |

Enfin, on pourrait définir des variables aléatoires comme le produit ou le quotient des deux lancers, l'inverse du premier lancer, le sinus du second, etc.

Nous retiendrons des exemples précédents le fait que:

$\boxed{\text{La loi d'une variable aléatoire discrète est la donnée de } \mathbb{P}[X=k] \text{ pour } k \in X(\Omega)}$

C'est la définition que nous prendrons de la loi d'une VA; on peut la définir de façon plus rigoureuse:

$$\begin{array}{rcl} \mathbb{P}_X:\mathcal{P}(\Omega) & \rightarrow & [0,1] \\ A & \mapsto & \mathbb{P}_X(A)=\mathbb{P}(X^{-1}(A))=\mathbb{P}[X\in A] \end{array}$$

Nous avons en fait "transporté" sur $\mathbb{R}$ la probabilité $\mathbb{P}$ initialement définie sur $\Omega$.

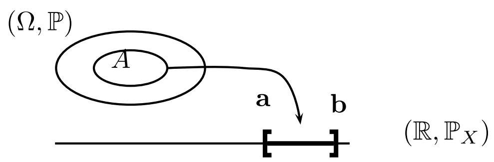


Figure 23.2: Loi d'une variable aléatoire

### 23.1.2 Caractéristique d'une loi

**Espérance (ou valeur moyenne)**

DÉFINITION 87

*Soit $X$ une variable aléatoire sur $\Omega$. On suppose que $X(\Omega)=\{x_1,x_2,...,x_n\}$*

*L'espérance de $X$ est le réel noté $\mathbb{E}[X]$ et défini par $\mathbb{E}[X]=\sum_{k=1}^{n} x_k \times \mathbb{P}[X=x_k]$*

$\mathbb{E}[X]$ représente la valeur moyenne de la variable aléatoire pondérée par les probas.

**Ex0:** Pile ou face équitable.

$\mathbb{E}[X]=0\times\mathbb{P}[X=0]+1\times\mathbb{P}[X=1]=1/2$ qui représente le gain moyen après un tirage.

**Ex1:** Loi de Bernoulli de paramètre $p$.

$\mathbb{E}[X]=0\times\mathbb{P}[X=0]+1\times\mathbb{P}[X=1]=p$

**Ex2:** Lancer d'un dé normal à six faces.

$\mathbb{E}[X]=\sum_{k=1}^{6} k\times\mathbb{P}[X=k]=1\times\mathbb{P}[X=1]+...+6\times\mathbb{P}[X=6]=7/2$

Rappelons au passage que $1+2+...+n = \dfrac{n(n+1)}{2}$ et $1^2+2^2+...+n^2 = \dfrac{n(n+1)(2n+1)}{6}$

PROPRIÉTÉ 81

*Soient $X$ et $Y$ deux variables aléatoires sur $\Omega$, soit $\lambda \in \mathbb{R}$*
*Alors $X+Y$ et $\lambda X$ sont des variables aléatoires et l'on a:*
- $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- $\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X]$
- $\mathbb{E}[\lambda] = \lambda$

DÉMO
La linéarité de l'espérance est une conséquence immédiate de la linéarité de la somme.
Par ailleurs, $\mathbb{E}[\lambda] = \sum_{\omega \in \Omega} \lambda \mathbb{P}[X=\omega] = \lambda \sum_{\omega \in \Omega} \mathbb{P}[X=\omega] = \lambda \times 1 = \lambda$
□

**Moments, variance et écart type.**

Nous commencerons par un théorème très important qui permet de déterminer la loi de la variable aléatoire $f(X)$ pour n'importe quelle fonction (mesurable) $f$.

THÉORÈME 76 (THÉORÈME DE TRANSFERT)

*Soit $X$ une variable aléatoire définie sur $(\Omega, \mathbb{P})$*
*Soit $f$ une fonction mesurable.*

$$\mathbb{E}[f(X)] = \sum_{k=1}^{n} f(x_k)\mathbb{P}[X = x_k]$$

En particulier, si $f(x) = x^k$, $\mathbb{E}[f(X)] = \mathbb{E}[X^k]$ s'appelle le moment d'ordre $k$ de la variable aléatoire.
Si $k=1$, le moment d'ordre 1 n'est rien d'autre que l'espérance.
Si $k=2$, $\mathbb{E}[X^2]$ est le moment d'ordre 2:

$$\mathbb{E}[X^2] = \sum_{k=1}^{n} x_k^2 \times P[X = x_k]$$

Cette quantité va nous servir à calculer la variance à l'aide d'une formule plus utile.

DÉFINITION 88

$$var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$
$$\sigma(X) = \sqrt{var(X)}$$

La variance $\text{var}(X)$ est la moyenne des carrés des écarts à la moyenne et caractérise la dispersion de la variable aléatoire. L'écart type $\sigma(X)$ mesure le même phénomène mais avec une valeur homogène aux quantités de départ ($\text{var}(X)$ et $\sigma(X)$ sont toujours $\geq 0$).

PROPRIÉTÉ 82

$$var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

DÉMO
$\text{var}(X) = \mathbb{E}\left[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2\right] = \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2$ En effet, $\mathbb{E}[X]$ est un nombre réel et son espérance est donc égal à lui-même. On en déduit la formule.
□

Propriété 83

*Soient $X$ et $Y$ deux variables aléatoires et soit $\lambda \in \mathbb{R}$. On a:*
- $var(X+\lambda) = var(X)$
- $var(\lambda X) = \lambda^2 \times var(X)$
- $var(X) \geq 0$
- $var(\lambda) = 0$
- *$X$ et $Y$ indépendants* $\Rightarrow var(X+Y) = var(X) + var(Y)$

Les propriétés de l'écart type découlent du fait qu'il s'agit de la racine carrée de la variance.

Démo

- $\mathrm{var}(X+\lambda) = \mathbb{E}[(X+\lambda)^2] - \mathbb{E}[(X+\lambda)]^2 = \mathbb{E}[X^2] + 2\lambda\mathbb{E}[X] + \lambda^2 - \mathbb{E}[X]^2 - 2\lambda\mathbb{E}[X] - \lambda^2 = \mathrm{var}(X)$
- $\mathrm{var}(\lambda X) = \mathbb{E}[(\lambda X)^2] - \mathbb{E}[\lambda X]^2 = \lambda^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2) = \lambda^2\mathbb{E}[X]$
- $\mathrm{var}(\lambda) = \mathbb{E}[\lambda^2] - \mathbb{E}[\lambda]^2 = \lambda^2 - \lambda^2 = 0$
- $\mathrm{var}(X+Y) = \mathbb{E}[(X+Y)^2] - \mathbb{E}[X+Y]^2 = \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[XY] - \mathbb{E}[X]^2 - \mathbb{E}[Y]^2 - 2\mathbb{E}[X]\mathbb{E}[Y]$

Mais si $X$ et $Y$ sont indépendants, $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \ \Rightarrow\ \mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)$

□

**Ex1:** Loi de Bernoulli de paramètre $p$. $\mathbb{E}[X^2] = p \Rightarrow \mathrm{var}(X) = p - p^2 = p(1-p) = pq$ et $\sigma(X) = \sqrt{pq}$

**Ex2:** Lancer d'un dé.

$$\mathbb{E}[X^2] = \sum_{k=1}^{6} k^2/6 = 6 \times 7 \times 13/36 = 91/6 \Rightarrow \mathrm{var}(X) = 91/6 - (7/2)^2 = 35/12$$

**Fonction de répartition**

Définition 89

*Soit $X$ une variable aléatoire sur un ensemble $\Omega$.*
*On note $F$ la fonction dé répartition de $X$. Elle est définie par:*

$$\begin{array}{rcl} F : \mathbb{R} & \longrightarrow & [0,1] \\ x & \longmapsto & F(x) = \mathbb{P}[X \leq x] \end{array}$$

Notons $x_1$, $x_2$, ..., $x_k$ les valeurs possibles de $X$ et $p_1$, $p_2$, ..., $p_k$ leur proba respective.
Si les $x_k$ sont rangés par ordre croissant, $X \leq x_k \iff X = x_1$ ou $X = x_2$ ou... $X = x_k$
de sorte que $[X \leq x_k] = [X = x_1] \cup [X = x_2] \cup ... \cup [X = x_k]$ et donc
$F(x_i) = \mathbb{P}[X \leq x_i] = p_1 + p_2 + ...p_i$

**Ex1:** Loi de Bernoulli

$X$ ne peut prendre que deux valeurs 0 et 1. $X$ ne peut donc être négatif donc $x < 0 \ \Rightarrow\ F(x) = 0$
De même, $X$ est forcément plus petit que 1 donc $F(x) = 1$ si $x \geq 1$
Enfin, si $0 \leq x < 1$, la seule valeur que peut prendre $X$ est 0 donc $F(x) = \mathbb{P}[X = 0] = q$
$F$ est donc une fonction en escalier à trois paliers.

