

Introduction à la théorie des probabilités et aux statistiques

Magistère de mécatronique

1^{re} année

ENS Rennes – Université Rennes 1

Année 2018-2019

Pour toutes remarques et corrections : florian.lemonnier@ens-rennes.fr

Table des matières

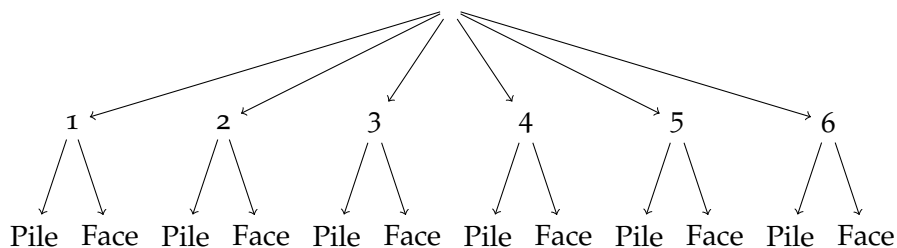
1	Rappels de dénombrement	3
1.1	Permutation d'un ensemble	3
1.2	Arrangement	4
1.3	Combinaison	4
2	Espaces probabilisés	7
2.1	Quelques définitions	7
2.2	Probabilité conditionnelle	10
3	Variables aléatoires réelles	12
3.1	Variables aléatoires discrètes	12
3.2	Variables aléatoires continues	13
3.3	Grandeurs associées aux variables aléatoires	14
3.3.1	Espérance	14
3.3.2	Variance	16
3.3.3	Fonction de répartition	17
3.4	Lois normales	19
3.4.1	Loi normale centrée réduite	19
3.4.2	Cas général	20
3.5	Couples de variables aléatoires	21
3.6	Variables aléatoires indépendantes	22
4	Théorèmes limites	25
4.1	Loi des grands nombres	25
4.2	Théorème central limite	25
5	Applications en statistiques	26
5.1	Modélisation statistique	26
5.1.1	Principe fondamental de la statistique	26
5.1.2	Modèle statistique	26
5.2	Estimation paramétrique	27
5.2.1	Estimateur	27
5.2.2	Estimateur du maximum de vraisemblance	27
5.2.3	Critères de performance des estimateurs	28
5.3	Intervalle de confiance	29
5.3.1	Intervalle de confiance par excès	29
5.3.2	Intervalle de confiance asymptotique	30
A	Lois de probabilités classiques	I
A.1	Lois discrètes	I
A.2	Lois à densité	II
B	Table de la loi normale centrée réduite	III

Chapitre 1

Rappels de dénombrement

En France, lorsqu'on joue au loto, l'objectif est de trouver les 5 bons numéros parmi 49 ainsi qu'un numéro chance (parmi 10 possibilités). Pour calculer les chances de gagner, il est naturel de compter le nombre N total de choix possibles, et on pourra ainsi dire qu'il y a une chance sur N de gagner le premier lot. Pour compter ce nombre de possibilités, on va faire intervenir la théorie du dénombrement.

La base du dénombrement est le fait que lorsqu'on effectue deux expériences successives ayant respectivement m et n possibilités, alors il existe $m \times n$ possibilités pour les deux expériences prises ensemble. Pour s'en convaincre, imaginons un jeu consistant à devoir deviner le résultat d'un dé équilibré à 6 faces et le résultat du lancer d'une pièce.



On a ainsi $6 \times 2 = 12$ possibilités.

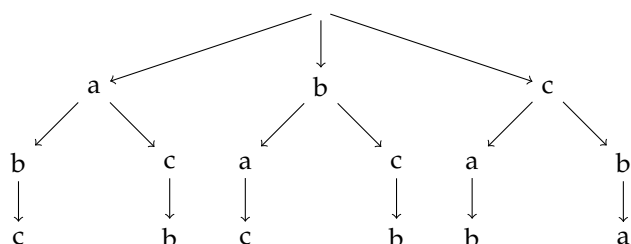
Bien sûr, sur le même principe, si on effectue m expériences, avec chacune un nombre n_i de possibilités, alors on a $n_1 \times n_2 \times \dots \times n_m$ possibilités pour les m expériences prises ensemble.

1.1 Permutation d'un ensemble

Définition 1.1.1 (Permutation)

On appelle permutation d'un ensemble composé de n éléments tout rangement ordonné de ces éléments.

Exemple 1.1.2. (a, b, c) , (b, c, a) et (c, a, b) sont trois permutations d'un ensemble à 3 éléments $\{a, b, c\}$. Combien a-t-on de façons de ranger un ensemble à 3 éléments $\{a, b, c\}$? On a trois possibilités pour le premier élément, deux pour le second, et enfin un seul choix pour le troisième, donc $3 \times 2 \times 1 = 6$ possibilités.



Proposition 1.1.3

Le nombre de permutations d'un ensemble à n éléments est $n!$.

Remarque 1.1.4. $n!$ se lit "factorielle n " et vaut $n! = n \times (n - 1) \times \dots \times 3 \times 2 \times 1$. Par convention, on pose $0! = 1$.

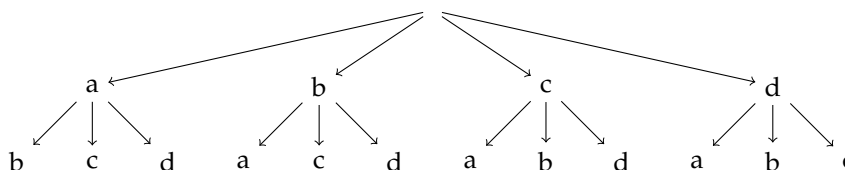
1.2 Arrangement

Définition 1.2.1 (Arrangement)

On appelle arrangement de k éléments pris parmi n éléments d'un ensemble E toute suite ordonnée de k éléments distincts de E .

Exemple 1.2.2. (a, b, c, f) , (f, a, e, d) et (e, a, f, d) sont trois arrangements distincts à 4 éléments d'un ensemble à 6 éléments $\{a, b, c, d, e, f\}$.

Combien y a-t-il d'arrangements de 2 éléments parmi 4 ? On a quatre possibilités pour le premier, puis trois pour le second, soit $4 \times 3 = 12$ possibilités.



Proposition 1.2.3

Le nombre d'arrangements de k éléments parmi n est :

$$n \times (n - 1) \times (n - 2) \times \dots \times (n - k + 1) = \frac{n!}{(n - k)!}$$

Exemple 1.2.4. Dans une course de chevaux, on compte 18 partants. Le nombre de résultats possibles au tiercé est donc $\frac{18!}{(18 - 3)!} = 4896$.

1.3 Combinaison

Revenons à l'exemple du loto. La remarque du début de ce chapitre nous invite à considérer séparément l'expérience relative aux 5 bons numéros (ordinaires) parmi 49 et l'expérience relative au numéro chance. L'ordre des éléments du tirage des numéros ordinaires n'importe pas, car $\{1, 12, 33, 42, 48\}$ et $\{33, 48, 1, 42, 12\}$ correspondent à une même grille. La méthode précédente de dénombrement est donc inadéquate.

Définition 1.3.1 (Combinaison)

On appelle combinaison de k éléments parmi n tout sous-ensemble de E (ensemble de cardinal n) comportant exactement k éléments.

Le nombre d'arrangements de k parmi n est $\frac{n!}{(n - k)!}$. On ne désire plus distinguer deux arrangements qui sont composés des mêmes éléments, c'est-à-dire les arrangements qui sont égaux à une permutation près. Le nombre d'arrangements qui sont égaux à une permutation près est $k!$.

Proposition 1.3.2

Le nombre de combinaisons de k éléments parmi n est le nombre noté :

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}$$

Ainsi, le nombre de grilles différentes de 5 numéros parmi les 49 numéros possibles est

$$\binom{49}{5} = \frac{49 \times 48 \times 47 \times 46 \times 45}{5 \times 4 \times 3 \times 2 \times 1} = 1\,906\,884.$$

En outre, il y a une chance sur 10 de choisir le bon numéro chance. On a donc une chance sur 19 068 840 de gagner le premier lot.

Remarque 1.3.3. Par convention, si $k > n$, on pose $\binom{n}{k} = 0$.

Proposition 1.3.4

Les coefficients binomiaux $\binom{n}{k}$ vérifient les relations suivantes :

1. $\binom{n}{k} = \binom{n}{n-k}$;
2. $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$.

Démonstration:

1. Par le calcul, il suffit d'écrire la formule dans les deux cas. En termes de dénombrement, on peut expliquer que "choisir k éléments parmi n " revient à "(ne pas) choisir $n - k$ éléments parmi n ".
2. Par le calcul, on a :

$$\begin{aligned} \binom{n}{k} + \binom{n}{k+1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-k-1)!} = \frac{n!(k+1) + n!(n-k)}{(k+1)!(n-k)!} \\ &= \frac{(n+1)!}{(k+1)!(n-k)!} = \binom{n+1}{k+1}. \end{aligned}$$

D'un point de vue dénombrement, considérons un ensemble à $n + 1$ éléments dans lequel on piochera $k + 1$ éléments simultanément :

$$\boxed{e_1, e_2, e_3, \dots, e_n \mid e_{n+1}}$$

On isole intentionnellement le dernier élément e_{n+1} . On va alors compter séparément les combinaisons qui comportent l'élément e_{n+1} et celles qui ne le comportent pas.

— Choisir une combinaison qui ne contient pas e_{n+1} revient exactement à choisir les $k + 1$ éléments parmi e_1, e_2, \dots, e_n .

$$\boxed{e_1, e_2, e_3, \dots, e_n \mid \times}$$

Le nombre de telles combinaisons est $\binom{n}{k+1}$ (le nombre de combinaisons de $k + 1$ éléments parmi n).

— Maintenant on cherche le nombre de combinaisons qui comportent l'élément e_{n+1} . Il ne reste plus que k éléments à choisir vu qu'il y en a déjà un de fixé.

$$\boxed{e_1, e_2, e_3, \dots, e_n \mid \surd}$$

Ces k éléments ne peuvent être choisis que parmi les e_1, e_2, \dots, e_n . Il y a donc autant de combinaisons de $k + 1$ éléments parmi $n + 1$ qui contiennent e_{n+1} que de combinaisons de k éléments parmi n , ce qui fait $\binom{n}{k}$. ■

Proposition 1.3.5 (Binôme de Newton)

Soient a et b deux réels quelconques et soit n un entier naturel. Alors on a la relation :

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

Démonstration: par récurrence.