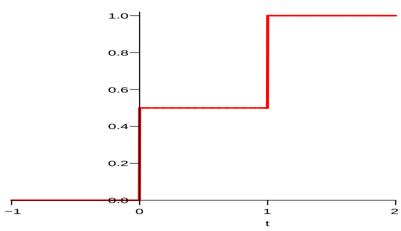


Figure 23.3: Fonction de répartition d'une loi de Bernoulli

**Ex2:** Lancer d'un dé à six faces

De même que dans l'exemple précédent, $X(\Omega) = \{1, 2, ..., 6\}$.
$F(x) = 0$ si $x < 1$, $F(x) = 1$ si $x \geq 6$ et $F(x) = \dfrac{k}{6}$ si $x \in [k, k+1[$, $k = 1, ..., 5$
là encore, il s'agit d'une fonction en escalier avec sept paliers.

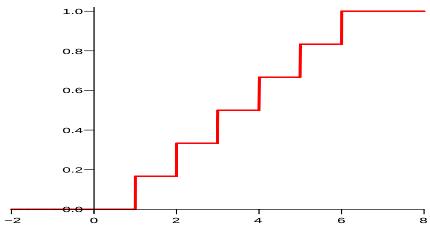


Figure 23.4: Fonction de répartition d'une loi équiprobable

Propriété 84

> *Soit $F$ la fonction de répartition d'une variable aléatoire discrète $X$:*
> - *$F$ est croissante sur $\mathbb{R}$*
> - *$F$ est continue à droite*
> - $\lim\limits_{x \to -\infty} F(x) = 0$
> - $\lim\limits_{x \to +\infty} F(x) = 1$

Démo

- Il s'agit de montrer que si $x < y$ alors $F(x) < F(y)$

Si $X \leq x$ alors $X \leq y$ donc $[X \leq x] \in [X \leq y]$ d'après les propriétés des probabilités.
Puisque $F(x) = \mathbb{P}[X \leq x]$ et $F(y) = \mathbb{P}[Y \leq y]$ on a bien $F(x) < F(y)$ et $F$ est croissante.

- $\lim\limits_{h \to 0^+} F(x+h) = \lim\limits_{h \to 0^+} \mathbb{P}(]-\infty, x+h]) = \mathbb{P}(]-\infty, x])$ d'où la continuité à droite

- On se limite au cas où la variable ne prend qu'un nombre fini de valeurs $x_1, x_2, ..., x_n$

Si $x < x_1$, alors $[X \leq x] = \emptyset$
Si $x \geq x_n$, alors $[X \leq x] = \Omega$
Donc, lorsque $x$ tend vers $-\infty$, $F(x) = \mathbb{P}[X \leq x]$ tend vers $\mathbb{P}(\emptyset) = 0$
Lorsque $x$ tend vers $+\infty$ $F(x)$ tend vers $\mathbb{P}(\Omega) = 1$.
□

Propriété 85 (Propriété fondamentale)

> $\mathbb{P}[X = k] = \lim\limits_{x \to k^+} F(x) - \lim\limits_{x \to k^-} F(x)$

Autrement dit, les points de discontinuité $k$ de la fonction de répartition sont ceux où la probabilité de $X$ est non nulle et cette proba est égale à la "hauteur" du palier.

Démo

$\mathbb{P}[X = k] = \mathbb{P}(]-\infty, k]) - \mathbb{P}(]-\infty, k[) = \lim\limits_{x \to k^+} F(x) - \lim\limits_{x \to k^-} F(x)$
□

## 23.2 Loi binomiale et loi de Poisson

### 23.2.1 Loi binomiale

Considérons $n$ variables aléatoires indépendantes $X_1, X_2, ..., X_n$ de même loi de Bernoulli de paramètre $p$. Cela revient à considérer $n$ lancers à pile ou face indépendants les uns des autres pour lesquels la probabilité de faire pile est $p$ (et celle de faire face est $q = 1 - p$).

Soit $X$ la variable aléatoire égale au nombre de piles obtenus au cours de ces $n$ tirages.
$X$ est une variable aléatoire dont les valeurs possibles sont $0, 1, 2, ..., n$ et l'on voit que
$X = X_1 + X_2 + ... + X_n$

L'évènement $[X = k]$ se lit "au cours des $n$ expériences, on a obtenu $k$ fois pile". Cherchons sa loi de probabilité:

La probabilité d'obtenir pile au cours d'un lancer est $p$.
Celle d'obtenir $k$ fois pile est donc $p^k$ par indépendance des lancers (cf. cours de l'anne passée).
Si pile est apparu $k$ fois, alors face est apparu $n - k$ fois ce qui se produit avec une probabilité $(1-p)^{n-k}$

Maintenant, les $k$ piles peuvent être placés de $C_n^k$ façons différentes (place de $k$ objets parmi $n$). On obtient ainsi le résultat suivant:

DÉFINITION 90

> *On appelle variable aléatoire binomiale de paramètres $n \in \mathbb{N}$ et $p \in [0,1]$,*
> *une variable aléatoire $X$ dont la loi est:*
> $\mathbb{P}[X = k] = C_n^k p^k q^{n-k}$ *si* $k = 0, 1, ...n$
> $\mathbb{P}[X = k] = 0$ *sinon*
> *Nous noterons* $X \sim \mathcal{B}(n,p)$

On a donc $X(\Omega) = \{0, 1, 2, ..., n\}$

**Ex5:** Loi binomiale $p = 1/2$ et $n = 4$

$\mathbb{P}[X = 1] = C_4^1(1/2)(1/2)^3 = 1/8$ et $\mathbb{P}[X = 4] = C_4^4(1/2)^4 = 1/16$

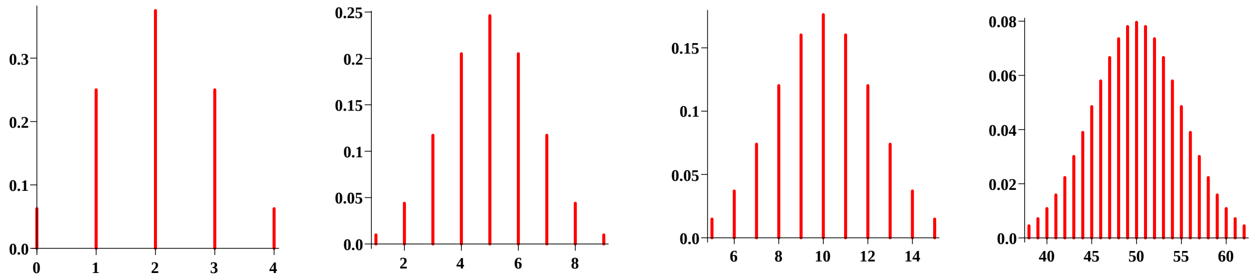


Figure 23.5: Distributions binomiales avec $p = 0.5$ et $n = 4, 10, 20$ et $100$.

⚠ Remarque importante:

$X \sim \mathcal{B}(n,p)$ est égal au nombre de succès au cours de $n$ tirages à pile ou face indépendants avec probabilité de succès $p$ à chaque tirage.

PROPRIÉTÉ 86

> *Soit* $X \sim \mathcal{B}(n,p)$ *et* $q = 1 - p$
> - $\mathbb{E}[X] = np$
> - $var(X) = npq$
> - $\sigma(X) = \sqrt{npq}$

DÉMO

- Si $X$ est une VA binomiale alors on peut écrire $X = X_1 + X_2 + ... + X_n$
où les $X_k$ sont des variables de Bernoulli indépendantes de même paramètre $p$.
Par linéarité de l'espérance on a donc:
$\mathbb{E}[X] = \mathbb{E}[X_1 + ... + X_n] = \mathbb{E}[X_1] + ...\mathbb{E}[X_n] = np$
- de même, par indépendance de ces variables, on a aussi:
$\text{var}(X) = \text{var}(X_1) + ... + \text{var}(X_n) = npq$
□

### 23.2.2 La loi de Poisson

DÉFINITION 91

> *On appelle variable aléatoire de Poisson de paramètre $\lambda > 0$ une variable $X$ telle que:*
> $\mathbb{P}[X = k] = e^{-\lambda}\dfrac{\lambda^k}{k\,!}\ \forall k \in \mathbb{N}$
> $\mathbb{P}[X = k] = 0$ *sinon*

Cette loi est (très) utilisée pour modéliser les phénomènes de files d'attente ou le nombre d'apparitions d'un phénomène dans un laps de temps donné ($\lambda$ représente alors la fréquence du phénomène).

Par exemple, le nombre d'appels reçus dans un central ou le nombre d'utilisateurs dans un réseau durant une période de l'année. Il s'agit d'une loi discrète, mais contrairement aux précédentes, elle peut prendre une infinité de valeurs puisque $X(\Omega) = \mathbb{N}$

Propriété 87

*Si $X$ suit une loi de Poisson alors $\mathbb{E}[X] = var(X) = \lambda$*

Démo

$\mathbb{E}[X] = \sum_{k\geq 0} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k\geq 0}\frac{\lambda^{k-1}}{(k-1)!} = \lambda$

$\mathbb{E}[X(X-1)] = \sum_{k\geq 2} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} = \lambda^2 e^{-\lambda}\sum_{k\geq 2}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$

Ainsi, $\mathbb{E}[X(X-1)] = \mathbb{E}[X^2] - \mathbb{E}[X] = \lambda^2 \ \Rightarrow\ \mathbb{E}[X^2] = \lambda^2 + \lambda \ \Rightarrow\ \text{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda$ □

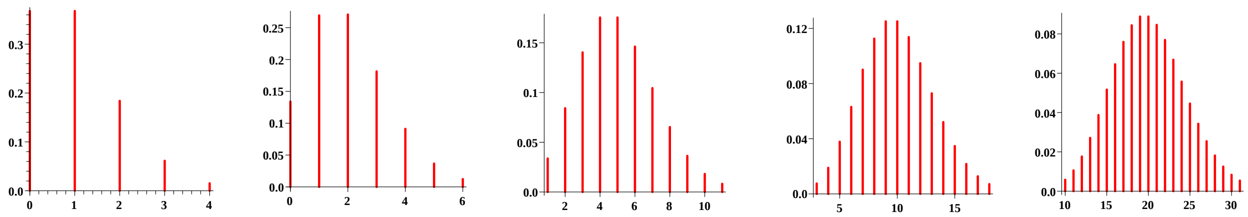

Figure 23.6: Distributions de Poisson avec $\lambda = 1, 2, 5, 10$ et 20.

Nous admettrons le fait que la loi de Poisson est la seule loi discrète dont l'espérance est égale à la variance.

**Ex6:**

Le nombre de personnes $X$ faisant la queue à une caisse de supermarché à un instant donné est une variable de Poisson de paramètre $\lambda = 4$.
On a $\mathbb{P}[X = 1] = e^{-4} \times 4^1/1! \approx 0,07$

### 23.2.3 Approximation entre lois discrètes

Soit $X$ une variable aléatoire suivant une loi binomiale de paramètres $n = 50$ et $p = 0,04$.
Soit $Y$ une variable suivant une loi de Poisson de paramètre $\lambda = 2$.
Comparons leurs probabilités respectives pour quelques valeurs de $k$:

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $\mathbb{P}[X = k]$ | 0,13 | 0,271 | 0,276 | 0,184 | 0,090 | 0,0035 |
| $\mathbb{P}[Y = k]$ | 0,135 | 0,271 | 0,271 | 0,18 | 0,09 | 0,0036 |

Le tableau montre que ces lois sont très voisines (cf. figure).

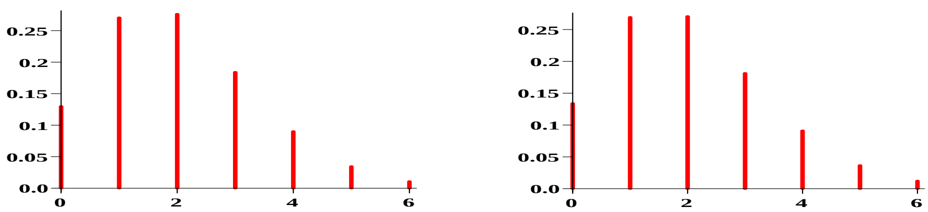

Figure 23.7: Distributions binomiale $n = 50$, $p = 0.04$ et poissonienne $\lambda = 2$

Pour de grandes valeurs de $n$, le calcul des coefs binomiaux est complexe et coûteux en temps de calcul et il peut être avantageux d'approcher la loi binomiale par une loi de Poisson plus facile à implémenter.

La loi de Poisson a d'ailleurs été découverte comme approximation de la binomiale lorqsue $n$ est grand et $p$ est petit devant $n$. Le théorème précisant cette approximation s'appelle la loi des évènements rares.

Théorème 77 (Loi des évènements rares)

*Soit* $(p_n)_{n\geq 0}$ *une suite de nombres de* $]0,1[$

*Soit* $S_n$ *une variable aléatoire de loi binomiale* $\mathcal{B}(n,p_n)$ *On suppose que* $\lim\limits_{n\to+\infty} np_n = \lambda$

*Alors* $\lim\limits_{n\to+\infty} \mathbb{P}[S_n = k] = e^{-\lambda}\dfrac{\lambda^k}{k!}$

Démo

$\mathbb{P}[S_n = k] = C_n^k p^k q^{n-k} = \dfrac{1}{k!}n(n-1)...(n-k+1)\dfrac{(np)^k}{n^k}q^{-k}q^n \sim \dfrac{\lambda^k}{k!}q^n$

avec $\ln(q^n) = n\ln(1-p) \sim n\dfrac{-\lambda}{n} = -\lambda \;\Rightarrow\; \mathbb{P}[S_n = k] \to \dfrac{\lambda^k}{k!}e^{-\lambda}$

□

Une autre démonstration sera effectuée en TD.

On peut même préciser la vitesse de convergence vers la loi de Poisson grâce à l'inégalité suivante (admise):

$$\sum_{k=0}^{+\infty} |\mathbb{P}[S_n = k] - e^{-\lambda}\frac{\lambda^k}{k!}| \leq \frac{2\lambda}{n} min(2,\lambda)$$

Dans la pratique, on appliquera ce théorème pour $p$ petit et $n$ grand. On retiendra donc le théorème d'approximation suivant:

Théorème 78

*Soit* $X$ *une variable aléatoire suivant une loi binomiale de paramètres* $n$ *et* $p$

*Soit* $Y$ *une loi de Poisson de paramètre* $\lambda = np$

*Si* $n \geq 50$ *et* $np \leq 5$ *alors on peut approcher la loi de* $X$ *par celle de* $Y$*, ie:*

$\mathbb{P}[X = k] \approx P[\mathbb{P} = k] \;\forall k \in \mathbb{N}^\star$

Les probabilités de la loi de Poisson sont données par des tables pour différentes valeurs de $\lambda$.

## 23.3 Lois conjointes

### 23.3.1 Vecteur aléatoire, loi conjointe et loi marginale

On souhaite examiner simultanément deux aspects d'une même expérience aléatoire. On définit un vecteur aléatoire comme un couple de variables aléatoires sur un même espace de probabilité $\Omega$

Définition 92

*Soient* $X$ *et* $Y$ *deux va réelles sur un espace de probabilité* $\Omega$

*On suppose que* $X(\Omega) = \{x_k, k \in \mathbb{N}\}$ *et* $Y(\Omega) = \{y_i, i \in \mathbb{N}\}$ *et l'on pose* $Z = (X,Y)$

$Z$ *est une variable aléatoire à valeurs dans* $\mathbb{R}^2$ *dont la loi est appelée loi conjointe de* $X$ *et* $Y$

*Les lois des coordonnées* $X$ *et* $Y$ *s'appellent lois marginales de* $Z$

La loi conjointe est caractérisée par la donnée de $\boxed{\mathbb{P}[X = x_k, Y = y_i] = \mathbb{P}([X = x_k] \cap [Y = y_i]) \;\forall k,i \in \mathbb{N}}$

de telle sorte que $\displaystyle\sum_{k\geq 0}\sum_{i\geq 0} \mathbb{P}[X = x_k, Y = y_i] = 1$

Si $\Omega$ est fini, $X$ et $Y$ prennent un nombre fini de valeurs et la loi peut être donnée par un tableau à deux dimensions.

Les lois marginales s'obtiennent grâce à la formule des probabilités totales.
**Ex:** $\Omega = \{0,1\}$

| $X \setminus Y$ | 0 | 1 | TOT |
|---|---|---|---|
| 0 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| 1 | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{2}$ |
| TOT | $\frac{1}{4}$ | $\frac{3}{4}$ | 1 |

On a ainsi $\mathbb{P}[X=0,Y=1]=\dfrac{1}{2}$ et $\mathbb{P}[X=1,Y=1]=\dfrac{1}{4}$
Les lois marginales en $X$ et $Y$ sont données par la dernière ligne et la dernière colonne.
On a $\mathbb{P}[X=0]=\dfrac{1}{2}$ et $\mathbb{P}[Y=1]=\dfrac{3}{4}$

**Ex2:** Sur $\Omega=\mathbb{N}$ on pose $\mathbb{P}[X=k,Y=i]=\dfrac{1}{e^2}\dfrac{1}{i!k!}$

Cette formule définit bien une loi de probabilité car $\displaystyle\sum_{k\geq 0}\sum_{i\geq 0}\mathbb{P}[X=k,Y=i]=\frac{1}{e^2}(\sum_{k\geq 0}\frac{1}{k!})(\sum_{i\geq 0}\frac{1}{i!})=1$

La loi marginale en $X$ est donnée par $\displaystyle\mathbb{P}[X=k]=\sum_{i\geq 0}\frac{1}{e^2k!i!}=\frac{1}{e\times k!}$ (loi de Poisson)

### 23.3.2 Covariance

Définition 93

> *La covariance de deux variables aléatoires $(X,Y)$ est le réel*
> $cov(X,Y)=\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]=\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]$
> *avec* $\displaystyle\mathbb{E}[XY]=\sum_{i,k\geq 0}x_kx_i\mathbb{P}[X=k,Y=i]$

la covariance est un terme qui mesure la dépendance, la corrélation entre les deux variables.