L'initialisation est immédiate : $(a + b)^0 = 1 = \binom{0}{0} a^0 b^0$.

On suppose maintenant que l'égalité soit vraie pour un certain rang n (c'est l'hypothèse de récurrence) et on veut vérifier qu'elle est vraie au rang $n + 1$.

On écrit :

$$\begin{aligned} (a + b)^{n+1} &= (a + b)(a + b)^n = (a + b) \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k, \text{ par hypothèse de récurrence} \\ &= \sum_{k=0}^n \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^n \binom{n}{k} a^{n-k} b^{k+1} \end{aligned}$$

On effectue alors un changement d'indice dans la deuxième somme, en posant $j = k + 1$:

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

Quant à la première somme, on la casse de la façon suivante :

$$\sum_{k=0}^n \binom{n}{k} a^{n+1-k} b^k = a^{n+1} + \sum_{k=1}^n \binom{n}{k} a^{n+1-k} b^k = a^{n+1} + \sum_{j=1}^n \binom{n}{j} a^{n+1-j} b^j.$$

Ainsi, en additionnant ces deux sommes, on obtient :

$$\begin{aligned} (a + b)^{n+1} &= a^{n+1} + \sum_{j=1}^n \binom{n}{j} a^{n+1-j} b^j + b^{n+1} + \sum_{j=1}^n \binom{n}{j-1} a^{n+1-j} b^j \\ &= \binom{n+1}{0} a^{n+1} b^0 + \binom{n+1}{n+1} a^0 b^{n+1} + \sum_{j=1}^n \left(\binom{n}{j} + \binom{n}{j-1} \right) a^{n+1-j} b^j \\ &= \binom{n+1}{0} a^{n+1} b^0 + \binom{n+1}{n+1} a^0 b^{n+1} + \sum_{j=1}^n \binom{n+1}{j} a^{n+1-j} b^j \\ &= \sum_{j=0}^{n+1} \binom{n+1}{j} a^{(n+1)-j} b^j. \end{aligned}$$

Si on réfléchit d'un point de vue dénombrement, on peut remarquer que chaque terme issu du développement de $(a + b)^n$ est un produit de n facteurs issus de $(a + b)$. Ainsi, tous les termes ont la forme $a^k b^{n-k}$. Ce qu'il faut savoir désormais, c'est le nombre de termes de la forme $a^k b^{n-k}$ (pour k allant de 0 à n). Cela revient à choisir k éléments a parmi les n parenthèses $(a + b)$, et le nombre de b sera alors $n - k$.

Le nombre de termes $a^k b^{n-k}$ est le nombre de combinaisons de k éléments parmi n , donc $\binom{n}{k}$. Ainsi, une fois les termes égaux regroupés :

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

Remarque 1.3.6. Cette formule permet notamment de montrer que $\sum_{k=0}^n \binom{n}{k} = 2^n$.

Chapitre 2

Espaces probabilisés

La théorie des probabilités se propose de calculer, de quantifier le hasard.

Par exemple, le jeu du *Pile ou Face*. Les équations déterminant le mouvement de la pièce ne sont pas dues au hasard. Le problème, c'est que l'on ne connaît pas les conditions initiales. Les propriétés de la pièce incitent à penser que la "moitié" des conditions initiales aboutissent au résultat *Pile*, l'autre "moitié" au résultat *Face*. **Le hasard apparaît donc lorsqu'il est impossible d'accéder à une information utilisable sur les conditions initiales. Cela se produit fréquemment : météorologie, finances ...**

Les techniques probabilistes sont beaucoup utilisées pour obtenir des informations sur des phénomènes d'ordres les plus divers : sondages, estimations diverses ...

L'efficacité de ces techniques est sans doute leur meilleure justification. Le formalisme couramment adopté en théorie des probabilités est dû à A.N. Kolmogorov (1903-1987). Ses ingrédients principaux sont les théories des ensembles et de la mesure.

2.1 Quelques définitions

En probabilités, pour modéliser une expérience aléatoire, on se donne un ensemble permettant de décrire toutes les issues possibles de l'expérience. On appelle cet ensemble *univers*, et on le note généralement Ω .

Exemple 2.1.1.

1. On jette un dé à six faces et on s'intéresse au numéro de la face supérieure. On choisit $\Omega = \{1, 2, 3, 4, 5, 6\}$.
2. On joue au *Pile ou Face*. On choisit $\Omega = \{0, 1\}$.
3. On jette une pièce sur une table carrée; on s'intéresse aux coordonnées du centre de masse de la pièce. On choisit $\Omega = [0, 1]^2$.

On appelle *événement* toute partie de Ω . Dans la suite, on va vouloir calculer la probabilité de certains événements.

Malheureusement, sauf lorsque Ω sera fini ou dénombrable (*id est* en bijection avec \mathbb{N} , comme par exemple \mathbb{Z} ou \mathbb{Q} , mais pas $[0, 1]$ ou \mathbb{R}), on ne pourra pas s'intéresser à l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω . On se restreindra donc à un sous-ensemble \mathcal{F} de $\mathcal{P}(\Omega)$, qui constituera l'ensemble des parties de Ω dont on peut calculer la probabilité.

Afin d'obtenir un modèle aussi cohérent que possible, on doit imposer à \mathcal{F} de vérifier certaines conditions de stabilité : par union, par intersection, par passage au complémentaire ...

Définition 2.1.2 (Tribu)

On appelle tribu sur l'ensemble Ω , tout sous-ensemble $\mathcal{F} \subset \mathcal{P}(\Omega)$ vérifiant les assertions suivantes :

1. l'univers est un élément de \mathcal{F} ,
2. l'ensemble \mathcal{F} est stable par passage au complémentaire,
3. l'ensemble \mathcal{F} est stable par réunion dénombrable,

Exemple 2.1.3. Quelques exemples classiques de tribus sont :

- la tribu pleine $\mathcal{F}_1 = \mathcal{P}(\Omega)$;
- la tribu triviale $\mathcal{F}_2 = \{\emptyset, \Omega\}$;
- $\mathcal{F}_3 = \{\emptyset, A, \bar{A}, \Omega\}$, avec $A \subset \Omega$ fixé;
- la tribu borélienne $\mathcal{F}_4 = \mathcal{B}(\Omega)$ est la plus petite tribu contenant tous les ouverts de Ω . Pour rappel, un ouvert est un ensemble qui ne contient aucun point de son bord, comme par exemple $]0, 1[$, dont le bord est $\{0, 1\}$. Dans le cas particulier $\Omega = \mathbb{R}^d$, la tribu borélienne est aussi la plus petite tribu contenant tous les ensembles de la forme $[a_1, b_1[\times \dots \times [a_d, b_d[$.

En pratique, si Ω est fini ou dénombrable, on choisira en général $\mathcal{F} = \mathcal{P}(\Omega)$. Si Ω n'est pas dénombrable, comme c'est le cas dans l'exemple d'une suite infinie de lancers ($\Omega = \{0, 1\}^{\mathbb{N}}$), on considérera en général la tribu $\mathcal{F} = \mathcal{B}(\Omega)$, qui est plus petite que $\mathcal{P}(\Omega)$.

Définition 2.1.4 (Probabilité)

Une probabilité est une fonction (généralement notée \mathbb{P}), définie sur une tribu \mathcal{F} telle que :

- 1.
- 2.
- 3.

Proposition 2.1.5 (Propriétés d'une probabilité)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Tous les ensembles sont supposés appartenir à \mathcal{F} .

1. Monotonie : si $A \subset B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$ et plus précisément :

$$\mathbb{P}(B) =$$

2. Additivité forte :

$$\mathbb{P}(A) + \mathbb{P}(B) =$$

3. Sous- σ -additivité :

$$\mathbb{P} \left(\bigcup_{n=0}^{\infty} A_n \right) \leq$$

4. Continuité monotone croissante : si $(A_n)_{n \in \mathbb{N}}$ est une suite d'événements croissante pour l'inclusion (id est $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$), alors

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) =$$

5. Continuité monotone décroissante : si $(A_n)_{n \in \mathbb{N}}$ est une suite d'événements décroissante pour l'inclusion (id est $A_{n+1} \subset A_n$ pour tout $n \in \mathbb{N}$), alors

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) =$$

Démonstration :

1. On applique la définition avec $A_0 = A$, $A_1 = B \setminus A$ et $A_n = \emptyset$ pour $n \geq 2$.
2. On décompose de façon disjointe :

$$A \cup B = (A \setminus (A \cap B)) \sqcup (A \cap B) \sqcup (B \setminus (A \cap B)).$$

On en déduit, toujours par le point 3 de la définition :

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

Par la propriété précédente, on aboutit par le calcul à :

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

3. On construit une suite d'événements $(B_n)_{n \in \mathbb{N}}$ par $B_0 = A_0$ et pour $n \geq 1$, $B_n = A_n \setminus \left(\bigcup_{k=0}^{n-1} A_k \right)$. Les

B_n sont deux à deux disjoints, $B_n \subset A_n$ et $\bigcup_{n=0}^{\infty} B_n = \bigcup_{n=0}^{\infty} A_n$. En conséquence :

$$\mathbb{P} \left(\bigcup_{n=0}^{\infty} A_n \right) = \mathbb{P} \left(\bigcup_{n=0}^{\infty} B_n \right) = \sum_{n=0}^{\infty} \mathbb{P}(B_n) \leq \sum_{n=0}^{\infty} \mathbb{P}(A_n).$$

4. On reprend la même suite $(B_n)_{n \in \mathbb{N}}$. On remarque que pour tout n , on a :

$$A_n = B_0 \cup B_1 \cup \dots \cup B_n.$$

Ainsi :

$$\mathbb{P} \left(\bigcup_{n=0}^{\infty} A_n \right) = \mathbb{P} \left(\bigcup_{n=0}^{\infty} B_n \right) = \sum_{n=0}^{\infty} \mathbb{P}(B_n) = \lim_{N \rightarrow \infty} \sum_{n=0}^N \mathbb{P}(B_n) = \lim_{N \rightarrow \infty} \mathbb{P}(A_N).$$

5. On considère la suite d'ensembles définie par $C_n = A_0 \setminus A_n$, qui vérifie $\mathbb{P}(C_n) = \mathbb{P}(A_0) - \mathbb{P}(A_n)$.