**Ex1:** $\text{cov}(X,Y)=-\dfrac{1}{8}$

Propriété 88

> *Soient $X$ et $Y$ deux variables aléatoires de carré intégrable*
> - $var(X+Y)=var(X)+var(Y)+2cov(X,Y)$
> - $cov(X,X)=var(X)$

Définition 94 (Rappels sur l'indépendance)

> *Deux évènements $A$ et $B$ sont indépendants si* $\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(B)$
> *Deux variables aléatoires sont indépendantes si* $\mathbb{P}([X\in A]\cap[Y\in B])=\mathbb{P}[X\in A]\times\mathbb{P}[Y\in B]\ \forall A,B\subset\mathbb{R}$

Lorsque deux évènements sont indépendants, la réalisation de l'un n'influe pas sur l'autre.

Propriété 89

> *$X$ et $Y$ stochastiquement indépendants* $\Rightarrow$ $\mathbb{E}[XY]=\mathbb{E}[X]\times\mathbb{E}[Y]$ *et* $cov(X,Y)=0$
> *La réciproque est fausse*

Démo
$\displaystyle\mathbb{E}[XY]=\sum_{i,k\geq 0}x_kx_i\mathbb{P}([X=x_k,Y=x_i])=\sum_{i,k}x_kx_i\mathbb{P}[X=x_k]\times\mathbb{P}[Y=x_i]$ par indépendance
$\displaystyle=\sum_{k\geq 0}x_k\mathbb{P}[X=x_k]\sum_{i\geq 0}x_i\mathbb{P}[Y=x_i]=\mathbb{E}[X]\mathbb{E}[Y]$
$\text{cov}(X,Y)=\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]=0$
□

---

# Chapitre 24

# Variables aléatoires continues

## 24.1 Densité d'une loi continue

### 24.1.1 Introduction

La touche rand d'une calculatrice ou la commande random de Matlab permettent d'obtenir un nombre aléatoire $X$ compris entre 0 et 1. En principe, n'importe quelle valeur de $[0,1]$ a, à priori, même probabilité d'apparaître (en pratique un ordinateur ne peut représenter qu'un nombre fini de nombres qui ne sont même pas réellement aléatoires. Nous reviendrons plus tard là-dessus). Puisque $[0,1]$ contient une infinité non dénombrable de réels, $X$ n'est pas discrète. On dit qu'il s'agit d'une variable aléatoire continue dont la loi sera caractérisée non plus par un tableau mais par une fonction appellée densité.

De la même façon, si l'on observe les désintégrations dans un atome radioactif, la durée entre deux désintégrations successives est aléatoire et prend ses valeurs dans l'intervalle $[0,+\infty[$. Elle se représente également à l'aide d'une variable aléatoire continue (dont la loi est exponentielle).

Si l'on choisit au hasard une personne dans une population donnée et que l'on note $X$ sa taille en cm, on obtient une variable aléatoire continue puisqu'elle peut prendre n'importe quelle valeur d'un intervalle, par exemple $[150,200]$, avec une plus forte probabilité de se trouver aux alentours de 167. Les statistiques montrent que $X$ suit une loi normale dont la densité est une fonction gaussienne que nous étudierons.

### 24.1.2 Définition

DÉFINITION 95

*Une fonction $f$ de $\mathbb{R}$ dans $\mathbb{R}$ est la densité de probabilité d'une loi continue si elle vérifie:*

- $f(x) \geq 0 \ \forall x \in \mathbb{R}$
- *$f$ est intégrable.*
- $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1$

Le support de la fonction de densité est l'ensemble des $x$ pour lesquels $f(x) \neq 0$

**Ex1:** Loi uniforme sur $[\alpha,\beta]$

On dit qu'une variable aléatoire $X$ suit une loi uniforme sur l'intervalle $[\alpha,\beta]$ si sa densité est $f(x) = \dfrac{1}{\beta-\alpha} \times \mathbb{1}_{[\alpha,\beta]}(x)$

Cette loi modélise le tirage d'un nombre compris entre $\alpha$ et $\beta$, tous les nombres ayant même probabilité d'apparaître. Le support de la fonction est l'intervalle $[\alpha,\beta]$

**Ex2:** Loi exponentielle

On dit que $X$ est une loi exponentielle de paramètre $\lambda > 0$ notée $\mathcal{E}(\lambda)$ si sa densité est $f(x) = \lambda e^{\lambda x}\mathbb{1}_{[0,+\infty[}(x)$

Cette loi est à support dans $\mathbb{R}_+$. Elle modélise par exemple la durée entre deux appels successifs dans un central téléphonique ou bien la durée entre deux désintégrations dans un atome.

**Ex3:** Loi de Laplace

$X$ suit une loi de Laplace si sa densité est $f(x) = \dfrac{1}{2}e^{-|x|}$
Il s'agit d'une loi définie sur $\mathbb{R}$ tout entier.

**Ex4:** Loi de Cauchy

Sa densité est $f(x) = \dfrac{1}{\pi}\dfrac{1}{1+x^2}$ et son support est donc aussi $\mathbb{R}$

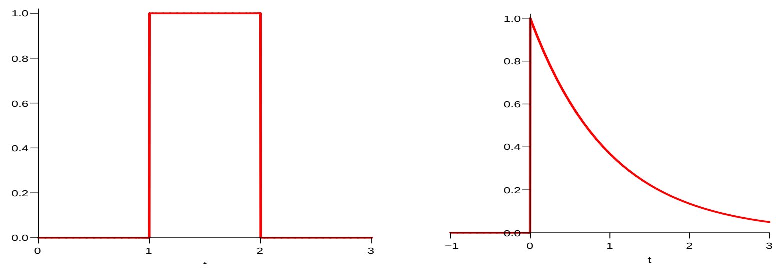

Figure 24.1: Densité uniforme et expontentielle

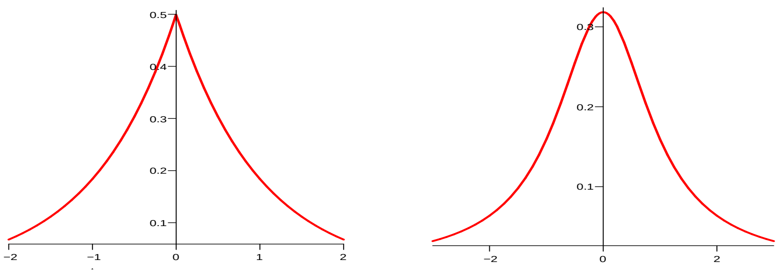

Figure 24.2: Densité de Laplace et de Cauchy

**Ex5:** Loi triangulaire

C'est une loi définie sur l'intervalle $[-1,1]$ de densité $f(x) = (1-|x|)\mathbb{1}_{[-1,1]}(x)$
A titre d'exercice, démontrer que cette loi définit bien une densité de probabilité.