La suite $(C_n)_{n \in \mathbb{N}}$ est croissante et $\bigcup_{n=0}^{\infty} C_n = A_0 \setminus \left(\bigcap_{n=0}^{\infty} A_n \right)$. Par continuité monotone croissante,

$$\lim_{n \rightarrow \infty} \mathbb{P}(C_n) = \mathbb{P} \left(A_0 \setminus \left(\bigcap_{n=0}^{\infty} A_n \right) \right) = \mathbb{P}(A_0) - \mathbb{P} \left(\bigcap_{n=0}^{\infty} A_n \right).$$

Mais, d'autre part,

$$\lim_{n \rightarrow \infty} \mathbb{P}(C_n) = \mathbb{P}(A_0) - \lim_{n \rightarrow \infty} \mathbb{P}(A_n).$$

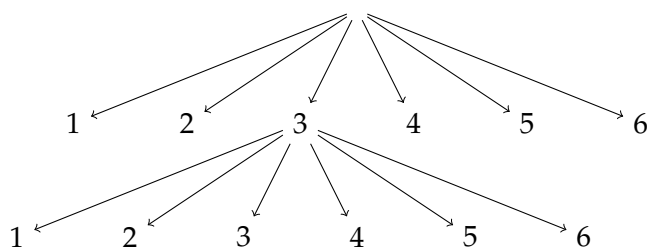
Ce qui permet de conclure. ■

Exemple 2.1.6.

1. On lance 3 fois de suite une pièce équilibrée et on compte le nombre de fois où *Pile* est apparu. On a donc $\Omega = \{0, 1, 2, 3\}$, mais il n’y a pas équiprobabilité puisque les probabilités élémentaires sont $\mathbb{P}(\{0\}) = \frac{1}{8}$, $\mathbb{P}(\{1\}) = \frac{3}{8}$, $\mathbb{P}(\{2\}) = \frac{3}{8}$ et $\mathbb{P}(\{3\}) = \frac{1}{8}$.
2. Si on veut construire une probabilité \mathbb{P} sur un ensemble infini dénombrable, typiquement sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, on ne peut plus avoir équiprobabilité des événements élémentaires $\{n\}$. Supposons en effet que pour tout $n \in \mathbb{N}$ on ait $\mathbb{P}(\{n\}) = p > 0$, alors on aurait $\mathbb{P}(\mathbb{N}) = \mathbb{P}(\bigcup_{n=0}^{\infty} \{n\}) = \sum_{n=0}^{\infty} \mathbb{P}(\{n\}) = \sum_{n=0}^{\infty} p = +\infty$, ce qui est en contradiction avec la condition $\mathbb{P}(\mathbb{N}) = 1$. Une façon de construire une probabilité sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ est de considérer une suite $(p_n)_{n \in \mathbb{N}}$ de nombres positifs telle que la série $\sum_{n=0}^{\infty} p_n$ soit convergente et de somme 1. On définit alors pour tout événement $A \in \mathcal{P}(\mathbb{N})$ sa probabilité par : $\mathbb{P}(A) = \sum_{i \in A} \mathbb{P}(\{i\})$.
3. On lance une pièce équilibrée jusqu’à ce que *Pile* apparaisse (en excluant le cas improbable où *Pile* n’apparaît jamais). On a donc $\Omega = \{1, 2, \dots\} = \mathbb{N}^*$. On a clairement $p_1 = \mathbb{P}(\{1\}) = \frac{1}{2}$, $p_2 = \frac{1}{4}$ et de façon générale $p_n = \frac{1}{2^n}$. On reconnaît dans les p_n les termes d’une suite géométrique dont la somme vaut bien 1.

2.2 Probabilité conditionnelle

Supposons que nous jetions deux dés à 6 faces. Les résultats d’une telle expérience sont des couples de la forme (a, b) , où a est le résultat du dé numéro 1 et b celui du numéro 2. On se place dans le cas où tous ces couples sont équiprobables de probabilité $\frac{1}{36}$. Une fois que l’on connaît le résultat du premier dé, mettons 3, quelle est la probabilité que la somme des deux résultats soit par exemple 8 ?



Étant donné que le premier dé est un 3, la probabilité (conditionnelle) de chacun des événements $(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)$ est $\frac{1}{6}$ vu que le deuxième dé est équilibré. Mais la probabilité (conditionnelle) d’obtenir un des 30 autres couples est nul, et un seul couple de la liste donnée a une somme de 8.

Si A est l’événement “le premier dé donne 3”, et B le résultat “la somme des deux résultats est 8”, la probabilité de B sachant A est ici $\frac{1}{6}$.

Définition 2.2.1 (Probabilité conditionnelle)

Soient A et B deux événements, avec $\mathbb{P}(B) > 0$.
 La probabilité conditionnelle de A sachant B , est :

Remarque 2.2.2. En utilisant successivement la définition de la probabilité conditionnelle, on obtient la formule des probabilités composées :

Définition 2.2.3 (Indépendance)

Un événement A est dit *indépendant* d’un deuxième événement B , si $\mathbb{P}(A|B) = \mathbb{P}(A)$.

Proposition 2.2.4

Les propriétés suivantes sont équivalentes :

$$\begin{array}{ll} \mathbb{P}(A|B) = \mathbb{P}(A) & \text{“}A \text{ est indépendant de } B\text{”} \\ \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) & \text{“}A \text{ et } B \text{ sont indépendants”} \\ \mathbb{P}(B|A) = \mathbb{P}(B) & \text{“}B \text{ est indépendant de } A\text{”} \\ \mathbb{P}(B|\bar{A}) = \mathbb{P}(B|A) & \text{“la réalisation ou non de } A \text{ ne modifie pas la probabilité de } B\text{”} \end{array}$$

Démonstration: Exercice. ■

Définition 2.2.5 (Indépendance mutuelle)

Une famille quelconque d'événements $(A_i)_{i \in I}$ est mutuellement indépendante si

Remarque 2.2.6. Si $(A_i)_{i \in I}$ est une famille d'événements mutuellement indépendants, alors ces événements sont indépendants deux à deux (au sens où pour tous i_1 et i_2 dans I , A_{i_1} et A_{i_2} sont indépendants). L'exemple suivant montre que **l'indépendance deux à deux n'implique pas l'indépendance mutuelle**.

Exemple 2.2.7. On s'intéresse à un jeu de dés (l'un est rouge, l'autre bleu). On considère les événements :

- A : “le résultat du dé rouge est impair” ;
- B : “le résultat du dé bleu est impair” ;
- C : “la somme des deux dés est impaire”.

Il est facile de constater que les événements A , B et C sont indépendants deux à deux. Mais ils ne sont pas mutuellement indépendants.

Proposition 2.2.8 (Probabilités totales)

Soit $(A_n)_{n \in \mathbb{N}}$ une famille d'événements deux à deux incompatibles (c'est à dire que $\mathbb{P}(A_i \cap A_j) = 0$, pour $i \neq j$) et tels que $\sum_{n=0}^{\infty} \mathbb{P}(A_n) = 1$.

Alors :

$$\mathbb{P}(B) =$$

Théorème 2.2.9 (Formule de Bayes)

Sous les mêmes conditions que la proposition précédente :

$$\mathbb{P}(A_i|B) =$$

Chapitre 3

Variables aléatoires réelles

Désormais, on se place dans un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ fixé.

3.1 Variables aléatoires discrètes

Définition 3.1.1 (Variable aléatoire discrète)

Soit E un ensemble au plus dénombrable.

On dit qu'une application $X : \Omega \rightarrow E$ est une *variable aléatoire discrète* si

$$\forall x_i \in E, X^{-1}(x_i) := \{\omega \in \Omega \mid X(\omega) = x_i\} \in \mathcal{F}.$$

Remarque 3.1.2. En pratique, si Ω est dénombrable, on prend $\mathcal{F} = \mathcal{P}(\Omega)$; alors, la seule condition pour qu'une application soit une variable aléatoire discrète est qu'elle soit à valeurs dans un ensemble dénombrable.

Exemple 3.1.3.

1. On lance un dé, on note X le numéro obtenu : c'est une variable aléatoire sur l'univers
 - (a)
 - (b)
2. On lance un dé trois fois, on observe le numéro obtenu au deuxième lancer,
3. On lance un dé trois fois, on fait la somme des résultats,

Définition 3.1.4 (Loi d'une variable aléatoire discrète)

Soit X une variable aléatoire discrète, à valeurs dans $E = \{x_i \mid i \in I\}$.

La loi de X est la famille des $(p_i)_{i \in I}$, définie par

$$\forall i \in I, p_i = \mathbb{P}(X = x_i).$$

Remarque 3.1.5. Par abus de notation, on écrit souvent des quantités comme $\mathbb{P}(X = x_i)$. Il s'agit d'un raccourci de notation pour désigner $\mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x_i\}) = \mathbb{P}(X^{-1}(x_i))$.

Exemple 3.1.6.

1. Lancer d'une pièce : la loi de X est donnée par $p_0 = p_1 = \frac{1}{2}$. On dit que X suit la loi uniforme sur l'ensemble $\{0, 1\}$.

2. Somme issue du lancer de deux dés : la loi de X est donnée par la famille $p = (p_2, \dots, p_{12})$ où

$$p = \left(\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36} \right).$$

3.2 Variables aléatoires continues

Définition 3.2.1 (Variable aléatoire réelle)

Une *variable aléatoire réelle* est une application $X : \Omega \rightarrow \mathbb{R}$ telle que pour tout intervalle I de \mathbb{R} , on ait $X^{-1}(I) := \{\omega \in \Omega \mid X(\omega) \in I\} \in \mathcal{F}$.

Remarque 3.2.2. On retrouve la définition d'une variable aléatoire discrète si X est une variable aléatoire réelle qui prend ses valeurs dans un sous-ensemble dénombrable de \mathbb{R} .

Expliquons en deux mots la définition générale ci-dessus. L'intérêt de supposer qu'un événement appartient à \mathcal{F} est d'être assuré qu'on pourra en calculer la probabilité. Mais dans le cas où X prend ses valeurs dans \mathbb{R} tout entier (ou dans un intervalle de \mathbb{R}), on ne veut plus se restreindre à des événements de la forme "X prend la valeur x " comme dans le cas discret, car ceux-ci seront généralement de probabilité nulle, donc sans grand intérêt. Le fait de supposer qu'on va pouvoir calculer la probabilité de n'importe quel intervalle est bien plus pertinent, car à eux seuls les intervalles engendrent une famille très riche de sous-ensembles de \mathbb{R} : la tribu borélienne.