Remarque 20
*Pour qu'une fonction positive et continue par morceaux soit une densité de probabilité, il est nécessaire qu'elle soit intégrable sur $\mathbb{R}$ (ce qui arrive, par exemple, lorsqu'elle est nulle à l'extérieur d'un intervalle donné). Si l'intégrale ne converge pas, on peut tout de même construire une loi de probabilité avec la fonction en la tronquant à un intervalle donné sur lequel son intégrale vaut 1.*

Propriété 90

*Soit $X$ une variable aléatoire de densité $f$ et soient $a$ et $b$ deux réels avec $a < b$*

*On pose $\mathbb{P}[a \leq X \leq b] = \displaystyle\int_a^b f(x)dx$*

*La probabilité est égale à l'aire comprise sous la courbe entre les droites $x = a$ et $x = b$*

Ainsi, on aura $\forall \alpha \in \mathbb{R},\ \mathbb{P}[X \leq \alpha] = \displaystyle\int_{-\infty}^{\alpha} f(x)dx$

En posant $a = b$ dans l'intégrale on en déduit le résultat paradoxal suivant:

$$\boxed{\text{Si } X \text{ est une loi continue de densité } f\text{, alors } \mathbb{P}[X = \alpha] = 0 \quad \forall \alpha \in \mathbb{R}}$$

**Ex1:** Loi uniforme

Soit $X$ suivant une loi uniforme sur $[\alpha, \beta]$. Calculons la probabilité que $X \in [0, \beta]$

$\mathbb{P}[0 < X < \beta] = \displaystyle\int_0^{\beta} f(x)dx = \frac{\beta}{\beta - \alpha}$

**Ex2:** Loi de Cauchy

Soit $X$ suivant une loi de Cauchy. Calculons $I = \mathbb{P}[X > 1]$ et $J = \mathbb{P}[-1 < X < 1]$

$$I = \int_1^{+\infty} f(x)dx = \frac{1}{\pi}\int_1^{+\infty} \frac{dx}{1+x^2} = \frac{1}{\pi}[\arctan(x)]_1^{+\infty} = \frac{1}{\pi}(\lim_{x\to+\infty} \arctan x - \frac{\pi}{4}) = \frac{1}{4}$$

$$J = \int_{-1}^{1} f(x)dx = \frac{2}{\pi}[\arctan(x)]_{-1}^{1} = \frac{2}{\pi}\times\frac{\pi}{4} = \frac{1}{2}$$

On peut aussi écrire $J = 1 - \mathbb{P}[X > 1] - \mathbb{P}[X < -1] = 1 - 2\mathbb{P}[X > 1] = \dfrac{1}{2}$

### 24.1.3 Sommes de deux variables continues

Théorème 79

*Soient $X$ et $Y$ deux variables aléatoires continues et indépendantes de densité $f$ et $g$*

*Soit $Z = X + Y$*

*La densité $h$ de $Z$ est la convolée de $f$ et $g$:* $h(s) = f * g(s) = \displaystyle\int_{-\infty}^{+\infty} f(s-x)g(x)dx$

**Rappels sur le produit de convolution**

On pourra se reporter à la fin du chapitre sur les transformées de Fourier pour les propriétés du produit de convolution. Rappelons simplement que:

• Si $f$ et $g$ sont des densités, alors $f * g$ existe et est aussi une densité de probabilité

• Le produit de convolution est linéaire en $f$ et en $g$ , commutatif et associatif

• $f * g$ représente un mélange entre $f$ et $g$, plus régulière que $f$ et $g$ et correspondant à une sorte de filtrage entre les deux fonctions.

**Ex:**

Soient $X$ et $Y$ deux variables indépendantes de même loi exponentielle de paramètre $\lambda$ et soit $Z = X + Y$

La densité de $Z$ est la fonction $h$ définie par

$$h(s) = \int_{-\infty}^{+\infty} f(s-x)f(x)dx = \int_{-\infty}^{+\infty} \lambda e^{-\lambda(s-x)}\mathbb{1}_{[0,+\infty[}(s-x)\lambda e^{-\lambda x}\mathbb{1}_{[0,+\infty[}(x)dx$$

$$= \lambda^2\mathbb{1}_{[0,+\infty[}(s)\int_0^s e^{-\lambda s}ds = \lambda^2 s e^{-\lambda s}\mathbb{1}_{[0,+\infty[}(s)$$

## 24.2 Caractéristiques d'une loi continue

### 24.2.1 Espérance et variance

Les formules et propriétés ci dessous sont les mêmes que pour des variables discrètes mais en remplaçant les sommes par des intégrales

Définition 96

*Si $X$ est une loi de densité $f$*

• $\mathbb{E}[X] = \displaystyle\int_{-\infty}^{+\infty} xf(x)dx$

• $\mathbb{E}[X^2] = \displaystyle\int_{-\infty}^{+\infty} x^2 f(x)dx$

• $var(X) = \displaystyle\int_{-\infty}^{+\infty} (x - \mathbb{E}[X])^2 f(x)dx$

**Ex1:** Loi uniforme sur $[\alpha, \beta]$

$$\mathbb{E}[X] = \int_\alpha^\beta \frac{x}{\beta-\alpha}dx = \frac{1}{\beta-\alpha}[\frac{1}{2}x^2]_\alpha^\beta = \frac{\alpha+\beta}{2} \quad \mathbb{E}[X^2] = \frac{1}{3}\frac{1}{\beta-\alpha}(\beta^3-\alpha^3) = \frac{\alpha^2+\alpha\beta+\beta^2}{3} \Rightarrow \text{var}(X) = \frac{(\beta-\alpha)^2}{12}$$

**Ex2:** Loi exponentielle $\mathcal{E}(\lambda)$

$\mathbb{E}[X] = \lambda \displaystyle\int_0^{+\infty} xe^{-\lambda x}dx = \frac{1}{\lambda}$

Par une double intégration par parties, on obtient $\mathbb{E}[X^2] = \dfrac{2}{\lambda^2}$ et donc $\text{var}(X) = \dfrac{1}{\lambda^2}$

**Ex3:** Loi de Cauchy

$\mathbb{E}[X] = \displaystyle\int_{-\infty}^{+\infty} \frac{x}{1+x^2}dx$ mais cette fonction n'est pas intégrable: $\mathbb{E}[X]$ et $\text{var}(X)$ n'existent pas

### 24.2.2 Fonction de répartition

Définition 97

*La fonction de répartition d'une variable continue $X$ de denisté $f$ est la fonction $F$ définie par:*

$$\begin{array}{rcl} F : \mathbb{R} & \longrightarrow & [0,1] \\ s & \longmapsto & F(s) = \mathbb{P}[X \leq s] = \int_{-\infty}^{s} f(x)dx \end{array}$$

Dans le cas d'une variable continue, $\mathbb{P}[X \leq s] = \mathbb{P}[X < s]$

Les propriétés sont les mêmes que pour une variable discrète. En outre,

Propriété 91

- *$F$ est une fonction dérivable et $F' = f$*
- *$F$ est croissante, tend vers 0 en $-\infty$ et vers 1 en $+\infty$*
- $\mathbb{P}[a < X < b] = \displaystyle\int_a^b f(x)dx = F(b) - F(a)$

Cette dernière propriété est la plus importante car elle va seule permettre le calcul de probabilités d'évènements.

**Ex1:** Loi uniforme

$F(s) = \displaystyle\int_{-\infty}^{+\infty} f(x)dx = \int_{\alpha}^{s} \frac{dx}{\beta-\alpha} = \frac{s-\alpha}{\beta-\alpha}$ si $\alpha < s < \beta$

$F(s) = 0$ si $s < \alpha$

$F(s) = 1$ si $s > \beta$

**Ex2:** Loi exponentielle $\mathcal{E}(\lambda)$

$F(s) = \displaystyle\int_{-\infty}^{s} \lambda e^{-\lambda x}dx = 1 - e^{-\lambda s}$ si $s > 0$ et $F(s) = 0$ si $s < 0$

**Ex4:** Loi de Cauchy

$F(s) = \displaystyle\int_{-\infty}^{s} \frac{dx}{\pi(1+x^2)} = \frac{1}{\pi}(\text{arctan}s + \frac{\pi}{2}) = \frac{1}{\pi}\text{arctan}s + \frac{1}{2}$

**Ex5:** Loi de Laplace

$F(s) = \displaystyle\int_{-\infty}^{s} \frac{1}{2}e^{-|x|}dx = \frac{e^s}{2}$ si $s < 0$

$F(s) = \displaystyle\int_{-\infty}^{0} \frac{1}{2}e^{x}dx + \int_0^s \frac{1}{2}e^{-x}dx = \frac{1}{2}(2 - e^{-s})$ si $s > 0$

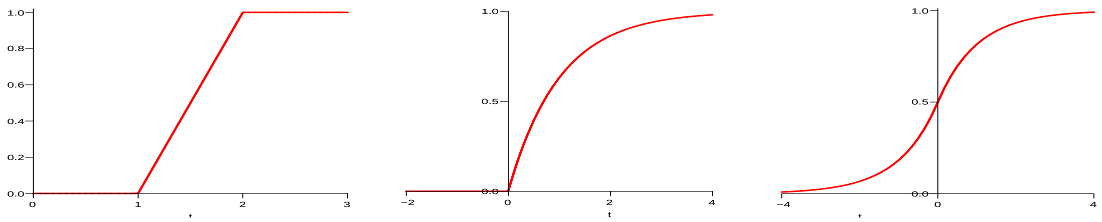

Figure 24.3: Fonctions de répartition: loi uniforme, exponentielle et de Laplace.

## 24.3 Loi normale et applications

### 24.3.1 Densité d'une loi normale

Il s'agit de la loi la plus importante des probabilités et statistiques.

Définition 98

*Une variable aléatoire $X$ suit une loi normale de paramètres $m \in \mathbb{R}$ et $\sigma > 0$ si sa densité est*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-m}{\sigma})^2}$$

On notera $\mathcal{N}(m, \sigma)$ cette loi (qui s'appelle aussi loi de Laplace Gauss). Sa densité est une fonction gaussienne et se représente par une courbe en cloche que nous allons maintenant étudier

**Etude de la fonction gaussienne**

$\mathcal{D}_f = \mathbb{R}$, $f$ est dérivable sur $\mathbb{R}$ et $f'(x) = \dfrac{m-x}{\sigma^3\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-m}{\sigma})^2}$

Ainsi, $f'(x) = 0 \iff x = m$ et $f'(x) > 0 \iff x < m$

| $x$ | $-\infty$ | | $m$ | | $+\infty$ |
|---|---|---|---|---|---|
| $f'$ | | $+$ | | $-$ | |
| $f$ | $0$ | $\nearrow$ | $\frac{1}{\sigma\sqrt{2\pi}}$ | $\searrow$ | $0$ |

Lorsque $m = 0$, on dit que la gaussienne est centrée et lorsque $\sigma = 1$ on dit qu'elle est réduite.

La droite d'équation $x = m$ est un axe de symétrie pour la courbe représentative de $f$. Ainsi, la fonction est paire si et seulement si la loi est centrée.

Le maximum de la courbe est atteint en $x = m$ (sommet de la cloche) et est inversement proportionnel à la valeur de $\sigma$. Plus $\sigma$ est grand, plus la courbe est basse et aplatie. Lorsque $\sigma$ tend vers 0, le pic tend vers l'infini et la courbe est de plus en plus étroite.
La limite d'une gaussienne lorsque $\sigma$ tend vers 0 n'est pas une fonction. C'est une masse de Dirac (cf. la leçon sur la théorie du signal).

La densité d'un variable aléatoire gaussienne centrée et réduite est la fonction $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$

On note habituellement $Z$ la variable aléatoire de densité $\Phi$.

La densité d'une variable $X$ de loi $\mathcal{N}(m, \sigma)$ s'obtient à partir de $Z$ en posant $X = m + \sigma Z$ ou encore

$$\boxed{Z = \frac{X-m}{\sigma}}$$

Sa densité est alors $f(x) = \dfrac{1}{\sigma}\Phi(\dfrac{x-m}{\sigma})$

Lorsque $m$ augmente le pic de la cloche se décale vers la droite et lorsque $\sigma$ augmente, le pic diminue et la courbe s'écrase en se dilatant.

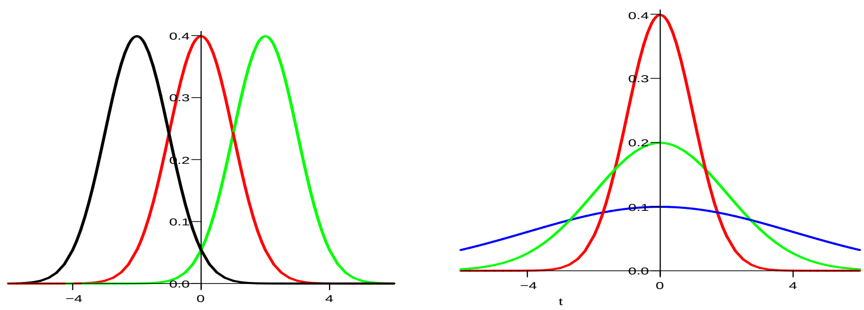


Figure 24.4: Densités normales: $m$ variable puis $\sigma$ variable.

Nous avons vu l'année dernière que $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}dx = \int_{-\infty}^{+\infty}\Phi(x)dx = 1$

ce qui montrer que $\Phi$ est bien une densité de loi de probabilité.

De même, si $X$ suit une loi $\mathcal{N}(m,\sigma)$ de densité $f$, $\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty}\Phi(\frac{x-m}{\sigma})\frac{dx}{\sigma} = 1$ en opérant le changement de variables $s = \frac{x-m}{\sigma}$ dans la dernière intégrale.

### 24.3.2 Caractéristique d'une loi normale

Propriété 92

*Soit $X$ une variable aléatoire de loi $\mathcal{N}(m,\sigma)$*
- $\mathbb{E}[X] = m$  • $var(X) = \sigma^2$

Démo

Il suffit de démontrer que $\mathbb{E}[Z] = 0$ et $\sigma(Z) = 1$ pour une va $Z$ centrée et réduite
- $\Phi$ étant paire, $x\Phi(x)$ est impaire et son intégrale sur $\mathbb{R}$ est nulle $\Rightarrow$ $\mathbb{E}[Z] = 0$
- $\text{var}(Z) = \mathbb{E}[Z^2]$ nous admettrons le résultat car le calcul de cette intégrale dépasse le cadre du cours $\Box$

Définition 99

*La fonction de répartition de la loi normale centrée réduite est:*

$\Pi(s) = \int_{-\infty}^{s}\Phi(x)dx$

On ne peut pas exprimer la primitive de $\Phi$ à l'aide de fonctions usuelles et par conséquent les valeurs de $\Pi$ sont données dans une table appelée table de la loi normale ou table de l'écart réduit (à apprendre par cœur):

| s | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.50000 | 0.50399 | 0.50798 | 0.51197 | 0.51595 | 0.51994 | 0.52392 | 0.52790 | 0.53188 | 0.53586 |
| 0.1 | 0.53983 | 0.54380 | 0.54776 | 0.55172 | 0.55567 | 0.55962 | 0.56356 | 0.56749 | 0.57142 | 0.57535 |
| 0.2 | 0.57926 | 0.58317 | 0.58706 | 0.59095 | 0.59483 | 0.59871 | 0.60257 | 0.60642 | 0.61026 | 0.61409 |
| 0.3 | 0.61791 | 0.62172 | 0.62552 | 0.62930 | 0.63307 | 0.63683 | 0.64058 | 0.64431 | 0.64803 | 0.65173 |
| 0.4 | 0.65542 | 0.65910 | 0.66276 | 0.66640 | 0.67003 | 0.67364 | 0.67724 | 0.68082 | 0.68439 | 0.68793 |
| 0.5 | 0.69146 | 0.69497 | 0.69847 | 0.70194 | 0.70540 | 0.70884 | 0.71226 | 0.71566 | 0.71904 | 0.72240 |
| 0.6 | 0.72575 | 0.72907 | 0.73237 | 0.73565 | 0.73891 | 0.74215 | 0.74537 | 0.74857 | 0.75175 | 0.75490 |
| 0.7 | 0.75804 | 0.76115 | 0.76424 | 0.76730 | 0.77035 | 0.77337 | 0.77637 | 0.77935 | 0.78230 | 0.78524 |
| 0.8 | 0.78814 | 0.79103 | 0.79389 | 0.79673 | 0.79955 | 0.80234 | 0.80511 | 0.80785 | 0.81057 | 0.81327 |
| 0.9 | 0.81594 | 0.81859 | 0.82121 | 0.82381 | 0.82639 | 0.82894 | 0.83147 | 0.83398 | 0.83646 | 0.83891 |
| 1.0 | 0.84134 | 0.84375 | 0.84614 | 0.84849 | 0.85083 | 0.85314 | 0.85543 | 0.85769 | 0.85993 | 0.8621 |
| 1.1 | 0.86433 | 0.86650 | 0.86864 | 0.87076 | 0.87286 | 0.87493 | 0.87698 | 0.87900 | 0.88100 | 0.88298 |
| 1.2 | 0.88493 | 0.88686 | 0.88877 | 0.89065 | 0.89251 | 0.89435 | 0.89617 | 0.89796 | 0.89973 | 0.90147 |
| 1.3 | 0.90320 | 0.90490 | 0.90658 | 0.90824 | 0.90988 | 0.91149 | 0.91309 | 0.91466 | 0.91621 | 0.91774 |
| 1.4 | 0.91924 | 0.92073 | 0.92220 | 0.92364 | 0.92507 | 0.92647 | 0.92785 | 0.92922 | 0.93056 | 0.93189 |
| 1.5 | 0.93319 | 0.93448 | 0.93574 | 0.93699 | 0.93822 | 0.93943 | 0.94062 | 0.94179 | 0.94295 | 0.94408 |
| 1.6 | 0.94520 | 0.94630 | 0.94738 | 0.94845 | 0.94950 | 0.95053 | 0.95154 | 0.95254 | 0.95352 | 0.95449 |
| 1.7 | 0.95543 | 0.95637 | 0.95728 | 0.95818 | 0.95907 | 0.95994 | 0.96080 | 0.96164 | 0.96246 | 0.96327 |
| 1.8 | 0.96407 | 0.96485 | 0.96562 | 0.96638 | 0.96712 | 0.96784 | 0.96856 | 0.96926 | 0.96995 | 0.97062 |
| 1.9 | 0.97128 | 0.97193 | 0.97257 | 0.97320 | 0.97381 | 0.97441 | 0.97500 | 0.97558 | 0.97615 | 0.97670 |
| 2.0 | 0.97725 | 0.97778 | 0.97831 | 0.97882 | 0.97932 | 0.97982 | 0.98030 | 0.98077 | 0.98124 | 0.98169 |
| 2.1 | 0.98214 | 0.98257 | 0.98300 | 0.98341 | 0.98382 | 0.98422 | 0.98461 | 0.98500 | 0.98537 | 0.98574 |
| 2.2 | 0.98610 | 0.98645 | 0.98679 | 0.98713 | 0.98745 | 0.98778 | 0.98809 | 0.98840 | 0.98870 | 0.98899 |
| 2.3 | 0.98928 | 0.98956 | 0.98983 | 0.99010 | 0.99036 | 0.99061 | 0.99086 | 0.99111 | 0.99134 | 0.99158 |
| 2.4 | 0.99180 | 0.99202 | 0.99224 | 0.99245 | 0.99266 | 0.99286 | 0.99305 | 0.99324 | 0.99343 | 0.99361 |
| 2.5 | 0.99379 | 0.99396 | 0.99413 | 0.99430 | 0.99446 | 0.99461 | 0.99477 | 0.99492 | 0.99506 | 0.99520 |
| 2.6 | 0.99534 | 0.99547 | 0.99560 | 0.99573 | 0.99585 | 0.99598 | 0.99609 | 0.99621 | 0.99632 | 0.99643 |
| 2.7 | 0.99653 | 0.99664 | 0.99674 | 0.99683 | 0.99693 | 0.99702 | 0.99711 | 0.99720 | 0.99728 | 0.99736 |
| 2.8 | 0.99744 | 0.99752 | 0.99760 | 0.99767 | 0.99774 | 0.99781 | 0.99788 | 0.99795 | 0.99801 | 0.99807 |
| 2.9 | 0.99813 | 0.99819 | 0.99825 | 0.99831 | 0.99836 | 0.99841 | 0.99846 | 0.99851 | 0.99856 | 0.99861 |