Définition 3.2.3 (Densité)

Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}_+$ admettant un nombre fini de discontinuités, et vérifiant $\int_{-\infty}^{+\infty} f(x) dx = 1$, est une densité de probabilité.

C'est une densité d'une variable aléatoire X si pour tout intervalle I de \mathbb{R} , on a :

$$\mathbb{P}(X \in I) = \int_I f(x) dx.$$

Remarque 3.2.4. Là encore, noter l'abus de notation $\mathbb{P}(X \in I)$ qui signifie $\mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in I\}) = \mathbb{P}(X^{-1}(I))$.

Remarque 3.2.5. Notons que si X admet une densité (notée f_X), alors $\mathbb{P}(X = x_0) = \int_{x_0}^{x_0} f_X(x) dx = 0$.

En particulier, pour une variable aléatoire X à densité, il est important de remarquer que :

$$\mathbb{P}(X \in]a, b[) = \mathbb{P}(X \in [a, b]) = \mathbb{P}(X \in]a, b]) = \mathbb{P}(X \in [a, b]).$$

Exemple 3.2.6.

1. Loi uniforme : une variable X se répartissant de façon uniforme dans le segment $[0, 1]$ est définie par la densité $f_X = \mathbb{1}_{[0,1]}$, c'est-à-dire $f_X(x) = 1$ si $x \in [0, 1]$ et $f_X(x) = 0$ sinon. Pour tous points $a \leq b$ de $[0, 1]$, la probabilité de tomber entre a et b est

2. Loi exponentielle : prenons maintenant $f(x) = e^{-x}$ si $x \geq 0$ et $f(x) = 0$ si $x < 0$. Alors f est discontinue en 0 seulement, à valeurs positives et

$$\int_{-\infty}^{+\infty} f(x) dx =$$

donc f définit bien une densité. Si X est une variable aléatoire de densité f , on dit que X suit la loi exponentielle de paramètre 1 et on note $X \sim \mathcal{E}(1)$.

3.3 Grandeurs associées aux variables aléatoires

3.3.1 Espérance

Définition 3.3.1 (Espérance)

1. On définit l'espérance d'une variable aléatoire discrète X par la formule

$$\mathbb{E}[X] =$$

dans le cas où

2. On définit l'espérance d'une variable aléatoire X à densité f_X par

$$\mathbb{E}[X] =$$

dans le cas où

Remarque 3.3.2. Si une variable aléatoire X ne prend que des valeurs positives, on peut néanmoins accepter de définir l'espérance de X lorsque celle-ci est infinie.

Remarque 3.3.3. L'espérance d'une variable aléatoire, c'est la moyenne des valeurs qu'elle prend, pondérée par ses probabilités.

Exemple 3.3.4.

1. Loi uniforme : si $X \sim \mathcal{U}_{[0,1]}$, alors son espérance vaut
2. Loi exponentielle : si $X \sim \mathcal{E}(1)$, alors une intégration par parties donne
3. Loi de Cauchy. On considère une variable aléatoire X de densité $f(x) = \frac{1}{\pi(1+x^2)}$; en effet, f est une fonction continue positive et $\int_{-\infty}^{+\infty} f(x) dx = 1$. On dit que X suit la loi de Cauchy de paramètre 1. Néanmoins, l'intégrale $\int_{-\infty}^{+\infty} xf(x) dx$ diverge car

Proposition 3.3.5

Soient X et Y deux variables aléatoires (discrètes ou à densité toutes les deux), et soit $\lambda \in \mathbb{R}$. Alors,

1. $\mathbb{E}[\lambda X + Y] = \lambda \mathbb{E}[X] + \mathbb{E}[Y]$;
2. si pour tout $\omega \in \Omega$, $X(\omega) \geq 0$, alors $\mathbb{E}[X] \geq 0$;
3. si pour tout $\omega \in \Omega$, $X(\omega) \geq Y(\omega)$, alors $\mathbb{E}[X] \geq \mathbb{E}[Y]$;
4. pour tout A intervalle de \mathbb{R} , on a $\mathbb{E}[\mathbb{1}_A(X)] = \int_A f_X(x) dx$.

Théorème 3.3.6 (Transfert)

1. Soit X une variable aléatoire discrète et $\varphi : X(\Omega) \rightarrow X(\Omega)$ une fonction, alors l'espérance de $\varphi(X)$ vaut :

$$\mathbb{E}[\varphi(X)] =$$

dans le cas où

2. Soit X une variable aléatoire à densité et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction, alors l'espérance de $\varphi(X)$ vaut :

$$\mathbb{E}[\varphi(X)] =$$

dans le cas où

Démonstration :

1. Puisque X prend ses valeurs dans un ensemble au plus dénombrable $E = (x_i)_{i \in I}$, Y prend elle aussi ses valeurs dans un ensemble au plus dénombrable $F = (y_j)_{j \in J} = \varphi(E)$. Pour tout indice j de J , notons : $E_j = \{i \in I \mid \varphi(x_i) = y_j\}$ et $q_j = \mathbb{P}(Y = y_j) = \sum_{i \in E_j} p_i$. Puisque les $(y_j)_{j \in J}$ et les $(q_j)_{j \in J}$ définissent la loi de Y , son espérance vaut :

$$\mathbb{E}[Y] = \sum_{j \in J} y_j q_j = \sum_{j \in J} y_j \left(\sum_{i \in E_j} p_i \right) = \sum_{j \in J} \sum_{i \in E_j} \varphi(x_i) p_i$$

et puisque les $(E_j)_{j \in J}$ forment une partition de I , on obtient en regroupant les deux symboles de sommation :

$$\mathbb{E}[Y] = \sum_{i \in I} \varphi(x_i) p_i.$$

Dans ce qui précède, toutes les manipulations de sommes sont justifiées par l'absolue convergence de la série $\sum_{i \in I} \varphi(x_i) p_i$.

2. Admis. ■

Théorème 3.3.7 (Inégalité de Markov)

On suppose que pour tout $\omega \in \Omega$, on ait $X(\omega) \geq 0$.

Alors pour tout $\lambda > 0$, on a :

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}[X]}{\lambda}.$$

Démonstration :

■

Remarque 3.3.8. Cette inégalité reste vraie si $\mathbb{E}[X] = +\infty$, mais est alors inutile.

3.3.2 Variance

L'espérance est un indicateur important sur une variable aléatoire, mais elle ne donne pas toutes les informations sur la variable ou la quantité observée. Nous allons donc introduire un nouvel indicateur, qui donnera une idée de la répartition des valeurs autour de la moyenne.

Définition 3.3.9 (Variance)

On définit la variance d'une variable aléatoire X (discrète ou à densité) par

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

lorsque $\mathbb{E} [X^2] < \infty$.

Exemple 3.3.10. Pour $X \sim \mathcal{U}_{[0,1]}$, on a déjà vu que $\mathbb{E}[X] = \frac{1}{2}$. Ainsi :

Proposition 3.3.11 (Formule de Koenig)

Lorsque la variance est bien définie, on a la formule suivante :

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

Démonstration:

■

Proposition 3.3.12

Soit X une variable aléatoire admettant une variance ; et soit λ un réel. On a :

1. $\mathbb{V}(X)$;
2. $\mathbb{V}(\lambda X) =$;
3. $\mathbb{V}(X + \lambda) =$.

Démonstration: Exercice.

■

Remarque 3.3.13. Lorsque X admet une variance, on définit également son écart-type, par $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Théorème 3.3.14 (Inégalité de Bienaymé-Tchebychev)

Soit X une variable aléatoire réelle ; on suppose que $\mathbb{V}[X]$ est bien définie. Alors, pour tout $\alpha > 0$, on a :

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \alpha) \leq \frac{\mathbb{V}(X)}{\alpha^2}.$$

Démonstration:

■

Exemple 3.3.15. Si on prend $\alpha = 10\sqrt{\mathbb{V}(X)}$; la probabilité d'obtenir à l'issue de l'expérience une valeur en dehors de l'intervalle $[\mathbb{E}[X] - 10\sqrt{\mathbb{V}(X)}; \mathbb{E}[X] + 10\sqrt{\mathbb{V}(X)}]$ est inférieure à 1%.

3.3.3 Fonction de répartition

Définition 3.3.16

Si X est une variable aléatoire réelle (discrète ou à densité), on définit sa fonction de répartition F_X par :

$$\forall x \in \mathbb{R}, F_X(x) = \mathbb{P}(X \leq x).$$

Proposition 3.3.17

La fonction de répartition F_X d'une variable aléatoire réelle vérifie :

1. F_X est croissante ;
2. F_X est continue à droite ;
3. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.

Démonstration :

1.

2. Soit $x_0 \in \mathbb{R}$; montrons que F_X est continue à droite en x_0 . Puisque F_X est croissante, elle admet une limite à droite en x_0 , notée $F_X(x_0^+) = \lim_{n \rightarrow \infty} F_X(x_0 + \frac{1}{n})$ (par ailleurs, il existe une limite à gauche, mais qui ne nous intéresse pas). Pour montrer la continuité à droite de F_X en x_0 , il nous suffit de montrer que $F_X(x_0) = F_X(x_0^+)$. On note $C_n = \{X \leq x_0 + \frac{1}{n}\}$. Ainsi :

$$F_X(x_0^+) = \lim_{n \rightarrow \infty} \mathbb{P}(C_n) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} C_n\right) = \mathbb{P}(X \leq x_0) = F_X(x_0).$$

3. Déjà, comme F_X est croissante sur \mathbb{R} , on sait que F_X admet des limites en $-\infty$ et en $+\infty$, éventuellement infinies. En particulier, cela signifie que $\lim_{x \rightarrow -\infty} F_X(x) = \lim_{\substack{n \rightarrow +\infty \\ n \in \mathbb{N}}} F_X(-n)$. Pour $n \in \mathbb{N}$, on définit

$A_n = \{X \leq -n\}$. Ainsi :

$$\lim_{n \rightarrow \infty} F_X(-n) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n=0}^{\infty} A_n\right) = \mathbb{P}(\emptyset) = 0.$$

Le cas de la limite en $+\infty$ est laissé en exercice. ■

Proposition 3.3.18

Si X est une variable aléatoire réelle discrète (prenant des valeurs notées x_k), alors F_X est une fonction en escalier avec un nombre au plus dénombrable de sauts, d'amplitude $\mathbb{P}(X = x_k)$ au point x_k .