Le tableau se lit de la façon suivante ($i$ donne l'unité et la première décimale de $s$, $j$ la seconde décimale):

| $s = i + j$ | ... | $j$ | ... |
|---|---|---|---|
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $i$ | ... | $\Pi(s)$ | ... |

Par exemple $\Pi(1.12) = \Pi(1.1 + 0.02) = 0.86864$

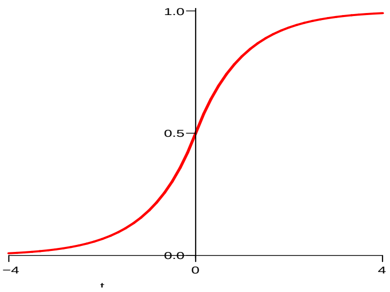


Figure 24.5: Fonction de répartition $\Pi(s)$ de la loi normale centrée réduite.

La fonction $\Pi$ étant une fonction de répartition, elle est continue, dérivable sur $\mathbb{R}$, croissante, tend vers 0 en $-\infty$ et vers 1 en $+\infty$. On a en outre les propriétés suivantes:

Propriété 93

- $\Pi(-s) = 1 - \Pi(s)\ \forall s \in \mathbb{R}$
- $\Pi(s) - \Pi(-s) = 2\Pi(s) - 1\ \forall s \in \mathbb{R}$
- *Si* $X \sim \mathcal{N}(m, \sigma)$ *, alors* $\mathbb{P}[a < X < b] = \Pi(\frac{b-m}{\sigma}) - \Pi(\frac{a-m}{\sigma})$

Démo

Les deux premières propriétés se voient facilement à partir de la fonction de densité $\phi(x)$ en se souvenant que $\pi(s)$ est une primitive de $\phi(x)$ et que sa valeur en $s$ représente donc l'aire comprise sous la courbe de $\phi(x)$ entre $-\infty$ et $s$. On sait par ailleurs que l'aire totale sous cette courbe est égale à 1. D'après la relation de Chasles,

$$\int_{-\infty}^{+\infty} \Phi(x)dx = \int_{-\infty}^{s} \Phi(x)dx + \int_{s}^{+\infty} \Phi(x)dx = \int_{-\infty}^{s} \Phi(x)dx + \int_{-\infty}^{-s} \Phi(x)dx \Rightarrow \Pi(-s) = 1 - \Pi(s)$$

$\Pi(s) - \Pi(-s) = \Pi(s) - 1 + \Pi(s) = 2\Pi(s) - 1$
Enfin, la dernière propriété est une conséquence directe de la définition de la fonction de répartition, après s'être ramené à une loi centrée réduite: En effet, si $Z \sim \mathcal{N}(0,1)$, $\mathbb{P}[a \leq Z \leq b] = \Pi(b) - \Pi(a)$. On remplace alors $Z$ par $(X - m)/\sigma$ dans la formule.
□

**Ex6:** Utilisation de la table de l'écart réduit

Si $Z$ suit une loi $\mathcal{N}(0,1)$, calculons $\mathbb{P}[Z \leq 1]$, $\mathbb{P}[Z \geq 2]$ et $\mathbb{P}[-1 < Z < 2]$

$\mathbb{P}[Z \leq 1] = \Pi(1) = 0.8413$
$\mathbb{P}[Z \geq 2] = 1 - \Pi(2) = 1 - 0.9772 = 0.023$
$\mathbb{P}[-1 < Z < 2] = \Pi(2) - \Pi(-1) = 0.8185$

Maintenant si $X$ suit une loi $\mathcal{N}(2,3)$, calculons $\mathbb{P}[X < 3]$
Pour pouvoir utiliser la table de l'écart réduit, il faut avoir à faire à une variable normale centrée réduite.

Posons alors $Z = \dfrac{X-2}{3}$

$\mathbb{P}[X < 3] = \mathbb{P}[X - 2 < 3 - 2] = \mathbb{P}[\frac{X-2}{3} < \frac{3-2}{3}] = \mathbb{P}[Z < \frac{1}{3}] = \Phi(\frac{1}{3}) = 0.6293$

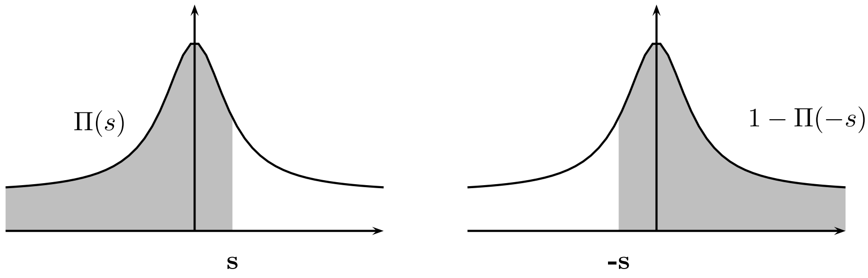


Figure 24.6: Propriétés de $\pi(s)$ lues à partir de la courbe $\phi(x)$

### 24.3.3 Opérations sur les variables gaussiennes

THÉORÈME 80

*Soient $X_1$ et $X_2$ deux va gaussiennes indépendantes $\mathcal{N}(m_1, \sigma_1)$ et $\mathcal{N}(m_2, \sigma_2)$*
*Soit $X$ la va définie par $X = X_1 + X_2$*
*Alors $X$ suit une loi normale $\mathcal{N}(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$*

Les moyennes s'ajoutent et les variances s'ajoutent. On peut généraliser ce résultat à $n$ variables aléatoires.

REMARQUE 21
*Sous les mêmes hypothèses, $X = X_1 - X_2$ suit une loi normale de paramètres $m = m_1 - m_2$ et $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$*

**Ex:**
Soit $X \sim \mathcal{N}(25, 3)$ et $Y \sim \mathcal{N}(35, 4)$ deux variables aléatoires indépendantes, alors $X + Y \sim \mathcal{N}(60, 5)$

Si $X \sim \mathcal{N}(1, 2)$ alors $-X \sim \mathcal{N}(-1, 2)$

THÉORÈME 81

*Soit $X$ une va de loi $\mathcal{N}(m, \sigma)$ alors $nX$ suit une loi $\mathcal{N}(nm, n\sigma)$*

REMARQUE 22
*Soient $X_1, X_2, ..., X_n$ $n$ va aléatoires indépendantes de même loi normale $\mathcal{N}(m, \sigma)$ et soit*
$\bar{X} = \dfrac{X_1 + X_2 + ... + X_n}{n}$
*$\bar{X}$ est la moyenne arithmétique des $X_k$*
*D'après ce qui précède, $\bar{X}$ suit une loi $\mathcal{N}(m, \dfrac{\sigma}{\sqrt{n}})$*

## 24.4 Théorèmes limites

### 24.4.1 Deux inégalités importantes

THÉORÈME 82 (INÉGALITÉ DE MARKOV)

*Soit $X$ une variable aléatoire dont l'espérance existe et $\epsilon > 0$:*
$\mathbb{P}[|X| > \epsilon] \leq \dfrac{\mathbb{E}[X]}{\epsilon}$

THÉORÈME 83 (INÉGALITÉ DE TCHEBYCHEV)

*Soit $X$ une variable aléatoire dont la variance existe et $\epsilon > 0$:*
$\mathbb{P}[|X - \mathbb{E}[X]| > \epsilon] \leq \dfrac{var(X)}{\epsilon^2}$

Une va dont l'espérance existe est dite d'ordre 1 et une va dont la variance existe est dite d'ordre 2

Ces inégalités (que nous admettons) donnent une précision sur la position des valeurs que peut prendre $X$ autour de sa moyenne. En particulier, la seconde montre que la probabilité qu'une variable aléatoire s'éloigne de sa moyenne diminue avec l'inverse du carré de la distance à cette moyenne

### 24.4.2 Loi faible des grands nombres

Rappelons que vaiid signifie variables aléatoires indépendantes identiquement distribuées (c.a.d. ayant toutes la même loi).

Théorème 84

*Soient $X_1, X_2, ..., X_n$ $n$ vaiid selon une loi de moyene $m$ et de variance $\sigma^2$*
*Soit $\bar{X} = \dfrac{X_1 + X_2 + ... + X_n}{n}$*
*Alors $\lim\limits_{n \to +\infty} \mathbb{P}[\bar{X} = m] = 1$*

Démo

Appliquons l'inégalité de Tchebychev à la variable aléatoire $\bar{X} = \dfrac{S_n}{n}$ avec $S_n = X_1 + ...X_n$

$\mathbb{P}[|\dfrac{S_n}{n} - m| \geq \epsilon] \leq \dfrac{1}{\epsilon^2} \mathrm{var}(\dfrac{S_n}{n})$

Or, $\mathrm{var}(\dfrac{S_n}{n}) = \dfrac{\sigma^2}{n} \Rightarrow \mathbb{P}[|\dfrac{S_n}{n} - m| \geq \epsilon] \leq \dfrac{K}{n} \Rightarrow \lim\limits_{n \to +\infty} \mathbb{P}[|\dfrac{S_n}{n} - m| \geq \epsilon] = 0$

Cette inégalité étant vraie pout tout $\epsilon$, $\lim\limits_{n \to +\infty} \mathbb{P}[\bar{X} = m] = 1$

□

**Applications**

Le résultat ci-dessus est théorique. En pratique, comment l'utilise-t-on ?
Lors d'observations statistiques où l'on fait $n$ expériences indépendantes, on étudie un évènement donné dont on relève la fréquence d'apparition. Lorsque $n$ tend vers l'infini, la fréquence tend vers l probabilité de l'évènement. Ceci constitue le résultat fondamental des statistiques.

### 24.4.3 Théorème de la limite centrale

Théorème 85

*Soit $X_1, ..., X_n$ $n$ vaiid selon une loi de moyenne $m$ et d'écart type $\sigma$*
*Soit $S_n = X_1 + X_2 + ... + X_n$*
*Alors $T_n = \dfrac{S_n - nm}{\sigma\sqrt{n}}$ converge en loi vers une variable aléatoire $\mathcal{N}(0, 1)$*
*C'est à dire: $\lim\limits_{n \to +\infty} \mathbb{P}[a < T_n < b] = \Pi(b) - \Pi(a)$*

Démo

admis

Remarquons néanmoins que $\mathbb{E}[S_n] = nm$ et $\sigma(S_n) = \sigma\sqrt{n}$, de sorte que $T_n = \dfrac{S_n - \mathbb{E}[S_n]}{\sigma(S_n)}$

Ainsi, $T_n$ n'est autre que la variable $S_n$ qui a été centrée et réduite.

□

On peut comprendre le théorème de la façon suivante: lorsque l'on prend un nombre suffisamment grand de variables aléatoire indépendantes, quelque soit leur loi, la somme (et également la moyenne) de ces variables aléatoires ressemble toujours à une loi normale.

En pratique, on peut appliquer cette loi si $X_k$ suit une loi de Poisson avec $\lambda$ grand ($\lambda > 15$) ou si $S_n$ suit une loi binomiale avec $n$ grand et $np \geq 18$.

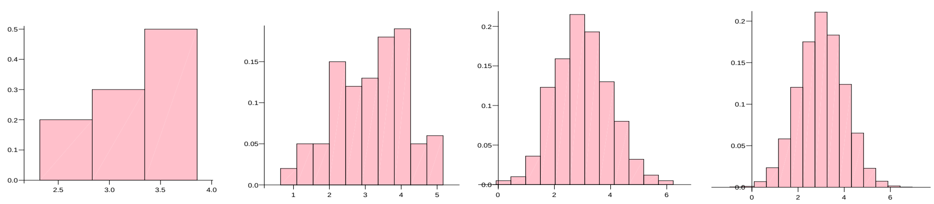

Figure 24.7: Tendance vers la loi normale $n = 10, 100, 1000$ et $10000$.

**Ex:**

Un caissier constate une erreur d'au plus 10F par jour. Quelle est la probabilité pour que l'erreur commise au bout de 220 jours soit inférieure à 200F?
Soit $X_k$ la variable aléatoire égale à l'erreur commise au jour n°k
$X_k$ suit une loi uniforme sur $[-10, 10]$
Comme cette loi admet une espérance nulle et une variance $\sigma^2 = \frac{400}{12} \approx 33.3$, on peut appliquer le théorème et poser $S_n = X_1 + X_2 + ... + X_{220}$ et l'on cherche $\mathbb{P}[-200 < X < 200]$

Si l'on suppose les $X_k$ indépendants alors $S_n$ suit à peu près une loi $\mathcal{N}(0, \dfrac{33.3}{\sqrt{220}})$

$\mathbb{P}[-200 < X < 200] = \mathbb{P}[-2.45 < Z < 2.45] = 2\Pi(2.45) - 1 = 0.99$

Voici un cas particulier important de ce théorème qui permet d'approcher une loi binomiale par une loi normale.

Théorème 86 (De Moivre Laplace)

*Soit $X$ une variable de loi binomiale de paramètre $n$ et $p$*
*Soit* $Y = \dfrac{X - np}{\sqrt{npq}}$
*Lorsque $n$ tend vers l'infini, $Y$ tend vers une loi normale $\mathcal{N}(0, 1)$:*
$\lim\limits_{n \to +\infty} \mathbb{P}[a < Y < b] = \Pi(b) - \Pi(a)$

On estime que l'approximation par la loi normale est valable lorsque $\boxed{npq \geq 18}$

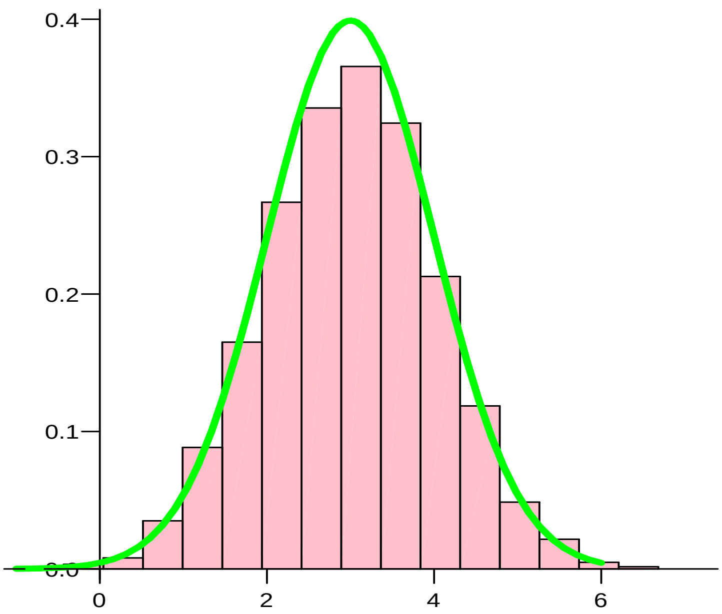


Figure 24.8: Convergence de la loi binomiale vers la loi normale.

Lorsque l'on veut calculer $\mathbb{P}[X = k]$ avec ce théorème (par exemple pour approcher une loi discrète par la loi normale), il faut poser $\mathbb{P}[X = k] = \mathbb{P}[k - 1/2 < X < k + 1/2]$ puisque $X$ ne prend que des valeurs entières. En effet, le calcul de $\mathbb{P}[X = k]$ pour une loi à densité donnerait 0 pour tout entier $k$. Nous verrons en TD des applications de ces théorèmes.