On a :

$$F_X(x) =$$

Démonstration:

On remarque que, pour tout $x \in \mathbb{R}$, $\{X = x\} = \bigcap_{n=1}^{\infty} \left\{x - \frac{1}{n} < X \leq x\right\}$. On note alors $D_n = \left\{x - \frac{1}{n} < X \leq x\right\}$.

Ainsi,

$$F_X(x) - F_X(x^-) = \lim_{n \rightarrow \infty} \left[F_X(x) - F_X\left(x - \frac{1}{n}\right) \right] = \lim_{n \rightarrow \infty} \mathbb{P}(D_n) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} D_n\right) = \mathbb{P}(X = x). \quad \blacksquare$$

Remarque 3.3.19. De façon plus générale, on a montré que si F_X est continue en x , alors x n'est pas atome de X , au sens où $\mathbb{P}(X = x) = 0$.

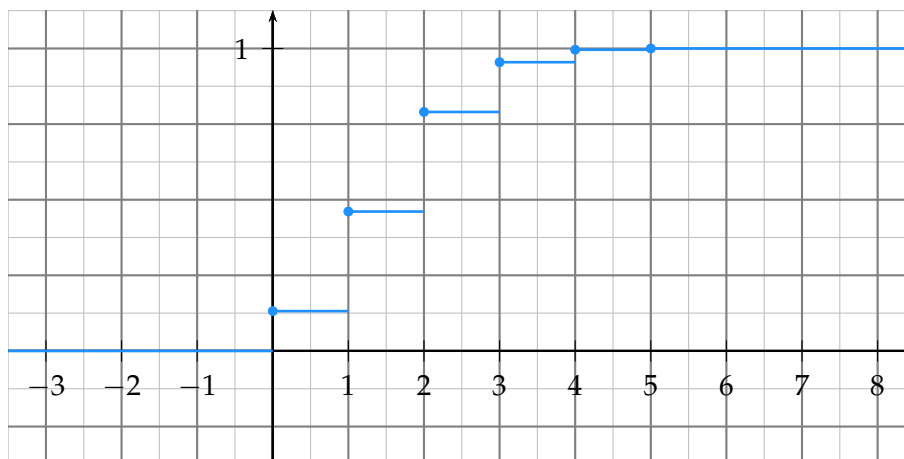


FIGURE 3.1 – Fonction de répartition d’une variable aléatoire suivant la loi binomiale $\mathcal{B}\left(5, \frac{1}{3}\right)$.

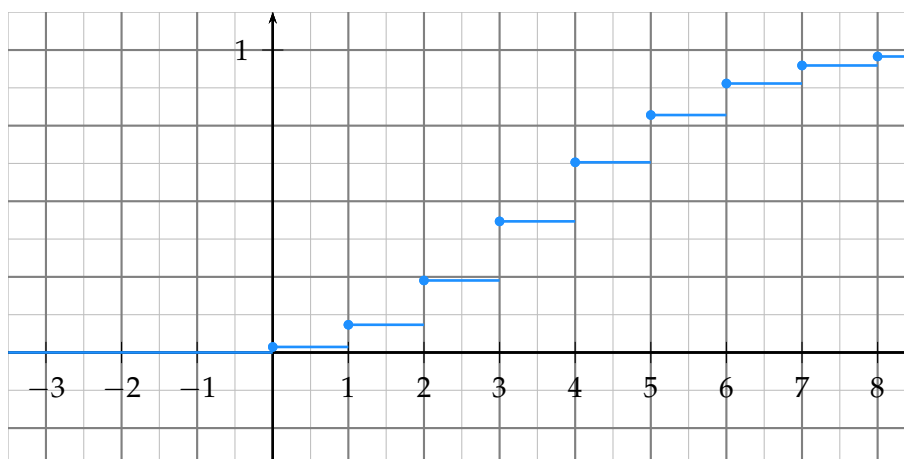


FIGURE 3.2 – Fonction de répartition d’une variable aléatoire suivant la loi de Poisson $\mathcal{P}(4)$.

Proposition 3.3.20

Si X est une variable aléatoire réelle à densité, alors F_X est une fonction continue, et qui admet une dérivée égale à $f_X(x)$ en tout point x où la densité f_X est continue.

On a :

$$F_X(x) =$$

Démonstration:

On a déjà évoqué la continuité de F_X , puisque les variables aléatoires à densité n'admettent pas d'atomes. Soit x_0 un point de continuité de f_X ; il s'agit de montrer que la limite du taux d'accroissement de F_X en x_0 existe et vaut $f_X(x_0)$.

Soit $\varepsilon > 0$; par continuité de f_X en x_0 , il existe $\delta > 0$, tel que $|x - x_0| < \delta \Rightarrow |f_X(x_0) - f_X(x)| < \varepsilon$. Ainsi :

$$\left| \frac{F_X(x_0 + \delta) - F_X(x_0)}{\delta} - f_X(x_0) \right| = \left| \frac{1}{\delta} \int_{x_0}^{x_0 + \delta} [f_X(x) - f_X(x_0)] dx \right| \leq \frac{1}{\delta} \int_{x_0}^{x_0 + \delta} |f_X(x) - f_X(x_0)| dx < \varepsilon. \quad \blacksquare$$

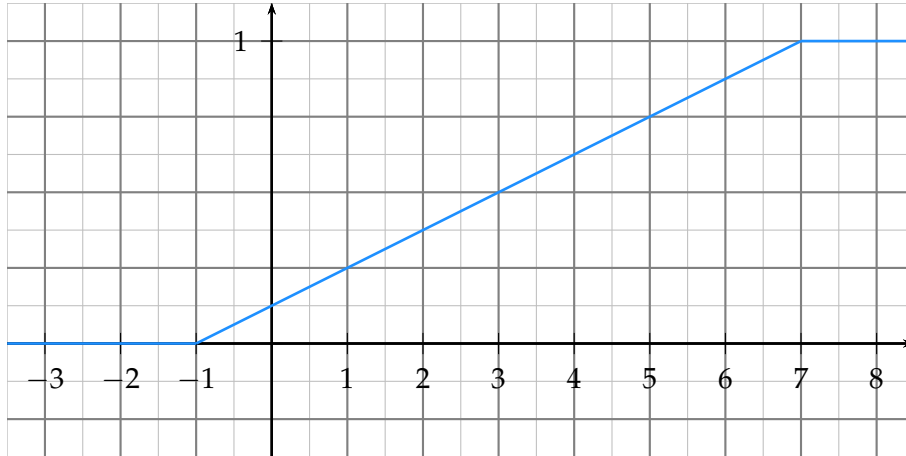


FIGURE 3.3 – Fonction de répartition d'une variable aléatoire suivant la loi uniforme $\mathcal{U}_{[-1,7]}$.

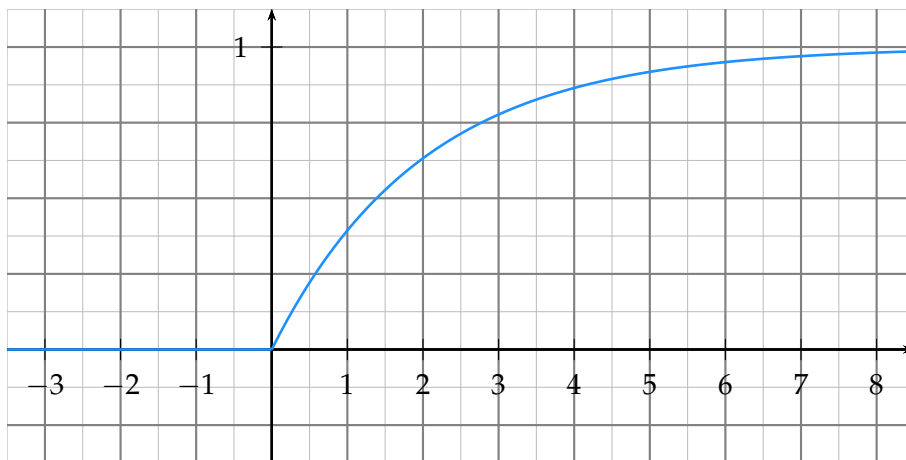


FIGURE 3.4 – Fonction de répartition d'une variable aléatoire suivant la loi exponentielle $\mathcal{E}\left(\frac{1}{2}\right)$.

3.4 Lois normales

3.4.1 Loi normale centrée réduite

Soit $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$. La fonction f est la densité de la loi normale $\mathcal{N}(0,1)$. On remarque que f est paire et qu'elle admet un maximum en 0, égal à $\frac{1}{\sqrt{2\pi}}$.

Soit $X \sim \mathcal{N}(0, 1)$; on note $F_X = \Phi$. Pour tous réels $a \leq b$, on a

$$\mathbb{P}(a \leq X \leq b) =$$

Géométriquement, la valeur de cette probabilité est représentée par l'aire sous la courbe d'équation $y = f(x)$, entre les droites $x = a$ et $x = b$. Dans la figure qui suit, l'aire colorée représente la probabilité que X se trouve entre $-1,5$ et $0,2$.

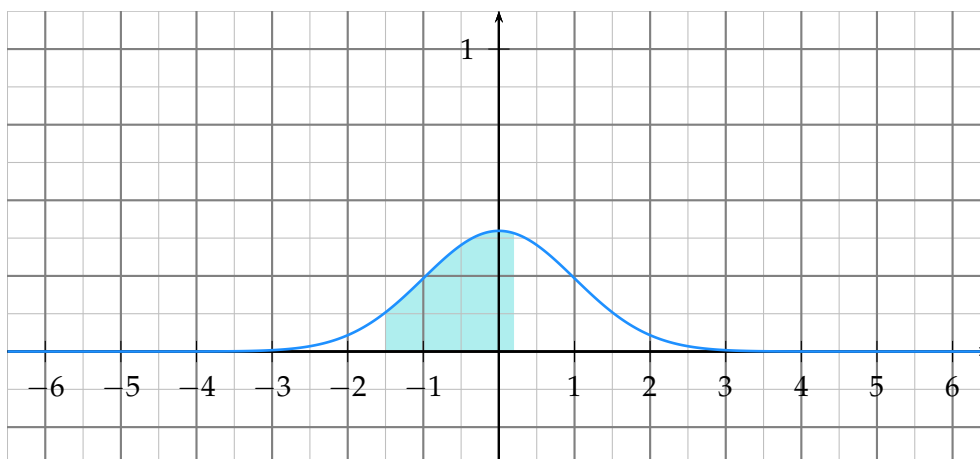


FIGURE 3.5 – Densité de la loi normale centrée réduite.

Compte-tenu de la parité de f , on a aussi, pour tout réel t :

Cette propriété permet de calculer les valeurs de Φ pour les t négatifs quand on les connaît pour les t positifs. Enfin, pour $t > 0$:

$$\mathbb{P}(|X| \leq t) =$$

Ces relations nous serviront en TD pour utiliser la table de la loi normale (voir en annexe).

3.4.2 Cas général

Pour calculer des probabilités de type $\mathbb{P}(a \leq Y \leq b)$ où Y suit une loi $\mathcal{N}(\mu, \sigma^2)$, on utilise la propriété qui suit.

Proposition 3.4.1

Une variable aléatoire Y suit une loi $\mathcal{N}(\mu, \sigma^2)$ si et seulement si $\frac{Y - \mu}{\sigma}$ suit une loi $\mathcal{N}(0, 1)$.

Ainsi, pour $Y \sim \mathcal{N}(\mu, \sigma^2)$:

$$\mathbb{P}(a \leq Y \leq b) =$$

3.5 Couples de variables aléatoires

Définition 3.5.1

Étant données deux variables aléatoires X et Y définies sur un même espace $(\Omega, \mathcal{F}, \mathbb{P})$, à valeurs respectivement dans E et F , on peut définir une variable aléatoire à valeurs dans $E \times F$ appelée *variable aléatoire couple* :

$$(X, Y) : \begin{cases} \Omega & \rightarrow & E \times F \\ \omega & \mapsto & (X(\omega), Y(\omega)) \end{cases}$$

Remarque 3.5.2. Si X et Y sont des variables aléatoires discrètes, déterminer la loi du couple (X, Y) , c'est donner la liste de nombres $(\mathbb{P}((X, Y) = (e, f)))_{(e, f) \in E \times F}$.

Remarque 3.5.3. On écrit :

$$\begin{aligned} \mathbb{P}((X, Y) = (e, f)) &= \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = e \text{ et } Y(\omega) = f\}) \\ &= \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = e\} \cap \{\omega \in \Omega \mid Y(\omega) = f\}) \\ &= \mathbb{P}((X = e) \cap (Y = f)) =: \mathbb{P}(X = e, Y = f), \end{aligned}$$

où on insiste sur le fait que la dernière égalité est une notation.

Exemple 3.5.4. On considère une urne contenant deux boules rouges et une noire. On tire successivement deux boules de cette urne. On note X la couleur de la première, et Y la couleur de la seconde. Quelle est la loi du couple (X, Y) ?

Définition 3.5.5 (Lois marginales)

Les lois marginales du couple (X, Y) sont les lois des variables X et Y étudiées séparément ; la première marginale est la loi de X , la seconde est celle de Y .

Remarque 3.5.6. Si X et Y sont discrètes, alors

$$\mathbb{P}(X = x_i) = \sum_{y_j \in Y(\Omega)} \mathbb{P}(X = x_i, Y = y_j) \quad \text{et} \quad \mathbb{P}(Y = y_j) = \sum_{x_i \in X(\Omega)} \mathbb{P}(X = x_i, Y = y_j).$$

Définition 3.5.7 (Densité d'un couple)

Une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ admettant un nombre fini de discontinuités, et vérifiant $\int_{\mathbb{R}^2} f(x, y) \, dx dy = 1$, est une densité de probabilité.

C'est la densité d'un couple aléatoire (X, Y) si pour tous intervalles I et J de \mathbb{R} , on a :

$$\mathbb{P}((X, Y) \in I \times J) = \int_{I \times J} f(x, y) \, dx dy.$$

Remarque 3.5.8. Si (X, Y) est à densité, alors X et Y sont des variables aléatoires réelles à densité ; elles valent :

$$f_X(x) = \int_{\mathbb{R}} f_{(X, Y)}(x, y) \, dy \quad \text{et} \quad f_Y(y) = \int_{\mathbb{R}} f_{(X, Y)}(x, y) \, dx.$$

Exemple 3.5.9. Attention, si X et Y admettent une densité, le couple (X, Y) n'admet pas forcément de densité. Par exemple, si X suit la loi $\mathcal{N}(0, 1)$ et si $Y = X$, on a : $\mathbb{P}((X, Y) \in \Delta) = \mathbb{P}(X \in [0, 1]) > 0$, où $\Delta = \{(x, y) \in [0, 1]^2 | x = y\}$; alors que le segment Δ est de surface nulle. Si le couple (X, Y) admettait une densité, on aurait ici une contradiction.

Remarque 3.5.10. On peut bien évidemment considérer des couples de variables aléatoires dont une composante est discrète et l'autre à densité, mais on ne se penchera pas sur ce genre de subtilité ici (le couple n'admettant plus de densité, le traitement mathématique de ces couples devient plus ardu).

3.6 Variables aléatoires indépendantes

Définition 3.6.1

Deux variables aléatoires réelles X et Y sont dites indépendantes quand,

Remarque 3.6.2. Comme pour les événements, on définit l'indépendance mutuelle d'une famille $(X_m)_{m \in M}$ de variables aléatoires réelles par :

Théorème 3.6.3

Une famille quelconque de variables aléatoires réelles $(X_m)_{m \in M}$ est mutuellement indépendante si et seulement si pour toute famille finie $J \subset M$ et toute famille de fonctions continues φ_m (avec $m \in J$) telles que les $\varphi_m(X_m)$ soient intégrables,

$$\mathbb{E} \left[\prod_{m \in J} \varphi_m(X_m) \right] = \prod_{m \in J} \mathbb{E} [\varphi_m(X_m)].$$

Démonstration: Admis. ■

Corollaire 3.6.4

Si X_1, \dots, X_n sont des variables aléatoires réelles mutuellement indépendantes, alors on a :

$$\mathbb{E} [X_1 \dots X_n] =$$

Remarque 3.6.5. Il est important de penser que ce corollaire n'admet pas de réciproque. Soit $X \sim \mathcal{N}(0, 1)$ et soit $Y = X^2$. Alors

Proposition 3.6.6 (Identité de Bienaymé)

Si X_1, \dots, X_n sont des variables aléatoires réelles indépendantes deux à deux, alors

$$\mathbb{V} \left(\sum_{i=1}^n X_i \right) =$$

Démonstration:

■

Proposition 3.6.7

Soient X et Y deux variables aléatoires réelles et indépendantes.

1. Si X et Y sont discrètes, alors $\mathbb{P}((X, Y) = (x_i, y_j)) =$
2. Si X et Y admettent des densités f_X et f_Y , alors le couple (X, Y) est à densité; celle-ci vaut $f_{(X,Y)} : (x, y) \mapsto$

Démonstration: Exercice.

■

Proposition 3.6.8

Soient X et Y deux variables aléatoires réelles et indépendantes. On suppose que X et Y admettent des densités f_X et f_Y .

Alors, la variable aléatoire $X + Y$ est à densité, et

$$\forall z \in \mathbb{R}, f_{X+Y}(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x) dx.$$

Démonstration:

■

Exemple 3.6.9. Soient $X_1 \sim \mathcal{N}(0, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(0, \sigma_2^2)$, deux variables aléatoires indépendantes. On a :

$$f_{X_1+X_2}(z) = \int_{\mathbb{R}} f_{X_1}(x) f_{X_2}(z-x) dx = \frac{1}{\sqrt{2\pi}\sigma_1} \frac{1}{\sqrt{2\pi}\sigma_2} \int_{\mathbb{R}} \exp\left(-\frac{x^2}{2\sigma_1^2} - \frac{(z-x)^2}{2\sigma_2^2}\right) dx$$

Or, on écrit :

$$\begin{aligned} \frac{x^2}{2\sigma_1^2} + \frac{(z-x)^2}{2\sigma_2^2} &= \left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}\right)x^2 - 2\frac{z}{2\sigma_2^2}x + \frac{z^2}{2\sigma_2^2} \\ &= \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2}x^2 - 2\frac{z}{2\sigma_2^2}x + \frac{z^2}{2\sigma_2^2} \\ &= \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x^2 - 2\frac{z\sigma_1^2}{\sigma_1^2 + \sigma_2^2}x\right) + \frac{z^2}{2\sigma_2^2} \\ &= \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x^2 - 2\frac{z\sigma_1^2}{\sigma_1^2 + \sigma_2^2}x + \frac{z^2\sigma_1^4}{(\sigma_1^2 + \sigma_2^2)^2} - \frac{z^2\sigma_1^4}{(\sigma_1^2 + \sigma_2^2)^2}\right) + \frac{z^2}{2\sigma_2^2} \\ &= \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x - \frac{z\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2 - \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \frac{z^2\sigma_1^4}{(\sigma_1^2 + \sigma_2^2)^2} + \frac{z^2}{2\sigma_2^2} \\ &= \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x - \frac{z\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2 + \left(\frac{1}{2\sigma_2^2} - \frac{\sigma_1^2}{2\sigma_2^2(\sigma_1^2 + \sigma_2^2)}\right)z^2 \\ &= \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x - \frac{z\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2 + \frac{1}{2(\sigma_1^2 + \sigma_2^2)}z^2 \end{aligned}$$

Ainsi, on obtient :

$$f_{X_1+X_2}(z) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{z^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \int_{\mathbb{R}} \exp\left(-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x - \frac{z\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2\right) dx$$

Mais on remarque que

$$\int_{\mathbb{R}} \exp\left(-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(x - \frac{z\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2\right) dx = \sqrt{2\pi} \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \mathbb{P}(N \in \mathbb{R}) = \sqrt{2\pi} \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}},$$

où $N \sim \mathcal{N}\left(\frac{z\sigma_1^2}{\sigma_1^2 + \sigma_2^2}, \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)$.

En fin de compte,

$$f_{X_1+X_2}(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{z^2}{2(\sigma_1^2 + \sigma_2^2)}\right).$$

En d'autres termes, $X_1 + X_2 \sim \mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$.

Maintenant, si $Y_1 \sim \mathcal{N}(m_1, \sigma_1^2)$ et $Y_2 \sim \mathcal{N}(m_2, \sigma_2^2)$ sont indépendantes, d'après ce qu'on vient de montrer, on a $(Y_1 - m_1) + (Y_2 - m_2) \sim \mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$. On en déduit alors immédiatement que $Y_1 + Y_2 \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Chapitre 4

Théorèmes limites

4.1 Loi des grands nombres

Théorème 4.1.1 (Loi faible des grands nombres)

Soient X_1, \dots, X_n, \dots des variables aléatoires indépendantes (deux à deux), de même loi de probabilité, et admettant une variance finie.

Alors la moyenne empirique $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$ converge (en probabilité) vers l'espérance $\mathbb{E}[X_1]$, au sens où :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mathbb{E}[X_1]| \geq \varepsilon) = 0.$$

Démonstration :

■

Remarque 4.1.2. Il existe une version "forte" de ce théorème, mais qui est beaucoup plus difficile à démontrer.

4.2 Théorème central limite

Théorème 4.2.1 (Théorème central limite)

Soient X_1, \dots, X_n, \dots des variables aléatoires indépendantes (deux à deux), de même loi de probabilité, et admettant une variance finie. On note $S_n = X_1 + \dots + X_n$.

Alors la variable aléatoire $\frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n\mathbb{V}(X_1)}}$ converge (en loi) vers une variable aléatoire de loi normale centrée réduite, au sens où :

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n\mathbb{V}(X_1)}} \leq x\right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

Démonstration : Admis.

■

Chapitre 5

Applications en statistiques

5.1 Modélisation statistique

5.1.1 Principe fondamental de la statistique

La démarche statistique est la suivante : on dispose de x_1, \dots, x_n , qui sont des valeurs observées de variables aléatoires X_1, \dots, X_n indépendantes et identiquement distribuées, de loi Q , inconnue. En ce basant sur ce n -uplet d'observations, on veut trouver la loi de probabilité Q .

Problème : est-ce possible ?

Q est appelée probabilité théorique ; et on suppose que les X_i prennent des valeurs réelles. On définit la probabilité δ_x , pour $x \in \mathbb{R}$, par $\delta_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$.

On définit alors la loi de probabilité empirique :

$$Q_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}.$$

On voit que $\delta_{X_i}(A) = \mathbb{1}_A(X_i)$; c'est une variable aléatoire. Par la loi des grands nombres, on a :

$$\lim_{n \rightarrow \infty} Q_n(A) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{X_i}(A) = \mathbb{E}[\mathbb{1}_A(X_1)] = Q(A).$$

Les lois empiriques Q_n convergent donc (en un certain sens) vers Q ; mais on peut encore se demander à quelle vitesse la convergence a lieu, ou vouloir estimer "l'écart" entre les lois Q_n et Q .

5.1.2 Modèle statistique

Définition 5.1.1 (Modèle statistique)

Soit $E \subset \mathbb{R}$. Un modèle statistique sur E est un couple (E^n, \mathcal{P}_n) , où \mathcal{P}_n est une famille de lois de probabilité sur E^n muni de sa tribu borélienne.

Remarque 5.1.2. La tribu borélienne de E est le plus petit ensemble contenant tous les ouverts de E , qui est stable par complémentaire et par union dénombrable.

Exemple 5.1.3.

Définition 5.1.4 (Modèle paramétré)

1. Le modèle statistique (E^n, \mathcal{P}_n) est *paramétré* par l'ensemble Θ si on peut écrire $\mathcal{P}_n = \{P_\theta^{\otimes n}\}_{\theta \in \Theta}$.
2. On dit que ce modèle statistique est *paramétrique* quand Θ est inclus dans un espace vectoriel de dimension finie. Il est *non-paramétrique* sinon.
3. Le modèle statistique est dit *identifiable* quand l'application $\theta \mapsto P_\theta$ est injective.

Exemple 5.1.5.

5.2 Estimation paramétrique

5.2.1 Estimateur

Définition 5.2.1 (Estimateur)

Un estimateur est une fonction sur E^n qui prend ses valeurs dans Θ (ou dans un ensemble qui contient Θ), que l'on note généralement $\hat{\theta}_n$.

Exemple 5.2.2. Voici quatre exemples d'estimateurs dans le cas du *Pile ou Face* :

5.2.2 Estimateur du maximum de vraisemblance

Définition 5.2.3 (Vraisemblance)

On se place dans le cadre d'un modèle statistique $(E^n, \{P_\theta^{\otimes n}\}_{\theta \in \Theta})$.
 La fonction de vraisemblance est la fonction $L_n : E^n \times \Theta \rightarrow \mathbb{R}_+$ telle que pour tout $\theta \in \Theta$,
 — si les P_θ sont des lois discrètes, $L_n(x_1, \dots, x_n; \theta) = P_\theta(\{x_1\}) \dots P_\theta(\{x_n\})$;
 — si les P_θ sont des lois à densité notées f_θ , $L_n(x_1, \dots, x_n; \theta) = f_\theta(x_1) \dots f_\theta(x_n)$.

Remarque 5.2.4. Plus la valeur prise par la fonction de vraisemblance d'une observation est élevée, et plus cette observation est vraisemblable sous la loi P_θ .

Définition 5.2.5 (Estimateur du maximum de vraisemblance)

Un EMV est un estimateur $\hat{\theta}_n$ qui vérifie :

$$\forall x_1, \dots, x_n \in E, L_n(x_1, \dots, x_n; \hat{\theta}_n(x_1, \dots, x_n)) = \sup_{\theta \in \Theta} L_n(x_1, \dots, x_n; \theta).$$

Remarque 5.2.6. Comme la fonction de vraisemblance est positive, la maximiser revient à maximiser son logarithme généralement noté $l_n(x_1, \dots, x_n; \theta) = \ln L_n(x_1, \dots, x_n; \theta)$. C'est en général plus pratique de maximiser une somme plutôt qu'un produit.

Exemple 5.2.7.

Remarque 5.2.8. En pratique, l'idée est d'obtenir un candidat pour être l'EMV. Si ce candidat n'est pas l'EMV, alors il y a des chances qu'il n'ait pas un bon comportement asymptotique ; on se rendra rapidement compte qu'il ne nous intéresse pas. Ce n'est donc pas vraiment la peine de vérifier s'il s'agit bien d'un maximum.

5.2.3 Critères de performance des estimateurs

Remarque 5.2.9. Quand on parle d'estimateurs, bien souvent, on peut désigner invariablement la fonction sur E^n (dans l'exemple précédent $\hat{\theta}_n = \bar{x}_n$) et la variable aléatoire associée (dans l'exemple $\hat{\theta}_n = \overline{X_n}$).

Définition 5.2.10 (Biais)

Un estimateur $\hat{\theta}_n$ est dit :

- sans biais, si $\mathbb{E}_\theta [\hat{\theta}_n] = \theta$;
- asymptotiquement sans biais, si $\mathbb{E}_\theta [\hat{\theta}_n] \xrightarrow{n \rightarrow \infty} \theta$.

Remarque 5.2.11. La notation \mathbb{E}_θ désigne l'espérance sous la loi P_θ . En d'autres termes, on se pose la question suivante : "supposons que le paramètre cherché soit θ ; sous cette hypothèse, l'espérance de $\hat{\theta}_n$ vaut-elle θ ?"

Exemple 5.2.12.

Définition 5.2.13 (Risque quadratique)

On définit le risque quadratique d'un estimateur $\hat{\theta}_n$ par

$$R_\theta (\hat{\theta}_n) = \mathbb{E}_\theta \left[(\hat{\theta}_n - \theta)^2 \right].$$

Exemple 5.2.14.

Définition 5.2.15 (Consistance)

L'estimateur $\hat{\theta}_n$ est dit consistant si il converge en probabilité vers θ , c'est-à-dire :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P_\theta (|\hat{\theta}_n - \theta| \geq \varepsilon) = 0.$$

Exemple 5.2.16.

Définition 5.2.17 (Vitesse)

Soit $(v_n)_{n \in \mathbb{N}}$ une suite de réels qui tend vers $+\infty$.

On dit que l'estimateur $\hat{\theta}_n$ est de vitesse v_n si $v_n (\hat{\theta}_n - \theta)$ converge en loi vers une variable aléatoire de loi $l(\theta)$ non-dégénérée (c'est-à-dire autre chose qu'une variable aléatoire constamment égale à 0 ou à l'infini); en d'autres termes :

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow \infty} P_\theta (v_n (\hat{\theta}_n - \theta) \leq x) = \mathbb{P} (Y \leq x),$$

où Y suit la loi $l(\theta)$ dans l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

Si $l(\theta)$ est une loi normale, on dit que $\hat{\theta}_n$ est asymptotiquement normal.

Exemple 5.2.18.

Théorème 5.2.19 (Delta-méthode)

On suppose que l'estimateur $\hat{\theta}_n$ est de vitesse v_n . Soit g une fonction de classe \mathcal{C}^1 , telle que $g'(\theta) \neq 0$. Alors, la variable $v_n (g(\hat{\theta}_n) - g(\theta))$ converge en loi vers $g'(\theta)l(\theta)$; en d'autres termes

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow \infty} P_\theta (v_n (g(\hat{\theta}_n) - g(\theta)) \leq x) = \mathbb{P} (g'(\theta)Y \leq x),$$

où Y suit la loi $l(\theta)$ dans l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

Démonstration: Admis. ■

Remarque 5.2.20. Ce résultat est important; en effet, si on a facilement accès à un estimateur $\hat{\theta}_n$ de θ , alors que la quantité qu'on souhaiterait estimer est $g(\theta)$, alors on peut obtenir la vitesse de $g(\hat{\theta}_n)$ à partir de celle de $\hat{\theta}_n$.

5.3 Intervalle de confiance

5.3.1 Intervalle de confiance par excès

Définition 5.3.1 (Intervalle de confiance par excès)

Soit $\alpha \in]0, 1[$; un intervalle de confiance par excès pour θ de niveau de confiance $(1 - \alpha)$ est un intervalle I de \mathbb{R} tel que :

$$\forall \theta \in \Theta, P_\theta (\theta \in I) \geq 1 - \alpha.$$

Exemple 5.3.2.

5.3.2 Intervalle de confiance asymptotique

Définition 5.3.3 (*Intervalle de confiance asymptotique*)

Soit $\alpha \in]0, 1[$; un intervalle de confiance asymptotique pour θ de niveau de confiance $(1 - \alpha)$ est un intervalle I de \mathbb{R} tel que :

$$\forall \theta \in \Theta, P_\theta (\theta \in I) \xrightarrow[n \rightarrow \infty]{} 1 - \alpha.$$

Exemple 5.3.4.

Lemme 5.3.5 (*Slutsky*)

Soit $\hat{\theta}_n$ un estimateur consistant de θ de vitesse notée v_n , et soit $g : \Theta \rightarrow \mathbb{R}^*$ une fonction continue.

Si

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow \infty} P_\theta \left(v_n \frac{\hat{\theta}_n - \theta}{g(\theta)} \leq x \right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{t^2}{2} \right) dt,$$

Alors

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow \infty} P_\theta \left(v_n \frac{\hat{\theta}_n - \theta}{g(\hat{\theta}_n)} \leq x \right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{t^2}{2} \right) dt.$$

Démonstration: Admis. ■

Exemple 5.3.6.

Annexe A

Lois de probabilités classiques

A.1 Lois discrètes

Loi uniforme (équiprobabilité) $\mathcal{U}(\{1, \dots, n\})$

Cette loi modélise une expérience à n issues possibles, chacune ayant la même probabilité $\frac{1}{n}$. Par exemple, X suit cette loi si elle représente la valeur obtenue par un lancer de dé équilibré, la valeur d'une carte tirée dans un jeu (où elles sont toutes indiscernables au toucher)...

$$\forall k \in \{1, \dots, n\}, \mathbb{P}(X = k) = \frac{1}{n}, \mathbb{E}[X] = \frac{n+1}{2}, \mathbb{V}(X) = \frac{n^2 - 1}{12}.$$

Loi de Bernoulli $b(p)$

Cette loi modélise une expérience à deux issues possibles (succès et échec), dont la probabilité de succès vaut p . Par exemple, X suit cette loi si elle représente le résultat d'un jeu du *Pile ou Face* où l'on n'a pas nécessairement la même chance de tomber sur l'un ou l'autre des côtés.

$$\mathbb{P}(X = 1) = p \text{ et } \mathbb{P}(X = 0) = 1 - p, \mathbb{E}[X] = p, \mathbb{V}(X) = p(1 - p).$$

Loi binomiale $\mathcal{B}(n, p)$

Cette loi modélise le nombre de succès obtenus au cours de n expériences indépendantes, chacune ayant une probabilité p de succès. Par exemple, X suit cette loi si elle représente le nombre de *Pile* obtenus sur n lancers avec une pièce éventuellement "faussée".

$$\forall k \in \{0, \dots, n\}, \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \mathbb{E}[X] = np, \mathbb{V}(X) = np(1 - p).$$

Loi géométrique $\mathcal{G}(p)$

Cette loi modélise le rang du premier succès dans une suite d'expériences indépendantes, chacune ayant une probabilité p de succès. Par exemple, X suit cette loi si elle représente la première fois qu'on obtient *Pile* dans une suite de *Pile ou Face* indépendants avec probabilité p de tomber sur *Pile*.

$$\forall k \in \mathbb{N}^*, \mathbb{P}(X = k) = (1 - p)^{k-1} p, \mathbb{E}[X] = \frac{1}{p}, \mathbb{V}(X) = \frac{1 - p}{p^2}.$$

Loi de Poisson $\mathcal{P}(\lambda)$

Elle modélise de manière relativement satisfaisante le nombre X de clients dans certains types de file d'attente. Elle sert aussi à modéliser les occurrences d'événements peu probables. Le paramètre λ est strictement positif.

$$\forall k \in \mathbb{N}, \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \mathbb{E}[X] = \lambda, \mathbb{V}(X) = \lambda.$$

A.2 Lois à densité

Loi uniforme $\mathcal{U}([a, b])$

Cette loi modélise des phénomènes dont les valeurs sont uniformément réparties sur un intervalle $[a, b]$ (avec ou sans les bornes), tel que $-\infty < a < b < +\infty$.

$$f_X(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x), F_X(x) = \frac{x-a}{b-a} \mathbb{1}_{[a,b]}(x) + \mathbb{1}_{]b,+\infty[}(x), \mathbb{E}[X] = \frac{a+b}{2}, \mathbb{V}(X) = \frac{(b-a)^2}{12}.$$

Loi exponentielle $\mathcal{E}(\lambda)$

Cette loi modélise généralement des temps d'attente ou des durées de vie. Le paramètre λ est strictement positif.

$$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x), F_X(x) = (1 - e^{-\lambda x}) \mathbb{1}_{\mathbb{R}_+}(x), \mathbb{E}[X] = \frac{1}{\lambda}, \mathbb{V}(X) = \frac{1}{\lambda^2}.$$

Loi normale $\mathcal{N}(m, \sigma^2)$

Cette loi apparaît notamment dans le cadre du théorème central limite : elle permet donc parfois d'approximer le comportement d'une grande somme de variables aléatoires indépendantes et identiquement distribuées.

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \mathbb{E}[X] = m, \mathbb{V}(X) = \sigma^2.$$

Annexe B

Table de la loi normale centrée réduite

La courbe ci-dessous représente la densité de la loi normale centrée réduite, $x \mapsto \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$.

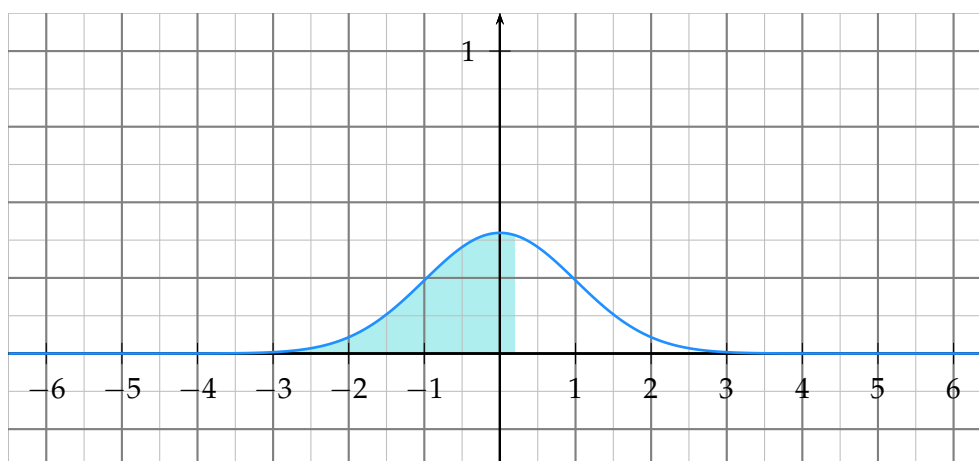


FIGURE B.1 – Densité de la loi normale centrée réduite.

L'aire coloriée sous la courbe est égale à la fonction de répartition prise au point t :

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx = \mathbb{P}(X \leq t),$$

où $X \sim \mathcal{N}(0, 1)$.

La table suivante donne, pour tout t allant de 0 jusqu'à 3,6 par pas de 0,01, la valeur de $10^5 \Phi(t)$. Ces valeurs sont arrondies à l'unité la plus proche. Par exemple, on peut y lire : $\Phi(1,73) \simeq 0,95818$.

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	50000	50399	50798	51197	51595	51994	52392	52790	53188	53586
0,1	53983	54380	54776	55172	55567	55962	56356	56749	57142	57535
0,2	57926	58317	58706	59095	59483	59871	60257	60642	61026	61409
0,3	61791	62172	62552	62930	63307	63683	64058	64431	64803	65173
0,4	65542	65910	66276	66640	67003	67364	67724	68082	68439	68793
0,5	69146	69497	69847	70194	70540	70884	71226	71566	71904	72240
0,6	72575	72907	73237	73565	73891	74215	74537	74857	75175	75490
0,7	75804	76115	76424	76730	77035	77337	77637	77935	78230	78524
0,8	78814	79103	79389	79673	79955	80234	80511	80785	81057	81327
0,9	81594	81859	82121	82381	82639	82894	83147	83398	83646	83891
1,0	84134	84375	84614	84849	85083	85314	85543	85769	85993	86214
1,1	86433	86650	86864	87076	87286	87493	87698	87900	88100	88298
1,2	88493	88686	88877	89065	89251	89435	89617	89796	89973	90147
1,3	90320	90490	90658	90824	90988	91149	91309	91466	91621	91774
1,4	91924	92073	92220	92364	92507	92647	92785	92922	93056	93189
1,5	93319	93448	93574	93699	93822	93943	94062	94179	94295	94408
1,6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449
1,7	95543	95637	95728	95818	95907	95994	96080	96164	96246	96327
1,8	96407	96485	96562	96638	96712	96784	96856	96926	96995	97062
1,9	97128	97193	97257	97320	97381	97441	97500	97558	97615	97670
2,0	97725	97778	97831	97882	97932	97982	98030	98077	98124	98169
2,1	98214	98257	98300	98341	98382	98422	98461	98500	98537	98574
2,2	98610	98645	98679	98713	98745	98778	98809	98840	98870	98899
2,3	98928	98956	98983	99010	99036	99061	99086	99111	99134	99158
2,4	99180	99202	99224	99245	99266	99286	99305	99324	99343	99361
2,5	99379	99396	99413	99430	99446	99461	99477	99492	99506	99520
2,6	99534	99547	99560	99573	99585	99598	99609	99621	99632	99643
2,7	99653	99664	99674	99683	99693	99702	99711	99720	99728	99736
2,8	99744	99752	99760	99767	99774	99781	99788	99795	99801	99807
2,9	99813	99819	99825	99831	99836	99841	99846	99851	99856	99861
3,0	99865	99869	99874	99878	99882	99886	99889	99893	99896	99900
3,1	99903	99906	99910	99913	99916	99918	99921	99924	99926	99929
3,2	99931	99934	99936	99938	99940	99942	99944	99946	99948	99950
3,3	99952	99953	99955	99957	99958	99960	99961	99962	99964	99965
3,4	99966	99968	99969	99970	99971	99972	99973	99974	99975	99976
3,5	99977	99978	99978	99979	99980	99981	99981	99982	99983	99983
3,6	99984	99985	99985	99986	99986	99987	99987	99988	99988	99989